Diophantine approximation and Dynamical systems

PhD Dissertation, July 2012

by Jonas Lindstrøm Jensen

Supervisor: Simon Kristensen

Department of Mathematics – Aarhus University

Abstract

In this dissertation we consider approximation of real numbers by rationals with certain restrictions on the denominators. First we consider numbers that are badly approximable from the left by fractions with denominator of the form q^n for a fixed integer $q \geq 2$ and all $n \geq 0$. More precisely we let $c \in (0,1)$ and consider the $x \in [0,1)$ such that

$$x - \frac{k}{q^n} \geq \frac{c}{q^n}$$

for all $n \geq 0$ and $\frac{k}{q^n} \leq x$. The set of these x is a nullset with respect to the Lebesgue measure, and we give a formula on how to calculate the Hausdorff dimension of this set. This formula is then generalized to the case where $q = \beta > 1$ is in a certain dense set of real numbers, namely the simple numbers.

Then, we let $p, q \geq 2$ be integers and consider approximation by fractions with denominators of the form $p^n q^m$ for all $n, m \geq 0$ and prove some versions of classical results in this settings.

Finally we show a result related to the famous Littlewood conjecture. We prove that there is a subset of the badly approximable numbers of full Hausdorff dimension such that a family of conjectures related to the Littlewood conjecture is simultaneously true on this set, namely the Littlewood conjecture, the mixed Littlewood conjecture and a hybrid of a conjecture by Cassel and Littlewoods conjecture.

Resumé

I denne afhandling undersøges det, hvor godt man kan approksimere reelle tal ved brøker, hvor vi sætter visse betingelser på nævnerne i brøkerne. Først betragter vi mængden af tal, der er dårligt approksimérbare fra venstre ved brøker med nævner på formen q^n, hvor $q \geq 2$ er et helt tal og $n \geq 0$. Mere præcist lader vi $c \in (0, 1)$ og ser vi på de $x \in [0, 1)$, der opfylder at

$$x - \frac{k}{q^n} \geq \frac{c}{q^n}$$

for alle $n \geq 0$ og $\frac{k}{q^n} \leq x$. Det viser sig, at det bliver en Lebesgue-nulmængde, og vi giver en formel for udregning af Hausdorffdimensionen af mængden af disse x. Resultatet generaliseres til også at gælde, når $q = \beta > 1$ tilhører en tæt mængde af reelle tal, nemlig de simple tal.

Derefter undersøges det, hvad der sker med approksimationen, hvis man lader $p, q \geq 2$ være forskellige hele tal og tillader af nævnerne kan være på formen $p^n q^m$. Vi ser her på, hvorledes nogle klassiske resultater i Diofantisk approksimation tager sig ud, når vi begrænser nævnerne på denne måde.

Til sidst vises et resultat relateret til Littlewoods formodning, nemlig at der findes en delmængde af de dårligt approximerbare tal af fuld Hausdorff dimension, således at en række formodninger relateret til Littlewoods formdning alle er sande simultant på denne mængde. Det drejer sig om Littlewoods formodning, mixed Littlewoods formodning og en hybrid mellem en formodning af Cassel og Littlewoods formodning.

Acknowledgements

I would like to thank my supervisor Simon Kristensen for helpful discussions and supervision, and for making the last three years a both educational and entertaining experience. I would also like to thank Stephen Harrap for carefully proof reading this dissertation and for many helpful comments and corrections.

Contents

Chapter 1

Introduction and summary

1.1 Diophantine approximation

Diophantine approximation is a research area that deals with approximation of real numbers by rational numbers – the rationals are dense in the reals, and we are interested in saying something about *how* dense they are. The most basic result is the following classical theorem by Dirichlet which is proved using his pigeon hole principle.

Theorem 1.1 (Dirichlet)**.** *Let $x \in \mathbb{R}$. For any $N \in \mathbb{N}$ there is a non-negative integer $0 \leq n \leq N$ and $m \in \mathbb{Z}$ such that*

$$\left|x - \frac{m}{n}\right| \leq \frac{1}{nN}.$$

Proof. Let $x \in \mathbb{R}$ and let $x = [x] + \{x\}$ where $[x] \in \mathbb{Z}$ and $\{x\} = [0, 1)$. By Dirichlet's pigeon hole principle, two elements from the set $\{\{nx\} \mid m = 0, 1, \ldots, N\}$ are in one of the intervals

$$\left[\frac{i}{N}, \frac{i+1}{N}\right)$$

for $i = 0, 1, \ldots, N-1$, say $\{n_1x\}$ and $\{n_2x\}$ with $0 \leq n_1 < n_2 \leq N$. Now

$$\frac{1}{N} \geq |\{n_2x\} - \{n_1x\}| = |(n_2 - n_1)x - ([n_2x] - [n_1x])|,$$

so if we let $m = [n_2x] - [n_1x] \in \mathbb{Z}$ and $n = (n_2 - n_1) \in \{1, 2, \ldots, N\}$ we have

$$\frac{1}{N} \geq |mx - n|$$

and hence

$$\left|x - \frac{m}{n}\right| \leq \frac{1}{nN}$$

as desired. $\square$

For $x \in \mathbb{R}$ we let

$$\|x\| = \min_{y \in \mathbb{Z}} |x - y|$$

denote the distance to the distance to the nearest integer. We then get the following corollary.

Corollary 1.2. *Let $x \in \mathbb{R}$. Then*

$$\liminf_{n \to \infty} n \|nx\| \leq 1. \tag{1.1}$$

The corollary above is not far from the best approximation possible. Hurwitz improved this to $\liminf_n n \|nx\| \leq \frac{1}{\sqrt{5}}$ and showed that this is best possible since the golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$ satisfies

$$\liminf_n n \|n\varphi\| = \frac{1}{\sqrt{5}}.$$

The numbers x such that $\liminf_n n \|nx\| > 0$ are the so called *Badly approximable numbers*,

$$\mathrm{Bad} = \{x \in \mathbb{R} \mid \liminf_n n \|nx\| > 0\}.$$

These numbers have attracted a lot of interest from researchers and they are the main topic of this dissertation. Recall the following classical theorem by Khinchine.

Theorem 1.3 (Khinchine)**.** *Let $\psi : \mathbb{R} \to (0, \infty)$ be a continuous, non-increasing function. Then the inequality*

$$\|nx\| < \psi(n) \tag{1.2}$$

has infinitely many solutions in integers $n > 0$ for almost all $x \in \mathbb{R}$ with respect to the Lebesgue measure if

$$\sum_{n=1}^{\infty} \psi(n) = \infty, \tag{1.3}$$

and if the series in (1.3) *converges then there are only finitely many $n > 0$ such that* (1.2) *is true.*

It follows from this theorem that Bad is a nullset with respect to the Lebesgue measure.This is because if $x \notin \mathrm{Bad}$ then for any $m \in \mathbb{N}$ we have

$$n \|nx\| < \frac{1}{m} \tag{1.4}$$

for infintely many n. Now let

$$E_m = \{x \in \mathbb{R} \mid n\,\|nx\| < \frac{1}{m} \text{ for infinitely many } n \in \mathbb{N}\}$$

then E_m is of full measure because

$$\sum_{n=1}^{\infty} \frac{1}{mn} = \infty,$$

Now

$$\text{Bad} = \mathbb{R} \setminus \bigcap_{m=1}^{\infty} E_m$$

so Bad is a nullset.

Throughout the thesis we will let $O(1)$ denote a bounded quantity and let $O(X) = O(1)X$. If $Y = O(X)$ we write $Y \ll X$ and if both $X \ll Y$ and $Y \ll X$ is true, we write $X \asymp Y$.

1.2 Hausdorff dimension

Many of the results in this thesis use the notion of *Hausdorff dimension*. Let $Y \subseteq X$ where (X, d) is a metric space. We define the *diameter* of a bounded set $U \subseteq X$ as

$$\operatorname{diam}(U) = \sup_{x,y \in U} d(x, y),$$

and we call a collection of sets $\{U_i\}$ a δ-cover of Y if the union of the sets cover Y and $\operatorname{diam}(U_i) \leq \delta$ for all i. Now let for $s \geq 0$

$$\mathscr{H}^s_\delta(Y) = \inf \left\{ \sum_i \operatorname{diam}(U_i)^s \mid \{U_i\} \text{ is a } \delta\text{-cover of } Y \right\}.$$

As δ decreases this value increases, and we let

$$\mathscr{H}^s(Y) = \lim_{\delta \to 0} \mathscr{H}^s_\delta(Y)$$

which may be a non-negative number or $+\infty$. We call this the *s-dimensional Hausdorff measure of Y*. This is in fact a measure, and it generalizes the Lebesgue measure λ^n on $\mathbb{R}^n$ since

$$\mathscr{H}^n(Y) = \frac{\lambda^n(Y)}{\lambda^n(B(0, \frac{1}{2}))},$$

where $B(0, \frac{1}{2}) \in \mathbb{R}^n$ is the n-dimensional ball of radius $\frac{1}{2}$.

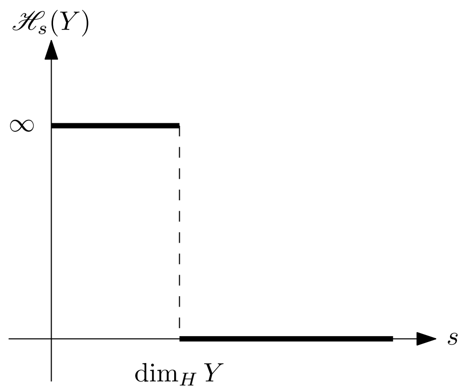


Figure 1.1: A plot of $\mathscr{H}^s(Y)$ as a function of s.

The Hausdorff dimension of Y is in some sense the appropriate s to pick when trying to measure Y with the Hausdorff measure. If we for a moment only consider integer values of s and let Y be a 2-dimensional set, it is clear that it does not make sense to measure neither the length which would be ∞, nor the volume which would be 0. So we are looking for the critical value, where we make this jump from ∞ to 0. This happens in a single point, because we see that for $t > s$ we have

$$\mathscr{H}^t_\delta(Y) \leq \delta^{t-s}\mathscr{H}^s_\delta(Y),$$

so letting $\delta \to 0$ we see that if $\mathscr{H}^s(Y) < \infty$ we must have $\mathscr{H}^t(Y) = 0$. This proves that the graph of $s \mapsto \mathscr{H}^s(Y)$ looks something like figure 1.1. The critical value s where the graph jumps from $+\infty$ to 0 is called the *Hausdorff dimension*, and is denoted $\dim_H Y$. In this jump the Hausdorff measure might be $0, \infty$, or any positive real number.

We will use Hausdorff dimension to measure the sizes of nullsets. Jarník proved [11] that $\dim_H \mathrm{Bad} = 1$, so even though Bad is a nullset, it is still large in the sense of Hausdorff dimension.

1.3 Shift spaces and words

Some of the variations of Bad we shall consider can be rephrased in terms of dynamical systems in form of shift spaces and words. We provide a brief introduction to both here. Let $q \geq 2$ be an integer and consider the set

$$\Sigma = \Sigma_q = \{0, 1, 2, \ldots, q-1\}^{\mathbb{N}},$$

of right-infinite words on q-digits. Throughout the dissertation we will write words in Σ in bold, that is $\mathbf{a_1 a_2} \cdots = (a_1, a_2, \ldots) \in \Sigma$. For simplicity we let a finite word denote the corresponding word with a tail of zeros,

$$\mathbf{a_1} \cdots \mathbf{a_n} = \mathbf{a_1} \cdots \mathbf{a_n} \mathbf{0} \cdots$$

When we compare two words we use the *lexicographical ordering* where the leftmost digit is the most significant, so for example

$$\mathbf{1221} > \mathbf{1212}.$$

When an expression should be considered as one digit we write it in square brackets, so

$$[\mathbf{1}+\mathbf{2}]\mathbf{345} = \mathbf{3345}.$$

We define powers of words as concatenation, so

$$(\mathbf{a_1} \cdots \mathbf{a_m})^{\infty} = \mathbf{a_1} \cdots \mathbf{a_m} \mathbf{a_1} \cdots \mathbf{a_m} \mathbf{a_1} \cdots$$

and

$$(\mathbf{a_1} \cdots \mathbf{a_m})^{n} = \mathbf{a_1} \cdots \mathbf{a_m} \mathbf{a_1} \cdots \mathbf{a_m} \cdots \mathbf{a_1} \cdots \mathbf{a_m}$$

where we on the right have n copies of $\mathbf{a_1} \cdots \mathbf{a_m}$ for $n \geq 0$.

Define the *shift-map* $\sigma : \Sigma \to \Sigma$ by

$$\sigma(\mathbf{a_1 a_2 a_3} \cdots) = \mathbf{a_2 a_3} \cdots.$$

and define $\theta = \theta_q : \Sigma \to [0,1]$ by

$$\theta(\mathbf{a_1 a_2 a_3} \cdots) = \sum_{i=1}^{\infty} \frac{a_i}{q^i},$$

The map θ is not one-to-one everywhere since for instance

$$\theta(\mathbf{0999} \cdots) = \theta(\mathbf{1000} \cdots) = 0.1$$

when $q = 10$. Given $x \in [0,1)$ we get

$$\theta(\mathbf{a_1 a_2} \cdots) = x$$

if we let

$$a_i = \lfloor q T^{i-1} x \rfloor$$

be the *q-expansion of* x where the *times-q map* $T = T_q : [0,1) \to [0,1)$ is defined by $Tx = \{qx\}$. We also get

$$T\theta(c) = \theta(\sigma(c))$$

if $\theta(c), \theta(\sigma(c)) < 1$.

1.4 Summary of new results

In Chapter 2 we will consider onesided approximation of real numbers by fractions of the form $\frac{k}{q^n}$ where $q \geq 2$ is a fixed integer. Studying the following variant of Bad will be our main focus. We will fix $c \in [0,1]$ and consider numbers $x \in \mathbb{R}$ such that

$$x - \frac{k}{q^n} \geq \frac{c}{q^n}$$

for all $k, n \geq 0$ with $\frac{k}{q^n} \leq x$. This is equivalent with

$$\{q^n x\} \geq c$$

for all $n \geq 0$ where where $\{\cdot\}$ denotes the fractional part of a number, and since this is invariant under integer translation, we let

$$F_c = \{x \in [0,1) \mid \{q^n x\} \geq c \text{ for all } n \geq 0\}$$

denote this variant of Bad. Urbanski [24] proved that the map $c \mapsto \dim_H F_c$ is continuous and constant almost everywhere with respect to the Lebesgue measure, and Nilsson [17] characterized the intervals where the map is constant. In Chapter 2 we prove the following theorem regarding the Hausdorff dimension of F_c.

Theorem 2.4. *Let $c = \sum_{i=1}^{m} \frac{c_i}{q^i} > 0$ with $0 \leq c_i < q$ and let*

$$n = \min\{1 \leq j \leq m \mid \mathbf{c_{j+1} \cdots c_m} \leq \mathbf{c_1 \cdots c_{m-j}}\}$$

and

$$a_i = \begin{cases} q - c_i - 1 & \text{for } 1 \leq i < n \\ q - c_i & \text{for } i = n \end{cases}.$$

Then

$$\dim_H F_c = \frac{\log \rho}{\log q}$$

where ρ is the largest real root of

$$x^n - a_1 x^{n-1} - a_2 x^{n-2} - \cdots - a_n.$$

Note that we let both $\mathbf{c_{j+1} \cdots c_m}$ and $\mathbf{c_1 \cdots c_{m-j}}$ denote the empty word when $j = m$, and hence

$$\mathbf{c_{j+1} \cdots c_m} = \mathbf{c_1 \cdots c_{m-j}}$$

in this case.

A corollary of this theorem is that the analogue of Bad where we let c depend on x has Hausdorff dimension one.

The results in Chapter 2 can be found in [12]. The proof given in Chapter 2 is different and simpler than the one in [12], but the paper in its full length is given in Appendix A.

One can generalize the base-q expansion of a real number in the following way. Given a real number $\beta > 1$ we can define the *β-expansion* $\mathbf{b_1 b_2} \cdots$ of a real number $x \in [0,1]$ by

$$b_1 = \lfloor \beta x \rfloor$$

and

$$b_i = \lfloor \beta \{\beta^i x\} \rfloor.$$

for $i \geq 2$. Evidently we have

$$x = \sum_{i=1}^{\infty} \frac{b_i}{\beta^i}.$$

The numbers $\beta > 1$ such that the β-expansion of 1 is finite are the *simple numbers.* In Chapter 3 we prove that the numbers $\rho > 1$ occuring in Theorem 2.4 are precisely the simple numbers, and we use this to give a new proof that the map $c \mapsto \dim_H F_c$ is continuous.

Then we use the theory of β expansions to generalize Theorem 2.4, and consider for $c \in [0,1)$ and a simple β the set of numbers badly approximable from the left by fractions of the form $\frac{k}{\beta^n}$ for integers k, n. By the same argument as above this set is given by

$$F_c(\beta) = \{x \in [0,1) \mid \{\beta^n x\} \geq c \text{ for all } n \geq 0\}.$$

Note that when $\beta = q$ is an integer, it is also simple because then the β-expansion of 1 is $\mathbf{q}$. We prove the following theorem.

Theorem 3.5. *Let $\beta > 1$ be a simple number such that the β expansion of 1 is $\mathbf{b_1} \cdots \mathbf{b_k}$ and let $c \in [0,1)$ have β-expansion $\mathbf{c_1} \cdots \mathbf{c_m}$. Let*

$$n = \min\{1 \leq j \leq m \mid \mathbf{c_{j+1}} \cdots \mathbf{c_m} \leq \mathbf{c_1} \cdots \mathbf{c_{m-j}}\}$$

and

$$t = \min\{1 \leq j \leq k \mid \mathbf{b_{j+1}} \cdots \mathbf{b_k} \leq \mathbf{c_1} \cdots \mathbf{c_m}\}.$$

Assume that $c_j < b_1$ for $j < n$. Let $d_j = b_j - c_1 - 1$ for $j = 2, 3, \ldots, t$,

$$a_i = \begin{cases} b_1 - c_i - 1 & \text{for } i = 1, 2, \ldots, n-1 \\ b_1 - c_i & \text{for } i = n \end{cases}.$$

Then

$$\dim_H F_c(\beta) = \frac{\log \rho}{\log \beta}$$

where $\rho > 0$ *is the spectral radius of*

$$\begin{pmatrix}
a_1 & a_2 & \cdots & a_{n-1} & a_n & d_2 & d_3 & \cdots & d_{t-1} & d_t \\
1 & & & & & 1 & 1 & \cdots & 1 & 1 \\
 & 1 & & & & & & & & \\
 & & \ddots & & & & & & & \\
 & & & 1 & & & & & & \\
1 & 1 & \cdots & 1 & 1 & & & & & \\
 & & & & & 1 & & & & \\
 & & & & & & 1 & & & \\
 & & & & & & & \ddots & & \\
 & & & & & & & & 1 &
\end{pmatrix}.$$

Note that the empty entries in the matrix are zeros. The technique used to prove Theorem 3.5 is the same as that used to prove Theorem 2.4, but the combinatorics involved is more complicated. Note that when $\beta = q$ is an integer, the theorem gives us that

$$\dim_H F_c = \frac{\log \rho}{\log q}$$

where ρ is the spectral radius of

$$\begin{pmatrix}
a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & a_n \\
1 & & & & & \\
 & 1 & & & & \\
 & & \ddots & & & \\
 & & & 1 & & \\
 & & & & 1 &
\end{pmatrix}.$$

The characteristic polynomial of this can be calculated to be $x^n - a_1 x^{n-1} - \cdots - a_n$, so Theorem 2.4 follows from Theorem 3.5.

In Chapter 4 we consider approximation where the denominators are restricted to be in the set

$$A = \{p^\ell q^m \mid \ell, m \geq 0\}.$$

Among other things we prove a version of Khinchine's Theorem.

Theorem 4.3. *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a monotonic function. For almost all $x \in \mathbb{R}$, the inequality*

$$\left| x - \frac{k}{n} \right| < \frac{\psi(n)}{n}$$

is true for infinitely many integers k, n with $n \in A$ if

$$\sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} = \infty,$$

and if

$$\sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} < \infty$$

then the inequality is false for almost all $x \in \mathbb{R}$.

We also prove versions of this theorem in higher dimensions. Badly approximable numbers in this setting are also considered. Here we prove that

$$\liminf_{n \in A} \log n^2 \|nx\| > 0$$

is true for x in a nullset, but

$$\liminf_{n \in A} \log n^{2+\varepsilon} \|nx\| > 0$$

is true for almost all x when $\varepsilon > 0$. Using an effective version of Furstenberg's orbit closure theorem by Bourgain, Lindenstrauss, Michel and Venkatesh [4] we prove the following theorem.

Theorem 4.8. *There is a constant $c > 0$ only dependent on p and q such that the set*

$$\{x \in [0, 1] : \liminf_{n \in A} (\log \log \log n)^c \|nx\| > 0\}$$

is a subset of the union of rationals and Liouville numbers. In particular it has Hausdorff dimension zero.

The results are generalized to the case where the denominator is of the form $q_1^{m_1} q_2^{m_2} \cdots q_k^{m_k}$. The results in this chapter are joint work with Sanju Velani during a stay at University of York in the spring of 2011.

A famous conjecture by Littlewood states that for all $x, y \in \mathbb{R}$, we have

$$\liminf_{n} n \|nx\| \|ny\| = 0. \tag{1.5}$$

If $x \notin \text{Bad}$ we have

$$\liminf_{n} n \|nx\| = 0,$$

and since $\|ny\| \leq \frac{1}{2}$, (1.5) is true in this case. So the conjecture is only interesting when $x, y \in \text{Bad}$.

Related to Littlewood's conjecture is the Mixed Littlewood Conjecture. Here we let $\mathscr{D} = \{q_k\}$ with $q_k \mid q_{k+1}$ be a *pseudo-absolute sequence* and let

$$|n|_{\mathscr{D}} = \min \left\{ \frac{1}{q_k} : q_k \mid n \right\}$$

be a *pseudo absolute value.* Then the Mixed Littlewood conjecture states that

$$\liminf_{n} n \, |n|_{\mathscr{D}} \, \|ny\| = 0 \tag{1.6}$$

for all $y \in \mathbb{R}$.

In Chapter 5 we prove that there is a set $G \subseteq \text{Bad}$ of full Hausdorff dimension such that both these conjectures and a version of a conjecture by Cassel is true on this set.

Theorem 5.1. *Fix $\varepsilon > 0$ and let $\{x_i\} \subseteq \text{Bad}$ be a countable set of badly approximable numbers, and $\{\mathcal{D}_j\}$ a countable set of pseudo-absolute value sequences. Then there is set of $G \subseteq \text{Bad}$ of Hausdorff dimension 1 such that for any $y \in G$, the following is true.*

(i) For any $i \in \mathbb{N}$ and $\gamma \in \mathbb{R}$ the inequality

$$n \|nx_i\| \, \|ny - \gamma\| < \frac{1}{(\log n)^{1/2 - \varepsilon}},$$

has infinitely many solutions $n \in \mathbb{N}$, and

(ii) For any $j \in \mathbb{N}$ and $\delta \in \mathbb{R}$ we have that

$$\liminf_{n \to \infty} n \, |n|_{\mathcal{D}_j} \, \|ny - \delta\| = 0.$$

Furthermore, for each j such that $\mathcal{D}_j = (q_k)$ satisfies the inequality $q_k \leq C^k$ for some $C > 1$, we may replace (5.2) by the stronger statement that for any $\delta \in \mathbb{R}$ the inequality

$$n \, |n|_{\mathcal{D}_j} \, \|ny - \delta\| < \frac{1}{(\log n)^{1/2 - \varepsilon}},$$

has infinitely many solutions $n \in \mathbb{N}$.

The results in Chapter 5 is done joint with Simon Kristensen and Alan Haynes and can be found in [10].

Chapter 2

One-sided Bad

2.1 Introduction

This chapter presents the results from [12], but with significantly simpler proofs. The paper [12] in its full length is given in Appendix A.

Recall that $x \in \mathbb{R}$ is a Badly approximable number if there is $c > 0$ such that

$$\left| x - \frac{k}{n} \right| \geq \frac{c}{n^2}$$

for all $n, k \in \mathbb{Z}$. We consider the problem where we fix an integer $q \geq 2$ and restrict the denominator to be of the form q^n for $n \geq 0$. We let $c \in [0, 1]$ and consider $x \in \mathbb{R}$ such that

$$x - \frac{k}{q^n} \geq \frac{c}{q^n}$$

for all $n \geq 0$ and $\frac{k}{q^n} \leq x$, or equivalently that

$$\{q^n x\} \geq c \tag{2.1}$$

for all $n \geq 0$ where $\{\cdot\}$ denotes the fractional part. Note that we approximate with fractions from the left and that c is fixed. The set of $x \in \mathbb{R}$ that satisfies (2.1) is invariant under integer translation so we consider the set

$$F_c = \{x \in [0, 1) \mid \{q^n x\} \geq c \text{ for all } n \geq 0\}.$$

We see that $F_0 = [0, 1)$ and $F_1 = \emptyset$ and that $F_{c_2} \subseteq F_{c_1}$ when $c_1 \leq c_2$ so if we define $\phi_q = \phi : [0, 1] \to [0, 1]$ by $\phi(c) = \dim_H F_c$ then ϕ is non-increasing.

Urbanski [24] proved that this map is continuous and constant almost everywhere with respect to the Lebesgue measure and Nilsson [17] completely characterized the intervals where ϕ is constant. The main result in this chapter is a formula for calculating the exact value of $\phi(c)$ when c has finite base q-expansion. Figure A.1 is a plot of ϕ for different values of q made using this result.

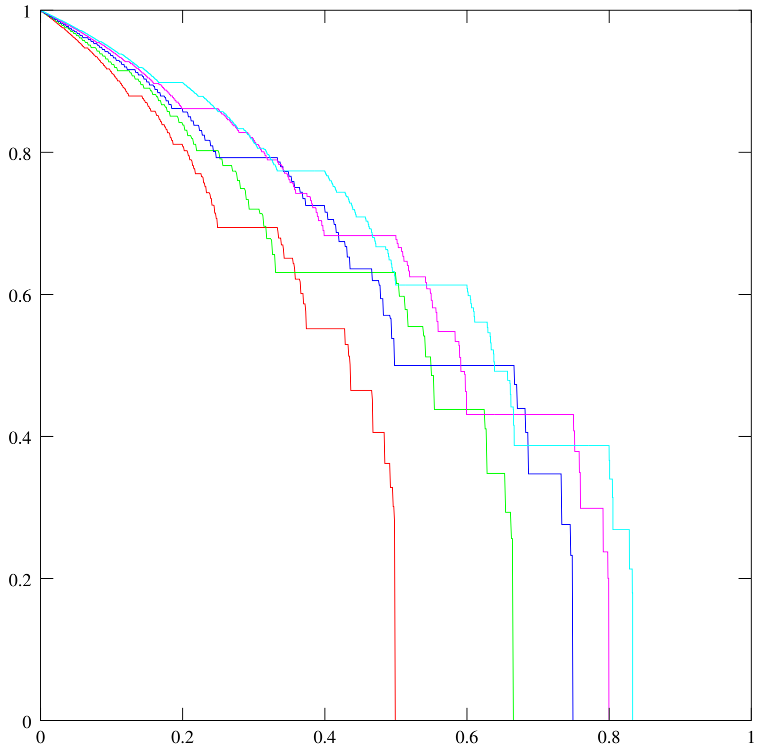


Figure 2.1: *Plots of ϕ for $q = 2$ (red), 3 (green), 4 (blue), 5 (magenta), 7 (light blue).*

2.2 A golden example

To illustrate the technique used to prove the main result in this chapter, we begin with a simple example. Here we fix $q = 2$ and $c = \frac{1}{4}$. Define $T : [0, 1) \to [0, 1)$ by

$$Tx = \{2x\}.$$

We have

$$F = F_{\frac{1}{4}} = \bigcap_{n=0}^{\infty} F^n.$$

where

$$F^n = \bigcap_{i=0}^{n} T^{-i}\left[\frac{1}{4}, 1\right),$$

for $n = 0, 1, \ldots$. We wish to prove that the Hausdorff dimension of F is given by

$$\dim_H F = \frac{\log \varphi}{\log 2},$$

where φ is the largest real solution of $x^2 - x - 1 = 0$, namely the Golden ratio,

$$\varphi = \frac{1+\sqrt{5}}{2}.$$

We can construct F^n as in figure 2.2. Here we draw $T^{-i}[1/4, 1)$ for $i = 0, 1, 2, \ldots$. The intersection of the first n of these is equal to F^n.

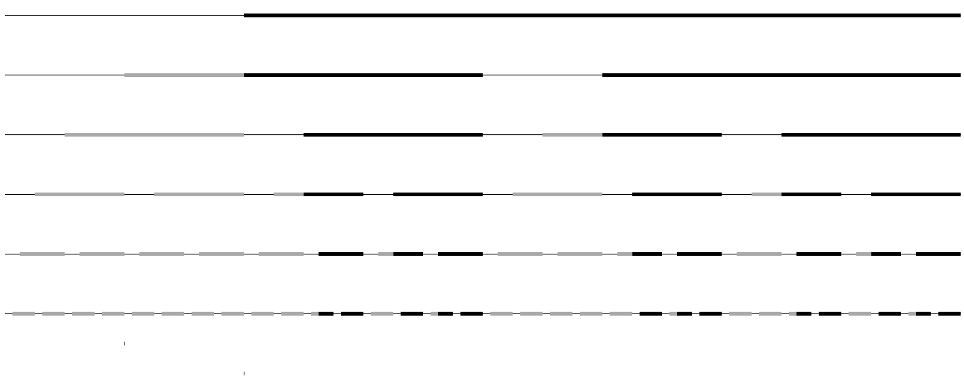


Figure 2.2: *Geometric construction of $F_{\frac{1}{4}}$ when $q = 2$. The grey lines are $T^{-i}[1/4, 1)$ for $i = 0, 1, 2, 3, 4, 5$ and the black lines are the intersections of these.*

We associate a word $C_n \in \Sigma_2$ to F^n for all $n = 0, 1, 2, \ldots$ in the following way. We claim that F^n is a disjoint union of two types of intervals: Some intervals I_1 such that $T^n I_1 = \left[\frac{1}{4}, 1\right)$, which we say is of type **1**, and some I_2 such that $T^n I_2 = \left[\frac{1}{2}, 1\right)$, which we say is of type **2**. We now define C_n by looking at the intervals in F^n from left to right, and for each interval of type **j**, we add a digit **j** to the right of C_n. We then get the following sequence of words.

$$
\begin{aligned}
C_0 &= \mathbf{1} \\
C_1 &= \mathbf{21} \\
C_2 &= \mathbf{121} \\
C_3 &= \mathbf{21121} \\
C_4 &= \mathbf{12121121} \\
C_5 &= \mathbf{2112112121121}
\end{aligned}
$$

We now prove the following lemma.

Lemma 2.1. *We can find C_n by letting $C_0 = \mathbf{1}$, and then iterating from C_i to C_{i+1} for $i \geq 0$ via the morphisms*

$$\mathbf{1} \to \mathbf{21}, \quad \mathbf{2} \to \mathbf{1}.$$

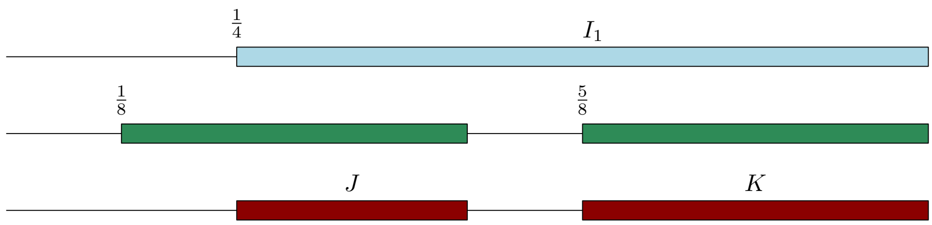
Proof. We prove this by induction. Recall that $F^{n+1} = T^{-n-1}[1/4, 1) \cap F^n$. Assume we have an interval I_1 of type $\mathbf{1}$ in F^n. Consider the intersection

$$I_1 \cap T^{-n-1}\left[\frac{1}{4}, 1\right).$$

Note that

$$T^n\left(I_1 \cap T^{-n-1}\left[\frac{1}{4}, 1\right)\right) = \left[\frac{1}{4}, 1\right) \cap T^{-1}\left[\frac{1}{4}, 1\right).$$

This set is illustrated in the figure below. Here the blue line is $I_1 = [1/4, 1)$, the green lines are $T^{-1}[1/4, 1)$ and the red lines are the intersection of the two. We see that this intersection is two intervals J, K where $TJ = \left[\frac{1}{2}, 1\right)$ and $TK = \left[\frac{1}{4}, 1\right)$.

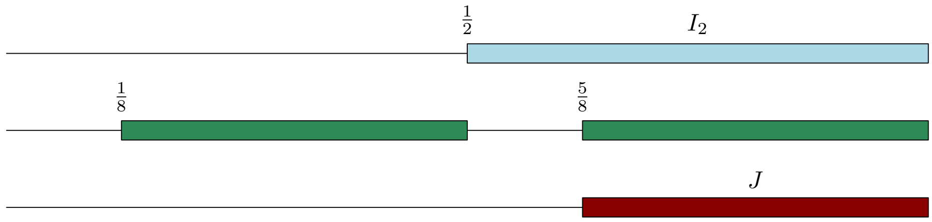


So the intersection $I_1 \cap T^{-n-1}\left[\frac{1}{4}, 1\right)$ consists of two intervals – one of type $\mathbf{1}$ and one of type $\mathbf{2}$, so for each interval of type $\mathbf{1}$ there is in F^n we get an interval of type $\mathbf{1}$ and one of type $\mathbf{2}$ in F^{n+1}.

Now assume that there is an interval I_2 of type $\mathbf{2}$ in F^n. Then

$$T^n\left(I_2 \cap T^{-n-1}\left[\frac{1}{4}, 1\right)\right) = \left[\frac{1}{2}, 1\right) \cap T^{-1}\left[\frac{1}{4}, 1\right).$$

as in the figure below. Here the blue line is $I_2 = [1/2, 1)$, the green lines are $T^{-1}[1/4, 1)$ and the red line is the intersection which is one interval J such that $TJ = \left[\frac{1}{4}, 1\right)$.



So for each interval of type **2** in F^n we get an interval of type **1** in F^{n+1}. This finishes the proof of the lemma. □

We define f_n to be the number of disjoint intervals in F^n and prove the following lemma.

Lemma 2.2. *For $n \geq 0$, f_n is the n'th Fibonacci number, that is $f_0 = 1, f_1 = 2$ and $f_n = f_{n-1} + f_{n-2}$ for $n \geq 2$.*

Proof. We see that f_n is equal to the number of digits of C_n, so it is enough to prove that the number of digits of C_n is the n'th Fibonacci number. We know that we can define C_n by $C_1 = \mathbf{1}$ and then use the morphisms

$$\mathbf{1} \to \mathbf{21}, \quad \mathbf{2} \to \mathbf{1}$$

to get from C_n to C_{n+1} for $n \geq 0$. Now let f_n^1 and f_n^2 denote the number of **1**'s and **2**'s respectively in C_n. Then we have for any $n > 1$ that

$$f_n^2 = f_{n-1}^1,$$

and

$$f_n^1 = f_{n-1}^1 + f_{n-1}^2 = f_{n-1}.$$

So

$$f_n = f_n^1 + f_n^2 = f_{n-1}^1 + f_{n-1} = f_{n-2} + f_{n-1},$$

as desired. □

Using this result we can now calculate the Hausdorff dimension of F. Here we give a heuristic argument where we assume that

$$0 < \mathscr{H}^s(F) < \infty$$

for some s and leave the rigorous calculation to the proof of the general case. We know that F^n consists of f_n intervals, namely $f_n^1 = f_{n-1}$ of type

1 which is of length $\frac{3}{2^{n+2}}$ and $f_n^2 = f_{n-2}$ of type **2** which is of length $\frac{2}{2^{n+2}}$. So we get for $s \geq 0$ that

$$\begin{aligned}
\mathscr{H}^s(F^n) &= f_{n-1}\left(\frac{3}{2^{n+2}}\right)^s + f_{n-2}\left(\frac{2}{2^{n+2}}\right)^s \\
&= (f_{n-2} + f_{n-3})\left(\frac{3}{2^{n+2}}\right)^s + (f_{n-3} + f_{n-4})\left(\frac{2}{2^{n+2}}\right)^s \\
&= \frac{1}{2^s}\left(f_{n-2}\left(\frac{3}{2^{n+1}}\right)^s + f_{n-3}\left(\frac{2}{2^{n+1}}\right)^s\right) \\
&\quad + \frac{1}{2^{2s}}\left(f_{n-3}\left(\frac{3}{2^n}\right)^s + f_{n-4}\left(\frac{2}{2^n}\right)^s\right) \\
&= \frac{1}{2^s}\mathscr{H}^s(F^{n-1}) + \frac{1}{2^{2s}}\mathscr{H}^s(F^{n-2}).
\end{aligned}$$

Letting $n \to \infty$ and assuming that $0 < \mathscr{H}^s(F) < \infty$ for some s we get

$$1 = \frac{1}{2^s} + \frac{1}{2^{2s}}.$$

Solving this in s gives

$$s = \frac{\log \varphi}{\log 2}.$$

So under the assumption that $0 < \mathscr{H}^s(F) < \infty$ for some s we get $\dim_H F = s$ in this example.

2.3 The general case

We now proceed to the general case. Here we need the Perron-Frobenius theorem. A square matrix A with non-negative entries is said to be *irreducible* if for any i, j there is n such that $(A^n)_{ij} > 0$. Recall that the *spectral radius* of a matrix is the eigenvalue that is largest in absolute value. We have the following theorem.

Theorem 2.3 (Perron-Frobenius)**.** *If A is an irreducible matrix, then the spectral radius ρ is real and $\rho > 0$.*

We are now ready to prove the following theorem which gives a simple way of calculating the value of $\phi(c)$ when $c = \frac{k}{q^m}$ for integers k, m.

Theorem 2.4. *Let $c = \sum_{i=1}^m \frac{c_i}{q^i} > 0$ with $0 \leq c_i < q$ and let*

$$n = \min\{1 \leq j \leq m \mid \mathbf{c_{j+1}} \cdots \mathbf{c_m} \leq \mathbf{c_1} \cdots \mathbf{c_{m-j}}\}$$

and

$$a_i = \begin{cases} q - c_i - 1 & \textit{for } 1 \le i < n \\ q - c_i & \textit{for } i = n \end{cases} .$$

Then

$$\dim_H F_c = \frac{\log \rho}{\log q}$$

where ρ is the largest real root of

$$x^n - a_1 x^{n-1} - a_2 x^{n-2} - \cdots - a_n.$$

Proof. First notice that since $x \in F_c$ if and only if $x \in T^{-i}[c, 1)$ for all $i = 0, 1, \ldots$ we have

$$F_c = \bigcap_{i=0}^{\infty} F_c^i$$

where $F_c^0 = [c, 1)$ and $F_c^i = F_c^{i-1} \cap T^{-i}[c, 1)$.

We claim that F_c^i can be written as a disjoint union,

$$F_c^i = \bigcup_{k=1}^{f_i} J_k \tag{2.2}$$

where $f_i \in \mathbb{N}$ and each J_k is an interval such that

$$T^i J_k = [\theta(\mathbf{c_{j_k}} \cdots \mathbf{c_m}), 1) \tag{2.3}$$

for some integer $j_k \le n$. Furthermore we claim that if we for $i \ge 0$ define $C_i \in \Sigma_n$ by

$$C_i = \mathbf{j_1} \cdots \mathbf{j_{f_i}}$$

with j_k as in (2.3), we have $C_0 = 1$ and for $i \ge 0$, we can find C_{i+1} in terms of C_i by the morphims

$$\mathbf{1} \to \mathbf{21}^{a_1}, \quad \mathbf{2} \to \mathbf{31}^{a_2}, \ \ldots \ , [\mathbf{n-1}] \to \mathbf{n1}^{a_{n-1}}, \quad \mathbf{n} \to \mathbf{1}^{a_n}. \tag{2.4}$$

We will prove this claim by induction. Now $F_c^0 = [c, 1)$ and $C_0 = 1$, so the claim is true for $i = 0$. Now assume that $i \ge 0$ and that F_c^i is a disjoint union as in (2.2) and define $j_1, \ldots, j_{f_i}$ as in (2.3). Then

$$F_c^{i+1} = \left(\bigcup_{k=1}^{f_i} J_k \right) \cap T^{-i-1}[c, 1) = \bigcup_{k=1}^{f_i} \left(J_k \cap T^{-i-1}[c, 1) \right).$$

Now let k with $1 \le k \le f_i$ be given, and let $J = J_k$ and $j = j_k$. If $j < n$ then

$$\mathbf{c_{j+1}} \cdots \mathbf{c_m} > \mathbf{c_1} \cdots \mathbf{c_{m-j}}$$

and so

$$\theta(\mathbf{c_j}\cdots\mathbf{c_m}) \in T^{-1}[c,1)$$

because $T\theta(\mathbf{c_j}\cdots\mathbf{c_m}) = \theta(\mathbf{c_{j+1}}\cdots\mathbf{c_m})$. We see that

$$T^i\big(J\cap T^{-i-1}[c,1)\big) = [\theta(\mathbf{c_j}\cdots\mathbf{c_m}),1)\cap T^{-1}[c,1)$$

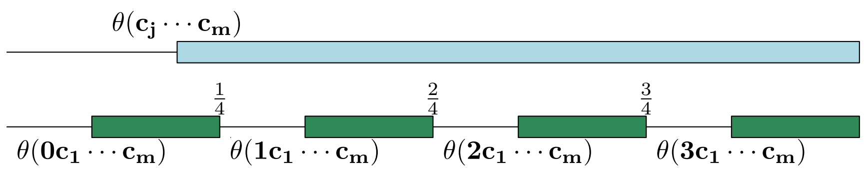
consists of the interval

$$\left[\theta(\mathbf{c_j}\cdots\mathbf{c_m}), \frac{c_j+1}{q}\right).$$

and if $c_j < q-1$ then furthermore of the $q-c_j-1 = a_j$ intervals

$$[\theta([\mathbf{q-1}]\mathbf{c_1}\cdots\mathbf{c_m}),1), \left[\theta([\mathbf{q-2}]\mathbf{c_1}\cdots\mathbf{c_m}), \frac{q-1}{q}\right), \ldots,$$
$$\left[\theta([\mathbf{c_j+1}]\mathbf{c_1}\cdots\mathbf{c_m}), \frac{c_j+2}{q}\right).$$

See the figure below to see an illustration of this where $q=4$. Here the blue line is $[\theta(\mathbf{c_j}\cdots\mathbf{c_m}),1)$ and the green lines are $T^{-1}[c,1)$.



Now

$$T\left[\theta(\mathbf{c_j}\cdots\mathbf{c_m}), \frac{c_j+1}{q}\right) = [\theta(\mathbf{c_{j+1}}\cdots\mathbf{c_m}),1)$$

and

$$T\left[\theta([\mathbf{q}-\ell]\mathbf{c_1}\cdots\mathbf{c_m}), \frac{q-\ell+1}{q}\right) = [c,1)$$

for $\ell = 1,2,\ldots,a_j$ so the intersection $J\cap T^{-i-1}[c,1)$ gives an interval I such that $T^{i+1}I = [\theta(\mathbf{c_{j+1}}\cdots\mathbf{c_m}),1)$ and a_j intervals K such that $T^{i+1}K = [c,1)$. So for each j in C_i we get $[\mathbf{j+1}]\mathbf{1_j^a}$ in C_{i+1} so the morphisms given in (2.4) are true when $j<n$.

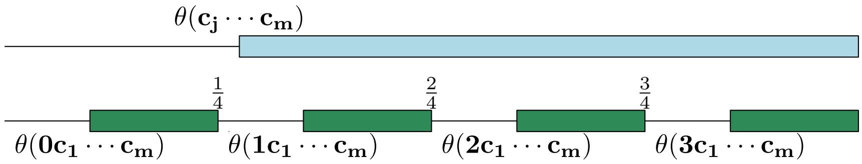
Assume now that $j=n$ then

$$\mathbf{c_{j+1}}\cdots\mathbf{c_m} \le \mathbf{c_1}\cdots\mathbf{c_{m-j}}$$

and so

$$\theta(\mathbf{c_j}\cdots\mathbf{c_m}) \notin T^{-1}[c,1).$$

as in the illustration below where the green lines are $T^{-1}[c,1)$.



Now $[\theta(\mathbf{c_j}\cdots\mathbf{c_m}),1)] \cap T^{-1}[c,1)$ consists only of $q-c_j = a_j$ intervals

$$[\theta([\mathbf{q-1}]\mathbf{c_1}\cdots\mathbf{c_m}),1), \left[\theta([\mathbf{q-2}]\mathbf{c_1}\cdots\mathbf{c_m}), \frac{q-1}{q}\right), \ldots,$$
$$\left[\theta(\mathbf{c_j}\mathbf{c_1}\cdots\mathbf{c_m}), \frac{c_j+1}{q}\right).$$

For each of these we have

$$T\left[\theta([\mathbf{q}-\ell]\mathbf{c_1}\cdots\mathbf{c_m}), \frac{q-\ell+1}{q}\right) = [c,1)$$

for $\ell = 1,2,\ldots,a_j$ so the intersection $J \cap T^{-i-1}[c,1)$ consists of a_j intervals K such that $T^{i+1}K = [c,1)$. So for each $\mathbf{j} = \mathbf{n}$ in C_i we get $\mathbf{1^{a_n}}$ in C_{i+1}. This proves that the morphisms in (2.4) are true and finishes the proof of the claim.

Let $g_i^j(k)$ denote the number of j's in a word after i iteratations, where we start with the word just consisting of the digit $\mathbf{k}$. Then we get from the morphisms that

$$\begin{aligned}
g_i^1(k) &= a_1 g_{i-1}^2(k) + a_2 g_{i-1}^2(k) + \cdots + a_n g_{i-1}^n(k)\\
g_i^2(k) &= g_{i-1}^1(k)\\
g_i^3(k) &= g_{i-1}^2(k)\\
&\vdots\\
g_i^n(k) &= g_{-1}i^{n-1}(k),
\end{aligned}$$

so if we let

$$\overline{g}_i(k) = \begin{pmatrix} g_i^1(k)\\ g_i^2(k)\\ \vdots\\ g_i^n(k)\end{pmatrix} \tag{2.5}$$

we have

$$\overline{g}_i(k) = M\overline{g}_{i-1}(k)$$

when $i \geq 1$ and M is the $n \times n$ matrix given by

$$M = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & a_n \\ 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 1 & \end{pmatrix}$$

where the empty entries are zeroes. We now wish to prove that the characteristic polynomial of M is

$$\det(xI - M) = x^n - a_1 x^{n-1} - a_2 x^{n-2} - \cdots - a_n. \tag{2.6}$$

By expanding along the first column we get

$$\begin{aligned} \det(xI - M) &= \det \begin{pmatrix} x - a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} & -a_n \\ -1 & x & & & & \\ & -1 & x & & & \\ & & \ddots & \ddots & & \\ & & & -1 & x & \\ & & & & -1 & x \end{pmatrix} \\ &= x^n - a_1 x^{n-1} + \det \begin{pmatrix} -a_2 & -a_3 & \cdots & -a_{n-2} & -a_{n-1} & -a_n \\ -1 & x & & & & \\ & -1 & x & & & \\ & & \ddots & \ddots & & \\ & & & -1 & x & \\ & & & & -1 & x \end{pmatrix} \end{aligned}$$

which is equal to

$$x^n - a_1x^{n-1} - a_2x^{n-2} + \det\begin{pmatrix} -a_3 & -a_4 & \cdots & -a_{n-2} & -a_{n-1} & -a_n \\ -1 & x & & & & \\ & -1 & x & & & \\ & & \ddots & \ddots & & \\ & & & -1 & x & \\ & & & & -1 & x \end{pmatrix}$$

$$= x^n - a_1x^{n-1} - a_2x^{n-2} - a_3x^{n-3} + \det\begin{pmatrix} -a_4 & -a_5 & \cdots & -a_{n-2} & -a_{n-1} & -a_n \\ -1 & x & & & & \\ & -1 & x & & & \\ & & \ddots & \ddots & & \\ & & & -1 & x & \\ & & & & -1 & x \end{pmatrix}$$

and continuing like this we get (2.6). Recall that the Cayley-Hamilton theorem states that a square matrix satisfies its own characteristic equation. This gives us

$$M^n = a_1M^{n-1} + a_2M^{n-2} + \cdots + a_n,$$

and hence by (2.5) that

$$\overline{g}_i(k) = a_1\overline{g}_{i-1}(k) + a_2\overline{g}_{i-2}(k) + \cdots + a_n\overline{g}_{i-n}(k)$$

for $i > n$. This gives us

$$g_i^j(k) = a_1g_{i-1}^j(k) + a_2g_{i-2}^j(k) + \cdots + a_ng_{i-n}^j(k)$$

for $j = 1, 2, \ldots, n$. Now let $g_i(k)$ denote the total number of digits after i iterations. Then $g_i(k)$ is the sum of the entries in $\overline{g}_i$, and we that

$$g_i(k) = a_1g_{i-1}(k) + a_2g_{i-2}(k) + \cdots + a_ng_{i-n}(k).$$

Recall that $a_n = q - c_n \geq 1$, so by entrywise comparsion we have

$$M \geq P$$

where P is the permutation matrix given by

$$P = \begin{pmatrix} & & & & 1 \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \end{pmatrix}.$$

We see that for each i, j there is n such that the (i, j)'th entry of P^n is strictly positive. Since $M \geq P$ this is also true for M, so M is irreducible, and by the Perron-Frobenius Theorem we get that the spectral radius ρ of M is real and $\rho > 0$.

We now claim that $g_i(k) \asymp \rho^i$ for $i \geq n$. This can be proved by induction. Assume that there are constants $K_1, K_2 > 0$ such that

$$K_1\rho^j \leq f_j \leq K_2\rho^j$$

for all k and all $j < i$. Now

$$\begin{aligned} f_i &= a_1 g_{i-1}(k) + \cdots + a_n g_{i-n}(k) \\ &\leq K_2(a_1\rho^{i-1} + \cdots + a_n\rho^{i-n}) \\ &= K_2\rho^{i-n}(a_1\rho^{n-1} + \cdots + a_n) \\ &= K_2\rho^i \end{aligned}$$

since $\rho^n = a_1\rho^{n-1} + a_2\rho^{n-2} + \cdots + a_n$. The lower bound is simlar, so we have now proved the claim.

We now wish to prove that

$$\dim_H F_c = \frac{\log \rho}{\log q}.$$

We consider the slightly larger set

$$\hat{F}_c = \bigcap_{i=0}^{\infty} \hat{F}_c^i,$$

where $\hat{F}_c^i$ is the union of F_c^i and the right end-points of the half-open disjoint intervals that is in F_c^i. This ensures that $\hat{F}_c$ is a compact set and that

$$\dim_H F_c = \dim_H \hat{F}_c$$

since $\hat{F}_c$ is the union of F_c and a countable set of points.

First we note that $\hat{F}_c^i$ consists of f_i closed intervals. Now $T^{-i}[c, 1]$ is the disjoint union of q^i intervals of length $(1 - c)q^{-i} < q^{-i}$ and spaced $cq^{-i} < q^{-i-m}$. So because $\hat{F}_c^i \subseteq T^{-i}[c, 1]$, the f_i intervals in f_i are also at most q^{-i} long and spaced at least q^{-i-m}. Recalling the definition of the s-dimensional Hausdorff measure we get

$$\mathscr{H}^s_{q^{-i}}(\hat{F}_c) \leq f_i q^{is} = f_i\rho^{-i} \leq K_2 < \infty$$

where K_2 is as above and

$$s = \frac{\log \rho}{\log q},$$

so letting $i \to \infty$ we get $\mathscr{H}^s(\hat{F}_c) \le K_2$. To prove a lower bound we consider any cover of $\hat{F}_c$. By the compactness it is enough to consider finite covers of closed intervals, so we assume that $\{U_h\}$ is such a cover. Let h be given. We can pick j such that

$$q^{-j-1} \le \operatorname{diam}(U_h) < q^{-j}.$$

Since the intervals in $\hat{F}_c^{j+m}$ are spaced at least q^{-j-m} we see that U_h intersects at most one of these intervals. Let $i \ge j+m$. Since any digit will give at most $K_2\rho^{i-j-m}$ digits after $i-j-m$ iterations, any interval in $\hat{F}_c^{j+m}$ and hence U_h intersects at most $K_2\rho^{i-j-m}$ intervals in $\hat{F}_c^i$. Now

$$K_2\rho^{i-j-m} \ll f_i\rho^{-j-m} = f_i q^{(-j-m)s} \le f_i q^s \operatorname{diam}(U_h)^s \tag{2.7}$$

Now pick j large enough to ensure that $q^{-j-1} \le \operatorname{diam}(U_h)$ for all h and let $i \ge j+m$. Since $\{U_h\}$ is a cover of $\hat{F}_c$ it intersects all f_i intervals from $\hat{F}_c^i$. By counting intervals and applying (2.7) we get

$$\sum_h f_i q^s \operatorname{diam}(U_h)^s \gg f_i$$

and hence

$$\sum_h \operatorname{diam}(U_h)^s \gg q^{-s}$$

for any cover $\{U_i\}$. This proves that $\mathscr{H}^s(\hat{F}_c) \gg q^{-s} = \rho^{-1} > 0$ and hence that

$$\dim_H \hat{F}_c = \frac{\log \rho}{\log q}$$

as desired. □

2.4 Asymptotics

We now wish to give show result on the asymptotic behavoir of $\phi_q(c)$ when $q \to \infty$. Let

$$\psi(c) = \psi_q(c) = \begin{cases} 1 + \frac{\log(1-c)}{\log q} & 0 \le c < \frac{q-1}{q} \\ 0 & \text{otherwise.} \end{cases}$$

See figure 2.3 for a plot of ϕ and ψ when $q = 6$. We now prove the following lemma.

Lemma 2.5. *For all $c \in [0,1)$ we have*

$$|\phi_q(c) - \psi_q(c)| \to 0 \text{ as } q \to \infty.$$

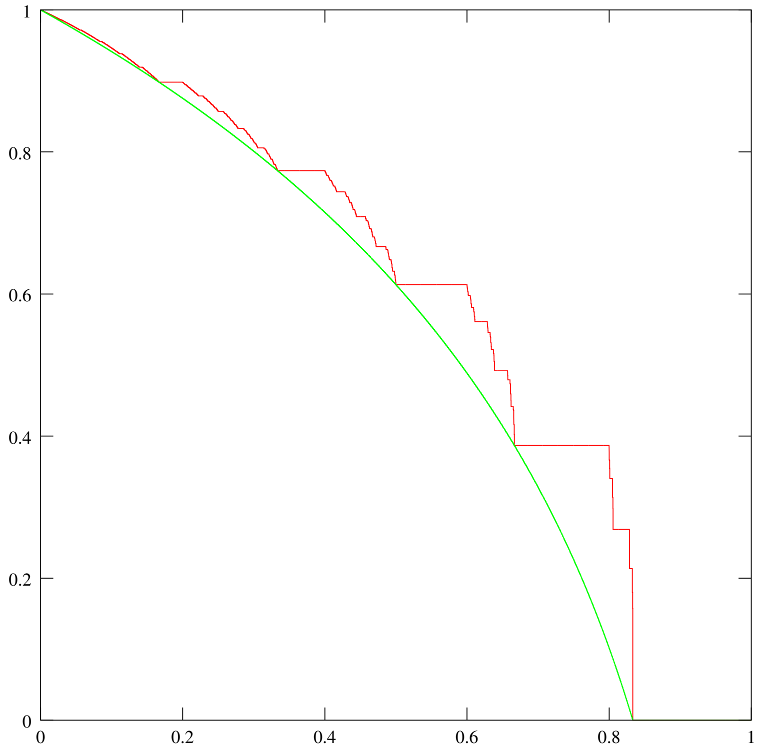


Figure 2.3: *Plots of ϕ_6 (red) and ψ_6 (green).*

Proof. Let $c \in [0,1)$ be given. Then if we let $i = \lfloor qc \rfloor$ we have

$$\frac{i}{q} \leq c \leq \frac{i+1}{q}.$$

Now

$$\phi_q\left(\frac{i}{q}\right) \geq \phi_q(c) \geq \phi_q\left(\frac{i+1}{q}\right)$$

and likewise for ψ since both functions are decreasing. Now from Theorem 2.4 we get $\psi_q(\frac{i}{q}) = \phi_q(\frac{i}{q}) = \frac{\log(q-i)}{\log q}$, so

$$\frac{\log(q-i)}{\log q} \geq \phi_q(c), \psi_q(c) \geq \frac{\log(q+1-i)}{\log q}$$

and recalling the definition of i we have

$$|\phi_q(c) - \psi_q(c)| \leq \frac{\log(q-i+1) - \log(q-i)}{\log q} \leq \frac{\log 2}{\log q} \to 0$$

as $q \to \infty$. $\square$

Since ϕ and ψ are asymptotically similar, and $\psi_q \to 1$ as $q \to \infty$, we get the following theorem.

Theorem 2.6. *For all $c \in [0,1)$ we have*

$$\phi_q(c) \to 1 \text{ as } q \to \infty.$$

Proof. Let $c \in [0,1)$. We now consider q large enough to ensure that

$$\frac{q-1}{q} > c.$$

For these q we have

$$\psi_q(c) = 1 + \frac{\log(1-c)}{\log q}$$

and hence that $\psi_q(c) \to 1$ as $q \to \infty$. From Lemma 2.5 we now get that

$$\phi_q(c) \to 1 \text{ as } q \to \infty$$

as desired. □

In figure 2.4 we see a plot of ϕ when $q = 500,000$ to illustrate the convergence.

2.5 The dimension is constant

We wish to prove that the map ϕ_q is constant almost everywhere. We follow the proof of Nilsson [17] with some minor alterations.

Define

$$l(\mathbf{c_1} \cdots \mathbf{c_m}) = \min\{1 \le j \le m \mid \mathbf{c_{j+1}} \cdots \mathbf{c_m} \le \mathbf{c_1} \cdots \mathbf{c_{m-j}}\}.$$

We now prove the following lemma.

Lemma 2.7 (Nilsson)**.** *Let $c = \theta(\mathbf{c_1}\mathbf{c_2}\cdots) \in [0,1)$ for some $\mathbf{c_1}\mathbf{c_2}\cdots \in \Sigma$. If there is a maximal $m \in \mathbb{N}$ such that*

$$l(\mathbf{c_1} \cdots \mathbf{c_m})) = m$$

then ϕ is constant on the interval $J = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta((\mathbf{c_1} \cdots \mathbf{c_m})^\infty]$ and $c \in J$.

Proof. Since $l(\mathbf{c_1} \cdots \mathbf{c_m}) = m$ we get that for any $n \ge 1$ we have $l((\mathbf{c_1} \cdots \mathbf{c_m})^n) = m$, because $l((\mathbf{c_1} \cdots \mathbf{c_m})^n) \le m$ and for any $j < m$ we have

$$\mathbf{c_{j+1}} \cdots \mathbf{c_m}(\mathbf{c_1} \cdots \mathbf{c_m})^{n-1} > (\mathbf{c_1} \cdots \mathbf{c_m})^{n-1}\mathbf{c_1} \cdots \mathbf{c_{m-j}}.$$

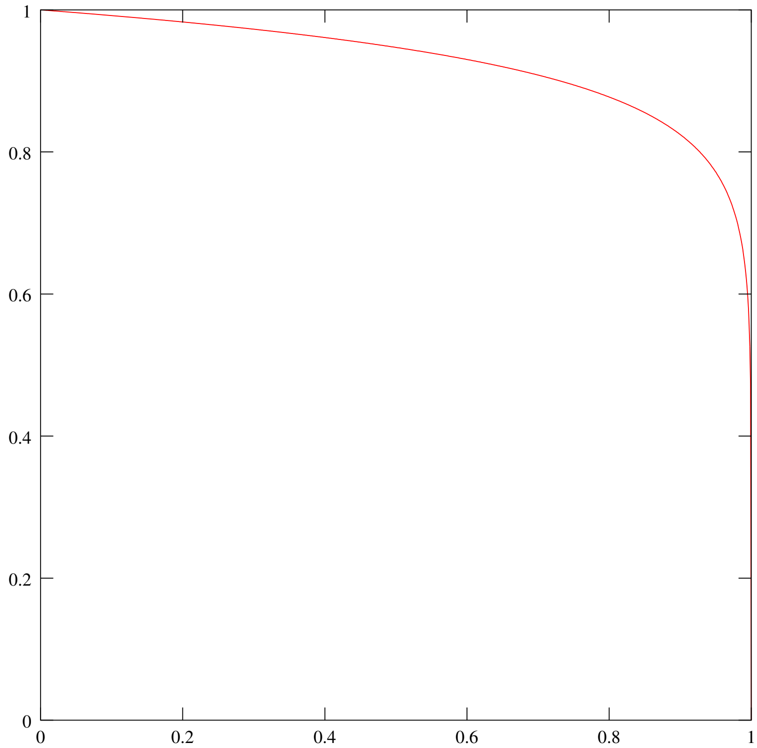


Figure 2.4: *Plot of* $\phi_{500{,}000}$.

So when $t = \theta((\mathbf{c_1} \cdots \mathbf{c_m})^n)$ for some $n \geq 1$, we get from Theorem 2.4 that

$$\phi(t) = \frac{\log \rho}{\log q}$$

where ρ is the largest real root of $x^m - a_1 x^{m-1} - \cdots - a_m$ where $a_1, \ldots, a_m$ is as in Theorem 2.4. This is independent of n, so ρ is the same for all n. This proves that ϕ is constant on J.

Now we wish to prove that $c \in J$. To prove this it is enough to show that $\mathbf{c_{m+1}} \mathbf{c_{m+2}} \cdots \leq (\mathbf{c_1} \cdots \mathbf{c_m})^\infty$, so we assume for contradiction that this is not the case. For any integer $k \in \mathbb{N}$ we let $k' \in \{1, 2, \ldots, m\}$ be the unique number such that $k' \equiv k \pmod{m}$. Since we have assumed $\mathbf{c_{m+1}} \mathbf{c_{m+2}} \cdots > (\mathbf{c_1} \cdots \mathbf{c_m})^\infty$ there is $n > m$ such that

$$\mathbf{c_{m+1}} \cdots \mathbf{c_{n-1}} = (\mathbf{c_1} \cdots \mathbf{c_m})^u \mathbf{c_1} \cdots \mathbf{c_{n'-1}}$$

for some integer $u \geq 0$ and $\mathbf{c_n} > \mathbf{c_{n'}}$. We now want to prove that this implies $l(\mathbf{c_1} \cdots \mathbf{c_n}) = n$ which will give us the desired contradiction.

Assume on the contrary that $l(\mathbf{c_1}\cdots\mathbf{c_n}) = j < n$, then

$$\mathbf{c_1}\cdots\mathbf{c_{n-j}} \geq \mathbf{c_{j+1}}\cdots\mathbf{c_n}. \tag{2.8}$$

The left hand side of (2.8) is

$$(\mathbf{c_1}\cdots\mathbf{c_m})^s\mathbf{c_1}\cdots\mathbf{c_{(n-j)'}}$$

for some integer $s \geq 0$ and the right hand side is

$$\mathbf{c_{j'+1}}\cdots\mathbf{c_m}(\mathbf{c_1}\cdots\mathbf{c_m})^t\mathbf{c_1}\cdots\mathbf{c_{n'-1}}\mathbf{c_n}$$

for some integer $t \geq 0$. If $j \not\equiv 0 \pmod m$ then considering just the first $m-j$ digits we get

$$\mathbf{c_1}\cdots\mathbf{c_{m-j'}} \geq \mathbf{c_{j'+1}}\cdots\mathbf{c_m}$$

so $l(\mathbf{c_1}\cdots\mathbf{c_m}) \leq j' < m$ which is a contradiction. If $j \equiv 0 \pmod m$ then (2.8) reduces to

$$(\mathbf{c_1}\cdots\mathbf{c_m})^s\mathbf{c_1}\cdots\mathbf{c_{n'}} \geq (\mathbf{c_1}\cdots\mathbf{c_m})^{t-1}\mathbf{c_1}\cdots\mathbf{c_{n'-1}}\mathbf{c_n},$$

but here all but the last digit are equal, so we must have $\mathbf{c_{n'}} \geq \mathbf{c_n}$ which is a contradiction. □

Using this lemma we can now prove that ϕ is constant almost everywhere.

Theorem 2.8 (Nilsson)**.** *The map ϕ is constant almost everywhere with respect to the Lebesgue measure.*

Proof. We assume that $c > 0$. From Lemma 2.7 we see that it is enough to prove that the set of points

$$c = \sum_{i=1}^{\infty} \frac{c_i}{q^i}$$

for which there are infinitely many $m \in \mathbb{N}$ such that $l(\mathbf{c_1}\cdots\mathbf{c_m}) = m$ is a nullset. Given such an c we see that

$$\mathbf{c_1}\cdots\mathbf{c_{m-n}} < \mathbf{c_{n+1}}\cdots\mathbf{c_m}$$

for infinitely many $m \in \mathbb{N}$ and all $n < m$. Now

$$\{q^n\theta(\mathbf{c_1}\mathbf{c_2}\cdots)\} = \theta(\mathbf{c_{n+1}}\mathbf{c_{n+2}}\cdots)$$

so in particular this gives us

$$c < \{q^n c\} = T^n c$$

for all $n \geq 1$. So to finish the proof we need to prove that

$$L = \{c \in [0,1) \mid T^n c > c \text{ for all } n \geq 1\}$$

is a nullset. Note that

$$L \subseteq \bigcup_{m=1}^{\infty} L_m$$

where

$$L_m = \left\{c \in [0,1) \mid T^n c \geq \frac{1}{q^m} \text{ for all } n \geq 1\right\},$$

so it is enough to prove that L_m is a nullset for all m. Now

$$L_m = \bigcap_{n=1}^{\infty} T^{-n}[q^{-m}, 1) \subseteq F_{q^{-m}}.$$

By Theorem 2.4 we get that since $q^m = \theta(\mathbf{0^{m-1}1})$ we have

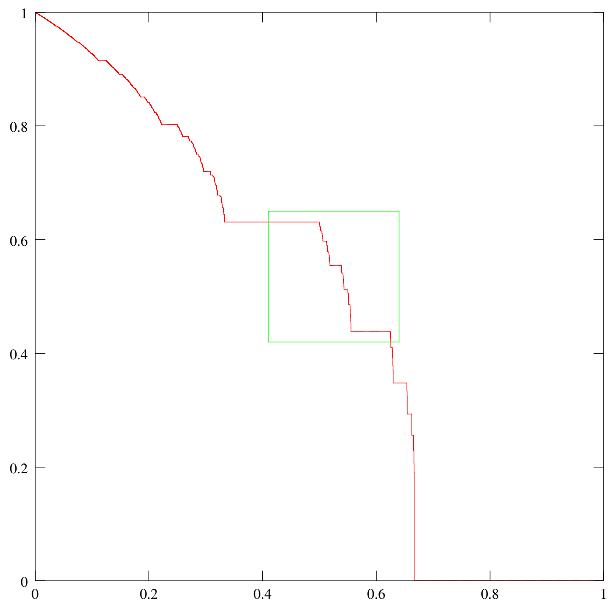
$$\dim_H F_{q^{-m}} = \frac{\log \rho}{\log q}$$

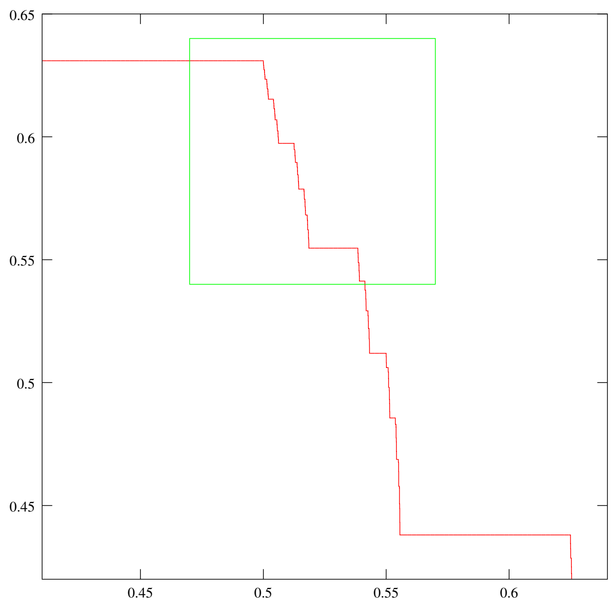
where ρ is the largest real root of $x^m - (q-1)x^{m-1} - \cdots - (q-1)$. But q is not a root of this and $\rho \leq q$, so $\rho < q$ and hence $\dim_H F_{q^{-m}} < 1$. In particular $F_{q^{-m}}$ is a nullset with respect to the Lebesgue measure, so L_m is also a nullset. This finishes the proof of the theorem. □

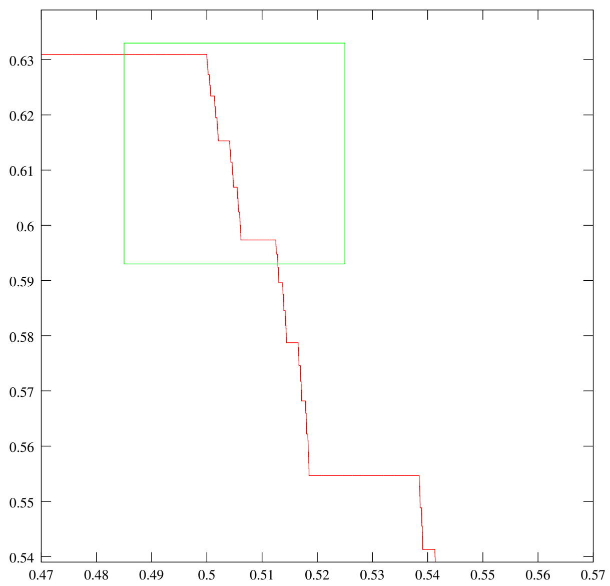
2.6 Fractal plots

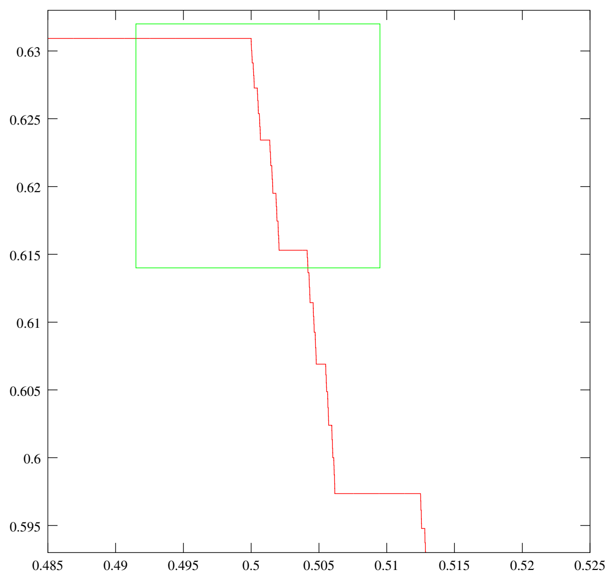
In the six plots below we zoom in on the graph of ϕ_3 – notice the self-similarity and the fractal structure of the graph. The green square on the first five plots indicates where we zoom in on the next plot.

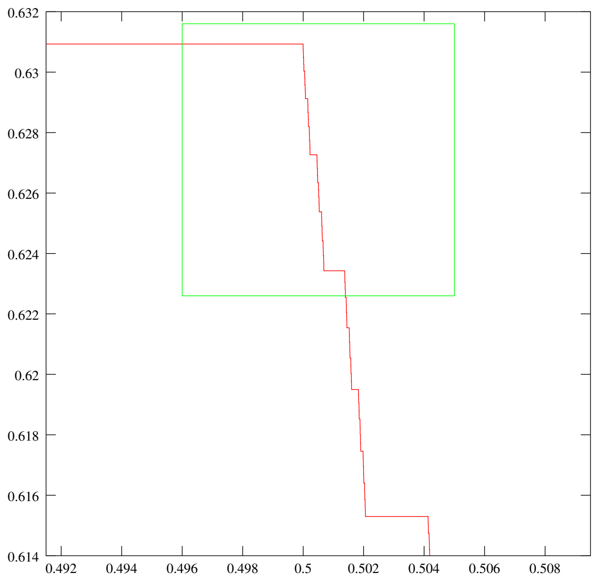
The Octave code used to generate the data for these plots as well as for the other plots in this chapter can be found in Appendix B.

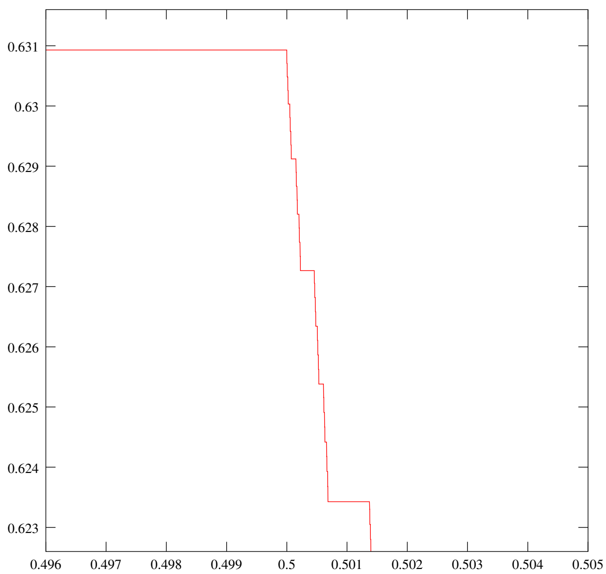

1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1


0.65
0.6
0.55
0.5
0.45
0.45
0.5
0.55
0.6


0.63
0.62
0.61
0.6
0.59
0.58
0.57
0.56
0.55
0.54
0.47
0.48
0.49
0.5
0.51
0.52
0.53
0.54
0.55
0.56
0.57


0.63
0.625
0.62
0.615
0.61
0.605
0.6
0.595
0.485
0.49
0.495
0.5
0.505
0.51
0.515
0.52
0.525


0.632
0.63
0.628
0.626
0.624
0.622
0.62
0.618
0.616
0.614
0.492
0.494
0.496
0.498
0.5
0.502
0.504
0.506
0.508


0.631
0.63
0.629
0.628
0.627
0.626
0.625
0.624
0.623
0.496
0.497
0.498
0.499
0.5
0.501
0.502
0.503
0.504
0.505

Chapter 3

Beta-shifts

3.1 Introduction

The theory of β-expansions goes back to Parry [18] and Rényi [22]. It is a generalization of base q-expansions of real numbers where we allow a non-integer base $\beta > 1$. The first result in this chapter is that we will describe a connection between Theorem 2.4, and the theory of β-expansions, and we will use this connection to give a new proof that ϕ as defined in the preceding chapter is continuous. Now we define

$$F_c(\beta) = \{x \in [0,1) \mid \{\beta^n x\} \geq c \text{ for all } n \geq 0\}$$

for $\beta > 1$. We will generalize Theorem 2.4 in order to calculate $\phi(c) = \phi_\beta(c) = \dim_H F_c(\beta)$.

First we define β-expansions and give some general results. Let $\beta > 0$ and define $T_\beta : [0,1) \to [0,1)$ by $T = T_\beta x = \{\beta x\}$. We now define a map $\gamma_\beta : [0,1] \to \Sigma$ by $\gamma_\beta(x) = \mathbf{a_1 a_2} \cdots$ where $a_1 = \lfloor \{\beta x\} \rfloor$ and

$$a_i = \lfloor \beta T_\beta^{i-1} x \rfloor$$

for $i = 2, 3, \ldots$. Note that $a_i \in \{0, 1, \ldots, \lceil \beta \rceil - 1\}$ for $i \geq 2$. If we let

$$\theta_\beta(\mathbf{a_1 a_2} \cdots) = \sum_{i=1}^{\infty} \frac{a_i}{\beta^i}$$

then

$$x = \theta_\beta(\gamma_\beta(x)).$$

The map T works as a shift operator since

$$T\theta(\mathbf{a_1 a_2} \cdots) = \theta(\mathbf{a_2 a_3} \cdots).$$

We call $\mathbf{a_1 a_2} \cdots$ the *β-expansion* of x. If the β-expansion of 1 is finite, we call β a *simple number*. If

$$\gamma_\beta(1) = \mathbf{a_1 a_2} \cdots \mathbf{a_n}$$

then

$$1 = \frac{a_1}{\beta} + \frac{a_2}{\beta^2} + \cdots + \frac{a_n}{\beta^n}$$

and hence

$$\beta^n = a_1\beta^{n-1} + a_2\beta^{n-2} + \cdots + a_n$$

so β is a solution of $0 = x^n - a_1x^{n-1} - \cdots - a_n$. We call this the *characteristic equation* for β.

The map θ preserves order, that is

$$\theta(\mathbf{a_1 a_2} \cdots) > \theta(\mathbf{b_1 b_2} \cdots) \iff \mathbf{a_1 a_2} \cdots > \mathbf{b_1 b_2} \cdots$$

where we when comparing words use the lexicographical ordering. This was proved by Parry [18]. He also proved the following two theorems.

Theorem 3.1 (Parry). *There is a $\beta > 1$ with $\gamma_\beta(1) = \mathbf{a_1} \cdots \mathbf{a_n 0} \cdots$ if and only if*

$$\mathbf{a_1} \cdots \mathbf{a_n} > \mathbf{a_{i+1}} \cdots \mathbf{a_n 0}^i$$

for all $0 \leq i < n$.

Theorem 3.2 (Parry). *The simple numbers are dense in $(1, \infty)$.*

3.2 Continuity of the dimension

Recall that for a finite word $\mathbf{c_1} \cdots \mathbf{c_n} \in \Sigma$ we let

$$l(\mathbf{c_1} \cdots \mathbf{c_n}) = \min\{1 \leq j \leq n \mid \mathbf{c_{j+1}} \cdots \mathbf{c_n} \leq \mathbf{c_1} \cdots \mathbf{c_{n-j}}\}.$$

We now prove the following lemma.

Lemma 3.3. *Let $\mathbf{c_1} \cdots \mathbf{c_n}, \mathbf{a_1} \cdots \mathbf{a_n} \in \Sigma$ and assume*

$$a_i = \begin{cases} q - c_i - 1 & \text{if } 1 \leq i < n \\ q - c_n & \text{if } i = n. \end{cases}$$

Then the following are equivalent.

(i) $\dim_H F_c = \frac{\log \beta}{\log q}$ *where $c = \theta(\mathbf{c_1} \cdots \mathbf{c_n})$ and $l(\mathbf{c_1} \cdots \mathbf{c_n}) = n$.*

(ii) β is simple with characteristic polynomial $x^n - a_1x^{n-1} - \cdots - a_n$.

Proof. Assume *(i)* is true. By Theorem 2.4 we see that β is the largest real root of $x^n - a_1x^{n-1} - \cdots - a_n$ so by Theorem 3.1 it is enough to prove that

$$\mathbf{a_1 \cdots a_n > a_{i+1} \cdots a_n 0}^i$$

for all $i < n$. Since $l(\mathbf{c_1 \cdots c_n}) = n$ we know that

$$\mathbf{c_1 \cdots c_{n-i} < c_{i+1} \cdots c_n}$$

for $i < n$. Let $i < n$ be given. If $\mathbf{c_1 \cdots c_{n-i-1} < c_{i+1} \cdots c_{n-1}}$ then

$$\mathbf{a_1 \cdots a_{n-i-1} > a_{i+1} \cdots a_{n-1}}$$

and we are done. If not then $\mathbf{c_1 \cdots c_{n-i-1} = c_{i+1} \cdots c_{n-1}}$ and $c_{n-i} < c_n$. In this case

$$\mathbf{a_1 \cdots a_{n-i-1} > a_{i+1} \cdots a_{n-1}}$$

and

$$a_{n-i} = q - c_{n-i} - 1 > q - c_n - 1 > q - c_n = a_n$$

and we are done.

Now assume that *(ii)* is true. Then

$$\mathbf{a_1 \cdots a_n > a_{i+1} \cdots a_n 0}^i$$

for all $i < n$. By an argument identical to that above we can prove that

$$\mathbf{c_1 \cdots c_{n-i} < c_{i+1} \cdots c_n}$$

for $i < n$ and hence that $l(\mathbf{c_1 \cdots c_n}) = n$. Using Theorem 2.4 now finishes the proof of the lemma. □

Combining this lemma with Theorem 3.2 we get the following result.

Theorem 3.4. *The map ϕ is continuous.*

Proof. Since ϕ is non-increasing it is enough to prove that the image of the map is dense, but this follows from Lemma 3.3 and Theorem 3.2. □

3.3 A real example

We now wish to study the generalization of the problem considered in Chapter 2, where we instead of an integer $q \geq 2$ consider a simple number $\beta > 1$ and consider the set

$$F_c(\beta) = \{x \in [0, 1) \mid \{\beta^n x\} \geq c \text{ for all } n \geq 0\}.$$

We begin with an example where $\beta \approx 2.83118$ has characteristic equation

$$0 = x^3 - 2x^2 - 2x - 1$$

and $c = \theta(\mathbf{02}) = \frac{2}{\beta^2} \approx 0.12476$. Note that $1 = \theta(\mathbf{221})$. Now

$$F = F_{\frac{2}{\beta^2}}(\beta) = \bigcap_{i=0}^{\infty} F^i$$

where $F^0(\beta) = [c, 1)$ and $F^i = F^{i-1} \cap T^{-i}[c, 1)$. For each $i \geq 0$ we now define four classes of intervals,

$$\begin{aligned}
I_1^i &= \{[a, b) \subseteq [0, 1) : T^i[a, b) = [\theta(\mathbf{02}), 1)\} \\
I_2^i &= \{[a, b) \subseteq [0, 1) : T^i[a, b) = [\theta(\mathbf{2}), 1)\} \\
K_1^i &= \{[a, b) \subseteq [0, 1) : T^i[a, b) = [\theta(\mathbf{02}), \theta(\mathbf{21}))\} \\
K_2^i &= \{[a, b) \subseteq [0, 1) : T^i[a, b) = [\theta(\mathbf{02}), \theta(\mathbf{1}))\}
\end{aligned}$$

We now claim that F^i is a disjoint union of intervals

$$F^i = \bigcup_{h=1}^{f_i} J_h$$

where $f_i \in \mathbb{N}$ and for each h, the interval J_h is in one of the four classes of intervals. We will prove this by induction. When $i = 0$ we see that

$$F^0 = [c, 1) = [\theta(\mathbf{02}), 1) \in I_1^0.$$

Now assume that it is true for some $i \geq 0$. Then we see that

$$F^{i+1} = F^i \cap T^{-i-1}[c, 1) = \bigcup_{h=1}^{f_i} (J_h \cap T^{-i-1}[c, 1))$$

for some $f_i \in \mathbb{N}$. So it is enough to prove that for each h, the intersection $J_h \cap T^{-i-1}[c, 1)$ is a union of intervals that each is in one of $I_1^{i+1}, I_2^{i+1}, K_1^{i+1}, K_2^{i+1}$. Let h with $1 \leq h \leq f_i$ be given and let $J = J_h$. First we assume that $J \in I_1^i$, so

$$T^i J = [\theta(\mathbf{02}, 1)).$$

Now

$$T^i(J \cap T^{-i-1}[c, 1)) = [\theta(\mathbf{02}, 1)) \cap T^{-1}[c, 1). \tag{3.1}$$

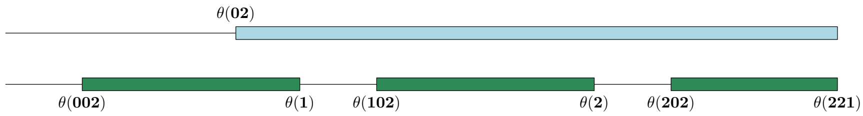
We see that

$$T^{-1}[c, 1) = [\theta(\mathbf{002}), \theta(\mathbf{1})) \cup [\theta(\mathbf{102}), \theta(\mathbf{2})) \cup [\theta(\mathbf{202}), \theta(\mathbf{221}))$$

From this we see that $[\theta(\mathbf{02}), 1) \cap T^{-1}[c, 1)$ consists of three intervals

$$\begin{aligned}
&[\theta(\mathbf{02}, 1)) \cap [\theta(\mathbf{002}), \theta(\mathbf{1})) = [\theta(\mathbf{02}), \theta(\mathbf{1})) \\
&[\theta(\mathbf{02}, 1)) \cap [\theta(\mathbf{102}), \theta(\mathbf{2})) = [\theta(\mathbf{102}), \theta(\mathbf{2})) \\
&[\theta(\mathbf{02}, 1)) \cap [\theta(\mathbf{202}), 1) = [\theta(\mathbf{202}), \theta(\mathbf{221})),
\end{aligned}$$

see the figure below for an illustration. Her the blue line is $[\theta(\mathbf{02}, 1]$ and the green lines are $T^{-1}[c, 1)$.



Note that we were interested in counting the intervals of $J \cap T^{-i-1}[c, 1)$ and instead found the intervals in $T^i(J \cap T^{-i-1}[c, 1))$. This is not a problem, because there is actually the same number of intervals in the two sets. To prove this, it is enough to prove that if $x, y \in J$ with $x < y$ then $T^i x < T^i y$. This implies in particular that $T^i : J \to [0, 1)$ is injective and that we do not have any 'wrap around' issues. Now assume for contradiction that $x, y \in J$ with $x < y$, but $T^i x > T^i y$. Let $\mathbf{x_1 x_2} \cdots$ and $\mathbf{y_1 y_2} \cdots$ denote the β-expansions of x and y respectively. Then we have

$$\mathbf{x_1 x_2} \cdots < \mathbf{y_1 y_2} \cdots$$

but

$$\mathbf{x_{i+1} x_{i+2}} \cdots > \mathbf{y_{i+1} y_{i+2}} \cdots$$

so

$$\mathbf{x_1 x_2} \cdots \mathbf{x_i} < \mathbf{y_1 y_2} \cdots \mathbf{y_i}.$$

Now let $z = \theta(\mathbf{y_1 y_2} \cdots \mathbf{y_i})$. Then $x < z < y$ so we have a number $z \in J$ with $T^i z = 0$. From the definition of the four types of intervals J can be in, we see that $z \notin J$, so this gives the desired contradiction.

We now see that

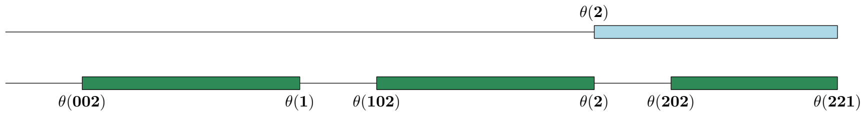
$$\begin{aligned}
T[\theta(\mathbf{02}), \theta(\mathbf{1})) &= [\theta(\mathbf{2}), 1) \\
T[\theta(\mathbf{102}), \theta(\mathbf{2})) &= [\theta(\mathbf{02}), 1) \\
T[\theta(\mathbf{202}), \theta(\mathbf{221})) &= [\theta(\mathbf{02}), \mathbf{21}.
\end{aligned}$$

So $J \cap T^{-i-1}[\theta(\mathbf{02}), 1)$ is a union of three intervals: One in I_2^{i+1}, one in I_1^{i+1}, and one in K_1^{i+1}.

Now we assume that $J \in I_2^i$. Then

$$T^i J = [\theta(\mathbf{2}), 1).$$

As above we consider the intersection of this interval with $T^{-1}[c,1) = [\theta(\mathbf{002}),\theta(\mathbf{1})) \cup [\theta(\mathbf{102}),\theta(\mathbf{2})) \cup [\theta(\mathbf{202}),1)$. Here we get one interval, $[\theta(\mathbf{202}),1)$ and $T[\theta(\mathbf{202}),1) = [\theta(\mathbf{02}),\theta(\mathbf{21}))$ so $J \cap T^{-i-1}[c,1)$ is an interval in K_1^{i+1}. This is illustrated in the figure below where the blue line is $[\theta(\mathbf{2}),1)$ and the green lines are $T^{-1}[c,1)$.



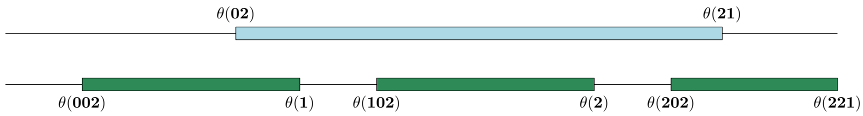
If $J \in K_1^i$ then

$$T^i J = [\theta(\mathbf{02}),\theta(\mathbf{21}))$$

and when intersected with $T^{-1}[c,1)$ we get three intervals

$$[\theta(\mathbf{02}),\theta(\mathbf{1})),\ [\theta(\mathbf{102}),\theta(\mathbf{2})),\ [\theta(\mathbf{202}),\theta(\mathbf{21})).$$

as seen below. Here the blue line is $[\theta(\mathbf{02}),\theta(\mathbf{21}))$ and the green lines are $T^{-1}[c,1)$.



Now

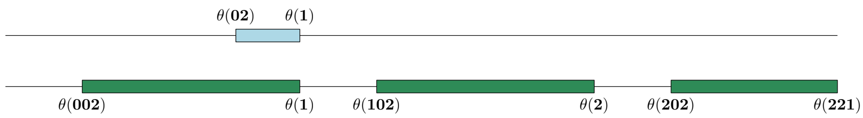
$$\begin{aligned} T[\theta(\mathbf{02}),\theta(\mathbf{1})) &= [\theta(\mathbf{2}),1) \\ T[\theta(\mathbf{102}),\theta(\mathbf{2})) &= [\theta(\mathbf{02}),1) \\ T[\theta(\mathbf{202}),\theta(\mathbf{21})) &= [\theta(\mathbf{02}),\theta(\mathbf{1})) \end{aligned}$$

so $J \cap T^{-i-1}[c,1)$ consists of three intervals: One in I_2^{i+1}, one in I_1^{i+1}, and one in K_2^{i+1}.

Finally, if $J \in K_2^i$ then

$$T^i J = [\theta(\mathbf{02}),\theta(\mathbf{1}))$$

and the intersection with $T^{-1}[c,1)$ gives us one interval $[\theta(\mathbf{02}),\theta(\mathbf{1}))$. The figure below illustrates this where the blue line is $[\theta(\mathbf{02}),\theta(\mathbf{1}))$ and the green lines are $T^{-1}[c,1)$.



Now

$$T[\theta(\mathbf{02}),\theta(\mathbf{1})) = [\theta(\mathbf{2}),1)$$

so $J \cap T^{-i-1}[c,1)$ is just one interval in I_2^{i+1}. This finishes the proof of the claim.

For each $i \geq 0$ we define a word C_i by putting a digit j in C_i for each interval in F^i that is in I_j^i for $j = 1,2$, and a digit $[j+n]$ for each interval in F^i that is in K_j^i for $j = 1,2$. From the above considerations we get that $C_0 = \mathbf{1}$ and that for $i \geq 0$ we can find C_{i+1} from C_i by the morphisms

$$\mathbf{1} \to \mathbf{213}, \quad \mathbf{2} \to \mathbf{3}, \quad \mathbf{3} \to \mathbf{214}, \quad \mathbf{4} \to \mathbf{2}.$$

So

$$\begin{aligned} C_0 &= \mathbf{1} \\ C_1 &= \mathbf{213} \\ C_2 &= \mathbf{3213214} \\ C_3 &= \mathbf{214321321432132} \end{aligned}$$

and so on.

Recall that we denoted the number of intervals in F^i by f_i. The number of digits in C_i equals f_i, and we now wish to compute these numbers. If we let f_i^j denote the number of j's in C_i then we see that for $i > 0$,

$$\begin{aligned} f_i^1 &= f_{i-1}^1 + f_{i-1}^3 \\ f_i^2 &= f_{i-1}^1 + f_{i-1}^3 + f_{i-1}^4 \\ f_i^3 &= f_{i-1}^1 + f_{i-1}^2 \\ f_i^4 &= f_{i-1}^3 \end{aligned}$$

or equivalently

$$\begin{pmatrix} f_i^1 \\ f_i^2 \\ f_i^3 \\ f_i^4 \end{pmatrix} = M \begin{pmatrix} f_{i-1}^1 \\ f_{i-1}^2 \\ f_{i-1}^3 \\ f_{i-1}^4 \end{pmatrix} \tag{3.2}$$

where

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

The characteristic polynomial of M is $x^4 - x^3 - 2x^2 - x - 1$ so from the Cayley-Hamilton Theorem we get that

$$M^4 = M^3 + 2M^2 + M + I$$

From (3.2) we now get that

$$\begin{pmatrix} f_i^1 \\ f_i^2 \\ f_i^3 \\ f_i^4 \end{pmatrix} = \begin{pmatrix} f_{i-1}^1 \\ f_{i-1}^2 \\ f_{i-1}^3 \\ f_{i-1}^4 \end{pmatrix} + 2 \begin{pmatrix} f_{i-2}^1 \\ f_{i-2}^2 \\ f_{i-2}^3 \\ f_{i-2}^4 \end{pmatrix} + \begin{pmatrix} f_{i-3}^1 \\ f_{i-3}^2 \\ f_{i-3}^3 \\ f_{i-3}^4 \end{pmatrix} + \begin{pmatrix} f_{i-4}^1 \\ f_{i-4}^2 \\ f_{i-4}^3 \\ f_{i-4}^4 \end{pmatrix}$$

for $i \geq 4$. Since $f_i = f_i^1 + f_i^2 + f_i^3 + f_i^4$ we get in particular that

$$f_i = f_{i-1} + 2f_{i-2} + f_{i-3} + f_{i-4}.$$

Using this we can calculate (see the details in the proof in the next section) that the Hausdorff dimension of $F_c(\beta)$ is

$$\dim_H F_c(\beta) = \frac{\log \rho}{\log \beta}$$

where $\rho > 0$ is the spectral radius of M and hence that

$$\dim_H F_c(\beta) \approx 0.70420.$$

3.4 The real general case

We now wish to prove the following theorem which is a a generalization of Theorem 2.4.

Theorem 3.5. *Let $\beta > 1$ have characteristic equation $0 = x^k - b_1 x^{k-1} - \cdots - b_k$ and let $c \in [0,1)$ have β-expansion $\mathbf{c_1} \cdots \mathbf{c_m}$ and*

$$F_c(\beta) = \{x \in [0,1) \mid \{\beta^i x\} \geq c \text{ for all } i \geq 0\}.$$

Let

$$n = \min\{1 \leq j \leq m \mid \mathbf{c_{j+1}} \cdots \mathbf{c_m} \leq \mathbf{c_1} \cdots \mathbf{c_{m-j}}\}$$

and

$$t = \min\{1 \leq j \leq k \mid \mathbf{b_{j+1}} \cdots \mathbf{b_k} \leq \mathbf{c_1} \cdots \mathbf{c_m}\}.$$

Assume that $c_j < b_1$ for $j < n$. Let $d_j = b_j - c_1 - 1$ for $j = 2, 3, \ldots, t$,

$$a_i = \begin{cases} b_1 - c_i - 1 & \text{for } i = 1, 2, \ldots, n-1 \\ b_1 - c_i & \text{for } i = n \end{cases}.$$

Then

$$\dim_H F_c(\beta) = \frac{\log \rho}{\log \beta}$$

where $\rho > 0$ is the spectral radius of

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n & d_2 & d_3 & \cdots & d_{t-1} & d_t \\ 1 & & & & & 1 & 1 & \cdots & 1 & 1 \\ & 1 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & 1 & & & & & & \\ 1 & 1 & \cdots & 1 & 1 & & & & & \\ & & & & & 1 & & & & \\ & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & \end{pmatrix}.$$

Proof. We see that when $k = 1$ the theorem is identical to Theorem 2.4. Because here $\beta = q$ is an integer with characteristic equation $x - q = 0$, so we need to find the spectral radius of

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \end{pmatrix}$$

which we found in the proof of Theorem 2.4 to be $x^n - a_1x^{n-1} - \cdots - a_n$. So we may assume that $k \geq 2$

Let $T : [0, 1) \to [0, 1)$ be defined by $Tx = \{\beta x\}$. Then we see that

$$F = F_c(\beta) = \bigcap_{i=0}^{\infty} F^i$$

where $F_0 = [c, 1)$ and

$$F^i = F^{i-1} \cap T^{-i}[c, 1).$$

For $i \geq 0$ we define

$$\begin{aligned}
I_1^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), 1)\} \\
I_2^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_2} \cdots \mathbf{c_m}), 1)\} \\
&\vdots \\
I_n^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_n} \cdots \mathbf{c_m}), 1)\} \\
K_1^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_2} \cdots \mathbf{b_k})\} \\
K_2^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_3} \cdots \mathbf{b_k})\} \\
&\vdots \\
K_{t-1}^i &= \{[a,b) \subseteq [0,1) : T^i[a,b) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_t} \cdots \mathbf{b_k})\}
\end{aligned}$$

and let

$$\mathscr{I}^i = I_1^i \cup \cdots \cup I_n^i \cup K_1^i \cup \cdots \cup K_{t-1}^i.$$

We now claim that F^i is the disjoint union of intervals from $\mathscr{I}^i$. We will prove this by induction. First we see that when $i = 0$ we have

$$F^0 = [c, 1) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), 1) \in I_1^0.$$

Now assume that it is true for some $i \geq 0$. So

$$F^i = \bigcup_{h=1}^{f_i} J_h$$

for some $f_i \in \mathbb{N}$ such that $J_h \in \mathscr{I}^i$ for all $h = 1, 2, \ldots, f_i$. Now

$$F^{i+1} = F^i \cap T^{-i-1}[c, 1) = \bigcup_{h=1}^{f_i} (J_h \cap T^{-i-1}[c, 1)).$$

So it is enough to prove that given one of the intervals J_h, the intersection $J_h \cap T^{-i-1}[c, 1)$ is the disjoint union of intervals that all are in $\mathscr{I}^{i+1}$. Let $J = J_h$ be one of the intervals in F^i. By the induction hypothesis we get that $J \in \mathscr{I}^i$.

First we prove a claim that we will use throughout the proof, nemely that there is the same number of intervals in $J \cap T^{-i-1}[c, 1)$ and $T^i(J \cap T^{-i-1}[c, 1))$ for any $J \in \mathscr{I}^i$. It is enough to prove that if $x, y \in J$ with $x < y$ then $T^i x < T^i y$. Now assume that $x, y \in J$ with $x < y$, but $T^i x > T^i y$. Let $\mathbf{x_1 x_2} \cdots$ and $\mathbf{y_1 y_2} \cdots$ denote the β-expansions of x and y respectively. Then we have

$$\mathbf{x_1 x_2} \cdots < \mathbf{y_1 y_2} \cdots$$

but

$$\mathbf{x_{i+1}x_{i+2}}\cdots > \mathbf{y_{i+1}y_{i+2}}\cdots$$

so

$$\mathbf{x_1x_2}\cdots\mathbf{x_i} < \mathbf{y_1y_2}\cdots\mathbf{y_i}.$$

Now let $z = \theta(\mathbf{y_1y_2}\cdots\mathbf{y_i})$. Then $x < z < y$ so we have a number $z \in J$ with $T^i z = 0$. By the definition of the intervals in $\mathscr{I}^i$ we see, that $z \notin J$ which gives the desired contradiction.

The first case we consider is when $J \in I^i_j$ for some $j < n$. Then

$$T^i(J \cap T^{-i-1}[c,1)) = [\theta(\mathbf{c_j}\cdots\mathbf{c_m}), 1) \cap T^{-1}[c,1).$$

Now $T^{-1}[c,1)$ is equal to

$$\begin{aligned}[]
[\theta(\mathbf{0c_1}\cdots\mathbf{c_m}), \theta(\mathbf{1})) \cup [\theta(\mathbf{1c_1}\cdots\mathbf{c_m}), \theta(\mathbf{2})) \cup \cdots \\
\cup [\theta([\mathbf{b_1}-\mathbf{1}]\mathbf{c_1}\cdots\mathbf{c_m}), \theta(\mathbf{b_1})) \cup [\theta(\mathbf{b_1c_1}\cdots\mathbf{c_m}), 1)
\end{aligned} \tag{3.3}$$

Now since $j < n$ we get by definition of n that

$$\mathbf{c_{j+1}}\cdots\mathbf{c_m} > \mathbf{c_1}\cdots\mathbf{c_{m-j}}.$$

So since $c_j < b_1$, the intersection $[\theta(\mathbf{c_j}\cdots\mathbf{c_m}), 1) \cap T^{-1}[c,1)$ is equal to

$$\begin{aligned}[]
[\theta(\mathbf{c_j}\cdots\mathbf{c_m}), \theta([\mathbf{c_j}+\mathbf{1}])) \cup [\theta([\mathbf{c_j}+\mathbf{1}]\mathbf{c_1}\cdots\mathbf{c_m}), \theta([\mathbf{c_j}+\mathbf{2}])) \\
\cup [\theta([\mathbf{c_j}+\mathbf{2}]\mathbf{c_1}\cdots\mathbf{c_m}), \theta([\mathbf{c_j}+\mathbf{3}])) \cup \cdots \\
\cup [\theta([\mathbf{b_1}-\mathbf{1}]\mathbf{c_1}\cdots\mathbf{c_m}), \theta(\mathbf{b_1})) \cup [\theta(\mathbf{b_1c_1}\cdots\mathbf{c_m}), 1)
\end{aligned}$$

For the first of these intervals we have

$$T[\theta(\mathbf{c_j}\cdots\mathbf{c_m}), \theta([\mathbf{c_j}+\mathbf{1}]) = [\theta(\mathbf{c_{j+1}}\cdots\mathbf{c_m}), 1).$$

For the following $b_1 - c_j - 1 = a_j$ intervals we have

$$T[\theta([\mathbf{c_j}+\mathbf{u}]\mathbf{c_1}\cdots\mathbf{c_m}), \theta([\mathbf{c_j}+\mathbf{u}+\mathbf{1}])) = [\theta(\mathbf{c_1}\cdots\mathbf{c_m}), 1)$$

for $u = 1, 2, \ldots, a_i$. For the last interval we have

$$T[\theta(\mathbf{b_1c_1}\cdots\mathbf{c_m}), 1) = [\theta(\mathbf{c_1}\cdots\mathbf{c_m}), \theta(\mathbf{b_2}\cdots\mathbf{b_k}))$$

since $1 = \theta(\mathbf{b_1}\cdots\mathbf{b_k})$. So the intersection $J \cap T^{-i-1}[c,1)$ is the union of an interval from I^{i+1}_{j+1}, a_j intervals from I^{i+1}_1, and an interval from K^{i+1}_1.

Now assume that $J \in I^i_n$. Then

$$T^i(J \cap T^{-i-1}[c,1)) = [\theta(\mathbf{c_n}\cdots\mathbf{c_m}), 1) \cap T^{-1}[c,1).$$

By definition of n we get that

$$\mathbf{c_{n+1}} \cdots \mathbf{c_m} \leq \mathbf{c_1} \cdots \mathbf{c_{m-n}}.$$

Now $T^{-1}[c,1)$ is as in (3.3) so the intersection $[\theta(\mathbf{c_n} \cdots \mathbf{c_m}), 1) \cap T^{-1}[c,1)$ is equal to

$$[\theta(\mathbf{c_n c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_n}+\mathbf{1}])) \cup [\theta([\mathbf{c_n}+\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_n}+\mathbf{2}])) \cup \cdots \\ \cup [\theta([\mathbf{b_1}-\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_1})) \cup [\theta(\mathbf{b_1 c_1} \cdots \mathbf{c_m}), 1).$$

For the first $b_1 - c_n = a_n$ intervals we have

$$T[\theta([\mathbf{c_n}+\mathbf{u}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_n}+\mathbf{u}+\mathbf{1}])) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), 1)$$

for $u = 0, 1, \ldots, a_n - 1$. For the last interval we have

$$T[\theta(\mathbf{b_1 c_1} \cdots \mathbf{c_m}), 1) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_2} \cdots \mathbf{b_k})).$$

So the intersection $J \cap T^{-i-1}[c,1)$ is the union of a_n intervals from I_1^{i+1} and an interval from K_1^{i+1}.

We now assume that $J \in K_j^i$ for $j < t-1$. Then

$$T^i(J \cap T^{-i-1}[c,1)) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}} \cdots \mathbf{b_m})) \cap T^{-1}[c,1).$$

Recall that $T^{-1}[c,1)$ is as in (3.3). Since $j < t-1$, we must have

$$\mathbf{b_{j+2}} \cdots \mathbf{b_k} > \mathbf{c_1} \cdots \mathbf{c_m}$$

and the intersection $[\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}} \cdots \mathbf{b_m})) \cap T^{-1}[c,1)$ is equal to

$$[\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{1}])) \cup [\theta([\mathbf{c_1}+\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{2}])) \cup \cdots \\ \cup [\theta([\mathbf{b_{j+1}}-\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}})) \cup [\theta(\mathbf{b_{j+1} c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}} \cdots \mathbf{b_m})).$$

For the first interval we have

$$T[\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{1}])) = [\theta(\mathbf{c_2} \cdots \mathbf{c_m}), 1).$$

For the next $b_{j+1} - c_1 - 1 = d_{j+1}$ intervals we have

$$T[\theta([\mathbf{c_1}+\mathbf{u}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{c_1}+\mathbf{u}+\mathbf{1})) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), 1)$$

for $u = 1, 2, \ldots, d_{j+1}$. And for the last interval we have

$$T[\theta(\mathbf{b_{j+1} c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}} \cdots \mathbf{b_m}) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+2}} \cdots \mathbf{b_m})).$$

So the intersection $J \cap T^{-i-1}[c,1)$ is the union of an interval from I_2^{i+1}, d_{j+1} intervals from I_1^{i+1} and an interval from K_{j+1}^{i+1}.

Finally we assume that $J \in K^i_{t-1}$. Then $j = t-1$ and

$$\mathbf{b_{j+2}} \cdots \mathbf{b_k} \leq \mathbf{c_1} \cdots \mathbf{c_m},$$

so the intersection $T^i(J \cap T^{-i-1}[c,1))$ is equal to

$$[\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{1}])) \cup [\theta([\mathbf{c_1}+\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{2}])) \cup \cdots$$
$$\cup\, [\theta([\mathbf{b_{j+1}}-\mathbf{1}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{b_{j+1}})).$$

For the first interval we have

$$T[\theta(\mathbf{c_1} \cdots \mathbf{c_m}), \theta([\mathbf{c_1}+\mathbf{1}])) = [\theta(\mathbf{c_2} \cdots \mathbf{c_m}), 1).$$

and for the next $b_{j+1} - c_1 - 1 = d_{j+1}$ intervals we have

$$T[\theta([\mathbf{c_1}+\mathbf{u}]\mathbf{c_1} \cdots \mathbf{c_m}), \theta(\mathbf{c_1}+\mathbf{u}+\mathbf{1})) = [\theta(\mathbf{c_1} \cdots \mathbf{c_m}), 1).$$

So $J \cap T^{-i-1}[c,1)$ is the union of d_{j+1} intervals from I^{i+1}_1 and an interval from K^{i+1}_{j+1}. This finishes the proof of the claim.

For each $i \geq 0$ we now define a word C_i by putting a digit $\mathbf{j}$ in C_i for each interval in F^i that is in I^i_j and a $[\mathbf{j}+\mathbf{n}]$ for each interval that is in K^i_j. From the above considerations we then get that $C_0 = \mathbf{1}$ and that for $i \geq 0$ we get C_{i+1} from C_i by the morphisms

$$\mathbf{j} \to \begin{cases} [\mathbf{j}+\mathbf{1}]\mathbf{1}^{a_j}[\mathbf{n}+\mathbf{1}] & \text{if } 1 \leq j < n \\ \mathbf{1}^{a_n}[\mathbf{n}+\mathbf{1}] & \text{if } j = n \\ \mathbf{2}\mathbf{1}^{d_{j-n+1}}[\mathbf{j}+\mathbf{1}] & \text{if } n < j < n+t-1 \\ \mathbf{2}\mathbf{1}^{d_t} & \text{if } j = n+t-1 \end{cases}$$

Now let $g^i_j(k)$ denote the number of j's in a word after i iterations where we begin with just the digit . From the morphisms we get for all $k = 1, 2, \ldots, n+t-1$ and $i \geq 1$ that

$$\begin{aligned} g^1_i(k) &= a_1 g^1_{i-1}(k) + \cdots + a_n g^n_{i-1}(k) + d_2 g^{n+1}_{i-1}(k) + \cdots d_t g^{n+t-1}_{i-1}(k) \\ g^2_i &= g^1_{i-1}(k) + g^{n+1}_{i-1}(k) + g^{n+2}_{i-1}(k) + \cdots + g^{n+t-1}_{i-1}(k) \\ g^3_i &= g^2_{i-1}(k) \\ g^4_i &= g^3_{i-1}(k) \\ &\vdots \\ g^n_i &= g^{n-1}_{i-1}(k) \\ g^{n+1}_i &= g^1_{i-1}(k) + \cdots + f^n_{i-1}(k) \\ g^{n+2}_i &= g^{n+1}_{i-1}(k) \\ g^{n+3}_i &= g^{n+2}_{i-1}(k) \\ &\vdots \\ g^{n+t-1}_i &= g^{n+t-2}_{i-1}(k), \end{aligned}$$

or equivalently

$$\begin{pmatrix} g_i^1(k) \\ g_i^2(k) \\ \vdots \\ g_i^{n+t-1}(k) \end{pmatrix} = M \begin{pmatrix} g_{i-1}^1(k) \\ g_{i-1}^2(k) \\ \vdots \\ g_{i-1}^{n+t-1}(k) \end{pmatrix} \tag{3.4}$$

for $i \geq 1$ where

$$M = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n & d_2 & d_3 & \cdots & d_{t-1} & d_t \\ 1 & & & & & 1 & 1 & \cdots & 1 & 1 \\ & 1 & & & & & & & & \\ & & \ddots & & & & & & & \\ & & & 1 & & & & & & \\ 1 & 1 & \cdots & 1 & 1 & & & & & \\ & & & & & 1 & & & & \\ & & & & & & 1 & & & \\ & & & & & & & \ddots & & \\ & & & & & & & & 1 & \end{pmatrix}.$$

Now let $x^{n+t-1} - e_1 x^{n+t-2} - e_2 x^{n+t-3} - \cdots - e_{n+t-1}$ be the characteristic polynomial of M. By the Cayley-Hamilton theorem we get that M is a root in its own characteristic polynomial,

$$0 = M^{n+t-1} - e_1 M^{n+t-2} - e_2 M^{n+t-3} - \cdots - e_{n+t-1}$$

and hence that

$$M^{n+t-1} = e_1 M^{n+t-2} + e_2 M^{n+t-3} + \cdots + e_{n+t-1}.$$

From (3.4) we get that for $i \geq n+t-1$,

$$\begin{pmatrix} g_i^1(k) \\ g_i^2(k) \\ \vdots \\ g_i^{n+t-1}(k) \end{pmatrix} = e_1 \begin{pmatrix} g_{i-1}^1(k) \\ g_{i-1}^2(k) \\ \vdots \\ g_{i-1}^{n+t-1}(k) \end{pmatrix} + \cdots + e_{n+t-1} \begin{pmatrix} g_{i-n-t+1}^1(k) \\ g_{i-n-t+1}^2(k) \\ \vdots \\ g_{i-n-t+1}^{n+t-1}(k) \end{pmatrix}.$$

Now let $g_i(k)$ denote the total number of digits after i iterations. Then $g_i(k) = g_i^1(k) + g_i^2(k) + \cdots + g_i^{n+t-1}(k)$ for all $i \geq 0$, and we get that

$$g_i(k) = e_1 g_{i-1}(k) + e_2 g_{i-2}(k) + \cdots + e_{n+t-1} g_{i-n-t+1}(k) \tag{3.5}$$

for $i \geq n+t-1$.

Now M is irreducible, so from the Perron-Frobenius theorem we get that the spectral radius ρ is real and that $\rho > 0$. We now claim that $g_i(k) \asymp \rho^i$ for $i \geq n$. This can be proved by induction. Assume that there are constants $K_1, K_2 > 0$ such that for all $k = 1, 2, \ldots, n+t-1$ we have

$$K_1 \rho^j \leq g_j(k) \leq K_2 \rho^j$$

for all $j < i$. Now

$$\begin{aligned} g_i(k) &= e_1 g_{i-1}(k) + \cdots + e_{n+t-1} g_{i-n-t+1}(k) \\ &\leq K_2(e_1 \rho^{i-1} + \cdots + e_{n+t-1} \rho^{i-n-t+1}) \\ &= K_2 \rho^{i-n-t+1}(e_1 \rho^{n+t-1} + \cdots + e_{n+t-1}) \\ &= K_2 \rho^i \end{aligned}$$

since $e_1 \rho^{n+t-2} + e_2 \rho^{n+t-3} + \cdots + e_{n+t-1} = \rho^{n+t-1}$. The lower bound is similar, so this proves the claim.

We now wish to prove that

$$\dim_H F = \frac{\log \rho}{\log \beta}.$$

We consider the slightly larger set

$$\hat{F} = \bigcap_{i=0}^{\infty} \hat{F}^i,$$

where $\hat{F}^i$ is the union of F^i and the right end-points of the half-open disjoint intervals that is in F^i. This ensures that $\hat{F}$ is a compact set and that

$$\dim_H F = \dim_H \hat{F}$$

since $\hat{F}$ is the union of F_c and a countable set of points.

First we note $T^{-i}[c, 1]$ consists of disjoint intervals of length at most $(1-c)\beta^i < \beta^i$ that are spaced $c\beta^{-i} < \beta^{-i-m}$. Now $\hat{F}^i \subseteq T^{-i}[c-1]$, so the same thing is trus for the f_i closed intervals that $\hat{F}_c^i$ consists of. We see that $\hat{F}^i$ covers $\hat{F}$, so recalling the definition of the s-dimensional Hausdorff measure we get

$$\mathscr{H}^s_{\beta^{-i}}(\hat{F}) \leq f_i \beta^{-is} = f_i \rho^{-i} \leq K_2 < \infty$$

where K_2 is as above and

$$s = \frac{\log \rho}{\log \beta},$$

so letting $i \to \infty$ we get $\mathscr{H}^s(\hat{F}) \leq K_2$. To prove a lower bound we consider any cover of $\hat{F}$. It is enough to consider finite covers of closed

intervals, so we assume that $\{U_h\}$ is such a cover. For each h we can pick j such that
$$\beta^{-j-1} \leq \operatorname{diam}(U_h) < \beta^{-j}.$$
Since the intervals in $\hat{F}^{j+m}$ are spaced at least β^{-j-m} we see that U_h intersects at most one of these intervals. Let $i \geq j+m$. Since each digit can become at most $K_2\rho^{i-j-m}$ digits after $i-j-m$ iterations, U_h intersects at most $K_2\rho^{i-j-m}$ intervals in $\hat{F}_c^i$. Now
$$K_2\rho^{i-j-m} \ll f_i\rho^{-j-m} = f_i\beta^{(-j-m)s} \leq f_i\beta^s \operatorname{diam}(U_h)^s. \tag{3.6}$$
Now pick j big enough to ensure that $\beta^{-j-1} \leq \operatorname{diam}(U_h)$ for all h, and let $i \geq j+m$. By counting intervals and applying (3.6) we get
$$\sum_h f_i\beta^s \operatorname{diam}(U_h)^s \gg f_i$$
and hence
$$\sum_h \operatorname{diam}(U_h)^s \gg \beta^{-s}$$
for any cover $\{U_h\}$. This proves that $\mathscr{H}^s(\hat{F}) \gg \beta^{-s} = \rho^{-1} > 0$ and hence that
$$\dim_H \hat{F} = \frac{\log\rho}{\log\beta}$$
as desired. This finishes the proof of the theorem. □

3.5 Two-sided approximation

In this section we give some heuristic arguments on how the result in the preceeding section can be used to give a way of calculating the Hausdorff dimension of
$$F_{c,b} = \{x \in [0,1) \mid c \leq \{q^n x\} < b \text{ for all } n \geq 0\}.$$

Note that if $x = \theta(\mathbf{x_1 x_2}\cdots)$ then
$$\{q^n x\} = \theta(\mathbf{x_{n+1} x_{n+2}}\cdots),$$
so if we let $c = \theta(\mathbf{c_1 c_2}\cdots\mathbf{c_m})$, we can see F_c as defined in chapter 2 as all words $\mathbf{x_1 x_2}\cdots$ such that
$$\mathbf{x_{n+1} x_{n+2}}\cdots\mathbf{x_{n+m}} \geq \mathbf{c_1}\cdots\mathbf{c_m}.$$
This is exactly this apporoach we used in the proof of Theorem 2.4 given in Appendix A.

Now assume that we are given a simple $\beta > 0$ with characteristic equation

$$0 = x^k - b_1 x^{k-1} - \cdots - b_k.$$

The β-expansion of 1 is then $\mathbf{b_1} \cdots \mathbf{b_k}$ and Parry [18] proved that there is $x \in [0,1)$ with β-expansion $\mathbf{x_1 x_2} \cdots$ if and only if

$$\mathbf{x_{n+1} x_{n+2}} \cdots \mathbf{x_{n+k}} < \mathbf{b_1 b_2} \cdots \mathbf{b_k}$$

for all $n = 0, 1, 2, \ldots$. So Theorem 3.5 can be seen as putting an upper bound in F_c where $q = \lceil \beta \rceil$, and changing the proof of Theorem 3.5 slightly we get the following theorem.

Theorem 3.6. *Let $q \geq 2$ be an integer. Let $c, b \in [0,1)$ have q-expansion $\mathbf{c_1} \cdots \mathbf{c_m}$ and $\mathbf{b_1} \cdots \mathbf{b_k}$ respectively. Let*

$$F_{c,b} = \{x \in [0,1) \mid c \leq \{q^i x\} < d \text{ for all } i \geq 0\}.$$

Let

$$n = \min\{1 \leq j \leq m \mid \mathbf{c_{j+1}} \cdots \mathbf{c_m} \leq \mathbf{c_1} \cdots \mathbf{c_{m-j}}\}$$

and

$$t = \min\{1 \leq j \leq k \mid \mathbf{b_{j+1}} \cdots \mathbf{b_k} \leq \mathbf{c_1} \cdots \mathbf{c_m}\}.$$

Assume that $c_j < b_1$ for $j < n$. Let $d_j = b_j - c_1 - 1$ for $j = 2, 3, \ldots, t$, and

$$a_i = \begin{cases} b_1 - c_i - 1 & \text{for } i = 1, 2, \ldots, n-1 \\ b_1 - c_i & \text{for } i = n \end{cases}.$$

Then

$$\dim_H F_{c,b} = \frac{\log \rho}{\log q}$$

where $\rho > 0$ is the spectral radius of

$$\begin{pmatrix}
a_1 & a_2 & \cdots & a_{n-1} & a_n & d_2 & d_3 & \cdots & d_{t-1} & d_t \\
1 & & & & & 1 & 1 & \cdots & 1 & 1 \\
 & 1 & & & & & & & & \\
 & & \ddots & & & & & & & \\
 & & & 1 & & & & & & \\
1 & 1 & \cdots & 1 & 1 & & & & & \\
 & & & & & 1 & & & & \\
 & & & & & & 1 & & & \\
 & & & & & & & \ddots & & \\
 & & & & & & & & 1 &
\end{pmatrix}.$$

Chapter 4

Restricted denominators, $p^n q^m$

4.1 Introduction

We now consider Diophantine approximation where we restrict the denominators to be in the set $A = \{p^n q^m \mid n, m \geq 0\}$ where p, q are two distinct primes. The work in this chapter has been done joint with Sanju Velani.

Throughout the chapter we let $\phi(n) = \#\{k \leq n : k \mid n\}$ denote the Euler totient function and let $l(\cdot)$ be the Lebesgue measure.

4.2 Khintchine's theorem

We now prove a version of Khintchine's theorem in this setting, which follows from the Duffin-Schaeffer theorem (Thm 2.5 in [8]). The Duffin-Schaeffer theorem is an attempt to remove the assumption of monotonicity from Khinchine's theorem.

Theorem 4.1 (Duffin-Schafer). *Let $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ and assume thats*

$$\limsup_{N \in \mathbb{N}} \left(\sum_{n=1}^{N} \frac{\Psi(n)\phi(n)}{n} \right) \left(\sum_{n=1}^{N} \Psi(n) \right)^{-1} > 0.$$

For almost all $x \in \mathbb{R}$ there are infinitely many $n \in \mathbb{N}$ such that

$$\|nx\| < \Psi(n) \tag{4.1}$$

if

$$\sum_{n=1}^{\infty} \frac{\Psi(n)\phi(n)}{n} = \infty,$$

and if the series converges, there are only finitely many $n \in \mathbb{N}$ such that (4.1) *is true.*

We also need the following lemma on the convergence of series of functions supported on A. We let $a(N) = \#\{n \in A \mid n \leq N\}$.

Lemma 4.2. *Let $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a non-increasing function. Then*

$$\sum_{n \in A} \psi(n) < \infty \iff \sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} < \infty.$$

Proof. Let $k \geq 2$. Then we have

$$\sum_{n \in A} \psi(n) = \sum_{t=0}^{\infty} \sum_{k^t < p^i q^j \leq k^{t+1}} \psi(p^i q^j).$$

Assume without loss of generality that $p < q$. We are trying to count the number of lattice points $(i, j) \in \mathbb{N}^2$ such that

$$k^t < p^i q^j \leq k^{t+1}$$

for $k \geq p$ and $t \geq 0$. Equivalently we wish to find solutions to the inequality

$$t \log k < i \log p + j \log q \leq (t+1) \log k.$$

For each j with $0 \leq j \leq \frac{(t+1)\log k}{\log q}$ we can find at least one i to make this true, this proves the lower bound. For each j we see that $\frac{\log k}{\log p} + 1$ is an upper bound to how many i's we can find, so we get that

$$\frac{(t+1)\log k}{\log q} + 1 \leq a(k^{t+1}) - a(k^t) \leq \left(\frac{(t+1)\log k}{\log q} + 1 \right) \left(\frac{\log k}{\log p} + 1 \right).$$

Figure 4.1 is an illustration of this.

In particular the above estimates show that $a(k^{t+1}) - a(k^t) \asymp t$ so

$$\sum_{t=0}^{\infty} \sum_{k^t < p^i q^j \leq k^{t+1}} \psi(p^i q^j) \asymp \sum_{t=0}^{\infty} \psi(k^t) t = \sum_{t=0}^{\infty} \sum_{n=k^t+1}^{k^{t+1}} \frac{\psi(k^t) t}{k^{t+1} - k^t}.$$

Since ψ is assumed to be non-increasing we have

$$\sum_{t=0}^{\infty} \sum_{n=k^t+1}^{k^{t+1}} \frac{\psi(k^t) t}{k^{t+1} - k^t} \asymp \sum_{t=0}^{\infty} \sum_{n=k^t+1}^{k^{t+1}} \frac{\psi(n) \log n}{n} = \sum_{n=1}^{\infty} \frac{\psi(n) \log(n)}{n},$$

which finishes the proof. $\square$

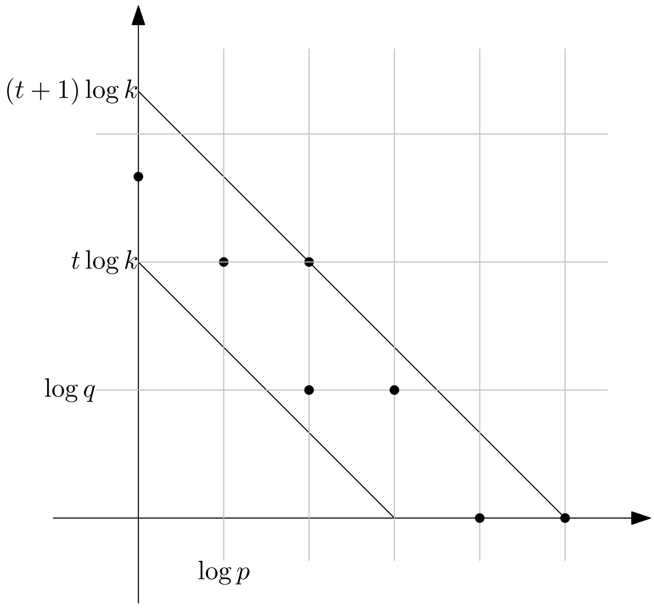


Figure 4.1: *We are counting lattice points in this strip.*

This gives us the following version of Khinchine's theorem.

Theorem 4.3. *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a monotonic function. For almost all $x \in \mathbb{R}$, the inequality*

$$\left| x - \frac{k}{n} \right| < \frac{\psi(n)}{n}$$

is true for infinitely many integers k, n with $n \in A$ if

$$\sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} = \infty,$$

and if

$$\sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} < \infty$$

then the inequality is false for almost all $x \in \mathbb{R}$.

Proof. Let

$$\Psi(n) = \begin{cases} \psi(n) & \text{if } n \in A \\ 0 & \text{otherwise} \end{cases}$$

Then

$$\limsup_N \left(\sum_{n=1}^{N} \frac{\Psi(n)\phi(n)}{n} \right) \left(\sum_{n=1}^{N} \Psi(n) \right)^{-1} = \frac{(p-1)(q-1)}{pq} > 0$$

and
$$\sum_{n=1}^{\infty} \frac{\Psi(n)\phi(n)}{n} = \frac{(p-1)(q-1)}{pq} \sum_{n=1}^{\infty} \Psi(n)$$
since $\phi(p^n q^m) = (p-1)p^{n-1}(q-1)q^{m-1}$. So the theorem follows from the above lemma. □

4.3 Badly approximable numbers

For $\delta \geq 0$, let
$$\mathrm{Bad}_A(\delta) = \{x \in [0,1] : \liminf_{n \in A} (\log n)^{\delta} \|nx\| > 0\}.$$
We claim that for $\delta > 2$ this set has full measure, and for $\delta \leq 2$ it is a nullset. To prove this we need the following consequence of Cauchy's condensation test.

Proposition 4.4. *We have*
$$\sum_{n=1}^{\infty} \frac{1}{n \log n} = \infty, \tag{4.2}$$
but for any $\varepsilon > 0$ we have
$$\sum_{n=1}^{\infty} \frac{1}{n (\log n)^{1+\varepsilon}} < \infty. \tag{4.3}$$

Proof. Recall that Cauchy's condensation test states that if $a_n \geq 0$ and $a_n \geq a_{n+1}$ for all n then
$$\sum_{n=1}^{\infty} a_n < \infty \iff \sum_{k=1}^{\infty} 2^k a_{2^k} < \infty.$$
This shows that the sum in (4.2) diverges as it is asymptotically equivalent to the sum
$$\sum_{k=1}^{\infty} \frac{1}{k \log 2}$$
which is divergent. Similarly, the sum in (4.3) converges as it is asymptotically equivalent to the sum
$$\sum_{k=1}^{\infty} \frac{1}{(k \log 2)^{1+\varepsilon}},$$
which is convergent. □

This gives the following consequence of Khinchine's theorem.

Theorem 4.5. *For $\delta \leq 2$, the set $\mathrm{Bad}_A(\delta)$ is a nullset, and if $\delta > 2$ then $\mathrm{Bad}_A(\delta)$ has full measure.*

Proof. Let $\delta \leq 2$, $m \in \mathbb{N}$ and define $\psi(n) = \frac{1/m}{(\log n)^{\delta-1}}$. Then

$$\sum_{n=1}^{\infty} \frac{\psi(n) \log n}{n} = \frac{1}{m} \sum_{n=1}^{\infty} \frac{1}{n(\log n)^{\delta-1}}$$

which diverges by the above proposition. From Theorem 4.3 we get that for almost all $x \in [0,1]$ there are infinitely many $n \in A, i \in \mathbb{N}$ such that

$$\left| x - \frac{i}{n} \right| < \frac{1/m}{n(\log n)^{\delta-1}}.$$

Let E_m denote these x and let

$$E = \bigcap_{m \in \mathbb{N}} E_m.$$

Since each E_m has full measure, the set E has full measure. Now if $x \in \mathrm{Bad}_A(\delta)$ then there is $c > 0$ such that for all sufficiently large $n \in A$

$$\left| x - \frac{i}{n} \right| > \frac{c}{n(\log n)^{\delta-1}}$$

for all $i \in \mathbb{N}$. For m large enough we see that $x \notin E_m$, so $x \notin E$ and hence $\mathrm{Bad}_A(\delta)$ must be a nullset.

Now let $\delta > 2$ and define ψ as above. Now the corresponding sum diverges, and hence must E be a nullset. If we take $x \notin E$ there must exist m such that $x \notin E_m$ and hence we have that for all but finitely many $n \in A$ and $i \in \mathbb{N}$ that

$$\left| x - \frac{i}{n} \right| > \frac{1/m}{n(\log n)^{\delta-1}}.$$

This implies that $x \in \mathrm{Bad}_A(\delta)$ and hence that $\mathrm{Bad}_A(\delta)$ has full measure. □

From now on we define $\mathrm{Bad}_A = \mathrm{Bad}_A(2)$, because this is the largest δ that gives a nullset.

4.4 Lower bound

Now we would like to prove that replacing the logarithm in the definition of Badly approximable numbers with something slower makes the set very small. In the classical case the set of badly approximable numbers

$$\mathrm{Bad} = \{x \in \mathbb{R} \mid \liminf_n n \|nx\| > 0\}$$

is a nullset, and if we let $\varepsilon > 0$ then $\{x \in \mathbb{R} \mid \liminf_n n^{1+\varepsilon} \|nx\| > 0\} = \emptyset$. The set will not become empty in our case, since if we let $x = \frac{s}{t}$ with $(t, p) = (t, q) = 1$, then

$$\|p^n q^m x\| = \left\| \frac{p^n q^m s}{t} \right\|.$$

The denominator and numerator of this fraction are coprime, so the distance to the nearest integer is at least $\frac{1}{t}$ and hence

$$\|p^n q^m x\| \geq \frac{1}{t}$$

for all $n, m \in \mathbb{N}$. So fractions with denominator coprime to p and q will always be badly approximable no matter what we replace the logarithm with.

It turns out that we also have to exclude Liouville Numbers. Recall that such a number is defined as follows.

Definition 4.6. We say that $x \in [0, 1]$ is a *Liouville number* if for any integer $n > 0$ there is $p, q \in \mathbb{N}$ with $q > 1$ such that

$$0 < \left| x - \frac{p}{q} \right| < \frac{1}{q^n}.$$

From a result of Bourgain, Lindenstrauss, Michel, and Venkatesh [4] we get the following effective version of Furstenberg's orbit closure theorem.

Theorem 4.7. *Suppose $p, q \in \mathbb{N}$ are multiplicative independent and $x \in [0, 1]$ is neither rational nor a Liouville number. Then there is a constant $\kappa = \kappa(p, q)$ such that for $N \geq N_0(p, q, x)$, the minimal distance from any number in $[0, 1]$ to an element from $\{p^s q^t x \mod 1 \mid s, t \leq N\}$ is $\frac{1}{(\log \log N)^\kappa}$.*

This theorem yields that there is $\kappa > 0$ such that for $x \in [0, 1]$ neither Liouville nor rational we have

$$\inf_{s,t \leq N} \left\| p^s q^t x \right\| \leq \frac{1}{(\log \log N)^\kappa}. \tag{4.4}$$

for sufficiently large N. Now since the set of Liouville numbers and rationals has zero Hausdorff dimension, we get the following lemma.

Theorem 4.8. *There is a constant $c > 0$ only dependent on p and q such that the set*

$$\mathrm{Bad}'_A = \{x \in [0,1] : \liminf_{n \in A} (\log\log\log n)^c \|nx\| > 0\}$$

is a subset of the union of rationals and Liouville numbers. In particular it has Hausdorff dimension zero.

Proof. We see that if we let $c = \kappa/2$ where $\kappa > 0$ is the constant from (4.4) we get

$$\inf_{s,t \le N} (\log\log N)^c \left\|p^s q^t x\right\| \le \frac{1}{(\log\log N)^{\kappa/2}}.$$

Without loss of generality we assume that $q > p$. Since $2N \ge s + t$ we have

$$N \ge \frac{\log(p^s q^t)}{2\log q}$$

and hence

$$\inf_{s,t \le N} \left(\log\log\left(\frac{\log(p^s q^t)}{2\log q}\right)\right)^c \left\|p^s q^t\right\| \to 0$$

as $N \to \infty$. So to prove the lemma we need to prove that there is a constant $\delta > 0$ such that for sufficiently large s, t we have

$$\log\log\left(\frac{\log(p^s q^t)}{2\log q}\right) \ge \delta \log\log\log(p^s q^t).$$

Now since the left hand side is $\log(\log\log(p^s q^t) - k)$ with $k = \log 2 + \log\log q$, it is enough to prove that for sufficiently large x, there is a constant $\delta > 0$ such that

$$\log(x - k) \ge \delta \log x.$$

So we need to find $\delta > 0$ such that $x^\delta \le x - k$ for sufficiently large x. Since this is true for any $\delta < 1$, this finishes the proof of the theorem. $\square$

4.5 Zero-one laws for Hausdorff measures

We now wish to prove a version of Khinchine's theorem for Hausdorff measures by using the Mass transference Principle from [2]. This is designed to transfer Lebesgue statements to statements about Hausdorff measures. For a ball $B \subseteq \mathbb{R}^k$ with center c and radius r we let $B^s \subseteq \mathbb{R}^k$ be the ball with center c and radius $r^{s/k}$.

Theorem 4.9 (Mass Transference Principle)**.** *Let $s < k$ and let $\{B_i\}_{i\in\mathbb{N}}$ be a sequence of balls in $\mathbb{R}^k$ with radii r_i such that $r_i \to 0$ as $i \to \infty$. Suppose that* $\limsup_i B_i^s$ *is of full Lebesgue measure. Then*

$$\mathscr{H}^s(\limsup_i B_i \cap B) = \infty$$

for all closed balls B.

We define for $k \geq 1, \psi : \mathbb{N} \to \mathbb{R}$ and an infinite subset $P \subseteq \mathbb{N}$,

$$S_k(\psi, P) = \limsup_{n\in P} \Big\{(x_1, \cdots, x_k) \in [0,1)^k : \prod_{i=1}^{k} \|nx\| < \psi(n)\Big\}$$

of multiplicatively well approximable points. Now Theorem 4.3 can be formulated as

$$l(S_1(\psi, A)) = \begin{cases} 0 & \text{if } \sum_{k=0}^{\infty} \frac{\psi(k)\log k}{k} < \infty \\ 1 & \text{if } \sum_{k=0}^{\infty} \frac{\psi(k)\log k}{k} = \infty \end{cases}$$

In this section we will consider the case $k = 1$ and in the next we will consider the general case $k \geq 1$. Note that

$$S_1(\psi, A) = \limsup_{n\in A} \bigcup_{k=0}^{n} \left[\frac{k}{n} - \frac{\psi}{n}, \frac{k}{n} + \frac{\psi n}{n}\right]. \tag{4.5}$$

We now use Mass Transference Principle to prove the following theorem.

Theorem 4.10. *Let $0 < s < 1$ and let $\psi : \mathbb{N} \to \mathbb{R}$ be monotonic. Then*

$$\mathscr{H}^s(S_1(\psi, A)) = \begin{cases} 0 & \textit{if } \sum_{k=0}^{\infty} \left(\frac{\psi(k)}{k}\right)^s \log k < \infty \\ \infty & \textit{if } \sum_{k=0}^{\infty} \left(\frac{\psi(k)}{k}\right)^s \log k = \infty \end{cases}$$

Proof. Suppose that

$$\sum_{n=0}^{\infty} \left(\frac{\psi(n)}{n}\right)^s \log n = \infty.$$

From (4.5) we see that we have to check if

$$\limsup_{n\in A} \bigcup_{k=0}^{n} \left[\frac{k}{n} - \left(\frac{\psi}{n}\right)^s, \frac{k}{n} + \left(\frac{\psi n}{n}\right)^s\right]$$

is of full Lebesgue measure. This set is equal to $S_q(\xi, A)$ where

$$\xi(n) = \frac{\psi(n)^s}{n^{s-1}},$$

so by Theorem 4.3 we get that $S_1(\xi, A)$ is of full measure since

$$\sum_{n=0}^{\infty} \frac{\xi(n) \log n}{n} = \sum_{n=0}^{\infty} \left(\frac{\psi(n)}{n}\right)^s \log n = \infty$$

and by the mass transference principle this gives us that $\mathscr{H}^s(S_1(\psi, A)) = \infty$.

Now suppose that

$$\sum_{n=0}^{\infty} \left(\frac{\psi(n)}{n}\right)^s \log n < \infty.$$

For each $N \geq 0$ we see that

$$S_1(\psi, A) \subseteq \bigcup_{n \geq N n \in A} \bigcup_{k=0}^{n} \left[\frac{k}{n} - \frac{\psi}{n}, \frac{k}{n} + \frac{\psi n}{n}\right],$$

so for all $N \geq 0$

$$\mathscr{H}^s(S_1(\psi, A)) \leq \sum_{n \geq N n \in A} (n+1) \left(\frac{2\psi(n)}{n}\right)^s \asymp \sum_{n \geq N n \in A} \frac{\psi(n)^s}{n^{s-1}}. \tag{4.6}$$

Now if we let

$$\Psi(n) = \begin{cases} \frac{\psi(n)^s}{n^{s-1}} & \text{if } n \in A \\ 0 & \text{otherwise} \end{cases}$$

then

$$\sum_{n=0}^{\infty} \Psi(n) = \sum_{n=0}^{N-1} \Psi + \sum_{n \geq N n \in A} \frac{\psi(n)^s}{n^{s-1}}$$

and using Lemma 4.2 we see that this series is convergent, so

$$\sum_{n \geq N n \in A} \frac{\psi(n)^s}{n^{s-1}} \to 0$$

as $N \to \infty$ so from (4.6) we get $\mathscr{H}^s(S_1(\psi, A)) = 0$. $\square$

4.6 Multiplicative approximation

We now consider $S_k(\psi, A)$ for general $k \geq 1$, and again we wish to prove a version of Khinchine's theorem. The proof is similar to that of Theorem 4.3, namely it is a direct consequence of the Duffin-Schaeffer theorem, but this time we need a multiplicative analogue of this theorem which is proved in [3].

Theorem 4.11 (Multiplicative Duffin-Schafer). *Let $k \geq 1$ and let $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$. Suppose that*

$$\sum_{n=1}^{\infty} \Psi(n) \log(n)^{k-1} = \infty$$

and

$$\limsup_N \left(\sum_{n=1}^{N} \left(\frac{\phi(n)}{n} \right)^k \Psi(n) \log(n)^{k-1} \right) \left(\sum_{n=1}^{N} \Psi(n) \log(n)^{k-1} \right)^{-1} > 0.$$

Then $l(S_k(\Psi, \mathbb{N})) = 1$.

This enables us to prove the following theorem.

Theorem 4.12. *Let $k \geq 1$ and $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a monotonic function. If*

$$\sum_{n=0}^{\infty} \frac{\psi(n) \log(n)^k}{n} = \infty$$

then $l(S_k(\psi, A)) = 1$.

Proof. Let

$$\Psi(n) = \begin{cases} \psi(n) & \text{if } n \in A \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $S_k(\Psi, \mathbb{N}) = S_k(\psi, A)$. Now

$$\limsup_N \left(\sum_{n=1}^{N} \left(\frac{\phi(n)}{n} \right)^k \Psi(n) \log(n)^{k-1} \right) \left(\sum_{n=1}^{N} \Psi(n) \log(n)^{k-1} \right)^{-1} = \left(\frac{(p-1)(q-1)}{pq} \right)^k > 0$$

so by Theorem 4.11 we have that if

$$\sum_{n=1}^{\infty} \Psi(n) \log(n)^{k-1} = \infty$$

then $l(S_k(\phi, A)) = 1$, and the theorem now follows from Lemma 4.2. □

4.7 Simultaneous approximation

We now consider the sets

$$T_k(\psi, P) = \limsup_{n\in P} = \{(x_1, \cdots, x_k) \in [0,1)^k : \max_i \|nx_i\| < \psi(n)\}$$

where $P \subseteq \mathbb{N}$ is an infinite set, $k \geq 2$ and $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$. Once again we wish to prove Khinchine's theorem in this case which here follows immediately from the fact that the Duffin-Schafer conjecture is true for dimension ≥ 2, namely that we from [21] get the following theorem.

Theorem 4.13 (Duffin-Schafer in higher dimension)**.** *Let $k \geq 2$ and $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$. We have $l(T_k(\Psi, \mathbb{N})) \in \{0, 1\}$ and if*

$$\sum_{n=1}^{\infty} \left(\frac{\Psi(n)\phi(n)}{n}\right)^k = \infty$$

then $l(T_k(\Psi, \mathbb{N})) = 1$.

Using this we can prove the following theorem.

Theorem 4.14. *Let $k \geq 1$ and let $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be monotonic. Then*

$$l(T_k(\psi, A)) = \begin{cases} 0 & \text{if } \sum_{n=0}^{\infty} \frac{\psi(n)^k \log n}{n} < \infty \\ 1 & \text{if } \sum_{n=0}^{\infty} \frac{\psi(n)^k \log n}{n} = \infty \end{cases}$$

Proof. We have already considered the case $k = 1$ so assume $k \geq 2$ and that

$$\sum_{n=0}^{\infty} \frac{\psi(n)^k \log n}{n} = \infty.$$

Let

$$\Psi(n) = \begin{cases} \psi(n) & \text{if } n \in A \\ 0 & \text{otherwise.} \end{cases}$$

Then $l(T_k(\Psi, \mathbb{N})) = l(T_k(\psi, A))$ so we need to prove that

$$\sum_{n=1}^{\infty} \left(\frac{\Psi(n)\phi(n)}{n}\right)^k = \infty.$$

The terms in this sum are only non-zero when $n = p^i q^j$ and here

$$\frac{\phi(n)}{n} = \frac{(p-1)(q-1)}{pq},$$

so

$$\sum_{n=1}^{\infty}\left(\frac{\Psi(n)\phi(n)}{n}\right)^k \asymp \sum_{n=1}^{\infty}\Psi(n)^k \asymp \sum_{n=0}^{\infty}\frac{\psi(n)^k\log n}{n} = \infty$$

using Lemma 4.2.

Now assume that

$$\sum_{n=0}^{\infty}\frac{\psi(n)^k\log n}{n} < \infty.$$

We see that

$$T_k(\psi, A) \subseteq \limsup_{n\in A}\bigcup_{m=0}^{n}\left[\frac{m}{n}-\frac{\psi(n)}{n}, \frac{m}{n}+\frac{\psi(n)}{n}\right]^k,$$

and by Lemma 4.2 we get

$$\sum_{n\in A}(n+1)\left(\frac{2\psi(n)}{n}\right)^k \asymp \sum_{n\in A}\psi(n)^k \asymp \sum_{n=0}^{\infty}\frac{\psi(n)^k\log n}{n} < \infty,$$

so an application of the Borel-Cantelli lemma finishes the proof. □

4.8 Generalizations

We should note that all the results here, except the lower bound result based on the theorem by Bourgain, Lindenstrauss, Michel and Venkatesh, can be generalized to the case of more than two primes p and q, ie. the set A can be replaced with

$$A_m = \{p_1^{n_1}\cdots p_m^{n_m} \mid n_i \in \mathbb{N}\},$$

for $m \geq 2$. Here Lemma 4.2 is as follows.

Lemma 4.15. *Let $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a non-increasing function. Then*

$$\sum_{n\in A_m} < \infty \iff \sum_{n=1}^{\infty}\frac{\psi(n)\log(n)^{m-1}}{n} < \infty.$$

The proofs in this case will be almost identical, except that the notation will be more tedious. The different results in this chapter will then be as follows.

Theorem 4.16. *Let $\psi : \mathbb{N} \to \mathbb{R}_+$ be a monotonic function. For almost all $x \in \mathbb{R}$, the inequality*

$$\left|x - \frac{k}{n}\right| < \frac{\psi(n)}{n}$$

is true for infinitely many integers k, n with $n \in A_m$ if

$$\sum_{n=1}^{\infty} \frac{\psi(n)\log(n)^{m-1}}{n} = \infty,$$

and if

$$\sum_{n=1}^{\infty} \frac{\psi(n)\log(n)^{m-1}}{n} < \infty$$

then the inequality is false for almost all $x \in \mathbb{R}$.

Theorem 4.17. *For $\delta \leq m$, the set $\mathrm{Bad}_{A_m}(\delta)$ is a nullset, and if $\delta > m$ then $\mathrm{Bad}_A(\delta)$ has full measure.*

Theorem 4.18. *Let $0 < s < 1$ and let $\psi : \mathbb{N} \to \mathbb{R}$ be monotonic. Then*

$$\mathscr{H}^s(S_1(\psi, A_m)) = \begin{cases} 0 & \text{if } \sum_{k=0}^{\infty} \left(\frac{\psi(k)}{k}\right)^s \log(k)^{m-1} < \infty \\ \infty & \text{if } \sum_{k=0}^{\infty} \left(\frac{\psi(k)}{k}\right)^s \log(k)^{m-1} = \infty \end{cases}$$

Theorem 4.19. *Let $k \geq 1$ and $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be a monotonic function. If*

$$\sum_{n=0}^{\infty} \frac{\psi(n)\log(n)^{k+m-1}}{n} = \infty$$

then $l(S_k(\psi, A_m)) = 1$.

Theorem 4.20. *Let $k \geq 1$ and let $\psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ be monotonic. Then*

$$l(T_k(\psi, A_m)) = \begin{cases} 0 & \text{if } \sum_{n=0}^{\infty} \frac{\psi(n)^k \log(n)^{m-1}}{n} < \infty \\ 1 & \text{if } \sum_{n=0}^{\infty} \frac{\psi(n)^k \log(n)^{m-1}}{n} = \infty \end{cases}$$

Chapter 5

Littlewood's Conjecture

5.1 Introduction

In the previous chapters we have discussed several ways of considering approximation by rationals. In this chapter we will discuss the Littlewood conjecture which deals with approximation of two real numbers x, y by rationals with the same denominator. More precisely, the conjecture states that

$$\liminf_n n \|nx\| \|ny\| = 0$$

for all $x, y \in \mathbb{R}$. Recall that if one of the real numbers, say x, is not a badly approximable number, then

$$\liminf_n n \|nx\| = 0,$$

and hence

$$\liminf_n n \|nx\| \|ny\| < \frac{1}{2} \liminf_n n \|nx\| = 0,$$

so the conjecture is only interesting when $x, y \in \mathrm{Bad}$. This conjecture, proposed by John Edensor Littlewood around 1930, has recently attracted a lot of attention. Einsiedler, Katok, and Lindenstrauss [5] proved that the set of exceptions to this conjecture has Hausdorff dimension zero. Preceding the paper of Einsiedler, Katok and Lindenstrauss is the result of Pollington and Velani [20], which states that for any $x \in \mathrm{Bad}$, there is a set $G \subseteq \mathrm{Bad}$ which is also Hausdorff dimension one, such that for $y \in G$, the Littlewood conjecture is true for x, y.

A problem related to the Littlewood conjecture is the so-called mixed Littlewood conjecture of de Mathan and Teulié [15]. Here we only approximate a single real number x and replace the condition on y by a condition of divisibility. The mixed Littlewood conjecture states that for a prime p we have

$$\liminf_n n \, |n|_p \, \|nx\| = 0$$

where

$$|n|_p = \min\left\{\frac{1}{p^k} : p^k \mid n\right\}$$

is the p-adic absolute value. A more general version of this conjecture comes about by letting $\mathcal{D} = (q_k)_{k=1}^{\infty}$ with $q_k \mid q_{k+1}$ be a *pseudo-absolute value sequence* and

$$|n|_{\mathcal{D}} = \min\left\{\frac{1}{q_k} : q_k \mid n\right\}$$

be a *pseudo-absolute value*. The conjecture then states that

$$\liminf_n n\,|n|_{\mathcal{D}}\,\|nx\| = 0.$$

Harrap and Haynes [9] has proved that adding an extra p-adic norm to the left hand side makes this conjecture true, that is

$$\liminf_n n\,|n|_p\,|n|_{\mathcal{D}}\,\|nx\| = 0$$

when there is some bound C such that $\frac{q_{k+1}}{q_k} < C$ for all $k \geq 1$.

A final problem related to the Littlewood conjecture is a conjecture of Cassels, which was recently resolved by Shapira [23]. This theorem states that for almost all $x, y \in \mathbb{R}$,

$$\liminf_n n\,\|nx - \gamma_1\|\,\|ny - \gamma_2\| = 0$$

for all $\gamma_1, \gamma_2 \in \mathbb{R}$. Here we consider the case where $\gamma_1 = 0$.

The work in this chapter is joint with Simon Kristensen and Alan Haynes [10]. We prove the following result.

Theorem 5.1. *Fix $\varepsilon > 0$ and let $\{x_i\} \subseteq \mathrm{Bad}$ be a countable set of badly approximable numbers, and $\{\mathcal{D}_j\}$ a countable set of pseudo-absolute value sequences. Then there is set of $G \subseteq \mathrm{Bad}$ of Hausdorff dimension 1 such that for any $y \in G$, the following is true.*

(i) For any $i \in \mathbb{N}$ and $\gamma \in \mathbb{R}$ the inequality

$$n\,\|nx_i\|\,\|ny - \gamma\| < \frac{1}{(\log n)^{1/2-\varepsilon}}, \tag{5.1}$$

has infinitely many solutions $n \in \mathbb{N}$, and

(ii) For any $j \in \mathbb{N}$ and $\delta \in \mathbb{R}$ we have that

$$\liminf_{n\to\infty} n\,|n|_{\mathcal{D}_j}\,\|ny - \delta\| = 0. \tag{5.2}$$

Furthermore, for each j such that $\mathcal{D}_j = (q_k)$ satisfies the inequality $q_k \leq C^k$ for some $C > 1$, we may replace (5.2) *by the stronger statement that for any $\delta \in \mathbb{R}$ the inequality*

$$n\,|n|_{\mathcal{D}_j}\,\|ny - \delta\| < \frac{1}{(\log n)^{1/2-\varepsilon}}, \tag{5.3}$$

has infinitely many solutions $n \in \mathbb{N}$.

This theorem relies on a discrepancy estimate for almost all points with respect the a certain measure, namely the Kaufman measure.

5.2 Kaufman's measure

Recall that given a real number $x \in [0, 1)$ one can find the continued fraction expansion,

$$x = [a_1, a_2, \ldots] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots}}}$$

The Badly approximable numbers Bad are exactly the numbers x such that the sequence (a_n) is bounded. As we did in Chapter 2, we fix a bound $M \geq 3$ and let

$$F_M = \{x \in [0, 1) : a_n(x) < M \text{ for all } n \geq 1\}.$$

Kaufman [13] introduced a measure μ_M supported on F_M with, among other things, the following properties.

(i) For any $s < \dim(F_M)$, there are positive constants $c, l > 0$ such that for any interval $I \subseteq [0, 1)$ of length $|I| \leq l$,

$$\mu_N(I) \leq c\,|I|^s.$$

(ii) For any M, there are positive constants $c, \eta > 0$ such that the Fourier transform $\hat{\mu}_M$ of the Kaufman measure μ_M satisfies

$$\hat{\mu}_M(u) \leq c\,|u|^{-\eta}.$$

5.3 Discrepancy

Another key ingredient in this proof is the notion of discrepancy which measure how well the fractional parts of a given sequence is distributed in the interval $[0, 1)$. Given an sequence $(x_n) \subseteq \mathbb{R}$, we define the discrepancy of the sequence is

$$D_N(x_n) = \sup_{I \subseteq [0,1)} \left| \sum_{n=1}^{N} \chi_I(\{x_n\}) - n\,|I| \right|,$$

where the supremum is over all intervals $I \subseteq [0, 1)$, χ_I is the indicator function on I, $\{\cdot\}$ denotes the fractional part of a number, and $|I|$ is the length of the interval I. If

$$\frac{D_N(x_n)}{N} \to 0 \text{ as } N \to \infty,$$

the sequence is uniformly distributed modulo 1.

An important result on discrepancy is the Erdős–Turán inequality, see e.g. [16]. Note that we let $e(t) = e^{2\pi i t}$.

Theorem 5.2 (Erdős–Turán inequality)**.** *For any positive integer K and any sequence $(x_n) \subseteq [0, 1)$,*

$$D_N(x_n) \leq \frac{N}{K+1} + 3 \sum_{k=1}^{K} \frac{1}{k} \left| \sum_{n=1}^{N} e(kx_n) \right|$$

We will also need the following lemma found in [8].

Lemma 5.3. *Let (X, μ) be a measure space with $\mu(X) < \infty$. Let $F(n, m, x), n = 0, 1, \ldots, m = 1, 2, \ldots$ be μ-measurable functions and let ϕ_n be a sequence of real numbers such that $|F(n-1, n, x)| \leq \phi_n$ forn $= 1, 2, \ldots$. Let $\Phi_N = \phi_1 + \cdots + \phi_N$ and assume that $\Phi_N \to \infty$. Suppose that for $0 \leq u < v$ we have*

$$\int_X |F(u, v, x)|^2 \, d\mu \ll \sum_{n=u}^{v} \phi_n.$$

Then for μ-almost all x, we have

$$F(0, N, x) \ll \Phi_N^{1/2} \log(\Phi_N)^{3/2+\varepsilon} + \max_{n \leq N} \phi_n.$$

Using these results we can now prove the following theorem, which will be crucial in proving the main result, but is also interesting in its own right

Theorem 5.4. *Let μ_M be a Kaufman measure and assume that for positive integers $u < v$ we have*

$$\sum_{n,m=u}^{v} |a_n - a_m|^{-\eta} \ll \frac{1}{\log v} \sum_{n=u}^{v} \psi_n$$

where (ψ_n) is a sequence of non-negative numbers and $\eta > 0$ is the constant from property (ii) *of the Kaufman measure. Then for μ_M-almost every $x \in [0,1]$ we have*

$$D_N(a_n x) \ll (N\log(N)^2 + \Psi_N)^{1/2} \log(N\log(N)^2 + \Psi_N)^{3/2+\varepsilon} + \max_{n\le N} \psi_n$$

where $\Psi_N = \psi_1 + \cdots + \psi_N$.

Proof. Let

$$F(u,v,x) = \sum_{h=1}^{v} \frac{1}{h} \left| \sum_{n=u}^{v} e(ha_n x) \right|.$$

By the Erdős-Turán inequality with $K = N$ we get $D_N(a_n x) \ll F(0,N,x)$. Let u, v be integers such that $0 \le u < v$, then applying the Cauchy-Schwartz inequality gives

$$\begin{aligned}
\int |F(u,v,x)|^2 \, d\mu &\le \sum_{h,k=1}^{v} \frac{1}{hk} \int \left| \sum_{n=u}^{v} e(ha_n x) \right|^2 d\mu \\
&= \sum_{h,k=1}^{v} \frac{1}{hk} \left(v - u + 1 + \sum_{\substack{n,m=u \\ n\neq m}}^{v} \int e(h(a_n - a_m)x) d\mu \right) \\
&= \sum_{h,k=1}^{v} \frac{1}{hk} \left(v - u + 1 + \sum_{\substack{n,m=u \\ n\neq m}}^{v} \hat{\mu}(h(a_n - a_m)) \right).
\end{aligned}$$

Using property *(ii)* of the Kaufman measure, we now get

$$\begin{aligned}
\int |F(u,v,x)|^2 \, d\mu &\ll \sum_{h,k=1}^{v} \frac{1}{hk} \left(v - u + 1 + Ch^{-\eta} \sum_{\substack{n,m=u \\ n\neq m}}^{v} |a_n - a_m|^{-\eta} \right) \\
&\ll \sum_{n=u}^{v} [\log(n)^2 + \psi_n].
\end{aligned}$$

We also see that $F(n-1,n,x) \ll \log(n)^2 + \psi_n$ for all $n \ge 1$ so the theorem now follows from Lemma 5.3. $\square$

Assuming nothing about the sequence (a_n) we get the following corollary.

Corollary 5.5. *Let μ be a Kaufman measure. For μ-almost every $x \in [0,1]$ we have $D_N(a_n x) \ll N^{1-\nu}$ for some $\nu > 0$. In particular $(a_n x)$ is uniformly distributed modulo 1.*

Proof. Without loss of generality we may assume that $a_n < a_{n+1}$ for all n and get the following estimate,

$$\sum_{n,m=u}^{v} |a_n - a_m|^{-\eta} \leq 2 \sum_{m=u}^{v-1} \sum_{n=m+1}^{v} |n-m|^{-\eta} \ll v^{2-\eta} - u^{2-\eta} \ll \frac{1}{\log v} \sum_{n=u}^{v} n^{1-2\nu'}$$

for some $\nu' > 0$, so for μ-a.e. x we have

$$D_N(a_n x) \ll (N \log(N)^2 + N^{2-2\nu'})^{1/2} (\log(N \log(N)^2 + N^{2-2\nu'}))^{3/2+\varepsilon} + N^{1-2\nu'} \ll N^{1-\nu}$$

for any $\nu > 0$ with $\nu < \nu'$. □

If we assume the sequence (a_n) to be lacunary, that is $\frac{a_{n+1}}{a_n} \geq \lambda > 1$ for all n, then we get the following corollary.

Corollary 5.6. *Let $\nu > 0$ and let μ be a Kaufman measure and (a_n) a lacunary sequence of integers. For μ-almost every $x \in [0,1]$ we have $D_N(a_n x) \ll N^{1/2} (\log N)^{5/2+\nu}$.*

Proof. We apply again Theorem 5.4. Using lacunarity of the sequence (a_n), we see that

$$\sum_{n,m=1}^{\infty} |a_n - a_m|^{-\eta} < \infty.$$

Consequently, we can absorb all occurrences of Ψ_N as well as the final term $\max_{n \leq N} \psi_n$ in the discrepancy estimate of Theorem 5.4 into the constant in the Vinogradov symbol. It follows that

$$D_N(a_n x) \ll (N \log(N)^2)^{1/2} \log(N \log(N)^2)^{3/2+\varepsilon} \ll N^{1/2} (\log N)^{5/2+\nu}$$

for μ-almost every x, where ν can be made as small as desired by picking ε small enough. □

5.4 Proof

We are now ready to prove the main theorem, Theorem 5.1. Let $\varepsilon > 0, \{x_i\} \subset \text{Bad}$ be a countable set, and let $\{\mathcal{D}\}$ be a countable set of pseudo-absolute value sequences. Let G be the set of $y \in \text{Bad}$ from Theorem 5.1 and assume for contradiction that $\dim_H G < 1$. By Jarník's theorem [11] we pick M such that $\dim_H F_M > \dim_H G$, and let $\mu = \mu_M$ be the corresponding Kaufman measure supported on F_M.

Pick one of the x_i, and let (n_k) be the sequence of the denominators of the convergents of the continued fraction expansion of x_i. Now this sequence grows faster than the Fibonacci sequence, in particular it is lacunary, and by Corollary 5.6 we get

$$D_N(n_k z) \ll N^{1/2}(\log N)^{5/2+\nu}$$

for almost every z.

Let $\psi(N) = N^{-1/2+\varepsilon}$ and consider the interval

$$I_N^\gamma = [\gamma - \psi(N), \gamma + \psi(N)].$$

By the definition of discrepancy we get that for every $\gamma \in [0, 1]$ and μ-almost every $y \in \mathbb{R}$ that

$$|\#\{k \le N : \{n_k y\} \in I_N^\gamma\} - 2N\psi(N)| \ll N^{1/2}(\log N)^{5/2+\nu}$$

and hence

$$\begin{aligned}\#\{k \le N : \{n_k y\} \in I_N^\gamma\} &\ge 2N\psi(N) - KN^{1/2}(\log N)^{5/2+\nu} \\ &= 2N^{1/2+\varepsilon} - KN^{1/2}(\log N)^{5/2+\nu}.\end{aligned}$$

where $K > 0$ is the implied constant from Corollary 5.6. Now let

$$N_h^\gamma = \min\{N \in \mathbb{N} : \#\{k \le N : \{n_k y\} \in I_N^\gamma\} = h\}$$

be an increasing sequence of integers. We claim that the sequence $q_{N_h^\gamma}$ satisfies (5.1) for the given x_i and every γ with μ-almost every y. Now since $n_{N_h^\gamma}$ is a convergent of x_i we have

$$n_{N_h^\gamma} \left\| n_{N_h^\gamma} x_i \right\| \le 1,$$

so

$$n_{N_h^\gamma} \left\| n_{N_h^\gamma} x_i \right\| \left\| n_{N_h^\gamma} y - \gamma \right\| \le \left\| n_{N_h^\gamma} y - \gamma \right\| \le (N_h^\gamma)^{-1/2+\varepsilon}.$$

Since $x_i \in \text{Bad}$, the sequence (n_k) is bounded between the Fibonacci sequence and $(2M)^k$ where M is the maximal partial quotient of the continued fraction expansion of x_i. Hence $k \asymp \log n_k$, so

$$n_{N_h^\gamma} \left\| n_{N_h^\gamma} x_i \right\| \left\| n_{N_h^\gamma} y - \gamma \right\| \le (\log n_{N_h^\gamma})^{-1/2+\varepsilon/2}$$

for sufficiently large h. This proves the claim, and proves that the set E_i for which (5.1) is not true satisfies $\mu(E_i) = 0$.

Now let $\mathcal{D} = \mathcal{D}_i = (q_k)$ be one of the pseudo-absolute value sequences. Recall that this implies $q_k \mid q_{k+1}$. Assume that

$$\liminf_n n\,|n|_{\mathcal{D}}\,\|ny - \gamma\| = \delta$$

for some $\delta > 0$. In particular this implies that

$$q_k\,|q_k|_{\mathcal{D}}\,\|q_k y - \gamma\| \geq \delta,$$

for some k, but $q_k\,|q_k|_{\mathcal{D}} = 1$, so

$$\|q_k y - \gamma\| \geq \delta.$$

From Corollary 5.5 we see that $\{q_k y\}$ is uniformly distributed for μ-almost all y, which the above contradicts. So the set E_i' of $y \in \text{Bad}$ for which (5.2) is not true must satisfy $\mu(E_i') = 0$.

Furthermore, assume that $q_k \leq C^k$ for some $C > 1$. Then $q_k \mid q_{k+1}$, so (q_k) is lacunary, and the upper bound implies that $k \asymp \log q_k$. So we can repeat the argument from before to get that (5.3) is true for μ-almost all y.

Now let E be the set of $y \in \text{Bad}$ such that there is an i or j such that (5.1) and (5.2) is false. Then

$$E = \bigcup_i E_i \cup \bigcup_j E_j',$$

and hence $\mu(E) = 0$.

Now $\mu(G)$ is maximal. Pick s such that $\dim_H G < s < \dim_H F_M$. Recall that $\mu(I) \leq c\,|I|^s$. Now suppose that $\{U_i\}$ is a cover of G. We may assume that all U_i are intervals. Then

$$\sum_i |U_i| \geq \frac{1}{c}\sum_i \mu(U_i) \geq \frac{1}{c}\mu\left(\bigcup_i U_i\right) \geq \frac{1}{c}\mu(G) > 0,$$

so $\mathscr{H}^s(G) > 0$ and hence

$$\dim_H G \geq s,$$

which is a contradiction. This proves that $\dim_H G = 1$.

Appendix A

On lower bounded orbits of the times-q map

Abstract

In this paper we consider the times-q map on the unit interval as a subshift of finite type by identifying each number with its base q expansion, and we study certain non-dense orbits of this system where no element of the orbit is smaller than some fixed parameter c.

The Hausdorff dimension of these orbits can be calculated using the spectral radius of the transition matrix of the corresponding subshift, and using simple methods based on Euclidean division in the integers, we completely characterize the characteristic polynomials of these matrices as well as give the value of the spectral radius for certain values of c. It is known through work of Urbanski and Nilsson that the Hausdorff dimension of the orbits mentioned above as a map of c is continuous and constant almost everywhere, and as a new result we give some asymptotic results on how this map behaves as $q \to \infty$.

A.1 Introduction

In this paper we study the set

$$F_c^q = \{x \in [0, 1) \mid q^n x \geq c \text{ for all } n \geq 0\}$$

where $q \geq 2$ is an integer. This set is related to badly approximable numbers in Diophantine approximation, and has been studied by Nilsson [17], who studied the Hausdorff dimension of the set as a map of c, and in more generality by Urbanski [24] who considered the orbit of an expanding map on the circle.

As Nilsson did we will consider F_c^q as a subshift of finite type which enables us to see it as a problem in dynamical systems. When studied as a subshift of finite type we can find the dimension of F_c^q using the spectral radius of the corresponding transition matrix, and this motivates the theorem of this paper which characterizes the characteristic polynomial of this matrix.

The author would like to his PhD supervisor Simon Kristensen and he would also like to than Johan Nilsson for reading and commenting on this paper.

A.2 Basic definitions

We begin with a definition of part and residue which comes from elementary integer division with residue. We let $q \geq 2$ be an integer throughout the paper and start with a well known result.

Proposition A.1. *For integers $n \in \mathbb{N}$ and $m \geq 0$ there are unique integers $\langle n, m\rangle \in \mathbb{N}$* (part) *and $0 \leq [n, m] < q^m$* (residue) *such that*

$$n = q^m \langle n, m\rangle + [n, m].$$

We note that if we write $n = n_k \cdots n_1$ in base q it is easy to find the part and the residue, since $[n, m] = n_m \cdots n_1$ and $\langle n, m\rangle = n_k \cdots n_{m+1}$.

The matrix we will consider in this paper is defined as follows.

Definition A.2. For $m \geq 1$ we define a 0-1 matrix A_m of size $q^m \times q^m$ by

$$(A_m)_{ij} = 1 \iff [i-1, m-1] = \langle j-1, 1\rangle.$$

We let $A_m(P)$ with $P \subseteq \{1, 2, \ldots, q^m\}$ be the $\#P \times \#P$ matrix made from picking only the rows and columns from A_m corresponding to the elements in P and for $0 \leq k \leq m$ we let $A_m(k)$ be the $m-k \times m-k$ matrix where we have removed the first k rows and columns from A_m.

We will often omit the dependency on m when it is not confusing. Considering i and j in base q we see that $(A_m)_{ij} = 1$ if and only if the first $m-1$ digits of $j-1$ are equal to the last $m-1$ digits of $i-1$. So when $c = \frac{i}{q^m}$ we see that the base q^m expansions of the numbers in F_c^q can be seen as a subshift of finite type with transition matrix $A_m(i)^m$. The metric of the subshift and the unit interval are equivalent so the dimensional properties are the same. In particular, finding the Hausdorff dimension of F_c^q now boils down to finding the spectral radius $\rho(A_m(k))$, since

$$\dim_H F(c) = \frac{\rho(A_m(i)^m)}{\log q^m} = \frac{\rho(A_m(i))}{\log q}. \tag{A.1}$$

For a proof of the first equality see [19]. This is why we were interested in finding the characteristic polynomials of $A_m(i)$. The main theorem of this paper is a complete characterization of these polynomials, and to formulate this theorem we need the following definition.

Definition A.3. For $n, m \geq 1$ with $0 \leq n < q^m$ we define

$$l_m(n) = \min\{1 \leq j \leq m \mid \langle n, j\rangle \geq [n, m-j]\}.$$

Using this definition we let

$$\overline{n}_m = n - [n, m - l_m(n)] = q^{m-l_m(n)}\langle n, m - l_m(n)\rangle$$

be the *minimal prefix* of n.

This is well defined since $[n, 0] = \langle n, m\rangle = 0$ for any n with $0 \leq n < q^m$. The notion of minimal prefix is taken from Nilsson [17], but is here defined somewhat different since we only consider finite sequences.

Let us consider some examples.

Example A.4. Let $q = 3, m = 3$. Then

$$\langle 11, 1\rangle = 3 \geq 2 = [11, 2]$$

so $l_3(11) = 1$ and

$$\overline{11}_3 = 11 - [11, 2] = 9.$$

If we let $n = 7$ we have

$$\langle 7, 1\rangle = 2 < 7 = [7, 2]$$

and

$$\langle 7, 2\rangle = 0 < 1 = [7, 1]$$

but

$$\langle 7, 3\rangle = 0 = [7, 0]$$

so $l_3(7) = 3$ and $\overline{7}_3 = 7$.

We are now ready to state the main theorem.

Theorem A.5. *Let $0 < k < q^m$ and let $f_k^m(x)$ be the characteristic polynomial of $A_m(k)$. Then*

$$f_k^m(x) = g_k^m(x)x^{q^m-m-k}$$

where

$$g_k^m(x) = x^m - a_1x^{m-1} - \cdots - a_m$$

and $a_1a_2\ldots a_m$ is the base q expansion of $q^m - \overline{k}_m$.

Notice that this implies the nice equality

$$g_k^m(q) = \overline{k}_m.$$

A.3 Proof outline

First recall that we can find the characteristic polynomial $f(x) = x^{q^m-k} - a_1 x^{q^m-k-1} - \cdots - a_{q^m-k}$ of $A_m(k)$ as

$$a_i = (-1)^i \sum_{\#P=i} \det A_m(P) \tag{A.2}$$

where we also require that $P \subseteq \{k+1, k+2, \ldots, q^m\}$, or as

$$a_i = \frac{1}{i}\big(\operatorname{trace} A_m(k)^i + a_1 \operatorname{trace} A_m(k)^{i-1} + \cdots + a_{i-1} \operatorname{trace} A_m(k)\big). \tag{A.3}$$

The first formula is sometimes used as the definition of the characteristic polynomial, and for a proof of the latter see [6]. We now try to outline the proof that essentially is the construction of an algorithm that calculates both the characteristic polynomial of $A_m(i)$ and $\bar{i}_m$.

- We prove that all the submatrices $A(P)$ that gives non-zero principal minors are permutations, so when removing rows and columns from the first to the last, we only change the characteristic polynomial when removing rows and columns corresponding to the smallest element of a cycle.
- If $l_m(i) = m$ then i is the smallest element of an m-cycle and this is the only permutation of size $\leq m$ that has i as an element. So removing i decreases the m'th coefficient of the characteristic polynomial by 1 and leaves all the preceding coefficients unchanged. On the other hand, if $l_m(i) = n < m$, then the nontrivial part of the characteristic polynomial, $g_i^m(x)$, can be found as $x^{m-n} g^n_{\langle i, m-n\rangle}(x)$ since we have (A.3) and can prove that

  $$\operatorname{trace} A_m(i)^k = \operatorname{trace} A_n(\langle i, m-n\rangle)^k$$

  for all $k \leq m$.
- If $l_m(i) = m$, then $\bar{i}_m = \overline{i+1}_m - 1$, and if $l_m(i) = n < m$ then $\bar{i}_m = q^{m-n}\overline{\langle i, m-n\rangle}_n$, so we see that $\bar{i}$ and the characteristic polynomials follow the same pattern.
- Since the theorem is true for $m = 1$, we can now use induction if $l_m(i) < m$. If not, we increase i until we have $l_m(i) < n$, which happens since $l_m(q^m - 1) = 1$.
- The $m+1$'st, $m+2$'nd, $\ldots$, q^m'th coefficient of $f_k^m(x)$ are all zero, because we have found the first M coefficients of the characteristic

polynomial for any M, so we pick K such that $l_M(K) = m$ and $\langle K, M-m\rangle = k$, then we see that $g_K^M(x)$ has its $m+1$'th, $m+2$'th, ..., M'th coefficients equal to zero, which will then also be true for $g_k^m(x)$. This finishes the proof of the theorem.

A.4 Part and residue

The results in this sections explain some properties of the part and residue functions and gives a characterization of the powers of A. We will use these results throughout the paper, often without specifically stating so. The proofs in this section are rather straightforward and may be skipped on a first read.

Proposition A.6. *1. For $j, k, n \geq 0$ we have $[[n, j], k] = [n, \min\{j, k\}]$ and*

$$\big\langle \langle n, k\rangle, j\big\rangle = \langle n, k+j\rangle.$$

2. For $j > k$ we have

$$\langle [n, j], k\rangle = [\langle n, k\rangle, j-k].$$

Proof. Let us first prove the two equalities in 1. Since $[n, k]$ is the same as $n \pmod q)^k$ we have the first equality. Now assume that $j + k \leq m$. Now $\langle n, k\rangle = q^j \big\langle \langle n, k\rangle, j\big\rangle + [\langle n, k\rangle, j]$, so

$$n = q^k \langle n, k\rangle + [n, k] = q^{k+j} \big\langle \langle n, k\rangle, j\big\rangle + q^k [\langle n, k\rangle, j] + [n, k],$$

but since $[\langle n, k\rangle, j] < q^j$ and $[n, k] < q^k$ we have

$$q^k [\langle n, k\rangle, j] + [n, k] \leq q^k (q^j - 1) + q^k - 1 = q^{k+j} - 1 < q^{k+j},$$

and by the uniqueness of the residue and parts we see that $\langle \langle n, k\rangle, j\rangle = \langle n, k+j\rangle$. Now consider 2., so let $j > k$. From 1. we have

$$\langle n, k\rangle = q^{j-k} \big\langle \langle n, k\rangle, j-k\big\rangle + [\langle n, k\rangle, j-k] = q^{j-k} \langle n, j\rangle + [\langle n, k\rangle, j-k]$$

and

$$[n, j] = q^k \langle [n, j], k\rangle + \big[[n, j], k\big] = q^k \langle [n, j], k\rangle + [n, k].$$

So

$$\begin{aligned} n &= q^k \langle n, k\rangle + [n, k] \\ &= q^j \langle n, j\rangle + q^k [\langle n, k\rangle, j-k] - q^k \langle [n, j], k\rangle + [n, j] \\ &= q^j \langle n, j\rangle + [n, j] + q^k ([\langle n, k\rangle, j-k] - \langle [n, j], k\rangle) \end{aligned}$$

Table A.1: Calculation of the characteristic polynomials of $A_3(i)$ when $q = 3$. We let $g_i(x) = x^3 - a_1x^2 - a_2x - a_3$. We also give the minimal prefix and the length of the minimal prefix. The numbers in bold indicates that we consider a minimal number with non-maximal prefix length $l(i) < 3$.

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
$\bar{i}$	0	1	2	3	4	5	6	7	8	9	9	9	9	9	14	15	15	17	18	18	18	18	18	18	18	18	18
a_1	3	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	0
a_2	0	2	2	2	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0
a_3	0	2	1	0	2	1	0	2	1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
$l(i)$	1	3	3	2	3	3	2	3	3	1	1	1	1	1	3	2	2	3	1	1	1	1	1	1	1	1	1
$A_{i+1,i+1}$	**1**	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	1
$A^2_{i+1,i+1}$	1	0	0	**1**	0	0	**1**	0	0	0	1	0	0	1	0	0	**1**	0	0	0	1	0	0	1	0	0	**1**

and since $n = q^j \langle n, j\rangle + [n, j]$ this implies that

$$[\langle n, k\rangle, j-k] = \langle [n, j], k\rangle.$$

$\square$

Lemma A.7. *Let* $1 \le k \le m$. *Then* $A^k_{ij} = 1$ *if and only if*

$$[i-1, m-k] = \langle j-1, k\rangle.$$

Proof. We will prove this by induction. For $k = 1$ it is the definition of A, so assume that $1 < k \le m$. We assume that the lemma is true for all smaller k. If $A^k_{ij} = 1$ there must exist some n with $0 \le n < q^m$ and $A_{nj} = 1$ and $A^{k-1}_{in} = 1$. Using the induction hypothesis we get

$$[i-1, m-k+1] = \langle n-1, k-1\rangle \quad \text{and} \quad [n-1, m-1] = \langle j-1, 1\rangle \tag{A.4}$$

for this n. Now by part 2. of the above proposition we have

$$[\langle n-1, k-1\rangle, m-k] = \langle [n-1, m-1], k-1\rangle,$$

and using (A.4) we get

$$\big[[i-1, m-k+1], m-k\big] = \big\langle \langle j-1, 1\rangle, k-1\big\rangle,$$

and using part 1. of the proposition we get

$$[i-1, m-k] = \langle j-1, k\rangle$$

as desired.

Now assume that $[i-1, m-k] = \langle j-1, k\rangle$. Let

$$n - 1 = q^{k-1}[i-1, m-k+1] + [\langle j-1, 1\rangle, k-1].$$

This is a positive integer smaller than q^m. By the uniqueness of the residue and parts we see that

$$[i-1, m-k+1] = \langle n-1, k-1\rangle \tag{A.5}$$

and

$$[\langle j-1, 1\rangle, k-1] = \langle n-1, k-1\rangle. \tag{A.6}$$

From (A.5) and the induction hypothesis we see that $A^{k-1}_{in} = 1$. We now want to prove that $A_{nj} = 1$. Recall that we assume $[i-1, m-k] = \langle j-1, k\rangle$, so

$$\begin{aligned}\langle [n-1, m-1], k-1\rangle &= [\langle n-1, k-1\rangle, m-k] \\ &= \big[[i-1, m-k+1], m-k\big] \\ &= [i-1, m-k] \\ &= \langle j-1, k\rangle.\end{aligned}$$

Using this and (A.6) we see that

$$\begin{aligned}
[n-1,m-1] &= q^{k-1}\langle [n-1,m-1],k-1\rangle + \big[[n-1,m-1],k-1\big] \\
&= q^{k-1}\langle j-1,k\rangle + [n-1,k-1] \\
&= q^{k-1}\langle j-1,k\rangle + [\langle j-1,1\rangle,k-1] \\
&= q^{k-1}\big\langle \langle j-1,1\rangle,k-1\big\rangle + [\langle j-1,1\rangle,k-1] \\
&= \langle j-1,1\rangle.
\end{aligned}$$

This proves that $A_{in}^{k-1}=1$ and $A_{nj}=1$ which implies that $A_{ij}^k>0$. Now assume that there is another n' such that $A_{in'}^{k-1}=1$ and $A_{n'j}=1$. Then

$$[i-1,m-k+1] = \langle n'-1,k-1\rangle$$

and

$$[\langle j-1,1\rangle,k-1] = \langle n'-1,k-1\rangle$$

so

$$\begin{aligned}
n'-1 &= q^{k-1}\langle n'-1,k-1\rangle + [n'-1,k-1] \\
&= q^{k-1}[i-1,m-k+1] + \big[[n'-1,m-1],k-1\big] \\
&= q^{k-1}[i-1,m-k+1] + [\langle j-1,1\rangle,k-1] \\
&= n-1,
\end{aligned}$$

which proves that there can be only one such n, so $A_{ij}^k=1$. $\square$

Lemma A.8. *If a,b,k is such that $[a,k]<[b,k]$ and $\langle a,k\rangle = \langle b,k\rangle$, then*

$$[a,k+j]<[b,k+j]$$

for all $0\le j\le m-k$.

Proof. If $\langle a,k\rangle=\langle b,k\rangle$ then

$$\big\langle\langle a,k\rangle,j\big\rangle = \big\langle\langle b,k\rangle,j\big\rangle,$$

and hence

$$\langle a,k+j\rangle = \langle b,k+j\rangle.$$

Since $a<b$ we thus have

$$[a,k+j]<[b,k+j]$$

as desired. $\square$

A.5 Minimality

We now prove the following rather simple lemma which states that the only non-zero principal minors can be found as submatrices of A who are permutations.

Lemma A.9. *If* $\det A(P) \neq 0$ *then the corresponding matrix is a permutation matrix.*

Proof. Assume that we choose P such that one of the rows of $A(P)$ has two ones. In other words there are $i, j_1, j_2 \in P$ such that

$$A_{ij_1} = A_{ij_2} = 1.$$

Using the definition of A this implies that

$$\langle j_1 - 1, 1\rangle = [i-1, m-1] = \langle j_2 - 1, 1\rangle.$$

Now let $k \in P$ be arbitrary. Then $A_{kj_1} = 1$ if and only if $[k-1, m-1] = \langle j_1 - 1, 1\rangle$, which is true if and only if

$$[k-1, m-1] = \langle j_2 - 1, 1\rangle,$$

so $A_{kj_1} = A_{kj_2}$ for all $k \in P$, so the j_1'th and j_2'nd column are equal and so $\det A(P) = 0$. The proof is similar when we assume that there are two ones in one column. □

Recall that if $A(P)$ is a permutation, then $P = P_1 \cup \cdots \cup P_n$ where $\cap_i P_i = \emptyset$ and $A(P_i)$'s are all cycles. This motivates the following two theorems, where we characterize the subsets P where $A(P)$ is a cycle. We are interested in the smallest elements of cycles, since the whole cycle are removed when we remove this element, which we will prove is exactly the numbers that are minimal.

Definition A.10. We say that $0 \leq n \leq q^m$ is m-minimal if

$$A^{l(n)}_{n+1,n+1} = 1,$$

or equivalently using Lemma A.7 if

$$[n, m - l(n)] = \langle n, l(n)\rangle.$$

Theorem A.11. *Let* $P \subset \{1, 2, \ldots, q^m\}$ *be such that* $A(P)$ *is a* k*-cycle for some* $1 \leq k \leq m$*. Then* $\min P - 1$ *is minimal with* $l_m(\min P - 1) = k$*.*

Proof. Let $P = \{i_1, i_2, \dots, i_k\}$ be a k-cycle with $A^j_{i_1 i_{j+1}} = 1$ for $1 \le j < k$ and $A^k_{i_1 i_1} = 1$. Without loss of generality we can assume that $\min P = i_1$. Using Lemma A.7 we get that

$$[i_1 - 1, m - j] = \langle i_{j+1} - 1, j\rangle,$$

for $1 \le j < k$ and

$$[i_1 - 1, m - k] = \langle i_1 - 1, k\rangle$$

so we need to prove that $\langle i_{j+1} - 1, j\rangle > \langle i_1 - 1, j\rangle$ for $j = 1, 2, k-1$. We have the non-strict inequality since $i_1 < i_j$. So assume for contradiction that

$$\langle i_1 - 1, j\rangle = \langle i_{j+1} - 1, j\rangle.$$

Now since $i_1 < i_{j+1}$ we have

$$[i_1 - 1, j] < [i_{j+1} - 1, j],$$

and due to Lemma A.8 we have

$$[i_1 - 1, m - k + j] < [i_{j+1} - 1, m - k + j] \tag{A.7}$$

since $k \le m$. Since $A^{k-j}_{i_{j+1} i_1} = 1$ we have $[i_{j+1} - 1, m - k + j] = \langle i_1 - 1, k - j\rangle$. Using (A.7) we get

$$[i_1 - 1, m - k + j] < \langle i_1 - 1, k - j\rangle.$$

Now consider i_{k-j+1}. Since $j < k$ we have $A^{k-j}_{i_1 i_{k-j+1}} = 1$ so

$$[i_1 - 1, m - k + j] = \langle i_{k-j+1} - 1, k - j\rangle,$$

and hence

$$\langle i_{k-j+1} - 1, k - j\rangle < \langle i_1 - 1, k - j\rangle.$$

This implies that $i_{k-j+1} < i_1$ which is a contradiction against i_1 being the least element in P. $\square$

Theorem A.12. *Assume that $i - 1$ is minimal. Then there is a unique $P \subseteq \{1, 2, \dots, q^m\}$ such that $\min P = i$ and $A(P)$ is a $l(i-1)$-cycle.*

Proof. We let $P = \{i, i_2, i_3, \dots, i_k\}$ where

$$\begin{aligned} i_2 - 1 &= q[i - 1, m - 1] + \langle i - 1, m - 1\rangle \\ i_3 - 1 &= q^2[i - 1, m - 2] + \langle i - 1, m - 2\rangle \\ &\vdots \\ i_k - 1 &= q^{k-1}[i - 1, m - k + 1] + \langle i - 1, m - k + 1\rangle. \end{aligned}$$

We now need to prove that $A^{n-1}_{ii_n} = 1$ and that $i < i_n$ for all $n = 2, 3, \ldots, k$. Using the uniqueness of the part and residue we see that

$$\langle i_n - 1, n - 1\rangle = [i - 1, m - n + 1]$$

and

$$[i_n - 1, n - 1] = \langle i - 1, m - n + 1\rangle$$

for $n = 2, 3, \ldots, k$. The first of these equations implies that $A^{n-1}_{ii_n} = 1$.

Since $l_m(i - 1) = k$ we know that

$$\langle i - 1, n\rangle < [i - 1, m - n]$$

for $n = 1, 2, \ldots, k - 1$. This implies that

$$i_{n+1} - 1 = q^n[i - 1, m - n] + \langle i - 1, m - n\rangle > q^n\langle i - 1, n\rangle + [i - 1, n] = i - 1$$

since both $\langle i - 1, m - n\rangle$ and $[i - 1, n]$ are smaller than q^n.

We now need to prove that this P is unique. Assume that we have $P' = \{i, i'_2, \ldots, i'_k\}$, where we order the elements such that $A^{n-1}_{ii'_n} = 1$. This implies that

$$[i - 1, m - n + 1] = \langle i'_n - 1, n - 1\rangle$$

for all $n = 2, 3, \ldots, k$. Since $A(P)$ is a k-cycle, we furthermore know that $A^{k-n+1}_{i'_n i} = 1$, so

$$[i'_n - 1, m - k + n - 1] = \langle i - 1, k - n + 1\rangle.$$

Now we want to prove that $i'_n = i_n$, so let $2 \leq n \leq k$ be given. We have

$$i'_n - 1 = q^{n-1}\langle i'_n - 1, n - 1\rangle + [i'_n - 1, n - 1]$$

and $\langle i'_n - 1, n - 1\rangle = [i - 1, m - n + 1]$, so we just need to prove that

$$[i'_n - 1, n - 1] = \langle i - 1, m - n + 1\rangle.$$

We have

$$\begin{aligned}
[i'_n - 1, n - 1] &= \big[[i'_n - 1, m - k + n - 1], n - 1\big] \\
&= [\langle i - 1, k - n + 1\rangle, n - 1] \\
&= \big[[i_n - 1, m - k + n - 1], n - 1\big] \\
&= [i_n - 1, n - 1] \\
&= \langle i - 1, m - n + 1\rangle
\end{aligned}$$

so $i_n = i'_n$ for all n, and so $P = P'$. $\square$

Corollary A.13. *If* $l_m(i-1) = m$ *then there is exactly one* $P \subseteq \{1, 2, \ldots, q^m\}$ *such that* $\min P = i$ *and* $A(P)$ *is a* m*-cycle.*

Proof. This follows from the fact that $A^m_{ij} = 1$ for all i, j. In particular we have $A^m_{ii} = 1$ for all i. □

Now compare this corollary with the following lemma.

Lemma A.14. *If* $l_m(i-1) = m$*, then* $\bar{i}_m = \overline{i-1}_m + 1$.

Proof. It is enough to prove that $\bar{i} = i$, since we certainly have $\overline{i-1} = i-1$. Using the definition we see that this is equivalent with $[i, m-l(i)] = 0$. If $l(i) = m$ we are done, so assume that $l(i) < m$. Now either $[i, m-l(i)] = 0$, in which case we are done, or $[i, m-l(i)] = [i-1, m-l(i)] + 1$. Now since $l(i-1) = m$ we have

$$[i-1, m-l(i)] < \langle i-1, l(i) \rangle,$$

since $l(i) < m = l(i-1)$, but

$$[i-1, m-l(i)] = [i, m-l(i)] - 1 \leq \langle i, l(i) \rangle - 1 \leq \langle i-1, m-l(i) \rangle,$$

which is a contradiction. □

Recalling the idea of the proof we here see that if $l_m(i-1) = m$ and we remove the i'th row and column of A_m, then we remove exactly one permutation of size $\leq m$, namely a m-cycle, which increases the m'th coefficient of the characteristic polynomial by one, and we also see that it increases the m'th digit of the base q expansion of $\bar{i}$ by one.

A.6 Induction mapping

In the following chapter we will no longer suppress the dependency on m, since we are interested in mapping permutations between matrices of different sizes while preserving cycles. We will illustrate the idea with an example. If $q = 3$, and we write all numbers in base 3 we see that

$$012, 120, 201 \tag{A.8}$$

is a 3-cycle in $A_3(012)$. We now map this up to

$$0120, 1201, 2012$$

which is a 3-cycle in $A_4(0120)$. On the other hand we could also map (A.8) down to

$$01, 12, 20$$

which is a 3-permutation in $A_2(01)$. In this section we will formally define these maps, and also prove that they map cycles to cycles. We begin with the 'down' map which is defined in the following way.

Definition A.15. If $0 \le i < q^{m+1}$ then we define

$$D_m(i) = \langle i, 1\rangle.$$

For $M > m$ and $0 \le i \le q^M$ we let

$$D_{m,M}(i) = D_m \circ \cdots \circ D_{M-1}(i) = \langle i, M-m\rangle.$$

We now prove the following lemma.

Lemma A.16. *If $l_M(i) = m < M$ we have*

$$l_m(D_{m,M}(i)) = m.$$

Proof. We have $[i, M-m] \ge \langle i, m\rangle$ and $[i, M-j] < \langle i, j\rangle$ for all $1 \le j < m$, and we need to prove that $[i, m-j] < \langle i, j\rangle$ for all $1 \le j \le m$. But this is clearly the case since $m < M$, so

$$[i, m-j] < [i, M-j] < \langle i, j\rangle$$

for all $1 \le j < m$. □

Corollary A.17. *Let $0 \le i < q^M$. If $l_M(i) = m < M$, then*

$$\bar{i}_M = q^{M-m}\overline{D_{m,M}(i)}_m.$$

Proof. This follows from the definition of the minimal prefix. □

We saw earlier that the characteristic polynomial of a matrix can be found by considering the trace of the powers of the matrix. So if we can map permutations bijectively between two transition matrices we must have the same characteristic polynomials. As before we only need to consider cycles as all permutations are products of cycles.

Definition A.18. An ordered k-tuple of distinct elements, $(i_1, \ldots, i_k)$ with $0 \le i_j \le q^m$ for all $j = 1, 2, \ldots, k$ is a k-cycle in $A_m(c)$ if $A_m(c)_{i_j, i_{j+1}} = 1$ for all $j = 1, 2, \ldots, k-1$, and $A_m(c)_{i_k, i_1} = 1$. In other words, if we have

$$[i_j, m-1] = \langle i_j, 1\rangle$$

for $j = 1, 2, \ldots, k-1$ and $[i_k, m-1] = \langle i_1, 1\rangle$ and $i_j \ge c$ for all $j = 1, 2, \ldots, k$.

We have a 'down' map, mapping from large matrices to smaller and we now define an 'up' map, mapping from smaller to larger.

Definition A.19. Let $P = (i_1, \ldots, i_k)$ be a k-cycle in $A_m(c)$. Then we let
$$U_m(P) = (qi_1 + [i_2, 1], \cdots, qi_k + [i_1, 1]),$$
and for $M > m$ we let $U_{m,M} = U_{M-1} \circ U_{M-2} \circ \cdots \circ U_m$.

Lemma A.20. *Let $m = l_M(c)$ and let $P = (i_1, i_2, \ldots, i_k)$ be a k-cycle in $A_M(c)$. Then*
$$D_{m,M}(P) = (D_{m,M}(i_1), \cdots, D_{m,M}(i_k))$$
is a k-cycle in $A_m(D_{m,M}(c))$. Furthermore, if $Q = (j_1, \ldots, j_k)$ is a k-cycle in $A_m(D_{m,M}(c))$, then $U_{m,M}(Q)$ is a k-cycle in $A_M(c)$.

Proof. To prove that $D_{m,M}(P)$ is a k-cycle in $A_m(D_{m,M}(c))$ can be done by straightforward calculations. We also get that $U_{m,M}(Q)$ is a k-cycle in $A_M(q^{M-m}\langle c, M-m\rangle)$ rather straightforward. The problem is to prove that it actually is a k-cycle in $A_M(c)$, or in other words that there are no k-cycles with its smallest element in the interval between $q^{M-m}\langle c, M-m\rangle$ and c. Recalling the definition of $\overline{c}_M$ and that the least element of a cycle always is minimal we thus need to prove that if we have $\overline{c}_M \leq n < c$, then n cannot be minimal.

We get that $\overline{n}_M = \overline{c}_M$ and $l_M(n) = l_M(c)$ so
$$[c, M-m] - [n, M-m] = c - n$$
so if we assume that n is minimal we get
$$\langle c, m\rangle \geq [c, M-m] = [n, M-m] + c - n = \langle n, m\rangle + c - n$$
which is a contradiction. This finishes the proof of the theorem. □

These two lemmas now lead to the following theorem regarding the invariance of the traces.

Theorem A.21. *Let $m, k \leq M$. Then*
$$\operatorname{trace} A_m(c)^k = \operatorname{trace} A_M(q^{M-m}c)^k.$$
More generally we have
$$\operatorname{trace} A_m(\langle c, M-m\rangle)^k = \operatorname{trace} A_M(c)^k$$
whenever $l_M(c) \geq m$.

Proof. Each k-cycle contributes to the trace, and since the maps used in the lemmas maps all k-cycles invectively, we get the theorem. □

Newton's formula for the characteristic polynomial gives us, that if

$$f_c^m(x) = x^n - a_1 x^{n-1} - \cdots - a_n = \det(xI - A_m(k))$$

is the characteristic polynomial of $A_m(k)$ where $n = q^m - k$, then

$$a_k = \frac{1}{k}\big(\operatorname{trace} A_m(c)^k - a_1 \operatorname{trace} A_m(c)^{k-1} - \cdots - a_{t-1} \operatorname{trace} A_m(c)\big)$$

so the above theorem gives us that

$$f_c^M(x) = x^{M-m} f^m_{q^{M-m}c}(x).$$

Combining this with the simple lemma below gives us the proof of the main theorem.

Lemma A.22. *Let $0 \le n < q^m$. Then*

$$q\overline{n}_m = \overline{qn}_{m+1}.$$

Proof. We see that

$$q\overline{n}_m = q(n - [n, m - l_m(n)]) = qn - [qn, m + 1 - l_m(n)],$$

so we just need to prove that $l_{m+1}(qn) = l_m(n)$. Assume that $j = l_m(n)$. Then

$$\langle qn, j\rangle \ge q\big\langle\langle qn, j\rangle, 1\big\rangle = q\langle qn, j+1\rangle = q\langle n, j\rangle \ge q[n, m-j] = [qn, m].$$

Now assume that $\langle qn, j\rangle \ge [qn, m+1-j]$ for some $j > l_m(n)$. Then

$$q[n, j] = [qn, j] \le \langle qn, m+1-j\rangle$$

so

$$[n, j] \le \langle\langle qn, m+1-j\rangle, 1\rangle = \langle n, m-j\rangle$$

which is a contradiction. □

We are now ready to prove the main theorem, so let us restate it.

Theorem A.23. *Let $1 \le k \le q^m$ and let $f_k^m(x)$ be the characteristic polynomial of $A_m(k)$. Then*

$$f_k^m(x) = g_k^m(x) x^{q^m - m - k}$$

where

$$g_k^m(x) = x^m - a_1 x^{m-1} - \cdots - a_m$$

and $a_1 a_2 \ldots a_m$ is the base q expansion of $q^m - \overline{k}_m$.

Proof. We prove this theorem using induction. If $m = 1$ it is certainly true since $\bar{i}_1 = i$ for all $0 \leq i < q$ and A_1 is the all one matrix of size $q \times q$.

We see that when choosing m and $i > 0$ we have two possibilities: Either we have $l(i-1) = m$ or $l(i-1) < m$. In the first case removing the i'th column and row only removes one non-zero minor, namely the unique m-cycle with i as its minimal element given in Theorem A.12. In this case we also have that the last digit of $\overline{i-1}_m$ is $[i-1, 1]$ which must be non-zero, so here we just decrease a_m with 1, so the first m coefficients of the characteristic polynomial changes in the right way due to Lemma A.14.

If we have $l(i-1) = n < m$ we see that we can find the characteristic polynomial of the smaller matrix of size q^n instead and multiply it by x^{m-n}. As we see in Corollary A.17 this is also the case for $\bar{k}$. So by induction we are done.

Now we need to prove that the remaining coefficients are all zero. To prove this we once again use Lemma A.21 to see that the M'th coefficient of f_k^m must be equal to the M'th coefficient of $f^M_{q^{M-m}c}$ for any $M > m$. And here we see that the $m+1$'th, $m+2$'th, $\ldots$, and M'th coefficient all are zero, since the M'th digit of the base q expansion of

$$q^M - \overline{q^{M-m}c}_M = q^{M-m}(q^m - \bar{c}_m)$$

is zero. This finishes the proof of the theorem. $\square$

A.7 Constant dimension

Now define $\phi : c \mapsto \dim_H F(c)$. Recall from (A.1) that when c has finite base q expansion we can calculate $\phi(c)$. Nilsson [17] proved that this function is continuous and constant almost everywhere. Using the theorem we see that if we have $0 \leq i < j < q^m$ such that $\bar{i}_m = \bar{j}_m$ then

$$\phi\left(\frac{i}{q^m}\right) = \phi\left(\frac{j}{q^m}\right)$$

and since ϕ is a decreasing function it must be constant on the interval

$$\left[\frac{i}{q^m}, \frac{j}{q^m}\right].$$

Now let $0 \leq i < q$ be given and let

$$j(m) = \sum_{n=1}^{m} iq^{n-1}.$$

We now claim that

$$\overline{q^{m-1}i}_m = \overline{j(m)}_m.$$

To prove this we see that $l_m(q^{m-1}i) = 1$ and so $\overline{q^{m-1}i}_m = q^{m-1}i$. Now $l_m(j(m)) = 1$ and

$$\overline{j(m)}_m = iq^{m-1}$$

which proves the claim. This gives us

$$\phi\left(\frac{i}{q}\right) = \phi\left(\frac{j(m)}{q^m}\right)$$

for all m and letting $m \to \infty$ we get that ϕ is constant on the interval

$$\left[\frac{i}{q}, \frac{i}{q-1}\right].$$

Now letting $m = 1$ we find

$$g_i^1(x) = x - \bar{i}_1 = x - i$$

which has one root, $x = i$, so we get

$$\phi\left(\frac{i}{q}\right) = \frac{\log i}{\log q}$$

on this interval.

A bit more work allows us to calculate $\phi(x)$ for $x = \frac{i}{q^n}$ for larger n since we here need to solve polynomial equations of degree n.

A.8 Numerical plot

Calculating the spectral radii of $A(k)$, we can make numerical plots of the function ϕ. The plot in figure A.1 was made using GNU Octave.

A.9 Asymptotics

We now want to consider ϕ as $q \to \infty$. We consider the function $\psi : [0, 1) \to [0, 1)$ where

$$\psi(c) = \begin{cases} 1 + \frac{\log(1-c)}{\log q} & 0 \le c < \frac{q-1}{q} \\ 0 & \text{otherwise.} \end{cases}$$

and wish to prove that ϕ and ψ are somewhat asymptotically similar. This can also be expressed by saying that $\rho(A_c)$ behaves somewhat like $q - qc$, which is true in the starting point of the intervals where ϕ is constant, so we get the following theorem.

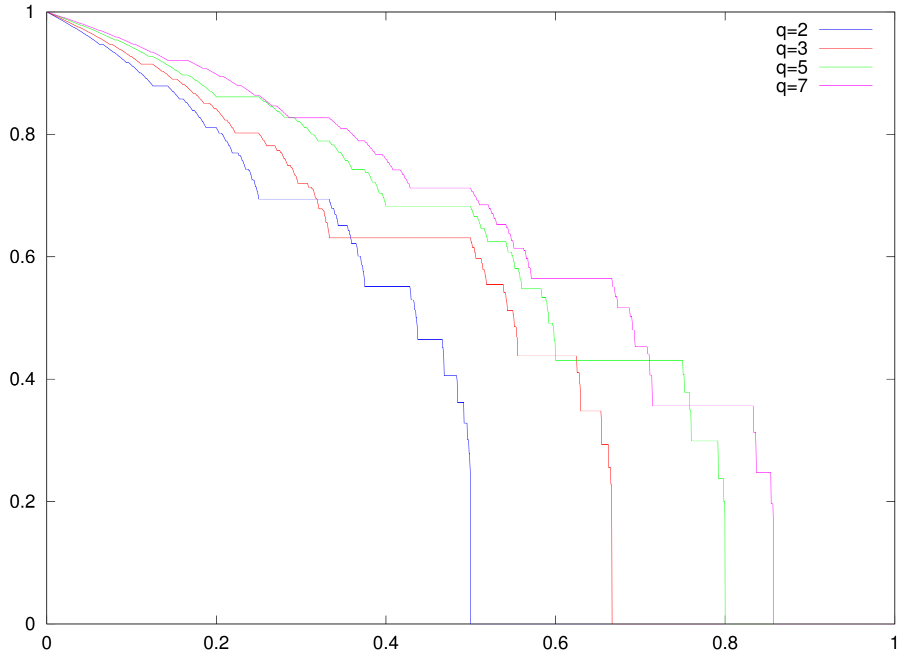


Figure A.1: Numerical plots of ϕ for $q \in \{2, 3, 5, 7\}$.

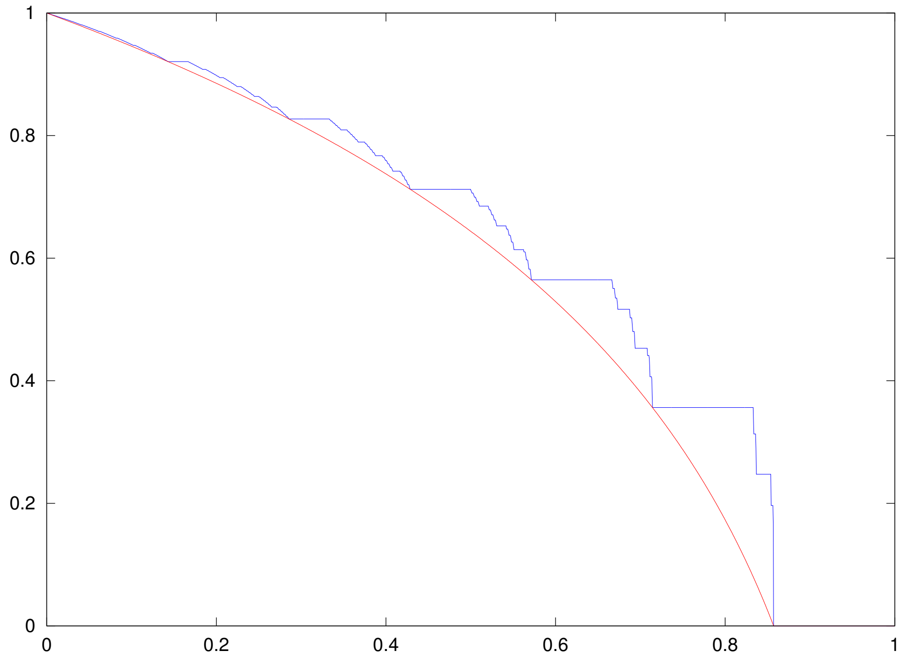


Figure A.2: Plots of ϕ and ψ when $q = 7$.

Theorem A.24. *For all $c \in [0,1)$ we have*

$$\frac{\phi(c)}{\psi(c)} \to 1$$

as $q \to \infty$.

Proof. Let $c \in [0,1)$ be given. Then if we let $i = \lfloor qc \rfloor$ we have

$$\frac{i}{q} \leq c \leq \frac{i+1}{q}.$$

Now

$$\phi\left(\frac{i}{q}\right) \geq \phi(c) \geq \phi\left(\frac{i+1}{q}\right)$$

and likewise for ψ since both functions are decreasing. Due to the result we got earlier on constant intervals we have

$$\frac{\log(q-i)}{\log q} \geq \phi(c), \psi(c) \geq \frac{\log(q+1-i)}{\log q}$$

so recalling the definition of i we have

$$\frac{\log(q-i)}{\log(q-i+1)} \geq \frac{\phi(c)}{\psi(c)} \geq \frac{\log(q-i+1)}{\log(q-i)}$$

and since $i \to \infty$ as $q \to \infty$, both the lower and upper bound converges to 1. This finishes the proof. $\square$

Since we also see that $\psi(c) \to 1$ as $q \to \infty$, we also have the following corollary.

Corollary A.25. *For all $c \in [0,1)$ we have*

$$\phi(c) \to 1 \text{ as } q \to \infty.$$

The convergence is very slow though – since ϕ and ψ are equal on q points we can just look at the convergence of

$$\frac{\log(1-c)}{\log q}$$

to zero which is easy to calculate.

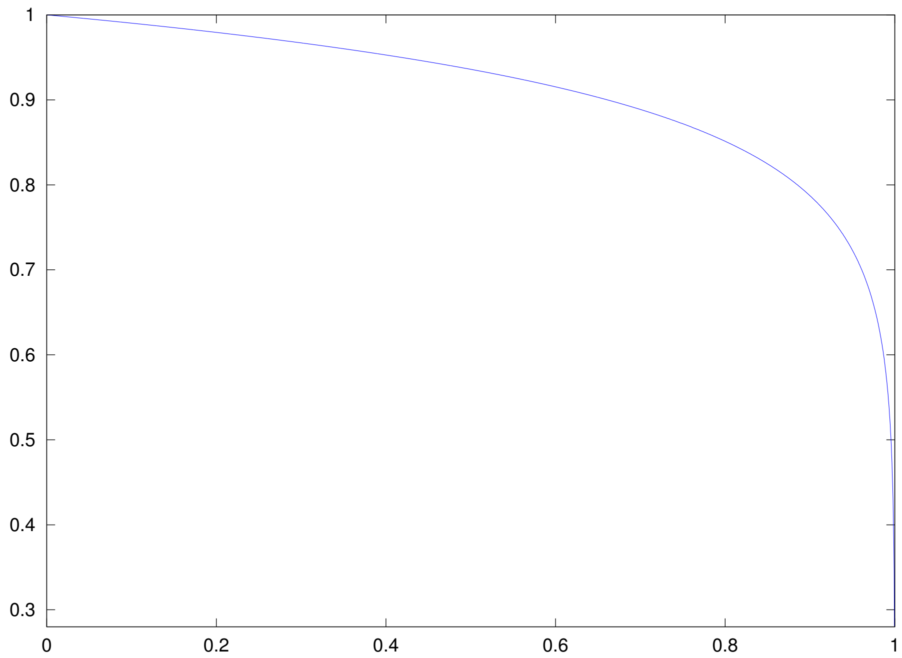


Figure A.3: Plot of ϕ when $q = 50000$.

Appendix B

Code

The code is done in GNU Octave, and can be downloaded from `http://jonaslindstrom.dk/thesis/hdim.zip`.

B.1 Brute force calculation of ϕ

The set F_c can be seen as a subshift of finite type, see Appendix A, and here we construct the transition matrix and find the spectral radius ρ. Then $\phi(c) = \frac{\log \rho}{\log q}$.

```
##
# Promt for base number q and length of the words considered, m
##
q = input("q=");
m = input("m=");


##
# Construct the transition matrix
##
A = zeros(q^m, q^m);

for i=1:q^m
  for j=1:q^m
    # If the last m-1 digits of i-1 are equal to to
    # the first m-1 of j-1 then set A(i,j) = 1
    if mod(i-1,q^(m-1)) == floor((j-1)/q);
      A(i,j) = 1;
    endif
  endfor
endfor
```

```
##
# Calculate the Hausdorff dimensions of F_c for c = (i-1)/q^m
# where i=1,2,...,q^m
##
x = zeros(1, q^m);
h = zeros(1, q^m);

for i=1:q^m
  # B is the transition matrix for F_c where c=(i-1)/q^m
  B = A(i:q^m, i:q^m);

  x(i) = (i-1)/q^m;
  h(i) = log(max(eig(B)))/log(q);
endfor

plot(x,h);
```

B.2 Smarter calculation of ϕ

Here we use Theorem 2.4 to calculate ϕ.

```
##
# Promt for base number q and length of the words considered, m
##
q = input("q=");
m = input("m=");

# h(i) will be the Hausdorff dimension of F_x(i)
x = zeros(1,q^m);
h = zeros(1,q^m);

# When c=0, the entropy is q
x(1) = 0;
h(1) = q;

for c=1:q^m
  ## Find n as in the theorem
  for j=m:-1:1
    ## If the last m-j digits of c are larger than the first m-j,
    ## then n=j.
```

```
    if mod(c, q^(m-j)) <= floor(c / q^j)
      n = j;
    endif
  endfor

  ## The coefficients of the polynomial is saved in P
  P = zeros(1,n+1);
  P(1) = 1;

  ## The base q-expansion of a is a_1 ... a_n as in the theorem where
  a = q^n - floor(c / q^(m-n));

  ## Find a_n, a_{n-1}, ..., a_1 and save them in P
  for j=n+1:-1:2
    b = mod(a, q);
    P(j) = -b;
    a = (a-b) / q;
  endfor

  h(c+1) = log(max(roots(P)))/log(q);
  x(c+1) = c/q^m;
endfor

plot(x,h);
```

Bibliography

[1] B. Adamczewski and Y. Bugeaud (2006), On the Littlewood conjecture in simultaneous Diophantine approximation, *J. London Math. Soc.* (2) **73**, no. 2, pp. 355–366.

[2] V. Beresnevich and S. Velani (2006), A Mass Transference Principle and The Duffin-Schafer conjecture for Hausdorff measures, *Ann. of Math.* (2) **164**, pp. 971–992.

[3] V. Beresnevich, A. Haynes and S. Velani (2010), Multiplicative zero-one laws and metric number theory, *arXiv:1012-0675v1*.

[4] Bourgain et. al (2009), Some effective results for $\times a \times b$, *Ergodic Theory Dynam. Systems* **29**, pp. 1705–1722.

[5] M. Einsiedler, A. Katok, and E. Lindenstrauss (2006), Invariant measures and the set of exceptions to Littlewood's conjecture, *Ann. of Math.* (2) **164**, no. 2, pp. 513–560.

[6] D. K. Faddev and V. N. Faddeeva (1963), *Computational Methods of Linear Algebra*, W.H. Freeman and Company.

[7] K. Falconer (2003), *Fractal Geometry*, Second Edition, Mathematical Foundations and Applications, Wiley.

[8] G. Harman (1998), *Metric number theory*, London mathematical society monographs, new series 18, Oxford science publications.

[9] S. Harrap and A. Haynes (2011), The mixed Littlewood conjecture for pseudo-absolute values, *arXiv:1012.0191v2*.

[10] A. Haynes, J. L. Jensen and S. Kristensen (2012), *arXiv:1204.0938v1*, to appear in *Proc. Amer. Math. Soc.*

[11] V. Jarník (1928-9), Zur metrischen Theorie der diophantischen Approximationen, *Prace Mat.-Fiz.*, pp. 91–106.

[12] J. L. Jensen (2011), On lower bounded orbits of the times-Q map, *Unif. Distrib. Theory* **6**, no. 2, pp. 157–175.

[13] R. Kaufman (1980), Continued fractions and Fourier transforms, *Mathematika* **27**, no. 2, pp. 262–267.

[14] A. Khintchine (1924), Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen, *Math. Ann.* **92**, no. 1-2, pp. 115–125.

[15] B. de Mathan and O. Teulié (2004), Problèmes diophantiens simultanés, *Monatsh. Math.* **143**, no. 3, pp. 229–245.

[16] H. L. Montgomery (1994), Ten lectures on the interface between analytic number theory and harmonic analysis, *CBMS Reg. Conf. Ser. Math.*, vol. **84**, Published for the Conference Board of the Mathematical Sciences, Washington, DC.

[17] J. Nilsson (2009), On numbers badly approximable by dyadic rationals, *Israel J. Math.* **171**, pp. 93–110.

[18] W. Parry (1960). On the β-Expansion of Real Numbers, *Acta Math. Acad. Sci. Hung.* **11**, pp. 401–416.

[19] Y. Pesin (1997), *Dimension theory in Dynamical Systems*, The University of Chicago Press.

[20] A. D. Pollington and S. L. Velani (2000), On a problem in simultaneous Diophantine approximation: Littlewood's conjecture, *Acta Math.* **185**, no. 2, pp. 287–306.

[21] A. D. Pollington and R. C. Vaughan (1990), The k-dimensional Duffin-Schafer conjecture, *Matematika* **37**, pp. 190–200.

[22] A. Rényi (1957), Representations for Real Numbers and their Ergodic Properties, *Acta Math. Acad. Sci. Hung.* **8**, pp. 477–493.

[23] U. Shapira (2011), A solution to a problem of Cassels and Diophantine properties of cubic numbers, *Ann. of Math.* (2) **173**, pp. 543–557.

[24] M. Urbanski (1986), On Hausdorff dimension of invariant sets for expanding maps of the circle, *Ergodic Theory Dynam. Systems* **6**, pp. 295–309.