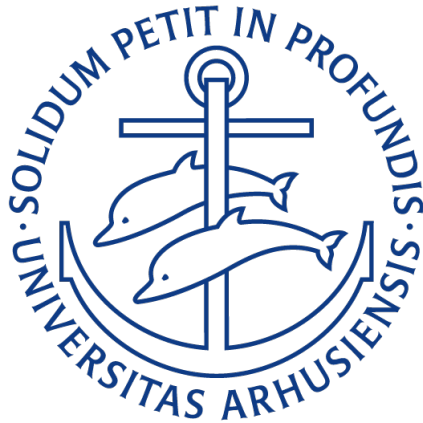


METRIC DIOPHANTINE
APPROXIMATION AND
MISSING-DIGITS SETS

or

DIMENSION, DIGITS, AND DIOPHANTINE
APPROXIMATION



PHD THESIS
MAIKEN BALMAN GRAVGAARD
201808974

SUPERVISOR: SIMON KRISTENSEN

JULY 2026

INSTITUT FOR MATEMATIK
AARHUS UNIVERSITET

Abstract

This thesis is concerned with number-theoretic properties of real numbers, approached from two directions: the first is representation of real numbers via numeration systems and the (measure-theoretic) size of sets of real numbers with prescribed Diophantine approximation properties and the second is numbers defined by digit restrictions.

The results presented stem from one paper and two preprints, written in collaboration with Ying Wai Lee, Jaroslav Hančl, and Simon Kristensen, concerning missing-digits sets, expressible sets, and quantitative Khintchine-type theorems.

In the first part of the thesis we study missing-digits sets arising from base- b expansions, continued fractions, and Lüroth expansions. We use their interpretations as Cantor sets to establish several results on decomposing a real number into a sum of numbers whose Lüroth expansion digits are restricted in particular ways. We also use the interpretation of missing-digits sets as Cantor sets to give a lower bound on the Hausdorff dimension of expressible sets.

In the second part, we turn to metric Diophantine approximation. We prove a quantitative version of the convergence part of Khintchine's theorem in metric Diophantine approximation, but where the approximated points lie on the standard parabola and the approximation function is not assumed to be monotonic. Also included in this section is a note on twisted inhomogeneous Dirichlet non-improvability.

Resumé

Denne afhandling omhandler talteoretiske egenskaber ved de reelle tal tilgået fra to retninger: først det at repræsentere reelle tal i forskellige talsystemer og størrelsen (i målteoretisk forstand) af mængder af reelle tal med foreskrevne diofantiske approksimationsegenskaber og dernæst reelle tal defineret ved begrænseder på cifrene.

Resultaterne, som præsenteres, stammer fra en artikel og to preprints, skrevet i samarbejde med Ying Wai Lee, Jaroslav Hančl og Simon Kristensen, og omhandler missing digits sets (tal repræsenteret ved talsystemer med manglende cifre), expressible sets (mængder af tal repræsenteret ved rækker på formen $\sum_{n=1}^{\infty} \frac{1}{a_n}$) og kvantitative Khintchine-type-sætninger.

I den første del af afhandlingen studerer vi mængder med restringerede cifre, der stammer fra base- b -udvidelser, kædebrøker og Lüroth-udvidelser. Vi bruger deres fortolkninger som Cantor-mængder til at vise adskillige resultater om dekomponering af et reelt tal som summen af to eller flere tal, hvis Lüroth-udvidelsescifre er begrænsede på den eller den anden måde. Vi bruger også deres fortolkning som Cantor-mængder til at give en nedre grænse for Hausdorff-dimensionen af expressible sets for forskellige følger.

I anden del vender vi os mod metrisk diofantisk approksimation. Vi beviser en kvantitativ version af konvergensdelen af Khintchines sætning i metrisk diofantisk approksimation, men hvor de approksimerede punkter ligger på standardparablen og approksimationsfunktionen ikke antages at være monoton. Dette afsnit indeholder tilmed en note om twisted inhomogeneous Dirichlet non-improvability.

Preface

This PhD thesis consists of three chapters followed by one paper and two preprints. All original results along with their proofs can be found in the papers and preprints appended at the end of the thesis. The purpose of the Chapters 1-3 is to give more thorough background for the material presented in the papers, to explain the relevance as well as the limitations of the results, and to serve as a more gentle walkthrough than that of the papers. Because the papers are inserted in their original preprint form, some information is inevitably repeated between this introduction and the individual preprints' introductions, all of which are available open access on arXiv.

All original results are included in this introduction, as well as the necessary background to understand the theorem statements. Readers should consult the appended papers for the full proofs, since these are only sketched here to give a flavour of the methods so as to not weigh the ideas down with details. One exception is an original proof of a result on twisted inhomogeneous Dirichlet non-improvability (§3.2). The result itself is not new, but the proof presented here is.

- Chapter 4 contains a preprint of the (accepted and online ready) paper [46] *Decomposition of real numbers into sums of Liuroth sets* by myself and Ying Wai Lee. This paper was written during my stay in York in the spring of 2025.
- Chapter 5 contains a preprint of the (submitted but not yet accepted) paper [44] *Lower bound for the Hausdorff dimension of expressible sets* by myself, Jaroslav Hančl and Simon Kristensen. This paper came about when Jaroslav visited Simon and me in Aarhus in October 2025.
- Finally, Chapter 6 contains a preprint of the paper [45] *Quantitative Khintchine on the parabola with non-monotonic approximation functions* by myself and Simon Kristensen.

The central thread of this thesis is Diophantine approximation. As the alternative title, *Dimension, Digits, and Diophantine Approximation*, suggests,

we approach real numbers from several angles: the Hausdorff measure and dimension of sets of numbers with prescribed number-theoretic properties; numeration systems; and alternative representations of real numbers, and classical methods in metric Diophantine approximation.

Original results are presented throughout, in the sections where they thematically make most sense, rather than grouped by paper.

Chapter 1 is first and foremost meant to serve as a foundational introduction to the kinds of real numbers we will be interested in. We start with an introduction to the irrational, transcendental, and Liouville numbers in §1.1, and then move on to different ways to represent real numbers in §1.2. The papers [46] and [44] both concern missing-digits sets and their interpretations as Cantor sets. In this section, our results on decomposition of real numbers into sums of Lüroth sets are presented.

This leads us into Chapter 2, which is largely about Cantor sets. We start Chapter 2 with an introduction to Hausdorff measure and dimension, after which we introduce general Cantor sets as well as three different ways of representing them — each with their own advantages when it comes to applications. The chapter ends with a review of various original results on Hausdorff dimension both from [44] and from [46].

Having established how to represent real numbers and how to measure the resulting sets, we turn in Chapter 3 to approximating them by rational numbers. We are not so concerned with each individual number's approximation properties as we are with the size (in terms of measure) of sets of numbers with certain approximation properties. §3.2 contains a proof of a result on the Hausdorff dimension of twisted inhomogeneous Dirichlet non-improvable real vectors. This proof was originally included in the Progress Report and has been inserted completely unchanged (aside from the reference numbers and the formatting). This is done in accordance with the GSNS rules and regulations. Lastly, the paper [45] is motivated and the results are presented alongside a sketch of the proof and a comparison of the four main theorems therein.

In terms of notation, the main aim has been to keep it consistent throughout the thesis, with the secondary aim of being consistent with the papers and lowest priority placed on consistency with other literature on the subjects. Perhaps controversially and regrettably, we write $\mathbb{N}$ for the natural numbers not including 0 and $\mathbb{N}_0$ for the natural numbers including 0. $\log$ is always the natural logarithm.

Maiken Balman Gravgaard
Aarhus, July 2026

Acknowledgements

As I put the final touches on this thesis, it is time to acknowledge the many people who have made this journey possible.

First and foremost, I would like to thank my supervisor, Simon Kristensen, without whom this project would not have happened at all. From my undergraduate thesis all the way to the final week of my PhD, I have always looked forward to our meetings. There has never been a dull moment. You have taught me an enormous amount, mathematically and otherwise, and your unwavering belief that everything would work out has been a great comfort.

I am also grateful to Sanju Velani and the wonderful number theory group at the University of York for welcoming me during my visit in the spring of 2025. The semester was both mathematically fruitful and enjoyable, and I am very happy to be returning for my postdoctoral position. In particular, I would like to thank Sanju and Ying Wai for the many discussions that ultimately led to my first paper.

I would also like to thank my collaborators, Ying Wai Lee and Jaroslav Hančl. It has been a pleasure to work with you, and I have learned a great deal from our collaborations.

To Nikolaj Lundbye and Tobias Overgaard: thank you for the last eight years. I could not have asked for better companions throughout my undergraduate studies, master's degree, and PhD. Thank you for the late nights studying until we were exhausted and sometimes a little delirious, the endless pros-and-cons lists and discussions of every major (and sometimes minor) decision, the faithful exchange of the latest gossip, the slightly too much chatter in the office, the anime club, the ongoing battle against screen addiction, and so much more. It would not have been the same without you!

I would also like to thank the revolving cast of PhD students and postdocs in the department for creating such a fun office culture. Thank you for being willing to socialize at unconventional academic hours, for the many tours de kontor, trips to the Friday bar, PhD and postdoc seminars, for conversations

about mathematics and everything else, and for making my time as a PhD student a fun one.

For this I would also like to thank my students. Teaching undergraduate courses has always been a welcome change of pace and something I could consistently look forward to each week.

Last but not least, I would like to thank my family and especially my parents, who have always been extremely supportive, encouraging and helpful and from whom I no doubt inherited my curiosity and the confidence to believe I can pursue whatever I set my mind to and to have fun doing it.

This PhD study was funded by the Aarhus University Research Foundation, grant no. AUFF-E-2021-9-20.

Contents

Abstract	i
Resumé	iii
Preface	v
Acknowledgements	vii
Contents	ix
1 Irrational numbers, numeration systems and missing-digits sets	1
1.1 Irrational, transcendental and Liouville numbers	1
1.2 Numeration systems	5
Base b	6
Continued fractions	8
Lüroth expansion	10
Expressible sets	12
1.3 Missing-digits sets	12
2 Hausdorff Measure and Dimension	19
2.1 Construction	19
Hausdorff measure and dimension	19
2.2 Cantor sets and iterated function systems	20
Examples	25
Thickness	34
2.3 Summary of my results on Hausdorff dimension	37
3 Metric Diophantine Approximation	41
3.1 Introduction and state of the art	41
3.2 Twisted inhomogeneous Dirichlet non-improvability	44
3.3 Quantitative Khintchine	49

Bibliography	59
4 Paper 1: Decomposition of real numbers into sums of Lüroth sets	63
5 Preprint 2: Lower bound for the Hausdorff dimension of expressible sets	101
6 Preprint 3: Quantitative Khintchine on the parabola with non-monotonic approximation functions	111

Irrational numbers, numeration systems and missing-digits sets

1.1 Irrational, transcendental and Liouville numbers

The purpose of this section is first and foremost to establish the number-theoretic background needed for the remainder of the thesis and to provide historical context for the study of irrational, transcendental, and Liouville numbers. For a more detailed historical overview, I recommend [37]. For a survey of results by Euler on this topic see [39].

We say that a real number a is *rational*, if it solves an equation of the form $qx - p = 0$, where p and q are integers with $q > 0$. On the other hand, if a number does not solve any such equation, we call it *irrational*. The idea of irrational numbers goes back to antiquity. It is thought that Hippasus of Metapontum, an early follower of Pythagoras, knew of the irrationality of $\sqrt{2}$. According to legend, this discovery led to his untimely death. Some claim, that the Vedic civilization in India implicitly described an idea of irrationality in the Shulba Sutra around or before 800 BC. By the end of the 1700s π and e were shown to be irrational by Euler in 1737 and Lambert in 1768 respectively.

Looking at the expression $qa - p = 0$, we can reformulate the above to say that a is rational because it is a root of a degree 1 polynomial with integer coefficients (where the leading coefficient q is positive). This phrasing suggests a natural extension to the rational numbers, namely, the numbers (real or complex) that are roots of a polynomial with integer coefficients of

arbitrary degree $d \geq 1$:

$$c_0 + c_1x + c_2x^2 + \cdots + c_dx^d \quad \text{with } c_d \geq 1 \text{ and } c_0, c_1, \dots, c_d \in \mathbb{Z}.$$

This defines the *algebraic numbers*, and we define the degree of an algebraic number as the smallest $d \geq 1$ such that this holds. For instance $\sqrt{2}$ is algebraic since it solves $x^2 - 2 = 0$ and its degree is 2.

If a (real or complex) number is not algebraic, it is called *transcendental*. The study of these is comparatively much younger. As far as we know, Euler was the first to define transcendental numbers. Additionally he conjectured that logarithms of numbers, which were not themselves powers of the base, were transcendental. Existence of transcendental numbers was not, however, proven until 1844, when Liouville proved:

Theorem (Liouville, 1844). *If a is an algebraic irrational number of degree $d \geq 2$, then there exists a positive real number A so that for any rational number $\frac{p}{q}$ in $[a - 1, a + 1]$*

$$\left| \frac{p}{q} - a \right| \geq \frac{1}{Aq^d}.$$

This result essentially says that rational numbers $\frac{p}{q}$ are bad at approximating irrational algebraic numbers, since it states that there is always a gap of size at least $\frac{1}{Aq^d}$. More notably, Liouville used his result to finally construct a transcendental number, namely

$$\sum_{n=1}^{\infty} \frac{1}{10^{n!}}.$$

To see why, pick rationals $\frac{p_N}{q_N} = \sum_{n=1}^N \frac{1}{10^{n!}}$ for each natural number N and notice that

$$\left| \frac{p_N}{q_N} - \sum_{n=1}^{\infty} \frac{1}{10^{n!}} \right| < \frac{2}{10^{(N+1)!}},$$

and by Liouville's theorem, if $\sum_{n=1}^{\infty} \frac{1}{10^{n!}}$ was irrational algebraic, then we would have, for some $d \geq 2$ and $A > 0$ that

$$0 < \frac{1}{A(10^{N!})^d} \leq \left| \frac{p_N}{q_N} - \sum_{n=1}^{\infty} \frac{1}{10^{n!}} \right|.$$

Thus the first¹ example of a transcendental number was born. In fact many transcendental numbers can be constructed using the strategy of Liouville's

¹I believe technically this was the second transcendental number ever constructed. Earlier in the same paper, Liouville constructs a transcendental number represented as a continued fraction with rapidly growing partial quotients.

theorem. Any non-periodic decimal number with sufficiently long blocks of zeros or any continued fraction (see §1.2) with rapidly increasing partial quotients is transcendental

Numbers constructed using this recipe possess the property that they have a sequence of rationals that approximate them extremely well. A real number a , which can be approximated very well by rational, is called a *Liouville number*. These are characterised by the property that for any $w > 1$, there is a rational number $\frac{p}{q}$ such that

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

It is one thing to invent a transcendental number, it is much more difficult to prove whether a specific number is or is not transcendental. Liouville's original intention had been to prove that e is transcendental. This was done by Hermite in 1873 and subsequently in 1882 Lindemann proved that π was transcendental as well.

In his 1874 paper [26], Cantor proved that the real numbers are uncountable, while the algebraic numbers are countable, thus the transcendental numbers must be uncountable, making it even more unintuitive that they are so hard to identify.

The class of examples of transcendental numbers given by Liouville's theorem are all Liouville numbers. Although the Liouville numbers make up an uncountable set (since any sequence of natural numbers $(a_n)_{n=0}^{\infty}$ gives rise to a Liouville number $a_0.a_1a_2(3 \text{ zeros})a_4(17 \text{ zeros})a_5(95 \text{ zeros})a_6 \dots$) the set of Liouville numbers has both Lebesgue measure and Hausdorff dimension zero (see Chapter 2 for definitions).

In an effort to identify more examples of irrational numbers, Erdős [38] famously showed that if $(a_n)_{n=1}^{\infty}$ is a sequence of natural numbers with $a_n \geq n^{1+\varepsilon}$ for some $\varepsilon > 0$ and with $\limsup_{n \rightarrow \infty} a_n^{1/2^n} = \infty$, then the real number

$$\sum_{n=1}^{\infty} \frac{1}{a_n}$$

is irrational. Furthermore, if we take any sequence of natural numbers $(d_n)_{n=1}^{\infty}$, then the number $\sum_{n=1}^{\infty} \frac{1}{a_n d_n}$ is irrational, so for each sequence $(a_n)_{n=1}^{\infty}$ satisfying the conditions of Erdős' theorem, we get a whole class of irrational numbers.

This was later strengthened in [48], with the $\limsup$ condition replaced by the condition $1 < \liminf_{n \rightarrow \infty} a_n^{1/2^n} < \limsup_{n \rightarrow \infty} a_n^{1/2^n}$.

In a sense, this is optimal. Consider Sylvester's sequence $(a_n)_{n=1}^\infty$ given by the recurrence $a_1 = 2$ and $a_{n+1} = a_n^2 - a_n + 1$ for $n \geq 1$. Aho and Sloane [2] showed that the sequence $(a_n^{1/2^n})_{n=1}^\infty$ converges with a limit that is > 1 and by [50] $\sum_{n=1}^\infty \frac{1}{a_n} = \frac{1}{a_1-1}$ is rational, so we cannot relax the condition that $(a_n^{1/2^n})_{n=1}^\infty$ diverges.

For a growth condition that guarantees transcendence, we look to [49]. If $a_n = \frac{p_n}{q_n}$ is rational such that for some $\alpha > \beta > 0$, the sequence of numerators satisfies $p_n \leq 2^{(3+\beta)n}$ and the denominators are $q_n \geq 2^{(3+\alpha)n}$. Then $\sum_{n=1}^\infty \frac{p_n}{q_n d_n}$ is transcendental for any choice of sequence $(d_n)_{n=1}^\infty$ in $\mathbb{N}$. Since we are considering sequences $(a_n)_{n=1}^\infty$ of rational numbers, in principle nothing is stopping us from considering algebraic numbers as well. For examples of such, see e.g. [3], [50], [65], and the references therein.

Erdős also proved that if the growth condition in his other result is strengthened to the condition that

$$\limsup_{n \rightarrow \infty} a_n^{1/t^n} = \infty$$

for every $t \in \mathbb{N}$, then the number $\sum_{n=1}^\infty \frac{1}{d_n a_n}$ is Liouville.

For a sequence $(a_n)_{n=1}^\infty$, we call the set of numbers of the form $\sum_{n=1}^\infty \frac{1}{d_n a_n}$, the *expressible set* of $(a_n)_{n=1}^\infty$, and denote it

$$\mathcal{K}((a_n)_{n=1}^\infty) = \left\{ x = \sum_{n=1}^\infty \frac{1}{d_n a_n} : d_n \in \mathbb{N} \text{ for each } n \in \mathbb{N} \right\}.$$

For a sequence of natural numbers $(a_n)_{n=1}^\infty$ Erdős' results imply that if

$$\limsup_{n \rightarrow \infty} a_n^{1/2^n} = \infty,$$

then the expressible set $\mathcal{K}((a_n)_{n=1}^\infty)$ contains only irrational numbers.

Moreover, if

$$\limsup_{n \rightarrow \infty} a_n^{1/t^n} = \infty$$

for every natural number t , then $\mathcal{K}((a_n)_{n=1}^\infty)$ consists only of Liouville numbers. In the preprint *Lower bound for the Hausdorff dimension of expressible sets* by the author, Jaroslav Hančl and Simon Kristensen [44], found in Chapter 5, Corollary 5 states that this condition is sharp, that is, if there exists a natural number k such that $\limsup_{n \rightarrow \infty} a_n^{1/k^n} < \infty$, then the expressible set $\mathcal{K}((a_n)_{n=1}^\infty)$ contains non-Liouville numbers. Thus we have found honest irrational numbers in our expressible set, not just Liouville numbers, which are easy to construct but take up very little space in the metric sense. The proof however is

not constructive, so of the elements of $\mathcal{K}((a_n)_{n=1}^{\infty})$ we cannot point out any specific irrational non-Liouville number.

For results on measure and dimension of expressible sets see also [51], [54] and [53],[55] respectively (again see Chapter 2 for definitions).

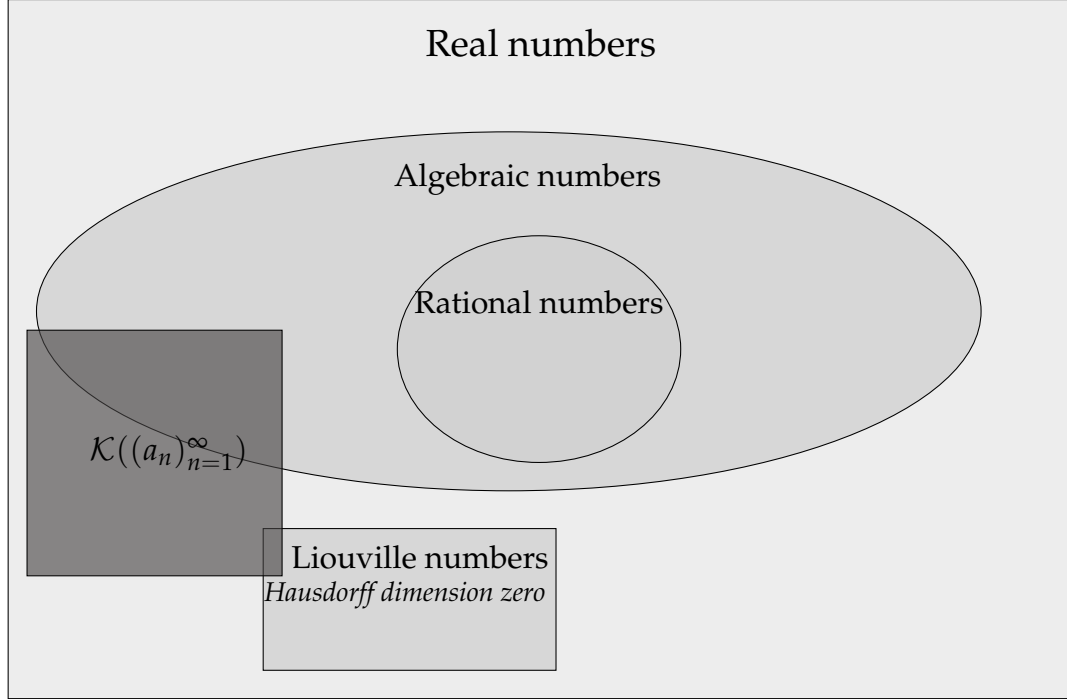


Figure 1.1: Sets represented by rectangles are uncountable, while those represented by ovals are countable. The irrational numbers and the transcendental numbers are precisely those numbers *not* belonging to the rational numbers and the algebraic numbers, respectively. The set $\mathcal{K}((a_n)_{n=1}^{\infty})$ is the expressible set, about which Corollary 5 in [44] is concerned. The placement of $\mathcal{K}((a_n)_{n=1}^{\infty})$ is illustrative and is intended only to show possible intersections with the sets above. The main takeaway is that it is a subset of the irrational numbers, not wholly contained in the Liouville numbers.

1.2 Numeration systems

Let x be a real number from the unit interval $[0, 1)$. The most common way to denote x is $0.a_1(x)a_2(x)\dots$, where $a_n(x) \in \{0, 1, \dots, 9\}$ for $n \geq 1$, by which

we mean

$$x = \sum_{n=1}^{\infty} \frac{a_n(x)}{10^n}.$$

This expansion of x is called *decimal expansion*, and it is unique, provided we do not allow infinite expansions with $a_n(x) = 9$ for all n greater than some n_0 .

When representing real numbers this way, the rational numbers can be characterised as exactly those whose decimal expansion is either finite or eventually periodic. Indeed if the decimal expansion is finite, then x is of the form

$$x = 0.a_1(x) \dots a_k(x) = \frac{a_1(x)10^{k-1} + \dots + a_{k-1}(x)10 + a_k(x)}{10^k}$$

for some natural number k , and this is clearly rational.

On the other hand, if x has an eventually periodic expansion, meaning it is of the form

$$x = 0.a_1(x) \dots a_k(x) \overline{a_{k+1} \dots a_{k+l}}$$

where k and l are chosen minimally then we can rewrite x as

$$x = 0.a_1(x) \dots a_k(x) + \frac{10^l}{10^l - 1} 0.a_{k+1}(x) \dots a_{k+l}(x)$$

which is again rational.

Conversely, suppose that $x = \frac{p}{q}$ is rational such that $x \in (0, 1)$ with $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ relatively prime. By Euclid's algorithm, there are unique integers q_1 and r_1 with $0 \leq r_1 < q$ such that $10p = q_1q + r_1$. Either $r_1 = 0$, in that case $\frac{p}{q} = 0.q_1$ and the decimal expansion terminates, otherwise $r_1 > 0$, in which case we repeat the procedure $10r_1 = q_2q + r_2$. Since the remainders r_i are restricted to the numbers $0, 1, \dots, q-1$, by the pigeonhole principle, at some point we must either get a remainder of 0, in which case the decimal expansion terminates, or the remainder repeats a previous value, after which the decimal expansion becomes periodic.

Base b

The choice of 10 as the basis for writing numbers is somewhat arbitrary. We might as well have chosen any integer $b \geq 1$ (or even a real number > 1 , see e.g. Chapter 2.4.1 in [30]) and "digits" in $\{0, 1, \dots, b-1\}$. In base b the characterisation of rationals as those numbers whose expansion is either finite or eventually periodic remains.

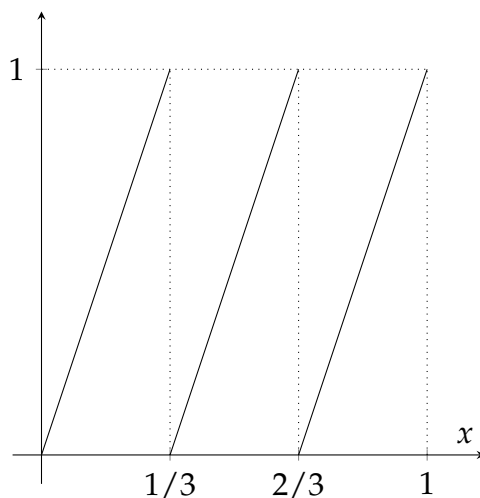


Figure 1.2: Times-3 Map

Fix a base $b \geq 2$. For a number $x \in [0, 1)$, let us run through how the "digits" $a_1(x), a_2(x), \dots \in \{0, 1, \dots, b-1\}$ can be obtained.

First partition $[0, 1)$ into intervals $[\frac{i}{b}, \frac{i+1}{b})$ for $0 \leq i \leq b-1$. Each interval $[\frac{i}{b}, \frac{i+1}{b})$ is labeled with the "digit" " i ". Then $a_1(x) = i$ for $x \in [\frac{i}{b}, \frac{i+1}{b})$. To find $a_2(x)$ we repeat the procedure: we subdivide each interval $[\frac{i}{b}, \frac{i+1}{b})$ into $b-1$ subintervals of equal length: $[\frac{i}{b} + \frac{j}{b^2}, \frac{i}{b} + \frac{j+1}{b^2})$ for $0 \leq i \leq b-1$ and $0 \leq j \leq b-1$. For $x \in [\frac{i}{b} + \frac{j}{b^2}, \frac{i}{b} + \frac{j+1}{b^2})$, $a_1(x) = i$ and $a_2(x) = j$. This repeats forever, so the digits are found by repeated application of the same procedure: partitioning intervals into $b-1$ subintervals of equal length.

The base b expansion is generated by the map $T : [0, 1) \rightarrow [0, 1)$ (illustrated in Figure 1.2 for $b = 3$):

$$Tx = bx \pmod{b} = \begin{cases} bx & \text{for } 0 \leq x < \frac{1}{b}, \\ bx - 1 & \text{for } \frac{1}{b} \leq x < \frac{2}{b}, \\ \vdots & \vdots \\ bx - (b-1) & \text{for } \frac{b-1}{b} \leq x < 1. \end{cases}$$

Continued fractions

A *regular continued fraction* is an expression of the form

$$x = [a_0; a_1, a_2, \dots] := a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}, \quad (1.1)$$

where a_0 is an integer and $a_1, a_2, \dots$ are natural numbers. If the fraction (1.1) is finite, then x is rational. Conversely, using Euclid's algorithm, if $x = \frac{r_1}{r_0}$ is rational (assume $x \in (0, 1)$ and $\gcd(r_0, r_1) = 1$), then it can be written in the form of a finite continued fraction.

Indeed, there are unique integers r_2, a_1 with $0 \leq r_2 < r_1$ such that

$$r_0 = a_1 r_1 + r_2$$

and continuing, we get a strictly decreasing sequence of integers $r_1 > r_2 > \dots r_n > r_{n+1} = 0$ and a sequence of integers $a_1, a_2, \dots, a_n$ such that

$$r_1 = a_2 r_2 + r_3$$

$$r_2 = a_3 r_3 + r_4$$

$$\vdots$$

$$r_{n-2} = a_{n-1} r_{n-1} + r_n$$

$$r_{n-1} = a_n r_n + 0.$$

From this we get

$$a_1 = \left\lfloor \frac{r_0}{r_1} \right\rfloor, a_2 = \left\lfloor \frac{r_1}{r_2} \right\rfloor, \dots, \text{ and } a_n = \left\lfloor \frac{r_{n-1}}{r_n} \right\rfloor.$$

Put $x = \frac{r_0}{r_1}$, $T_1 = \frac{r_2}{r_1}$, $\dots$, and $T_{n-1} = \frac{r_n}{r_{n-1}}$. Then

$$x = \frac{1}{a_1 + T_1} = \frac{1}{a_1 + \frac{1}{a_2 + T_2}} = \dots = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}},$$

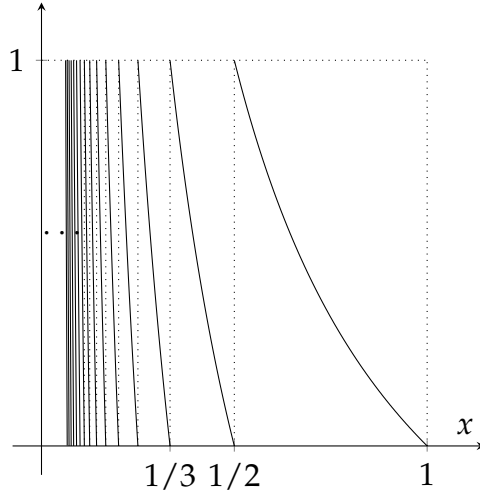


Figure 1.3: Gauss map

which is a finite continued fraction. By the algorithm $a_n \geq 2$. For each rational number we have two regular continued fraction expansions, namely $[a_1; a_1, \dots, a_n] = [a_1; a_1, \dots, a_n - 1, 1]$

Based on this procedure we define the regular continued fraction map

$$T : [0, 1) \rightarrow [0, 1)$$

(also called the Gauss map, pictured in Figure 1.3), which generates the continued fraction expansion for all real numbers, by

$$Tx := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor$$

for $x \neq 0$ and $T0 := 0$.

For an irrational number $x \in \mathbb{R} \setminus \mathbb{Q}$ define a_0 to be the unique integer such that $x - a_0 \in [0, 1)$ and let $T_0 := x - a_0$ and $T_n := T^n(x - a_0)$ for $n \geq 1$. By definition of T , T_n is irrational and $T_n \in [0, 1)$. Put

$$a_n = a_n(x) := \left\lfloor \frac{1}{T_{n-1}} \right\rfloor$$

for $n \geq 1$. Then

$$x = a_0 + \frac{1}{a_1 + T_1} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + T_2}} = \dots = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n + T_n}}}}.$$

We call the "digits" a_i , the *partial quotients* of x . Taking the limit as $n \rightarrow \infty$ gives us the regular continued fraction for x :

$$x = [a_0; a_1, a_2, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots}}}$$

(For existence of the limit see either [30], [60], or [61]). If we truncate the continued fraction after finitely many steps, we get what is called the *convergents* of x

$$[a_0; a_1, a_2, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n}}}} =: \frac{p_n}{q_n}.$$

To approximate an irrational number x by fractions, it is natural to consider its convergents $\frac{p_n(x)}{q_n(x)}$, and in fact it turns out that the convergents constitute the *best rational approximations* to x , better than for instance truncations of the decimal expansion.

Continued fractions have certain number theoretic advantages over base b expansion. As we saw above, the rational numbers are characterised as exactly the numbers with finite continued fraction expansion. Furthermore, as mentioned, the continued fraction convergents provide very good rational approximations. We can tell even more about a real number from its continued fraction expansion, namely Lagrange discovered that the partial quotients of a number are periodic if and only if that number is quadratic irrational.

Lüroth expansion

The map generating the base b expansion consists of linear maps defined on $b - 1$ subintervals of equal length partitioning $[0, 1)$, each with range $[0, 1)$.

For continued fractions, we instead subdivided $[0, 1)$ into intervals of the form $[\frac{1}{n}, \frac{1}{n-1})$ for $n \in \mathbb{N}$ but the transformation was not linear.

Combining these two constructions, partition $[0, 1)$ into the intervals $[\frac{1}{2}, 1]$, $[\frac{1}{3}, \frac{1}{2})$, $[\frac{1}{4}, \frac{1}{3})$, ... and let the transformation be linear on each interval with range $[0, 1)$ (see Figure 1.4).

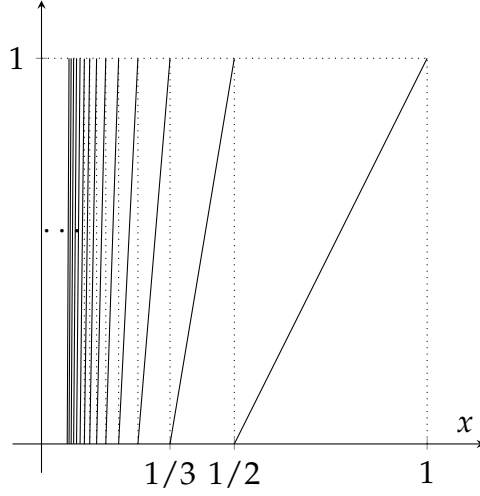


Figure 1.4: Lüroth Map

This map is defined by $T : [0, 1) \rightarrow [0, 1)$:

$$Tx = \begin{cases} n(n+1)x - n & \text{for } x \in (\frac{1}{n+1}, \frac{1}{n}], \\ 0 & \text{for } x = 0. \end{cases}$$

Every real number $x \in [0, 1)$ can be written as a *Lüroth series expansion*

$$\begin{aligned} & [0; d_1(x), d_2(x), d_3(x), \dots] \\ &= \frac{1}{d_1(x)} + \frac{1}{d_1(x)(d_1(x)-1)d_2(x)} + \frac{1}{d_1(x)(d_1(x)-1)d_2(x)(d_2(x)-1)d_3(x)} + \dots \end{aligned}$$

with $d_n(x) \geq 2$ for each n . We use the same bracket notation for Lüroth expansion as we used for continued fractions. To avoid confusion, the Lüroth expansion "digits" will be represented by d 's, whereas the continued fraction partial quotients are a 's. The digits $d_n(x)$ are defined by

$$d_n(x) = \left\lfloor \frac{1}{T^{n-1}x} \right\rfloor.$$

We can rewrite T as

$$Tx = \begin{cases} \left\lfloor \frac{1}{x} \right\rfloor \left(\left\lfloor \frac{1}{x} + 1 \right\rfloor x - 1 \right) & \text{for } x \neq 0, \\ 0 & \text{for } x = 0. \end{cases}$$

This map generates a unique infinite Lüroth expansion for each real number $x \in [0, 1)$. Alternatively the Lüroth expansion for a rational number can be

allowed to be finite, but the setup where all expansions are infinite fits the notation of the paper [46] *Decomposition of real numbers into sums of Lüroth sets* by the author and Ying Wai Lee (see Chapter 4).

More reading For more details, proofs and properties of the Lüroth series expansion see [30], as well as for more on the above expansions of real numbers, additional series expansions and their ergodic properties. For a phenomenal introduction to continued fractions and their properties, I recommend Khintchine's book [61]. See also [60].

Expressible sets

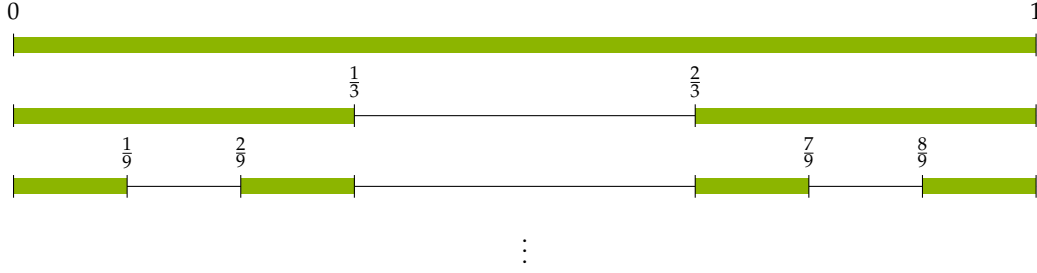
Like in Erdős' theorem [38] (see §1.1), for an increasing sequence of natural numbers $(a_n)_{n=1}^\infty$, consider $\sum_{n=1}^\infty \frac{1}{d_n a_n}$ for "digits" d_n in $\mathbb{N}$. For $x \in \mathcal{K}((a_n))$, this is a numeration system, but unlike for base b , continued fractions or Lüroth expansions, we might have $\mathcal{K}((a_n)) \neq \mathbb{R}$ and the representation may not be unique! Nevertheless, it is useful to think of the sequence $(d_n)_{n=1}^\infty$ as the "digits" of an element x of the expressible set $\mathcal{K}((a_n)_{n=1}^\infty)$ represented by the infinite series $x = \sum_{n=1}^\infty \frac{1}{d_n a_n}$.

1.3 Missing-digits sets

A natural way to obtain interesting subsets of the real numbers is by imposing restrictions on the digits appearing in a numeration system. For a numeration system with digits set $D \subseteq \mathbb{N}$ one specifies a subset $D' \subset D$ of admissible digits. The resulting sets are called *missing-digits sets*.

Base 3 Perhaps the most famous example is considering base 3 expansion (or ternary). Here the possible digits are 0, 1 and 2, so let us look at the set, where we only allow the digits 0 and 2. For this to make sense, we need the base 3 expansion to be unique, to avoid situations like $0.1 = 0.0\bar{2}$. For our purposes, let us adopt the convention that we forbid a 1 followed by infinitely many 0's or 2's. This means that e.g. $\frac{1}{3}$ is represented by $0.0\bar{2}$ (as opposed to 0.1) and $\frac{2}{3}$ is represented by 0.2 (not $0.1\bar{2}$).

This missing-digits set is a subset of the unit interval $[0, 1]$ and can be represented as follows: Consider the partition $[0, 1] = [0, \frac{1}{3}] \cup (\frac{1}{3}, \frac{2}{3}) \cup [\frac{2}{3}, 1]$. Elements of the middle interval $(\frac{1}{3}, \frac{2}{3})$ are exactly those numbers that in base 3 have 1 as the first digit, so we will remove these, leaving us with $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$ all of the numbers that start with 0.0 or 0.2. On each of these intervals we can

Figure 1.5: $C_{\frac{1}{3}}$

repeat the procedure: $[0, \frac{1}{3}] = [0, \frac{1}{9}] \cup (\frac{1}{9}, \frac{2}{9}) \cup [\frac{2}{9}, \frac{1}{3}]$, here the middle interval $(\frac{1}{9}, \frac{2}{9})$ consists exactly of those numbers which have 0 as their first digit and 1 as their second digit. Similarly we remove $(\frac{7}{9}, \frac{8}{9})$ from $[\frac{2}{3}, 1]$, since those are exactly the numbers with first digit 2 and second digit 1. Now we have $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$ the numbers that start with 0.00, 0.02, 0.20 and 0.22. Repeat ad infinitum. This procedure is illustrated in Figure 1.5. Let us denote the resulting set, called the *Cantor middle thirds set*, by $C_{\frac{1}{3}}$. This set has the same cardinality as $[0, 1]$, since every element has digits in $\{0, 2\}$, so can be mapped one-to-one to the real numbers written in binary. It is however also a Lebesgue null set (see Chapter 2). Despite being very small in the metric sense, Steinhaus [73] proved that every real number between 0 and 2 can be written as a sum of two numbers, whose ternary expansions do not contain any 1s.

Continued fractions A priori, the digit set for continued fraction expansion is all of $\mathbb{N}$. We consider two ways of restricting the digits (also called the partial quotients). The first is to bound them from above. For a natural number k denote by $F_{\leq k}$ the real numbers with continued fractions expansion $[0; a_1, a_2, \dots]$ such that $1 \leq a_n \leq k$ for all $n \in \mathbb{N}$. Same as for the missing-digit set above, for $k \geq 2$, this produces a set, which is uncountable but which at the same time has Lebesgue measure zero. Similarly to the result by Steinhaus on ternary expansion, in 1947 Hall [47] proved that every real number is representable as a sum of two elements from $F_{\leq 4}$ and an integer (See Figure 2.6.). We employ the notation

$$F_{\leq 4} + F_{\leq 4} \equiv \mathbb{R} \bmod 1$$

to mean exactly $F_{\leq 4} + F_{\leq 4} + \mathbb{Z} = \mathbb{R}$, where "+" denotes the Minkowski sum.

Later, using a refined version of Hall's idea, Hlavka [56] extended the result to

$$F_{\leq 3} + F_{\leq 4} \equiv \mathbb{R} \bmod 1$$

as well as proving the following four statements:

$$\begin{aligned}
 F_{\leq 2} + F_{\leq 7} &\equiv \mathbb{R} \bmod 1, \\
 F_{\leq 3} + F_{\leq 3} + F_{\leq 3} &\equiv \mathbb{R} \bmod 1, \\
 F_{\leq 2} + F_{\leq 2} + F_{\leq 4} &\equiv \mathbb{R} \bmod 1 \\
 \text{and} \\
 F_{\leq 2} + F_{\leq 2} + F_{\leq 2} + F_{\leq 2} &\equiv \mathbb{R} \bmod 1.
 \end{aligned}$$

Some of Hlavka's results are optimal. Hlavka himself [56] proved that $F_{\leq 2} + F_{\leq 4} \not\equiv \mathbb{R} \bmod 1$ and Divis [33] supplied with $F_{\leq 3} + F_{\leq 3} \not\equiv \mathbb{R} \bmod 1$ and $F_{\leq 2} + F_{\leq 2} + F_{\leq 2} \not\equiv \mathbb{R} \bmod 1$.

Using the notion of *thickness* of a Cantor set (details in §2.2) Astels [5] improved Hlavka's results to

$$F_{\leq 2} + F_{\leq 5} \equiv \mathbb{R} \bmod 1 \quad \text{and} \quad F_{\leq 2} + F_{\leq 2} + F_{\leq 3} \equiv \mathbb{R} \bmod 1.$$

This more or less put the question of decomposing real numbers into sums of continued fractions with partial quotients bounded from above to rest.

It is perhaps just as natural to consider partial quotients bounded from below. Define for any natural number k the set $F_{\geq k}$ of continued fractions $[0; a_1, a_2, \dots]$ (finite or infinite) with $a_n \geq k$ for each $n \in \mathbb{N}$.

Cusick and Lee [29] proved that

$$\underbrace{F_{\geq k} + \dots + F_{\geq k}}_{k \text{ times}} \equiv \mathbb{R} \bmod 1$$

Good [43] and later Das–Fishman–Simmons–Urbański [31] gave upper and lower bounds for the Hausdorff dimension of $F_{\geq k}$ and $F_{\leq k}$ (see §2.1 for definition).

In some cases, where $F_{\geq k_1} + F_{\geq k_2} \not\equiv \mathbb{R} \bmod 1$, we might still be interested in the size of the sum set. For this purpose Shulga [72] developed an algorithm, which proved that for $k \geq 2$

$$\left[0, \frac{1}{k-1}\right] \subseteq F_{\geq k} + F_{\geq 2}.$$

And for $k_1 \neq k_2$, if $3 \leq k_1 < k_2 \leq (k_1 - 1)^2$

$$\left[0, \frac{1}{k_1}\right] \subseteq F_{\geq k_1} + F_{\geq k_2}.$$

More recently Gayfulin and Nesharim [42] proved that in fact every number in $[0, 1]$ can be written as a sum of numbers, which are either rational or whose partial quotients diverge.²

Lüroth expansion The set-up here is very similar to that of continued fractions, so we consider exactly the analogue missing-digits sets, namely for $k \geq 2$

$$L_{\leq k} := \{[0; d_1, d_2, \dots] : 2 \leq a_n \leq k \text{ for each } n \in \mathbb{N}\}$$

and

$$L_{\geq k} := \{[0; d_1, d_2, \dots] : k \leq a_n \text{ for each } n \in \mathbb{N}\}.$$

In the undergraduate thesis of Brouwer [20], the following is stated as a Lüroth analogue to Hall's theorem, and a proof is sketched:

$$L_{\geq 4} + L_{\leq 4} \equiv \mathbb{R} \pmod{1}.$$

To the best of our knowledge this is the only statement regarding sums of Lüroth sets in the literature. The paper *Decomposition of real numbers into sums of Lüroth sets* by the author and Ying Wai Lee [46] in Chapter 4 is concerned with generalisations of this. We prove results analogous to those of Hall, Hlavka, Cusick-Lee and Good on continued fractions, but for Lüroth expansions.

The first result is on sums of sets where the digits are bounded from above. Here we also provide a different proof of Brouwer's result:

Theorem (G.-Lee, 2026, Theorem 2.1 in [46]).

$$\begin{aligned} L_{\leq 3} + L_{\leq 5} &\equiv \mathbb{R} \pmod{1} \\ L_{\leq 4} + L_{\leq 4} &\equiv \mathbb{R} \pmod{1} \\ L_{\leq 3} + L_{\leq 3} + L_{\leq 3} &\equiv \mathbb{R} \pmod{1} \\ L_{\leq 3} + L_{\leq 3} &\not\equiv \mathbb{R} \pmod{1}. \end{aligned}$$

Moreover $L_{\leq 3} + L_{\leq 3}$ and $L_{\leq 3} + L_{\leq 4}$ are not intervals.

It is shown that $L_{\leq 3} + L_{\leq 3}$ does not contain an interval of length ≥ 1 , therefore, it cannot be $\equiv \mathbb{R} \pmod{1}$. We also show that $L_{\leq 3} + L_{\leq 4}$ is not an interval, but are unable to decide whether or not it is $\equiv \mathbb{R} \pmod{1}$, so we

²Diverging partial quotients imply that a number is well-approximable, whereas bounded partial quotients imply badly approximable (see Chapter 3).

cannot confirm whether the result is optimal (note here that $L_{\leq 2}$ only contains one element, as opposed to $F_{\leq 2}$, so it is of no interest to consider this as a summand).

For digits bounded from below, we show the following result

Theorem (G.-Lee, 2026, Theorem 2.3 in [46]). *Let $n \in \mathbb{N}$ and $k_1, k_2, \dots, k_n \in \mathbb{N} \setminus \{1\}$. Suppose*

$$\sum_{j=1}^n \frac{k_j - 1}{k_j(k_j - 1) - 1} \geq 1,$$

and suppose for any $j \in \{2, \dots, n\}$ we have

$$k_{j-1} \leq k_j \leq k_{j-1}^2 + 2k_{j-1} + 2.$$

Then

$$L_{\geq k_1} + L_{\geq k_2} + \dots + L_{\geq k_n} \equiv \mathbb{R} \pmod{1}.$$

As a consequence, we get for instance that

$$L_{\geq 3} + L_{\geq 3} + L_{\geq 5} \equiv \mathbb{R} \pmod{1} \quad \text{and} \quad L_{\geq 3} + L_{\geq 4} + L_{\geq 5} + L_{\geq 6} \equiv \mathbb{R} \pmod{1}.$$

We also get (similar to Cusick-Lee in the continued fraction case):

$$\underbrace{L_{\geq k} + \dots + L_{\geq k}}_{k \text{ times}} \equiv \mathbb{R} \pmod{1} \quad \text{while} \quad \underbrace{L_{\geq k} + \dots + L_{\geq k}}_{(k-1) \text{ times}} \not\equiv \mathbb{R} \pmod{1}.$$

It is easy to see that $\underbrace{L_{\geq k} + \dots + L_{\geq k}}_{(k-1) \text{ times}} \not\equiv \mathbb{R} \pmod{1}$, since the diameter of the set

$$\underbrace{L_{\geq k} + \dots + L_{\geq k}}_{(k-1) \text{ times}} \text{ is } 1 - \frac{k-2}{k^2-k-1} < 1.$$

We also prove a mixed result:

Theorem (G.-Lee, 2026, Theorem 2.5 in [46]).

$$L_{\leq k+2} + L_{\geq k} \equiv \mathbb{R} \pmod{1}.$$

Additionally we provide a lower bound for the Hausdorff dimension of $L_{\geq k}$ as well as for the sum $L_{\geq k} + L_{\geq k}$. These will be discussed in Chapter 2. We were unable to modify Shulga's algorithm to provide results on e.g. intervals in $L_{\geq k} + L_{\geq k}$, whenever it is not $\equiv \mathbb{R} \pmod{1}$.

Expressible sets It is both a natural instinct and in fact a helpful tool to put restrictions on the "digits" in the expressible set defined in §1.2. Fix a sequence of natural numbers $(a_n)_{n=1}^\infty$ with $a_n \geq n^{1+\varepsilon}$ for each n , for some $\varepsilon > 0$. Take for each n a subset $\mathbb{D}_n$ of the natural numbers, this set is the set of allowable "digits" preceding a_n . We define the missing-digits set by:

$$\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty) := \left\{ \sum_{n=1}^\infty \frac{1}{d_n a_n} : d_n \in \mathbb{D}_n \text{ for each } n \in \mathbb{N} \right\}.$$

The only previous results we are aware of regarding the restricted digits set of an expressible set are due to Hančl and Šustek [52], who considered integer sequences $(a_n)_{n=1}^\infty$ with digit sets $\mathbb{D}_n = \{1, 2, \dots, K\}$ for some fixed natural number K .

In the preprint *Lower bound for the Hausdorff dimension of expressible sets* by the author, Hančl and Kristensen [44] we consider the following missing-digits expressible sets $\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty)$:³

$$a_n = b^n \text{ for some } b > 4$$

$$\text{with digits set } \mathbb{D}_n = \{1, 2, \dots, K\} \text{ for some natural number } K < \sqrt{b},$$

$$a_n = 2^{N^n} \text{ for some natural number } N \geq 2$$

$$\text{with digits set } \mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots < d_{n,m_n} \leq 2^{rN^n}\}$$

$$\text{where } m_n \leq 2^{sN^n} \text{ for } 0 < s \leq r < \frac{N-1}{2},$$

and

$$a_n = N^{2^n} \text{ for some natural number } N \geq 2$$

$$\text{with digits set } \mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots < d_{n,m_n} \leq N^{r2^n}\}$$

$$\text{where } m_n \leq N^{s2^n} \text{ for } 0 < s \leq r < \frac{1}{2}.$$

³or a slight strengthening of the last two, where we allow multiplication by some function, which does not alter the growth speed of a_n . See Theorems 3 and 4 in [44] included in Chapter 5.

Hausdorff Measure and Dimension

2.1 Construction

Many of the sets about which this thesis is concerned have Lebesgue measure zero. In an effort to assign some more meaningful value to their size, we will resort to using the Hausdorff dimension. Hausdorff dimension is defined in terms of Hausdorff measures. These were defined by Hausdorff based on a construction by Carathéodory. In this section we will briefly recount the construction.

For a more detailed delve into measure theory see [75] and for books on fractal geometry and Hausdorff measure and dimension, see [40] and [66] (or chapter 5 in [21]).

Hausdorff measure and dimension

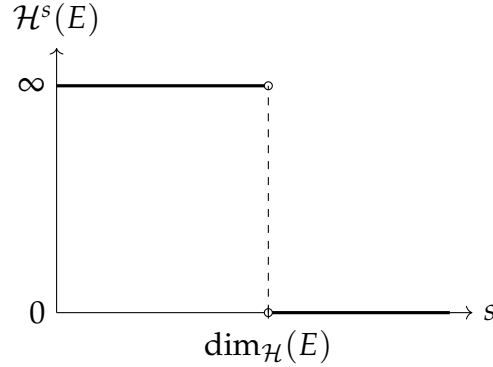
Intuitively the Hausdorff measure assigns a size to a set by covering it with a countable family of sets of bounded diameter δ (called a δ -cover), and then letting the diameter of the covers go to zero.

By diameter of a set $E \subseteq \mathbb{R}^n$, we mean

$$\text{diam}(E) := \sup\{\|x - y\| : x, y \in E\}.$$

Suppose E is a subset of $\mathbb{R}^n$ and let $s \geq 0$. For each $\delta > 0$ define

$$\mathcal{H}_\delta^s(E) = \inf \left\{ \sum_{n=1}^{\infty} \text{diam}(U_n)^s : 0 < \text{diam}(U_n) \leq \delta, \forall n \text{ and } E \subseteq \bigcup_{n=1}^{\infty} U_n \right\}$$

Figure 2.1: $s \mapsto \mathcal{H}^s(E)$

As δ decreases the collection of admissible covers becomes smaller, so the infimum is non-decreasing and thus the limit is well defined, so we may write

$$\mathcal{H}^s(E) := \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(E) = \sup_{\delta > 0} \mathcal{H}_\delta^s(E),$$

which lies in $[0, \infty]$ and can be shown to be a measure. This is called the *Hausdorff s -measure*. Hausdorff measures, like Lebesgue measures, generalise the notions of length, area and volumes, and in fact for any natural number n , the Hausdorff n -measure is a constant multiple of the n -dimensional Lebesgue measure.

For a fixed set E the function $s \mapsto \mathcal{H}^s(E)$ decreases, and if $\mathcal{H}^s(E) < \infty$, then $\mathcal{H}^t(E) = 0$ for $t > s$ and similarly, if $\mathcal{H}^s(E) > 0$, then $\mathcal{H}^t(E) = \infty$ for $t < s$.

Plotting $s \mapsto \mathcal{H}^s(E)$ (see Figure 2.1) we see, that it takes the value ∞ until a critical point, where it jumps to 0. That critical point is called the *Hausdorff dimension*, is denoted $\dim_{\mathcal{H}}(E)$ and is defined as

$$\dim_{\mathcal{H}}(E) := \inf\{s \geq 0 : \mathcal{H}^s(E) = 0\} = \sup\{s \geq 0 : \mathcal{H}^s(E) = \infty\}.$$

$\mathcal{H}^{\dim_{\mathcal{H}}(E)}(E)$ can take any value in $[0, \infty]$.

2.2 Cantor sets and iterated function systems

In this section we generalise the notion of a Cantor set. The simplest example is the middle third Cantor set $C_{\frac{1}{3}}$, which we constructed in §1.3.

We define a (generalised) Cantor set C as a subset of the real numbers of the form

$$C = I \setminus \bigcup_{i=1}^{\infty} O_i,$$

where I is a bounded closed interval and $\{O_i : i \in \mathbb{N}\}$ is a countable collection of disjoint open subintervals of I .

In this section we will introduce three different ways to construct a Cantor set. Each type of representation will be used in the papers *Lower bound for the Hausdorff dimension of expressible sets* by the author, Hančl and Kristensen [44] (see Chapter 5) and in *Decomposition of real numbers into sums of Liuroth sets* by the author and Lee [46] (see Chapter 4) in order to apply various lemmata. In §2.2 we construct all relevant Cantor sets, using the most appropriate construction. In §2.2 we define the notion of *thickness* of a Cantor set represented by *Construction 2* below.

Construction 1 Let E_0 be a bounded interval. Set $\mathcal{E}_0 = \{E_0\}$ and assume for any $n \geq 1$, that there exists a finite family of disjoint closed subintervals of E_0 , denoted $\mathcal{E}_n$, such that any interval in $\mathcal{E}_n$ is contained in exactly one interval from $\mathcal{E}_{n-1}$ and such that any interval in $\mathcal{E}_n$ contains at least two intervals from $\mathcal{E}_{n+1}$. Assume furthermore, that the maximum length of the intervals in $\mathcal{E}_n$ goes to zero as $n \rightarrow \infty$. Then the set

$$\mathcal{K} := \bigcap_{n=1}^{\infty} \underbrace{\bigcup_{U \in \mathcal{E}_n} U}_{E_n} = \bigcap_{n=1}^{\infty} E_n$$

is a Cantor set. See Figure 2.2 for a visualisation of this construction.

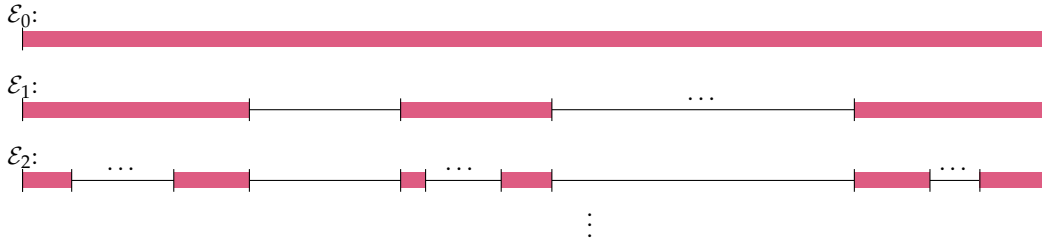


Figure 2.2: Generalised Cantor set with m_n sub-intervals in layer n and gaps of size $\geq \varepsilon_n$

The middle thirds Cantor set $C_{\frac{1}{3}}$ can be realised as a set of this form, by putting $\mathcal{E}_0 = \{[0, 1]\}$, $\mathcal{E}_1 = \{[0, \frac{1}{3}], [\frac{2}{3}, 1]\}$, $\mathcal{E}_2 = \{[0, \frac{1}{9}], [\frac{2}{9}, \frac{1}{3}], [\frac{2}{3}, \frac{7}{9}], [\frac{8}{9}, 1]\}$, etc.

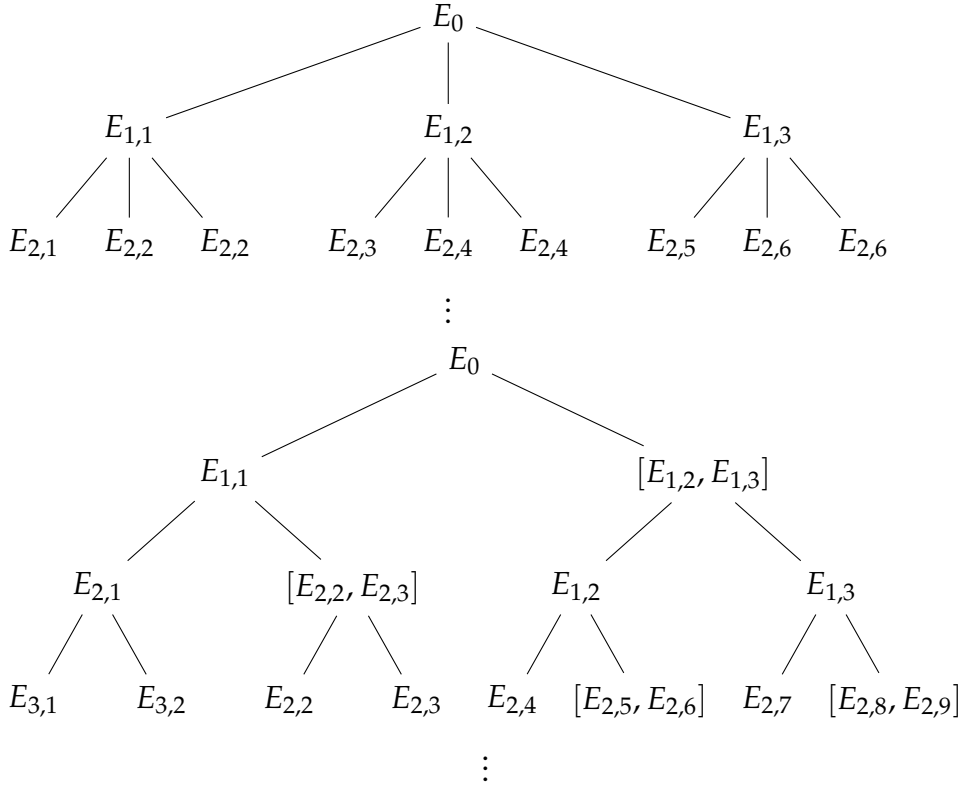


Figure 2.3: Here we consider a Cantor set (illustrated as a tree, where each node is a closed subinterval of its parent-node) first constructed using *Construction 1* with $\mathcal{E}_1 = \{E_{1,1}, E_{1,2}, E_{1,3}\}$, $\mathcal{E}_2 = \{E_{2,1}, E_{2,2}, E_{2,3}, E_{2,4}, E_{2,5}, E_{2,6}\}$ etc. Next we have the same Cantor set represented via *Construction 2*. By e.g. $[E_{1,2}, E_{1,3}]$ we mean the closed interval $[\min(E_{1,2}), \max(E_{1,3})]$.

Construction 2 In [47], Hall constructs Cantor sets in the following way. Let $I^1 = [x, x + a]$ be a closed and bounded interval. Remove some open subinterval $O^1 = (x + a_1, x + a_1 + a_{1,2})$, then two closed intervals remain $I^2 \cup I^3 = [x, x + a_1] \cup [x + a_1 + a_{1,2}, x + a]$. Divide I^2 and I^3 and every subsequent closed interval in the same way. The resulting set is a Cantor set, where the intervals making up level n are $\{I^{2^n}, I^{2^n+1}, \dots, I^{2^{n+1}-1}\}$ and where for each i , O^i is the immediate gap in I^i (see Figure 2.14)

At first glance, Hall's construction is more limited than the former, since the former allows the removal of several (but finitely many) open subintervals from each closed interval, where the latter allows only the removal of one. However, both constructions result in exactly the same sets (see Figure 2.3 for a demonstration of this phenomenon). We introduce both, because the

lemmata we will apply later use different constructions.

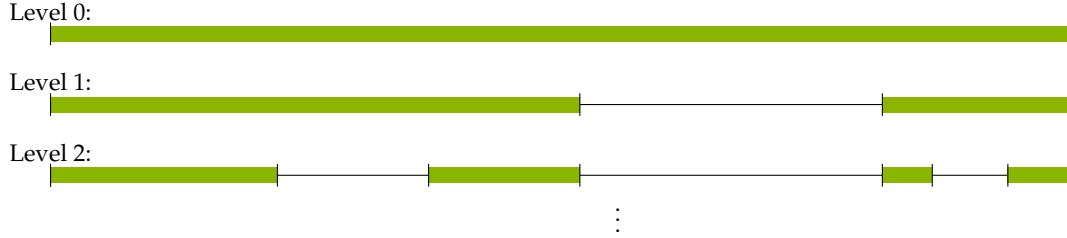


Figure 2.4: Cantor set with 2^n sub-intervals in layer n .

Both Hlavka and Astels [4, 5, 56] use a construction very similar to this – forming level $n + 1$ by removing *one* open subinterval from each closed interval in level n .

Iterated function systems Some Cantor sets can also be realised as an iterated function system. If a fractal is self-similar, meaning that it is made up of copies of itself (for instance the middle third Cantor set is a union of two smaller copies of itself $C_{\frac{1}{3}} = f_1(C_{\frac{1}{3}}) \cup f_2(C_{\frac{1}{3}})$, where $f_1(x) = \frac{1}{3}x$ and $f_2(x) = \frac{1}{3}x + \frac{2}{3}$) it can be described as an iterated function system. This is not limited to fractals on the real line, but also includes sets such as the von Koch curve, the Sierpinski gasket, Cantor dust, (see Figure 2.2) and many more.

A map $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called a *similitude* if there exists a constant $0 < c < 1$ such that $|f(x) - f(y)| = c|x - y|$ for all $x, y \in \mathbb{R}^n$.

A finite family of similitudes $\{f_1, \dots, f_m\}$ with $m \geq 2$ is called an *iterated function system*.

There exists a unique compact set $\mathcal{K} \subseteq \mathbb{R}^n$ (called the *attractor* for the iterated function system) that is invariant under $\{f_1, \dots, f_m\}$, meaning that

$$\mathcal{K} = \bigcup_{i=1}^m f_i(\mathcal{K}).$$

If the iterated function system satisfies, the *open set condition*, namely that there is a open set O such that

$$\bigcup_{i=1}^m f_i(O) \subset O \quad \text{and} \quad f_i(O) \cap f_j(O) = \emptyset \quad \text{for } i \neq j,$$

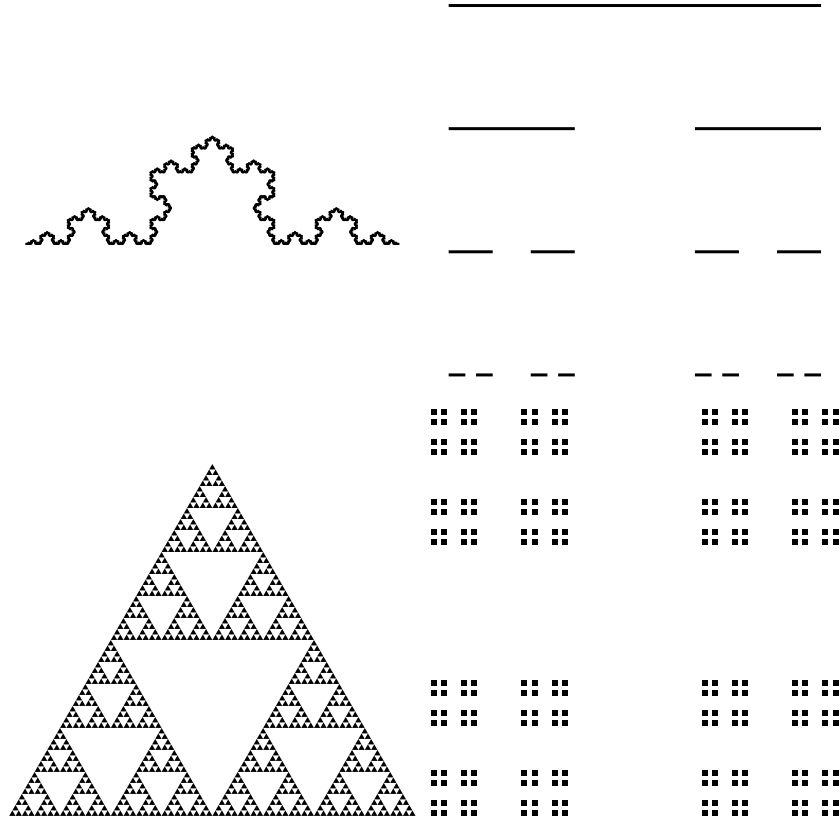


Figure 2.5: Examples of self-similar fractals defined by iterated function systems.

then the invariant set $\mathcal{K}$ is self-similar and $0 < \mathcal{H}^s(\mathcal{K}) < \infty$. Here s is the unique number that solves

$$\sum_{i=1}^n r_i^s = 1,$$

where for each i , r_i is the contraction ratio of f_i . For a proof of this see e.g. Falconer's book [40].

For instance, since

$$\left(\frac{1}{3}\right)^{\frac{\log(2)}{\log(3)}} + \left(\frac{1}{3}\right)^{\frac{\log(2)}{\log(3)}} = 1,$$

we know that $\dim_{\mathcal{H}} C_{\frac{1}{3}} = \frac{\log(2)}{\log(3)}$.

If we know how to calculate the dimension of the attractor of an iterated function system, we can consider the sum of two such fractals. For this we have the following result from [70]:

Theorem 1. *Let $\mathcal{K}$ be the attractor of the iterated function system $\{r_i x + t_i\}_{i=1}^n$ and $\mathcal{K}'$ the attractor of the iterated function system $\{r'_i x + t'_i\}_{i=1}^{n'}$. If there exists j and j' such that $\frac{\log(|r_j|)}{\log(|r'_{j'}|)}$ is irrational, then*

$$\dim_{\mathcal{H}}(\mathcal{K} + \mathcal{K}') = \min\{\dim_{\mathcal{H}}(\mathcal{K}) + \dim_{\mathcal{H}}(\mathcal{K}'), 1\}.$$

If the Cantor set in question is not given as the attractor of an iterated function system, we still have a shot at a lower bound for the Hausdorff dimension. The following application of the mass distribution principle is stated as Example 4.6 in [40], but is stated here and in [44] as a Theorem.

Theorem 2. *With the notation from Construction 1, assume that there exists a positive integer n_0 such that for all $n \geq n_0$, each interval from $\mathcal{E}_{n-1}$ contains at least $m_n \geq 2$ intervals from $\mathcal{E}_n$ and so that every gap in level n is $\geq \varepsilon_n$. Then*

$$\dim_{\mathcal{H}}(\mathcal{K}) \geq \liminf_{n \rightarrow \infty} \frac{\log(m_1 \dots m_n)}{-\log(m_{n+1} \varepsilon_{n+1})}.$$

Examples

In this section we construct the Cantor sets from the paper *Lower bound for the Hausdorff dimension of expressible sets* [44] and the preprint *Decomposition of real numbers into sums of Lüroth sets* [46].

Continued fractions For a finite set $\mathcal{B} = \{b_1, \dots, b_n\}$ of natural numbers in ascending order, with $n \geq 2$, we construct the set

$$F_{\mathcal{B}} := \{x = [a_1, a_2, \dots] : a_i \in \mathcal{B} \ \forall i \in \mathbb{N}\}$$

of numbers with partial quotients restricted to $\mathcal{B}$. We use the construction from Astels' thesis [6], which results in a derivation of the same type as *Construction 2*.

The 0th level of the Cantor set is the interval from the smallest to the largest element of $F_{\mathcal{B}}$, i.e.

$$I^1 = \left[\overline{[b_n, b_1]}, \overline{[b_1, b_n]} \right].$$

For two real numbers a, b , by $[[a, b]]$ we mean

$$[\min\{a, b\}, \max\{a, b\}].$$

In general, for $i < n$, any interval of the form

$$\left[[a_1, \dots, a_r, b_i, \overline{b_n, b_1}], [a_1, \dots, a_r, b_n, \overline{b_1, b_n}] \right] \quad (2.1)$$

splits into two intervals

$$\begin{aligned} & \left[[a_1, \dots, a_r, b_i, \overline{b_n, b_1}], [a_1, \dots, a_r, b_i, \overline{b_1, b_n}] \right] \\ &= \left[[a_1, \dots, a_r, b_i, \overline{b_1, b_n}], [a_1, \dots, a_r, b_i, \overline{b_n, b_1}] \right] \end{aligned}$$

and

$$\left[[a_1, \dots, a_r, b_{i+1}, \overline{b_n, b_1}], [a_1, \dots, a_r, b_n, \overline{b_1, b_n}] \right].$$

Both of these are again of the form (2.1), so the procedure continues.

Using this recipe, let us find a derivation of $F_{\leq 4}$. Here $\mathcal{B} = \{1, 2, 3, 4\}$ and

$$I^1 = [[4, 1], [1, 4]].$$

The next level then consists of

$$I^2 \cup I^3 = [[4, 1], [2, 4, 1]] \cup [[1, 1, 4], [1, 4]]$$

and level 2 is

$$\begin{aligned} & I^4 \cup I^5 \cup I^6 \cup I^7 \\ &= [[4, 1], [3, 4, 1]] \cup [[2, 1, 4], [2, 4, 1]] \cup [[1, 1, 4], [1, 1, 1, 4]] \cup [[1, 2, 4, 1], [1, 4]], \end{aligned}$$

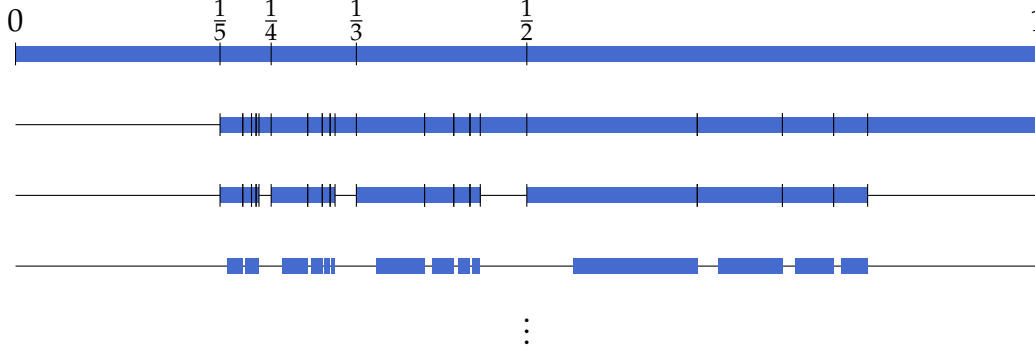
and so on.

$F_{\mathcal{B}}$ can also be realised as the attractor of the iterated function system consisting of the similitudes

$$f_{b_i}(x) = \frac{1}{b_i + x}.$$

For a fixed natural number k , we present a construction of $F_{\geq k}$ which appears e.g. in [28]. Level 0 of the Cantor set is

$$[[], [k]] = \left[0, \frac{1}{k} \right].$$

Figure 2.6: $F_{\leq 4}$ as the attractor of an IFS.

From here on, every interval of type 1 (where $n \geq 0$ is even), which are of the form

$$[[a_1, \dots, a_n], [a_1, \dots, a_n, a_{n+1}]],$$

is split into the following two intervals:

$$[[a_1, \dots, a_n], [a_1, \dots, a_n, a_{n+1} + 1]] \cup [[a_1, \dots, a_n, a_{n+1}, k], [a_1, \dots, a_n, a_{n+1}]].$$

Here the left-most interval is again of type 1. The right-most interval is of type 2, meaning of the form

$$[[a_1, \dots, a_{n-1}, a_n], [a_1, \dots, a_{n-1}]],$$

where again $n \geq 0$ is even. Intervals of this type split in the following way:

$$[[a_1, \dots, a_{n-1}, a_n], [a_1, \dots, a_{n-1}, a_n, k]] \cup [[a_1, \dots, a_{n-1}], [a_1, \dots, a_{n-1}, a_n + 1]].$$

Again these are of type 1 and type 2 respectively.

The first few levels look like (see also Figure 2.7 for $F_{\geq 2}$):

$$[[], [k]]$$

$$[[], [k + 1]] \cup [[k, k], [k]]$$

$$[[], [k + 2]] \cup [[k + 1, k], [k + 1]] \cup [[k, k], [k, k, k]] \cup [[k, k + 1], [k]]$$

Good [43] proved that for any $k \geq 20$

$$\frac{1}{2} + \frac{1}{2 \log(k + 2)} \leq \dim_{\mathcal{H}} F_{\geq k} < \frac{1}{2} + \frac{\log(\log(k - 1))}{2 \log(k - 1)}.$$

For $F_{\leq k}$ see [31].

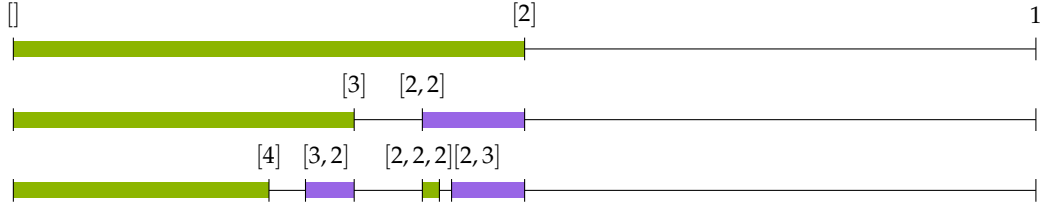


Figure 2.7: $F_{\geq 2}$ using *Construction 2* (intervals of type 1 are marked in green and type 2 are marked in purple).

Lüroth expansion We start by constructing $L_{\leq 3}$, since this is the simplest of the Lüroth sets. We construct it in two different ways, for two different purposes. First as a Cantor set using *Construction 2* (See Figure 2.8) and then as the attractor of an iterated function system.

For an integer k , by $\bar{k}$ we mean $k, k, \dots$ repeated infinitely. With this notation we have e.g. $[\bar{2}] = [2, 2, \dots] = \sum_{n=1}^{\infty} \frac{1}{2^n} = 1$.

Let $I^1 = [[\bar{3}], [\bar{2}]] = [\frac{2}{3}, 1]$. In level 1 we split into the intervals

$$I^2 \cup O^1 \cup I^3 = [[\bar{3}], [3, \bar{2}]] \cup ([3, \bar{2}], [2, \bar{3}]) \cup [[2, \bar{3}], [\bar{2}]].$$

The leftmost interval spans from the smallest element of $L_{\leq 3}$ with first digit 3 to the largest element of $L_{\leq 3}$ with first digit 3. The second interval does the same, but with 2 as first digit.

The rest of the construction goes the same way (see Figure 2.8).

An interval

$$I^{2^n+i} = [[d_1, \dots, d_n, \bar{3}], [d_1, \dots, d_n, \bar{2}]]$$

in level n (for $0 \leq i \leq 2^n - 1$) splits into

$$I^{2^{n+1}+2i} \cup I^{2^{n+1}+2i+1} = [[d_1, \dots, d_n, 3, \bar{3}], [d_1, \dots, d_n, 3, \bar{2}]] \cup [[d_1, \dots, d_n, 2, \bar{3}], [d_1, \dots, d_n, 2, \bar{2}]].$$

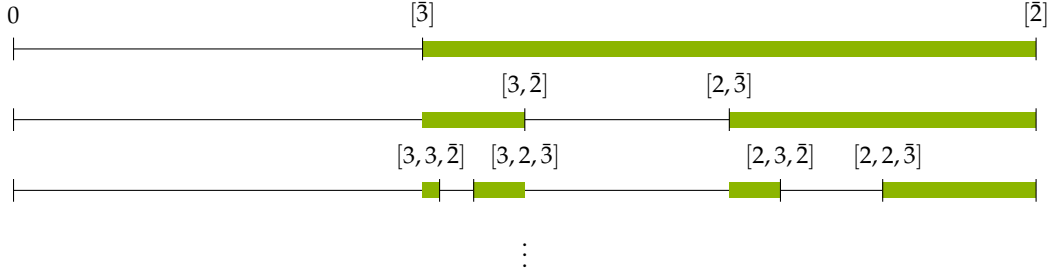
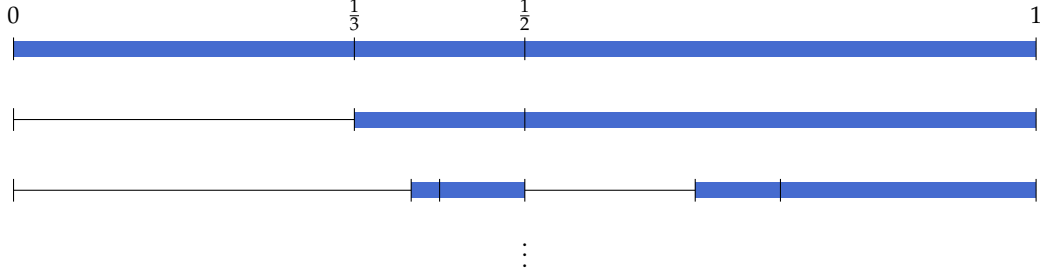
Alternatively we can construct $L_{\leq 3}$ as the attractor of an iterated function system (see Figure 2.9). Consider the similitudes

$$f_2(x) = \frac{1}{2}x + \frac{1}{2} \quad \text{and} \quad f_3(x) = \frac{1}{6}x + \frac{1}{3}.$$

Then

$$L_{\leq 3} = f_2(L_{\leq 3}) \cup f_3(L_{\leq 3}).$$

This IFS construction can easily be generalised to sets of real numbers in $[0, 1]$ whose Lüroth expansion has digits in some finite subset of the natural

Figure 2.8: $L_{\leq 3}$ using Construction 2.Figure 2.9: $L_{\leq 3}$ as the attractor of an IFS.

number, say $\mathcal{A} \subset \mathbb{N} \setminus \{1\}$. Consider for $d \in \mathcal{A}$ the similitudes

$$f_d(x) := \frac{1}{d(d-1)}x + \frac{1}{d}.$$

Then

$$L_{\mathcal{A}} = \bigcup_{d \in \mathcal{A}} f_d(L_{\mathcal{A}}).$$

For the *Construction 2* approach of e.g. $L_{\leq 4}$ and $L_{\leq 5}$ etc., we have to be more deliberate. Consider the following Lüroth set for $N_1 < N_2 \in \mathbb{N} \setminus \{1\}$:

$$\{[d_1, d_2, \dots] : d_i \in \{N_1, N_1 + 1, \dots, N_2\} \text{ for each } i \in \mathbb{N}\}$$

Level 0 of the Cantor set is then the following interval

$$I^1 = [[\bar{N}_2], [\bar{N}_1]]$$

Level 1 consists of

$$I^2 = [[\bar{N}_2], [N_1 + 1, \bar{N}_1]] \quad \text{and} \quad I^3 = [[N_1, \bar{N}_2], [\bar{N}_1]]$$

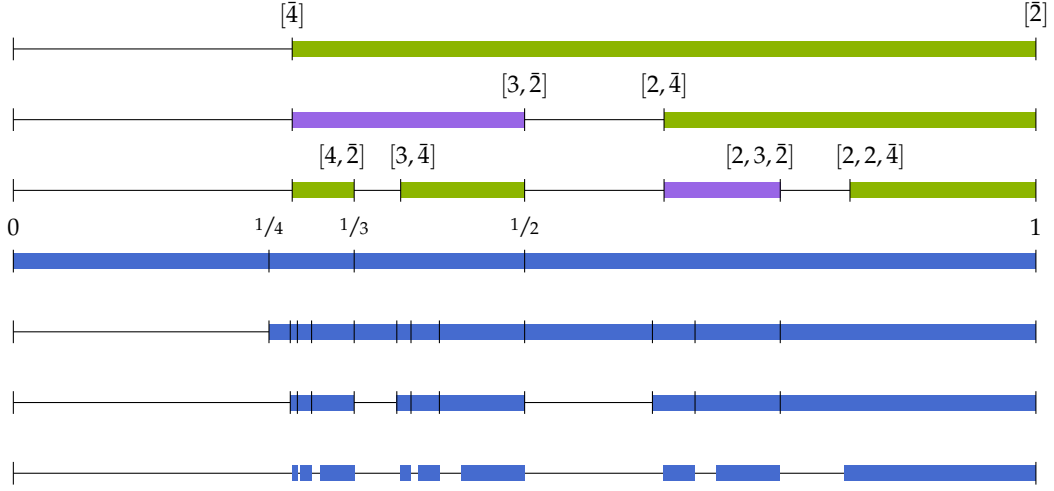


Figure 2.10: $L_{\leq 4}$ using *Construction 2* and as the attractor of an IFS.

If $N_1 + 1 \neq N_2$, then level 2 consists of

$$\begin{aligned} I^4 &= [[\bar{N}_2], [N_1 + 2, \bar{N}_1]], \\ I^5 &= [[N_1 + 1, \bar{N}_2], [N_1 + 1, \bar{N}_1]], \\ I^6 &= [[N_1, \bar{N}_2], [N_1, N_1 + 1, \bar{N}_1]], \\ \text{and } I^7 &= [[N_1, N_1, \bar{N}_2], [\bar{N}_1]]. \end{aligned}$$

For each $1 \leq i \leq N_2 - N_1$, the left-most intervals of level i are

$$I^{2^i} = [[\bar{N}_2], [N_1 + i, \bar{N}_1]]$$

and

$$I^{2^{i+1}} = [[N_1 + i - 1, \bar{N}_2], [N_1 + i - 1, \bar{N}_1]].$$

When we have an interval of the form

$$[[d_1, d_2, \dots, d_n, \bar{N}_2], [d_1, d_2, \dots, d_n, \bar{N}_1]],$$

we treat it exactly the same as $I^1 = [[\bar{N}_2], [\bar{N}_1]]$, i.e. it splits into

$$[[d_1, \dots, d_n, \bar{N}_2], [d_1, \dots, d_n, N_1 + 1, \bar{N}_1]] \cup [[d_1, \dots, d_n, N_1, \bar{N}_2], [d_1, \dots, d_n, \bar{N}_1]].$$

Figures 2.10 and 2.11 show this construction for $L_{\leq 4}$ and $L_{\leq 5}$.

For the sets $L_{\geq k}$ ($k \in \mathbb{N} \setminus \{1\}$) we do not represent these as Cantor sets. Instead we look at a subset $L_{\{k, \dots, N\}} \subset L_{\geq k}$ for some $N > k$, consisting of the real numbers in $[0, 1]$, whose Lüroth expansion digits lie between k and

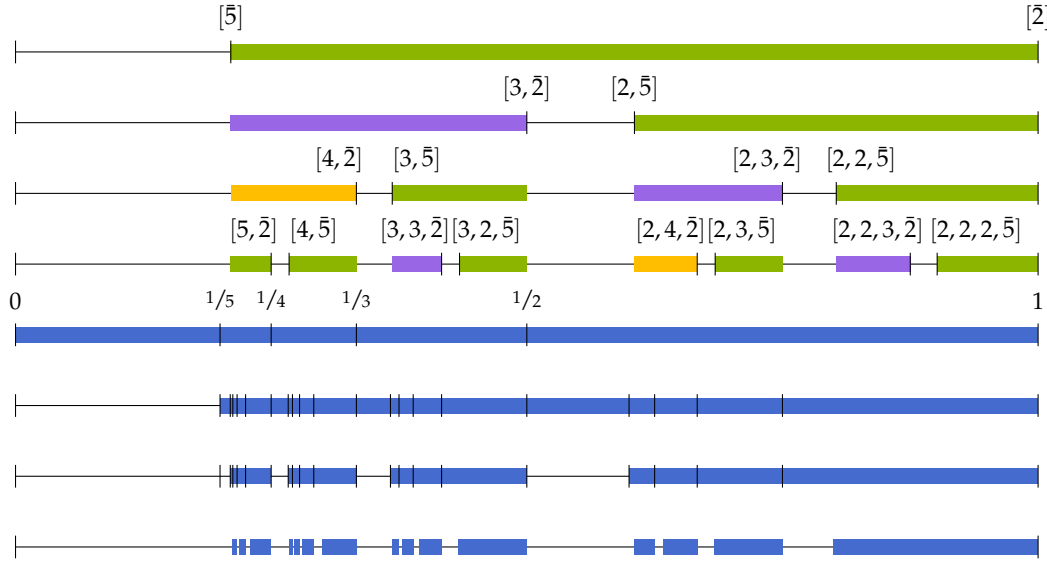


Figure 2.11: $L_{\leq 5}$ using *Construction 2* and as the attractor of an IFS.

N . This set we have already constructed in both ways above. It suffices to consider this set, since if we show for some values

$$N_1(k_1), \dots, N_n(k_n) \in \mathbb{N} \setminus \{1\},$$

that

$$L_{\{k_1, \dots, N_1\}} + \dots + L_{\{k_n, \dots, N_n\}} \equiv \mathbb{R} \pmod{1},$$

then

$$L_{\geq k_1} + \dots + L_{\geq k_n} \equiv \mathbb{R} \pmod{1}$$

as well.

The constructions described above of $L_{\leq 3}$, $L_{\leq 4}$ and $L_{\leq 5}$ are good enough to apply Hlavka's [56] result to obtain Theorem 2.1 in [46]. For Theorem 2.3 and 2.5 in [46] on sums of $L_{\geq k}$ this construction also suffices.

For each $k \geq 3$ we consider $L_{\{k, k+1, \dots, k^3\}} \subset L_{\geq k}$ derived using the above construction, as well as $L_{\{3, \dots, 16\}}$, $L_{\{4, \dots, 45\}}$, $L_{\{5, \dots, 96\}}$, and $L_{\{6, \dots, 175\}}$ again using the construction above. Here we apply instead Astels [4] (see §2.2).

The construction is not unique, and we could have made different choices and ended up with the same set. Figure 9 (a) and (b) in [46] (Chapter 4) shows a derivation of $L_{\{3, \dots, 26\}}$, which is different than the one detailed above. Using this allows us to show $L_{\geq 3} + L_{\geq 3} + L_{\geq 3} \equiv \mathbb{R} \pmod{1}$ using a weaker result than the one needed for $L_{\{3, \dots, 16\}}$.

Expressible sets In the preprint *Lower bound for the Hausdorff dimension of expressible sets* [44] we consider various missing digits sets, but their constructions are all the same, and of the same type as *Construction 1*.

For a positive increasing sequence $(a_n)_{n=1}^\infty$ and a sequence of finite subsets of the natural numbers $(\mathbb{D}_n)_{n=1}^\infty$, with $\mathbb{D}_n = \{1 = d_{n,1} < \dots < d_{n,m_n}\}$, the expressible set

$$\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty) = \left\{ \sum_{n=1}^\infty \frac{1}{d_n a_n} : d_n \in \mathbb{D}_n \text{ for each } n \in \mathbb{N} \right\},$$

is constructed as follows.

Level 0 is the interval

$$\left[\sum_{n=1}^\infty \frac{1}{d_{n,m_n} a_n}, \sum_{n=1}^\infty \frac{1}{a_n} \right]$$

There are m_1 intervals in level 1, each spanning from the smallest to the largest element of $\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty)$ with first digit equal to $d_{1,i}$ for each $1 \leq i \leq m_1$. Each interval in level 1 is of the form

$$\left[\frac{1}{d_{1,i} a_1} + \sum_{n=2}^\infty \frac{1}{d_{n,m_n} a_n}, \frac{1}{d_{1,i} a_1} + \sum_{n=2}^\infty \frac{1}{a_n} \right]$$

for $1 \leq i \leq m_1$. We repeat this procedure, so the intervals in level 2 are of the form

$$\left[\frac{1}{d_{1,i} a_1} + \frac{1}{d_{2,j} a_2} + \sum_{n=3}^\infty \frac{1}{d_{n,m_n} a_n}, \frac{1}{d_{1,i} a_1} + \frac{1}{d_{2,j} a_2} + \sum_{n=3}^\infty \frac{1}{a_n} \right]$$

for $1 \leq i \leq m_1$ and $1 \leq j \leq m_2$. And so on (See Figure 2.12).

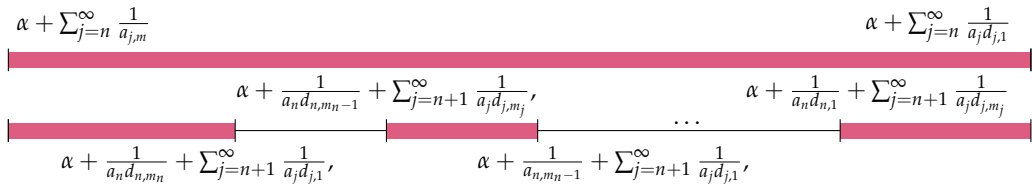


Figure 2.12: Splitting an interval in level $n - 1$ into subintervals in level n of the Cantor set $\mathcal{K}$ (Here α is of the form $\alpha = \frac{1}{a_1 d_{1,k_1}} + \dots + \frac{1}{a_{n-1} d_{1,k_{n-1}}}$,

$1 \leq k_i \leq m_i$ for each $1 \leq i \leq n - 1$).

Of course, we are not guaranteed, that the end-points of these intervals do not overlap. This is a problem that should be dealt with, by balancing the size of the digits $d_{n,i}$ with the growth of the sequence $(a_n)_{n=1}^\infty$.

For an exponential sequence $a_n = b^n$ with $b > 4$, take the digit sets to be $\mathbb{D}_n = \{1, 2, \dots, K\}$ for some integer K with $K^2 < b$. This guarantees gaps between the intervals in each layer, meaning the above procedure does indeed produce a Cantor set, where each interval in level n splits into K intervals in level $n + 1$. Furthermore, we can find a lower bound on the gap size, which we need in order to apply Theorem 2.

Consider now a doubly exponential sequence of the form $a_n = 2^{N^n} f(n)$ for some $N \in \mathbb{N}_{\geq 2}$ and some function $1 \leq f(n) \leq 2^{(N-1)^n}$ and the set of digits $\mathbb{D}_n = \{1 = d_{n,1} < \dots < d_{n,m_n}\} \subset \mathbb{N}$, where we assume that for $0 < s \leq r \leq \frac{N-1}{2}$ and for two functions $1 \leq g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq 2^{(N-1)^n}$, that m_n is of the form

$$m_n = \left\lfloor 2^{sN^n} g(n) \right\rfloor$$

and that

$$d_{n,m_n} \leq 2^{rN^n} h(n).$$

These assumptions on the number and size of the digits guarantee, that the expected gaps do indeed occur in the Cantor set.

The last type of sequence considered in [44] are doubly exponential sequences of the form $a_n = N^{2^n} f(n)$ for some $N \in \mathbb{N}_{\geq 2}$ and $1 \leq f(n) \leq N^{(2-\eta)^n}$, for some $0 < \eta < 1$. Similarly to above, let $0 < s \leq r \leq \frac{1}{2}$ and $1 \leq g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq N^{(2-\eta)^n}$ and assume, that the digit sets looks like

$$\mathbb{D}_n = \{1 = d_{n,1} < \dots < d_{n,m_n} \leq N^{r2^n} h(n)\} \subset \mathbb{N}$$

where

$$m_n \leq \left\lfloor N^{s2^n} g(n) \right\rfloor.$$

In [52] Hančl and Šustek similarly consider boundedly expressible sets, here of the form

$$\mathcal{K} \left((b_n + 2^{2^{\log_2(n)}})_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty} \right).$$

If $(b_n)_{n=1}^{\infty}$ is a non-decreasing sequence of natural numbers such that $b_n = O\left(2^{2^{\frac{n}{2}}}\right)$ and $\mathbb{D}_n = \{1, \dots, K\}$, then the Hausdorff dimension of the expressible set is zero.

In [53] Hančl and Šustek give growth conditions on a sequence of real numbers $(a_n)_{n=1}^{\infty}$, which guarantees Hausdorff dimension zero of the expressible set $\mathcal{K}((a_n)_{n=1}^{\infty})$. For rational sequences $\left(\frac{a_n}{b_n}\right)_{n=1}^{\infty}$ we refer to [55] for an upper bound for the dimension.

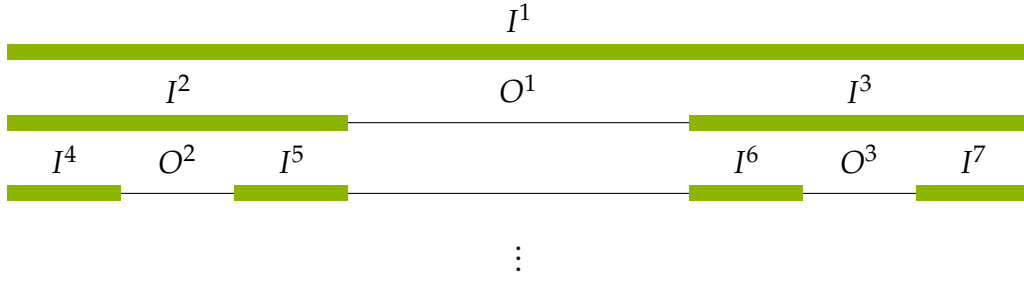


Figure 2.13: A Construction of a General Cantor Set

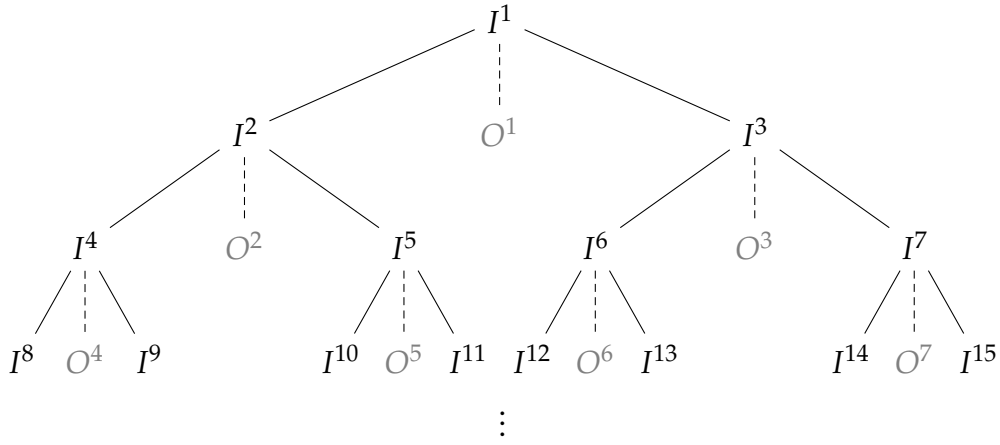


Figure 2.14: The same construction, represented as a tree

Thickness

Using the notation of *Construction 2*, for a Cantor set with construction consisting of intervals in level n given by $\{I^{2^n}, I^{2^n+1}, \dots, I^{2^{n+1}-1}\}$ and gaps given by $I^n = I^{2^n} \cup O^n \cup I^{2^{n+1}}$ (see Figure 2.14), we define some quantities associated with a Cantor set that say something about the size or "thickness" of it. Using these different quantities, we state some results from [4, 5, 47, 56] on sums of Cantor sets.

In [47], Hall uses just $|O^n|$, $|I^{2^n}|$, and $|I^{2^{n+1}}|$ (for $n \in \mathbb{N}$) to prove the following result:

Theorem 3 (Theorem 2.2 in [47]). *Let $\mathcal{K}$ and $\mathcal{K}'$ be two Cantor sets with constructions given by Construction 2 with intervals and gaps given by $((I^n)_{n=1}^\infty, (O^n)_{n=1}^\infty)$ and $((I'^n)_{n=1}^\infty, (O'^n)_{n=1}^\infty)$ respectively. Write $I^1 = [x, x + a]$, $I'^1 = [y, y + b]$, and $e = \min\{a, b\}$. Suppose for every $n \in \mathbb{N}$, that*

$$|O^n| \leq \min\{|I^{2^n}|, |I^{2^{n+1}}|\} \quad \text{and} \quad |O'^n| \leq \min\{|I'^{2^n}|, |I'^{2^{n+1}}|\}.$$

Then

$$[x + y, x + y + 2e] \cup [x + y + a + b - 2e, x + y + a + b] \subseteq \mathcal{K} + \mathcal{K}'.$$

In particular if $\frac{1}{3} \leq \frac{a}{b} \leq 3$, then

$$\mathcal{K} + \mathcal{K}' = I^1 + I'^1.$$

In [56], Hlavka defines two quantities

$$g(\mathcal{K}) := \sup_{n \in \mathbb{N}} \frac{|O^n|}{|I^n|}$$

and

$$h(\mathcal{K}) := \inf_{n \in \mathbb{N}} \min \left\{ \frac{|I^{2n}|}{|I^n|}, \frac{|I^{2n+1}|}{|I^n|} \right\}.$$

Both of these potentially depend on the specific construction of a Cantor set. g measures the maximal ratio of the gaps and h measures the minimal ratio of the subintervals compared to the interval in the previous layer (so in a sense they measure how much we throw away and how little we keep, which should give some idea of the "size" of the Cantor set).

Theorem 4 (Theorem 3 in [56]). *Let $\mathcal{K}$ and $\mathcal{K}'$ be two Cantor sets with constructions given by Construction 2 with intervals and gaps given by $((I^n)_{n=1}^\infty, (O^n)_{n=1}^\infty)$ and $((I'^n)_{n=1}^\infty, (O'^n)_{n=1}^\infty)$ respectively. If*

$$g(\mathcal{K})g(\mathcal{K}') \leq h(\mathcal{K})h(\mathcal{K}'), \quad g(\mathcal{K})|I^1| \leq |I'^1|, \quad \text{and} \quad g(\mathcal{K}')|I'^1| \leq |I^1|,$$

then

$$\mathcal{K} + \mathcal{K}' = I^1 + I'^1.$$

For sums of more than two Cantor sets, Hlavka also proved

Theorem 5 (Theorem 10 in [56]). *Let $\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_m$ be Cantor sets with constructions given by Construction 2 with intervals and gaps given by*

$$((I_i^n)_{n=1}^\infty, (O_i^n)_{n=1}^\infty)$$

for $1 \leq i \leq m$. If for all $i, j \in \{1, \dots, m\}$

$$h(\mathcal{K}_i) \leq \frac{|I_i^1|}{|I_j^1|}$$

and for any $i \in \{1, \dots, m\}$

$$g(\mathcal{K}_i) + h(\mathcal{K}_i) \leq \sum_{j=1}^m h(\mathcal{K}_j),$$

then

$$\mathcal{K}_1 + \mathcal{K}_2 + \dots + \mathcal{K}_m = I_1^1 + I_2^1 + \dots + I_m^1.$$

Astels [4] introduces the *thickness* of a Cantor set $\mathcal{K}$:

$$\tau(\mathcal{K}) := \sup \inf_{n \in \mathbb{N}} \min \left\{ \frac{|I^{2n}|}{|O^n|}, \frac{|I^{2n+1}|}{|O^n|} \right\}$$

where the sup is taken over all constructions of $\mathcal{K}$, as well as the *normalised thickness*:

$$\gamma(\mathcal{K}) := \frac{\tau(\mathcal{K})}{1 + \tau(\mathcal{K})}.$$

Both of these quantities are independent of the construction of $\mathcal{K}$.

The expression $\inf_{n \in \mathbb{N}} \min \left\{ \frac{|I^{2n}|}{|O^n|}, \frac{|I^{2n+1}|}{|O^n|} \right\}$ from the thickness looks similar to $\frac{h(\mathcal{K})}{g(\mathcal{K})}$, but they are not always equal. $\frac{h(\mathcal{K})}{g(\mathcal{K})}$ might be smaller than or equal to $\inf_{n \in \mathbb{N}} \min \left\{ \frac{|I^{2n}|}{|O^n|}, \frac{|I^{2n+1}|}{|O^n|} \right\}$, since the largest gap and the smallest remaining subinterval may occur in different levels.

Using these Astels states the following result in [4] concerning the size, dimension and thickness of sums of Cantor sets.

Theorem 6 (Theorem 2.4 in [4]). *Let $\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_m$ be Cantor sets with constructions given by Construction 2.*

- (1) *If $\gamma(\mathcal{K}_1) + \dots + \gamma(\mathcal{K}_m) \geq 1$, then $\mathcal{K}_1 + \mathcal{K}_2 + \dots + \mathcal{K}_m$ contains an interval. Otherwise it contains a Cantor set of thickness at least*

$$\frac{\gamma(\mathcal{K}_1) + \dots + \gamma(\mathcal{K}_m)}{1 - (\gamma(\mathcal{K}_1) + \dots + \gamma(\mathcal{K}_m))}.$$

Furthermore,

$$\dim_{\mathcal{H}}(\mathcal{K}_1 + \mathcal{K}_2 + \dots + \mathcal{K}_m) \geq \min \left\{ 1, \frac{\log(2)}{\log(1 + \frac{1}{\gamma(\mathcal{K}_1) + \dots + \gamma(\mathcal{K}_m)})} \right\}.$$

- (2) Let, for each $1 \leq i \leq m$, O_i be the gap of maximal size in $\mathcal{K}_i$ (that is $|O_i| = \sup_{n \in \mathbb{N}} |O_i^n|$). Suppose that for any $2 \leq i \leq m$ and $1 \leq j \leq i-1$ we have

$$|I_i^1| \geq |O_j|$$

and that for any $1 \leq i \leq m-1$

$$|I_1^1| + \cdots + |I_i^1| \geq |O_{i+1}|.$$

Then if $\gamma(\mathcal{K}_1) + \cdots + \gamma(\mathcal{K}_m) \geq 1$, we have

$$\mathcal{K}_1 + \mathcal{K}_2 + \cdots + \mathcal{K}_m = I_1^1 + I_2^1 + \cdots + I_m^1.$$

Otherwise

$$\tau(\mathcal{K}_1 + \mathcal{K}_2 + \cdots + \mathcal{K}_m) \geq \frac{\gamma(\mathcal{K}_1) + \cdots + \gamma(\mathcal{K}_m)}{1 - (\gamma(\mathcal{K}_1) + \cdots + \gamma(\mathcal{K}_m))}.$$

Theorems 3, 4, 5, and 6 can and have been used to prove a majority of the results on sums of missing digits set for continued fractions in §1.2. Furthermore Theorems 4, 5, and 6 were used by the author and Ying Wai Lee in [46] to prove the results on sums of missing digits sets for Lüroth expansion, detailed in §1.2. To apply these results to the relevant Cantor sets, use the constructions detailed in §2.2. Most of the work lies in finding a clever construction – one where the ratio of gaps to remaining subintervals is good. Moreover the construction should be such that we have a certain amount of self-similarity, otherwise calculating g , h and τ is very hard. After finding a good construction, the job is to calculate either g and h or τ , depending on which result is being applied.

2.3 Summary of my results on Hausdorff dimension

In this section we summarise the results on Hausdorff dimension found in the papers [44] and [46].

Lüroth sets The main results of the paper [46] are the ones on decomposition of real numbers as sums of numbers with bounded Lüroth digits (see §1.3). In the cases where we either know, that not all real numbers can be expressed as a sum or where we cannot determine if they can, we can instead try to decide the Hausdorff dimension of the sumset.

Consider sums of Lüroth sets with digits in some set $\mathcal{A}$, $L_{\mathcal{A}}$.

Theorem (G.-Lee, 2026, Theorem 2.9 in [46]). *For any finite $\mathcal{A}, \mathcal{B} \subset \mathbb{N} \setminus \{1\}$, we have*

$$\dim_{\mathcal{H}}(L_{\mathcal{A}} + L_{\mathcal{B}}) = \min(1, \dim_{\mathcal{H}} L_{\mathcal{A}} + \dim_{\mathcal{H}} L_{\mathcal{B}}).$$

This result is quite general, and can be used on both sums of $L_{\leq k}$ and $L_{\{k, \dots, N\}}$, if we find their dimensions.

We have for instance for $L_{\leq 3}$, that

$$\dim_{\mathcal{H}} L_{\leq 3} = 0.600967 \dots,$$

so by the above

$$\dim_{\mathcal{H}} L_{\leq 3} + L_{\leq 3} = 1.$$

Even though $L_{\leq 3} + L_{\leq 3} \not\equiv \mathbb{R} \pmod{1}$, it still has dimension 1. The same goes for $L_{\leq 3} + L_{\leq 4}$.

For $L_{\geq k}$ we prove

Theorem (G.-Lee, 2026, Theorem 2.10 in [46]). *For each $k \in \mathbb{N} \setminus \{1\}$*

$$\dim_{\mathcal{H}} L_{\geq k} > \frac{1}{2} + \frac{1}{2 \log(\max\{16, k\})}.$$

Apply a similar strategy to [43] to estimate a values s_k such that

$$\sum_{d=k}^N \left(\frac{1}{d(d-1)} \right)^{s_k} \geq 1.$$

For smaller values of k we can improve this. We have

$$\begin{aligned} \dim_{\mathcal{H}} L_{\geq 3} &> 0.8209, & \dim_{\mathcal{H}} L_{\geq 4} &> 0.7740, & \dim_{\mathcal{H}} L_{\geq 5} &> 0.7500, \\ \dim_{\mathcal{H}} L_{\geq 6} &> 0.7347, & \dim_{\mathcal{H}} L_{\geq 7} &> 0.7239, & \dim_{\mathcal{H}} L_{\geq 8} &> 0.7157. \end{aligned}$$

So the sumset $L_{\geq k} + L_{\geq k}$ also has full dimension:

Theorem (G.-Lee, 2026, Theorem 2.11 in [46]). *For each $k \in \mathbb{N} \setminus \{1\}$*

$$\dim_{\mathcal{H}} L_{\geq k} + L_{\geq k} = 1.$$

Expressible sets This section details the main results of the paper [44].

For the expressible sets defined in §2.2 we prove lower bounds for the Hausdorff dimension. A consequence of this, is that each expressible set must contain non-Liouville numbers, since those have dimension 0.

Firstly for an exponential sequence $(a_n)_{n=1}^{\infty}$ we have:

Theorem (G.-Hančl-Kristensen, 2026+, Theorem 2 in [44]). *Assume $a_n = b^n$ for some $b > 4$ and $\mathbb{D}_n = \{1, 2, \dots, K\}$ for every n , where $K^2 < b$. Then*

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty})) \geq \frac{\log(K)}{\log(b)}.$$

We consider also doubly exponential sequences 2^{N^n} and N^{2^n} :

Theorem (G.-Hančl-Kristensen, 2026+, Theorem 3 in [44]). *Let $N \in \mathbb{N}_{\geq 2}$ and $0 < s \leq r < \frac{N-1}{2}$. Take functions $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ satisfying*

$$1 \leq f(n), g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq 2^{(N-1)^n}$$

and $2^{sN^n} g(n) \leq 2^{rN^n} h(n)$. Define for each $n \in \mathbb{N}$

$$a_n = 2^{N^n} f(n),$$

$$m_n = \lfloor 2^{sN^n} g(n) \rfloor,$$

and

$$\mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots, d_{n,m_n} \leq 2^{rN^n} h(n)\} \subset \mathbb{N}.$$

Then

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty})) \geq \frac{s}{(N-1)(N-s)}.$$

Theorem (G.-Hančl-Kristensen, 2026+, Theorem 4 in [44]). *Let $N \in \mathbb{N}_{\geq 2}$ and $0 < s \leq r < \frac{1}{2}$. Let $0 < \eta < 1$. Take functions $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ satisfying $1 \leq f(n), g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq N^{(2-\eta)^n}$ and $N^{s2^n} g(n) \leq N^{r2^n} h(n)$. Define for each $n \in \mathbb{N}$*

$$a_n = N^{2^n} f(n),$$

$$m_n = \lfloor N^{s2^n} g(n) \rfloor,$$

and

$$\mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots, d_{n,m_n} \leq N^{r2^n} h(n)\} \subset \mathbb{N}.$$

Then

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty})) \geq \frac{s}{2-s}.$$

We do not go any higher than doubly exponential, since a triply exponential sequence guarantees that all elements of the expressible set are Liouville by Erdős' result, since $\limsup_{n \rightarrow \infty} a_n^{1/t^n} = \infty$ for each $t \in \mathbb{N}$ (see §1.1).

Each of these results are proved by determining a lower bound for the gap sizes in each level and applying Theorem 2.

Metric Diophantine Approximation

3.1 Introduction and state of the art

In Chapter 1 we discussed different types of real numbers. This chapter concerns Diophantine approximation, i.e. the quality of rational approximations of real numbers. Metric Diophantine approximation concerns the size (measure or dimension, see Chapter 2) of sets of real numbers with certain approximation properties.

First a bit of notation. In this chapter, by $|\cdot|$ we will mean the sup-norm and by $\|\cdot\|$ we mean the distance to the nearest integer, i.e. for a real vector $x \in \mathbb{R}^m$, we have $\|x\| = \min_{p \in \mathbb{Z}^m} |x - p|$.

The first result on approximating real numbers by rationals we will mention is Dirichlet's approximation theorem (see [27]).

Theorem 7 (Dirichlet's Theorem). *Let α and $Q > 1$ be real numbers. Then there exists an integer p and a natural number q such that*

$$0 < q < Q \quad \text{and} \quad \left| \alpha - \frac{p}{q} \right| \leq \frac{1}{qQ}.$$

This tells us how well we can expect to approximate real numbers by rationals with restricted denominators. In general Dirichlet's Theorem cannot be improved. Consider for instance $Q > 1$ an integer and $\alpha = Q^{-1}$, then for all $0 < q < Q$ and for all integers p we have $\left| \alpha - \frac{p}{q} \right| = \left| \frac{1}{Q} - \frac{p}{q} \right| \geq \frac{1}{qQ}$.

Dirichlet's Theorem generalises to approximating real matrices with integer vectors:

Theorem 8 (Dirichlet's Theorem in $\text{Mat}_{m \times n}(\mathbb{R})$). *For two natural numbers n and m let $A \in \text{Mat}_{m \times n}(\mathbb{R})$ and $Q > 1$. Then there exists integer vectors $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n \setminus \{\mathbf{0}\}$ such that*

$$|\mathbf{q}| < Q \quad \text{and} \quad |A\mathbf{q} - \mathbf{p}| \leq \frac{1}{Q^{\frac{n}{m}}}.$$

For $m = 1$ the type of approximation above is called *dual approximation*, while it is called *simultaneous approximation* for $n = 1$. In the case of simultaneous approximation, we have the following corollary of Dirichlet's Theorem:

Corollary 9. *For each $\mathbf{a} \in \mathbb{R}^m$ there are infinitely many integer vectors $\mathbf{p} \in \mathbb{Z}^m$ and natural numbers q such that*

$$|q\mathbf{a} - \mathbf{p}| \leq \frac{1}{q^{\frac{1}{m}}}.$$

Replacing $\frac{1}{q^{\frac{1}{m}}}$ by $\frac{c}{q^{\frac{1}{m}}}$ for some $0 < c < 1$ we get the *badly approximable* real vectors. These will make an appearance in §3.2¹.

Instead, replacing the function $q \mapsto q^{-\frac{1}{m}}$ with some other positive function ψ of q (called the *approximation function*), which decays faster, we are no longer guaranteed that all real vectors can be approximated well by rationals in terms of ψ . We call the set of real vectors $\mathbf{a}$ in $[0, 1]^m$ which do have infinitely many rational approximations $(\mathbf{p}, q) \in \mathbb{Z}^m \times \mathbb{N}$ such that

$$|q\mathbf{a} - \mathbf{p}| \leq \psi(q),$$

the set of *simultaneously ψ -approximable numbers*, which we denote by $\mathcal{S}_m(\psi)$. By Corollary 9 we have $\mathcal{S}_m(q \mapsto q^{-\frac{1}{m}}) = [0, 1]^m$.

For a rapidly decreasing function ψ , $\mathcal{S}_m(\psi)$ not only fails to contain every real number, but is in fact a Lebesgue null set:

Theorem 10 (Khintchine's Theorem). *If $\sum_{q=1}^{\infty} \psi(q)^m$ converges, then $\mathcal{S}_m(\psi)$ is a Lebesgue null set. If we additionally assume, that ψ is monotonic, then if $\sum_{q=1}^{\infty} \psi(q)^m$ diverges, then $\mathcal{S}_m(\psi)$ has full Lebesgue measure.*

If ψ is monotonic, then either almost all or almost no points in $\mathbb{R}^m$ can be approximated well in terms of ψ . Notice that for the convergence case, we

¹going back to Chapter 1, the badly approximable numbers are exactly those whose partial quotients are bounded.

do not assume monotonicity, while we do for the divergence case. This is no coincidence, since monotonicity is indeed needed for the divergence case if $m = 1$ (see [36]). For $m \geq 2$ we may remove the monotonicity condition (see [12] and [41]).

If instead of considering points in $\mathbb{R}^m$, we consider points on some submanifold $\mathcal{M}$, we can ask, if a Khintchine-type theorem still holds.

If $\dim \mathcal{M} < m$, then the measure of $\mathcal{M}$ is zero, so the classical Khintchine's Theorem does not guarantee ψ -approximable points on $\mathcal{M}$.

On planar curves, we have Bernik [17] showing the convergence part for the parabola (x, x^2) , assuming monotonicity of the approximation function and later Huang [57] without monotonicity. There are very few results not assuming monotonicity even for the convergence case. The rest of the results we mention all assume monotonicity of ψ .

For non-degenerate C^3 planar curves Beresnevich, Dickinson and Velani [11] showed the divergence part while Vaughan and Velani [77] proved the convergence part for C^2 non-degenerate planar curves. See also [7, 8, 14].

In higher dimension see e.g. Bernik [15, 16] and Dodson-Rynne-Vickers [35, 34] for both convergence and divergence results on some quite specific manifolds. Beresnevich proved the divergence part for analytic manifolds [9]

The convergence part was proven somewhat recently by Beresnevich and Yang [13] for non-degenerate (see [64]) manifolds. This is quite general, meaning the convergence case has more or less been put to rest.

If instead of approximating by rationals, we translate by some real vector $\theta \in \mathbb{R}^m$ and estimate $|qa - p + \theta|$, this is called *inhomogeneous* approximation. Both the convergence and divergence part of Khintchine's theorem was shown by Beresnevich and Datta [10] last year. This also improved the divergence case for homogeneous approximation, since that was previously only known for analytic manifolds.

Going back to Dirichlet's Theorem in $\text{Mat}_{m \times n}(\mathbb{R})$, consider for an approximation function ψ , uniform approximation of sets of affine forms $Ax + b$ for $A \in \text{Mat}_{m \times n}(\mathbb{R})$ and $b \in \mathbb{R}^m$. We call a pair (A, b) ψ -Dirichlet improvable if the system

$$|Aq + b - p|^m < \psi(T) \quad \text{and} \quad |q|^n < T$$

has a solution $(p, q) \in \mathbb{Z}^m \times \mathbb{Z}^n$ for all sufficiently large T . Denote the set of ψ -Dirichlet improvable pairs in $\text{Mat}_{m \times n}([0, 1]) \times [0, 1]^m$ by $\hat{D}_{m,n}(\psi)$.

In [63] Kleinbock and Wadleigh showed that for non-increasing ψ , we have

$$\lambda^{mn+m}(\hat{D}_{m,n}(\psi)^c) = \begin{cases} 0 & \text{if } \sum_{j=1}^{\infty} \frac{1}{j^2 \psi(j)} < \infty, \\ \text{FULL} & \text{if } \sum_{j=1}^{\infty} \frac{1}{j^2 \psi(j)} = \infty, \end{cases}$$

where λ^{mn+m} is the $(mn + m)$ -dimensional Lebesgue measure.

In the divergence case this implies that for almost all fixed $A \in \text{Mat}_{m \times n}([0, 1])$, the set of $\mathbf{b}$ such that $(A, \mathbf{b}) \in \hat{D}_{m,n}(\psi)$ has full Lebesgue measure and therefore also full Hausdorff dimension.

Kim and Kim [62] showed a zero-full law for Hausdorff measures both of the set of pairs $(A, \mathbf{b})$ and for the set where $\mathbf{b}$ is fixed. Neckrasov [68] showed a Hausdorff dimension result, where the approximation is allowed to be weighted as well as a result on the size of the set of twisted inhomogeneous ψ -Dirichlet improvable vectors $\hat{D}_{m,n,A}(\psi)$ (meaning for a fixed matrix A , the set of $\mathbf{b}$ such that $(A, \mathbf{b}) \in \hat{D}_{m,n}(\psi)$).

In §3.2 we give an alternative proof to a result, which is also implied by Neckrasov's Theorem 1.5, namely that for certain ψ , the dimension of the set $\hat{D}_{m,n,A}(\psi)^c$, is full for *all* fixed $A \in \text{Mat}_{m \times n}([0, 1])$. The proof is original, but the result is not, which is why it does not appear in any preprints or published papers by the author.

Theorem 11 also works for fewer functions ψ than Neckrasov's result, so it is not as strong. However in Theorem 12 we allow intersection with certain fractals, which is not necessarily included in any previous results.

3.2 Twisted inhomogeneous Dirichlet non-improvability

All text in §3.2 is copied from the progress report by the author which was defended at the qualifying examination in August 2024. The text has not been edited, the only change is the numbers in the references and the formatting. Both have been set to match the remainder of the dissertation. This is done in accordance with GSNS rules.

Theorem 11. *Let $\psi(t) = \frac{1}{t\eta(t)}$, with η increasing to infinity, let $m, n \in \mathbb{N}$ and let $A \in \text{Mat}_{m \times n}(\mathbb{R})$, then the set $\hat{D}_{m,n,A}(\psi)^c$ has full Hausdorff dimension.*

We will show a more general result than the above, for which we introduce the following notions:

For (X, d) a metric space let (Ω, d) be a compact subspace of X which is the support of a finite atom-free measure μ . As in [23] we call μ *δ -Ahlfors regular* if there are $\delta, r_0 > 0$ and constants $0 < a \leq 1 \leq b$ such that for x in Ω and $r < r_0$

$$ar^\delta \leq \mu(\overline{B(x, r)}) \leq br^\delta.$$

Let $X = \mathbb{R}^m$ and let $\mathcal{L}$ be an $(m - 1)$ -dimensional hyperplane. For $\epsilon > 0$ put $\mathcal{L}^{(\epsilon)} := \{x \in \mathbb{R}^m : d(x, \mathcal{L}) < \epsilon\}$. Again as in [23] we call μ *absolutely*

α -decaying if there are $C, \alpha, r_0 > 0$ such that for any hyperplane $\mathcal{L}$, any $\epsilon > 0$, any $x \in \Omega$, and any $r < r_0$ we have

$$\mu(\overline{B(x, r)} \cap \mathcal{L}^{(\epsilon)}) \leq C \left(\frac{\epsilon}{r} \right)^\alpha \mu(\overline{B(x, r)}).$$

Theorem 12. *Let $K \subseteq [0, 1]^m$ be a compact set which supports an absolutely α -decaying, δ -Ahlfors regular measure μ such that $\delta > m - 1$. Then for $\psi(t) = \frac{1}{t\eta(t)}$, with η increasing to infinity and any real $m \times n$ matrix A , we have*

$$\dim(\hat{D}_{m,n,A}(\psi)^c \cap K) = \delta.$$

So within K , the set $\hat{D}_{m,n,A}(\psi)^c$ is of maximal dimension. Theorem 11 follows from Theorem 12 by setting $K = [0, 1]^m$ and μ the m -dimensional Lebesgue measure.

Preliminaries In order to prove our main result, we need a result by Bugeaud, Harrap, Kristensen and Velani from [23]. A sequence $(\mathbf{y}_i)_{i=1}^\infty$ in $\mathbb{Z}^m \setminus \{\mathbf{0}\}$ is called *lacunary* if there exists a constant $\lambda > 1$ such that

$$|\mathbf{y}_{i+1}| \geq \lambda |\mathbf{y}_i| \quad \text{for all } i \in \mathbb{N}.$$

Given a sequence $(\mathbf{y}_i)_{i=1}^\infty$ in $\mathbb{Z}^m$ define

$$\text{Bad}_{\{\mathbf{y}_i\}} := \{\mathbf{x} \in [0, 1]^m : \exists c(\mathbf{x}) > 0 \text{ s.t. } \|\mathbf{y}_i \cdot \mathbf{x}\| \geq c(\mathbf{x}) \quad \forall i \in \mathbb{N}\}$$

We have the following theorem from [23]

Theorem 13 ([23], Theorem 3). *Let $(\mathbf{y}_i)_{i=1}^\infty$ be a lacunary sequence in $\mathbb{Z}^m$. Furthermore, let $K \subseteq [0, 1]^m$ be a compact set which supports an absolutely α -decaying, δ -Ahlfors regular measure μ such that $\delta > m - 1$. Then*

$$\dim(\text{Bad}_{\{\mathbf{y}_i\}} \cap K) = \delta.$$

The proof of theorem 12 uses the above theorem 13 as well as the existence of a sequence of best approximations constructed in [22]. In [[22], §3] Bugeaud and Laurent show that for each matrix A , in $\text{Mat}_{m \times n}(\mathbb{R})$ with $A^t \mathbb{Z}^m + \mathbb{Z}^n$ of full rank, there is a sequence $(\mathbf{y}_i)_{i=1}^\infty$ in $\mathbb{Z}^m$ satisfying the following:

1. $1 = |\mathbf{y}_1| < |\mathbf{y}_2| < \dots$
2. $\|A^t \mathbf{y}_1\| > \|A^t \mathbf{y}_2\| > \|A^t \mathbf{y}_3\| > \dots$
3. $\mathbf{y} \in \mathbb{Z}^m \setminus \{\mathbf{0}\}$ with $|\mathbf{y}| < |\mathbf{y}_{i+1}| \implies \|A^t \mathbf{y}\| \geq \|A^t \mathbf{y}_i\|$.

called a *sequence of best approximations* to A . In the one-dimensional case, when A is an irrational number α , this sequence is exactly what we know to be the sequence of best approximations, i.e. the convergents of the continued fraction of α .

Lemma 14. *From ([22], §3) we have*

$$\|A^t \mathbf{y}_i\| \leq |\mathbf{y}_{i+1}|^{-\frac{m}{n}}. \quad (3.1)$$

Where $(\mathbf{y}_i)_{i=1}^\infty$, $\mathbf{y}_i \in \mathbb{Z}^m$ for every $i \in \mathbb{N}$ is the sequence of best approximations to the matrix A .

Furthermore from ([22], §5) we get a subsequence $(\mathbf{y}_{\varphi(i)})_{i=1}^\infty$ of $(\mathbf{y}_i)_{i=1}^\infty$ that satisfies

$$\frac{|\mathbf{y}_{\varphi(i)+1}|}{|\mathbf{y}_{\varphi(i+1)}|} \geq \frac{1}{9m}. \quad (3.2)$$

So the subsequence $(\mathbf{y}_{\varphi(i)})_{i=1}^\infty$ is lacunary by (3.2), which means we can use Theorem 13. Furthermore the sequence $(\mathbf{y}_{\varphi(i)})_{i=1}^\infty$ still satisfies the inequality (3.1), so we get

$$\|A^t \mathbf{y}_{\varphi(i)}\| \leq |\mathbf{y}_{\varphi(i)+1}|^{-\frac{m}{n}}. \quad (3.3)$$

Note the following norm inequality: For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^k$

$$\|\mathbf{x} \cdot \mathbf{y}\| < k\|\mathbf{x}\|\|\mathbf{y}\|. \quad (3.4)$$

Proof of Theorem 12 Let A in $\text{Mat}_{m \times n}(\mathbb{R})$ be such that the group $A^t \mathbb{Z}^m + \mathbb{Z}^n$ has rank $m + n$. Let $(\mathbf{y}_i)_{i=1}^\infty$ be a sequence of best approximations to A . From Lemma 14 there is a lacunary subsequence $(\mathbf{y}_{\varphi(i)})_{i=1}^\infty$ of the sequence of best approximations.

For $c > 0$ let

$$B_{\{\mathbf{y}_{\varphi(i)}\}}(c) := \{\mathbf{x} \in K : \|\mathbf{y}_{\varphi(i)} \cdot \mathbf{x}\| \geq c \text{ for all } i \in \mathbb{N}\}$$

From [23] we have

$$\text{Bad}_{\{\mathbf{y}_{\varphi(i)}\}} \cap K = \bigcup_{c>0} B_{\{\mathbf{y}_{\varphi(i)}\}}(c)$$

and

$$\dim B_{\{\mathbf{y}_{\varphi(i)}\}}(c) \xrightarrow{c \rightarrow 0} \dim(\text{Bad}_{\{\mathbf{y}_{\varphi(i)}\}} \cap K)$$

Suppose that we had a sequence $(c_j)_{j=1}^\infty$ with $c_j > 0$ and $c_j \rightarrow 0$ such that

$$B_{\{\mathbf{y}_{\varphi(i)}\}}(c_j) \subseteq \hat{D}_{m,n,A}(\psi)^c \cap K \text{ for all } j. \quad (3.5)$$

From Theorem 3 in [23] we get $\dim(\text{Bad}_{\{\mathbf{y}_{\varphi(i)}\}} \cap K) = \delta$, so

$$\dim(\hat{D}_{m,n,A}(\psi)^c \cap K) \geq \dim(B_{\{\mathbf{y}_{\varphi(i)}\}}(c_j) \xrightarrow{j \rightarrow \infty} \dim(\text{Bad}_{\{\mathbf{y}_{\varphi(i)}\}} \cap K) = \delta$$

Thus

$$\delta \leq \dim(\hat{D}_{m,n,A}(\psi)^c \cap K) \leq \dim(K) = \delta$$

where the last equality is [(6), [23]]. So it suffices to show the inclusion (3.5).

For $T > 0$ define the set

$$\hat{D}_{m,n,A,T}(\psi)^c := \{\mathbf{b} \in [0,1]^m : \forall \mathbf{q} \in \mathbb{Z}^n, |\mathbf{q}|^n < T \Rightarrow \|A\mathbf{q} + \mathbf{b}\|^m \geq \psi(T)\}.$$

Then for any sequence $(T_j)_{j=1}^\infty$ with $T_j > 0$ and $T_j \rightarrow \infty$ we have

$$\limsup_{j \rightarrow \infty} \hat{D}_{m,n,A,T_j}(\psi)^c \subseteq \hat{D}_{m,n,A}(\psi)^c.$$

For $(T_j)_{j=1}^\infty$ any sequence as above let

$$c_j := m \left(\frac{\log(\eta(T_j))}{\eta(T_j)} \right)^{1/m} + \frac{n \cdot (9m)^{m/n}}{\log(\eta(T_j))^{1/n}}.$$

Since the sequence $(|\mathbf{y}_{\varphi(i)}|)_{i=1}^\infty$ is increasing to infinity, we can find a sequence $(T_j)_{j=1}^\infty$ also increasing to infinity such that for each i

$$|\mathbf{y}_{\varphi(i)}| \leq (T_i \log(\eta(T_i)))^{1/m} < |\mathbf{y}_{\varphi(i+1)}|. \quad (3.6)$$

We will show the inclusion

$$B_{\{\mathbf{y}_{\varphi(i)}\}}(c_j) \subseteq \hat{D}_{m,n,A,T_j}(\psi)^c \cap K \quad \text{for all } j \in \mathbb{N}. \quad (3.7)$$

Fix a point $\mathbf{b} \in B_{\{\mathbf{y}_{\varphi(i)}\}}(c_j)$, then

$$\|\mathbf{y}_{\varphi(i)} \cdot \mathbf{b}\| \geq c_j \text{ for all } i \in \mathbb{N}. \quad (3.8)$$

Let $\mathbf{q} \in \mathbb{Z}^n$ with

$$|\mathbf{q}|^n < T_j \text{ i.e } |\mathbf{q}| < T_j^{\frac{1}{n}}. \quad (3.9)$$

We write

$$\mathbf{y}_{\varphi(i)} \cdot \mathbf{b} = \mathbf{q} \cdot A^t \mathbf{y}_{\varphi(i)} - \mathbf{y}_{\varphi(i)} \cdot (A\mathbf{q} - \mathbf{b})$$

Then from (3.8) we get that

$$0 < c_j \leq \|\mathbf{y}_{\varphi(i)} \cdot (-\mathbf{b})\| = \|\mathbf{q} \cdot A^t \mathbf{y}_{\varphi(i)} - \mathbf{y}_{\varphi(i)} \cdot (A\mathbf{q} + \mathbf{b})\|$$

Using the triangle inequality and the norm inequality (3.4) we get

$$\|A\mathbf{q} + \mathbf{b}\| > \frac{c_j - n|\mathbf{q}|\|A^t \mathbf{y}_{\varphi(i)}\|}{m|\mathbf{y}_{\varphi(i)}|}$$

For the numerator we get

$$\begin{aligned} c_j - n|\mathbf{q}|\|A^t \mathbf{y}_{\varphi(i)}\| &\stackrel{(3.9)}{>} c_j - sT_j^{\frac{1}{s}}\|A^t \mathbf{y}_{\varphi(i)}\| \\ &\stackrel{(3.3)}{\geq} c_j - nT_j^{\frac{1}{n}}\|\mathbf{y}_{\varphi(i)+1}\|^{-\frac{m}{n}} \\ &= c_j - nT_j^{\frac{1}{n}}|\mathbf{y}_{\varphi(i+1)}|^{-\frac{m}{n}} \left(\frac{|\mathbf{y}_{\varphi(i)+1}|}{|\mathbf{y}_{\varphi(i+1)}|} \right)^{-\frac{m}{n}} \\ &\stackrel{(3.2)}{\geq} c_j - nT_j^{\frac{1}{n}}|\mathbf{y}_{\varphi(i+1)}|^{-\frac{m}{n}} \left(\frac{1}{9m} \right)^{-\frac{m}{n}} \end{aligned}$$

using Lemma 14 and the choice of $\mathbf{q}$. Now employing our inequality (3.6) defining T_j , we get

$$\begin{aligned} &> c_j - nT_j^{\frac{1}{n}} \left((T_j \log(\eta(T_j)))^{1/m} \right)^{-\frac{m}{n}} \left(\frac{1}{9m} \right)^{-\frac{m}{n}} \\ &= c_j - nT_j^{\frac{1}{n}} (T_j \log(\eta(T_j)))^{-1/n} \left(\frac{1}{9m} \right)^{-\frac{m}{n}}. \end{aligned}$$

For the denominator we have from (3.6)

$$|\mathbf{y}_{\varphi(i)}| \leq (T_j \log(\eta(T_j)))^{1/m}$$

So in total, inserting the above and the definition of c_j , we get

$$\begin{aligned} \|A\mathbf{q} + \mathbf{b}\| &\geq \frac{c_j - nT_j^{\frac{1}{n}} (T_j \log(\eta(T_j)))^{-1/n} \left(\frac{1}{9m} \right)^{-\frac{m}{n}}}{m (T_j \log(\eta(T_j)))^{1/m}} \\ &= \frac{m \left(\frac{\log(\eta(T_j))}{\eta(T_j)} \right)^{1/m} + \frac{n \cdot (9m)^{m/n}}{\log(\eta(T_j))^{1/n}} - nT_j^{\frac{1}{n}} (T_j \log(\eta(T_j)))^{-1/n} \left(\frac{1}{9m} \right)^{-\frac{m}{n}}}{m (T_j \log(\eta(T_j)))^{1/m}} \\ &= \frac{m \log(\eta(T_j))^{1/m} \eta(T_j)^{-1/m}}{m T_j^{1/m} \log(\eta(T_j))^{1/m}} \\ &= \left(\frac{1}{T_j \eta(T_j)} \right)^{1/m} =: \psi(T_j)^{1/m} \end{aligned}$$

So $\mathbf{b}$ is in $\hat{D}_{m,n,A,T_j}(\psi)^c \cap K$ and thus the inclusion (3.7) is shown, which completes the proof of Theorem 12 in the case where the group $A^t\mathbb{Z}^m + \mathbb{Z}^n$ has full rank.

If we on the other hand assume that $A^t\mathbb{Z}^m + \mathbb{Z}^n$ has rank strictly smaller than $m + n$, then by [[23], §5] we have that $\{A\mathbf{q} : \mathbf{q} \in \mathbb{Z}^n\}$ is restricted to at most a countable family of positively separated, parallel hyperplanes in $\mathbb{R}^m$. Denoting these by F we get

$$K \setminus F = \hat{D}_{m,n,A}(\psi)^c \cap K$$

since $\hat{D}_{m,n,A}(\psi)^c \cap K$ is exactly the $\mathbf{b} \in K$ that have some positive distance to $\{A\mathbf{q} : \mathbf{q} \in \mathbb{Z}^n\}$ (since $\psi(T)$ gets arbitrarily close to 0 for T large enough).

Again by [[23], §5] we get, that $\dim(K \setminus F) = \dim(K)$ so $\dim(\hat{D}_{m,n,A}(\psi)^c \cap K) = \dim(K) = \delta$, which completes the proof.

3.3 Quantitative Khintchine

As pointed out in [1], whenever you have a Khintchine convergence result, like e.g. Beresnevich and Yang's result:

Theorem (Beresnevich-Yang 2023 [13]). *Let $m \geq 2$ and $\mathcal{M} \subseteq \mathbb{R}^m$ be a non-degenerate submanifold. Assume ψ is monotonic and that the series $\sum_{q=1}^{\infty} \psi(q)^m$ converges. Then almost no points on $\mathcal{M}$ are simultaneously ψ -well approximable.*

An immediate consequence is, that if you have ψ as above, then for almost every point $\mathbf{x} \in \mathcal{M}$, there exists a $\kappa(\mathbf{x}) > 0$ such that

$$\|q\mathbf{x}\| > \kappa(\mathbf{x})\psi(q) \quad \forall q \in \mathbb{N}.$$

This type of estimate is very useful, but applications usually ask that $\kappa(\mathbf{x})$ be independent of $\mathbf{x}$. Unfortunately this is not possible to guarantee with probability 1, leading Adiceam, Beresnevich, Levesley, Velani and Zorin [1] to state the following problem:

Problem 15 (Problem in [1]). *Investigate the dependency between $\kappa > 0$ and the probability of the set of vectors $\mathbf{x} \in \mathcal{M}$ where $\kappa(\mathbf{x}) = \kappa$.*

Answering this question is the subject of [32], where Datta proves a quantitative Khintchine's theorem. Assume that $\mathcal{M}$ is a compact l -non-degenerate (at every point) submanifold of $\mathbb{R}^n$ (for $n \geq 2$) of the form

$$\{f(\mathbf{x}) \mid \mathbf{x} \in V = U^{n+d+3}\}$$

for some C^{l+1} function f (we will not specify the precise details, but rather aim to give a flavour of the result). Assume that ψ is monotonic and that $\sum_{q=1}^{\infty} \psi(q)^n < \infty$. Then given any $0 < \delta < 1$, she gives an explicit expression for $\kappa = \kappa(\delta) > 0$ such that

$$\lambda(\mathcal{B}(\kappa, \psi)) \geq (1 - \delta)\lambda(U),$$

where

$$\mathcal{B}(\kappa, \psi) := \{x \in U : \|qf(x)\| > \kappa\psi(q) \quad \forall q \in \mathbb{N}\}.$$

The assumption that ψ is monotonic is integral to the proof, but since monotonicity is not needed for the convergence part of Khintchine's theorem on $\mathbb{R}^m$ one might be tempted to conjecture, that it is not needed on manifolds either. In general removing the monotonicity assumption is a very hard problem and few results are known.

As mentioned in §3.1 Huang [57] has such a result, namely the convergence part of Khintchine's theorem on the standard parabola without monotonicity of the approximation function ψ :

Theorem 16 (Theorem 1 in [57]). *Let ψ be any approximation function such that $\sum_{q=1}^{\infty} \psi(q)^2$ converges. Then almost no points on the parabola are simultaneously ψ -well approximable.*

(The measure implicit in the statement of the theorem is $\mathcal{H}^1$.)

Same as in [1] and [32], one can state a problem analogous to Problem 15. Answering this problem is exactly the subject of the preprint *Quantitative Khintchine on the parabola with non-monotonic approximation functions* by the author and Simon Kristensen [45] attached in Chapter 6.

For any function $\psi : \mathbb{N} \rightarrow (0, \infty)$ and any $\kappa > 0$ define

$$\mathcal{B}'(\kappa, \psi) := \left\{ x \in [0, 1] : \max(\|qx\|, \|qx^2\|) > \kappa\psi(q) \text{ for all } q \in \mathbb{N} \right\}.$$

[45] consists of four main quantitative Khintchine results on the parabola (analogous to Datta's theorem), each assuming different conditions on the function ψ (however we will never assume monotonicity, since that is covered by Datta's result) and thus resulting in different values of $\kappa(\delta)$. We state all four below, and afterwards compare them, sketch the proofs and the lemmata needed for said proofs.

Restricting to the parabola is of course much less general than the setup for Datta's result on non-degenerate manifolds, however the methods used in the proof rely on knowing more specifically how the space looks in order to count rational points. In general no one has managed to remove monotonicity assumptions on very large classes of manifolds.

Theorem 17 (G.-Kristensen, Theorem 4 in [45]). *All q . Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^{\infty} \psi(q)^2$ converges. Given any $0 < \delta < 1$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{2\sqrt{6}(5\log(2))^4}{3\pi^3 \sqrt{\zeta\left(\frac{12}{11}\right)} 11^4}, \sqrt{\frac{\delta}{S_\psi}} \frac{1}{\sqrt{18\pi}}, \right. \\ &\quad \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(\frac{1}{2\pi^2 \sqrt{12}} \right)^{\frac{2}{3}} \left(99684 \cdot 4.19253564643679 \right. \\ &\quad \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(e^{e^8} - 10^5 \right) e^{\frac{5\log(2)e^8}{2(8-1.39177)}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} \right. \\ &\quad \left. + 5 \cdot 4^{\frac{3}{16} + \frac{0.00218819 + 0.00218819^2}{2} + \frac{5}{2.57}} \left(\frac{57}{\log(2)} \right)^{\frac{7.176846309399143769411680}{2}} \right. \\ &\quad \left. \cdot \left(\frac{1}{e \cdot 0.00218819} \right)^{1 + \frac{0.00218819}{2}} \right)^{\frac{-2}{3}} \left. \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.00015248265228, \right. \\ &\quad \left. \sqrt{\frac{\delta}{S_\psi}} 0.07502635967975, \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 10^{-10^{23.89775276641198}} \right\}. \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

The first result is the most general. Here we assume nothing except convergence of the series $\sum_{q=1}^{\infty} \psi(q)^2$. The tradeoff is, that the constant

$$\simeq 10^{-10^{23.89775276641198}},$$

which appears in the expression for $\kappa(\delta)$, is extremely small. This unfortunately cannot be helped if one wants to maintain the structure of the proof. This stems from an explicit bound on the divisor function $d(q)$, which has to show up in the proof unless we want to weaken the statement significantly by avoiding denominators q with many divisors (see further discussion of this below).

The three other results represent different strategies to remedy the tiny constants in the above while still maintaining the original proof idea.

Since asymptotically for each $\varepsilon > 0$ we have $d(q) \ll_\varepsilon q^\varepsilon$, in principle we need only to remove small values of q to get as good a bound on $d(q)$ as we please. This is the approach in the next result:

Theorem 18 (G.-Kristensen, 2026+, Theorem 5 in [45]. Large q). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^{\infty} \psi(q)^2$ converges. Assume additionally, that $\psi(q) = 0$ for all $q \leq \exp(\exp(41))$. Given any $\delta > 0$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{2}{3\pi^2 \sqrt{\frac{\pi^2}{6} \zeta\left(\frac{112}{57}\right)}}, \sqrt{\frac{\delta}{S_\psi}} \sqrt{\frac{1}{18\pi^2}}, \right. \\ &\quad \left. \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left(4.19253564643679 + 27 \cdot 4^{\frac{3}{16} + \frac{5}{2 \cdot 57} + \frac{41 + \left(\frac{41}{e^{41}}\right)^2}{2}} \right) \right)^{\frac{-2}{3}} \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.04064400, \sqrt{\frac{\delta}{S_\psi}} 0.07502635, \right. \\ &\quad \left. \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 0.00499238 \right\}. \end{aligned}$$

Then

$$\lambda(B'(\psi, \kappa)) \geq 1 - \delta.$$

We see that the constants that appear in $\kappa(\delta)$ are much more reasonably sized this time around, but the price is, that we need to assume, that $q > \exp(\exp(41))$ which is astronomical!

Rather than finding a good bound for $d(q)$, one could just assume, that q has at most e.g. 2^n divisors (for some fixed n). By assuming q has at most n prime divisors, we get $d(q) \leq 2^n$ for free (this time at the cost, that we are approximating by rationals, where the denominators only have at most n prime divisors).

Theorem 19 (G.-Kristensen, 2026+, Theorem 6 in [45]. $q = p_1 \dots p_k$). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^{\infty} \psi(q)^2$ converges. Let $k \in \mathbb{N}$ and assume additionally, that $\psi(q) = 0$ for all q with more than k prime*

divisors. Given any $0 < \delta < 1$, let

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{4}{3\pi^4} \frac{1}{2^n}, \right. \\ &\quad \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left[4.19253564643679 \cdot 99684 \right. \right. \\ &\quad \left. \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} 2^{\frac{7k}{2}} \right] \right)^{\frac{-2}{3}}, \\ &\quad \left. \sqrt{\frac{\delta}{S_\psi}} \sqrt{\frac{1}{18\pi^2}} \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.0410639 \frac{1}{2^n}, \right. \\ &\quad \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 0.0598025 \left(417928.72 + 2082.9292 \cdot 2^{\frac{7k}{2}} \right)^{\frac{-2}{3}}, \\ &\quad \left. \sqrt{\frac{\delta}{S_\psi}} 0.0750263 \right\}. \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

Taking this to the extreme suppose, that $q = p$ itself is a prime. This is of course a special case of Theorem 19 above, but taking $q = p$ to be a prime can be used to our advantage elsewhere in the proof, so the case $q = p$ is stated separately:

Theorem 20 (Theorem 7 in [45]. $q = p$ prime). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^{\infty} \psi(q)^2$ converges. Assume additionally, that $\psi(q) =$*

0 whenever q is not a prime. Given any $0 < \delta < 1$, let

$$\begin{aligned} \kappa(\delta) &< \min \left(\frac{1}{6 \max_p \psi(p)}, \frac{2\delta}{3\pi^2 S_\psi}, \left(\frac{\delta(e \cdot 0.249)^{\frac{1}{4}}}{2\pi^2 \sqrt{12} \cdot 10.0366 S_\psi} \right)^{\frac{2}{3}}, \sqrt{\frac{\delta}{18\pi^2 S_\psi}} \right) \\ &\simeq \min \left(\frac{1}{6 \max_p \psi(p)}, \frac{\delta}{S_\psi} 0.0675474558, \frac{\delta^{\frac{2}{3}}}{S_\psi^{\frac{2}{3}}} 0.0120432671, \right. \\ &\quad \left. \sqrt{\frac{\delta}{S_\psi}} 0.0750263597 \right). \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

The expressions for κ in Theorem 19 for $n = 1$ and in Theorem 20 are not easily comparable, since they contain different powers of S_ψ and δ , so determining which one is larger depends on ψ and δ .

Proof idea All four proofs follow the same structure. Rather than bounding the set $\mathcal{B}'(\kappa, \psi)$ from below, we bound its complement from above. The set

$$\mathcal{B}'(\psi, \kappa)^c = \left\{ x \in [0, 1] : \exists q \in \mathbb{N} \text{ s.t. } \|qx\|, \|qx^2\| \leq \kappa\psi(q) \right\}$$

is contained in

$$\bigcup_{q=1}^{\infty} \bigcup_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} \left[\frac{a}{q} - \frac{\kappa\psi(q)}{q}, \frac{a}{q} + \frac{\kappa\psi(q)}{q} \right]$$

so finding an upper bound for $\lambda(\mathcal{B}'(\kappa, \psi))$ reduces to bounding the counting function

$$\sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} 1,$$

which roughly counts the number of rational points with fixed denominator close to the parabola.

The counting function is estimated with exponential sums, which are in turn estimated by sums of Dirichlet characters. To make the result completely explicit, great care is taken throughout to preserve all constants and lower order terms (however not all constants are optimal). In order to bound the characters sums the obvious choice would be Burgess' bound for short character sums [24, 25], however this is only asymptotic.

Burgess's bound for short character sums. *Let χ be a non-principal character modulo q . Then for any $\varepsilon > 0$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \ll_{\varepsilon} \sqrt{N} q^{\frac{11}{16} + \varepsilon},$$

where the implied constant only depends on ε .

There are several explicit versions of Burgess' bound. See for instance [18, 58, 67, 76] for $q = p$ primes and [19, 59] for composite moduli. For primes we have quite good results, but for composite moduli, we only have bounds for q of approximate size $q \geq e^{e^8}$. In order to prove Theorem 17 we need it for *all* q , and since e^{e^8} is too large a number to calculate the constant for $1 \leq q \leq e^{e^8}$ by hand, something else is needed.

Explicit Burgess bound In [19] Bordignon uses his Theorem 3.4 to prove his main result. This will be our lifeline as well. Using different values of the variables appearing in Theorem 3.4 we can prove an explicit Burgess bound specifically for $10^5 \leq q \leq e^{e^8}$.

For the remaining $1 \leq q \leq 10^5$, we use Python to calculate an explicit constant.

What results are the following three versions of explicit Burgess bounds for different sizes of q :

Lemma 21 (Theorem 1.3 in [19]. Large q). *Let $q > \exp(\exp(8))$. Let χ be a primitive Dirichlet character mod q . For any integers M and $N < q$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 5d(q)^{\frac{3}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))}$$

Lemma 22 (G.-Kristensen, 2026+, Lemma 9 in [45]. Small q). *For $1 \leq q \leq 10^5$ and $q^{\frac{3}{8}} \leq N \leq q^{\frac{5}{8}}$ we have*

$$\left| \sum_{j=1}^N \chi(j) \right| \leq 4.19253564643679 \sqrt{N} q^{\frac{3}{16}},$$

where $\chi(j) = \left(\frac{j}{q}\right)$ or $\chi(j) = \left(\frac{q}{j}\right)$.

(See Appendix in [45]).

Lemma 23 (G.-Kristensen, 2026+, Lemma 10 in [45]. Medium q). *Let $10^5 \leq q \leq \exp(\exp(8))$ and let χ be a primitive Dirichlet character mod q . For any integers M and $N < q$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 27d(q)^{\frac{3}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))}$$

Lemma 21 and Lemma 23 still have the problem that the character is assumed to be primitive, which is not always the case for us. To fix this issue we use a remark by Burgess [25]. In using this remark another $d(q)$ is introduced into the expression. It is mainly due to this argument (which we do not see a way to avoid) that there is no good way to avoid getting either extremely small constants in the expression for $\kappa(\delta)$ or restricting to q astronomically large in order to get bound for general composite numbers. The problem lies with the available bounds for the divisor function $d(q)$, which we need to use in order for this approach to work.

For Theorem 20 McGown's result suffices, since it only assumes $p \geq 2 \cdot 10^4$, so for small values we can calculate separately.

Lemma 24 (Theorem 7.1 in [67]. $q = p$). *Suppose χ is a non-principal Dirichlet character modulo a prime $p \geq 2 \cdot 10^4$. For $M \in \mathbb{N}$ and $N \leq p^{\frac{5}{8}}$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 10.0366 \log(p)^{\frac{1}{4}} p^{\frac{3}{16}} \sqrt{N} = 10.0366 \sqrt[4]{C_\varepsilon} \sqrt{N} p^{\frac{3}{16} + \frac{\varepsilon}{4}},$$

with $C_\varepsilon = \log(p) / p^\varepsilon$.

For $1 \leq q \leq 20000$ use the constant 3.41105405758246 (see Appendix in [45]).

Divisor function bounds Asymptotically we have for any $\varepsilon > 0$, that

$$d(q) \ll_\varepsilon q^\varepsilon,$$

but we need something explicit.

For all q use the following bound:

Lemma 25 (Divisor function bound (from Terence Tao's blog [74])). *For any $\beta > 0$, there is an explicit constant $K_\beta > 0$ such that*

$$d(q) \leq K_\beta q^\beta,$$

Specifically we have $K_{\frac{5}{11}} \leq \left(\frac{11}{5 \log(2)}\right)^4$ and $K_{\frac{1}{57}} \leq \left(\frac{57}{\log(2)}\right)^{176846309399143769411680}$.

The constant K_β gets very large if we need β to be small. For q very large, we can do better. For instance we have:

Lemma 26 (From Robin's thesis [71], also mentioned in the survey paper [69]). *For $q \geq 56$*

$$\log(d(q)) \leq \frac{\log(q) \log(2)}{\log(\log(q)) - 1.39177}$$

In particular, for $q \geq \exp(\exp(41))$, we get $d(q) \leq q^{\frac{1}{57}}$.

Bibliography

- [1] F. Adiceam et al., *Diophantine approximation and applications in interference alignment*, Advances in mathematics (New York. 1965) 302 (2016), pp. 231–279.
- [2] A. V. Aho and N. Sloane, *Some Doubly Exponential Sequences*, The Fibonacci Quarterly 11.4 (1973), pp. 429–437.
- [3] S. B. Andersen and S. Kristensen, *Arithmetic properties of series of reciprocals of algebraic integers*, Monatsh. Math. 190.4 (2019), pp. 641–656.
- [4] S. Astels, *Cantor sets and numbers with restricted partial quotients*, Trans. Am. Math. Soc. 352.1 (2000), pp. 133–170.
- [5] ———, *Sums of numbers with small partial quotients. II*, J. Number Theory 91.2 (2001), pp. 187–205.
- [6] S. Astels, *Cantor sets and numbers with restricted partial quotients*, Thesis, 1999.
- [7] D. Badziahin and J. Levesley, *A note on simultaneous and multiplicative Diophantine approximation on planar curves*, Glasg. Math. J. 49.2 (2007), pp. 367–375.
- [8] V. V. Beresnevich, R. C. Vaughan, and S. L. Velani, *Inhomogeneous Diophantine approximation on planar curves*, Math. Ann. 349.4 (2011), pp. 929–942.
- [9] V. Beresnevich, *Rational points near manifolds and metric Diophantine approximation*, Ann. of Math. (2) 175(1) (2012), pp. 187–235.
- [10] V. Beresnevich and S. Datta, *Rational points near manifolds and Khintchine theorem*, Preprint, arXiv:2505.01227 [math.NT] (2025), 2025.
- [11] V. Beresnevich, D. Dickinson, and S. Velani, *Diophantine approximation on planar curves and the distribution of rational points. With an appendix: Sums of two squares near perfect squares by R. C. Vaughan*, Ann. Math. (2) 166.2 (2007), pp. 367–426.
- [12] V. Beresnevich and S. Velani, *Classical metric Diophantine approximation revisited: the Khintchine-Groshev theorem*, Int. Math. Res. Not. 2010.1 (2010), pp. 69–86.
- [13] V. Beresnevich and L. Yang, *Khintchine’s theorem and Diophantine approximation on manifolds*, Acta Math. 231.1 (2023), pp. 1–30.
- [14] V. V. Beresnevich and S. L. Velani, *A note on simultaneous Diophantine approximation on planar curves*, Math. Ann. 337.4 (2007), pp. 769–796.
- [15] V. I. Bernik, *Asymptotisches Verhalten der Zahl der Lösungen einiger Systeme von Ungleichungen in der Theorie der diophantischen Approximationen abhängiger Größen*, Izv. Akad. Nauk BSSR, Ser. Fiz.-Mat. Nauk 1973.1 (1973), pp. 10–17.
- [16] ———, *Induced extremal surfaces*, Mat. Sb., Nov. Ser. 103 (1977), pp. 480–489.

- [17] V. I. Bernik, *On the exact order of approximation of almost all points of a parabola*, Mat. Zametki 26 (1979), pp. 657–665.
- [18] A. R. Booker, *Quadratic class numbers and character sums*, Math. Comp. (75 2006), pp. 1481–1492.
- [19] M. Bordignon, *Partial Gaussian sums and the Pólya–Vinogradov inequality for primitive characters*, Rev. Mat. Iberoam 38.4 (2022), pp. 1101–1127.
- [20] B. Brouwer, *Cantor sets for regular continued fractions and Lüroth series*, Bachelor thesis, Mentor: C. Kraaikamp, 2024, URL: <https://repository.tudelft.nl/record/uuid:c59b330d-04ad-4aa8-887f-bb8926959d83>.
- [21] Y. Bugeaud, *Approximation by algebraic numbers*, Cambridge tracts in mathematics ; 160, Cambridge: Cambridge University Press, 2004.
- [22] Y. Bugeaud and M. Laurent, *On exponents of homogeneous and inhomogeneous Diophantine approximation*, Moscow mathematical journal 5.4 (2005), pp. 747–766.
- [23] Y. Bugeaud et al., *On shrinking targets for $\mathbb{Z}^m$ actions on tori*, Mathematika 56 (2 2010), pp. 193–202.
- [24] D. A. Burgess, *On Character Sums and L-Series*, Proceedings of the London Mathematical Society s3-12 (1962), pp. 193–206.
- [25] ———, *On Character Sums and L-Series. II*, Proceedings of the London Mathematical Society s3-13.1 (Jan. 1963), pp. 524–536.
- [26] Cantor, *Über eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen*, Journal für die reine und angewandte Mathematik 1874.77 (1874), pp. 258–262.
- [27] J. W. S. Cassels, *An Introduction to Diophantine Approximation*, Cambridge University Press, 1957.
- [28] T. W. Cusick, *Sums and products of continued fractions*, Proc. Am. Math. Soc. 27 (1971), pp. 35–38.
- [29] T. W. Cusick and R. A. Lee, *Sums of sets of continued fractions*, Proc. Am. Math. Soc. 30 (1971), pp. 241–246.
- [30] K. Dajani and C. Kraaikamp, *Ergodic theory of numbers*. The Carus mathematical monographs ; no. 29. Washington, DC: «The» Mathematical Association of America, 2002.
- [31] T. Das et al., *Hausdorff dimensions of perturbations of a conformal iterated function system via thermodynamic formalism*, Sel. Math., New Ser. 29.2 (2023), Id/No 19, p. 56.
- [32] S. Datta, *Quantitative Khintchine in simultaneous approximation*, Discrete and Continuous Dynamical Systems 44 (Jan. 2023).
- [33] B. Divis, *On the sums of continued fractions*, Acta Arith. 22 (1973), pp. 157–173.
- [34] M. M. Dodson, B. P. Rynne, and J. A. G. Vickers, *Simultaneous diophantine approximation and asymptotic formulae on manifolds*, J. Number Theory 58.2 (1996), pp. 298–316.
- [35] ———, *Khintchine-type theorems on manifolds*, Acta arithmetica 57.2 (1991), pp. 115–130.
- [36] R. J. Duffin and A. C. Schaeffer, *Khintchine’s problem in metric Diophantine approximation*, Duke Mathematical Journal 8.2 (1941), pp. 243–255.
- [37] W. Dunham, *The calculus gallery: masterpieces from newton to Lebesgue*, Princeton University Press, 2015.

- [38] P. Erdős, *Some problems and results on the irrationality of the sum of infinite series*, Journal of Mathematical Sciences 10 (1975), pp. 1–7.
- [39] P. Erdős and U. Dudley, *Some remarks and problems in number theory related to the work of Euler*, Math. Mag. 56 (1983), pp. 292–298.
- [40] K. Falconer, *Fractal Geometry : Mathematical Foundations and Applications*. 3rd ed., New York: John Wiley & Sons, Incorporated, 2014.
- [41] P. X. Gallagher, *Metric simultaneous diophantine approximation. II*, Mathematika 12 (1965), pp. 123–127.
- [42] D. Gayfulin and E. Nesharim, *Every real number is a sum of two real numbers with diverging partial quotients*, Preprint, arXiv:2507.04521 [math.NT] (2025), 2025.
- [43] I. J. Good, *The fractional dimensional theory of continued fractions.*, Proc. Camb. Philos. Soc. 37 (1941), pp. 199–228.
- [44] M. Gravgaard, J. Hančl, and S. Kristensen, *Lower bound for the Hausdorff dimension of expressible sets*, preprint, 2026, arXiv: 2605.26839 [math.NT], URL: <https://arxiv.org/abs/2605.26839>.
- [45] M. Gravgaard and S. Kristensen, *Quantitative Khintchine on the parabola with non-monotonic approximation functions*, preprint, 2026.
- [46] M. Gravgaard and Y. W. Lee, *Decomposition of real numbers into sums of Lüroth sets*, International Journal of Number Theory (2026), pp. 1–38.
- [47] M. j. Hall, *On the sum and product of continued fractions*, Ann. Math. (2) 48 (1947), pp. 966–993.
- [48] J. Hančl, *A criterion for linear independence of series*, The Rocky Mountain journal of mathematics 34.1 (2004), pp. 173–186.
- [49] ———, *Transcendental sequences*, Mathematica Slovaca 46.2-3 (1996), pp. 177–179.
- [50] J. Hančl and R. Nair, *On the irrationality of infinite series of reciprocals of square roots*, Rocky Mountain Journal of Mathematics 47.5 (2017), pp. 1525–1538.
- [51] J. Hančl, R. Nair, and J. Šustek, *On the Lebesgue measure of the expressible set of certain sequences*, Indag. Math., New Ser. 17.4 (2006), pp. 567–581.
- [52] J. Hančl and J. Šustek, *Boundedly expressible sets.*, Czech. Math. J. 59.3 (2009), pp. 649–654.
- [53] ———, *Expressible sets of sequences with Hausdorff dimension zero*, Monatsh. Math. 152.4 (2007), pp. 315–319.
- [54] ———, *Sequences of Cantor type and their expressibility*, Math. Slovaca 67.1 (2017), pp. 41–50.
- [55] J. Hančl et al., *On the Hausdorff dimension of the expressible set of certain sequences*, Acta Arith. 155.1 (2012), pp. 85–90.
- [56] J. L. Hlavka, *Results on sums of continued fractions*, Trans. Am. Math. Soc. 211 (1975), pp. 123–134.
- [57] J.-J. Huang, *Diophantine approximation on the parabola with non-monotonic approximation functions*, Mathematical Proceedings of the Cambridge Philosophical Society 168.3 (2020), pp. 535–542.

- [58] H. Iwaniec, E. Kowalski, and A. M. Society, *Analytic Number Theory*, vol. 53, American Mathematical Society Colloquium Publications, American Mathematical Society, 2004.
- [59] N. Jain-Sharma, T. Khale, and M. Liu, *Explicit Burgess bound for composite moduli*, International Journal of Number Theory 17.10 (2021), pp. 2207–2219.
- [60] O. N. Karpenkov, *Geometry of continued fractions*, Second edition, vol. 26, Algorithms and Computation in Mathematics, Springer, 2022.
- [61] A. Y. Khintchine, *Continued Fractions*, Dover Books on Mathematics, Dover Publications, 1997.
- [62] T. Kim and W. Kim, *Hausdorff measure of sets of Dirichlet non-improvable affine forms*, Advances in Mathematics 403 (2022), p. 108353.
- [63] D. Kleinbock and N. Wadleigh, *An inhomogeneous Dirichlet theorem via shrinking targets*, Compositio Mathematica 155.7 (2019), pp. 1402–1423.
- [64] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and diophantine approximation on manifolds*, Annals of mathematics (1998), pp. 339–360.
- [65] M. Løkkegaard Laursen, *Algebraic degree of series of reciprocal algebraic integers*, Rocky Mt. J. Math. 53.2 (2023), pp. 517–529.
- [66] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and rectifiability*, Cambridge studies in advances mathematics, 1995.
- [67] K. J. McGown, *Norm-Euclidean cyclic fields of prime degree*, International Journal of Number Theory 08.01 (2012), pp. 227–254.
- [68] V. Neckrasov, *Khintchine-type theorems for weighted uniform inhomogeneous approximations via transference principle*, Mathematika 72.2 (2026), p. 36.
- [69] J.-L. Nicolas, *On highly composite numbers*, Ramanujan Revisited: Proceedings of the Conference, Urbana-Champaign, Illinois, 1987, 1988, pp. 215–244.
- [70] Y. Peres and P. Shmerkin, *Resonance between Cantor sets*, Ergodic Theory Dyn. Syst. 29.1 (2009), pp. 201–221.
- [71] G. Robin, *Thèse d'état: Grandes valeurs de fonctions arithmétiques et problèmes d'optimisation en nombres entiers*, Thesis, Université de Limoges, 1983.
- [72] N. Shulga, *On a conjecture of Cusick on a sum of Cantor sets*, Trans. Am. Math. Soc. 379.6 (2026), pp. 4169–4190.
- [73] H. Steinhaus, *Nowa własność mnogości G. Cantora. (Eine neue Eigenschaft der Cantorschen Menge.)*, Polish, Wektor 7. 1917.
- [74] T. Tao, *The divisor bound*, <https://terrytao.wordpress.com/2008/09/23/the-divisor-bound/>, Accessed: 2026-04-28, 2008.
- [75] S. Thorbjørnsen, *Grundlæggende mål- og integralteori*, dan, 1st ed., Aarhus: Aarhus Universitetsforlag, 2014.
- [76] E. Treviño, *The Burgess inequality and the least k th power non-residue*, International Journal of Number Theory 11.5 (2015), pp. 1653–1678.
- [77] R. Vaughan and S. Velani, *Diophantine approximation on planar curves: the convergence theory*, Invent. Math. 166 (1 2006), pp. 103–124.

CHAPTER 4

Paper 1: Decomposition of real numbers into sums of Lüroth sets

This chapter contains a preprint of the paper *Decomposition of real numbers into sums of Lüroth sets* by the author and Ying Wai Lee. The paper has been accepted and is online ready on the website of the journal International Journal of Number Theory, but is at the time of writing yet to be published.

The preprint is also available on arXiv at <https://arxiv.org/abs/2506.12513v2>. The pages are numbered 1 through 36. The references within are selfcontained and do not match the numbering in the remainder of the thesis.

Preprint of an article accepted and online ready in [International Journal of Number Theory, 2026] [10.1142/S1793042126500958] © [copyright World Scientific Publishing Company] [<https://doi.org/10.1142/S1793042126500958>]

International Journal of Number Theory
 © World Scientific Publishing Company

Decomposition of real numbers into sums of Lüroth sets

Maiken Gravgaard*

*Department of Mathematics, Aarhus University, Ny Munkegade 118
 Aarhus, 8000, Denmark
 maiken@math.dk*

Ying Wai Lee

*Department of Mathematics, University of York
 York, YO10 5DD, United Kingdom
 yingwai.lee@york.au.dk*

Received (14/06/2025)

Revised (18/11/2025)

Accepted (01/02/2026)

We study the decomposition of real numbers into sums of Lüroth sets, which are defined by numbers whose Lüroth expansions have prescribed digit constraints. We establish several results on the congruence modulo 1 of sums of Lüroth sets, including summands with digits bounded above, below, and combinations of the two. We also analyse the Hausdorff dimension of Lüroth sets and their sums. The results extend classical findings on continued fractions to Lüroth expansions.

Keywords: Lüroth expansions; Sumsets; General Cantor sets; Hausdorff dimension; Fractals.

Mathematics Subject Classification 2020: 11J70, 11J83, 11K55, 28A80, 37A05

1. Introduction

Continued fractions and dynamical systems are fundamental concepts in mathematics. Connections between number theory and fractal geometry can be studied via decomposing numbers as sums of elements from specific sets, such as those defined by restricted digit representations.

Let $A, B \subseteq \mathbb{R}$ be two subsets of the real numbers. The sumset of A and B , denoted $A + B$, is defined as the Minkowski sum of A and B :

$$A + B := \{a + b : a \in A \text{ and } b \in B\}.$$

*Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus, Denmark

2 *Maiken Gravgaard, Ying Wai Lee*

The above definition extends naturally to the sumset of multiple subsets of $\mathbb{R}$. For any $n \in \mathbb{N}$, the n -fold sumset of A with itself, denoted nA , is defined as:

$$nA := \{a_1 + a_2 + \cdots + a_n : a_1, a_2, \dots, a_n \in A\}.$$

We adopt the following notation for congruence modulo 1: $A \equiv B \pmod{1}$, if $A + \mathbb{Z} = B + \mathbb{Z}$; that is every element of B can be written as the sum of an element of A and an integer, and vice versa. Otherwise, we write $A \not\equiv B \pmod{1}$.

Let $C_{1/3}$ be the middle-thirds Cantor set. A classical result^{23,26} states that

$$C_{1/3} + C_{1/3} = [0, 2]. \quad (1.1)$$

Although $C_{1/3}$ is a fractal with Lebesgue measure 0 and Hausdorff dimension $\log 2 / \log 3$, two of them sum up to an interval of length 2. The middle-thirds Cantor set and the result (1.1) can be understood from a number theory point of view. Notice that

$$C_{1/3} = \left\{ \sum_{n=1}^{\infty} \frac{a_n}{3^n} \in [0, 1] : a_n \in \{0, 2\} \text{ for all } n \in \mathbb{N} \right\},$$

that is, $C_{1/3}$ is the set of numbers whose ternary expansion does not contain any ones. With this interpretation, (1.1) implies that $C_{1/3} + C_{1/3} \equiv \mathbb{R} \pmod{1}$. In other words, every real number can be written as a sum of an integer and two numbers whose ternary expansions do not contain any ones.

This connection between base expansions and additive structure illustrates how fractal sets can possess rich number-theoretic properties. In particular, it motivates the study of sums of fractal sets and their role in problems concerning the decomposition of real numbers into sums of elements with restricted digit representations.

1.1. Results for Continued Fractions

With the above example in mind, one could choose a different representation for the real numbers than ternary expansions, for example, continued fractions. Define, for any $k \in \mathbb{N}$,

$$F_{\leq k} := \{[a_1, a_2, a_3, \dots] : 1 \leq a_n \leq k \text{ for all } n \in \mathbb{N}\},$$

where $[a_1, a_2, a_3, \dots]$ denotes the continued fraction for the partial quotients $(a_n)_{n \in \mathbb{N}}$, that is,

$$[a_1, a_2, a_3, \dots] := \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}.$$

Sumsets that are induced by continued fractions are studied by various mathematicians. In a work of Hall [15, Theorem 3.1], it is proved that every real number

can be expressed as a sum of an integer and two continued fractions with partial quotients not greater than 4; in other words,

$$F_{\leq 4} + F_{\leq 4} \equiv \mathbb{R} \pmod{1}.$$

In particular, every real number can be decomposed into the sum of two badly approximable numbers. Several papers have since appeared that extend this classical result. Below we mention the most relevant for our purposes. By employing refined versions of the techniques used by Hall, Hlavka [16, Theorems 6 & 9] proved that

$$F_{\leq 3} + F_{\leq 4} \equiv \mathbb{R} \pmod{1}; \quad F_{\leq 2} + F_{\leq 7} \equiv \mathbb{R} \pmod{1}.$$

Hlavka [16, Theorems 11] also extended the study into the sums of multiple sets:

$$\begin{aligned} F_{\leq 2} + F_{\leq 3} + F_{\leq 3} &\equiv \mathbb{R} \pmod{1}, \\ F_{\leq 2} + F_{\leq 2} + F_{\leq 4} &\equiv \mathbb{R} \pmod{1}, \\ F_{\leq 2} + F_{\leq 2} + F_{\leq 2} + F_{\leq 2} &\equiv \mathbb{R} \pmod{1}. \end{aligned}$$

Hlavka¹⁶ also proved that $F_{\leq 2} + F_{\leq 4} \not\equiv \mathbb{R} \pmod{1}$. Diviš [11, Theorems 1 & 2] proved that some of these results of Hlavka are optimal, in the sense that

$$F_{\leq 3} + F_{\leq 3} \not\equiv \mathbb{R} \pmod{1}; \quad F_{\leq 2} + F_{\leq 2} + F_{\leq 2} \not\equiv \mathbb{R} \pmod{1}.$$

Astels [2, Theorem 1.1] [3, Theorem 1] later invented another tool to prove that

$$F_{\leq 2} + F_{\leq 5} \equiv \mathbb{R} \pmod{1}; \quad F_{\leq 2} + F_{\leq 2} + F_{\leq 3} \equiv \mathbb{R} \pmod{1}.$$

The sets in which the partial quotients are bounded from below are studied by Cusick among others. Define for any $k \in \mathbb{N}$,

$$F_{\geq k} := \{[a_1, a_2, a_3, \dots] : a_n \geq k \text{ for all } n \in \mathbb{N}\} \cup F'_{\geq k},$$

where $F'_{\geq k}$ is the set consisting of 0 and all continued fractions of finite lengths with partial quotients at least k ; that is, continued fractions of finite lengths are also allowed in $F_{\geq k}$. Cusick [7, Theorem 1] proved that

$$F_{\geq 2} + F_{\geq 2} \equiv \mathbb{R} \pmod{1}.$$

Cusick–Lee⁸ extend this result by showing that for any $k \in \mathbb{N}$,

$$kF_{\geq k} \equiv \mathbb{R} \pmod{1}. \tag{1.2}$$

Good [14, Theorem 2] proved that for any $k \in \mathbb{N}$, if $k \geq 20$ then

$$\frac{1}{2} + \frac{1}{2 \log(k+2)} < \dim F_{\geq k} < \frac{1}{2} + \frac{\log \log(k-1)}{2 \log(k-1)},$$

where $\dim$ denotes the Hausdorff dimension. Further refinements of Good's bounds, along with an analysis of the Hausdorff dimension of the sets $F_{\leq k}$, were recently shown by Das–Fishman–Simmons–Urbański in.¹⁰ Shulga [25, Theorem 1.2] invented an algorithm and proved that for any $k \in \mathbb{N}$, if $k \geq 2$ then

$$\left[0, \frac{1}{k-1}\right] \subseteq F_{\geq k} + F_{\geq k};$$

4 Maiken Gravgard, Ying Wai Lee

as well as some other results on sets of the form $F_{\geq k_1} + F_{\geq k_2}$. In a recent work, Gayfulin–Nesharim¹³ derive more results for the summands in Shulga’s algorithm.

Difference and product sets derived from continued fraction expansions are also studied by Astels.^{1–3} However, we shift our focus to the Lüroth expansions and seek to establish analogous results for their corresponding sumsets.

1.2. Lüroth Expansion

In this subsection, we explain what the Lüroth expansion (or the Lüroth representation of numbers) is, and why its fractal sumsets are natural to study, due to the dynamical systems involved. For any $x \in (0, 1]$, there exists a unique sequence $(d_n)_{n \in \mathbb{N}}$ of positive integers not equal to 1, referred to as digits, such that

$$\begin{aligned} x &= [d_1, d_2, d_3, \dots] \\ &:= \sum_{n=1}^{+\infty} \frac{1}{d_n \prod_{j=1}^{n-1} d_j (d_j - 1)} \\ &= \frac{1}{d_1} + \frac{1}{d_1(d_1 - 1)d_2} + \frac{1}{d_1(d_1 - 1)d_2(d_2 - 1)d_3} + \dots, \end{aligned}$$

where for any $n \in \mathbb{N}$, $d_n \in \mathbb{N} \setminus \{1\}$. The notation $[d_1, d_2, d_3, \dots]$ denotes the Lüroth representation when the brackets contain d ’s instead of a ’s, distinguishing it from the continued fraction representation $[a_1, a_2, a_3, \dots]$.

The middle-thirds Cantor set $C_{1/3}$ has an invariant property under the times-3 map (see Figure 1). The map shifts base-3 digits left (a symbolic shift), so the map restricted to $C_{1/3}$ is a dynamical system conjugate to the full shift on two symbols, meaning it mimics the dynamics of all possible binary sequences (that is, using symbols 0 and 2). Thus, it preserves the structure of $C_{1/3}$, making $C_{1/3}$ invariant under the map and a rich object in dynamical systems theory.

Just as the times-3 map preserves the middle-thirds Cantor set, a more complicated map preserves continued fraction expansions with restricted partial quotients. The Gauss map (also known as the Gauss–Kuzmin–Wirsing operator) generates the continued fraction partial quotients and dynamically shifts the continued fraction expansions. Hence, the sets $F_{\leq k}$ and $F_{\geq k}$ are both invariant under this map. The Gauss map restricted to $F_{\leq k}$ (or $F_{\geq k}$) models a symbolic dynamical system on the finite alphabet $\{1, 2, \dots, k\}$ (or infinite alphabet $\{k, k + 1, \dots\}$). The set $F_{\leq k}$ is also well-studied in Diophantine approximation (see [4, Theorem 1.4]) and Fourier analysis (see^{17, 22, 24}).

While the times-3 map consists of finitely many piecewise linear segments and the Gauss map consists of infinitely many piecewise non-linear segments (see Figure 2), both play fundamental roles in dynamical systems and number theory. Each generates fractal structures in the interval $[0, 1]$ through their invariant sets, and the sumsets of these fractals have been extensively studied. It is therefore natural to consider their synthesis, known as the *linearised Gauss map* or the *Lüroth map*, combining linearity and infinite segments (see Figure 3). Like its predecessors, it

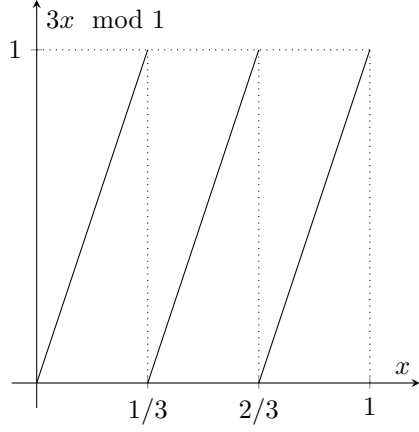


Fig. 1: Times-3 Map

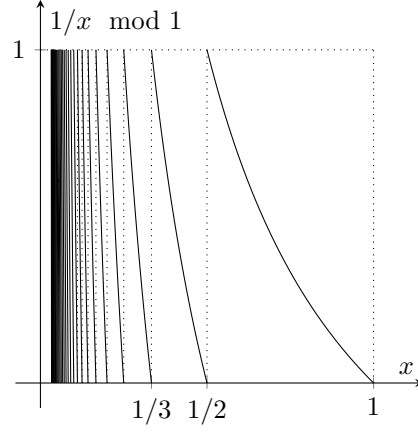


Fig. 2: Gauss–Kuzmin–Wirsing Operator

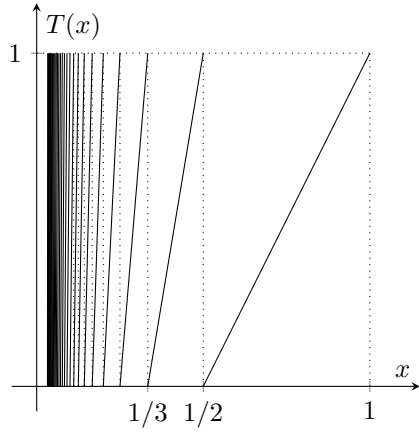
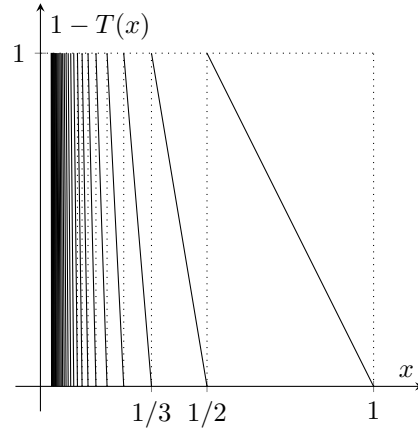
Fig. 3: Lüroth Map $T : (0, 1] \rightarrow (0, 1]$ 

Fig. 4: Alternating Lüroth Map

exhibits rich arithmetic properties and gives rise to fractals in $[0, 1]$ whose sumsets display intriguing behaviour.

The Lüroth map, denoted $T : (0, 1] \rightarrow (0, 1]$, is defined by, for any $x \in (0, 1]$,

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

The Lüroth map, as well as the other two mentioned maps above, can generate the digits or the partial quotients of numbers (see⁹). Recent studies on the Lüroth map and Lüroth expansion include the metric theory of Diophantine approximation.^{6, 19, 27, 28}

Define, for any $\mathcal{A} \subset \mathbb{N} \setminus \{1\}$, the Lüroth set $L_{\mathcal{A}}$ to be the set of numbers whose

6 *Maiken Gravgard, Ying Wai Lee*

digits in their Lüroth representations are restricted to $\mathcal{A}$; that is

$$L_{\mathcal{A}} := \{[d_1, d_2, d_3, \dots] : d_n \in \mathcal{A} \text{ for all } n \in \mathbb{N}\}.$$

These kinds of sets have also been considered in previous work.^{5,18} For any $\mathcal{A} \subseteq \mathbb{N} \setminus \{1\}$, the Lüroth set $L_{\mathcal{A}}$ is invariant under the Lüroth map T .

Our focus is on the following sets, as we aim to draw analogous conclusions to the continued fraction sets we discussed. Define for any $k \in \mathbb{N} \setminus \{1\}$, Lüroth sets of digits bounded from above $L_{\leq k}$ and respectively below $L_{\geq k}$ by:

$$L_{\leq k} := L_{\{2,3,\dots,k\}}, \quad L_{\geq k} := L_{\{k,k+1,\dots\}},$$

and for any $N_1, N_2 \in \mathbb{N} \setminus \{1\}$,

$$L_{N_1, N_2} := L_{\geq N_1} \cap L_{\leq N_2}.$$

In the work of Brouwer [5, (6.1)], an attempt is made to apply Hall's tool [15, Theorem 2.2] in establishing the following statement on the sum of Lüroth sets:

$$L_{\leq 4} + L_{\leq 4} \equiv \mathbb{R} \pmod{1}.$$

The idea of a proof is outlined but not fully completed in.⁵ We verify that the above claim is true by applying a tool developed by Hlavka [16, Theorem 3] (the tools are summarised in Section 3), alongside this, we present further results on the sums of Lüroth sets in this paper.

2. Main Results

In this paper, we provide two primary types of results concerning Lüroth sets. The first type is on the equivalence modulo 1 of sums of Lüroth sets, as presented in Section 2.1. The second type is on the Hausdorff dimension of Lüroth sets and their sums, as presented in Section 2.2.

2.1. Congruence

Theorem 2.1 can be seen as an analogue to the results of Hall, Hlavka, and Astels.

Theorem 2.1. *We have*

$$\begin{aligned} L_{\leq 3} + L_{\leq 5} &\equiv \mathbb{R} \pmod{1}; \\ L_{\leq 4} + L_{\leq 4} &\equiv \mathbb{R} \pmod{1}; \\ L_{\leq 3} + L_{\leq 3} + L_{\leq 3} &\equiv \mathbb{R} \pmod{1}; \\ L_{\leq 3} + L_{\leq 3} &\not\equiv \mathbb{R} \pmod{1}. \end{aligned}$$

Both $L_{\leq 3} + L_{\leq 3}$ and $L_{\leq 3} + L_{\leq 4}$ are not intervals.

However, we are unable to determine whether $L_{\leq 3} + L_{\leq 4} + \mathbb{Z}$ equals $\mathbb{R}$. Figure 8 indicates that many sums in $L_{\leq 3} + L_{\leq 4}$ cluster densely enough to suggest the

presence of a genuine interval, say $[1, 2]$. This observation naturally leads to the following conjecture.

Conjecture 2.2.

$$L_{\leq 3} + L_{\leq 4} \equiv \mathbb{R} \pmod{1}.$$

By applying a result of Astels [1, Theorem 2.4] (see Proposition 3.4 in Section 3), we give Theorem 2.3 which corresponds to the continued fraction setting as in [1, Theorem 1.5].

Theorem 2.3. *Let $n \in \mathbb{N}$ and $k_1, k_2, \dots, k_n \in \mathbb{N} \setminus \{1\}$. Suppose*

$$\sum_{j=1}^n \frac{k_j - 1}{k_j(k_j - 1) - 1} \geq 1, \quad (2.1)$$

and, for any $j \in \{2, \dots, n\}$,

$$k_{j-1} \leq k_j \leq k_{j-1}^2 + 2k_{j-1} + 2. \quad (2.2)$$

Then

$$L_{\geq k_1} + L_{\geq k_2} + \dots + L_{\geq k_n} \equiv \mathbb{R} \pmod{1}.$$

Theorem 2.3 gives, for example,

$$L_{\geq 3} + L_{\geq 3} + L_{\geq 5} \equiv \mathbb{R} \pmod{1}, \quad L_{\geq 3} + L_{\geq 4} + L_{\geq 5} + L_{\geq 6} \equiv \mathbb{R} \pmod{1}.$$

Condition (2.1) guarantees that the total length of the Lüroth sets involved is at least 1. More precisely, the sum in condition (2.1) is exactly the diameter of the sumset. In particular, if this sum is strictly less than 1, then

$$L_{\geq k_1} + L_{\geq k_2} + \dots + L_{\geq k_n} \not\equiv \mathbb{R} \pmod{1},$$

simply because the total length is insufficient. The only case in which the sum in condition (2.1) is equal to 1 is $n = 1$ and $k_1 = 2$, where the conclusion of Theorem 2.3 is trivially true.

Condition (2.2) plays a different role: it ensures that the sequence $(k_j)_j$ is not too widely spread, so that the maximal gap of each Lüroth set can be filled by the next one in the sum. Note that condition (2.1) does not imply condition (2.2). For example, Theorem 2.3 does not apply to the sumset

$$L_{\geq 3} + L_{\geq 4} + L_{\geq 5} + L_{\geq 9} + L_{\geq 245}.$$

It remains unknown whether a condition of the form (2.2) is actually necessary in Theorem 2.3 when the sum in condition (2.1) is greater than 1.

Corollary 2.4 below is a direct application of Theorem 2.3, and can be seen as an analogue to the result (1.2) of Cusick–Lee. We do, however, give a separate proof of this corollary as well.

Corollary 2.4. *For any $k \in \mathbb{N} \setminus \{1\}$,*

$$kL_{\geq k} \equiv \mathbb{R} \pmod{1}. \quad (2.3)$$

While (2.3) can be compared with (1.2), it is worth noting that in the definition of $L_{\geq k}$, we do not include Lüroth representations of finite lengths. We note that for $k = 2$, (2.3) is trivially true, as $L_{\geq 2} = (0, 1] \equiv \mathbb{R} \pmod{1}$. (2.3) is optimal in the sense that for $k \in \mathbb{N}$, if $k \geq 3$ then

$$(k-1)L_{\geq k} \not\equiv \mathbb{R} \pmod{1},$$

as the diameter of $L_{\geq k}$ is not sufficient:

$$(k-1) \operatorname{diam} L_{\geq k} = \frac{(k-1)^2}{(k-1)k-1} = 1 - \frac{k-2}{k^2-k-1} < 1.$$

We also suggest a new type of sumsets to study: mixing Lüroth sets whose digits are both bounded above and below, as in Theorem 2.5.

Theorem 2.5. *For any $k \in \mathbb{N} \setminus \{1\}$,*

$$L_{\leq k+2} + L_{\geq k} \equiv \mathbb{R} \pmod{1}.$$

One can observe that $\operatorname{diam}(L_{\leq k} + L_{\geq k}) = 1$ and suggest the following conjecture as a stronger statement than Theorem 2.5.

Conjecture 2.6. *For any $k \in \mathbb{N} \setminus \{1\}$,*

$$L_{\leq k} + L_{\geq k} \equiv \mathbb{R} \pmod{1}.$$

Theorems 2.1 and 2.3, together with Corollary 2.4, are motivated by their analogues in the continued-fraction setting. This naturally raises the question of whether the converse direction also holds; namely, whether Theorem 2.5 and Conjecture 2.6 admit counterparts for continued fractions. One may notice that for any $k \in \mathbb{N}$, if $k \geq 2$ then

$$\operatorname{diam}(F_{\leq k} + F_{\geq k}) = \frac{\sqrt{k^2 + 4k} - k}{2} - \frac{\sqrt{k^2 + 4k} - k}{2k} + \frac{1}{k} < 1;$$

in particular, $F_{\leq k} + F_{\geq k} \not\equiv \mathbb{R} \pmod{1}$, as the diameter of the sum is not sufficient to make it contain an interval of length at least 1. One can formulate the following alternative conjecture instead.

Conjecture 2.7. *For any $k \in \mathbb{N}$,*

$$F_{\leq 2k} + F_{\geq k} \equiv \mathbb{R} \pmod{1}.$$

In the context of Diophantine approximation, the result of Hall implies that every real number can be decomposed as a sum of two badly approximable numbers. Erdős¹² invented an algorithm based on binary representation and proved that every real number is a sum of two Liouville numbers. In an upcoming work, Gayfulin–Nesharim,¹³ claims that in Shulga’s algorithm, the partial quotients of

both summands are unbounded for all irrational numbers. Both results of Erdős and Gayfulin–Nesharim imply that every real number is a sum of two well approximable numbers.

The next natural question would be whether every real number can be decomposed as a sum of a badly approximable number and a well approximable number. Conjecture 2.7 above is in the spirit of this question, but does not capture it exactly: while $F_{\leq k}$ is a proper subset of the badly approximable numbers, and there exist numbers in $F_{\geq k}$ that are not well approximable. Nevertheless, Conjecture 2.8 below asks this as a bigger question.

Conjecture 2.8.

$$\bigcup_{k \in \mathbb{N}} F_{\leq k} + [0, 1] \setminus \bigcup_{k \in \mathbb{N}} F_{\leq k} \equiv \mathbb{R} \pmod{1}.$$

2.2. Dimension

We give some results on the Hausdorff dimension of the Lüroth sets and their sums. Theorem 2.9 gives the Hausdorff dimension of sums of Lüroth sets in terms of the respective Hausdorff dimensions.

Theorem 2.9. *For any finite $\mathcal{A}, \mathcal{B} \subseteq \mathbb{N} \setminus \{1\}$,*

$$\dim(L_{\mathcal{A}} + L_{\mathcal{B}}) = \min\{1, \dim L_{\mathcal{A}} + \dim L_{\mathcal{B}}\}.$$

By an application of [20, Theorem II], the Hausdorff dimension of $L_{\mathcal{A}}$ is the unique solution $s \in [0, 1]$ to

$$\sum_{d \in \mathcal{A}} \left(\frac{1}{d(d-1)} \right)^s = 1.$$

We can approximate that $\dim L_{\leq 3} = 0.600967\dots > 1/2$ and consequently $\dim(L_{\leq 3} + L_{\leq 3}) = 1$.

Theorem 2.10 is an analogy to the lower bound estimate to the result of Good [14, Theorem 2].

Theorem 2.10. *For any $k \in \mathbb{N} \setminus \{1\}$,*

$$\dim L_{\geq k} > \frac{1}{2} + \frac{1}{2 \log \max\{16, k\}}.$$

Notice that for any $k, N \in \mathbb{N} \setminus \{1\}$, if $k < N$ then we have $\dim L_{\geq k} \geq \dim L_{k, N}$. Hence, we can find a lower bound of $\dim L_{\geq k}$ by finding the supremum of $s_{k, N} > 0$ such that

$$\sum_{d=k}^N \left(\frac{1}{d(d-1)} \right)^{s_{k, N}} > 1.$$

10 *Maiken Gravgaard, Ying Wai Lee*

For some small values of $k \in \mathbb{N} \setminus \{1\}$, we have,

$$\begin{array}{lll} \dim L_{\geq 3} > 0.8209; & \dim L_{\geq 4} > 0.7740; & \dim L_{\geq 5} > 0.7500; \\ \dim L_{\geq 6} > 0.7347; & \dim L_{\geq 7} > 0.7239; & \dim L_{\geq 8} > 0.7157; \end{array}$$

and so on. The improvement is not significant for larger values of k .

Shulga's algorithm does not directly extend to the setting of Lüroth expansions, for example in the case of $k = 3$ and

$$x = \frac{12}{41} = [\dot{7}] + [\dot{7}] \in (L_{\geq 3} + L_{\geq 3}) \cap \left[0, \frac{1}{2}\right],$$

where the dot indicates that, in the Lüroth expansion, all subsequent digits following the initial digit are identical to the digit marked with a dot. Direct application of Shulga's algorithm to this case does not produce two Lüroth expansions with digits bounded below by 3. Instead, the algorithm yields the following Lüroth representations:

$$c = [4, 25, 12, 5, 5, 6, 6, 6, 5, 5, 5, \dots], \quad b = [11, 16, 8, 2, 4, 2, 2, 2, 2, 2, 3, \dots].$$

However, we propose a result and a conjecture, as analogies to the result of Shulga.

Theorem 2.11. *For any $k \in \mathbb{N} \setminus \{1\}$,*

$$\dim(L_{\geq k} + L_{\geq k}) = 1.$$

Conjecture 2.12. *For any $k \in \mathbb{N} \setminus \{1\}$, $L_{\geq k} + L_{\geq k}$ contains an interval of positive length in $\mathbb{R}$.*

3. Results on General Cantor Sumsets

In this section, we introduce the definition and some known results from the literature on general Cantor sets. They are useful in proving our main results. General Cantor sets are defined in a similar construction to the middle-thirds Cantor set. It starts from a closed and bounded interval in $\mathbb{R}$, and proceeds by removing open intervals at each level. The fractals of $F_{\leq k}$ and $F_{\geq k}$ are examples of general Cantor sets, so are $L_{\leq k}$, $L_{\geq k}$ and many other self-similar sets.

We generalise the middle-thirds Cantor set $C_{1/3}$ into the general Cantor set $C_{\mathcal{A}}$ by the following construction. Let I be the minimal closed interval that contains the targeted fractal set. Put $I^1 := I$, which is Level 0 in the construction. Let $O^1 \subseteq I^1$ be an open interval. Then, at Level 1, the two closed intervals I^2 and I^3 are formed by removing O^1 from I^1 :

$$I^1 = I^2 \cup O^1 \cup I^3.$$

To construct a general Cantor set, at each step $i \in \mathbb{N}$, we remove an open interval O^i from the closed, non-empty interval I^i , retaining the pair of closed, non-empty intervals I^{2i} and I^{2i+1} , and continue this process iteratively across each level. Figure 5 is an illustration.

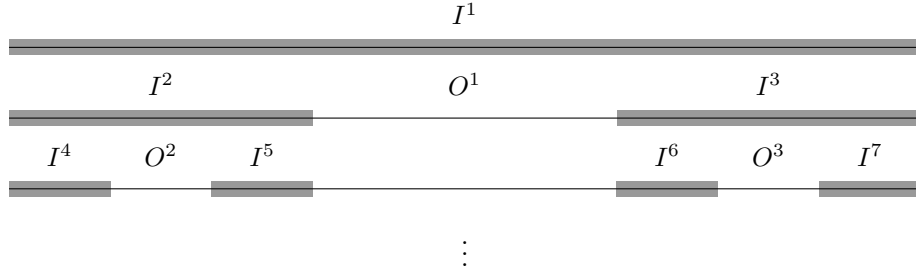


Fig. 5: A Construction of a General Cantor Set

The general Cantor set is obtained as $C := I \setminus \bigcup_{i \in \mathbb{N}} O^i$, and $D := ((I^i)_{i \in \mathbb{N}}, (O^i)_{i \in \mathbb{N}})$ is referred to as a construction (or a derivation) of C . Note that there can be more than one construction that results in the same general Cantor set. For any interval $I \subseteq \mathbb{R}$, denote by $|I|$ the length of I . Hall¹⁵ proved the following result.

Proposition 3.1 (Hall [15, Theorem 2.2], 1947). *Let C_A and C_B be two general Cantor sets with constructions D_A and D_B . Write $I_A^1 = [x, x + a]$, $I_B^1 = [y, y + b]$ and $e = \min\{a, b\}$. Suppose for any interval I^i in either construction, with $I^i = I^{2i} \cup O^i \cup I^{2i+1}$, that*

$$|O^i| \leq \min\{|I^{2i}|, |I^{2i+1}|\}.$$

Then $[x + y, x + y + 2e] \cup [x + y + a + b - 2e, x + y + a + b] \subseteq C_A + C_B$. In particular, if $1/3 \leq a/b \leq 3$, then $C_A + C_B = I_A^1 + I_B^1$.

Two quantities are defined by Hlavka in¹⁶ regarding a construction of a general Cantor set. Define g_A and h_A on the construction D_A of a general Cantor set.

$$g_A = g(D_A) := \sup_{j \in \mathbb{N}} \frac{|O_A^j|}{|I_A^j|}, \quad h_A = h(D_A) := \inf_{j \in \mathbb{N}} \min \left\{ \frac{|I_A^{2j}|}{|I_A^j|}, \frac{|I_A^{2j+1}|}{|I_A^j|} \right\},$$

where g_A captures the *maximal removal ratio* and h_A captures the *minimal retaining ratio* of the construction D_A . By the quantities defined above, Hlavka proved the following results regarding the sum of two or more general Cantor sets.

Proposition 3.2 (Hlavka [16, Theorem 3], 1975). *Suppose there exist Cantor set constructions of C_A and C_B with all of the following conditions are satisfied:*

$$g_A g_B \leq h_A h_B, \quad g_A |I_A| \leq |I_B|, \quad g_B |I_B| \leq |I_A|.$$

Then $C_A + C_B = I_A + I_B$.

Proposition 3.3 (Hlavka [16, Theorem 10], 1975). *Suppose there exist Cantor set constructions of $C_{A_1}, \dots, C_{A_n}$ such that both of the following conditions are*

12 Maiken Gravggaard, Ying Wai Lee

satisfied: for any $i, j \in \{1, \dots, n\}$,

$$h_{\mathcal{A}_i} \leq \frac{|I_{\mathcal{A}_i}|}{|I_{\mathcal{A}_j}|},$$

and for any $i \in \{1, \dots, n\}$,

$$g_{\mathcal{A}_i} + h_{\mathcal{A}_i} \leq \sum_{j=1}^n h_{\mathcal{A}_j}.$$

Then $C_{\mathcal{A}_1} + \dots + C_{\mathcal{A}_n} = I_{\mathcal{A}_1} + \dots + I_{\mathcal{A}_n}$.

The concept of *thickness* was introduced by Astels in.¹ For a construction $D_{\mathcal{A}}$ of the general Cantor set $C_{\mathcal{A}}$, we define

$$\tau(D_{\mathcal{A}}) := \inf_{i \in \mathbb{N}} \left\{ \min \left\{ \frac{|I_{\mathcal{A}}^{2i}|}{|O_{\mathcal{A}}^i|}, \frac{|I_{\mathcal{A}}^{2i+1}|}{|O_{\mathcal{A}}^i|} \right\} \right\} \geq \frac{h_{\mathcal{A}}}{g_{\mathcal{A}}}.$$

Define the *thickness* of the general Cantor set $C_{\mathcal{A}}$ by the supremum of $\tau(D_{\mathcal{A}})$ over all constructions $D_{\mathcal{A}}$ for $C_{\mathcal{A}}$, that is

$$\tau_{\mathcal{A}} = \tau(C_{\mathcal{A}}) := \sup_{D_{\mathcal{A}}} \tau(D_{\mathcal{A}}),$$

and also another quantity γ on the general Cantor set $C_{\mathcal{A}}$:

$$\gamma_{\mathcal{A}} = \gamma(C_{\mathcal{A}}) := \frac{\tau_{\mathcal{A}}}{1 + \tau_{\mathcal{A}}} \in [0, 1).$$

A construction $D = ((I^i)_{i \in \mathbb{N}}, (O^i)_{i \in \mathbb{N}})$ is said to be ordered, if for any $i \in \mathbb{N}$,

$$|O^i| \geq \max \{|O^{2i}|, |O^{2i+1}|\};$$

that is, the lengths of removing intervals are always not greater than those at the previous level. One can observe that the following statements are equivalent to each other:

- $D = ((I^i)_{i \in \mathbb{N}}, (O^i)_{i \in \mathbb{N}})$ is ordered;
- for any $i \in \mathbb{N}$, $|O^i| \geq \sup\{|O^{2i}|, |O^{2i+1}|, |O^{2(2i)}|, |O^{2(2i)+1}|, |O^{2(2i+1)}|, |O^{2(2i+1)+1}|, \dots\}$;
- for any $i, j \in \mathbb{N}$, if $I^i \subseteq I^j$ then $|O^i| \leq |O^j|$.

D is said to be unordered if D is not ordered. It is not hard to see that an ordered construction always exists. If $D_{\mathcal{A}}$ is ordered, then the thickness of the general Cantor set is exactly the thickness of $D_{\mathcal{A}}$; that is

$$\tau_{\mathcal{A}} = \tau(D_{\mathcal{A}}). \quad (3.1)$$

With the newly defined quantities, Astels established the following useful result concerning the sum of general Cantor sets. They are also applied in proving our main results.

Proposition 3.4 (Astels [1, Theorem 2.4], 2000). *Let $S_{\gamma} = \gamma_{\mathcal{A}_1} + \dots + \gamma_{\mathcal{A}_n}$.*

- (1) If $S_\gamma \geq 1$, then $C_{\mathcal{A}_1} + \cdots + C_{\mathcal{A}_n}$ contains an interval. Otherwise, it contains a Cantor set of thickness at least

$$\frac{S_\gamma}{1 - S_\gamma}.$$

Furthermore,

$$\dim(C_{\mathcal{A}_1} + \cdots + C_{\mathcal{A}_n}) \geq \min \left\{ 1, \frac{\log 2}{\log(1 + 1/S_\gamma)} \right\}$$

- (2) Let $O_{\mathcal{A}_j}$ be a gap of maximal size in $C_{\mathcal{A}_j}$ (that is $|O_{\mathcal{A}_j}| = \sup_{i \in \mathbb{N}} |O_{\mathcal{A}_j}^i|$). Suppose for any $i \in \{2, \dots, n\}$ and $j \in \{1, \dots, i-1\}$,

$$|I_{\mathcal{A}_i}| \geq |O_{\mathcal{A}_j}|, \quad (3.2)$$

and for any $i \in \{1, \dots, n-1\}$,

$$|I_{\mathcal{A}_1}| + \cdots + |I_{\mathcal{A}_i}| \geq |O_{\mathcal{A}_{i+1}}|. \quad (3.3)$$

Then if $S_\gamma \geq 1$ we have $C_{\mathcal{A}_1} + \cdots + C_{\mathcal{A}_n} = I_{\mathcal{A}_1} + \cdots + I_{\mathcal{A}_n}$. Otherwise,

$$\tau_{C_{\mathcal{A}_1} + \cdots + C_{\mathcal{A}_n}} \geq \frac{S_\gamma}{1 - S_\gamma}.$$

4. Constructions of Lüroth Sets

In this section, we give the constructions of the Lüroth sets as general Cantor sets. The constructions below can be compared to the construction given in¹⁸ to observe the differences.

Let $N_1, N_2 \in \mathbb{N} \setminus \{1\}$. Suppose $N_1 < N_2$. Define $I_{N_1, N_2} := [\langle N_2, N_1 \rangle]$, where for any $d \in \mathbb{N} \setminus \{1\}$, the left chevron and the right chevron are respectively defined by

$$\begin{aligned} \langle d := [d, \dot{N}_2] &= \frac{1}{d} + \frac{1}{d(d-1)} \frac{N_2 - 1}{(N_2 - 1)N_2 - 1}, \\ d \rangle := [d, \dot{N}_1] &= \frac{1}{d} + \frac{1}{d(d-1)} \frac{N_1 - 1}{(N_1 - 1)N_1 - 1}, \end{aligned}$$

where the dot indicates that, in the Lüroth expansion, all subsequent digits following the initial digit are identical to the digit marked with a dot. When an expression includes at least one chevron, the order of operations is as follows: first, perform multiplication, division, and fractional operations; then, proceed with addition, chevron operations, and finally subtraction.

We construct the Lüroth set L_{N_1, N_2} in the following *Stepwise Complete Construction (SCC)* as a general Cantor set. At Level 0, we start with the following closed interval:

$$I_{N_1, N_2}^1 := I_{N_1, N_2} = [\langle N_2, N_1 \rangle].$$

For any $i \in \{1, \dots, N_2 - N_1\}$, at Level i , we have

$$I_{N_1, N_2}^{2^i} = [\langle N_2, N_1 + i \rangle], O_{N_1, N_2}^{2^{i-1}} = (N_1 + i, \langle (N_1 + i - 1) \rangle), I_{N_1, N_2}^{2^i + 1} = [\langle (N_1 + i - 1), (N_1 + i - 1) \rangle].$$

The intervals $I_{N_1, N_2}^{2^i+1}$ and $I_{N_1, N_2}^{2^{(N_2-N_1)}}$ are all treated proportionally to $I_{\mathcal{A}}^1$ to give all further levels. We have $L_{N_1, N_2} = I_{N_1, N_2} \setminus \bigcup_{i \in \mathbb{N}} O_{N_1, N_2}^i$. The SCC for $L_{\leq 3}$, $L_{\leq 4}$, and $L_{\leq 5}$ illustrated in Figure 6 as examples.

To understand why the SCC yields L_{N_1, N_2} , observe that the numbers $\langle N_2$ and $N_1 \rangle$ are respectively the smallest and the greatest elements in L_{N_1, N_2} . For each digit $d \in \{N_1, \dots, N_2\}$, the number $\langle d = [d, \dot{N}_2] \rangle$ is the smallest element in L_{N_1, N_2} whose Lüroth expansion begins with d , while $\langle d = [d, \dot{N}_1] \rangle$ is the largest such element. Thus, the interval $[\langle d, d \rangle]$ contains all numbers whose Lüroth expansions have first digit d , and all onward digits behave the same as those in I_{N_1, N_2} . Hence, the restriction on the first digit to d is completed at each step (the last step, that is also at the last level, completes the restriction for digits first digit being $N_2 - 1$ and N_2 at the same time). The gap $\langle d+1 \rangle, \langle d \rangle$ consists of numbers whose second digit is less than N_1 or greater than N_2 , and thus are excluded. This process is recursively applied to each retained subinterval through affine similarity, resulting in the same digit constraints on all further positions. Therefore, the SCC yields exactly the set L_{N_1, N_2} , realised as a general Cantor set.

For example, in $L_{\leq 3}$, we have at Level 0:

$$I_3^1 = I_3 = [\langle 3, 2 \rangle] = \left[\frac{2}{5}, 1 \right];$$

and at Level 1:

$$I_3^2 = [\langle 3, 3 \rangle] = \left[\frac{2}{5}, \frac{1}{2} \right], \quad O_3^1 = (3), \langle 2 \rangle = \left(\frac{1}{2}, \frac{7}{10} \right), \quad I_3^3 = [\langle 2, 2 \rangle] = \left[\frac{7}{10}, 1 \right];$$

I_3^2 and I_3^3 are both treated proportionally to I_3^1 to give all further levels.

In general, for Lüroth sets of the form L_{N_1, N_2} , the g - and h - values of the SCCs are given by their finite steps due to self-similarity. The following proposition gives the lengths of the intervals in the SCCs; we can also see that all SCCs are ordered.

Proposition 4.1. *Let $N_1, N_2 \in \mathbb{N} \setminus \{1\}$. Suppose $N_1 < N_2$. Under SCC, for any $i \in \{1, \dots, N_2 - N_1\}$, at Level i ,*

- *the main interval of length (the interval on the left at Level $i - 1$):*

$$\left| I_{N_1, N_2}^{2^{i-1}} \right| = (N_1 + i - 1) - \langle N_2 \rangle;$$

- *the interval on the left of length:*

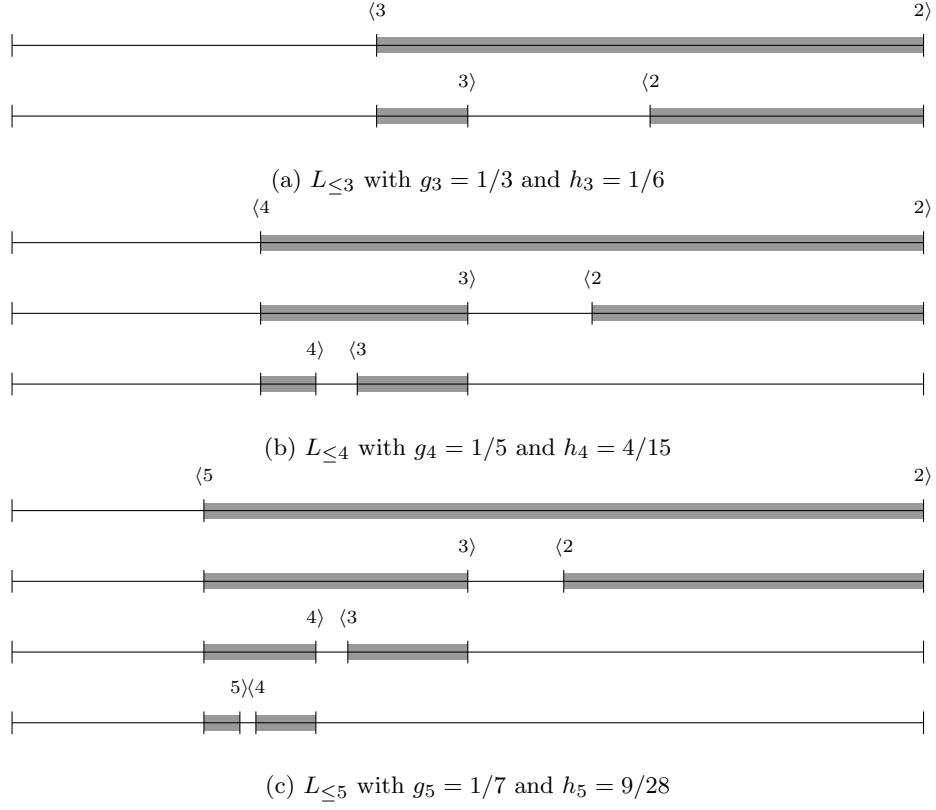
$$\left| I_{N_1, N_2}^{2^i} \right| = N_1 + i - \langle N_2 \rangle;$$

- *the gap of length (strictly decreasing):*

$$\left| O_{N_1, N_2}^{2^{i-1}} \right| = \langle (N_1 + i - 1) - N_1 + i \rangle; \tag{4.1}$$

- *the interval on the right of length:*

$$\left| I_{N_1, N_2}^{2^i+1} \right| = (N_1 + i - 1) - \langle (N_1 + i - 1) \rangle.$$

Fig. 6: Stepwise Complete Constructions (SCCs) for $L_{\leq 3}$, $L_{\leq 4}$, and $L_{\leq 5}$ in $[0, 1]$

Proposition 4.2. *Let $N_1, N_2 \in \mathbb{N} \setminus \{1\}$. Suppose $N_1 < N_2$. The SCC for L_{N_1, N_2} is ordered.*

Proof. By self-similarity, it suffices to prove that $i \in \{1, \dots, N_2 - N_1 - 1\}$,

$$\left| O_{N_1, N_2}^{2^{i-1}} \right| \geq \max \left\{ \left| O_{N_1, N_2}^{2^i} \right|, \left| O_{N_1, N_2}^{2^{i+1}} \right| \right\}.$$

On the one hand, from (4.1), one has that for any $i \in \{1, \dots, N_2 - N_1 - 1\}$,

$$\left| O_{N_1, N_2}^{2^{i-1}} \right| = \frac{1}{N_1 + i - 1} + \frac{\langle N_2 \rangle}{(N_1 + i - 1)(N_1 + i - 2)} - \frac{1}{N_1 + i} - \frac{N_1}{(N_1 + i)(N_1 + i - 1)}.$$

16 *Maiken Gravgard, Ying Wai Lee*

One can see the following difference is positive:

$$\begin{aligned}
\Delta_{N_1, N_2, i, 0} &:= \left| O_{N_1, N_2}^{2^{i-1}} \right| - \left| O_{N_1, N_2}^{2^i} \right| \\
&= \frac{1}{N_1 + i - 1} - \frac{2}{N_1 + i} + \frac{1}{N_1 + i + 1} + \frac{\langle N_2 \rangle}{N_1 + i - 1} \left(\frac{1}{N_1 + i - 2} - \frac{1}{N_1 + i} \right) \\
&\quad - \frac{N_1 \rangle}{N_1 + i} \left(\frac{1}{N_1 + i - 1} - \frac{1}{N_1 + i + 1} \right) \\
&> \frac{1}{N_1 + i - 1} - \frac{2}{N_1 + i} + \frac{1}{N_1 + i + 1} - \frac{N_1 \rangle}{N_1 + i} \left(\frac{1}{N_1 + i - 1} - \frac{1}{N_1 + i + 1} \right) \\
&\geq N_1 \rangle \left(\frac{1}{N_1 + i - 1} - \frac{2}{N_1 + i} + \frac{1}{N_1 + i + 1} \right) - \frac{N_1 \rangle}{N_1 + i} \left(\frac{1}{N_1 + i - 1} - \frac{1}{N_1 + i + 1} \right) \\
&= N_1 \rangle \left(\frac{1}{N_1 + i - 1} - \frac{2}{N_1 + i} + \frac{1}{N_1 + i + 1} - \frac{1}{N_1 + i} \left(\frac{1}{N_1 + i - 1} - \frac{1}{N_1 + i + 1} \right) \right) \\
&= 0.
\end{aligned}$$

Hence, (4.1) is strictly decreasing, that is for any $i \in \{1, \dots, N_2 - N_1 - 1\}$,

$$\left| O_{N_1, N_2}^{2^{i-1}} \right| \geq \left| O_{N_1, N_2}^{2^i} \right|.$$

On the other hand, from similarity, one has that for any $i \in \{1, \dots, N_2 - N_1 - 1\}$,

$$\left| O_{N_1, N_2}^{2^i+1} \right| = \frac{|O_{N_1, N_2}^1|}{|I_{N_1, N_2}^1|} \left| I_{N_1, N_2}^{2^i+1} \right| = \alpha_{N_1, N_2} ((N_1 + i - 1) \rangle - \langle (N_1 + i - 1)),$$

where

$$\alpha_{N_1, N_2} := \frac{|O_{N_1, N_2}^1|}{|I_{N_1, N_2}^1|} = \frac{\langle N_1 - N_1 + 1 \rangle}{N_1 \rangle - \langle N_2 \rangle}.$$

One can see that the following difference is positive:

$$\begin{aligned}
\Delta_{N_1, N_2, i, 1} &:= \left| O_{N_1, N_2}^{2^{i-1}} \right| - \left| O_{N_1, N_2}^{2^i+1} \right| \\
&= (1 + \alpha_{N_1, N_2}) \langle (N_1 + i - 1) - N_1 + i \rangle - (N_1 + i - 1) \rangle \alpha_{N_1, N_2} \\
&= \frac{1}{N_1 + i - 1} \left(1 + \frac{\langle N_2 - \alpha_{N_1, N_2} (N_1) - \langle N_2 \rangle}{N_1 + i - 2} - \frac{N_1 \rangle}{N_1 + i} \right) - \frac{1}{N_1 + i} \\
&= \frac{1}{N_1 + i - 1} \left(1 + \frac{N_1 + 1 \rangle - \langle N_1 + \langle N_2 \rangle}{N_1 + i - 2} - \frac{N_1 \rangle}{N_1 + i} \right) - \frac{1}{N_1 + i} \\
&\geq \frac{1}{N_1 + i - 1} \left(1 + \frac{N_1 + 1 \rangle - 1/N_1}{N_1 + i - 2} - \frac{N_1 \rangle}{N_1 + i} \right) - \frac{1}{N_1 + i} \\
&= \frac{1}{N_1 + i - 1} \left(1 + \frac{1}{N_1 + i - 2} \frac{2 - N_1}{N_1^3 - 2N_1 - 1} - \frac{1}{N_1 + i} \frac{N_1 - 1}{N_1 (N_1 - 1) - 1} \right) - \frac{1}{N_1 + i} \\
&= \frac{(N_1 - 2) (N_1^3 + (i - 1) N_1^2 + (i - 3) N_1 - i)}{(N_1 + i - 2) (N_1 + i - 1) (N_1 + i) (N_1 + 1) (N_1^2 - N_1 - 1)} \\
&\geq 0.
\end{aligned}$$

Hence, for any $i \in \{1, \dots, N_2 - N_1 - 1\}$,

$$\left| O_{N_1, N_2}^{2^{i-1}} \right| \geq \left| O_{N_1, N_2}^{2^i+1} \right|. \quad \square$$

Since SCC are ordered, Proposition 4.1 also helps to compute the thickness of Lüroth sets of the form L_{N_1, N_2} and their γ -values. This is discussed in Section 5.

5. Thickness of Lüroth Sets

In this section, we establish a closed expression of thickness τ_{N_1, N_2} for Lüroth sets of the form L_{N_1, N_2} . The construction SCC from the previous section is ordered, so by (3.1) it suffices to find the thickness of the SCC for L_{N_1, N_2} . The following proposition serves as a foundation for most of the main results in this paper:

Proposition 5.1. *Let $N_1, N_2 \in \mathbb{N} \setminus \{1\}$. Suppose $N_1 < N_2$. Then the thickness of L_{N_1, N_2} is*

$$\begin{aligned} \tau_{N_1, N_2} &= \frac{N_2 - \langle N_2 \rangle}{\langle (N_2 - 1) - N_2 \rangle} \\ &= \frac{(N_2 - N_1)(N_2 - 2)(N_1 N_2 - N_1 - N_2 + 2)}{N_1^2(N_2^3 - 2N_2^2 + 2) - N_1(2N_2^3 - 5N_2^2 + N_2 + 4) - (N_2^2 - N_2)}. \end{aligned} \quad (5.1)$$

In other words, the supremum τ_{N_1, N_2} is realised at the left interval at the last step in the SCC. By the definition of SCC and since it is ordered, we have:

$$\tau_{N_1, N_2} = \min_{x \in \{0, 1, \dots, N_2 - N_1 - 1\}} \left\{ \min \left\{ \frac{N_1 + x + 1 - \langle N_2 \rangle}{\langle N_1 + x - N_1 + x + 1 \rangle}, \frac{N_1 + x - \langle N_1 + x \rangle}{\langle N_1 + x - N_1 + x + 1 \rangle} \right\} \right\}.$$

Define $f_0, f_1 : [0, +\infty) \rightarrow \mathbb{R}$ by, for any $x \geq 0$,

$$\begin{aligned} f_0(x) &= f_{N_1, N_2, 0}(x) := \frac{N_1 + x + 1 - \langle N_2 \rangle}{\langle N_1 + x - N_1 + x + 1 \rangle}, \\ f_1(x) &= f_{N_1, N_2, 1}(x) := \frac{N_1 + x - \langle N_1 + x \rangle}{\langle N_1 + x - N_1 + x + 1 \rangle}. \end{aligned}$$

To prove Proposition 5.1, it suffices to prove the following lemmas as together they conclude that

$$\tau_{N_1, N_2} = f_0(N_2 - N_1 - 1).$$

The first lemma takes care of the thickness of $L_{\leq 3}$, $L_{\leq 4}$ and $L_{\leq 5}$.

Lemma 5.2. *For $N_1 = 2$ and any $N_2 \in \{3, 4, 5\}$, (5.1) holds.*

The next four lemmas combine to take care of $L_{\leq k}$ for $k \geq 6$ and L_{N_1, N_2} for $N_1 \geq 3$.

Lemma 5.3. *For $N_1 = 2$ and any $N_2 \geq 6$,*

$$f_0(0) > f_0(N_2 - 3),$$

18 Maiken Gravgaard, Ying Wai Lee

and for any $x \in [1, N_2 - 4]$,

$$f_0(x) \geq f_1(x).$$

Lemma 5.4. For any $N_1 \geq 3$, $N_2 \geq N_1 + 2$ and $x \in [0, N_2 - N_1 - 2]$,

$$f_0(x) \geq f_1(x).$$

Lemma 5.5. For any $N_1, N_2 \in \mathbb{N} \setminus \{1\}$, if $N_1 < N_2$ then f_1 is strictly decreasing.

Lemma 5.6. For any $N_1, N_2 \in \mathbb{N} \setminus \{1\}$, if $N_1 < N_2$ then

$$f_0(N_2 - N_1 - 1) < f_1(N_2 - N_1 - 1).$$

Lemmas 5.3, 5.5, 5.6 work together to find the thickness for $L_{\leq k}$ for $k \geq 6$; Lemmas 5.4, 5.5, 5.6 work together to find the thickness of L_{N_1, N_2} for $N_1 \geq 3$. All of that combined proves Proposition 5.1.

Proof of Lemma 5.2. We calculate the exact values of τ_3 , τ_4 and τ_5 . Based on the SCC for $L_{\leq 3}$, $L_{\leq 4}$ and $L_{\leq 5}$, we have

$$\begin{aligned} \tau_3 &= \min \left\{ \frac{3}{\langle 2-3 \rangle}, \frac{2}{\langle 2-3 \rangle} \right\} &= \frac{1}{7} = \frac{3}{\langle 2-3 \rangle}; \\ \tau_4 &= \min \left\{ \frac{3}{\langle 2-3 \rangle}, \frac{2}{\langle 2-3 \rangle}, \frac{4}{\langle 3-4 \rangle}, \frac{3}{\langle 3-4 \rangle} \right\} &= \frac{4}{7} = \frac{4}{\langle 3-4 \rangle}; \\ \tau_5 &= \min \left\{ \frac{3}{\langle 2-3 \rangle}, \frac{2}{\langle 2-3 \rangle}, \frac{4}{\langle 3-4 \rangle}, \frac{5}{\langle 3-4 \rangle}, \frac{5}{\langle 4-5 \rangle}, \frac{4}{\langle 4-5 \rangle} \right\} &= \frac{9}{13} = \frac{5}{\langle 4-5 \rangle}. \end{aligned}$$

The result follows. $\square$

In Lemmas 5.3, 5.4 and 5.6, we are comparing f_0 and f_1 , so let us look at the numerators. We have, for any $x \geq 0$,

$$\begin{aligned} N_1 + x + 1 - \langle N_2 \rangle &= \frac{1}{N_1 + x + 1} + \frac{1}{(N_1 + x + 1)(N_1 + x)} N_1 - \langle N_2 \rangle \\ &= \frac{1}{N_1 + x + 1} \left(1 + \frac{N_1}{N_1 + x} \right) - \langle N_2 \rangle; \end{aligned}$$

and

$$\begin{aligned} N_1 + x - \langle N_1 + x \rangle &= \frac{1}{N_1 + x} + \frac{1}{(N_1 + x)(N_1 + x - 1)} N_1 - \frac{1}{N_1 + x} - \frac{1}{(N_1 + x)(N_1 + x - 1)} \langle N_2 \rangle \\ &= \frac{N_1}{(N_1 + x)(N_1 + x - 1)} - \frac{\langle N_2 \rangle}{(N_1 + x)(N_1 + x - 1)} \\ &= \frac{N_1 - \langle N_2 \rangle}{(N_1 + x)(N_1 + x - 1)}. \end{aligned}$$

Define $D : [0, +\infty) \rightarrow \mathbb{R}$ to be the difference of the numerators of f_0 and f_1 , that is for any $x \geq 0$,

$$D(x) := \frac{1}{N_1 + x + 1} \left(1 + \frac{N_1}{N_1 + x} \right) - \left\langle N_2 - \frac{N_1}{(N_1 + x)(N_1 + x - 1)} \right\rangle. \quad (5.2)$$

Proof of Lemma 5.3. Suppose $N_1 = 2$ and $N_2 \geq 6$. To prove $f_0(0) - f_0(N_2 - 3) > 0$, it suffices to show that

$$(3) - \langle N_2 \rangle (\langle (N_2 - 1) - N_2 \rangle) - (N_2) - \langle N_2 \rangle (\langle 2 - 3 \rangle) > 0$$

Writing out the full expression of the left-hand side above yields

$$\begin{aligned} & \left(\frac{1}{3} + \frac{1}{6} - \frac{N_2 - 1}{N_2^2 - N_2 - 1} \right) \left(\frac{1}{N_2 - 1} + \frac{1}{(N_2 - 1)(N_2 - 2)} \frac{N_2 - 1}{N_2^2 - N_2 - 1} - \frac{1}{N_2} - \frac{1}{N_2(N_2 - 1)} \right) \\ & - \left(\frac{1}{N_2} + \frac{1}{N_2(N_2 - 1)} - \frac{N_2 - 1}{N_2^2 - N_2 - 1} \right) \left(\frac{1}{2} + \frac{1}{2} \frac{N_2 - 1}{N_2^2 - N_2 - 1} - \frac{1}{3} - \frac{1}{6} \right) \\ & = \frac{N_2^2 - 3N_2 + 1}{2(N_2^2 - N_2 - 1)} \frac{1}{(N_2 - 2)(N_2^2 - N_2 - 1)} - \frac{N_2 - 2}{(N_2 - 1)(N_2^2 - N_2 - 1)} \frac{N_2 - 1}{2(N_2^2 - N_2 - 1)} \\ & = \frac{N_2 - 3}{2(N_2 - 2)(N_2^2 - N_2 - 1)^2} > 0, \end{aligned}$$

which is positive since we assumed $N_2 \geq 6$.

For the second claim we show that for any $1 \leq x \leq N_2 - 4$, $D(x) \geq 0$. To see this, we have

$$\begin{aligned} D(x) &= \frac{1}{N_1 + x + 1} \left(1 + \frac{N_1}{N_1 + x} \right) - \left\langle N_2 - \frac{N_1}{(N_1 + x)(N_1 + x - 1)} \right\rangle \\ &= \frac{1}{2 + x} - \left\langle N_2 - \frac{1 - \langle N_2 \rangle}{(2 + x)(1 + x)} \right\rangle \\ &= \frac{-x^2 \langle N_2 + x(1 - 3\langle N_2 \rangle) \rangle N_2}{(1 + x)(2 + x)}, \end{aligned}$$

where the numerator is non-negative if and only if

$$\frac{(1 - 3\langle N_2 \rangle) - \sqrt{(1 - 3\langle N_2 \rangle)^2 - 4(\langle N_2 \rangle)^2}}{2\langle N_2 \rangle} \leq x \leq \frac{(1 - 3\langle N_2 \rangle) + \sqrt{(1 - 3\langle N_2 \rangle)^2 - 4(\langle N_2 \rangle)^2}}{2\langle N_2 \rangle}. \quad (5.3)$$

Notice that for any $N_2 \geq 6$,

$$\begin{aligned} \Delta_{N_2,0} &:= 1 - \frac{(1 - 3\langle N_2 \rangle) - \sqrt{(1 - 3\langle N_2 \rangle)^2 - 4(\langle N_2 \rangle)^2}}{2\langle N_2 \rangle} \\ &= \frac{-4 + 6N_2 - N_2^2 + \sqrt{(-4 + 6N_2 - N_2^2)^2 + 4N_2^3 - 28N_2^2 + 40N_2 - 16}}{2(N_2 - 1)} \geq 0, \end{aligned}$$

20 *Maiken Gravggaard, Ying Wai Lee*

and

$$\begin{aligned}\Delta_{N_2,1} &:= \frac{(1 - 3\langle N_2 \rangle + \sqrt{(1 - 3\langle N_2 \rangle)^2 - 4(\langle N_2 \rangle)^2})}{2\langle N_2 \rangle} - (N_2 - 4) \\ &= \frac{-6 + 6N_2 - N_2^2 + \sqrt{(-6 + 6N_2 - N_2^2)^2 + 4N_2^3 - 32N_2^2 + 64N_2 - 36}}{2(N_2 - 1)} \geq 0.\end{aligned}$$

Thus, (5.3) holds for all $1 \leq x \leq N_2 - 4$, which also completes the proof. $\square$

Now we handle the cases when N_1 is large.

Proof of Lemma 5.4. Suppose $N_1 \geq 3$ and $N_2 - N_1 - 2 \geq 0$. For any $x \geq 0$,

$$D(x) = \frac{c_0 + c_1x + c_2x^2 + c_3x^3}{C_{N_1, N_2}(N_1 + x - 1)(N_1 + x)(N_1 + x + 1)},$$

where the coefficients c_0, c_1, c_2, c_3 and the constant $C_{N_1, N_2} > 0$ are in terms of N_1 and N_2 ,

$$\begin{aligned}c_0 &= -N_1^5 N_2 + N_1^5 + N_1^4 N_2^2 - 2N_1^4 - 2N_1^3 N_2^2 + 5N_1^3 N_2 - N_1^3 - N_1^2 N_2 + N_1^2 - N_1 N_2^2 \\ &\quad - 2N_1 N_2 + 4N_1 + 2N_2^2 - 3N_2 - 1, \\ c_1 &= -3N_1^4 N_2 + 3N_1^4 + 2N_1^3 N_2^2 + N_1^3 N_2 - 5N_1^3 - 3N_1^2 N_2^2 + 8N_1^2 N_2 - 2N_1^2 - N_1 N_2^2 \\ &\quad - N_1 N_2 + 3N_1 + N_2^2 - 3N_2 + 1, \\ c_2 &= -3N_1^3 N_2 + 3N_1^3 + N_1^2 N_2^2 + 2N_1^2 N_2 - 4N_1^2 - N_1 N_2^2 + 4N_1 N_2 - 2N_1 - N_2^2 + N_2 + 1, \\ c_3 &= -N_1^2 N_2 + N_1^2 + N_1 N_2 - N_1 + N_2 - 1 \\ &= -(N_1^2 - N_1 - 1)(N_2 - 1) < 0, \\ C_{N_1, N_2} &= (N_1^2 - N_1 - 1)(N_2^2 - N_2 - 1).\end{aligned}$$

Define $E : \mathbb{R} \rightarrow \mathbb{R}$ to be the cubic equation in the numerator of D , that is for any $x \in \mathbb{R}$,

$$E(x) := c_0 + c_1x + c_2x^2 + c_3x^3.$$

It now suffices to prove the following properties for E :

- (1) $c_3 < 0$; hence $\lim_{x \rightarrow -\infty} E(x) = +\infty$ and $\lim_{x \rightarrow +\infty} E(x) = -\infty$;
- (2) $E(-N_1) < 0$.
- (3) $E(0) > 0$;
- (4) $E(N_2 - N_1 - 2) > 0$;

Indeed the above implies that the cubic polynomial E has two negative roots and one positive root that is strictly greater than $N_2 - N_1 - 2$, thus for any $x \in [0, N_2 - N_1 - 1]$, $E(x) > 0$ and consequently $D(x) > 0$.

- (1) It is clear from the definition of the coefficient c_3 above.

(2) By taking $N_2 \geq N_1 + 2$ we have,

$$\begin{aligned}
 E(-N_1) &= c_0 - c_1 N_1 + c_2 N_1^2 - c_3 N_1^3 \\
 &= (-1 + 3N_1 - N_1^2) + (-3 + N_1 + N_1^2) N_2 + (2 - 2N_1) N_2^2 \\
 &\leq (-1 + 3N_1 - N_1^2) + (1 - N_1 - N_1^2) N_2 \\
 &\leq 1 + 2N_1 - 4N_1^2 - N_1^3 \\
 &< 0.
 \end{aligned}$$

(3) $E(0)$ is given by

$$\begin{aligned}
 E(0) &= c_0 \\
 &= -N_1^5 N_2 + N_1^5 + N_1^4 N_2^2 - 2N_1^4 - 2N_1^3 N_2^2 + 5N_1^3 N_2 - N_1^3 \\
 &\quad - N_1^2 N_2 + N_1^2 - N_1 N_2^2 - 2N_1 N_2 + 4N_1 + 2N_2^2 - 3N_2 - 1 \\
 &= N_2^2 (N_1^4 - 2N_1^3 - N_1 + 2) + N_2 (-N_1^5 + 5N_1^3 - N_1^2 - 2N_1 - 3) \\
 &\quad + (N_1^5 - 2N_1^4 - N_1^3 + N_1^2 + 4N_1 - 1).
 \end{aligned}$$

Suppose $N_1 \geq 3$ and $N_2 \geq N_1 + 2$. the constant term and the coefficients of N_2 and N_2^2 are positive, negative and positive respectively. Thus, we have

$$\begin{aligned}
 c_0 &\geq N_2^2 (N_1^4 - 2N_1^3 - N_1 + 2) + N_2 (-N_1^5 + 5N_1^3 - N_1^2 - 2N_1 - 3) \\
 &= N_2 (N_2 (N_1^4 - 2N_1^3 - N_1 + 2) + (-N_1^5 + 5N_1^3 - N_1^2 - 2N_1 - 3)) \\
 &\geq (N_1 + 2) (N_1^4 - 2N_1^3 - N_1 + 2) + (-N_1^5 + 5N_1^3 - N_1^2 - 2N_1 - 3) \\
 &= (N_1 + 1) (N_1^2 - 3N_1 + 1) \\
 &> 0.
 \end{aligned}$$

(4) Suppose $N_1 \geq 3$ and $N_2 \geq N_1 + 2 \geq 5$. Calculating $E(N_2 - N_1 - 2)$ we get

$$\begin{aligned}
 E(N_2 - N_1 - 2) &= c_0 + c_1(N_2 - N_1 - 2) + c_2(N_2 - N_1 - 2)^2 + c_3(N_2 - N_1 - 2)^3 \\
 &= (9 - 16N_2 + 8N_2^2 - N_2^3) \\
 &\quad + (13 - 12N_2 + 4N_2^2 - N_2^3)N_1 + (-11 + 14N_2 - 6N_2^2 + N_2^3)N_1^2
 \end{aligned}$$

Now we consider the polynomial below for $x \geq 3$ and $y \geq 5$

$$P(x, y) = a_y x^2 + b_y x + c_y,$$

where

$$a(y) = -11 + 14y - 6y^2 + y^3, \quad b(y) = 13 - 12y + 4y^2 - y^3, \quad c(y) = 9 - 16y + 8y^2 - y^3.$$

Note that a is increasing and that for $y \geq 5$,

$$a(y) = (y - 2)^3 + 2(y - 2) + 1 \geq 34.$$

Thus, for any fixed $y \geq 5$, $P(x, y)$ is quadratic in x and forms an upward parabola. The partial derivative of P with respect to x is given by

$$\frac{\partial P}{\partial x}(x, y) = 2a(y)x + b(y).$$

22 *Maiken Gravgaard, Ying Wai Lee*

Evaluating at $x = 3$, we see that

$$\begin{aligned}\frac{\partial P}{\partial x}(3, y) &= 5y^3 - 32y^2 + 72y - 53 \\ &= 5 \left(y - \frac{32}{15} \right)^3 + \frac{56}{15} \left(y - \frac{32}{15} \right) + \frac{2369}{675} > 0.\end{aligned}$$

This tells us that for fixed $y \geq 5$, $P(x, y)$ is increasing on $x \in [3, \infty)$. Hence, for any $x \geq 3$ and $y \geq 5$

$$\begin{aligned}P(x, y) &\geq P(3, y) \\ &= (y - 3)(5y^2 - 19y + 17) \\ &= (y - 3)(5(y - 5)^2 + 31(y - 5) + 47) \\ &> 0.\end{aligned}\quad \square$$

Proof of Lemma 5.5. Let $N_1, N_2 \in \mathbb{N} \setminus \{1\}$ with $N_1 < N_2$. To show that f_1 is decreasing, we write for any $x \geq 0$,

$$f_1(x) = \frac{\langle N_1 + x \rangle - \langle N_1 + x \rangle}{\langle N_1 + x - N_1 + x + 1 \rangle} = \frac{N_1 - \langle N_2 \rangle}{(1 - N_1) \frac{N_1 + x - 1}{N_1 + x + 1} + \langle N_2 \rangle},$$

it remains to notice that the following expression is strictly increasing on $[0, +\infty)$ with respect to x :

$$\frac{N_1 + x - 1}{N_1 + x + 1} = 1 - \frac{2}{N_1 + x + 1}.\quad \square$$

Proof of Lemma 5.6. It suffices to prove that $D(N_2 - N_1 - 1) \leq 0$. By $2 \leq N_1 < N_2$, we have

$$\begin{aligned}D(N_2 - N_1 - 1) &= \frac{1}{N_2} \left(1 + \frac{N_1}{N_2 - 1} \right) - \langle N_2 - \frac{N_1 - \langle N_2 \rangle}{(N_2 - 1)(N_2 - 2)} \rangle \\ &= \frac{4N_1 - 2N_1^2 - 4N_2 + 2N_1^2N_2 + 2N_2^2 - 2N_1N_2^2}{(N_1^2 - N_1 - 1)N_2(N_2 - 1)(N_2 - 2)(N_2^2 - N_2 - 1)} \\ &< \frac{(N_1 - N_2)(N_1N_2 - N_1 - N_2)}{(N_1^2 - N_1 - 1)N_2(N_2 - 1)(N_2 - 2)(N_2^2 - N_2 - 1)} \\ &< 0.\end{aligned}$$

Hence, $f_0(N_2 - N_1 - 1) \leq f_1(N_2 - N_1 - 1)$, which completes the proof of Lemma 5.6, which in turn completes the proof of Proposition 5.1. $\square$

6. Proofs on Congruence

In this section, we prove the main results on whether sums of Lüroth sets are congruent to $\mathbb{R} \bmod 1$, namely Theorems 2.1, 2.3, Corollary 2.4, and Theorem 2.5.

6.1. Proof of Theorem 2.1

In Section 4, we construct the set $L_{\leq 3}$ under the SCC and we have that,

$$g_3 = \frac{\langle 2-3 \rangle}{2) - \langle 3} = \frac{7/10 - 1/2}{1 - 2/5} = \frac{1}{3};$$

$$h_3 = \min \left\{ \frac{3) - \langle 3}{2) - \langle 3}, \frac{2) - \langle 2}{2) - \langle 3} \right\} = \min \left\{ \frac{1/2 - 2/5}{1 - 2/5}, \frac{1 - 7/10}{1 - 2/5} \right\} = \frac{1}{6}.$$

Similarly calculating g and h for $L_{\leq 4}$ and $L_{\leq 5}$ yields that,

$$g_4 = \frac{\langle 3-4 \rangle}{3) - \langle 4} = \frac{1}{5}; \quad h_4 = \frac{\langle 4-4 \rangle}{3) - \langle 4} = \frac{4}{15};$$

$$g_5 = \frac{\langle 4-5 \rangle}{4) - \langle 5} = \frac{1}{7}; \quad h_5 = \frac{\langle 5-5 \rangle}{4) - \langle 5} = \frac{9}{28}.$$

To complete the proof, we apply Propositions 3.2 and 3.3.

To prove that $L_{\leq 3} + L_{\leq 5} \equiv \mathbb{R} \pmod{1}$, we notice that $|I_3| = 3/5$ and $|I_5| = 15/19$. We verify the three conditions in Proposition 3.2:

$$g_3 g_5 = \frac{1}{21} \leq \frac{3}{56} = h_3 h_5, \quad g_3 |I_3| = \frac{1}{5} \leq \frac{15}{19} = |I_5|, \quad g_5 |I_5| = \frac{15}{133} \leq \frac{3}{5} = |I_3|.$$

By Proposition 3.2, we see that $L_{\leq 3} + L_{\leq 5}$ is an interval of length $132/95 \geq 1$, as

$$L_{\leq 3} + L_{\leq 5} = I_3 + I_5 = \left[\frac{2}{5}, 1 \right] + \left[\frac{4}{19}, 1 \right] = \left[\frac{58}{95}, 2 \right].$$

To prove that $L_{\leq 4} + L_{\leq 4} \equiv \mathbb{R} \pmod{1}$, we notice that $|I_4| = 8/11$. We verify the conditions in Proposition 3.2:

$$g_4 g_4 = \frac{1}{25} \leq \frac{16}{225} = h_4 h_4, \quad g_4 |I_4| = \frac{8}{55} \leq \frac{8}{11} = |I_4|.$$

By Proposition 3.2, we see that $L_{\leq 4} + L_{\leq 4}$ is an interval of length $16/11 \geq 1$, as

$$L_{\leq 4} + L_{\leq 4} = I_4 + I_4 = \left[\frac{3}{11}, 1 \right] + \left[\frac{3}{11}, 1 \right] = \left[\frac{6}{11}, 2 \right].$$

To prove that $L_{\leq 3} + L_{\leq 3} + L_{\leq 3} \equiv \mathbb{R} \pmod{1}$, we apply Proposition 3.3 with g_3 and h_3 obtained above. The first condition is clearly satisfied since all the I_A 's are the same and h is always not greater than 1. For the second condition, we have

$$g_3 + h_3 = 3h_3 = \frac{1}{2}.$$

Proposition 3.3 then tells us that $L_{\leq 3} + L_{\leq 3} + L_{\leq 3}$ is an interval of length $9/5 \geq 1$, namely

$$L_{\leq 3} + L_{\leq 3} + L_{\leq 3} = I_3 + I_3 + I_3 = \left[\frac{2}{5}, 1 \right] + \left[\frac{2}{5}, 1 \right] + \left[\frac{2}{5}, 1 \right] = \left[\frac{6}{5}, 3 \right].$$

Next, we prove $L_{\leq 3} + L_{\leq 3} \not\equiv \mathbb{R} \pmod{1}$. Notice that at Level 1 of SCC for $L_{\leq 3}$, we have

$$L_{\leq 3} \subset [\langle 3, 3 \rangle] \cup [\langle 2, 2 \rangle] = \left[\frac{2}{5}, \frac{1}{2} \right] \cup \left[\frac{7}{10}, 1 \right].$$

24 *Maiken Gravggaard, Ying Wai Lee*

Thus, we have

$$L_{\leq 3} + L_{\leq 3} \subset \left[\frac{4}{5}, 1 \right] \cup \left[\frac{11}{10}, 2 \right];$$

hence,

$$L_{\leq 3} + L_{\leq 3} + \mathbb{Z} \subset \left[\frac{1}{10}, 1 \right] + \mathbb{Z}.$$

In particular, $L_{\leq 3} + L_{\leq 3} \not\equiv \mathbb{R} \pmod{1}$.

If we go down further levels in SCC we reveal more gaps in $L_{\leq 3} + L_{\leq 3}$. Note that

$$\begin{aligned} [\dot{3}] &= \frac{2}{5}, & [3, 3, \dot{2}] &= \frac{5}{12}, & [3, 2, \dot{3}] &= \frac{9}{20}, & [3, \dot{2}] &= \frac{1}{2}; \\ [2, \dot{3}] &= \frac{7}{10}, & [2, 3, \dot{2}] &= \frac{3}{4}, & [2, 2, \dot{3}] &= \frac{17}{20}, & [\dot{2}] &= 1. \end{aligned}$$

Thus at Level 2 of the SCC for $L_{\leq 3}$, we have

$$L_{\leq 3} \subseteq \left[\frac{2}{5}, \frac{5}{12} \right] \cup \left[\frac{9}{20}, \frac{1}{2} \right] \cup \left[\frac{7}{10}, \frac{3}{4} \right] \cup \left[\frac{17}{20}, 1 \right],$$

and hence,

$$L_{\leq 3} + L_{\leq 3} \subseteq \left[\frac{4}{5}, \frac{5}{6} \right] \cup \left[\frac{17}{20}, 1 \right] \cup \left[\frac{11}{10}, \frac{3}{2} \right] \cup \left[\frac{31}{20}, 2 \right].$$

This in turn shows that

$$\begin{aligned} L_{\leq 3} + L_{\leq 3} + \mathbb{Z} &\subseteq \left[\frac{4}{5}, \frac{5}{6} \right] \cup \left[\frac{17}{20}, 1 \right] \cup \left[\frac{1}{10}, \frac{1}{2} \right] \cup \left[\frac{11}{20}, 1 \right] + \mathbb{Z} \\ &= \left[\frac{1}{10}, \frac{1}{2} \right] \cup \left[\frac{11}{20}, 1 \right] + \mathbb{Z} \end{aligned}$$

hence revealing another gap in $L_{\leq 3} + L_{\leq 3}$.

This also provides an upper bound for the lengths of the potential intervals contained in $L_{\leq 3} + L_{\leq 3}$, but does not indicate whether it contains an interval or not. Figure 7 shows where these gaps occur in $L_{\leq 3} + L_{\leq 3}$, as well as indicating where the gaps lie in $[0, 1]$.

Finally, we prove that $L_{\leq 3} + L_{\leq 4}$ is not an interval. Note that Level 3 of the SCC for $L_{\leq 3}$ is given by:

$$\left[\frac{2}{5}, \frac{29}{72} \right] \cup \left[\frac{49}{120}, \frac{5}{12} \right] \cup \left[\frac{9}{20}, \frac{11}{24} \right] \cup \left[\frac{19}{40}, \frac{1}{2} \right] \cup \left[\frac{7}{10}, \frac{17}{24} \right] \cup \left[\frac{29}{40}, \frac{3}{4} \right] \cup \left[\frac{17}{20}, \frac{7}{8} \right] \cup \left[\frac{37}{40}, 1 \right].$$

Level 4 of the SCC for $L_{\leq 4}$ is given by:

$$\begin{aligned} &\left[\frac{3}{11}, \frac{5}{18} \right] \cup \left[\frac{223}{792}, \frac{7}{24} \right] \cup \left[\frac{10}{33}, \frac{5}{16} \right] \cup \left[\frac{7}{22}, \frac{1}{3} \right] \cup \left[\frac{25}{66}, \frac{7}{18} \right] \cup \left[\frac{157}{396}, \frac{5}{12} \right] \cup \left[\frac{29}{66}, \frac{11}{24} \right] \cup \left[\frac{31}{66}, \frac{1}{2} \right] \cup \\ &\left[\frac{7}{11}, \frac{31}{48} \right] \cup \left[\frac{43}{66}, \frac{2}{3} \right] \cup \left[\frac{91}{132}, \frac{17}{24} \right] \cup \left[\frac{95}{132}, \frac{3}{4} \right] \cup \left[\frac{9}{11}, \frac{5}{6} \right] \cup \left[\frac{223}{264}, \frac{7}{8} \right] \cup \left[\frac{10}{11}, \frac{15}{16} \right] \cup \left[\frac{21}{22}, 1 \right]. \end{aligned}$$

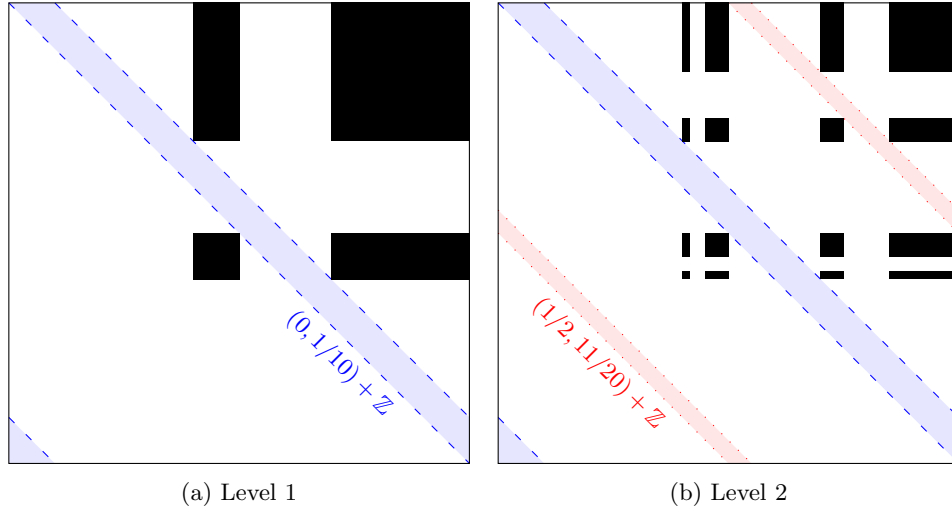


Fig. 7: Different Levels of $L_{\leq 3}$ under SCC product with itself on $[0, 1]^2$

Consequently, we get

$$L_{\leq 3} + L_{\leq 4} \subseteq \left[\frac{37}{55}, \frac{49}{72} \right] \cup \left[\frac{899}{1320}, 2 \right],$$

and for each interval, there is an element belonging to $L_{\leq 3} + L_{\leq 4}$, namely

$$[\dot{3}] + [\dot{4}] = \frac{2}{5} + \frac{3}{11} = \frac{37}{55} \in L_{\leq 3} + L_{\leq 4},$$

and

$$[\dot{2}] + [\dot{2}] = 1 + 1 = 2 \in L_{\leq 3} + L_{\leq 4}.$$

Figure 8 in appendix illustrates this gap in $L_{\leq 3} + L_{\leq 4}$.

6.2. Proof of Theorem 2.3

The following result is useful to prove Theorem 2.3. It basically says that $L_{\geq k}$ can be estimated arbitrarily well by $L_{k,N}$ for sufficiently large N .

Proposition 6.1. *For any $k \in \mathbb{N} \setminus \{1\}$ and $\varepsilon > 0$, there exists $N_{k,\varepsilon} \in \mathbb{N}$ such that*

$$\begin{aligned} \gamma_{k,N_{k,\varepsilon}} &> \frac{k-1}{k(k-1)-1} - \varepsilon, \\ |I_{k,N_{k,\varepsilon}}| &> \frac{k-1}{k(k-1)-1} - \varepsilon, \\ |O_{k,N_{k,\varepsilon}}| &< \frac{k-2}{k(k^2-2)-1} + \varepsilon. \end{aligned}$$

26 *Maiken Gravgard, Ying Wai Lee*

Proof. Pick any $k \in \mathbb{N} \setminus \{1\}$ and $\varepsilon > 0$. Define

$$N_{k,\varepsilon} := \max \left\{ k, \frac{1}{\varepsilon} \left(\frac{k-1}{k(k-1)-1} + 1 \right) + 1 \right\} + 1 > k.$$

Then for $N_1 = k$ and $N_2 = N = N_{k,\varepsilon}$, using SCC we have,

$$\begin{aligned} |I_{k,N}| &= \langle k \rangle - \langle N \rangle \\ &= \frac{k-1}{k(k-1)-1} - \frac{N-1}{N(N-1)-1} \\ &\geq \frac{k-1}{k(k-1)-1} - \frac{1}{N-1} \\ &> \frac{k-1}{k(k-1)-1} - \varepsilon; \end{aligned}$$

and

$$\begin{aligned} |O_{k,N}| &= \langle k - k + 1 \rangle \\ &= \left(\frac{1}{k} + \frac{1}{k(k-1)} \frac{N-1}{N(N-1)-1} \right) - \left(\frac{1}{k+1} + \frac{1}{(k+1)k} \frac{k-1}{k(k-1)-1} \right) \\ &= \left(\frac{1}{k} - \frac{1}{k+1} - \frac{1}{(k+1)k} \frac{k-1}{k(k-1)-1} \right) + \frac{1}{k(k-1)} \frac{N-1}{N(N-1)-1} \\ &\leq \left(\frac{1}{k} - \frac{1}{k+1} - \frac{1}{(k+1)k} \frac{k-1}{k(k-1)-1} \right) + \frac{1}{N-1} \\ &< \frac{k-2}{k(k^2-2)-1} + \varepsilon. \end{aligned}$$

By applying Proposition 5.1, we get the thickness:

$$\tau_{k,N} = \frac{N\rangle - \langle N}{\langle (N-1) - N \rangle}.$$

The numerator is given by,

$$\begin{aligned} N\rangle - \langle N &= \left(\frac{1}{N} + \frac{1}{N(N-1)} \frac{k-1}{k(k-1)-1} \right) - \frac{N-1}{N(N-1)-1} \\ &= \frac{1}{N(N-1)} \left(\frac{k-1}{k(k-1)-1} - \frac{N-1}{N(N-1)-1} \right), \end{aligned}$$

and the denominator is given by,

$$\begin{aligned} \langle (N-1) - N \rangle &= \left(\frac{1}{N-1} + \frac{1}{(N-2)} \frac{1}{N(N-1)-1} \right) - \left(\frac{1}{N} + \frac{1}{N(N-1)} \frac{k-1}{k(k-1)-1} \right) \\ &= \frac{1}{N(N-1)} \left(\frac{N^3 - 2N^2 + 2}{(N-2)(N^2 - N - 1)} - \frac{k-1}{k(k-1)-1} \right). \end{aligned}$$

Hence we have

$$\tau_{k,N} = \left(\frac{k-1}{k(k-1)-1} - \frac{N-1}{N(N-1)-1} \right) \Bigg/ \left(\frac{N^3 - 2N^2 + 2}{(N-2)(N^2 - N - 1)} - \frac{k-1}{k(k-1)-1} \right),$$

Lastly we can calculate γ ,

$$\begin{aligned}
\gamma_{k,N} &= \frac{1}{1 + 1/\tau_{k,N}} \\
&= \left(\frac{k-1}{k(k-1)-1} - \frac{N-1}{N(N-1)-1} \right) \bigg/ \left(\frac{N^3 - 2N^2 + 2}{(N-2)(N^2 - N - 1)} - \frac{N-1}{N(N-1)-1} \right) \\
&= \left(\frac{k-1}{k(k-1)-1} - \frac{N-1}{N(N-1)-1} \right) \frac{N^3 - 3N^2 + N + 2}{N^3 - 3N^2 + 3N} \\
&> \left(\frac{k-1}{k(k-1)-1} - \frac{1}{N-1} \right) \left(1 - \frac{1}{N-1} \right) \\
&> \frac{k-1}{k(k-1)-1} - \frac{1}{N-1} \left(\frac{k-1}{k(k-1)-1} + 1 \right) \\
&> \frac{k-1}{k(k-1)-1} - \varepsilon,
\end{aligned}$$

which completes the proof. $\square$

Proof of Theorem 2.3. It suffices to prove that there exists $\varepsilon > 0$ such that

$$L_{k_1, N_{k_1, \varepsilon}} + L_{k_2, N_{k_2, \varepsilon}} + \cdots + L_{k_n, N_{k_n, \varepsilon}} \equiv \mathbb{R} \pmod{1}.$$

We first require

$$0 < \varepsilon < \frac{1}{n} \left(\sum_{j=1}^n \frac{k_j - 1}{k_j(k_j - 1) - 1} - 1 \right).$$

then we get, by Proposition 6.1, that

$$S_\gamma = \sum_{j=1}^n \gamma_{k_j, N_{k_j, \varepsilon}} > 1, \quad \sum_{j=1}^n |I_{k_j, N_{k_j, \varepsilon}}| = \sum_{j=1}^n \text{diam } L_{k_j, N_{k_j, \varepsilon}} > 1.$$

It remains to verify the conditions (3.2) and (3.3), where $C_{\mathcal{A}_i} = L_{k_{n+1-i}, N_{k_{n+1-i}, \varepsilon}}$.

For (3.2), for any $i < j$,

$$\left| I_{k_i, N_{k_i, \varepsilon}} \right| > \frac{k_i - 1}{k_i(k_i - 1) - 1} - \varepsilon > \frac{k_j - 2}{k_j(k_j^2 - 2) - 1} + \varepsilon > \left| O_{k_j, N_{k_j, \varepsilon}} \right|,$$

where we also assume

$$0 < \varepsilon < \frac{1}{2} \min_{i < j} \left\{ \frac{k_i - 1}{k_i(k_i - 1) - 1} - \frac{k_j - 2}{k_j(k_j^2 - 2) - 1} \right\}.$$

For (3.3), we need (2.2), for any $j = 2, \dots, n$,

$$\begin{aligned}
\left| I_{k_j, N_{k_j, \varepsilon}} \right| &> \frac{k_j - 1}{k_j(k_j - 1) - 1} - \varepsilon \\
&> \frac{1}{k_j} \geq \frac{1}{k_{j-1}^2 + 2k_{j-1} + 2} \\
&> \frac{k_{j-1} - 2}{k_{j-1}(k_{j-1}^2 - 2) - 1} + \varepsilon > \left| O_{k_{j-1}, N_{k_{j-1}, \varepsilon}} \right|,
\end{aligned}$$

28 *Maiken Gravgaard, Ying Wai Lee*

here we are furthermore requiring that

$$0 < \varepsilon < \min_{j \geq 2} \left\{ \frac{k_j - 1}{k_j(k_j - 1) - 1} - \frac{1}{k_j}, \frac{1}{k_{j-1}^2 + 2k_{j-1} + 2} - \frac{k_{j-1} - 2}{k_{j-1}(k_{j-1}^2 - 2) - 1} \right\}.$$

The result then follows by applying Proposition 3.4. $\square$

6.3. Proof of Corollary 2.4

One can observe that Corollary 2.4 follows directly from Theorem 2.3, by taking $k_j = k$ for each $j \in \{1, \dots, n\}$. Alternatively, it is possible to prove it directly, by applying Proposition 3.4. For $k = 3$, it is even possible to prove it by applying Proposition 3.3.

By Proposition 5.1 and the definition of γ , we see that for $L_{3,16}$,

$$\gamma_{3,16} = \frac{2821}{8440} > \frac{1}{3}.$$

By Proposition 3.4 since $3\gamma_{3,16} > 1$, we have

$$3L_{3,16} = 3I_{3,16} = 3 \left[\frac{15}{239}, \frac{2}{5} \right] = \left[\frac{45}{239}, \frac{6}{5} \right],$$

which is of length $1209/1195 > 1$. Therefore,

$$3L_{3,16} \equiv \mathbb{R} \pmod{1},$$

and consequently $3L_{\geq 3} \equiv \mathbb{R} \pmod{1}$. Similarly,

$$4L_{4,45} \equiv \mathbb{R} \pmod{1}, \quad 5L_{5,96} \equiv \mathbb{R} \pmod{1}, \quad 6L_{6,175} \equiv \mathbb{R} \pmod{1},$$

and so on. In general, for any $k \in \mathbb{N} \setminus \{1\}$, it can be verified that,

$$kL_{k,k^3} \equiv \mathbb{R} \pmod{1}.$$

By the formula for τ_{k,k^3} derived in Proposition 5.1, we have the value of $k\gamma_{k,k^3}$ being at least 1:

$$\begin{aligned} k\gamma_{k,k^3} &= \frac{k(k-1)(k+1)(k^3-2)(k^4-k^3-k+2)}{k^2(k^2-k-1)(k^6-3k^3+3)} \\ &= 1 + \frac{k^6+k^5-5k^3-k^2+k+4}{k^9-k^8-k^7-3k^6+3k^5+3k^4+3k^3-3k^2-3k} \\ &\geq 1. \end{aligned}$$

Lastly for $k = 3$, we can apply Proposition 3.3, by considering an unordered construction of $L_{3,26}$. For the unordered construction, given in Figure 9 in appendix, we have

$$g_{3,26} = \frac{\langle 25 - 26 \rangle}{25} - \langle 26 \rangle = \frac{24989}{54314}, \quad h_{3,26} = \frac{6}{\langle 5 - 6 \rangle} - \langle 6 \rangle = \frac{391}{1689},$$

where the g - and h - values are obtained from Level 6 and 4 respectively. Note that

$$g_{3,26} + h_{3,26} = \frac{63443195}{91736346} \leq \frac{391}{563} = 3h_{3,26}.$$

By Proposition 3.3, we see that

$$3L_{3,26} = 3I_{3,26} = 3 \left[\frac{25}{649}, \frac{2}{5} \right] = \left[\frac{75}{649}, \frac{6}{5} \right],$$

which is of length $3519/3245 > 1$. Hence,

$$3L_{3,26} \equiv \mathbb{R} \pmod{1},$$

and consequently $3L_{\geq 3} \equiv \mathbb{R} \pmod{1}$.

6.4. Proof of Theorem 2.5

Let $k \in \mathbb{N} \setminus \{1\}$. By Proposition 5.1, the thickness and the γ -value of $L_{\leq k+2}$ are

$$\tau_{k+2} = \frac{k^2}{k+1}, \quad \gamma_{k+2} = \frac{1}{1 + 1/\tau_{k+2}} = \frac{k^2}{k^2 + k + 1}.$$

Let

$$\varepsilon = \min \left\{ \frac{2k}{k^4 - k^2 - 2k - 1}, \frac{2k^2}{k^4 + 2k^3 - 3k^2 - 4k - 1}, \frac{1}{2} \right\} > 0.$$

By Proposition 6.1, there exists $N_{k,\varepsilon} \in \mathbb{N} \setminus \{1\}$ such that

$$\begin{aligned} \gamma_{k,N_{k,\varepsilon}} &> \frac{k-1}{k(k-1)-1} - \varepsilon, \\ |I_{k,N_{k,\varepsilon}}| &> \frac{k-1}{k(k-1)-1} - \varepsilon, \\ |O_{k,N_{k,\varepsilon}}| &< \frac{k-2}{k(k^2-2)-1} + \varepsilon. \end{aligned}$$

Now we observe, that

$$\gamma_{k+2} + \gamma_{k,N_{k,\varepsilon}} > \frac{k^4 - k^2 - 1}{k^4 - k^2 - 2k - 1} - \varepsilon \geq 1,$$

and

$$|I_{k+2}| = \frac{k(k+2)}{k^2 + 3k + 1} \geq \frac{k-2}{k(k^2-2)-1} + \frac{1}{2} \geq |O_{k,N_{k,\varepsilon}}|.$$

By applying Proposition 3.4 we get,

$$L_{\leq k+2} + L_{k,N_{k,\varepsilon}} = I_{k+2} + I_{k,N_{k,\varepsilon}}.$$

On the other hand,

$$\begin{aligned} |I_{k+2}| + |I_{k,N_{k,\varepsilon}}| &> \frac{k(k+2)}{k^2 + 3k + 1} + \frac{k-1}{k(k-1)-1} - \varepsilon \\ &= \frac{k^4 + 2k^3 - k^2 - 4k - 1}{k^4 + 2k^3 - 3k^2 - 4k - 1} - \varepsilon \\ &\geq 1. \end{aligned}$$

Hence

$$L_{\leq k+2} + L_{k,N_{k,\varepsilon}} \equiv \mathbb{R} \pmod{1},$$

and consequently $L_{\leq k+2} + L_{\geq k} \equiv \mathbb{R} \pmod{1}$.

7. Proofs on Dimension

In this section, we prove the main results on the Hausdorff dimensions of Lüroth sets and sums of Lüroth sets, namely Theorems 2.9, 2.10, and 2.11.

7.1. Proof of Theorem 2.9

In the degenerate case, that is, at least one of $\mathcal{A}$ and $\mathcal{B}$ is a singleton, the result holds trivially. Indeed, without loss of generality, there exists $a \in \mathbb{N} \setminus \{1\}$ such that $\mathcal{A} = \{a\}$. Then we have,

$$L_{\mathcal{A}} = \left\{ \frac{a-1}{(a-1)a-1} \right\},$$

and consequently $\dim L_{\mathcal{A}} = 0$. Hence, $\dim(L_{\mathcal{A}} + L_{\mathcal{B}}) = \dim L_{\mathcal{B}} = \dim L_{\mathcal{A}} + \dim L_{\mathcal{B}} \leq 1$.

Suppose that both $\mathcal{A}$ and $\mathcal{B}$ are finite and not singletons. We aim to apply the result of Peres–Shmerkin [21, Theorem 2] on the Hausdorff dimension of the sumset of two iterative function systems (IFS).

Proposition 7.1 (Peres–Shmerkin [21, Theorem 2], 2009). *Let K and K' be the attractors of the IFS's $\{r_i x + t_i\}_{i=1}^n$ and $\{r'_i x + t'_i\}_{i=1}^{n'}$ respectively. If there exist j, j' such that $\log |r_j| / \log |r'_{j'}|$ is irrational, then*

$$\dim(K + K') = \min\{1, \dim K + \dim K'\}.$$

Since $\mathcal{A}$ and $\mathcal{B}$ are finite, $L_{\mathcal{A}}$ and $L_{\mathcal{B}}$ can be expressed as IFS's:

$$L_{\mathcal{A}} = \bigcup_{d \in \mathcal{A}} f_d(L_{\mathcal{A}}), \quad L_{\mathcal{B}} = \bigcup_{d \in \mathcal{B}} f_d(L_{\mathcal{B}}),$$

where for any $d \in \mathbb{N} \setminus \{1\}$ and $x \in [0, 1]$,

$$f_d(x) := \frac{1}{d(d-1)}x + \frac{1}{d}.$$

Since both $\mathcal{A}$ and $\mathcal{B}$ are not singletons, there exist $a \in \mathcal{A}$ and $b \in \mathcal{B}$ such that $a \neq b$. By a previous result by the second author,¹⁸ we have that

$$\frac{\log(a(a-1))}{\log(b(b-1))}$$

is irrational. By applying Proposition 7.1, the proof of Theorem 2.9 is completed.

7.2. Proof of Theorem 2.10

The proof applies a similar technique to that of Good. It suffices to prove the result for $k \geq 16$, since for any $k \leq 15$ we have $L_{\geq 16} \subseteq L_{\geq k}$, so

$$\dim L_{\geq k} \geq \dim L_{\geq 16} > \frac{1}{2} + \frac{1}{2 \log 16} = \frac{1}{2} + \frac{1}{2 \log \max\{16, k\}}.$$

Let $k \in \mathbb{N}$. Suppose $k \geq 16$. By the Intermediate Value Theorem, there exists $x_k > 0$ such that $k^{-x_k} = x_k$. Let $\varepsilon > 0$ and pick

$$N := \max \left\{ k + 1, \left\lceil (x_k (k^\varepsilon - 1))^{-1/(x_k - \varepsilon)} \right\rceil \right\}.$$

Define $F = F_{k,\varepsilon} : \mathbb{R} \rightarrow \mathbb{R}$ by, for any $x \in \mathbb{R}$,

$$F(x) := k^{-x} - N^{-x} - x.$$

Note that F is continuous on $\mathbb{R}$, $F(x_k) < 0$, and

$$\begin{aligned} F(x_k - \varepsilon) &= k^{-x_k + \varepsilon} - N^{-x_k + \varepsilon} - (x_k - \varepsilon) \\ &= x_k (k^\varepsilon - 1) + \varepsilon - N^{-x_k + \varepsilon} \geq \varepsilon > 0. \end{aligned}$$

By the Intermediate Value Theorem, there exists $x_{k,\varepsilon} \in (x_k - \varepsilon, x_k)$ such that $F(x_{k,\varepsilon}) = 0$. Thus, by writing $s_k = (x_k + 1)/2 \in \mathbb{R}$, there exists $s_{k,\varepsilon} > s_k - \varepsilon/2$ such that

$$\frac{1}{2s_{k,\varepsilon} - 1} \left(k^{-(2s_{k,\varepsilon} - 1)} - N^{-(2s_{k,\varepsilon} - 1)} \right) = 1,$$

and $k^{-(2s_k - 1)} = 2s_k - 1$. Thus, for any $\varepsilon > 0$,

$$\sum_{d=k}^N \left(\frac{1}{d(d-1)} \right)^{s_{k,\varepsilon}} > \int_k^N \frac{dx}{x^{2s_{k,\varepsilon}}} = 1;$$

and

$$\dim L_{\geq k} \geq \dim L_{k,N} > s_{k,\varepsilon} > s_k - \varepsilon/2.$$

Since this $\varepsilon > 0$ is arbitrary, we have

$$\dim L_{\geq k} \geq s_k.$$

Since $k \geq 16 > e^e$, the equality $k^{-(2s_k - 1)} = 2s_k - 1$ implies that

$$\begin{aligned} 2s_k - 1 &> \frac{1}{\log k} \\ s_k &> \frac{1}{2} + \frac{1}{2 \log k}. \end{aligned}$$

7.3. Proof of Theorem 2.11

It suffices to prove the following result.

Proposition 7.2. *For any $k \in \mathbb{N} \setminus \{1\}$, $\dim L_{k,3k} > 1/2$.*

Proof. By, for example [18, Theorem 2], the Hausdorff dimension of $L_{k,3k}$ is the unique solution $s = s_{k,3k} \in [0, 1]$ to

$$\sum_{d=k}^{3k} \left(\frac{1}{d(d-1)} \right)^{s_{k,3k}} = 1.$$

32 *Maiken Gravggaard, Ying Wai Lee*

Notice that

$$\sum_{d=k}^{3k} \left(\frac{1}{d(d-1)} \right)^1 = \frac{2k+1}{3k(k-1)} < 1.$$

and

$$\sum_{d=k}^{3k} \left(\frac{1}{d(d-1)} \right)^{1/2} > \int_k^{3k} \frac{dx}{x} = \log 3 > 1.$$

By the Intermediate Value Theorem, $s_{k,3k} > 1/2$. □

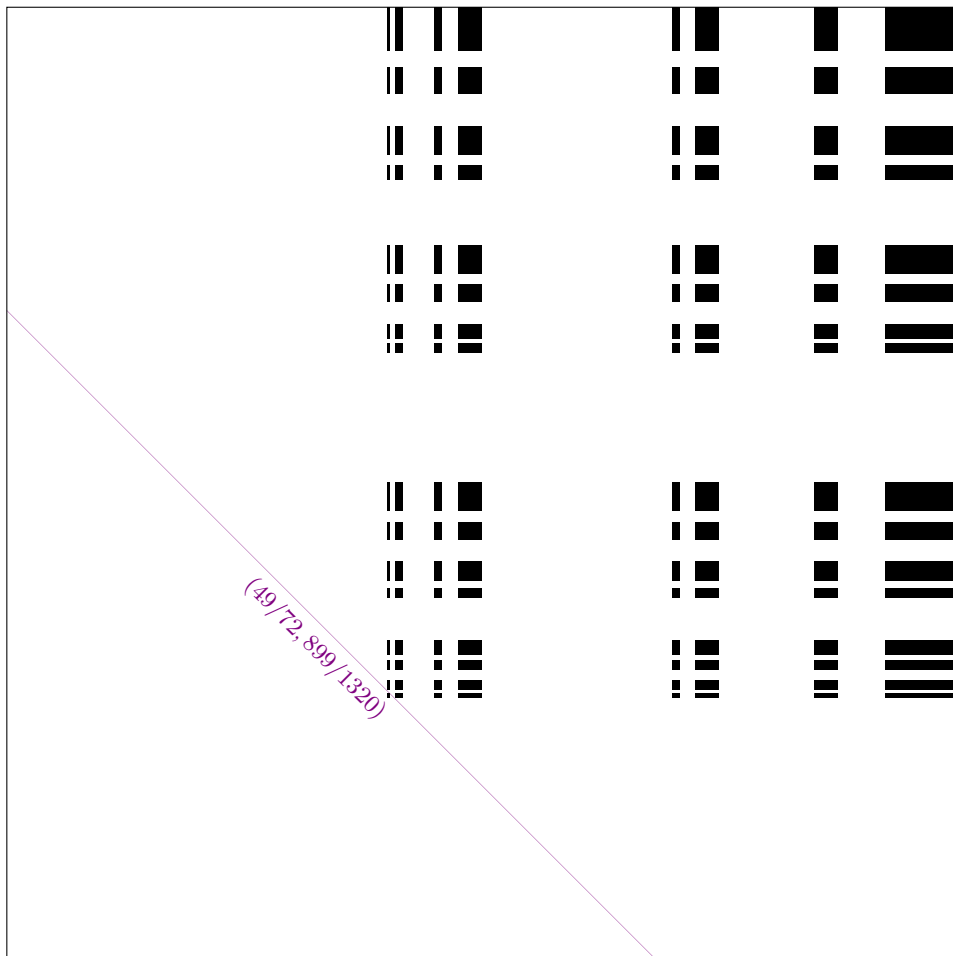
Proof of Theorem 2.11. By Theorem 2.9 and Proposition 7.2, for any $k \in \mathbb{N} \setminus \{1\}$,

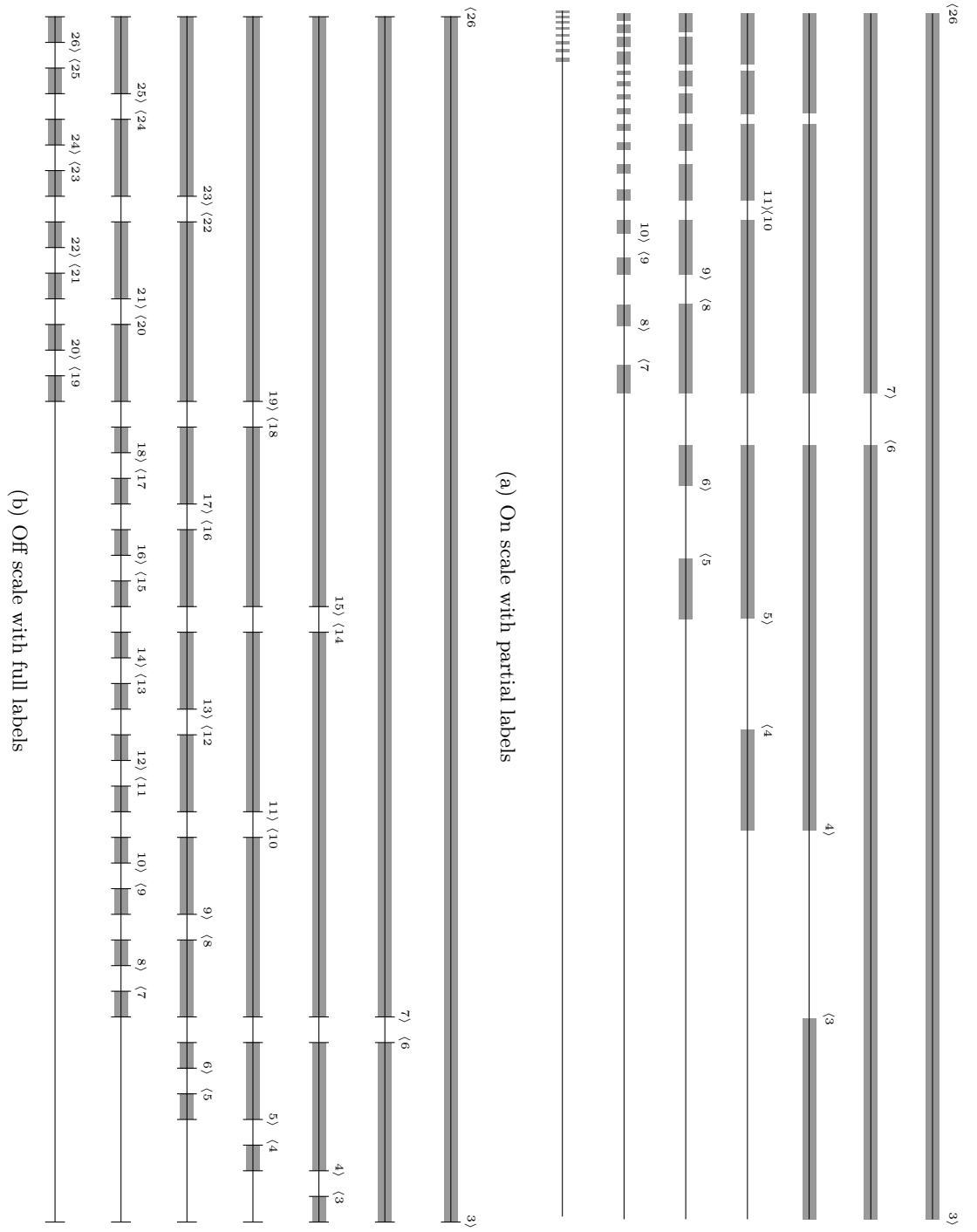
$$\begin{aligned} \dim(L_{\geq k} + L_{\geq k}) &\geq \dim(L_{k,3k} + L_{k,3k}) \\ &= \min\{1, \dim L_{k,3k} + \dim L_{k,3k}\} \\ &= 1. \end{aligned} \quad \square$$

Acknowledgements

The first author would like to express sincere gratitude to Professor Sanju Velani and the Department of Mathematics at the University of York for hosting the research environment change. Both authors also express their appreciation to Dr. Raphael Bennett-Tennenhaus for providing valuable and insightful feedback on the manuscript, and to the anonymous referees for their comments, which improved the clarity of the paper.

Appendix A.

Fig. 8: SCC $L_{\leq 3}$ Level 3 product with SCC $L_{\leq 4}$ Level 4 on $[0, 1]^2$

Fig. 9: An Unordered Construction of $L_{3,26}$

References

- [1] S. Astels. Cantor sets and numbers with restricted partial quotients. *Transactions of the American Mathematical Society*, 352(1):133–170, 2000.
- [2] S. Astels. Sums of numbers with small partial quotients ii. *Journal of Number Theory*, 91(2):187–205, 2001.
- [3] S. Astels. Sums of numbers with small partial quotients. *Proceedings of the American Mathematical Society*, 130(3):637–642, 2002.
- [4] V. Beresnevich, F. Ramírez, and S. Velani. Metric Diophantine approximation: aspects of recent work, 2016.
- [5] B. Brouwer. Cantor sets for regular continued fractions and Lüroth series, 2024. Mentor: C. Kraaikamp.
- [6] C. Cao, J. Wu, and Z. Zhang. The efficiency of approximating real numbers by Lüroth expansion. *Czechoslovak mathematical journal*, 63:497–513, 2013.
- [7] T. Cusick. Sums and products of continued fractions. *Proceedings of the American Mathematical Society*, 27(1):35–38, 1971.
- [8] T. Cusick and R. Lee. Sums of sets of continued fractions. *Proceedings of the American Mathematical Society*, 30(2):241–246, 1971.
- [9] K. Dajani and C. Kraaikamp. *Ergodic Theory of Numbers*. Carus Mathematical Monographs. Mathematical Association of America, 2002.
- [10] T. Das, L. Fishman, D. Simmons, and M. Urbański. Hausdorff dimensions of perturbations of a conformal iterated function system via thermodynamic formalism. *Selecta Mathematica*, 29(2), Jan. 2023.
- [11] B. Diviš. On the sums of continued fractions. *Acta Arithmetica*, 22:157–173, 1973.
- [12] P. Erdős. Representations of real numbers as sums and products of Liouville numbers. *Michigan Math. J.*, 9(1):59–60, 1962.
- [13] D. Gayfulin and E. Nesharim. Every real number is a sum of two real numbers with diverging partial quotients, 2025.
- [14] I. J. Good. The fractional dimensional theory of continued fractions. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(3):199–228, 1941.
- [15] M. Hall. On the sum and product of continued fractions. *Annals of Mathematics*, 48(4):966–993, 1947.
- [16] J. L. Hlavka. Results on sums of continued fractions. *Transactions of the American Mathematical Society*, 211:123–134, 1975.
- [17] R. Kaufman. Continued fractions and Fourier transforms. *Mathematika*, 27(2):262–267, 1980.
- [18] Y. W. Lee. Explicit upper bounds on decay rates of Fourier transforms of self-similar measures on self-similar sets, 2024.
- [19] Y. W. Lee. Lüroth expansions in diophantine approximation: Metric properties and conjectures, 2025.
- [20] P. A. P. Moran. Additive functions of intervals and Hausdorff measure. *Mathematical Proceedings of the Cambridge Philosophical Society*, 42(1):15–23, 1946.
- [21] Y. Peres and P. Shmerkin. Resonance between Cantor sets. *Ergodic Theory and Dynamical Systems*, 29(1):201–221, 2009.
- [22] M. Queffélec and O. Ramaré. Analyse de Fourier des fractions continues à quotients restreints. *Enseignement Mathématique*, 49(3/4):335–356, 2003.
- [23] J. F. Randolph. Distances between points of the Cantor set. *The American Mathematical Monthly*, 47(8):549–551, 1940.
- [24] T. Sahlsten and C. Stevens. Fourier transform and expanding maps on Cantor sets. *American Journal of Mathematics*, 146(4):945–982, 2024.
- [25] N. Shulga. On a conjecture of Cusick on a sum of Cantor sets. *Transactions of the*

- American Mathematical Society*, 2024. Early View.
- [26] H. Steinhaus. Sur les distances des points dans les ensembles de mesure positive. *Fundamenta Mathematicae.*, 1(1), 1920.
 - [27] B. Tan and Q. Zhou. Approximation properties of Lüroth expansions. *Discrete & Continuous Dynamical Systems: Series A*, 41(6), 2021.
 - [28] B. Tan and Q. Zhou. Dimension theory of the product of partial quotients in Lüroth expansions. *International Journal of Number Theory*, 17(05):1139–1154, 2021.

Preprint 2: Lower bound for the Hausdorff dimension of expressible sets

This chapter contains a preprint of the paper *Lower bound for the Hausdorff dimension of expressible sets* by the author, Jaroslav Hančl and Simon Kristensen. The paper is currently under review and has not yet been accepted for publication. The preprint can be accessed through arXiv on <https://arxiv.org/abs/2605.26839>. The pages are numbered 1 through 9. The references within are selfcontained and do not match the numbering in the remainder of the thesis.

LOWER BOUNDS FOR THE HAUSDORFF DIMENSION OF EXPRESSIBLE SETS

MAIKEN GRAVGAARD, JAROSLAV HANČL, AND SIMON KRISTENSEN

ABSTRACT. We obtain positive lower bounds on the Hausdorff dimension of sets of real numbers given by expressions of the form $\sum_{n=1}^{\infty} \frac{1}{a_n b_n}$, where b_n satisfies some growth condition and a_n lies in some set, possibly depending on n . As a consequence of our results, some of the irrational numbers arising from Erdős' celebrated construction from 1976 are not Liouville numbers.

1. INTRODUCTION

In his seminal paper from 1975 [4], Erdős famously showed that if $(a_n)_{n=1}^{\infty}$ is an increasing sequence of integers with $a_n \geq n^{1+\varepsilon}$ for some $\varepsilon > 0$ and with $\limsup_{n \rightarrow \infty} a_n^{1/2^n} = \infty$, then the real number

$$\sum_{n=1}^{\infty} \frac{1}{a_n}$$

is irrational. This was later strengthened by Hančl [7], who proved that the condition $\limsup_{n \rightarrow \infty} a_n^{1/2^n} = \infty$ may be replaced with the condition that the sequence $(a_n^{1/2^n})_{n=1}^{\infty}$ is divergent with lower limit > 1 .

It was seen, for instance in [8], that this condition cannot be weakened further in terms of growth. Indeed, if a_n is given by the recurrence $a_{n+1} = a_n^2 - a_n + 1$, then

$$\sum_{n=1}^{\infty} \frac{1}{a_n} = \frac{1}{a_1 - 1}.$$

On the other hand, it is similarly easy to show that the sequence with terms $a_n^{1/2^n}$ is convergent, and indeed it was shown by Aho and Sloane [1] that for $a_1 = 2$, the limit is about 1.264, which is certainly greater than 1.

At a first glance, it seems reasonable to attempt to prove results of this kind by approximating the sum by its truncations. These will of course be rationals, and as the denominators increase very rapidly indeed, in the absence of cancellation, one would suspect them to converge rapidly enough to provide a proof of the irrationality of the sum. However, as the example above shows, cancellation does appear and for some sequences in a fairly dramatic way at the limit of the method.

It is of some interest to know whether the method – when it works – yields numbers which are not extremely well approximated by rationals. By this, we mean numbers which are not Liouville numbers. We recall that a Liouville number is a real number ℓ such that for any $w > 1$, there is a rational $p/q \in \mathbb{Q}$ for which

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

The set consists only of transcendental numbers, and is a set of Hausdorff dimension (see below for a definition) zero.

The questions whether or not a sum of the above form is irrational leads us to the notion of an expressible set. Given a sequence $A = (a_n)_{n=1}^{\infty}$, we will say that the expressible set

of the sequence is

$$\mathcal{K}(A) = \left\{ x = \sum_{n=1}^{\infty} \frac{1}{d_n a_n} : d_n \in \mathbb{N} \right\}.$$

An alternative way of stating Erdős' result [4] is that if $\limsup_{n \rightarrow \infty} a_n^{1/2^n} = \infty$, then $\mathcal{K}(A)$ contains only irrational numbers. However in the same paper, Erdős proved that if the growth condition is strengthened to the condition that $\limsup_{n \rightarrow \infty} a_n^{1/t^n} = \infty$ for any $t \in \mathbb{N}$, then every element in $\mathcal{K}(A)$ is a Liouville number. A consequence of Theorem 3 below is that this is sharp. If there is a $t \in \mathbb{N}$ so that $\limsup_{n \rightarrow \infty} a_n^{1/t^n} < \infty$, then $\mathcal{K}(A)$ contains non-Liouville numbers.

Another natural question is then what growth conditions of a sequence ensure that the expressible set contains only transcendental numbers. An answer to this was obtained by Hančl [6] in more generality. Note that there is nothing in our set-up requiring the sequence $(a_n)_{n=1}^{\infty}$ to consist of integers. Hančl proved that if $a_n = p_n/q_n$ is rational with the sequence of numerators $p_n \leq 2^{(3+\beta)^n}$ and denominators $q_n \geq 2^{(3+\alpha)^n}$ for some choice of $\alpha > \beta > 0$, then the expressible set $\mathcal{K}(A)$ consists exclusively of transcendental numbers. We will give a growth classification in terms of Mahler's and Koksma's classification of transcendental numbers below in Corollary 5. For sequences A of algebraic numbers, see Hančl and Nair [8], Andersen and Kristensen [2] and Laursen [15] for related results.

From the point of view of measure, it was shown by Hančl, Nair and Šustek [10] that of $a_n = p_n/q_n$, $\limsup_{n \rightarrow \infty} q_n^{1/3^n} = \infty$, $q_n \geq n^{1+\varepsilon}$ and $p_n \leq 2^{\log_2^\alpha a_n}$ for some $\alpha \in (0, 1)$, then $\mathcal{K}(A)$ has Lebesgue measure zero. Under the same conditions, but with the first replaced by $\limsup_{n \rightarrow \infty} q_n^{1/(3+\delta)^n} = \infty$, it was shown by Hančl, Nair, Novotný and Šustek [9] that the Hausdorff dimension of the set is at most $2/(2+\delta)$.

Hančl and Šustek [14] proved that $\mathcal{K}(A)$ has zero Lebesgue measure for the sequence with $a_n = 2^{3^n}$. Furthermore, if $a_{n+1} \geq n2^{2^{a_n}}$, Hančl and Šustek [12] showed that $\mathcal{K}(A)$ has Hausdorff dimension zero. This of course follows from Erdős' result [4] if we assume the a_n to be integers, since in this case we would see only Liouville numbers in $\mathcal{K}(A)$, but they make no assumption on the integrality of the sequence.

All the above results are concerned with sequences of doubly exponential growth. As we will also be dealing with sequences of the form $a_n = A^{n-1}$, we will briefly mention some results on the expressible set of these sequences. In this case, A can be assumed to be any real number. It was shown by Hančl, Schintzel and Šustek [11] that for $0 < A \leq 1$, $\mathcal{K}(A) = \mathbb{R}_+$. For $A > 1$, the series $\sum_{n=1}^{\infty} \frac{1}{A^{n-1}} = \frac{1}{A-1}$, so there is no hope for this to remain true, but for $1 < A \leq 3$, $\mathcal{K}(A) = (0, \frac{1}{A-1}]$, the maximal possible. However, for $3 < A$, they were only able to prove that $\mathcal{K}(A) \supseteq (0, \frac{1}{(A-1)(\lceil A \rceil - 2)}]$.

In the present paper, we will be considering expressible sets, but with additional restrictions. Concretely, we will consider various growth restrictions on the sequence $A = (a_n)_{n=1}^{\infty}$, but also restrictions on the 'digits', d_n , which will be required to lie in some set $\mathbb{D}_n$, possibly depending on n . We will assume them to be natural numbers, but unless otherwise stated, we can assume the a_n to be real numbers.

For a sequence $A = (a_n)_{n=1}^{\infty}$ and a sequence of sets of natural numbers, $(\mathbb{D}_n)_{n=1}^{\infty}$, we define

$$(1) \quad \mathcal{K}(A, (\mathbb{D}_n)_{n=1}^{\infty}) = \left\{ x = \sum_{n=1}^{\infty} \frac{1}{d_n a_n} : d_n \in \mathbb{D}_n \right\}.$$

Putting restrictions on the 'digits' seems natural, but the only previous results we are aware of are due to Hančl and Šustek [13], who considered expressible sets of integer sequences, where the digits were assumed to be bounded. Criteria implying Lebesgue

measure zero as well as upper bounds on the Hausdorff dimension of the expressible set with these restrictions were derived. In all cases, the sequence A was assumed to consist of integers.

2. BACKGROUND

We will be concerned with the Hausdorff dimension of expressible sets. Hausdorff dimension is defined in terms of Hausdorff measures. We briefly define these concepts. For a set $E \subseteq \mathbb{R}^n$ and real numbers $s \geq 0$ and $\delta > 0$, let

$$\mathcal{H}_\delta^s(E) = \inf \left\{ \sum_{U \in \mathcal{C}_\delta} \text{diam}(U)^s : \mathcal{C}_\delta \text{ is a cover of } E \text{ with sets of diameter } \leq \delta \right\}.$$

As δ decreases, there are fewer covers at our disposal, so the infimum can only increase. Thus, allowing for the limit to be ∞ , we can define

$$\mathcal{H}^s(E) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(E),$$

the Hausdorff s -measure of E . This construction yields an outer measure on $\mathbb{R}^n$, for which the Borel sets are measurable and which is even Borel regular. For a fixed set E , it is a decreasing function of s , and can take a positive and finite value for at most one value of s . We thus define the Hausdorff dimension of E to be

$$\dim_{\mathcal{H}}(E) = \inf \{s \geq 0 : \mathcal{H}^s(E) = 0\}$$

The notion is well-defined due to the remarks preceding its definition. It has most of the properties one would expect from a dimension. We refer to Falconer's book [5] for details.

A plethora of examples of sets in $\mathbb{R}$ for which the Hausdorff dimension may be estimated are (generalised) Cantor sets. One constructs such a set by starting with the closed unit interval at level 0. This is split into m_1 closed intervals, which are kept and where the distance between them is at most ε_1 . The procedure is repeated, so that each of the level 1 intervals is split into m_2 closed intervals, and so that the gaps between the remaining intervals is at least ε_2 . This procedure is repeated *ad infinitum*, and the resulting set denoted by $\mathcal{C}$, see Figure 1.

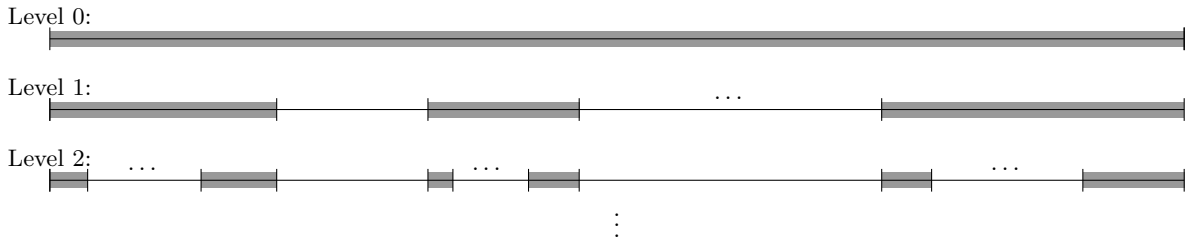


FIGURE 1. Generalised Cantor set with m_n sub-intervals in layer n and gaps of size $\geq \varepsilon_n$

The following is stated as Example 4.6 in [5]. We state it here as a theorem.

Theorem 1. *Let $\mathcal{C}$ be a generalised Cantor set (see Figure 1). Assume that there exists a positive integer n_0 such that, for any $n \geq n_0$, each interval in level n splits into m_n subintervals and each gap in level n is $\geq \varepsilon_n$. Then*

$$\dim_{\mathcal{H}}(\mathcal{C}) \geq \limsup_{n \rightarrow \infty} \frac{\log(m_1 \dots m_n)}{-\log(m_{n+1}\varepsilon_{n+1})}$$

In a corollary below, we will make use of Koksma's classification of transcendental numbers, see e.g. Bugeaud's book [3]. We briefly define the classes here.

First, for $x \in \mathbb{R}$ and $n \in \mathbb{N}$, let

$$w_n^*(x) = \sup \{w > 0 : |x - \alpha| < H(\alpha)^{-w-1} \text{ for infinitely many } \alpha \text{ with } \deg(\alpha) \leq n\}.$$

Here, $H(\alpha)$ denotes the naive height of α , i.e. the maximum absolute value of the coefficients of the minimal integer polynomial of α . Furthermore, define

$$w^*(x) = \limsup_{n \rightarrow \infty} \frac{w_n(x)}{n}.$$

The classes in Koksma's classification are:

- A^* -numbers, which are the algebraic numbers.
- S^* -numbers, which are the numbers x such that $w_n^*(x) < \infty$ and $w^*(x) < \infty$.
- T^* -numbers, which are the numbers x such that $w_n^*(x) < \infty$ but $w^*(x) = \infty$.
- U^* -numbers, which are numbers such that $w_n^*(x) = \infty$ for n large enough.

It is a consequence of the algebraic invariance of the classes and Lebesgue's density theorem that one of them must contain almost all numbers, and indeed almost all numbers are S -numbers. U -numbers form a set of Hausdorff dimension 0, which contains the Liouville numbers. For further details on this classification and the related classification of Mahler, see [3].

3. THEOREMS AND PROOFS

We will first describe the overall strategy of our proofs. It is helpful to think of the series as a numeration system, so that the sequence A is thought of as the 'base' and the varying elements from the $\mathbb{D}_n$ are thought of as the 'digits'.

For an increasing non-negative sequence $A = (a_n)_{n \in \mathbb{N}}$ and a finite set of 'digits'

$$\mathbb{D}_n = \{1 = d_{n,1} < \dots < d_{n,m_n}\} \subset \mathbb{N}$$

for each n , consider the set

$$\mathcal{K}(A, (\mathbb{D}_n)_{n=1}^\infty) = \left\{ \sum_{n=1}^\infty \frac{1}{a_n d_n} : d_n \in \mathbb{D}_n \text{ for each } n \in \mathbb{N} \right\}.$$

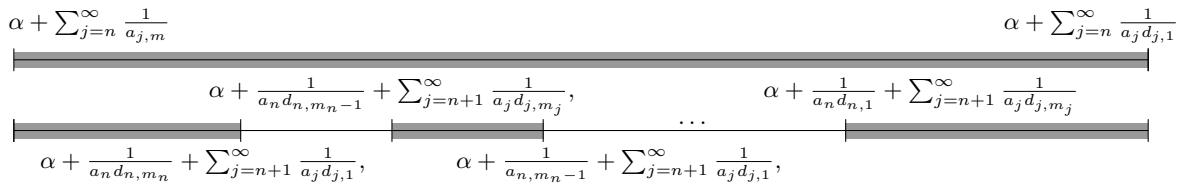


FIGURE 2. Splitting an interval in level $n - 1$ into subintervals in level n of the Cantor set $\mathcal{K}$ (Here α is of the form $\alpha = \frac{1}{a_1 d_{1,k_1}} + \dots + \frac{1}{a_{n-1} d_{1,k_{n-1}}}$, $1 \leq k_i \leq m_i$ for each $1 \leq i \leq n - 1$), so that the expression in the figure are indeed the end point of bigger interval.

We wish to apply Theorem 1. As such, it is important for us to get estimates on the involved quantities, m_n and ε_n . In the proofs, m_n will be (one less than) the number of digits, and so be part of the assumptions. This leaves us with the gap sizes, which must be lower bounded by some ε_n . The main difficulty in the proofs is to ensure that the conditions stated ensure that gaps do indeed occur, and that their lengths may be lower bounded by some function.

We now state and prove three theorems, where the sequence A is allowed different growth rates. In the first theorem, we consider the case when the growth is geometric, i.e. if $a_n = b^n$ for some real number $b > 1$. In this case, the result of Hančl, Schintzel and Šustek [11] implies that the full expressible set $\mathcal{K}(A)$ is of maximal Hausdorff dimension and in fact contains an interval. However, restricting the digits to lie in some small set will remove this property. We obtain the following.

Theorem 2. *Assume $a_n = b^n$ for some $b > 4$ and $\mathbb{D}_n = \{1, 2, \dots, K\}$ for every n , where $K^2 < b$. Then*

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty})) \geq \frac{\log(K)}{\log(b)}.$$

Note that the theorem remains valid, but trivial, for $1 < b \leq 4$. Indeed, the condition that $K^2 < b$ will in this case imply that $K = 1$, so that the lower bound obtained on the dimension is equal to zero.

Proof. In this case $m_n = K$, if we make sure that all of the expected gaps occur. Let $n \in \mathbb{N}$ and $1 \leq k \leq K - 1$

$$\begin{aligned} \text{gap in level } n &= \left(\alpha + \frac{1}{b^n k} + \sum_{j=n+1}^{\infty} \frac{1}{K b^j} \right) - \left(\alpha + \frac{1}{b^n (k+1)} + \sum_{j=n+1}^{\infty} \frac{1}{b^j} \right) \\ &= \frac{1}{b^n} \left(\frac{1}{k(k+1)} \right) + \left(\frac{1}{K} - 1 \right) \frac{1}{b^n} \frac{1}{b-1} \\ &\geq \frac{1}{b^n} \frac{1}{K} \left(\frac{1}{K+1} + (1-K) \frac{1}{b-1} \right) =: \varepsilon_n > 0 \text{ since } K^2 < b, \end{aligned}$$

so we have a valid Cantor set (meaning that there are in fact gaps everywhere we expect there to be) and expressions for m_n and ε_n . Now we plug it into Theorem 1

$$\begin{aligned} \dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^{\infty}, (\mathbb{D}_n)_{n=1}^{\infty})) &\geq \limsup_{n \rightarrow \infty} \frac{\log(K^n)}{-\log \left(K \frac{1}{b^n} \frac{1}{K} \left(\frac{1}{K+1} + (1-K) \frac{1}{b-1} \right) \right)} \\ &= \limsup_{n \rightarrow \infty} \frac{n \log(K)}{n \log(b) - \log \left(\left(\frac{1}{K+1} + (1-K) \frac{1}{b-1} \right) \right)} \\ &= \frac{\log(K)}{\log(b)}. \end{aligned}$$

□

In our second result, we consider sequences of doubly exponential growth.

Theorem 3. *Let $N \in \mathbb{N}_{\geq 2}$ and $0 < s \leq r < \frac{N-1}{2}$. Take functions $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ satisfying $1 \leq f(n), g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq 2^{(N-1)n}$ and $2^{sN^n} g(n) \leq 2^{rN^n} h(n)$. Define for each $n \in \mathbb{N}$*

$$\begin{aligned} a_n &= 2^{N^n} f(n), \\ m_n &= \lfloor 2^{sN^n} g(n) \rfloor, \end{aligned}$$

and

$$\mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots, d_{n,m_n} \leq 2^{rN^n} h(n)\} \subset \mathbb{N}.$$

Then

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n \in \mathbb{N}}, (\mathbb{D}_n)_{n=1}^{\infty})) \geq \frac{s}{(N-1)(N-s)}.$$

Proof. First we prove that there exists an $\varepsilon > 0$ and an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have

$$(2) \quad \min_{x \neq y \in \mathbb{D}_n} \left| \frac{1}{xa_n} - \frac{1}{ya_n} \right| \geq (1 + \varepsilon) \sum_{k=n+1}^{\infty} \frac{1}{a_k}.$$

Consider first the left hand side of (2):

$$\begin{aligned} \min_{x \neq y \in \mathbb{D}_n} \left| \frac{1}{xa_n} - \frac{1}{ya_n} \right| &\geq \frac{1}{a_n} \left(\frac{1}{d_{n,m_n} - 1} - \frac{1}{d_{n,m_n}} \right) \\ &\geq \frac{1}{2^{N^n} f(n)} \frac{1}{2^{2rN^n} h(n)^2} \\ &\geq \frac{1}{2^{(1+2r)(N^n + (N-1)^n)}} \end{aligned}$$

As for the right hand side of (2), we have

$$\begin{aligned} (1 + \varepsilon) \sum_{k=n+1}^{\infty} \frac{1}{a_k} &\leq (1 + \varepsilon) \frac{2}{2^{N^{n+1}} f(n)} \\ &\leq (1 + \varepsilon) \frac{2}{2^{N^{n+1}}}, \end{aligned}$$

so we just need to show, that there is an $\varepsilon > 0$ and an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have

$$\frac{1}{2^{(1+2r)(N^n + (N-1)^n)}} \geq (1 + \varepsilon) \frac{2}{2^{N^{n+1}}},$$

i.e.

$$\varepsilon \leq \frac{2^{(N-2r-1)N^n - (1+2r)(N-1)^n} - 2}{2},$$

In the above expression, the right hand side goes to ∞ as $n \rightarrow \infty$ (since $N - 2r - 1 > 0$), so for any $\varepsilon > 0$, there is an $n_0(\varepsilon) \in \mathbb{N}$ such that this holds, so the claim is true.

Now to look at the gaps in the Cantor set. Let $n \geq n_0$. For $j \in \{2, \dots, m_n\}$ the j 'th gap in layer n is:

$$\begin{aligned} \text{gap}_n^j &= \frac{1}{d_{n,j-1}a_n} + \sum_{k=n+1}^{\infty} \frac{1}{d_{k,m_k}a_k} - \frac{1}{d_{n,j}a_n} - \sum_{k=n+1}^{\infty} \frac{1}{a_k} \\ &\stackrel{(2)}{\geq} (1 + \varepsilon) \sum_{k=n+1}^{\infty} \frac{1}{a_k} - \sum_{k=n+1}^{\infty} \frac{1}{a_k} + \sum_{k=n+1}^{\infty} \frac{1}{d_{k,m_k}a_k} \\ &\geq \varepsilon \sum_{k=n+1}^{\infty} \frac{1}{a_k} \geq \frac{\varepsilon}{2^{N^{n+1}} f(n+1)} =: \varepsilon_n > 0. \end{aligned}$$

So after level n_0 all of the expected gaps do indeed occur, and we have a lower bound for the size of the gaps in layer n , namely $\varepsilon_n = \frac{\varepsilon}{2^{N^{n+1}} f(n+1)}$.

For each $n \geq n_0$ we have, since $\lfloor 2^{sN^k} g(n) \rfloor \geq \frac{1}{2} 2^{sN^k} g(n) \geq \frac{1}{2} 2^{sN^k}$,

$$\begin{aligned} m_1 \cdot \dots \cdot m_{n-1} &= \lfloor 2^{N^1} g(1) \rfloor \cdot \dots \cdot \lfloor 2^{N^{n-1}} g(n-1) \rfloor \\ &\geq \left(\frac{1}{2} \right)^{n-1} 2^{s(N^1 + \dots + N^{n-1})} \\ &= 2^{s \frac{N^n - N}{N-1} - (n-1)}. \end{aligned} \tag{3}$$

And for each $n \geq n_0$ we have

$$\begin{aligned}
 \varepsilon_n m_n &= \frac{\varepsilon}{2^{N^{n+1}} f(n+1)} \lfloor 2^{sN^n} g(n) \rfloor \\
 &\geq \frac{\varepsilon}{2} \cdot \frac{1}{2^{(N-s)N^n}} \frac{g(n)}{f(n+1)} \\
 (4) \quad &\geq \frac{\varepsilon}{2} \cdot \frac{1}{2^{(N-s)N^n + (N-1)^{n+1}}},
 \end{aligned}$$

so

$$-\log(\varepsilon_n m_n) \leq -\log\left(\frac{\varepsilon}{2}\right) + \log(2) [(N-s)N^n + (N-1)^{n+1}].$$

To find a lower bound for the Hausdorff dimension of $\mathcal{K}((a_n)_{n \in \mathbb{N}}, (\mathbb{D}_n)_{n=1}^\infty)$ we apply Theorem 1 as well as (3) and (4):

$$\begin{aligned}
 \dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n \in \mathbb{N}}, (\mathbb{D}_n)_{n=1}^\infty)) &\geq \limsup_{n \rightarrow \infty} \frac{\log(m_1 \cdots m_{n-1})}{-\log(\varepsilon_n m_n)} \\
 &\geq \limsup_{n \rightarrow \infty} \frac{\log(2) \left(s \frac{N^n - N}{N-1} - (n-1)\right)}{-\log\left(\frac{\varepsilon}{2}\right) + \log(2) [(N-s)N^n + (N-1)^{n+1}]} \\
 &= \frac{s}{(N-1)(N-s)}
 \end{aligned}$$

□

Our third theorem again concerns doubly exponential growth, but this time slower than that of Theorem 3.

Theorem 4. *Let $N \in \mathbb{N}_{\geq 2}$ and $0 < s \leq r < \frac{1}{2}$. Let $0 < \eta < 1$. Take functions $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ satisfying $1 \leq f(n), g(n)^{\frac{1}{s}}, h(n)^{\frac{1}{r}} \leq N^{(2-\eta)^n}$ and $N^{s2^n} g(n) \leq N^{r2^n} h(n)$. Define for each $n \in \mathbb{N}$*

$$\begin{aligned}
 a_n &= N^{2^n} f(n), \\
 m_n &= \lfloor N^{s2^n} g(n) \rfloor,
 \end{aligned}$$

and

$$\mathbb{D}_n = \{1 = d_{n,1} < d_{n,2} < \dots, d_{n,m_n} \leq N^{r2^n} h(n)\} \subset \mathbb{N}.$$

Then

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty)) \geq \frac{s}{2-s}.$$

Proof. Again we have for some $\varepsilon > 0$ and some $n_0 \in \mathbb{N}$, that for any $n \geq n_0$:

$$(5) \quad \min_{x \neq y \in \mathbb{D}_n} \left| \frac{1}{xa_n} - \frac{1}{ya_n} \right| \geq (1+\varepsilon) \sum_{k=n+1}^\infty \frac{1}{a_k}.$$

Indeed

$$\min_{x \neq y \in \mathbb{D}_n} \left| \frac{1}{xa_n} - \frac{1}{ya_n} \right| \geq \frac{1}{d_{n,m_n}^2 a_n} \geq \frac{1}{N^{(1+2r)2^n} N^{(1+2r)(2-\eta)^n}}$$

and

$$(1+\varepsilon) \sum_{k=n+1}^\infty \frac{1}{a_k} \leq (1+\varepsilon) \frac{2}{N^{2^{n+1}}}$$

and since $r < \frac{1}{2}$, for any $\varepsilon > 0$ there is an $n_0(\varepsilon) \in \mathbb{N}$ such that for all $n \geq n_0(\varepsilon)$

$$\varepsilon > \frac{N^{(1-2r)2^n - (1+2r)(2-\eta)^n} - 2}{2}.$$

Exactly as in the proof of Theorem 3, this implies that the gaps in layer n of the Cantor set are $\geq \frac{\varepsilon}{N^{2^n+1}}$. Furthermore we have

$$(6) \quad \begin{aligned} m_1 \cdot \dots \cdot m_{n-1} &= \lfloor N^{2^1} g(1) \rfloor \cdot \dots \cdot \lfloor N^{2^{n-1}} g(n-1) \rfloor \\ &\geq \left(\frac{1}{2}\right)^{n-1} N^{s(2^n-2)} \end{aligned}$$

and

$$(7) \quad \begin{aligned} \varepsilon_n m_n &= \frac{\varepsilon}{N^{2^n+1}} \lfloor N^{s2^n} g(n) \rfloor \\ &\geq \frac{\varepsilon}{2} \frac{1}{N^{(2-s)2^n}}. \end{aligned}$$

We find our bound for the Hausdorff dimension of $\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty)$ by applying Theorem 1 as well as (6) and (7):

$$\dim_{\mathcal{H}}(\mathcal{K}((a_n)_{n=1}^\infty, (\mathbb{D}_n)_{n=1}^\infty)) \geq \limsup_{n \rightarrow \infty} \frac{\log(N)s2^n - \log(N)s2 - \log(2)(n-1)}{-\log\left(\frac{\varepsilon}{2}\right) + (2-s)2^n \log(N)} = \frac{s}{2-s}.$$

□

In the case $N = 2$ Theorem 3 and Theorem 4 agree. It is worth noting that the bound obtained in Theorem 4 is independent of N , whereas the bound in Theorem 3 tends to zero as N increases. This is natural, as the result of Erdős [4] mentioned in the introduction says that if the a_n are integers, then the limiting set would contain only Liouville numbers, which is a set of Hausdorff dimension 0.

We note however that in the above results, the sequence $A = (a_n)_{n=1}^\infty$ is not assumed to be a sequence of integers. Any sequence of real numbers satisfying the required growth conditions will suffice.

Assuming that the sequence consists of rational numbers or of algebraic numbers brings us into the realms of the works of Hančl [6], Hančl and Nair [8], Andersen and Kristensen [2] and Laursen [15]. The following result is mostly relevant for these settings. For sequences of integers, it shows in particular that Erdős' criterion on when an expressible set contains only Liouville numbers is sharp.

Corollary 5. *Suppose that the a_n are all integers with $\limsup_{n \rightarrow \infty} a_n^{1/t^n} < \infty$ for some $t \in \mathbb{N}$. Then $\mathcal{K}(A)$ contains numbers which are not U^* -numbers.*

Proof. The growth condition on a_n implies that $a_n \leq 2^{N^n}$ for some $N \in \mathbb{N}_{\geq 2}$, which is the same as that in Theorem 3. Since $\mathbb{D}_n \subseteq \mathbb{N}$ for any choice of digit sets, $\mathcal{K}(A, (\mathbb{D}_n)_{n=1}^\infty) \subseteq \mathcal{K}(A)$, so choosing digits so that the Hausdorff dimension of the former is positive immediately implies the positivity of the Hausdorff dimension of $\mathcal{K}(A)$. Since the Hausdorff dimension of the set of U^* -numbers is equal to 0, there must be other types of numbers in $\mathcal{K}(A)$. □

4. CONCLUDING REMARKS AND OPEN PROBLEMS

We end the paper with some remarks on the sharpness and limitations of our results.

For sequences of exponential growth, Theorem 2 appears to be reasonably sharp. Indeed, the lower bound would tend to 1 for K tending to b , which at least for integral b seems a natural restriction on the digits. However, the proof requires us to assume that $K^2 < n$ in order to ensure sufficiently large gaps in the construction, so values of K approaching b are not allowed. Removing this restriction is an open problem.

Concerning Theorems 3 and 4, we suspect that the bounds obtained here are not best possible. We are not aware of work calculating upper bounds on sets of the exact same

form as those considered here, but for related sets there are upper bounds by Hančl and Šustek [13]. However, their sets assume a uniform bound on the ‘digits’ of any individual member of the set, which is not the case here. Getting an upper bound on the Hausdorff dimension for the sets of this paper remains an open problem.

Funding. MG’s and SK’s research is supported by Aarhus University Research Foundation, grant no. AUFF-E-2021-9-2. JH is supported by the University of Ostrava.

Data Availability. Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Declaration. The authors declare that they have no conflict of interest.

REFERENCES

- [1] A. V. AHO AND N. J. A. SLOANE, *Some doubly exponential sequences*, Fibonacci Q., 11 (1973), pp. 429–437.
- [2] S. B. ANDERSEN AND S. KRISTENSEN, *Arithmetic properties of series of reciprocals of algebraic integers*, Monatsh. Math., 190 (2019), pp. 641–656.
- [3] Y. BUGEAUD, *Approximation by algebraic numbers*, vol. 160 of Camb. Tracts Math., Cambridge: Cambridge University Press, 2007.
- [4] P. ERDŐS, *Some problems and results on the irrationality of the sum of infinite series*. J. Math. Sci. 10, 1-7 (1975)., 1975.
- [5] K. FALCONER, *Fractal geometry. Mathematical foundations and applications*, Chichester: Wiley, 2nd ed. ed., 2003.
- [6] J. HANČL, *Transcendental sequences*, Math. Slovaca, 46 (1996), pp. 177–179.
- [7] ———, *A criterion for linear independence of series*, Rocky Mt. J. Math., 34 (2004), pp. 173–186.
- [8] J. HANČL AND R. NAIR, *On the irrationality of infinite series of reciprocals of square roots*, Rocky Mt. J. Math., 47 (2017), pp. 1525–1538.
- [9] J. HANČL, R. NAIR, L. NOVOTNÝ, AND J. ŠUSTEK, *On the Hausdorff dimension of the expressible set of certain sequences*, Acta Arith., 155 (2012), pp. 85–90.
- [10] J. HANČL, R. NAIR, AND J. ŠUSTEK, *On the Lebesgue measure of the expressible set of certain sequences*, Indag. Math., New Ser., 17 (2006), pp. 567–581.
- [11] J. HANČL, A. SCHINZEL, AND J. ŠUSTEK, *On expressible sets of geometric sequences*, Funct. Approximatio, Comment. Math., 39 (2008), pp. 71–95.
- [12] J. HANČL AND J. ŠUSTEK, *Expressible sets of sequences with Hausdorff dimension zero*, Monatsh. Math., 152 (2007), pp. 315–319.
- [13] ———, *Boundedly expressible sets.*, Czech. Math. J., 59 (2009), pp. 649–654.
- [14] ———, *Sequences of Cantor type and their expressibility*, Math. Slovaca, 67 (2017), pp. 41–50.
- [15] M. LØKKEGAARD LAURSEN, *Algebraic degree of series of reciprocal algebraic integers*, Rocky Mt. J. Math., 53 (2023), pp. 517–529.

MAIKEN GRAVGAARD, DEPARTMENT OF MATHEMATICS, AARHUS UNIVERSITY, NY MUNKEGADE 118, 8000 AARHUS C, DENMARK.

Email address: maiken@math.au.dk

JAROSLAV HANČL, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSTRAVA, 30. DUBNA 22, 701 03 OSTRAVA 1, CZECH REPUBLIC.

Email address: hancl@osu.cz

SIMON KRISTENSEN, DEPARTMENT OF MATHEMATICS, AARHUS UNIVERSITY, NY MUNKEGADE 118, 8000 AARHUS C, DENMARK.

Email address: sik@math.au.dk

Preprint 3: Quantitative Khintchine on the parabola with non-monotonic approximation functions

This chapter contains a preprint of the paper *Quantitative Khintchine on the parabola with non-monotonic approximation functions* by the author and Simon Kristensen. The pages are numbered 1 through 19. The references within are selfcontained and do not match the numbering in the remainder of the thesis.

QUANTITATIVE KHINTCHINE ON THE PARABOLA WITH NON-MONOTONIC APPROXIMATION FUNCTIONS

MAIKEN GRAVAARD AND SIMON KRISTENSEN

ABSTRACT. We prove a quantitative version of the convergence case of Khintchine's celebrated theorem in metric Diophantine approximation, but where the approximated points are restricted to lying on the parabola. A novel feature of our result is that unlike other results in literature, the approximating function is no longer required to be monotonic. This requires us to obtain explicit constants in classical number theoretic results, most notably in Burgess' bound for character sums in short intervals.

1. INTRODUCTION

Begin by recalling the notion of ψ -approximable points. Let ψ be a function $\psi : \mathbb{N} \rightarrow [0, \infty)$, which we will hereby refer to as the *approximation function*. Let $n \in \mathbb{N}$, then we define the set of *simultaneously ψ -approximable points* in $\mathbb{R}^n$ as follows

$$\mathcal{S}_n(\psi) := \left\{ (x_1, \dots, x_n) \in \mathbb{R}^n \mid \exists_{\infty} (p_1, \dots, p_n, q) \in \mathbb{Z}^n \times \mathbb{N} \text{ s.t. } \left| x_i - \frac{p_i}{q} \right| < \frac{\psi(q)}{q}, 1 \leq i \leq n \right\}.$$

By Dirichlet's Theorem [10], $\mathcal{S}_n(q \mapsto q^{-\frac{1}{n}}) = \mathbb{R}^n$, and for functions decaying faster than $q \mapsto q^{-\frac{1}{n}}$ Khintchine proved his famous 0–1-law, which we state a version of below:

Theorem (Khintchine, 1924). *λ_n -almost no points are simultaneously ψ -well approximable if the series $\sum_{q=1}^{\infty} \psi(q)^n$ converges. If we additionally assume, that the approximation function ψ is monotonic, then λ_n -almost every point is simultaneously ψ -well approximable, provided the series $\sum_{q=1}^{\infty} \psi(q)^n$ diverges.*

Here, λ_n is the n -dimensional Lebesgue measure. Note here, that we do not assume monotonicity of the approximation function in the convergence case! (It is however both assumed and indeed needed in the divergence case, see [12]).

The present work deals with Diophantine approximation on manifolds. Here, one restricts the point $\mathbf{x}$ to be on some manifold, which is itself a Lebesgue null set with respect to the Lebesgue measure on the ambient space, but which nonetheless carries a natural measure. In this case, geometric conditions on the manifold plays an important role. For instance, if the manifold in question is in a coordinate hyperplane, one would only be observing Diophantine phenomena from the lower dimensional space. The usual way of avoiding this is to assume that the manifold ‘curves out of every affine subspace, a notion which can be formalised as an explicit non-degeneracy condition.

In [5] Beresnevich and Yang show the convergence case for non-degenerate submanifolds $\mathcal{M} \subset \mathbb{R}^n$ for monotonic ψ , and it was recently shown by Beresnevich, Datta and Yang [4] that a full (and even more general) version of Khintchine's theorem holds for such manifolds. However, it is critical for these results that the approximation function is monotonic, even in the convergence case.

As pointed out in [1], an immediate consequence of Beresnevich and Yang's theorem is

Corollary 1 (Corollary 1 in [1]). *Let ψ be monotonic such that $\sum_{q=1}^{\infty} \psi(q)^n$ converges. Then for almost every (with respect to the natural measure) $\mathbf{x} \in \mathcal{M}$, there exists a $\kappa(\mathbf{x}) > 0$ such that*

$$\|q\mathbf{x}\| > \kappa(\mathbf{x})\psi(q),$$

for all $q \in \mathbb{N}$.

Applications of this often require that $\kappa(\mathbf{x})$ is independent of $\mathbf{x}$, which is impossible to guarantee on a set of full Lebesgue measure. In [1] Adiceam, Beresnevich, Levesley, Velani and Zorin state the following problem

Problem 2 (Problem in [1]). *Investigate the dependency between $\kappa > 0$ and the probability of the set of vectors $\mathbf{x} \in \mathcal{M}$ where $\kappa(\mathbf{x}) = \kappa$.*

This problem was addressed by Datta in [11], who found a estimate lower estimate on the measure of the set of points in $\mathcal{M}$ satisfying the conditions of Corollary 1, but with κ being a fixed, small number. Again, for her proof to work, it is imperative that the approximation function ψ is monotonic.

Removing monotonicity from Khintchine's theorem is in general a hard problem, and in fact it is strictly necessary for $n = 1$ where a different volume sum is needed in the non-monotonic case, see the celebrated paper of Koukoulopoulos and Maynard [16]. For manifolds, even removing it in the case of convergence is hard and often requires the manifold to be very special indeed. To the best of our knowledge, the only result of its kind is the following, due to Huang [13].

Theorem 3 (Theorem 1 in [13]). *Let ψ be any approximation function, such that $\sum_{q=1}^{\infty} \psi(q)^2$ converges. Then almost no points on the parabola $\mathcal{P} := \{(x, x^2) \mid x \in [0, 1]\}$ are simultaneously ψ -well approximable.*

As in [1] and [11], one can formulate statements analogous to Corollary 1 and Problem 2, which is exactly what we aim to address.

Define

$$(1) \quad \mathcal{B}'(\psi, \kappa) := \{x \in [0, 1] : \max(\|qx\|, \|qx^2\|) > \kappa\psi(q) \text{ for all } q \in \mathbb{N}\}$$

We give two versions of a quantitative Khintchine's Theorem on the parabola with non-monotonic approximation functions for general composite q . The first, Theorem 4, holds for all q , at the cost of the constants appearing in $\kappa(\delta)$ being *very* small! The other version, Theorem 5, has more reasonable constants, but this comes at the cost of only considering *very* large q with large prime divisors.

Additionally we give two versions, where we assume that q does not have too many divisors (namely $q = p_1 \cdots p_n$ for a fixed n in Theorem 6 and $q = p$ in Theorem 7).

Theorem 4 (All q). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^{\infty} \psi(q)^2$ converges. Given any $0 < \delta < 1$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{2\sqrt{6}(5 \log(2))^4}{3\pi^3 \sqrt{\zeta\left(\frac{12}{11}\right)} 11^4}, \sqrt{\frac{\delta}{S_\psi}} \frac{1}{\sqrt{18\pi}}, \right. \\ &\quad \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(\frac{1}{2\pi^2 \sqrt{12}} \right)^{\frac{2}{3}} \left(99684 \cdot 4.19253564643679 \right. \\ &\quad \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(e^{e^8} - 10^5 \right) e^{\frac{5 \log(2) e^8}{2(8-1.39177)}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} \right. \\ &\quad \left. + 5 \cdot 4^{\frac{3}{16} + \frac{0.00218819 + 0.00218819^2}{2} + \frac{5}{2 \cdot 57}} \left(\frac{57}{\log(2)} \right)^{\frac{7.176846309399143769411680}{2}} \left(\frac{1}{e \cdot 0.00218819} \right)^{1 + \frac{0.00218819}{2}} \right)^{\frac{-2}{3}} \Bigg\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.00015248265228, \sqrt{\frac{\delta}{S_\psi}} 0.07502635967975, \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 10^{-10^{23.89775276641198}} \right\}. \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

Theorem 5 (Large q). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^\infty \psi(q)^2$ converges. Assume additionally, that $\psi(q) = 0$ for all $q \leq \exp(\exp(41))$. Given any $\delta > 0$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{2}{3\pi^2 \sqrt{\frac{\pi^2}{6} \zeta\left(\frac{112}{57}\right)}}, \sqrt{\frac{\delta}{S_\psi}} \sqrt{\frac{1}{18\pi^2}} \right. \\ &\quad \left. \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left(4.19253564643679 + 27 \cdot 4^{\frac{3}{16} + \frac{5}{2.57} + \frac{41}{e^{41}} + \left(\frac{41}{e^{41}}\right)^2} \right) \right)^{\frac{-2}{3}} \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.04064400, \sqrt{\frac{\delta}{S_\psi}} 0.07502635, \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 0.00499238 \right\}. \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

Theorem 6 ($q = p_1 \dots p_k$). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^\infty \psi(q)^2$ converges. Let $k \in \mathbb{N}$ and assume additionally, that $\psi(q) = 0$ for all q with more than k prime divisors. Given any $0 < \delta < 1$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} \frac{4}{3\pi^4} \frac{1}{2^n}, \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left[4.19253564643679 \cdot 99684 \right. \right. \right. \\ &\quad \left. \left. \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} 2^{\frac{7k}{2}} \right] \right)^{\frac{-2}{3}}, \sqrt{\frac{\delta}{S_\psi}} \sqrt{\frac{1}{18\pi^2}} \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_\psi}} 0.0410639 \frac{1}{2^n}, \right. \\ &\quad \left. \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 0.0598025 \left(417928.72 + 2082.9292 \cdot 2^{\frac{7k}{2}} \right)^{\frac{-2}{3}}, \sqrt{\frac{\delta}{S_\psi}} 0.0750263 \right\}. \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

Theorem 7 ($q = p$ prime). *Let $\psi : \mathbb{N} \rightarrow [0, \infty)$ be any function such that $S_\psi := \sum_{q=1}^\infty \psi(q)^2$ converges. Assume additionally, that $\psi(q) = 0$ whenever q is not a prime. Given any $0 < \delta < 1$, let*

$$\begin{aligned} \kappa(\delta) &< \min \left(\frac{1}{6 \max_p \psi(p)}, \frac{2\delta}{3\pi^2 S_\psi}, \left(\frac{\delta(e \cdot 0.249)^{\frac{1}{4}}}{2\pi^2 \sqrt{12} \cdot 10.0366 S_\psi} \right)^{\frac{2}{3}}, \sqrt{\frac{\delta}{18\pi^2 S_\psi}} \right) \\ &\simeq \min \left(\frac{1}{6 \max_p \psi(p)}, \frac{\delta}{S_\psi} 0.0675474558, \frac{\delta^{\frac{2}{3}}}{S_\psi^{\frac{2}{3}}} 0.0120432671, \sqrt{\frac{\delta}{S_\psi}} 0.0750263597 \right). \end{aligned}$$

Then

$$\lambda(\mathcal{B}'(\psi, \kappa)) \geq 1 - \delta.$$

Note that Theorem 6 of course applies here as well with $k = 1$, but the bounds are not easily comparable. It is a consequence of the versions of Burgess' theorem used in the different proofs. We could of course put the bounds from Theorem 6 into the minimum of Theorem 7, but we have opted for the above version which is already complicated to state.

All four proofs follow the same structure. For this reason, we will first prove Theorem 4 in detail, after which Theorems 5, 6, and 7 are shown with references to the main proof. The structure of the proof is as follows: Reduce the problem to counting rational points with fixed denominator near the parabola. Then this counting function is estimated by exponential sums, which in turn are rewritten into sums of Dirichlet characters. Up to this point we have followed the same steps as [13], making sure to keep all constants and lower order terms along the way. As in [13] we want to estimate our character sums using Burgess's bound

for short sums of Dirichlet characters, the problem is, that there is no effective version that holds for all composite moduli q . In Section 2 we take care of this problem.

2. NOTES ON EXPLICIT BURGESS BOUNDS

Before we get into the proofs of Theorems 4 and 5 we need an explicit version of Burgess's famous bound for short sums of Dirichlet characters [8].

Burgess's bound for short character sums. *Let χ be a non-principal character modulo q . Then for any $\varepsilon > 0$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \ll_{\varepsilon} \sqrt{N} q^{\frac{11}{16} + \varepsilon},$$

where the implied constant only depends on ε .

We will need an explicit version, where the implied constant is known. In addition, we will need versions both for composite and prime moduli.

For composite moduli q we split into three cases. For large moduli, we refer to a result of Bordignon [7] (See also [15]).

Lemma 8 (Large q . Theorem 1.3 in [7]). *Let $q > \exp(\exp(8))$. Let χ be a primitive Dirichlet character mod q . For any integers M and $N < q$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 5d(q)^{\frac{3}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))}$$

Here and throughout, $d(q)$ denotes the number of divisors of q .

For small values of q , the constant may be calculated by brute force, which we have done using **Python**.

Lemma 9 (Small q . See Appendix A). *For $1 \leq q \leq 10^5$ and $q^{\frac{3}{8}} \leq N \leq q^{\frac{5}{8}}$ we have*

$$\left| \sum_{j=1}^N \chi(j) \right| \leq 4.19253564643679 \sqrt{N} q^{\frac{3}{16}},$$

where $\chi(j) = \left(\frac{j}{q}\right)$ or $\chi(j) = \left(\frac{q}{j}\right)$.

This leaves the intermediate range. For this, we use again results of Bordignon [7], but in this case additional work is needed.

Lemma 10 (Medium q). *Let $10^5 \leq q \leq \exp(\exp(8))$ and let χ be a primitive Dirichlet character mod q . For any integers M and $N < q$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 27d(q)^{\frac{3}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))}$$

Proof. The statement follows from Theorem 3.4 in [7], which we state a simplified version of. Below we take $\tau(q)$ to be the number of prime divisors of q , C to be the Euler–Mascheroni constant, and $\varphi(q)$ to be Euler's totient function.

Let q be an integer. Let $g \geq 2$ be a real number and let m be a positive real number. Let χ be a primitive Dirichlet character modulo q . Assume that

$$(2) \quad q > \max \left\{ \left(\max \left\{ 29, 2^{\tau(q)+1} \frac{q}{\varphi(q)} + 1 \right\} \frac{g}{m^2 \log(q) \log(\log(q))} \right)^8, \left(\frac{12}{g} \right)^4, h^4, \left(\frac{16 \log(q)}{m^2 \log(\log(q))} \right)^8 \right\}.$$

Define

$$v_1(m, q) = \frac{2 \left(1 + \frac{2}{\varepsilon \log(q)} \right)}{m},$$

$$v_2(m, q, g) = \frac{2v_1(m, q)^4}{g} + \frac{2 \log(\log(q))^2 \log\left(1.85v_1(m, q)^2 q^{\frac{3}{8}} \frac{\log(q)}{g \log(\log(q))}\right)}{\log(q)^2},$$

and

$$v_3(m, q, g, h) = 2g \left(1 - \frac{1}{h} - \frac{g}{m^2 q^{\frac{1}{8}} \log(q) \log(\log(q))}\right)^{-1} \left(\frac{17v_2(m, q, g)}{4g^3}\right)^{\frac{1}{4}} \left(e^C + \frac{2.51}{\log(\log(q))^2}\right) + \frac{2}{\sqrt{g}}.$$

If

$$(3) \quad v_3(m, q, g, h) \leq m$$

holds, then for any integers M, N , we have

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq md(q)^{\frac{3}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))}$$

We claim, that for $g = 2$, $h = 17$, $m = 27$ and $10^5 \leq q \leq \exp(\exp(8))$, equations (2) and (3) hold. Indeed

$$\left(29 \frac{2}{27^2 \log(q) \log(\log(q))}\right)^8 \frac{1}{q} \leq \left(29 \frac{2}{27^2 \log(10^5) \log(\log(10^5))}\right)^8 \frac{1}{10^5} < 1.$$

We now state two lemmas from literature giving bounds on arithmetic functions. The first is a lower bound for the Euler function.

Lemma (Theorem 8.8.7 in [3]).

$$\phi(q) > \frac{q}{e^C \log(\log(q)) + \frac{3}{\log(\log(q))}}.$$

We also need an upper bound for the τ -function.

Lemma (Theorem 12 in [20]).

$$\tau(q) \leq \frac{\log(q)}{\log(\log(q))} \left(1 + \frac{1.4573}{\log(\log(q))}\right).$$

Using these, we obtain

$$\begin{aligned} & \left[\left(2^{\tau(q)+1} \frac{q}{\psi(q)} + 1 \right) \frac{2}{m^2 \log(q) \log(\log(q))} \right]^8 \frac{1}{q} \\ & \leq \left[\left(2^{\frac{\log(q)}{\log(\log(q))} \left(1 + \frac{1.4573}{\log(\log(q))}\right) + 1} \left(1.7811 \log(\log(q)) + \frac{3}{\log(\log(q))} \right) + 1 \right) \frac{2}{m^2 \log(q) \log(\log(q))} \right]^8 \frac{1}{q} \\ & \leq \left[\left(2^{\frac{\log(10^5)}{\log(\log(10^5))} \left(1 + \frac{1.4573}{\log(\log(10^5))}\right) + 1} \left(1.7811 \log(\log(10^5)) + \frac{3}{\log(\log(10^5))} \right) + 1 \right) \frac{2}{m^2 \log(10^5) \log(\log(10^5))} \right]^8 \frac{1}{10^5} \\ & < 1. \end{aligned}$$

as well as

$$\begin{aligned} \left(\frac{12}{2}\right)^4 &< q, \\ 17^4 &< q \end{aligned}$$

and

$$\left(\frac{16 \log(q)}{m^2 \log(\log(q))}\right)^8 \frac{1}{q} < 1.$$

So (2) holds.

Checking (3), we have for $v_1(m, q)$:

$$v_1(m, q) \leq \frac{2 \left(1 + \frac{2}{e \log(q)}\right)}{27} \leq \frac{2 \left(1 + \frac{2}{e \log(10^5)}\right)}{27} \leq 0.0789,$$

and for $v_2(m, q, g)$:

$$\begin{aligned} v_2(m, q, g) &\leq 0.0789^4 + \frac{2 \log(\log(q))^2}{\log(q)^2} \log \left(1.85 \cdot 0.0789^2 \frac{q^{\frac{3}{8}} \log(q)}{2 \log(\log(q))} \right) \\ &\leq 0.0789^4 + \frac{2 \log(\log(10^5))^2}{\log(10^5)^2} \log \left(1.85 \cdot 0.0789^2 \frac{(\exp(\exp(8)))^{\frac{3}{8}} \exp(8)}{16} \right) \leq 100.776. \end{aligned}$$

Notice here, that the above is absolutely not optimal, since we are estimating quite brutally using both the upper and lower limit for q . This means that the m value 27 likely is not optimal. The function above is eventually decreasing in q , so for intervals of large q , an upper bound can be obtained from plugging in just the lower bound for q .

And finally for $v_3(m, q, g, h)$:

$$\begin{aligned} v_3(m, q, g, h) &\leq 4 \left(\frac{16}{17} - \frac{2}{q^{\frac{1}{8}} \log(q) \log(\log(q))} \right)^{-1} \left(\frac{17 \cdot 100.776}{4 \cdot 2^3} \right)^{\frac{1}{4}} \left(e^C + \frac{2.51}{\log(\log(q))^2} \right) + \frac{1}{\sqrt{2}} \\ &\leq 4 \left(\frac{16}{17} - \frac{2}{10^{\frac{5}{8}} \log(10^5) \log(\log(10^5))} \right)^{-1} \left(\frac{17 \cdot 100.776}{4 \cdot 2^3} \right)^{\frac{1}{4}} \left(e^C + \frac{2.51}{\log(\log(10^5))^2} \right) + \frac{1}{\sqrt{2}} \\ &\leq 26.7236... < 27. \end{aligned}$$

□

We will also need to consider Dirichlet characters which are not primitive. To extend to this case, we employ the following, as remarked by Burgess in [8]:

Remark 11. *If $\chi \bmod q$ is not primitive, it is induced by a primitive character, so we may write $\chi = \chi_1 \chi_2$, where χ_1 is primitive mod k_1 and χ_2 is principal mod k_2 ($k_1 + k_2 = q$), and then*

$$\begin{aligned} \sum_{M < n \leq M+N} \chi(n) &= \sum_{\substack{M < n \leq M+N \\ \gcd(n, k_2)=1}} \chi_1(n) \\ &= \sum_{M < n \leq M+N} \chi_1(n) \sum_{t | \gcd(n, k_2)} \mu(t) \end{aligned}$$

where μ is the Möbius function. Applying the relevant lemma above (depending on q), we get

$$\begin{aligned} \left| \sum_{M < n \leq M+N} \chi(n) \right| &\leq \sum_{t | k_2} \left| \sum_{M < ut \leq M+N} \chi(ut) \right| = \sum_{t | k_2} \left| \sum_{u=\frac{M+1}{t}}^{\frac{M+N}{t}} \chi(u) \right| \\ (4) \quad &\leq \begin{cases} 27 \cdot d(q)^{\frac{5}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))} & \text{for } 10^5 < q \leq \exp(\exp(8)), \\ 5 \cdot d(q)^{\frac{5}{2}} \sqrt{N} q^{\frac{3}{16}} \sqrt{\log(q) \log(\log(q))} & \text{for } \exp(\exp(8)) < q. \end{cases} \end{aligned}$$

For $q = p$ a prime, we have better explicit Burgess bounds (see also [6], [14] and [23])

Lemma 12 ($q = p$. Theorem 7.1 in [17]). *Suppose χ is a non-principal Dirichlet character modulo a prime $p \geq 2 \cdot 10^4$. For $M \in \mathbb{N}$ and $N \leq p^{\frac{5}{8}}$ we have*

$$\left| \sum_{M < n \leq M+N} \chi(n) \right| \leq 10.0366 \log(p)^{\frac{1}{4}} p^{\frac{3}{16}} \sqrt{N} = 10.0366 \sqrt[4]{C_\varepsilon} \sqrt{N} p^{\frac{3}{16} + \frac{\varepsilon}{4}},$$

with $C_\varepsilon = \log(p)/p^\varepsilon$.

Running the Python code in Appendix A up to $q = 20000$ yields the constant 3.41105405758246.

3. PROOFS

3.1. Proof of Theorem 4. Without loss of generality, assume for any fixed $0 < \eta < \frac{1}{8}$, that

$$\psi(q) \geq q^{-5/8+\eta} \quad \text{for all } q.$$

Otherwise, we replace ψ with $\tilde{\psi}(q) = \max\{\psi(q), q^{-5/8+\eta}\}$ and note that the associated series still converges. Note also, that $\mathcal{B}'(\tilde{\psi}, \kappa) \subseteq \mathcal{B}'(\psi, \kappa)$, so a lower bound on the measure of the former is automatically a lower bound on the measure of the set of interest. Write

$$\mathcal{B}'(\psi, \kappa)^c = \{x \in [0, 1] : \exists q \in \mathbb{N} \text{ s.t. } \|qx\|, \|qx^2\| \leq \kappa\psi(q)\}.$$

Note that if $\|qx\|, \|qx^2\| \leq \kappa\psi(q)$, there are integers $1 \leq a, b \leq q$ such that

$$\begin{aligned} |qx - a| &\leq \kappa\psi(q) \quad \text{and} \\ |qx^2 - b| &\leq \kappa\psi(q). \end{aligned}$$

Then

$$\begin{aligned} \left| \frac{a^2}{q^2} - \frac{b}{q} \right| &\leq \left| \frac{a^2}{q^2} - x^2 \right| + \left| x^2 - \frac{b}{q} \right| \\ &= \left| x - \frac{a}{q} \right| \left| x + \frac{a}{q} \right| + \left| x^2 - \frac{b}{q} \right| \leq 3\kappa\psi(q). \end{aligned}$$

From this we get, that

$$\begin{aligned} \lambda(\mathcal{B}'(\psi, \kappa)^c) &\leq \lambda \left(\bigcup_{q=1}^{\infty} \bigcup_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} \left[\frac{a}{q} - \frac{\kappa\psi(q)}{q}, \frac{a}{q} + \frac{\kappa\psi(q)}{q} \right] \right) \\ (5) \quad &\leq 2\kappa \sum_{q=1}^{\infty} \frac{\psi(q)}{q} \sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} 1 \end{aligned}$$

Our goal is now to estimate the counting function $\sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} 1$. This function roughly counts the number of rational points with fixed denominator q lying close to the parabola.

Put $J = \left\lfloor \frac{1}{6\kappa\psi(q)} \right\rfloor$, whenever $\psi(q) \neq 0$. Consider the Fejér kernel (see e.g. [18])

$$\mathcal{F}_J(x) := \sum_{j=-J}^J \frac{J-|j|}{J^2} \exp(2\pi i j x) = \frac{1}{J^2} \left(\frac{\sin(2\pi J x)}{\sin(2\pi x)} \right)^2 \geq 0 \quad (x \in \mathbb{R}).$$

Using the local bound

$$\frac{2}{\pi}|x| \leq |\sin(x)| \leq |x| \quad \text{for } |x| \leq \frac{\pi}{2},$$

we get for $\| \frac{a^2}{q} \| \leq 3\kappa\psi(q)$, that $\left| \pi \left\| \frac{a^2}{q} \right\| \right| \leq \left| J\pi \left\| \frac{a^2}{q} \right\| \right| \leq |3J\pi\kappa\psi(q)| \leq \frac{\pi}{2}$ and thus

$$\mathcal{F}_J \left(\frac{a^2}{q} \right) \geq \frac{1}{J^2} \left(\frac{\frac{2}{\pi} 2\pi J \left\| \frac{a^2}{q} \right\|}{2\pi \left\| \frac{a^2}{q} \right\|} \right)^2 = \frac{4}{\pi^2}.$$

This in turn gives us, that

$$\begin{aligned}
 \sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} 1 &\leq \frac{\pi^2}{4} \sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} \sum_{j=-J}^J \frac{J-|j|}{J^2} \exp\left(2\pi i j \frac{a^2}{q}\right) \\
 &\leq \frac{\pi^2}{4} \left[\sum_{j=1}^J \frac{J-j}{J^2} \sum_{a=1}^q \left(\exp\left(2\pi i j \frac{a^2}{q}\right) + \exp\left(-2\pi i j \frac{a^2}{q}\right) \right) + \frac{q}{J} \right] \\
 (6) \quad &\leq \frac{\pi^2}{4} \left[\sum_{d|q} d \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^{J/d} \frac{J-j_1 d}{J^2} \sum_{a=1}^{q_1} \left(\exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) \right) + \frac{q}{J} \right].
 \end{aligned}$$

To estimate this sum, we use Abel summation. Recall the following lemma.

Lemma 13 (Abel's identity, Thm. 4.2 in [2]). *Let $(a_n)_{n=0}^\infty$ be a sequence and for $N \in (0, \infty)$ ψ be a continuously differentiable function on $[0, N]$. Then*

$$\sum_{n=0}^N a_n \phi(n) = \sum_{n=0}^N a_n \phi(N) - \int_0^N \sum_{n=0}^u a_n \phi'(u) du.$$

Taking in our case $N = \frac{J}{d}$, $a_{j_1} = \mathbb{1}_{\gcd(j_1, q_1)=1} \sum_{a=1}^{q_1} \left(\exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) - \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) \right)$ and $\phi(j_1) = \frac{J-j_1 d}{J^2}$ and using $\phi(\frac{J}{d}) = \frac{J-\frac{J}{d}d}{J^2} = 0$ and $\phi'(u) = \frac{-d}{J^2}$ yields

$$(6) = \frac{\pi^2}{4} \left[\sum_{d|q} \frac{d^2}{J^2} \int_1^{J/d} \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^u \sum_{a=1}^{q_1} \left(\exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) \right) du + \frac{q}{J} \right].$$

As in [13], to evaluate the quadratic Gauss sums, we state the following lemma, originally from [14] §3.5.

Lemma 14. *Suppose $\gcd(j_1, q_1) = 1$, then*

$$\sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) = \begin{cases} 0 & \text{when } q_1 \equiv 2 \pmod{4}, \\ \left(\frac{j_1}{q_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 1 \pmod{4}, \\ i \left(\frac{j_1}{q_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 3 \pmod{4}, \\ (1+i) \left(\frac{q_1}{j_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 0 \pmod{4} \text{ and } j_1 \equiv 1 \pmod{4}, \\ (1+i)^{\frac{1}{i}} \left(\frac{q_1}{j_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 0 \pmod{4} \text{ and } j_1 \equiv 3 \pmod{4}. \end{cases}$$

Here $(\cdot)$ is the Jacobi symbol.

This immediately gives us,

$$\begin{aligned}
& \sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \overline{\sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right)} \\
&= \begin{cases} 0 & \text{when } q_1 \equiv 2 \pmod{4}, \\ \left(\frac{j_1}{q_1}\right) \sqrt{q_1} + \overline{\left(\frac{j_1}{q_1}\right) \sqrt{q_1}} & \text{when } q_1 \equiv 1 \pmod{4}, \\ i \left(\frac{j_1}{q_1}\right) \sqrt{q_1} + i \overline{\left(\frac{j_1}{q_1}\right) \sqrt{q_1}} & \text{when } q_1 \equiv 3 \pmod{4}, \\ (1+i) \left(\frac{q_1}{j_1}\right) \sqrt{q_1} + (1+i) \overline{\left(\frac{q_1}{j_1}\right) \sqrt{q_1}} & \text{when } q_1 \equiv 0 \pmod{4} \text{ and } j_1 \equiv 1 \pmod{4}, \\ (1+i)^{\frac{1}{i}} \left(\frac{q_1}{j_1}\right) \sqrt{q_1} + (1+i)^{\frac{1}{i}} \overline{\left(\frac{q_1}{j_1}\right) \sqrt{q_1}} & \text{when } q_1 \equiv 0 \pmod{4} \text{ and } j_1 \equiv 3 \pmod{4}, \end{cases} \\
(8) \quad &= \begin{cases} 0 & \text{when } q_1 \equiv 2, 3 \pmod{4}, \\ 2 \left(\frac{j_1}{q_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 1 \pmod{4}, \\ 2 \left(\frac{q_1}{j_1}\right) \sqrt{q_1} & \text{when } q_1 \equiv 0 \pmod{4}. \end{cases}
\end{aligned}$$

The Legendre symbols, $\left(\frac{\cdot}{q_1}\right)$ and $\left(\frac{q_1}{\cdot}\right)$, are Dirichlet characters of modulus at most q_1 resp. $4q_1$.

Lemma 15 (Divisor function bound (from Terence Tao's blog [22])). *For any $\beta > 0$, there is an explicit constant $K_\beta > 0$ such that*

$$d(q) \leq K_\beta q^\beta,$$

Specifically we have $K_{\frac{5}{11}} \leq \left(\frac{11}{5 \log(2)}\right)^4$ and $K_{\frac{1}{57}} \leq \left(\frac{57}{\log(2)}\right)^{176846309399143769411680}$.

We will repeat Tao's proof below.

Proof. If q has the prime factorisation $q = p_1^{a_1} \cdots p_k^{a_k}$, then

$$\frac{d(q)}{q^\beta} = \prod_{j=1}^k \frac{d(p_j^{a_j})}{p_j^{a_j \beta}} = \prod_{j=1}^k \frac{a_j + 1}{p_j^{a_j \beta}}$$

The numerator is linear, while the denominator is exponential in a_j , so one would expect the denominator to dominate. However, β might be small enough that this only happens once p_j reaches a certain size. We therefore consider small and large primes separately.

If p_j is large, say $p_j \geq e^{\frac{1}{\beta}}$, then

$$p_j^{a_j \beta} \geq e^{a_j} \geq a_j + 1$$

so indeed the denominator dominates the numerator, and the term contributes ≤ 1 to the product, and may be ignored. For a fixed $\beta > 0$ we need only consider primes $p_j < e^{\frac{1}{\beta}}$.

By Taylor expansion

$$p_j^{a_j \beta} = e^{a_j \beta \log(p_j)} \geq 1 + \beta a_j \log(p_j)$$

so

$$(9) \quad \frac{a_j + 1}{p_j^{a_j \beta}} \leq \frac{a_j + 1}{1 + \beta a_j \log(p_j)} \leq \frac{1}{\beta \log(2)}.$$

So in the end we get

$$d(q) \leq q^\beta \left(\frac{1}{\beta \log(2)} \right)^{\pi(e^{\frac{1}{\beta}})},$$

where $\pi(e^{\frac{1}{\beta}})$ denotes the number of primes smaller than $e^{\frac{1}{\beta}}$. We will need it for $\beta = \frac{5}{11}$ and $\beta = \frac{1}{57}$, where $\pi(e^{\frac{11}{5}}) = \pi(9) = 4$ and $\pi(e^{\frac{57}{1}}) \leq \pi(10^{25}) = 176846309399143769411680$ ([9]). For β very small use the crude bound $\pi(e^{\frac{1}{\beta}}) < e^{\frac{1}{\beta}}$.

Note that in (9) we technically use that $\beta \leq \frac{1}{\log(2)}$, but this is okay since we are interested in the result for small β . For large β use e.g. the trivial bound $d(q) \leq q$, which immediately yields the result for $\beta \geq 1$. $\square$

Lemma 16 (Log bound). *For any $\varepsilon > 0$*

$$\log(q) \leq C_\varepsilon q^\varepsilon,$$

where

$$C_\varepsilon = \frac{1}{\varepsilon e}.$$

Proof. This is a simple matter of maximising $\frac{\log(q)}{q^\varepsilon}$:

$$\frac{d}{dq} \frac{\log(q)}{q^\varepsilon} = \frac{1 - \varepsilon \log(q)}{q^{1+\varepsilon}} = 0$$

for $q = e^{\frac{1}{\varepsilon}}$. Plugging in this value in gives us the maximum of $\frac{\log(q)}{q^\varepsilon}$:

$$\frac{\log(q)}{q^\varepsilon} \leq \frac{1}{\varepsilon e} \quad \forall q.$$

□

Combining our explicit Burgess bound (see Section 2) with (8), we prove a version of Huang's Lemma 3 from [13]:

Lemma 17. *For any $\varepsilon_1, \varepsilon_2 > 0$*

$$\sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^N \sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) \leq \begin{cases} 2N\sqrt{q_1} & \text{if } q_1 \text{ is a square,} \\ 2 \cdot 4.19253564643679 \cdot \sqrt{N} q_1^{\frac{11}{16}} & \text{if } q_1 \text{ is not a square and } 1 \leq q_1 \leq 10^5, \\ 2 \cdot 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_2}{2}} d(4q_1)^{\frac{5}{2}} C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} \sqrt{N} q_1^{\frac{11}{16} + \frac{\varepsilon_1 + \varepsilon_2}{2}} & \text{if } q_1 \text{ is not a square and } 10^5 < q_1 \leq \exp(\exp(8)), \\ 2 \cdot 5 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_2 + \varepsilon_2}{2}} d(4q_1)^{\frac{5}{2}} C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} \sqrt{N} q_1^{\frac{11}{16} + \frac{\varepsilon_2 + \varepsilon_2}{2}} & \text{if } q_1 \text{ is not a square and } \exp(\exp(8)) < q_1. \end{cases}$$

Proof. If q_1 is a square, then applying Lemma 14 gives us

$$\begin{aligned} \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^N \sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) &\leq 2\sqrt{q_1} \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^N \chi(j_1) \\ &\leq 2\sqrt{q_1} N, \end{aligned}$$

where $\chi(j_1) = \left(\frac{j_1}{q_1}\right)$ if $q_1 \equiv 1 \pmod{4}$ and $\chi(j_1) = \left(\frac{q_1}{j_1}\right)$ if $q_1 \equiv 0 \pmod{4}$.

For q_1 not a square and $q_1 > 10^5$, using Lemma 14, we have

$$\sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^N \sum_{a=1}^{q_1} \exp\left(2\pi i j_1 \frac{a^2}{q_1}\right) + \exp\left(-2\pi i j_1 \frac{a^2}{q_1}\right) \leq 2\sqrt{q_1} \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^N \chi(j_1),$$

where $\chi(j_1) = \left(\frac{j_1}{q_1}\right)$ if $q_1 \equiv 1 \pmod{4}$, $\chi(j_1) = \left(\frac{q_1}{j_1}\right)$ if $q_1 \equiv 0 \pmod{4}$ and is zero otherwise. Either way χ is a Dirichlet character of modulus $\leq 4q_1$. Applying Remark 11 as well as Lemma 16, we get for q_1 not a square and $10^5 < q_1 \leq \exp(\exp(8))$:

$$\leq 2 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_1}{2}} 27 C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} \sqrt{N} d(4q_1)^{\frac{5}{2}} q_1^{\frac{11}{16} + \frac{\varepsilon_1 + \varepsilon_1}{2}}$$

and for q_1 not a square and $\exp(\exp(8)) < q_1$:

$$\leq 2 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_2 + \varepsilon_2}{2}} 5 C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} \sqrt{N} d(4q_1)^{\frac{5}{2}} q_1^{\frac{11}{16} + \frac{\varepsilon_2 + \varepsilon_2}{2}}.$$

For q_1 not a square and $1 \leq q_1 \leq 10^5$ apply Lemma 14 and then Lemma 9. □

We now split the sum over $d \mid q$ in (7) into the four cases,

- (1) $q = q_1 d$ and q_1 is a square, denoted $q_1 = \square$ below,

- (2) $q = q_1 d$, q_1 is not a square, denoted $q_1 \neq \square$, and $1 \leq q_1 \leq 10^5$,
- (3) $q = q_1 d$, q_1 is not a square, and $10^5 < q_1 \leq \exp(\exp(8))$ and
- (4) $q = q_1 d$, q_1 is not a square, and $\exp(\exp(8)) < q_1$.

We then apply Lemma 17, and in the case $q = q_1 d$, q_1 is not a square, and $\exp(\exp(8)) < q_1$ Lemma 15 for some $\beta_2 > 0$ to obtain

$$\begin{aligned}
(7) &\leq \frac{\pi^2}{4} \left[2 \sum_{q=q_1 d, q_1=\square} \frac{d^2}{J^2} \sqrt{q_1} \int_1^{\frac{J}{d}} u \, du + 2 \cdot 4.19253564643679 \sum_{q=q_1 d, q_1 \neq \square, 1 \leq q_1 \leq 10^5} \frac{d^2}{J^2} q_1^{\frac{11}{16}} \int_1^{\frac{J}{d}} \sqrt{u} \, du \right. \\
&\quad + 2 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} 27 C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} \sum_{q=q_1 d, q_1 \neq \square, 10^5 < q_1 \leq \exp(\exp(8))} \frac{d^2}{J^2} q_1^{\frac{11}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} d(4q_1)^{\frac{5}{2}} \int_1^{\frac{J}{d}} \sqrt{u} \, du \\
&\quad \left. + 2 \cdot 4^{\frac{3}{16} + \frac{5\beta_2}{2} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} 5K_{\beta_2}^{\frac{5}{2}} C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} \sum_{q=q_1 d, q_1 \neq \square, \exp(\exp(8)) < q_1} \frac{d^2}{J^2} q_1^{\frac{11}{16} + \frac{5\beta_2}{2} + \frac{\varepsilon_2 + \varepsilon_2^2}{2}} \int_1^{\frac{J}{d}} \sqrt{u} \, du + \frac{q}{J} \right] \\
&\leq \frac{\pi^2}{4} \left[\sum_{q=q_1 d, q_1=\square} \sqrt{q_1} \frac{d^2}{J^2} \frac{J^2}{d^2} + 2 \cdot 4.19253564643679 \frac{2}{3} \sum_{q=q_1 d, q_1 \neq \square, 1 \leq q_1 \leq 10^5} q_1^{\frac{11}{16}} \frac{d^2}{J^2} \left(\frac{J}{d} \right)^{\frac{3}{2}} \right. \\
&\quad + 2 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} 27 C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} \frac{2}{3} \sum_{q=q_1 d, q_1 \neq \square, 10^5 < q_1 \leq \exp(\exp(8))} q_1^{\frac{11}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} d(4q_1)^{\frac{5}{2}} \frac{d^2}{J^2} \left(\frac{J}{d} \right)^{\frac{3}{2}} \\
&\quad \left. + 2 \cdot 4^{\frac{3}{16} + \frac{5\beta_2}{2} + \frac{\varepsilon_2 + \varepsilon_2^2}{2}} 5K_{\beta_2}^{\frac{5}{2}} C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} \frac{2}{3} \sum_{q=q_1 d, q_1 \neq \square, \exp(\exp(8)) < q_1} q_1^{\frac{11}{16} + \frac{5\beta_2}{2} + \frac{\varepsilon_2 + \varepsilon_2^2}{2}} \frac{d^2}{J^2} \left(\frac{J}{d} \right)^{\frac{3}{2}} + \frac{q}{J} \right].
\end{aligned}
\tag{10}$$

We have $d(4q_1) \leq 3d(q_1)$. In the case $q = q_1 d$, q_1 is not a square, and $10^5 < q_1 \leq \exp(\exp(8))$, applying Lemma 18 to $d(q_1)$ yields (for $10^5 \leq q_1 \leq \exp(\exp(8))$):

$$\log(d(q_1)) \leq \frac{\log(2) \log(q)}{\log(\log(q)) - 1.39177} \leq \frac{\log(2)e^8}{8 - 1.39177},$$

so

$$d(4q_1)^{\frac{5}{2}} \leq 3^{\frac{5}{2}} e^{\frac{5 \log(2)e^8}{2(8-1.39177)}}.$$

Employing the bounds

$$\sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ q_1 \leq 10^5}} 1 \leq 99684, \quad \sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ 10^5 < q_1 \leq \exp(\exp(8))}} 1 \leq (\exp(\exp(8)) - 10^5), \quad \text{and} \quad \sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ \exp(\exp(8)) < q_1}} 1 \leq d(q) \leq K_{\beta_2} q^{\beta_2}.$$

gives us the following upper bound for the counting function:

$$\begin{aligned}
(10) &\leq \frac{\pi^2}{4} \left[\sum_{q=q_1 d, q_1=\square} \sqrt{q_1} + 2 \cdot 4.19253564643679 \frac{2}{3} J^{-\frac{1}{2}} q^{\frac{11}{16}} 99684 \right. \\
&\quad + 2 \cdot 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} \frac{2}{3} J^{-\frac{1}{2}} q^{\frac{11}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} 3^{\frac{5}{2}} e^{\frac{5 \log(2)e^8}{2(8-1.39177)}} (\exp(\exp(8)) - 10^5) \\
&\quad \left. + 2 \cdot 5 \cdot 4^{\frac{3}{16} + \frac{5\beta_2}{2} + \frac{\varepsilon_2 + \varepsilon_2^2}{2}} K_{\beta_2}^{\frac{7}{2}} C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} \frac{2}{3} J^{-\frac{1}{2}} q^{\frac{11}{16} + \frac{7\beta_2}{2} + \frac{\varepsilon_2 + \varepsilon_2^2}{2}} + \frac{q}{J} \right]
\end{aligned}$$

For r the biggest square dividing q , applying Lemma 15 for some $\beta_3 > 0$ yields

$$\sum_{\substack{q=q_1 d \\ q_1=\square}} \sqrt{q_1} = \sum_{q=r_1^2 d} r_1 \leq r \sum_{r_1|r} 1 = rd(r) \leq r^{1+\beta_3} K_{\beta_3}.$$

Furthermore, since $J = \left\lfloor \frac{1}{6\kappa\psi(q)} \right\rfloor (\neq 0)$, we have $J^{-1} \leq 12\kappa\psi(q)$. So

$$\begin{aligned} \lambda(\mathcal{B}'(\psi, \kappa)^c) &\leq \kappa K_{\beta_3} \frac{\pi^2}{2} \sum_{q=1}^{\infty} \frac{\psi(q)}{q} r^{1+\beta_3} \\ &\quad + \kappa^{\frac{3}{2}} \left(\frac{2\pi^2 4.19253564643679 \sqrt{1299684}}{3} \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16}} \right. \\ &\quad + \frac{2\pi^2 4^{\frac{3}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} 27 \sqrt{12} C_{\varepsilon_1}^{1 + \frac{\varepsilon_1}{2}} 3^{\frac{5}{2}} e^{\frac{5 \log(2) e^8}{2(8-1.39177)}} (\exp(\exp(8)) - 10^5)}{3} \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}} \\ &\quad \left. + \frac{2\pi^2 4^{\frac{3}{16} + \frac{\varepsilon_2 + \varepsilon_2^2}{2} + \frac{5\beta_2}{2}} 5 \sqrt{12} C_{\varepsilon_2}^{1 + \frac{\varepsilon_2}{2}} K_{\beta_2}^{\frac{7}{2}}}{3} \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\varepsilon_2 + \varepsilon_2^2}{2} + \frac{7\beta_2}{2}} \right) \\ &\quad + \kappa^2 6\pi^2 \sum_{q=1}^{\infty} \psi(q)^2. \end{aligned}$$

Choose $\beta_2 = \frac{1}{57}$, $\beta_3 = \frac{5}{11}$, $\varepsilon_1 = 0.11237$ and $\varepsilon_2 = 0.00218819$. For the first series, we have

$$\begin{aligned} \sum_{q=1}^{\infty} \frac{\psi(q)}{q} r^{1+\beta_3} &\leq \sqrt{\sum_{q=1}^{\infty} \psi(q)^2} \sqrt{\sum_{q=1}^{\infty} r^{2(1+\beta_3)} q^{-2}} \\ &\leq \sqrt{S_{\psi}} \sqrt{\sum_{r=1}^{\infty} \sum_{t=1}^{\infty} |\mu(t)| r^{2 + \frac{10}{11}} (r^2 t)^{-2}} \\ &\leq \sqrt{S_{\psi}} \sqrt{\sum_{r=1}^{\infty} r^{-\frac{12}{11}} \sum_{t=1}^{\infty} t^{-2}} \\ &\leq \sqrt{S_{\psi}} \sqrt{\zeta\left(\frac{12}{11}\right) \sqrt{\frac{\pi^2}{6}}}, \end{aligned}$$

where μ is the Möbius function. Note that we estimate $|\mu(t)|$ by 1 here. A slightly sharper estimate would follow if one only considered square free t , since the function is 0 elsewhere, but the improvement is so marginal that we decided not to pursue this further. For the second, third and fourth, we have

$$\sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16}}, \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\varepsilon_1 + \varepsilon_1^2}{2}}, \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\varepsilon_2 + \varepsilon_2^2}{2} + \frac{7\beta_2}{2}} \leq S_{\psi}$$

since $\psi(q) \geq q^{\frac{-5}{8} + \eta}$. So $\lambda(\mathcal{B}'(\psi, \kappa)^c)$ is

$$\begin{aligned} &\leq \kappa \frac{\pi^2 \sqrt{\zeta\left(\frac{12}{11}\right)} \sqrt{\frac{\pi^2}{6}} \left(\frac{11}{5 \log(2)}\right)^4}{2} \sqrt{S_{\psi}} \\ &\quad + \kappa^{\frac{3}{2}} \left(\frac{2\pi^2 \sqrt{1299684} \cdot 4.192423564643679}{3} \right. \\ &\quad + \frac{2\pi^2 27 \sqrt{12} \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} e^{\frac{5 \log(2) e^8}{2(8-1.39177)}} (e^{e^8} - 10^5) \left(\frac{1}{e \cdot 0.11237}\right)^{1 + \frac{0.11237}{2}}}{3} \\ &\quad \left. + \frac{2\pi^2 5 \sqrt{12} \cdot 4^{\frac{3}{16} + \frac{0.00218819 + 0.00218819^2}{2} + \frac{5}{2 \cdot 57}} \left(\frac{57}{\log(2)}\right)^{\frac{7}{2}} 176846309399143769411680 \left(\frac{1}{e \cdot 0.00218819}\right)^{1 + \frac{0.00218819}{2}}}{3} \right) S_{\psi} \\ &\quad + \kappa^2 6\pi^2 S_{\psi} \end{aligned}$$

which is $< \delta$ for

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{\delta}{\sqrt{S_\psi}} \frac{2\sqrt{6}(5\log(2))^4}{3\pi^3\sqrt{\zeta\left(\frac{12}{11}\right)}11^4}, \sqrt{\frac{\delta}{S_\psi}} \frac{1}{\sqrt{18\pi}}, \right. \\ &\quad \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} \left(\frac{1}{2\pi^2\sqrt{12}} \right)^{\frac{2}{3}} \left(99684 \cdot 4.19253564643679 \right. \\ &\quad \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(e^{e^8} - 10^5 \right) e^{\frac{5\log(2)e^8}{2(8-1.39177)}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} \right. \\ &\quad \left. + 5 \cdot 4^{\frac{3}{16} + \frac{0.00218819 + 0.00218819^2}{2} + \frac{5}{2 \cdot 57}} \left(\frac{57}{\log(2)} \right)^{\frac{7}{2} \cdot 176846309399143769411680} \left(\frac{1}{e \cdot 0.00218819} \right)^{1 + \frac{0.00218819}{2}} \right)^{\frac{-2}{3}} \Bigg\} \\ &\simeq \min \left\{ \frac{\delta}{\sqrt{S_\psi}} 0.00015248265228, \sqrt{\frac{\delta}{S_\psi}} 0.07502635967975, \left(\frac{\delta}{S_\psi} \right)^{\frac{2}{3}} 10^{-10^{23.89775276641198}} \right\}. \end{aligned}$$

□

Remark. The constant $\sim 10^{-10^{23.89775276641198}}$ is very small. The main problem here is our divisor function bound from Lemma 15. There are two ways to address the problem. Large values of $d(q)$ occur at smooth numbers, i.e. numbers with only small prime divisors, so one could consider e.g., for a fixed n , only q of the form $q = p_1 p_2 \dots p_n$, so $d(q) \leq 2^n$. Alternatively by looking at only very large values of q , we also get a more reasonable constant in our divisor function bound. The former is explored in Theorem 6 and Theorem 7 (for $n = 1$), while latter is Theorem 5.

3.2. Proof of Theorem 5. If we consider only very large q , we can improve Lemmas 15 and 16 as follows:

Lemma 18 (From Robin's thesis [21], also mentioned in the survey paper [19]). For $q \geq 56$

$$\log(d(q)) \leq \frac{\log(q) \log(2)}{\log(\log(q)) - 1.39177}$$

In particular, for $q \geq \exp(\exp(41))$, we get $d(q) \leq q^{\frac{1}{57}}$.

Lemma 19. For $q \geq \exp(\exp(41))$

$$\log(q) \leq q^{\frac{41}{e^{41}}}.$$

Proof. We have $\log(q) \leq q^{\frac{41}{e^{41}}}$ if and only if $\frac{41}{e^{41}} \geq \frac{\log(\log(q))}{\log(q)}$.

□

Aside from having these improved lemmas, the proof is exactly the same as before, only with K_β and C_ε both replaced by 1 and with β and ε replaced by $\frac{1}{57}$ respectively $\frac{41}{e^{41}}$. Picking up from (7) we have

$$\begin{aligned} \sum_{\substack{a \leq q \\ \|a^2/q\| \leq 3\kappa\psi(q)}} 1 &\leq \frac{\pi^2}{4} \left[\sum_{d|q} \frac{d^2}{J^2} \int_1^{\frac{J}{d}} \sum_{\substack{j_1=1 \\ \gcd(j_1, q_1)=1}}^u \sum_{a=1}^{q_1} e^{2\pi i j_1 \frac{a^2}{q_1}} + e^{-2\pi i j_1 \frac{a^2}{q_1}} du + \frac{q}{J} \right] \\ &\leq \frac{\pi^2}{4} \left[\sum_{\substack{q=q_1 d \\ q_1=\square}} \frac{d^2}{J^2} \int_1^{\frac{J}{d}} 2\sqrt{q_1} u \, du \right. \\ &\quad + \sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ 1 \leq q_1 \leq 10^5}} \frac{d^2}{J^2} \int_1^{\frac{J}{d}} 2 \cdot 4.19253564643679 q_1^{\frac{11}{16}} \sqrt{u} \, du \\ &\quad \left. + \sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ 10^5 < q_1 \leq \exp(\exp(8))}} \frac{d^2}{J^2} \int_1^{\frac{J}{d}} 2 \cdot 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon + \varepsilon^2}{2}} d(4q_1)^{\frac{5}{2}} q_1^{\frac{11}{16} + \frac{\varepsilon + \varepsilon^2}{2}} \sqrt{u} \, du + \frac{q}{J} \right] \end{aligned}$$

This time we use the estimates

$$\sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ 1 \leq q_1 \leq 10^5}} 1, \sum_{\substack{q=q_1 d \\ q_1 \neq \square \\ 10^5 < q_1}} 1 \leq d(q) < q^{\frac{1}{57}},$$

as well as

$$q^{\frac{1}{57}} q_1^{\frac{11}{16}} \sqrt{d} \leq q^{\frac{11}{16} + \frac{1}{57}}$$

and

$$q_1^{\frac{11}{16} + \frac{7}{2 \cdot 57} + \frac{\frac{41}{e^{41}} + \left(\frac{41}{e^{41}}\right)^2}{2}} \sqrt{d} \leq q^{\frac{11}{16} + \frac{7}{2 \cdot 57} + \frac{\frac{41}{e^{41}} + \left(\frac{41}{e^{41}}\right)^2}{2}}.$$

Apply also

$$\sum_{\substack{q=q_1 d \\ q_1=\square}} \sqrt{q_1} = \sum_{q=r_1^2 d} r_1 \leq r \sum_{r_1|r} 1 \leq r d(q) \leq r q^{\frac{1}{57}},$$

where r is the largest square dividing q . Putting everything together, exactly like in the proof of Theorem 4, we get

$$\begin{aligned} \lambda(\mathcal{B}'(\kappa, \psi)^c) &\leq \kappa \frac{\pi^2}{2} \sum_{q=1}^{\infty} \psi(q) q^{-1} r^{\frac{58}{57}} \\ &\quad + \kappa^{\frac{3}{2}} \frac{2\pi^2 \sqrt{12}}{3} \left(4.19253564643679 \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{1}{57}} \right. \\ &\quad \left. + 27 \cdot 4^{\frac{3}{16} + \frac{\frac{41}{e^{41}} + \left(\frac{41}{e^{41}}\right)^2}{2} + \frac{5}{2 \cdot 57}} \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\frac{41}{e^{41}} + \left(\frac{41}{e^{41}}\right)^2}{2} + \frac{7}{2 \cdot 57}} \right) \\ &\quad + \kappa^2 6\pi^2 \sum_{q=1}^{\infty} \psi(q)^2. \end{aligned}$$

For the first series we have

$$\begin{aligned} \sum_{q=1}^{\infty} \psi(q) q^{-1} r^{\frac{58}{57}} &\leq \sqrt{\sum_{q=1}^{\infty} \psi(q)} \sqrt{\sum_{r=1}^{\infty} \sum_{t=1}^{\infty} |\mu(t)| r^{2\frac{58}{57}} (r^2 t)^2} \\ &\leq \sqrt{S_{\psi}} \sqrt{\sum_{r=1}^{\infty} r^{-\frac{112}{57}} \sum_{t=1}^{\infty} t^{-2}} \\ &\leq \sqrt{S_{\psi}} \sqrt{\zeta\left(\frac{112}{57}\right) \frac{\pi^2}{6}}. \end{aligned}$$

Since $\psi(q) \geq q^{-\frac{5}{8}+\eta}$, the other three series are $\leq S_{\psi}$. Thus

$$\begin{aligned} \lambda(\mathcal{B}'(\kappa, \psi)^c) &\leq \kappa \frac{\pi^2}{2} \sqrt{S_{\psi}} \sqrt{\frac{\pi^2}{6} \zeta\left(\frac{112}{57}\right)} + \kappa^2 6\pi^2 S_{\psi} \\ &\quad + \kappa^{\frac{3}{2}} \frac{2\pi^2 \sqrt{12}}{3} S_{\psi} \left(4.19253564643679 + 27 \cdot 4^{\frac{3}{16} + \frac{41}{2 \cdot 57} + \frac{(\frac{41}{e^{41}})^2}{2}} + \frac{5}{2 \cdot 57} \right). \end{aligned}$$

This is $< \delta$ for

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{\delta}{\sqrt{S_{\psi}}} \frac{2}{3\pi^2 \sqrt{\frac{\pi^2}{6} \zeta\left(\frac{112}{57}\right)}}, \sqrt{\frac{\delta}{S_{\psi}}} \sqrt{\frac{1}{18\pi^2}} \right. \\ &\quad \left. \left(\frac{\delta}{S_{\psi}} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left(4.19253564643679 + 27 \cdot 4^{\frac{3}{16} + \frac{5}{2 \cdot 57} + \frac{41}{e^{41}} + \frac{(\frac{41}{e^{41}})^2}{2}} \right) \right)^{\frac{-2}{3}} \right\} \\ &\simeq \min \left\{ \frac{1}{6 \max_{n \in \mathbb{N}} \psi(n)}, \frac{\delta}{\sqrt{S_{\psi}}} 0.04064400, \sqrt{\frac{\delta}{S_{\psi}}} 0.07502635, \left(\frac{\delta}{S_{\psi}} \right)^{\frac{2}{3}} 0.00499238 \right\}. \end{aligned}$$

□

Remark. The size of the constants in $\kappa(\delta)$ are much more managable here, but it comes at the cost of assuming $q > \exp(\exp(41))$, which is again very large. Once more, this is due to our divisor function bound!

3.3. Proof of Theorem 6. Since q has at most n prime divisors, we have $d(q) \leq 2^n$. As in §3.1 equation (10) (but not using the divisor function bound from §3.1) we have

$$\begin{aligned} \sum_{\substack{a \leq q \\ \|\frac{a^2}{q}\| \leq 3\kappa\psi(q)}} 1 &\leq \frac{\pi^2}{4} \left[\sum_{\substack{q=q_1 d \\ q_1 = \square}} \sqrt{q_1} + 2 \cdot 4.192535.64643679 \cdot \frac{2}{3} J^{-\frac{1}{2}} q^{\frac{11}{16}} 99684 \right. \\ &\quad \left. + 2 \cdot 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon + \varepsilon^2}{2}} \left(\frac{1}{e\varepsilon} \right)^{1 + \frac{\varepsilon}{2}} \frac{2}{3} d(4q)^{\frac{5}{2}} d(q) q^{\frac{11}{16} + \frac{\varepsilon + \varepsilon^2}{2}} J^{-\frac{1}{2}} + \frac{q}{J} \right]. \end{aligned}$$

Using $d(q) \leq 2^n$, $J^{-1} \leq 12\kappa\psi(q)$, and

$$\sum_{\substack{q=q_1 d \\ q_1 = \square}} \sqrt{q_1} \leq rd(r) \leq rd(q) \leq r2^n$$

gives us

$$\begin{aligned} &\leq \frac{\pi^2}{4} \left[r2^n + \frac{4 \cdot 4.19253564643679 \sqrt{12} 99684}{3} \sqrt{\kappa} \sqrt{\psi(q)} q^{\frac{11}{16}} \right. \\ &\quad \left. + \frac{4 \cdot 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon + \varepsilon^2}{2}} \sqrt{12} 3^{\frac{5}{2}} 2^{\frac{7n}{2}} \left(\frac{1}{e\varepsilon} \right)^{1 + \frac{\varepsilon}{2}}}{3} \sqrt{\kappa} \sqrt{\psi(q)} q^{\frac{11}{16} + \frac{\varepsilon + \varepsilon^2}{2}} + 12\kappa\psi(q) \right]. \end{aligned}$$

Plugging into (5)

$$\begin{aligned} \lambda(\mathcal{B}'(\kappa, \psi)^c) &\leq \kappa \frac{\pi^2}{2} 2^n \sum_{q=1}^{\infty} \frac{\psi(q)}{q} r \\ &\quad + \kappa^{\frac{3}{2}} \frac{2\pi^2 \sqrt{12}}{3} \left(4.19253564643679 \cdot 99684 \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16}} \right. \\ &\quad \left. + 27 \cdot 4^{\frac{3}{16} + \frac{\varepsilon + \varepsilon^2}{2}} 3^{\frac{5}{2}} \left(\frac{1}{e\varepsilon} \right)^{1 + \frac{\varepsilon}{2}} 2^{\frac{7n}{2}} \sum_{q=1}^{\infty} \psi(q)^{\frac{3}{2}} q^{\frac{-5}{16} + \frac{\varepsilon + \varepsilon^2}{2}} \right) \\ &\quad + \kappa^2 6\pi^2 \sum_{q=1}^{\infty} \psi(q)^2. \end{aligned}$$

Let $\varepsilon = 0.11237$, then

$$\sum_{q=1}^{\infty} \frac{\psi(q)}{q} r \leq \sqrt{\sum_{q=1}^{\infty} \psi(q)^2} \sqrt{\sum_{r=1}^{\infty} \sum_{t=1}^{\infty} r^{-2} t^{-2}} \leq \sqrt{S_{\psi}} \frac{\pi^2}{6},$$

and the other sums are $\leq S_{\psi}$. All together $\lambda(\mathcal{B}'(\kappa, \psi)^c)$ is $< \delta$ for

$$\begin{aligned} \kappa(\delta) &< \min \left\{ \frac{\delta}{\sqrt{S_{\psi}}} \frac{4}{3\pi^4} \frac{1}{2^n}, \left(\frac{\delta}{S_{\psi}} \right)^{\frac{2}{3}} \left(2\pi^2 \sqrt{12} \left[4.19253564643679 \cdot 99684 \right. \right. \right. \\ &\quad \left. \left. + 27 \cdot 4^{\frac{3}{16} + \frac{0.11237 + 0.11237^2}{2}} 3^{\frac{5}{2}} \left(\frac{1}{e \cdot 0.11237} \right)^{1 + \frac{0.11237}{2}} 2^{\frac{7n}{2}} \right] \right)^{\frac{-2}{3}}, \sqrt{\frac{\delta}{S_{\psi}}} \sqrt{\frac{1}{18\pi^2}} \right\} \\ &\simeq \min \left\{ \frac{\delta}{\sqrt{S_{\psi}}} 0.0410639 \frac{1}{2^n}, \left(\frac{\delta}{S_{\psi}} \right)^{\frac{2}{3}} 0.0598025 \left(417928.72 + 2082.9292 \cdot 2^{\frac{7n}{2}} \right)^{\frac{-2}{3}}, \sqrt{\frac{\delta}{S_{\psi}}} 0.0750263 \right\}. \end{aligned}$$

□

3.4. Proof of Theorem 7. We go a bit more into detail here, since $q = p$ being a prime affects a lot. As in §3.1 (6), we have

$$\sum_{\substack{a \leq p \\ \left\| \frac{a^2}{p} \right\| \leq 3\kappa\psi(p)}} 1 \leq \frac{\pi^2}{4} \left(\sum_{j=1}^J \frac{J-j}{J^2} \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right) + \frac{p}{J} \right).$$

If $J < p$, then for $1 \leq j \leq J$, $\gcd(j, p)$ is guaranteed to be 1, but if $J \geq p$, there is a chance, that $1 \leq j \leq J$ is a multiple of p . For $J \geq p$ we have

$$\begin{aligned} &\sum_{j=1}^J \frac{J-j}{J^2} \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right) + \frac{p}{J} \\ &\leq 1 \cdot \sum_{\substack{j=1 \\ \gcd(j,p)=1}}^J \frac{J-j}{J^2} \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right) + p \cdot \sum_{\substack{j_1=1 \\ \gcd(j_1, \frac{p}{p})=1}}^{\frac{J}{p}} \frac{J-j_1 p}{J^2} \sum_{a=1}^{\frac{p}{p}} 2 + \frac{p}{J} \\ &\leq \sum_{\substack{j=1 \\ \gcd(j,p)=1}}^J \frac{J-j}{J^2} \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right) + 1. \end{aligned}$$

In either case, to the expression

$$\sum_{\substack{j=1 \\ \gcd(j,p)=1}}^J \frac{J-j}{J^2} \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right)$$

apply first Abel's identity (Lemma 13)

$$= \int_1^J \frac{1}{J^2} \sum_{\substack{j=1 \\ \gcd(j,p)=1}}^u \sum_{a=1}^p \left(e^{2\pi i j \frac{a^2}{p}} + e^{-2\pi i j \frac{a^2}{p}} \right) du$$

then apply Lemma 14

$$\leq \frac{2}{J^2} \sqrt{p} \int_1^J \sum_{\substack{j=1 \\ \gcd(j,p)=1}}^u \left(\frac{j}{p} \right) du.$$

To estimate the sum of Dirichlet characters, we apply Lemma 12 as well as Lemma 16 for $\varepsilon = 0.249$ (Since $\frac{10.0366}{(e \cdot 0.249)^{\frac{1}{4}}} \geq 3.41105405758246$, we use the bigger constant in both the $p \leq 2 \cdot 10^4$ and the $2 \cdot 10^4 < p$ case):

$$\leq \frac{4 \cdot 10.0366}{3(e \cdot 0.249)^{\frac{1}{4}}} J^{-\frac{1}{2}} p^{\frac{11}{16} + \frac{0.249}{4}}.$$

Recall $J = \left\lfloor \frac{1}{6\kappa\psi(p)} \right\rfloor \neq 0$, so $J^{-\frac{1}{2}} \leq \sqrt{12\kappa\psi(p)}$ and $\psi(p) \geq p^{-\frac{5}{8} + \eta}$. In total, we get

$$\begin{aligned} \lambda(\mathcal{B}'(\kappa, \psi)^c) &\leq 2\kappa \sum_{p \text{ prime}} \frac{\psi(p)}{p} \sum_{\substack{a=1 \\ \left\| \frac{a^2}{p} \right\| \leq 3\kappa\psi(p)}} 1 \\ &\leq \kappa \frac{\pi^2}{2} \left(\frac{4\sqrt{12} \cdot 10.0366}{3(e \cdot 0.249)^{\frac{1}{4}}} \sqrt{\kappa} \sum_{p \text{ prime}} \psi(p)^{\frac{3}{2}} p^{-\frac{5}{16} + \frac{0.249}{4}} + \sum_{\substack{p \text{ prime} \\ p \leq \left\lfloor \frac{1}{6\kappa\psi(p)} \right\rfloor}} \frac{\psi(p)}{p} + 12\kappa \sum_{\substack{p \text{ prime} \\ p > \left\lfloor \frac{1}{6\kappa\psi(p)} \right\rfloor}} \psi(p)^2 \right) \\ &\leq \kappa^{\frac{3}{2}} \frac{\pi^2 2\sqrt{12} \cdot 10.0366}{3(e \cdot 0.249)^{\frac{1}{4}}} S_\psi + \kappa \frac{\pi^2}{2} S_\psi + \kappa^2 6\pi^2 S_\psi, \end{aligned}$$

which is $< \delta$ for

$$\begin{aligned} \kappa(\delta) &< \min \left(\left(\frac{\delta(e \cdot 0.249)^{\frac{1}{4}}}{2\pi^2 \sqrt{12} \cdot 10.0366 S_\psi} \right)^{\frac{2}{3}}, \frac{2\delta}{3\pi^2 S_\psi}, \sqrt{\frac{\delta}{18\pi^2 S_\psi}} \right) \\ &\simeq \min \left(\frac{\delta^{\frac{2}{3}}}{S_\psi^{\frac{2}{3}}} 0.0120432671, \frac{\delta}{S_\psi} 0.0675474558, \sqrt{\frac{\delta}{S_\psi}} 0.0750263597 \right). \end{aligned}$$

□

4. CONCLUDING REMARKS

The results of this paper give quantitative Khintchine theorems for convergence under four different conditions on the approximating rationals. The varying κ is very small indeed for the most general of the results. The problem is in the explicit constants arising from the various variants of Burgess' theorem applied, but more specifically from the occurrence of the divisor function in these settings. It occurs twice, once in the theorem for primitive characters and once in the passage from primitive to general characters. While a better estimate is desirable, it is in our opinion unlikely to be attainable with the methods of this paper.

APPENDIX A. PYTHON CODE

```
from sympy import kronecker_symbol
import math

FILENAME = "BurgessKonstantLogV2.txt"

# Read the last saved values from the log file so we can pick up where we left off:
with open(FILENAME, "r") as file:
```

```

last = file.readlines()[-2:]
qstart = int(last[0])
maks = float(last[1])

# Runs until manually interrupted with Ctrl+C
while True:
    for q in range(qstart, qstart + 500):
        lower = math.floor(q ** (3 / 8))
        upper = math.ceil(q ** (5 / 8))
        q_factor = q ** (3 / 16) # Constant for this q, so computed once outside the inner loop

        # Build the running sum up to lower:
        running_sum = sum(kronecker_symbol(N, q) for N in range(1, lower))

        # Check all N values in the relevant range [lower, upper]
        for N in range(lower, upper):
            running_sum += kronecker_symbol(N, q) # Extend the sum by one term rather than
                                                    recomputing from scratch

            VS = abs(running_sum)
            HS = math.sqrt(N) * q_factor
            maks = max(maks, VS / HS) # Update the maximum if we found a larger value

    # After completing a batch of 500 q values, save progress to the log file:
    qstart += 500
    with open(FILENAME, "a") as file:
        file.write("\n" + str(qstart)) # Save where to start next time
        file.write("\n" + str(maks)) # Save the largest value found so far

```

Running this results in Lemma 9 as well as the constant mentioned after Lemma 12. Included is a snippet of the txt-file, BurgessKonstantLogV2.txt, the program produces (The first two lines containing zeroes were added manually before running the program).

```

0
0
500
1.96397891158409
1000
2.29264220735264
1500
2.40876073940135
2000
2.51290888600514
2500
2.57511197155484
3000
2.66040431048477
3500
2.73314435040598
4000
2.75794837745823
4500
2.82596689219766
5000
2.82596689219766
...
19000
3.39672051063381
19500
3.41105405758246
20000
3.41105405758246
20500
3.41105405758246
21000

```

```

3.41105405758246
...
99500
4.19253564643679
100000
4.19253564643679
100500
4.20598018618317
101000
4.20598018618317

```

REFERENCES

- [1] F. ADICEAM, V. BERESNEVICH, J. LEVESLEY, S. VELANI, AND E. ZORIN, *Diophantine approximation and applications in interference alignment*, Advances in Mathematics, 302 (2016), pp. 231–279.
- [2] T. M. APOSTOL, *Introduction to Analytic Number Theory*, Springer-Verlag, 1979.
- [3] E. BACH AND J. SHALLIT, *Algorithmic Number Theory.*, MIT Press, Cambridge, MA, 1996.
- [4] V. BERESNEVICH, S. DATTA, AND L. YANG, *Weighted Diophantine approximation on manifolds*. Preprint, arXiv:2602.11045 [math.NT] (2026), 2026.
- [5] V. BERESNEVICH AND L. YANG, *Khinchine’s theorem and Diophantine approximation on manifolds*, Acta Math., 231 (2023), pp. 1–30.
- [6] A. R. BOOKER, *Quadratic class numbers and character sums*, Math. Comput., 75 (2006), pp. 1481–1492.
- [7] M. BORDIGNON, *Partial Gaussian sums and the Pólya–Vinogradov inequality for primitive characters*, Rev. Mat. Iberoam, (2022), pp. 1101–1127.
- [8] D. A. BURGESS, *On character sums and L-series. ii*, Proceedings of the London Mathematical Society, s3-13 (1963), pp. 524–536.
- [9] J. BÜTKE, *An improved analytic method for calculating $\pi(x)$* , Manuscr. Math., 151 (2016), pp. 329–352.
- [10] J. W. S. CASSELS, *An Introduction to Diophantine Approximation*, Cambridge University Press, 1965.
- [11] S. DATTA, *Quantitative Khinchine in simultaneous approximation*, Discrete and Continuous Dynamical Systems, 44(2) (2024), pp. 365–377.
- [12] R. J. DUFFIN AND A. C. SCHAEFFER, *Khinchine’s problem in metric Diophantine approximation*, Duke Mathematical Journal, 8 (1941), pp. 243 – 255.
- [13] J. HUANG, *Diophantine approximation on the parabola with non-monotonic approximation functions*, Mathematical Proceedings of the Cambridge Philosophical Society, 168 (2020), pp. 535–542.
- [14] H. IWANIEC, E. KOWALSKI, AND A. M. SOCIETY, *Analytic Number Theory*, vol. 53 of American Mathematical Society Colloquium Publications, American Mathematical Society, 2004.
- [15] N. JAIN-SHARMA, T. KHALE, AND M. LIU, *Explicit Burgess bound for composite moduli*, International Journal of Number Theory, 17 (2021), pp. 2207–2219.
- [16] D. KOUKOULOPOULOS AND J. MAYNARD, *On the Duffin-Schaeffer conjecture*, Ann. Math. (2), 192 (2020), pp. 251–307.
- [17] K. J. MCGOWN, *Norm-euclidean cyclic fields of prime degree*, International Journal of Number Theory, 08 (2012), pp. 227–254.
- [18] H. L. MONTGOMERY, *Ten Lectures on the Interface Between Analytic Number Theory and Harmonic Analysis*, American Mathematical Society Providence Rhode Island, American Mathematical Society, 1991.
- [19] J.-L. NICOLAS, *On highly composite numbers*, (1988).
- [20] G. ROBIN, *Estimate of the Chebyshev theta function on the k th prime number and large values of the number of prime divisors function $\omega(n)$ of n* , Acta arithmetica, 42 (1982), pp. 367–389.
- [21] G. ROBIN, *Thèse d’état: Grandes valeurs de fonctions arithmétiques et problèmes d’optimisation en nombres entiers*, PhD thesis, Université de Limoges, 1983.
- [22] T. TAO, *The divisor bound*. <https://terrytao.wordpress.com/2008/09/23/the-divisor-bound/>, 2008. Accessed: 2026-04-28.
- [23] E. TREVIÑO, *The Burgess inequality and the least k th power non-residue*, International Journal of Number Theory, 11 (2015), pp. 1653–1678.

MAIKEN GRAVGAARD, DEPARTMENT OF MATHEMATICS, AARHUS UNIVERSITY, NY MUNKEGADE 118, 8000 AARHUS C, DENMARK.

Email address: maiken@balmangravgaard.dk

SIMON KRISTENSEN, DEPARTMENT OF MATHEMATICS, AARHUS UNIVERSITY, NY MUNKEGADE 118, 8000 AARHUS C, DENMARK.

Email address: sik@math.au.dk