NEW GAPS ON THE LAGRANGE AND MARKOV SPECTRA

LUKE JEFFREYS, CARLOS MATHEUS, AND CARLOS GUSTAVO MOREIRA

ABSTRACT. Let L and M denote the Lagrange and Markov spectra, respectively. It is known that $L \subset M$ and that $M \setminus L \neq \varnothing$. In this work, we exhibit new gaps of L and M using two methods. First, we derive such gaps by describing a new portion of $M \setminus L$ near to 3.938: this region (together with three other candidates) was found by investigating the pictures of L recently produced by V. Delecroix and the last two authors with the aid of an algorithm explained in one of the appendices to this paper. As a by-product, we also get the largest known elements of $M \setminus L$ and we improve upon a lower bound on the Hausdorff dimension of $M \setminus L$ obtained by the last two authors together with M. Pollicott and P. Vytnova (heuristically, we get a new lower bound of 0.593 on the dimension of $M \setminus L$). Secondly, we use a renormalisation idea and a thickness criterion (reminiscent from the third author's PhD thesis) to detect infinitely many maximal gaps of M accumulating to Freiman's gap preceding the so-called Hall's ray $[4.52782956616..., \infty) \subset L$.

1. INTRODUCTION

The classical theory of Diophantine approximation is concerned with how well irrational numbers can be approximated by rational numbers. Given a positive real number α we define its *best constant of Diophantine approximation* to be

$$L(\alpha) := \limsup_{p,q\to\infty} \frac{1}{|q(q\alpha - p)|}.$$

In a sense, $L(\alpha)$ is the largest constant so that the inequality

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{L(\alpha)q^2}$$

has infinitely many solutions $p, q \in \mathbb{N}, q \neq 0$. The *Lagrange spectrum* is defined to be the set

$$L := \{L(\alpha) \mid \alpha \in \mathbb{R} \setminus \mathbb{Q}\}.$$

Perron [Pe21] proved that if we have the continued fraction expansion

$$\alpha = [a_0; a_1, a_2, \ldots] := a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots}}},$$

then we have

$$L(\alpha) = \limsup_{n\to\infty} ([a_n; a_{n-1}, \ldots, a_0] + [0; a_{n+1}, a_{n+2}, \ldots]).$$

As such, we are also able to define the Lagrange spectrum in terms of the bi-infinite shift space $\Sigma := \{1, 2, 3, \ldots\}^{\mathbb{Z}}$. More specifically, for $(a_i)_{i\in\mathbb{Z}} \in \Sigma$ we define

$$\lambda_0((a_i)_{i\in\mathbb{Z}}) := [a_0; a_{-1}, a_{-2}, \ldots] + [0; a_1, a_2, \ldots],$$

2020 *Mathematics Subject Classification.* Primary: 11J06. Secondary: 11A55.
The first author is thankful for support from the Heilbronn Institute for Mathematical Research.

and, for $j \in \mathbb{Z}$,

$$\lambda_j((a_i)_{i\in\mathbb{Z}}) := \lambda_0(\sigma^j((a_i)_{i\in\mathbb{Z}})) = \lambda_0((a_{i+j})_{i\in\mathbb{Z}}),$$

where $\sigma : \Sigma \to \Sigma$ is the left-shift sending $(a_i)_{i\in\mathbb{Z}}$ to $(a_{i+1})_{i\in\mathbb{Z}}$. We can now define the Lagrange spectrum to be

$$L := \{\limsup_{j\to\infty} \lambda_j(\underline{a}) \mid \underline{a} \in \Sigma\}.$$

Similarly, given $(a_i)_{i\in\mathbb{Z}} \in \Sigma$ we define

$$m((a_i)_{i\in\mathbb{Z}}) := \sup_{n\in\mathbb{Z}} \lambda_n((a_i)_{i\in\mathbb{Z}}).$$

Then the *Markov spectrum* is defined to be the set

$$M := \{m(\underline{a}) \mid \underline{a} \in \Sigma\}.$$

In the sequel, we will write a sequence $(a_i)_{i\in\mathbb{Z}}$ as the string $\dots a_{-2}a_{-1}a_0^*a_1a_2\dots$ where the asterisk denotes the 0th position. We will also use an overline to denote periodicity so that, for example, the sequence $a_i = (i \mod 3) + 1$ is denoted $\overline{1^*23} = \dots 1231231^*23123123\dots$. This notation should be clear from the context as we will mostly restrict to the subshift $\{1,2,3,4\}^{\mathbb{Z}}$ so, in particular, all a_i will be single digits.

Markov [Ma79, Ma80] first studied the spectra L and M around 1880. It is known that $L \subset M \subset \mathbb{R}^+$ with $L \cap (0,3) = M \cap (0,3)$ an explicit discrete set. In 1975, Freiman [Fr75] showed that $[\mu,\infty) \subset L \subset M$, and $(\nu,\mu) \cap M = \varnothing$ with $\nu, \mu \in M$, where

$$\nu = \lambda_0(\overline{323444}313134^*31312113\overline{3313121}) = 4.52782953841\dots$$

and

$$\mu = \lambda_0(\overline{12131}322344^*3211\overline{313121}) = 4.52782956616\dots.$$

The ray $[\mu,\infty)$ is known as Hall's ray after earlier work of Hall [Ha47] (see also the intermediate results of Freiman-Judin [FJ66], Hall [Ha71], Freiman [Fr73] and Schecker [Sc77]).

Freiman [Fr68] also showed that $M \setminus L \neq \varnothing$. In fact, the second and third authors together with M. Pollicott and P. Vytnova [MMPV22] recently proved that the Hausdorff dimension $HD(M \setminus L)$ of $M \setminus L$ satisfies

$$0.537152 < HD(M \setminus L) < 0.796445.$$

We direct the reader to the survey [MM21] and the textbooks of Cusick-Flahive [CF89] and Lima-Matheus-Moreira-Romaña [L+20] for more details on these spectra.

1.1. **A new portion of $M \setminus L$.** Our first result finds a new portion of $M \setminus L$ and gives an improved lower bound for its Hausdorff dimension.

Theorem 1.1. *The intersection of $M \setminus L$ with the interval $(3.938, 3.939)$ is non-empty. The largest known element of $M \setminus L$ is*

$$m(\overline{12}3311133113212311333111212113 33^*\overline{11121211333}) = 3.938776241989784909\dots.$$

Remark 1.2. *Our proof of this result yields that the local dimension of $M \setminus L$ near 3.938 coincides with the dimension of a dynamically defined Cantor set which is richer than the Cantor set Ω considered in [MMPV22, §4.6.5]. In particular, this improves the lower bound on $HD(M \setminus L)$ and, in fact, a heuristic computation (based on the so-called Jenkinson–Pollicott method) indicates that $HD(M \setminus L) > 0.593$: see the next section.*

The proof of this result is contained in Section 2. We also, in Appendix A, give some additional newly discovered portions of $M \setminus L$. We do not give the proof of these claims as they do not lead to significantly better estimates of the Hausdorff dimension of $M \setminus L$.

1.2. New maximal gaps of M.

Our second result concerns maximal gaps in the Markov spectrum M. Recall that Freiman proved that the gap (ν, μ) is a maximal gap of M. We find infinitely many new maximal gaps of M accumulating to Freiman's gap. Specifically, we prove the following.

Theorem 1.3. *There is a sequence (α_n, β_n) of maximal gaps of M such that $\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \beta_n = \nu$.*

In Section 3, we give a proof of Freiman's result that (ν, μ) is a maximal gap since the contributing lemmas are used in Section 4 in which we prove Theorem 1.3 via a renormalisation idea (leading to a sort of "recurrence on scales") and a thickness criterion in the spirit of the discussion of [Mo96].

1.3. Computational assistance in the investigations of $M \setminus L$.

The candidate sequence giving rise to elements of $M \setminus L$ analysed in Section 2 and those discussed in the appendix were discovered with the assistance of a computer search. The code was essentially running the arguments we will give in Section 2 which are themselves similar to those given in previous work of the second and third authors concerning elements of $M \setminus L$ near to 3.7096 [MM20].

We now describe the ideas behind the computer search. Firstly, for a candidate finite sequence a we determine the Markov value of the periodic sequence $s = \overline{a}$ determined by a. We then consider modifications of this sequence s where we force the sequence to instead terminate by $\overline{21}$ to the right or by $\overline{12}$ to the left. We find the modification that gives the smallest increase in the corresponding Markov value. Call this modified sequence w. Next, we try to determine the central portions of sequences that could give rise to Markov values in the range $[m(s), m(w) + \epsilon]$, for some small (possibly negative) ϵ. By searching for central portions of larger and larger length we can observe evidence for the one-sided periodicity we hope to make use of in the arguments given in Section 2. If we see no evidence for such one-sided periodicity after searching for central portions of a reasonable length then we throw out the candidate a and try for a new finite sequence. The pseudocode describing the algorithm used to determine the central portions of candidate sequences is given in Appendix B.

In practice the candidate finite sequences a are chosen to be odd length non-semi-symmetric words, where a word is semi-symmetric if it is a palindrome or a concatenation of two palindromes. We direct the reader to [MM20, Subsection 1.3] for a discussion of why odd length non-semi-symmetric words are natural candidates for finding elements of $M \setminus L$.

Acknowledgements. We are thankful to the referee whose comments helped to improve this paper.

2. A NEW PORTION OF $M \setminus L$ NEAR 3.938

We consider the word of odd length 11121211333. Note that it is non-semi-symmetric (in the sense of Flahive), i.e., it is not a palindrome nor a concatenation of two palindromes.

The Markov value of the associated periodic sequence is

$$\lambda_0(\overline{11121211333^*}) = 3.93877624198\textcolor{red}{10}28026\dots$$

Generally speaking, our goal below is to show that a portion of $M \setminus L$ occurs near

$$\lambda_0(\overline{1}2121211333111212113 33^*\overline{11121211333}) = 3.93877624198\textcolor{red}{11}39302\dots$$

In the sequel, we shall study a sequence $(\dots, x_{-m}, \dots, x_{-1}, x_0^*, x_1, \dots, x_n, \dots) \in \{1,2,3\}^{\mathbb{Z}}$ with a Markov value $m(x) = \lambda_0(x)$ nearby $3.93877624198\textcolor{red}{11}$.

For a finite sequence a, inequalities of the form $\lambda_0(...a...) > v$, say, mean that we have $\lambda_0(w) > v$ for all bi-infinite sequences w that are obtained by extending a on both sides.

2.1. **Local uniqueness.** Note that $x_0 = 3$. Moreover, the possible vicinities of x_0^* (up to transposition) are 13^*1, 13^*2, 13^*3, 23^*2, 23^*3, 33^*3.

Lemma 2.1. (i) $\lambda_0(...13^*1...) > 4.11$
(ii) $\lambda_0(...33^*3...) \leq \lambda_0(...33^*2...) \leq \lambda_0(...23^*2...) < 3.884$

By the previous lemma, up to transposition, it suffices to analyse the extensions to the right of 23^*1 and 33^*1, i.e., 23^*11, 23^*12, 23^*13, 33^*11, 33^*12, 33^*13.

Lemma 2.2. $\lambda_0(...3^*13...) > \lambda_0(...3^*12...) > 3.957$.

By the previous lemma, it suffices to analyse the extensions to the left of 23^*11 and 33^*11, i.e., 123^*11, 223^*11, 323^*11, 133^*11, 233^*11, 333^*11.

Lemma 2.3. (i) $\lambda_0(...323^*11...) > \lambda_0(...223^*11...) > 3.9678$
(ii) $\lambda_0(...133^*11...) < 3.9228$

By the previous lemma, it suffices to analyse the extensions to the right of 123^*11, 233^*11, 333^*11, i.e., 123^*111, 123^*112, 123^*113, 233^*111, 233^*112, 233^*113, 333^*111, 333^*112, 333^*113.

Lemma 2.4. (i) $\lambda_0(...123^*111...) > 3.9673$
(ii) *if* 131 *and* 312 *are forbidden, then* $\lambda_0(...233^*113...) < \lambda_0(...233^*112...) < \lambda_0(...233^*111...) \leq \lambda_0(...21233^*11132...) < 3.93676$
(iii) $\lambda_0(...333^*113...) < \lambda_0(...333^*112...) < 3.8969$

By the previous lemma, it suffices to analyse the extensions to the left of 123^*112, 123^*113, 333^*111, i.e., 1123^*112, 2123^*112, 3123^*112, 1123^*113, 2123^*113, 3123^*113, 1333^*111, 2333^*111, 3333^*111.

Lemma 2.5. (i) $\lambda_0(...1123^*112...) > \lambda_0(...2123^*112...) > 3.9414$*; in particular,* 123^*112 *is forbidden if* 312 *is forbidden*
(ii) $\lambda_0(...2123^*113...) < 3.93768$
(iii) *if* 131 *is forbidden, then* $\lambda_0(...1123^*113...) \geq \lambda_0(...1123^*11323...) > 3.9419$

By the previous lemma, it suffices to analyse the extensions to the right of 1333^*111, 2333^*111, 3333^*111, i.e., 1333^*1111, 1333^*1112, 1333^*1113, 2333^*1111, 2333^*1112, 2333^*1113, 3333^*1111, 3333^*1112, 3333^*1113.

Lemma 2.6. (i) $\lambda_0(...333^*1113...) > 3.94084$
(ii) $\lambda_0(...3333^*1111...) < \lambda_0(...2333^*1111...) < \lambda_0(...1333^*1111...) < 3.92786$
(iii) $\lambda_0(...3333^*1112...) < \lambda_0(...2333^*1112...) < 3.93844$

By the previous lemma, it suffices to analyse the extensions to the left of 1333^*1112, i.e., 11333^*1112, 21333^*1112, 31333^*1112. Since 213 and 313 are forbidden (cf. Lemma 2.2), our task is reduced to study the extensions to the right of 11333^*1112, i.e., 11333^*11121, 11333^*11122, 11333^*11123.

Lemma 2.7. $\lambda_0(...11333^*11123...) < \lambda_0(...11333^*11122...) < 3.93631$

By the previous lemma, it suffices to analyse the extensions to the left and right of 11333^*11121 (while taking into account that 213 is forbidden), i.e., $111333^*111211$, $211333^*111211$, $311333^*111211$, $111333^*111212$, $211333^*111212$, $311333^*111212$.

Lemma 2.8. $\lambda_0(...311333^*111211...) < \lambda_0(...211333^*111211...) < \lambda_0(...111333^*111211...) < 3.938464$

By the previous lemma (and after recalling that 131 and 3111333 are forbidden, cf. Lemmas 2.1 and 2.6 (i)), it suffices to analyse the extensions to the left of $111333^*111212$, $211333^*111212$, $311333^*111212$, i.e., $1111333^*111212$, $1211333^*111212$, $2111333^*111212$, $2211333^*111212$, $2311333^*111212$, $3211333^*111212$, $3311333^*111212$.

Lemma 2.9. $\lambda_0(...2111333^*111212...) > 3.93889$

By the previous lemma, it suffices to analyse the extensions to the right of $1111333^*111212$, $1211333^*111212$, $2211333^*111212$, $2311333^*111212$, $3211333^*111212$, $3311333^*111212$, i.e.,

- $1111333^*1112121$, $1111333^*1112122$, $1111333^*1112123$
- $1211333^*1112121$, $1211333^*1112122$, $1211333^*1112123$
- $2211333^*1112121$, $2211333^*1112122$, $2211333^*1112123$
- $2311333^*1112121$, $2311333^*1112122$, $2311333^*1112123$
- $3211333^*1112121$, $3211333^*1112122$, $3211333^*1112123$
- $3311333^*1112121$, $3311333^*1112122$, $3311333^*1112123$

Lemma 2.10. (i) $\lambda_0(...1111333^*1112121...) > \lambda_0(...1111333^*1112122...) > 3.938835$
(ii) $\max\{\lambda_0(...1211333^*1112123...), \lambda_0(...1211333^*1112122...), \lambda_0(...2211333^*1112123...)\} < \lambda_0(...2211333^*1112122...) < 3.938751$
(iii) $\lambda_0(...3211333^*1112121...) > \lambda_0(...2211333^*1112121...) > 3.938824$
(iv) $\lambda_0(...3211333^*1112123...)$, $\lambda_0(...2311333^*1112122...)$, $\lambda_0(...2311333^*1112123...)$, $\lambda_0(...3311333^*1112122...)$, $\lambda_0(...3311333^*1112123...) < \lambda_0(...3211333^*1112122...) < 3.9387718$

By the previous lemma (and after recalling that 312, 22311 and 32311 are forbidden, cf. Lemmas 2.2 and 2.3 (i)), it suffices to analyse the extensions to the left of $1111333^*1112123$, $1211333^*1112121$, $2311333^*1112121$, $3311333^*1112121$, i.e.,

- $11111333^*1112123$, $21111333^*1112123$, $31111333^*1112123$
- $11211333^*1112121$, $21211333^*1112121$
- $12311333^*1112121$
- $13311333^*1112121$, $23311333^*1112121$, $33311333^*1112121$

Lemma 2.11. (i) $\lambda_0(...11111333^*1112123...) > 3.9388049$
(ii) $\lambda_0(...11211333^*1112121...) > 3.9387855$

(iii) *if* 312 *and* 313 *are forbidden, then* $\lambda_0(...21111333^*1112123...) \geq \lambda_0(...21111333^*111212311...) > 3.93877973$

By the previous lemma (and after recalling that 213 is forbidden), it suffices to analyse the extensions to the right of $31111333^*1112123$, $21211333^*1112121$, $12311333^*1112121$, $13311333^*1112121$, $23311333^*1112121$, $33311333^*1112121$, i.e.,

- $31111333^*11121231$, $31111333^*11121232$, $31111333^*11121233$
- $21211333^*11121211$, $21211333^*11121212$
- $12311333^*11121211$, $12311333^*11121212$
- $13311333^*11121211$, $13311333^*11121212$
- $23311333^*11121211$, $23311333^*11121212$
- $33311333^*11121211$, $33311333^*11121212$

Lemma 2.12. (i) *if* 312 *and* 313 *are forbidden, then* $\lambda_0(...31111333^*11121231...) \leq \lambda_0(...31111333^*111212311...) < 3.938775326$
(ii) *if* 131 *is forbidden, then* $\lambda_0(...31111333^*11121233...) > \lambda_0(...31111333^*11121232...) \geq \lambda_0(...231111333^*11121232...) > 3.9387807$
(iii) $\lambda_0(...21211333^*11121212...) > \lambda_0(...3311333^*11121212...) > 3.938783$
(iv) $\lambda_0(...12311333^*11121211...) < \lambda_0(...3311333^*11121211...) < 3.9387521$

By the previous lemma (and after recalling that 312 and 1123113 are forbidden, cf. Lemmas 2.2 and 2.5 (iii)), it suffices to analyse the extensions to the left of $21211333^*11121211$, $12311333^*11121212$, i.e., $121211333^*11121211$, $221211333^*11121211$, $321211333^*11121211$, $212311333^*11121212$.

Lemma 2.13. *If* 131 *is forbidden, then* $\lambda_0(...321211333^*11121211...) > \lambda_0(...221211333^*11121211...) \geq \lambda_0(...221211333^*1112121132...) > 3.9387772$

By the previous lemma, it suffices to analyse the extensions to the right of $121211333^*11121211$, $212311333^*11121212$, i.e., $121211333^*111212111$, $121211333^*111212112$, $121211333^*111212113$, $212311333^*111212121$, $212311333^*111212122$, $212311333^*111212123$.

Lemma 2.14. (i) $\lambda_0(...121211333^*111212111...) > \lambda_0(...121211333^*111212112...) > 3.9387821$
(ii) *if* 312 *and* 313 *are forbidden, then* $\lambda_0(...212311333^*11121212...) \geq \lambda_0(...212311333^*11121212311...) > 3.938776505$

By the previous lemma (and after recalling that 312 is forbidden), it suffices to analyse the extensions to the left of $121211333^*111212113$, i.e., $1121211333^*111212113$, $2121211333^*111212113$.

Lemma 2.15. $\lambda_0(...2121211333^*111212113...) < 3.93877609$

By the previous lemma (and after recalling that 131 is forbidden), it suffices to analyse the extensions to the right of $1121211333^*111212113$, i.e., $1121211333^*1112121132$, $1121211333^*1112121133$.

Lemma 2.16. *If* 131 *and* 211321 *are forbidden*[1], *then* $\lambda_0(...1121211333^*1112121132...) \leq \lambda_0(...231121211333^*11121211322...) < 3.938775922$

By the previous lemma, we are led to investigate the extensions of $1121211333^*1112121133$. More concretely, the following statement is an immediate corollary of our discussions so far:

[1] Compare with Lemma 2.5 (i)

Corollary 2.17. *Let* $x \in \{1,2,3\}^{\mathbb{Z}}$ *be a sequence such that* $3.93877609 < m(x) = \lambda_0(x) < 3.938776505$. *Then,*

$$\dots x_{-1}x_0^* x_1 \dots = \dots 1121211333^* 1112121133 \dots$$

2.2. **Self-replication.** Our current goal is to describe the extensions of the string 1121211333*1112121133 leading to a Markov value strictly smaller than 3.93877624198**14**43.

For this sake, note that the extensions to the left of 1121211333*1112121133 are 11121211333*1112121133, 21121211333*1112121133, 31121211333*1112121133.

Lemma 2.18. $\lambda_0(...31121211333^*1112121133...) > \lambda_0(...21121211333^*1112121133...) > 3.93877687$

By the previous lemma, it suffices to analyse the extensions to the right of 11121211333*1112121133, i.e., 11121211333*11121211331, 11121211333*11121211332, 11121211333*11121211333.

Lemma 2.19. $\lambda_0(...11121211333^*11121211331...) > \lambda_0(...11121211333^*11121211332...) > 3.938776301$

By the previous lemma, it suffices to analyse the extensions to the left of 11121211333*11121211333, i.e., 111121211333*11121211333, 211121211333*11121211333, 311121211333*11121211333.

Lemma 2.20. $\lambda_0(...111121211333^*11121211333...) > \lambda_0(...211121211333^*11121211333...) > 3.938776282$

By the previous lemma (and the fact that 312 and 313 are forbidden), it suffices to analyse the extensions to the right of 311121211333*11121211333, i.e., 311121211333*1112121133311, 311121211333*111212113332, 311121211333*111212113333.

Lemma 2.21. *If* 131 *is forbidden, then* $\lambda_0(...311121211333^*111212113333...) > \lambda_0(...311121211333^*111212113332...) \geq \lambda_0(...2311121211333^*111212113332...) > 3.938776248$

By the previous lemma (and after recalling that 131, 22311, 32311, 123111 are forbidden, cf Lemmas 2.1 (i), 2.3 (i), 2.4 (i)), it suffices to analyse the extensions to the left of 311121211333*1112121133311, i.e., 3311121211333*1112121133311. Now, we observe that the extensions to the left of 3311121211333*1112121133311 are 13311121211333*1112121133311, 23311121211333*1112121133311, 33311121211333*1112121133311.

Lemma 2.22. *If* 213 *and* 3331113 *are forbidden, then* $\lambda_0(...13311121211333^*1112121133311...) > \lambda_0(...23311121211333^*1112121133311...) \geq \lambda_0(\overline{21}23311121211333^*11121211333111\overline{12}) = 3.938776242699$

By the previous lemma, it suffices to analyse the extensions to the right of 33311121211333*1112121133311, i.e., 33311121211333*11121211333111, 33311121211333*11121211333112, 33311121211333*11121211333113.

Lemma 2.23. $\lambda_0(...33311121211333^*11121211333113...) > \lambda_0(...33311121211333^*11121211333112...) > 3.93877624592$

By the previous lemma (and after recalling that 213 and 313 are forbidden), it suffices to analyse the extensions to the left of 33311121211333*11121211333111, i.e., 1133311121211333*11121211333111, 233311121211333*11121211333111, 333311121211333*11121211333111.

Lemma 2.24. *If* 213 *and* 3331113 *are forbidden, then* $\lambda_0(...333311121211333^*11121211333111...) > \lambda_0(...233311121211333^*11121211333111...) \geq \lambda_0(...233311121211333^*111212113331112\overline{1}) > 3.93877624206$

By the previous lemma (and after recalling that 3331113 is forbidden), it suffices to analyse the extensions to the right of 1133311121211333*11121211333111, i.e., 1133311121211333*111212113331111, 1133311121211333*111212113331112.

Lemma 2.25. $\lambda_0($...1133311121211333*111212113331111...$) > 3.93877624309$

By the previous lemma, it suffices to analyse the extensions to the right of 1133311121211333*111212113331112, i.e,

- 1133311121211333*1112121133311121,
- 1133311121211333*1112121133311122, 1133311121211333*1112121133311123

Lemma 2.26. $\lambda_0($...1133311121211333*1112121133311123...$) >$ $\lambda_0($...1133311121211333*1112121133311122...$) > 3.938776242211$

By the previous lemma (and after recalling that 213 is forbidden), it suffices to analyse the extensions to the right of 1133311121211333*1112121133311121, i.e., 1133311121211333*11121211333111211, 1133311121211333*11121211333111212.

Lemma 2.27. $\lambda_0($...1133311121211333*11121211333111211...$) > 3.93877624201$

By the previous lemma (and after recalling that 3111333, 2111333111212, 11113331112121 are forbidden, cf Lemmas 2.6 (i), 2.9, 2.10 (i)), it suffices to analyse the extensions to the left of 1133311121211333*11121211333111212, i.e., 21133311121211333*11121211333111212, 31133311121211333*11121211333111212. As it turns out, the extensions to the right of these two words are:

- 21133311121211333*111212113331112121, 31133311121211333*111212113331112121
- 21133311121211333*111212113331112122, 31133311121211333*111212113331112122
- 21133311121211333*111212113331112123, 31133311121211333*111212113331112123

Lemma 2.28. $\min\{\lambda_0($...21133311121211333*111212113331112123...$),$ $\lambda_0($...31133311121211333*111212113331112123...$),$ $\lambda_0($...31133311121211333*111212113331112122...$)\} >$ $\lambda_0($...21133311121211333*111212113331112122...$) \geq$ $\lambda_0($...12121133311121211333*111212113331112122...$) > 3.938776241990046$, *since* 32113331112121 *and* 22113331112121 *are forbidden by Lemma 2.10*, 11211333111212 *is forbidden by Lemma 2.11, and* 321211333111121211 *and* 221211333111121211 *forbidden by Lemma 2.13.*

By the previous lemma (and after recalling that 213 and 2121133311121212 are forbidden, cf. Lemmas 2.2 and 2.12 (iii)), it suffices to analyse the extensions to the right of 21133311121211333*111212113331112121, 31133311121211333*111212113331112121, i.e., 21133311121211333*1112121133311121211, 31133311121211333*1112121133311121211. As it turns out, the extensions to the right of these two words are

21133311121211333*11121211333111212113,

and

31133311121211333*11121211333111212113

because the strings 121211333111212111, 121211333111212112 are forbidden (cf. Lemma 2.14 (i)). Finally, the resulting words extend to the right as

21133311121211333*1112121133311121211 33

and

31133311121211333*111212113331112121133

because 131 and 11323, 11322, 211321 are forbidden (cf. Lemmas 2.3 (i) and 2.5 (i)).

In summary, our discussion so far yields the following statement:

Corollary 2.29. *Let $x \in \{1,2,3\}^{\mathbb{Z}}$ be a sequence with Markov value $m(x) < 3.938776241990046$. If x contains the string $1121211333^*1112121133$, say,*

$$x = \dots x_{i-9} \dots x_i^* \dots x_{i+10} \dots = \dots 1121211333^*1112121133 \dots,$$

then one has

$$x = \dots x_{i-15} \dots x_i^* \dots x_{i+21} = \dots 1133311121211333^*11121211333^{**}1112121133 \dots$$

and the vicinity of x_{i+11}^{**} is $1121211333^{**}1112121133$. In particular, by recursively analysing the positions x_{i+11k}, $k \in \mathbb{N}$, one actually has

$$x = \dots x_{i-15} \dots x_i^* \dots = \dots 1133311121211333^*\overline{11121211333}$$

Let

$$j_0 := \lambda_0(\overline{11121211333^*}) = 3.938776241981028026... \in L$$

and

$$\begin{aligned} j_1 &:= \lambda_0(\overline{21}233111331132123113331112121133311121211333^*11121211333111212 2\overline{32}) \\ &= 3.9387762419905494786868 7... \in L \end{aligned}$$

Proposition 2.30. *If $j_0 \leq m(a) = \lambda_0(a) < 3.9387762419922$ then (up to transposition) either*

- *$a = ...2113331112121133 3^*111212113331112122...$;*
- *$a = ...21133311121211333^*\overline{11121211333}$; or*
- *$a = ...31133311121211333^*\overline{11121211333}$.*

Proof. Since $j_0 \leq m(a) = \lambda_0(a) < 3.9387762419922$, we can use Corollary 2.17 and all of the results from Lemma 2.18 up to Lemma 2.27. Because

$$\begin{aligned} &\min\{\lambda_0(...21133311121211333^*111212113331112123...), \\ &\lambda_0(...31133311121211333^*111212113331112123...), \\ &\lambda_0(...31133311121211333^*111212113331112122...)\} > 3.9387762419922, \end{aligned}$$

we can partly use Lemma 2.28 together with the subsequent analysis to derive that either

$$a = ...21133311121211333^*111212113331112122...,$$

$$a = ...21133311121211333^*\overline{11121211333},$$

or

$$a = ...31133311121211333^*\overline{11121211333}.$$

□

Proposition 2.31. *If $j_0 < m(a) < 3.9387762419922$ and a contains*

$$21133311121211333^*111212113331112122,$$

then $m(a) \geq j_1$.

Proof. As in Lemma 2.28, we are forced to have

$$m(a) = \lambda_0(...12121133311121211333^*111212113331112122...).$$

Therefore, our task is reduced to check that if

$$m(a) = \lambda_0(...12121133311121211333^*111212113331112122...),$$

then one actually has $m(a) \geq j_1$. For this sake, observe that

$$\lambda_0(a) \geq \lambda_0(...112121133311121211333^*111212113331112122...).$$

At this point, Lemmas 2.18, 2.20, 2.22 and 2.24 force us to have

$$\lambda_0(a) \geq \lambda_0(...113331112121133311121211333^*111212113331112122...).$$

Hence,

$$\lambda_0(a) \geq \lambda_0(...123113331112121133311121211333^*111212113331112122...)$$

since 131, 32311 and 22311 are forbidden (cf. Lemmas 2.1 and 2.3). It follows from Lemma 2.5 (iii) that

$$\lambda_0(a) \geq \lambda_0(...132123113331112121133311121211333^*111212113331112122...).$$

After Lemmas 2.2, 2.4 (i), 2.5 (i), one has

$$\lambda_0(a) \geq \lambda_0(...311133113212311333111212113331112121 1333^*111212113331112122...).$$

By Lemmas 2.1(i), 2.3 (i), 2.4 (i), 2.6 (i), the strings 131, 23111 and 3331113 are forbidden, so that

$$\lambda_0(a) \geq \lambda_0(\overline{21}2331113311321231133311121211333111212 11333^*111212113331112122...).$$

We also have that

$$\lambda_0(a) \geq \lambda_0(\overline{21}233111331132123113331112121133311121211333^*111212113331121223...).$$

We claim that a cannot contain 2231. Indeed, Lemma 2.2 forbids 22313 and 22312 since both contain 313 or 312, while Lemma 2.3 forbids 22311. So we see that 2231 can never be extended.

We also claim that a cannot contain 3231. Indeed, Lemma 2.2 forbids 32313 and 32312 since both contained 313 or 312, while Lemma 2.3 forbids 32311. So we see that 3231 can never be extedned.

Therefore, since 2231 is forbidden,

$$\lambda_0(a) \geq \lambda_0(\overline{21}233111331132123113331112121133311121211333^*1112121133311121 2232...).$$

We also have that 3231 is forbidden and so we find that

$$\lambda_0(a) \geq \lambda_0(\overline{21}233111331132123113331112121133311121211333^*11121211333111212 2\overline{32}) = j_1.$$

□

Proposition 2.32. *The open interval $J = (j_0, j_1)$ is a maximal gap of L.*

Proof. If a is periodic and $j_0 \leq m(a) \leq j_1 < 3.9387762419922$, then Proposition 2.30 tells us that $a = \overline{11121211333}$ in which case $m(a) = j_0 \notin J$, or a contains 21133311121211333*111212113331112122. In the latter case, Proposition 2.31 then tells us that $m(a) \geq j_1$ and so again $m(a) \notin J$. Therefore, J does not contain the Markov value of any periodic sequence and so, since the Lagrange spectrum is the closure of the set of Markov values of periodic sequences, we conclude that J is indeed a maximal gap of L. □

Proposition 2.33. *Let $a \in \{1,2,3\}^{\mathbb{Z}}$ be a sequence with Markov value $j_0 < m(a) = \lambda_0(a) < j_1$ then $m_1 \leq m(a) \leq m_4$, where*

$$m_1 = m(\overline{12}2331113311321212113331112121133 3^*\overline{11121211333})$$

$$= 3.9387762419810960597...$$

and

$$m_4 = m(\overline{12}23311133113212311333111212113 33^*\overline{11121211333})$$

$$= 3.938776241989784909....$$

Proof. By Propositions 2.30 and 2.31, we have that

$$a = ...2113331112121133 3^*\overline{11121211333}$$

or

$$a = ...3113331112121133 3^*\overline{11121211333}$$

.

We begin by analysing the former. Since 32113331112121 and 22113331112121 are forbidden by Lemma 2.10, 11211333111212 is forbidden by Lemma 2.11, and 32121133311121211 is forbidden by Lemma 2.13, we have

$$a = \ldots 1212113331112121133 3^*\overline{11121211333}.$$

Since 312 is forbidden, this sequence extends to the left with 1 or 2. Suppose that it extends by a 1. By Corollary 2.29, and the same arguments we just made, we see that

$$a = \ldots 1212113331112121133 3^{***}11121211333^*\overline{11121211333}$$

and, once again, this word could extend on the left with 1 or 2. Here, the triple *** indicates the neighbourhood in which Corollary 2.29 is being applied. However, an extension with 2 is not possible because this would force $\lambda_{-11}(a) > \lambda_0(a) = m(a)$, a contradiction. Continuing would leave us with $a = \overline{11121211333}$, so $m(a) = j_0$, which is also a contradiction. So we must have

$$a = \ldots 21212113331112121133 3^*\overline{11121211333}.$$

Now

$$m(a) \geq m(...1321212113331112121133 3^*\overline{11121211333}).$$

By Lemma 2.2, 313 and 213 are forbidden in a and so

$$m(a) \geq m(...11321212113331112121133 3^*\overline{11121211333}).$$

Lemmas 2.4 and 2.5 forbid 111321 and 2113212, so we must have

$$m(a) \geq m(...311321212113331112121133 3^*\overline{11121211333}).$$

Similar arguments allow us to show that

$$m(a) \geq m(...3111331132121211333111212113 33^*\overline{11121211333}).$$

Lemma 2.1 forbids 131. We claim that 23111 is also forbidden. Lemma 2.3 forbids 223111 and 323111 while Lemma 2.4 forbids 123111 and so 23111 is never extendible and so must be forbidden. Therefore,

$$m(a) \geq m(...33111331132121211333111212113 33^*\overline{11121211333}).$$

Lemma 2.6 prevents 3331113 and so

$$m(a) \geq m(...233111331132121211333111212113 33^*\overline{11121211333}).$$

From here on, 312 being forbidden by Lemma 2.2 gives us that

$$m(a) \geq m(\overline{12}3311133113212121133311121211333^*\overline{11121211333}) = m_1.$$

Now analysing the possibility that $a = ...3113331112121133 3^*\overline{11121211333}$. Since 131 is forbidden, we have

$$m(a) \leq m(...2311333111212113 33^*\overline{11121211333}).$$

Now, we are forbidden to have 32311 and 22311 so we must have

$$m(a) \leq m(...12311333111212113 33^*\overline{11121211333}).$$

Next, since 1123113 is forbidden, we must have

$$m(a) \leq m(...212311333111212113 33^*\overline{11121211333}).$$

Then

$$m(a) \leq m(...3212311333111212113 33^*\overline{11121211333}).$$

Now we have

$$m(a) \leq m(...13212311333111212113 33^*\overline{11121211333}).$$

Since 313 and 213 are forbidden, we must have

$$m(a) \leq m(...113212311333111212113 33^*\overline{11121211333}).$$

Now 111321 and 211321 are forbidden so we must have

$$m(a) \leq m(...3113212311333111212113 33^*\overline{11121211333}).$$

Then

$$m(a) \leq m(...133113212311333111212113 33^*\overline{11121211333}).$$

Since 313 and 213 are forbidden we get

$$m(a) \leq m(...1133113212311333111212113 33^*\overline{11121211333}).$$

Then

$$m(a) \leq m(...31113311321231133311121211333^*\overline{11121211333}).$$

Now 131 is forbidden and extending by 2 would lead to one of 32311, 22311, or 123111 all of which are forbidden. So we obtain

$$m(a) \leq m(...331113311321231133311121211333^*\overline{11121211333}).$$

We have that 3331113 is forbidden and so we must have

$$m(a) \leq m(...2331113311321231133311121211333^*\overline{11121211333}).$$

From here we obtain

$$m(a) \leq m(\overline{12}331113311321231133311121211333^*\overline{11121211333}) = m_4.$$

This completes the proof. □

An immediate consequence of our discussion so far is the following statement:

Corollary 2.34. *$HD((M \setminus L) \cap (j_0, j_1)) = HD(K)$ where K is the Gauss–Cantor set of continued fractions with entries* 1, 2, 3 *not containing the following forbidden strings (nor their transposes):*

- 131, 312, 313, 22311, 32311, 123111, 123112, 1123113, 3331113, 2111333111212,
- 11113331112121, 11113331112122, 22113331112121, 32113331112121,
- 111113331112123, 112113331112121, 211113331112123, 3111133311121232,
- 3111133311121233, 2121133311121212, 331133311121212, 22121133311121211,

- 321211333111121211, 121211333111212111, 121211333111212112,
- 21231133311121212, 11212113331112121133.

Proof. Denote by $\mathcal{F}$ the set consisting of the strings above and their transposes. By Corollary 2.17, if $x \in \{1,2,3\}^{\mathbb{Z}}$ and $j_0 < m(x) < j_1$, then $\dots x_{-1}x_0^*x_1 \dots = \dots 1121211333^*1112121133\dots$ (up to transposition). Furthermore, the discussion before Corollary 2.17 says that x doesn't contain the strings in $\mathcal{F} \setminus \{\gamma, \gamma^t\}$, where $\gamma = 11212113331112121133$ is the "self-replicating" word and γ^t is its transpose.

By Propositions 2.30 and 2.31, one actually has that

$$x = y^t 1133311121211333^* \overline{11121211333}$$

where $y \in \{1,2,3\}^{\mathbb{N}}$ doesn't contain strings from $\mathcal{F} \setminus \{\gamma, \gamma^t\}$. By Proposition 2.33 and Corollary 2.29, either y has the form $y = \delta\overline{11121211333}$ where δ is a *finite* string or y doesn't contain a string from $\mathcal{F}$. In particular, $M \cap (j_0, j_1)$ is included in the union of a countable set and a set which is bi-Lipschitz homeomorphic to K, so that $\mathrm{HD}((M \setminus L) \cap (j_0, j_1)) = \mathrm{HD}(M \cap (j_0, j_1)) \leq \mathrm{HD}(K)$. Since it is not hard to see that $(M \setminus L) \cap (j_0, j_1)$ contains the set

$$\{m(y^t 2121211333111121211333^* \overline{11121211333}) : y^t 21212 \text{ doesn't contain strings from } \mathcal{F}\}$$

which is bi-Lipschitz homeomorphic to K, the argument is now complete. □

Performing calculations using the methods of Jenkinson-Pollicot [JP01], we obtained heuristics suggesting that $0.593 < \mathrm{HD}(K') < \mathrm{HD}(K'') < 0.595$, where K' is the Gauss–Cantor set of continued fractions with entries 1, 2, 3 not containing the forbidden strings 131, 312, 313, 22311, 32311, 123111, 123112, 1123113, 3331113, and 11333111212 (nor their transposes), and K'' is the Gauss–Cantor set of continued fractions with entries 1, 2, 3 not containing the forbidden strings 131, 312, 313, 22311, 32311, 123111, 123112, 1123113, 3331113 (nor their transposes). Since the every forbidden string for K has a subword that is a forbidden string for K', we see that $K' \subset K$. Similarly, since the forbidden strings for K'' are a strict subset of those for K, we have $K \subset K''$. Hence we expect the heuristic

$$0.593 < \mathrm{HD}(K) < 0.595$$

to be true which would also give us that $\mathrm{HD}(M \setminus L) > 0.593$ - an improved lower bound.

3. Freiman's gap

In [Fr75, Section 10, pp.66–71], G. Freiman proved the following result:

Theorem 3.1. *One has $M \cap (\nu, \mu) = \varnothing$ where*

$$\nu = [4; 3,1,3,1,3,\overline{4,4,4,3,2,3}] + [0; 3,1,3,1,2,1,1,3,3,\overline{3,1,3,1,2,1}]$$

and

$$\mu = [4; 4,3,2,2,\overline{3,1,3,1,2,1}] + [0; 3,2,1,1,\overline{3,1,3,1,2,1}]$$

In this section, we extract key parts of the proof of this theorem. For this sake, we restrict from now on our attention to the sequences $\underline{a} = (a_n)_{n \in \mathbb{Z}} \in (\mathbb{N}^*)^{\mathbb{Z}}$ such that

$$4 < m(\underline{a}) = \lambda_0(\underline{a}) < 5.$$

Note that these inequalities imply that

$$\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}} \quad \text{and} \quad a_0 \in \{3,4\}.$$

3.1. **Preliminaries.** We require the following results the proofs of which can be found in [L+20, Appendix D]. The first determine that the central portion of a candidate sequence giving rise to Markov values in the range (ν, μ) must be (up to transposition) ...34*3... or ...34*4....

Lemma 3.2. *If* $m(\underline{a}) < 4.55$*, then* $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ *can not contain the subwords* 41, 42 *or their transposes.*

Lemma 3.3. *If* $m(\underline{a}) < 4.52786$*, then* $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ *can not contain the subwords* 313133, 443131344 *or their transposes.*

Corollary 3.4. *Suppose that* $4.5278 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.52786$*. Then,* $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ *has the form* $\dots a_{-1}a_0a_1 \dots = \dots 343 \dots$ *or* $\dots 344 \dots$ *(up to transposition).*

3.2. **Extensions of the word 343.** The following results analyse possible extensions of ...34*3....

Lemma 3.5. *If* $m(\underline{a}) < 4.52786$*, then* $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ *can not contain the subwords* 3432, 134312, 31343132, 21313431312 *or their transposes.*

Corollary 3.6. *If* $4.5278295 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.5278296$ *and* $a_{-1}a_0a_1 = 343$*, then* $a_{-9} \dots a_0 \dots a_7 = 33112131343131344$ *(up to transposition).*

Lemma 3.7. *If* $m(\underline{a}) < 4.528$*, then* $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ *can not contain the subwords* 334, 223444 *or their transposes.*

We include the proof of the following corollary as we will make use of the details in the next section.

Corollary 3.8. *If* $4.5278295 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.5278296$ *and* $a_{-1}a_0a_1 = 343$*, then* $m(\underline{a}) \leq \nu$*.*

Proof. By Corollary 3.6, we have that $a_{-9} \dots a_0 \dots a_7 = 33112131343131344$ (up to transposition). We want to *maximize* $4.5278295 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.5278296$. By Lemma 3.2, this means that $a_{-9} \dots a_0 \dots a_9 = 3311213134313134443$. By Lemma 3.7, we have $a_{-9} \dots a_0 \dots a_{11} = 331121313431313444323$. By Lemma 3.5, we derive $a_{-9} \dots a_0 \dots a_{11} = 331121313431313444432344$. By repeating this argument, we conclude that $a_{-9} \dots a_0 \dots a_7 \dots = 33112131343131344\overline{432344}$. Similarly, we have from Lemma 3.7 that $a_{-10} \dots a_0 \dots a_7 = 333112131343131344$. By Lemma 3.2, we get $a_{-13} \dots a_0 \dots a_7 = 1313331121313431313144$. By Lemma 3.3, $a_{-15} \dots a_0 \dots a_7 = 121313331121313431313144$. By repeating this argument, we get $\dots a_{-9} \dots a_0 \dots a_7 = \overline{1213133}31121313431313144$.

In summary, our assumptions imply the maximal value of $m(\underline{a})$ is ν. □

3.3. **Extensions of the word 344.** The following corollary results from an analysis of possible extensions of ...34*4....

Corollary 3.9. *If* $4.5278291 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.527832$ *and* $a_{-1}a_0a_1 = 344$*, then* $m(\underline{a}) \geq \mu$*.*

3.4. **End of the proof of Theorem 3.1.** The desired result follows directly from Corollaries 3.4, 3.8 and 3.9.

4. Gaps of the spectra nearby Freiman's gap

In this section we prove Theorem 1.3. The proof of this theorem begins with the following lemmas.

Lemma 4.1. *If* $4.5278295 < m(\underline{a}) = \lambda_0(\underline{a}) < 4.5278296$, *then either* $m(\underline{a}) \geq \mu > \nu$ *or* $m(\underline{a}) \leq \nu$ *and, up to transposition,*

$$\underline{a} = \dots 3311213134^*3131344\dots$$

Proof. This is a direct consequence of Corollaries 3.4, 3.6, 3.8 and 3.9. □

Define, for $n, m \in \mathbb{N}$,

$$\underline{\theta}_n := (444323)^n = \underbrace{444323\dots444323}_{n \text{ times}} \quad \text{and} \quad \underline{\theta}'_m := (313121)^m = \underbrace{313121\dots313121}_{m \text{ times}}.$$

Lemma 4.2. *The family of sets*

$$\begin{aligned} W_{n,m} \quad = \quad & \{m(\underline{a}) = \lambda_0(\underline{a}) \in (4.5278295, \mu) : \underline{a} = \underline{\theta}^t 323444313134^*313121133313121\underline{\theta}' \\ & \text{with } \underline{\theta} = \underline{\theta}_n \widehat{\underline{\theta}},\ \underline{\theta}' = \underline{\theta}'_m \widetilde{\underline{\theta}}, \text{ and } \widehat{\underline{\theta}}, \widetilde{\underline{\theta}} \in \{1,2,3,4\}^{\mathbb{N}}\} \end{aligned}$$

indexed by $n, m \in \mathbb{N}$ *is a basis of neighborhoods of* ν *in* M.

Proof. This follows directly from Lemma 4.1 and the proof of Corollary 3.8. □

Lemma 4.3. *Let*

$$\begin{aligned} K \quad = \quad & \{[0;\underline{\theta}] : \underline{\theta} \in \{1,2,3,4\}^{\mathbb{N}} \text{ doesn't contain the strings } 14, 24, 433, 434, \\ & 131313, 2343, 223444, 123444 \text{ or their transposes}\}, \end{aligned}$$

$$K_1 = \{[0;3,1,3,1,2,1,\widetilde{\underline{\theta}}] \in K\},$$
$$K_2 = \{[0;4,4,4,3,2,3,\widehat{\underline{\theta}}] \in K\},$$

and define

$$g(x) = [0;3,1,3,1,2,1+x] \quad \text{and} \quad h(y) = [0;4,4,4,3,2,3+y].$$

Then, for each $n, m \in \mathbb{N}$, *one has*

$$W_{n,m} \subset A_n + B_m$$

where

$$A_n = \{[4;3,1,3,1,2,1,1,3,3,3,1,3,1,2,1+g^{n-1}(x)] : x \in K_1\},$$

and

$$B_m = \{[0;3,1,3,1,3,4,4,4,3,2,3+h^{m-1}(y)] : y \in K_2\}.$$

Proof. This is an immediate consequence of Lemma 4.2, and the fact that Lemmas 3.2, 3.3, 3.5, 3.7 ensure that $\underline{a} \in \{1,2,3,4\}^{\mathbb{N}}$ with $m(\underline{a}) < \mu$ can't contain the strings 14, 24, 433, 434, 131313, 2343, 223444, 123444 or their transposes. □

In view of Lemma 4.3, our task is reduced to find gaps in the arithmetic sums $A_n + B_m$ for infinitely many pairs of indices n, m. In this direction, we observe that K_1 and K_2 are dynamical Cantor sets which are invariant under the contractions

$$g(x) = [0;3,1,3,1,2,1+x] \quad \text{and} \quad h(y) = [0;4,4,4,3,2,3+y]$$

whose fixed points are

$$\alpha = [0;\overline{313121}] \quad \text{and} \quad \beta = [0;\overline{444323}].$$

For subsequent reference, we note that g and h can be rewritten as

$$g(x) = \frac{14x+19}{53x+72}, \quad h(y) = \frac{127y+436}{538y+1847}.$$

In particular,

$$g'(x) = \frac{1}{(53x+72)^2}, \qquad h'(y) = \frac{1}{(538y+1847)^2}$$

and

$$\alpha = \frac{2\sqrt{462}-29}{53}, \qquad \beta = \frac{\sqrt{243542}-430}{269}.$$

Lemma 4.4. *One has* $\alpha = \min K_1$, $\beta = \min K_2$, *and*

$$\frac{\log|g'(\alpha)|}{\log|h'(\beta)|} \in \mathbb{R}\setminus\mathbb{Q}.$$

Proof. The fact that $\alpha = \min K_1$, $\beta = \min K_2$ follows from the definition of K_1, K_2 and the constraint on the continued fraction expansions of the elements of K. Furthermore, a straightforward computation yields

$$g'(\alpha) = \frac{1}{(43+2\sqrt{462})^2} \quad \text{and} \quad h'(\beta) = \frac{1}{(987+2\sqrt{243542})^2}.$$

Since $462 = 2\cdot 3\cdot 7\cdot 11$ and $243542 = 2\cdot 13\cdot 17\cdot 19\cdot 29$, their square roots generate distinct quadratic extensions of $\mathbb{Q}$ and

$$g'(\alpha)^m = \frac{1}{(43+2\sqrt{462})^{2m}} \neq \frac{1}{(987+2\sqrt{243542})^{2n}} = h'(\beta)^n$$

for all $n, m \in \mathbb{N}^*$. Hence, $\frac{\log|g'(\alpha)|}{\log|h'(\beta)|} \in \mathbb{R}\setminus\mathbb{Q}$. This ends the proof of the lemma. $\square$

Also for later use, let us recall the following bound on the distortion of certain inverse branches of the Gauss map:

Lemma 4.5. *Let* $f(x) = [0; a_1, \dots, a_k + x]$ *be the inverse branch of the Gauss map associated to a finite word* $(a_1, \dots, a_k) \in \{1,2,3,4\}^k$, $k \geq 1$. *Then,*

$$\frac{1}{2.3} < \frac{|f'(x)|}{|f'(y)|} < 2.3$$

for any $\frac{\sqrt{2}-1}{2} \leq x, y \leq 2\sqrt{2}-2$.

Proof. Since $f(z) = \frac{p_{k-1}z+p_k}{q_{k-1}z+q_k}$ and $|f'(z)| = \frac{1}{(q_{k-1}z+q_k)^2}$, where $\frac{p_j}{q_j} = [0; a_1, \dots, a_j]$ for all $1 \leq j \leq k$, we have

$$\frac{1}{2.3} < \left(\frac{1+\sqrt{2}}{2(2\sqrt{2}-1)}\right)^2 \leq \frac{|f'(x)|}{|f'(y)|} = \left(\frac{\frac{q_{k-1}}{q_k}y+1}{\frac{q_{k-1}}{q_k}x+1}\right)^2 \leq \left(\frac{2(2\sqrt{2}-1)}{1+\sqrt{2}}\right)^2 < 2.3$$

for $\frac{\sqrt{2}-1}{2} \leq x, y \leq 2\sqrt{2}-2$ (as $1/5 \leq q_{k-1}/q_k \leq 1$). $\square$

An interesting consequence of this lemma is the fact that the sets A_n and B_m (cf. Lemma 4.3) are mildly distorted "copies" of K_1 and K_2. For this reason, the next lemma about the "thickness" of K_1 and K_2 at their minima will be useful later.

Lemma 4.6. *Consider the intervals* $R_0 = [\alpha, \alpha_1]$, $U_0 = (\alpha_1, \alpha_2)$, $L_0 = [\beta, \beta_1]$ *and* $V_0 = (\beta_1, \beta_2)$, *where*

- α_1 *is the largest element of* K_1 *of the form* $[0; 3, 1, 3, 1, 2, 1, 3, \underline{\widetilde{\theta}}]$,
- α_2 *is the smallest element of* K_1 *of the form* $[0; 3, 1, 3, 1, 2, 1, 2, \underline{\widetilde{\theta}}]$,
- β_1 *is the largest element of* K_2 *of the form* $[0; 4, 4, 4, 3, 2, 3, 4, \underline{\widehat{\theta}}]$,
- β_2 *is the smallest element of* K_2 *of the form* $[0; 4, 4, 4, 3, 2, 3, 3, \underline{\widehat{\theta}}]$.

Then,

$$\frac{|R_0|}{|U_0|} < 1 \quad \text{and} \quad \frac{|L_0|}{|V_0|} < \frac{1}{100}.$$

Proof. Since the strings 41, 42 and 2343 are forbidden in continued fraction expansions in K, we have that $\beta_1 \leq [0; 4, 4, 4, 3, 2, 3, 4, \overline{4, 3}]$ and $\beta_2 \geq [0; 4, 4, 4, 3, 2, 3, \overline{3, 1}]$, and

$$\frac{|L_0|}{|V_0|} = \frac{\beta_1 - \beta}{\beta_2 - \beta_1} < 0.008565 < \frac{1}{100}.$$

Similarly, we have $\alpha_1 \leq [0; 3, 1, 3, 1, 2, 1, \overline{3, 4}]$ and $\alpha_2 \geq [0; 3, 1, 3, 1, 2, 1, 2, \overline{1, 3}]$, and

$$\frac{|R_0|}{|U_0|} = \frac{\alpha_1 - \alpha}{\alpha_2 - \alpha_1} < 0.98479 < 1.$$

This completes the argument. □

At this point, we are ready to complete the proof of Theorem 1.3. In fact, Lemmas 4.2 and 4.3 reduce our task to find gaps in $A_n + B_m$ for infinitely many $n, m \in \mathbb{N}^*$. Since $A_n = f_0 \circ g^n(K_1)$ and $B_m = f_1 \circ h^m(K_2)$, where

$$f_0(x) = [4; 3, 1, 3, 1, 2, 1, 1, 3, 3 + x] \quad \text{and} \quad f_1(x) = [0; 3, 1, 3, 1, 3 + x],$$

and Lemma 4.4 ensures the denseness of $\{|g'(\alpha)|^n / |h'(\beta)|^m : n, m \in \mathbb{N}^*\}$ in $\mathbb{R}_+$, we get[2], for any $c \in \mathbb{R}_+$, there are infinitely many $n, m \in \mathbb{N}^*$ such that

$$\frac{c}{2} < \frac{|R_n|}{|L_m|} < 2c,$$

where $R_n = f_0 \circ g^n(R_0)$ and $L_m = f_1 \circ h^m(L_0)$. Because Lemma 4.5 also says that

$$\frac{|L_m|}{|V_m|} < \frac{2.3}{100} \quad \text{and} \quad \frac{|R_n|}{|U_n|} < 2.3,$$

where $U_n = f_0 \circ g^n(U_0)$, $V_m = f_1 \circ h^m(V_0)$ are gaps of A_n and B_m (as U_0 and V_0 are gaps of K_1 and K_2), we conclude that

$$\frac{|L_m|}{|U_n|} = \frac{|L_m|}{|R_n|} \cdot \frac{|R_n|}{|U_n|} < \frac{2}{c} \cdot 2.3 \quad \text{and} \quad \frac{|R_n|}{|V_m|} = \frac{|R_n|}{|L_m|} \cdot \frac{|L_m|}{|V_m|} < 2c \cdot \frac{2.3}{100}.$$

Thus, if we take $c = 5$, then

$$\frac{|L_m|}{|U_n|} < 0.92 < 1 \quad \text{and} \quad \frac{|R_n|}{|V_m|} < 0.23 < 1.$$

This ends the proof of Theorem 1.3 because the inequalities above imply that $A_n + B_m$ has a gap: indeed, these estimates say that any parameter $t \in \mathbb{R}$ such that $t - U_n$ contains L_m and their right endpoints are sufficiently close also satisfies

[2] Actually, using the general distortion bound statement in Chapter 4 of Palis–Takens book, it is possible to show that for any $c \in \mathbb{R}_+$ and $0 < \varepsilon < 1$, one has $c(1 - \varepsilon) < \frac{|R_n|}{|L_m|} < c(1 + \varepsilon)$ for infinitely many $n, m \in \mathbb{N}^*$.

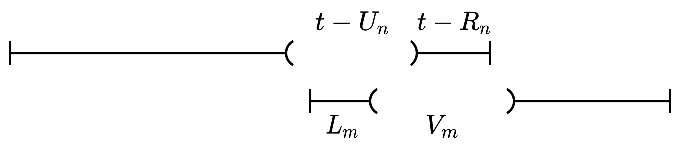

FIGURE 1. Producing gaps in $A_n + B_m$.

$t - R_n \subset V_m$ and, *a fortiori*, $(t - A_n) \cap B_m = \varnothing$ (see Figure 1); hence, $A_n + B_m$ misses an entire open interval of parameters. Furthermore, in the language of the statement of Theorem 1.3, each maximal gap (α_n, β_n) in the infinite sequence is contained in $A_j + B_k$, for some $j, k \in \mathbb{N}$, with the diameter of $A_j + B_k$ tending to 0 as $j, k \to \infty$. Note that $j, k \to \infty$ as $n \to \infty$. It then follows from Lemma 4.2 that $\lim_{n\to\infty} \alpha_n = \lim_{n\to\infty} \beta_n = \nu$, as claimed.

APPENDIX A. ADDITIONAL ELEMENTS OF $M \setminus L$

Here we present new elements of $M \setminus L$ that are less than those discussed in Section 2. We only give the definitions of the sequences and Cantor sets involved and leave the proofs to the interested reader. These new sequences were also discovered using the computational search technique discussed in the introduction.

A.1. **Elements of $M \setminus L$ near to 3.676.** Computer investigations lead us to believe that there is a portion of $M \setminus L$ near to 3.676 given by an analysis of the subset of the real line near to

$$m(\overline{3^*21112123}) = 3.676699417246755742\dots.$$

A.2. **Elements of $M \setminus L$ near to 3.726.** Computer investigations lead us to believe that there is a portion of $M \setminus L$ near to 3.726 given by an analysis of the subset of the real line near to

$$m(\overline{3322211121223^*}) = 3.726146224233042720\dots.$$

Computer investigations also lead us to believe that there is a portion of $M \setminus L$ near to 3.726 given by an analysis of the subset of the real line near to

$$m(\overline{33222121223^*}) = 3.726278993734881116\dots.$$

A.3. **Elements of $M \setminus L$ near to 3.942.** Computer investigations lead us to believe that there is a portion of $M \setminus L$ near to 3.942 given by an analysis of the subset of the real line near to

$$m(\overline{33211121232331113^*}) = 3.942001159911341469\dots.$$

Note that this value is higher than the elements near to 3.938 that we rigorously considered in this paper. We chose not to analyse this sequence since, given its length, it would require a more involved analysis of the combinatorics without (in heuristic calculations) giving rise to an appreciable increase in the Hausdorff dimension estimates of $M \setminus L$.

APPENDIX B. PSEUDO-CODE FOR COMPUTER SEARCH

Below is the pseudo-code for the part of the computer search that determines the central portion of sequences $\underline{a} \in \{1,2,3,4\}^{\mathbb{Z}}$ for which $m(\underline{a}) = \lambda_0(\underline{a}) \in [l,n]$, for some interval $[l,n]$. The algorithm was implemented using the SageMath mathematical software [Sage].

Algorithm 1 - Find sequences whose Markov values could lie in the range $[l,n]$

```
candidates ← [1*, 2*, 3*, 4*]
forbidden_words ← []
alphabet ← {_, 1, 2, 3, 4}                                   # _ is the empty string
extensions ← (alphabet × alphabet) \ {(_, _)}
l ← l
n ← n
length_limit ← maximum length of sequences to search up to
min_seq_len ← minimum length of all sequences in candidates
while min_seq_len < length_limit and candidates ≠ [] do
    for sequence in candidates do
        allowable ← True
        remove sequence from candidates
        for (x, y) in extensions do
            trial_sequence ← concatenation(x, sequence, y)
            if trial_sequence contains any words from forbidden_words then
                continue            # the sequence is forbidden so move on to the next
            end if
            λ_max ← maximum possible value of λ_0(trial_sequence)
            if λ_max < l then
                continue            # λ_0 is too small so move on to the next sequence
            end if
            for z in trial_sequence do
                j ← position of z in trial_sequence
                λ_min ← minimum possible value of λ_j(trial_sequence)
                if λ_min > n then
                    append trial_sequence to forbidden_words
                    allowable ← False                # the Markov value is too large
                end if
            end for
            if allowable then                        # the Markov value can lie in [l, n]
                append trial_sequence to candidates
            end if
        end for
    end for
    if candidates ≠ [] then
        min_seq_len ← minimum length of all sequences in candidates
    end if
end while
return candidates
```

The code can also be used to 'confirm' results about gaps in the spectra. For example, when running the code on intervals like $(0, \sqrt{5})$, $(\sqrt{12}, \sqrt{13})$ or other known gaps the code terminates and returns an empty list of candidate sequences. On closed intervals, if the endpoints correspond to unique sequences, the code will return a two element list of finite sequences approaching the sequences corresponding to the endpoints.

References

[Be73] A. A. Berstein. *The connections between the Markov and Lagrange spectra*, Number-theoretic studies in the Markov spectrum and in the structural theory of set addition, pp. 16–49, 121–125. Kalinin. Gos. Univ., Moscow, 1973.

[CF89] T. Cusick and M. Flahive. *The Markoff and Lagrange spectra*, Mathematical Surveys and Monographs, 30. American Mathematical Society, Providence, RI, 1989.

[L+20] Davi Lima, Carlos Matheus, Carlos G Moreira, and Sergio Romaña. *Classical and dynamical markov and lagrange spectra*, WORLD SCIENTIFIC, 2020.

[Fr68] G. A. Freiman. *Noncoincidence of the Markoff and Lagrange spectra*, Mat. Zametki 3 (1968), 195–200; English transl., Math. Notes 3 (1968),1 25–128.

[Fr73] G. A. Freiman. *The initial point of Hall's ray, Number-theoretic studies in the Markov spectrum and in the structural theory of set addition*, pp. 87–120, 121–125. Kalinin. Gos. Univ., Moscow, 1973.

[Fr75] G. A. Freiman. *Diophantine approximation and the geometry of numbers (The Markoff spectrum)*, Kalininskii Gosudarstvennyi Universitet, Moscow, 1975.

[FJ66] G.A. Freiman and A.A. Judin. *Über das Markoff-Spektrum (Russian)* Litovsk Mat. Sb., 6 (1966), pp. 443–447.

[Ha47] M. Hall Jr. *On the sum and product of continued fractions*, Ann. of Math., 48 (1947), pp. 966–993.

[Ha71] M. Hall Jr. *The Markoff spectrum*, Acta Arith., 18 (1971), pp. 387–399.

[JP01] O. Jenkinson and M. Pollicott, *Computing the dimension of dynamically defined sets: E2 and bounded continued fractions*, Ergodic Theory Dynam. Systems 21 (2001), no. 5, 1429–1445.

[Ma79] A. Markov. *Sur les formes quadratiques binaires indéfinies*, Math. Ann. 17 (1879), no. 3, 381–406.

[Ma80] A. Markov. *Sur les formes quadratiques binaires indéfinies II*, Math. Ann. 17 (1880), no. 3, 379–399.

[MM20] C. Matheus and C. G. Moreira. *Fractal geometry of the complement of Lagrange spectrum in Markov spectrum*, Comment. Math. Helv. 95 (2020), no. 3, 593–633.

[MM21] C. Matheus and C. G. Moreira. *Diophantine approximation, Lagrange and Markov spectra, and dynamical Cantor sets*, Notices Amer. Math. Soc. 68 (2021), no. 8, 1301–1311.

[MMPV22] C. Matheus, C. G. Moreira, M. Pollicott and P. Vytnova. *Hausdorff dimension of Gauss-Cantor sets and two applications to classical Lagrange and Markov spectra*, Adv. Math. 409 (2022), Paper No. 108693.

[Mo96] C. G. Moreira. *Stable intersections of Cantor sets and homoclinic bifurcations*. Ann. Inst. H. Poincaré C Anal. Non Linéaire 13 (1996), no. 6, 741–781.

[Mo18] C. G. Moreira. *Geometric properties of the Markov and Lagrange spectra*, Ann. of Math. 188 (2018), 145–170.

[Pe21] O. Perron. *Über die approximation irrationaler Zahlen durch rationale II*, S.-B. Heidelberg Akad. Wiss. 8 (1921).

[PT93] J. Palis and F. Takens. *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations, fractal dimensions and infinitely many attractors*, Cambridge Studies in Advanced Mathematics, 35. Cambridge University Press, Cambridge, 1993.

[Sage] SageMath, the Sage Mathematics Software System (Version 9.4), The Sage Developers, 2021, `https://www.sagemath.org`.

[Sc77] H. Schecker, *Über die Menge der Zahlen, die als Minima quadratischer Formen auftreten*, J. Number Theory 9 (1977), no. 1, 121–141.

SCHOOL OF MATHEMATICS , UNIVERSITY OF BRISTOL, FRY BUILDING, WOODLAND ROAD, BRISTOL BS8 1UG
Email address: luke.jeffreys@bristol.ac.uk

CENTRE DE MATHÉMATIQUES LAURENT SCHWARTZ, ÉCOLE POLYTECHNIQUE, 91128 PALAISEAU CEDEX, FRANCE
Email address: carlos.matheus@math.cnrs.fr

SCHOOL OF MATHEMATICAL SCIENCES, NANKAI UNIVERSITY, TIANJIN 300071, P. R. CHINA, AND IMPA, ESTRADA DONA CASTORINA 110, CEP 22460-320, RIO DE JANEIRO, BRAZIL
Email address: gugu@impa.br