New gaps on the Lagrange and Markov spectra

par Luke JEFFREYS, Carlos MATHEUS et Carlos Gustavo MOREIRA

Résumé. On note respectivement $L$ et $M$ les spectres de Lagrange et de Markov. Il est connu que $L \subset M$ et que $M \setminus L \neq \emptyset$. Dans ce travail, on détecte de nouvelles lacunes dans $L$ et $M$ en utilisant les deux méthodes suivantes. Premièrement, on obtient de telles lacunes en décrivant une nouvelle partie de $M \setminus L$ proche de 3.938 : cette région (avec trois autres candidats) a été trouvée en étudiant les images de $L$ récemment produites par V. Delecroix et les deux derniers auteurs à l’aide de l’algorithme expliqué dans l’un des appendices de cet article. En outre, on obtient les plus grands éléments connus de $M \setminus L$ et on améliore la minoration de la dimension de Hausdorff de $M \setminus L$ obtenue par les deux derniers auteurs avec M. Pollicott et P. Vytnova (heuristiquement, on obtient une nouvelle minoration de la dimension de $M \setminus L$ par 0,593). Deuxièmement, on utilise une idée de renormalisation et un critère d’épaisseur (issu de la thèse de doctorat du troisième auteur) pour détecter une infinité de lacunes maximales de $M$ s’accumulant près de la lacune de Freiman précédant le célèbre rayon de Hall $[4.52782956616 \ldots, \infty) \subset L$.

Abstract. Let $L$ and $M$ denote the Lagrange and Markov spectra, respectively. It is known that $L \subset M$ and that $M \setminus L \neq \emptyset$. 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 $[4.52782956616 \ldots, \infty) \subset L$.

Manuscrit reçu le 21 octobre 2022, révisé le 3 avril 2023, accepté le 26 mai 2023.

2020 Mathematics Subject Classification. 11J06, 11A55.

Mots-clés. Lagrange and Markov spectra, maximal gaps, Hausdorff dimension.

The first author is thankful for support from the Heilbronn Institute for Mathematical Research.
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 $\alpha$ 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

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

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 [17] proved that if we have the continued fraction expansion

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

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

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

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

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

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

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

In the sequel, we will write a sequence $(a_i)_{i \in \mathbb{Z}}$ as the string

$$\ldots a_{-2}a_{-1}a_0a_1a_2 \ldots$$
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 $\overline{123} = \ldots 123123123123 \ldots$. 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 [10, 11] 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 [4] showed that $[\mu, \infty) \subset L \subset M$, and $(\nu, \mu) \cap M = \emptyset$ with $\nu, \mu \in M$, where

$$\nu = \lambda_0(323443\overline{13134}3131211333\overline{13121}) = 4.52782953841 \ldots$$

and

$$\mu = \lambda_0(12131322344\overline{32113}313121) = 4.52782956616 \ldots .$$

The ray $[\mu, \infty)$ is known as Hall’s ray after earlier work of Hall [6] (see also the intermediate results of Freiman–Judin [5], Hall [7], Freiman [3] and Schecker [19]).

Freiman [2] also showed that $M \setminus L \neq \emptyset$. In fact, the second and third authors together with M. Pollicott and P. Vytnova [14] recently proved that the Hausdorff dimension $\text{HD}(M \setminus L)$ of $M \setminus L$ satisfies

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

We direct the reader to the survey [13] and the textbooks of Cusick–Flahive [1] and Lima–Matheus–Moreira–România [9] 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

$$m(\overline{123311133113212311333111211333}3211) = 3.938776241989784909 \ldots .$$

**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 $\Omega$ considered in [14, §4.6.5]. In particular, this improves the lower bound on $\text{HD}(M \setminus L)$ and, in fact, a heuristic computation (based on the so-called Jenkinson–Pollicott method [8]) indicates that $\text{HD}(M \setminus L) > 0.593$: 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 $(\nu, \mu)$ 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 $(\alpha_n, \beta_n)$ 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 $(\nu, \mu)$ 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 [15].

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 [12].

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 $21$ to the right or by $12$ 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) $\epsilon$. 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 pseudo-code 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 [12, 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. Below, we use boldface to highlight decimal places of importance.

The Markov value of the associated periodic sequence is

$$\lambda_0(11121211333^*) = 3.938776241981028026\ldots$$

Generally speaking, our goal below is to show that a portion of $M \setminus L$ occurs near $\lambda_0(1212121133311121211333^*) = 3.938776241981139302\ldots$

In the sequel, we shall study a sequence

$$(\ldots, x_{-m}, \ldots, x_{-1}, x_0^*, x_1, \ldots, x_n, \ldots) \in \{1, 2, 3\}^\mathbb{Z}$$

with a Markov value $m(x) = \lambda_0(x)$ nearby $3.93877624198111$.

For a finite sequence $a$, inequalities of the form $\lambda_0(\ldots a \ldots) > v$, 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) $\lambda_0(\ldots 13^*1 \ldots) > 4.11$

(ii) $\lambda_0(\ldots 33^*3 \ldots) \leq \lambda_0(\ldots 33^*2 \ldots) \leq \lambda_0(\ldots 23^*2 \ldots) < 3.884$.

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

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) $\lambda_0(\ldots 323^*11 \ldots) > \lambda_0(\ldots 223^*11 \ldots) > 3.9678$

(ii) $\lambda_0(\ldots 133^*11 \ldots) < 3.9228$.

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) \( \lambda_0(\ldots123*111\ldots) > 3.9673 \)
(ii) if 131 and 312 are forbidden, then \( \lambda_0(\ldots233*111\ldots) < \lambda_0(\ldots233*112\ldots) < \lambda_0(\ldots233*111\ldots) \leq \lambda_0(\ldots21233*11132\ldots) < 3.93676 \)
(iii) \( \lambda_0(\ldots333*111\ldots) < \lambda_0(\ldots333*112\ldots) < 3.8969. \)

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) \( \lambda_0(\ldots1123*111\ldots) > \lambda_0(\ldots2123*112\ldots) > 3.9414; \) in particular, 123*112 is forbidden if 312 is forbidden
(ii) \( \lambda_0(\ldots2123*113\ldots) < 3.93768 \)
(iii) if 131 is forbidden, then \( \lambda_0(\ldots1123*113\ldots) \geq \lambda_0(\ldots1123*1132\ldots) > 3.9419. \)

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) \( \lambda_0(\ldots333*111\ldots) > 3.94084 \)
(ii) \( \lambda_0(\ldots3333*1111\ldots) < \lambda_0(\ldots2333*1111\ldots) < \lambda_0(\ldots1333*1111\ldots) < 3.92786 \)
(iii) \( \lambda_0(\ldots3333*1112\ldots) < \lambda_0(\ldots2333*1112\ldots) < 3.93844. \)

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. \( \lambda_0(\ldots11333*11123\ldots) < \lambda_0(\ldots11333*11122\ldots) < 3.93631. \)

By the previous lemma, it suffices to analyse the extensions to the left and right of 13333*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. \( \lambda_0(\ldots311333*111211\ldots) < \lambda_0(\ldots211333*111211\ldots) < \lambda_0(\ldots111333*111211\ldots) < 3.938464. \)

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. \( \lambda_0(\ldots 2111333^*111212\ldots) > 3.93889. \)

By the previous lemma, it suffices to analyse the extensions to the right of
1111333*111212, 2111333*111212, 2211333*111212, 2311333*111212, 3211333*111212, 3311333*111212, i.e.,
- 1111333*111212, 1111333*1112122, 1111333*1112123
- 1211333*111212, 1211333*1112122, 1211333*1112123
- 2211333*111212, 2211333*1112122, 2211333*1112123
- 2311333*111212, 2311333*1112122, 2311333*1112123
- 3211333*111212, 3211333*1112122, 3211333*1112123
- 3311333*111212, 3311333*1112122, 3311333*1112123.

Lemma 2.10.

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

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.,
- 1111333*1112123, 2111333*1112123, 3111333*1112123
- 1211333*1112121, 1211333*1112121
- 1311333*1112121, 2311333*1112121, 3311333*1112121.

Lemma 2.11.

(i) \( \lambda_0(\ldots 1111333^*1112123\ldots) > 3.9388049 \)
(ii) \( \lambda_0(\ldots 11211333^*1112121\ldots) > 3.9387855 \)
(iii) if 312 and 313 are forbidden, then \( \lambda_0(\ldots 21111333^*1112123\ldots) \geq \lambda_0(\ldots 21111333^*11121231\ldots) > 3.9387797. \)

By the previous lemma (and after recalling that 213 is forbidden),
it suffices to analyse the extensions to the right of 3111333*1112123, 21211333*1112121, 12311333*1112121, 13311333*1112121, 2311333*1112121, 33311333*1112121, i.e.,
- 3111333*11121231, 3111333*11121232, 3111333*11121233
- 21211333*11121211, 21211333*11121212
Lemma 2.12.

(i) if 312 and 313 are forbidden, then $\lambda_0(\ldots 3111333*11121231\ldots) \leq \lambda_0(\ldots 31111333*111212311\ldots) < 3.938775326$

(ii) if 131 is forbidden, then $\lambda_0(\ldots 3111333*11121232\ldots) > \lambda_0(\ldots 31111333*11121232\ldots) > 3.9387807$

(iii) $\lambda_0(\ldots 21211333*11121212\ldots) > \lambda_0(\ldots 3111333*11121212\ldots) > 3.938783$

(iv) $\lambda_0(\ldots 1311333*11121211\ldots) < \lambda_0(\ldots 3111333*11121211\ldots) < 3.9387521$.

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 $21211333*11121211$, $3111333*11121212$, i.e., $121211333*11121211$, $21211333*11121211$, $12311333*11121212$, $212311333*11121212$.

Lemma 2.13. If 131 is forbidden, then

$\lambda_0(\ldots 321211333*11121211\ldots) > \lambda_0(\ldots 221211333*11121212\ldots) > 3.9387772.$

By the previous lemma, it suffices to analyse the extensions to the right of $121211333*11121211$, $21211333*11121212$, i.e., $121211333*11121211$, $121211333*11121212$, $21211333*11121212$, $212311333*11121213$.

Lemma 2.14.

(i) $\lambda_0(\ldots 121211333*11121211\ldots) > \lambda_0(\ldots 121211333*11121212\ldots) > 3.9387821$

(ii) if 312 and 131 are forbidden, then $\lambda_0(\ldots 212311333*11121212\ldots) > \lambda_0(\ldots 212311333*1112121231\ldots) > 3.938776505$.

By the previous lemma (and after recalling that 312 is forbidden), it suffices to analyse the extensions to the left of $121211333*11121211$, i.e., $1121211333*111212113$, $21211333*111212113$, $212311333*11121212$, $212311333*11121212$.

Lemma 2.15. $\lambda_0(\ldots 212111333*11121213\ldots) < 3.93877609.$

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
\[
\lambda_0(\ldots1121211333^*1112121132\ldots) \leq \lambda_0(\ldots231121211333^*11121211322\ldots) < 3.938775922.
\]

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

**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,
\[
\ldots x_{-1}x_0^*x_1^*\ldots = \ldots1121211333^*1112121133\ldots.
\]

### 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 3.938776241981443.

For this sake, note that the extensions to the left of

1121211333\(^*\)1112121133

are

11121211333\(^*\)1112121133, 21121211333\(^*\)1112121133

and 31121211333\(^*\)1112121133.

**Lemma 2.18.**
\[
\lambda_0(\ldots31121211333^*1112121133\ldots) > \lambda_0(\ldots21121211333^*1112121133\ldots) > 3.93877687.
\]

By the previous lemma, it suffices to analyse the extensions to the right of 11121211333\(^*\)1112121133, i.e.,

11121211333\(^*\)11121211331, 11121211333\(^*\)11121211332,

and 11121211333\(^*\)11121211333.

**Lemma 2.19.**
\[
\lambda_0(\ldots11121211333^*11121211331\ldots) > \lambda_0(\ldots11121211333^*11121211332\ldots) > 3.938776301.
\]

By the previous lemma, it suffices to analyse the extensions to the left of 11121211333\(^*\)11121211333, i.e.,

111121211333\(^*\)11121211333, 211121211333\(^*\)11121211333,

and 311121211333\(^*\)11121211333.

---

\(^1\)Compare with Lemma 2.5(i)
Lemma 2.20.

\[ \lambda_0(\ldots 111121211333*11121211333\ldots ) \]
\[ > \lambda_0(\ldots 211121211333*11121211333\ldots ) \]
\[ > 3.938776282. \]

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.,

\[ 311121211333*1112121133311, \ 311121211333*111212113332, \]

and 311121211333*11121211333.

Lemma 2.21. If 131 is forbidden, then

\[ \lambda_0(\ldots 311121211333*11121211333\ldots ) \]
\[ > \lambda_0(\ldots 311121211333*111212113332\ldots ) \]
\[ \geq \lambda_0(\ldots 311121211333*111212113333\ldots ) > 3.938776248. \]

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, 23311121211333*1112121133311, 33311121211333*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

\[ \lambda_0(\ldots 3311121211333*1112121133311\ldots ) \]
\[ > \lambda_0(\ldots 3311121211333*111212113332\ldots ) \]
\[ \geq \lambda_0(\ldots 3311121211333*111212113333\ldots ) = 3.938776242699. \]

By the previous lemma, it suffices to analyse the extensions to the right of 3311121211333*1112121133311, i.e., 33311121211333*1112121133311, 33311121211333*1112121133312, 33311121211333*1112121133313.

Lemma 2.23.

\[ \lambda_0(\ldots 3311121211333*111212113333\ldots ) \]
\[ > \lambda_0(\ldots 3311121211333*111212113333\ldots ) > 3.93877624592. \]

By the previous lemma (and after recalling that 213 and 313 are forbidden), it suffices to analyse the extensions to the left of

3311121211333*1112121133311,
New gaps on the Lagrange and Markov spectra

i.e.,

\[1133311121211333*11121211333111, \quad 233311121211333*11121211333111,\]

and \[333311121211333*11121211333111.\]

**Lemma 2.24.** If 213 and 3331113 are forbidden, then

\[\lambda_0(\ldots 333311121211333*11121211333111\ldots ) > \lambda_0(\ldots 233311121211333*11121211333111\ldots ) \geq \lambda_0(\ldots 233311121211333*11121211333111^{23}) > 3.93877624206.\]

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.,

\[1133311121211333*111212113331111,\]

and

\[1133311121211333*111212113331112.\]

**Lemma 2.25.**

\[\lambda_0(\ldots 1133311121211333*111212113331111\ldots ) > 3.93877624309\]

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

- \[1133311121211333*1112121133311121,\]
- \[1133311121211333*1112121133311122,\]
- \[1133311121211333*11121211333111123.\]

**Lemma 2.26.**

\[\lambda_0(\ldots 1133311121211333*1112121133311123\ldots ) > \lambda_0(\ldots 1133311121211333*1112121133311122\ldots ) > 3.93877624221.\]

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.,

\[1133311121211333*11121211333111211,\]

and

\[1133311121211333*11121211333111212.\]

**Lemma 2.27.**

\[\lambda_0(\ldots 1133311121211333*11121211333111211\ldots ) > 3.93877624201.\]
By the previous lemma (and after recalling that $3111333, 2111333111212, 1113331112121$ are forbidden, cf. Lemmas 2.6(i), 2.9, 2.10(i)), it suffices to analyse the extensions to the left of $113331112111333*11121211333111212$, i.e.,

$$2113331112111333*11121211333111212,$$

and

$$3113331112111333*11121211333111212.$$  

As it turns out, the extensions to the right of these two words are:

- $2113331112111333*11121211133311112121,$
- $3113331112111333*11121211133311112121,$
- $2113331112111333*11121211133311112122,$
- $3113331112111333*11121211133311112122,$
- $2113331112111333*11121211133311112123,$
- $3113331112111333*11121211133311112123.$

**Lemma 2.28.**

$$\min\{\lambda_0(\ldots 2113331112111333*11121211133311112123 \ldots),$$

$$\lambda_0(\ldots 3113331112111333*11121211133311112123 \ldots),$$

$$\lambda_0(\ldots 3113331112111333*11121211133311112123 \ldots)\}$$

$$> \lambda_0(\ldots 2113331112111333*11121211133311112122 \ldots)$$

$$\geq \lambda_0(\ldots 121133311112111333*11121211133311112122 \ldots)$$

$$> 3.938776241990046,$$

since $32113331112121$ and $22113331112121$ are forbidden by Lemma 2.10, $11211333111212$ is forbidden by Lemma 2.11, and $32112113331112121$ and $22112113331112121$ forbidden by Lemma 2.13.

By the previous lemma (and after recalling that $2121133311121212$ and $213$ are forbidden, cf. Lemmas 2.12(iii) and 2.2), it suffices to analyse the extensions to the right of $2113331112111333*11121211133311112121$, i.e.,

$$21133311121211333*11121211133311112121,$$

and

$$31133311121211333*11121211133311112121.$$  

As it turns out, the extensions to the right of these two words are

$$21133311121211333*11121211133311112113,$$

and

$$31133311121211333*11121211133311112113.$$
because the strings $121211333111212111, 12121133311212112$ are forbidden (cf. Lemma 2.14(i)). Finally, the resulting words extend to the right as

$$21133311121211333*111212113331112121133$$

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 $1121211333*11121211333, \text{say,}$

$$x = \ldots x_{i-9} \ldots x_i^* \ldots x_{i+10} \ldots = \ldots 1121211333*11121211333 \ldots,$$

then one has

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

and the vicinity of $x_{i+11}^*$ is $1121211333*11121211333$. In particular, by recursively analysing the positions $x_{i+11k}, k \in \mathbb{N}$, one actually has

$$x = \ldots x_{i-15} \ldots x_i^* \ldots = \ldots 1133311121211333*11121211333.$$

Let

$$j_0 := \lambda_0(1121211333^\tau) = 3.938776241981028026 \cdots \in L$$

and

$$j_1 := \lambda_0(212331113311321311333111211333111212110333^*111212113331112122132) = 3.9387762419905494786867 \cdots \in L.$$

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

- $a = \ldots 21133311121211333*111212113331112122 \ldots$;
- $a = \ldots 21133311121211333*11121211333; \text{ or}$
- $a = \ldots 31133311121211333*11121211333.$

**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

$$\min\{\lambda_0(\ldots 21133311121211333*111212113331112123 \ldots),$$

$$\lambda_0(\ldots 31133311121211333*111212113331112123 \ldots),$$

$$\lambda_0(\ldots 31133311121211333*111212113331112122 \ldots)\}$$

$$> 3.9387762419922,$$
we can partly use Lemma 2.28 together with the subsequent analysis to derive that either

- \( a = \ldots 21133311121211333^{*}111212113331112122 \ldots ; \)
- \( a = \ldots 21133311121211333^{*}1112111333; \) or
- \( a = \ldots 31133311121211333^{*}1112111333. \)

\[ \square \]

**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

\[ m(a) = \lambda_0(\ldots 1212113331121211333^{*}111212113331112122 \ldots ). \]

Therefore, our task is reduced to check that if

\[ m(a) = \lambda_0(\ldots 1212113331121211333^{*}111212113331112122 \ldots ), \]

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

\[ \lambda_0(a) \geq \lambda_0(\ldots 112121133311212113331112122 \ldots ). \]

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

\[ \lambda_0(a) \geq \lambda_0(\ldots 11333111212113331121211333^{*}111212113331112122 \ldots ). \]

Hence,

\[ \lambda_0(a) \geq \lambda_0(\ldots 1231133311212113331121211333^{*}111212113331112122 \ldots ), \]

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

\[ \lambda_0(a) \geq \lambda_0(\ldots 1321231133311212113331121211333^{*}111212113331112122 \ldots ). \]

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

\[ \lambda_0(a) \geq \lambda_0(\ldots 3113331123121133311212113331121211^{*}111212113331112122 \ldots ). \]

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

\[ \lambda_0(a) \geq \lambda_0(\ldots 233113331121231133311212113331121211^{*}111212113331112122 \ldots ). \]

We also have that

\[ \lambda_0(a) \geq \lambda_0(\ldots 233113331121231133311212113331121211^{*}1112121133311121223 \ldots ). \]
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 contain $313$ or $312$, while Lemma 2.3 forbids $32311$. So we see that $3231$ can never be extended.

Therefore, since $2231$ is forbidden,

$$
\lambda_0(a) \geq \lambda_0(2T233111311321231133311121113331111211112222) = j_1.
$$

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

$$
\lambda_0(a) \geq \lambda_0(2T233111311321231133311121113331111211112222) = j_1.
$$

\[\square\]

**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 $a = 1121211333$ in which case $m(a) = j_0 \not\in J$, or $a$ contains $21133311121211333*11121211333111211222$. In the latter case, Proposition 2.31 then tells us that $m(a) \geq j_1$ and so again $m(a) \not\in J$. 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 $L$. \[\square\]

**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

$$
\begin{align*}
 m_1 &= m(2T23311131132123113331112111333*11121211333) \\
 &= 3.9387762419810960597...
\end{align*}
$$

and

$$
\begin{align*}
 m_4 &= m(T23311131132123113331112111333*11121211333) \\
 &= 3.938776241989784909...
\end{align*}
$$

**Proof.** By Propositions 2.30 and 2.31, we have that

$$
a = \ldots 21133311121111333*11121211333
$$

or

$$
a = \ldots 31133311121111333*11121211333.
$$

We begin by analysing the former. Since $32113331112121$ and $22113331112121$ are forbidden by Lemma 2.10, $112113331111212$ is forbidden by Lemma 2.11, and $3212113331112111$ is forbidden by Lemma 2.13,
we have

\[ a = \ldots 121211333111121211333^*11121211333^*11121211333. \]

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

\[ a = \ldots 121211333111121211333^*11121211333^*11121211333^*11121211333^*11121211333. \]

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 \( a = 11121211333 \), so \( m(a) = j_0 \), which is also a contradiction. So we must have

\[ a = \ldots 2121211333111121211333^*11121211333. \]

Now

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

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

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

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

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

Similar arguments allow us to show that

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

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,

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

Lemma 2.6 prevents 3331113 and so

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333). \]

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

\[ m(a) \geq m(\ldots 12121211333111121211333^*11121211333) = m_1. \]

Now analysing the possibility that

\[ a = \ldots 31133311112111333^*11121211333. \]

Since 131 is forbidden, we have

\[ m(a) \leq m(\ldots 31133311112111333^*11121211333). \]
Now, we are forbidden to have 32311 and 22311 so we must have
\[ m(a) \leq m(\ldots 1231133311121211333^*11121211333). \]
Next, since 1123113 is forbidden, we must have
\[ m(a) \leq m(\ldots 21231133311121211333^*11121211333). \]
Then
\[ m(a) \leq m(\ldots 321231133311121211333^*11121211333). \]
Now we have
\[ m(a) \leq m(\ldots 1321231133311121211333^*11121211333). \]
Since 313 and 213 are forbidden, we must have
\[ m(a) \leq m(\ldots 11321231133311121211333^*11121211333). \]
Now 111321 and 211321 are forbidden so we must have
\[ m(a) \leq m(\ldots 311321231133311121211333^*11121211333). \]
Then
\[ m(a) \leq m(\ldots 13311321231133311121211333^*11121211333). \]
Since 313 and 213 are forbidden we get
\[ m(a) \leq m(\ldots 113311321231133311121211333^*11121211333). \]
Then
\[ m(a) \leq m(\ldots 3113311321231133311121211333^*11121211333). \]
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
\[ m(a) \leq m(\ldots 33113311321231133311121211333^*11121211333). \]
We have that 3331113 is forbidden and so we must have
\[ m(a) \leq m(\ldots 33111133113212311133311121211333^*11121211333). \]
From here we obtain
\[ m(a) \leq m(T2331113311321231133311121211333^*11121211333) = m_4. \]
This completes the proof. \( \square \)

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

**Corollary 2.34.** \( \text{HD}((M \setminus L) \cap (j_0, j_1)) = \text{HD}(K) \) 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,
  11113331112123, 12113331112121, 21113331112123,
  311133311121232, 311133311121233, 21211331112121,
  331133311121212, 221133311121211, 321133311121211,
  1212133311121211, 1211333111212112,
  2123133311121212, 1121133311121211.

\[ \text{Proof.} \] Denote by \( F \) 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 \( \ldots x_{-1} x_0 x_1 \cdots = \ldots 11211333^3111211333 \ldots \) (up to transposition). Furthermore, the discussion before Corollary 2.17 says that \( x \) doesn’t contain the strings in \( F \setminus \{\gamma, \gamma^t\} \), where \( \gamma = 112113331112121133 \) is the “self-replicating” word and \( \gamma^t \) is its transpose.

By Propositions 2.30 and 2.31, one actually has that \( x = y^t 1133311121211333^3 \Pi 1211333 \) where \( y \in \{1,2,3\}^\mathbb{N} \) doesn’t contain strings from \( F \setminus \{\gamma, \gamma^t\} \). By Proposition 2.33 and Corollary 2.29, either \( y \) has the form \( y = \delta \Pi 1211333 \) where \( \delta \) is a finite string or \( y \) doesn’t contain a string from \( F \). 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 \( \text{HD}((M \setminus L) \cap (j_0, j_1)) = \text{HD}(M \cap (j_0, j_1)) \leq \text{HD}(K) \). Since it is not hard to see that \( (M \setminus L) \cap (j_0, j_1) \) contains the set

\[ \{m(y^t 21211333111211333^3 \Pi 1211333) : y^t 21212 \text{ doesn’t contain strings from } F\} \]

which is bi-Lipschitz homeomorphic to \( K \), the argument is now complete. \( \square \)

Performing calculations using the methods of Jenkinson–Pollicot \[8\], we obtained heuristics suggesting that \( 0.593 < \text{HD}(K') < \text{HD}(K'') < 0.595 \), where \( K' \) is the Gauss–Cantor set of continued fractions with entries 1, 2, 3 not containing the forbidden strings 131, 312, 313, 2311, 32311, 12311, 123112, 1123113, 3331113, and 113331111212 (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, 2311, 32311, 12311, 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

\[ 0.593 < \text{HD}(K) < 0.595 \]

to be true which would also give us that \( \text{HD}(M \setminus L) > 0.593 \), an improved lower bound.
3. Freiman’s gap

In [4, Section 10, pp. 66–71], G. Freiman proved the following result:

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

\[
\nu = [4; 3, 1, 3, 1, 3, 4, 4, 4, 3, 2, 3] + [0; 3, 1, 3, 1, 2, 1, 3, 3, 3, 1, 3, 1, 2, 1]
\]

and

\[
\mu = [4; 4, 3, 2, 3, 1, 3, 1, 2, 1] + [0; 3, 2, 1, 1, 3, 1, 2, 1].
\]

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 \( a = (a_n)_{n \in \mathbb{Z}} \in (\mathbb{N}^*)^\mathbb{Z} \) such that

\[
4 < m(a) = \lambda_0(a) < 5.
\]

Note that these inequalities imply that \( a \in \{1, 2, 3, 4\}^\mathbb{Z} \) and \( a_0 \in \{3, 4\} \).

### 3.1. Preliminaries

We require the following results the proofs of which can be found in [9, Appendix D]. The first determine that the central portion of a candidate sequence giving rise to Markov values in the range \((\nu, \mu)\) must be (up to transposition) \( \ldots 34^*3 \ldots \) or \( \ldots 34^*4 \ldots \).

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

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

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

### 3.2. Extensions of the word 343

The following results analyse possible extensions of \( \ldots 34^*3 \ldots \).

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

**Corollary 3.6.** If \( 4.5278295 < m(a) = \lambda_0(a) < 4.5278296 \) and \( a_{-1}a_0a_1 = 343 \), then \( a_{-9} \ldots a_0 \ldots a_7 = 331121313431344 \) (up to transposition).

**Lemma 3.7.** If \( m(a) < 4.528 \), then \( a \in \{1, 2, 3, 4\}^\mathbb{Z} \) 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 \( 4.5278295 < m(a) = \lambda_0(a) < 4.5278296 \) and \( a_{-1}a_0a_1 = 343 \), then \( m(a) \leq \nu \).
Proof. By Corollary 3.6, we have that
\[ a_{-9} \ldots a_0 \ldots a_7 = 33112131343131344 \]
(up to transposition). We want to maximize \( 4.5278295 < m(a) = \lambda_0(a) < 4.5278296 \). By Lemma 3.2, this means that
\[ a_{-9} \ldots a_0 \ldots a_9 = 3311213134313134443. \]
By Lemma 3.7, we have \( a_{-9} \ldots a_0 \ldots a_{11} = 33112131343131344432344 \). By repeating this argument, we conclude that
\[ a_{-9} \ldots a_0 \ldots a_7 \ldots = 33112131343131344432344. \]
Similarly, we have from Lemma 3.7 that
\[ a_{-10} \ldots a_0 \ldots a_7 = 33112131343131344. \]
By Lemma 3.2, we get \( a_{-13} \ldots a_0 \ldots a_7 = 131333112131343131344 \). By Lemma 3.3, \( a_{-15} \ldots a_0 \ldots a_7 = 1213333112131343131344 \). By repeating this argument, we get \( a_{-9} \ldots a_0 \ldots a_7 = 121333112131343131344 \).

In summary, our assumptions imply the maximal value of \( m(a) \) is \( \nu \). □

3.3. Extensions of the word 344. The following corollary results from an analysis of possible extensions of \( \ldots 34^*4 \ldots \).

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

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 \( 4.5278295 < m(a) = \lambda_0(a) < 4.5278296 \), then either \( m(a) \geq \mu > \nu \) or \( m(a) \leq \nu \) and, up to transposition,
\[ a = \ldots 3311213134^*3131344 \ldots . \]

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

Define, for \( n, m \in \mathbb{N} \),
\[ \theta_n := (444323)^n = \underbrace{444323 \ldots 444323}_{n \text{ times}} \]
and
\[ \theta'_m := (313121)^m = \underbrace{313121 \ldots 313121}_{m \text{ times}}. \]
Lemma 4.2. The family of sets

\( W_{n,m} := \{ m(\mathcal{g}) = \lambda_0(\mathcal{g}) \in (4.5278295, \mu) : \mathcal{g} = \theta^n \mathcal{g}' 323444313134^{*}313121133313121\theta' \text{ with} \theta = \theta^n \hat{\theta}, \theta' = \theta'_m \tilde{\theta}, \text{ and } \hat{\theta}, \tilde{\theta} \in \{1,2,3,4\}^N \} \)

indexed by \( n,m \in \mathbb{N} \) is a basis of neighbourhoods of \( \nu \) in \( M \).

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

\[ \square \]

Lemma 4.3. Let

\( K = \{ [0; \theta] : \theta \in \{1,2,3,4\}^\mathbb{N} \text{ doesn’t contain the strings } 14, 24, 433, 434, 131313, 2343, 223444, 123444 \text{ or their transposes} \} \),

\[ K_1 = \{ [0; 3,1,3,1,2,1,\hat{\theta}] \in K \}, \]
\[ K_2 = \{ [0; 4,4,4,3,2,3,\hat{\theta}] \in K \}, \]

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

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

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 \( \mathcal{g} \in \{1,2,3,4\}^\mathbb{N} \) with \( m(\mathcal{g}) < \mu \) can’t contain the strings 14, 24, 433, 434, 131313, 2343, 223444, 123444 or their transposes.

\[ \square \]

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

\[ \alpha = [0; 3T3T2T] \quad \text{and} \quad \beta = [0; 4T4323]. \]
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,

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

and

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

**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

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

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. \hfill \Box

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

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

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

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

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

- \( \alpha_1 \) is the largest element of \( K_1 \) of the form \( [0; 3, 1, 3, 1, 2, 1, 3, \hat{\theta}] \),
- \( \alpha_2 \) is the smallest element of \( K_1 \) of the form \( [0; 3, 1, 3, 1, 2, 1, 2, \hat{\theta}] \),
- \( \beta_1 \) is the largest element of \( K_2 \) of the form \( [0; 4, 4, 4, 3, 2, 3, 4, \hat{\theta}] \),
- \( \beta_2 \) is the smallest element of \( K_2 \) of the form \( [0; 4, 4, 4, 3, 2, 3, 3, \hat{\theta}] \).

Then,

\[
\left| \frac{R_0}{U_0} \right| < 1 \quad \text{and} \quad \left| \frac{L_0}{V_0} \right| < \frac{1}{100}.
\]

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

\[
\left| \frac{L_0}{V_0} \right| = \frac{\beta_1 - \beta}{\beta_2 - \beta_1} < 0.008565 < \frac{1}{100}.
\]

Similarly, we have

\[
\alpha_1 \leq [0; 3, 1, 3, 1, 2, 1, 3, 4] \quad \text{and} \quad \alpha_2 \geq [0; 3, 1, 3, 1, 2, 1, 2, \overline{1, 3}],
\]

and

\[
\left| \frac{R_0}{U_0} \right| = \frac{\alpha_1 - \alpha}{\alpha_2 - \alpha_1} < 0.98479 < 1.
\]

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 \( t - R_n \subset V_m \) and, \textit{a fortiori}, \( \{t - A_n \cap B_m\} = \emptyset \) (see Figure 4.1); 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 \( (\alpha_n, \beta_n) \) 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

---

\(^2\)Actually, using the general distortion bound statement in Chapter 4 of Palis–Takens book \[16\], 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 \( n, m \in \mathbb{N}^* \).
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 

\[
m(3^{*21}112123) = 3.676699417246755742 \ldots
\]

**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 

\[
m(332221121223^*) = 3.726146224233042720 \ldots
\]

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 

\[
m(33222121223^*) = 3.726278993734881116 \ldots
\]

**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 

\[
m(3321112123233113^*) = 3.942001159911341469 \ldots
\]

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 \( a \in \{1, 2, 3, 4\}^\mathbb{Z} \) for which \( m(a) = \lambda_0(a) \in [l, n] \), for some interval \([l, n] \). The algorithm was implemented using the SageMath mathematical software [18].

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.

\[
\begin{align*}
\text{Appendix B. Pseudo-code for computer search} \\
\end{align*}
\]
Algorithm 1 - Find sequences whose Markov values could lie in the range $[l, n]$

1. $\textit{candidates} \leftarrow [1^*, 2^*, 3^*, 4^*]$
2. $\textit{forbidden} \_ \textit{words} \leftarrow []$
3. $\textit{alphabet} \leftarrow \{\_, 1, 2, 3, 4\}$  # $\_ \text{ is the empty string}$
4. $\textit{extensions} \leftarrow (\textit{alphabet} \times \textit{alphabet}) \setminus \{(\_, \_)\}$
5. $l \leftarrow l$
6. $n \leftarrow n$
7. $\textit{length\_limit} \leftarrow \text{maximum length of sequences to search up to}$
8. $\textit{min\_seq\_len} \leftarrow \text{minimum length of all sequences in } \textit{candidates}$
9. \textbf{while} $\textit{min\_seq\_len} < \textit{length\_limit} \text{ and } \textit{candidates} \neq [] \textbf{ do}$
   10. \textbf{for} sequence in $\textit{candidates}$ \textbf{do}$
       11. \textbf{remove} sequence from $\textit{candidates}$
       12. \textbf{for} $(x, y)$ in $\textit{extensions}$ \textbf{do}$
           13. $\textit{allowable} \leftarrow \textbf{True}$
           14. $\textit{trial\_sequence} \leftarrow \text{concatenation}(x, \textit{sequence}, y)$
           15. \textbf{if} $\textit{trial\_sequence}$ contains any words from $\textit{forbidden\_words}$\textbf{then}$
               16. \textbf{continue}  \# the sequence is forbidden so move on to the next
           17. \textbf{end if}$
           18. $\lambda_{\text{max}} \leftarrow \text{maximum possible value of } \lambda_0(\textit{trial\_sequence})$
           19. \textbf{if} $\lambda_{\text{max}} < l$ \textbf{then}$
               20. \textbf{continue}  \# $\lambda_0$ is too small so move on to the next sequence
           21. \textbf{end if}$
           22. \textbf{for} $z$ in $\textit{trial\_sequence}$ \textbf{do}$
               23. $j \leftarrow \text{position of } z \text{ in } \textit{trial\_sequence}$
               24. $\lambda_{\text{min}} \leftarrow \text{minimum possible value of } \lambda_j(\textit{trial\_sequence})$
               25. \textbf{if} $\lambda_{\text{min}} > n$ \textbf{then}$
                   26. append $\textit{trial\_sequence}$ to $\textit{forbidden\_words}$
                   27. $\textit{allowable} \leftarrow \textbf{False}$  \# the Markov value is too large
               28. \textbf{end if}$
           29. \textbf{end for}$
           30. \textbf{if} $\textit{allowable}$ \textbf{then}$  \# the Markov value can lie in $[l, n]$
               31. append $\textit{trial\_sequence}$ to $\textit{candidates}$
           32. \textbf{end if}$
       33. \textbf{end for}$
    34. \textbf{end for}$
35. \textbf{if} $\textit{candidates} \neq []$ \textbf{then}$
   36. $\textit{min\_seq\_len} \leftarrow \text{minimum length of all sequences in } \textit{candidates}$
37. \textbf{end if}$
38. \textbf{end while}$
39. \textbf{return} $\textit{candidates}$
References

[1] T. W. Cusick & M. E. Flahive, The Markoff and Lagrange spectra, Mathematical Surveys and Monographs, vol. 30, American Mathematical Society, 1989, ix+97 pages.
[2] G. A. Freiman, “Noncoincidence of the Markov and Lagrange spectra”, Mat. Zametki 3 (1968), p. 195-200, English translation in Math. Notes 3 (1968), p. 125-128.
[3] ———, “The initial point of Hall’s ray”, in Number-theoretic studies in the Markov spectrum and in the structural theory of set addition (Russian), Kalininiskii Gos. Univ., 1973, p. 87-120.
[4] ———, Diophantine approximations and geometry of numbers. (Markov problem) (Russian), Kalininiskii Gos. Univ., 1975, 144 pages.
[5] G. A. Freiman & A. A. Judin, “On the Markov spectrum (Russian)”, Litov. Mat. Sb. 6 (1966), p. 443-447.
[6] M. Hall, Jr, “On the sum and product of continued fractions”, Ann. Math. 48 (1947), p. 966-993.
[7] ———, “The Markoff spectrum”, Acta Arith. 18 (1971), p. 387-399.
[8] O. Jenkinson & M. Pollicott, “Computing the dimension of dynamically defined sets: $E_2$ and bounded continued fractions”, Ergodic Theory Dyn. Syst. 21 (2001), no. 5, p. 1429-1445.
[9] D. Lima, C. Matheus, C. G. Moreira & S. Romaña, Classical and dynamical Markov and Lagrange spectra. Dynamical, fractal and arithmetic aspects, World Scientific, 2020, xiii+213 pages.
[10] A. Markov, “Sur les formes quadratiques binaires indéfinies”, Math. Ann. 15 (1879), p. 381-406.
[11] ———, “Sur les formes quadratiques binaires indéfinies II”, Math. Ann. 17 (1880), p. 379-399.
[12] C. Matheus & C. G. Moreira, “Fractal geometry of the complement of Lagrange spectrum in Markov spectrum”, Comment. Math. Helv. 95 (2020), no. 3, p. 593-633.
[13] ———, “Diophantine approximation, Lagrange and Markov spectra, and dynamical Cantor sets”, Notices Am. Math. Soc. 68 (2021), no. 8, p. 1301-1311.
[14] C. Matheus, C. G. Moreira, M. Pollicott & P. Vytnova, “Hausdorff dimension of Gauss-Cantor sets and two applications to classical Lagrange and Markov spectra”, Adv. Math. 409 (2022), article no. 108693 (70 pages).
[15] C. G. Moreira, “Stable intersections of Cantor sets and homoclinic bifurcations”, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 13 (1996), no. 6, p. 741-781.
[16] J. Palis & F. Takens, Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations, fractal dimensions and infinitely many attractors, Cambridge Studies in Advanced Mathematics, vol. 35, Cambridge University Press, 1993.
[17] O. Perron, “Über die Approximation irrationaler Zahlen durch rationale II”, Sitzungsber. Heidelb. Akad. Wiss. 1921 (1921), no. 4.
[18] The Sage Developers, “SageMath, the Sage Mathematics Software System (Version 9.4)”, 2021, http://www.sagemath.org/.
[19] H. Schick, “Über die Menge der Zahlen, die als Minima quadratischer Formen auftreten”, J. Number Theory 9 (1977), no. 1, p. 121-141.
Luke Jeffreys
School of Mathematics, University of Bristol, Fry Building,
Woodland Road, Bristol BS8 1UG, UK
E-mail: luke.jeffreys@bristol.ac.uk
URL: https://seis.bristol.ac.uk/~pk20963/

Carlos Matheus
Centre de Mathématiques Laurent Schwartz, École Polytechnique,
91128 Palaiseau Cedex, France
E-mail: carlos.matheus@math.cnrs.fr
URL: http://carlos.matheus.perso.math.cnrs.fr/index.html

Carlos Gustavo Moreira
SUSTech International Center for Mathematics
Shenzhen, Guangdong, People's Republic of China
IMPA, Estrada Dona Castorina 110,
CEP 22460-320, Rio de Janeiro, Brazil
E-mail: gugu@impa.br
URL: https://w3.impa.br/~gugu/