ACTA ARITHMETICA
Online First version

The Markov and Lagrange spectra on the Hecke group $\mathbf{H}_4$

by

DONG HAN KIM and DEOKWON SIM

Abstract. We consider the Markov spectrum and the Lagrange spectrum on the Hecke group $\mathbf{H}_4$. They are identical to the Markov and Lagrange spectra on the unit circle. The Markov spectrum on $\mathbf{H}_4$ is termed the Markov spectrum on index 2 sublattices by Vulakh and the Markov spectrum on 2-minimal forms or C-minimal forms by Schmidt. They characterized the spectrum up to the first accumulation point, independently. We show that, after the first accumulation point, both spectra have positive Hausdorff dimension. Then we find gaps in the spectra and give a bound on Hall's ray.

1. Introduction. For an irrational number ξ, the *Lagrange number* $L(\xi)$ is defined as the supremum of all L such that

$$\left|\xi - \frac{p}{q}\right| < \frac{1}{Lq^2}$$

for infinitely many rational numbers p/q. The classical *Lagrange spectrum* is the set of Lagrange numbers, i.e.,

$$\mathscr{L}_0 := \left\{ \limsup_{p/q\in\mathbb{Q}} \left(q^2 \left|\xi - \frac{p}{q}\right| \right)^{-1} \,\middle|\, \xi \in \mathbb{R}\setminus\mathbb{Q} \right\} = \{L(\xi) \mid \xi \in \mathbb{R}\setminus\mathbb{Q}\}. \tag{1.1}$$

The *Markov spectrum* is defined as the set of reciprocals of the infimum of the non-zero values of indefinite quadratic forms $f(x,y) = ax^2 + bxy + cy^2$ with real coefficients, normalized by the square root of their discriminants $\delta(f) = b^2 - 4ac > 0$, i.e.,

$$\mathscr{M}_0 := \left\{ \left(\inf_{(x,y)\in\mathbb{Z}^2\setminus\{(0,0)\}} \frac{|f(x,y)|}{\sqrt{\delta(f)}} \right)^{-1} \,\middle|\, \delta(f) > 0 \right\}. \tag{1.2}$$

2020 *Mathematics Subject Classification*: Primary 11J06; Secondary 11J70.
Key words and phrases: Lagrange spectrum, Markov spectrum, Hecke group, Diophantine approximation on the circle.
Received 27 August 2025; revised 11 February 2026.
Published online 9 June 2026.

DOI: 10.4064/aa250827-16-2

It is well known [31] that $\mathscr{L}_0 \subset \mathscr{M}_0$. The classical results of Markov [19, 20] show that

$$\mathscr{L}_0 \cap [0,3) = \mathscr{M}_0 \cap [0,3) = \left\{ \sqrt{9 - \frac{4}{x^2}} \,\middle|\, x \in \mathcal{M}_0 \right\},$$

where $\mathcal{M}_0 = \{1, 2, 5, 13, 29, 34, 89, 169, \dots\}$ is the set of elements of positive integer triples (x_1, x_2, x_3) satisfying

$$x_1^2 + x_2^2 + x_3^2 = 3x_1x_2x_3.$$

Therefore, the smallest accumulation point of the spectra is 3. Moreira [21] showed that the two spectra have positive Hausdorff dimension right after the first accumulation point 3 and they have full dimension starting at $\sqrt{12} - \delta$ for some $\delta > 0$. There are gaps in $\mathscr{L}_0$ and $\mathscr{M}_0$ like $(\sqrt{12}, \sqrt{13})$, which was found by Perron [25]. Note that $\sqrt{13}$ is an isolated point of both spectra. Eventually, there exists a half-infinite interval contained in the Lagrange and Markov spectra, which is called *Hall's ray* [12]. Hall showed that $(6, \infty) \subset \mathscr{L}_0$ and Freĭman [10] gave the smallest possible value $c = \frac{2221564069 + 283748\sqrt{462}}{491993569} = 4.5278\dots$ for which $[c, \infty)$ is contained in $\mathscr{L}_0$. For a detailed discussion of the Markov and Lagrange spectra, see [3, 8, 17].

The Lagrange and Markov spectra have been generalized to discrete subgroups of $\mathrm{PSL}_2(\mathbb{R})$, called *Fuchsian groups*. Let $\mathbf{G}$ be a finitely generated Fuchsian group acting on the upper half-plane $\mathbb{H}$ and its boundary $\hat{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ via the linear fractional transformation

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.$$

We further assume that ∞ is a fixed point of a parabolic element of $\mathbf{G}$ and let $\mathbb{Q}(\mathbf{G})$ be the set of orbits of ∞ under the action of $\mathbf{G}$. For a real number ξ not in $\mathbb{Q}(\mathbf{G})$, we define the *Lagrange number* $L_{\mathbf{G}}(\xi)$ to be the supremum of L satisfying

$$|\xi - M \cdot \infty| = \left| \xi - \frac{a}{c} \right| < \frac{1}{Lc^2} \qquad \text{for infinitely many } M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{G}.$$

Since

$$|M^{-1} \cdot \xi - M^{-1} \cdot \infty| = \frac{1}{c^2 |\xi - a/c|},$$

$L_{\mathbf{G}}(\xi)$ is the limit superior of $|M^{-1} \cdot \xi - M^{-1} \cdot \infty|$, which is the Euclidean diameter of the image of the geodesic from ∞ to ξ under the action of $M^{-1} \in \mathbf{G}$. We define the *Lagrange spectrum* on $\mathbf{G}$ as

$$\mathscr{L}(\mathbf{G}) = \{ L_{\mathbf{G}}(\xi) \mid \xi \in \mathbb{R} \setminus \mathbb{Q}(\mathbf{G}) \}. \tag{1.3}$$

Let $f(x, y) = ax^2 + bxy + cy^2$ be an indefinite quadratic form with real coefficients. For each quadratic form f, we associate a geodesic in $\mathbb{H}$ with endpoints $\xi, \eta \in \hat{\mathbb{R}}$, $\xi \neq \eta$, satisfying

$$\frac{|f(x,y)|}{\sqrt{\delta(f)}} = \frac{|(x-\xi y)(x-\eta y)|}{|\xi-\eta|}.$$

For a matrix $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{G}$, we set $f(M) := f(a,c)$ and check that

$$M\cdot\xi - M\cdot\eta = \frac{\xi-\eta}{(c\xi+d)(c\eta+d)} \qquad \text{for } \xi,\eta\in\hat{\mathbb{R}}. \tag{1.4}$$

Therefore, we have

$$\frac{\sqrt{\delta(f)}}{|f(M)|} = |M^{-1}\cdot\xi - M^{-1}\cdot\eta|.$$

We define the *Markov spectrum* on $\mathbf{G}$ as

$$\begin{aligned}\mathscr{M}(\mathbf{G}) :=& \left\{ \sup_{M\in\mathbf{G}} \frac{\sqrt{\delta(f)}}{|f(M)|} \;\middle|\; \delta(f)>0 \right\}\\ =& \left\{ \sup_{M\in\mathbf{G}} \left|M^{-1}\cdot\xi - M^{-1}\cdot\eta\right| \;\middle|\; \xi,\eta\in\hat{\mathbb{R}},\ \xi\neq\eta \right\}.\end{aligned}$$

In other words, $\mathscr{M}(\mathbf{G})$ is the set of the supremums of the Euclidean diameters of the geodesics in $\mathbb{H}$ under the action of $\mathbf{G}$. Note that the Lagrange spectrum on $\mathbf{G}$ satisfies

$$\mathscr{L}(\mathbf{G}) = \left\{ \limsup_{M\in\mathbf{G}} \left|M^{-1}\cdot\xi - M^{-1}\cdot\infty\right| \;\middle|\; \xi\in\mathbb{R}\setminus\mathbb{Q}(\mathbf{G}) \right\}.$$

For the modular group $\mathrm{PSL}_2(\mathbb{Z})$, we have

$$\mathscr{M}(\mathrm{PSL}_2(\mathbb{Z})) = \mathscr{M}_0, \qquad \mathscr{L}(\mathrm{PSL}_2(\mathbb{Z})) = \mathscr{L}_0.$$

Some closed geodesics in $\mathbb{H}/\mathrm{PSL}_2(\mathbb{Z})$ with low heights are shown in Figure 1.

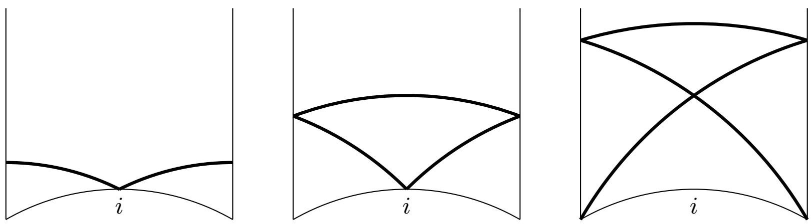

Fig. 1. Several closed geodesics in the fundamental domain of the modular group on the upper-half plane. They have maximal heights $\sqrt{5}$, $2\sqrt{2}$, $2\sqrt{3}$ (from left to right).

In this paper, we consider the Lagrange and Markov spectra on the *Hecke group* $\mathbf{H}_4$ or the hyperbolic triangle group $(2,4,\infty)$. The Hecke group $\mathbf{H}_q$ is the subgroup of $\mathrm{PSL}_2(\mathbb{R})$ generated by $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & \lambda_q \\ 0 & 1 \end{smallmatrix}\right)$, where $\lambda_q = 2\cos\frac{\pi}{q}$ and $q \geq 3$ is an integer. Note that $\mathbf{H}_q$ has the presentation

$$\mathbf{H}_q \cong \langle S, T \mid S^2 = I,\ (ST)^q = I\rangle,$$

where I is the identity 2 by 2 matrix. When $q = 3$, we have $\lambda_3 = 1$ and $\mathbf{H}_3$ is the modular group $\mathrm{PSL}_2(\mathbb{Z})$. If $q = 4$, then $\lambda_4 = \sqrt{2}$. Moreover, it is known [24] that

$$\mathbf{H}_4 = \left\{ \begin{pmatrix} a & \sqrt{2}\,b \\ \sqrt{2}\,c & d \end{pmatrix} \,\middle|\, ad - 2bc = 1,\ a, b, c, d \in \mathbb{Z} \right\} \cup \left\{ \begin{pmatrix} \sqrt{2}\,a & b \\ c & \sqrt{2}\,d \end{pmatrix} \,\middle|\, 2ad - bc = 1,\ a, b, c, d \in \mathbb{Z} \right\}.$$

Therefore, $\mathbb{Q}(\mathbf{H}_4) = \sqrt{2}\,\mathbb{Q}$. The Diophantine approximation on the Hecke group $\mathbf{H}_q$ has also been studied using the Rosen continued fraction [27] (see e.g. [16, 28, 22, 5, 23, 4]). The three geodesics of lowest heights in $\mathbb{H}/\mathbf{H}_4$ are shown in Figure 2.

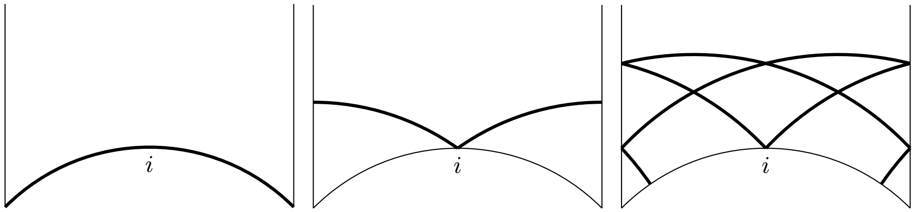

Fig. 2. Three closed geodesics in the fundamental domain of the group $\mathbf{H}_4$ on the upper half-plane with lowest heights.

The minimum of the Lagrange spectrum, which is called *Hurwitz's constant*, for the Hecke group $\mathbf{H}_q$ was studied in [16] and [11]. In particular, if q is even, then the minimum of the Lagrange spectrum $\mathscr{L}(\mathbf{H}_q)$ is always equal to 2. Series [30] examined the discrete part of the Markov spectrum on $\mathbf{H}_5$. The Markov spectra on general Hecke groups were studied in [33].

The discrete part of the Markov spectrum on the Hecke group $\mathbf{H}_4$ has been studied by Schmidt and Vulakh independently. It is termed the Markov spectrum on 2-minimal forms by Schmidt [29] and the Markov spectrum on the sublattice of index 2 by Vulakh ([32]; see also [18]). It is also identical to the Markov spectrum on the unit circle S^1, related to the Diophantine approximation of the points of $\mathbb{R}^2 \cap S^1$ by points of $\mathbb{Q}^2 \cap S^1$ (see [15, 7]). We will call $(x; y_1, y_2)$ a *Vulakh–Schmidt triple* if $(x; y_1, y_2)$ is a positive integer triple satisfying

$$2x^2 + y_1^2 + y_2^2 = 4xy_1y_2.$$

We set

$$\mathcal{N} = \{1, 5, 29, 65, 169, 349, \dots\} \quad \text{and} \quad \mathcal{M} = \{1, 3, 11, 17, 41, 59, \dots\}$$

as the sets of x's and y_i's ($i = 1, 2$) in the Vulakh–Schmidt triples respectively. The spectral values less than $2\sqrt{2}$ are given in [29] (see also [18]) as

$$\mathscr{M}(\mathbf{H}_4) \cap [0, 2\sqrt{2}) = \left\{ \sqrt{8 - \frac{2}{x^2}} \;\middle|\; x \in \mathcal{N} \right\} \cup \left\{ \sqrt{8 - \frac{4}{y^2}} \;\middle|\; y \in \mathcal{M} \right\}.$$

Therefore, the first accumulation point of $\mathscr{M}(\mathbf{H}_4)$ is $2\sqrt{2}$. The discrete part of the Lagrange spectrum $\mathscr{L}(\mathbf{H}_4) \cap [0, 2\sqrt{2})$ coincides with the discrete part of the Markov spectrum $\mathscr{M}(\mathbf{H}_4) \cap [0, 2\sqrt{2})$ (see also [6]). Using a method similar to the classical case, we show the first theorem.

Theorem 1.1. *The Markov spectrum $\mathscr{M}(\mathbf{H}_4)$ is closed and the Lagrange spectrum $\mathscr{L}(\mathbf{H}_4)$ is contained in $\mathscr{M}(\mathbf{H}_4)$.*

We show that, after the first accumulation point, the Lagrange spectrum has positive Hausdorff dimension.

Theorem 1.2. *For any $\varepsilon > 0$, we have*

$$\dim_H \left(\mathscr{M}(\mathbf{H}_4) \cap [0, 2\sqrt{2} + \varepsilon) \right) \geq \dim_H \left(\mathscr{L}(\mathbf{H}_4) \cap [0, 2\sqrt{2} + \varepsilon) \right) > 0.$$

We call an open interval (a, b) a *maximal gap of the spectrum* if it does not intersect the spectrum and is not a proper subset of a larger gap. We find two maximal gaps in $\mathscr{M}(\mathbf{H}_4)$ and $\mathscr{L}(\mathbf{H}_4)$ after the first accumulation point (see Figure 3).

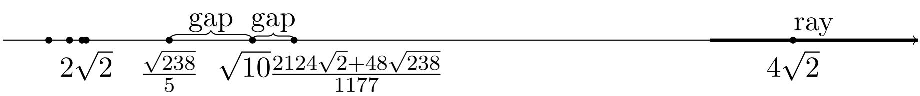

Fig. 3. Gaps and a ray in $\mathscr{M}(\mathbf{H}_4)$

Theorem 1.3. *The intervals $\left(\frac{\sqrt{238}}{5}, \sqrt{10}\right)$ and $\left(\sqrt{10}, \frac{2124\sqrt{2}+48\sqrt{238}}{1177}\right)$ are maximal gaps in $\mathscr{M}(\mathbf{H}_4)$ and $\mathscr{L}(\mathbf{H}_4)$.*

We note that $\sqrt{10}$ is an isolated point. The two gaps in Theorem 1.3 seem to be similar to the gaps $(\sqrt{12}, \sqrt{13})$ and $(\sqrt{13}, \frac{1}{22}(9\sqrt{3} + 65))$ in the classical Markov and Lagrange spectra $\mathscr{M}_0$ and $\mathscr{L}_0$ [8, Chapter 1, Lemmas 7 and 9].

After a certain point the Lagrange spectrum $\mathscr{L}(\mathbf{H}_4)$ contains a half-line, and so does $\mathscr{M}(\mathbf{H}_4)$, which is called *Hall's ray* (see Figure 3). The existence of Hall's ray in $\mathscr{L}(\mathbf{H}_4)$ is established [1] in general groups. We give a bound of Hall's ray as follows.

Theorem 1.4. *The Lagrange spectrum $\mathscr{L}(\mathbf{H}_4)$ contains every real number greater than or equal to $4\sqrt{2}$, i.e. $[4\sqrt{2}, \infty) \subset \mathscr{L}(\mathbf{H}_4) \subset \mathscr{M}(\mathbf{H}_4)$.*

In Section 2, we introduce a symbolic coding for a geodesic and its endpoints in $\hat{\mathbb{R}}$ using the Hecke group $\mathbf{H}_4$. A geodesic in the hyperbolic space is determined by a doubly infinite expansion and we deduce the formula for the spectral value of the Markov and Lagrange spectra by the doubly infinite expansion of the geodesic. We then prove Theorems 1.1–1.4 in Sections 3–6 respectively.

2. Symbolic coding of a geodesic and the Perron formula. In this section, we introduce a symbolic coding for a geodesic and its endpoints using the Hecke group $\mathbf{H}_4$, following the work of Hass and Serie [11] and Series [30]. We then derive the Perron formula (Theorems 2.5 and 2.6) for $\mathbf{H}_4$ using this expansion. The $\mathbf{H}_4$-expansion is closely related to the digit expansion on the unit circle introduced by Romik [26], which is also related to the even integer continued fraction or continued fractions of specific parities (see [13, 14]). For the connection between the $\mathbf{H}_4$-expansion and the Rosen continued fraction, consult [2].

Let

$$T = \begin{pmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad K = ST^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & \sqrt{2} \end{pmatrix}.$$

Note that $K^4 = I$. We consider a fundamental domain Ω for $\mathbf{H}_4$ surrounded by the geodesics given by $x = 0$, $x = \sqrt{2}$, $|z| = 1$ and $|z - \sqrt{2}| = 1$ (Figure 4 (left)). Let $\boldsymbol{\delta}_0$ be the geodesic given by the imaginary axis and let $\boldsymbol{\delta}_d = K^d(\boldsymbol{\delta}_0)$ for $d = 1, 2, 3$. Let $\Delta = \Omega \cup K(\Omega) \cup K^2(\Omega) \cup K^3(\Omega)$ be the ideal quadrilateral with edges $\boldsymbol{\delta}_d$ for $d = 0, 1, 2, 3$ (Figure 4 (right)). Let $\boldsymbol{\Gamma}_4$ be the subgroup of $\mathbf{H}_4$ generated by $K^d S K^{-d}$, $d = 0, 1, 2, 3$. Then Δ is a fundamental domain of $\boldsymbol{\Gamma}_4$ (see [11]).

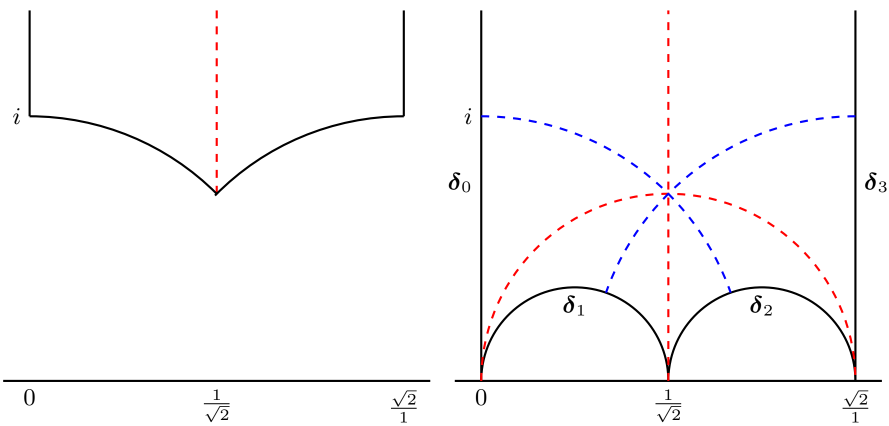

Fig. 4. A fundamental domain Ω (left) and the ideal quadrilateral Δ (right)

Let $\boldsymbol{\gamma}$ be an oriented geodesic with endpoints $\boldsymbol{\gamma}^-, \boldsymbol{\gamma}^+ \in \hat{\mathbb{R}}$. We assume that neither $\boldsymbol{\gamma}^-$ nor $\boldsymbol{\gamma}^+$ belongs to $\mathbb{Q}(\mathbf{H}_4)$. Let $\mathscr{T} = \bigcup_{G \in \mathbf{H}_4} G(\boldsymbol{\delta}_0)$. Then, by cutting $\mathscr{T}$, the oriented geodesic $\boldsymbol{\gamma}$ is divided into geodesic segments

$\dots, \boldsymbol{\gamma}_{-2}, \boldsymbol{\gamma}_{-1}, \boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots$ along the orientation. Let $\gamma_n^-, \gamma_n^+ \in \mathscr{T}$ be the two endpoints of the geodesic segment $\boldsymbol{\gamma}_n$ along the orientation of $\boldsymbol{\gamma}$.

For each $n \in \mathbb{Z}$, there exists $M_n \in \boldsymbol{\Gamma}_4$ such that $\boldsymbol{\gamma}_n$ belongs to $M_n(\Delta)$. Let $e_n \in \{0,1,2,3\}$ be such that $\gamma_n^- \in M_n(\boldsymbol{\delta}_{e_n})$ and define $G_n = M_n K^{e_n}$. Then $G_n \in \mathbf{H}_4$ and

$$\gamma_n^- \in G_n(\boldsymbol{\delta}_0) \quad \text{and} \quad \gamma_n^+ \in G_n(\boldsymbol{\delta}_{d_n}) \quad \text{for some } d_n \in \{1,2,3\}. \tag{2.1}$$

Since

$$\gamma_{n+1}^- = \gamma_n^+ \in G_n(\boldsymbol{\delta}_{d_n}) = G_n K^{d_n}(\boldsymbol{\delta}_0) = G_n K^{d_n} S(\boldsymbol{\delta}_0),$$

we deduce that for all $n \in \mathbb{Z}$,

$$G_{n+1} = G_n K^{d_n} S = G_n N_{d_n}, \tag{2.2}$$

where we denote

$$N_d := K^d S \quad \text{for } d = 1,2,3.$$

Note that

$$N_1 = \begin{pmatrix} 1 & 0 \\ \sqrt{2} & 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} \sqrt{2} & 1 \\ 1 & \sqrt{2} \end{pmatrix}, \quad N_3 = \begin{pmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{pmatrix}. \tag{2.3}$$

For each oriented geodesic $\boldsymbol{\gamma}$ on $\mathbb{H}$, we define a two-sided infinite sequence $(d_n)_{n\in\mathbb{Z}} \in \{1,2,3\}^{\mathbb{Z}}$. We define an equivalence relation $\sim$ in $\{1,2,3\}^{\mathbb{Z}}$ by $(a_n) \sim (b_n)$ if and only if there exists some $m \in \mathbb{Z}$ such that $a_{n+m} = b_n$ for all $n \in \mathbb{Z}$. Then an equivalence class of $\{1,2,3\}^{\mathbb{Z}}$ under $\sim$ is called a *doubly infinite* $\mathbf{H}_4$*-sequence*. A *section* of a doubly infinite $\mathbf{H}_4$-sequence is an element $(d_n)_{n\in\mathbb{Z}} \in \{1,2,3\}^{\mathbb{Z}}$ in the equivalence class. With each oriented geodesic $\boldsymbol{\gamma}$ on $\mathbb{H}$, we associate a doubly infinite $\mathbf{H}_4$-sequence. Figure 5 shows an example of an oriented geodesic $\boldsymbol{\gamma}$ with $G_1 = I$ and

$$\dots,\ d_{-1} = 1,\ d_0 = 2,\ d_1 = 3,\ d_2 = 1,\ d_3 = 3,\ \dots$$

thus,

$$\begin{aligned} G_{-1} &= N_2^{-1} N_1^{-1}, & G_1 &= I, & G_3 &= N_3 N_1, \\ G_0 &= N_2^{-1}, & G_2 &= N_3, & G_4 &= N_3 N_1 N_3. \end{aligned}$$

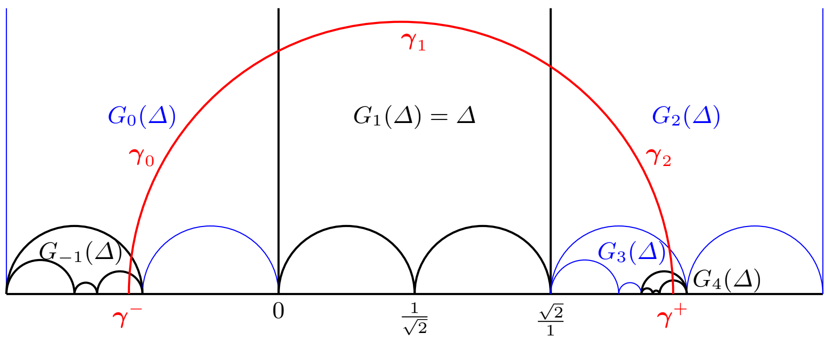

Fig. 5. An oriented geodesic $\boldsymbol{\gamma}$ with a sequence of geodesic segments $\boldsymbol{\gamma}_n$.

From (2.1), we deduce that for each $n \in \mathbb{Z}$ the oriented geodesic $G_n^{-1}(\boldsymbol{\gamma})$ intersects the imaginary axis $\boldsymbol{\delta}_0$ and satisfies

$$G_n^{-1}(\boldsymbol{\gamma}^-) \in (-\infty, 0) \quad \text{and} \quad G_n^{-1}(\boldsymbol{\gamma}^+) \in (0, \infty). \tag{2.4}$$

Suppose that $G_1 = I$. Then we have $\boldsymbol{\gamma}^+ \in (0, \infty)$, and by (2.2), we obtain $G_{n+1} = N_{d_1} \dots N_{d_n}$ for $n \geq 0$. Therefore, (2.4) implies that for all $n \geq 1$,

$$\boldsymbol{\gamma}^+ \in N_{d_1} \dots N_{d_n} \cdot (0, \infty).$$

Using the symbolic coding of the geodesic, we have an expansion of a positive real number by the one-sided infinite sequence $(d_n)_{n \in \mathbb{N}}$. Let $f : [0, \infty] \to [0, \infty]$ be the map given by

$$f(x) = \begin{cases} N_1^{-1} \cdot x & \text{if } x \in \left[0, \frac{1}{\sqrt{2}}\right] = N_1 \cdot [0, \infty], \\ N_2^{-1} \cdot x & \text{if } x \in \left[\frac{1}{\sqrt{2}}, \sqrt{2}\,\right] = N_2 \cdot [0, \infty], \\ N_3^{-1} \cdot x & \text{if } x \in \left[\sqrt{2}, \infty\right] = N_3 \cdot [0, \infty]. \end{cases}$$

For a real number $\alpha \in [0, \infty]$, there exists an infinite sequence $(d_n)_{n \in \mathbb{N}}$ satisfying

$$f^{n-1}(x) \in N_{d_n} \cdot [0, \infty] \qquad \text{for all } \; n \geq 1,$$

thus

$$x \in N_{d_1} \cdots N_{d_n} \dots [0, \infty] \qquad \text{for all } \; n \geq 1.$$

We define the $\mathbf{H}_4$-*expansion* of α as

$$\alpha = [d_1, d_2, \dots].$$

Remark 2.1. For $q = 3$ we have $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\mathbf{H}_3 = \mathrm{PSL}_2(\mathbb{Z})$. In this case,

$$N_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Thus, the $\mathbf{H}_3$-expansion of α is

$$\alpha = [\overbrace{2, \dots, 2}^{a_0}, \overbrace{1, \dots 1}^{a_1}, \overbrace{2, \dots, 2}^{a_2}, \dots] \quad \text{for } \alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}.$$

By an *infinite* $\mathbf{H}_4$-*sequence* we mean an element of $\{1, 2, 3\}^{\mathbb{N}}$. For an infinite $\mathbf{H}_4$-sequence $P = (d_n)_{n \geq 1}$, we write $[P] = [d_1, d_2, \dots]$. For $d_1, \dots, d_k \in \{1, 2, 3\}$, we define the *cylinder set*

$$[d_1, \dots, d_k] := N_{d_1} \dots N_{d_k} \cdot [0, \infty].$$

For $\alpha = [d_1, d_2, \dots]$, we have $\alpha \in [d_1, \dots, d_n]$ for all $n \geq 1$. Some cylinder sets of the $\mathbf{H}_4$-expansion are shown in Figure 6. We note that for each $k \geq 1$,

$$[d_1, d_2, \dots] = N_{d_1} \dots N_{d_k} \cdot [d_{k+1}, d_{k+2}, \dots].$$

In particular, we check

$$[1, P] = N_1 \cdot [P], \quad [2, P] = N_2 \cdot [P], \quad [3, P] = N_3 \cdot [P], \tag{2.5}$$

and deduce that

$$0 \le [1, P] \le \frac{1}{\sqrt{2}}, \quad \frac{1}{\sqrt{2}} \le [2, P] \le \sqrt{2}, \quad \sqrt{2} \le [3, P]$$

for $P \in \{1, 2, 3\}^{\mathbb{N}}$.

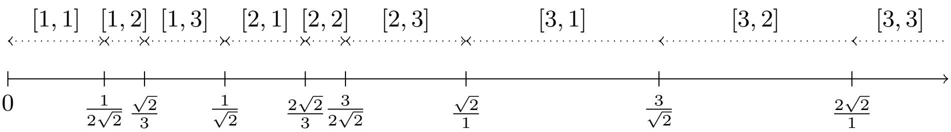

Fig. 6. Cylinder sets on $\mathbb{R}$

Since $\mathbf{H}_4$ is generated by S and K, any $M \in \mathbf{H}_4$ takes one of the following forms:

$$N_{d_1} \dots N_{d_k} \quad \text{or} \quad N_{d_1} \dots N_{d_k} S \quad \text{or} \quad S N_{d_1} \cdots N_{d_k} \quad \text{or} \quad S N_{d_1} \dots N_{d_k} S.$$

Therefore, $\alpha \in [0, \infty]$ belongs to $\mathbb{Q}(\mathbf{H}_4)$ if and only if

$$\alpha = N_{d_1} \cdots N_{d_k} \cdot 0 \quad \text{or} \quad \alpha = N_{d_1} \cdots N_{d_k} \cdot \infty,$$

which is equivalent to α being a boundary point of the cylinder set $[d_1, \dots, d_k]$. If $\alpha \in \mathbb{R} \setminus \mathbb{Q}(\mathbf{H}_4)$, then α has a unique $\mathbf{H}_4$-expansion $[d_1, d_2, \dots]$. For the boundary points of the cylinder set, we have

$$0 = [1, 1, \dots] =: [1^\infty], \quad \infty = [3, 3, \dots] =: [3^\infty]$$

and

$$[d_1, \dots, d_k] = \big[[d_1, \dots, d_k, 1^\infty],\, [d_1, \dots, d_k, 3^\infty]\big].$$

Therefore, if $\alpha \in \mathbb{Q}(\mathbf{H}_4)$, then there exist up to two expressions of α. For instance,

$$\frac{1}{\sqrt{2}} = [1, 3^\infty] = [2, 1^\infty], \quad \sqrt{2} = [2, 3^\infty] = [3, 1^\infty].$$

EXAMPLE 2.2. As $[2^\infty] = N_2 \cdot [2^\infty]$, $[(1,3)^\infty] = N_1 N_3 \cdot [(1,3)^\infty]$, we have

$$[2^\infty] = 1, \quad [(1,3)^\infty] = \frac{\sqrt{3} - 1}{\sqrt{2}}. \tag{2.6}$$

Similarly, we check

$$[(1,2,3)^\infty] = \frac{\sqrt{17} - 2\sqrt{2}}{3}, \quad [(1,1,2)^\infty] = \frac{1}{\sqrt{7} + \sqrt{2}}. \tag{2.7}$$

For infinite $\mathbf{H}_4$-sequences $P = (a_n)_{n\ge 1}$ and $Q = (b_n)_{n\ge 1}$, we define a combined two-sided sequence

$$P^*|Q := (c_n)_{n\in\mathbb{Z}}, \qquad c_n = \begin{cases} b_n & \text{if } n \ge 1, \\ a_{-n+1} & \text{if } n \le 0, \end{cases}$$

which is an element of $\{1,2,3\}^{\mathbb{Z}}$. Let

$$d^\vee = \begin{cases} 3 & \text{if } d = 1, \\ 2 & \text{if } d = 2, \\ 1 & \text{if } d = 3. \end{cases}$$

Then we have the identities

$$N_d^{-1} = S N_{d^\vee} S \quad \text{and} \quad N_{d^\vee} = J N_d J, \quad \text{where } J = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.8}$$

For a given infinite $\mathbf{H}_4$-sequence $P = (a_n)_{n\ge 1}$, let $P^\vee = (a_n^\vee)_{n\ge 1}$. For a doubly infinite $\mathbf{H}_4$-sequence U with a section $P^*|Q$, we define $U^\vee$ and U^* as the doubly infinite $\mathbf{H}_4$-sequences with a section $(P^\vee)^*|Q^\vee$ and $Q^*|P$ respectively. Using (2.8), we have

$$[d_1^\vee, \dots, d_k^\vee] = J N_{d_1} \dots N_{d_k} J \cdot [0,\infty] = \left[\frac{1}{[d_1,\dots,d_k,3^\infty]}, \frac{1}{[d_1,\dots,d_k,1^\infty]} \right]$$

and

$$[P^\vee] = \frac{1}{[P]}.$$

For an example, from (2.6), we have

$$[(3,1)^\infty] = \frac{1}{[(1,3)^\infty]} = \frac{\sqrt{2}}{\sqrt{3}-1} = \frac{\sqrt{3}+1}{\sqrt{2}}.$$

We also note that

$$[(3,1)^\infty] = N_3 \cdot [(1,3)^\infty] = \sqrt{2} + [(1,3)^\infty] = \frac{\sqrt{3}+1}{\sqrt{2}}.$$

LEMMA 2.3. *Suppose that $\boldsymbol{\gamma}$ is an oriented geodesic on $\mathbb{H}$ with an associated doubly infinite $\mathbf{H}_4$-sequence U.*

(i) *If $\boldsymbol{\gamma}^- \in (-\infty, 0)$ and $\boldsymbol{\gamma}^+ \in (0,\infty)$, then there exists a section $P^*|Q$ of U with $P = (a_n)_{n\in\mathbb{N}}$, $Q = (b_n)_{n\in\mathbb{N}}$ satisfying*

$$\boldsymbol{\gamma}^- = -[P] = -[a_1, a_2, \dots] \quad \text{and} \quad \boldsymbol{\gamma}^+ = [Q] = [b_1, b_2, \dots].$$

(ii) *If $\boldsymbol{\gamma}^- \in (0,\infty)$ and $\boldsymbol{\gamma}^+ \in (-\infty, 0)$, then there exists a section $P^*|Q$ of U such that*

$$\boldsymbol{\gamma}^+ = -[P^\vee] = -[a_1^\vee, a_2^\vee, \dots] \quad \text{and} \quad \boldsymbol{\gamma}^- = [Q^\vee] = [b_1^\vee, b_2^\vee, \dots].$$

Proof. (i) We choose geodesic segments $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$ in $S(\varDelta)$ and $\varDelta$ respectively, thus, $\boldsymbol{\gamma}_1^- \in \boldsymbol{\delta}_0$ and $G_1 = I$. By (2.2) and (2.8) we obtain

$$G_n = \begin{cases} N_{d_1} \dots N_{d_{n-1}} & \text{if } n \geq 2, \\ SN_{d_0^\vee} N_{d_{-1}^\vee} \dots N_{d_n^\vee} S & \text{if } n \leq 0. \end{cases}$$

Therefore, by (2.4), for all $m \geq 1$ we have

$$\boldsymbol{\gamma}^+ \in N_{d_1} \dots N_{d_m} \cdot (0, \infty), \qquad S \cdot \boldsymbol{\gamma}^- \in N_{d_0^\vee} N_{d_{-1}^\vee} \dots N_{d_{-m}^\vee} \cdot (0, \infty),$$

which yields

$$\boldsymbol{\gamma}^+ = [d_1, d_2, \dots], \qquad \boldsymbol{\gamma}^- = -\frac{1}{[d_0^\vee, d_{-1}^\vee, d_{-2}^\vee, \dots]} = -[d_0, d_{-1}, d_{-2}, \dots].$$

(ii) We choose geodesic segments $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_1$ in $\varDelta$ and $S(\varDelta)$ respectively, thus, $\boldsymbol{\gamma}_1^- \in S(\boldsymbol{\delta}_0) = \boldsymbol{\delta}_0$ and $G_1 = S$. By (2.2) and (2.8) we get

$$G_n = \begin{cases} SN_{d_1} \dots N_{d_{n-1}} & \text{if } n \geq 2, \\ N_{d_0^\vee} N_{d_{-1}^\vee} \dots N_{d_n^\vee} S & \text{if } n \leq 0. \end{cases}$$

Therefore, by (2.4), we have for all $m \geq 1$

$$S \cdot \boldsymbol{\gamma}^+ \in N_{d_1} \dots N_{d_m} \cdot (0, \infty), \qquad \boldsymbol{\gamma}^- \in N_{d_0^\vee} N_{d_{-1}^\vee} \dots N_{d_{-m}^\vee} \cdot (0, \infty),$$

which implies that

$$\boldsymbol{\gamma}^+ = -\frac{1}{[d_1, d_2, \dots]} = -[d_1^\vee, d_2^\vee, \dots], \qquad \boldsymbol{\gamma}^- = [d_0^\vee, d_{-1}^\vee, d_{-2}^\vee, \dots]. \ \blacksquare$$

Lemma 2.4. *Let $\boldsymbol{\gamma}$ be an oriented geodesic on $\mathbb{H}$ with two endpoints $\boldsymbol{\gamma}^-$, $\boldsymbol{\gamma}^+$ and let U be the doubly infinite $\mathbf{H}_4$-sequence associated to $\boldsymbol{\gamma}$. If $|M \cdot \boldsymbol{\gamma}^+ - M \cdot \boldsymbol{\gamma}^-| > \sqrt{2}$ for some $M \in \mathbf{H}_4$, then there exists a section $P^*|Q$ of U such that*

$$|M \cdot \boldsymbol{\gamma}^+ - M \cdot \boldsymbol{\gamma}^-| = [P] + [Q] \quad \text{or} \quad [P^\vee] + [Q^\vee].$$

Proof. Suppose that

$$|M \cdot \boldsymbol{\gamma}^+ - M \cdot \boldsymbol{\gamma}^-| > \sqrt{2}.$$

By writing $M' = T^m M$ for some $m \in \mathbb{Z}$, we may assume that

$$M' \cdot \boldsymbol{\gamma}^- < 0,\ M' \cdot \boldsymbol{\gamma}^+ > 0 \quad \text{or} \quad M' \cdot \boldsymbol{\gamma}^- > 0,\ M' \cdot \boldsymbol{\gamma}^+ < 0.$$

Let $\tilde{\boldsymbol{\gamma}} = M'(\boldsymbol{\gamma})$. Then U is also the associated doubly infinite $\mathbf{H}_4$-sequence of $\tilde{\boldsymbol{\gamma}}$. By Lemma 2.3, there exists a section $P^*|Q$ such that

$$M' \cdot \boldsymbol{\gamma}^- = -[P],\ M' \cdot \boldsymbol{\gamma}^+ = [Q] \quad \text{or} \quad M' \cdot \boldsymbol{\gamma}^- = [Q^\vee],\ M' \cdot \boldsymbol{\gamma}^+ = -[P^\vee]. \ \blacksquare$$

Let $\xi, \eta \in \hat{\mathbb{R}}$ be two distinct points on the boundary of $\mathbb{H}$, and U be the associated doubly infinite $\mathbf{H}_4$-sequence of the oriented geodesic $\boldsymbol{\gamma}$ with $\boldsymbol{\gamma}^- = \xi$, $\boldsymbol{\gamma}^+ = \eta$. For any section $P^*|Q$ of U, there exists $M \in \mathbf{H}_4$ such that

$$\tilde{\boldsymbol{\gamma}}^- = M \cdot \xi = -[P], \quad \tilde{\boldsymbol{\gamma}}^+ = M \cdot \eta = [Q] \quad \text{for } \tilde{\boldsymbol{\gamma}} = M(\boldsymbol{\gamma}).$$

Since

$$SM \cdot \xi = [P^\vee], \quad SM \cdot \eta = -[Q^\vee] \quad \text{and} \quad SM \in \mathbf{H}_4,$$

we have

$$\sup_{M \in \mathbf{H}_4} |M \cdot \xi - M \cdot \eta| \geq \sup_{P^*|Q} \max \{[Q] + [P], [Q^\vee] + [P^\vee]\} \geq 2,$$

where $P^*|Q$ runs over all sections of U, and the second inequality comes from $([P] + [Q])([P^\vee] + [Q^\vee]) = 2 + \frac{[P]}{[Q]} + \frac{[Q]}{[P]} \geq 4$.

Therefore, Lemma 2.4 implies that

$$\sup_{M \in \mathbf{H}_4} |M \cdot \xi - M \cdot \eta| = \sup_{P^*|Q} \max \{[Q] + [P], [Q^\vee] + [P^\vee]\}.$$

Let

$$L(P^*|Q) := [P] + [Q].$$

Then we have Perron's formula for the Hecke group $\mathbf{H}_4$ as follows.

THEOREM 2.5. *Let U be a doubly infinite $\mathbf{H}_4$-sequence. We define $\mathcal{M}(U)$ by the maximum of two supremum values as follows:*

$$\mathcal{M}(U) := \sup_{P^*|Q} \max \{L(P^*|Q), L((P^\vee)^*|Q^\vee)\},$$

where $P^*|Q$ runs over all sections of U. The Markov spectrum is the set of $\mathcal{M}(U)$ as U runs over all doubly infinite $\mathbf{H}_4$-sequences:

$$\mathscr{M}(\mathbf{H}_4) = \{\mathcal{M}(U) \in \mathbb{R} \mid U \text{ is a doubly infinite } \mathbf{H}_4\text{-sequence}\}.$$

THEOREM 2.6. *Let U be a doubly infinite $\mathbf{H}_4$-sequence. We define $\mathcal{L}(U)$ by the maximum of two limit superior values as follows:*

$$\mathcal{L}(U) := \limsup_{P^*|Q} \max \{L(P^*|Q), L((P^\vee)^*|Q^\vee)\},$$

where $P^*|Q$ runs over all sections of U. The Lagrange spectrum is the set of $\mathcal{L}(U)$ as U runs over all doubly infinite $\mathbf{H}_4$-sequences:

$$\mathscr{L}(\mathbf{H}_4) = \{\mathcal{L}(U) \in \mathbb{R} \mid U \text{ is a doubly infinite } \mathbf{H}_4\text{-sequence}\}.$$

For a finite sequence W, we denote the k-repeated sequence $W \dots W$ by W^k. We also denote an infinite sequence with period W and a doubly infinite sequence with period W by W^∞ and ${}^\infty W^\infty$.

EXAMPLE 2.7. The associated doubly infinite $\mathbf{H}_4$-sequences of the three closed geodesics in Figure 2 are $U_1 = {}^\infty 2^\infty$ (left), $U_2 = {}^\infty(13)^\infty$ (center) and

$U_3 = {}^\infty(123)^\infty$ (right). From (2.6) and (2.7), we check

$$\mathcal{M}(U_1) = L(\dots 22|22\dots) = 2[2^\infty] = 2,$$
$$\mathcal{M}(U_2) = L(\dots 3131|3131\dots) = \sqrt{2} + 2[(13)^\infty] = \sqrt{6},$$
$$\mathcal{M}(U_3) = L(\dots 123|123123\dots) = [(123)^\infty] + \frac{1}{[(123)^\infty]} = \frac{2\sqrt{17}}{3}.$$

Hereafter, commas in $\mathbf{H}_4$-sequences may occasionally be omitted for simplicity of notation.

3. Closedness of the Markov spectrum. We prove Theorem 1.1. First we note that given the discrete topology on $\{1,2,3\}$, the product space $\{1,2,3\}^{\mathbb{Z}}$ is compact due to Tikhonov's theorem.

LEMMA 3.1. *Let U be a doubly infinite $\mathbf{H}_4$-sequence. If $\mathcal{M}(U)$ is finite, then there exists a doubly infinite $\mathbf{H}_4$-sequence $\tilde{U}$ with a section $P^*|Q$ such that $\mathcal{M}(U) = \mathcal{M}(\tilde{U}) = L(P^*|Q)$.*

Proof. By Theorem 2.5, there exists a sequence $\{P_n^*|Q_n\}_{n\in\mathbb{N}}$ of sections of U or $U^\vee$, say U, such that $\lim_{n\to\infty} L(P_n^*|Q_n) = \mathcal{M}(U)$. Since the product space $\{1,2,3\}^{\mathbb{Z}}$ is compact, there exists a subsequence $\{P_{n_k}^*|Q_{n_k}\}_{k\in\mathbb{N}}$ which converges to a section $P^*|Q$ of a doubly infinite $\mathbf{H}_4$-sequence $\tilde{U}$. By the continuity of L, we have $L(P^*|Q) = \mathcal{M}(U) \le \mathcal{M}(\tilde{U})$.

If $\tilde{P}^*|\tilde{Q}$ is another section of $\tilde{U}$, then $\tilde{P}^*|\tilde{Q}$ is a limit of $\{\tilde{P}_{n_k}^*|\tilde{Q}_{n_k}\}_{k\in\mathbb{N}}$, which is a shifted subsequence of $\{P_{n_k}^*|Q_{n_k}\}$. Thus $L(\tilde{P}^*|\tilde{Q}) \le \mathcal{M}(U)$, which implies that $\mathcal{M}(\tilde{U}) \le \mathcal{M}(U)$. ∎

Proof of Theorem 1.1. We first show that the Markov spectrum $\mathscr{M}(\mathbf{H}_4)$ is closed. Choose a convergent sequence $\{m_n\}_{n\in\mathbb{N}}$ in $\mathscr{M}(\mathbf{H}_4)$. By Lemma 3.1, there exists a sequence of doubly infinite $\mathbf{H}_4$-sequences $\{U_n\}_{n\in\mathbb{N}}$ with a sequence $\{P_n^*|Q_n\}_{n\in\mathbb{N}}$ of sections such that $m_n = \mathcal{M}(U_n) = L(P_n^*|Q_n)$ for all $n \in \mathbb{N}$. By the compactness of $\{1,2,3\}^{\mathbb{Z}}$, we have a subsequence $\{P_{n_k}^*|Q_{n_k}\}_{k\in\mathbb{N}}$ converging to the limit $P^*|Q$ which is a section of a doubly infinite $\mathbf{H}_4$-sequence U. By the continuity of L, we have $\lim_{n\to\infty} m_n = L(P^*|Q)$, thus $\lim_{n\to\infty} m_n \le \mathcal{M}(U)$.

Let $\tilde{P}^*|\tilde{Q}$ be another section of U. Then $\tilde{P}^*|\tilde{Q}$ is a limit of finite shifts of the subsequence $\{P_{n_k}^*|Q_{n_k}\}_{k\in\mathbb{N}}$. Therefore, $L(\tilde{P}^*|\tilde{Q}) \le \lim_{n\to\infty} \mathcal{M}(U_n)$ and $\mathcal{M}(U) \le \lim_{n\to\infty} m_n$. Hence, $\mathcal{M}(U) = \lim_{n\to\infty} m_n$ and we conclude that the Markov spectrum is closed.

Now we show that $\mathscr{L}(\mathbf{H}_4) \subset \mathscr{M}(\mathbf{H}_4)$. By Theorem 2.6, for a doubly infinite $\mathbf{H}_4$-sequence U, there exists a sequence $\{P_n^*|Q_n\}_{n\in\mathbb{N}}$ of sections of U or $U^\vee$, say U, such that $\mathcal{L}(U) = \lim_{n\to\infty} L(P_n^*|Q_n)$. Since the product space $\{1,2,3\}^{\mathbb{Z}}$ is compact, there exists a subsequence $\{P_{n_k}^*|Q_{n_k}\}_{k\in\mathbb{N}}$ which converges to an element $P^*|Q \in \{1,2,3\}^{\mathbb{Z}}$, which is a section of a doubly infinite sequence $\tilde{U}$. By the continuity of L, we deduce that $\mathcal{L}(U) \le \mathcal{M}(\tilde{U})$.

For another section $\tilde{P}^*|\tilde{Q}$ of $\tilde{U}$, we have $L(\tilde{P}^*|\tilde{Q}) \le \mathcal{L}(U)$ since $\tilde{P}^*|\tilde{Q}$ is a limit of a sequence of sections of U. Therefore, $\mathcal{M}(\tilde{U}) \le \mathcal{L}(U)$. Hence, $\mathcal{L}(U) = \mathcal{M}(\tilde{U}) \in \mathscr{M}(\mathbf{H}_4)$. ∎

4. Hausdorff dimension of the Lagrange spectrum. In this section, we prove Theorem 1.2. By (1.4), for any $\mathbf{H}_4$-sequences P, Q, we have

$$|[1P]-[1Q]| \le |[P]-[Q]|, \qquad |[2P]-[2Q]| \le \frac{|[P]-[Q]|}{\sqrt{2}}.$$

In this section, assume that $\varepsilon > 0$ is given. Then we choose $m \ge 1$ such that

$$[(12^m3)^\infty] - [12^\infty] \le \frac{[3(12^m3)^\infty] - [2^\infty]}{(\sqrt{2})^m} < \varepsilon.$$

For any $\mathbf{H}_4$-sequence P we have

$$[32^{m+1}1P] + [(12^m3)^\infty] < [32^\infty] + [12^\infty] + \varepsilon = 2\sqrt{2} + \varepsilon. \tag{4.1}$$

Let $A = 32^{m+1}1$, $B = 32^m1$. Define

$$\Sigma := \{P \in \{1,2,3\}^{\mathbb{N}} \mid P = B^{m_1}A^{n_1}B^{m_2}A^{n_2}\ldots,\ n_i, m_i \in \{1,2\} \text{ for all } i\}.$$

LEMMA 4.1. *Let $\mathcal{F} = \{[P] \in \mathbb{R} \mid P \in \Sigma\}$. Then we have*

$$\dim_H(\mathcal{F}) > 0.$$

Proof. Let

$$\alpha := [(B^2A)^\infty], \qquad \beta := [(BA^2)^\infty].$$

Then for each $P \in \Sigma$ we have

$$\alpha \le [P] \le \beta.$$

Let

$$M_A = N_3N_2^{m+1}N_1, \qquad M_B = N_3N_2^mN_1.$$

Define $f_i : [\alpha,\beta] \to [\alpha,\beta]$, $i = 1,2,3,4$, to be

$$f_1(x) = M_B^2M_A \cdot x, \qquad f_3(x) = M_BM_A \cdot x,$$
$$f_2(x) = M_B^2M_A^2 \cdot x, \qquad f_4(x) = M_BM_A^2 \cdot x.$$

Then $\{f_1, f_2, f_3, f_4\}$ is a family of contracting functions, which is called an *iterated function system* (see e.g. [9]) satisfying

$$\mathcal{F} = f_1(\mathcal{F}) \cup f_2(\mathcal{F}) \cup f_3(\mathcal{F}) \cup f_4(\mathcal{F}), \qquad f_i(\mathcal{F}) \cap f_j(\mathcal{F}) = \emptyset \quad \text{for } i \ne j.$$

Using (1.4), we check that there are $c_i > 0$ for each $i = 1,2,3,4$ such that $|f_i(x) - f_i(y)| \ge c_i|x-y|$ for $x, y \in [\alpha,\beta]$ since all elements of the matrices $M_B^2M_A$, $M_B^2M_A^2$, M_BM_A, $M_BM_A^2$ are positive. By [9, Proposition 9.7], we conclude that

$$\dim_H(\mathcal{F}) \ge s,$$

where $s > 0$ is the constant satisfying

$$c_1^s + c_2^s + c_3^s + c_4^s = 1. \ \blacksquare$$

Choose

$$R = B^{m_1}A^{n_1}B^{m_2}A^{n_2}\ldots \in \Sigma, \qquad n_i, m_i \in \{1,2\},$$

and let

$$W_k^R := B^{m_1}A^{n_1}B^{m_2}A^{n_2}B^{m_3}\ldots B^{m_k}A^{n_k},$$
$$U_R := {}^{\infty}BW_1^RB^2A^3W_2^RB^3A^3W_3^RB^4\ldots B^kA^3W_k^RB^{k+1}\ldots.$$

LEMMA 4.2. *We have*

$$\mathcal{L}(U_R) = [(B^{\vee})^{\infty}] + [A^3R] = \frac{1}{[B^{\infty}]} + [A^3R].$$

Proof. Let $P^*32^k|2^\ell 1Q$ be a section of U_R for some $k, \ell \geq 0$. Then we have, for $k \geq 1, \ell \geq 0$,

$$L(P^*32^k|2^\ell 1Q) = [2^k3P] + [2^\ell 1Q] < [23^{\infty}] + [2^{\infty}] = \sqrt{2}+1,$$

and for $k = 0$,

$$L(P^*32^k|2^\ell 1Q) = L(P^*|32^\ell 1Q).$$

On the other hand, if $P^*32^k1|32^\ell 1Q$ is a section of U, then

$$L(P^*32^k1|32^\ell 1Q) = [12^k3P]+[32^\ell 1Q] > [12^{\infty}]+[321^{\infty}] = \frac{5}{\sqrt{2}}-1 > \sqrt{2}+1.$$

Therefore,

$$\mathcal{L}(U_R) = \limsup_{P^*|Q} \max\{L(P^*|Q), L((P^{\vee})^*|Q^{\vee})\},$$

where $P^*|Q$ runs over all sections of U_R such that $P^*|Q = \tilde{P}^*A|A\tilde{Q}$, $\tilde{P}^*A|B\tilde{Q}$, $\tilde{P}^*B|A\tilde{Q}$, or $\tilde{P}^*B|B\tilde{Q}$ for some $\tilde{P}$ and $\tilde{Q}$. Using the facts that $[AP] > [BQ]$ for any infinite sequences P, Q and that W_k^R does not contain A^3, we conclude that

$$\begin{aligned}\mathcal{L}(U_R) &= \limsup_{k\to\infty} L(\ldots B_{k-1}A^3W_{k-1}^RB^k|A^3W_k^RB^{k+1}A^3W_{k+1}^R\ldots)\\ &= \limsup_{k\to\infty}\left(\frac{1}{[B^kW_{k-1}^RA^3B_{k-1}\ldots]} + [A^3W_k^RB^{k+1}A^3W_{k+1}^R\ldots]\right)\\ &= L({}^{\infty}B|A^3R) = [(B^{\vee})^{\infty}] + [A^3R] = \frac{1}{[B^{\infty}]} + [A^3R]. \ \blacksquare\end{aligned}$$

Let

$$\mathcal{H} := \{\mathcal{L}(U_R) \mid R \in \Sigma\}.$$

Then Lemma 4.2 and (4.1) yield

$$\mathcal{H} = \left\{\frac{1}{[B^{\infty}]} + [A^3R] \mid R \in \Sigma\right\} \subset \mathscr{L}(\mathbf{H}_4) \cap (0, 2\sqrt{2}+\varepsilon). \tag{4.2}$$

Since all elements of the matrix M_A^3 are positive, $[R] \mapsto [A^3R] = M_A^3 \cdot [R]$ is a bi-Lipschitz function on the closed interval $[\alpha, \beta]$. Therefore, Lemma 4.1 implies that $\dim_H(\mathcal{H}) > 0$, which completes the proof of Theorem 1.2.

5. Gaps of the Markov spectrum. In this section, we investigate the gaps of $\mathscr{M}(\mathbf{H}_4)$ above the first limit point $2\sqrt{2}$. We prove Theorem 1.3 through Theorems 5.2 and 5.3.

We check that

$$[(21)^\infty] = \frac{\sqrt{2}}{\sqrt{7}-1}, \qquad [(2131)^\infty] = \frac{\sqrt{119}+3}{11\sqrt{2}} \tag{5.1}$$

and let

$$\begin{aligned} m_0 :=& \mathcal{M}({}^\infty(3132)123(2131)^\infty) \\ =& L({}^\infty(3132)12 \mid 3(2131)^\infty) = \frac{2124\sqrt{2}+48\sqrt{238}}{1177} = 3.181\ldots. \end{aligned}$$

Lemma 5.1. *Let $\mathcal{M}(U) \le m_0$. Then U satisfies one of the following:*

(i) $U = {}^\infty(1232)^\infty$, *or*
(ii) U *or* $U^\vee = {}^\infty(3132)123(2131)^\infty$, *or*
(iii) U *does not contain* $11, 33, 212, 232$.

Proof. First, if U or $U^\vee$, say U, contains 333, then

$$\mathcal{M}(U) \ge L(P^*|333Q) = [P] + [333Q] = [P] + [Q] + 3\sqrt{2} \ge 3\sqrt{2} > m_0$$

for some infinite $\mathbf{H}_4$-sequences P, Q with $U = P^*333Q$. Therefore, U and $U^\vee$ do not contain 333 or 111.

Next, assume that U or $U^\vee$, say U, contains 33. Let $U = P^*33Q$ for some infinite $\mathbf{H}_4$-sequences P, Q starting with 1 or 2. Then, by (2.7), we have

$$\begin{aligned} \mathcal{M}(U) &\ge L(P^*|33Q) = [Q] + [P] + 2\sqrt{2} \\ &\ge [(112)^\infty] + [(112)^\infty] + 2\sqrt{2} = \frac{2}{\sqrt{7}+\sqrt{2}} + 2\sqrt{2} > m_0. \end{aligned}$$

Hence, U and $U^\vee$ do not contain 33 or 11.

We claim that U and U^* do not contain 2322 or 2323. Let $U = P^*232Q$ for some infinite $\mathbf{H}_4$-sequences P,Q with Q beginning with 2 or 3. Then, by (5.1), we have

$$\begin{aligned} \mathcal{M}(U) &\ge L(P^*2|32Q) = [2P] + [2Q] + \sqrt{2} \\ &\ge [2(12)^\infty] + [2(21)^\infty] + \sqrt{2} = \frac{\sqrt{2}}{\sqrt{7}-1} + \frac{\sqrt{7}+1}{\sqrt{14}} + \sqrt{2} > m_0. \end{aligned}$$

Therefore, U does not contain 11, 33, 2323, 2121, 3232, 1212, 2232, 2212, 2322, 2122. Thus, any infinite $\mathbf{H}_4$-sequence P of U or U^* satisfies

$$[(1213)^\infty] \le [P] \le [(3231)^\infty]. \tag{5.2}$$

Suppose that $U \neq {}^{\infty}(1232)^{\infty}$ and U contains 232. Then U also contains (a) 3123213, or (b) 2123212, or (c) 3123212 or 2123213, say 2123213.

(a) In this case, $U = P^*3123213Q$ for some infinite $\mathbf{H}_4$-sequences P, Q and by (5.2) and (5.1) we have

$$\begin{aligned}\mathcal{M}(U) &\geq L(P^*312|3213Q) = [213P] + [213Q] + \sqrt{2} \\ &\geq [(2131)^{\infty}] + [(2131)^{\infty}] + \sqrt{2} = 2\frac{\sqrt{119}+3}{11\sqrt{2}} + \sqrt{2} > m_0.\end{aligned}$$

Hence, U does not contain 3123213 or 1321231.

(b) In this case, since $U \neq {}^{\infty}(1232)^{\infty}$, there exists an infinite $\mathbf{H}_4$-sequence P, not beginning with 32, for which $U, U^{\vee}, U^*$, or $(U^{\vee})^*$, say $U^{\vee}$, is $P^*2123212Q$ for some Q. Hence, by (5.2), we have

$$\begin{aligned}\mathcal{M}(U) &\geq L((P^{\vee})^*2|321232Q^{\vee}) = [2P^{\vee}] + [21232Q^{\vee}] + \sqrt{2} \\ &> [2(1312)^{\infty}] + [212(3132)^{\infty}] + \sqrt{2} \\ &= \mathcal{M}({}^{\infty}(3132)123(2131)^{\infty}) = m_0.\end{aligned}$$

(c) If U contains 2123213, but does not contain 2123212 or 2321232, then we have $U = P^*21232Q$ for some infinite $\mathbf{H}_4$-sequences P, Q, where P does not begin with 32 and Q does not start with 12. Thus, (5.2) implies

$$[P^{\vee}],\ [Q] \geq [13(1213)^{\infty}] = [(1312)^{\infty}].$$

By the elementary calculus, we check that $[2P] + [212P^{\vee}]$ is an increasing function of $[P]$ on the interval $([(1312)^{\infty}], [(3132)^{\infty}])$. Therefore, we have

$$\begin{aligned}\mathcal{M}(U) &\geq \frac{1}{2}(L(P^*212|32Q) + L((P^{\vee})^*23|212Q^{\vee})) \\ &= \frac{1}{2}([212P] + [2P^{\vee}]) + \frac{1}{2}([2Q] + [212Q^{\vee}]) + \sqrt{2} \\ &\geq \frac{[2(1312)^{\infty}] + [212(3132)^{\infty}]}{2} + \frac{[2(1312)^{\infty}] + [212(3132)^{\infty}]}{2} + \sqrt{2} \\ &= [(2131)^{\infty}] + [21(2313)^{\infty}] + \sqrt{2} = \mathcal{M}({}^{\infty}(3132)123(2131)^{\infty}) = m_0.\end{aligned}$$

Moreover, if equality holds, then we see that $U = {}^{\infty}(3132)123(2131)^{\infty}$ or $U^{\vee} = {}^{\infty}(3132)123(2131)^{\infty}$. ∎

THEOREM 5.2. *The interval*

$$\left(\sqrt{10}, \frac{2124\sqrt{2}+48\sqrt{238}}{1177}\right) = (3.162\ldots, 3.181\ldots)$$

is a maximal gap in $\mathscr{M}(\mathbf{H}_4)$. *The two boundary points of the interval correspond to* $\mathcal{M}({}^{\infty}(1232)^{\infty}) = \sqrt{10}$ *and* $\mathcal{M}(U) = \frac{2124\sqrt{2}+48\sqrt{238}}{1177}$ *for* $U = {}^{\infty}(3132)123(2131)^{\infty}$. *Moreover,* $\frac{2124\sqrt{2}+48\sqrt{238}}{1177}$ *is a limit point of* $\mathscr{M}(\mathbf{H}_4)$.

Proof. By direct calculation, we check that $\mathcal{M}({}^\infty(1232)^\infty) = \sqrt{10}$ and $\mathcal{M}({}^\infty(3132)123(2131)^\infty) = \frac{2124\sqrt{2}+48\sqrt{238}}{1177} =: m_0$.

Suppose that U is a doubly infinite $\mathbf{H}_4$-sequence with $\mathcal{M}(U) < m_0$ and $U \neq {}^\infty(1232)^\infty$. Then, by Lemma 5.1, U does not contain 11, 33, 212, 232. If $a, b \in \{1, 2\}$, then

$$L(P^* a \,|\, bQ) = [aP] + [bQ] \le [2P] + [2Q] \le \sqrt{2} + \sqrt{2} < \sqrt{10}$$

for any infinite $\mathbf{H}_4$-sequences P, Q. Moreover, if $P^*1|3aQ$ is a section of U for $a \in \{1, 2\}$, then

$$\begin{aligned} L(P^*1 \mid 3aQ) &= [1P] + [3aQ] \le [(1323)^\infty] + [(3231)^\infty] \\ &= \frac{\sqrt{119}-7}{5\sqrt{2}} + \frac{\sqrt{119}+7}{5\sqrt{2}} = \frac{\sqrt{238}}{5} < \sqrt{10}. \end{aligned}$$

Therefore, we deduce that $\mathcal{M}(U) < \sqrt{10}$. Hence, $(\sqrt{10}, m_0)$ is a maximal gap in $\mathscr{M}(\mathbf{H}_4)$.

Finally, let us show that m_0 is a limit point of $\mathscr{M}(\mathbf{H}_4)$. For $k \ge 1$, let

$$U_k := {}^\infty(1312)321A_k123(2131)^\infty, \quad \text{where} \quad A_k := (2313)^k 2 = 2(3132)^k.$$

Let $P_k^*|Q_k$ be a section of the doubly infinite sequence U_k. Then there exists a section $P^*|Q$ of the doubly infinite sequence U such that at least the first $4k+2$ digits of P and P_k and those of Q and Q_k are identical. Therefore, we have $\lim_{k\to\infty} \mathcal{M}(U_k) = \mathcal{M}(U) = m_0$. By Lemma 5.1, we have $\mathcal{M}(U_k) > m_0$ for all k. Hence, m_0 is a limit point of $\mathscr{M}(\mathbf{H}_4)$. ∎

THEOREM 5.3. *The interval*

$$\left(\frac{\sqrt{238}}{5}, \sqrt{10}\right) = (3.085\ldots, 3.162\ldots)$$

is a maximal gap in $\mathscr{M}(\mathbf{H}_4)$. *The lower endpoint satisfies* $\mathcal{M}({}^\infty(1312)^\infty) = \frac{\sqrt{238}}{5}$.

Proof. Suppose that $\mathcal{M}(U) < \sqrt{10} = \mathcal{M}({}^\infty(1232)^\infty)$. By Lemma 5.1, U does not contain 33, 11, 212, 232. Therefore, for any infinite $\mathbf{H}_4$-sequence R appearing in U, if R does not start with 32 or 12, then

$$[(1312)^\infty] \le [R] \le [(3132)^\infty].$$

If $U = P^*2Q$ for infinite $\mathbf{H}_4$-sequences P, Q, then neither P nor Q starts with 32. Therefore,

$$L(P^*|2Q) = [P] + [2Q] \le [(3132)^\infty] + [2(3132)^\infty] = \mathcal{M}({}^\infty(3132)^\infty).$$

If $U = P^*13Q$, then $[Q] \le [(2313)^\infty]$ and $[31P] \le [(3132)^\infty]$. Thus,

$$\begin{aligned} L(P^*1|3Q) = L(P^*13|Q) &= [31P] + [Q] \\ &\le [(2313)^\infty] + [(3132)^\infty] = \mathcal{M}({}^\infty(3132)^\infty). \end{aligned}$$

Therefore, for any section $P^*|Q$ of U, we have $L(P^*|Q) \le \mathcal{M}({}^\infty(3132)^\infty) = \frac{\sqrt{238}}{5}$. Thus, by Theorem 2.5, the interval $\left(\frac{\sqrt{238}}{5}, \sqrt{10}\right)$ is a maximal gap in $\mathscr{M}(\mathbf{H}_4)$. ∎

6. A bound of Hall's ray. In this section, we give the bound of Hall's ray stated in Theorem 1.4. Let

$$\mathcal{K} = \{[d_1, d_2, \dots] \mid d_1 \neq 3,\ d_k d_{k+1} d_{k+2} \neq 111, 333 \text{ for all } k \geq 1\}$$

and let $R = (332)^\infty$. We note that $[R] = \sqrt{7} + \sqrt{2}$ and

$$\min \mathcal{K} = [R^\vee] = \frac{\sqrt{7}-\sqrt{2}}{5}, \qquad \max \mathcal{K} = [2R] = \sqrt{7} - \sqrt{2}.$$

Let $\mathcal{E}(c_1, \dots, c_n)$ be the smallest closed interval containing $\{[d_1, d_2, \dots] \in \mathcal{K} \mid d_1 = c_1, \dots, d_n = c_n\}$ and let $\mathcal{E} = \big[[R^\vee], [2R]\big] = \big[\frac{\sqrt{7}-\sqrt{2}}{5}, \sqrt{7} - \sqrt{2}\big]$. Then we have

$$\mathcal{E}(c_1, \dots, c_n) = \begin{cases} \big[[c_1 \dots c_n 2R^\vee],\ [c_1 \dots c_n R]\big] & \text{if } c_{n-1}c_n = 11, \\ \big[[c_1 \dots c_n 12R^\vee],\ [c_1 \dots c_n R]\big] & \text{if } c_{n-1} \neq 1, c_n = 1, \\ \big[[c_1 \dots c_n R^\vee],\ [c_1 \dots c_n R]\big] & \text{if } c_n = 2, \\ \big[[c_1 \dots c_n R^\vee],\ [c_1 \dots c_n 32R]\big] & \text{if } c_{n-1} \neq 3, c_n = 3, \\ \big[[c_1 \dots c_n R^\vee],\ [c_1 \dots c_n 2R]\big] & \text{if } c_{n-1}c_n = 33. \end{cases}$$

We also define $\mathcal{E}_*(c_1, \dots, c_n)$ to be the smallest closed interval containing

$$\{[d_1, d_2, \dots] \in \mathcal{K} \mid d_1 = c_1, \dots, d_n = c_n, d_{n+1} \neq 3\}.$$

First, let us verify that $\mathcal{K}$ can be obtained by applying the Cantor dissection process. In the dissection process, each type of interval is divided by the following rules:

- The interval $\mathcal{E}(c_1, \dots, c_n)$ is divided into the union of two intervals

$$T \begin{cases} \mathcal{E}(c_1, \dots, c_n, 2) \cup \mathcal{E}(c_1, \dots, c_n, 3) & \text{if } c_{n-1}c_n = 11, \\ \mathcal{E}(c_1, \dots, c_n, 1) \cup \mathcal{E}(c_1, \dots, c_n, 2) & \text{if } c_{n-1}c_n = 33, \\ \mathcal{E}_*(c_1, \dots, c_n) \cup \mathcal{E}(c_1, \dots, c_n, 3) & \text{otherwise.} \end{cases}$$

- The interval $\mathcal{E}_*(c_1, \dots, c_n)$ with $c_{n-1}c_n \neq 11, 33$, is divided into the union of two intervals

$$\mathcal{E}(c_1, \dots, c_n, 1) \cup \mathcal{E}(c_1, \dots, c_n, 2).$$

Each type of interval is subdivided into two intervals. Starting from $\mathcal{E}$, we continue the dissection process according to the above rules. Thus, we obtain the Cantor set $\mathcal{K}$.

Lemma 6.1. *Let $\mathcal{I}$ be the closed interval $\mathcal{E}(c_1, \dots, c_n)$ or $\mathcal{E}_*(c_1, \dots, c_n)$. In the Cantor dissection process, we have closed intervals $\mathcal{I}_1, \mathcal{I}_2$ in $\mathcal{I}$ satisfying $\mathcal{I} \setminus \mathcal{J} = \mathcal{I}_1 \cup \mathcal{I}_2$ for an open interval $\mathcal{J}$. Then*

$$|\mathcal{I}_i| \geq |\mathcal{J}| \quad \textit{for } i = 1, 2.$$

Proof. Let

$$M = N_{c_1} \dots N_{c_n} = \begin{pmatrix} p & r \\ q & s \end{pmatrix}.$$

Then from (1.4) we have

$$[c_1 \dots c_n P] - [c_1 \dots c_n Q] = M \cdot [P] - M \cdot [Q] = \frac{[P] - [Q]}{(q[P] + s)(q[Q] + s)}.$$

Using (2.8), we get

$$\frac{s}{q} = -M^{-1} \cdot \infty = -N_{c_n}^{-1} \dots N_{c_1}^{-1} \cdot \infty = N_{c_n} \dots N_{c_1} \cdot \infty = [c_n \dots c_1 3^\infty],$$

and moreover

$$\frac{s}{q} \leq [113^\infty] = \frac{1}{2\sqrt{2}} \quad \text{if } c_{n-1}c_n = 11 \tag{6.1}$$

and

$$\frac{s}{q} > [113^\infty] = \frac{1}{2\sqrt{2}} \quad \text{if } c_{n-1}c_n \neq 11. \tag{6.2}$$

Let $\mathcal{I} = \mathcal{E}(c_1, \dots, c_n)$ with $c_{n-1}c_n \neq 11, 33$. Then

$$\mathcal{I}_1 = \mathcal{E}_*(c_1, \dots, c_n), \quad \mathcal{I}_2 = \mathcal{E}(c_1, \dots, c_n, 3).$$

Therefore,

$$\begin{aligned} |\mathcal{J}| &= [c_1 \dots c_n 3R^\vee] - [c_1 \dots c_n 2R] = \frac{[3R^\vee] - [2R]}{(q[3R^\vee] + s)(q[2R] + s)}, \\ |\mathcal{I}_1| &\geq [c_1 \dots c_n 2R] - [c_1 \dots c_n 12R^\vee] = \frac{[2R] - [12R^\vee]}{(q[2R] + s)(q[12R^\vee] + s)}, \\ |\mathcal{I}_2| &\geq [c_1 \dots c_n 32R] - [c_1 \dots c_n 3R^\vee] = \frac{[32R] - [3R^\vee]}{(q[32R] + s)(q[3R^\vee] + s)}, \end{aligned}$$

thus,

$$\begin{aligned} \frac{|\mathcal{J}|}{|\mathcal{I}_1|} &\leq \frac{(q[12R^\vee] + s)([3R^\vee] - [2R])}{(q[3R^\vee] + s)([2R] - [12R^\vee])} < \frac{[3R^\vee] - [2R]}{[2R] - [12R^\vee]} = 0.5025\ldots < 1, \\ \frac{|\mathcal{J}|}{|\mathcal{I}_2|} &\leq \frac{(q[32R] + s)([3R^\vee] - [2R])}{(q[2R] + s)([32R] - [3R^\vee])} < \frac{[32R]([3R^\vee] - [2R])}{[2R]([32R] - [3R^\vee])} = 0.9354\ldots < 1. \end{aligned}$$

Let $\mathcal{I} = \mathcal{E}(c_1, \dots, c_n)$ with $c_{n-1}c_n = 11$. Then we have

$$\mathcal{I}_1 = \mathcal{E}(c_1, \dots, c_n, 2), \quad \mathcal{I}_2 = \mathcal{E}(c_1, \dots, c_n, 3).$$

Therefore,

$$|\mathcal{J}| = [c_1 \dots c_n 3R^\vee] - [c_1 \dots c_n 2R] = \frac{[3R^\vee] - [2R]}{(q[3R^\vee] + s)(q[2R] + s)},$$
$$|\mathcal{I}_1| = [c_1 \dots c_n 2R] - [c_1 \dots c_n 2R^\vee] = \frac{[2R] - [2R^\vee]}{(q[2R] + s)(q[2R^\vee] + s)},$$
$$|\mathcal{I}_2| = [c_1 \dots c_n R] - [c_1 \dots c_n 3R^\vee] = \frac{[R] - [3R^\vee]}{(q[R] + s)(q[3R^\vee] + s)}.$$

By (6.1), we have $q \geq 2\sqrt{2}s$, thus

$$\frac{|\mathcal{J}|}{|\mathcal{I}_1|} = \frac{(q[2R^\vee] + s)([3R^\vee] - [2R])}{(q[3R^\vee] + s)([2R] - [2R^\vee])}$$
$$\leq \frac{2\sqrt{2}[2R^\vee] + 1}{2\sqrt{2}[3R^\vee] + 1} \cdot \frac{[3R^\vee] - [2R]}{[2R] - [2R^\vee]} = 0.5917\ldots < 1,$$
$$\frac{|\mathcal{J}|}{|\mathcal{I}_2|} = \frac{(q[R] + s)([3R^\vee] - [2R])}{(q[2R] + s)([R] - [3R^\vee])} \leq \frac{[R]([3R^\vee] - [2R])}{[2R]([R] - [3R^\vee])} = 0.5893\ldots < 1.$$

Let $\mathcal{I} = \mathcal{E}_*(c_1, \dots, c_n)$ with $c_{n-1}c_n \neq 11, 33$ or $\mathcal{I} = \mathcal{E}(c_1, \dots, c_n)$ with $c_{n-1}c_n = 33$. Then we have

$$\mathcal{I}_1 = \mathcal{E}(c_1, \dots, c_n, 1), \qquad \mathcal{I}_2 = \mathcal{E}(c_1, \dots, c_n, 2).$$

Therefore,

$$|\mathcal{J}| = [c_1 \dots c_n 2R^\vee] - [c_1 \dots c_n 1R] = \frac{[2R^\vee] - [1R]}{(q[2R^\vee] + s)(q[1R] + s)},$$
$$|\mathcal{I}_1| \geq [c_1 \dots c_n 1R] - [c_1 \dots c_n 12R^\vee] = \frac{[1R] - [12R^\vee]}{(q[1R] + s)(q[12R^\vee] + s)},$$
$$|\mathcal{I}_2| = [c_1 \dots c_n 2R] - [c_1 \dots c_n 2R^\vee] = \frac{[2R] - [2R^\vee]}{(q[2R] + s)(q[2R^\vee] + s)}.$$

Together with the condition that $c_{n-1}c_n \neq 11$, (6.1) implies $q < 2\sqrt{2}s$. Thus,

$$\frac{|\mathcal{J}|}{|\mathcal{I}_1|} \leq \frac{(q[12R^\vee] + s)([2R^\vee] - [1R])}{(q[2R^\vee] + s)([1R] - [12R^\vee])} < \frac{[2R^\vee] - [1R]}{[1R] - [12R^\vee]} = 0.9354\ldots < 1,$$
$$\frac{|\mathcal{J}|}{|\mathcal{I}_2|} = \frac{(q[2R] + s)([2R^\vee] - [1R])}{(q[1R] + s)([2R] - [2R^\vee])}$$
$$< \frac{2\sqrt{2}[2R] + 1}{2\sqrt{2}[1R] + 1} \cdot \frac{[2R^\vee] - [1R]}{[2R] - [2R^\vee]} = 0.8292\ldots < 1. \blacksquare$$

For $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}$, we write $\mathcal{X} + \mathcal{Y} := \{x + y \mid x \in \mathcal{X},\ y \in \mathcal{Y}\}$.

Lemma 6.2 ([8, Chapter 4, Lemma 3]). *Let $\mathcal{B}$ be the union of disjoint closed intervals $\mathcal{A}_1, \dots, \mathcal{A}_r$. Given an open interval $\mathcal{I}$ in $\mathcal{A}_1$, let $\mathcal{A}_{r+1}, \mathcal{A}_{r+2}$ be the disjoint closed intervals such that $\mathcal{A}_1 \setminus \mathcal{I} = \mathcal{A}_{r+1} \cup \mathcal{A}_{r+2}$. Let $\mathcal{B}'$ be the* *union of* $\mathcal{A}_2, \dots, \mathcal{A}_{r+2}$. *If* $|\mathcal{A}_i| \ge |\mathcal{I}|$ *for* $i = 2, \dots, r+2$, *then*

$$\mathcal{B} + \mathcal{B} = \mathcal{B}' + \mathcal{B}'.$$

LEMMA 6.3 ([8, Chapter 4, Lemma 4]). *If* $\mathcal{C}_1, \mathcal{C}_2, \dots$ *is a sequence of bounded closed sets such that* $\mathcal{C}_i$ *contains* $\mathcal{C}_{i+1}$ *for all* $i \ge 1$, *then*

$$\bigcap_{i=1}^{\infty} \mathcal{C}_i + \bigcap_{i=1}^{\infty} \mathcal{C}_i = \bigcap_{i=1}^{\infty} (\mathcal{C}_i + \mathcal{C}_i).$$

THEOREM 6.4. *We have* $\mathcal{K} + \mathcal{K} = \left[\frac{2\sqrt{7}-2\sqrt{2}}{5}, 2\sqrt{7} - 2\sqrt{2}\right]$.

Proof. Let $\mathcal{K}_0 := \mathcal{E} = \left[\frac{\sqrt{7}-\sqrt{2}}{5}, \sqrt{7} - \sqrt{2}\right]$ and construct a sequence of sets $\{\mathcal{K}_k\}_{k=0}^{\infty}$ satisfying the following four properties:

(1) Each $\mathcal{K}_k$ is closed and bounded.
(2) $\mathcal{K}_k \supset \mathcal{K}_{k+1}$ for all $k \ge 0$.
(3) $\bigcap_{k=0}^{\infty} \mathcal{K}_k = \mathcal{K}$.
(4) $\mathcal{K}_k + \mathcal{K}_k = \mathcal{K}_{k+1} + \mathcal{K}_{k+1}$ for $k \ge 0$.

We already verified that $\mathcal{K}$ is obtained from $\mathcal{K}_0$ by removing an infinite number of disjoint open intervals. Now, let us arrange these removed open intervals in decreasing order of length and denote them by $\mathcal{J}_0, \mathcal{J}_1, \dots$. For $k \ge 0$, we set $\mathcal{K}_{k+1} = \mathcal{K}_k \setminus \mathcal{J}_k$. By the definition of $\mathcal{K}_k$, properties (1)–(3) are satisfied. It remains to show (4).

Let $\mathcal{I}$ be the closed interval from which $\mathcal{J}_k$ is removed, and $\mathcal{I}_1$, $\mathcal{I}_2$ be the disjoint closed intervals such that $\mathcal{I} \setminus \mathcal{J}_k = \mathcal{I}_1 \cup \mathcal{I}_2$. By Lemma 6.1, $|\mathcal{I}_1|, |\mathcal{I}_2| \ge |\mathcal{J}_k|$. Since the intervals $(\mathcal{J}_k)$ are arranged in decreasing order of length, by Lemma 6.1, each closed interval in $\mathcal{K}_k$ has length greater than or equal to $|\mathcal{J}_{k-1}|$. Hence, each closed interval in $\mathcal{K}_{k+1}$ has length equal to or greater than $|\mathcal{J}_k|$. By Lemma 6.2, $\mathcal{K}_k + \mathcal{K}_k = \mathcal{K}_{k+1} + \mathcal{K}_{k+1}$. Therefore, by Lemma 6.3, $\mathcal{K} + \mathcal{K} = (\bigcap_{i=1}^{\infty} \mathcal{K}_i) + (\bigcap_{i=1}^{\infty} \mathcal{K}_i) = \bigcap_{i=1}^{\infty}(\mathcal{K}_i + \mathcal{K}_i) = \mathcal{K}_0 + \mathcal{K}_0$. ∎

Since the length of $\mathcal{K}_0 + \mathcal{K}_0 = \left[\frac{2\sqrt{7}-2\sqrt{2}}{5}, 2\sqrt{7} - 2\sqrt{2}\right]$ is greater than $\sqrt{2}$, Theorem 6.4 implies the following corollary.

COROLLARY 6.5. *Any real number can be expressed as* $\sqrt{2}n + [P] + [Q]$ *for* $n \in \mathbb{Z}$, $[P], [Q] \in \mathcal{K}$.

Now, we obtain the bound of Hall's ray:

Proof of Theorem 1.4. Let $\alpha \ge 4\sqrt{2}$. By Corollary 6.5, there exist two $\mathbf{H}_4$-sequences $P, Q \in \mathcal{K}$ and $n \in \mathbb{Z}$ such that $\alpha = \sqrt{2}n + [P] + [Q]$. Since $[P], [Q] \le \sqrt{7} - \sqrt{2} < \sqrt{2}$, we have $n \ge 3$. We set $P = (a_1, a_2, \dots)$ and $Q = (b_1, b_2, \dots)$. Let m_k and ℓ_k be increasing sequences satisfying $a_{\ell_k} \ne 3$ and $b_{m_k} \ne 3$. Put $A_k = a_1 \dots a_{\ell_k}$ and $B_k = b_1 \dots b_{m_k}$. Define a doubly infinite sequence

$$U = {}^{\infty}2A_1^* 3^n B_1 A_2^* 3^n B_2 A_3^* 3^n B_3 A_4^* 3^n B_4 \dots.$$

Note that A_k, B_k do not contain 333 and the first and the last digit of A_k, B_k are not 3. Since $L(\tilde{P}^*23|32\tilde{Q}) \leq 4\sqrt{2}$ for any section $\tilde{P}^*23|32\tilde{Q}$ of U, we have

$$\begin{aligned}\mathcal{L}(U) &= \limsup_{k\to\infty} L({}^{\infty}2A_1^*3^nB_1\ldots A_{k-1}^*3^nB_{k-1}A_k^*3^n \mid B_kA_{k+1}^*3^nB_{k+2}\ldots) \\ &= n\sqrt{2} + \limsup_{k\to\infty}([A_kB_{k-1}^*3^n\ldots B_1^*3^nA_12^{\infty}] + [B_kA_{k+1}^*3^nB_{k+2}\ldots]) \\ &= n\sqrt{2} + [P] + [Q] = \alpha.\end{aligned}$$

Hence, $\mathscr{L}(\mathbf{H}_4)$ contains every real number greater than or equal to $4\sqrt{2}$. ∎

Acknowledgements. The authors would like to thank the anonymous referee for a careful reading and for helpful comments. They also thank Byungchul Cha for helpful discussions and comments, and Mauro Artigiani and Luca Marchese for drawing their attention to the reference [1].

Funding. D.H.K. was supported by the National Research Foundation of Korea (NRF-2018R1A2B6001624, RS-2023-00245719).

D.S. was supported by the National Research Foundation of Korea (NRF-2020R1A2C1A01011543).

References

[1] M. Artigiani, L. Marchese and C. Ulcigrai, *Persistent Hall rays for Lagrange spectra at cusps of Riemann surfaces*, Ergodic Theory Dynam. Systems 40 (2020), 2017–2072.

[2] A. Bakhtawar, D. H. Kim and S. B. Lee, *Uniform Diophantine approximation on the Hecke group* $\mathbf{H}_4$, Int. Math. Res. Notices 2015, no. 16, art. rnaf257, 22 pp.

[3] E. Bombieri, *Continued fractions and the Markov tree*, Expo. Math. 25 (2007), 187–213.

[4] Y. Bugeaud, P. Hubert and T. A. Schmidt, *Transcendence with Rosen continued fractions*, J. Eur. Math. Soc. 15 (2013), 39–51.

[5] R. M. Burton, C. Kraaikamp and T. A. Schmidt, *Natural extensions for the Rosen fractions*, Trans. Amer. Math. Soc. 352 (2000), 1277–1298.

[6] B. Cha and D. H. Kim, *Intrinsic Diophantine approximation on the unit circle and its Lagrange spectrum*, Ann. Inst. Fourier (Grenoble) 73 (2023), 101–161.

[7] B. Cha and D. H. Kim, *Intrinsic Diophantine approximation on circles and spheres*, Mathematika 70 (2024), no. 1, art. e12228, 31 pp.

[8] T. W. Cusick and M. E. Flahive, *The Markov and Lagrange Spectra*, Amer. Math. Soc., Providence, RI, 1989.

[9] K. Falconer, *Fractal Geometry*, 2nd ed., Wiley, Hoboken, NJ, 2003.

[10] G. A. Freĭman, *Diophantine Approximations and the Geometry of Numbers (Markov's Problem)*, Kalinin. Gos. Univ., Kalinin, 1975 (in Russian).

[11] A. Haas and C. Series, *The Hurwitz constant and Diophantine approximation on Hecke groups*, J. London Math. Soc. 34 (1986), 219–234.

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

[13] D. H. Kim, S. B. Lee and L. Liao, *Odd-odd continued fraction algorithm*, Monatsh. Math. 198 (2022), 323–344.

[14] D. H. Kim, S. B. Lee and L. Liao, *Diophantine approximation by rational numbers of certain parity types*, Mathematika 71 (2025), no. 1, art. e12285, 22 pp.

[15] H. G. Kopetzky, *Über das Approximationsspektrum des Einheitskreises*, Monatsh. Math. 100 (1985), 211–213.

[16] J. Lehner, *Diophantine approximation on Hecke groups*, Glasgow Math. J. 27 (1985), 117–127.

[17] D. Lima, C. Matheus, C. G. Moreira and S. Romaña, *Classical and Dynamical Markov and Lagrange Spectra. Dynamical, Fractal and Arithmetic Aspects*, World Sci., Hackensack, NJ, 2021.

[18] A. V. Malyshev, *Markov and Lagrange spectra* (*survey of the literature*), Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. 67 (1977), 5–38 (in Russian); English transl.: J. Soviet Math. 16 (1981), 767–788.

[19] A. Markoff, *Sur les formes quadratiques binaires indéfinies*, Math. Ann. 15 (1879), 381–406.

[20] A. Markoff, *Sur les formes quadratiques binaires indéfinies. II*, Math. Ann. 17 (1880), 379–399.

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

[22] H. Nakada, *Continued fractions, geodesic flows and Ford circles*, in: Algorithms, Fractals, and Dynamics (Okayama/Kyoto, 1992), Plenum Press, New York, 1995, 179–191.

[23] H. Nakada, *On the Lenstra constant associated to the Rosen continued fractions*, J. Eur. Math. Soc. 12 (2010), 55–70.

[24] L. A. Parson, *Normal congruence subgroups of the Hecke groups $G(2^{(1/2)})$ and $G(3^{(1/2)})$*, Pacific J. Math. 70 (1977), 481–487.

[25] O. Perron, *Über die Approximation irrationaler Zahlen durch rationale II*, Sitzungsber. Heidelberg. Akad. Wiss. 1921, art. 8, 12 pp.

[26] D. Romik, *The dynamics of Pythagorean triples*, Trans. Amer. Math. Soc. 360 (2008), 6045–6064.

[27] D. Rosen, *A class of continued fractions associated with certain properly discontinuous groups*, Duke Math. J. 21 (1954), 549–563.

[28] D. Rosen and T. A. Schmidt, *Hecke groups and continued fractions*, Bull. Austral. Math. Soc. 46 (1992), 459–474.

[29] A. L. Schmidt, *Minimum of quadratic forms with respect to Fuchsian groups. I*, J. Reine Angew. Math. 286/287 (1976), 341–368.

[30] C. Series, *The Markov spectrum in the Hecke group G_5*, Proc. London Math. Soc. 57 (1988), 151–181.

[31] L. Tornheim, *Asymmetric minima of quadratic forms and asymmetric Diophantine approximation*, Duke Math. J. 22 (1955), 287–294.

[32] L. Ya. Vulakh, *The Markov spectrum of imaginary quadratic fields $Q(i\sqrt{D})$, where* $D \not\equiv 3 \pmod 4$, Vestnik Moskov. Univ. Ser. I Mat. Mekh. 26 (1971), no. 6, 32–41 (in Russian).

[33] L. Ya. Vulakh, *The Markov spectra for triangle groups*, J. Number Theory 67 (1997), 11–28.

Dong Han Kim
Department of Mathematics Education
Dongguk University – Seoul
Seoul, 04620 Korea
E-mail: kim2010@dgu.ac.kr

Deokwon Sim
Machine Learning TU
Samsung Advanced Institute of Technology
Suwon, 16678 Korea
E-mail: deokwon.sim@snu.ac.kr