ERGODIC OPTIMIZATION FOR GAUSS'S CONTINUED FRACTION MAP

YINYING HUANG, OLIVER JENKINSON, AND ZHIQIANG LI

Abstract. We extend part of the theory of ergodic optimization for distance-expanding maps to Gauss's continued fraction map. A structure theorem for the closure of the set of invariant measures and the Mañé lemma are established. We prove the typical periodic optimization (TPO) conjecture for α-Hölder essentially compact potentials and the typical finite optimization (TFO) property for rationally maximized potentials.

Contents

1. Introduction

The purpose of this work is to study the properties of maximizing measures for Gauss's continued fraction map $G\colon [0,1] \to [0,1]$, defined by

$$G(x) = \begin{cases} \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor & \text{if } x \in (0,1], \\ 0 & \text{if } x = 0. \end{cases} \tag{1.1}$$

For a general Borel measurable map $T\colon X \to X$ on a metric space X, let $\mathcal{M}(X,T)$ denote the set of T-invariant Borel probability measures on X, let $\overline{\mathcal{M}(X,T)}$ denote the weak* closure of $\mathcal{M}(X,T)$

Key words and phrases. Gauss map, continued fractions, maximizing measure, ergodic optimization.

in the space $\mathcal{P}(X)$ of Borel probability measures on X, and define the *maximum ergodic average* of a bounded Borel measurable function $\psi\colon X \to \mathbb{R}$ to be

$$Q(T,\psi) \coloneqq \sup_{\mu\in\mathcal{M}(X,T)} \int \psi \,\mathrm{d}\mu. \tag{1.2}$$

Any measure $\mu \in \mathcal{M}(X,T)$ that attains the supremum in (1.2) is called a (T,ψ)*-maximizing measure*, or simply a ψ*-maximizing measure*, and the set of such measures is denoted by

$$\mathcal{M}_{\max}(T,\psi) \coloneqq \biggl\{\mu \in \mathcal{M}(X,T) : \int \psi \,\mathrm{d}\mu = Q(T,\psi)\biggr\}. \tag{1.3}$$

If μ is a weak* accumulation point of $\mathcal{M}(X,T)$ with $\int \psi \,\mathrm{d}\mu = Q(T,\psi)$, then μ is called a (T,ψ)-*limit-maximizing measure* and the set of such measures is denoted by

$$\mathcal{M}^*_{\max}(T,\psi) \coloneqq \biggl\{\mu \in \overline{\mathcal{M}(X,T)} : \int \psi \,\mathrm{d}\mu = Q(T,\psi)\biggr\}. \tag{1.4}$$

The *ergodic optimization* problem in this article will be concerned with the study of maximizing, and limit-maximizing, measures for the Gauss map G and suitable real-valued functions ψ. By contrast with previous investigations of ergodic optimization for the Gauss map (see e.g. [Je08, JeSt10, Pi23]), or more general countable branch maps (see e.g. [BF13, BG10, GG24, GGS25, Io07, JMU06, JMU07]), here we shall be concerned with those properties of maximizing measures that are *typical* in spaces of Hölder functions on I. Problems of this kind had attracted the interest of Yuan & Hunt [YH99], who conjectured that for an expanding map, or an Axiom A diffeomorphism, there is an open and dense subset of Lipschitz functions ψ such that the ψ-maximizing measure is supported on a single periodic orbit.

The Yuan-Hunt conjecture stimulated work by various authors (see [Bou01, Bou08, CLT01, Mo08, QS12]), and was eventually proved by Contreras [Co16] in the case of expanding maps, and by Huang, Lian, Ma, Xu & Zhang [HLMXZ25] for Axiom A diffeomorphisms; going beyond the setting of uniform hyperbolicity, Li & Zhang [LZ25] proved the analogous result for expanding Thurston maps from complex dynamics. The so-called typical periodic optimization (TPO) conjecture (cf. [Boc18, Je19]) posits that for more general (suitably chaotic) maps, and spaces of (suitably regular) functions, maximizing measures are supported on a single periodic orbit, for an open dense subset of functions.

While the Gauss map shares certain features of expanding maps, the fact that it has infinitely many branches complicates matters, and the lack of compactness of $\mathcal{M}(I,G)$ makes it appropriate to study limit-maximizing measures rather than maximizing measures. This leads to a new phenomenon: while it can be shown that typical *periodic* optimization is false, it seems to be the case that *typical finite optimization* (TFO) holds; in other words, it is conjectured that there is an open dense subset U of Lipschitz functions on I such that if $\psi \in U$ then the (G,ψ)-limit-maximizing measure is unique and of finite support.

1.1. **Main results.** Henceforth we write $I \coloneqq [0,1]$. The Gauss map $G : I \to I$ defined by (1.1) is such that the set $\mathcal{M}(I,G)$ of G-invariant measures is not weak* compact. Any G-invariant measure $\mu \in \mathcal{M}(I,G)$ satisfying $\mu(I \cap \mathbb{Q}) = 0$ will be called $\mathbb{Q}$-*null*, and the set of such measures will be denoted by $\mathcal{M}_{\mathrm{irr}}(I,G)$, the notation reflecting the fact that the set of irrational members of I is given full mass by measures in $\mathcal{M}_{\mathrm{irr}}(I,G)$.

The noncompactness of $\mathcal{M}(I,G)$ means that certain continuous functions ψ do not have a maximizing measure; a satisfactory theory of ergodic optimization for G will therefore require an understanding of the closure $\overline{\mathcal{M}(I,G)}$.

Each rational number r_0 in $(0,1]$ has a finite G-orbit r_0, r_1, ..., r_n, 0 terminating at the fixed point 0. A related sequence r_0, r_1, ..., $r_n - 1$, 1, 0 corresponds to the *alternative* form of the (finite) continued fraction representation of r_0. A discrete probability measure concentrating equal mass on each of the points in a such sequence of either of the two forms above will be called a *finite-continued-fraction measure* (abbreviated as *FCF measure*). Although such measures are never G-invariant, they are precisely the ingredient needed to understand the closure $\overline{\mathcal{M}(I,G)}$, as described by our first main theorem:

Theorem 1.1. *The closure $\overline{\mathcal{M}(I,G)}$ is equal to the convex hull of the union of $\mathcal{M}_{\mathrm{irr}}(I,G)$ and the set of finite-continued-fraction measures.*

A well-known feature of ergodic optimization, that is useful both for the identification of specific maximizing measures, and for establishing typical properties of such measures in various function spaces, is a result known as a *Mañé lemma*. Our next main result establishes the following Mañé lemma for the Gauss map, in the context of the space $C^{0,\alpha}(I)$ of α-Hölder functions:

Theorem 1.2 (Mañé Lemma). *Suppose $\alpha \in (0,1]$ and $\phi \in C^{0,\alpha}(I)$. There exists $u_\phi \in C^{0,\alpha}(I)$ satisfying the functional equation*

$$u_\phi(x) = \sup_{n\in\mathbb{N}} \Big\{ \overline{\phi}\Big(\frac{1}{n+x}\Big) + u_\phi\Big(\frac{1}{n+x}\Big) \Big\} \quad \text{for all } x \in I, \tag{1.5}$$

where $\overline{\phi} \coloneqq \phi - Q(G,\phi)$.

Let $T\colon X \to X$ be a Borel measurable map on a metric space X, and $\alpha \in (0,1]$. We define $\mathscr{P}(T)$ to be the set of those continuous functions $\phi\colon X \to \mathbb{R}$ with a (T,ϕ)-maximizing measure supported on a periodic orbit, and define $\mathscr{P}^\alpha(T)$ to be those α-Hölder functions in $\mathscr{P}(T)$. If a function $\phi \in \mathscr{P}^\alpha(T)$ satisfies $\operatorname{card} \mathcal{M}_{\max}(T,\phi) = 1$ and $\mathcal{M}_{\max}(T,\phi) = \mathcal{M}_{\max}(T,\psi)$ for all $\psi \in C^{0,\alpha}(X)$ sufficiently close to ϕ in $C^{0,\alpha}(X)$, we say that ϕ has the *(periodic) locking*[1] *property* in $C^{0,\alpha}(X)$ (with respect to T). The set $\mathrm{Lock}^\alpha(T)$ is defined to consist of all $\phi \in \mathscr{P}^\alpha(T)$ satisfying the periodic locking property in $C^{0,\alpha}(X)$.

Similarly, if a function $\phi \in C^{0,\alpha}(X)$ satisfies $\operatorname{card} \mathcal{M}^*_{\max}(T,\phi) = 1$, $\mathcal{M}^*_{\max}(T,\phi) = \mathcal{M}^*_{\max}(T,\psi)$ for all $\psi \in C^{0,\alpha}(X)$ sufficiently close to ϕ in $C^{0,\alpha}(X)$, and the unique limit-maximizing measure is uniformly distributed[2] on a finite set, we say that ϕ has the *finite locking property* in $C^{0,\alpha}(X)$, and define ${\mathrm{Lock_F}}^\alpha(T)$ to consist of those $\phi \in C^{0,\alpha}(T)$ with the finite locking property in $C^{0,\alpha}(X)$.

As we shall see (cf. Example 7.6), there exist functions $\phi \in C^{0,\alpha}(I)$ that are not in the closure of $\mathrm{Lock}^\alpha(G)$, which implies that the Gauss map G does not have the TPO property for α-Hölder potentials.

Inspired by the structure of $\overline{\mathcal{M}(I,G)}$ as revealed by Theorem 1.1, and the guiding philosophy that maximizing measures are generically unique, and should be of low complexity, we state the following *typical finite optimization (TFO) conjecture* for the Gauss map G: for a (topologically) generic α-Hölder function, there is a unique ψ-limit-maximizing measure and it is supported on a finite set.

By Theorem 1.1, the extreme points of $\overline{\mathcal{M}(I,G)}$ are the ergodic invariant probability measures in $\mathcal{M}_{\mathrm{irr}}(I,G)$ together with the set of FCF measures. Any measure $\mu \in \overline{\mathcal{M}(I,G)}$ that is the unique ϕ-limit-maximizing measure for some continuous potential ϕ must be an extreme point of $\mu \in \overline{\mathcal{M}(I,G)}$. It is easy to obtain that any extreme point of $\overline{\mathcal{M}(I,G)}$ supported on a finite set must be a periodic measure or a FCF measure.

Let $\mathfrak{E}^\alpha(G)$ denote the set of α-Hölder functions with a maximizing measure whose support does not contain the point 0. The set $\mathfrak{E}^\alpha(G)$ turns out to be contained in the closure of $\mathrm{Lock}^\alpha(G)$:

[1] The terminology follows [Boc19, BZ15] (see also e.g. [Bou00, Je00]).

[2] A consequence of Theorem 1.1 is that for the Gauss map G, all finitely supported extremal points of $\overline{\mathcal{M}(I,G)}$ are *equidistributions* on their support; by contrast, for more general maps $T\colon I \to I$ with discontinuities, finitely supported extremal points of $\overline{\mathcal{M}(I,T)}$ need not give equal mass to their atoms. So it is natural to replace "uniformly distributed" by "supported" in our definition of the *finite locking property* in such general settings.

Theorem 1.3 (TPO for essentially compact potentials). *For $\alpha \in (0,1]$, the set $\operatorname{Lock}^{\alpha}(G)$ contains an open dense subset of $\mathfrak{E}^{\alpha}(G)$ (in the α-Hölder topology).*

Let $\mathfrak{R}^{\alpha}(G)$ denote the set of those α-Hölder functions such that either the Dirac measure δ_0 or some FCF measure is limit-maximizing. We call functions in $\mathfrak{R}^{\alpha}(G)$ *rationally maximized potentials* due to their close connection to rational orbits (cf. Definition 5.1).

We prove the following typical finite optimization result:

Theorem 1.4 (TFO for rationally maximized potentials). *For $\alpha \in (0,1]$, the set ${\operatorname{Lock}_{\mathrm{F}}}^{\alpha}(G)$ contains an open dense subset of $\mathfrak{R}^{\alpha}(G)$ (in the α-Hölder topology).*

It follows immediately from the definition that ${\operatorname{Lock}_{\mathrm{F}}}^{\alpha}(G) \cap \mathfrak{R}^{\alpha}(G)$ is open in $C^{0,\alpha}(I)$.

1.2. **Organisation of the article.** In Section 2, some frequently used notation and assumptions are fixed. Section 3 consists of a summary of continued fractions (in particular bounded continued fractions, which are used in the proof of Theorem 1.3), as well as the Gauss map, its invariant measures, its inverse branches, and its symbolic coding. In Section 4, we introduce the notion of limit-maximizing measures, and establish some basic properties of maximizing measures. In Section 5 we introduce finite-continued-fraction measures, and prove a structural theorem for the closure of the set of invariant measures of the Gauss map (Theorem 1.1). In Section 6 we discuss the Bousch operator for the Gauss map and prove the existence of a fixed point of the normalised Bousch operator; we are then able to establish the Mañé Lemma (Theorem 1.2). In Section 7 we prove Theorems 1.3 and 1.4. In an appendix we establish the periodic locking property for the Gauss map, which is used in the proof of Theorem 1.3.

2. Notation and assumptions

Let I denote the interval $[0,1]$, equipped with the Euclidean metric. Throughout this article, the golden ratio is denoted by $\theta := \frac{\sqrt{5}+1}{2}$, we set $c_0 := \frac{2\sqrt{5}}{5}$, and for each $\alpha \in (0,1]$ we define

$$K_{\alpha} := \frac{c_0^{-2\alpha}}{1-\theta^{-2\alpha}} = c_0^{-2\alpha} \sum_{n=0}^{+\infty} \theta^{-2\alpha n}. \tag{2.1}$$

We follow the convention that $\mathbb{N} := \{1, 2, 3, \dots\}$, $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$, and $\widehat{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$. The sets $\mathbb{N}$, $\mathbb{N}_0$, and $\widehat{\mathbb{N}}$ will all be equipped with the order relations $<$, $\leqslant$, $>$, $\geqslant$, defined in the obvious way. We endow $\mathbb{N}$ with the discrete topology and $\mathbb{N}^n$, $\mathbb{N}^{\mathbb{N}}$ with the product topology for each $n \in \mathbb{N}$. Let ρ be the metric on $\mathbb{N}$ defined by

$$\rho(a,b) := \begin{cases} \left|\frac{1}{a} - \frac{1}{b}\right| & \text{if } a, b \in \mathbb{N} \text{ and } a \neq b, \\ 0 & \text{if } a = b. \end{cases}$$

Observe that ρ is totally bounded, and the completion of $\mathbb{N}$ with respect to ρ is a compact metric space that can be identified with $\widehat{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$. The metric ρ extends to $\widehat{\mathbb{N}}$, and we denote the extension by $\widehat{\rho}$. Note that $\widehat{\rho}(n,\infty) = \frac{1}{n}$ for every $n \in \mathbb{N}$, and $\widehat{\rho}(\infty,\infty) = 0$. We also endow $\widehat{\mathbb{N}}^n$ with the product topology for each $n \in \mathbb{N}$.

For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer $\leqslant x$. As usual, log will denote the natural logarithm, and $\log_c$ the logarithm to base c for $c > 0$. The cardinality of a set A is denoted by $\operatorname{card} A$. For sets A, B, denote $A \triangle B = (A \smallsetminus B) \cup (B \smallsetminus A)$.

The collection of all maps from a set X to a set Y is denoted by Y^X. We denote the restriction of a map $g\colon X \to Y$ to a subset Z of X by $g|_Z$. Given a map $f\colon X \to X$ and a real-valued function $\psi\colon X \to \mathbb{R}$, we write $S_n\psi(x) := \sum_{j=0}^{n-1} \psi\bigl(f^j(x)\bigr)$ for $x \in X$ and $n \in \mathbb{N}_0$. Note that by definition we always have $S_0\psi = 0$. We denote the set of bounded functions from X to $\mathbb{R}$ by $B(X)$.

Let (X,d) be a metric space. For subsets A, $B \subseteq X$, we set $d(A,B) := \inf\{d(x,y) : x \in A,\, y \in B\}$, and $d(A,x) = d(x,A) := d(A,\{x\})$ for $x \in X$. For each subset $Y \subseteq X$, we denote the diameter of Y by $\mathrm{diam}_d(Y) \coloneqq \sup\{d(x,y) : x, y \in Y\}$, the interior of Y by $\mathrm{int}\, Y$, and the max-plus characteristic function of Y by $\mathbb{0}_Y$, which maps each $x \in Y$ to $0 \in \mathbb{R}$ and vanishes otherwise. We use the convention that $\mathbb{0} = \mathbb{0}_X$ when the space X is clear from the context. For each $r > 0$, we define $B^r_d(A)$ to be the open r-neighbourhood $\{y \in X : d(y,A) < r\}$ of A, and $\overline{B}^r_d(A)$ the closed r-neighbourhood $\{y \in X : d(y,A) \leqslant r\}$ of A. For $x \in X$, we denote the open (resp. closed) ball of radius r centered at x by $B_d(x,r)$ (resp. $\overline{B}_d(x,r)$). For $\mu \in \mathcal{M}(X)$ and a μ-integrable function $\phi\colon X \to \mathbb{R}$, we write

$$\langle \mu, \phi \rangle \coloneqq \int_I \phi \,\mathrm{d}\mu.$$

We set $C(X)$ (resp. $B(X)$) to be the space of continuous (resp. bounded) functions from X to $\mathbb{R}$, $\mathcal{M}(X)$ the set of finite signed Borel measures, and $\mathcal{P}(X)$ the set of Borel probability measures on X. If we do not specify otherwise, we equip $C(X)$ with the uniform norm $\|\cdot\|_\infty$. For a Borel measurable map $g\colon X \to X$, $\mathcal{M}(X,g)$ is the set of g-invariant Borel probability measures on X and $\mathcal{M}_{\mathrm{erg}}(X,g)$ is the set of ergodic measures in $\mathcal{M}(X,g)$. For each $x \in X$, we denote by δ_x the Dirac delta measure on x given by $\delta_x(A) = 1$ if $x \in A$ and 0 otherwise for all Borel measurable set $A \subseteq X$.

The space of real-valued Hölder continuous functions with an exponent $\alpha \in (0,1]$ is denoted by $C^{0,\alpha}(X,d)$. For each $\psi \in C^{0,\alpha}(X,d)$, we denote

$$|\psi|_{\alpha,X} \coloneqq \sup\{|\psi(x) - \psi(y)|/d(x,y)^\alpha : x, y \in X, x \neq y\},$$

and the Hölder norm is given by $\|\psi\|_{\alpha,X} \coloneqq |\psi|_{\alpha,X} + \|\psi\|_\infty$. We omit the subscript X when it does not cause confusion.

For a nonempty set E and $n \in \mathbb{N}$, let us denote $E^{\mathbb{N}} \coloneqq \{(a_k)_{k\in\mathbb{N}} : a_k \in E \text{ for all } k \in \mathbb{N}\}$, $E^n \coloneqq \{(a_k)_{k=1}^n : a_k \in \mathcal{A} \text{ for all } 1 \leqslant k \in n\}$, and $E^* \coloneqq \bigcup_{k\in\mathbb{N}} E^k$.

Define the (left) shift map

$$\sigma = \sigma_{E^{\mathbb{N}}}\colon E^{\mathbb{N}} \to E^{\mathbb{N}}$$

by $\sigma(A) \coloneqq (a_{n+1})_{n\in\mathbb{N}}$ for all $A = (a_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$. We often omit the subscript $E^{\mathbb{N}}$ in $\sigma_{E^{\mathbb{N}}}$ when it is clear from the context. Sometimes we also consider this shift as defined on words of finite length. Given $A \in E^*$, by $|A|$ we denote the length of the word A, i.e., the unique n such that $A \in E^n$. If $A \in E^{\mathbb{N}}$ and $n \in \mathbb{N}$, then denote $A|_n = a_1a_2\dots a_n$.

We denote $\widehat{\Sigma} \coloneqq \widehat{\mathbb{N}}^{\mathbb{N}}$ and $\Sigma_m \coloneqq \{1, 2, \dots, m\}^{\mathbb{N}}$.

For a sequence of natural numbers $\{a_i\}_{i\in\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}}$, for each $n \in \mathbb{N}$ we denote by $[a_1,\dots,a_n]$ the rational number

$$[a_1,\dots,a_n] \coloneqq \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}},$$

and denote $[a_1,a_2,\dots] \coloneqq \lim_{n\to\infty}[a_1,\dots,a_n]$, the real number in I with continued fraction $(a_n)_{n\in\mathbb{N}}$. We denote by $[\overline{a_1,\dots,a_n}]$ the number with periodic continued fraction.

3. The Gauss map and continued fractions

3.1. **Continued fractions.**

Definition 3.1 (Continued fraction transformation). The *continued fraction transformation* (or *Gauss*[3] *map*), is the map $G\colon I \to I$ defined by

$$G(x) = \begin{cases} \frac{1}{x} - \lfloor \frac{1}{x} \rfloor & \text{if } x \in (0,1], \\ 0 & \text{if } x = 0. \end{cases}$$

[3]The common convention of referring to G as the Gauss map reflects the foundational work of Gauss [Ga12] (cf. e.g. [Br91]) on the statistical properties of continued fractions (having worked on probabilistic aspects of continued fractions in 1799–1800 (cf. [Ga12]), his understanding of the so-called Gauss measure was mentioned in an 1812 letter to Laplace (see [Br91, pp. 147–148]). As a self-map of $I = [0,1]$, the value $G(0)$ is habitually defined to be 0 (see e.g. [FSU14, Ma87, PW99]).

Note that $G^{-1}(1) = \emptyset$, $G^{-1}(0) = \{0\} \cup \{1/n\}_{n\in\mathbb{N}}$, and

$$G^{-1}(x) = \{1/(a+x) : a \in \mathbb{N}\} \qquad \text{for } x \in (0,1). \tag{3.1}$$

The connection between G and continued fractions is that for each irrational $x \in I$, there are unique natural numbers $a_i(x) := \lfloor 1/G^{i-1}(x) \rfloor$ such that

$$x = [a_1(x), a_2(x), \dots] := \lim_{n\to+\infty} [a_1(x), a_2(x), \dots, a_n(x)], \tag{3.2}$$

and (3.2) is the *continued fraction expansion* of x (cf. [Sc80, Lemma 4D]).

Lemma 3.2. *Every $x \in (0,1) \cap \mathbb{Q}$ has precisely two finite continued fraction expansions, one of the form $[a_1, a_2, \dots, a_n]$ with $a_n \geqslant 2$, and the other of the form $[a_1, a_2, \dots, a_n - 1, 1]$.*

Proof. See [Sc80, Lemma 4C]. □

To accommodate the nonuniqueness of continued fraction expansions of rational numbers, it will be convenient to distinguish between those expansions which do, or do not, contain the digit 1, in the following way:

Notation 3.3. We let

$$R_1 := \{1/a : a \in \mathbb{N},\ a \geqslant 2\} = \{[a] : a \in \mathbb{N},\ a \geqslant 2\}$$

denote the set of unit fractions with the value 1 excluded, or in other words

$$R_1 = G^{-1}(0) \smallsetminus \{0,\ 1\}.$$

For $n \geqslant 2$, set

$$R_n := G^{-(n-1)}(R_1) = \{[a_1, a_2, \dots, a_n] : (a_1, \dots, a_n) \in \mathbb{N}^n,\ a_n \geqslant 2\}. \tag{3.3}$$

Note that

$$R_n = G^{-n}(0) \smallsetminus G^{-(n-1)}(0),$$

in other words R_n is the set of those points in I whose G-orbit lands at 0 for the first time after precisely n iterates.[4]

For each $n \in \mathbb{N}$, define $\mathcal{A}_n \subseteq \mathbb{N}^n$ by

$$\mathcal{A}_n := \{(a_1, \dots, a_n) \in \mathbb{N}^n : a_n \geqslant 2\}. \tag{3.4}$$

Denote $\mathcal{A} := \bigcup_{n=1}^{+\infty} \mathcal{A}_n \subseteq \mathbb{N}^*$. For each $n \in \mathbb{N}$, define $\mathcal{B}_n \subseteq \mathbb{N}^n$ by

$$\mathcal{B}_n := \{(a_1, \dots, a_n) \in \mathbb{N}^n : a_n = 1\}.$$

Denote $\mathcal{B} := \bigcup_{n=1}^{+\infty} \mathcal{B}_n \subseteq \mathbb{N}^*$. Clearly, for each $n \in \mathbb{N}$, the sets $\mathcal{A}_n$ and $\mathcal{B}_n$ together form a partition of $\mathbb{N}^n$.

For each $n \in \mathbb{N}$, define the bijection $f_n \colon \mathcal{A}_n \to \mathcal{B}_{n+1}$ by

$$(a_1, a_2, \dots, a_n) \mapsto (a_1, a_2, \dots, a_n - 1, 1).$$

Let $f \colon \mathcal{A} \to \mathcal{B}$ be the function defined by $f|_{\mathcal{A}_n} := f_n$ for all $n \in \mathbb{N}$, Define $g_n \colon \mathcal{A}_n \to R_n$ by

$$g_n : (a_1, a_2, \dots, a_n) \mapsto [a_1, a_2, \dots, a_n]. \tag{3.5}$$

It is easy to see that g_n maps $\mathcal{A}_n$ bijectively to R_n (see Lemma 5.5 (iv)). Let $g \colon \mathcal{A} \to (0,1)$ be the function defined by $g|_{\mathcal{A}_n} := g_n$ for all $n \in \mathbb{N}$.

[4] Note that the point 1 is not in the image of G, so the only eventually fixed orbit containing 1 is the two-element orbit $\{0,\ 1\}$.

Let a_1, a_2, a_3, ... be variables. We define polynomials p_0, q_0, p_1, q_1, p_2, q_2, ..., with p_n and q_n being polynomials in a_1, ..., a_n, as follows:

$$p_0 = 0, \quad p_{-1} = 1, \quad p_n = a_n p_{n-1} + p_{n-2}, \quad n = 1, 2, \dots$$
$$q_0 = 1, \quad q_{-1} = 0, \quad q_n = a_n q_{n-1} + q_{n-2}, \quad n = 1, 2, \dots$$

Definition 3.4 (Continuants). Fix $n \in \mathbb{N}$. For any finite word $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}^n$, the integers $p_n(\mathbf{a})$ and $q_n(\mathbf{a})$ satisfy

$$[a_1, a_2, \dots, a_n] = \frac{p_n(\mathbf{a})}{q_n(\mathbf{a})}.$$

The denominator $q_n(\mathbf{a})$ is called a *continuant* of the continued fraction $[a_1, \dots, a_n]$. We also define the $q_k(\mathbf{a})$ for $k < n$, even if the word $\mathbf{a}$ has length larger than k; in this case $q_k(\mathbf{a})$ is just the continuant $q_k(\mathbf{a}|_k)$, where $\mathbf{a}|_k$ is the restriction $(a_1, \dots, a_k)$.

The following lower bound on the growth of continuants is expressed in terms of the constants c_0 and θ (defined in Section 2):

Lemma 3.5. *If $n \in \mathbb{N}$ and $\mathbf{a} \in \mathbb{N}^n$, then $q_n(\mathbf{a}) \geqslant c_0\theta^n$.*

Proof. See [JoSa16, Lemma 1]. □

3.2. **Invariant measures.**

Notation 3.6. Fix $\mu \in \mathcal{M}(I)$. Define $P_1(\mu)$, $P_2(\mu) \in \mathcal{M}(I)$ by

$$P_1(\mu)(A) \coloneqq \mu(A \cap \mathbb{Q}) \quad \text{and} \quad P_2(\mu)(A) \coloneqq \mu(A \smallsetminus \mathbb{Q})$$

for all Borel sets $A \subseteq I$.

Lemma 3.7. *We have $G(I \cap \mathbb{Q}) = [0,1) \cap \mathbb{Q}$, $G^{-1}(I \cap \mathbb{Q}) = I \cap \mathbb{Q}$, and $G(I \smallsetminus \mathbb{Q}) = G^{-1}(I \smallsetminus \mathbb{Q}) = I \smallsetminus \mathbb{Q}$. The set $\mathcal{M}(I,G)$ is the convex hull of $\mathcal{M}_{\mathrm{irr}}(I,G) \cup \{\delta_0\}$ and $\overline{\mathcal{M}(I,G)} = \overline{\mathcal{M}_{\mathrm{irr}}(I,G)}$.*

Proof. The first part follows immediately from Definition 3.1. Clearly, $\delta_0 \in \mathcal{M}(I,G)$ and $\mathcal{M}_{\mathrm{irr}}(I,G) \subseteq \mathcal{M}(I,G)$.

Fix $\mu \in \mathcal{M}(I,G)$. For each Borel set $A \subseteq I$, since $G^{-1}(I \cap \mathbb{Q}) = I \cap \mathbb{Q}$ and $G^{-1}(I \smallsetminus \mathbb{Q}) = I \smallsetminus \mathbb{Q}$, we have

$$G^{-1}(A \cap \mathbb{Q}) = G^{-1}(A) \cap G^{-1}(I \cap \mathbb{Q}) = G^{-1}(A) \cap \mathbb{Q} \quad \text{and}$$
$$G^{-1}(A \smallsetminus \mathbb{Q}) = G^{-1}(A) \cap G^{-1}(I \smallsetminus \mathbb{Q}) = G^{-1}(A) \smallsetminus \mathbb{Q}.$$

Thus, for each Borel set $A \subseteq I$, we have

$$P_1(\mu)\big(G^{-1}(A)\big) = \mu\big(G^{-1}(A) \cap \mathbb{Q}\big) = \mu\big(G^{-1}(A \cap \mathbb{Q})\big) = \mu(A \cap \mathbb{Q}) = P_1(\mu)(A),$$
$$P_2(\mu)\big(G^{-1}(A)\big) = \mu\big(G^{-1}(A) \smallsetminus \mathbb{Q}\big) = \mu\big(G^{-1}(A \smallsetminus \mathbb{Q})\big) = \mu(A \smallsetminus \mathbb{Q}) = P_2(\mu)(A).$$

Hence $P_1(\mu)$ and $P_2(\mu)$ are G-invariant. Note that

$$\mu(\{0\}) = \mu\big(G^{-1}(0)\big) = \mu(\{0\}) + \sum_{n=1}^{+\infty} \mu(\{1/n\}).$$

Since μ is G-invariant, we have $\mu(\{1/n\}) = 0$ for all $n \in \mathbb{N}$ and $\mu\big(G^{-k}(1/n)\big) = 0$, for all n, $k \in \mathbb{N}$. Consequently, we obtain $\mu((I \cap \mathbb{Q}) \smallsetminus \{0\}) = 0$ and $P_1(\mu) = \mu(\{0\})\delta_0$. When $\mu(\{0\}) = 1$, we have $\mu = P_1(\mu) = \delta_0$. When $\mu(\{0\}) < 1$, we have $\frac{1}{\mu(I \smallsetminus \mathbb{Q})} P_2(\mu) \in \mathcal{M}_{\mathrm{irr}}(I,G)$. Therefore, μ is contained in the convex hull of $\mathcal{M}_{\mathrm{irr}}(I,G) \cup \{\delta_0\}$. But μ was an arbitrary member of $\mathcal{M}(I,G)$, so we see that $\mathcal{M}(I,G)$ is the convex hull of $\mathcal{M}_{\mathrm{irr}}(I,G) \cup \{\delta_0\}$.

Now $\mathcal{M}_{\mathrm{irr}}(I,G) \subseteq \mathcal{M}(I,G)$, so $\overline{\mathcal{M}_{\mathrm{irr}}(I,G)} \subseteq \overline{\mathcal{M}(I,G)}$, and the reverse inclusion $\overline{\mathcal{M}(I,G)} \subseteq \overline{\mathcal{M}_{\mathrm{irr}}(I,G)}$ follows from $\delta_0 \in \overline{\mathcal{M}_{\mathrm{irr}}(I,G)}$, since $\lim_{n \to +\infty} [\overline{n}] = 0$ and $\delta_{[\overline{n}]} \in \mathcal{M}_{\mathrm{irr}}(I,G)$, so the second part now follows. □

Remark 3.8. The set $\mathcal{M}(I,G)$ is not closed with respect to the weak* topology. Write $x_n := [2,n,\overline{2,n}]$ for $n \in \mathbb{N}$. Then $\{x_n\}_{n\in\mathbb{N}}$ are the periodic points of G satisfying $G^2(x_n) = x_n$, $\lim_{n\to+\infty} x_n = \frac{1}{2}$, and $\lim_{n\to+\infty} G(x_n) = 0$. Define $\mu_n := \frac{1}{2}(\delta_{x_n} + \delta_{G(x_n)})$. Then the weak* limit of the sequence $\{\mu_n\}_{n\in\mathbb{N}}$ is $\mu = \frac{1}{2}(\delta_0 + \delta_{1/2})$, which is not G-invariant (see Lemma 3.7).

The following lemma allows us to abuse notation by identifying $\mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}})$ with $\mathcal{M}_{\mathrm{irr}}(I,G) \subseteq \mathcal{M}(I,G)$. Define $j\colon I\smallsetminus\mathbb{Q} \to I$ to be the inclusion map and $j_*\colon \mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}}) \to \mathcal{M}(I,G)$ to be the pushforward of j, i.e., for each Borel subset $A \subseteq I$ and $\mu \in \mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}})$, we define

$$j_*(\mu)(A) := \mu\big(j^{-1}(A)\big) = \mu(A\smallsetminus\mathbb{Q}). \tag{3.6}$$

Lemma 3.9. *The map j_* is a continuous bijection from $\mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}})$ to $\mathcal{M}_{\mathrm{irr}}(I,G)$.*

Proof. By the fact that j is continuous, j_* is well-defined and continuous. Assume that $\mu_1, \mu_2 \in \mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}})$ with $j_*(\mu_1) = j_*(\mu_2)$. Then by (3.6), $\mu_1(A\smallsetminus\mathbb{Q}) = \mu_2(A\smallsetminus\mathbb{Q})$ for every Borel measurable subset $A \subseteq I$. So $\mu_1 = \mu_2$ and consequently j_* is injective.

Fix $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$. Define $\nu \in \mathcal{M}(I\smallsetminus\mathbb{Q})$ by $\nu(A) := \mu(A)$ for each Borel subset $A \subseteq I\smallsetminus\mathbb{Q}$. By the assumption that $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$ and the definition of $\mathcal{M}_{\mathrm{irr}}(I,G)$, we get $\nu(I\smallsetminus\mathbb{Q}) = 1$. For each Borel subset $A \subseteq I\smallsetminus\mathbb{Q}$, from the fact that $G^{-1}(I\smallsetminus\mathbb{Q}) = I\smallsetminus\mathbb{Q}$, the definition of ν, and the fact that $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$,

$$\nu\big((G|_{I\smallsetminus\mathbb{Q}})^{-1}(A)\big) = \nu\big(G^{-1}(A)\big) = \mu\big(G^{-1}(A)\big) = \mu(A) = \nu(A).$$

Hence $\nu \in \mathcal{M}(I\smallsetminus\mathbb{Q}, G|_{I\smallsetminus\mathbb{Q}})$. For each Borel subset $B \subseteq I$, by (3.6), the definition of ν, and the fact that $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$, we have $j_*(\nu)(B) = \nu(B\smallsetminus\mathbb{Q}) = \mu(B\smallsetminus\mathbb{Q}) = \mu(B)$. So $j_*(\nu) = \mu$ and consequently j_* is a surjective map from $\mathcal{M}(I\smallsetminus\mathbb{Q}, T|_{I\smallsetminus\mathbb{Q}})$ to $\mathcal{M}_{\mathrm{irr}}(I,G)$. This lemma now follows. ☐

3.3. **Inverse branches.** The following notational conventions for inverse branches of the Gauss map follow Jordan & Sahlsten [JoSa16].

Definition 3.10 (Inverse branches). For each $a \in \mathbb{N}$, define $G_a\colon I \to I$ by

$$G_a(x) := \frac{1}{a+x} \quad \text{for all } x \in I, \tag{3.7}$$

so that (3.1) becomes

$$G^{-1}(x) = \{G_a(x) : a \in \mathbb{N}\} \quad \text{for all } x \in (0,1).$$

Let us denote $I_a := G_a(I)$. For each $n \in \mathbb{N}$ and $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{N}^n$, define $G_{\mathbf{a}}\colon I \to I$ to be

$$G_{\mathbf{a}} := G_{a_1} \circ G_{a_2} \circ \cdots \circ G_{a_n}. \tag{3.8}$$

In other words, $G_{\mathbf{a}}(x) = [a_1, a_2, \dots, a_n + x]$ for each $\mathbf{a} \in \mathbb{N}^n$ and each $x \in I$. Denote $I_{\mathbf{a}} := G_{\mathbf{a}}(I)$.

Notation 3.11. It will be convenient to define $G_\infty\colon I \to I$ by setting $G_\infty(x) = 0$ for each $x \in I$.

Recalling that $\widehat{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$, for $n \in \mathbb{N}$ and $\widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n$ we define $G_{\widehat{\mathbf{a}}} = (\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_n)\colon I \to I$ by

$$G_{\widehat{\mathbf{a}}} := G_{\widehat{a}_1} \circ G_{\widehat{a}_2} \circ \cdots \circ G_{\widehat{a}_n}, \tag{3.9}$$

and set

$$I_{\widehat{\mathbf{a}}} := G_{\widehat{\mathbf{a}}}(I).$$

For $\widehat{\mathbf{a}} = a_1 a_2 \cdots \in \widehat{\mathbb{N}}^{\mathbb{N}} \cup \bigcup_{n=1}^{+\infty} \widehat{\mathbb{N}}^n$, we define its ∞-*index* $\iota(\widehat{\mathbf{a}})$ to be the smallest k such that $\widehat{a}_k = \infty$, so that $\iota(\widehat{\mathbf{a}}) = +\infty$ precisely when $\widehat{A} \in \mathbb{N}^{\mathbb{N}} \cup \bigcup_{n=1}^{+\infty} \mathbb{N}^n$.

Lemma 3.12. *If $n \in \mathbb{N}$ and $\mathbf{a} \in \mathbb{N}^n$, then*

$$q_n(\mathbf{a})^{-2}/4 \leqslant |G'_{\mathbf{a}}| \leqslant q_n(\mathbf{a})^{-2},$$

and in particular, the length $\operatorname{diam} I_{\mathbf{a}}$ *satisfies* $\frac{1}{4} q_n(\mathbf{a})^{-2} \leqslant \operatorname{diam} I_{\mathbf{a}} \leqslant q_n(\mathbf{a})^{-2}$.

Proof. See [JoSa16, Lemma 2]. □

Proposition 3.13. *If $n \in \mathbb{N}$, then*

(i) $|G'_{\mathbf{a}}(x)| \leqslant c_0^{-2}\theta^{-2n}$ *for all $x \in I$ and $\mathbf{a} \in \mathbb{N}^n$.*

(ii) *For $\widehat{\mathbf{a}} = a_1a_2\dots a_n \in \widehat{\mathbb{N}}^n \smallsetminus \mathbb{N}^n$,*

$$I_{\widehat{\mathbf{a}}} = \begin{cases} \{0\} & \text{if } \iota(\widehat{\mathbf{a}}) = 1, \\ \{[a_1, a_2, \dots, a_{\iota(\widehat{\mathbf{a}})-1}]\} & \text{if } \iota(\widehat{\mathbf{a}}) \geqslant 2. \end{cases}$$

(iii) *For $\mathbf{a} \in \mathbb{N}^n$, the map $G_{\mathbf{a}}$ is strictly increasing for n even, and strictly decreasing for n odd.*

(iv) *For each $x \in I$, the map $\widehat{\mathbb{N}}^n \to I$, $\widehat{\mathbf{a}} \mapsto G_{\widehat{\mathbf{a}}}(x)$, is continuous.*

(v) *For each $x \in I$, the closure of $\{G_{\mathbf{a}}(x) : \mathbf{a} \in \mathbb{N}^n\}$ is the set $\{G_{\widehat{\mathbf{a}}}(x) : \widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n\}$.*

(vi) *If $x \in I \smallsetminus \mathbb{Q}$ and $\mathbf{a} \in \mathbb{N}^n$, then $G_{\sigma^i(\mathbf{a})}(x) = G^i(G_{\mathbf{a}}(x))$ for all integer $1 \leqslant i \leqslant n-1$.*

Proof. (i) Lemmas 3.5 and 3.12 immediately yield $|G'_{\mathbf{a}}(x)| \leqslant q_n(\mathbf{a})^{-2} \leqslant c_0^{-2}\theta^{-2n}$.

(ii) follows from the definition of $G_{\widehat{\mathbf{a}}}$ (see (3.9)) and the fact that $G_\infty \equiv 0$ on I.

(iii) is immediate from the fact that G_a is strictly decreasing for each $a \in \mathbb{N}$ (cf. (3.7)), and $G_{\mathbf{a}}$ is the composition of n such maps (cf. (3.8)).

(iv) Suppose as m tends to infinity, $\widehat{\mathbf{a}}_m = a_{m,1}a_{m,2}\dots a_{m,n}$ converges to $\widehat{\mathbf{a}} = a_1a_2\dots a_n$ in $\widehat{\mathbb{N}}^n$.

If $\widehat{\mathbf{a}} \in \mathbb{N}^n$, then $a_{m,k} = a_k$ for all $k \in \{1, 2, \dots, n\}$ when m is large enough. So, $G_{\widehat{\mathbf{a}}_m}(x) = T_{\widehat{\mathbf{a}}}(x)$ when m is large enough.

If $\widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n \smallsetminus \mathbb{N}^n$, let $l := \iota(\widehat{\mathbf{a}})$ be the ∞-index of $\widehat{\mathbf{a}}$. When $l = 1$, we get $a_1 = \infty$ and $\lim_{m\to+\infty} a_{m,1} = +\infty$. So $G_{\widehat{\mathbf{a}}}(x) = 0 = \lim_{m\to+\infty} G_{\widehat{\mathbf{a}}_m}(x)$. When $l \geqslant 2$, we obtain $G_{\widehat{\mathbf{a}}}(x) = [a_1, a_2, \dots, a_{l-1}]$, $a_{m,k} = a_k$ for all $k \in \{1, 2, \dots, l-1\}$ when m is large enough, and $\lim_{m\to+\infty} a_{m,l} = +\infty$. Hence, $\lim_{m\to+\infty} G_{\widehat{\mathbf{a}}_m}(x) = G_{\widehat{\mathbf{a}}}(x)$, and (iv) follows.

(v) Denote $W := \{G_{\mathbf{a}}(x) : \mathbf{a} \in \mathbb{N}^n\}$ and $\widehat{W} := \{G_{\widehat{\mathbf{a}}}(x) : \widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n\}$. Fix $y \in \widehat{W} \smallsetminus W$. Then $y = G_{\widehat{\mathbf{b}}}(x)$ for some $\widehat{\mathbf{b}} = b_1b_2\dots b_n \in \widehat{\mathbb{N}}^n \smallsetminus \mathbb{N}^n$. Let $k := \iota(\widehat{\mathbf{b}})$ be the ∞-index of $\widehat{\mathbf{b}}$. By (ii), we get $y = G_{\widehat{\mathbf{b}}}(x) = 0$ when $k = 1$ and $y = G_{\widehat{\mathbf{b}}}(x) = [b_1, \dots, b_{k-1}]$ when $k \geqslant 2$. For each m, define $\mathbf{b}_m := (m, \dots, m) \in \mathbb{N}^n$ when $k = 1$ and define $\mathbf{b}_m := (b_1, \dots, b_{k-1}, m, \dots, m) \in \mathbb{N}^n$ when $k \geqslant 2$. It is easy to check $\lim_{m\to+\infty} G_{\mathbf{b}_m}(x) = y$. Hence, $\widehat{W} \subseteq \overline{W}$.

Assume that $z \in \overline{W}$. Then there exists a sequence $\{\mathbf{a}_m\}_{m\in\mathbb{N}}$ in $\mathbb{N}^n$ with $\lim_{m\to+\infty} G_{\mathbf{a}_m}(x) = z$. Since $\widehat{\mathbb{N}}^n$ is compact, there exists $\mathbf{c} \in \widehat{\mathbb{N}}^n$ and a subsequence $\{\mathbf{a}_{m_k}\}_{k\in\mathbb{N}}$ of $\{\mathbf{a}_m\}_{m\in\mathbb{N}}$ such that $\{\mathbf{a}_{m_k}\}_{k\in\mathbb{N}}$ converges to $\mathbf{c}$ as k tends to infinity. By (iv), we obtain that

$$G_{\mathbf{c}}(x) = \lim_{k\to+\infty} G_{\mathbf{a}_{m_k}}(x) = \lim_{m\to+\infty} G_{\mathbf{a}_m}(x) = z,$$

so $z \in \widehat{W}$. But z was an arbitrary member of $\overline{W}$, so we have shown that $\overline{W} \subseteq \widehat{W}$, and (v) follows.

(vi) Assume that $x = [b_1, b_2, \dots, b_n, \dots] \in I \smallsetminus \mathbb{Q}$ and $\mathbf{a} = [a_1, \dots, a_n]$. Then, by the definition of $G_{\mathbf{a}}$ and G, we see that $G_{\sigma^i(\mathbf{a})}(x) = [a_{i+1}, \dots, a_n, b_1, b_2, \dots, b_n, \dots] = G^i(G_{\mathbf{a}}(x))$, as required. □

Notation 3.14. Given $\psi \in \mathbb{R}^I$, $n \in \mathbb{N}$, and $\widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n$, define the function $S_{n,\widehat{\mathbf{a}}}\psi \colon I \to \mathbb{R}$ by

$$S_{n,\widehat{\mathbf{a}}}\psi := \sum_{i=0}^{n-1} \psi \circ G_{\sigma^i(\widehat{\mathbf{a}})}. \tag{3.10}$$

We need the following standard lemma (cf. [Wa78, pp. 144–147], and see also [MU03, Lemma 3.1.2] in the context of conformal graph directed Markov systems).

Lemma 3.15. *Suppose $\alpha \in (0,1]$ and $\phi \in C^{0,\alpha}(I)$. For all $n \in \mathbb{N}$, $\mathbf{a} \in \mathbb{N}^n$, and $x, y \in I$,*

$$|S_{n,\mathbf{a}}\phi(x) - S_{n,\mathbf{a}}\phi(y)| \leqslant K_\alpha |\phi|_\alpha |x-y|^\alpha.$$

Proof. The inequality clearly holds if $x = y$, while if $x \neq y$ then the α-Hölder assumption, together with the intermediate value theorem, implies that there exists some ξ_i in between x and y for each $0 \leqslant i \leqslant n-1$ such that

$$|S_{n,\mathbf{a}}\phi(x) - S_{n,\mathbf{a}}\phi(y)| \leqslant |\phi|_\alpha \sum_{i=0}^{n-1} |G_{\sigma^i(\mathbf{a})}(x) - G_{\sigma^i(\mathbf{a})}(y)|^\alpha = |\phi|_\alpha \sum_{i=0}^{n-1} \bigl|G'_{\sigma^i(\mathbf{a})}(\xi_i)\bigr|^\alpha |x-y|^\alpha,$$

and the result follows readily from Proposition 3.13 (i) and the definition of K_α (see (2.1). □

3.4. **Symbolic dynamics.**

Notation 3.16. Define the $\mathbb{N}$-*valued full shift* Σ by

$$\Sigma := \mathbb{N}^{\mathbb{N}},$$

and as usual define the shift map $\sigma \colon \Sigma \to \Sigma$ by $\sigma((a_n)_{n\in\mathbb{N}}) := (a_{n+1})_{n\in\mathbb{N}}$.

Recall that $\rho(a,b) := \bigl|\frac{1}{a} - \frac{1}{b}\bigr|$ for a, $b \in \mathbb{N}$, and $\widehat{\rho}$ denotes the extension of ρ to $\widehat{\mathbb{N}}$ given by defining $\widehat{\rho}(a,\infty) := \frac{1}{a}$ for $a \in \mathbb{N}$, and $\widehat{\rho}(\infty,\infty) := 0$ (cf. Section 2).

Recalling that $\theta := \frac{\sqrt{5}+1}{2}$, define the metric d_ρ on Σ by

$$d_\rho(A,B) := \sum_{n\in\mathbb{N}} \theta^{-2n} \rho(a_n, b_n),$$

where $A = (a_n)_{n\in\mathbb{N}}$, $B = (b_n)_{n\in\mathbb{N}} \in \Sigma$.

Definition 3.17 (Compactification of (Σ, d_ρ)). Since the metric d_ρ is totally bounded, its metric completion, denoted by $\widehat{\Sigma}$, is a compact metric space. In particular, $\widehat{\Sigma}$ is a compactification of Σ. The compactification $\widehat{\Sigma}$ can be described more explicitly. More precisely, $\widehat{\Sigma}$ can be identified with $\widehat{\mathbb{N}}^{\mathbb{N}}$ equipped with the extended metric $d_{\widehat{\rho}}$, where

$$d_{\widehat{\rho}}\bigl(\widehat{A}, \widehat{B}\bigr) := \sum_{n\in\mathbb{N}} \theta^{-2n} \widehat{\rho}\bigl(\widehat{a}_n, \widehat{b}_n\bigr), \tag{3.11}$$

for $\widehat{A} = (\widehat{a}_n)_{n\in\mathbb{N}}$, $\widehat{B} = \bigl(\widehat{b}_n\bigr)_{n\in\mathbb{N}} \in \widehat{\mathbb{N}}^{\mathbb{N}}$.

The shift map σ extends to a continuous self-map $\sigma : \widehat{\Sigma} \to \widehat{\Sigma}$ given by $\sigma((\widehat{a}_n)_{n\in\mathbb{N}}) := (\widehat{a}_{n+1})_{n\in\mathbb{N}}$.

For each $n \in \mathbb{N}$, define the *cylinder set* $\mathcal{C}(\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_n)$ to consist of those sequences $\bigl(\widehat{b}_i\bigr)_{i\in\mathbb{N}} \in \widehat{\Sigma}$ such that $\widehat{b}_1\widehat{b}_2 \dots \widehat{b}_n = \widehat{a}_1\widehat{a}_2 \dots \widehat{a}_n$.

Notation 3.18. Define the homeomorphism (cf. [Mi17, Theorem 1.1]) $\pi \colon \Sigma \to I \smallsetminus \mathbb{Q}$ by

$$\pi((a_i)_{i\in\mathbb{N}}) := [a_1, a_2, \dots],$$

and note that

$$\pi \circ \sigma = G \circ \pi. \tag{3.12}$$

Let

$$\widehat{\pi} \colon \widehat{\Sigma} \to I$$

denote the continuous extension of π to $\widehat{\Sigma}$.

Remark 3.19. It is readily checked that $\widehat{\pi}$ satisfies

$$\widehat{\pi}(\widehat{A}) = \begin{cases} [\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_n, \dots] & \text{if } \widehat{A} \in \Sigma, \text{ i.e., } \iota(\widehat{A}) = +\infty, \\ 0 & \text{if } \widehat{A} \in \widehat{\Sigma} \smallsetminus \Sigma \text{ with } \iota(\widehat{A}) = 1, \\ \bigl[\widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_{\iota(\widehat{A})-1}\bigr] & \text{if } \widehat{A} \in \widehat{\Sigma} \smallsetminus \Sigma \text{ with } 2 \leqslant \iota(\widehat{A}) < +\infty, \end{cases} \tag{3.13}$$

where $\widehat{A} = (\widehat{a}_i)_{i\in\mathbb{N}} \in \widehat{\Sigma}$.

We will need the following lemma:

Lemma 3.20. *If $a, b \in \widehat{\mathbb{N}}$ with $a \neq b$, then $|x-y| < 2\rho(a,b)$ for all $x \in I_a$, $y \in I_b$.*

Proof. Without loss of generality, assume that $a < b$. If $b = \infty$, then $|x-y| = \frac{1}{x} \leqslant \frac{1}{a} < 2\rho(a,b)$.

If $b \neq \infty$, then $a+1 \leqslant b$. Since $x \in \left[\frac{1}{a+1}, \frac{1}{a}\right]$ and $y \in \left[\frac{1}{b+1}, \frac{1}{b}\right]$, we get $|x-y| = x - y \leqslant \frac{1}{a} - \frac{1}{b+1} < \frac{2}{a} - \frac{2}{b} = 2\rho(a,b)$, where the second inequality, which is equivalent to $\frac{1}{a} > \frac{2}{b} - \frac{1}{b+1} = \frac{b+2}{b(b+1)}$, follows from $\frac{1}{a} \geqslant \frac{1}{b-1}$ and $\frac{1}{b-1} > \frac{b+2}{b(b+1)}$. □

The following Lemma 3.21 allows us to abuse notation by identifying the two sets $\mathcal{M}(\Sigma, \sigma_\Sigma)$ and $\left\{\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right) : \mu(\Sigma) = 1\right\}$. Let $k\colon \Sigma \to \widehat{\Sigma}$ be the inclusion map and $k_*\colon \mathcal{M}(\Sigma, \sigma_\Sigma) \to \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$ the pushforward of k, i.e., for each Borel subset $W \subseteq \widehat{\Sigma}$ and $\mu \in \mathcal{M}(\Sigma, \sigma_\Sigma)$, define

$$k_*(\mu)(W) := \mu\left(k^{-1}(W)\right) = \mu(A \cap \Sigma). \tag{3.14}$$

Lemma 3.21. *The map k_* is a continuous bijection from $\mathcal{M}(\Sigma, \sigma_\Sigma)$ to $\left\{\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right) : \mu(\Sigma) = 1\right\}$.*

Proof. Since k is continuous, k_* is well-defined and continuous. If $\mu_1, \mu_2 \in \mathcal{M}(\Sigma, \sigma_\Sigma)$ are such that $k_*(\mu_1) = k_*(\mu_2)$, then by (3.14), $\mu_1(W \cap \Sigma) = \mu_2(W \cap \Sigma)$ for every Borel subset $W \subseteq \widehat{\Sigma}$, so $\mu_1 = \mu_2$, therefore k_* is injective.

Fix $\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$ such that $\mu(\Sigma) = 1$. So $\mu(\mathcal{C}(\infty)) = 0$. Define $\nu \in \mathcal{P}(\Sigma)$ by $\nu(W) := \mu(W)$ for each Borel subset $W \subseteq \Sigma$. By definition of ν, since $\mu(\mathcal{C}(\infty)) = 0$ and $\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$, for each Borel subset $W \subseteq \Sigma$ we see that

$$\nu\left(\sigma_\Sigma^{-1}(W)\right) = \nu\left(\sigma_{\widehat{\Sigma}}^{-1}(W) \smallsetminus \mathcal{C}(\infty)\right) = \mu\left(\sigma_{\widehat{\Sigma}}^{-1}(W) \smallsetminus \mathcal{C}(\infty)\right) = \mu\left(\sigma_{\widehat{\Sigma}}^{-1}(W)\right) = \mu(W) = \nu(W),$$

hence $\nu \in \mathcal{M}(\Sigma, \sigma_\Sigma)$. For each Borel subset $V \subseteq \widehat{\Sigma}$, by (3.14), we have $k_*(\nu)(V) = \nu(V \cap \Sigma) = \mu(V \cap \Sigma) = \mu(V)$. So $k_*(\nu) = \mu$. Therefore, k_* is a bijection from $\mathcal{M}(\Sigma, \sigma_\Sigma)$ to $\left\{\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right) : \mu(\Sigma) = 1\right\}$, as required. □

The following lemma collects some basic properties of $\left(\widehat{\Sigma}, d_{\widehat{\rho}}\right)$.

Lemma 3.22. *The following statements are true:*

(i) *The map $\sigma\colon \left(\widehat{\Sigma}, d_{\widehat{\rho}}\right) \to \left(\widehat{\Sigma}, d_{\widehat{\rho}}\right)$ is Lipschitz.*

(ii) *$\mathcal{M}(\Sigma, \sigma_\Sigma)$ is a dense subset of $\mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$, and $\mathcal{M}_{\mathrm{erg}}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right) \smallsetminus \mathcal{M}(\Sigma, \sigma_\Sigma)$ is a dense subset of $\mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$.*

(iii) *If $\mu \in \mathcal{M}_{\mathrm{erg}}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$, then either $\mu(\Sigma) = 1$ or $\mu(\Sigma) = 0$.*

(iv) *$\mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right)$ is equal to the convex hull of $\mathcal{M}(\Sigma, \sigma_\Sigma) \cup \left\{\mu \in \mathcal{M}\left(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\right) : \mu(\Sigma) = 0\right\}$.*

Proof. Suppose $\widehat{A} = (\widehat{a}_n)_{n\in\mathbb{N}} \in \widehat{\Sigma}$ and $\widehat{B} = \left(\widehat{b}_n\right)_{n\in\mathbb{N}} \in \widehat{\Sigma}$.

(i) It is immediate from (3.11) that

$$d_{\widehat{\rho}}\left(\widehat{A}, \widehat{B}\right) = \frac{\widehat{\rho}(\widehat{a}_1, \widehat{b}_1)}{\theta^2} + \sum_{n=2}^{+\infty} \frac{\widehat{\rho}(\widehat{a}_n, \widehat{b}_n)}{\theta^{2n}} = \frac{\widehat{\rho}(\widehat{a}_1, \widehat{b}_1)}{\theta^2} + \frac{d_{\widehat{\rho}}(\sigma(\widehat{A}), \sigma(\widehat{B}))}{\theta^2} \geqslant \frac{d_{\widehat{\rho}}(\sigma(\widehat{A}), \sigma(\widehat{B}))}{\theta^2},$$

so $d_{\widehat{\rho}}\left(\sigma\left(\widehat{A}\right), \sigma\left(\widehat{B}\right)\right) \leqslant \theta^2 d_{\widehat{\rho}}\left(\widehat{A}, \widehat{B}\right)$, and statement (i) follows.

(ii) This follows from [IV25, Theorem 1.1].

(iii) Clearly, $\Sigma \triangle \sigma^{-1}(\Sigma) = \sigma^{-1}(\Sigma) \cap \mathcal{C}(\infty) = \sigma^{-1}(\Sigma) \smallsetminus \Sigma$. So $\mu\left(\Sigma \triangle \sigma^{-1}(\Sigma)\right) = \mu\left(\sigma^{-1}(\Sigma)\right) - \mu(\Sigma) = 0$. But μ is ergodic, so statement (ii) follows (cf. [Wa82, Theorem 1.5 (ii)]).

(iv) Assume that $\mu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big)$. By the ergodic decomposition theorem (see e.g. [EW11, Theorem 4.8]), writing $\widehat{\mathcal{M}}_1 \coloneqq \mathcal{M}_{\mathrm{erg}}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) \cap \mathcal{M}(\Sigma, \sigma_\Sigma)$ and $\widehat{\mathcal{M}}_2 \coloneqq \mathcal{M}_{\mathrm{erg}}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) \smallsetminus \mathcal{M}(\Sigma, \sigma_\Sigma)$,

$$\mu = \int_{\mathcal{M}_{\mathrm{erg}}(\widehat{\Sigma},\sigma)} m \, \mathrm{d}\alpha_\mu(m) = \int_{\widehat{\mathcal{M}}_1} m \, \mathrm{d}\alpha_\mu(m) + \int_{\widehat{\mathcal{M}}_2} m \, \mathrm{d}\alpha_\mu(m),$$

for some probability measure α_μ on $\mathcal{M}_{\mathrm{erg}}\big(\widehat{\Sigma}, \sigma\big)$. Define $p \coloneqq \mu(\Sigma)$, $\mu_1 \coloneqq \int_{\widehat{\mathcal{M}}_1} m \, \mathrm{d}\alpha_\mu(m)$, and $\mu_2 \coloneqq \int_{\widehat{\mathcal{M}}_2} m \, \mathrm{d}\alpha_\mu(m)$. Then by statement (ii), we conclude that

(a) when $p = 0$, $\mu = \mu_2 \in \big\{\mu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \mu(\Sigma) = 0\big\}$,
(b) when $p = 1$, $\mu = \mu_1 \in \mathcal{M}(\Sigma, \sigma_\Sigma)$,
(c) when $p \in (0, 1)$, $\mu = \mu_1 + \mu_2$, with $\frac{\mu_1}{p} \in \mathcal{M}(\Sigma, \sigma_\Sigma)$ and $\frac{\mu_2}{1-p} \in \mu \in \big\{\mu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \mu(\Sigma) = 0\big\}$.

Statement (iv) now follows. □

Remark 3.23. The map $\sigma \colon \big(\widehat{\Sigma}, d_{\widehat{\rho}}\big) \to \big(\widehat{\Sigma}, d_{\widehat{\rho}}\big)$ is not expansive: to see this note, for example, that if $A_n \coloneqq \overline{n}$ and $B \coloneqq \overline{\infty}$ for $n \in \mathbb{N}$, then $d_{\widehat{\rho}}\big(\sigma^k(A_n), \sigma^k(B)\big) = \sum_{i=1}^{+\infty} \frac{1}{n\theta^{2i}} = \frac{1}{(\theta^2-1)n}$ for each $k \in \mathbb{N}$.

3.5. **Bounded continued fractions.** Recall the following notion from Diophantine approximation (see e.g. [Sc80, Chapter 1]):

Definition 3.24 (Badly approximable number). An irrational number x is *badly approximable* if there is a constant $c = c(x) > 0$ such that $\big|x - \frac{p}{q}\big| > \frac{c}{q^2}$ for every rational number $\frac{p}{q}$.

An irrational number $x \in I$ is badly approximable if and only if the partial quotients in its continued fraction expansion are bounded (see e.g. [Sc80, Theorem 5F]).

Here we recall some properties for the set of bounded continued fractions. For a nonempty subset $A \subseteq \mathbb{N}$, let E_A denote the set of all irrational $x \in (0, 1)$ such that the digits $a_1(x)$, $a_2(x)$, ... in the continued fraction expansion $x = [a_1(x), a_2(x), a_3(x), \dots]$ all belong to A.

If $\operatorname{card} A < +\infty$, sets of the form E_A are said to be of bounded type and in particular they are Cantor sets. Of particular interest have been the sets $E_n \coloneqq E_{\{1, \dots, n\}}$. The restriction $G|_{E_n} : E_n \to E_n$ is conjugate to the one-sided full shift on a finite alphabet $\{1, 2, \dots, n\}$.

We recall the following notion (see e.g. [PU10, Chapter 4]):

Definition 3.25 (Distance-expanding map). For (X, d) a compact metric space, $T \colon X \to X$ is called a *distance-expanding map* if there exist constants $\lambda > 1$ and $\eta > 0$ such that for all $x, y \in X$ with $d(x, y) < 2\eta$,

$$d(T(x), T(y)) \geqslant \lambda d(x, y).$$

Definition 3.26. Suppose $n \in \mathbb{N}$. For $\psi \in C(I)$, define the corresponding *restricted ergodic supremum* $Q_n(G, \psi)$ by

$$Q_n(G, \psi) \coloneqq Q(G|_{E_n}, \psi|_{E_n}) = \sup\{\langle \mu, \psi \rangle : \mu \in \mathcal{M}(I, G), \operatorname{supp} \mu \subseteq E_n\}.$$

The following lemma collects some basic properties of the sets E_m.

Lemma 3.27. *Suppose $m \in \mathbb{N}$. If we write*

$$P \coloneqq \big\{G_{\widehat{a}}(0) : \widehat{a} \in \widehat{\mathbb{N}}\big\} = \{0,\, 1,\, 1/2,\, \dots,\, 1/n,\, \dots\},$$
$$\eta_m \coloneqq (m+2)^{-3}/2 \in (0, 1), \quad \lambda_m \coloneqq (1 - \eta_m)^{-2} > 1,$$

and denote the closed η_m-neighbourhood of E_m by

$$F_m \coloneqq \overline{B}_d^{\eta_m}(E_m) = \{x \in I : d(x, E_m) \leqslant \eta_m\}, \tag{3.15}$$

then the following hold:

(i) $d(E_m, P) > 2\eta_m$.
(ii) $G|_{F_m}$ *is Lipschitz, and in particular* $G|_{E_m}$ *is Lipschitz.*
(iii) *If* $x, y \in F_m$ *with* $|x-y| < \eta_m$*, then* $|G(x)-G(y)| \geqslant \lambda_m|x-y|$*, so* $G|_{E_m}$ *is distance-expanding.*
(iv) $G|_{E_m}$ *is an open map.*

Proof. (i) By Proposition 3.13 (iii),

$$\min E_m = [\overline{m,1}] > [m,1] = 1/(m+1) \text{ and } \max E_m = [\overline{1,m}] < [1,m+1] = (m+1)/(m+2). \quad (3.16)$$

But for each $1 \leqslant k \leqslant m$, the map G_k is strictly decreasing (see Proposition 3.13 (iii)), so (3.16) yields

$$G_k(E_m) \subseteq (G_k((m+1)/(m+2)), G_k(1/(m+1))) \subseteq (1/(k+1), 1/k),$$

and hence

$$\begin{aligned} d(P, G_k(E_m)) &> \min\{G_k((m+1)/(m+2)) - G_k(1),\, G_k(0) - G_k(1/(m+1))\} \\ &= \min\{|G_k'(z_1)|/(m+2),\, |G_k'(z_2)|/(m+1)\} \end{aligned}$$

for some $z_1 \in ((m+1)/(m+2), 1)$ and $z_2 \in (0, 1/(m+1))$ by the intermediate value theorem. Since $|G_k'(x)| = (k+x)^{-2} \geqslant (m+1)^{-2}$ for all $x \in I$, we deduce that

$$d(P, G_k(E_m)) > \min\big\{(m+1)^{-2}(m+2)^{-1},\, (m+1)^{-3}\big\} > 2\eta_m.$$

Now $E_m = \bigcup_{k=1}^{m} G_k(E_m)$ (an immediate consequence of the definition of E_m), so (i) follows.

(ii) Assume that $x, y \in F_m$. When $|x-y| < \eta_m$, by (i), we get $x, y \in (1/(k+1), 1/k)$ for some $1 \leqslant k \leqslant m$. Then

$$|G(x) - G(y)| = |k + 1/x - k - 1/y| = z^{-2}|x-y| \leqslant (m+1)^2|x-y|,$$

for some z between x and y by the intermediate value theorem, where the last inequality follows from $z \geqslant 1/(m+1)$. When $|x-y| \geqslant \eta_m$, we have $|G(x)-G(y)| < 1 < \frac{1}{\eta_m}|x-y|$, so (ii) follows.

(iii) Assume that $x, y \in F_m$ with $|x-y| < \eta_m$. By (i), we get $x, y \in (1/(k+1), 1/k)$ for some $1 \leqslant k \leqslant m$, and then

$$|G(x) - G(y)| = |k + 1/x - k - 1/y| = w^{-2}|x-y| \geqslant \lambda_m|x-y|$$

for some w in between x and y by the intermediate value theorem, where the inequality follows from the fact that $w < 1 - \eta_m$.

(iv) Evidently $G|_{E_m} = (\pi|_{\Sigma_m})^{-1} \circ \sigma_{\Sigma_m} \circ \pi|_{\Sigma_m}$, so the result follows from the fact that $\pi|_{\Sigma_m}$ is a homeomorphism and σ_{Σ_m} is open. □

4. Maximizing measures

Here we introduce the notion of *limit-maximizing measure*, which will be useful for a dynamical system, such as G, whose set of invariant measures is not weak* compact.

Notation 4.1. The pushforward $\pi_* \colon \mathcal{M}(\Sigma, \sigma) \to \mathcal{M}(I \smallsetminus \mathbb{Q}, G|_{I \smallsetminus \mathbb{Q}})$ of π is defined by

$$\pi_*(\mu)(V) := \mu\big(\pi^{-1}(V)\big),$$

for all Borel subsets $V \subseteq I \smallsetminus \mathbb{Q}$. Similarly, the pushforward $\widehat{\pi}_* \colon \mathcal{M}\big(\widehat{\Sigma}, \sigma\big) \to \mathcal{P}(I)$ is defined by

$$\widehat{\pi}_*(\mu)(W) := \mu\big(\widehat{\pi}^{-1}(W)\big)$$

for all Borel subsets $W \subseteq I$.

Remark 4.2. Since $\pi = \widehat{\pi} \circ k$, Lemma 3.21 allows us to abuse notation by writing

$$\widehat{\pi}_*|_{\mathcal{M}(\Sigma, \sigma_\Sigma)} = \pi_*.$$

Definition 4.3. Let $T\colon X \to X$ be a Borel measurable map on a compact metric space X. For a Borel measurable function $\psi\colon X \to \mathbb{R}$, a probability measure μ is called a (T,ψ)*-limit-maximizing measure*, or simply a ψ*-limit-maximizing measure*, if it is a weak* accumulation point of $\mathcal{M}(X,T)$ and $\int \psi \,\mathrm{d}\mu = Q(T,\psi)$. We denote the set of (T,ψ)-limit-maximizing measures by $\mathcal{M}^*_{\max}(T,\psi)$.

Clearly, $\mathcal{M}_{\max}(T,\psi) \subseteq \mathcal{M}^*_{\max}(T,\psi)$. The following lemma collects some basic properties of π and $\widehat{\pi}$.

Lemma 4.4. *If $n \in \mathbb{N}$, then the following hold:*

(i) $\widehat{\pi}^{-1}(0) = \mathcal{C}(\infty)$, $\widehat{\pi}^{-1}(1) = \mathcal{C}(1,\infty)$, *and for each* $x = [a_1,\dots,a_n] \in R_n$,
$$\widehat{\pi}^{-1}(x) = \mathcal{C}(a_1,\dots,a_n,\infty) \cup \mathcal{C}(a_1,\dots,a_{n-1},a_n - 1, 1, \infty).$$

(ii) *If* $\widehat{\mathbf{a}} = (\widehat{a}_1,\widehat{a}_2,\dots,\widehat{a}_n) \in \widehat{\mathbb{N}}^n$, *then the equality* $G_{\widehat{\mathbf{a}}} \circ \widehat{\pi} \circ \sigma^n = \widehat{\pi}$ *holds on* $\mathcal{C}(\widehat{a}_1,\dots,\widehat{a}_n)$.

(iii) *The map* $\widehat{\pi}\colon \big(\widehat{\Sigma}, d_{\widehat{\rho}}\big) \to (I,d)$ *is Lipschitz.*

(iv) $\pi_*\colon \mathcal{M}(\Sigma,\sigma) \to \mathcal{M}_{\mathrm{irr}}(I,G)$ *is a homeomorphism.*

(v) $\widehat{\pi}_*\colon \mathcal{M}\big(\widehat{\Sigma},\sigma\big) \to \overline{\mathcal{M}(I,G)}$ *is a continuous surjection.*

(vi) *If* $\widehat{\mathbf{a}} = (\widehat{a}_1,\dots,\widehat{a}_n) \in \widehat{\mathbb{N}}^n$, $\phi \in C(I)$, $\widehat{A} = (\widehat{a}_i)_{i\in\mathbb{N}} \in \mathcal{C}(\widehat{a}_1,\dots,\widehat{a}_n)$, *and* $y_n \coloneqq \widehat{\pi}\big(\sigma^n\big(\widehat{A}\big)\big)$, *then* $S_{n,\widehat{\mathbf{a}}}\phi(y_n) = S^\sigma_n(\phi\circ\widehat{\pi})\big(\widehat{A}\big)$.

Proof. (i) $\widehat{\pi}^{-1}(0) = \mathcal{C}(\infty)$ and $\widehat{\pi}^{-1}(1) = \mathcal{C}(1,\infty)$ follow from (3.13), while the equality
$$\widehat{\pi}^{-1}(x) = \mathcal{C}(a_1,\dots,a_n,\infty) \cup \mathcal{C}(a_1,\dots,a_{n-1},a_n - 1,1,\infty)$$
follows from (3.13) and Lemma 3.2.

(ii) Fix $\widehat{\mathbf{a}} = (\widehat{a}_1,\widehat{a}_2,\dots,\widehat{a}_n) \in \widehat{\mathbb{N}}^n$. Consider an arbitrary $\widehat{A} = \widehat{a}_1\widehat{a}_2\dots\widehat{a}_n\widehat{a}_{n+1}\dots \in \mathcal{C}[\widehat{a}_1,\widehat{a}_2,\dots,\widehat{a}_n]$. Write $k \coloneqq \iota\big(\widehat{A}\big)$. If $\widehat{A}$ belongs to Σ, it follows immediately from the definitions of π, $G_{\widehat{\mathbf{a}}}$, and σ that
$$G_{\widehat{\mathbf{a}}}\big(\pi\big(\sigma^n\big(\widehat{A}\big)\big)\big) = G_{\widehat{\mathbf{a}}}([\widehat{a}_{n+1},\widehat{a}_{n+2},\dots]) = [\widehat{a}_1,\widehat{a}_2,\dots] = \widehat{\pi}\big(\widehat{A}\big).$$
If on the other hand $\widehat{A} \in \widehat{\Sigma} \smallsetminus \Sigma$, in the case that $k = 1$, then by Proposition 3.13 (ii),
$$G_{\widehat{\mathbf{a}}}\big(\widehat{\pi}\big(\sigma^n\big(\widehat{A}\big)\big)\big) = 0 = \widehat{\pi}\big(\widehat{A}\big),$$
while in the case that $2 \leqslant k \leqslant n$, then by Proposition 3.13 (ii) and (3.13),
$$G_{\widehat{\mathbf{a}}}\big(\widehat{\pi}\big(\sigma^n\big(\widehat{A}\big)\big)\big) = [\widehat{a}_1,\widehat{a}_2,\dots,\widehat{a}_{k-1}] = \widehat{\pi}\big(\widehat{A}\big).$$
In the case that $k \geqslant n+1$, we obtain $\widehat{\pi}\big(\sigma^n\big(\widehat{A}\big)\big) = [\widehat{a}_n,\dots,\widehat{a}_{k-1}]$ and
$$G_{\widehat{\mathbf{a}}}\big(\widehat{\pi}\big(\sigma^n\big(\widehat{A}\big)\big)\big) = [\widehat{a}_1,\widehat{a}_2,\dots,\widehat{a}_n,\dots,\widehat{a}_{k-1}] = \widehat{\pi}\big(\widehat{A}\big).$$

(iii) Assume that $\widehat{B} = \big(\widehat{b}_i\big)_{i\in\mathbb{N}}$, $\widehat{C} = (\widehat{c}_i)_{i\in\mathbb{N}} \in \widehat{\Sigma}$. Let k be the smallest integer such that $\widehat{b}_k \neq \widehat{c}_k$. If $k = 1$, since $\widehat{\pi}\big(\widehat{B}\big) \in I_{\widehat{b}_1}$, $\widehat{\pi}\big(\widehat{C}\big) \in I_{\widehat{c}_1}$, Lemma 3.20, together with the definition of $d_{\widehat{\rho}}$ (cf. (3.11)), implies that
$$\big|\widehat{\pi}\big(\widehat{B}\big) - \widehat{\pi}\big(\widehat{C}\big)\big| \leqslant 2\rho\big(\widehat{b}_1,\widehat{c}_1\big) \leqslant 2\theta^2 d_{\widehat{\rho}}\big(\widehat{B},\widehat{C}\big).$$
If $k \geqslant 2$, the definition of $\widehat{\pi}$ guarantees that $\widehat{\pi}\big(\widehat{B}\big), \widehat{\pi}\big(\widehat{C}\big) \in I_{\widehat{\mathbf{a}}}$ for some $\widehat{\mathbf{a}} \coloneqq \widehat{a}_1\widehat{a}_2\dots\widehat{a}_{k-1} \in \widehat{\mathbb{N}}^{k-1}$. Combining statement (ii), Proposition 3.13 (i), Lemma 3.20, (3.11), and the fact that $\widehat{\pi}\big(\sigma^{k-1}\big(\widehat{B}\big)\big) \in I_{\widehat{b}_k}$ and $\widehat{\pi}\big(\sigma^{k-1}\big(\widehat{C}\big)\big) \in I_{\widehat{c}_k}$, we have
$$\begin{aligned}\big|\widehat{\pi}\big(\widehat{B}\big) - \widehat{\pi}\big(\widehat{C}\big)\big| &= \big|G_{\widehat{\mathbf{a}}}\big(\widehat{\pi}\big(\sigma^{k-1}\big(\widehat{B}\big)\big)\big) - G_{\widehat{\mathbf{a}}}\big(\widehat{\pi}\big(\sigma^{k-1}\big(\widehat{C}\big)\big)\big)\big| \\ &\leqslant \frac{|\widehat{\pi}(\sigma^{k-1}(\widehat{B})) - \widehat{\pi}(\sigma^{k-1}(\widehat{C}))|}{c_0^2\theta^{2(k-1)}} \leqslant \frac{2\rho\big(\widehat{b}_k,\widehat{c}_k\big)}{c_0^2\theta^{2(k-1)}} \leqslant \frac{2\theta^2 d_\rho(A,B)}{c_0^2},\end{aligned}$$

so statement (iii) follows.

(iv) Since $\pi\colon \Sigma \to I \smallsetminus \mathbb{Q}$ is a homeomorphism (cf. [Mi17, Theorem 1.1]), the relation (3.12) implies that $\sigma\colon \Sigma \to \Sigma$ and $G\colon I \smallsetminus \mathbb{Q} \to I \smallsetminus \mathbb{Q}$ are topologically conjugate, so their spaces of invariant probability measures are homeomorphic under π_*.

(v) Fix $\mu \in \overline{\mathcal{M}(I,G)}$. By Lemma 3.7, there exists $\{\mu_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_{\mathrm{irr}}(I,G)$ such that μ_n converges to μ in the weak* topology as $n \to +\infty$. Let us denote $\nu_n \coloneqq \pi_*^{-1}(\mu_n)$ for each $n \in \mathbb{N}$. By the weak* compactness of $\mathcal{M}\bigl(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\bigr)$, there exists an accumulation point ν of $\{\nu_n\}_{n\in\mathbb{N}}$. Since $\widehat{\pi}_*$ is continuous (see statement (iii)), $\widehat{\pi}_*(\nu) = \mu$. Consequently $\widehat{\pi}_*$ is surjective.

(vi) From the definition of $S_{n,\widehat{\mathbf{a}}}$ (see (3.10)), it suffices to note that $G_{\sigma^i(\widehat{\mathbf{a}})}(y_n) = \widehat{\pi}\bigl(\sigma^i(A)\bigr)$ for each $0 \leqslant i \leqslant n-1$, by statement (ii). □

Remark 4.5.
(i) The continuous map $\pi^{-1}\colon I \smallsetminus \mathbb{Q} \to \Sigma$ is not Lipschitz: for example if $x_n \coloneqq [3, \overline{n}]$ and $y_n \coloneqq [2,1,\overline{n}]$, then $\lim_{n\to+\infty} |x_n - y_n| = 0$ but $d_{\widehat{\rho}}\bigl(\pi^{-1}(x_n), \pi^{-1}(y_n)\bigr) \geqslant \theta^{-2}\widehat{\rho}(3,2) = \frac{1}{6\theta^2}$.

(ii) The pushforward $\widehat{\pi}_*$ of the extension $\widehat{\pi}$ is not injective: for example let μ_1 be the periodic measure supported on the periodic orbit of $\overline{1\infty2\infty}$, let μ_2 be the periodic measure supported on the periodic orbit of $\overline{1\infty}$, and let μ_3 be the periodic measure supported on the periodic orbit of $\overline{2\infty}$. Then by the definition of $\widehat{\pi}_*$, we get $\widehat{\pi}_*(\mu_1) = \frac{1}{4}\bigl(2\delta_0 + \delta_1 + \delta_{1/2}\bigr)$, $\widehat{\pi}_*(\mu_2) = \frac{1}{2}(\delta_0 + \delta_1)$, and $\widehat{\pi}_*(\mu_3) = \frac{1}{2}\bigl(\delta_0 + \delta_{1/2}\bigr)$. Clearly, $\widehat{\pi}_*(\mu_1) = \widehat{\pi}_*\bigl(\frac{1}{2}(\mu_2+\mu_3)\bigr)$ but $\mu_1 \neq \frac{1}{2}(\mu_2+\mu_3)$.

(iii) The dynamics of G and σ are not intertwined by $\widehat{\pi}$, in other words, $\widehat{\pi} \circ \sigma \neq G \circ \widehat{\pi}$: to see this note, for example, that if $A = \infty\overline{1}$ then $\widehat{\pi}(\sigma(A)) = [\overline{1}]$, but $G(\widehat{\pi}(A)) = 0$.

Proposition 4.6. *If $\phi \in C(I)$, then*

(i) $Q\bigl(\sigma_{\widehat{\Sigma}}, \phi\circ\widehat{\pi}\bigr) = Q(G,\phi)$ *and*

(ii) $\widehat{\pi}_*\bigl(\mathcal{M}_{\max}\bigl(\sigma_{\widehat{\Sigma}}, \phi\circ\widehat{\pi}\bigr)\bigr) = \mathcal{M}^*_{\max}(G,\phi) \neq \emptyset$.

Proof. (i) From the definition of $Q\bigl(\sigma_{\widehat{\Sigma}}, \phi\circ\widehat{\pi}\bigr)$, and the fact that $\mathcal{M}(\Sigma,\sigma_\Sigma)$ is dense in $\mathcal{M}\bigl(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\bigr)$, we see that

$$Q\bigl(\sigma_{\widehat{\Sigma}}, \phi\circ\widehat{\pi}\bigr) = \sup\{\langle \mu, \phi\circ\widehat{\pi}\rangle : \mu \in \mathcal{M}(\Sigma,\sigma_\Sigma)\} = \sup\{\langle \mu, \phi\circ\pi\rangle : \mu \in \mathcal{M}(\Sigma,\sigma_\Sigma)\}.$$

Since $\mathcal{M}_{\mathrm{irr}}(I,G)$ is dense in $\mathcal{M}(I,G)$ (see Lemma 3.7), we have

$$Q(G,\phi) = \sup\{\langle \nu, \phi\ \rangle : \nu \in \mathcal{M}_{\mathrm{irr}}(I,G)\}.$$

Combining the above two identities with Lemma 4.4 (iv), (i) follows.

(ii) The first identity comes from the fact that $\widehat{\pi}_*$ is a surjection from $\mathcal{M}\bigl(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\bigr)$ to $\overline{\mathcal{M}(I,G)}$ (cf. Lemma 4.4 (v)), and $\mathcal{M}^*_{\max}(G,\phi) \neq \emptyset$ follows from the fact that $\overline{\mathcal{M}(I,G)}$ is compact with respect to the weak* topology, and the assumption that $\phi \in C(I)$. □

Proposition 4.7. *If $\phi \in C(I)$, then*

(i) $\lim_{m\to+\infty} Q_m(G,\phi) = Q(G,\phi)$ *and*

(ii) $Q(G,\phi) = \sup\bigl\{\liminf_{n\to+\infty} \frac{S_n\phi(x)}{n} : x \in I \smallsetminus \mathbb{Q}\bigr\}$.

Proof. (i) The set of periodic measures is known to be dense in $\mathcal{M}(\Sigma,\sigma)$ (see [IV21, Theorem 3.8]), and evidently π^* gives a one-to-one correspondence between the set of periodic measures in $\mathcal{M}(\Sigma,\sigma)$ and the set of periodic measures in $\mathcal{M}_{\mathrm{irr}}(I,G)$. Combining this with Lemma 4.4 (iv), it follows that the set of periodic measures is dense in $\mathcal{M}_{\mathrm{irr}}(I,G)$, so for any $\epsilon > 0$ there is a periodic measure $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$ with $\int \phi\,\mathrm{d}\mu \geqslant Q(G,\phi) - \epsilon$. Since μ is periodic, there exists $m \in \mathbb{N}$ such that $\mu \in \mathcal{M}(E_m, G|_{E_m})$, and clearly $Q_m(G,\phi) \geqslant \int\phi\,\mathrm{d}\mu \geqslant Q(G,\phi) - \epsilon$. But $\epsilon > 0$ was arbitrary, and $\{Q_m(G,\phi)\}_{m\in\mathbb{N}}$ is nondecreasing and bounded above by $Q(G,\phi)$, so (i) follows.

(ii) Now $Q(\sigma_{\widehat{\Sigma}}, \phi \circ \widehat{\pi}) = Q(G, \phi)$ by Proposition 4.6 (i), the space $(\widehat{\Sigma}, d_{\widehat{\rho}})$ is compact, $\sigma : \widehat{\Sigma} \to \widehat{\Sigma}$ is continuous, and $\pi \colon \Sigma \to I \smallsetminus \mathbb{Q}$ is a homeomorphism (cf. [Mi17, Theorem 1.1]), so [Je19, Proposition 2.2] gives

$$\begin{aligned} Q(G,\phi) = Q(\sigma_{\widehat{\Sigma}}, \phi \circ \widehat{\pi}) &= \sup\bigl\{\liminf_{n\to+\infty} n^{-1} S_n^{\sigma}(\phi\circ\widehat{\pi})(A) : A \in \widehat{\Sigma}\bigr\} \\ &\geqslant \sup\bigl\{\liminf_{n\to+\infty} n^{-1} S_n^{\sigma}(\phi\circ\pi)(A) : A \in \Sigma\bigr\} = \sup\bigl\{\liminf_{n\to+\infty} n^{-1} S_n\phi(x) : x \in I \smallsetminus \mathbb{Q}\bigr\}. \end{aligned}$$

For each $m \in \mathbb{N}$, the restriction $G|_{E_m}$ is a continuous map on the compact metric space E_m, so (i) and [Je19, Proposition 2.2] together give

$$Q(G,\phi) \leqslant \sup_{x \in \bigcup_{m\in\mathbb{N}} E_m} \liminf_{n\to+\infty} \frac{S_n\phi(x)}{n} \leqslant \sup_{x\in I\smallsetminus\mathbb{Q}} \liminf_{n\to+\infty} \frac{S_n\phi(x)}{n},$$

and (ii) follows. □

5. Structure of the closure of $\mathcal{M}(I,G)$

Definition 5.1 (Finite-continued-fraction measures and rational orbits). Fix $n \in \mathbb{N}$ and $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}^n$. Let us denote

$$\begin{aligned} &p_0 = p_0(\mathbf{a}) \coloneqq 0 \quad \text{ and} \\ &p_k = p_k(\mathbf{a}) \coloneqq [a_{n-k+1}, \dots, a_n] \quad \text{ for each } 1 \leqslant k \leqslant n. \end{aligned} \tag{5.1}$$

We define the corresponding *rational orbit* of length

$$l_{\mathbf{a}} \coloneqq n+1 \tag{5.2}$$

to be

$$\mathcal{O}_{\mathbf{a}} \coloneqq \{p_0(\mathbf{a}),\, p_1(\mathbf{a}),\, p_2(\mathbf{a}),\, \dots,\, p_n(\mathbf{a})\} = \{0,\, p_1,\, p_2,\, \dots,\, p_n\}. \tag{5.3}$$

Note that each rational orbit $\mathcal{O}_{\mathbf{a}}$ contains only rational numbers.

The corresponding *finite-continued-fraction measure* (abbreviated as *FCF measure*) $\mu_{\mathbf{a}}$ (of *length* $l_{\mathbf{a}}$) is defined by

$$\mu_{\mathbf{a}} \coloneqq \frac{1}{l_{\mathbf{a}}}\bigl(\delta_{p_0(\mathbf{a})} + \delta_{p_1(\mathbf{a})} + \cdots + \delta_{p_n(\mathbf{a})}\bigr) = \frac{1}{n+1}\bigl(\delta_0 + \delta_{p_1} + \cdots + \delta_{p_n}\bigr). \tag{5.4}$$

Note that $\mu_{\mathbf{a}}$ is a probability measure, but is never G-invariant.

Define $\mathcal{F}$ to be the convex hull of $\{\delta_0\} \cup \{\mu_{\mathbf{a}} : n \in \mathbb{N},\, \mathbf{a} \in \mathbb{N}^n\}$. Define

$$\mathcal{F}_{[0,1)} \coloneqq \{\mu \in \mathcal{F} : \mu(\{1\}) = 0\} \subseteq \mathcal{F}. \tag{5.5}$$

Remark 5.2. Fix $n \in \mathbb{N}$ and $\mathbf{a} \in \mathbb{N}^n$. When $\mathbf{a} \in \mathcal{A}^n$, the rational orbit $\mathcal{O}_{\mathbf{a}} = \mathcal{O}(p_n(\mathbf{a}))$ is a G-orbit that is eventually fixed, in the sense that $G(p_0(\mathbf{a})) = p_0(\mathbf{a})$.

When $\mathbf{a} \in \mathcal{B}^n$, the rational orbit $\mathcal{O}_{\mathbf{a}} = \mathcal{O}(p_n(\mathbf{a})) \cup \{1\}$ is not a G-orbit, but is an orbit under the map that is equal to 1 on R_1, and equal to G elsewhere.

Lemma 5.3. *Every FCF measure is the limit of a sequence of periodic measures. Moreover,*

$$\mathcal{F} \subseteq \overline{\mathcal{M}(I,G)} = \overline{\mathcal{M}_{\mathrm{irr}}(I,G)}. \tag{5.6}$$

Proof. Fix $n \in \mathbb{N}$ and $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}^n$. For each $m \in \mathbb{N}$, define

$$r_m \coloneqq [\overline{a_1, \dots, a_n, m}] \in \operatorname{Fix}\bigl(G^{n+1}\bigr). \tag{5.7}$$

Evidently $\lim_{m\to+\infty} G^{n-k}(r_m) = p_k$ for each $0 \leqslant k \leqslant n$, hence the sequence $\{\mu_{\mathcal{O}(r_m)}\}_{m\in\mathbb{N}}$ of periodic measures converges to the FCF measure $\mu_{\mathbf{a}}$, as required.

The sequence $\{\mu_{\mathcal{O}(r_m)}\}_{m\in\mathbb{N}}$ is contained in $\mathcal{M}(I,G)$, so $\mu_{\mathbf{a}} \in \overline{\mathcal{M}(I,G)}$. But $\overline{\mathcal{M}(I,G)}$ is convex, and $\mathbf{a} \in \mathbb{N}^n$ was arbitrary, so $\mathcal{F} \subseteq \overline{\mathcal{M}(I,G)}$. By Lemma 3.7, we get $\overline{\mathcal{M}(I,G)} = \overline{\mathcal{M}_{\mathrm{irr}}(I,G)}$. Therefore, we obtain (5.6). □

The sets R_n, defined in Notation 3.3, have the following simple properties:

Lemma 5.4. $\bigcup_{n\in\mathbb{N}} R_n = (0,1)\cap\mathbb{Q}$, *and if* $m\neq n$ *then* $R_n\cap R_m=\emptyset$.

Proof. By (3.3), $\bigcup_{n\in\mathbb{N}} R_n \subseteq (0,1)\cap\mathbb{Q}$. By Lemma 3.2, $(0,1)\cap\mathbb{Q}\subseteq\bigcup_{n\in\mathbb{N}} R_n$ and $R_n\cap R_m=\emptyset$ when $n\neq m$. □

The following lemma collects some basic properties of measures in $\mathcal{F}$.

Lemma 5.5. *Suppose* $\mu\in\mathcal{F}$, $n\in\mathbb{N}$, $\mathbf{a}=(a_1,\dots,a_n)\in\mathcal{A}_n$, *and* $\mathbf{b}=(b_1,\dots,b_n)\in\mathcal{B}_n$.

(i) *Then* $\mu(\{1\})\leqslant\frac{1}{2}$.

(ii) *If* $\mu\in\mathcal{F}_{[0,1)}$, *then* $\mu(R_n)\leqslant\frac{1}{n+1}$.

(iii) $(l_{\mathbf{a}}+1)\mu_{f(\mathbf{a})}=l_{\mathbf{a}}\mu_{\mathbf{a}}+\delta_1$.

(iv) *The map* $g_n\colon\mathcal{A}_n\to R_n$ *is bijective.*

(v) *For every* $y\in I$, *we have* $\mu_{\mathbf{a}}(\{y\})=1/l_{\mathbf{a}}$ *if* $y\in\mathcal{O}(g_n(\mathbf{a}))$, *and* $\mu_{\mathbf{a}}(\{y\})=0$ *otherwise.*

(vi) $\mu_{\mathbf{a}}(\{1\})=0$ *and* $\mu_{\mathbf{a}}(\{0\})=\mu_{\mathbf{a}}(R_1)$.

(vii) *For all* $x\in(0,1)\cap\mathbb{Q}$, $\mu_{\mathbf{a}}(\{x\})\geqslant\mu_{\mathbf{a}}\big(G^{-1}(x)\big)$.

(viii) $\mu_{\mathbf{b}}(\{1\})=1/l_{\mathbf{b}}$ *and* $\mu_{\mathbf{b}}(\{0\})\geqslant\mu_{\mathbf{b}}(R_1)$.

(ix) *For all* $x\in(0,1)\cap\mathbb{Q}$, $\mu_{\mathbf{b}}(\{x\})\geqslant\mu_{\mathbf{b}}\big(G^{-1}(x)\big)$.

Proof. (i) follows immediately from the fact that $\mu_{\mathbf{b}}(\{1\})\leqslant\frac{1}{2}$ for all $\mathbf{b}\in\mathbb{N}^*$ (see (5.4)), the fact that $\delta_0(\{1\})=0$, and the definition of $\mathcal{F}$ (cf. Definition 5.1).

(ii) This follows from the fact that $\mu_{\mathbf{b}}(\{R_n\})\leqslant\frac{1}{n+1}$ for each $\mathbf{b}\in\mathcal{A}$ (see (5.4) and (3.3)), the fact that $\delta_0(\{R_n\})=0$, and the definition of $\mathcal{F}_{[0,1)}$ (see (5.5)).

(iii) Since $[a_{n-k+1},\dots,a_n]=[a_{n-k+1},\dots,a_n-1,1]$ for each $1\leqslant k\leqslant n$, using (5.4) and (5.1) we get $(l_{\mathbf{a}}+1)\mu_{f(\mathbf{a})}=\delta_0+\delta_1+\sum_{k=1}^{n}\delta_{p_k}=l_{\mathbf{a}}\mu_{\mathbf{a}}+\delta_1$.

(iv) By the definition of $\mathcal{A}_n$ (see (3.4)) and the definition of g_n (see (3.5)), we have $G^{n-1}(g_n(\mathbf{a}))=\frac{1}{a_n}\in R_1$. By Lemma 3.2, g_n is injective. By the definition of R_n (see (3.3)), g_n maps $\mathcal{A}_n$ surjectively to R_n.

(v) For $\mathbf{a}=(a_1,a_2,\dots,a_n)\in\mathcal{A}_n$, denote $x\coloneqq[a_1,a_2,\dots,a_n]=g_n(\mathbf{a})$. The definition of the rational orbit $\mathcal{O}_{\mathbf{a}}$ (cf. (5.3)) gives that $\mathcal{O}_{\mathbf{a}}=\mathcal{O}(x)$, and then (v) follows from the definition of $\mu_{\mathbf{a}}$ (cf. (5.4)).

(vi) We have $\mu_{\mathbf{a}}(\{1\})=0$ since the support of $\mu_{\mathbf{a}}$ is $\mathcal{O}_{\mathbf{a}}$ contained in $[0,1)\cap\mathbb{Q}$ (cf. (5.3)). The point 0 is an atom of $\mu_{\mathbf{a}}$, with $\mu_{\mathbf{a}}(\{0\})=1/l_{\mathbf{a}}$, and precisely one element of R_1, namely the point $1/a_n$, is an atom of $\mu_{\mathbf{a}}$, also with weight $\mu_{\mathbf{a}}(\{1/a_n\})=1/l_{\mathbf{a}}$, so in particular $\mu_{\mathbf{a}}(\{R_1\})=1/l_{\mathbf{a}}=\mu_{\mathbf{a}}(\{0\})$.

(vii) Note that the support of $\mu_{\mathbf{a}}$ is an eventually fixed G-orbit, so if x is not an atom of $\mu_{\mathbf{a}}$ then nor is any element of $G^{-1}(x)$, so $\mu(\{x\})=0=\mu\big(G^{-1}(x)\big)$. If x is an atom of $\mu_{\mathbf{a}}$ then $x=p_k(\mathbf{a})$ for some $0\leqslant k\leqslant n$: if $k=n$ then $G^{-1}(x)$ does not contain any atoms of $\mu_{\mathbf{a}}$, so $\mu_{\mathbf{a}}(\{x\})=1/l_{\mathbf{a}}>0=\mu_{\mathbf{a}}\big(G^{-1}(x)\big)$, while if $k<n$ then $G^{-1}(x)$ contains precisely one atom of $\mu_{\mathbf{a}}$, namely $p_{k+1}(\mathbf{a})$, so $\mu_{\mathbf{a}}(\{x\})=1/l_{\mathbf{a}}=\mu_{\mathbf{a}}\big(G^{-1}(x)\big)$, therefore in both cases we see that (vii) holds.

(viii) We have $\mu_{\mathbf{b}}(\{1\})=1/l_{\mathbf{b}}$ since the support of $\mu_{\mathbf{b}}$ is $\mathcal{O}_{\mathbf{b}}$ containing 1 (cf. (5.3)). When $n=1$, $\mu_{\mathbf{b}}=1/2(\delta_0+\delta_1)$ and $\mu_{\mathbf{b}}(\{0\})\geqslant0=\mu_{\mathbf{b}}(R_1)$. When $n\geqslant2$, note that $\mathcal{O}_{\mathbf{b}}=\mathcal{O}_{f^{-1}(\mathbf{b})}\cup\{1\}$ and $f^{-1}(\mathbf{b})\in\mathcal{A}^{n-1}$, so by (iii) and (vi) we get $\mu_{\mathbf{b}}(\{0\})\geqslant\mu_{\mathbf{b}}(R_1)$.

(ix) Note that the support of $\mu_{\mathbf{b}}$ is the union of an eventually fixed G-orbit and 1 (see Remark 5.2), so by the fact that $G^{-1}(1)=\emptyset$, if x is not an atom of $\mu_{\mathbf{b}}$ then nor is any element of $G^{-1}(x)$, so $\mu_{\mathbf{b}}(\{x\})=0=\mu_{\mathbf{b}}\big(G^{-1}(x)\big)$. If $x\in(0,1)$ is an atom of $\mu_{\mathbf{b}}$, then $x=p_k(\mathbf{b})$ for some $2\leqslant k\leqslant n$: if $k=n$ then $G^{-1}(x)$ does not contain any atoms of $\mu_{\mathbf{b}}$, so $\mu_{\mathbf{b}}(\{x\})=1/l_{\mathbf{b}}>0=\mu_{\mathbf{a}}\big(G^{-1}(x)\big)$, while if $k<n$ then $G^{-1}(x)$ contains precisely one atom of $\mu_{\mathbf{b}}$, namely $p_{k+1}(\mathbf{b})$, so $\mu_{\mathbf{b}}(\{x\})=1/l_{\mathbf{b}}=\mu_{\mathbf{b}}\big(G^{-1}(x)\big)$, therefore in both cases we see that (ix) holds. □

Lemma 5.6. *If $\nu \in \mathcal{F}_{[0,1)}$ and $r \in [0,1]$ satisfy $(1-r)\nu(\{0\}) \geqslant r$, then $(1-r)\nu + r\delta_1 \in \mathcal{F}$.*

Proof. Since $\nu \in \mathcal{F}_{[0,1)}$, we can write

$$\nu = r_0\delta_0 + \sum_{\mathbf{a}\in\mathcal{A}} r(\mathbf{a})\mu_{\mathbf{a}}, \tag{5.8}$$

where $r_0 \geqslant 0$, and $r(\mathbf{a}) \geqslant 0$ for each $\mathbf{a} \in \mathcal{A}$, and

$$r_0 + \sum_{\mathbf{a}\in\mathcal{A}} r(\mathbf{a}) = 1. \tag{5.9}$$

Combining (5.8), (5.2), and (5.4), we see that

$$\nu(\{0\}) = r_0 + \sum_{\mathbf{a}\in\mathcal{A}} r(\mathbf{a})\mu_{\mathbf{a}}(\{0\}) = r_0 + \sum_{\mathbf{a}\in\mathcal{A}} \frac{r(\mathbf{a})}{l_{\mathbf{a}}},$$

and combining this with the assumption that $(1-r)\nu(\{0\}) \geqslant r$ gives

$$r_0 + \sum_{\mathbf{a}\in\mathcal{A}} \frac{r(\mathbf{a})}{l_{\mathbf{a}}} \geqslant \frac{r}{1-r}. \tag{5.10}$$

Let us denote $\lambda \coloneqq \frac{r}{(1-r)\nu(\{0\})} \in [0,1]$. Then by (5.8), (5.10), and (5.9),

$$\begin{aligned}\lambda(1-r)\nu + r\delta_1 &= \frac{r}{\nu(\{0\})}\Big(r_0\delta_0 + \sum_{\mathbf{a}\in\mathcal{A}} r(\mathbf{a})\mu_{\mathbf{a}}\Big) + \frac{r}{\nu(\{0\})}\Big(r_0 + \sum_{\mathbf{a}\in\mathcal{A}} \frac{r(\mathbf{a})}{l_{\mathbf{a}}}\Big)\delta_1 \\ &= \frac{r}{\nu(\{0\})}\Big(r_0(\delta_0+\delta_1) + \sum_{\mathbf{a}\in\mathcal{A}} \frac{r(\mathbf{a})}{l_{\mathbf{a}}}(l_{\mathbf{a}}\mu_{\mathbf{a}} + \delta_1)\Big).\end{aligned}$$

This, together with $l_{\mathbf{a}}\mu_{\mathbf{a}} + \delta_1 = (l_{\mathbf{a}}+1)\mu_{f(\mathbf{a})}$ (see Lemma 5.5 (iii)), gives us

$$\begin{aligned}(1-r)\nu + r\delta_1 &= (1-\lambda)(1-r)\nu + \lambda(1-r)\nu + r\delta_1 \\ &= (1-\lambda)(1-r)\nu + \frac{r}{\nu(\{0\})}\Big(r_0(\delta_0+\delta_1) + \sum_{\mathbf{a}\in\mathcal{A}} \frac{r(\mathbf{a})}{l_{\mathbf{a}}}(l_{\mathbf{a}}+1)\mu_{f(\mathbf{a})}\Big).\end{aligned}$$

Then since $\mathcal{F}$ is convex, and each of the measures ν, $\frac{\delta_0+\delta_1}{2}$, and $\mu_{f(\mathbf{a})}$ belongs to $\mathcal{F}$, we conclude that $(1-r)\nu + r\delta_1 \in \mathcal{F}$, as required. ☐

Proposition 5.7. *Suppose $\mu \in \mathcal{P}(I)$ with $\mu(I\cap\mathbb{Q}) = 1$. Then $\mu \in \mathcal{F}_{[0,1)}$ if and only if μ satisfies the following conditions:*

(i) *$\mu(\{1\}) = 0$ and $\mu(\{0\}) \geqslant \mu(R_1)$.*
(ii) *$\mu(\{x\}) \geqslant \mu\big(G^{-1}(x)\big)$ for all $x \in (0,1)\cap\mathbb{Q}$.*

Proof. First we assume that $\mu \in \mathcal{F}_{[0,1)}$, and will show that μ satisfies conditions (i) and (ii). Note that these properties are closed under convex combination. From the definition of $\mathcal{F}_{[0,1)}$ (see (5.5)), and the fact that δ_0 satisfies conditions (i) and (ii), to prove that μ satisfies conditions (i) and (ii) it suffices to note that $\mu_{\mathbf{a}}$ satisfies conditions (i) and (ii) for all $\mathbf{a} \in \mathcal{A}$, by (vi) and (vii) of Lemma 5.5.

Now we assume that μ satisfies conditions (i) and (ii), and will show that this implies that $\mu \in \mathcal{F}_{[0,1)}$. Define a function $\phi_\mu \colon [0,1)\cap\mathbb{Q} \to [0,1]$ by

$$\phi_\mu(0) \coloneqq \mu(\{0\}) - \mu(R_1), \tag{5.11}$$

$$\phi_\mu(x) \coloneqq (m+1)\big(\mu(\{x\}) - \mu\big(G^{-1}(x)\big)\big) \quad \text{for } x \in R_m,\ m \in \mathbb{N}, \tag{5.12}$$

noting that conditions (i) and (ii) ensure that ϕ_μ is everywhere nonnegative. Define measures

$$\nu_0 \coloneqq \phi_\mu(0)\delta_0, \qquad \nu_n \coloneqq \sum_{x\in R_n} \phi_\mu(x)\mu_{g^{-1}(x)}, \qquad \nu \coloneqq \sum_{n=0}^{+\infty} \nu_n. \tag{5.13}$$

We note that by the constructions in (5.13), ν is a sum of positive measures. Thus, ν is a nonnegative combination of the base elements δ_0 and $\bigcup_{n\in\mathbb{N}}\{\mu_{g^{-1}(x)} : x \in R_n\}$ of $\mathcal{F}_{[0,1)}$, since $g^{-1}(x) \in \mathcal{A}$ (as $g_n\colon \mathcal{A}_n \to R_n$ is a bijection (cf. Lemma 5.5 (iv)), $x \in R_n$ guarantees $g^{-1}(x) \in \mathcal{A}_n \subseteq \mathcal{A}$, and thus $\mu_{g^{-1}(x)} \in \mathcal{F}_{[0,1)}$).

We claim that $\mu = \nu$, so in particular ν is a probability measure. From the above it will therefore follow that $\nu \in \mathcal{F}_{[0,1)}$, by the convexity of $\mathcal{F}_{[0,1)}$, and hence the required result that $\mu \in \mathcal{F}_{[0,1)}$.

To verify that $\mu = \nu$ it suffices to show that

$$\mu(\{y\}) = \nu(\{y\}) \quad \text{for all } y \in I \cap \mathbb{Q}.$$

For all $x \in (0,1) \cap \mathbb{Q}$, from the fact that $g^{-1}(x) \in \mathcal{A}$, the definition of $\mathcal{A}$ (see (3.4)), and the definition of FCF measures (see (5.4)), we obtain that $\mu_{g^{-1}(x)}(\{1\}) = 0$, and hence that

$$\nu(\{1\}) = 0 = \mu(\{1\}). \tag{5.14}$$

Now suppose $y \in (0,1)\cap\mathbb{Q}$; by Lemma 5.4, there exists $n \in \mathbb{N}$ such that $y \in R_n$. By Lemma 5.5 (v), for each $m \in \mathbb{N}$ and $x \in R_m$, we have

$$(m+1)\mu_{g^{-1}(x)}(\{y\}) = \begin{cases} 1 & \text{if } y \in \mathcal{O}(x), \\ 0 & \text{otherwise.} \end{cases}$$

Combining this with (5.12) gives

$$\phi_\mu(x)\mu_{g^{-1}(x)}(\{y\}) = \begin{cases} \mu(\{x\}) - \mu\bigl(G^{-1}(x)\bigr) & \text{if } y \in \mathcal{O}(x), \\ 0 & \text{otherwise.} \end{cases} \tag{5.15}$$

Fix $m \in \mathbb{N}$. When $m < n$, we have $y \notin \mathcal{O}(x)$ for all $x \in R_m$, and combining this with (5.13) and (5.15) gives

$$\nu_m(\{y\}) = \sum_{x\in R_m} \phi_\mu(x)\mu_{g^{-1}(x)}(y) = 0. \tag{5.16}$$

When $m \geqslant n$, for each $x \in R_m$, we have $y \in \mathcal{O}(x)$ if and only if $G^{m-n}(x) = y$, and combining this with (5.13) and (5.15) gives

$$\begin{aligned} \nu_m(\{y\}) = \sum_{x\in R_m} \phi_\mu(x)\mu_{g^{-1}(x)}(y) &= \sum_{x\in G^{-(m-n)}(y)} \mu(\{x\}) - \mu\bigl(G^{-1}(x)\bigr) \\ &= \mu\bigl(G^{-(m-n)}(y)\bigr) - \mu\bigl(G^{-(m-n+1)}(y)\bigr). \end{aligned} \tag{5.17}$$

By (5.13), (5.16), and (5.17),

$$\begin{aligned} \nu(\{y\}) = \sum_{m=n}^{+\infty} \nu_m(\{y\}) &= \sum_{m=n}^{+\infty} \bigl(\mu\bigl(G^{-(m-n)}(y)\bigr) - \mu\bigl(G^{-(m-n+1)}(y)\bigr)\bigr) \\ &= \sum_{j=0}^{+\infty} \bigl(\mu\bigl(G^{-j}(y)\bigr) - \mu\bigl(G^{-(j+1)}(y)\bigr)\bigr). \end{aligned}$$

By condition (ii), $\mu \in \mathcal{P}(I)$, and the fact that $G^{-j}(y) \cap G^{-k}(y) = \emptyset$ if $0 \leqslant j < k$, the series on the right-hand side of the above has nonnegative entries and is convergent. Thus by telescoping, we get

$$\nu(\{y\}) = \mu(\{y\}). \tag{5.18}$$

Combining (5.13) and (5.15), for each $m \in \mathbb{N}$, we get

$$\nu_m(\{0\}) = \sum_{x\in R_m} \phi_\mu(x)\mu_{g^{-1}(x)}(\{0\}) = \sum_{x\in R_m} \mu(\{x\}) - \mu\bigl(G^{-1}(x)\bigr) = \mu(R_m) - \mu(R_{m+1}).$$

Combining this with (5.13), (5.11), and a similar argument as above on the convergence of the series gives

$$\nu(\{0\}) = \sum_{n=0}^{+\infty} \nu_n(\{0\}) = \mu(\{0\}) - \mu(R_1) + \sum_{n=1}^{+\infty} (\mu(R_n) - \mu(R_{n+1})) = \mu(\{0\}). \tag{5.19}$$

Then by (5.14), (5.18), and (5.19), we conclude that $\mu = \nu \in \mathcal{F}_{[0,1)}$, as required. □

Combining Lemma 5.6 and Proposition 5.7 together gives the following corollary, which is important in the proof of Theorem 1.1.

Corollary 5.8. *Suppose $\mu \in \mathcal{P}(I)$ with $\mu(I \cap \mathbb{Q}) = 1$. Then $\mu \in \mathcal{F}$ if and only if μ satisfies the following conditions:*

(i) *$\mu(\{0\}) \geqslant \mu(\{1\})$ and $\mu(\{0\}) \geqslant \mu(R_1)$,*
(ii) *$\mu(\{x\}) \geqslant \mu\big(G^{-1}(x)\big)$ for all $x \in (0,1) \cap \mathbb{Q}$.*

Proof. First assume that $\mu \in \mathcal{F}$. Using the fact that conditions (i) and (ii) are closed under convex combination, the definition of $\mathcal{F}$ (see Definition 5.1), and the fact that δ_0 satisfies conditions (i) and (ii), it suffices to show that $\mu_{\mathbf{a}}$ satisfies conditions (i) and (ii) for all $\mathbf{a} \in \mathbb{N}^*$. When $\mathbf{a} \in \mathcal{A}$, it follows from Lemma 5.5 (vi) and (vii) that $\mu_{\mathbf{a}}$ satisfies conditions (i) and (ii). When $\mathbf{a} \in \mathcal{B}$, it follows from Lemma 5.5 (viii) and (ix) that $\mu_{\mathbf{a}}$ satisfies conditions (i) and (ii).

To prove the converse, let us assume that μ satisfies conditions (i) and (ii). Denoting $r \coloneqq \mu(\{1\}) \in [0, 1/2]$ by condition (i), define

$$\nu \coloneqq (\mu - r\delta_1)/(1-r), \tag{5.20}$$

and note that (5.20) gives $\nu(\{1\}) = 0 \leqslant \nu(\{0\})$ and $\nu(\{0\}) = \mu(\{0\})/(1-r) \geqslant \mu(\{R_1\})/(1-r) = \nu(R_1)$, which is condition (i) of Proposition 5.7, and condition (ii) gives, for all $x \in (0,1) \cap \mathbb{Q}$,

$$\nu(\{x\}) = \mu(\{x\})/(1-r) \geqslant \mu\big(G^{-1}(x)\big)/(1-r) = \nu\big(G^{-1}(x)\big),$$

which is condition (ii) of Proposition 5.7. From Proposition 5.7 it follows that $\nu \in \mathcal{F}_{[0,1)}$.

Now (5.20) can be written as $(1-r)\nu = \mu - r\delta_1$, so by condition (i),

$$(1-r)\nu(\{0\}) = \mu(\{0\}) \geqslant \mu(\{1\}) = r. \tag{5.21}$$

But (5.21) means, by Lemma 5.6, that $\mu = (1-r)\nu + r\delta_1 \in \mathcal{F}$, as required. □

Finally, we are able to prove the following main theorem.

Theorem 1.1. *The closure $\overline{\mathcal{M}(I,G)}$ is equal to the convex hull of the union of $\mathcal{M}_{\mathrm{irr}}(I,G)$ and the set of finite-continued-fraction measures.*

Proof. By Lemma 3.22 (iv), $\mathcal{M}\big(\widehat{\Sigma}, \sigma\big)$ is the convex hull of

$$\mathcal{M}(\Sigma, \sigma) \cup \big\{\nu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \nu(\Sigma) = 0\big\}.$$

But $\pi_* \colon \mathcal{M}(\Sigma,\sigma) \to \mathcal{M}_{\mathrm{irr}}(I,G)$ is a homeomorphism, according to Lemma 4.4 (iv), and the push-forward $\widehat{\pi}_* \colon \mathcal{M}\big(\widehat{\Sigma}, \sigma\big) \to \mathcal{P}(I)$ (cf. Notation 4.1) is affine, so by Lemma 4.4 , it suffices to show that $\widehat{\pi}_*(\mu) \in \mathcal{F}$ for all $\mu \in \big\{\nu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \nu(\Sigma) = 0\big\}$.

Fix $\mu \in \big\{\nu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \nu(\Sigma) = 0\big\}$. We want to apply Corollary 5.8 to $\widehat{\pi}_*(\mu)$.

Fix $x = [a_1, a_2, \dots, a_n]$ with $a_1, \dots, a_n \in \mathbb{N}$, $a_n \geqslant 2$. By Lemma 4.4 (i), we obtain

$$\widehat{\pi}^{-1}(0) = \mathcal{C}(\infty), \qquad \widehat{\pi}^{-1}(1) = \mathcal{C}(1,\infty),$$
$$\widehat{\pi}^{-1}(x) = \mathcal{C}(a_1, \dots, a_n, \infty) \cup \mathcal{C}(a_1, \dots, a_n - 1, 1, \infty).$$

Thus,

$$\widehat{\pi}_*(\mu)(\{0\}) = \mu\big(\widehat{\pi}^{-1}(0)\big) = \mu(\mathcal{C}(\infty)), \qquad \widehat{\pi}_*(\mu)(\{1\}) = \mu\big(\widehat{\pi}^{-1}(1)\big) = \mu(\mathcal{C}(1,\infty)),$$
$$\widehat{\pi}_*(\mu)(\{x\}) = \mu\big(\widehat{\pi}^{-1}(x)\big) = \mu(\mathcal{C}(a_1,\dots,a_n,\infty)) + \mu(\mathcal{C}(a_1,\dots,a_n-1,1,\infty)).$$

As a consequence, $\widehat{\pi}_*(\mu)(R_1) = \sum_{n=2}^{+\infty}\mu(\mathcal{C}(n,\infty)) + \sum_{n=1}^{+\infty}\mu(\mathcal{C}(n,1,\infty))$. Hence, from the fact that $\mu \in \mathcal{M}\big(\widehat{\Sigma},\sigma_{\widehat{\Sigma}}\big)$ and $\mu(\Sigma)=0$, we get

$$\begin{aligned}
&\widehat{\pi}_*(\mu)(I\smallsetminus\mathbb{Q}) = \mu\big(\widehat{\Sigma}\smallsetminus\Sigma\big) = 1,\\
&\quad\widehat{\pi}_*(\mu)(\{0\}) = \mu(\mathcal{C}(\infty)) \geqslant \mu(\mathcal{C}(1,\infty)) = \widehat{\pi}_*(\mu)(\{1\}),\\
&\quad\widehat{\pi}_*(\mu)(\{0\}) = \mu(\mathcal{C}(\infty)) \geqslant \sum_{n=1}^{+\infty}\mu(\mathcal{C}(n,\infty))\\
&\qquad\qquad\qquad \geqslant \sum_{n=1}^{+\infty}\mu(\mathcal{C}(n,1,\infty)) + \sum_{n=2}^{+\infty}\mu(\mathcal{C}(n,\infty)) = \widehat{\pi}_*(\mu)\big(\{1/n\}_{n=2}^{+\infty}\big),\\
&\quad\widehat{\pi}_*(\mu)(\{x\}) = \mu(\mathcal{C}(a_1,\dots,a_n,\infty)) + \mu(\mathcal{C}(a_1,\dots,a_n-1,1,\infty))\\
&\qquad\qquad\qquad \geqslant \sum_{m=1}^{+\infty}\mu(\mathcal{C}(m,a_1,\dots,a_n,\infty)) + \sum_{m=1}^{+\infty}\mu(\mathcal{C}(m,a_1,\dots,a_n-1,1,\infty)) = \widehat{\pi}_*(\mu)\big(G^{-1}(x)\big).
\end{aligned}$$

The last identity above follows from (3.1). Therefore, applying Corollary 5.8 to $\widehat{\pi}_*(\mu)$, we conclude $\widehat{\pi}_*(\mu) \in \mathcal{F}$, as required. □

6. The Mañé lemma

In this section we will prove a version of the Mañé lemma for the Gauss map G, and then use this to derive a revelation theorem. The approach, by analogy with [Bou00], will involve a certain nonlinear operator which can be shown (cf. Proposition 6.6) to have a fixed point function (a so-called *calibrated sub-action*, in the terminology of [GLT09]) with certain regularity properties.

For a Borel measurable map $T\colon I \to I$, and bounded Borel measurable function $\psi\colon I \to \mathbb{R}$, to study the (T,ψ)-maximizing measures it is convenient, whenever possible, to consider a cohomologous function $\widetilde{\psi}$ satisfying $\widetilde{\psi} \leqslant Q(T,\psi)$. We recall the following (cf. [Je19, p. 2601]):

Definition 6.1. Suppose $T\colon I \to I$ is Borel measurable, and $\psi \in C(I)$. If $\psi \leqslant Q(T,\psi)$ and $\psi^{-1}(Q(T,\psi))$ contains $\operatorname{supp}\mu$ for some $\mu \in \mathcal{M}(I,T)$, then ψ is said to be *revealed*. If $Q(T,\psi)=0$ then ψ is said to be *normalised*; in particular, a normalised function ψ is revealed if and only if $\psi \leqslant 0$ and $\psi^{-1}(0)$ contains $\operatorname{supp}\mu$ for some $\mu \in \mathcal{M}(I,T)$.

Lemma 6.2. *Suppose $T\colon I \to I$ is Borel measurable, $\phi\colon I \to \mathbb{R}$ is bounded and Borel measurable, and $\mathcal{M}_{\max}(T,\phi) \neq \emptyset$. Denote $\overline{\phi} \coloneqq \phi - Q(T,\phi)$, and suppose $\widetilde{\phi} \coloneqq \overline{\phi} + u - u\circ T$ for some bounded Borel measurable function $u\colon I \to \mathbb{R}$. Then:*

(i) $Q\big(T,\widetilde{\phi}\big) = Q\big(T,\overline{\phi}\big) = 0$.

(ii) $\mathcal{M}_{\max}(T,\phi) = \mathcal{M}_{\max}\big(T,\overline{\phi}\big) = \mathcal{M}_{\max}\big(T,\widetilde{\phi}\big)$.

Proof. (i) and (ii) follow from (1.2), (1.3), and the fact that $\int\widetilde{\phi}\,\mathrm{d}\mu = \int(\overline{\phi}+u-u\circ T)\,\mathrm{d}\mu = \int\overline{\phi}\,\mathrm{d}\mu$ for all $\mu \in \mathcal{M}(I,T)$. □

6.1. **Bousch operator.** The following operator $\mathcal{L}_\psi$ is an analogue of the one used by Bousch in [Bou00]. Instead of preimages used by Bousch, we use inverse branches in the definition to address the irregular behavior of the Gauss map at 0 and 1.

Definition 6.3. Let $\psi\colon I \to \mathbb{R}$ be bounded and Borel measurable. Define $\mathcal{L}_\psi\colon B(I) \to B(I)$ by

$$\mathcal{L}_\psi(u)(x) \coloneqq \sup\{(u+\psi)(G_a(x)) : a \in \mathbb{N}\} = \sup\{u(1/(a+x)) + \psi(1/(a+x)) : a \in \mathbb{N}\}.$$

Since ψ and u are bounded, $\mathcal{L}_\psi(u)$ is well-defined. If ψ and u are continuous, by Proposition 3.13 (v) we have

$$\mathcal{L}_\psi(u)(x) \coloneqq \max\bigl\{(u+\psi)(G_{\widehat{a}}(x)) : \widehat{a} \in \widehat{\mathbb{N}}\bigr\}. \tag{6.1}$$

Lemma 6.4. *If $\psi \in C(I)$ and $\overline{\psi} \coloneqq \psi - Q(G,\psi)$, then the following hold:*

(i) *If $x \in I$ and $u \in C(I)$, then $\mathcal{L}_\psi(u+c) = c + \mathcal{L}_\psi(u)$.*

(ii) *If $x \in I$, $n \in \mathbb{N}$, and $u \in C(I)$, then*

$$\mathcal{L}^n_{\overline{\psi}}(u)(x) + nQ(G,\psi) = \mathcal{L}^n_\psi(u)(x) = \sup\{u(G_{\mathbf{a}}(x)) + S_{n,\mathbf{a}}\psi(x) : \mathbf{a} \in \mathbb{N}^n\}.$$

(iii) *If $x \in I$, $n \in \mathbb{N}$, and $u \in C(I)$, then*

$$\mathcal{L}^n_\psi(u)(x) = \max\bigl\{u(G_{\widehat{\mathbf{a}}}(x)) + S_{n,\widehat{\mathbf{a}}}\psi(x) : \widehat{\mathbf{a}} \in \widehat{\mathbb{N}}^n\bigr\}.$$

(iv) *$\mathcal{L}_\psi(\sup_{v\in\mathcal{H}} v) = \sup_{v\in\mathcal{H}} \mathcal{L}_\psi(v)$ for any collection $\mathcal{H}$ of bounded real-valued functions on I.*

(v) *If $\{u_n\}_{n\in\mathbb{N}}$ is a pointwise convergent sequence of equicontinuous functions on I, then the identity $\lim_{n\to+\infty} \mathcal{L}_\psi(u_n) = \mathcal{L}_\psi\bigl(\lim_{n\to+\infty} u_n\bigr)$ holds, where the limits are pointwise.*

Proof. (i) By Definition 6.3, for any $x \in I$ and $u \in C(I)$,

$$\mathcal{L}_\psi(u+c)(x) = \sup\{\psi(G_a(x)) + u(G_a(x)) + c : a \in \mathbb{N}\} = \mathcal{L}_\psi(u)(x) + c.$$

(ii) The first identity immediately follows from the second identity. We use induction to prove the second identity: the case $n = 1$ follows from Definition 6.3, and assuming it is satisfied for some $n = m \in \mathbb{N}$, then

$$\begin{aligned}\mathcal{L}^{m+1}_\psi(u)(x) &= \sup\bigl\{\psi(G_a(x)) + \mathcal{L}^m_\psi(u)(G_a(x)) : a \in \mathbb{N}\bigr\}\\ &= \sup\{\psi(G_a(x)) + \sup\{u(y) + S_{m,\mathbf{a}}\psi(G_a(x)) : y = G_{\mathbf{a}}(G_a(x)),\, \mathbf{a} \in \mathbb{N}^m\} : a \in \mathbb{N}\}\\ &= \sup\bigl\{u(G_{\mathbf{b}}(x)) + S_{m+1,\mathbf{b}}\psi(G_{\mathbf{b}}(x)) : \mathbf{b} \in \mathbb{N}^{m+1}\bigr\}.\end{aligned}$$

(iii) By Proposition 3.13 (iv), and the fact that $u,\, \psi \in C(I)$, if $x \in I$ then $\widehat{\mathbf{a}} \mapsto u(G_{\widehat{\mathbf{a}}}(x)) + S_{n,\widehat{\mathbf{a}}}\psi(x)$ can be seen as a continuous function on $\widehat{\mathbb{N}}^n$; so (iii) follows from (ii) and the fact that $\mathbb{N}^n$ is dense in $\widehat{\mathbb{N}}^n$.

(iv) follows readily from the definition.

(v) Let v be the pointwise limit of $\{u_n\}_{n\in\mathbb{N}}$ as n tends to infinity. Fix arbitrary $x \in I$ and $\epsilon > 0$. Since $\{u_n\}_{n\in\mathbb{N}}$ is equicontinuous, there exists $\delta \in (0,1)$ such that for each $y \in [0,\delta)$ and each $n \in \mathbb{N}$,

$$|u_n(y) - u_n(0)| < \epsilon/3. \tag{6.2}$$

Letting n tend to infinity, we have

$$|v(y) - v(0)| \leqslant \epsilon/3. \tag{6.3}$$

We can find $N_1 \in \mathbb{N}$ such that if $n > N_1$ then $|u_n(0) - v(0)| < \epsilon/3$. When $n > N_1$, for each $y \in \{G_a(x)\}_{a\in\mathbb{N}} \cap [0,\delta)$, by (6.2) and (6.3), we obtain

$$|u_n(y) - v(y)| \leqslant |u_n(y) - u_n(0)| + |u_n(0) - v(0)| + |v(0) - v(y)| < 3 \cdot (\epsilon/3) < \epsilon. \tag{6.4}$$

Since $\{G_a(x) : a \in \mathbb{N}\} \cap [\delta,1]$ is finite, we can find $N_2 \in \mathbb{N}$ such that for each $n \geqslant N_2$ and each $y \in \{G_a(x) : a \in \mathbb{N}\} \cap [\delta,1]$, we obtain

$$|u_n(y) - v(y)| < \epsilon. \tag{6.5}$$

Let $N \coloneqq \max\{N_1, N_2\}$. For each integer $n > N$, by (6.4) and (6.5), we have $|u_n(y) - v(y)| < \epsilon$ for each $y \in \{G_a(x) : a \in \mathbb{N}\}$. Fix an arbitrary integer $n > N$. We choose $z_1, z_2 \in \{G_a(x) : a \in \mathbb{N}\}$ satisfying $\mathcal{L}_\psi(u_n)(x) < \psi(z_1) + u_n(z_1) + \epsilon$ and $\mathcal{L}_\psi(v)(x) < \psi(z_2) + v(z_2) + \epsilon$. Then by Definition 6.3,

$$\mathcal{L}_\psi(u_n)(x) - \mathcal{L}_\psi(v)(x) < \psi(z_1) + u_n(z_1) + \epsilon - \psi(z_1) - v(z_1) = u_n(z_1) - v(z_1) + \epsilon < 2\epsilon,$$
$$\mathcal{L}_\psi(u_n)(x) - \mathcal{L}_\psi(v)(x) > \psi(z_2) + u_n(z_2) - \psi(z_2) - v(z_2) - \epsilon = u_n(z_2) - v(z_2) - \epsilon > -2\epsilon.$$

Statement (v) now follows. □

Lemma 6.5. *Suppose $\alpha \in (0, 1]$ and $\phi \in C^{0,\alpha}(I)$. Then for each $u \in C^{0,\alpha}(I)$ and each $n \in \mathbb{N}$, we have $\mathcal{L}^n_\phi(u) \in C^{0,\alpha}(I)$ and*

$$\left|\mathcal{L}^n_\phi(u)\right|_\alpha \leqslant K_\alpha(|\phi|_\alpha + |u|_\alpha). \tag{6.6}$$

Proof. Suppose $u \in C^{0,\alpha}(I)$ and $x, y \in I$. Fix $\epsilon > 0$. By Lemma 6.4 (ii), there exists $\mathbf{a} \in \mathbb{N}^n$ such that

$$\mathcal{L}^n_\phi(u)(x) < u(G_{\mathbf{a}}(x)) + S_{n,\mathbf{a}}\phi(x) + \epsilon. \tag{6.7}$$

By Lemma 6.4 (ii), we have

$$\mathcal{L}^n_\phi(u)(y) \geqslant u(G_{\mathbf{a}}(y)) + S_{n,\mathbf{a}}\phi(y). \tag{6.8}$$

Combining (6.7) and (6.8) gives

$$\mathcal{L}^n_\phi(u)(x) - \mathcal{L}^n_\phi(u)(y) \leqslant S_{n,\mathbf{a}}\phi(x) + u(G_{\mathbf{a}}(x)) - S_{n,\mathbf{a}}\phi(y) - u(G_{\mathbf{a}}(y)) + \epsilon. \tag{6.9}$$

Lemma 3.15 and (2.1) gives

$$S_{n,\mathbf{a}}\phi(x) - S_{n,\mathbf{a}}\phi(y) \leqslant K_\alpha|\phi|_\alpha|x - y|^\alpha. \tag{6.10}$$

From the fact that $u \in C^{0,\alpha}(I)$, the intermediate value theorem, and Proposition 3.13 (i), there exists ξ in between x and y such that

$$u(G_{\mathbf{a}}(x)) - u(G_{\mathbf{a}}(y)) \leqslant |u|_\alpha|G_{\mathbf{a}}(x) - G_{\mathbf{a}}(y)|^\alpha = |u|_\alpha|x - y|^\alpha|G'_{\mathbf{a}}(\xi)|^\alpha \leqslant c_0^{-2\alpha}\theta^{-2n\alpha}|u|_\alpha|x - y|^\alpha.$$

But $c_0^{-2\alpha}\theta^{-2n\alpha} < K_\alpha$ (see (2.1)), so

$$u(G_{\mathbf{a}}(x)) - u(G_{\mathbf{a}}(y)) \leqslant K_\alpha|u|_\alpha|x - y|^\alpha. \tag{6.11}$$

Combining (6.9), (6.10), and (6.11) gives

$$\mathcal{L}^n_\phi(u)(x) - \mathcal{L}^n_\phi(u)(y) \leqslant K_\alpha(|\phi|_\alpha + |u|_\alpha)|x - y|^\alpha + \epsilon. \tag{6.12}$$

Since (6.12) is satisfied for all $x, y \in I$, by swapping the positions of x and y, we obtain

$$\mathcal{L}^n_\phi(u)(x) - \mathcal{L}^n_\phi(u)(y) \geqslant -K_\alpha(|\phi|_\alpha + |u|_\alpha)|x - y|^\alpha - \epsilon. \tag{6.13}$$

Finally, $\mathcal{L}^n_\phi(u) \in C^{0,\alpha}(I)$ and (6.6) follows from (6.12), (6.13), and the fact that $\epsilon > 0$ was arbitrary. □

We are now able to find a fixed point u_ϕ of the normalised Bousch operator $\mathcal{L}_{\overline{\phi}}$:

Proposition 6.6. *Suppose $\alpha \in (0, 1]$ and $\phi \in C^{0,\alpha}(I)$. Then the function $u_\phi \colon I \to \mathbb{R}$ given by*

$$u_\phi(x) \coloneqq \limsup_{n \to +\infty} \mathcal{L}^n_{\overline{\phi}}(\mathbb{0})(x), \qquad \text{for } x \in I, \tag{6.14}$$

where $\overline{\phi} \coloneqq \phi - Q\big(G, \phi\big)$, satisfies the following properties:

(i) $|u_\phi(x)| \leqslant K_\alpha|\phi|_\alpha$ *for each* $x \in I$,

(ii) $u_\phi \in C^{0,\alpha}(I)$ *with* $|u_\phi|_\alpha \leqslant K_\alpha|\phi|_\alpha$,

(iii) $\mathcal{L}_{\overline{\phi}}(u_\phi) = u_\phi$.

Proof. For each $n \in \mathbb{N}$ and each $x \in I$, we write

$$r_n(x) \coloneqq \mathcal{L}^n_{\overline{\phi}}(\mathbb{0})(x) \qquad \text{and} \qquad s_n(x) \coloneqq \sup\{r_m(x) : m \geqslant n\}. \tag{6.15}$$

Note that, for each $x \in I$, the sequence $\{s_n(x)\}_{n\in\mathbb{N}}$ is nonincreasing and by (6.14) and (6.15),

$$u_\phi(x) = \lim_{n\to+\infty} s_n(x) = \limsup_{n\to+\infty} r_n(x).$$

(i) Fix $x \in I$ and $n \in \mathbb{N}$. For each $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{N}^n$, denote $p_{\mathbf{a}} \coloneqq [\overline{a_1, \dots, a_n}]$ which satisfies $G_{\mathbf{a}}(p_{\mathbf{a}}) = G^n(p_{\mathbf{a}}) = p_{\mathbf{a}}$. By Proposition 3.13 (vi), (3.10), and the fact that $Q\big(G, \overline{\phi}\big) = 0$, we get

$$S_{n,\mathbf{a}}\overline{\phi}(p_{\mathbf{a}}) = S_n\overline{\phi}(p_{\mathbf{a}}) = n\int_I \overline{\phi}\,\mathrm{d}\mu_{\mathcal{O}(p_{\mathbf{a}})} \leqslant 0 \tag{6.16}$$

Combining this with Lemma 3.15 gives

$$S_{n,\mathbf{a}}\overline{\phi}(x) = S_{n,\mathbf{a}}\overline{\phi}(x) - S_{n,\mathbf{a}}\overline{\phi}(p_{\mathbf{a}}) + S_{n,\mathbf{a}}\overline{\phi}(p_{\mathbf{a}}) \leqslant K_\alpha|\phi|_\alpha. \tag{6.17}$$

So by (6.15), Lemma 6.4 (ii), and (6.17), we have

$$r_n(x) = \mathcal{L}^n_{\overline{\phi}}(\mathbb{0})(x) = \sup\big\{S_{n,\mathbf{a}}\overline{\phi}(x) : \mathbf{a} \in \mathbb{N}^n\big\} \leqslant K_\alpha|\phi|_\alpha.$$

Combining this with (6.15) and (6.14) gives

$$u_\phi(x) = \limsup_{n\to+\infty} r_n(x) \leqslant K_\alpha|\phi|_\alpha.$$

Next, we will show that $u_\phi(x) \geqslant -K_\alpha|\phi|_\alpha$. Fix $n \in \mathbb{N}$. We choose a point $A_n = (a_i)_{i\in\mathbb{N}} \in \widehat{\Sigma}$ on which $S^\sigma_n\big(\overline{\phi} \circ \widehat{\pi}\big)$ attains its maximum value. Denote $y_n \coloneqq \widehat{\pi}(\sigma^n(A_n))$ and $\mathbf{a} \coloneqq (a_1, a_2, \dots, a_n)$. By Lemma 4.4 (vi), we get

$$S_{n,\widehat{\mathbf{a}}}\overline{\phi}(y_n) = S^\sigma_n\big(\overline{\phi} \circ \widehat{\pi}\big)(A_n). \tag{6.18}$$

Choose $\mu \in \mathcal{M}_{\max}\big(\sigma_{\widehat{\Sigma}}, \overline{\phi} \circ \widehat{\pi}\big)$. By Proposition 4.6 (i), we have $Q\big(\sigma_{\widehat{\Sigma}}, \overline{\phi} \circ \widehat{\pi}\big) = Q(T, \overline{\phi}) = 0$, and then we have $\int_{\widehat{\Sigma}} S^\sigma_n\big(\overline{\phi} \circ \widehat{\pi}\big)\,\mathrm{d}\mu = 0$. So for all $x \in I$, combining (6.15), Lemma 6.5, and (6.18) gives

$$\begin{aligned} r_n(x) &\geqslant r_n(y_n) - K_\alpha|\phi|_\alpha \geqslant S_{n,\widehat{\mathbf{a}}}\overline{\phi}(y_n) - K_\alpha|\phi|_\alpha = S^\sigma_n\big(\overline{\phi} \circ \widehat{\pi}\big)(A_n) - K_\alpha|\phi|_\alpha \\ &\geqslant \int_{\widehat{\Sigma}} S^\sigma_n\big(\overline{\phi} \circ \widehat{\pi}\big)\,\mathrm{d}\mu - K_\alpha|\phi|_\alpha = -K_\alpha|\phi|_\alpha. \end{aligned}$$

Combining this with (6.15) and (6.14) gives $u_\phi(x) \geqslant -K_\alpha|\phi|_\alpha$ for all $x \in I$, so (i) follows.

(ii) Suppose $x, y \in I$ and fix $\epsilon > 0$. By (6.15) and (6.14), there exists $N \in \mathbb{N}$ such that $|r_N(x) - u_\phi(x)| < \epsilon$ and $s_N(y) - u_\phi(y) < \epsilon$. So by (6.15) and Lemma 6.5,

$$u_\phi(x) - u_\phi(y) < r_N(x) - s_N(y) + 2\epsilon \leqslant r_N(x) - r_N(y) + 2\epsilon \leqslant K_\alpha|\phi|_\alpha|x - y|^\alpha + 2\epsilon, \tag{6.19}$$

where the final inequality uses (6.6). Similarly, there exists $M \in \mathbb{N}$ such that $|r_M(y) - u_\phi(y)| < \epsilon$ and $s_M(x) - u_\phi(x) < \epsilon$, and an analogous calculation gives

$$u_\phi(x) - u_\phi(y) \geqslant -K_\alpha|\phi|_\alpha|x - y|^\alpha - 2\epsilon. \tag{6.20}$$

Since $\epsilon > 0$ was arbitrary, (ii) follows from (6.19) and (6.20).

(iii) First we prove that $\{s_n\}_{n\in\mathbb{N}}$ is equicontinuous. Fix arbitrary $\epsilon > 0$ and $m \in \mathbb{N}$. By (6.6) and (6.15), $\{r_n\}_{n\in\mathbb{N}}$ is equicontinuous. Hence there exists $\delta > 0$ such that if $|x - y| < \delta$, we have

$$|r_n(x) - r_n(y)| < \epsilon/2$$

for all $n \in \mathbb{N}$. Then fix arbitrary $x, y \in I$ satisfying $|x - y| < \delta$.

Since $s_m(x) = \sup_{k\geqslant m}\{r_k(x)\}$, we can find $N_1 > m$ such that $s_m(x) < r_{N_1}(x) + \frac{\epsilon}{2}$. Then we have

$$s_m(x) - s_m(y) < r_{N_1}(x) + \frac{\epsilon}{2} - s_m(y) \leqslant r_{N_1}(x) + \frac{\epsilon}{2} - r_{N_1}(y) \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Similarly, we can find $N_2 > m$ such that $s_m(y) < r_{N_2}(y) + \frac{\epsilon}{2}$. Then we have

$$s_m(x) - s_m(y) > s_m(x) - r_{N_2}(y) - \frac{\epsilon}{2} \geqslant r_{N_2}(x) - r_{N_2}(y) - \frac{\epsilon}{2} \geqslant -\frac{\epsilon}{2} - \frac{\epsilon}{2} = -\epsilon.$$

Therefore, $\{s_n\}_{n\in\mathbb{N}}$ is equicontinuous.

If $x \in I$, then by Lemma 6.4 (iv), (v), and (6.15),

$$\begin{aligned}\mathcal{L}_{\overline{\phi}}(u_\phi)(x) &= \mathcal{L}_{\overline{\phi}}\big(\lim_{n\to+\infty} s_n\big)(x) = \lim_{n\to+\infty} \mathcal{L}_{\overline{\phi}}\big(\sup\big\{\mathcal{L}^m_{\overline{\phi}}(\mathbb{0})(x) : m \geqslant n\big\}\big) \\ &= \lim_{n\to+\infty}\big(\sup\big\{\mathcal{L}^{m+1}_{\overline{\phi}}(\mathbb{0})(x) : m \geqslant n\big\}\big) = \lim_{n\to+\infty} s_{n+1}(x) = u_\phi(x).\end{aligned}$$

□

Definition 6.7. Suppose $\phi \in C^{0,\alpha}(I)$, and u_ϕ is the calibrated sub-action defined by (6.14). We define the *revealed version* $\widetilde{\phi}$ by

$$\widetilde{\phi} := \overline{\phi} + u_\phi - u_\phi \circ G.$$

We are now able to prove Theorem 1.2, a Mañé lemma for the Gauss map, which resembles the form of [Bou00, Lemma A].

Theorem 1.2 (Mañé Lemma). *Suppose $\alpha \in (0, 1]$ and $\phi \in C^{0,\alpha}(I)$. There exists $u_\phi \in C^{0,\alpha}(I)$ satisfying the functional equation*

$$u_\phi(x) = \sup_{n\in\mathbb{N}}\Big\{\overline{\phi}\Big(\frac{1}{n+x}\Big) + u_\phi\Big(\frac{1}{n+x}\Big)\Big\} \quad \text{for all } x \in I, \tag{1.5}$$

where $\overline{\phi} := \phi - Q(G, \phi)$.

Proof. This follows immediately from Definition 6.3 and Proposition 6.6. □

6.2. **Maximizing set.** Now, by analogy with the set Z' of [Bou00, p. 495], and the admissible words defined in [BM02, Definition 2.3], we wish to relate each $\phi \in C^{0,\alpha}(I)$ to a *maximizing set* $\mathcal{K}(\phi) \subseteq \widehat{\Sigma}$.

Definition 6.8. Given $\alpha \in (0, 1]$, $\phi \in C^{0,\alpha}(I)$, and u_ϕ the calibrated sub-action defined by (6.14), define

$$\Phi := \phi \circ \widehat{\pi}, \tag{6.21}$$

$$\overline{\Phi} := \Phi - Q(G, \phi) = \overline{\phi} \circ \widehat{\pi}, \tag{6.22}$$

$$U_\Phi := u_\phi \circ \widehat{\pi}, \text{ and} \tag{6.23}$$

$$\Psi := \Phi - Q(G, \phi) + U_\Phi - U_\Phi \circ \sigma, \tag{6.24}$$

and define the *maximizing set* for ϕ by $\mathcal{K}(\phi) := \bigcap_{n=1}^{+\infty} \sigma^{-n}_{\widehat{\Sigma}}\big(\Psi^{-1}(0)\big)$.

Lemma 6.9. *If $\phi \in C^{0,\alpha}(I)$ with $\alpha \in (0, 1]$, then the following hold:*

(i) Φ, U_Φ, $\Psi \in C^{0,\alpha}\big(\widehat{\Sigma}, d_{\widehat{\rho}}\big)$.

(ii) U_Φ *satisfies the functional equation*

$$U_\Phi(A) = \max\big\{\overline{\Phi}(B) + U_\Phi(B) : B \in \sigma^{-1}_{\widehat{\Sigma}}(A)\big\}, \quad \text{for all } A \in \widehat{\Sigma}, \tag{6.25}$$

and consequently $\Psi \leqslant 0$.

(iii) *$\mathcal{K}(\phi)$ is a nonempty compact closed subset of $\widehat{\Sigma}$, with $\sigma(\mathcal{K}(\phi)) \subseteq \mathcal{K}(\phi)$.*

(iv) $\mathcal{M}_{\max}\big(\sigma_{\widehat{\Sigma}}, \Phi\big) = \mathcal{M}_{\max}\big(\sigma_{\widehat{\Sigma}}, \Psi\big) = \big\{\mu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big) : \operatorname{supp} \mu \subseteq \mathcal{K}(\phi)\big\} \neq \emptyset.$

Proof. (i) follows immediately from the fact that ϕ, $u_\phi \in C^{0,\alpha}(I)$, and since $\widehat{\pi}$, σ are Lipschitz (see Lemmas 4.4 (iii) and 3.22 (i)).

(ii) Fix $A \in \widehat{\Sigma}$. By definition of Φ and U_Φ (see (6.21) and (6.23)), Proposition 6.6 (iii), (6.1), and Lemma 4.4 (ii), we obtain

$$\begin{aligned} U_\Phi(A) = u_\phi(\widehat{\pi}(A)) &= \max\bigl\{\overline{\phi}(G_{\widehat{a}}(\widehat{\pi}(A))) + u_\phi(G_{\widehat{a}}(\widehat{\pi}(A))) : \widehat{a} \in \widehat{\mathbb{N}}\bigr\} \\ &= \max\bigl\{\overline{\phi}(G_{\widehat{a}}(\widehat{\pi}(\sigma(\widehat{a}A)))) + u_\phi(G_{\widehat{a}}(\widehat{\pi}(\sigma(\widehat{a}A)))) : \widehat{a} \in \widehat{\mathbb{N}}\bigr\} \\ &= \max\bigl\{\overline{\phi}(\widehat{\pi}(\widehat{a}A)) + u_\phi(\widehat{\pi}(\widehat{a}A)) : \widehat{a} \in \widehat{\mathbb{N}}\bigr\} \\ &= \max\bigl\{\overline{\Phi}(B) + U_\Phi(B) : B \in \sigma_{\widehat{\Sigma}}^{-1}(A)\bigr\}. \end{aligned}$$

(iii) By definition of $\mathcal{K}(\phi)$, it is immediate from the continuity of σ and Ψ that $\mathcal{K}(\phi)$ is compact. By definition of $\mathcal{K}(\phi)$, it is also clear that $\sigma(\mathcal{K}(\phi)) \subseteq \mathcal{K}(\phi)$. The fact that $\mathcal{K}(\phi)$ is nonempty will follow directly from (iv) and the fact that $\mathcal{M}_{\max}\bigl(\sigma_{\widehat{\Sigma}}, \Phi\bigr)$ is nonempty.

(iv) The first identity follows from (6.24), (6.21), and Proposition 4.6 (i). To establish the second identity, we first note that by the first identity, (1.3), and (ii), every $\mu \in \mathcal{M}\bigl(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\bigr)$ with $\operatorname{supp}\mu \subseteq \mathcal{K}(\phi) \subseteq \Psi^{-1}(0)$ is in $\mathcal{M}_{\max}\bigl(\sigma_{\widehat{\Sigma}}, \Psi\bigr)$. Conversely, by (ii) and Proposition 4.6 (i), every $\mu \in \mathcal{M}_{\max}\bigl(\sigma_{\widehat{\Sigma}}, \Psi\bigr)$ satisfies $\int_{\widehat{\Sigma}} \Psi \,\mathrm{d}\mu = 0$. By (ii), $\operatorname{supp}\mu$ is a subset of the compact set $\Psi^{-1}(0)$. It now follows from the $\sigma_{\widehat{\Sigma}}$-invariance of μ that $\operatorname{supp}\mu \subseteq \bigcap_{n=0}^{+\infty} \sigma^{-n}\bigl(\Psi^{-1}(0)\bigr) = \mathcal{K}(\phi)$. The inequality follows from the fact that $\mathcal{M}\bigl(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\bigr)$ is compact with respect to the weak* topology. □

7. Typical finite optimization

It will be convenient to classify the potential functions $\phi \in C^{0,\alpha}(I)$ as follows:

Definition 7.1 (Classification of potentials). For $\alpha \in (0,1]$, a function $\phi \in C^{0,\alpha}(I)$ is said to be

(i) *essentially compact*[5] if $Q(G,\phi) = Q_m(G,\phi)$ (cf.Definition 3.26) for some $m \in \mathbb{N}$. Let $\mathfrak{E}^\alpha(G)$ denote the set of α-Hölder essentially compact functions;

(ii) *rationally maximized* if there exists $n \in \mathbb{N}$ and $\mathbf{a} \in \mathbb{N}^n$ such that $\int_I \phi \,\mathrm{d}\mu_{\mathbf{a}} = Q(G,\phi)$ or $\phi(0) = Q(G,\phi)$. Let $\mathfrak{R}^\alpha(G)$ denote the set of α-Hölder rationally maximized functions;

(iii) let $\mathfrak{Z}^\alpha(I)$ denote the set of α-Hölder functions satisfying neither (i) nor (ii) above.

In this section we will firstly establish the *typical finite optimization* for rationally maximized potentials (see Theorem 1.4), and secondly prove that the set $\mathfrak{E}^\alpha(G)$ of essentially compact functions is contained in the closure of $\mathrm{Lock}^\alpha(G)$ (see Theorem 1.3),

In the proof of Theorem 1.3, we show that for an arbitrary $\phi \in \mathfrak{E}^\alpha(G)$, any perturbation of the form $\phi' = \phi - \epsilon d(\cdot, \mathcal{O})^\alpha$, with $\epsilon > 0$ sufficiently small, belongs to $\mathscr{P}^\alpha(G)$, where $\mathcal{O}$ is a particular periodic orbit. The perturbation argument in our proof of Theorem 1.3 is mainly inspired by ideas appearing in [Co16], [Boc19], [HLMXZ25], and [LZ25].

Contreras [Co16] established the TPO conjecture for expanding maps using a proof by contradiction. Huang, Lian, Ma, Xu, and Zhang [HLMXZ25] proved the TPO conjecture in a wider uniformly hyperbolic setting, notably for Axiom A diffeomorphisms; their argument provides a direct proof based on a closing lemma. Bochi [Boc19] gave a simplified version of the argument in [HLMXZ25], valid for expanding maps. Li and Zhang [LZ25] proved an analogous result for a class of postcritically-finite maps (namely, expanding Thurston maps) arising in complex dynamics, where local closing lemmas are established and a local perturbation argument utilized, in order to avoid finitely many "bad points".

[5] The terminology follows [JMU06, JMU07]

Compared with the ideas and techniques in the aforementioned works, our approach is to apply the closing lemma from [HLMXZ25] in a neighbourhood of a maximizing measure supported on E_m for some $m \in \mathbb{N}$, and then to carry out a local analysis following the perspective discussed in [Boc19] near E_m. The main difficulties in our setting arise from the fact that G has *countably* many inverse branches, is discontinuous, and is not expanding at 1.

The technical ingredients of the proof consist of (1) quantitatively avoiding the discontinuities of G, using the fact that G is Lipschitz and distance-expanding in a small neighbourhood of E_m for each $m \in \mathbb{N}$ (cf. Lemma 3.27) and (2) handling the perturbation argument with constants that are necessarily local (such as η_m, λ_m, L_1, L_2, and L_3 in the proof of Theorem 1.3).

The proof of Theorem 1.4 is inspired by the proof of the periodic locking property (cf. [BZ15] and [YH99, Remark 4.5]). We first prove a technical lemma (Lemma 7.4). The technical part in the proof of Lemma 7.4 is the construction of a *transport sequence*[6] in a given rational orbit. Since there exists more than one FCF measure in a given rational orbit, the perturbation in the proof of Theorem 1.4 is also more sophisticated than the one for the periodic locking property.

7.1. **TFO for rationally maximized potentials.** In this subsection we will establish the following theorem.

Theorem 1.4 (TFO for rationally maximized potentials). *For $\alpha \in (0,1]$, the set $\mathrm{Lock_F}^{\alpha}(G)$ contains an open dense subset of $\mathfrak{R}^{\alpha}(G)$ (in the α-Hölder topology).*

Notation 7.2. Recall that $\bigcup_{n\in\mathbb{N}} R_n \coloneqq (0,1)\cap\mathbb{Q}$ (see Lemma 5.4). For each $x \in I \smallsetminus \mathbb{Q}$, define $\widetilde{\mathcal{O}}(x)$ as follows:

$$\widetilde{\mathcal{O}}(x) \coloneqq \begin{cases} \{0\} & \text{if } x = 0, \\ \{0,\, 1\} & \text{if } x = 1, \\ \bigl\{x,\, G(x),\, \dots,\, G^{n-1}(x),\, 0,\, 1\bigr\} & \text{if } x \in R_n \text{ for some } n \in \mathbb{N}. \end{cases}$$

Define

$$\mathcal{M}_0 \coloneqq \{\delta_0\} \quad \text{and} \quad \mathcal{M}_1 \coloneqq \{\delta_0,\, (\delta_0+\delta_1)/2\}.$$

Suppose $n \in \mathbb{N}$ and $x = [a_1, a_2, \dots, a_n] \in R_n$ with $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}_n$ and $\mathbf{b} \coloneqq g(\mathbf{a}) = (a_1, \dots, a_{n-1}, a_n - 1, 1) \in \mathcal{B}_{n+1}$. Define

$$\mathcal{M}_x \coloneqq \{\mu_{\mathbf{a}},\, \mu_{\sigma(\mathbf{a})} \dots,\, \mu_{\sigma^{n-1}(\mathbf{a})},\, \mu_{\mathbf{b}},\, \mu_{\sigma(\mathbf{b})} \dots,\, \mu_{\sigma^{n}(\mathbf{b})},\, \delta_0\}. \tag{7.1}$$

Denote $R_\epsilon(x) \coloneqq G_{\mathbf{a}}(\epsilon) = [a_1, \dots, a_n + \epsilon]$ and $L_\epsilon(x) \coloneqq G_{\mathbf{b}}(\epsilon) = [a_1, \dots, a_n - 1, 1+\epsilon]$.

Let $\mathrm{conv}(\mathcal{M}_x)$ denote the convex hull of $\mathcal{M}_x$.

The following two technical lemmas will be used in the proof of Theorem 1.4.

Lemma 7.3. *Suppose $\epsilon \in (0,1)$, $n \in \mathbb{N}$, and $x = [a_1, \dots, a_n] \in R_n$ with $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}_n$.*

(i) *When n is odd, $R_\epsilon(x) < x < L_\epsilon(x)$. When n is even, $L_\epsilon(x) < x < R_\epsilon(x)$.*

(ii) *Suppose n is odd. If $y \in (R_\epsilon(x), x]$ then $\bigl|G^i(y) - G^i(x)\bigr| \leqslant \epsilon$ for all $0 \leqslant i \leqslant n$. If $y \in (x, L_\epsilon(x))$ then $\bigl|G^i(y) - G^i(x)\bigr| \leqslant \epsilon$ for all $0 \leqslant i \leqslant n-1$, and moreover $|G^n(y) - 1| \leqslant \epsilon$ and $\bigl|G^{n+1}(y)\bigr| \leqslant \epsilon$.*

(iii) *Suppose n is even. If $y \in [x, R_\epsilon(x))$ then $\bigl|G^i(y) - G^i(x)\bigr| \leqslant \epsilon$ for all $0 \leqslant i \leqslant n$. If $y \in (L_\epsilon(x), x)$ then $\bigl|G^i(y) - G^i(x)\bigr| \leqslant \epsilon$ for all $0 \leqslant i \leqslant n-1$, and moreover $|G^n(y) - 1| \leqslant \epsilon$ and $\bigl|G^{n+1}(y)\bigr| \leqslant \epsilon$.*

(iv) *If $\delta \in (0,1)$ then $|R_\epsilon(x) - x| \leqslant \bigl|R_\epsilon\bigl(G^i(x)\bigr) - G^i(x)\bigr| \leqslant \epsilon$ and $|L_\epsilon(x) - x| \leqslant \bigl|L_\epsilon\bigl(G^i(x)\bigr) - G^i(x)\bigr| \leqslant \epsilon/(1+\epsilon) < \epsilon$ for all $0 \leqslant i \leqslant n-1$.*

[6]This terminology follows [BZ15]

Proof. We will write $\mathbf{b} \coloneqq g(\mathbf{a}) = [a_1, \dots, a_{n-1}, a_n - 1, 1]$ throughout this proof.

(i) Note that $R_\epsilon(x) = G_{\mathbf{a}}(\epsilon)$, $L_\epsilon(x) = G_{\mathbf{b}}(\epsilon)$, and $x = G_{\mathbf{a}}(0) = G_{\mathbf{b}}(0)$. When n is odd, by Proposition 3.13 (iii) $G_{\mathbf{a}}$ is strictly decreasing and $G_{\mathbf{b}}$ is strictly increasing. So $R_\epsilon(x) = G_{\mathbf{a}}(\epsilon) < G_{\mathbf{a}}(0) = x = G_{\mathbf{b}}(0) < G_{\mathbf{b}}(\epsilon) = L_\epsilon(x)$. Similarly, when n is even, $G_{\mathbf{a}}$ is strictly increasing and $G_{\mathbf{b}}$ is strictly decreasing. Then $L_\epsilon(x) = G_{\mathbf{b}}(\epsilon) < G_{\mathbf{b}}(0) = x = G_{\mathbf{a}}(0) < G_{\mathbf{a}}(\epsilon) = R_\epsilon(x)$.

(ii) If $y \in (R_\epsilon(x), x]$, and denoting $z \coloneqq G^n(y) \in [0, \epsilon)$, we have that $y = [a_1, \dots, a_n + z]$, so for all $0 \leqslant i \leqslant n$,

$$\left|G^i(y) - G^i(x)\right| = |G_{\sigma^i(\mathbf{a})}(z) - G_{\sigma^i(\mathbf{a})}(0)| \leqslant \epsilon,$$

since $\left|G'_{\sigma^i(\mathbf{a})}\right| \leqslant 1$, so the first part of (ii) follows.

Now assuming that $y \in (x, L_\epsilon(x))$, and denoting $w \coloneqq G^{n+1}(y) \in (0, \epsilon)$, we see that $y = [a_1, \dots, a_{n-1}, a_n - 1, 1 + w]$, so we have that, for all $0 \leqslant i \leqslant n-1$,

$$\left|G^i(y) - G^i(x)\right| = |G_{\sigma^i(\mathbf{b})}(w) - G_{\sigma^i(\mathbf{b})}(0)| \leqslant \epsilon,$$

since $\left|G'_{\sigma^i(\mathbf{b})}\right| \leqslant 1$. Moreover, $|G^n(y) - 1| = \frac{w}{1+w} < w < \epsilon$, so the second part of (ii) follows.

(iii) The proof is very similar to the one for (ii).

(iv) For each $a \in \mathbb{N}$ and $x, y \in I$, clearly $|G_a(x) - G_a(y)| = \left|\frac{1}{a+x} - \frac{1}{a+y}\right| = \frac{|x-y|}{(a+x)(a+y)} \leqslant |x - y|$. So $\left|R_\delta\big(G^{n-1}(x)\big) - G^{n-1}(x)\right| = |G_{a_n}(\delta) - G_{a_n}(0)| \leqslant \delta$, and for each $0 \leqslant i \leqslant n-1$ we obtain

$$\left|R_\delta\big(G^i(x)\big) - G^i(x)\right| = |G_{\sigma^i(\mathbf{a})}(\delta) - G_{\sigma^i(\mathbf{a})}(0)| \geqslant |G_{\sigma^{i-1}(\mathbf{a})}(\delta) - G_{\sigma^{i-1}(\mathbf{a})}(0)| = \left|R_\delta\big(G^{i-1}(x)\big) - G^{i-1}(x)\right|,$$

and so the first part of (iv) follows.

Similarly, we obtain $\left|L_\delta\big(G^{n-1}(x)\big) - G^{n-1}(x)\right| = |G_{a_n - 1}(1/(1+\delta)) - G_{a_n-1}(1)| \leqslant \delta/(1+\delta) < \delta$ and so for each $1 \leqslant i \leqslant n-1$ we have

$$\left|L_\delta\big(G^i(x)\big) - G^i(x)\right| = |G_{\sigma^i(\mathbf{b})}(\delta) - G_{\sigma^i(\mathbf{b})}(0)| \geqslant |G_{\sigma^{i-1}(\mathbf{b})}(\delta) - G_{\sigma^{i-1}(\mathbf{b})}(0)| = \left|L_\delta\big(G^{i-1}(x)\big) - G^{i-1}(x)\right|,$$

and the second part of (iv) follows. ☐

Lemma 7.4. *Suppose $\alpha \in (0, 1]$ and $x \in I \cap \mathbb{Q}$. There exists $C_x > 0$ such that for all $\nu \in \overline{\mathcal{M}(I, G)}$ and $\phi \in C^{0,\alpha}(I)$,*

$$\langle \nu, \phi \rangle \leqslant \max\{\langle \mu, \phi \rangle : \mu \in \mathcal{M}_x\} + C_x |\phi|_{\alpha, I} \big\langle \nu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle. \tag{7.2}$$

Proof. Denote $p \coloneqq \operatorname{card} \widetilde{\mathcal{O}}(x)$. Suppose $\nu \in \overline{\mathcal{M}(I, G)}$ and $\phi \in C^{0,\alpha}(I)$. Define

$$\eta \coloneqq \max\{\langle \mu, \phi \rangle : \mu \in \mathcal{M}_x\}. \tag{7.3}$$

If $p = 1$ then $x = 0$, $\widetilde{\mathcal{O}}(x) = \{0\}$, and (7.2) holds with $C_x = 1$ because

$$\langle \nu, \phi \rangle \leqslant \langle \nu, \phi(0) + |\phi|_\alpha d(\cdot, 0)^\alpha \rangle = \langle \delta_0, \phi \rangle + |\phi|_\alpha \big\langle \nu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle.$$

If $p \geqslant 2$, then by the ergodic decomposition theorem and the fact that $\mathcal{M}_{\mathrm{irr}}(I, G)$ is dense in $\overline{\mathcal{M}(I, G)}$ (see Lemma 5.3), it suffices to prove (7.2) for every ergodic measure $\nu \in \mathcal{M}_{\mathrm{irr}}(I, G)$. Fixing an arbitrary ergodic $\nu \in \mathcal{M}_{\mathrm{irr}}(I, G)$, the Birkhoff ergodic theorem implies that there exists $a \in I$ with

$$\langle \nu, \phi \rangle = \lim_{k \to +\infty} \frac{1}{k} S_k \phi(w) \quad \text{and} \tag{7.4}$$

$$\big\langle \nu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle = \lim_{k \to +\infty} \frac{1}{k} \sum_{i=0}^{k-1} d\big(G^i(w), \widetilde{\mathcal{O}}(x)\big)^\alpha. \tag{7.5}$$

Claim. There exists $C_x > 0$ and a *transport sequence* $\underline{y} = \{y_i\}_{i=-1}^{+\infty}$ with entries from $\widetilde{\mathcal{O}}(x)$ satisfying

$$\limsup_{k\to+\infty} \frac{1}{n}\sum_{i=0}^{k-1}\phi(y_i) \leqslant \eta \quad \text{and} \tag{7.6}$$

$$\left|G^i(w) - y_i\right| \leqslant C_x^{1/\alpha} d\big(G^i(w), \widetilde{\mathcal{O}}(x)\big) \quad \text{for all } i \in \mathbb{N}_0. \tag{7.7}$$

Note that a consequence of this claim is, by (7.4), (7.7), (7.6), and (7.5), that

$$\begin{aligned}\langle \nu, \phi\rangle = \lim_{k\to+\infty}\frac{1}{k}\sum_{i=0}^{k-1}\phi\big(G^i(w)\big) &\leqslant \limsup_{k\to+\infty}\frac{1}{k}\sum_{i=0}^{k-1}\big(\phi(y_i) + |\phi|_\alpha\left|G^i(w)-y_i\right|^\alpha\big)\\ &\leqslant \limsup_{k\to+\infty}\frac{1}{k}\sum_{i=0}^{k-1}\phi(y_i) + \lim_{k\to+\infty}\frac{1}{k}\sum_{i=0}^{k-1}C_x|\phi|_\alpha d\big(G^i(w),\widetilde{\mathcal{O}}(x)\big)^\alpha\\ &\leqslant \eta + C_x|\phi|_\alpha\big\langle \nu, d\big(\cdot,\widetilde{\mathcal{O}}(x)\big)^\alpha\big\rangle,\end{aligned}$$

so the required inequality (7.2) will hold, and the lemma will follow.

Proof of Claim. Define

$$\delta \coloneqq \Delta\big(\widetilde{\mathcal{O}}(x)\big)/3, \tag{7.8}$$

where $\Delta\big(\widetilde{\mathcal{O}}(x)\big) \coloneqq \min\big\{|y-z| : y, z \in \widetilde{\mathcal{O}}(x),\, y \neq z\big\}$.

Define

$$\epsilon \coloneqq \begin{cases}\delta/(1+\delta) & \text{if } x = 1,\\ \min\{|R_\delta(x)-x|,\,|L_\delta(x)-x|\} & \text{if } x \in (0,1)\cap\mathbb{Q}.\end{cases} \tag{7.9}$$

Define

$$C_x \coloneqq \epsilon^{-\alpha} > 1.$$

The sequence $\{y_i\}_{i=-1}^{+\infty}$ is constructed recursively as follows.

Base step. Define $y_{-1} \coloneqq 0$.

Recursive step. For some $t \in \mathbb{N}_0$, assume that y_{-1}, y_0, …, y_{t-1} are defined.

Case A. Assume that $G^t(w) \in [0, \epsilon)$. Define $y_t \coloneqq 0$. Then by (7.3) and (7.1),

$$\phi(y_t) = \phi(0) \leqslant \eta. \tag{7.10}$$

By (7.9) and (7.8), we have $\left|G^t(w) - y_t\right| < \epsilon \leqslant \delta \leqslant (1/3)\Delta\big(\widetilde{\mathcal{O}}(x)\big)$ and so

$$\left|G^t(w) - y_t\right| = d\big(G^t(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.11}$$

Case B. Assume that $G^t(w) \in (1-\epsilon, 1]$. Define $y_t \coloneqq 1$ and $y_{t+1} = 0$. Then by (7.3) and (7.1),

$$(1/2)(\phi(y_t) + \phi(y_{t+1})) = (1/2)(\phi(1) + \phi(0)) \leqslant \eta. \tag{7.12}$$

By (7.9) and (7.8), we have $\left|G^{t+i}(w) - y_{t+i}\right| \leqslant \delta \leqslant (1/3)\Delta\big(\widetilde{\mathcal{O}}(x)\big)$ for $i \in \{0,\,1\}$ and so for $i \in \{0,\,1\}$,

$$\left|G^{t+i}(w) - y_{t+i}\right| = d\big(G^{t+i}(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.13}$$

Case C. Assume that $G^t(w) \in (z-\epsilon, z+\epsilon)$ for some $z \in \widetilde{\mathcal{O}}(x) \smallsetminus \{0,\,1\}$. Suppose $x = [a_1, \dots, a_{p-1}] \in R_{p-1}$ and $z = [a_{p-m}, \dots, a_{p-1}] \in R_m$ for some $0 \leqslant m \leqslant p-2$. Let us write $\mathbf{a} \coloneqq (a_1, \dots, a_{p-2}, a_{p-1})$ and $\mathbf{b} \coloneqq (a_1, \dots, a_{p-2}, a_{p-1}-1, 1)$.

Subcase (i). Assume that m is odd. By (7.9) and Lemma 7.3 (i)(iv), we have $(z-\epsilon, z+\epsilon) \subseteq (R_\delta(z), L_\delta(z))$.

If $G^t(w) \in (R_\delta(z), z]$, we define $y_{t+i} \coloneqq G^i(z)$ for all $0 \leqslant i \leqslant m$. Then by (7.3) and (7.1),

$$(m+1)^{-1}S_{m+1}\phi(y_t) = \langle \mu_{\sigma^{p-m-1}(\mathbf{a})}, \phi\rangle \leqslant \eta. \tag{7.14}$$

By (7.9), (7.8), and Lemma 7.3 (ii), we see that for $0 \leqslant i \leqslant m$,

$$\big|G^{t+i}(w) - y_{t+i}\big| = d\big(G^{t+i}(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.15}$$

When $G^t(w) \in (z, L_\delta(z))$, we define $y_{t+i} \coloneqq G^i(z)$ for all $0 \leqslant i \leqslant m-1$, and $y_{t+m} \coloneqq 1$, and $y_{t+m+1} \coloneqq 0$. Then by (7.3) and (7.1),

$$(m+2)^{-1} S_{m+2}\phi(y_t) = \langle \mu_{\sigma^{p-m-1}(\mathbf{b})}, \phi \rangle \leqslant \eta. \tag{7.16}$$

By (7.9), (7.8), and Lemma 7.3 (ii), we get for $0 \leqslant i \leqslant m+1$ that

$$\big|G^{t+i}(w) - y_{t+i}\big| = d\big(G^{t+i}(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.17}$$

Subcase (ii). Assume that m is even. Using (7.9) and Lemma 7.3 (i)(iv), we see that $(z-\epsilon, z+\epsilon) \subseteq (L_\delta(z), R_\delta(z))$.

When $G^t(w) \in [z, R_\delta(z))$, we define $y_{t+i} \coloneqq G^i(z)$ for all $0 \leqslant i \leqslant m$. Then by (7.3) and (7.1),

$$(m+1)^{-1} S_{m+1}\phi(y_t) = \langle \mu_{\sigma^{p-m-1}(\mathbf{a})}, \phi \rangle \leqslant \eta. \tag{7.18}$$

By (7.9), (7.8), and Lemma 7.3 (iii), we get that for $0 \leqslant i \leqslant m$,

$$\big|G^{t+i}(w) - y_{t+i}\big| = d\big(G^{t+i}(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.19}$$

When $G^t(w) \in (L_\delta(z), z)$, we define $y_{t+i} \coloneqq G^i(z)$ for all $0 \leqslant i \leqslant m-1$ and $y_{t+m} \coloneqq 1$, and $y_{t+m+1} \coloneqq 0$. Then by (7.3) and (7.1),

$$(m+2)^{-1} S_{m+2}\phi(y_t) = \langle \mu_{\sigma^{p-m-1}(\mathbf{b})}, \phi \rangle \leqslant \eta. \tag{7.20}$$

By (7.9), (7.8), and Lemma 7.3 (iii), we get for $0 \leqslant i \leqslant m+1$,

$$\big|G^{t+i}(w) - y_{t+i}\big| = d\big(G^{t+i}(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.21}$$

Case D. Assume that $d(G^t(w), \widetilde{\mathcal{O}}) \geqslant \epsilon$. Define $y_t \coloneqq 0$. Then by (7.3) and (7.1),

$$\phi(y_t) = \phi(0) \leqslant \eta. \tag{7.22}$$

By the definition of C_x and (7.9), we get

$$\big|G^t(w) - y_t\big| \leqslant C_x^{1/\alpha}\epsilon \leqslant C_x^{1/\alpha} d\big(G^t(w), \widetilde{\mathcal{O}}(x)\big). \tag{7.23}$$

The recursive step is now complete, and combining (7.10), (7.12), (7.14), (7.16), (7.18), (7.20) and (7.22) gives (7.6). Combining (7.11), (7.13), (7.15), (7.17), (7.19), (7.21), and (7.23) gives (7.7), thereby completing the proof of the claim. □

Lemma 7.5. *Suppose $x \in I \cap \mathbb{Q}$ and $\alpha \in (0,1]$. Then $\big\langle \mu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle > 0$ for all $\mu \in \overline{\mathcal{M}(I,G)} \smallsetminus \operatorname{conv}(\mathcal{M}_x)$.*

Proof. By Theorem 1.1 and the ergodic decomposition theorem, it suffices to show that

$$\big\langle \mu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle > 0 \tag{7.24}$$

for all measures μ that either are ergodic and in $\mathcal{M}_{\mathrm{irr}}(I,G)$, or are FCF measures not contained in $\mathcal{M}_x$.

Assuming that $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$ is ergodic, if (7.24) were false, then $\big\langle \mu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle = 0$ would give $\operatorname{supp}\mu \subseteq \widetilde{\mathcal{O}}(x)$, which contradicts the fact that $\mu(I \smallsetminus \mathbb{Q}) = 1$.

If, on the other hand, $\mathbf{a} \in \mathbb{N}^*$ and $\mu = \mu_{\mathbf{a}}$ is an FCF measure not contained in $\mathcal{M}_x$, then $g(\mathbf{a}) \notin \widetilde{\mathcal{O}}(x)$, since otherwise $\mathbf{a}$ is equal to the continued fraction expansion of some point in $\widetilde{\mathcal{O}}(x)$, contradicting the definition of $\mathcal{M}_x$. So $\operatorname{supp}\mu_{\mathbf{a}}$ is not contained in $\widetilde{\mathcal{O}}(x)$, so we get $\big\langle \mu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle > 0$, and the result follows. □

With the preceding lemmas in hand, we can now prove Theorem 1.4.

Proof of Theorem 1.4. Fix arbitrary $\phi \in \mathfrak{R}^\alpha(G)$ and real numbers $s > 0$ and $t > 0$.

When $\delta_0 \in \mathcal{M}_{\max}(T, \phi)$, denote $x := 0$. When $\delta_0 \notin \mathcal{M}_{\max}(T, \phi)$, let n be the smallest integer such that there exists $\mathbf{a} \in \mathbb{N}^n$ with $\mu_{\mathbf{a}} \in \mathcal{M}^*_{\max}(G, \phi)$, and choose $\mathbf{a} = (a_1, a_2, \dots, a_n) \in \mathbb{N}^n$ satisfying $\mu_{\mathbf{a}} \in \mathcal{M}^*_{\max}(G, \phi)$, denote $x := [a_1, \dots, a_n]$ and $\mathbf{b} := f(\mathbf{a})$.

Case A. Assume that $\mathbf{a} \in \mathcal{A}_n$ or $\delta_0 \in \mathcal{M}_{\max}(T, \phi)$. Define

$$\Phi_s := \phi - sd(\cdot, \mathcal{O}(x))^\alpha. \tag{7.25}$$

By (7.25), we have $Q(G, \phi) \geqslant Q(G, \Phi_s)$. Combining this with the fact that $\langle \mu_{\mathbf{a}}, \Phi_s \rangle = \langle \mu_{\mathbf{a}}, \phi \rangle = Q(G, \phi)$, we obtain $Q(G, \phi) \leqslant Q(G, \Phi_s)$ and $\mu_{\mathbf{a}} \in \mathcal{M}^*_{\max}(G, \Phi_s)$. By the choice of n, we get

$$\langle \mu_{\mathbf{a}}, \phi \rangle > \langle \nu, \phi \rangle,$$

for each $\nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}, \mu_{\mathbf{b}}, \mu_{\sigma(\mathbf{b})}\}$. Since $1 \notin \mathcal{O}(x)$, $1 \in \mathcal{O}_{\mathbf{b}}$, and $1 \in \mathcal{O}_{\sigma(\mathbf{b})}$ (see Remark 5.2), we get $\langle \mu_{\mathbf{a}}, \Phi_s \rangle = \langle \mu_{\mathbf{a}}, \phi \rangle \geqslant \langle \mu_{\mathbf{b}}, \phi \rangle > \langle \mu_{\mathbf{b}}, \Phi_s \rangle$ and $\langle \mu_{\mathbf{a}}, \Phi_s \rangle = \langle \mu_{\mathbf{a}}, \phi \rangle \geqslant \langle \mu_{\sigma(\mathbf{b})}, \phi \rangle > \langle \mu_{\sigma(\mathbf{b})}, \Phi_s \rangle$, and so

$$\langle \mu_{\mathbf{a}}, \Phi_s \rangle > \langle \nu, \Phi_s \rangle,$$

for each $\nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}\}$. So we can define

$$\delta_s := \langle \mu_{\mathbf{a}}, \Phi_s \rangle - \max\{\langle \nu_{\mathbf{a}}, \Phi_s \rangle : \nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}\}\} > 0. \tag{7.26}$$

Define

$$\Phi_{t,s} := \Phi_s - td\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha = \phi - sd(\cdot, \mathcal{O}(x))^\alpha - td\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha.$$

Then for each $\nu \in \overline{\mathcal{M}(I, G)} \smallsetminus \mathrm{conv}(M_x)$, when $|\psi|_\alpha < t/C_x$ (where $C_x > 0$ is the constant from Lemma 7.4), by Lemmas 7.4 and 7.5,

$$\begin{aligned}\langle \nu, \Phi_{t,s} + \psi \rangle &= \langle \nu, \Phi_s \rangle + \langle \nu, \psi \rangle - \big\langle \nu, td\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle \\ &\leqslant \max_{\mu \in \mathcal{M}_x} \langle \mu, \Phi_s \rangle + \max_{\mu \in \mathcal{M}_x} \langle \mu, \psi \rangle + (C_x|\psi|_\alpha - t)\big\langle \nu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle < \max_{\mu \in \mathcal{M}_x} \langle \mu, \Phi_{t,s} + \psi \rangle.\end{aligned}$$

For each $\nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}\}$, when $\|\psi\|_\alpha < \delta_s/4$, by (7.26),

$$\begin{aligned}\langle \nu, \Phi_{t,s} + \psi \rangle &= \langle \nu, \Phi_s \rangle + \langle \mu_{\mathbf{a}}, \psi \rangle - \langle \mu_{\mathbf{a}}, \psi \rangle + \langle \nu, \psi \rangle \\ &< \langle \mu_{\mathbf{a}}, \Phi_s \rangle - (1/2)\delta_s + \langle \mu_{\mathbf{a}}, \psi \rangle + 2\|\psi\|_\infty < \langle \mu_{\mathbf{a}}, \Phi_{t,s} + \psi \rangle.\end{aligned}$$

So by Theorem 1.1, $\mathcal{M}^*_{\max}(G, \Phi_{t,s} + \psi) = \{\mu_{\mathbf{a}}\}$ when $\|\psi\|_\alpha < \min\{t/C_x, \delta_s/4\}$, and consequently, $\Phi_{t,s} \in \mathfrak{R}^\alpha(G) \cap \mathrm{Lock_F}^\alpha(G)$.

Case B. Assume that $\mathbf{a} \in \mathcal{B}_n$. Then $a_n = 1$. By the choice of n, we get

$$\langle \mu_{\mathbf{a}}, \phi \rangle > \langle \nu, \phi \rangle$$

for each $\nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}\}$. So we can define

$$\delta := \langle \mu_{\mathbf{a}}, \phi \rangle - \max\{\langle \nu, \phi \rangle : \nu \in \mathcal{M}_x \smallsetminus \{\mu_{\mathbf{a}}\}\} > 0. \tag{7.27}$$

For each $t > 0$, define

$$\phi_t := \phi - td\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha.$$

Then for each $\nu \in \overline{\mathcal{M}(I, G)} \smallsetminus \mathrm{conv}(M_x)$, when $|\psi|_\alpha < t/C_x$, by Lemmas 7.4 and 7.5,

$$\begin{aligned}\langle \nu, \phi_t + \psi \rangle &= \langle \nu, \phi \rangle + \langle \nu, \psi \rangle - \big\langle \nu, td\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle \\ &\leqslant \max_{\mu \in \mathcal{M}_x} \langle \mu, \phi \rangle + \max_{\mu \in \mathcal{M}_x} \langle \mu, \psi \rangle + (C_x|\psi|_\alpha - t)\big\langle \nu, d\big(\cdot, \widetilde{\mathcal{O}}(x)\big)^\alpha \big\rangle < \max_{\mu \in \mathcal{M}_x} \langle \mu, \phi_t + \psi \rangle.\end{aligned}$$

For each $\nu \in \mathrm{conv}(M_x) \smallsetminus \{\mu_{\mathbf{a}}\}$, when $\|\psi\|_\infty < \delta/4$, by (7.27),

$$\begin{aligned}\langle \nu, \phi_t + \psi \rangle &= \langle \nu, \phi_t \rangle + \langle \mu_{\mathbf{a}}, \psi \rangle - \langle \mu_{\mathbf{a}}, \psi \rangle + \langle \nu, \psi \rangle \\ &< \langle \mu_{\mathbf{a}}, \phi \rangle - (1/2)\delta + \langle \mu_{\mathbf{a}}, \psi \rangle + 2\|\psi\|_\infty < \langle \mu_{\mathbf{a}}, \phi_t + \psi \rangle.\end{aligned}$$

So by Theorem 1.1, $\mathcal{M}^*_{\max}(G, \phi_t + \psi) = \{\mu_{\mathbf{a}}\}$ when $\|\psi\|_\alpha < \min\{t/C_x, \delta/4\}$, and consequently, $\phi_t \in \mathfrak{R}^\alpha(G) \cap \mathrm{Lock_F}^\alpha(G)$.

From the preceding two cases we deduce that $\mathfrak{R}^\alpha(G) \cap \mathrm{Lock_F}^\alpha(G)$ is dense in $\mathfrak{R}^\alpha(G)$.

It follows immediately from the definition that $\mathrm{Lock_F}^\alpha(G) \cap \mathfrak{R}^\alpha(G)$ is open in $C^{0,\alpha}(I)$. Therefore the theorem is proved. □

7.2. **A counterexample to TPO for Hölder continuous potentials.**

Example 7.6. For $\alpha \in (0,1]$, let $\phi \in C^{0,\alpha}(I)$ satisfy $\phi(0) = -1$, $\phi(1) = 1$, $\phi(1/3) = \phi(3/4) = -2$, and be affine on each of the intervals $[0,1/3]$, $[1/3,3/4]$, $[3/4,1]$. More precisely, ϕ is given by

$$\phi(x) \coloneqq \begin{cases} -3x-1 & \text{if } 0 \leqslant x \leqslant 1/3, \\ -2 & \text{if } 1/3 \leqslant x \leqslant 3/4, \\ 12x-11 & \text{if } 3/4 \leqslant x \leqslant 1. \end{cases}$$

We claim that $Q(G,\phi) = 0$, $\phi \in \mathfrak{R}^\alpha(G)$, and $\phi \notin \mathfrak{E}^\alpha(G)$.

Using that $(1/2)(\phi(0)+\phi(1)) = 0$, and that $(1/2)(\delta_0+\delta_1) \in \overline{\mathcal{M}(I,G)}$ (see Lemma 5.3), we see that $Q(G,\phi) \geqslant 0$.

Fix $m \in \mathbb{N}$ and $x = [a_1,\dots,a_n,\dots] \in E_m$. We will recursively construct a sequence $\{x_k\}_{k\in\mathbb{N}}$ in $\mathcal{O}(x)$ and a sequence $\{n_k\}_{k\in\mathbb{N}}$ in $\{1,2\}$.

Base step. If $a_1 \geqslant 2$ or both $a_1 = 1$ and $a_2 \leqslant 3$ hold, define $n_1 \coloneqq 1$. Then by the definition of ϕ, we have $\phi(x_1) \leqslant -1$.

If $a_1 = 1$ and $3 \leqslant a_2 \leqslant m$, define $n_1 \coloneqq 2$. Then by the definition of ϕ,

$$\begin{aligned} S_2\phi(x_1) &= (1/2)(\phi(x_1)+\phi(x_2)) \leqslant (1/2)(\phi([1,m+1]) + \phi(1/(m+2))) \\ &= (1/2)\Big(\frac{12(m+1)}{(m+2)} - 11 - \frac{3}{m+2} - 1\Big) = -\frac{15}{2(m+2)}. \end{aligned}$$

Recursive step. Assume that for some $t \in \mathbb{N}$, $\{x_k\}_{k=1}^t$ and $\{n_k\}_{k=1}^t$ are defined. Denote $N_t \coloneqq \sum_{i=1}^t n_i$.

If $a_{N_t+1} \geqslant 2$ or both $a_{N_t+1} = 1$ and $a_{N_t+2} \leqslant 3$ hold, define $n_{t+1} \coloneqq 1$. Then by the definition of ϕ, we have $\phi(x_{N_t+1}) \leqslant -1$.

If $a_{N_t+1} = 1$ and $3 \leqslant a_{N_t+2} \leqslant m$, define $n_{t+1} \coloneqq 2$. Then by the definition of ϕ,

$$\begin{aligned} S_2\phi(x_{N_t+1}) &= (1/2)(\phi(x_{N_t+1})+\phi(x_{N_t+2})) \leqslant (1/2)(\phi([1,m+1]) + \phi(1/(m+2))) \\ &= (1/2)\Big(\frac{12(m+1)}{(m+2)} - 11 - \frac{3}{m+2} - 1\Big) = -\frac{15}{2(m+2)}. \end{aligned}$$

Denote $N_k \coloneqq \sum_{i=1}^k n_k$. Hence, we get

$$\liminf_{n\to+\infty} \frac{1}{n} S_n\phi(x) \leqslant \limsup_{k\to+\infty} \frac{1}{N_k} S_{N_k}\phi(x) \leqslant \max\Big\{-1, -\frac{15}{2(m+2)}\Big\}.$$

Combining this with [Je19, Proposition 2.2] gives $Q_m(G,\phi) \leqslant \max\big\{-1, -\frac{15}{2(m+2)}\big\}$. Letting m tend to infinity, from Proposition 4.7 (ii) we see that $Q(G,\phi) \leqslant 0$. Consequently we conclude that $Q(G,\phi) = 0$ and $Q(G,\phi) > Q_m(G,\phi)$ for all $m \in \mathbb{N}$. So $\phi \in \mathfrak{R}^\phi(G)$ but $\phi \notin \mathfrak{E}^\alpha(G)$.

By using methods similar to the proof of Theorem 1.4, we can prove that

$$\Phi = \phi - d(\cdot,\{0,1\})^\alpha = \phi - (1/2 - |x-1/2|)^\alpha \in \mathfrak{R}^\alpha(G) \cap \mathrm{Lock_F}^\alpha(G).$$

More precisely, let C_1 be the constant from Lemma 7.4, by Case B in the proof of Theorem 1.4, we conclude that $\mathcal{M}^*_{\max}(G,\,\Phi+\psi) = \mathcal{M}^*_{\max}(G,\,\Phi) = \{(1/2)(\delta_0+\delta_1)\}$ when $\|\psi\|_\alpha < \min\big\{\frac{1}{C_1}, \frac{1}{4}\big\}$. In particular, this implies that the Gauss map G does not enjoy the typical periodic optimization property in the space $C^{0,\alpha}(I)$.

Recall E_m from Subsection 3.5. For the set $\mathfrak{Z}^\alpha(G)$ (cf. Definition 7.1), we have the following:

Proposition 7.7. *Suppose $\alpha \in (0,1]$ and $\phi \in \mathfrak{Z}^\alpha(G)$. There exists $\mu \in \mathcal{M}_{\max}(G,\,\phi)$ satisfying $\mu(\mathbb{Q}) = 0$, $0 \in \operatorname{supp}\mu$, and $\operatorname{supp}\mu \cap E_m = \emptyset$ for all $m \in \mathbb{N}$.*

Proof. Since $\overline{\mathcal{M}(I,G)}$ is weak* compact, $\mathcal{M}^*_{\max}(G,\,\phi) \neq \emptyset$. Fix $\mu \in \mathcal{M}^*_{\max}(G,\phi)$. From the fact that $\mu \notin \mathfrak{R}^\alpha(I)$, and Theorem 1.1, we get $\mu \in \mathcal{M}_{\mathrm{irr}}(I,G)$ and consequently $\mu(\mathbb{Q}) = 0$. From the fact that $\phi \notin \mathfrak{E}^\alpha(I)$ and Lemma 7.8, we get $0 \in \operatorname{supp}\mu$ and $\operatorname{supp}\mu \cap E_m = \emptyset$ for all $m \in \mathbb{N}$. □

7.3. **TPO for essentially compact potentials.** The goal of this subsection is to prove that the set $\mathfrak{E}^\alpha(G)$ is contained in the closure of $\mathscr{P}^\alpha(G)$.

Theorem 1.3 (TPO for essentially compact potentials). *For $\alpha \in (0,1]$, the set $\mathrm{Lock}^\alpha(G)$ contains an open dense subset of $\mathfrak{E}^\alpha(G)$ (in the α-Hölder topology).*

Recall the maximizing set defined in Definition 6.8. To prove Theorem 1.3, we first require:

Lemma 7.8. *Suppose $\alpha \in (0,1]$ and $\phi \in C^{0,\alpha}(I)$. The following are equivalent:*

(i) $\phi \in \mathfrak{E}^\alpha(G)$.

(ii) *There exists $\mu \in \mathcal{M}_{\max}(G,\phi)$ with $\operatorname{supp}\mu \subseteq E_m$ for some $m \in \mathbb{N}$.*

(iii) *$\mathcal{K}(\phi) \cap \Sigma_m \neq \emptyset$ for some $m \in \mathbb{N}$.*

Proof. (i) $\Longrightarrow$ (ii): Assume that $\phi \in \mathfrak{E}^\alpha(G)$. Then there exists $m \in \mathbb{N}$ with $Q(G,\phi) = Q_m(G,\phi)$ by Definition 7.1. Since $G|_{E_m}$ is continuous, and E_m is compact (see Lemma 3.27), there exists $\mu \in \mathcal{M}_{\max}(G|_{E_m}, \phi|_{E_m}) \neq \emptyset$. Since μ can be seen as a measure in $\mathcal{M}(I,G)$, we obtain that $\mu \in \mathcal{M}_{\max}(G,\phi)$ satisfies $\operatorname{supp}\mu \subseteq E_m$.

(ii) $\Longrightarrow$ (iii): Assume that there exists $\mu \in \mathcal{M}_{\max}(G,\phi)$ satisfying $\operatorname{supp}\mu \subseteq E_m$ for some $m \in \mathbb{N}$. By Proposition 4.6 (ii), there exists $\nu \in \mathcal{M}_{\max}\big(\sigma_{\widehat{\Sigma}}, \phi \circ \widehat{\pi}\big)$ with $\widehat{\pi}_*(\nu) = \mu$. Since $\widehat{\pi}(\operatorname{supp}\nu) = \operatorname{supp}\mu$ (see e.g. [Ak93, p. 156]), $\widehat{\pi}^{-1}(E_m) = \Sigma_m$, and $\operatorname{supp}\mu \subseteq E_m$, we see that $\operatorname{supp}\nu \subseteq \Sigma_m$. By Lemma 6.9 (iv), $\operatorname{supp}\nu \subseteq \mathcal{K}(\phi)$.

(iii) $\Longrightarrow$ (i): Assume that $\mathcal{K}(\phi) \cap \Sigma_m \neq \emptyset$ for some $m \in \mathbb{N}$. Since $\sigma(\mathcal{K}(\phi)) \subseteq \mathcal{K}(\phi)$, $\sigma(\Sigma_m) \subseteq \Sigma_m$, and both $\mathcal{K}(\phi)$ and Σ_m are compact, there exists $\nu \in \mathcal{M}\big(\widehat{\Sigma}, \sigma_{\widehat{\Sigma}}\big)$ with $\operatorname{supp}\nu \subseteq \Sigma_m$. By Lemma 6.9 (iv), we see that $\nu \in \mathcal{M}_{\max}\big(\sigma_{\widehat{\Sigma}}, \phi \circ \widehat{\pi}\big)$. By Proposition 4.6 (ii), $\widehat{\pi}_*(\nu) \in \mathcal{M}_{\max}(G,\phi)$. Combining this with the fact that $\operatorname{supp}\widehat{\pi}_*(\nu) \subseteq E_m$ (see e.g. [Ak93, p. 156]), we obtain that $\phi \in \mathfrak{E}^\alpha(G)$, as required. □

Notation 7.9. For a compact metric space (X,d), a map $T\colon X \to X$, and a periodic orbit $\mathcal{O}$ of T, the corresponding *gap* is defined as

$$\Delta(\mathcal{O}) = \Delta^d(\mathcal{O}) \coloneqq \min\{d(x,y) : x,\, y \in \mathcal{O},\, x \neq y\}, \tag{7.28}$$

with the convention that $\min\emptyset = +\infty$. For $r,\,\theta > 0$, define the *(r,θ)-gap* of $\mathcal{O}$ by

$$\Delta_{r,\,\theta}(\mathcal{O}) = \Delta^d_{r,\,\theta}(\mathcal{O}) \coloneqq \min\{r,\, \theta \cdot \Delta(\mathcal{O})\}. \tag{7.29}$$

We will need the following closing lemma, which was first introduced in [HLMXZ25].

Lemma 7.10. *Let G be the Gauss map, d the Euclidean metric on I, $\alpha \in (0,1]$, and $m \in \mathbb{N}$. Let $\mathcal{K}$ be a nonempty compact subset of E_m with $G(\mathcal{K}) \subseteq \mathcal{K}$. For $r > 0$, $\theta > 0$, and $\tau > 0$, there exists a periodic orbit $\mathcal{O} \subseteq E_m$ of G with*

$$\sum_{x \in \mathcal{O}} d(x,\mathcal{K})^\alpha \leqslant \tau \cdot (\Delta_{r,\,\theta}(\mathcal{O}))^\alpha.$$

Proof. By Lemma 3.27, $G|_{E_m}$ is open, Lipschitz, and distance-expanding, so the result is immediate from [HLMXZ25, Proposition 2.1]. □

Proof of Theorem 1.3. By Proposition A.2, it suffices to prove that $\mathfrak{E}^\alpha(G)$ is contained in the closure of $\mathscr{P}^\alpha(G)$.

For each periodic orbit $\mathcal{O}$ of G, define the measure $\mu_{\mathcal{O}}$ by

$$\mu_{\mathcal{O}} \coloneqq \frac{1}{\operatorname{card}\mathcal{O}} \sum_{x\in\mathcal{O}} \delta_x \in \mathcal{M}(I,G). \tag{7.30}$$

Fix $\phi \in \mathfrak{E}^\alpha(G)$ with no ϕ-maximizing measure in $\mathcal{M}_{\max}(G,\phi)$ supported on a periodic orbit of G. Let $u_\phi \in C^{0,\alpha}(I)$ be the calibrated sub-action for ϕ and G (i.e., a fixed point of $\mathcal{L}_{\overline{\phi}}$) from Proposition 6.6. Define

$$\widetilde{\phi} \coloneqq \phi - Q(G,\phi) + u_\phi - u_\phi\circ G, \tag{7.31}$$

where $Q(G,\phi)$ is defined in (1.2). Then by Theorem 1.2,

$$\widetilde{\phi}(x) \leqslant 0 \quad \text{ for all } x \in I. \tag{7.32}$$

By Lemma 7.8, there exists $\mu \in \mathcal{M}_{\max}(G,\phi)$ and $m\in\mathbb{N}$ satisfying $\operatorname{supp}\mu \subseteq E_m$. Let $\eta_m > 0$ and $\lambda_m > 1$ be the constants defined in Lemma 3.27 and denote $\mathcal{K} \coloneqq \operatorname{supp}\mu$. Since $\widetilde{\phi}$ is continuous on $\mathcal{K}$ (note that $\mathcal{K} \subseteq E_m \subseteq I \smallsetminus \mathbb{Q}$), by (7.32) we obtain that

$$\widetilde{\phi}|_{\mathcal{K}} \equiv 0. \tag{7.33}$$

Without loss of generality, assume that $\mathcal{K}$ contains no periodic orbits of G. Recall (cf. (3.15)) that the closed η_m-neighbourhood of E_m is denoted by

$$F_m \coloneqq \overline{B}_d^{\eta_m}(E_m) = \{x\in I : d(x,E_m) \leqslant \eta_m\}. \tag{7.34}$$

Now ϕ, $u_\phi \in C^{0,\alpha}(I)$ (see Proposition 6.6 (ii)), and $G|_{F_m}$ is Lipschitz (see Lemma 3.27 (ii)), with Lipschitz constant $|G|_{\mathrm{LIP},\,F_m} \coloneqq |G|_{1,\,F_m}$, so using (7.31) we see that $\widetilde{\phi}|_{F_m} \in C^{0,\alpha}(F_m)$. Let us write

$$L_1 \coloneqq \bigl|\widetilde{\phi}\bigr|_{\alpha,\,F_m}. \tag{7.35}$$

Fix $\epsilon \in (0,1)$, and define constants

$$L_2 \coloneqq |G|_{\mathrm{LIP},\,F_m}, \tag{7.36}$$

$$r \coloneqq \eta_m, \tag{7.37}$$

$$\theta \coloneqq \min\{1/3,\ 1/(3L_2)\}, \tag{7.38}$$

$$L_3 \coloneqq \frac{L_1+1}{1-\lambda_m^{-\alpha}} > 0, \tag{7.39}$$

$$\tau \coloneqq \min\biggl\{1,\ \frac{\epsilon}{2L_1},\ \frac{\epsilon}{(1+L_3\epsilon^{-1})L_1}\biggr\} \leqslant 1. \tag{7.40}$$

By Lemma 7.10, there exists a periodic orbit $\mathcal{O} \subseteq E_m$ of G, of period $p \coloneqq \operatorname{card}\mathcal{O}$, satisfying

$$\sum_{x\in\mathcal{O}} d(x,\mathcal{K})^\alpha \leqslant \tau\cdot(\Delta_{r,\theta}(\mathcal{O}))^\alpha. \tag{7.41}$$

Define functions

$$\phi' \coloneqq \phi - Q(G,\phi) - \epsilon d(\cdot,\mathcal{O})^\alpha \in C^{0,\alpha}(I) \text{ and} \tag{7.42}$$

$$\psi \coloneqq \widetilde{\phi} - \epsilon d(\cdot,\mathcal{O})^\alpha = \phi' - Q(G,\phi) + u_\phi - u_\phi\circ G. \tag{7.43}$$

By (7.42), (7.43), and Lemma 6.2,

$$Q(G,\phi') = Q(G,\psi) \text{ and } \mathcal{M}_{\max}(G,\phi') = \mathcal{M}_{\max}(G,\psi). \tag{7.44}$$

Claim. The measure $\mu_{\mathcal{O}}$ (cf. (7.30)) belongs to $\mathcal{M}_{\max}(G,\psi)$, i.e., $Q(G,\psi)=\gamma$, where

$$\gamma \coloneqq \int \psi\,\mathrm{d}\mu_{\mathcal{O}} = \frac{1}{p}\sum_{x\in\mathcal{O}}\psi(x) = \frac{1}{p}\sum_{x\in\mathcal{O}}\widetilde{\phi}(x) < 0. \tag{7.45}$$

Note that the equality $\frac{1}{p}\sum_{x\in\mathcal{O}}\psi(x) = \frac{1}{p}\sum_{x\in\mathcal{O}}\widetilde{\phi}(x)$ in (7.45) follows from (7.43), whereas the inequality in (7.45) follows from (7.32) and the assumption that $\phi \notin \mathscr{P}^\alpha(G)$.

Note that if the claim holds then (7.44) gives that $\phi' \in \mathscr{P}^\alpha(G)$, and since ϵ can be chosen arbitrarily small we see that ϕ belongs to the closure of $\mathscr{P}^\alpha(G)$ in $C^{0,\alpha}(I)$, which is the required conclusion of Theorem 1.3.

So to prove Theorem 1.3 it suffices to establish the claim. From the definitions of γ (cf. (7.45)) and $Q(G,\psi)$ (cf. (1.2)), we see that $Q(G,\psi) \geqslant \gamma$, so it remains to show that $Q(G,\psi) = Q(G,\phi') \leqslant \gamma$. By Proposition 4.7 (ii), we only need to prove that for all $x \in I \smallsetminus \mathbb{Q}$,

$$\liminf_{n\to+\infty}\frac{1}{n}S_n\phi'(x) = \liminf_{n\to+\infty}\frac{1}{n}S_n\psi(x) \leqslant \gamma, \tag{7.46}$$

where the first identity follows from (7.42), (7.43), and the fact that u_ϕ is bounded (see Proposition 6.6 (i)).

Fix $x \in I \smallsetminus \mathbb{Q}$. In the remainder of this proof, we will divide the orbit of x into segments such that the average on each segment is less than γ, i.e., we will recursively construct a sequence $\{x_k\}_{k\in\mathbb{N}}$ in $\mathcal{O}(x)$ and a sequence $\{n_k\}_{k\in\mathbb{N}}$ in $\mathbb{N}$ satisfying $x_{k+1} = G^{n_k}(x_k)$ and $S_{n_k}\psi(x_k) \leqslant n_k\gamma$.

We first observe that by (7.45), (7.32), (7.33) (7.35), and (7.41),

$$p|\gamma| = \sum_{x\in\mathcal{O}}\bigl|\widetilde{\phi}(x) - 0\bigr| \leqslant \sum_{x\in\mathcal{O}} L_1 d(x,\mathcal{K})^\alpha \leqslant L_1\tau\cdot(\Delta_{r,\theta}(\mathcal{O}))^\alpha. \tag{7.47}$$

By (7.47), (7.40), and (7.29),

$$\rho := \epsilon^{-1/\alpha}|\gamma|^{1/\alpha} \leqslant (L_1\tau/\epsilon)^{1/\alpha}\Delta_{r,\theta}(\mathcal{O}) < \Delta_{r,\theta}(\mathcal{O}) \leqslant r. \tag{7.48}$$

Let us denote $U := \overline{B}^\rho_d(\mathcal{O}) = \{x\in I : d(x,\mathcal{O}) \leqslant \rho\}$ (cf. Section 2).

Base step. Define $x_1 := x$.

Recursive step. Assume that for some $t\in\mathbb{N}$, $\{x_k\}_{k=1}^t$ and $\{n_k\}_{k=1}^{t-1}$ are defined. We now divide our discussion into three cases, the third of which requires some delicate analysis.

Case A. Assume $x_t \in \mathcal{O}$. Then define $n_t := p$ and $x_{t+1} := G^{n_t}(x_t) = x_t$. Thus, by (7.45) and (7.43), we have

$$S_{n_t}\psi(x_t) = n_t\gamma. \tag{7.49}$$

Case B. Assume $x_t \notin U$. Then define $n_t := 1$ and $x_{t+1} := G(x_t)$, so that combining (7.43), (7.32), and (7.48) gives

$$S_{n_t}\psi(x_t) = \psi(x_t) = \widetilde{\phi}(x_t) - \epsilon d(x_t,\mathcal{O})^\alpha \leqslant -\epsilon d(x_t,\mathcal{O})^\alpha < -\epsilon\rho^\alpha = \gamma. \tag{7.50}$$

Case C. Assume $x_t \in U \smallsetminus \mathcal{O}$, then $0 < d(x_t,\mathcal{O}) \leqslant \rho$. By (7.48), (7.29), and (7.38), $\rho \leqslant \Delta_{r,\theta}(\mathcal{O}) \leqslant \frac{1}{3}\Delta(\mathcal{O})$. So by (7.28), there is a unique point $y\in\mathcal{O}$ which is closest to x_t among points in the periodic orbit $\mathcal{O}$. By (7.29) and (7.48), $|x_t - y| \leqslant \rho < \Delta_{r,\theta}(\mathcal{O}) \leqslant r$.

Let $N\in\mathbb{N}$ be the smallest positive integer satisfying

$$\bigl|G^N(x_t) - G^N(y)\bigr| > \Delta_{r,\theta}(\mathcal{O}). \tag{7.51}$$

Such a positive integer always exists; otherwise, we have $|G^m(x_t) - G^m(y)| \leqslant \Delta_{r,\theta}(\mathcal{O}) \leqslant r = \eta_m$ for all $m\in\mathbb{N}$ by (7.29) and (7.37). Then for each $m\in\mathbb{N}$, by Lemma 3.27 (i), $G^m(x_t)$ and $G^m(y)$ are contained in the same interval $(1/(l+1), 1/l)$ for some $l\in\mathbb{N}$. So x_t and y have the same continued fraction digits, and consequently $x_t = y \in \mathcal{O}$, contradicting the assumption that $x_t \notin \mathcal{O}$.

From the definition of N, together with (7.29) and (7.38), we see that for each $0 \leqslant j \leqslant N-1$,

$$d\bigl(G^j(x_t),\mathcal{O}\bigr) \leqslant \bigl|G^j(x_t) - G^j(y)\bigr| \leqslant \Delta_{r,\theta}(\mathcal{O}) \leqslant \Delta(\mathcal{O})/3. \tag{7.52}$$

By (7.52), (7.29), and (7.37), for each $0 \leqslant j \leqslant N-1$,

$$\bigl|G^j(x_t) - G^j(y)\bigr| \leqslant \Delta_{r,\theta}(\mathcal{O}) \leqslant r = \eta_m. \tag{7.53}$$

So by (7.53), the fact that $\mathcal{O} \subseteq E_m$, and (7.34), we have that $G^j(x_t) \in F_m$ for each $0 \leqslant j \leqslant N-1$. Therefore, by (7.52), (7.53), and Lemma 3.27 (iii), for each $0 \leqslant j \leqslant N-1$,

$$d\big(G^j(x_t), \mathcal{O}\big) = \big|G^j(x_t) - G^j(y)\big| \geqslant \lambda_m^j |x_t - y|. \tag{7.54}$$

By (7.36), (7.52), (7.29), and (7.38),

$$\big|G^N(x_t) - G^N(y)\big| \leqslant L_2 \big|G^{N-1}(x_t) - G^{N-1}(y)\big| \leqslant L_2 \Delta_{r,\theta}(\mathcal{O}) \leqslant L_2 \theta \Delta(\mathcal{O}) \leqslant \Delta(\mathcal{O})/3. \tag{7.55}$$

Let $n_t \coloneqq N+1$ and $x_{t+1} = G^{n_t}(x_t)$. We now aim to show that $S_{n_t}\psi(x_t) \leqslant n_t \gamma$.

Let $n \in \mathbb{N}$ be the smallest positive integer satisfying

$$|G^n(x_t) - G^n(y)| > \rho. \tag{7.56}$$

Such an integer n exists and satisfies $1 \leqslant n \leqslant N$ since $|x_t - y| \leqslant \rho < \Delta_{r,\theta}(\mathcal{O})$ (see (7.48)), and by definition of N. Moreover, we have

$$\big|G^{n-1}(x_t) - G^{n-1}(y)\big| \leqslant \rho. \tag{7.57}$$

We will separately estimate two parts of the sum

$$S_{n_t}(\gamma - \psi)(x_t) = S_n(\gamma - \psi)(x_t) + S_{n_t - n}(\gamma - \psi)(G^n(x_t)) =: \mathrm{I} + \mathrm{II}\,. \tag{7.58}$$

For each $j \in \mathbb{N}$ with $n \leqslant j \leqslant N$, by (7.54), (7.56), and the fact that $\lambda_m > 1$ (see Lemma 3.27), we have

$$d\big(G^j(x_t), \mathcal{O}\big) = \big|G^j(x_t) - G^j(y)\big| \geqslant \lambda_m^{j-n} |G^n(x_t) - G^n(y)| > \rho.$$

Thus $G^j(x_t) \notin U$, and by (7.50), $\gamma - \psi\big(G^j(x_t)\big) > 0$ for each $n \leqslant j \leqslant N$. Hence by (7.43), (7.32), (7.55), (7.51), we have

$$\begin{aligned} \mathrm{II} &\geqslant \gamma - \psi\big(G^N(x_t)\big) = \gamma - \widetilde{\phi}\big(G^N(x_t)\big) + \epsilon d\big(G^N(x_t), \mathcal{O}\big)^\alpha \\ &\geqslant \gamma + \epsilon \big|G^N(x_t) - G^N(y)\big|^\alpha \geqslant \gamma + \epsilon (\Delta_{r,\theta}(\mathcal{O}))^\alpha. \end{aligned}$$

To estimate I, we write

$$\mathrm{I} = (n\gamma - S_n\psi(y)) + (S_n\psi(y) - S_n\psi(x_t)) =: \mathrm{III} + \mathrm{IV} \tag{7.59}$$

and will bound each of the parts III and IV below.

We write $n = pq + r$ for $q, r \in \mathbb{N}_0$ with $0 \leqslant r \leqslant p-1$. Then by (7.43), (7.32), and (7.47), we have $S_n\psi(y) \leqslant S_n\widetilde{\phi}(y) = S_{pq}\widetilde{\phi}(y) + S_r\widetilde{\phi}(y) \leqslant pq\gamma$. Thus, considering $\gamma < 0$ (see (7.47)), we obtain

$$\mathrm{III} \geqslant r\gamma \geqslant (p-1)\gamma.$$

Next, by (7.43), (7.54), (7.35), Lemma 3.27 (iii), (7.57), and (7.48), we have

$$\begin{aligned}
|\mathrm{IV}| &\leqslant \sum_{j=0}^{n-1}\bigl|\psi\bigl(G^j(x_t)\bigr)-\psi\bigl(G^j(y)\bigr)\bigr| \\
&\leqslant \sum_{j=0}^{n-1}\Bigl(\bigl|\widetilde{\phi}\bigl(G^j(x_t)\bigr)-\widetilde{\phi}\bigl(G^j(y)\bigr)\bigr|+\epsilon d\bigl(G^j(x_t),\mathcal{O}\bigr)^{\alpha}\Bigr) \\
&= \sum_{j=0}^{n-1}\Bigl(\bigl|\widetilde{\phi}\bigl(G^j(x_t)\bigr)-\widetilde{\phi}\bigl(G^j(y)\bigr)\bigr|+\epsilon\bigl|G^j(x_t)-G^j(y)\bigr|^{\alpha}\Bigr) \\
&\leqslant \sum_{j=0}^{n-1}(L_1+\epsilon)\bigl|G^j(x_t)-G^j(y)\bigr|^{\alpha} \\
&\leqslant \sum_{j=0}^{n-1}(L_1+\epsilon)\lambda_m^{-(n-1-j)\alpha}\bigl|G^{n-1}(x_t)-G^{n-1}(y)\bigr|^{\alpha} \\
&\leqslant \sum_{j=0}^{n-1}(L_1+\epsilon)\lambda_m^{-(n-1-j)\alpha}\rho^{\alpha} \\
&\leqslant \rho^{\alpha}(L_1+\epsilon)/(1-\lambda_m^{-\alpha}) \\
&\leqslant \epsilon^{-1}|\gamma|(L_1+1)/(1-\lambda_m^{-\alpha}) \\
&= \epsilon^{-1}|\gamma|L_3.
\end{aligned}$$

Combining the above estimates for II, III, and IV, we obtain from (7.58), (7.59), (7.47), (7.40) the final estimate

$$\begin{aligned}
n_t\gamma-S_{n_t}\psi(y) &= \mathrm{II}+\mathrm{III}+\mathrm{IV} \\
&\geqslant \gamma+\epsilon(\Delta_{r,\theta}(\mathcal{O}))^{\alpha}-(p-1)|\gamma|-L_3\epsilon^{-1}|\gamma| \\
&\geqslant \epsilon(\Delta_{r,\theta}(\mathcal{O}))^{\alpha}-\bigl(1+L_3\epsilon^{-1}\bigr)p|\gamma| \\
&\geqslant \Bigl(\epsilon-\bigl(1+L_3\epsilon^{-1}\bigr)L_1\tau\Bigr)(\Delta_{r,\theta}(\mathcal{O}))^{\alpha} \\
&\geqslant 0.
\end{aligned} \tag{7.60}$$

Denote $N_k \coloneqq \sum_{i=1}^{k} n_i$. Combining (7.49), (7.50), and (7.60) now gives

$$\liminf_{n\to+\infty}\frac{1}{n}S_n\psi(x)\leqslant\liminf_{k\to+\infty}\frac{1}{N_k}\sum_{i=1}^{k}S_{n_i}\psi(x_i)\leqslant\liminf_{k\to+\infty}\frac{1}{N_k}\sum_{i=1}^{k}n_i\gamma=\gamma.$$

Therefore, since $x\in I\smallsetminus\mathbb{Q}$ was arbitrary, Proposition 4.7 (ii), and (7.46) give $Q(G,\psi)\leqslant\gamma$, thereby completing the proof of the claim. □

Appendix A. Periodic locking property

In this appendix we prove the periodic locking property for the Gauss map. The proof of Lemma A.1 and Proposition A.2 is similar to [BZ15].

Lemma A.1. *Let $G\colon I\to I$ be the Gauss map. Let $\mu\in\mathcal{M}(I,G)$ be a measure supported on a periodic orbit $\mathcal{O}$. Then there exists a constant $C_\mu\geqslant 1$ such that for all $\nu\in\mathcal{M}(I,G)$ and $\phi\in C^{0,\alpha}(I)$,*

$$\langle\nu,\phi\rangle\leqslant\langle\mu,\phi\rangle+C_\mu|\phi|_{\alpha}\langle\nu,d(\cdot,\mathcal{O})^{\alpha}\rangle. \tag{A.1}$$

Proof. We will write $p := \operatorname{card}\mathcal{O}$ throughout this proof. If $p = 1$, i.e., if $\mathcal{O}$ consists of a single point x_0, then for every $\nu \in \mathcal{M}(I, G)$ and $\phi \in C^{0,\alpha}(I)$,

$$\langle \nu, \phi \rangle \leqslant \phi(x_0) + C_\mu |\phi|_\alpha \langle \nu, d(\cdot, x_0)^\alpha \rangle.$$

So (A.1) holds with $C_\mu = 1$.

From now on assume that $p \geqslant 2$. Clearly, $\mathcal{O} \cap \mathbb{Q} = \emptyset$. Let

$$\delta := \min\{|x - y| : x, y \in \mathcal{O}, x \neq y\}$$

be the corresponding gap, i.e., the smallest distance between distinct points in $\mathcal{O}$. Fix an arbitrary $y \in \mathcal{O}$. Since G is continuous at each irrational number, there exists $\epsilon_y > 0$ such that $\left|G^i(x) - G^i(y)\right| < \frac{\delta}{2}$ for all $x \in (y - \epsilon_y, y + \epsilon_y)$ and $0 \leqslant i \leqslant p - 1$. Moreover, since $\mathcal{O}$ is a finite set, there exists $\epsilon := \min_{y \in \mathcal{O}} \epsilon_y$ such that if $x \in I$, $y \in \mathcal{O}_\mu$, and $d(x, y) < \epsilon$, then $\left|G^i(x) - G^i(y)\right| < \frac{\delta}{2}$ for all $0 \leqslant i \leqslant p - 1$.

Define $C_\mu = 1/\epsilon$. We aim to check that the inequality (A.1) is satisfied for every $\nu \in \mathcal{M}(I, G)$. It is sufficient to consider ergodic ν; the general case will follow using ergodic decomposition.

Fix $x \in I$ and $\phi \in C^{0,\alpha}(I)$ such that the Birkhoff averages of the continuous function ϕ along the orbit of x converge to $\int_I \phi \, \mathrm{d}\mu$. We will inductively define a transport sequence $\{y_i\}_{i \geqslant -1}$ in $\mathcal{O}$. As an auxiliary device for the definition of the sequence, each integer $i \geqslant -1$ will be labelled as good or bad. The definition is as follows: the point $y_{-1} \in \mathcal{O}$ is chosen arbitrarily, and the time -1 is labelled bad. As inductive hypothesis let us suppose that $y_{-1}, y_0, \cdots, y_i$ are already defined and that the times $-1, \cdots, i$ are already labelled. Then

(a) If $d\big(G^{i+1}(x), \mathcal{O}\big) < \epsilon$ then each time $j \in \{i + 1, i + 1, \cdots, i + p\}$ is labelled good, and y_j is defined as the unique point in $\mathcal{O}$ that is closest to $G^j(x)$. Note that $y_j = G^{j-i}(y_i)$, and in particular each point of $\mathcal{O}$ appears exactly once in the sequence $y_{i+1}, y_{i+2}, \cdots, y_{i+p}$;

(b) If $d\big(G^{i+1}(x), \mathcal{O}\big) \geqslant \epsilon$ then the time $i + 1$ is labelled bad, and we define y_{i+1} as $G(y_k)$, where k is the largest bad time less than or equal to i.

This completes the definition of the transport sequence. Notice that $\lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi(y_i) = \langle \mu, \phi \rangle$. On the other hand for all $i \geqslant 0$, the distance $d\big(G^i(x), \mathcal{O}\big)$ is equal to $\left|G^i(x) - y_i\right|$ if i is a good time, and otherwise is at least ϵ. In either case we have

$$\left|G^i(x) - y_i\right| \leqslant C_\mu d\big(G^i(x), \mathcal{O}\big).$$

Using these properties we obtain, for every $\phi \in C^{0,\alpha}(I)$,

$$\begin{aligned}
\langle \nu, \phi \rangle &= \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi\big(G^i(x)\big) \leqslant \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \big(\phi(y_i) + |\phi|_\alpha \left|G^i(x) - y_i\right|^\alpha\big) \\
&\leqslant \lim_{n \to +\infty} \frac{1}{n} \sum_{i=0}^{n-1} \big(\phi(y_i) + C_\mu |\phi|_\alpha d\big(G^i(x), \mathcal{O}\big)^\alpha\big) = \langle \mu, \phi \rangle + C_\mu |\phi|_\alpha \langle \nu, d(\cdot, \mathcal{O})^\alpha \rangle.
\end{aligned}$$

The inequality (A.1) follows. □

Proposition A.2 (Locking property for periodic orbits). *The set* $\operatorname{Lock}^\alpha(G)$ *is an open dense subset of* $\mathscr{P}^\alpha(G)$.

Proof. By definition, the set $\operatorname{Lock}^\alpha(I)$ is open and contained in $\mathscr{P}^\alpha(G)$, so we only need to prove that it is dense.

Let $\phi \in \mathscr{P}(I)$, and suppose $\mu \in \mathcal{M}_{\max}(G, \phi)$ is supported on a periodic orbit $\mathcal{O}$. For each $t > 0$, consider $\phi_t := \phi - t d(\cdot, \mathcal{O})^\alpha$. The functions ϕ_t belong to the Banach space $C^{0,\alpha}(I)$, and converge to

ϕ as $t \to 0$. Moreover, for any $\psi \in C^{0,\alpha}(I)$ and $\nu \in \mathcal{M}(I,G)$, by Lemma A.1 we obtain

$$\begin{aligned}\langle \nu, \phi_t + \psi\rangle \leqslant \langle \nu, \phi\rangle + \langle \nu, \psi\rangle - t\langle \nu, d(\cdot,\mathcal{O})^\alpha\rangle \leqslant \langle \mu, \phi\rangle + \langle \mu, \psi\rangle + (C_\mu|\psi|_\alpha - t)\langle \nu, d(\cdot,\mathcal{O})^\alpha\rangle \\ = \langle \mu, \phi_t + \psi\rangle + (C_\mu|\psi|_\alpha - t)\langle \nu, d(\cdot,\mathcal{O})^\alpha\rangle.\end{aligned}$$

Therefore, if $|\psi|_\alpha < t/C_\mu$ then μ is the unique maximizing measure for $\phi_t + \psi$. This shows that $\phi_t \in \mathrm{Lock}^\alpha(G)$ for each $t > 0$. So ϕ belongs to the closure of $\mathrm{Lock}^\alpha(G)$. Hence we conclude that $\mathrm{Lock}^\alpha(G)$ is dense in $\mathscr{P}^\alpha(G) \cap C^{0,\alpha}(I)$. □

References

[Ak93] Akin, E., *The General Topology of Dynamical Systems*, volume 1 of *Grad. Stud. Math.*, Amer. Math. Soc., Providence, RI, 1993.

[BF13] Bissacot, R. and Freire, F., On the existence of maximizing measures for irreducible countable Markov shifts: a dynamical proof. *Ergodic Theory Dynam. System* **34** (2013), 1103–1115.

[BG10] Bissacot, R. and Garibaldi, E., Weak KAM methods and ergodic optimal problems for countable Markov shifts. *Bull. Braz. Math. Soc. (N.S.)* **41** (2010), 321–338.

[Boc18] Bochi, J., Ergodic optimization of Birkhoff averages and Lyapunov exponents. In *Proc. Internat. Congr. Math. (Rio de Janeiro 2018)*, Volume III, World Sci. Publ., Singapore, 2018, pp. 1843–1866.

[Boc19] Bochi, J., Genericity of periodic maximization: proof of Contreras' theorem following Huang, Lian, Ma, Xu, and Zhang. 2019. Available at `https://personal.science.psu.edu/jzd5895/docs/Contreras_dapres_HLMXZ.pdf`.

[BZ15] Bochi, J. and Zhang, Yiwei, Note on robustness of periodic measures in ergodic optimization. 2015. Available at `http://www.personal.psu.edu/jzd5895/`.

[Bou00] Bousch, T., Le poisson n'a pas d'arêtes. *Ann. Inst. Henri Poincaré Probab. Stat.* **36** (2000), 489–508.

[Bou01] Bousch, T., La condition de Walters. *Ann. Sci. Éc. Norm. Supér. (4)* **34** (2001), 287–311.

[Bou08] Bousch, T., Nouvelle preuve d'un théorème de Yuan et Hunt. *Bull. Soc. Math. France* **126** (2008), 227–242.

[BM02] Bousch, T. and Mairesse, J., Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. *J. Amer. Math. Soc.* **15** (2002), 77–111.

[Br91] Brezinski, C., *History of Continued Fractions and Padé Approximants*, Springer, Berlin, 1991.

[Co16] Contreras, G., Ground states are generically a periodic orbit. *Invent. Math.* **205** (2016), 383–412.

[CLT01] Contreras, G., Lopes, A.O., and Thieullen, Ph., Lyapunov minimizing measures for expanding maps of the circle. *Ergodic Theory Dynam. System* **21** (2001), 1379–1409.

[EW11] Einsiedler, M. and Ward, T., *Ergodic Theory with a View towards Number Theory*, Springer, London, 2011.

[FSU14] Fishman, L., Simmons, D., and Urbański, M., Diophantine properties of measures invariant with respect to the Gauss map. *J. Anal. Math.* **122** (2014), 289–315.

[Ga12] Gauss, K.F., *Werke X. 1.*

[GG24] Garibaldi, E. and Gomes, J., Dense periodic optimization for countable Markov shift via Aubry points. Preprint, (arXiv:2410.06464), 2024.

[GGS25] Garibaldi, E., Gomes, J., and Sobottka, M., Maximizing measures for countable alphabet shifts via blur shift spaces. Preprint, (arXiv:2507.18736), 2025.

[GLT09] Garibaldi, E., Lopes, A.O., and Thieullen, Ph., On calibrated and separating sub-actions. *Bull. Braz. Math. Soc. (N.S.)* **40** (2009), 577–602.

[HLMXZ25] Huang, Wen, Lian, Zeng, Ma, Xiao, Xu, Leiye, and Zhang, Yiwei, Ergodic optimization theory for a class of typical maps. *J. Eur. Math. Soc.* doi:10.4171/JEMS/1652

[Io07] Iommi, G., Ergodic optimization for renewal type shifts. *Monatsh. Math.* **150** (2007), 91–95.

[IV21] Iommi, G. and Velozo, A., The space of invariant measures for countable Markov shifts. *J. Anal. Math.* **143** (2021), 461–501.

[IV25] Iommi, G. and Velozo, A., Non-compact space of invariant measures. Preprint, (arXiv:2503.09836), 2025.

[Je00] Jenkinson, O., Frequency locking on the boundary of the barycentre set. *Exp. Math.* **9** (2000), 309–317.

[Je08] Jenkinson, O., On sums of powers of inverse complete quotients. *Proc. Amer. Math. Soc.* **136** (2008), 1023–1027.

[Je19] Jenkinson, O., Ergodic optimization in dynamical systems. *Ergodic Theory Dynam. System* **39** (2019), 2593–2618.

[JMU06] Jenkinson, O., Mauldin, R.D., and Urbański, M., Ergodic optimization for countable alphabet subshifts of finite type. *Ergodic Theory Dynam. System* **26** (2006), 1791–1803.

[JMU07] Jenkinson, O., Mauldin, R.D., and Urbański, M., Ergodic optimization for noncompact dynamical systems. *Dyn. Syst.* **22** (2007), 379–388.

[JeSt10] Jenkinson, O. and Steel, J., Majorization of invariant measures for orientation-reversing maps. *Ergodic Theory Dynam. System* **30** (2010), 1471–1483.

[JoSa16] Jordan, T and Sahlsten, T, Fourier transforms of Gibbs measures for the Gauss map. *Math. Ann.* **364** (2016), 983–1023.

[LZ25] Li, Zhiqiang and Zhang, Yiwei, Ground states and periodic orbits for expanding Thurston maps. *Math. Ann.* **391** (2025), 3913–3985.

[Ma87] Mañé, R., *Ergodic Theory and Differentiable Dynamics*, Springer, Berlin, 1987.

[Mi17] Miller, A.W., *Descriptive Set Theory and Forcing: How to Prove Theorems about Borel Sets the Hard Way*, Cambridge Univ. Press, Cambridge, 2017.

[Mo08] Morris, I.D., Maximizing measures of generic Hölder functions have zero entropy. *Nonlinearity* **21** (2008), 993–1000.

[MU03] Mauldin, R.D. and Urbański, M., *Graph Directed Markov Systems: Geometry and Dynamics of Limit Sets*, Cambridge Univ. Press, Cambridge, 2003.

[Pa66] Parry, W., Symbolic dynamics and transformations of the unit interval. *Trans. Amer. Math. Soc.* **122** (1966), 368–378.

[Pi23] Pinto, N., Zero temperature limits for quotients of potentials in countable Markov shifts. *J. Stat. Phys.* **190** (2023), article no. 94.

[PW99] Pollicott, M. and Weiss, H., Multifractal analysis of Lyapunov exponent for continued fraction and Manneville–Pomeau transformations and applications to Diophantine approximation. *Comm. Math. Phys.* **207** (1999), 145–171.

[PU10] Przytycki, F. and Urbański, M., *Conformal Fractals: Ergodic Theory Methods*, Cambridge Univ. Press, Cambridge, 2010.

[QS12] Quas, A. and Siefken, J., Ergodic optimization of supercontinuous functions on shift spaces. *Ergodic Theory Dynam. Systems* **32** (2012), 2071–2082.

[Sc80] Schmidt, W.M., *Diophantine Approximation*, volume 785 of *Lecture Notes in Math.*, Springer, Berlin, 1980.

[Wa78] Walters, P., Invariant measures and equilibrium states for some mappings which expand distances. *Trans. Amer. Math. Soc.* **236** (1978), 121–153.

[Wa82] Walters, P., *An Introduction to Ergodic Theory*, Springer, New York, 1982.

[YH99] Yuan, Guocheng and Hunt, B.R., Optimal orbits of hyperbolic systems. *Nonlinearity* **12** (1999), 1207–1224.

Yinying Huang: School of Mathematical Sciences, Peking University, Beijing 100871, China.
Email address: miaoyan@stu.pku.edu.cn

Oliver Jenkinson, School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 4NS, United Kingdom
Email address: o.jenkinson@qmul.ac.uk

Zhiqiang Li, School of Mathematical Sciences & Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China
Email address: zli@math.pku.edu.cn