PERIODIC POINTS AND THE MEASURE OF MAXIMAL ENTROPY OF AN EXPANDING THURSTON MAP

ZHIQIANG LI

ABSTRACT. In this paper, we show that each expanding Thurston map $f\colon S^2 \to S^2$ has $1 + \deg f$ fixed points, counted with appropriate weight, where $\deg f$ denotes the topological degree of the map f. We then prove the equidistribution of preimages and of (pre)periodic points with respect to the unique measure of maximal entropy μ_f for f. We also show that (S^2, f, μ_f) is a factor of the left shift on the set of one-sided infinite sequences with its measure of maximal entropy, in the category of measure-preserving dynamical systems. Finally, we prove that μ_f is almost surely the weak* limit of atomic probability measures supported on a random backward orbit of an arbitrary point.

1. INTRODUCTION

A basic theme in dynamics is the investigation of the measure-theoretic entropy and its maximizing measures known as the measures of maximal entropy. By the pioneering work of R. Bowen, D. Ruelle, P. Walters, Ya. Sinai, M. Lyubich, R. Mañé and many others, existence and uniqueness results for the measure of maximal entropy are known for uniformly expansive continuous dynamical systems, distance expanding continuous dynamical systems, uniformly hyperbolic smooth dynamical systems and rational maps on the Riemann sphere. In many cases, the measure of maximal entropy is also the asymptotic distribution of the period points (see [Pa64, Si72, Bo75, Ly83, FLM83, Ru89, PU10]).

In this paper, we discuss a class of dynamical systems that are not among the classical dynamical systems mentioned above, namely, expanding Thurston maps on a topological 2-sphere S^2. Thurston maps

are branched covering maps on a sphere S^2 that generalize rational maps with finitely many post-critical points on the Riemann sphere. More precisely, a branched covering map $f \colon S^2 \to S^2$ is a *Thurston map* if its topological degree is at least 2 and if each of its finitely many critical points is preperiodic. These maps arose in W. P. Thurston's study of a characterization of rational maps on the Riemann sphere in a general topological context (see [DH93]). We will give a more detailed introduction to Thurston maps in Section 3.

In order to obtain the existence and uniqueness of the measure of maximal entropy of a Thurston map, one has to impose a condition of expansion for the map. More generally, P. Haïssinsky and K. Pilgrim introduced a notion of expansion for any finite branched coverings between two suitable topological spaces (see [HP09, Section 2.1 and Section 2.2]). We will use an equivalent definition in the context of Thurston maps formulated by M. Bonk and D. Meyer in [BM10]. We will discuss the precise definition in Section 3. We call Thurston maps with this property *expanding Thurston maps*. For a list of equivalent definitions of expanding Thurston maps, we refer to [BM10, Proposition 8.2].

As mentioned earlier, words like "expanding" and "expansive" have been used in different contexts to describe various notions of expansion. Our notion of expansion differs from all of the classical notions (except that of [HP09]), with the closest notion being that of *piecewise expanding maps* from [BS03]. Even though an expanding Thurston map f is expanding, in the sense of [BS03], when restricted to any 1-tile X in the cell decompositions that we will discuss in Section 3, it is still not clear why f is piecewise expanding in their sense.

As a consequence of their general results in [HP09], P. Haïssinsky and K. Pilgrim proved that for each expanding Thurston map, there exists a measure of maximal entropy and that the measure of maximal entropy is unique for an expanding Thurston map without periodic critical points. M. Bonk and D. Meyer then proved the existence and uniqueness of the measure of maximal entropy for all expanding Thurston maps using an explicit combinatorial construction [BM10].

In [BM10], M. Bonk and D. Meyer studied various properties of expanding Thurston maps and gave a wealth of combinatorial and analytical tools for such maps. Using the framework set in [BM10], we investigate in this paper the numbers and locations of the fixed points, periodic points, and preperiodic points of such maps. We establish equidistribution results for preimages of any point, for preperiodic points, and for periodic points, with respect to the measure of

maximal entropy. We also generalize some of the results from [BM10] in the development of this paper.

We will now give a brief description of the structure and main results of this paper.

After fixing notation in Section 2, we introduce Thurston maps in Section 3 and record, in some cases generalize, a few key concepts and results from [BM10].

In Section 4, we study the fixed points, periodic points, and preperiodic points of the expanding Thurston maps. We set out to prove the first main theorem of this paper.

Theorem 1.1. *Every expanding Thurston map $f \colon S^2 \to S^2$ has $1 + \deg f$ fixed points, counted with weight given by the local degree of the map at each fixed point.*

Here $\deg f$ denotes the topological degree of the map f. The local degree is a natural weight for points on S^2 for expanding Thurston maps. P. Haïssinsky and K. Pilgrim also used the same weight in the general context they considered in [HP09]. For a more detailed discussion on the local degree, we refer to Section 3.

We first observe that the statement of Theorem 1.1 agrees with what can be concluded from the Lefschetz fixed-point theorem (see for example, [GP10, Chapter 3]) if the map f is smooth and the graph of f intersects the diagonal of $S^2 \times S^2$ transversely at each fixed point of f. However, an expanding Thurston map may not satisfy either of these conditions. It is not clear how to give a proof by using the Lefschetz fixed-point theorem.

The proof of Theorem 1.1 uses the correspondence between the fixed points of f and the 1-tiles in some cell decomposition of S^2 induced by f and its invariant Jordan curve $\mathcal{C} \subseteq S^2$, for the special case when f has a special invariant Jordan curve $\mathcal{C}$. In fact, f may not have such a Jordan curve, but by a main result of [BM10], for each n large enough there exists an f^n-invariant Jordan curve. We will need a slightly stronger result as formulated in Lemma 3.11. Then the general case follows from an elementary number-theoretic argument. One of the advantages of this proof is that we also exhibit an almost one-to-one correspondence between the fixed points and the 1-tiles in the cell decomposition of S^2, which leads to precise information on the location of each fixed point. This information is essential later in the proof of the equidistribution of preperiodic and periodic points of expanding Thurston maps in Section 5. As a corollary of Theorem 1.1, we give a formula in Corollary 4.8 for the number of preperiodic points when counted with the corresponding weight.

Section 5 is based on the concepts of topological entropy, measure-theoretic entropy, and the measure of maximal entropy. Here we prove a number of equidistribution results. More precisely, we first prove in Theorem 5.7 the equidistribution of the n-tiles in the tile decompositions discussed in Section 3 with respect to the measure of maximal entropy μ_f of an expanding Thurston map f. The proof uses a combinatorial characterization of μ_f due to M. Bonk and D. Meyer [BM10] that we will state explicitly in Theorem 5.6.

We then formulate the equidistribution of preimages with respect to the measure of maximal entropy μ_f in Theorem 1.2 below. Here we denote by δ_x the Dirac measure supported on a point x in S^2.

Theorem 1.2 (Equidistribution of preimages). *Let $f\colon S^2 \to S^2$ be an expanding Thurston map with its measure of maximal entropy μ_f. Fix $p \in S^2$ and define the Borel probability measures*

$$(1.1) \qquad \nu_i = \frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(p)} \deg_{f^i}(q)\delta_q, \qquad \widetilde{\nu}_i = \frac{1}{Z_i} \sum_{q \in f^{-i}(p)} \delta_q,$$

for each $i \in \mathbb{N}_0$, where $Z_i = \operatorname{card}\left(f^{-i}(p)\right)$. Then

$$(1.2) \qquad\qquad\qquad \nu_i \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty,$$

$$(1.3) \qquad\qquad\qquad \widetilde{\nu}_i \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty.$$

Here $\deg_{f^i}(x)$ denotes the local degree of the map f^i at a point $x \in S^2$. In (1.2), (1.3), and similar statements below, the convergence of Borel measures is in the weak* topology, and we use w^* to denote it.

After generalizing Lemma 5.5, which is due to M. Bonk and D. Meyer [BM10, Lemma 20.2], in Lemma 5.12, we prove the equidistribution of preperiodic points with respect to μ_f. Note that Theorem 1.1 is used here.

Theorem 1.3 (Equidistribution of preperiodic points). *Let $f\colon S^2 \to S^2$ be an expanding Thurston map with its measure of maximal entropy μ_f. For each $m \in \mathbb{N}_0$ and each $n \in \mathbb{N}$ with $m < n$, we define the Borel probability measures*

$$(1.4) \qquad \xi_n^m = \frac{1}{s_n^m} \sum_{f^m(x)=f^n(x)} \deg_{f^n}(x)\delta_x, \qquad \widetilde{\xi}_n^m = \frac{1}{\widetilde{s}_n^m} \sum_{f^m(x)=f^n(x)} \delta_x,$$

where $s_n^m, \widetilde{s}_n^m$ are the normalizing factors defined in (4.6) and (4.7). If $\{m_n\}_{n\in\mathbb{N}}$ is a sequence in $\mathbb{N}_0$ such that $m_n < n$ for each $n \in \mathbb{N}$, then

$$(1.5) \qquad\qquad\qquad \xi_n^{m_n} \xrightarrow{w^*} \mu_f \text{ as } n \longrightarrow +\infty,$$

$$(1.6) \qquad \widetilde{\xi}_n^{m_n} \xrightarrow{w^*} \mu_f \ as \ n \longrightarrow +\infty.$$

We prove in Corollary 4.8 that $s_n^m = (\deg f)^n + (\deg f)^m$ for $m \in \mathbb{N}_0$, $n \in \mathbb{N}$ with $m < n$.

As a special case of Theorem 1.3, we get the following corollary.

Corollary 1.4 (Equidistribution of periodic points). *Let $f \colon S^2 \to S^2$ be an expanding Thurston map with its measure of maximal entropy μ_f. Then*

$$(1.7) \qquad \frac{1}{1 + (\deg f)^n} \sum_{x = f^n(x)} \deg_{f^n}(x)\delta_x \xrightarrow{w^*} \mu_f \ as \ n \longrightarrow +\infty,$$

$$(1.8) \qquad \frac{1}{\operatorname{card}\{x \in S^2 \,|\, x = f^n(x)\}} \sum_{x = f^n(x)} \delta_x \xrightarrow{w^*} \mu_f \ as \ n \longrightarrow +\infty,$$

$$(1.9) \qquad \frac{1}{(\deg f)^n} \sum_{x = f^n(x)} \delta_x \xrightarrow{w^*} \mu_f \ as \ n \longrightarrow +\infty.$$

The equidistribution (1.2), (1.3), (1.7), and (1.8) are analogs of corresponding results for rational maps on the Riemann sphere by M. Lyubich [Ly83]. Some ideas from [Ly83] are used in the proofs of Theorem 1.2 and Theorem 1.3 as well. P. Haïssinsky and K. Pilgrim also proved (1.2) and (1.7) in their general context [HP09], which includes expanding Thurston maps.

The equidistribution (1.5) and (1.6) are inspired by the recent work of M. Baker and L. DeMarco [BD11]. They used some equidistribution result of preperiodic points of rational maps on the Riemann sphere in the context of arithmetic dynamics.

We show in Corollary 5.15 that for each expanding Thurston map f, the exponential growth rate of the cardinality of the set of fixed points of f^n is equal to the topological entropy $h_{\text{top}}(f)$ of f, which is known to be equal to $\log(\deg f)$ (see for example, [BM10, Corollary 20.8]). This is analogous to the corresponding result for expansive homeomorphisms on compact metric spaces with the *specification property* (see for example, [KH95, Theorem 18.5.5]).

In Section 6, we prove in Theorem 6.2 that for each expanding Thurston map f with its measure of maximal entropy μ_f, the measure-preserving dynamical system (S^2, f, μ_f) is a factor, in the category of measure-preserving dynamical systems, of the measure-preserving dynamical system of the left-shift operator on the one-sided infinite sequences of $\deg f$ symbols together with its measure of maximal entropy.

This generalizes the corresponding result in [BM10] in the category of topological dynamical systems, reformulated in Theorem 6.1.

Finally, in Section 7, we follow the idea of J. Hawkins and M. Taylor [HT03] to prove in Theorem 7.1 that for each $p \in S^2$, the measure of maximal entropy μ_f of an expanding Thurston map f is almost surely the limit of

$$\frac{1}{n} \sum_{i=0}^{n-1} \delta_{q_i}$$

as $n \longrightarrow +\infty$ in the weak* topology, where q_i is one of the points in $f^{-1}(q_{i-1})$, chosen with probability proportional to the weight given by the local degree of f at each point in $f^{-1}(q_{i-1})$, for all $i \in \mathbb{N}$, and $q_0 = p$. A similar result for certain hyperbolic rational maps on the Riemann sphere was proved by M. Barnsley [Bar88]. J. Hawkins and M. Taylor generalized it to any rational map on the Riemann sphere of degree $d \geq 2$ [HT03].

2. Notation

Let $\mathbb{C}$ be the complex plane and $\widehat{\mathbb{C}}$ be the Riemann sphere. Let $\mathbb{D}$ be the open unit disk on $\mathbb{C}$. We use the convention that $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. We always use base e for logarithm unless otherwise specified.

The cardinality of a set A is denoted by $\operatorname{card} A$. For each $x \in \mathbb{R}$, we define $\lfloor x \rfloor$ as the greatest integer smaller than or equal to x, and $\lceil x \rceil$ the smallest integer greater than or equal to x.

Let (X, d) be a metric space. For subsets $A, B \subseteq X$, we set $d(A, B) = \inf\{d(x, y) \mid x \in A, y \in B\}$, and $d(A, x) = d(x, A) = d(A, \{x\})$ for $x \in X$. For each subset $S \subseteq X$, we denote the diameter of S by $\operatorname{diam}_d(S) = \sup\{d(x, y) \mid x, y \in S\}$, and the interior of S by $\operatorname{int} S$. For each $r > 0$, we set $N_d^r(A)$ to be the open r-neighborhood $\{y \in X \mid d(y, A) < r\}$ of A, and $\overline{N_d^r}(A)$ the closed r-neighborhood $\{y \in X \mid d(y, A) \leq r\}$ of A. The identity map $\operatorname{id}_X \colon X \to X$ maps each $x \in X$ to x itself. We denote by $C(X)$ the space of continuous functions from X to $\mathbb{R}$, by $\mathcal{M}(X)$ the set of finite signed Borel measures, and $\mathcal{P}(X)$ the set of Borel probability measures on X. We use $\|\cdot\|$ to denote the total variation norm on $\mathcal{M}(X)$. For a point $x \in X$, we define δ_x as the Dirac

measure supported on $\{x\}$. For a continuous map $g\colon X \to X$ we set $\mathcal{M}(X, g)$ to be the set of g-invariant Borel probability measures on X.

3. Thurston maps

In this section, we briefly review some key concepts and results on Thurston maps, and expanding Thurston maps in particular. For a more thorough treatment of the subject, we refer to [BM10]. Towards the end of this section, we state and prove a slightly stronger version of one of the main theorems in [BM10], which we will repeatedly use in the following sections.

Let S^2 denote an oriented topological 2-sphere. A continuous map $f\colon S^2 \to S^2$ is called a *branched covering map* on S^2 if for each point $x \in S^2$, there exists $d \in \mathbb{N}$, open neighborhoods U of x and V of $y = f(x)$, open neighborhoods U' and V' of 0 in $\widehat{\mathbb{C}}$, and orientation-preserving homeomorphisms $\varphi\colon U \to U'$ and $\eta\colon V \to V'$ such that $\varphi(x) = 0$, $\eta(y) = 0$, and

$$(\eta \circ f \circ \varphi^{-1})(z) = z^d$$

for each $z \in U'$. The positive integer d above is called the *local degree* of f at x and is denoted by $\deg_f(x)$. The *(global) degree* of f is defined as

$$(3.1) \qquad \deg f = \sum_{x \in f^{-1}(y)} \deg_f(x)$$

for each $y \in S^2$. It is independent of $y \in S^2$. It is true that if $f\colon S^2 \to S^2$ and $g\colon S^2 \to S^2$ are two branched covering maps on S^2, then

$$(3.2) \qquad \deg_{f \circ g}(x) = \deg_g(x) \deg_f(g(x)), \qquad \text{for } x \in S^2,$$

and moreover,

$$(3.3) \qquad \deg(f \circ g) = (\deg f)(\deg g).$$

A point $x \in S^2$ is a *critical point* of f if $\deg_f(x) \geq 2$. The set of critical points of f is denoted by $\operatorname{crit} f$. A point $y \in S^2$ is a *postcritical point* of f if $y \in \bigcup_{n \in \mathbb{N}} \{f^n(x) \mid x \in \operatorname{crit} f\}$. The set of postcritical points of f is denoted by $\operatorname{post} f$. Note that $\operatorname{post} f = \operatorname{post} f^n$ for all $n \in \mathbb{N}$.

Definition 3.1 (Thurston maps). A Thurston map is a branched covering map $f\colon S^2 \to S^2$ on S^2 with $\deg f \geq 2$ and $\operatorname{card}(\operatorname{post} f) < +\infty$.

We now set up the notation for cell decompositions of S^2. A *cell of dimension n* in S^2, $n \in \{1, 2\}$, is a subset $c \subseteq S^2$ that is homeomorphic to the closed unit ball $\overline{\mathbb{B}^n}$ in $\mathbb{R}^n$. We define the *boundary of c*, denoted by ∂c, to be the set of points corresponding to $\partial \mathbb{B}^n$ under such

a homeomorphism between c and $\overline{\mathbb{B}^n}$. The *interior of c* is defined to be $\mathrm{inte}(c) = c \setminus \partial c$. A cell c of dimension 0 is a singleton set $\{x\}$ for some point $x \in S^2$. For cells c of dimension 0, we adopt the convention that $\partial c = \emptyset$ and $\mathrm{inte}(c) = c$.

The following three definitions are from [BM10].

Definition 3.2 (Cell decompositions). Let $\mathbf{D}$ be a collection of cells in S^2. We say that $\mathbf{D}$ is a *cell decomposition of S^2* if the following conditions are satisfied:

(i) the union of all cells in $\mathbf{D}$ is equal to S^2,

(ii) for $c_1, c_2 \in \mathbf{D}$ with $c_1 \neq c_2$, we have $\mathrm{inte}(c_1) \cap \mathrm{inte}(c_2) = \emptyset$,

(iii) if $c \in \mathbf{D}$, then ∂c is a union of cells in $\mathbf{D}$,

(iv) every point in S^2 has a neighborhood that meets only finitely many cells in $\mathbf{D}$.

Definition 3.3 (Refinements). Let $\mathbf{D}'$ and $\mathbf{D}$ be two cell decompositions of S^2. We say that $\mathbf{D}'$ is a *refinement* of $\mathbf{D}$ if the following conditions are satisfied:

(i) for every cell $c' \in \mathbf{D}'$ there exists a cell $c \in \mathbf{D}$ with $c' \subseteq c$,

(ii) every cell $c \in \mathbf{D}$ is the union of all cells $c' \in \mathbf{D}'$ with $c' \subseteq c$.

Definition 3.4 (Cellular maps and cellular Markov partitions). Let $\mathbf{D}'$ and $\mathbf{D}$ be two cell decompositions of S^2. We say that a continuous map $f \colon S^2 \to S^2$ is *cellular* for $(\mathbf{D}', \mathbf{D})$ if for every cell $c \in \mathbf{D}'$, the restriction $f|_c$ is a homeomorphism of c onto a cell in $\mathbf{D}$. We say that $(\mathbf{D}', \mathbf{D})$ is a *cellular Markov partition* for f if f is cellular for $(\mathbf{D}', \mathbf{D})$ and $\mathbf{D}'$ is a refinement of $\mathbf{D}$.

Let $f \colon S^2 \to S^2$ be a Thurston map, and $\mathcal{C} \subseteq S^2$ be a Jordan curve such that $\mathrm{post}\, f \subseteq \mathcal{C}$. Then the pair $(f, \mathcal{C})$ induces natural cell decompositions $\mathbf{D}^n(f, \mathcal{C})$ of S^2, for $n \in \mathbb{N}_0$, in the following way:

By the Jordan curve theorem, the set $S^2 \setminus \mathcal{C}$ has two connected components. We call the closure of one of them the *white 0-tile* for $(f, \mathcal{C})$, denoted by X_w^0, and the closure of the other one the *black 0-tile* for $(f, \mathcal{C})$, denoted by X_b^0. The set of 0-*tiles* is $\mathbf{X}^0(f, \mathcal{C}) = \{X_b^0, X_w^0\}$. The set of 0-*vertices* is $\mathbf{V}^0(f, \mathcal{C}) = \mathrm{post}\, f$. We define $\overline{\mathbf{V}}^0(f, \mathcal{C})$ to be $\{\{x\} \mid x \in \mathbf{V}^0(f, \mathcal{C})\}$. The set of 0-*edges* $\mathbf{E}^0(f, \mathcal{C})$ is the set of the closures of the connected components of $\mathcal{C} \setminus \mathrm{post}\, f$. Then we get a cell decomposition

$$\mathbf{D}^0(f, \mathcal{C}) = \mathbf{X}^0(f, \mathcal{C}) \cup \mathbf{E}^0(f, \mathcal{C}) \cup \overline{\mathbf{V}}^0(f, \mathcal{C})$$

of S^2 consisting of *cells of level 0*, or *0-cells*.

One can recursively define unique cell decompositions $\mathbf{D}^n(f,\mathcal{C})$, $n \in \mathbb{N}$, consisting of *n-cells* such that f is cellular for $(\mathbf{D}^{n+1}(f,\mathcal{C}), \mathbf{D}^n(f,\mathcal{C}))$. For details, we refer to [BM10, Lemma 5.4]. We denote by $\mathbf{X}^n(f,\mathcal{C})$ the set of *n-cells* of dimension 2, called *n-tiles*; by $\mathbf{E}^n(f,\mathcal{C})$ the set of *n-cells* of dimension 1, called *n-edges*; by $\overline{\mathbf{V}}^n(f,\mathcal{C})$ the set of *n-cells* of dimension 0; and by $\mathbf{V}^n(f,\mathcal{C})$ the set $\{x \,|\, \{x\} \in \overline{\mathbf{V}}^n(f,\mathcal{C})\}$, called the set of *n-vertices*. The *k-skeleton*, for $k \in \{0,1,2\}$, of $\mathbf{D}^n(f,\mathcal{C})$ is the union of all *n-cells* of dimension k in this cell decomposition.

For the convenience of the reader, we record Proposition 6.1 of [BM10] here in order to summarize properties of the cell decompositions $\mathbf{D}^n(f,\mathcal{C})$.

Proposition 3.5. *Let $k, n \in \mathbb{N}_0$, let $f: S^2 \to S^2$ be a Thurston map, $\mathcal{C} \subseteq S^2$ be a Jordan curve with $\operatorname{post} f \subseteq \mathcal{C}$, and $m = \operatorname{card}(\operatorname{post} f)$.*

  (i) *The map f^k is cellular for $(\mathbf{D}^{n+k}(f,\mathcal{C}), \mathbf{D}^n(f,\mathcal{C}))$. In particular, if c is any $(n+k)$-cell, then $f^k(c)$ is an n-cell, and $f^k|_c$ is a homeomorphism of c onto $f^k(c)$.*

  (ii) *Let c be an n-cell. Then $f^{-k}(c)$ is equal to the union of all $(n+k)$-cells c' with $f^k(c') = c$.*

  (iii) *The 0-skeleton of $\mathbf{D}^n(f,\mathcal{C})$ is the set $\mathbf{V}^n(f,\mathcal{C}) = f^{-n}(\operatorname{post} f)$, and we have $\mathbf{V}^n(f,\mathcal{C}) \subseteq \mathbf{V}^{n+k}(f,\mathcal{C})$. The 1-skeleton of $\mathbf{D}^n(f,\mathcal{C})$ is equal to $f^{-n}(\mathcal{C})$.*

  (iv) *$\operatorname{card}(\mathbf{V}^n(f,\mathcal{C})) \leq m(\deg f)^n$, $\operatorname{card}(\mathbf{E}^n(f,\mathcal{C})) = m(\deg f)^n$, and $\operatorname{card}(\mathbf{X}^n(f,\mathcal{C})) = 2(\deg f)^n$.*

  (v) *The n-edges are precisely the closures of the connected components of $f^{-n}(\mathcal{C}) \setminus f^{-n}(\operatorname{post} f)$. The n-tiles are precisely the closures of the connected components of $S^2 \setminus f^{-n}(\mathcal{C})$.*

  (vi) *Every n-tile is an m-gon, i.e., the number of n-edges and the number of n-vertices contained in its boundary are equal to m.*

For $n \in \mathbb{N}_0$, we define the *set of black n-tiles* as

$$\mathbf{X}_b^n(f,\mathcal{C}) = \{X \in \mathbf{X}^n(f,\mathcal{C}) \,|\, f^n(X) = X_b^0\},$$

and the *set of white n-tiles* as

$$\mathbf{X}_w^n(f,\mathcal{C}) = \{X \in \mathbf{X}^n(f,\mathcal{C}) \,|\, f^n(X) = X_w^0\}.$$

Moreover, for $n \in \mathbb{N}$, we define the *set of black n-tiles contained in a white $(n-1)$-tile* as

$$\mathbf{X}_{bw}^n(f,\mathcal{C}) = \{X \in \mathbf{X}_b^n(f,\mathcal{C}) \,|\, \exists X' \in \mathbf{X}_w^{n-1}(f,\mathcal{C}),\, X \subseteq X'\},$$

and similarly, the *set of black n-tiles contained in a black $(n-1)$-tile* is denoted by $\mathbf{X}_{bb}^n(f,\mathcal{C})$, the *set of white n-tiles contained in a black*

$(n-1)$-*tile* by $\mathbf{X}_{wb}^n(f, \mathcal{C})$, and the *set of white n-tiles contained in a while $(n-1)$-tile* by $\mathbf{X}_{ww}^n(f, \mathcal{C})$.

From now on, we will say the cell decompositions induced by the pair $(f, \mathcal{C})$ and induced by f and $\mathcal{C}$ interchangeably. If the pair $(f, \mathcal{C})$ is clear from the context, we will sometimes omit $(f, \mathcal{C})$ in the notation above.

We now define two notions of expansion by M. Bonk and D. Meyer [BM10].

It is proved in [BM10, Corollary 6.4] that for each expanding Thurston map f (see Definition 3.8 below), we have card(post f) ≥ 3.

Definition 3.6 (Joining opposite sides). Fix a Thurston map f with card(post f) ≥ 3 and an f-invariant Jordan curve $\mathcal{C}$ containing post f. A set $K \subseteq S^2$ *joins opposite sides* of $\mathcal{C}$ if K meets two disjoint 0-edges when card(post f) ≥ 4, or K meets all three 0-edges when card(post f) $= 3$.

Definition 3.7 (Combinatorial expansion). Let f be a Thurston map. We say that f is *combinatorially expanding* if card(post f) ≥ 3, and there exists an f-invariant Jordan curve $\mathcal{C} \subseteq S^2$ (i.e., $f(\mathcal{C}) \subseteq \mathcal{C}$) with post $f \subseteq \mathcal{C}$, and there exists a number $n_0 \in \mathbb{N}$ such that none of the n_0-tiles in $\mathbf{X}^{n_0}(f, \mathcal{C})$ joins opposite sides of $\mathcal{C}$.

Definition 3.8 (Expansion). A Thurston map $f \colon S^2 \to S^2$ is called *expanding* if there exist a metric d on S^2 that induces the standard topology on S^2 and a Jordan curve $\mathcal{C} \subseteq S^2$ containing post f such that
$$\lim_{n\to+\infty} \max\{\mathrm{diam}_d(X) \,|\, X \in \mathbf{X}^n(f, \mathcal{C})\} = 0.$$

Remarks 3.9. We observe that being expanding is a purely topological property of a Thurston map and independent of the choice of the metric d. It is also independent of the choice of the Jordan curve $\mathcal{C}$ [BM10, Lemma 8.1]. From the definition, it is clear that if f is an expanding Thurston map, so is f^n for each $n \in \mathbb{N}$.

P. Haïssinsky and K. Pilgrim developed a more general notion of expansion for finite branched coverings between two Hausdorff, locally compact, locally connected topological spaces (see [HP09, Section 2.1 and Section 2.2]). When restricted to Thurston maps, their notion of expansion is equivalent to our notion defined above (see [BM10, Proposition 8.2]). Such notions of expansion are the natural analogs in the context of finite branched coverings and Thurston maps to some of the more classical notions of expansion, such as expansive homeomorphisms and forward-expansive continuous maps between compact metric spaces (see for example, [KH95, Definition 3.2.11]), and

distance-expanding maps between compact metric spaces (see for example, [PU10, Chapter 4]). Our notion of expansion is not equivalent to any of such classical notions in the context of Thurston maps.

For an expanding Thurston map f, we can fix a metric d for f on S^2 called a visual metric. For the existence and properties of such metrics, see [BM10, Chapter 8]. In particular, we will need the fact that d induces the standard topology on S^2 ([BM10, Proposition 8.9]). For a visual metric d for f, there exists a unique constant $\Lambda > 1$ called the *expansion factor* of d (see [BM10, Chapter 8] for more details). One major advantage of visual metrics d is that in (S^2, d) we have good quantitative control for the sizes of the cells in the cell decompositions discussed above, see [BM10, Lemma 8.10]. More precisely,

Lemma 3.10 (M. Bonk & D. Meyer, 2010). *Let $f\colon S^2 \to S^2$ be an expanding Thurston map, and $\mathcal{C} \subseteq S^2$ be a Jordan curve containing* post f. *Let d be a visual metric on S^2 for f with expansion factor $\Lambda > 1$. Then there exists a constant $C \geq 1$ such that for all n-edges and all n-tiles τ with $n \in \mathbb{N}_0$, we have $C^{-1}\Lambda^{-n} \leq \operatorname{diam}_d(\tau) \leq C\Lambda^{-n}$.*

A Jordan curve $\mathcal{C} \subseteq S^2$ is f-invariant if $f(\mathcal{C}) \subseteq \mathcal{C}$. We are interested in f-invariant Jordan curves that contain post f, since for such a curve $\mathcal{C}$, the partition $(\mathbf{D}^1(f, \mathcal{C}), \mathbf{D}^0(f, \mathcal{C}))$ is then a cellular Markov partition for f. According to Example 15.5 in [BM10], f-invariant Jordan curves containing post f need not exist. However, M. Bonk and D. Meyer [BM10, Theorem 1.2] proved that there exists an f^n-invariant Jordan curve $\mathcal{C}$ containing post f for each sufficiently large n depending on f. We will need a slightly stronger version in this paper. Its proof is almost the same as that of [BM10, Theorem 1.2]. For the convenience of the reader, we include the proof here.

Lemma 3.11. *Let $f\colon S^2 \to S^2$ be an expanding Thurston map, and $\widetilde{\mathcal{C}} \subseteq S^2$ be a Jordan curve with* post $f \subseteq \widetilde{\mathcal{C}}$. *Then there exists an integer $N(f, \widetilde{\mathcal{C}}) \in \mathbb{N}$ such that for each $n \geq N(f, \widetilde{\mathcal{C}})$ there exists an f^n-invariant Jordan curve $\mathcal{C}$ isotopic to $\widetilde{\mathcal{C}}$ rel.* post f *such that no n-tile in $\mathbf{D}^n(f, \mathcal{C})$ joins opposite sides of $\mathcal{C}$.*

Proof. By [BM10, Lemma 15.9], there exists an integer $N(f, \widetilde{\mathcal{C}}) \in \mathbb{N}$ such that for each $n \geq N(f, \widetilde{\mathcal{C}})$, there exists a Jordan curve $\mathcal{C}' \subseteq f^{-n}(\widetilde{\mathcal{C}})$ that is isotopic to $\widetilde{\mathcal{C}}$ rel. post f, and no n-tile for $(f, \widetilde{\mathcal{C}})$ joins opposite sides of $\mathcal{C}'$. Let $H\colon S^2 \times [0, 1] \to S^2$ be this isotopy rel. post f. We set $H_t(x) = H(x, t)$ for $x \in S^2, t \in [0, 1]$. We have $H_0 = \operatorname{id}_{S^2}$ and $\mathcal{C}' = H_1(\widetilde{\mathcal{C}}) \subseteq f^{-n}(\widetilde{\mathcal{C}})$.

If we set $F = f^n$, then $\operatorname{post} F = \operatorname{post} f$ and F is also an expanding Thurston map ([BM10, Lemma 8.4]). Note that F is cellular for $(\mathbf{D}^n(f, \widetilde{\mathcal{C}}), \mathbf{D}^0(f, \widetilde{\mathcal{C}}))$. So $\mathbf{D}^1(F, \widetilde{\mathcal{C}}) = \mathbf{D}^n(f, \widetilde{\mathcal{C}})$ (see [BM10, Lemma 5.4]). Thus no 1-cell for $(H_1 \circ F, \mathcal{C}')$ joins opposite sides of $\mathcal{C}'$, and thus $H_1 \circ F$ is combinatorially expanding for $\mathcal{C}'$. Note that $\mathcal{C}'$ contains $\operatorname{post}(H_1 \circ F) = \operatorname{post} F = \operatorname{post} f$. By Corollary 13.18 in [BM10], there exists a homeomorphism $\phi\colon S^2 \to S^2$ that is isotopic to the identity rel. $\operatorname{post}(H_1 \circ F)$ such that $\phi(\mathcal{C}') = \mathcal{C}'$ and $G = \phi \circ H_1 \circ F$ is an expanding Thurston map. Since $\phi \circ H_1$ is isotopic to the identity on S^2 rel. $\operatorname{post} F$, the pair F and G are Thurston equivalent. By Theorem 10.4 in [BM10], there exists a homeomorphism $h\colon S^2 \to S^2$ that is isotopic to the identity on S^2 rel. $F^{-1}(\operatorname{post} F)$ with $F \circ h = h \circ G$. Set $\mathcal{C} = h(\mathcal{C}')$. Then $\mathcal{C}$ is a Jordan curve in S^2 that is isotopic to $\mathcal{C}'$ rel. $F^{-1}(\operatorname{post} F)$ and thus isotopic to $\widetilde{\mathcal{C}}$ rel. $\operatorname{post} F$. Since $F(\mathcal{C}) = F(h(\mathcal{C}')) = h(G(\mathcal{C}')) = h(\phi(H_1(F(\mathcal{C}')))) \subseteq h(\phi(\mathcal{C}')) = h(\mathcal{C}') = \mathcal{C}$, we get that $\mathcal{C}$ is F-invariant.

Moreover, since no 1-cell for $(H_1 \circ F, \mathcal{C}')$ joins opposite sides of $\mathcal{C}'$, $H_1 \circ F(\mathcal{C}') \subseteq H_1(\widetilde{\mathcal{C}}) = \mathcal{C}'$, $\phi\colon S^2 \to S^2$ is a homeomorphism isotopic to the identity rel. $\operatorname{post}(H_1 \circ F)$ with $\phi(\mathcal{C}') = \mathcal{C}'$, $G = \phi \circ H_1 \circ F$, we can conclude that $G(\mathcal{C}') \subseteq \mathcal{C}'$ and no 1-cell for $(G, \mathcal{C}')$ joins opposite sides of $\mathcal{C}'$. Since $h\colon S^2 \to S^2$ is a homeomorphism, $\mathcal{C} = h(\mathcal{C}')$, and $F \circ h = h \circ G$, we can finally conclude that no 1-cell for $(F, \mathcal{C})$ joins opposite sides of $\mathcal{C}$. Therefore no n-cell for $(f, \mathcal{C})$ joins opposite sides of $\mathcal{C}$.   □

Compared with [BM10, Lemma 1.2], the above lemma carries additional combinatorial information of $\mathcal{C}$, i.e., no n-tile joins opposite sides of $\mathcal{C}$. In fact, we will only need the following corollary of Lemma 3.11 in the following sections.

Corollary 3.12. *Let* $f\colon S^2 \to S^2$ *be an expanding Thurston map. Then there exists a constant* $N(f) > 0$ *such that for each* $n \geq N(f)$, *there exists an* f^n-*invariant Jordan curve* $\mathcal{C}$ *containing* $\operatorname{post} f$ *such that no* n-*tile in* $\mathbf{D}^n(f, \mathcal{C})$ *joins opposite sides of* $\mathcal{C}$.

Proof. We can choose an arbitrary Jordan curve $\widetilde{\mathcal{C}} \subseteq S^2$ containing $\operatorname{post} f$ and set $N(f) = N(f, \widetilde{\mathcal{C}})$, and $\mathcal{C}$ an f^n-invariant Jordan curve containing $\operatorname{post} f$ as in Lemma 3.11.   □

Lemma 3.13. *Let* $f\colon S^2 \to S^2$ *be an expanding Thurston map. Then for each* $p \in S^2$, *the set* $\bigcup\limits_{n=1}^{+\infty} f^{-n}(p)$ *is dense in* S^2, *and*

$$(3.4) \qquad \lim_{n \to +\infty} \operatorname{card}(f^{-n}(p)) = +\infty.$$

Proof. Let $\mathcal{C} \subseteq S^2$ be a Jordan curve containing post f. Let d be any metric on S^2 that generates the standard topology on S^2.

Without loss of generality, we assume that $p \in X_w^0$ where $X_w^0 \in \mathbf{X}_w^0(f, \mathcal{C})$ is the white 0-tile in the cell decompositions induced by $(f, \mathcal{C})$. The proof for the case when $p \in X_b^0$ where $X_b^0 \in \mathbf{X}_b^0(f, \mathcal{C})$ is the black 0-tile is similar.

By Proposition 3.5(ii), for each $n \in \mathbb{N}$ and each white n-tile $X_w^n \in \mathbf{X}_w^n(f, \mathcal{C})$, there is a point $q \in X_w^n$ with $f^n(q) = p$. Since f is an expanding Thurston map,

$$(3.5) \qquad \lim_{n \to +\infty} \max\{\operatorname{diam}_d(X) \mid X \in \mathbf{X}^n(f, \mathcal{C})\} = 0.$$

Then the density of the set $\bigcup_{n=1}^{+\infty} f^{-n}(p)$ follows from the observation that for each $n \in \mathbb{N}$, each black n-tile $X_b^n \in \mathbf{X}_b^n(f, \mathcal{C})$ intersects nontrivially with some white n-tile $X_w^n \in \mathbf{X}_w^n(f, \mathcal{C})$.

By the above observation, the triangular inequality, and the fact that $\operatorname{diam}_d(S^2) > 0$ and S^2 is connected in the standard topology, the equation (3.4) follows from (3.5). $\qquad\square$

4. Fixed points of expanding Thurston maps

The main goal of this section is to prove Theorem 1.1; namely, that the number of fixed points, counted with an appropriate weight, of an expanding Thurston map f is exactly $1 + \deg f$. In order to prove Theorem 1.1, we first establish in Lemma 4.1 and Lemma 4.2 an almost one-to-one correspondence between fixed points and 1-tiles in the cell decomposition $\mathbf{D}^1(f, \mathcal{C})$ for an expanding Thurston map f with an f-invariant Jordan curve $\mathcal{C}$ containing post f. As a consequence, we establish in Corollary 4.8 an exact formula for the number of preperiodic points, counted with appropriate weight. We end this section by establishing a formula for the exact number of periodic points with period n, $n \in \mathbb{N}$, for expanding Thurston maps without periodic critical points.

Let f be a Thurston map and $p \in S^2$ a periodic point of f of period $n \in \mathbb{N}$, we define the *weight of p (with respect to f)* as the local degree $\deg_{f^n}(p)$ of f^n at p. When f is understood from the context and p is a fixed point of f, we abbreviate it as the *weight of p*. We will prove in this section that each expanding Thurston map f has exactly $1 + \deg f$ fixed points, counted with weight.

Lemma 4.1. *Let f be an expanding Thurston map with an f-invariant Jordan curve $\mathcal{C}$ containing* post f. *If $X \in \mathbf{X}_{ww}^1(f, \mathcal{C}) \cup \mathbf{X}_{bb}^1(f, \mathcal{C})$ is a white 1-tile contained in the while 0-tile X_w^0 or a black 1-tile contained*

in the black 0*-tile* X_b^0*, then* X *contains at least one fixed point of* f*. If* $X \in \mathbf{X}_{wb}^1(f, \mathcal{C}) \cup \mathbf{X}_{bw}^1(f, \mathcal{C})$ *is a white* 1*-tile contained in the black* 0*-tile* X_b^0 *or a black* 1*-tile contained in the white* 0*-tile* X_w^0*, then* inte(X) *contains no fixed points of* f*.*

Recall that cells in the cell decompositions are by definition closed sets, and the set of 0-tiles $\mathbf{X}^0(f, \mathcal{C})$ consists of the white 0-tile X_w^0 and the black 0-tile X_b^0.

Proof. If $X \in \mathbf{X}_{ww}^1(f, \mathcal{C}) \cup \mathbf{X}_{bb}^1(f, \mathcal{C})$, then $X \subseteq f(X)$. By Proposition 3.5(i), $f|_X$ is a homeomorphism from X to $f(X)$, which is one of the two 0-tiles. Hence, $f(X)$ is homeomorphic to the closed unit disk. So by Brouwer's fixed point theorem, $(f|_X)^{-1}$ has a fixed point p. Thus p is also a fixed point of f.

If $X \in \mathbf{X}_{wb}^1(f, \mathcal{C})$, then inte$(X) \subseteq$ inte(X_b^0) and $f(X) = X_w^0$. Since $X_w^0 \cap$ inte$(X_b^0) = \emptyset$, the map f has no fixed points in inte(X). The case when $X \in \mathbf{X}_{bw}^1(f, \mathcal{C})$ is similar. $\square$

Lemma 4.2. *Let* f *be an expanding Thurston map with an* f*-invariant Jordan curve* $\mathcal{C}$ *containing* post f *such that no* 1*-tile in* $\mathbf{D}^1(f, \mathcal{C})$ *joins opposite sides of* $\mathcal{C}$*. Then for every* $n \in \mathbb{N}$*, each* n*-tile* $X^n \in \mathbf{X}^n(f, \mathcal{C})$ *contains at most one fixed point of* f^n*.*

Proof. Fix an arbitrary $n \in \mathbb{N}$. We denote $F = f^n$ and consider the cell decompositions induced by F and $\mathcal{C}$ in this proof. Note that F is also an expanding Thurston map and there is no 1-tile in $\mathbf{D}^1(F, \mathcal{C})$ joining opposite sides of $\mathcal{C}$.

It suffices to prove that each 1-tile $X^1 \in \mathbf{X}^1$ contains at most one fixed point of F.

Suppose that there are two distinct fixed points p, q of F in a 1-tile X^1. We prove that there is a contradiction in each of the following cases.

Case 1: at least one of the fixed points, say p, is in inte(X^1). Then $X^1 \in \mathbf{X}_{ww}^1 \cup \mathbf{X}_{bb}^1$ by Lemma 4.1. Since p is contained in the interior of $X_1 \cap F(X_1)$, we get that $X_1 \subseteq F(X_1)$. Since $F|_{X^1}$ is a homeomorphism from X^1 to $F(X^1)$ (see Proposition 3.5(i)), we define a 2-tile $X^2 = (F|_{X^1})^{-1}(X^1) \subseteq X^1$. Then we get that $p \in$ inte(X^2) and $F(X^2) = X^1$. On the other hand, the point q must be in X^2 as well for otherwise there exists $q' \neq q$ such that $q' \in X^2$ and $F(q') = q$, thus q' and q are two distinct points in X^1 whose images under F are q, contradicting the fact that $F|_{X^1}$ is a homeomorphism from X^1 to $F(X^1)$ and $X^1 \subseteq F(X^1)$. Similarly we can inductively construct an $(n+1)$-cell $X^{n+1} \subseteq X^n$ such that $F(X^{n+1}) = X^n$, $p \in$ inte(X^{n+1}), and $q \in X^{n+1}$, for each $n \in \mathbb{N}$.

This contradicts the fact that F is an expanding Thurston map, see Remark 3.9.

Case 2: there exists a 1-edge $e \in \mathbf{E}^1$ such that $p, q \in e$. Note that $e \subseteq X^1$. Then one of the fixed points p and q, say p, must be contained in the interior of e, for otherwise p, q are distinct 1-vertices that are fixed by F, thus they are both 0-vertices, hence X^1 joins opposite sides, a contradiction. Since $F(e)$ is a 0-edge by Proposition 3.5, and $p \in F(e)$, there exists a 1-edge $e' \subseteq F(e)$ with $p \in e'$. Thus e' intersects with e at the point p, which is an interior point of e. So $e' = e$, and $e \subseteq F(e)$. Then by the same argument as when $p \in \mathrm{inte}(X^1)$ in Case 1, we can get a contradiction to the fact that F is an expanding Thurston map.

Case 3: the points p, q are contained in two distinct 1-edges e_1, e_2 of X^1, respectively, and $e_1 \cap e_2 \neq \emptyset$. Since F is an expanding Thurston map, we have $m = \mathrm{card}(\mathrm{post}\, F) \geq 3$ (see [BM10, Corollary 6.4]). So X^1 is an m-gon (see Proposition 3.5(vi)). Since $e_1 \cap e_2 \neq \emptyset$, we get $\mathrm{card}(e_1 \cap e_2) = 1$, say $e_1 \cap e_2 = \{v\}$. By Case 2, we get that $v \neq p$ and $v \neq q$. Note that $p \in F(e_1)$, $q \in F(e_2)$, and $F(e_1), F(e_2)$ are 0-edges. If at least one of p and q is a 1-vertex, thus a 0-vertex as well, then since Proposition 3.5(i) implies that $F(e_1) \neq F(e_2)$, we can conclude that X^1 touches at least three 0-edges, thus joins opposite sides of $\mathcal{C}$, a contradiction. Hence $p \in \mathrm{inte}(e_1)$ and $q \in \mathrm{inte}(e_2)$. So $e_1 \subseteq F(e_1)$, $e_2 \subseteq F(e_2)$, and

$$\{v\} = e_1 \cap e_2 \subseteq F(e_1) \cap F(e_2) = F(e_1 \cap e_2) = F(\{v\}),$$

by Proposition 3.5(i). Thus $F(v) = v$. Then e_1 contains two distinct fixed points p and v of F, which is impossible by Case 2.

Case 4: the points p, q are contained in two distinct 1-edges e_1, e_2 of X^1, respectively, and $e_1 \cap e_2 = \emptyset$. Thus $\mathrm{card}(\mathrm{post}\, F) \geq 4$ by Proposition 3.5(vi), and $F(e_1)$ and $F(e_2)$ are a pair of disjoint edges of $F(X^1)$ by Proposition 3.5(i). But $p = F(p) \in F(e_1)$, $q = F(q) \in F(e_2)$, so X^1 joins opposite sides of $\mathcal{C}$, a contradiction.

Combining all cases above, we can conclude, therefore, that each 1-tile $X^1 \in \mathbf{X}^1$ contains at most one fixed point of F. $\square$

We can immediately get an upper bound of the number of periodic points of an expanding Thurston map from Lemma 4.2.

Corollary 4.3. *Let f be an expanding Thurston map. Then for each $n \in \mathbb{N}$ sufficiently large, the number of fixed points of f^n is $\leq 2(\deg f)^n$. In particular, the number of fixed points of f is finite.*

The proof is straightforward and left as an exercise for the reader.

Lemma 4.4. *Let f be an expanding Thurston map with an f-invariant Jordan curve $\mathcal{C}$ containing* post f. *Then*

$$(4.1) \qquad \deg(f|_{\mathcal{C}}) = \mathrm{card}(\mathbf{X}^1_{ww}(f, \mathcal{C})) - \mathrm{card}(\mathbf{X}^1_{bw}(f, \mathcal{C}))$$
$$= \mathrm{card}(\mathbf{X}^1_{bb}(f, \mathcal{C})) - \mathrm{card}(\mathbf{X}^1_{wb}(f, \mathcal{C})).$$

Here $\deg(f|_{\mathcal{C}})$ is the *degree* of the map $f|_{\mathcal{C}} \colon \mathcal{C} \to \mathcal{C}$. Roughly speaking, it measures the total number of times the image of $\mathcal{C}$ under f winds around $\mathcal{C}$ along the orientation of $\mathcal{C}$. See for example, [Ha02, Section 2.2] for a precise definition.

Recall that for each continuous path $\gamma \colon [a, b] \to \mathbb{C} \setminus \{0\}$ on the Riemann sphere $\widehat{\mathbb{C}}$, with $a, b \in \mathbb{R}$ and $a < b$, we can define the *variation of the argument along* γ, denoted by $V(\gamma)$, as the change of the imaginary part of the logarithm along γ. Note that $V(\gamma)$ is invariant under an orientation-preserving reparametrization of γ and if $\widetilde{\gamma} \colon [a, b] \to \widehat{\mathbb{C}}$ reverses the orientation of γ, i.e., $\widetilde{\gamma}(t) = \gamma(a+b-t)$, then $V(\widetilde{\gamma}) = -V(\gamma)$. We also note that if γ is a loop, then $V(\gamma) = 2\pi \, \mathrm{Ind}_{\gamma}(0)$, where $\mathrm{Ind}_{\gamma}(0)$ is the *winding number of* γ *with respect to* 0 [Bu79, Chapter IV].

Proof. Consider the cell decompositions induced by $(f, \mathcal{C})$. Let X^0_w be the white 0-tile.

We start with proving the first equality in (4.1).

By the Schoenflies theorem (see, for example, [Mo77, Theorem 10.4]), we can assume that S^2 is the Riemann sphere $\widehat{\mathbb{C}}$, and X^0_w is the unit disk with the center 0 disjoint from $f^{-1}(\mathcal{C})$.

For each 1-edge $e \in \mathbf{E}^1$, we choose a parametrization $\gamma^+_e \colon [0, 1] \to \mathbb{C} \setminus \{0\}$ of e with positive orientation (i.e., with the white 1-tile on the left), and a parametrization $\gamma^-_e \colon [0, 1] \to \mathbb{C} \setminus \{0\}$ of e with negative orientation. Then $f \circ \gamma^+_e$ and $f \circ \gamma^-_e$ are parametrizations of one of the 0-edges on the unit circle $\mathcal{C}$, with positive orientation and negative orientation, respectively.

We claim that

$$(4.2) \qquad \sum_{X \in \mathbf{X}^1_{ww}} \sum_{e \in \mathbf{E}^1, e \subseteq \partial X} V(f \circ \gamma^+_e) - \sum_{X \in \mathbf{X}^1_{bw}} \sum_{e \in \mathbf{E}^1, e \subseteq \partial X} V(f \circ \gamma^+_e)$$
$$= \sum_{e \in \mathbf{E}^1, e \subseteq \mathcal{C}} V(f \circ \gamma_e),$$

where on the right-hand side, $\gamma_e = \gamma^+_e$ if $e \subseteq \mathcal{C} \cap X$ for some $X \in \mathbf{X}^1_{ww}$ and $\gamma_e = \gamma^-_e$ if $e \subseteq \mathcal{C} \cap X$ for some $X \in \mathbf{X}^1_{bw}$, or equivalently, γ_e parametrizes e in such a way that X^0_w is always on the left of e for each $e \in \mathbf{E}^1$ with $e \subseteq \mathcal{C}$.

We observe that the left-hand side of (4.2) is the sum of $V(f \circ \gamma_e^+)$ over all 1-edges e in the boundary of a white 1-tile $X \subseteq X_w^0$ plus the sum of $V(f \circ \gamma_e^-)$ over all 1-edges e in the boundary of a black 1-tile $X \subseteq X_w^0$. Since each 1-edge e with $\text{inte}(e) \subseteq X_w^0$ is the intersection of exactly one 1-tile in $\mathbf{X}_{ww}^1$ and one 1-tile in $\mathbf{X}_{bw}^1$, the two terms corresponding to a 1-edge e that is not contained in $\mathcal{C}$ cancel each other. Moreover, there is exactly one term for each 1-edge $e \subseteq X_w^0$ that is contained in $\mathcal{C}$, and e that corresponds to such a term is parametrized in such a way that X_w^0 is on the left of e. The claim now follows.

We then note that by Proposition 3.5(i), the left-hand side of (4.2) is equal to

$$\sum_{X \in \mathbf{X}_{ww}^1} 2\pi - \sum_{X \in \mathbf{X}_{bw}^1} 2\pi = 2\pi \left(\text{card}(\mathbf{X}_{ww}^1) - \text{card}(\mathbf{X}_{bw}^1) \right),$$

and the right-hand side of (4.2) is equal to

$$2\pi \, \text{Ind}_{f \circ \gamma_{\mathcal{C}}}(0) = 2\pi \deg(f|_{\mathcal{C}}),$$

where $\gamma_{\mathcal{C}}$ is a parametrization of $\mathcal{C}$ with positive orientation. Hence the first equality in (4.1) follows.

The second equality in (4.1) follows by symmetry. $\qquad\square$

Let f be an expanding Thurston map with an f-invariant Jordan curve $\mathcal{C}$ containing post f. We orient $\mathcal{C}$ in such a way that the white 0-tile lies on the left of $\mathcal{C}$. Let $p \in \mathcal{C}$ be a fixed point of f. We say that $f|_{\mathcal{C}}$ *preserves the orientation at p* (resp. *reverses the orientation at p*) if there exists an open arc $l \subseteq \mathcal{C}$ with $p \in l$ such that f maps l homeomorphically to $f(l)$ and $f|_{\mathcal{C}}$ preserves (resp. reverses) the orientation on l. Note that it may happen that $f|_{\mathcal{C}}$ neither preserves nor reverses the orientation at p, because $f|_{\mathcal{C}}$ need not be a local homeomorphism near p, where it may behave like a "folding map".

Lemma 4.5. *Let f be an expanding Thurston map with an f-invariant Jordan curve $\mathcal{C}$ containing post f. Then the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ preserves the orientation minus the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ reverses the orientation is equal to $\deg(f|_{\mathcal{C}}) - 1$.*

Proof. Let $\psi \colon [0, 1] \to \mathcal{C}$ be a continuous map such that $\psi|_{(0,1)} \colon (0, 1) \to \mathcal{C} \setminus \{x_0\}$ is an orientation-preserving homeomorphism, and $\psi(0) = \psi(1) = x_0$ for some $x_0 \in \mathcal{C}$ that is not a fixed point of $f|_{\mathcal{C}}$. Note that for each $x \in \mathcal{C}$ with $x \neq x_0$, $\psi^{-1}(x)$ is a well-defined number in $(0, 1)$. In particular, $\psi^{-1}(y)$ is a well-defined number in $(0, 1)$ for each fixed point y of $f|_{\mathcal{C}}$. Define $\pi \colon \mathbb{R} \to \mathcal{C}$ by $\pi(x) = \psi(x - \lfloor x \rfloor)$. Then π is a covering map. We lift $f|_{\mathcal{C}} \circ \psi$ to $G \colon [0, 1] \to \mathbb{R}$ such that $\pi \circ G = f|_{\mathcal{C}} \circ \psi$
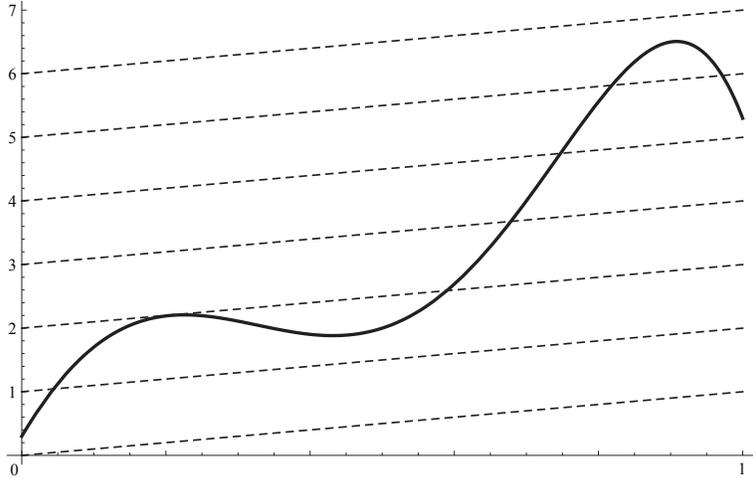
FIGURE 4.1. The lines l_k for $k \in \mathbb{Z}$ and an example of the graph of G.

and $G(0) = \psi^{-1}(f(x_0)) \in (0, 1)$. So we get the following commutative diagram:

$$
\begin{array}{ccc}
 & & \mathbb{R} \\
 & {\scriptstyle G} \nearrow & \downarrow {\scriptstyle \pi} \\
[0, 1] & \xrightarrow{\;f|_{\mathcal{C}} \circ \psi\;} & \mathcal{C}.
\end{array}
$$

Then $G(1) - G(0) \in \mathbb{Z}$ and

$$(4.3) \qquad\qquad \deg(f|_{\mathcal{C}}) = G(1) - G(0).$$

Observe that $y \in \mathcal{C}$ is a fixed point of $f|_{\mathcal{C}}$ if and only if $G(\psi^{-1}(y)) - \psi^{-1}(y) \in \mathbb{Z}$. Indeed, if $y \in \mathcal{C}$ is a fixed point of $f|_{\mathcal{C}}$, then $\pi \circ G \circ \psi^{-1}(y) = f|_{\mathcal{C}}(y) = y$. Thus $G \circ \psi^{-1}(y) - \psi^{-1}(y) \in \mathbb{Z}$. Conversely, if $G \circ \psi^{-1}(y) - \psi^{-1}(y) \in \mathbb{Z}$, then $y \neq x_0$ since $G(\psi^{-1}(x_0)) - \psi^{-1}(x_0) = G(0) - 0 \in (0, 1)$, thus

$$f|_{\mathcal{C}}(y) = f|_{\mathcal{C}} \circ \psi \circ \psi^{-1}(y) = \pi \circ G \circ \psi^{-1}(y) = \pi \circ \psi^{-1}(y) = y.$$

For each $m \in \mathbb{Z}$, we define the line l_m to be the graph of the function $x \mapsto x + m$ from $\mathbb{R}$ to $\mathbb{R}$.

Let $y \in \mathcal{C}$ be any fixed point of $f|_{\mathcal{C}}$. Since by Corollary 4.3 fixed points of f are isolated, there exists a neighborhood $(s, t) \subseteq (0, 1)$ such that $\psi^{-1}(y) \in (s, t)$ and for each fixed point $z \in \mathcal{C} \setminus \{y\}$ of $f|_{\mathcal{C}}$, $\psi^{-1}(z) \notin (s, t)$. Define $k = G(\psi^{-1}(y)) - \psi^{-1}(y)$; then $k \in \mathbb{Z}$. Moreover, $z \in \mathcal{C}$ is a fixed point of $f|_{\mathcal{C}}$ if and only if the graph of G intersects with l_m at the point $(\psi^{-1}(z), G(\psi^{-1}(z)))$ for some $m \in \mathbb{Z}$.

Depending on the orientation of $f|_{\mathcal{C}}$ at the fixed point $y \in \mathcal{C}$, we get one of the following cases:

(1) If $f|_{\mathcal{C}}$ preserves the orientation at y, then the graph of $G|_{(s,\psi^{-1}(y))}$ lies strictly between the lines l_{k-1} and l_k, and the graph of $G|_{(\psi^{-1}(y),t)}$ lies strictly between the lines l_k and l_{k+1}.

(2) If $f|_{\mathcal{C}}$ reverses the orientation at y, then the graph of $G|_{(s,\psi^{-1}(y))}$ lies strictly between the lines l_k and l_{k+1}, and the graph of $G|_{(\psi^{-1}(y),t)}$ lies strictly between the lines l_{k-1} and l_k.

(3) If $f|_{\mathcal{C}}$ neither preserves nor reverses the orientation at y, then the graph of $G|_{(s,t)\setminus\{\psi^{-1}(y)\}}$ either lies strictly between the lines l_{k-1} and l_k or lies strictly between the lines l_k and l_{k+1}.

Thus the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ preserves the orientation is exactly the number of intersections between the graph of G and the lines l_m with $m \in \mathbb{Z}$, where the graph of G crosses the lines from below, and the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ reserves the orientation is exactly the number of intersections between the graph of G and the lines l_m with $m \in \mathbb{Z}$, where the graph of G crosses the lines from above. Therefore the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ preserves the orientation minus the number of fixed points of $f|_{\mathcal{C}}$ where $f|_{\mathcal{C}}$ reverses the orientation is equal to $G(1) - G(0) - 1 = \deg(f|_{\mathcal{C}}) - 1$.   □

For each $n \in \mathbb{N}$ and each expanding Thurston map $f\colon S^2 \to S^2$, we denote by

$$(4.4) \quad P_{n,f} = \{x \in S^2 \mid f^n(x) = x, f^k(x) \neq x, k \in \{1, 2, \ldots, n-1\}\}$$

the *set of periodic points of f with period n*, and by

$$(4.5) \qquad p_{n,f} = \sum_{x \in P_{n,f}} \deg_{f^n}(x), \qquad \widetilde{p}_{n,f} = \operatorname{card} P_{n,f}$$

the numbers of periodic points x of f with period n, counted with and without weight $\deg_{f^n}(x)$, respectively, at each x. In particular, $P_{1,f}$ is the set of fixed points of f and $p_{1,f} = 1 + \deg f$ as we will see in the proof of Theorem 1.1 below. More generally, for all $m \in \mathbb{N}_0$ and $n \in \mathbb{N}$ with $m < n$, we denote by

$$(4.6) \qquad\qquad S_n^m = \{x \in S^2 \mid f^m(x) = f^n(x)\}$$

the *set of preperiodic points of f with parameters m, n* and by

$$(4.7) \qquad\qquad s_n^m = \sum_{x \in S_n^m} \deg_{f^n}(x), \qquad \widetilde{s}_n^m = \operatorname{card} S_n^m$$

the numbers of preperiodic points of f with parameters m, n, counted with and without weight $\deg_{f^n}(x)$, respectively, at each x. Note that

in particular, for each $n \in \mathbb{N}$, $S_n^0 = P_{1,f^n}$ is the set of fixed points of f^n.

Proposition 4.6. *Let $F \colon S^2 \to S^2$ be an expanding Thurston map with an F-invariant Jordan curve $\mathcal{C}$ containing* post F *such that no 1-tile in $\mathbf{D}^1(F, \mathcal{C})$ joins opposite sides of $\mathcal{C}$. Then F has $1 + \deg F$ fixed points, counted with weight given by the local degree $\deg_F(x)$ of the map at each fixed point x.*

Proof. We consider the cell decompositions induced by $(F, \mathcal{C})$ in this proof. Let $b_w = \operatorname{card} \mathbf{X}_{bw}^1$ be the number of black 1-tiles contained in the white 0-tile, and similarly, $w_w = \operatorname{card} \mathbf{X}_{ww}^1$, $w_b = \operatorname{card} \mathbf{X}_{wb}^1$, and $b_b = \operatorname{card} \mathbf{X}_{bb}^1$. Note that $w_w + w_b = b_w + b_b = \deg F$.

By Corollary 4.3, we know that fixed points of F are isolated.

Note that

(4.8) $$w_w + b_b = \deg F + \deg(F|_{\mathcal{C}}),$$

which follows from the equation $w_w - b_w = \deg(F|_{\mathcal{C}})$ by Lemma 4.4, and the equation $b_w + b_b = \deg F$.

We define sets

$$A = \{X \in \mathbf{X}_{ww}^1 \mid \text{there exists } p \in \mathcal{C} \cap X \text{ with } F(p) = p\},$$

$$B = \{X \in \mathbf{X}_{bw}^1 \mid \text{there exists } p \in \mathcal{C} \cap X \text{ with } F(p) = p\},$$

and let $a = \operatorname{card} A$, $b = \operatorname{card} B$.

We then claim that

(4.9) $$a - b = \deg(F|_{\mathcal{C}}) - 1.$$

In order to prove this claim, we will first prove that $a - b$ is equal to the number of fixed points of $F|_{\mathcal{C}}$ where $F|_{\mathcal{C}}$ preserves the orientation minus the number of fixed points of $F|_{\mathcal{C}}$ where $F|_{\mathcal{C}}$ reverses the orientation.

So let $p \in \mathcal{C}$ be a fixed point of $F|_{\mathcal{C}}$.

(1) If p is not a critical point of F, then either $F|_{\mathcal{C}}$ preserves or reverses the orientation at p. In this case, the point p is contained in exactly one white 1-tile and one black 1-tile.

   (a) If $F|_{\mathcal{C}}$ preserves the orientation at p, then p is contained in exactly one white 1-tile that is contained in the white 0-tile, and p is not contained in any black 1-tile that is contained in the while 0-tile.
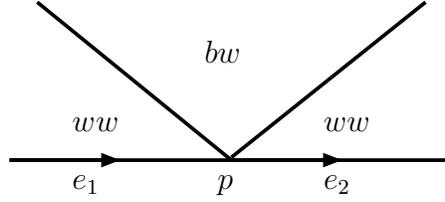
FIGURE 4.2. Case (2)(a) where $F(e_1) \supseteq e_1$ and $F(e_2) \supseteq e_2$.
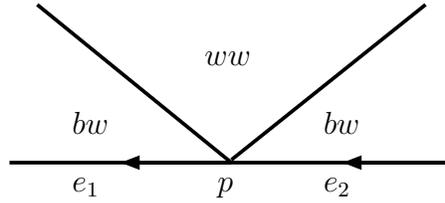

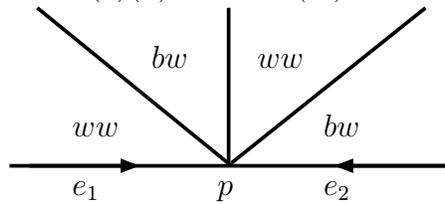FIGURE 4.3. Case (2)(b) where $F(e_1) \supseteq e_2$ and $F(e_2) \supseteq e_1$.

FIGURE 4.4. Case (2)(c) where $F(e_1) = F(e_2) \supseteq e_1$.
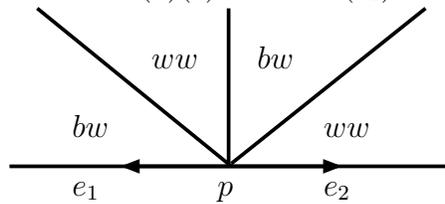
FIGURE 4.5. Case (2)(d) where $F(e_1) = F(e_2) \supseteq e_2$.

(b) If $F|_\mathcal{C}$ reverses the orientation at p, then p is contained in exactly one black 1-tile that is contained in the white 0-tile, and p is not contained in any white 1-tile that is contained in the while 0-tile.

(2) If p is a critical point of F, then $p = F(p) \in \text{post } f$ and so there are two distinct 1-edges $e_1, e_2 \subseteq \mathcal{C}$ such that $\{p\} = e_1 \cap e_2$. We refer to Figures 4.2 to 4.5.

(a) If $e_1 \subseteq F(e_1)$ and $e_2 \subseteq F(e_2)$, then p is contained in exactly k white and $k - 1$ black 1-tiles that are contained in the white 0-tile, for some $k \in \mathbb{N}$. Note that in this case $F|_\mathcal{C}$ preserves the orientation at p.

(b) If $e_2 \subseteq F(e_1)$ and $e_1 \subseteq F(e_2)$, then p is contained in exactly $k - 1$ white and k black 1-tiles that are contained in the

white 0-tile, for some $k \in \mathbb{N}$. Note that in this case $F|_{\mathcal{C}}$ reverses the orientation at p.

(c) If $e_1 \subseteq F(e_1) = F(e_2)$, then p is contained in exactly k white and k black 1-tiles that are contained in the white 0-tile, for some $k \in \mathbb{N}$. Note that in this case $F|_{\mathcal{C}}$ neither preserves nor reverses the orientation at p.

(d) If $e_2 \subseteq F(e_1) = F(e_2)$, then p is contained in exactly k white and k black 1-tiles that are contained in the white 0-tile, for some $k \in \mathbb{N}$. Note that in this case $F|_{\mathcal{C}}$ neither preserves nor reverses the orientation at p.

It follows then that $a - b$ is equal to the number of fixed points of $F|_{\mathcal{C}}$ where $F|_{\mathcal{C}}$ preserves the orientation minus the number of fixed points of $F|_{\mathcal{C}}$ where $F|_{\mathcal{C}}$ reverses the orientation.

Then the claim follows from Lemma 4.5.

Next, we are going to prove that the number of fixed points of F, counted with weight given by the local degree, is equal to

$$(4.10) \qquad\qquad w_w + b_b - a + b,$$

which, by (4.8) and the claim above, is equal to

$$\deg F + \deg(F|_{\mathcal{C}}) - (\deg(F|_{\mathcal{C}}) - 1) = 1 + \deg F.$$

Indeed, by Lemma 4.1 and Lemma 4.2, each 1-tile that contributes in (4.10), i.e., each 1-tile in $\mathbf{X}^1_{ww} \cup \mathbf{X}^1_{bb} \cup B \cup A$, contains exactly one fixed point (not counted with weight) of F. On the other hand, each fixed point is contained in at least one of the 1-tiles in $\mathbf{X}^1_{ww} \cup \mathbf{X}^1_{bb} \cup B \cup A$. Let p be a fixed point of F, then one of the following happens:

(1) If $p \notin \mathcal{C}$, then p is not contained in any 1-edges e since $F(e) \subseteq \mathcal{C}$. So $p \in \mathrm{inte}(X)$ for some $X \in \mathbf{X}^1_{ww} \cup \mathbf{X}^1_{bb} \setminus (A \cup B)$, by Lemma 4.1. So each such p contributes 1 to (4.10).

(2) If $p \in \mathcal{C}$ but $p \notin \mathrm{crit}\, F$, then p is not a 1-vertex, so either p is contained in exactly two 1-tiles $X \in \mathbf{X}^1_{ww}$ and $X' \in \mathbf{X}^1_{bb}$, or p is contained in exactly two 1-tiles $X \in \mathbf{X}^1_{bw}$ and $X' \in \mathbf{X}^1_{wb}$. In either case, p contributes 1 to (4.10).

(3) If $p \in \mathcal{C}$ and $p \in \mathrm{crit}\, F$, then p is a 0-vertex, so the part that p contributes in (4.10) counts the number of black 1-tiles that contains p, which is exactly the weight $\deg_F(p)$ of p.

The proof is now complete. $\square$

We can use some elementary number-theoretic argument to generalize Proposition 4.6 to all expanding Thurston maps, thus proving Theorem 1.1.

Proof of Theorem 1.1. We first observe that by Proposition 4.6, the theorem holds for f replaced by $F = f^n$ for each $n \geq N(f)$ where $N(f)$ is a constant as given in Corollary 3.12 depending only on f. Indeed, let $\mathcal{C}$ be an f^n-invariant Jordan curve containing post f such that no n-tile in $\mathbf{D}^n(f, \mathcal{C})$ joins opposite sides of $\mathcal{C}$ as given in Corollary 3.12. So $\mathcal{C}$ is an F-invariant Jordan curve containing post F such that no 1-tile in $\mathbf{D}^1(F, \mathcal{C})$ joins opposite sides of $\mathcal{C}$. Proposition 4.6 now applies.

Next, choose a prime $r \geq N(f)$. Note that the set of fixed points of f^r can be decomposed into periodic orbits under f of length r or 1, since r is a prime. Let p be a fixed point of f^r. By using the following formula derived from (3.2),

$$(4.11) \quad \deg_{f^r}(p) = \deg_f(p) \deg_f(f(p)) \deg_f(f^2(p)) \cdots \deg_f(f^{r-1}(p)),$$

we can conclude that

(1) if $p \notin \mathrm{crit}(f^r)$, or equivalently, $\deg_{f^r}(p) = 1$, and

    (i) if p is in a periodic orbit of length r, then $p, f(p), \ldots, f^{r-1}(p) \notin \mathrm{crit} f$, or equivalently, the local degrees of f^r at these points are all 1;

    (ii) if p is in a periodic orbit of length 1, then $p \notin \mathrm{crit} f$, or equivalently, $\deg_f(p) = 1$;

(2) if $p \in \mathrm{crit}(f^r)$, and

    (i) if p is in a periodic orbit of length r, then all $p, f(p), \ldots, f^{r-1}(p)$ are fixed points of f^r with the same weight $\deg_{f^r}(p) = \deg_{f^r}(f^k(p))$ for each $k \in \mathbb{N}$;

    (ii) if p is in a periodic orbit of length 1, then $p \in \mathrm{crit} f$ and the weight of f^r at p is $\deg_{f^r}(p) = (\deg_f(p))^r$.

Note that a fixed point $p \in S^2$ of f^r is a fixed point of f if and only if p is in a periodic orbit of length 1 under f. So by first summing the weight of the fixed points of f^r in the same periodic orbit then summing over all such orbits and applying Fermat's Little Theorem, we can conclude that

$$
\begin{aligned}
p_{1,f^r} &= \sum_{x \in P_{1,f^r}} \deg_{f^r}(x) \\
&= \underbrace{\sum r}_{(1)(i)} + \underbrace{\sum 1}_{(1)(ii)} + \underbrace{\sum r \deg_{f^r}(p)}_{(2)(i)} + \underbrace{\sum (\deg_f(p))^r}_{(2)(ii)} \\
&\equiv \underbrace{\sum 1}_{(1)(ii)} + \underbrace{\sum \deg_f(p)}_{(2)(ii)} \\
&= p_{1,f} \pmod{r},
\end{aligned}
$$

where on the second line, the first sum ranges over all periodic orbits in Case (1)(i), the second sum ranges over all periodic orbits in Case (1)(ii), the third sum ranges over all periodic orbits of the form $\{p, f(p), \ldots, f^{r-1}(p)\}$ in Case (2)(i), the last sum ranges over all periodic orbits of the form $\{p\}$ in Case (2)(ii). Thus by (3.3) and Fermat's Little Theorem again, we have

$$
\begin{aligned}
0 &= \deg(f^r) + 1 - p_{1,f^r} \\
&\equiv (\deg f)^r + 1 - p_{1,f} \\
&\equiv 1 + \deg f - p_{1,f} \pmod{r}.
\end{aligned}
$$

By choosing the prime r larger than $|1 + \deg f - p_{1,f}|$, we can conclude that $p_{1,f} = 1 + \deg f$. $\qquad\square$

In particular, we have the following corollary in which the weight for all points are trivial. The proof is left as a simple exercise for the reader.

Corollary 4.7. *If f is an expanding Thurston map with no critical fixed points, then there are exactly $1 + \deg f$ distinct fixed points of f. Moreover, if f is an expanding Thurston map with no periodic critical points, then there are exactly $1 + (\deg f)^n$ distinct fixed points of f^n, for each $n \in \mathbb{N}$.*

We recall the definition of s_n^m in (4.6) and (4.7).

Corollary 4.8. *Let f be an expanding Thurston map. For each $m \in \mathbb{N}_0$ and $n \in \mathbb{N}$ with $m < n$, we have*

$$(4.12) \qquad\qquad s_n^m = (\deg f)^n + (\deg f)^m.$$

Proof. For all $m \in \mathbb{N}_0$ and $n \in \mathbb{N}$ with $m < n$, we have

$$
\begin{aligned}
s_n^m &= \sum_{x \in S_n^m} \deg_{f^n}(x) = \sum_{y = f^{n-m}(y)} \sum_{x \in f^{-m}(y)} \deg_{f^n}(x) \\
&= \sum_{y = f^{n-m}(y)} \deg_{f^{n-m}}(y) \sum_{x \in f^{-m}(y)} \deg_{f^m}(x) \\
&= \left((\deg f)^{n-m} + 1\right)(\deg f)^m.
\end{aligned}
$$

The last equality follows from (3.1), (3.3), and Theorem 1.1. $\qquad\square$

Finally, for expanding Thurston maps with no periodic critical points, we derive a formula for $p_{n,f}$, $n \in \mathbb{N}$, from Theorem 1.1 and the Möbius inversion formula (see for example, [Bak85, Section 2.4]).

Definition 4.9. The *Möbius function*, $\mu(u)$, is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1; \\ (-1)^r & \text{if } n = p_1 p_2 \dots p_r, \text{ and } p_1, \dots, p_r \text{ are distinct primes;} \\ 0 & \text{otherwise.} \end{cases}$$

Corollary 4.10. *Let f be an expanding Thurston map without any periodic critical points. Then for each $n \in \mathbb{N}$, we have*

$$p_{n,f} = \sum_{d|n} \mu(d) p_{1,f^{n/d}} = \begin{cases} \sum_{d|n} \mu(d)(\deg f)^{n/d} & \text{if } n > 1; \\ 1 + \deg f & \text{if } n = 1. \end{cases}$$

Proof. The first equality follows from the Möbius inversion formula and the equation $p_{1,f^n} = \sum_{d|n} p_{d,f}$, for $n \in \mathbb{N}$. The second equality follows from Theorem 1.1 and the following fact (see for example, [Bak85, Section 2.4]):

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

$\square$

5. Equidistribution

In this section, we derive various equidistribution results as stated in Theorem 1.2, Theorem 1.3, and Corollary 1.4. We prove these results by first establishing a general statement in Theorem 5.7 on the convergence of the distributions of the white n-tiles in the tile decompositions discussed in Section 3, in the weak* topology, to the unique measure of maximal entropy of an expanding Thurston map. Theorem 1.1 is used in several places in this section.

We start with a review of the concepts of entropy for dynamical systems. Then we prove in Theorem 5.7 that the distributions of the points, each of which is located "near" its corresponding white n-tile where the correspondence is a bijection, converges in the weak* topology to μ_f as $n \longrightarrow +\infty$. Then Theorem 1.2 follows from Theorem 5.7. Theorem 1.3 finally follows after we prove a technical bound in Lemma 5.12 generalizing a corresponding lemma from [BM10]. As a special case, we obtain Corollary 1.4.

We start with recalling concepts of entropy for dynamical systems. See for example, [Wa82, Chapter 4 and Chapter 7] for a more detailed introduction to such concepts. We follow closely the notation from [BM10, Chapter 20].

Let (X, d) be a compact metric space and $g \colon X \to X$ a continuous map. One can define the *topological entropy* $h_{\mathrm{top}}(g)$ of g (see [BM10, Chapter 20] for a quick summary of relevant concepts and facts in our context, and refer to [KH95, Section 3.1.b] and [Wa82, Chapter 7] for more details). Note that $h_{\mathrm{top}}(g)$ is independent of d as long as the topology on X defined by d remains the same [KH95, Proposition 3.1.2]. Let $\mu \in \mathcal{M}(X, g)$ be a g-invariant Borel probability measure. We denote the *measure-theoretic entropy* of g for μ by $h_\mu(g)$ (see [BM10, Chapter 20] for a quick summary of relevant concepts and facts in our context, and refer to [KH95, Section 4.3] and [Wa82, Section 4.4] for more details).

By the variational principle (see [Wa82, Theorem 8.6]), we have

$$h_{\mathrm{top}}(g) = \sup\{h_\mu(g) \mid \mu \in \mathcal{M}(X, g)\}.$$

A measure μ that achieves the supreme above is called *a measure of maximal entropy* of g.

If f is an expanding Thurston map, then

$$(5.1) \qquad\qquad h_{\mathrm{top}}(f) = \log(\deg f),$$

and there exists a unique measure of maximal entropy μ_f for f (see [BM10, Theorem 20.9] and [HP09, Section 3.4 and Section 3.5]). Moreover, for each $n \in \mathbb{N}$, the unique measure of maximal entropy μ_{f^n} of the expanding Thurston map f^n is equal to μ_f (see [BM10, Theorem 20.7 and Theorem 20.9]).

We recall that in a compact metric space (X, d), a sequence of finite Borel measures μ_n converges in the weak* topology to a finite Borel measure μ, or $\mu_n \xrightarrow{w^*} \mu$, as $n \longrightarrow +\infty$ if and only if $\lim\limits_{n \to +\infty} \int u \, \mathrm{d}\mu_n = \int u \, \mathrm{d}\mu$ for each $u \in C(X)$.

We need the following lemmas for weak* convergence.

Lemma 5.1. *Let X and $\widetilde{X}$ be two compact metric spaces and $\phi \colon X \to \widetilde{X}$ a continuous map. Let μ and μ_i, for $i \in \mathbb{N}$, be finite Borel measures on X. If*

$$\mu_i \xrightarrow{w^*} \mu \ \text{as } i \longrightarrow +\infty,$$

then $\phi_*(\mu)$ and $\phi_*(\mu_i)$, $i \in \mathbb{N}$, are finite Borel measures on $\widetilde{X}$, and

$$\phi_*(\mu_i) \xrightarrow{w^*} \phi_*(\mu) \ \text{as } i \longrightarrow +\infty.$$

Recall for a continuous map $\phi \colon X \to \widetilde{X}$ between two metric spaces and a Borel measure ν on X, the *push-forward* $\phi_*(\nu)$ of ν by ϕ is defined to be the unique Borel measure that satisfies $(\phi_*(\nu))(B) = \nu(\phi^{-1}(B))$ for each Borel set $B \subseteq \widetilde{X}$.

Proof. By the Riesz representation theorem (see for example, [Fo99, Chapter 7]), the lemma follows if we observe that for each $h \in C(\widetilde{X})$, we have

$$\int_{\widetilde{X}} h \, d\phi_*(\mu_i) = \int_X (h \circ \phi) \, d\mu_i \xrightarrow{i \to +\infty} \int_X (h \circ \phi) \, d\mu = \int_{\widetilde{X}} h \, d\phi_*(\mu).$$

$\square$

Lemma 5.2. *Let (X, d) be a compact metric space, and I be a finite set. Suppose that μ and $\mu_{i,n}$, for $i \in I$ and $n \in \mathbb{N}$, are finite Borel measures on X, and $w_{i,n} \in [0, +\infty)$, for $i \in I$ and $n \in \mathbb{N}$ such that*

(1) $\mu_{i,n} \xrightarrow{w^*} \mu$, *as $n \longrightarrow +\infty$, for each $i \in I$,*

(2) $\lim\limits_{n \to +\infty} \sum\limits_{i \in I} w_{i,n} = r$ *for some $r \in \mathbb{R}$.*

Then $\sum\limits_{i \in I} w_{i,n} \mu_{i,n} \xrightarrow{w^*} r\mu$ as $n \longrightarrow +\infty$.

Proof. For each $u \in C(X)$ and each $n \in \mathbb{N}$,

$$\left| \int u \, d\left(\sum_{i \in I} w_{i,n} \mu_{i,n} \right) - r \int u \, d\mu \right|$$

$$\leq \sum_{i \in I} w_{i,n} \left| \int u \, d\mu_{i,n} - \int u \, d\mu \right| + \left| r - \sum_{i \in I} w_{i,n} \right| \int |u| \, d\mu.$$

Since $\mu_{i,n} \xrightarrow{w^*} \mu$, as $n \longrightarrow +\infty$, for each $i \in I$, and $\lim\limits_{n \to +\infty} \sum\limits_{i \in I} w_{i,n} = r$, we can conclude that the right-hand side of the inequality above tends to 0 as $n \longrightarrow +\infty$. $\square$

We record the following well-known lemma, sometimes known as the Portmanteau Theorem, and refer the reader to [Bi99, Theorem 2.1] for the proof.

Lemma 5.3. *Let (X, d) be a compact metric space, and μ and μ_i, for $i \in \mathbb{N}$, be Borel probability measures on X. Then the following are equivalent:*

(1) $\mu_i \xrightarrow{w^*} \mu$ *as $i \longrightarrow +\infty$;*

(2) $\limsup\limits_{i \to +\infty} \mu_i(F) \leq \mu(F)$ *for each closed set $F \subseteq X$;*

(3) $\liminf\limits_{i \to +\infty} \mu_i(G) \geq \mu(G)$ *for each open set $G \subseteq X$;*

(4) $\lim\limits_{i \to +\infty} \mu_i(B) = \mu(B)$ *for each Borel set $B \subseteq X$ with $\mu(\partial B) = 0$.*

Lemma 5.4. *Let (X, d) be a compact metric space. Suppose that $A_i \subseteq X$, for $i \in \mathbb{N}$, are finite subsets of X with maps $\phi_i \colon A_i \to X$ such that*

$$\lim_{i \to +\infty} \max\{d(x, \phi_i(x)) \mid x \in A_i\} = 0.$$

Let $m_i \colon A_i \to \mathbb{R}$, for $i \in \mathbb{N}$, be functions that satisfy

$$\sup_{i \in \mathbb{N}} \|m_i\|_1 = \sup_{i \in \mathbb{N}} \sum_{x \in A_i} |m_i(x)| < +\infty.$$

Define for each $i \in \mathbb{N}$,

$$\mu_i = \sum_{x \in A_i} m_i(x) \delta_x, \quad \widetilde{\mu}_i = \sum_{x \in A_i} m_i(x) \delta_{\phi_i(x)}.$$

If $\mu_i \xrightarrow{w^*} \mu$ as $i \longrightarrow +\infty$, for some finite Borel measure μ on X, then $\widetilde{\mu}_i \xrightarrow{w^*} \mu$ as $i \longrightarrow +\infty$.

Proof. It suffices to prove that for each continuous function $g \in C(X)$,

$$\int g \, \mathrm{d}\mu_i - \int g \, \mathrm{d}\widetilde{\mu}_i \longrightarrow 0 \text{ as } i \longrightarrow +\infty.$$

Indeed, g is uniformly continuous, so for each $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for each $n > N$ and for each $x \in A_n$, we have $|g(x) - g(\phi_n(x))| < \epsilon$. Thus

$$\left| \int g \, \mathrm{d}\mu_n - \int g \, \mathrm{d}\widetilde{\mu}_n \right| \leq \sum_{x \in A_n} |g(x) - g(\phi_n(x))| \, |m_n(x)| \leq \epsilon \sup_{n \in \mathbb{N}} \|m_n\|_1.$$

$\square$

The following lemma is a reformulation of Lemma 20.2 in [BM10]. We will later generalize it in Lemma 5.12.

Lemma 5.5 (M. Bonk & D. Meyer, 2010). *Let f be an expanding Thurston map, and $\mathcal{C} \subseteq S^2$ be an f^N-invariant Jordan curve containing post f for some $N \in \mathbb{N}$. Then there exists a constant $L_0 \in [1, \deg f)$ with the following property:*

For each $m \in \mathbb{N}_0$ with $m \equiv 0 \pmod{N}$, there exists a constant $C_0 > 0$ such that for each $k \in \mathbb{N}_0$ with $k \equiv 0 \pmod{N}$ and each m-edge e, there exists a collection M_0 of $(m+k)$-tiles with $\operatorname{card} M_0 \leq C_0 L_0^k$ and $e \subseteq \operatorname{int}\left(\bigcup_{X \in M_0} X \right)$.

Let F be an expanding Thurston map with an F-invariant Jordan curve $\mathcal{C} \subseteq S^2$ containing post F. As before, we let $b_w = \operatorname{card} \mathbf{X}_{bw}^1$ be

the number of black 1-tiles contained in the white 0-tile, and similarly, $w_w = \operatorname{card} \mathbf{X}_{ww}^1$, $w_b = \operatorname{card} \mathbf{X}_{wb}^1$, and $b_b = \operatorname{card} \mathbf{X}_{bb}^1$. We define

$$(5.2) \qquad w = \frac{b_w}{b_w + w_b}, \quad b = \frac{w_b}{b_w + w_b}.$$

Note that (see the discussion in [BM10] proceeding Lemma 20.1 in Chapter 20) $b_w, w_b, w, b > 0$, $w + b = 1$, and

$$(5.3) \qquad |w_w - b_w| < \deg F.$$

M. Bonk and D. Meyer gave the following characterization of the unique measure of maximal entropy of F (see [BM10, Proposition 20.7 and Theorem 20.9]):

Theorem 5.6 (M. Bonk & D. Meyer, 2010). *Let F be an expanding Thurston map with an F-invariant Jordan curve $\mathcal{C} \subseteq S^2$. Then there is a unique measure of maximal entropy μ_F of F, which is characterized among all Borel probability measures by the following property:*
For each $n \in \mathbb{N}_0$ and each n-tile $X^n \in \mathbf{X}^n(F, \mathcal{C})$,

$$(5.4) \qquad \mu(X^n) = \begin{cases} w(\deg F)^{-n} & \text{if } X^n \in \mathbf{X}_w^n(F, \mathcal{C}), \\ b(\deg F)^{-n} & \text{if } X^n \in \mathbf{X}_b^n(F, \mathcal{C}). \end{cases}$$

We now state our first characterization of the measure of maximal entropy μ_f of an expanding Thurston map f.

Theorem 5.7. *Let f be an expanding Thurston map with its measure of maximal entropy μ_f. Let $\mathcal{C} \subseteq S^2$ be an f^n-invariant Jordan curve containing $\operatorname{post} f$ for some $n \in \mathbb{N}$. Fix a visual metric d for f. Consider any sequence of non-negative numbers $\{\alpha_i\}_{i \in \mathbb{N}_0}$ with $\lim_{i \to +\infty} \alpha_i = 0$, and any sequence of maps $\{\beta_i\}_{i \in \mathbb{N}_0}$ with β_i sending each white i-tile $X^i \in \mathbf{X}_w^i(f, \mathcal{C})$ to a point $\beta_i(X^i) \in N_d^{\alpha_i}(X^i)$. Let*

$$\mu_i = \frac{1}{(\deg f)^i} \sum_{X^i \in \mathbf{X}_w^i(f, \mathcal{C})} \delta_{\beta_i(X^i)}, \quad i \in \mathbb{N}_0.$$

Then $\mu_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$.

Recall that $N_d^{\alpha_i}(X^i)$ denotes the open α_i-neighborhood of X^i in (S^2, d). This theorem says that a sequence of probability measures $\{\mu_i\}_{i \in \mathbb{N}}$, with μ_i assigning the same weight to a point near each white i-tile, converges in the weak* topology to the measure of maximal entropy. In some sense, it asserts the equidistribution of the white i-tiles with respect to the measure of maximal entropy.

We first prove a weaker version of the above theorem.

Proposition 5.8. *Let F be an expanding Thurston map with its measure of maximal entropy μ_F and an F-invariant Jordan curve $\mathcal{C} \subseteq S^2$ containing $\mathrm{post}\,F$. Consider any sequence of maps $\{\beta_i\}_{i \in \mathbb{N}_0}$ with β_i sending each white i-tile $X^i \in \mathbf{X}_w^i(F,\mathcal{C})$ to a point $\beta_i(X^i) \in \mathrm{inte}(X^i)$ for each $i \in \mathbb{N}_0$. Let*

$$\mu_i = \frac{1}{(\deg F)^i} \sum_{X^i \in \mathbf{X}_w^i(F,\mathcal{C})} \delta_{\beta_i(X^i)}, \quad i \in \mathbb{N}_0.$$

Then $\mu_i \xrightarrow{w^*} \mu_F$ as $i \longrightarrow +\infty$.

Proof. Note that $\mathrm{card}\,\mathbf{X}_w^i = (\deg F)^i$, so μ_i is a probability measure for each $i \in \mathbb{N}_0$. Thus by Alaoglu's theorem, it suffices to prove that for each Borel measure μ which is a subsequential limit of $\{\mu_i\}_{i \in \mathbb{N}_0}$ in the weak* topology, we have $\mu = \mu_F$.

Let $\{i_n\}_{n \in \mathbb{N}} \subseteq \mathbb{N}$ be an arbitrary strictly increasing sequence such that $\mu_{i_n} \xrightarrow{w^*} \mu$ as $n \longrightarrow +\infty$, for some Borel measure μ. Clearly μ is also a probability measure.

Recall the definitions of $w, b \in (0,1)$ and w_w, b_w, w_b, b_b (see (5.2)). For each $m, i \in \mathbb{N}_0$ with $0 \le m \le i$, each white m-tile $X_w^m \in \mathbf{X}_w^m$, and each black m-tile $X_b^m \in \mathbf{X}_b^m$, by the formulas in Lemma 20.1 in [BM10], we have

$$
\begin{aligned}
\mu_i(X_w^m) &= \frac{1}{(\deg F)^i}\,\mathrm{card}\{X^i \in \mathbf{X}_w^i \mid X^i \subseteq X_w^m\} \\
(5.5) \qquad &= \frac{1}{(\deg F)^i}\left(w(\deg F)^{i-m} + b(w_w - b_w)^{i-m}\right),
\end{aligned}
$$

and similarly,

$$(5.6) \qquad \mu_i(X_b^m) = \frac{1}{(\deg F)^i}\left(w(\deg F)^{i-m} - b(w_w - b_w)^{i-m}\right).$$

We claim that for each m-tile $X^m \in \mathbf{X}^m$ with $m \in \mathbb{N}_0$, we have $\mu(\partial X^m) = 0$.

To establish the claim, by Proposition 3.5(vi), it suffices to prove that $\mu(e) = 0$ for each m-edge e with $m \in \mathbb{N}_0$. Applying Lemma 5.5 in the case $f = F$ and $n = 1$, we get that there exist constants $1 < L_0 < \deg F$ and $C_0 > 0$ such that for each $k \in \mathbb{N}_0$, there is a collection M_0^k of $(m + k)$-tiles with $\mathrm{card}\,M_0^k \le C_0 L_0^k$ such that e is contained in the interior of the set $\bigcup_{X \in M_0^k} X$. So by (5.3), (5.5), (5.6),

and Lemma 5.3, we get

$$\mu(e) \leq \mu\left(\text{int}\left(\bigcup_{X \in M_0^k} X \right) \right) \leq \limsup_{l \to +\infty} \mu_{m+k+l}\left(\text{int}\left(\bigcup_{X \in M_0^k} X \right) \right)$$

$$\leq \limsup_{l \to +\infty} \sum_{X \in M_0^k} \mu_{m+k+l}(X) \leq \sum_{X \in M_0^k} \limsup_{l \to +\infty} \mu_{m+k+l}(X)$$

$$\leq C_0 L_0^k \frac{w + b}{(\deg F)^{m+k}}.$$

By letting $k \longrightarrow +\infty$, we get $\mu(e) = 0$, proving the claim.

Thus by (5.3), (5.5), (5.6), the claim, and Lemma 5.3, we can conclude that for each $m \in \mathbb{N}_0$, and each white m-tile $X_w^m \in \mathbf{X}_w^m$, each black m-tile $X_b^m \in \mathbf{X}_b^m$, we have that

$$\mu(X_w^m) = \lim_{n \to +\infty} \mu_{i_n}(X_w^m) = w(\deg F)^{-m},$$

$$\mu(X_b^m) = \lim_{n \to +\infty} \mu_{i_n}(X_b^m) = b(\deg F)^{-m}.$$

By Theorem 5.6, therefore, the measure μ is equal to the unique measure of maximal entropy μ_F of F. $\qquad\square$

As a consequence of the above proposition, we have

Corollary 5.9. *Let f be an expanding Thurston map with its measure of maximal entropy μ_f. Let $\mathcal{C} \subseteq S^2$ be an f^n-invariant Jordan curve containing $\operatorname{post} f$ for some $n \in \mathbb{N}$. Fix an arbitrary $p \in \operatorname{inte}(X_w^0)$ where X_w^0 is the white 0-tile for $(f, \mathcal{C})$. Define, for $i \in \mathbb{N}$,*

$$\nu_i = \frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(p)} \delta_q.$$

Then $\nu_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$.

Proof. First observe that since p is contained in the interior of the white 0-tile, each $q \in f^{-n}(p)$ is contained in the interior of one of the white n-tiles, and each white n-tile contains exactly one q with $f^n(q) = p$. So by Proposition 5.8,

$$(5.7) \qquad\qquad \nu_{ni} \xrightarrow{w^*} \mu_{f^n} \text{ as } i \longrightarrow +\infty,$$

where μ_{f^n} is the unique measure of maximal entropy of f^n, which is equal to μ_f (see [BM10, Theorem 20.7 and Theorem 20.9]).

Then note that for $k > 1$,

$$(5.8) \quad f_* \nu_k = \frac{1}{(\deg f)^k} \sum_{q \in f^{-k}(p)} \delta_{f(q)} = \frac{1}{(\deg f)^{k-1}} \sum_{q \in f^{-k+1}(p)} \delta_q = \nu_{k-1}.$$

So by (5.7), (5.8), Lemma 5.1, and the fact that μ_f is invariant under pushforward of f from Theorem 20.9 in [BM10], for each $k \in \{0, 1, \ldots, n-1\}$, we get

$$\nu_{ni-k} = (f_*)^k \nu_{ni} \xrightarrow{w^*} (f_*)^k \mu_f = \mu_f \text{ as } i \longrightarrow +\infty.$$

Therefore $\nu_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$. $\square$

Proof of Theorem 5.7. Fix an arbitrary $p \in \text{inte}(X_w^0)$ in the interior of the while 0-tile X_w^0 for the cell decomposition induced by $(f, \mathcal{C})$.

As in the proof of Corollary 5.9, for each $i \in \mathbb{N}_0$, there is a bijective correspondence between points in $f^{-i}(p)$ and the set of white i-tiles, namely, each $q \in f^{-i}(p)$ corresponds to the unique white i-tile, denoted by X_q, containing q. Then we define functions $\phi_i \colon f^{-i}(p) \to S^2$ by setting $\phi_i(q) = \beta_i(X_q)$.

Let $\Lambda > 1$ be the expansion factor of our fixed visual metric d. There exists $C \geq 1$ such that for each $n \in \mathbb{N}_0$ and each n-tile $X^n \in \mathbf{X}^n$, $\text{diam}_d(X^n) \leq C\Lambda^{-n}$ (see Lemma 3.10). So for each $i \in N_0$ and each $q \in f^{-i}(p)$, we have

$$d(q, \phi_i(q)) \leq d(\phi_i(q), X_q) + \text{diam}_d(X_q) \leq \alpha_i + C\Lambda^{-i}.$$

Thus $\lim\limits_{i \to +\infty} \max\{d(x, \phi_i(x)) \mid x \in f^{-i}(p)\} = 0$.

For $i \in \mathbb{N}_0$, define

$$\widetilde{\mu}_i = \frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(p)} \delta_q.$$

Note that for $i \in \mathbb{N}_0$,

$$\mu_i = \frac{1}{(\deg f)^i} \sum_{X^i \in \mathbf{X}_w^i(f, \mathcal{C})} \delta_{\beta_i(X^i)} = \frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(p)} \delta_{\phi_i(q)}.$$

Then by Corollary 5.9, $\widetilde{\mu}_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$. Therefore, by Lemma 5.4 with $A_i = f^{-i}(p)$ and $m_i(x) = \frac{1}{(\deg f)^i}$, $i \in \mathbb{N}_0$, we can conclude that $\mu_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$. $\square$

Remarks 5.10. We can replace "white" by "black", $\mathbf{X}_w^i$ by $\mathbf{X}_b^i$, and X_w^0 by X_b^0 in the statements of Theorem 5.7, Proposition 5.8, and Corollary 5.9. The proofs are essentially the same.

We are now ready to prove the equidistribution of preimages of an arbitrary point with respect to the measure of maximal entropy μ_f.

Proof of Theorem 1.2. By Theorem 1.2 in [BM10] or Corollary 3.12, we can fix an f^n-invariant Jordan curve $\mathcal{C} \subseteq S^2$ containing post f for some $n \in \mathbb{N}$. We consider the cell decompositions induced by $(f, \mathcal{C})$.

We first prove (1.2).

We assume that p is contained in the (closed) white 0-tile. The proof for the case when p is contained in the black 0-tile is exactly the same except that we need to use a version of Theorem 5.7 for black tiles instead of using Theorem 5.7 literally, see Remark 5.10.

Observe that for each $i \in \mathbb{N}_0$ and each $q \in f^{-i}(p)$, the number of white i-tiles that contains q is exactly $\deg_{f^i}(q)$. On the other hand, each white i-tile contains exactly one point q with $f^i(q) = p$. So we can define $\beta_i \colon \mathbf{X}_w^i \to S^2$ by mapping a white i-tile to the point q in it that satisfies $f^i(q) = p$. Define $\alpha_i \equiv 0$. Theorem 5.7 applies, and thus (1.2) is true.

Next, we prove (1.3). The proof breaks into three cases.

Case 1. Assume that $p \notin \operatorname{post} f$. Then $\deg_f(x) = 1$ for all $x \in \bigcup_{n=1}^{+\infty} f^{-n}(p)$. So $\widetilde{\nu}_i = \nu_i$ for each $i \in \mathbb{N}$. Then (1.3) follows from (1.2) in this case.

Case 2. Assume that $p \in \operatorname{post} f$ and p is not periodic. Then there exists $N \in \mathbb{N}$ such that $f^{-N}(p) \cap \operatorname{post} f = \emptyset$. For otherwise, there exists a point $z \in \operatorname{post} f$ which belongs to $f^{-c}(p)$ for infinitely many distinct $c \in \mathbb{N}$. In particular, there exist two integers $a > b > 0$ such that $z \in f^{-a}(p) \cap f^{-b}(p)$. Then $f^{a-b}(p) = p$, a contradiction. So $\deg_f(q) = 1$ for each $q \in \bigcup_{x \in f^{-N}(p)} \bigcup_{i=1}^{+\infty} f^{-i}(x)$. Note that for each $x \notin \operatorname{post} f$ and each $i \in \mathbb{N}$, the number of preimages of x under f^i is exactly $(\deg f)^i$. Then for each $i \in \mathbb{N}$, $Z_{i+N} = Z_N(\deg f)^i$, and

$$\widetilde{\nu}_{i+N} = \frac{1}{Z_{i+N}} \sum_{q \in f^{-(i+N)}(p)} \delta_q = \frac{1}{Z_N} \sum_{x \in f^{-N}(p)} \left(\frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(x)} \delta_q \right).$$

For each $x \in f^{-N}(p)$, by Case 1,

$$\frac{1}{(\deg f)^i} \sum_{q \in f^{-i}(x)} \delta_q \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty.$$

Thus each term in the sequence $\{\widetilde{\nu}_{i+N}\}_{i \in \mathbb{N}}$ is a convex combination of the corresponding terms in sequences of measures, each of which converges to μ_f in the weak* topology. Hence by Lemma 5.2, the sequence $\{\widetilde{\nu}_{i+N}\}_{i \in \mathbb{N}}$ also converges to μ_f in the weak* topology in this case.

Case 3. Assume that $p \in \operatorname{post} f$ and p is periodic with period $k \in \mathbb{N}$. Let $l = \operatorname{card}(\operatorname{post} f)$. We first note that for each $m, N \in \mathbb{N}$, the

inequality $Z_{m+N} \geq (Z_m - l)(\deg f)^N$, and equivalently,

$$\frac{Z_m}{Z_{m+N}} \leq \frac{1}{(\deg f)^N} + \frac{l}{Z_{m+N}}$$

hold, since there are at most l points in $Z_m \cap \text{post } f$. So by Lemma 3.13, for each $\epsilon > 0$ and each N large enough such that $1/(\deg f)^N < \epsilon/2$ and $l/Z_{m+N} < \epsilon/2$, we get $Z_m/Z_{m+N} < \epsilon$ for each $m \in \mathbb{N}$. We fix $j \in \mathbb{N}$ large enough such that $Z_{m-jk}/Z_m < \epsilon$ for each $m > jk$. Observe that for each $m > jk$,

$$\widetilde{\nu}_m = \frac{1}{Z_m} \sum_{q \in f^{-m}(p)} \delta_q$$

$$(5.9) \qquad = \frac{1}{Z_m} \left(\sum_{q \in f^{-(m-jk)}(p)} \delta_q + \sum_{x \in f^{-jk}(p) \backslash \{p\}} \sum_{q \in f^{-(m-jk)}(x)} \delta_q \right)$$

$$= \frac{Z_{m-jk}}{Z_m} \left(\frac{1}{Z_{m-jk}} \sum_{q \in f^{-(m-jk)}(p)} \delta_q \right) + \frac{1}{Z_m} \sum_{x \in f^{-jk}(p) \backslash \{p\}}$$

$$\text{card}\left(f^{-(m-jk)}(x) \right) \left(\frac{1}{\text{card}\left(f^{-(m-jk)}(x) \right)} \sum_{q \in f^{-(m-jk)}(x)} \delta_q \right).$$

Note that no point $x \in f^{-jk}(p) \backslash \{p\}$ is periodic. Indeed, if $x \in f^{-jk}(p) \backslash \{p\}$ were periodic, then $x \in \bigcup_{i=0}^{k-1} f^i(p)$, and so x would have period k as well. Thus $x = f^{jk}(x) = p$, a contradiction. Hence by Case 1 and Case 2, for each $x \in f^{-jk}(p) \backslash \{p\}$,

$$\frac{1}{\text{card}\left(f^{-(m-jk)}(x) \right)} \sum_{q \in f^{-(m-jk)}(x)} \delta_q \xrightarrow{w^*} \mu_f \text{ as } m \longrightarrow +\infty.$$

Let $\mu \in \mathcal{P}(S^2)$ be an arbitrary subsequential limit of $\{\widetilde{\nu}_m\}_{m \in \mathbb{N}}$ in the weak* topology. For each strictly increasing sequence $\{m_i\}_{i \in \mathbb{N}}$ in $\mathbb{N}$ that satisfies $\widetilde{\nu}_{m_i} \xrightarrow{w^*} \mu$ as $i \longrightarrow +\infty$, we can assume, due to Alaoglu's Theorem, by choosing a subsequence if necessary, that

$$\frac{Z_{m_i-jk}}{Z_{m_i}} \left(\frac{1}{Z_{m_i-jk}} \sum_{q \in f^{-(m_i-jk)}(p)} \delta_q \right) \xrightarrow{w^*} \eta \text{ as } i \longrightarrow +\infty,$$

for some Borel measure η with total variation $\|\eta\| \leq \epsilon$. Observe that for each $i \in \mathbb{N}$,

$$\frac{1}{Z_{m_i}} \sum_{x \in f^{-jk}(p) \backslash \{p\}} \text{card}\left(f^{-(m_i-jk)}(x) \right) = 1 - \frac{Z_{m_i-jk}}{Z_{m_i}},$$

since $p \in f^{-jk}(p)$ and $\mathrm{card}\left(f^{-(m_i-jk)}(p)\right) = Z_{m_i-jk}$. By choosing a subsequence of $\{m_i\}_{i \in \mathbb{N}}$ if necessary, we can assume that there exists $r \in [0, \epsilon]$ such that

$$\lim_{i \to +\infty} \frac{Z_{m_i-jk}}{Z_{m_i}} = r.$$

So by taking the limits of both sides of (5.9) in the weak* topology along the subsequence $\{m_i\}_{i \in \mathbb{N}}$, we get from Lemma 5.2 that $\mu = \eta + (1-r)\mu_f$. Thus

$$\|\mu - \mu_f\| \leq \|\eta\| + r \|\mu_f\| \leq 2\epsilon.$$

Since ϵ is arbitrary, we can conclude that $\mu = \mu_f$. We have proven in this case that each subsequential limit of $\{\widetilde{\nu}_m\}_{m \in \mathbb{N}}$ in the weak* topology is equal to μ_f. Therefore (1.3) is true in this case.   □

In order to prove Theorem 1.3, we will need Lemma 5.12 which is a generalization of Lemma 5.5.

Lemma 5.11. *Let f be an expanding Thurston map and $d = \deg f$. Then there exist constants $C > 0$ and $\alpha \in (0, 1]$ such that for each nonempty finite subset M of S^2 and each $n \in \mathbb{N}$, we have*

$$(5.10) \qquad \frac{1}{d^n} \sum_{x \in M} \deg_{f^n}(x) \leq C \max\left\{ \left(\frac{\mathrm{card}\, M}{d^n} \right)^\alpha, \frac{\mathrm{card}\, M}{d^n} \right\}.$$

Note that when $\mathrm{card}\, M \leq d^n$, the right-hand side of (5.10) becomes $C \left(\frac{\mathrm{card}\, M}{d^n} \right)^\alpha$.

Proof. Let $m = \mathrm{card}\, M$. Set

$$D = \prod_{x \in \mathrm{crit}\, f} \deg_f(x).$$

In order to establish the lemma, we consider the following three cases.

Case 1: Suppose that f has no periodic critical points. Then since for each $x \in S^2$ and each $n \in \mathbb{N}$,

$$(5.11) \qquad \deg_{f^n}(x) = \deg_f(x)\deg_f(f(x)) \cdots \deg_f(f^{n-1}(x)),$$

it is clear that $\deg_{f^n}(x) \leq D$. So

$$\frac{1}{d^n} \sum_{x \in M} \deg_{f^n}(x) \leq D\frac{m}{d^n}.$$

Thus in this case, $C = D$ and $\alpha = 1$.

Case 2: Suppose that f has periodic critical points, but all periodic critical points are fixed points of f.

Let $T_0 = \{x \in \operatorname{crit} f \mid f(x) = x\}$ be the set of periodic critical points of f. Then define recursively for each $i \in \mathbb{N}$,

$$T_i = f^{-1}(T_{i-1}) \setminus \bigcup_{j=0}^{i-1} T_j.$$

Define $T_{-1} = S^2 \setminus \bigcup\limits_{j=0}^{+\infty} T_j$, and $\widetilde{T}_i = S^2 \setminus \bigcup\limits_{j=0}^{i} T_j$ for each $i \in \mathbb{N}_0$. Set $t_0 = \operatorname{card} T_0$. Since $T_0 \subseteq \operatorname{post} f$, we have $1 \leq t_0 < +\infty$. Then for each $i \in \mathbb{N}$, we have

$$\operatorname{card} T_i \leq d^i t_0.$$

We note that if $\deg_f(x) = d$ for some $x \in T_0$, then $f^{-i}(x) = \{x\}$ for each $i \in \mathbb{N}$, contradicting Lemma 3.13. So $\deg_f(x) \leq d - 1$ for each $x \in T_0$. Thus for each $x \in T_0$ and each $m \in \mathbb{N}$, we have

$$\deg_{f^m}(x) \leq (d - 1)^m.$$

Moreover, for each $i, m \in \mathbb{N}$ with $i < m$ and each $x \in T_i$, we get

$$\deg_{f^m}(x) = \deg_f(x) \deg_f(f(x)) \cdots \deg_f(f^{i-1}(x)) \deg_{f^{m-i}}(f^i(x))$$
$$\leq D(d - 1)^{m-i}.$$

Similarly, for each $i, m \in \mathbb{N}$ with $i \geq m$ and each $x \in \widetilde{T}_i$, we have

$$\deg_{f^m}(x) \leq D.$$

Thus for each $n \in \mathbb{N}$,

$$\frac{1}{d^n} \sum_{x \in M} \deg_{f^n}(x)$$

$$= \frac{1}{d^n} \sum_{j=-1}^{+\infty} \sum_{x \in M \cap T_j} \deg_{f^n}(x)$$

$$\leq \frac{1}{d^n} \left(\sum_{j=0}^{n} \sum_{x \in M \cap T_j} D(d - 1)^{n-j} + \sum_{x \in M \cap \widetilde{T}_n} D \right).$$

Note that the more points in M lie in T_j with $j \in [0, n]$ as small as possible, the larger the right-hand side of the last inequality is. So the right-hand side of the last inequality is

$$\leq \frac{1}{d^n} \left(\sum_{j=0}^{\lceil \log_d \lceil \frac{m}{t_0} \rceil \rceil} (\operatorname{card} T_j) D(d - 1)^{n-j} + mD \right)$$

$$\leq \frac{Dt_0}{d^n} \sum_{j=0}^{\lceil \log_d \lceil \frac{m}{t_0} \rceil \rceil} d^j (d-1)^{n-j} + \frac{mD}{d^n}$$

$$\leq Dt_0 \left(\frac{d-1}{d} \right)^n \sum_{j=0}^{\lceil \log_d m \rceil} \left(\frac{d}{d-1} \right)^j + \frac{mD}{d^n}$$

$$= Dt_0 \left(\frac{d-1}{d} \right)^n \frac{\left(\frac{d}{d-1} \right)^{\lceil \log_d m \rceil + 1} - 1}{\frac{d}{d-1} - 1} + \frac{mD}{d^n}$$

$$\leq Dt_0 \left(\frac{d-1}{d} \right)^n \left(\frac{d}{d-1} \right)^{2 + \log_d m} (d-1) + \frac{mD}{d^n}$$

$$\leq Dt_0 \frac{d^2}{d-1} \left(\left(\frac{d-1}{d} \right)^{n - \log_d m} + \frac{m}{d^n} \right)$$

$$= \frac{1}{2} E_f \left(d^{(n - \log_d m) \log_d \frac{d-1}{d}} + \frac{m}{d^n} \right)$$

$$\leq E_f \max \left\{ \left(\frac{m}{d^n} \right)^{\log_d \frac{d}{d-1}}, \frac{m}{d^n} \right\},$$

where $E_f = 2Dt_0 \frac{d^2}{d-1}$ is a constant that only depends on f. Thus in this case, $C = E_f$ and $\alpha = \log_d \frac{d}{d-1} \in (0, 1]$.

Case 3: Suppose that f has periodic critical points that may not be fixed points of f.

Set κ to be the product of the periods of all periodic critical points of f. It is easy to see that each periodic critical point of f^κ is a fixed point of f^κ.

Note that for each $n \in \mathbb{N}$,

$$\frac{1}{d^n} \sum_{x \in M} \deg_{f^n}(x) \leq d^\kappa \frac{1}{d^{\kappa \lceil \frac{n}{\kappa} \rceil}} \sum_{x \in M} \deg_{f^{\kappa \lceil \frac{n}{\kappa} \rceil}}(x).$$

Hence by applying Case 2 for f^κ, we get a constant E_{f^κ} that depends only on f, such that the right-hand side of the above inequality is

$$\leq d^\kappa E_{f^\kappa} \max \left\{ \left(\frac{m}{d^{\kappa \lceil \frac{n}{\kappa} \rceil}} \right)^{\log_{d^\kappa} \frac{d^\kappa}{d^\kappa - 1}}, \frac{m}{d^{\kappa \lceil \frac{n}{\kappa} \rceil}} \right\}$$

$$\leq d^\kappa E_{f^\kappa} \max \left\{ \left(\frac{m}{d^n} \right)^{\log_{d^\kappa} \frac{d^\kappa}{d^\kappa - 1}}, \frac{m}{d^n} \right\}.$$

Thus in this case $C = d^\kappa E_{f^\kappa}$ and $\alpha = \log_{d^\kappa} \frac{d^\kappa}{d^\kappa - 1} \in (0, 1]$. $\square$

Now we formulate a generalization of Lemma 5.5.

Lemma 5.12. *Let f be an expanding Thurston map, and $\mathcal{C} \subseteq S^2$ be an f^N-invariant Jordan curve containing* post f *for some $N \in \mathbb{N}$. Then there exists a constant $L \in [1, \deg f)$ with the following property:*

For each $m \in \mathbb{N}_0$, there exists a constant $D > 0$ such that for each $k \in \mathbb{N}_0$ and each m-edge e, there exists a collection M of $(m + k)$-tiles with card $M \leq DL^k$ *and $e \subseteq$ int $\left(\bigcup\limits_{X \in M} X \right)$.*

Proof. We denote $d = \deg f$, and consider the cell decompositions induced by $(f, \mathcal{C})$ in this proof.

Step 1: We first assume that for some $m \in \mathbb{N}$, there exist constants $L \in [1, d)$ and $D > 0$ such that for each $k \in \mathbb{N}_0$ and each m-edge e, there exists a collection M of $(m + k)$-tiles with card $M \leq DL^k$ and $e \subseteq$ int $\left(\bigcup\limits_{X \in M} X \right)$. Then by Proposition 3.5(i), for each $(m - 1)$-edge e, we can choose an m-edge e' such that $f(e') = e$. For each $k \in \mathbb{N}_0$, there exists a collection M' of $(m+k)$-tiles with card $M' \leq DL^k$ and $e' \subseteq$ int $\left(\bigcup\limits_{X \in M'} X \right)$. We set M to be the collection $\{ f(X) \mid X \in M' \}$ of $(m - 1 + k)$-tiles. Then card $M \leq$ card $M' \leq DL^k$ and $e \subseteq$ int $\left(\bigcup\limits_{X \in M} X \right)$. Hence, it suffices to prove the lemma for "each $m \in \mathbb{N}_0$ with $m \equiv 0 \pmod{N}$" instead of "each $m \in \mathbb{N}_0$".

Step 2: Let $i \in \mathbb{N}_0$ and $p \in S^2$ be an i-vertex. We define the *i-flower* $W^i(p)$ as in [BM10] by

$$W^i(p) = \bigcup \{ \text{inte}(c) \mid c \in \mathbf{D}^i, p \in c \},$$

By definition, the only i-vertex contained in $W^i(p)$ is p, and the interior inte(e) of an i-edge e is a subset of $W^i(p)$ if and only if $p \in e$. Note that the number of i-tiles in $W^i(p)$ is $2 \deg_{f^i}(p)$, i.e.,

$$(5.12) \qquad \text{card}\{ X \in \mathbf{X}^i \mid p \in X \} = 2 \deg_{f^i}(p).$$

By [BM10, Lemma 7.11], there exists a constant $\beta \in \mathbb{N}$, which depends only on f and $\mathcal{C}$, such that for each $i \in \mathbb{N}$ and each i-tile $X \in \mathbf{X}^i$, X can be covered by a union of at most β $(i + 1)$-flowers.

Let $C > 0$ and $\alpha \in (0, 1]$ be constants from Lemma 5.11, both of which depend only on f.

We will prove the following statement by induction on κ:

> Let f be an expanding Thurston map, and $\mathcal{C} \subseteq S^2$ be an f^N-invariant Jordan curve containing post f for some $N \in \mathbb{N}$. For each $\kappa \in \{0, 1, \dots, N - 1\}$, there exists a constant $L_\kappa \in [1, d)$, depending only on f, $\mathcal{C}$, α, and κ, with the following property:

For each $m \in \mathbb{N}_0$ with $m \equiv 0 \pmod{N}$, there exists a constant $D_\kappa > 0$, depending only on f, $\mathcal{C}$, L_0, β, α, C, κ, and m, such that for each $k \in \mathbb{N}_0$ with $k \equiv \kappa \pmod{N}$ and each m-edge e, there exists a collection $M_{m,k,e}$ of $(m+k)$-tiles that satisfies $\operatorname{card} M_{m,k,e} \leq D_\kappa L_\kappa^k$ and $e \subseteq \operatorname{int}\left(\bigcup_{X \in M_{m,k,e}} X\right)$.

Lemma 5.5 gives the case for $\kappa = 0$. The constant L_0 is the one from Lemma 5.5 that depends only on f and $\mathcal{C}$. For the induction step, we assume the above statement for some $\kappa \in [0, N-1]$.

Fix an arbitrary $m \in \mathbb{N}_0$ with $m \equiv 0 \pmod{N}$, and fix an arbitrary m-edge e.

By the induction hypothesis, there exist constants $D_\kappa > 0$ and $L_\kappa \in [1, d)$ such that for each $k \in \mathbb{N}_0$ with $k \equiv \kappa + 1 \pmod{N}$, there exists a collection $M_{m,k-1,e}$ of $(m+k-1)$-tiles with $\operatorname{card} M_{m,k-1,e} \leq D_\kappa L_\kappa^{k-1}$ and $e \subseteq \operatorname{int}\left(\bigcup_{X \in M_{m,k-1,e}} X\right)$. Each $X \in M_{m,k-1,e}$ can be covered by β $(m+k)$-flowers $W^{m+k}(p)$. We can then construct a set $F \subseteq \mathbf{V}^{m+k}$ of $(m+k)$-vertices such that

$$\text{(5.13)} \qquad \operatorname{card} F \leq \beta D_\kappa L_\kappa^{k-1}$$

and

$$\text{(5.14)} \qquad \bigcup_{X \in M_{m,k-1,e}} X \subseteq \bigcup_{p \in F} W^{m+k}(p).$$

We define

$$\text{(5.15)} \qquad M_{m,k,e} = \{X \in \mathbf{X}^{m+k} \mid X \cap F \neq \emptyset\}.$$

Then $e \subseteq \operatorname{int}\left(\bigcup_{X \in M_{m,k,e}} X\right)$, and by (5.12),

$$\text{(5.16)} \qquad \operatorname{card} M_{m,k,e} \leq \sum_{p \in F} 2 \deg_{f^{m+k}}(p).$$

Since $L_\kappa \in [1, d)$, there exists $K \in \mathbb{N}$, depending only on f, $\mathcal{C}$, L_0, β, α, C, m, and κ, such that for each $i \geq K$, we have $\beta D_\kappa L_\kappa^{i-1} \leq d^{m+i}$. We choose K to be the smallest such number.

Thus by (5.13), (5.16), and Lemma 5.11, for each $k \geq K$ with $k \equiv \kappa + 1 \pmod{N}$, we have

$$
\begin{aligned}
\text{(5.17)} \qquad \operatorname{card} M_{m,k,e} &\leq 2 \sum_{p \in F} \deg_{f^{m+k}}(p) \\
&\leq 2 C d^{(m+k)(1-\alpha)} \left(\beta D_\kappa L_\kappa^{k-1}\right)^\alpha \\
&= 2 C d^{m(1-\alpha)} \beta^\alpha D_\kappa^\alpha L_\kappa^{-\alpha} \left(d^{1-\alpha} L_\kappa^\alpha\right)^k.
\end{aligned}
$$

Recall the constants α and C are from Lemma 5.11 that depend only on f. Let $L_{\kappa+1} = d^{1-\alpha} L_\kappa^\alpha$. Since $L_\kappa \in [1, d)$, we get $L_{\kappa+1} \in [L_\kappa, d) \subseteq [1, d)$. Note that $L_{\kappa+1}$ only depends on f, $\mathcal{C}$, L_0, α, and κ. It is easy to see by induction that

$$\sum_{p\in\mathbf{V}^i} \deg_{f^i}(p) = \operatorname{card}(\operatorname{post} f) d^i$$

for $i \in \mathbb{N}$. We set

(5.18) $D_{\kappa+1} = \max\{2\operatorname{card}(\operatorname{post} f)d^{m+K}, 2Cd^{m(1-\alpha)}\beta^\alpha D_\kappa^\alpha L_\kappa^{-\alpha}\}.$

Then by (5.16), (5.17), and (5.18), we get that for each $k \in \mathbb{N}_0$ with $k \equiv \kappa + 1 \pmod{N}$,

(5.19) $\operatorname{card} M_{m,k,e} \leq \sum_{p\in F} 2 \deg_{f^{m+k}}(p) \leq D_{\kappa+1} L_{\kappa+1}^k.$

We note that $D_{\kappa+1}$ only depends on f, $\mathcal{C}$, L_0, β, α, C, m, and κ.

This completes the induction.

Step 3: For given f and $\mathcal{C}$, we now define

$$L = \max\{L_\kappa \,|\, \kappa \in \{0, 1, \ldots, N-1\}\}.$$

For each fixed $m \in \mathbb{N}_0$ with $m \equiv 0 \pmod{N}$, we set

$$D = \max\{D_\kappa \,|\, \kappa \in \{0, 1, \ldots, N-1\}\},$$

and for each given $k \in \mathbb{N}_0$ and $e \in \mathbf{E}^m$, let $M = M_{m,k,e}$. Then we have $\operatorname{card} M \leq DL^k$ and $e \subseteq \operatorname{int}\left(\bigcup_{X\in M} X\right)$. We note that here L only depends on f, $\mathcal{C}$, L_0, and α, and on the other hand, D only depends on f, $\mathcal{C}$, L_0, β, α, C, and m. Recall that the constants β from Lemma 7.11 in [BM10] and L_0 from Lemma 5.5 depend only on f and $\mathcal{C}$, and that the constants α and C from Lemma 5.11 depend only on f. The proof is now complete. $\square$

Remarks 5.13. It is also possible to prove the previous lemma by observing that $\mathcal{C}$ equipped with the restriction of a visual metric d for f is a quasicircle (see [BM10, Theorem 1.8]), and S^2 equipped with d is linearly locally connected (see [BM10, Proposition 16.3]). A metric space X, that is homeomorphic to the plane and with $\overline{X}$ linearly locally connected and ∂X a Jordan curve, has the property that ∂X is porous in $\overline{X}$ (see [Wi07, Theorem IV.14]). Then we can mimic the original proof of Lemma 20.2 in [BM10].

We are finally ready to prove the equidistribution of the preperiodic points with respect to the measure of maximal entropy μ_f.

Proof of Theorem 1.3. Fix an arbitrary $N \geq N(f)$ where $N(f)$ is an constant as given in Corollary 3.12 depending only on f. We also fix an f^N-invariant Jordan curve $\mathcal{C}$ containing post f such that no N-tile in $\mathbf{D}^N(f, \mathcal{C})$ joins opposite sides of $\mathcal{C}$ as given in Corollary 3.12. In the proof below, we consider the cell decompositions $\mathbf{D}^i(f, \mathcal{C}), i \in \mathbb{N}_0$, induced by $(f, \mathcal{C})$, and denote $d = \deg f$.

Since ξ_n^m and $\widetilde{\xi}_n^m$ are Borel probability measures for all $m \in \mathbb{N}_0$ and $n \in \mathbb{N}$ with $m < n$, by Alaoglu's Theorem, it suffices to prove that in the weak* topology, every convergent subsequence of $\{\xi_n^{m_n}\}_{n \in \mathbb{N}}$ and $\{\widetilde{\xi}_n^{m_n}\}_{n \in \mathbb{N}}$ converges to μ_f.

Proof of (1.5):

Let $\{n_i\}_{i \in \mathbb{N}}$ be a strictly increasing sequence with

$$\xi_{n_i}^{m_{n_i}} \xrightarrow{w^*} \mu \text{ as } i \longrightarrow +\infty,$$

for some measure μ.

Case 1 for (1.5): We assume in this case that there is no constant $K \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} \leq K$. Then by choosing a subsequence of $\{n_i\}_{i \in \mathbb{N}}$ if necessary, we can assume that $n_i - m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$.

Here is the idea of the proof in this case. By the spirit of Lemma 4.1 and Lemma 4.2, there is an almost bijective correspondence between the fixed points of f^{n-m_n} and the $(n-m_n)$-tiles containing such points. The correspondence is particularly nice away from $\mathcal{C}$. Thus there is almost a bijective correspondence between the preperiodic points in $S_n^{m_n}$ and the n-tiles containing such points. So if we can control the behavior near $\mathcal{C}$, then Theorem 5.7 applies and we finish the proof in this case. Finally the control we need is provided by Lemma 5.12.

Now we start to implement this idea. We fix a 0-edge $e_0 \subseteq \mathcal{C}$. Observe that for each $i \in \mathbb{N}$, we can pair a white i-tile $X_w^i \in \mathbf{X}_w^i$ and a black i-tile $X_b^i \in \mathbf{X}_b^i$ whose intersection $X_w^i \cap X_b^i$ is an i-edge contained in $f^{-i}(e_0)$. There are a total of d^i such pairs and each i-tile is in exactly one such pair. We denote by $\mathbf{P}_i$ the collection of the unions $X_w^i \cup X_b^i$ of such pairs, i.e.,

$$\mathbf{P}_i = \{X_w^i \cup X_b^i \mid X_w^i \in \mathbf{X}_w^i, X_b^i \in \mathbf{X}_b^i, X_w^i \cap X_b^i \cap f^{-i}(e_0) \in \mathbf{E}^i\}.$$

We denote $\mathbf{P}_i' = \{A \in \mathbf{P}_i \mid A \cap \mathcal{C} = \emptyset\}$.

By Lemma 5.12, there exists $1 \leq L < d$ and $C > 0$ such that for each $i \in \mathbb{N}$ there exists a collection M of i-tiles with card $M \leq CL^i$ such that $\mathcal{C}$ is contained in the interior of the set $\bigcup_{X \in M} X$. Note that L and C are constants independent of i. Observe that for each $A \in \mathbf{P}_i$ that does

not contain any i-tile in the collection M, we have $A \cap X \subseteq \partial \left(\bigcup_{X \in M} X \right)$
for each $X \in M$, so $A \cap \text{int} \left(\bigcup_{X \in M} X \right) = \emptyset$. Since the number of distinct $A \in \mathbf{P}_i$ that contains an i-tile in M is bounded above by CL^i, we get

$$(5.20) \qquad \text{card}(\mathbf{P}_i') \geq d^i - CL^i.$$

Note that for each $i \in \mathbb{N}$ and each $A \in \mathbf{P}_i'$, either $A \subseteq X_w^0$ or $A \subseteq X_b^0$ where X_w^0 (resp. X_b^0) is the white (resp. black) 0-tile for $(f, \mathcal{C})$. So by Proposition 3.5(i) and Brouwer's Fixed Point Theorem, there is a map $\tau \colon \mathbf{P}_i' \to P_{1,f^i}$ from $\mathbf{P}_i'$ to the set of fixed points of f^i such that $\tau(A) \in A$. Note if a fixed point x of f^i has weight $\deg_{f^i}(x) > 1$, then x has to be contained in post $f \subseteq \mathcal{C}$. Thus $\deg_{f^i}(\tau(A)) = 1$ for all $A \in \mathbf{P}_i'$.

If for some $A \in \mathbf{P}_i'$, the point $\tau(A)$ were on the boundaries of the two i-tiles whose union is A, then $\tau(A)$ would have to be contained in $\mathcal{C}$ since the boundaries are mapped into $\mathcal{C}$ under f^i. Thus for each $A \in \mathbf{P}_i'$, the point $\tau(A)$ is contained in the interior of one of the two i-tiles whose union is A. Hence τ is injective. Moreover,

$$(5.21) \quad \deg_{f^{i+j}}(x) = 1 \text{ for each } j \in \mathbb{N}_0 \text{ and each } x \in \bigcup_{A \in \mathbf{P}_i'} f^{-j}(\tau(A)).$$

For each $i \in \mathbb{N}$, we choose a map $\beta_{n_i} \colon \mathbf{X}_w^{n_i} \to S^2$ by letting $\beta_{n_i}(X)$ be the unique point in $f^{-m_{n_i}}(\tau(A)) \cap B$ where $B \in \mathbf{P}_{n_i}$ with $X \subseteq B$, if there exists $A \in \mathbf{P}_{n_i - m_{n_i}}'$ with $f^{m_{n_i}}(X) \subseteq A$; and by letting $\beta_{n_i}(X)$ be an arbitrary point in X if there exists no $A \in \mathbf{P}_{n_i - m_{n_i}}'$ with $f^{m_{n_i}}(X) \subseteq A$.

We fix a visual metric d for f with expansion factor $\Lambda > 1$. Note that Λ depends only on f and d. Then $\text{diam}_d(A) < c\Lambda^{-i}$ for each $i \in \mathbb{N}$, where $c \geq 1$ is a constant depending only on f, d, and $\mathcal{C}$ (See Lemma 3.10). Define $\alpha_n = c\Lambda^{-n}$ for each $n \in \mathbb{N}$. Thus α_{n_i} and β_{n_i} satisfy the hypothesis in Theorem 5.7. Define μ_{n_i} as in Theorem 5.7. Then

$$(5.22) \qquad \mu_{n_i} \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty,$$

by Theorem 5.7.

We claim that the total variation $\left\| \mu_{n_i} - \xi_{n_i}^{m_{n_i}} \right\|$ of $\mu_{n_i} - \xi_{n_i}^{m_{n_i}}$ converges to 0 as $i \longrightarrow +\infty$.

Assuming the claim, then by (5.22), we can conclude that (1.5) holds in this case.

To prove the claim, by Corollary 4.8, we observe that for each $i \in \mathbb{N}$,

$$
\left\| \mu_{n_i} - \xi_{n_i}^{m_{n_i}} \right\| \leq \left\| \mu_{n_i} - \frac{1}{d^{n_i - m_{n_i}}} \sum_{A \in \mathbf{P}'_{n_i - m_{n_i}}} \frac{1}{d^{m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q \right\|
$$

$$
(5.23) \qquad + \left\| \left(\frac{1}{d^{n_i}} - \frac{1}{d^{n_i} + d^{m_{n_i}}} \right) \sum_{A \in \mathbf{P}'_{n_i - m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q \right\|
$$

$$
+ \left\| \frac{1}{d^{n_i - m_{n_i}} + 1} \sum_{A \in \mathbf{P}'_{n_i - m_{n_i}}} \frac{1}{d^{m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q - \xi_{n_i}^{m_{n_i}} \right\|.
$$

In the first term on the right-hand side of (5.23), each δ_q in the summations cancels with the corresponding term in the definition of μ_{n_i}. So the first term on the right-hand side of (5.23) is equal to the difference of the total variations of the two measures, which by (5.20), is

$$
\leq 1 - \frac{(d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}})d^{m_{n_i}}}{d^{n_i}} = C \left(\frac{L}{d} \right)^{n_i - m_{n_i}}.
$$

In the second term on the right-hand side of (5.23), the total number of terms in the summations is bounded above by d^{n_i}. So the second term on the right-hand side of (5.23) is

$$
\leq \left| \frac{1}{d^{n_i}} - \frac{1}{d^{n_i} + d^{m_i}} \right| d^{n_i}.
$$

In the third term on the right-hand side of (5.23), by (5.21), $\deg_{f^{n_i}}(q) = 1$ for each $A \in \mathbf{P}'_{n_i - m_{n_i}}$ and each $q \in f^{-m_{n_i}}(\tau(A))$. So by (1.4) and Corollary 4.8, each δ_q in the summations cancels with the corresponding δ_q in $\xi_{n_i}^{m_{n_i}}$. So the third term on the right-hand side of (5.23) is equal to the difference of the total variations of the two measures, which by (5.20) and Corollary 4.8, is

$$
\leq 1 - \frac{(d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}})d^{m_{n_i}}}{(d^{n_i - m_{n_i}} + 1)d^{m_{n_i}}} = \frac{1 + CL^{n_i - m_{n_i}}}{d^{n_i - m_{n_i}} + 1}.
$$

Since $n_i - m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$, each term on the right-hand side of (5.23) converges to 0 as $i \longrightarrow +\infty$. So the claim is established.

Case 2 for (1.5): We assume in this case that there is a constant $K \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} \leq K$. Then by choosing a subsequence of $\{n_i\}_{i \in \mathbb{N}}$ if necessary, we can assume that there exists some constant $l \in [0, K]$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} = l$. Note that in this case, $m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$.

Then by Corollary 4.8 and Theorem 1.1,

$$\xi_{n_i}^{m_{n_i}} = \frac{1}{d^{m_{n_i}}(d^l+1)} \sum_{x \in S_{n_i}^{m_{n_i}}} \deg_{f^{n_i}}(x)\delta_x$$

$$= \frac{1}{d^l+1} \sum_{y=f^l(y)} \deg_{f^l}(y) \left(\frac{1}{d^{m_{n_i}}} \sum_{x \in f^{-m_{n_i}}(y)} \deg_{f^{m_{n_i}}}(x)\delta_x \right).$$

By Theorem 1.2, for each $y \in S^2$,

$$\frac{1}{d^{m_{n_i}}} \sum_{x \in f^{-m_{n_i}}(y)} \deg_{f^{m_{n_i}}}(x)\delta_x \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty.$$

So each term in the sequence $\{\xi_{n_i}^{m_{n_i}}\}_{i \in \mathbb{N}}$ is a convex combination of the corresponding terms in sequences of measures, each of which converges in the weak* topology to μ_f. Hence by Lemma 5.2, $\{\xi_{n_i}^{m_{n_i}}\}_{i \in \mathbb{N}}$ also converges to μ_f in the weak* topology. It then follows that $\mu = \mu_f$. Thus (1.5) follows in this case.

Proof of (1.6):

Let $\{n_i\}_{i \in \mathbb{N}}$ be a strictly increasing sequence with

$$\widetilde{\xi}_{n_i}^{m_{n_i}} \xrightarrow{w^*} \widetilde{\mu} \text{ as } i \longrightarrow +\infty,$$

for some measure $\widetilde{\mu}$.

Case 1 for (1.6): We assume in this case that there is no constant $K \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} \leq K$. Then by choosing a subsequence of $\{n_i\}_{i \in \mathbb{N}}$ if necessary, we can assume that $n_i - m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$.

The idea of the proof in this case is similar to that of the proof of Case 1 for (1.5).

We use the same notation as in the proof of Case 1 for (1.5). Then (1.6) follows in this case if we can prove that $\left\| \mu_{n_i} - \widetilde{\xi}_{n_i}^{m_{n_i}} \right\|$ converges to 0 as $i \longrightarrow +\infty$.

As before, we observe that

$$\left\| \mu_{n_i} - \widetilde{\xi}_{n_i}^{m_{n_i}} \right\| \leq \left\| \mu_{n_i} - \frac{1}{d^{n_i-m_{n_i}}} \sum_{A \in \mathbf{P}'_{n_i-m_{n_i}}} \frac{1}{d^{m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q \right\|$$

$$(5.24) \qquad + \left\| \left(\frac{1}{d^{n_i}} - \frac{1}{\widetilde{s}_{n_i}^{m_{n_i}}} \right) \sum_{A \in \mathbf{P}'_{n_i-m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q \right\|$$

$$+ \left\| \frac{1}{\widetilde{s}_{n_i}^{m_{n_i}}} \sum_{A \in \mathbf{P}'_{n_i-m_{n_i}}} \sum_{q \in f^{-m_{n_i}}(\tau(A))} \delta_q - \widetilde{\xi}_{n_i}^{m_{n_i}} \right\|.$$

As the first term on the right-hand side of (5.23) discussed before, the first term on the right-hand side of (5.24) is

$$\leq 1 - \frac{(d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}})d^{m_{n_i}}}{d^{n_i}} = C\left(\frac{L}{d}\right)^{n_i - m_{n_i}}.$$

In the second term on the right-hand side of (5.24), the total number of terms in the summations is bounded above by d^{n_i}. By (5.20), (5.21), and Corollary 4.8, we have

$$(5.25) \qquad d^{m_{n_i}}(d^{n_i - m_{n_i}} + 1) = s_{n_i}^{m_{n_i}} \geq \widetilde{s}_{n_i}^{m_{n_i}}$$

$$\geq d^{m_{n_i}} \operatorname{card}(\mathbf{P}'_{n_i - m_{n_i}}) \geq d^{m_{n_i}}(d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}}).$$

So the second term on the right-hand side of (5.24) is

$$\leq \left| \frac{1}{d^{n_i}} - \frac{1}{\widetilde{s}_{n_i}^{m_{n_i}}} \right| d^{n_i} = \left| 1 - \frac{d^{n_i}}{\widetilde{s}_{n_i}^{m_{n_i}}} \right| \leq \max\left\{ \frac{1}{d^{n_i - m_{n_i}}}, \frac{CL^{n_i - m_{n_i}}}{d^{n_i - m_{n_i}}} \right\}.$$

In the third term on the right-hand side of (5.24), by (5.21), $\deg_{f^{n_i}}(q) = 1$ for each $A \in \mathbf{P}'_{n_i - m_{n_i}}$ and each $q \in f^{-m_{n_i}}(\tau(A))$. So by (1.4), each δ_q in the summations cancels with the corresponding δ_q in $\widetilde{\xi}_{n_i}^{m_{n_i}}$. So the third term on the right-hand side of (5.24) is equal to the difference of the total variations of the two measures, which by (5.25) and (5.20), for $n_i - m_{n_i}$ large enough, is

$$\leq \frac{d^{n_i} + d^{m_{n_i}} - (d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}})d^{m_{n_i}}}{\widetilde{s}_{n_i}^{m_{n_i}}} \leq \frac{1 + CL^{n_i - m_{n_i}}}{d^{n_i - m_{n_i}} - CL^{n_i - m_{n_i}}}.$$

Since $n_i - m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$, each term on the right-hand side of (5.24) converges to 0 as $i \longrightarrow +\infty$. So we can conclude that

$$\left\| \mu_{n_i} - \widetilde{\xi}_{n_i}^{m_{n_i}} \right\| \longrightarrow 0 \text{ as } i \longrightarrow +\infty.$$

So $\widetilde{\mu} = \mu_f$. Thus (1.6) follows in this case.

Case 2 for (1.6): We assume in this case that there is a constant $K \in \mathbb{N}$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} \leq K$. Then by choosing a subsequence of $\{n_i\}_{i \in \mathbb{N}}$ if necessary, we can assume that there exists some constant $l \in [0, K]$ such that for all $i \in \mathbb{N}$, $n_i - m_{n_i} = l$. Note that in this case, $m_{n_i} \longrightarrow +\infty$ as $i \longrightarrow +\infty$.

Then for each $i \in \mathbb{N}$, we have

$$\widetilde{\xi}_{n_i}^{m_{n_i}} = \frac{1}{\widetilde{s}_{n_i}^{m_{n_i}}} \sum_{x \in S_{n_i}^{m_{n_i}}} \delta_x = \frac{1}{\widetilde{s}_{n_i}^{m_{n_i}}} \sum_{y = f^l(y)} Z_{m_{n_i}, y}\left(\frac{1}{Z_{m_{n_i}, y}} \sum_{x \in f^{-m_{n_i}}(y)} \delta_x \right),$$

where $Z_{m,y} = \operatorname{card}(f^{-m}(y))$ for each $y \in S^2$ and each $m \in \mathbb{N}_0$. Note that for each $i \in \mathbb{N}$, we have

$$\widetilde{s}_{n_i}^{m_{n_i}} = \sum_{y=f^l(y)} Z_{m_{n_i},y}.$$

Denote, for each $i \in \mathbb{N}$ and each $y \in S^2$, the Borel probability measure $\mu_{i,y} = \frac{1}{Z_{m_{n_i},y}} \sum_{x \in f^{-m_{n_i}}(y)} \delta_x$. Then by Theorem 1.2, we have

$$\mu_{i,y} \xrightarrow{w^*} \mu_f \text{ as } i \longrightarrow +\infty.$$

So each term in $\{\widetilde{\xi}_{n_i}^{m_{n_i}}\}_{i \in \mathbb{N}}$ is a convex combination of the corresponding terms in sequences of measures, each of which converges in the weak* topology to μ_f. Hence by Lemma 5.2, $\{\widetilde{\xi}_{n_i}^{m_{n_i}}\}_{i \in \mathbb{N}}$ also converges to μ_f in the weak* topology. It then follows that $\widetilde{\mu} = \mu_f$. Thus (1.6) follows in this case. $\qquad\square$

The proof of Theorem 1.3 also gives us the following corollary.

Corollary 5.14. *Let f be an expanding Thurston map. If $\{m_n\}_{n \in \mathbb{N}}$ is a sequence in $\mathbb{N}_0$ such that $m_n < n$ for each $n \in \mathbb{N}$ and $\lim\limits_{n \to +\infty} n - m_n = +\infty$, then*

$$(5.26) \qquad\qquad \lim_{n \to +\infty} \frac{\widetilde{s}_n^{m_n}}{s_n^{m_n}} = 1.$$

Proof. By the proof of Theorem 1.3, especially (5.25), we get that for each $n \in \mathbb{N}$,

$$(5.27) \qquad \frac{d^{n-m_n} - CL^{n-m_n}}{d^{n-m_n} + 1} \leq \frac{\widetilde{s}_n^{m_n}}{s_n^{m_n}} \leq 1,$$

where $d = \deg f$. Then (5.26) follows from the fact that $1 \leq L < d$ and the condition that $\lim\limits_{n \to +\infty} n - m_n = +\infty$. $\qquad\square$

By (5.1), Theorem 1.1, and Corollary 5.14 with $m_n = 0$ for each $n \in \mathbb{N}$, we get the following corollary, which is an analog of the corresponding result for expansive homeomorphisms on compact metric spaces with the *specification property* (see, for example, [KH95, Theorem 18.5.5]).

Corollary 5.15. *Let f be an expanding Thurston map. Then for each constant $c \in (0,1)$, there exists a constant $N \in \mathbb{N}$ such that for each $n \geq N$,*

$$c e^{n h_{\mathrm{top}}(f)} = c(\deg f)^n < \operatorname{card}\{x \in S^2 \mid f^n(x) = x\}$$

$$\leq \sum_{x=f^n(x)} \deg_{f^n}(x) = (\deg f)^n + 1 < \frac{1}{c} e^{nh_{\text{top}}(f)}.$$

In particular,

$$\lim_{n \to +\infty} \frac{\text{card}\{x \in S^2 \mid f^n(x) = x\}}{\exp(nh_{\text{top}}(f))} = \lim_{n \to +\infty} \frac{\text{card}\{x \in S^2 \mid f^n(x) = x\}}{(\deg f)^n} = 1.$$

Finally, we get the equidistribution of the periodic points with respect to the measure of maximal entropy μ_f as an immediate corollary.

Proof of Corollary 1.4. We get (1.7) and (1.8) from Theorem 1.3 with $m_n = 0$ for all $n \in \mathbb{N}$. Then (1.9) follows from (1.8) and Corollary 5.15.
$\square$

6. Expanding Thurston maps as factors of the left-shift

M. Bonk and D. Meyer [BM10] proved that for an expanding Thurston map f, the topological dynamical system (S^2, f) is a factor of a certain classical topological dynamical system, namely, the left-shift on the one-sided infinite sequences of $\deg f$ symbols. The goal of this section is to generalize this result to the category of measure-preserving dynamical systems. The invariant measure for each measure-preserving dynamical system considered in this section is going to be the unique measure of maximal entropy of the corresponding system.

Let X and $\widetilde{X}$ be topological spaces, and $f \colon X \to X$ and $\widetilde{f} \colon \widetilde{X} \to \widetilde{X}$ be continuous maps. We say that the topological dynamical system (X, f) is a *factor of the topological dynamical system* $(\widetilde{X}, \widetilde{f})$ if there is a surjective continuous map $\varphi \colon \widetilde{X} \to X$ such that $\varphi \circ \widetilde{f} = f \circ \varphi$. For measure-preserving dynamical systems (X, g, μ) and $(\widetilde{X}, \widetilde{g}, \widetilde{\mu})$ where X and $\widetilde{X}$ are measure spaces, $g \colon X \to X$ and $\widetilde{g} \colon \widetilde{X} \to \widetilde{X}$ measurable maps, and $\mu \in \mathcal{M}(X, g)$ and $\widetilde{\mu} \in \mathcal{M}(\widetilde{X}, \widetilde{g})$, we say that the measure-preserving dynamical system $(\widetilde{X}, \widetilde{g}, \widetilde{\mu})$ is a *factor of the measure-preserving dynamical system* (X, g, μ) if there is a measurable map $\varphi \colon \widetilde{X} \to X$ such that $\varphi \circ \widetilde{g} = g \circ \varphi$ and $\varphi_* \widetilde{\mu} = \mu$.

We recall a classical example of symbolic dynamical systems, namely (J_k^ω, Σ), where the *alphabet* $J_k = \{0, 1, \dots, k-1\}$ for some $k \in \mathbb{N}$, the *set of infinite words* $J_k^\omega = \prod_{i=1}^{+\infty} J_k$, and Σ is the left-shift operator with

$$\Sigma(i_1, i_2, \dots) = (i_2, i_3, \dots)$$

for each $(i_i, i_2, \dots) \in J_k^\omega$. We equip J_k^ω with a metric d such that the distance between two distinct infinite words $(i_1, i_2, \dots)$ and $(j_1, j_2, \dots)$ is $\frac{1}{m}$, where $m = \min\{n \in \mathbb{N} \mid i_n \neq j_n\}$.

Define the *set of words of length n* as $J_k^n = \prod_{i=1}^n J_k$, for $n \in \mathbb{N}$ and $J_k^0 = \{\emptyset\}$ where $\emptyset$ is considered as the word of length 0, which is also denoted by (). Denote the *set of finite words* by $J_k^* = \bigcup_{n=0}^{+\infty} J_k^n$. Then the left-shift operator Σ is defined on $J_k^* \setminus J_k^0$ naturally by

$$\Sigma(i_1, i_2, \dots, i_n) = (i_2, i_3, \dots, i_n).$$

It is well-known that the dynamical system (J_k^ω, Σ) has a unique measure of maximal entropy η_Σ, which is characterized by the property that

$$\eta_\Sigma \left(C(j_1, j_2, \dots, j_n) \right) = k^{-n},$$

for $n \in \mathbb{N}$ and $j_1, j_2, \dots, j_n \in J_k$, where

(6.1)
$$C(j_1, j_2, \dots, j_n) = \{(i_1, i_2, \dots) \in J_k^\omega \mid i_1 = j_1, i_2 = j_2, \dots, i_n = j_n\}$$

is the *cylinder set* determined by $j_1, j_2, \dots, j_n$ (see for example, [KH95, Section 4.4]).

We will prove that for each expanding Thurston map f with $\deg f = k$ and its measure of maximal entropy μ_f, the measure-preserving dynamical system (S^2, f, μ_f) is a factor of the system $(J_k^\omega, \Sigma, \eta_\Sigma)$.

We now review a construction from [BM10].

Let $f \colon S^2 \to S^2$ be an expanding Thurston map, and $\mathcal{C} \subseteq S^2$ a Jordan curve with post $f \subseteq \mathcal{C}$. Consider the cell decompositions induced by the pair $(f, \mathcal{C})$. Let $k = \deg f$. Fix an arbitrary point $p \in \operatorname{inte}(X_w^0)$. Let $q_1, q_2, \dots, q_k$ be the distinct points in $f^{-1}(p)$. For $i = 1, 2, \dots, k$, we pick a continuous path $\alpha_i \colon [0, 1] \to S^2 \setminus \operatorname{post} f$ with $\alpha_i(0) = p$ and $\alpha_i(1) = q_i$.

We construct $\psi \colon J_k^* \to S^2$ inductively such that $\psi(I) \in f^{-n}(p)$, for each $n \in \mathbb{N}_0$ and $I \in J_k^n$, in the following way:

Define $\psi(\emptyset) = p$, and $\psi((i)) = q_i$ for each $(i) \in J_k^1$. Suppose that ψ has been defined for all $I \in \bigcup_{j=0}^n J_k^j$, where $n \in \mathbb{N}$. Now for each $(i_1, i_2, \dots, i_{n+1}) \in J_k^{n+1}$, the point $\psi((i_1, i_2, \dots, i_n)) \in f^{-n}(p)$ has already been defined. Since $f^n(\psi((i_1, i_2, \dots, i_n))) = p$ and $f^n \colon S^2 \setminus f^{-n}(\operatorname{post} f) \to S^2 \setminus \operatorname{post} f$ is a covering map, the path $\alpha_{i_{n+1}}$ has a unique lift $\widetilde{\alpha}_{i_{n+1}} \colon [0, 1] \to S^2$ with $\widetilde{\alpha}_{i_{n+1}}(0) = \psi((i_1, i_2, \dots, i_n))$ and $f^n \circ \widetilde{\alpha}_{i_{n+1}} = \alpha_{i_{n+1}}$. We now define $\psi((i_1, i_2, \dots, i_{n+1})) = \widetilde{\alpha}_{i_{n+1}}(1)$. Note that then

$$f^{n+1}(\psi((i_1, i_2, \dots, i_{n+1})))$$
$$= f^{n+1}(\widetilde{\alpha}_{i_{n+1}}(1)) = f(\alpha_{i_{n+1}}(1)) = f(q_{i_{n+1}}) = p.$$

Hence $\psi((i_1, i_2, \ldots, i_{n+1})) \in f^{-(n+1)}(p)$. This completes the inductive construction of ψ.

Note that $\psi \colon J_k^* \to S^2$ induces a map $\widetilde{\psi} \colon J_k^* \to \bigcup\limits_{n=0}^{+\infty} \mathbf{X}_w^n$ by mapping each $(i_1, i_2, \ldots, i_n) \in J_k^n$ to the unique white n-tile $X_w^n \in \mathbf{X}_w^n$ containing $\psi((i_1, i_2, \ldots, i_n)) \in f^{-n}(p)$.

By the proof of Theorem 1.6 in Chapter 9 of [BM10], for each $n \in \mathbb{N}$, $\psi|_{J_k^n} \colon J_k^n \to f^{-n}(p)$ is a bijection. Hence $\widetilde{\psi}|_{J_k^n} \colon J_k^n \to \mathbf{X}_w^n$ for $n \in \mathbb{N}_0$, and $\widetilde{\psi} \colon J_k^* \to \bigcup\limits_{n=0}^{+\infty} \mathbf{X}_w^n$ are also bijections. Moreover, by the proof of Theorem 1.6 in [BM10], we have that for each $(i_1, i_2, \ldots) \in J_k^\omega$, $\{\psi((i_1, i_2, \ldots, i_n))\}_{n \in \mathbb{N}}$ is a Cauchy sequence in (S^2, d), for each visual metric d for f. So as shown in the proof of Theorem 1.6 in [BM10], the map $\varphi \colon J_k^\omega \to S^2$ defined by

$$(6.2) \qquad \varphi((i_1, i_2, \ldots)) = \lim_{n \to +\infty} \psi((i_1, i_2, \ldots, i_n))$$

satisfies

(1) φ is continuous,

(2) $f \circ \varphi = \varphi \circ \Sigma$,

(3) $\varphi \colon J_k^\omega \to S^2$ is surjective.

We now reformulate Theorem 1.6 from [BM10] in the following way.

Theorem 6.1 (M. Bonk & D. Meyer 2010). *Let $f \colon S^2 \to S^2$ be an expanding Thurston map with $\deg f = k$. Then (S^2, f) is a factor of the topological dynamical system (J_k^ω, Σ). More precisely, the surjective continuous map $\varphi \colon J_k^\omega \to S^2$ defined above satisfies $f \circ \varphi = \varphi \circ \Sigma$.*

We will strengthen Theorem 6.1 in the following theorem.

Theorem 6.2. *Let $f \colon S^2 \to S^2$ be an expanding Thurston map with $\deg f = k$. Then (S^2, f, μ_f) is a factor of the measure-preserving dynamical system $(J_k^\omega, \Sigma, \eta_\Sigma)$, where μ_f and η_Σ are the unique measures of maximal entropy of (S^2, f) and (J_k^ω, Σ), respectively. More precisely, the surjective continuous map $\varphi \colon J_k^\omega \to S^2$ defined above satisfies $f \circ \varphi = \varphi \circ \Sigma$ and $\varphi_* \eta_\Sigma = \mu_f$.*

Proof. Let $\mathcal{C} \subseteq S^2$ be a Jordan curve containing $\operatorname{post} f$. Let d be a visual metric on S^2 for f with expansion factor $\Lambda > 1$. Note that Λ depends only on f and d. Consider the cell decompositions induced by $(f, \mathcal{C})$.

By Theorem 6.1, it suffices to prove that $\varphi_* \eta_\Sigma = \mu_f$.

For each $n \in \mathbb{N}$, we fix a function $\widetilde{\beta}_n \colon J_k^n \to J_k^\omega$ which maps each $(i_1, i_2, \ldots, i_n) \in J_k^n$ to $(i_1, i_2, \ldots, i_n, i_{n+1}, \ldots) \in J_k^\omega$, for some arbitrarily chosen $i_{n+1}, i_{n+2}, \cdots \in J_k$ depending on $i_1, i_2, \ldots, i_n$.

Define $\beta_n = \varphi \circ \widetilde{\beta}_n \circ \widetilde{\psi}^{-1}$, for each $n \in \mathbb{N}$, where $\widetilde{\psi}$ is defined earlier in this section.

We claim that the maps $\beta_n \colon \mathbf{X}_w^n \to S^2$ with $n \in \mathbb{N}$ satisfy the hypothesis for β_n in Theorem 5.7, namely,

$$\max\{d(\beta_n(X_w^n), X_w^n) \mid X_w^n \in \mathbf{X}_w^n\} \longrightarrow 0 \text{ as } n \longrightarrow +\infty.$$

Indeed, by the construction of $\varphi, \widetilde{\beta}_n, \psi$ and $\widetilde{\psi}$ above, we have that β_n maps a white n-tile X_w^n to the limit of a Cauchy sequence

$$(\psi((j_1, j_2, \ldots, j_m)))_{m \in \mathbb{N}}$$

such that $\psi((j_1, j_2, \ldots, j_n)) \in X_w^n$. Since for each $m \in \mathbb{N}$, the points $\psi((j_1, j_2, \ldots, j_m))$ and $\psi((j_1, j_2, \ldots, j_{m+1}))$ are joined by a lift of one of the paths $\alpha_1, \alpha_2, \ldots, \alpha_k$ (defined above) by f^m, by Lemma 8.11 in [BM10], we have that

$$d\left(\psi((j_1, j_2, \ldots, j_m)), \psi((j_1, j_2, \ldots, j_{m+1}))\right) \leq C\Lambda^{-m},$$

for all $m \in \mathbb{N}$, where $C > 0$ is a constant depending only on f, d, and the curves α_i, $i \in \{1, 2, \ldots, k\}$, in the construction of ψ. In particular, both C and Λ are independent of m and $(j_1, j_2, \ldots) \in J_k^\omega$. So $d(\beta_n(X_w^n), X_w^n) \leq C\frac{\Lambda^n}{1-\Lambda}$ for each $n \in \mathbb{N}$ and each $X_w^n \in \mathbf{X}_w^n$. The above claim follows.

For $i \in \mathbb{N}$, define

$$\eta_i = \frac{1}{k^i} \sum_{I \in J_k^i} \delta_{\widetilde{\beta}_i(I)}.$$

Observe that for all $n \in \mathbb{N}$ and $m \in \mathbb{N}$ with $m \geq n$, and each $(i_1, i_2, \ldots, i_n) \in J_k^n$, we have

$$\eta_m(C(i_1, i_2, \ldots, i_n)) = \eta_\Sigma(C(i_1, i_2, \ldots, i_n)),$$

where $C(i_1, i_2, \ldots, i_n)$ is defined in (6.1). So by the uniform continuity of each continuous function on J_k^ω, it is easy to see that

(6.3) $$\eta_i \xrightarrow{w^*} \eta_\Sigma \text{ as } i \longrightarrow +\infty.$$

Note that since $\widetilde{\psi}|_{J_k^n} \colon J_k^n \to \mathbf{X}_w^n$ is a bijection for each $n \in \mathbb{N}_0$, we have for each $i \in \mathbb{N}$,

$$\varphi_*\eta_i = \frac{1}{k^i} \sum_{I \in J_k^i} \delta_{\varphi \circ \widetilde{\beta}_i(I)} = \frac{1}{k^i} \sum_{X^i \in \mathbf{X}_w^i} \delta_{\varphi \circ \widetilde{\beta}_i \circ \widetilde{\psi}^{-1}(X^i)} = \frac{1}{k^i} \sum_{X^i \in \mathbf{X}_w^i} \delta_{\beta_i(X^i)}.$$

Hence, by Theorem 5.7, $\varphi_* \eta_i \xrightarrow{w^*} \mu_f$ as $i \longrightarrow +\infty$. Combining this with (6.3) and Lemma 5.1, therefore, we can conclude that $\varphi_* \eta_\Sigma = \mu_f$. $\qquad\qquad\square$

7. A RANDOM ITERATION ALGORITHM FOR PRODUCING THE MEASURE OF MAXIMAL ENTROPY

In this section, we follow the idea of [HT03] to prove that for each $p \in S^2$, the measure of maximal entropy μ_f of an expanding Thurston map f is almost surely the limit of

$$\frac{1}{n} \sum_{i=0}^{n-1} \delta_{q_i}$$

as $n \longrightarrow +\infty$ in the weak* topology, where $q_0 = p$, and q_i is one of the points x in $f^{-1}(q_{i-1})$, chosen with probability $\frac{\deg_f(x)}{\deg f}$, for each $i \in \mathbb{N}$.

To give a more precise formulation, we will use the language of Markov process from the probability theory (see, for example, [Du10] for an introduction).

Let (X, d) be a compact metric space. Equip the space $\mathcal{P}(X)$ of Borel probability measures with the weak* topology. A continuous map $X \to \mathcal{P}(X)$ assigning to each $x \in X$ a measure μ_x defines a *random walk* on X. We define the corresponding *Markov operator* $Q \colon C(X) \to C(X)$ by

$$(7.1) \qquad\qquad Q\phi(x) = \int \phi(y) \, \mathrm{d}\mu_x(y).$$

Let Q^* be the adjoint operator of Q, i.e., for each $\phi \in C(X)$ and $\rho \in \mathcal{P}(X)$,

$$(7.2) \qquad\qquad \int Q\phi \, \mathrm{d}\rho = \int \phi \, \mathrm{d}(Q^*\rho).$$

Consider a stochastic process $(\Omega, \mathcal{F}, P)$, where

(1) $\Omega = \{(\omega_0, \omega_1, \dots) \mid \omega_i \in X, i \in \mathbb{N}_0\} = \prod_{i=0}^{+\infty} X$, equipped with the product topology,

(2) $\mathcal{F}$ is the Borel σ-algebra on Ω,

(3) $P \in \mathcal{P}(\Omega)$.

This process is a *Markov process with transition probabilities* $\{\mu_x\}_{x \in X}$ if

$$(7.3) \qquad P\{\omega_{n+1} \in A \mid \omega_0 = z_0, \omega_1 = z_1, \dots, \omega_n = z_n\} = \mu_{z_n}(A)$$

for all $n \in \mathbb{N}_0$, Borel subsets $A \subseteq X$, and $z_0, z_1, \dots, z_n \in X$.

The transition probabilities $\{\mu_x\}_{x\in X}$ are determined by the operator Q and so we can speak of a *Markov process determined by Q*.

Let $f\colon S^2 \to S^2$ be an expanding Thurston map with $\deg f = k$. The continuous map $S^2 \to \mathcal{P}(S^2)$ assigning to each $x \in S^2$ the probability measure

$$(7.4) \qquad \mu_x = \frac{1}{k} \sum_{y\in f^{-1}(x)} \deg_f(y)\delta_y$$

induces the Markov operator $Q\colon C(S^2) \to C(S^2)$ which satisfies

$$(7.5) \qquad Q\phi(x) = \frac{1}{k} \sum_{y\in f^{-1}(x)} \deg_f(y)\phi(y)$$

for all $\phi \in C(S^2)$ and $x \in S^2$. To show Q is well-defined, we need to prove $Q\phi(x)$ is continuous in $x \in S^2$ for each $\phi \in C(S^2)$. Indeed, by fixing an arbitrary Jordan curve $\mathcal{C} \subseteq S^2$ containing post f, we know for each x in the white 0-tile X_w^0,

$$Q\phi(x) = \frac{1}{k} \sum_{X\in \mathbf{X}_w^1} \phi(y_X),$$

where y_X is the unique point contained in the white 1-tile X with the property that $f(y_X) = x$. If we move x around continuously within X_w^0, then each y_X moves around continuously within X. Thus $Q\phi(x)$ restricted to X_w^0 is continuous in x. Similarly, $Q\phi(x)$ restricted to X_b^0 is also continuous. Hence $Q\phi(x)$ is continuous in $x \in S^2$.

Fix an arbitrary $z \in S^2$. Then there exists a unique Markov process $(\Omega, \mathcal{F}, P_z)$ determined by Q with

(1) $\Omega = \prod\limits_{i=0}^{+\infty} S^2$, equipped with the product topology,

(2) $\mathcal{F}$ being the Borel σ-algebra on Ω,

(3) P_z being a Borel probability measure on Ω satisfying

   $$P_z\{\omega_{n+1} \in A \,|\, \omega_0 = z, \omega_1 = z_1, \dots, \omega_n = z_n\} = \mu_{z_n}(A)$$

   for all $n \in \mathbb{N}$, Borel subset $A \subseteq S^2$, and $z_1, z_2, \dots, z_n \in S^2$.

The existence and uniqueness of P_z follows from [Lo77, Theorem 1.4.2]. Since the Markov process $(\Omega, \mathcal{F}, P_z)$ is determined by f as well, we will also call $(\Omega, \mathcal{F}, P_z)$ the *Markov process determined by f*.

Now we can formulate our main theorem for this section.

Theorem 7.1. *Let $f\colon S^2 \to S^2$ be an expanding Thurston map with its measure of maximal entropy μ_f. Let $(\Omega, \mathcal{F}, P_z)$ be the Markov process*

determined by f. Then for each $z \in S^2$, we have that P_z-almost surely,

$$(7.6) \qquad \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\omega_j} \xrightarrow{w^*} \mu_f \ \text{as } n \longrightarrow +\infty.$$

In other words, if we fix a point $z \in S^2$ and set it as the first point in an infinite sequence, and choose each of the following points randomly according to the Markov process determined by f, then P_z-almost surely, the probability measure equally distributed on the first n points in the sequence converges in the weak* topology to μ_f as $n \longrightarrow +\infty$.

In order to prove Theorem 7.1, we need a theorem of H. Furstenberg and Y. Kifer from [FK83].

Theorem 7.2 (H. Furstenberg & Y. Kifer 1983). *Let $\Omega = \{\omega_n \in X \mid n \in \mathbb{N}_0\}$ be the Markov process determined by the operator Q. Assume that there exists a unique Borel probability measure μ that is invariant under the adjoint operator Q^* on $\mathcal{P}(X)$. Then for each $\omega_0 \in X$, we have that P_{ω_0}-almost surely,*

$$(7.7) \qquad \frac{1}{n} \sum_{j=0}^{n-1} \delta_{\omega_j} \xrightarrow{w^*} \mu \ \text{as } n \longrightarrow +\infty.$$

Theorem 7.1 follows immediately from Theorem 7.2 and the following lemma.

Lemma 7.3. *Let $f \colon S^2 \to S^2$ be an expanding Thurston map. Then the unique measure of maximal entropy μ_f for f is the only Borel probability measure on S^2 that is invariant under the adjoint operator $Q^* \colon \mathcal{P}(S^2) \to \mathcal{P}(S^2)$ of $Q \colon C(S^2) \to C(S^2)$, where Q is defined in (7.5).*

Proof. Let $k = \deg f$. Fix a Jordan curve $\mathcal{C} \subseteq S^2$ with post $f \subseteq \mathcal{C}$. Let d be a visual metric on S^2 for f with expansion factor $\Lambda > 1$. Consider the cell decompositions induced by $(f, \mathcal{C})$.

Recall ν_n defined in (1.1) for a fixed $p \in S^2$. Observe that $Q^* \nu_n = \nu_{n+1}$ for all $n \in \mathbb{N}_0$. By Theorem 1.2,

$$\int \varphi \, d(Q^* \mu_f) = \int Q\varphi \, d\mu_f = \lim_{n \to +\infty} \int Q\varphi \, d\nu_n = \lim_{n \to +\infty} \int \varphi \, dQ^* \nu_n$$

$$= \lim_{n \to +\infty} \int \varphi \, d\nu_{n+1} = \int \varphi \, d\mu_f.$$

Thus $Q^* \mu_f = \mu_f$, and so μ_f is indeed invariant under Q^*.

By (7.5) and (3.2), for each $x \in S^2$, each $n \in \mathbb{N}$ and each $\varphi \in C(S^2)$, we have

$$(7.8) \qquad Q^n \varphi(x) = \frac{1}{k^n} \sum_{y \in f^{-n}(x)} \deg_{f^n}(y) \varphi(y).$$

So by Theorem 1.2, we get

$$(7.9) \qquad Q^n \varphi(x) - \int \varphi \, \mathrm{d}\mu_f \longrightarrow 0 \text{ as } n \longrightarrow +\infty.$$

We claim that the convergence in (7.9) is uniform in x.

To prove the claim, we first assume that x is in the (closed) white 0-tile X_w^0. If we move x around continuously within X_w^0, then each preimage of x under f^n moves around continuously within one of the white n-tiles $X_w^n \in \mathbf{X}_w^n$, for each $n \in \mathbb{N}$. By Lemma 3.10, there exists a constant $C \geq 1$ depending only on f, $\mathcal{C}$, and d such that $\operatorname{diam}_d(X_w^n) \leq C\Lambda^{-n}$ for each $n \in \mathbb{N}$ and each $X_w^n \in \mathbf{X}_w^n$. Then by the uniform continuity of φ and (7.8), we have that $Q^n \varphi(x)$ converges uniformly to $\int \varphi \, \mathrm{d}\mu_f$ over X_w^0 as $n \longrightarrow +\infty$. Similarly, we have that the convergence in (7.9) is uniform over the black 0-tile X_b^0. Hence, the convergence in (7.9) is uniform over S^2. The claim is proved.

Suppose that $\mu \in \mathcal{P}(S^2)$ satisfies $Q^* \mu = \mu$. Then for each $\varphi \in C(S^2)$, by the claim above, we have

$$\int \varphi(x) \, \mathrm{d}\mu(x) = \lim_{n \to +\infty} \int \varphi(x) \, \mathrm{d}(Q^*)^n \mu(x)$$

$$= \lim_{n \to +\infty} \int Q^n \varphi(x) \, \mathrm{d}\mu(x) = \int \varphi(x) \, \mathrm{d}\mu_f(x).$$

Hence $\mu = \mu_f$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

As a special case of [HP09, Theorem 3.4.11], the next corollary follows immediately from the uniform convergence in (7.9):

Corollary 7.4. *Let $f \colon S^2 \to S^2$ be an expanding Thurston map with its measure of maximal entropy μ_f. Then for each Borel probability measure μ on S^2, we have*

$$(7.10) \qquad (Q^*)^n \mu \xrightarrow{w^*} \mu_f \text{ as } n \longrightarrow +\infty,$$

where $Q^* \colon \mathcal{P}(S^2) \to \mathcal{P}(S^2)$ is the adjoint operator of $Q \colon C(S^2) \to C(S^2)$ defined in (7.5).

Proof. By the claim proved in the proof of Lemma 7.3, the convergence in (7.9) is uniform in x for each $\varphi \in C(S^2)$. Thus by integrating (7.9) over S^2 with respect to μ, we get $\int Q^n \varphi \, \mathrm{d}\mu \longrightarrow \int \varphi \, \mathrm{d}\mu_f$ as $n \longrightarrow +\infty$, from which (7.10) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Remark 7.5. The operator Q as defined in (7.5) is actually the Ruelle operator for an expanding Thurston map f multiplied by a factor $\frac{1}{\deg f}$, in the special case when the potential is identically 0. For some background of the Ruelle operator and the thermodynamical formalism, see, for example, [Ru89, PU10], and in the context of expanding Thurston maps, see [Li14]. More generally, we prove in [Li14] that for each expanding Thurston map and each Hölder continuous (with respect to any visual metric) potential ϕ, there exists a unique equilibrium state, which is exact, and in particular, mixing. As a generalization of the measure of maximal entropy, an equilibrium state is an invariant Borel probability measure that maximizes the pressure, which in turn is a generalization of the topological entropy. Moreover, we prove in [Li14] that the equilibrium state is the unique Borel probability measure invariant under the adjoint $\mathcal{L}_{\widetilde{\phi}}^*$ of the Ruelle operator $\mathcal{L}_{\widetilde{\phi}}$ corresponding to the Hölder continuous potential $\widetilde{\phi}$ determined by ϕ. In the case when $\phi = 0$, we have $\widetilde{\phi} = -h_{\mathrm{top}}(f)$, and $\mathcal{L}_{\widetilde{\phi}}^* = \exp(-h_{\mathrm{top}}(f))\mathcal{L}_{\phi}^* = \frac{1}{\deg f}\mathcal{L}_{\phi}^*$. Thus Lemma 7.3 follows from the more general result in [Li14]. The direct proof of Lemma 7.3 we included above is much simpler though.

References

[Bak85] BAKER, A., *A concise introduction to the theory of numbers*, Cambridge Univ. Press, Cambridge, 1985.

[Bar88] BARNSLEY, M., *Fractals everywhere*, Academic Press Professional, San Diego, 1988.

[BD11] BAKER, M. and DeMARCO, L., Preperiodic points and unlikely intersections. *Duke Math. J.* 159 (2011), 1–29.

[Bi99] BILLINGSLEY, P., *Convergence of probability measures*, John Wiley & Sons, New York, 1999.

[BM10] BONK, M. and MEYER, D., Expanding Thurston maps. Preprint, (arXiv:1009.3647), 2010.

[Bo75] BOWEN, R., *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.*, Springer, Berlin, 1975.

[Br10] BRACCI, F., Local holomorphic dynamics of diffeomorphisms in dimension one. In CONTRERAS, M. D. and DÍAZ-MADRIGAL, S. (Eds.), *Five lectures in complex analysis* (pp. 1–42). Amer. Math. Soc., Providence, RI, 2010.

[Bu79] BURCKEL, R., *An introduction to classical complex analysis*, vol. 1, Academic Press, New York, 1979.

[BS03] BUZZI, J. and SARIG, O., Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps. *Ergod. Th. & Dynam. Sys.* 23 (2003), 1383–1400.

[CG93] CARLESON, L. and GAMELIN, T.H., *Complex dynamics*, Springer, New York, 1993.

[DH93] DOUADY, A. and HUBBARD, J.H., A proof of Thurston's topological characterization of rational functions. *Acta Math.* 171 (1993), 263–297.

[Du10] DURRETT, R., *Probability: theory and examples*, Cambridge Univ. Press, Cambridge, 2010.

[Fo99] FOLLAND, G.B., *Real analysis: modern techniques and their applications*, 2nd ed., Wiley, New York, 1999.

[FK83] FURSTENBERG, H. and KIFER, Y., Random matrix products and measures on projective spaces. *Israel J. Math.* 46 (1983), 12–32.

[FLM83] FREIRE, A., LOPES, A. and MAÑÉ, R., An invariant measure for rational maps. *Bol. Bras. Mat. Soc.* 14 (1983), 45–62.

[GP10] GUILLEMIN, V. and POLLACK, A., *Differential topology*, Amer. Math. Soc., Providence, RI, 2010.

[HP09] HAÏSSINSKY, P. and PILGRIM, K., Coarse expanding conformal dynamics. *Astérisque* 325 (2009).

[Ha02] HATCHER, A., *Algebraic topology*, Cambridge Univ. Press, Cambridge, 2002.

[HT03] HAWKINS, J. and TAYLOR, M., Maximal entropy measure for rational maps and a random iteration algorithm for Julia sets. *Intl. J. of Bifurcation and Chaos* 13 (6) (2003), 1442–1447.

[KH95] KATOK, A. and HASSELBLATT, B., *Introduction to the modern theory of dynamical systems*, Cambridge Univ. Press, Cambridge, 1995.

[Li14] LI, Z., Equilibrium states for expanding Thurston maps. Preprint, (arXiv:1410.4920), 2014.

[Lo77] LOÈVE, M., *Probability theory*, vol. 1, 4th ed., Springer, New York, 1977.

[Ly83] LYUBICH, M., Entropy properties of rational endomorphisms of the Riemann sphere. *Ergod. Th. & Dynam. Sys.* 3 (1983), 351–385.

[Mi06] MILNOR, J., *Dynamics in one complex variable*, 3rd ed., Princeton Univ. Press, Princeton, 2006.

[Mo77] MOISE, E., *Geometric topology in dimensions* 2 *and* 3, Springer, New York, 1977.

[Pa64] PARRY, W., On intrinsic Markov chains. *Trans. Amer. Math. Soc.* 112 (1964), 55–66.

[PU10] PRZYTYCKI, F. and URBAŃSKI, M., *Conformal fractals: ergodic theory methods*, Cambridge Univ. Press, Cambridge, 2010.

[Ru89] RUELLE, D., The thermodynamical formalism for expanding maps. *Comm. Math. Phys.* 125 (1989), 239–262.

[Si72] SINAI, YA., Gibbs measures in ergodic theory. *Russian Math. Surveys* 27 (1972), 21–69.

[Wa82] WALTERS, P., *An introduction to ergodic theory*, Springer, New York, 1982.

[Wi07] WILDRICK, K.M., Quasisymmetric parameterizations of two-dimensional metric spaces. PhD thesis, University of Michigan, 2007.

DEPARTMENT OF MATHEMATICS, UCLA, LOS ANGELES CA 90095-1555
E-mail address: lizq@math.ucla.edu