Abstract. For a hyperbolic rational map $f$ of degree at least two on the Riemann sphere, we obtain estimates for the number of primitive periodic orbits of $f$ ordered by their multiplier, and establish equidistribution of the associated holonomies, both with power saving error terms.

1. Introduction

The prime number theorem states that the number of primes of size at most $T$ grows like $\text{Li}(t) = \int_2^t \frac{dt}{\log t}$. It was proved by Hadamard and de la Vallée Poussin in 1896, and is equivalent to the non-vanishing of the Riemann zeta function $\zeta(s)$ on the plane $\Re(s) \geq 1$. For Gaussian primes, that is, prime ideals in $\mathbb{Z}[i]$, not only does the number of Gaussian primes of norm less than $t$ grow like $\text{Li}(t)$ but also the angular components of Gaussian primes are equidistributed in the circle $S^1$. This follows from the non-vanishing of the Hecke $L$-functions on $\Re(s) \geq 1$ established by Hecke in 1920. No power-saving error terms have been obtained for the prime number theorems.

Geometric analogues of these profound facts have been of great interest over the years, and our aim in this work is to prove a version of the theorem for hyperbolic rational maps, providing a power saving error term. These counting questions are studied through their associated dynamical zeta functions, twisted by characters of $S^1$, which are analogues of Hecke $L$-functions. We establish that these zeta functions are non-vanishing (except for a necessary pole at $s = \delta$) on a half plane $\Re(s) > \delta - \epsilon$ for some $\epsilon > 0$, where $\delta$ is the Hausdorff dimension of the Julia set of the rational map.

More precisely let $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a rational map on the Riemann sphere of degree $d$ at least 2 and assume that $f$ is hyperbolic, that is, the post critical orbit closure of $f$ is disjoint from the Julia set $J$. We seek to understand the dynamics of iterates of $f$ acting on $\hat{\mathbb{C}}$. All but finitely many periodic orbits lie on the Julia set $J$. For a primitive periodic orbit $\hat{x} = \{x, fx, f^2x, \ldots, f^{n-1}x\}$ on $J$ of period $n$, the multiplier is given by

$$\lambda(\hat{x}) := (f^n)'(x) \in \mathbb{C}$$

Supported in parts by the NSF.
and the holonomy is given by
\[ \lambda_\theta(\hat{x}) := \frac{\lambda(\hat{x})}{|\lambda(\hat{x})|} \in S^1. \]

Primitive periodic orbits naturally play the role of primes in our prime number theorem, while the multiplier will be a measure of the size of the periodic orbit (or “prime”).

We denote by \( P \) the set of all primitive periodic orbits of \( f \) on \( J \). The hyperbolicity assumption implies that for a given \( t > 1 \), there are only finitely many primitive periodic orbits of multiplier bounded by \( t \), i.e.,
\[ N_t(P) := \# \{ \hat{x} \in P : |\lambda(\hat{x})| < t \} < \infty. \]

We will prove the following estimate which both provides asymptotics for \( N_t(P) \) and establishes equidistribution of their holonomies with power saving error terms.

**Theorem 1.1.** Let \( f : \hat{\mathbb{C}} \to \hat{\mathbb{C}} \) be a hyperbolic rational map of degree at least 2.

1. Suppose \( f \) is not conjugate to a monomial \( f(x) = x^\pm d \) for any \( d \in \mathbb{N} \). Then there exists \( \epsilon > 0 \) such that
\[ \# \{ \hat{x} \in P : |\lambda(\hat{x})| < t \} = \text{Li}(t^\delta) + O(t^{\delta-\epsilon}) \]
where \( \delta \) is the Hausdorff dimension of the Julia set of \( f \).

2. Suppose that the Julia set of \( f \) is not contained in any circle in \( \hat{\mathbb{C}} \). Then there exists \( \epsilon > 0 \) such that for any \( \psi \in C^4(S^1) \),
\[ \sum_{\hat{x} \in P : |\lambda(\hat{x})| < t} \psi(\lambda_\theta(\hat{x})) = \int_0^1 \psi(e^{2\pi i \theta}) \, d\theta \cdot \text{Li}(t^\delta) + O(t^{\delta-\epsilon}) \]
where the implied constant depends only on the Sobolev norm of \( \psi \).

In both statements above, the assumptions can easily be seen to be necessary. The special case of \( f(x) = x^2 + c \) with \( c \in (-\infty, -2) \) was studied by Naud [9], who proved Theorem 1.1(1) in that setting. By conjugation, Naud’s work covers all quadratic maps with Julia set contained in a circle.

We reduce the study of periodic orbits to the study of appropriate zeta functions. If \( x \) is a periodic point of period \( n \), the multiplier \( \lambda(x) \) and holonomy \( \lambda_\theta(x) \) of \( x \) are defined to be the multiplier and holonomy of the orbit \( \hat{x} = \{ x, f(x), \ldots, f^{n-1}(x) \} \) respectively. For each character \( \chi \) of \( S^1 \), define the weighted dynamical zeta function twisted by \( \chi \):

\[ \zeta(s, \chi) = \prod_{\hat{x} \in P} (1 - \chi(\lambda_\theta(\hat{x}))|\lambda(\hat{x})|^{-s})^{-1} \]
which is holomorphic and non-vanishing on \( \Re(s) > \delta \). Note that this reduces to the usual dynamical zeta function

\[ \zeta(s) = \prod_{\hat{x} \in P} (1 - |\lambda(\hat{x})|^{-s})^{-1} \]
when the character is trivial. The point is to establish a zero free half plane beyond the critical line $\Re(s) = \delta$ as follows.

**Theorem 1.2.** Let $f$ be a hyperbolic rational map of degree at least 2.

1. If $f$ is not conjugate to $x \pm d$ for any $d \in \mathbb{N}$, then there exists $\epsilon > 0$ such that $\zeta(s)$ is analytic and non-vanishing on $\Re(s) \geq \delta - \epsilon$ except for the simple pole at $s = \delta$.

2. If the Julia set of $f$ is not contained in a circle, then there exists $\epsilon > 0$ such that for any non-trivial character $\chi$ of $S^1$, $\zeta(s, \chi)$ is analytic and non-vanishing on $\Re(s) \geq \delta - \epsilon$.

1.1. **On the proof of Theorem 1.1(2).** We identify the group of all characters of $S^1$ with $\mathbb{Z}$ via the map $\chi_\ell(\alpha) = \alpha^\ell$. By the arguments of [8], Theorem 1.1 follows from Theorem 1.2. Theorem 1.2 then follows from spectral bounds for a family of Ruelle transfer operators

$$\{L_{s,\ell} : s \in \mathbb{C} \text{ and } \ell \in \mathbb{Z}\}$$

which act on $C^1(U)$ for some neighbourhood $U$ of the Julia set; see Theorem 2.7. These spectral bounds will be our main object of study.

Our approach is to adapt Dolgopyat’s ideas from his work on exponential mixing of Anosov flows [5]. A similar approach was carried through by Pollicott-Sharp [8] for negatively curved compact manifolds, and by Naud [9] for convex cocompact hyperbolic surfaces as well as for some class of quadratic polynomials as remarked above.

In our setting of hyperbolic rational maps, the main technical difficulties in implementing this approach arise from verifying the Non-Local-Integrability (NLI) condition of the distortion function $\tau = \log |f'|$, the introduction of holonomy parameters, the fractal nature of the Julia set, and understanding how all of these interact with the intricate Dolgopyat machinery.

Here is a rough outline of our proof of the spectral bounds for $L_{s,\ell}$.

1. For a test function $\phi$, the transfer operator $(L_{s,\ell}\phi)(x)$ is given as a sum of complex valued summand functions (see subsection 2.3).

2. Each of the summand functions has two sources of oscillation, coming from $b := \Im(s)$ and $\ell$ respectively; using the Non-Local-Integrability condition for $f$ (see Section 3) we see that these two sources of oscillation are complimentary rather than cancelling each other out, so the summand functions are rapidly oscillating.

3. When we add up our collection of rapidly oscillating summand functions we expect to get cancellation, at least away from some smooth curve where all the summand functions might line up perfectly.

4. Since the Julia set $J$ of a hyperbolic rational map does not concentrate near smooth curves (see the Non-Concentration-Property in Section 4) it must meet the set where we have cancellation in the sum.
(5) Since the sum defining the transfer operator always experiences cancellation, its spectral radius must be smaller than expected; see Section 5.

Acknowledgement We would like to thank Dennis Sullivan for bringing our attention to this problem, and Curt McMullen for telling us about Zdunik’s work. We would also like to thank Ralf Spatzier and Mark Pollicott for helpful discussions.

2. Hyperbolic rational maps and Ruelle operators

2.1. Iterates of rational maps. Let \( f : \hat{\mathbb{C}} \to \hat{\mathbb{C}} \) be a rational map of degree \( d \) at least two. Below we list some basic definitions and facts on dynamics of the iterates of \( f \). Useful references are [16], [4] and [15].

The domain of normality for the collection of iterates \( \{ f^n : n = 1, 2, \cdots \} \) is called the Fatou set for \( f \), and its complement is called the Julia set, which we will denoted by \( J = J(f) \). The post critical set \( P = P(f) \) means the union of all forward images of the critical points of \( f \).

A periodic orbit \( \hat{x} \) is called repelling, attracting or indifferent according as its multiplier satisfies \( |\lambda| > 1 \) or \( |\lambda| < 1 \) or \( |\lambda| = 1 \).

Throughout the paper, we assume \( f \) is hyperbolic, that is, \( f \) is eventually expanding on the Julia set \( J \) in the sense that there are constants \( \kappa_1 > \kappa > 1 \) such that

\[
\kappa^n c_0 \leq |(f^n)'(x)| \leq \kappa^n
\]

for all \( x \in J \) and for all \( n \geq 1 \). Equivalently, \( f \) is hyperbolic if the post critical closure \( \overline{P} \) is disjoint from the Julia set \( J \), or if the orbit of every critical point converges to an attracting periodic orbit. Julia sets of hyperbolic rational functions are known to have zero area. The Generic Hyperbolicity Conjecture says that every rational map can be approximated arbitrarily closely by a hyperbolic map.

Every periodic orbit of a hyperbolic map is either attracting or repelling and there are at most \( 2d - 2 \) attracting periodic cycles [16, Coro. 10.16]. Repelling cycles, on the other hand, are dense in the Julia set. For this reason we are interested exclusively in repelling cycles for the rest of the paper.

We will say that two rational functions are conjugate if they are conjugate by a holomorphic automorphism of \( \hat{\mathbb{C}} \), that is, an element of \( PSL_2(\mathbb{C}) \).

2.2. Markov partitions and symbolic dynamics for rational maps. Replacing \( f \) by a conjugate we may assume that \( \infty \notin J \). We will identify \( \hat{\mathbb{C}} - \{ \infty \} \) with \( \mathbb{C} \) throughout.

It is well understood that hyperbolic rational maps can be studied by means of symbolic dynamics. We now recall the essential features of this approach, and refer to [11] for further details. For any small \( \varepsilon > 0 \), we can
find a Markov partition $P_1, \ldots, P_{k_0}$ for $J$, that is, $P_i$'s are compact subsets of $J$ of diameter at most $\epsilon$ such that

- $J = \bigcup_{j=1}^{k_0} P_j$,
- each $P_j$ is the closure of its interior (relative to $J$), that is $\text{int}_J P_j = P_j$ for each $j$;
- the interiors are disjoint in the sense that $\text{int}_J P_i \cap \text{int}_J P_j = \emptyset$ whenever $i \neq j$;
- for each $j$, the image $f(P_j)$ is a union of partition elements $P_i$.

In fact, $f(P_j)$ is given as the union of all $P_i$ with $\text{int}_J P_j \cap f(P_i) \neq \emptyset$. Fixing a Markov parition $P_1 \ldots P_{k_0}$ of a small diameter, we may also assume the existence of open neighbourhoods $U_j$ of $P_j$ such that:

1. the map $f$ is injective on the closure of $U_j$ for each $j$;
2. $f$ is injective on the union $U_i \cup U_j$ whenever $U_i \cap U_j \neq \emptyset$;
3. for every $j$, $P_j$ is not contained in $\cup_{k \neq j} U_k$;
4. for every pair $i, j$ with $f(P_i) \supset P_j$ we have a local inverse $g_{ij} : U_j \to U_i$ for $f$.

Associated to $f$ we have the $k_0 \times k_0$ transition matrix

$$M_{ij} = \begin{cases} 1 & \text{if } f(P_i) \supset P_j, \\ 0 & \text{else.} \end{cases}$$

Note that $M$ is topologically mixing, as $f$ images of any open set meeting $J$ eventually cover $J$ (see [16, Corollary 14.2]). Denote by $\Sigma^+$ the shift space of admissible sequences

$$\Sigma^+ = \{ \omega = (\omega_1, \omega_2, \ldots) \in \{1, \ldots, k_0\}^\mathbb{N} : M_{\omega_k \omega_{k+1}} = 1 \text{ for all } k \},$$

and by $\sigma : \Sigma^+ \to \Sigma^+$ the shift map $(\sigma \omega)_k = \omega_{k+1}$. The relationship between this symbolic space $(\Sigma^+, \sigma)$ and the complex dynamics of $(J, f)$ is described by [11, Proposition 2.2]; there is a bounded to one surjective continuous map

$$\Sigma^+ \to J$$

given by

$$(\omega_1, \omega_2, \ldots) \mapsto \cap_{i=1}^{\infty} f^{1-k} P_{\omega_k}.$$

This map intertwines $\sigma$ and $f$.

**Notation 2.1.** For an admissible multi-index $I = (i_r, \ldots, i_1) \in \{1, \ldots, k_0\}^{r}$ we write

$$g_I := g_{i_r, i_{r-1}} \circ \cdots \circ g_{i_2, i_1} : U_{i_1} \to U_{i_r}.$$

We record two elementary remarks about Markov partitions for later use.

**Lemma 2.2.** If $f(x) = y$ and $y$ is in the interior of some $P_k$, then $x$ is in the interior of some $P_j$.

**Proof.** Suppose that $x \in P_j$. Consider any sequence $x_\ell \to x$ in $U_j \cap J$. Then $f(x_\ell) \to y$, so for some $\ell_0$, $f(x_\ell) \in P_k$ for all $\ell > \ell_0$. As $f(P_j) \supset P_k$, there exists $z_\ell \in P_j$ such that $f(x_\ell) = f(z_\ell)$ for all $\ell > \ell_0$. And since $f$ is injective
on \( U_j \) we have that \( x_\ell = z_\ell \in P_j \) eventually. Therefore \( x \) is in the interior of \( P_j \).

\[ \square \]

**Lemma 2.3.** Suppose that we have two admissible sequences \( I, J \) of same length \( m \geq 2 \) and that \( i_1 = j_1 = j \). If \( I \neq J \) then \( g_I(U_j) \cap g_J(U_j) = \emptyset \).

**Proof.** We argue by induction on \( m \). Suppose first that \( m = 2 \). Choose a point \( z \) in the interior of \( P_j \). Note that \( g_{j_2,j}(z) \) and \( g_{i_2,j}(z) \) are in the interiors of \( P_{j_2}, P_{i_2} \) respectively. If they are equal, then we conclude \( P_{j_2} = P_{i_2} \) and \( j_2 = i_2 \). If they are unequal, then we see that \( f \) is not injective on \( U_{j_2} \cup U_{i_2} \), and so that \( U_{j_2} \cap U_{i_2} = \emptyset \) by condition (2) on the \( U_i \) in Subsection 2.2. This provides our base case.

Now suppose that we have the lemma for sequences of length \( m - 1 \). Let \( I, J \) be sequences of length \( m \geq 3 \) as above, and suppose that \( x \in g_I(U_j) \cap g_J(U_j) \). Then

\[ f(x) \in g_{i_{m-1},...,i_2,j}(U_j) \cap g_{j_{m-1},...,j_2,j}(U_j) \subset U_{i_{m-1}} \]

hence \( (i_{m-1},\ldots,i_2,j) = (j_{m-1},\ldots,j_2,j) \). Thus

\[ x \in g_{j_{m-j_{m-1}}}(U_{j_{m-1}}) \cap g_{i_{m-j_{m-1}}}(U_{j_{m-1}}) \]

and so \( i_m = j_m \) again by the inductive assumption. This completes the proof. \[ \square \]

Associated to \( f \), we have the distortion function

\[ \tau(z) = \log |f'(z)| \]

and a rotation

\[ \theta(z) = \arg(f'(z)) \in \mathbb{R}/2\pi\mathbb{Z}, \]

which are both defined and analytic away from the critical set. We write

\[ \tau_N(z) := \sum_{j=0}^{N-1} \tau(f^j z) \text{ and } \theta_N(z) := \sum_{j=0}^{N-1} \theta(f^j z), \]

and denote \( \alpha(z) = e^{i\theta(z)} \) and \( \alpha_N(z) = e^{i\theta_N(z)} \). Note that \( \tau \) is eventually positive on \( J \), that is \( \tau_N > 0 \) for some \( N \).

We denote by \( \delta = \delta(f) \) the unique positive zero of the pressure

\[ P(-s\tau) := \sup_\nu \{ h_f(\nu) - \int s\tau d\nu \} \]

where the supremum is taken over all \( f \) invariant probability measures on \( J \) and \( h_f(\nu) \) is the measure-theoretic entropy of \( f \).

There exists a unique \( f \)-invariant probability measure \( \nu = \nu_{-\delta\tau} \) on \( J \), called the equilibrium state of the potential \(-\delta\tau\), such that \( P(-\delta\tau) = h_f(\nu) - \delta \int \tau d\nu = 0 \). The measure \( \nu \) is equivalent to a \( \delta \) dimensional Hausdorff measure on \( J \) (see [2, 17]). It follows that \( \delta \) is the Hausdorff dimension of the Julia set, and that \( 0 < \delta < 2 \).
Notation 2.4. We write $U$ for the disjoint union of the $U_j$’s. Abusing notation we regard $\tau$ and $\theta$ as functions on $U$. We regard $\nu$ as a measure on $U$ by taking $\nu = \sum \nu_j$ where $\nu_j$ is the restriction of $\nu$ to the copy of $P_j$ sitting inside $U_j$. Since the boundary points of $\nu$ have zero mass this gives a probability measure on $U$.

2.3. Ruelle transfer operators. Our main results will follow from spectral estimates on twisted transfer operators, which we now introduce. We write $C^1(U)$ for $C^1(U, \mathbb{C})$. For a character $\chi : S^1 \to S^1$ and a complex number $s \in \mathbb{C}$, we consider the following transfer operators

$$L_{s,\chi} : C^1(U) \to C^1(U)$$

$$(L_{s,\chi} h)(x) := \sum_{i:M_{ij}=1} e^{-s\tau(g_{ij}x)} \chi(\alpha(g_{ij}x)) h(g_{ij}x) \text{ when } x \in U_j.$$ 

We will sometimes write $\chi_\ell$ for the character $\chi_\ell(\alpha) = \alpha^\ell$, and

$$L_{s,\ell} := L_{s,\chi_\ell}.$$ 

We will denote $L_s := L_{s,0}$.

Our main technical result in this paper is the following:

Theorem 2.5.  
(1) Suppose that $f$ is not conjugate to a monomial. Then, for any $\epsilon > 0$, there exist $C_\epsilon > 0$, $0 < \epsilon_0 = \epsilon_0(\epsilon) < 1$ and $0 < \rho_\epsilon < 1$ such that for any $n \in \mathbb{N}$, for all $\Re(s) > \delta - \epsilon_0$ and $|\Im(s)| > 1$,

$$\|L_s^n\|_{C^1} \leq C_\epsilon |\Im(s)|^{1+\epsilon} \rho_\epsilon^n.$$ 

(2) Suppose that the Julia set of $f$ is not contained in a circle. Then, for any $\epsilon > 0$, there exist $C_\epsilon > 0$, $0 < \epsilon_0 = \epsilon_0(\epsilon) < 1$ and $0 < \rho_\epsilon < 1$ such that for any $n \in \mathbb{N}$, for all $\Re(s) > \delta - \epsilon_0$ and $|\Im(s)| + |\ell| > 1$,

$$\|L_{s,\ell}^n\|_{C^1} \leq C_\epsilon (|\Im(s)| + |\ell|)^{1+\epsilon} \rho_\epsilon^n.$$ 

Notation 2.6. For a holomorphic function $g : \mathbb{C} \to \mathbb{C}$ we denote the derivative by $g'$ as usual.

Often we will have to deal with real valued functions instead, in which case we denote the gradient of $h : \mathbb{C} = \mathbb{R}^2 \to \mathbb{R}$ by $\nabla h : \mathbb{C} = \mathbb{R}^2 \to \mathbb{R}^2$. We will denote the Euclidean norm in $\mathbb{R}^2$ by $| \cdot |$.

Sometimes we will be driven to consider functions that are complex valued but not necessarily holomorphic. For these we use the notation $\nabla h$ to mean the Jacobian of $h$ regarded as a map $\mathbb{R}^2 \to \mathbb{R}^2$, and write $|\nabla h|$ to mean the operator norm of the matrix $\nabla h$. We write $\|h\|_{C^1(W)}$ for the supremum of $|\nabla h(u)|$ as $u$ ranges over $W$.

In two instances we shall be forced to consider $C^2$ norms of a function $h$. By this we simply mean the max of the $C^1$ norm and the supremum of $X^2 h$ as $X$ ranges over unit tangent vectors.
We define a modified $C^1$ norm on $C^1(U)$ by taking
\begin{equation}
||h||_r := \begin{cases} 
||h||_{\infty} + ||\nabla h||_{\infty} & \text{if } r \geq 1 \\
||h||_{\infty} + ||\nabla h||_{\infty} & \text{if } 0 < r < 1.
\end{cases}
\end{equation}
(The point of this is that the operators $L_{s,\ell}$ are uniformly bounded in $||\cdot||_{|b|+|\ell|}$ norm at least for large $|b| + |\ell|$, whereas they are not uniformly bounded in the usual $C^1$ norm.)

By the arguments of Ruelle, Theorem 2.5 is a consequence of the following estimates (see [8] and [9, Section 5] for readable accounts).

**Theorem 2.7.**

(1) Suppose that $f$ is not conjugate to a monomial. Then there exist $C > 0$, $\rho \in (0, 1)$ and $\epsilon_0 > 0$ such that for any $h \in C^1(U)$ and any $n \in \mathbb{N}$,
\begin{equation}
||L^n_s h||_{L^2(\nu)} \leq C\rho^n ||h||_{3(s)}
\end{equation}
for all $s \in \mathbb{C}$ with $\Re(s) > \delta - \epsilon_0$ and $|\Im(s)| \geq 1$.

(2) Suppose that the Julia set of $f$ is not contained in a circle. Then there exist $C > 0$, $\rho \in (0, 1)$ and $\epsilon_0 > 0$ such that for any $h \in C^1(U)$ and any $n \in \mathbb{N}$,
\begin{equation}
||L^n_{s,\ell} h||_{L^2(\nu)} \leq C\rho^n ||h||_{3(s)+|\ell|}
\end{equation}
for all $s \in \mathbb{C}$ and $\ell \in \mathbb{Z}$ with $\Re(s) > \delta - \epsilon_0$ and $|\Im(s)| + |\ell| \geq 1$.

After preparations in Sections 3 and 4, the proof of Theorem 2.7 will be completed in Section 5.

3. **Non-Local-Integrability and hyperbolic rational maps**

The goal of this section is to establish the Non-Local-Integrability condition for hyperbolic rational maps which are not conjugate to monomials: see Theorem 3.4. We begin with a precise formulation of the NLI condition.

3.1. **The NLI condition.** We retain the notation of subsection 2.2. Let $f$ be a hyperbolic rational map of degree $d \geq 2$. Consider an admissible sequence $(\ldots, \xi_{-2}, \xi_{-1}, \xi_0)$ with local inverses $g_{\xi_{-k},\xi_{-(k+1)}} : U_{\xi_{-k}} \to U_{\xi_{-(k+1)}}$, $k \geq 0$. For any fixed $n \geq 0$, we have the section
g^n_{\xi} := g_{(\xi_{-n}, \ldots, \xi_{-1}, \xi_0)} : U_{\xi_0} \to U_{\xi_{-n}}
of $f^n$ defined on $U_{\xi_0}$. The sum $\sum \tau_n(g^n_{\xi}(x))$ always diverges for any $x \in U_{\xi_0}$ as a consequence of the eventual positivity of $\tau$. On the other hand
\begin{equation}
\tau_\infty(\xi, x, y) := \sum_{1}^{\infty} (\tau(g^n_{\xi}(x)) - \tau(g^n_{\xi}(y)))
\end{equation}
always converges for any pair $x, y \in U_{\xi_0}$, just using the contraction properties of $g^n_{\xi}$ and the Lipschitz property of $\tau$. 
Definition 3.1 (Non-Local-Integrability). The function $\tau$ satisfies the NLI property if there exist $j \in \{1, \ldots, k_0\}$, points $x_0, x_1 \in P_j$, and admissible sequences $(\ldots \xi_{-2}, \xi_{-1}, j), (\ldots \xi_{-2}, \xi_{-1}, j)$ with property that the gradient of
\begin{equation}
\hat{\tau}(x) := \tau_\infty(\xi, x, x_0) - \tau_\infty(\xi, x, x_0)
\end{equation}
is non-zero at $x_1$; note if this holds for any single choice of $x_0 \in P_j$, it must hold for all choices of $x_0 \in P_j$.

In fact it will be more convenient to use the following reformulation of the NLI property.

Lemma 3.2 (Non-Local-Integrability II). Suppose that $\tau$ satisfies the NLI property with respect to sequences $\xi, \hat{\xi}$ and the points $x_0, x_1 \in P_j$. Then there exists an open neighbourhood $U_0$ of $x_1$ and constants $\delta_2 \in (0, 1), N \in \mathbb{N}$ such that the following holds: for any $n \geq N$, the map
\begin{equation}
(\tilde{\tau}, \tilde{\theta}) := (\tau_n \circ g_n^\xi - \tau_n \circ g_n^{\hat{\xi}}, \theta_n \circ g_n^\xi - \theta_n \circ g_n^{\hat{\xi}}) : U_0 \to \mathbb{R} \times \mathbb{R}/2\pi \mathbb{Z}
\end{equation}
is a local diffeomorphism satisfying $\|\tilde{\tau}\|_{C^2} < \frac{1}{2\delta_2}$ and
\[ \inf_{u \in U_0} |\nabla (\tilde{\tau}, \tilde{\theta})(u) \cdot v| \geq \delta_2 |v| \quad \text{for all } v \in \mathbb{R}^2. \]

Proof. Define
\begin{equation}
\theta_\infty(\xi, x, y) := \sum_1^\infty \theta(g_n^\xi(x)) - \theta(g_n^{\hat{\xi}}(y)) \in \mathbb{R}/2\pi \mathbb{Z},
\end{equation}
which is again convergent on $U_{\xi_0} \times U_{\xi_0}$.

Since $f$ is holomorphic we see that $h := e^{\tilde{\tau} + i\tilde{\theta}}$ is holomorphic. The NLI property for $\tau$ implies then that the derivative of $h$ is non-zero, and so that $h$ defines a local diffeomorphism from a neighbourhood of $x_1$ to some open subset of $\mathbb{C}$. Since the derivatives of (3.3) converge locally uniformly and in $C^1$ norm to log $h$, we are finished. \qed

We will also need the following observation later:

Proposition 3.3. If the Julia set $J$ of $f$ is contained in $\mathbb{R}$, then
\[ \tilde{\tau}|_{U_0 \cap \mathbb{R}} : U_0 \cap \mathbb{R} \to \mathbb{R} \]
is a local diffeomorphism satisfying $\|\tilde{\tau}|_{U_0 \cap \mathbb{R}}\|_{C^2} < \frac{1}{2\delta_2}$ and $|\nabla \tilde{\tau}| > \delta_2$ on $U_0 \cap \mathbb{R}$.

Proof. Since $f$ preserves $J \subset \mathbb{R}$, it also preserves $\mathbb{R}$. Thus $\alpha := e^{2\pi i \theta}$ is locally constant on $\mathbb{R}$ away from the critical set and takes values only in $\{1, -1\}$. Hence $\theta$ is locally constant on $\mathbb{R} \cap J$. It follows that
\[ \frac{\partial}{\partial x} \tilde{\tau} \]
is non-vanishing at some point of $J$; in the notation of Lemma 3.2, this means that
\[
\tau : U_0 \cap \mathbb{R} \to \mathbb{R}
\]
is a local diffeomorphism with controlled $C^2$-norm as above. \hfill \square

For a hyperbolic rational map conjugate to $f(x) = x^d$, $\tau$ is cohomologous
to a constant function on $J$, and hence does not have the NLI property. We
show that this is the only obstruction:

**Theorem 3.4.** For a hyperbolic rational function $f : \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ of degree at
least 2, the distortion function $\tau = \log |f'|$ on $J$ satisfies the NLI property if
and only if $f$ is not conjugate to $f(x) = x^d$ for any $d \in \mathbb{N}$.

**Remark 3.5.** In the above theorem, $x_1$ can be chosen to be any point of
$P_j$ with at most finitely many exceptions. This is because of the fact that
he collection of critical points for $\hat{\tau}$ is either discrete or everything, since $\hat{\tau}$
is the real part of a holomorphic function. Thus if the NLI property holds
for some $x_1$ in $P_j$ it holds for almost every $x_1$ in $P_j$.

The rest of this section is devoted to the proof of Theorem 3.4. Through-
out we assume that
\[
f \text{ is not conjugate to } x^d \text{ for any } d \in \mathbb{N}.
\]

3.2. **Inverse branches for $f$ and universal covers.** Recall that the hyper-
bolicity of $f$ implies that the post critical closure $\mathcal{P}$ of $f$ is equal to the
union of the countable set $P$ of the forward orbits of critical points with the
finite set $A$ of attracting critical points for $f$. Note also that $\mathcal{P}$ contains at
least three points; otherwise $f$ is conjugate to a map $x^d$ for some $d \in \mathbb{N}$.
We write
\[
\Omega := \hat{\mathbb{C}} - \mathcal{P}
\]
and note that $\Omega$ is a connected Riemann surface, admitting a hyperbolic
metric.

Let $\hat{\Omega}$ be the universal cover of $\Omega$. We write $\hat{d}$ for the hyperbolic metric
on $\hat{\Omega}$, and $\pi : \hat{\Omega} \to \Omega$ for the covering map.

**Definition** We say that a holomorphic function $\hat{g} : \hat{\Omega} \to \hat{\Omega}$ is an inverse
branch for $f$ if
\[
f(\pi(\hat{g}(z))) = \pi(z) \quad \text{for all } z \in \hat{\Omega}.
\]

**Lemma 3.6.** For any pair $\hat{z}_1, \hat{z}_2 \in \hat{\Omega}$ with the property $f(\pi(\hat{z}_1)) = \pi(\hat{z}_2)$
there exists a unique inverse branch $\hat{g}$ of $f$ with $\hat{g}(\hat{z}_2) = \hat{z}_1$. The image of
any inverse branch is $\hat{\Omega} - \pi^{-1}(f^{-1}(\mathcal{P}))$.

**Proof.** We note that the maps
\[
f : \Omega - f^{-1}(\mathcal{P}) \to \Omega
\]
and
\[
\pi : \hat{\Omega} - \pi^{-1}(f^{-1}(\mathcal{P})) \to \Omega - f^{-1}(\mathcal{P})
\]
are covering maps. It follows that their composition is also a covering map, and the lemma follows by universality of $\pi : \tilde{\Omega} \to \Omega$. \hfill \Box

If $\tilde{g}$ is an inverse branch for $f$, then it is a strict contraction for the hyperbolic metric $\tilde{d}$ and has a unique fixed point $\tilde{\beta}$ in $\tilde{\Omega}$ by [16, Theorem 5.2]. Note that the projection $\pi(\tilde{\beta})$ of a fixed point of an inverse branch is necessarily a fixed point of $f$.

3.3. Normalized distortion functions and the $\pi_1$ action. For inverse branches $\tilde{g}_0, \tilde{g}_1$ of $f$ with fixed points $\tilde{\beta}_i$ and $\beta_i := \pi(\tilde{\beta}_i)$, we form normalized distortion functions $\tilde{h}_0, \tilde{h}_1$ on $\tilde{\Omega}$ as follows:

$$\tilde{h}_j(z) := \prod_{k=1}^\infty \frac{f'(\pi(\tilde{g}_j^k z))}{\lambda(\beta_j)}.$$

These are convergent, holomorphic and non-vanishing on $\tilde{\Omega}$ and satisfy functional equations

$$\frac{\tilde{h}_j(z)}{\tilde{h}_j(\tilde{g}_j(z))} = \frac{f'(\pi(\tilde{g}_j z))}{\lambda(\beta_j)}.$$

This is the moment at which NLI comes into play.

**Lemma 3.7.** Suppose that NLI fails. Then $\tilde{h}_0$ satisfies both of these functional equations:

$$\frac{\tilde{h}_0(z)}{\tilde{h}_0(\tilde{g}_j(z))} = \frac{f'(\pi(\tilde{g}_j z))}{\lambda(\beta_j)} \quad \text{for } j = 0, 1.$$

**Proof.** It suffices to show that the ratio $H(z) := \tilde{h}_0(z)/\tilde{h}_1(z)$ is constant on $\tilde{\Omega}$. We work with the contrapositive. Assume $H(z)$ is non-constant. As $H$ is holomorphic and $J$ is a perfect set, it follows that there exists a point, say $z_0$, in the cover $\pi^{-1}(J)$ where $H'(z_0) \neq 0$. It will be convenient to renormalize once more and choose

$$\hat{H}(z) := \frac{H(z)\tilde{h}_1(z_0)}{\tilde{h}_0(z_0)}.$$

Note that $\hat{H}'(z_0) \neq 0$, too.

Without loss of generality we may assume that $\pi(z_0)$ is in the interior of some element $P_\xi$ of the Markov partition. Now the sequence $\pi(\tilde{g}_0^k z_0)$ is a sequence of pre-images for $\pi(z_0)$ under $f$ and lands in the interior of a partition element $P_{\xi-k}$. Similarly we choose $\xi$ such that $\pi(\tilde{g}_1^k z_0) \in P_{\xi-k}$. Unwrapping the definitions we have that

$$\hat{H}(z) = \exp(\tau_\infty(\xi, \pi(z), \beta_0) - \tau_\infty(\xi, \pi(z), \beta_1))$$

$$\times \exp(2\pi i \theta_\infty(\xi, \pi(z), \beta_0) - 2\pi i \theta_\infty(\xi, \pi(z), \beta_1)).$$

The left hand side of this equation is a holomorphic function with non-zero derivative at $z = z_0$, and hence automatically a local diffeomorphism of a small neighbourhood of $z_0$ into a small neighbourhood of it’s image $1 \in \mathbb{C}$. 
This contradicts the assumption that NLI fails and hence completes the proof. □

Fix an inverse branch \( \tilde{g}_0 \) for \( f \) and write \( \tilde{h}_0, \tilde{\beta}_0, \beta_0 \) for the associated normalized distortion function, fixed point in \( \tilde{\Omega} \) and projected fixed point in \( \Omega \) respectively. The fundamental group \( \pi_1(\Omega) \) has a tautological action on the universal cover \( \tilde{\Omega} \); this extends to an action on the set of inverse branches for \( f \) by choosing

\[
(\alpha \cdot \tilde{g})(z) := \alpha(\tilde{g}(z)).
\]

**Proposition 3.8.** Suppose that NLI fails. Then
1. \( |\tilde{h}_0| \) is invariant for the \( \pi_1(\Omega) \)-action on \( \pi^{-1}(J) \);
2. If \( |h_0| \) denotes the projection of \( |\tilde{h}_0| \) to \( J \), then
   \[
   \frac{|h_0(x)||f'(x)|}{|h_0(f(x))|} = d^{1/\delta} \text{ for all } x \in J.
   \]

That is, \( \tau \) is cohomologous to the constant function \( \log d^{1/\delta} \) on \( J \).

**Proof.** Let \( \alpha \in \pi_1(\Omega) \). We denote the fixed point of \( \alpha \cdot \tilde{g}_0 \) by \( \tilde{\beta}_0 \) and write \( \beta_0 := \pi(\tilde{\beta}(\alpha)) \). By considering the functional equations (3.6) for \( \tilde{g}_0 \) and \( \alpha \cdot \tilde{g}_0 \) we deduce from Lemma 3.7 that for all \( z \in \tilde{\Omega} \),

\[
(3.10) \quad \frac{\tilde{h}_0(\alpha \tilde{g}_0(z))}{\tilde{h}_0(\tilde{g}_0(z))} = \frac{\lambda(\beta_0)}{\lambda(\beta(z))}.
\]

We claim now that \( |\tilde{h}_0| \) is invariant for the \( \pi_1(\Omega) \) action on \( \tilde{\Omega} \); if not then we may choose \( \alpha \in \pi_1(\Omega) \) and \( z \in \tilde{\Omega} \) such that \( \frac{|h_0(\alpha \tilde{g}_0(z))|}{|h_0(\tilde{g}_0(z))|} < 1 \). By Lemma 3.6, \( \tilde{g}_0 \) is surjective, and hence we can choose \( z_j \in \Omega \) such that

\[
\tilde{g}_0(z_j) = \alpha^{j-1} \tilde{g}_0(z).
\]

But then for any \( k \geq 1 \),

\[
\left| \frac{\lambda(\beta(\alpha^k))}{\lambda(\beta_0)} \right| = \left| \frac{\tilde{h}_0(\alpha^k \tilde{g}_0(z))}{\tilde{h}_0(\tilde{g}_0(z))} \right| \text{ by (3.10)}
\]

\[
= \prod_{1}^{k} \left| \frac{\tilde{h}_0(\alpha^j \tilde{g}_0(z))}{\tilde{h}_0(\alpha^j \tilde{g}_0(z))} \right| = \prod_{1}^{k} \left| \frac{\tilde{h}_0(\alpha^j \tilde{g}_0(z_j))}{\tilde{h}_0(\tilde{g}_0(z_j))} \right| = \left| \frac{\lambda(\beta(\alpha))}{\lambda(\beta_0)} \right|^k \text{ by (3.10)}.
\]

The final terms here clearly converges to \( \infty \) with \( k \to \infty \), but the collection of fixed points for \( f \) is finite, so there are finitely many multipliers to choose from for \( \lambda(\beta(\alpha^k)) \); this is contradiction, proving the claim (1).
To prove (2), choose \( x \in J \) and set \( y = fx \). Choose lifts \( \tilde{x}, \tilde{y} \) of \( x, y \) to \( \pi^{-1}(J) \) and an inverse branch \( \tilde{g} \) of \( f \) such that \( \tilde{g}(\tilde{y}) = \tilde{x} \). Let \( \tilde{\beta} \) be the fixed point of \( \tilde{g} \) and \( \beta := \pi(\tilde{\beta}) \). Then

\[
|f'(x)| = |f'(\pi(\tilde{g}(\tilde{y})))| = \frac{|\lambda(\beta)| |h_0(\tilde{g})|}{|h_0(\tilde{g}\tilde{y})|} \quad \text{by (3.6)}
\]

\[
= \frac{|\lambda(\beta)| |h_0(\tilde{g})|}{|h_0(\tilde{x})|} = \frac{|\lambda(\beta)| |h_0(y)|}{|h_0(x)|} \quad \text{by (1)}
\]

\[
= \frac{|\lambda(\beta)| |h_0(f(x))|}{|h_0(x)|}.
\]

In other words, for any \( x \in J \),

\[
|\lambda(\beta)| = \frac{|h_0(x)||f'(x)|}{|h_0(f(x))|}.
\]

Therefore

\[
\log |f'(x)| = \log |h_0(f(x))| - \log |h_0(x)| + \log |\lambda(\beta)|,
\]

that is, \( \tau = \log |f'| \) is cohomologous to the constant \( \log |\lambda(\beta)| \). Since the equilibrium state \( \nu \) for \( -\delta \tau \) must be the measure of maximal entropy for \( f \), and the topological entropy of \( f \) is given as the logarithm of the degree of \( f \), it follows that \( \log |\lambda(\beta)| = \log d^{1/\delta} \). This finishes the proof. \( \square \)

We now recall work of Zdunik [18, Corollary in section 7 and Proposition 8].

**Theorem 3.9.** Suppose that \( f \) is hyperbolic and \( \delta \tau \) is cohomologous to \( \log d \). Then \( f \) is conjugate to the map \( x \pm d \).

**Proof of Theorem 3.4.** Suppose \( f \) is a hyperbolic rational function with degree at least 2 and not conjugate to \( z^{\pm d} \). If NLI fails then Proposition 3.8 shows that \( \delta \tau \) is cohomologous (on \( J \)) to \( \log d \); however this is impossible by Zdunik’s result. \( \square \)

4. **Non-Concentration and doubling for hyperbolic Julia sets**

4.1. **Non-concentration.** As before, let \( f \) be a hyperbolic rational map of degree \( d \geq 2 \), and keep the notation from Section 2.2. In this section we shall address non-concentration properties for hyperbolic Julia sets. We will also recall that the associated measures have the doubling property on each cylinder.

**Notation 4.1.** We write cylinders of length \( r \in \mathbb{N} \) as

\[
C([i_1, \ldots, i_r]) := \{ x \in J : f^{j-1}x \in P_{i_j} \text{ for } 1 \leq j \leq r \}.
\]
Note that we can regard cylinders either as subsets of $J$ or as subsets of $P_i \subset U_i \subset U$.

Let $J$ denote the Julia set of $f$.

**Definition 4.2** (The Non-Concentration Property). The Julia set $J$ has the Non-Concentration Property (NCP) if, for each cylinder $C$ of $J$, there exists $0 < \delta_1 < 1$ such that, for all $x \in C$, all $w \in \mathbb{C}$ of unit length, and all $\epsilon \in (0, 1)$,

\begin{equation}
B_\epsilon(x) \cap \{y \in C : |\langle y - x, w \rangle| > \delta_1 \epsilon\} \neq \emptyset
\end{equation}

where $\langle a + bi, c + di \rangle = ac + bd$ for $a, b, c, d \in \mathbb{R}$.

It is clear that the NCP must fail whenever $J$ is contained in a circle, in which case we refer to Theorem 4.3(2) for the required modification.

The NCP is a consequence of quasi-self-similarity of the Julia set: in some precise sense $J$ looks the same at every length scale. We will describe the required aspects of this notion as they appear.

We start by choosing some constants to be used for the rest of the section. For each partition element $P_j \subset U_j$, choose a neighbourhood $P_j \subset \Omega_j \subset U_j$ with $\Omega_j \subset U_j$. Note that we can find positive integers $K_1, K_2, K_3, K_4, L$ such that

1. for any $x \in J$ and any $\epsilon \in (0, 1)$, there exists $k \in \mathbb{N}$ with
   \[ \frac{1}{K_1 \epsilon} \leq |(f^k)'(x)| \leq \frac{1}{\epsilon}; \]
2. for any $j$ and any $x \in P_j$, we have $B_{3/K_3}(x) \subset \Omega_j$;
3. for any $j$ and any $x \in \Omega_j$, we have $B_{3/K_2}(x) \subset U_j$;
4. for any admissible sequence $I = (i_1, \ldots, i_r)$ and associated map $g_I : U_{i_r} \to U_{i_1}$, we have
   \[ \frac{|g_I'(x)|}{|g_I'(y)|} \leq K_3 \quad \text{for all } x, y \in \Omega_{i_r}; \]
5. for any $x \in \Omega_j$, any $r \in (0, 1/K_2)$, and any univalent map $T : U_j \to \mathbb{C}$, we have that
   \[ B_{|T'(x)|r/K_4}(T(x)) \subset T(B_r(x)) \subset B_{|T'(x)|rK_4}(T(x)); \]
6. every cylinder of length $L$ has diameter less than $\frac{1}{2K_1K_2K_3K_4}$.

The statement (1) follows from the hyperbolicity of $f$, and the statements (4) and (5) follow from Koebe’s Distortion theorem (cf. [4]). Other statements are clear.

**Theorem 4.3** (NCP). (1) If $J$ is not contained in a circle, then the NCP holds for $J$.

(2) If $J$ is contained in $\mathbb{R}$, then for each cylinder $C$ of $J$, there exists $0 < \delta_1 < 1$ such that, for all $x \in C$, and all $\epsilon > 0$,

\[ B_\epsilon(x) \cap \{y \in C : |y - x| > \delta_1 \epsilon\} \neq \emptyset. \]
Proof. We address part (1) first. Suppose that \( J \) is not contained in a circle. Fix now some cylinder \( C = C([i_1, \ldots, i_r]) \) of length \( r \) and we assert that the outcome of NCP holds for \( C \). If not, we may choose \( x_n \in C, \epsilon_n \to 0 \) and \( w_n \in \mathbb{C} \) of modulus one with which the NCP property fails, that is

\[
\frac{1}{\epsilon_n}(y_n, -x_n, w_n) \to 0
\]

for any sequence \( y_n \in B_{\epsilon_n}(x_n) \cap C \). Without loss of generality, we assume that \( \epsilon_n \in (0, 1) \). Choose \( k_n \in \mathbb{N} \) such that \( \frac{1}{K_{1\epsilon_n}} \leq |(f^{k_n})'(x_n)| \leq \frac{1}{\epsilon_n} \). We may also assume that \( k_n > r + 1 \).

Let \( C(I_n) = C([i^n_1, \ldots, i^n_{k_n}]) \) be a length \( k_n \)-subcylinder of \( C \) containing \( x_n \), and set \( y_n = f^{k_n-1}(x_n) \in P_{i_{k_n}} \). We renormalize via the map

\[
\phi_n : B_{\epsilon_n}(x_n) \to B_1(0), \quad y \mapsto \frac{y - x_n}{\epsilon_n}.
\]

Now choose a subcylinder \( D_n \) of length \( L \) in \( P_{i_{k_n}} \) containing \( y_n \). We then have that \( g_{I_n} \) maps \( y_n \in D_n \subset B_{1/(K_1K_2K_4)}(y_n) \) to \( x_n \in C \cap B_{\epsilon_n/K_2}(x_n) \) which then maps via \( \phi_n \) to \( B_1(0) \).

We now consider the composition

\[
\phi_n \circ g_{I_n} : B_{1/(K_1K_2K_4)}(y_n) \to B_1(0).
\]

Note that the derivative of \( \phi_n \circ g_{I_n} \) at \( y_n \) is bounded both above and below independent of \( n \). Now, by passing to a subsequence, we may assume that \( D_n \) is constant independent of \( n \), that \( y_n \) converges to some \( y_\infty \in D_n \), that \( w_n \) converges to some \( w_\infty \in S^1 \) and that \( \phi_n \circ g_{I_n} \) converges locally uniformly to a non-constant univalent function

\[
g_\infty : B_{1/(2K_1K_2K_4)}(y_\infty) \to B_1(0).
\]

We see that the non-empty open subset \( D_n \cap B_{1/(2K_1K_2K_4)}(y_\infty) \) is contained in the smooth curve \( g_\infty^{-1}(L_\infty \cap \text{Image}(g_\infty)) \) where \( L_\infty \) is the perpendicular line to \( w_\infty \). But then for some \( N \geq 1 \), \( f^N(D_n \cap B_{1/(2K_1K_2K_4)}(y_\infty)) \) contains \( J \) [16, Corollary 14.2]. Therefore \( J \) is contained in a smooth curve. Work of Eremenko-Von Strien [13] implies that \( J \) must then be contained in a circle.

Now address part (2). Now suppose \( J \) is contained in the real line. Then the above argument shows that the failure of (2) implies that \( x \) is an isolated point in \( J \), which is a contradiction as \( J \) is a perfect set. \( \square \)

4.2. Doubling of the conformal measure on cylinders. Recall the equilibrium measure \( \nu \) on \( J \) and its restriction \( \nu_j \) to \( P_j \) for each \( j \) from (2.4). The doubling property for \( \nu \) itself is a straightforward consequence of the fact that for all small \( \epsilon > 0 \), \( \nu(B_\epsilon(x)) \) is equivalent to \( \epsilon^\delta \) up to bounded constants [17]. It is however important for later arguments that the measures \( \nu_j \) also have the doubling property.
Proposition 4.4 (Doubling). For each \( j \), the measure \( \nu_j \) has the doubling property. That is, there exists \( C > 1 \) (called the doubling constant) such that for any \( x \in P_j \) and any \( \epsilon > 0 \):

\[
\nu_j(B_{2\epsilon}(x)) \leq C \cdot \nu_j(B_{\epsilon}(x)).
\]

(4.3)

It does not seem clear a priori that the doubling property of \( \nu \) descends to the restrictions \( \nu_j \). For this reason we provide an argument which, again, is based on quasi self similarity of the Julia set. Proposition 4.4 follows from the following:

Proposition 4.5. There is a constant \( c > 0 \) with the following property. For any \( j \), any \( x \in P_j \) and any \( \epsilon > 0 \), we have

\[
\frac{\nu(B_{\epsilon}(x) \cap P_j)}{\nu(B_{\epsilon}(x))} > c.
\]

In particular, each \( \nu_j \) inherits the doubling property from \( \nu \) itself.

Proof. We retain our choices of constants \( K_1, K_2, K_3, K_4, \) and \( L \). Fix \( 1 \leq j \leq k_0 \), \( x \in P_j \) and \( \epsilon > 0 \). We choose \( k = k(x, \epsilon) \) such that

\[
\frac{1}{2K_1K_2K_3K_4\epsilon} \leq |(f^k)'(x)| \leq \frac{1}{2K_2K_3K_4\epsilon}.
\]

We claim that this \( k \) satisfies a Goldilocks property: it is neither too large nor too small.

Claim 1: \( k \) is not too large in the sense that

\[
B_{\epsilon}(x) \subset g_I(\Omega_{k+1})
\]

for every length \( k + 1 \) cylinder \( C(I) = C([i_1, \ldots, i_{k+1}]) \) meeting \( B_{\epsilon}(x) \). In particular \( f^k \) is injective on \( B_{\epsilon}(x) \).

This follows from the following calculation: suppose \( y \in C(I) \cap B_{\epsilon}(x) \), and let \( \hat{y} = f^k y \in P_{i_{k+1}} \), then

\[
g_I(\Omega_{k+1}) \supset g_I(B_{\epsilon/2K_2}(\hat{y})) \supset B_{|g_I'(\hat{y})/(K_2K_4)|}(y) \supset B_{1/(K_2K_4)(f^k)'(y))}(y) \supset B_{1/(K_2K_3K_4(f^k)'(x))}(y) \supset B_{2\epsilon}(y) \supset B_{\epsilon}(x).
\]

The fourth inclusion here requires some comment: we want to use condition (4), that is

\[
|g_I'(f^k y)/g_I'(f^k x)| < K_3;
\]

this only makes sense if \( C(I) \) contains \( x \). To get around this we first run the argument for \( y' = x \) contained in some cylinder \( C(J) \) and so obtain \( B_{\epsilon}(x) \subset g_J(\Omega_{i_{k+1}}) \). For any \( y \in B_{\epsilon}(x) \) we then have that

\[
|(f^k)'(x)/(f^k)'(y)| = |g_I'(f^k y)/g_I'(f^k x)| < K_3,
\]
as required.

**Claim 2:** \( k \) is not too small in that sense that any subcylinder \( C(I) \) of \( P_j \) of length \( k + L \) that contains \( x \) is contained in \( B_e(x) \).

For the proof, suppose \( I = (i_1, \ldots, i_{k+L}) \). Note that \( f^k(C(I)) \) is a length \( L \) cylinder, so has small diameter by the choice of \( L \), and the statement follows by applying condition (5) to \( g' \) where \( I' = (i_1, \ldots, i_{k+1}) \).

**Claim 3:** If \( r \) is the number of length \( k \) cylinders meeting \( B_e(x) \), then \( r \leq k_0 \), the number of elements in our Markov partition

Let \( I_1, \ldots, I_r \) be the length \( k \) cylinders meeting \( B_e(x) \). The claim follows from the pigeonhole principle, injectivity of \( f \) on \( B_e(x) \), and injectivity of \( f^k \) on each set \( g_I(\Omega_{j,k+1}) \). If \( r > k_0 \) then we may assume without loss of generality that \( f^{k-1}(C(I_1)) = f^{k-1}(C(I_2)) \) both give the same partition element. Choose \( y \in B_e(x) \) to be an interior point for \( C(I_1) \). Then \( f^{k-1}(y) \in f^{k-1}(C(I_1)) \). But it is then an easy exercise in injectivity to see that \( f^{k-1}(y) \notin f^{k-1}(C(I_2)) \). This gives the required contradiction.

We are now ready to prove our Proposition. Let \( C(I) \) be a length \( k + L \) subcylinder of \( P_j \) contained in \( B_e(x) \). Then

\[
\frac{\nu(B_e(x) \cap P_j)}{\nu(B_e(x))} \geq \frac{\nu(C(I))}{\nu(B_e(x))} = \frac{\nu(f^{k}C(I))}{\nu(f^{k}B_e(x))} \geq \frac{\nu(f^{k}C(I))}{\sum_1 \nu(f^{k}I_j)} \geq \frac{\min\{\nu(C) : C \text{ a cylinder of length } L\}}{k_0 \cdot \max_i \nu(P_i)}
\]

where the second inequality here follows by the \( f \)-conformality of the measure \( \nu \) and uniform bounds on \( \frac{|f^{k}(z)|}{|f^{k}(w)|} \) for \( z, w \in B_e(x) \) given by the Koebe distortion theorem.

\[\square\]

5. **Spectral bounds for transfer operators**

In the entire section, we assume that \( f \) is a hyperbolic rational function of degree at least 2 and that \( f \) is not conjugate to \( x^{\pm d} \) for any \( d \in \mathbb{N} \). Our goal in this section is to prove Theorem 2.7.

Spectral bounds for the transfer operators

\[
(L_{s,\ell}h)(x) := \sum_{i : M_{ij} = 1} e^{-s \tau(g_{ij}x)} \chi_{\ell}(\alpha(g_{ij}x)) h(g_{ij}x), \quad s \in \mathbb{C}, \ell \in \mathbb{Z}
\]

on \( C^1(U) \) will follow from the oscillatory nature of the summand terms. The role of the NLI condition is to ensure that the summand terms are rapidly oscillating (relative to one another). This ensures that we will see some cancellation in the summation for \( (L_{s,\ell}h)(x) \) at least for many \( x \in U \). We then need to ensure that some of this cancellation happens on the Julia set; that follows from the NCP, which says that the Julia set is 'too big' to avoid
the cancellation. Finally this cancellation on the Julia set must be used to prove spectral bounds for \( \hat{L} \); that is the topic of this section.

Our task for this section is to make the heuristic above precise.

5.1. Setup for the construction of Dolgopyat operators. The aim of this subsection is to establish notation and to recall some standard results that will be needed to prove Theorem 2.7. We retain the notation of the previous sections.

By the NLI property of \( \tau \) shown in Theorem 3.4 and Remark 3.5, we may choose constants as in Lemma 3.2; in other words we fix a partition element \( P \), points \( x_0, x_1 \in P \), admissible sequences \( \xi, \hat{\xi} \), a neighbourhood \( U_0 \) of \( x_1 \), and \( \delta_2 > 0 \) satisfying the conditions of that Lemma. As a notational convenience we shall assume that the \( x_1, U_0 \) described above satisfy that

\[ x_1 \in P_1 \text{ and } U_0 \subset U_1 \text{ with } \overline{U_0} \cap \overline{U_k} = \emptyset \text{ for all } k \neq 1. \]

It will be convenient to normalize our transfer operators. Ruelle showed in [11, Theorem 3.6] that \( \mathcal{L}_{\delta,0} \) has leading eigenvalue 1 with associated positive \( C^1 \) eigenfunction \( h_\delta \). We choose to work with the normalised transfer operator

\[ (\hat{\mathcal{L}}_{s,\ell}h)(x) := \frac{\mathcal{L}_{s,\ell}(h \cdot h_\delta)}{h_\delta}. \]

The convenience of this setup is that we get to assume \( \hat{\mathcal{L}}_{\delta,0}1 = 1 \). We note that it suffices to prove Theorem 2.7 for the operators \( \hat{\mathcal{L}}_{s,\ell} \). When \( \chi = 0 \) is the trivial character, we sometimes write \( \hat{\mathcal{L}}_s \) for \( \hat{\mathcal{L}}_{s,0} \).

Our next task is to fix a large number of parameters which will be needed throughout the rest of the argument. Though some of these parameters will not appear in our proofs for a good while yet, it is not an arbitrary choice to fix them now; almost the entire technical difficulty of this proof is to understand how to fix coherent choices of these parameters, and to understand that they can be chosen independent of the variables \( s, \ell \) that we are studying.

All constants we choose below are positive real numbers. Recall the expansion constants \( 1 < \kappa < \kappa_1 \) and \( c_0 \) from (2.1).

Without loss of generality we assume that \( \kappa < 2 \). Choose now \( n_1 \in \mathbb{N} \) and, for each \( 1 \leq i \leq k_0 \), a length \( n_1 + 1 \) cylinder \( X_i \) contained in \( U_0 \) such that

\[ f^{n_1}X_i = P_i. \]

Let \( \delta_i \in (0,1) \) be a constant with respect to which the cylinders \( X_1, \ldots, X_{k_0} \) satisfy the NCP as in Definition 4.2 and Theorem 4.3. Denote the minimal doubling constant for any \( \nu_j \) by \( C_3 > 1 \) given by Proposition 4.4.
Let
\[ A_0 > \frac{8}{c_0(\kappa - 1)} \max(||\tau||_{C^1}, ||h\delta||_{C^1}, ||\log h\delta||_{C^1}, ||\theta||_{C^1}) + \frac{1}{c_0} + \frac{2}{\delta_2}, \]
and
\[ E \geq 2A_0 + 1, \]
Choose \( N_0 \in \mathbb{N} \) large enough that the NLI condition from Lemma 3.2 holds, and such that
\[ 4(E + 1) < \kappa^{N_0}, \quad 160E < c_0 \delta_1 \delta_2 \kappa^{N_0}, \quad 4A_0 < \kappa^{N_0}. \]
We write
\[ (5.2) \quad v_1 := g_{\xi}^{N_0} \text{ and } v_2 := g_{\hat{\xi}}^{N_0} \]
and note that they satisfy the conclusion of Lemma 3.2. We write
\[ N = N_0 + n_1. \]
Choose
\[ \epsilon_1 \leq \min \left\{ \log 2, \frac{1}{20E}, \frac{c_0 \log 2}{160E}, \frac{\delta_1 \delta_2}{200 \kappa^{N_1} E}, \frac{\delta_2}{100} \right\}. \]
In addition we assume that \( \epsilon_1 \) is less than one tenth the distance from \( U_0 \) to the complement of \( U_1 \), that \( \epsilon_1 \kappa^{N_1} \) is less than the minimum distance from any \( P_i \) to the complement \( U_i^c \), and that \( 2\epsilon_1 \) is less than the distance from \( U_0 \) to any \( U_k, k > 1 \).
Recall from Lemma 2.3 that \( v_1(U_1) \cap v_2(U_1) = \emptyset \).
Choose
\[ \eta < \min \left\{ \frac{c_0 \epsilon_1 \delta_3}{4k_0 \kappa^{N_0}}, \frac{\delta_1 \delta_2 \epsilon_1}{512k_0}, \frac{1}{4k_0} \right\}. \]
Let \( C_1 := \exp(\log C_3 \log_2 \frac{200\kappa^{N_1}}{c_0 \delta_1 \delta_2 \kappa^{N_0}}) \), and choose \( a_0 \) so that
\[ a_0 \leq \min \left\{ \frac{\log \left( \frac{1 - \eta e^{-N A_0}}{2} \right) - \log \left( 1 - \eta e^{-N A_0} \right)}{2 N A_0}, \frac{\log \left( \frac{1 + \eta e^{-N A_0}}{8C_1} \right)}{2 N A_0}, 1 \right\}. \]
Our final choice of constant is to choose
\[ 0 < \epsilon_2 \text{ satisfying that } \left( 1 - \frac{\eta^2 e^{-2N A_0}}{64C_1^2} \right) \leq (1 - \epsilon_2)^2. \]
Consider
\[ (5.3) \quad (\tilde{\tau}, \tilde{\theta}) := (\tau_{N_0} \circ g_{\xi}^{N_0} - \tau_{N_0} \circ g_{\xi}^{N_0}, \theta_{N_0} \circ g_{\xi}^{N_0} - \theta_{N_0} \circ g_{\xi}^{N_0}), \]
as in Lemma 3.2. By a geometric sums argument and the choice of \( A_0 \), we have
\[ ||\tilde{\tau}||_{C^1} < \frac{A_0}{8} \text{ and } ||\tilde{\theta}||_{C^1} < \frac{A_0}{8}. \]
Definition. For a positive real $R$, we write $K_R(U)$ for the set of all positive functions $h \in C^1(U)$ satisfying
\begin{equation}
|\nabla h(u)| \leq Rh(u) \quad \text{for all } u \in U.
\end{equation}

5.2. Preparatory Lemmas. We recall some standard lemmas, which may be proven by direct calculation.

Lemma 5.1 (Lasota-Yorke). For all $a, b \in \mathbb{R}$ and $\ell \in \mathbb{Z}$ with $|\delta - a| < 1$, all $|b| + |\ell| \geq 1$, the following hold: fixing $B > 0$,
- if $H \in K_B(U)$, then
  \[ |\nabla \hat{L}^m_a H(u) \leq A_0 \left( 1 + \frac{B}{\kappa^m} \right) |\hat{L}^m_a H(u)| \]
  for all $m \geq 0$ and all $u \in U$;
- if $h \in C^1(U)$ and $H \in C^1(U, \mathbb{R})$ satisfy
  \[ |h(u)| < H(u) \text{ and } |\nabla h(u)| \leq BH(u) \quad \text{for all } u \in U, \]
  then
  \[ |\nabla \hat{L}^m_{a+ib,\ell} h(u) \leq A_0 \left( \frac{B}{\kappa^m} (\hat{L}^m_a H)(u) + (|b| + |\ell|)(\hat{L}^m_a |h|)(u) \right) \]
  for all $u \in U$ and any $m \in \mathbb{N}$.

Lemma 5.2. (cf. [9, Lemma 5.12]) Suppose that $z_1, z_2 \in \mathbb{C}$ are non-zero with $|z_1| \leq |z_2|$ and that the magnitude of the argument of $\frac{z_1}{z_2}$ is at least $\theta \in [0, \pi]$. Then
\begin{equation}
|z_1 + z_2| \leq \left( 1 - \frac{\theta^2}{8} \right) |z_1| + |z_2|;
\end{equation}

Lemma 5.3. (cf. [9, Lemma 5.11]) Let $R > 1$. Suppose that $h \in C^1(U)$, and $H \in K_{ER}(U)$ satisfy
\begin{equation}
|h(u)| \leq H(u) \text{ and } |\nabla h(u)| \leq ER \cdot H(u) \quad \text{for all } u \in U.
\end{equation}
Then for any $x \in U_0$ and any section $v$ of $f^{N_0}$ that is defined on $U_1$, we have either
\begin{enumerate}
  \item $|h \circ v| \leq \frac{3H_{ov}}{4}$ on $B_{10\epsilon_1/R}(x)$, or
  \item $|h \circ v| \geq \frac{H_{ov}}{4}$ on $B_{10\epsilon_1/R}(x)$.
\end{enumerate}

5.3. Construction of the Dolgopyat operators. With preliminaries out of the way we now proceed to construct our Dolgopyat operators. Throughout this section, we fix a complex number $s = a + ib$ and a character $\chi_\ell$ with $\ell \in \mathbb{Z}$. We assume that
\[ |a - \delta| < a_0 \quad \text{and} \quad |b| + |\ell| \geq 1, \]
and write $\tilde{\epsilon} = \epsilon_1/(|b| + |\ell|)$.

In order to prove Theorem 2.7, there are two cases depending on whether $J$ is contained in a circle or not. In the case when $J$ is contained in a
circle, we will assume that $J$ is contained in the real line (this is achieved by conjugation with a fractional linear transformation, which does not effect holonomies or periods), and will specialize to the case $\ell = 0$ for the entire remaining argument and hence prove Theorem 2.7 for $\ell = 0$. When $J$ is not contained in any circle, we deal with an arbitrary integer $\ell$ as required in Theorem 2.7.

Recall the length $n_1+1$ subcylinders $X_1, \ldots, X_{k_0}$ of $P_1$. For each $k$, consider a cover of $U_0 \cap X_k$ by finitely many balls $B_{5\epsilon}(x_r^k), j = 1, \ldots, r_0 = r_0(k)$ with $x_r^k \in U_0 \cap X_k$ and $B_{10\epsilon}(x_r^k)$ pairwise disjoint; this is provided by a Vitali covering argument. For each $x_r^k$, we consider the gradient vector at $x_r^k$:

$$w_r^k = b\nabla \tau(x_r^k) + \ell\nabla \theta(x_r^k)$$

and its normalization

$$\hat{w}_r^k = \frac{w_r^k}{|w_r^k|}.$$  

Note that the NLI condition implies that

$$|w_r^k| > \frac{\delta_2(|b| + |\ell|)}{2}. \quad (5.7)$$

Applying Theorem 4.3 (NCP), we choose, for each $x_r^k$, a partner point $y_r^k \in B_{5\epsilon}(x_r^k) \cap X_k$ with

$$|\langle y_r^k - x_r^k, \hat{w}_r^k \rangle| > 5\delta_1\epsilon. \quad (5.8)$$

For each point $x \in \mathbb{C}$ and each $\epsilon > 0$, we choose a smooth bump function $\psi_{x,\epsilon}$ taking the value zero on the exterior of the $B_{\epsilon}(x)$ and the value one on $B_{\epsilon/2}(x)$. We may assume that

$$||\psi_{x,\epsilon}||_{C^1} \leq \frac{4}{\epsilon}.$$

For each $j = 1, 2$, we will associate $v_j(x_r^k)$ with $(j, 1, r, k)$ and $v_j(y_r^k)$ with $(j, 2, r, k)$, so that we parameterize the set

$$\{v_j(x_r^k), v_j(y_r^k) : 1 \leq j \leq 2, 1 \leq r \leq r_0, 1 \leq k \leq k_0\}$$

by $\{1, 2\} \times \{1, 2\} \times \{1, \ldots, r_0\} \times \{1, \ldots, k_0\}$. For a subset $\Lambda \subset \{1, 2\} \times \{1, 2\} \times \{1, \ldots, r_0\} \times \{1, \ldots, k_0\}$, we define the function $\beta_\Lambda$ on $U$ as

$$\begin{cases}
1 - \eta \left( \sum_{v_1(x_r^k) \in \Lambda} \psi_{x_r^k, 2\delta_\epsilon} \circ f_{N_0} + \sum_{v_1(y_r^k) \in \Lambda} \psi_{y_r^k, 2\delta_\epsilon} \circ f_{N_0} \right) & \text{on } v_1(U_1) \\
1 - \eta \left( \sum_{v_2(x_r^k) \in \Lambda} \psi_{x_r^k, 2\delta_\epsilon} \circ f_{N_0} + \sum_{v_2(y_r^k) \in \Lambda} \psi_{y_r^k, 2\delta_\epsilon} \circ f_{N_0} \right) & \text{on } v_2(U_1) \\
1 & \text{elsewhere.}
\end{cases}$$

**Definition 5.4.** We will say that $\Lambda \subset \{1, 2\} \times \{1, 2\} \times \{1, \ldots, r_0\} \times \{1, \ldots, k_0\}$ is *full* if for every $1 \leq r \leq r_0$ and $1 \leq k \leq k_0$, there is $j \in \{1, 2\}$ such that $v_j(x_r^k)$ or $v_j(y_r^k)$ belongs to $\Lambda$. We write $\mathcal{F}$ for the collection of all full subsets.
Fullness implies that the set of $x^k_r$'s and $y^k_r$'s indicated by $\Lambda$ forms a 100$\tilde{\epsilon}$ net for $X_k$.

**Definition 5.5.** Set $N := N_0 + n_1$. For each $\Lambda \in \mathcal{F}$, we define the Dolgopyat operator $\mathcal{M}_{\Lambda,a}$ on $C^1(U)$ by

$$\mathcal{M}_{\Lambda,a} h := \hat{L}^N_{\alpha,0}(h \beta_{\Lambda}).$$

**5.4. Properties of the Dolgopyat operators.** Our next task is to establish two key properties of the Dolgopyat operators.

**Theorem 5.6.** Fix $\Lambda \in \mathcal{F}$. If $H \in \mathcal{K}_{E([|b|+|\ell|]}(U)$, then

1. $\mathcal{M}_{\Lambda,a} H \in \mathcal{K}_{E([|b|+|\ell|]}(U)$ and
2. $||\mathcal{M}_{\Lambda,a} H||_{L^2(\nu)} \leq (1 - \epsilon_2)||H||_{L^2(\nu)}$.

**Proof.** Suppose that $H \in \mathcal{K}_{E([|b|+|\ell|]}(U)$. Then direct calculation yields that

$$\beta_{\Lambda} H \in K_{\frac{2n_{k_0}N_0}{\epsilon_0 \delta_3} + E}([|b|+|\ell|]}(U).$$

By our choice of $\eta$, we therefore have $\beta_{\Lambda} H \in K_2(E+1)([|b|+|\ell|]}(U)$. In conjunction with Lemma 5.1 and the definition of the Dolgopyat operator, this yields that for all $u \in U$,

$$|\nabla \mathcal{M}_{\Lambda,a} H(u)| \leq A_0 \left(1 + \frac{2(E+1)(|b|+|\ell|]}{K^{N}}\right)|\mathcal{M}_{\Lambda,a} H(u)|.$$

Our choices of $N$ and $E$ now yield

$$\mathcal{M}_{\Lambda,a} H \in \mathcal{K}_{E([|b|+|\ell|]}(U)$$

as desired in (1).

We now address part (2). We work first with the case that $a = \delta$, and then use continuity in $a$ to conclude the result; details are below.

A direct calculation in Cauchy Schwartz gives

$$(\mathcal{M}_{\Lambda,\delta} H)^2 = (\hat{L}^N_{\delta}(H \beta_{\Lambda}))^2 \leq (\hat{L}^N_{\delta} H^2)(\hat{L}^N_{\delta} \beta_{\Lambda}^2).$$

We are therefore interested in finding a large subset of $U$ where $\hat{L}^N_{\delta}(\beta_{\Lambda}^2)$ is strictly less than 1. For each $k$, consider now the subset $\tilde{S}_k$ which is the union of

$$\{x^k_r \in X_k : v_j(x^k_r) \in \Lambda \text{ for some } j = 1, 2\}$$

and

$$\{y^k_r \in X_k : v_j(y^k_r) \in \Lambda \text{ for some } j = 1, 2\}.$$ 

Recall that $f^{n_1} X_k = P_k$. Write $S_k = f^{n_1} \tilde{S}_k \subset P_k$. Consider the following neighbourhood of $S_k$:

$$\tilde{S}_k := \cup_{z \in S_k} B_{\epsilon_0 \delta_3 N}(z) \subset U_k.$$ 

We bound $\hat{L}^N_{\delta}(\beta_{\Lambda}^2)$ away from one on $\tilde{S}_k$: if $y \in \tilde{S}_k$, then for some “section” $\hat{v}$ of $f^N$ defined on $P_k$ we have

$$\beta_{\Lambda}^2(\hat{v}(y)) \leq \beta_{\Lambda}(\hat{v}(y)) < 1 - \eta$$

PRIME NUMBER THEOREMS FOR HYPERBOLIC RATIONAL FUNCTIONS 22
by the definition of $\beta_{\Lambda}$.

By (5.13), the normalization $\hat{L}_\delta^N(1) = 1$, and the choice of $A_0$, we compute for any $y \in \tilde{S}_k$,

$$\hat{L}_\delta^N(\beta_{\Lambda}^2)(y) = \sum_{\text{sections } v \text{ of } f^N} \frac{e^{-\delta \tau_N(v(y))} \beta_{\Lambda}^2(v(y))}{h_\delta(y)}$$

$$\leq \left( \sum_{\text{sections } v \text{ of } f^N} \frac{e^{-\delta \tau_N(v(y))} h_\delta(v(y))}{h_\delta(y)} \right) - \eta \frac{e^{-\delta \tau_N(\hat{v}(y))} h_\delta(\hat{v}(y))}{h_\delta(y)}$$

$$\leq 1 - \eta \frac{e^{-\delta \tau_N(\hat{v}(y))} h_\delta(\hat{v}(y))}{h_\delta(y)}$$

$$= 1 - \eta e^{-2N||\tau||_\infty - 2|| \log h_\delta ||_\infty}$$

(5.14) $\leq 1 - \eta e^{-NA_0}$.

Note that

(5.15) $P_k \subset \cup_{z \in \tilde{S}_k} B_{100\tilde{\epsilon}k_1/\epsilon_0}(z)$;

indeed, if $x \in P_k$, then $x = f^{n_1}y$ for some $y \in X_k$. We can then choose either an $x^k$ or a $y^k$ in $\tilde{S}_k$ such that one of $|x^k - y|$ and $|y^k - y|$ is less than $100\tilde{\epsilon}$. We then have $|f^{n_1}(x^k) - x| < 100k_1\tilde{\epsilon}/\epsilon_0$ and $|f^{n_1}(y^k) - x| < 100k_1\tilde{\epsilon}/\epsilon_0$ respectively, as required.

For ease of notation, we will write

$$\tilde{H} := \hat{L}_\delta^N H^2$$

for the rest of this proof. We now use regularity of $\tilde{H}$ together with the doubling properties of the measure $\nu_j$ (Theorem 4.5) to bound the integral $\int_{\tilde{S}_k} \tilde{H} \nu$ from below by a constant multiple of $\int_{\tilde{S}_k} \tilde{H} \nu$. By the Lasota-Yorke Lemma 5.1, we have $\tilde{H} \in K_{\tilde{H}([k]+|\tilde{e}|)}(U)$ and hence $\log \tilde{H}$ is Lipschitz. So by the small size of $\epsilon_1$, we have that

(5.16) $\sup\{\tilde{H}(x) : x \in B_{100k_1\tilde{\epsilon}/\epsilon_0}(z)\} \leq 2\inf\{\tilde{H}(x) : x \in B_{100k_1\tilde{\epsilon}/\epsilon_0}(z)\}$.

Setting $C_1 = \exp(\log C_3 \log_2 \frac{200k_1\tilde{\epsilon}}{c_0^3k_1\tilde{\epsilon}^3})$, we deduce

(5.17) $\int_{P_k} \tilde{H} \nu_k \leq 2C_1 \int_{\tilde{S}_k} \tilde{H} \nu$. 
as follows:
\[
\int_{P_k} \tilde{H} \, d\nu_k \leq \sum_{z \in S_k} \int_{B_{100\tilde{\epsilon}_1 n_1/c_0}(z)} \tilde{H} \, d\nu_k
\]
\[
\leq \sum_{z \in S_k} \sup \{ \tilde{H}(x) : x \in B_{100\tilde{\epsilon}_1 n_1/c_0}(z) \} \nu_k(B_{100\tilde{\epsilon}_1 n_1/c_0}(z))
\]
\[
\leq 2 \sum_{z \in S_k} \inf \{ \tilde{H}(x) : x \in B_{100\tilde{\epsilon}_1 n_1/c_0}(z) \} \nu_k(B_{c_1 \tilde{\epsilon}_3 n_1 c_0}(z)) \text{ by (5.16)}
\]
\[
\leq 2C_1 \sum_{z \in S_k} \inf \{ \tilde{H}(x) : x \in B_{100\tilde{\epsilon}_1 n_1/c_0}(z) \} \nu_k(B_{c_1 \tilde{\epsilon}_3 n_1 c_0}(z)) \text{ by choice of } C_3
\]
\[
\leq 2C_1 \int_{S_k} \tilde{H} \, d\nu_k.
\]

Putting all these together we have
\[
\int_{P_k} \mathcal{L}_\delta^N(H^2) \, d\nu_k - \int_{P_k} (\mathcal{M}_\Lambda,\delta H)^2 \, d\nu_k
\]
\[
\geq \eta e^{-NA_0} \int_{S_k} (\hat{\mathcal{L}}_\delta^N H^2) \, d\nu_k \text{ by (5.12)}
\]
\[
\geq \eta e^{-NA_0} \int_{S_k} (\hat{\mathcal{L}}_\delta^N H^2) \, d\nu_k \text{ as } 0 \leq \beta \leq 1
\]
\[
\geq \frac{\eta e^{-NA_0}}{2} \int_{S_k} (\hat{\mathcal{L}}_\delta^N H^2) \, d\nu_k \text{ by (5.14)}
\]
\[
\geq \frac{\eta e^{-NA_0}}{8C_1} \int_{P_k} (\hat{\mathcal{L}}_\delta^N H^2) \, d\nu_k \text{ by (5.17)}.
\]

Summing over all \(k\) and using that \(\hat{\mathcal{L}}_\delta\) preserves \(\nu\), we have
\[
||H||_{L^2(\nu)}^2 - ||\mathcal{M}_\Lambda,\delta H||_{L^2(\nu)}^2 = \int_X \mathcal{L}_\delta^N(H^2) \, d\nu - \int_X (\mathcal{M}_\Lambda,\delta H)^2 \, d\nu
\]
\[
\geq \frac{\eta e^{-NA_0}}{2} \int_X \hat{\mathcal{L}}_\delta^N(H^2) \, d\nu
\]
\[
= \frac{\eta e^{-NA_0}}{8C_1} \int_X H^2 \, d\nu.
\]

We now have, for \(|a - \delta| < a_0\), that
\[
||\mathcal{M}_{\Lambda,a} H||_{L^2(\nu)}^2 \leq e^{2N||\tau||_{\infty,a_0}} ||\mathcal{M}_{\Lambda,\delta} H||_{L^2(\nu)}^2
\]
\[
\leq e^{2N||\tau||_{\infty,a_0}} \left(1 - \frac{\eta e^{-NA_0}}{8C_1} \right) ||H||_{L^2(\nu)}^2
\]
\[
\leq \left(1 - \frac{\eta^2 e^{-2NA_0}}{64C_1^2} \right) ||H||_{L^2(\nu)}^2
\]
so we are done by choice of \(\epsilon_2\). \(\square\)
5.5. The iterative argument. Our final technical challenge is to relate the Dolgopyat operator to the twisted transfer operators in preparation for an iterative argument.

**Theorem 5.7.** There exists \( a_0 > 0 \) such that for all \(| \delta - a | < a_0 \), we have the following: for every \( h \in C^1(U) \) and \( H \in K_E(|\ell| + |b|)(U) \) satisfying
\[
|h| \leq H \quad \text{and} \quad |\nabla h| \leq E(|\ell| + |b|)H \quad \text{pointwise on } U,
\]
there is a choice of \( \Lambda \in \mathcal{F} \) such that
\[
|\hat{\mathcal{L}}^N_{s,e} h| \leq M_{\Lambda,a} H
\]
and
\[
|\nabla (\hat{\mathcal{L}}^N_{s,e} h)| \leq E(|\ell| + |b|)M_{\Lambda,a} H
\]
both pointwise on \( U \) (here \( s = a + ib \)).

The second part, (5.19) is a direct calculation using the Lasota-Yorke bounds (Lemma 5.1). We now address the first part. The actual procedure is to make a clever choice of \( \Lambda \) such that
\[
|\hat{\mathcal{L}}^N_{a+ib,e} h| \leq \hat{\mathcal{L}}^N_a (\beta \Lambda H)
\]
pointwise. Once we do this, it will be easy to deduce (5.18) by applying \( \hat{\mathcal{L}}^N_a \) to both sides, recalling \( N = N_0 + n_1 \). The next Lemma will guide us in that choice. Recall the sections \( v_1, v_2 \) of \( f^{N_0} \) chosen in (5.2).

**Lemma 5.8.** Define \( \Delta_1 \) and \( \Delta_2 \) on \( U_0 \) as follows:
\[
\Delta_1(x) = \frac{\sum_{k=1,2} e^{-(\delta + \epsilon)\tau_{N_0}(v_k(x)) + i\theta N_0(v_k(x))} h(v_k(x)) h_\delta(v_k(x))}{(1 - k_3\delta)e^{-(\delta + \epsilon)\tau_{N_0}(v_1(x))} h(v_1(x)) h_\delta(v_1(x)) + e^{-(\delta + \epsilon)\tau_{N_0}(v_2(x))} h(v_2(x)) h_\delta(v_2(x))}
\]
\[
\Delta_2(x) = \frac{\sum_{k=1,2} e^{-(\delta + \epsilon)\tau_{N_0}(v_k(x)) + i\theta N_0(v_k(x))} h(v_k(x)) h_\delta(v_k(x))}{e^{-(\delta + \epsilon)\tau_{N_0}(v_1(x))} h(v_1(x)) h_\delta(v_1(x)) + (1 - k_3\delta)e^{-(\delta + \epsilon)\tau_{N_0}(v_2(x))} h(v_2(x)) h_\delta(v_2(x))}
\]
For each \( (r, k) \in \{1, \ldots, r_0\} \times \{1, \ldots, k_0\} \), at least one of \( \Delta_1, \Delta_2 \) is less than or equal to one on at least one of the discs \( B_{2\delta\delta_3}(x^{k_0}_r) \) or \( B_{2\delta\delta_3}(y^{k_0}_r) \).

**Proof.** Fix \( (r, k) \in \{1, \ldots, r_0\} \times \{1, \ldots, k_0\} \), and consider the alternative described in Lemma 5.3 for the sections \( v_1, v_2 \). If the first alternative holds on \( B_{10\delta}(x^{k_0}_r) \) for either \( v_1 \) or \( v_2 \), then we are finished, so we shall assume the converse. For \( x \in B_{10\delta}(x^{k_0}_r) \), we write \( \text{Arg}(x) \) for the argument (taking values in \( [-\pi, \pi] \) of the ratio of the summands in the numerator of \( \Delta_2 \):
\[
\text{Arg}(x) := \text{arg} \left( \frac{e^{-ib\tau_{N_0}(v_1(x)) + i\theta N_0(v_1(x))} h(v_1(x))}{e^{-ib\tau_{N_0}(v_2(x)) + i\theta N_0(v_2(x))} h(v_2(x))} \right)
\]
with \( s = a + ib \). Our aim is to exclude the possibility that \( \text{Arg}(x) \) is small both near \( x^{k_0}_r \) and near \( y^{k_0}_r \); that will imply some cancellation in the numerator and give the lemma.

We will do this simply by estimating derivatives from below.
Remark A priori one might worry that the arg(x) function has discontinuities where it jumps from −π to π. However, a straightforward calculation using the Lipschitz bound (5.27) and the small diameter of $B_{10κ}(x_0^k)$ establishes that the required bound (5.28) follows immediately if $|\text{Arg}(x)| > \frac{π}{2}$ anywhere on $B_{10κ}(x_0^k)$.

We therefore assume that $|\text{Arg}(x)| \leq \frac{π}{2}$ on that ball, and that the arg function is correspondingly continuous. Recall the functions $\tilde{τ}, \tilde{θ}$ from Lemma 3.2 and calculate

\begin{equation}
\text{Arg}(x) = b\tilde{τ}(x) + ℓ\tilde{θ}(x) + \arg \left( \frac{h(v_1(x))}{h(v_2(x))} \right)
\end{equation}

\begin{equation}
\text{Arg}(x) = b\tilde{τ}(x) + ℓ\tilde{θ}(x) + \arg(h(v_1(x))) - \arg(h(v_2(x)))
\end{equation}

where we think of these numbers as elements of $S^1 = \mathbb{R}/2π\mathbb{Z}$ where necessary. Direct calculation yields

\begin{equation}
|\nabla \text{arg}(h(v_j(x)))| \leq \frac{4E(|b| + |ℓ|)}{c_0κN_0}
\end{equation}

for $j = 1, 2$.

Applying Taylor’s expansion to the map $y \mapsto b\tilde{τ}(y) + ℓ\tilde{θ}(y)$ at $x_0^k$, and using the condition $||{(\tilde{τ}, \tilde{θ})}||_{C^2} \leq 1/(2δ_2)$ by the NLI condition, we get

\begin{align*}
|b\tilde{τ}(y_r^k) + ℓ\tilde{θ}(y_r^k) - b\tilde{τ}(x_r^k) - ℓ\tilde{θ}(x_r^k)|
\geq |\langle y_r^k - x_r^k, \hat{w}_r^k \rangle||w_r^k| - \frac{(|b| + |ℓ|)|x_r^k - y_r^k|^2}{δ_2}.
\end{align*}

Therefore

\begin{align*}
|\text{Arg}(y_r^k) - \text{Arg}(x_r^k)| &\geq |b\tilde{τ}(y_r^k) + ℓ\tilde{θ}(y_r^k) - b\tilde{τ}(x_r^k) - ℓ\tilde{θ}(x_r^k)| - |y_r^k - x_r^k| \left( \frac{8E(|b| + |ℓ|)}{c_0κN_0} \right) \\
&\geq |\langle y_r^k - x_r^k, \hat{w}_r^k \rangle||w_r^k| - \frac{(|b| + |ℓ|)|x_r^k - y_r^k|^2}{δ_2} - |y_r^k - x_r^k| \left( \frac{8E(|b| + |ℓ|)}{c_0κN_0} \right) \\
&\geq \frac{5δ_1δ_2\epsilon}{2} \left( |b| + |ℓ| \right) - \frac{25ε_1^2}{δ_2} - \frac{40Eε_1}{c_0κN_0} \\
&\geq 2δ_1δ_2\epsilon_1,
\end{align*}

where we used the NCP condition (5.8) for the third inequality.

The next step is to show that $|\text{Arg}(y) - \text{Arg}(x)| > δ_1δ_2\epsilon_1$ for every $x \in B_{2δ_1\epsilon}(x_r^k)$ and $y \in B_{2δ_1\epsilon}(y_r^k)$. First estimate the $C^1$ norm of Arg(x):

\begin{equation}
|\nabla \text{Arg}(x)| \leq A_0(|b| + |ℓ|) + \frac{8E(|b| + |ℓ|)}{c_0κN_0} \leq E(|b| + |ℓ|).
\end{equation}
This gives
\[ |\text{Arg}(x_k^r) - \text{Arg}(x)| \leq 2E\delta_3\delta(|b| + |\ell|) \leq \frac{\delta_1\delta_2\epsilon_1}{2}. \]

A similar calculation shows \( |\text{Arg}(y_k^r) - \text{Arg}(y)| \leq \delta_1\delta_2\epsilon_1/2. \) It follows that for every \( x \in B_{2\delta_3\delta}(x_k^r) \) and \( y \in B_{2\delta_3\delta}(y_k^r) \), we have
\[
|\text{Arg}(y) - \text{Arg}(x)| \\
> |\text{Arg}(y_k^r) - \text{Arg}(x_k^r)| - |\text{Arg}(y) - \text{Arg}(y_k^r)| - |\text{Arg}(x) - \text{Arg}(x_k^r)| \\
\geq \delta_1\delta_2\epsilon_1
\]
as expected. We now claim that one of the following holds:

1. \( \text{Arg}(x) > \frac{\delta_1\delta_2\epsilon_1}{4} \) on \( B_{2\delta_3\delta}(x_k^r) \) or
2. \( \text{Arg}(y) > \frac{\delta_1\delta_2\epsilon_1}{4} \) on \( B_{2\delta_3\delta}(y_k^r) \).

To see this, suppose that the first statement fails; but then we may choose \( x \in B_{2\delta_3\delta}(x_k^r) \) with \( |\text{Arg}(x)| \leq \frac{\delta_1\delta_2\epsilon_1}{4} \). But then
\[ (5.28) \quad |\text{Arg}(y)| \geq |\text{Arg}(x) - \text{Arg}(y)| - |\text{Arg}(x)| \geq \frac{\delta_1\delta_2\epsilon_1}{2} \]
for all \( y \in B_{2\delta_3\delta}(y_k^r) \).

This claim implies Lemma 5.8 by means of Lemma 5.2. Suppose, for example, that the argument condition holds on \( B_{2\delta_3\delta}(y_k^r) \), that \( x \in B_{2\delta_3\delta}(y_k^r) \) has
\[ |e^{-(\delta+a)\tau_N(o_1(x))}H(v_2(x))h_\delta(v_2(x))| \leq |e^{-(\delta+a)\tau_N(o_1(x))}h(v_1(x))h_\delta(v_1(x))|, \]
and that
\[ e^{-(\delta+a)\tau_N(o_1(x))}H(v_1(y_k^r))h_\delta(v_1(y_k^r)) \leq e^{-(\delta+a)\tau_N(o_2(x))}H(v_2(y_k^r))h_\delta(v_2(y_k^r)). \]

Then Lemma 5.2 gives
\[
\left| \sum_{k=1,2} e^{-(\delta+a)\tau_N(o_1(x)) + id\theta N_0(v_k(x))}h(v_k(x))h_\delta(v_k(x)) \right| \\
\leq e^{-(\delta+a)\tau_N(o_1(x))}h(v_1(x))|h_\delta(v_1(x)) + \eta' e^{-(\delta+a)\tau_N(o_2(x))}h(v_2(x))|h_\delta(v_2(x)) \\
\leq e^{-(\delta+a)\tau_N(o_1(x))}H(v_1(x))h_\delta(v_1(x)) + \eta' e^{-(\delta+a)\tau_N(o_2(x))}H(v_2(x))h_\delta(v_2(x))
\]
where \( \eta' := (1 - 4k_0\eta) \). Now we notice that the logarithmic derivative of
\[ e^{-(\delta+a)\tau_N(o_1(x))}H(v_j(x))h_\delta(v_j(x)) \]
is bounded by \( E(|b| + |\ell|) \), so the small choice of \( \epsilon_1 \) yields
\[
4k_0\eta e^{-(\delta+a)\tau_N(o_2(x))}H(v_2(x))h_\delta(v_2(x)) \\
\geq 2k_0\eta e^{-(\delta+a)\tau_N(o_2(y_k^r))}H(v_2(y_k^r))h_\delta(v_2(y_k^r)) \\
\geq 2k_0\eta e^{-(\delta+a)\tau_N(o_1(y_k^r))}H(v_1(y_k^r))h_\delta(v_1(y_k^r)) \\
\geq k_0\eta e^{-(\delta+a)\tau_N(o_1(x))}H(v_1(x))h_\delta(v_1(x)).
\]
Hence

\[(5.29) \quad \sum_{k=1,2} e^{-(\delta+s)\tau N_0(v_k(x)) + i\theta N_0(v_k(x))} h(v_k(x)) \leq (1 - k_0\eta) e^{-(\delta+a)\tau N_0(v_1(x))} H(v_1(x)) + e^{-(\delta+a)\tau N_0(v_2(x))} H(v_2(x))\]
on B_{2\delta_5}(y^k_r). The other cases follow similarly.

**Proof of Theorem 5.7.** We need to construct an appropriate set \( \Lambda \in \mathcal{F} \). For each \( (r, k) \in \{1, \ldots, r_0\} \times \{1, \ldots, k_0\} \) proceed as follows:

1. If \( \Delta_1 \leq 1 \) on \( B_{2\delta_5}(x^k_r) \), then include \( v_1(x^k_r) \) in \( \Lambda \), otherwise
2. If \( \Delta_2 \leq 1 \) on \( B_{2\delta_5}(x^k_r) \), then include \( v_2(x^k_r) \) in \( \Lambda \), otherwise
3. If \( \Delta_1 \leq 1 \) on \( B_{2\delta_5}(y^k_r) \), then include \( v_1(y^k_r) \) in \( \Lambda \), otherwise
4. If \( \Delta_2 \leq 1 \) on \( B_{2\delta_5}(y^k_r) \), then include \( v_2(y^k_r) \) in \( \Lambda \).

By Lemma 5.8, at least one of these situations occurs for each \( (r, k) \); we therefore obtain a full set \( \Lambda \in \mathcal{F} \). The inequality (5.20) then follows from the inequalities (5.21) or (5.22) as applicable. In order to deduce (5.18) from (5.20), it suffices to observe that

\[ |\hat{\mathcal{L}}^N_{s,\ell}h| \leq \hat{\mathcal{L}}^N_a \left( |\hat{\mathcal{L}}^{N_0}_a h| \right) \leq \hat{\mathcal{L}}^N_a \left( \hat{\mathcal{L}}^{N_0}_a (\beta_\Lambda H) \right) = \hat{\mathcal{L}}^N_a (\beta_\Lambda H).\]

**Proof of Theorem 2.7.** The deduction of Theorem 2.7 from Theorems 5.6 and 5.7 is now standard as in [9, Section 5]. For completeness we recall the argument. Fix \( a, b \in \mathbb{R} \) and \( \ell \in \mathbb{Z} \) with \( |a - \delta| < a_0 \) as in Theorem 5.7 and \( |b| + |\ell| > 1 \). Let \( h \in C^1(U) \).

Using Theorems 5.6 and 5.7 we inductively choose functions

\[ H_k \in K_{E(|b| + |\ell|)}(U) \subset C^1(U, \mathbb{R})\]

with the properties:

1. \( H_0 \) is the constant function \(|h||_{(|b| + |\ell|)}\),
2. \( H_{k+1} = M_{\Lambda_k, a} H_k \) for some \( \Lambda_k \in \mathcal{F} \),
3. \( |\hat{\mathcal{L}}^{kN}_s h| \leq H_k \) pointwise, and
4. \( |\nabla \hat{\mathcal{L}}^{kN}_s h| \leq E(|b| + |\ell|) H_k \) pointwise.

From this it follows by Theorem 5.6 that

\[ ||\hat{\mathcal{L}}^{kN}_s h||_{L^2(\nu)} \leq (1 - \epsilon_2)^k ||h||_{(|b| + |\ell|)}\]

for any \( h \in C^1(U) \) and any \( k \geq 0 \). For general \( n = kN + r \) with \( 0 \leq r \leq N - 1 \), we then have

\[(5.30) \quad ||\hat{\mathcal{L}}^N_{s,\ell} h||_{L^2(\nu)} = ||\hat{\mathcal{L}}^{kN} \hat{\mathcal{L}}^N_{s,\ell} h||_{L^2(\nu)} \leq (1 - \epsilon_2)^k ||\hat{\mathcal{L}}^r_{s,\ell} h||_{(|b| + |\ell|)}\]

\[ \leq (1 - \epsilon)^n \frac{(||\hat{\mathcal{L}}_{s,\ell}||_{(|b| + |\ell|)} + 1)^N}{(1 - \epsilon)^N} ||h||_{(|b| + |\ell|)} \]

\[ \leq (1 - \epsilon)^n \frac{(||\hat{\mathcal{L}}_{s,\ell}||_{(|b| + |\ell|)} + 1)^N}{(1 - \epsilon)^N} ||h||_{(|b| + |\ell|)} \]
where we have chosen $\epsilon > 0$ such that $(1 - \epsilon)^N \geq (1 - \epsilon_2)$. Since there is a uniform bound on $|\tilde{L}^r_{s,\ell}||_{((b + |\ell|))}$, $0 \leq r \leq N - 1$, independent of $b$ and $\ell$ we are now finished.

\section{L-functions, Transfer Operators and Counting Estimates for Hyperbolic Polynomials}

In this section we describe the relationship between transfer operators and zeta functions. This relationship, together with the bounds we have established, will allow us to deduce our main equidistribution theorem. The approach described in this section is well established; the interested reader can find a clear account in [8]. We include an outline of the argument for the convenience of readers.

We recall the shift space $\Sigma^+$ from Section 2 together with its shift operator $\sigma$. We have a map $\pi: \Sigma^+ \to J$ given by $x = (i_0, i_1, i_2, \cdots) \mapsto \cap_{j=0}^\infty f^{-j}(P_{i_j})$; it is a semiconjugacy between $\sigma$ and $f$. Abusing notation we think of $\tau, \tau_n, \alpha, \alpha_n$ as functions on $\Sigma^+$ by pulling them back via $\pi$.

Define the symbolic zeta function, also called the Ruelle zeta function for $\Sigma^+$, by

$$\tilde{\zeta}(s) = \exp \left( \sum_{n=1}^\infty \frac{1}{n} \sum_{\sigma^n x = x} e^{-st_n(x)} \right) = \prod_{\hat{x} \in \hat{\mathcal{P}}} (1 - |\lambda(\hat{x})|^s)^{-1}$$

where $\hat{\mathcal{P}}$ is the collection of primitive periodic orbits of $\sigma$. This is convergent and analytic for $\Re(s) > \delta$.

For each $\ell \in \mathbb{Z}$, let $\chi_\ell(x) = x^\ell$ be a unitary character $S^1 \to S^1$, and define

$$\tilde{\zeta}(s, \ell) = \exp \left( \sum_{n=1}^\infty \frac{1}{n} \sum_{\sigma^n x = x} \chi_\ell(\alpha_n(x))e^{-st_n(x)} \right).$$

By comparison with $\tilde{\zeta}(s)$, $\tilde{\zeta}(s, \ell)$ converges for $\Re(s) > \delta$.

Let

$$\tilde{Z}_n(s, \ell) := \sum_{\sigma^n x = x} \chi_\ell(\alpha_n(x))e^{-st_n(x)}$$

so that

$$\tilde{\zeta}(s, \ell) = \exp \left( \sum_{n=1}^\infty \frac{1}{n} \tilde{Z}_n(s, \ell) \right).$$

These functions $\tilde{Z}_n$ are related to our transfer operators by the following proposition, which is originally due to Ruelle [12] (see also [8], [9, Appendix]).

For each $1 \leq j \leq k_0$, let $\phi_j \in C^1(U)$ be the characteristic function of $U_j$, recalling that $U$ is the disjoint union of $U_j$. 
**Proposition 6.1.** Fix $a_0 > 0, b_0 > 0$. There exists $x_j \in P_j$, $j = 1, \cdots, k_0$, such that for any $\epsilon > 0$, there exists $C_\epsilon > 0$ such that for all $n \geq 2$ and for any $\ell \in \mathbb{Z}$,

\[
\left| \tilde{Z}_n(s, \ell) - \sum_{j=1}^{k_0} L_{s, \ell}^n(\phi_j)(x_j) \right| \leq C_\epsilon (|\Im(s)| + |\ell|) \sum_{m=2}^{n} \|L_{s, \ell}^{n-m}\|_{C^1} \left( \kappa^{-1-\epsilon + P(-\Re(s)\tau)} \right)^m
\]

for all $|\Im(s)| + |\ell| > b_0$ and $|\Re(s) - \delta| \leq a_0$ (here $\kappa > 1$ is the expansion rate of $f$).

In the rest of this section, suppose that $f$ is not conjugate to a monomial $x^{\pm d}$ for $d \in \mathbb{N}$. If the Julia set of $f$ is contained in a circle, we only consider $\ell = 0$; and otherwise, $\ell$ is any integer.

Fix $\epsilon > 0$. Let $\epsilon_0 = \epsilon_0(\epsilon) > 0$ be as given by Theorem 2.5. Then by Proposition 6.1 and Theorem 2.5, we have, for some $C_\epsilon > 1$,

\[
|\tilde{Z}_n(s, \ell)| \leq C_\epsilon (|\Im(s)| + |\ell|)^{2+\epsilon} \rho_\epsilon^n
\]

for all $|\Im(s)| + |\ell| > 1$ and $|\Re(s) - \delta| < \epsilon_0$; here we adjust $\rho_\epsilon$ and $\epsilon_0$ slightly if needed.

Since

\[
\log \tilde{\zeta}(s, \ell) = \sum_{n=1}^{\infty} \frac{1}{n} \tilde{Z}_n(s, \ell),
\]

we deduce that for all $|\Im(s)| + |\ell| > 1$ and $|\Re(s) - \delta| < \epsilon_0$,

\[(6.1) \quad |\log \tilde{\zeta}(s, \ell)| \leq C_\epsilon (|\Im(s)| + |\ell|)^{2+\epsilon}
\]

for some $C_\epsilon > 1$.

The following follows from the estimates (6.1) for $\ell = 0$ (cf. [9]):

**Corollary 6.2.** The zeta function $\tilde{\zeta}(s)$ is non-vanishing and analytic on $\Re(s) \geq \delta$ except for the simple pole at $s = \delta$.

**Proof.** The estimates (6.1), together with the RPF theorem, implies that the non-lattice property holds for $\tau$, in the sense that there exist no function $L : J \rightarrow m\mathbb{Z}$ for some $m \in \mathbb{R}$ and a Lipschitz function $u : J \rightarrow \mathbb{R}$ such that

\[(6.2) \quad \tau = L + u - u \circ f.
\]

Indeed, if $\tau$ satisfies (6.2), than $\tilde{\zeta}$ has poles at $\delta + 2p\pi i/m$ for all $p \in \mathbb{Z}$, which contradicts (6.1) (see [9] for details). The claim then follows from [7].

In view of this corollary, by making $\epsilon_0$ smaller if necessary, we deduce from (6.1):

**Proposition 6.3.** (1) For $\ell = 0$, $\tilde{\zeta}(s) = \tilde{\zeta}(s, 0)$ is analytic and non-vanishing on $\Re(s) \geq \delta - \epsilon_0$ except for the simple pole at $s = \delta$. 
(2) For $\ell \neq 0$, $\tilde{\zeta}(s, \ell)$ is analytic and non-vanishing on $\Re(s) \geq \delta - \epsilon_0$.

(3) For any $\epsilon > 0$, there exists $C_\epsilon > 1$ such that for any $\ell \in \mathbb{Z}$, we have

$$\log |\tilde{\zeta}(s, \ell)| \leq C_\epsilon \cdot (|\ell| + 1)^{2+\epsilon} \cdot |\Im(s)|^{2+\epsilon}$$

for all $\Re(s) > \delta - \epsilon_0$ and $|\Im(s)| \geq 1$.

Due to a result of Manning [14], these analytic properties of $\tilde{\zeta}(s, \ell)$ can be transferred to those of $\zeta(s, \ell)$:

$$\zeta(s, \ell) = \exp \left( \sum_{n=1}^{\infty} \frac{1}{n} \sum_{f^n x = x} \chi_\ell(\alpha_n(x)) \cdot e^{-s \tau_n(x)} \right).$$

**Theorem 6.4** (Manning [14]). There exists $\epsilon_1 > 0$ such that for any $\ell \in \mathbb{Z}$, the ratio

$$\frac{\tilde{\zeta}(s, \ell)}{\zeta(s, \ell)}$$

is holomorphic, bounded, and non-vanishing on $\Re(s) > \delta - \epsilon_1$.

Therefore Proposition 6.3 holds for $\zeta(s, \ell)$ as well as for $\tilde{\zeta}(s, \ell)$, which finishes the proof of Theorem 1.2.

We can convert this into a bound on the logarithmic derivative of $\zeta(s, \ell)$. By an extra application of Phragmen-Lindelof theorem (as in [8]), for any $\epsilon > 0$, there exist $C_\epsilon > 1$, $0 < \epsilon_1 < \epsilon_0$, and $0 < \beta < 1$ such that for all $\Re(s) > \delta - \epsilon_1$ and $|\Im(s)| \geq 1$,

$$|\zeta(s, \ell)'| \leq C_\epsilon \cdot (|\ell| + 1)^{2+\epsilon} \cdot |\Im(s)|^\beta.$$

Define the counting function:

$$\pi_\ell(t) := \sum_{\hat{x} \in \mathcal{P}_t} \chi_\ell(\lambda_\theta(\hat{x}))$$

where $\mathcal{P}_t$ is the set of all primitive periodic orbits of $f$ with $|\lambda(\hat{x})| < t$ in the Julia set $J$.

Now using Proposition 6.3 and (6.3), the arguments of [8] deduce the following from (6.3):

**Proposition 6.5.** There exists $\eta > 0$ such that

1. $\pi_0(t) = \text{Li}(t^\delta) + O(t^{(\delta-\eta)})$;

2. for any $\epsilon > 0$ and any $\ell \neq 0$, we have

$$\pi_\ell(t) = O((|\ell| + 1)^{2+\epsilon} t^{(\delta-\eta)})$$

where the implied constant depends only on $\epsilon$. 

Now let $\psi \in C^4(S^1)$. Then we can write $\psi = \sum a_\ell \chi_\ell$ by the Fourier expansion where $a_0 = \int \psi dm$, and $a_\ell = O(|\ell|^{-4})$. Then we deduce from Proposition 6.5 that

$$\sum_{\hat{x} \in \mathcal{P}_2} \psi(\lambda_\theta(\hat{x})) = \int \psi dm \cdot \text{Li}(t^\delta) + O(t^{(\delta - \eta)}).$$

Recalling that there are at most $2d - 2$ periodic orbits of $f$ which do not lie on $J$, Proposition 6.5 and (6.4) provide Theorem 1.1.

**Remark on equidistribution** We can also deduce from Proposition 6.3 the joint equidistribution of pairs

$$\{(\hat{x}, \lambda_\theta(\hat{x})) : \hat{x} \text{ is a primitive periodic orbit with } |\lambda(\hat{x})| < T\}$$

in $J \times S^1$ with respect to the product measure $\nu \times m$, using arguments of Bowen [1] and Parry-Pollicott [7]. It is likely that a combination of our approach with their techniques would provide an effective joint equidistribution theorem, but we shall not address that here.

**References**


Mathematics department, Yale university, New Haven, CT 06511 and Korea Institute for Advanced Study, Seoul, Korea

E-mail address: hee.oh@yale.edu

IAS, Princeton

E-mail address: dale.alan.winter@gmail.edu