Ergod. Th. & Dynam. Sys. (2017), **37**, 176–192 © *Cambridge University Press, 2015*
doi:10.1017/etds.2015.36

Complex dynamics with focus on the real part

JOHN ERIK FORNÆSS† and HAN PETERS‡

† *Department for Mathematical Sciences, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway*
(e-mail: john.fornass@math.ntnu.no)
‡ *KdV Institute for Mathematics, University of Amsterdam, Science Park 105–107, 1098 XG Amsterdam, The Netherlands*
(e-mail: h.peters@uva.nl)

(*Received* 29 *October* 2013 *and accepted in revised form* 5 *April* 2015)

Abstract. We consider the dynamics of holomorphic polynomials in $\mathbb{C}$. We show that the ergodic properties of the map can be seen already from the real parts of the orbits.

1. *Introduction*

Science is concerned with describing the world and mathematics is an important tool. One can use mathematics to obtain equations for how a system will change over time. To get manageable equations, the usual procedure is to suppress some of the parameters which play a role in the system. This simplifies the mathematical equations, but it is important to ask whether the obtained results still accurately describe the original situation. In this paper, we will investigate rigorously whether one can recover precise results when one suppresses some variables. We will do this in the case of one-dimensional complex dynamics, where the exact theory is highly developed.

Questions similar to those addressed above were studied by Takens in [**7**], where the following result was proved.

THEOREM 1.1. (Takens) *Let M be a compact manifold of dimension m. For pairs (ϕ, y), $\phi : M \to M$ a $\mathcal{C}^2$ diffeomorphism and $y : M \to \mathbb{R}$ a $\mathcal{C}^2$ function, it is a generic property that the map $\Phi_{(\phi,y)} : M \to \mathbb{R}^{2m+1}$, defined by*

$$\Phi_{(\phi,y)}(x) = (y(x), y(\phi(x)), \ldots, y(\phi^{2m}(x)))$$

is an embedding.

Hence all information about the original dynamical system can be retrieved from the suppressed dynamical system. When the map ϕ is not injective one should not expect the Takens' theorem to hold. For example, one could have distinct points $z, w \in X$ with $y(z) = y(w)$ and $\phi(z) = \phi(w)$. The points z, w will be identified once the other variables are suppressed. In fact, such identifications will always occur in our complex analytic setting.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Let $P : \mathbb{C} \to \mathbb{C}$ be a polynomial of degree $d \geq 2$. Let $z = x + iy$ denote coordinates in $\mathbb{C}$. We use the notation $z_n = x_n + iy_n$ to denote an orbit, $z_n = P^n(x_0 + iy_0)$. We will consider the real orbits $\{x_n\}_{n\geq 0}$.

We prove that it will suffice to consider only the first $N(P)$ terms.

LEMMA 1.2. *Let $P(z)$ be a complex polynomial of degree $d \geq 2$. Then there exists an integer $N = N(P)$ so that if $\{x_n + iy_n\}$ and $\{u_n + iv_n\}$ are two orbits and $x_n = u_n$ for $n \leq N$, then $x_n = u_n$ for all n.*

Let $\Phi : \mathbb{C} \to \mathbb{R}^{N+1}$ denote the map $\Phi(z_0) = (x_0, \ldots, x_N)$, and let $S = \Phi(\mathbb{C})$. For $z_0 \in \mathbb{C}$ we define the map $Q : S \to S$ by $Q \circ \Phi = \Phi \circ P$, which gives

$$Q(x_0, \ldots, x_N) = (x_1, \ldots, x_{N+1}).$$

Even though our map Φ will not be an embedding, it is possible that many properties of the dynamical system $(\mathbb{C}, P)$ can still be observed for the system (S, Q). We will focus on ergodic theoretic aspects of the dynamical systems. The following is a classical result by Brolin, Lyubich and Mañé, see [**2**, **4**, **5**].

THEOREM 1.3. (Brolin, Lyubich and Mañé) *There is a unique invariant, ergodic probability measure μ on $\mathbb{C}$ of maximal entropy,* $\log d$.

We define a probability measure ν on S by $\nu = \Phi_*(\mu)$. Our main result is the following.

THEOREM 1.4. *Let P be a non-exceptional complex polynomial of degree $d \geq 2$. Then the probability measure ν is invariant and ergodic. Moreover, it is the unique measure of maximal entropy,* $\log d$.

The plan of the paper is the following. In the next section we introduce the concept of non-exceptional polynomials. We also prove Lemma 1.2 and provide some basic estimates on orbits. In §3 we investigate mirrored orbits, i.e. points z_0, w_0 for which $x_n = u_n$ for all n. We prove the useful Lemma 3.9, which states that for a non-exceptional polynomial of degree d, a point can only be mirrored by $d^2 - 1$ points. A more detailed analysis gives that a non-exceptional polynomial can only have an open set of mirrored points if all coefficients are real (Theorem 3.8). In §4 we consider the entropy of the map Q and prove our main result Theorem 1.4.

2. *Preliminary results*

LEMMA 2.1. *Let $P(z)$ denote a complex polynomial of degree $d \geq 2$. Then there can be at most one vertical line which is mapped to a vertical line. For all other lines, the number of points in any given vertical line which is mapped to any other given vertical line is at most d.*

Proof. For notational simplicity we work with horizontal lines instead. If a horizontal line is mapped to a horizontal line, then after translations we can arrange that both lines are equal to the x axis. Let us write $P(z) = \sum_{n=0}^{d} a_n z^n$. The invariance of the real axis is equivalent to all a_n being real. We show that no other horizontal line is mapped to a horizontal line. We may assume that if there exists another horizontal line which is mapped

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

to a horizontal line, then it is the line given by $y = 1$. Let $f(x) = \sum a_n(x+i)^n$ as a function of $x \in \mathbb{R}$. Since the imaginary part is constant, the derivative must be real. Hence $\sum_{n=1}^{d} na_n(x+i)^{n-1}$ is real valued. But then $\sum_{n=1}^{d} na_n(x+i)^{n-1} - \sum_{n=1}^{d} na_n x^{n-1} \in \mathbb{R}$. By looking at terms of order $d-2$ in x we see that $i(d(d-1)a_d x^{d-2}) + \mathcal{O}(x^{d-3})$ is purely real. This is not possible for large x.

It remains to be shown that if the image A of a line is not included in another line B, then at most d points in A are mapped to B. We may assume both lines are the x axis, and the map has the form $P = \sum_{n=0}^{d} a_n z^n$. We write $a_n = b_n + ic_n$, so on the x axis the map looks like

$$P(x) = \sum b_n x^n + i \sum c_n x^n.$$

The imaginary part is a non-zero polynomial of degree at most d, so can have at most d real zeros (counted with multiplicity). □

For polynomials of degree three, all polynomials with real coefficients must map some vertical line to a vertical line. We note that this is not the case for any other degree greater or equal to two.

LEMMA 2.2. *Let P be a polynomial of degree three with real coefficients. Then it must send some vertical line to some vertical line.*

Proof. After real scaling and real translations in the domain and range, we can write $P(z) = z^3 + az$ for some real a. But then the imaginary axis is mapped to itself. □

We prove Lemma 1.2.

LEMMA 2.3. *Let P be a polynomial of degree $d \geq 2$. There exists an integer $N(P)$ so that, for any two real orbits $\{x_n\}_{n\geq 0}$ and $\{u_n\}_{n\geq 0}$ with $x_n = u_n$ for $n = 0, 1, \ldots, N(P)$, we have $x_n = u_n$ for all $n \geq 0$.*

Proof. Let $Q_n(x, y, u, v) = \mathrm{Re}(P^n(x+iy) - P^n(u+iv))$, and set

$$Z_n = \{Q_0 = \cdots = Q_n = 0\}.$$

The sequence (Z_n) is a decreasing sequence of algebraic sets. It follows from Hilbert's Nullstellensatz that, for some $N(P) \in \mathbb{N}$, we have that $Z_n = Z_{N(P)}$ for all $n \geq N(P)$. □

We let S, Φ, Q be as in the introduction. Then $S \subset \mathbb{R}^{N+1} \subset \mathbb{R}^{N+1} \cup \infty$ which is the $N+1$ sphere. We extend P to $\mathbb{C} \cup \{\infty\} = \overline{\mathbb{C}}$, i.e. the Riemann sphere, by mapping the point at infinity to itself. We call the extension $\overline{P}$.

Note that we can conjugate with any affine map of the form $az + b$ with a non-zero real, because these maps preserve vertical lines. Hence, we can assume the polynomial is of the form

$$P(z) = e^{i\theta} z^d + \mathcal{O}(z^{d-2}).$$

Definition 2.4. If P maps a vertical line to itself, we say that P is strongly exceptional.

We give an example of a strongly exceptional map.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Example 2.5. Let $P(z) = -iz^2 + ia$ for a real number a. Then $P(0 + iy) = i[y^2 + a]$. Hence, the imaginary axis is invariant and the dynamics on the y axis is $y \to y^2 + a$. For large a the Julia set is contained in the y axis. For this map the measure μ of maximal entropy is supported on the y axis, so the push forward ν to S is the Dirac mass at zero. In particular, the result in Theorem 1.4 does not hold.

The following is clear.

LEMMA 2.6. *A polynomial $P(z) = a_d z^d + \sum_{j=0}^{d-2} a_j z^j$ is strongly exceptional if and only if the y axis is mapped to itself, which is equivalent to $a_j i^{j-1}$ being real for all $j \le d$.*

Definition 2.7. We say that $P(z) = a_d z^d + \sum_{j=0}^{d-2} a_j z^j$ is weakly exceptional if $a_d i^{d-1}$ is real, but there is at least one $0 \le j \le d-2$ for which $a_j i^{j-1}$ is not real.

If P is weakly exceptional, then the set $\overline{S} := S \cup \infty$ is not compact. We will exclude this case from consideration. In fact, we have the following.

THEOREM 2.8. *If P is exceptional, then the image $\Phi(\mathbb{C})$ is not closed in $\mathbb{R}^{N+1}$.*

Proof. First suppose that P is strongly exceptional. Without loss of generality we may assume that P maps the imaginary axis to itself. Let $x \in \mathbb{R}\backslash\{0\}$ and consider points $z_k = x + ik$. Then, for $k \in \mathbb{N}$ large enough, we obtain a sequence of pre-images $w_1, \ldots, w_N$, with $P(w_{n+1}) = w_n$ and $P(w_1) = z_k$, for which $w_1, \ldots, w_N$ converge to the imaginary axis as $k \to \infty$. Therefore, the points $(\mathrm{Re}(w_N), \ldots, \mathrm{Re}(w_1), x)$ lie in $\Phi(\mathbb{C})$ and converge to $(0, \ldots, 0, x)$ as $k \to \infty$. But $(0, \ldots, 0, x) \notin \Phi(\mathbb{C})$, hence $\Phi(\mathbb{C})$ is not closed in $\mathbb{R}^{N+1}$.

Now suppose that P is weakly exceptional. We use horizontal lines instead of vertical lines for convenience. Note that, since the real axis is not mapped to itself, there can be at most d points on the real axis for which $y_0 = y_1 = 0$. Hence, if we can show if there are more than d values for t such that $(0, 0, \ldots, 0, t)$ is in the closure of S, then we have shown that S is not closed. We can assume that $P(z) = z^d + a_{d-2}z^{d-2} + \cdots$. Let w denote Böttcher coordinates near infinity. We then easily see that $w = z + (\alpha/z) + \mathcal{O}(1/z^2) + \cdots$ or equivalently that $z = w - (\alpha/w) + \cdots$. The real line is invariant in the Böttcher coordinate, so there is an invariant curve for $P(z)$ of the form $y = u - (\alpha/u) + \cdots$. Given an initial point z_0 corresponding to $w_0 = re^{i\epsilon}$, we get $w_k = r^{d^k} e^{d^k i\epsilon}$ and then $z_k = r^{d^k} e^{d^k i\epsilon} - (\alpha/r^{d^k} e^{d^k i\epsilon}) + \cdots$. We see then that $y_k = r^{d^k} d^k \epsilon + \cdots$. We choose ϵ so that $r^{d^N} d^N \epsilon = t$. The conclusion follows. □

We next study polynomials which are neither strongly nor weakly exceptional. We call these non-exceptional. They are characterized by the condition that $a_d i^d$ is not purely imaginary.

THEOREM 2.9. *If P is a non-exceptional polynomial, then $S \cup \infty = \overline{S}$ is compact. Moreover, the map Φ extends to a continuous map from $\mathbb{P}^1$ to $\overline{S}$ by sending infinity to infinity. Similarly, Q extends to a continuous map from $\overline{S}$ to itself by sending infinity to infinity.*

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

To prove the theorem it is sufficient to show that if the sequence $\{z_n\} \subset \mathbb{C}$ converges to infinity, then the images $\{\Phi(z_n)\}$ converge to infinity in $\mathbb{R}^{N+1}$. We do this by estimating orbits.

We write $a_d i^d = A + iB$. Then the assumption that P is non-exceptional gives $A \neq 0$. Pick some small $\eta > 0$ so that $\eta(1+\eta)^{d-1} \leq |A|/(4d|a_d|)$. Note that by the mean value theorem $|(1+z)^d - 1| \leq d(1+|z|)^{d-1}|z|$ for any complex number z.

LEMMA 2.10. *There exists an $R > 0$ so that if $|x| \leq \eta|y|$ and $|y| \geq R$, then $|x_1| \geq |Ay^d|/2$.*

Proof.

$$P(z) - a_d(iy)^d = a_d(iy)^d(1 + (x/(iy))^d - 1) + \sum_{n<d} a_n(x+iy)^n,$$

$$|x_1 - Ay^d| \leq |a_d y^d| d(1+\eta)^{d-1}\eta + \frac{|Ay^d|}{4}$$

$$\leq \frac{|Ay^d|}{2}. \qquad \square$$

We set $\sigma := \eta/(1+\eta)$, i.e. $(1-\sigma)/\sigma = 1/\eta$.

LEMMA 2.11. *There exists an $R' > 0$ so that if $|x| \geq \eta|y|$ and $|x| \geq R'$, then*
(i) $|x_1| \geq (\sigma|a_d||x|^d)/2$ *or*
(ii) $|x_1| \leq \eta|y_1|$ *and* $|y_1| \geq ((1-\sigma)|a_d||x^d|)/2$.

Proof. We have $x_1 + iy_1 = a_d(x+iy)^d + \sum_{n<d} a_n(x+iy)^n$. Hence if R' is large enough, $|x_1| + |y_1| \geq |x_1 + iy_1| \geq (|a_d||x+iy|^d)/2$. It follows that if Conclusion (i) fails, then $|y_1| \geq ((1-\sigma)|a_d||x^d|)/2$, but then it also follows that $|y_1| \geq (1-\sigma)/\sigma|x_1|$. $\square$

LEMMA 2.12. *There exist $R'', \lambda > 0$ so that if $x + iy \in \mathbb{C}$ and $|x| \geq R''$, then $|x_1| \geq \lambda|x|^d$ or $|x_2| \geq \lambda^{d+1}|x|^{d^2}$.*

Proof. Assume that $|x| \geq R''$ where R'' will be chosen sufficiently large. The first possibility is that $|x| \leq \eta|y|$. Then $|y| \geq R''/\eta \geq R$. Hence by Lemma 2.10 we have $|x_1| \geq |Ay^d|/2 \geq (|A||x|^d)/(2\eta^d)$. The other possibility is that $|x| \geq \eta|y|$. Then we can conclude Case (i) or (ii) in Lemma 2.11. In Case (i) we are done, so assume we are in Case (ii). Then the point (x_1, y_1) satisfies the condition of Lemma 2.10. Hence

$$|x_2| \geq \frac{|Ay_1^d|}{2} \geq \frac{1}{2} \cdot |A| \left(\frac{(1-\sigma)|a_d||x^d|}{2} \right)^d. \qquad \square$$

More generally, it follows that $|x_n| \geq \lambda^{(d^n-1)/(d-1)}|x|^{d^n}$ cannot fail for two consecutive positive integers n.

Theorem 2.9 now follows from the combination of Lemmas 2.10 and 2.12.

3. *Mirrored orbits*

The goal of this section is to carefully study which points z, w in $\mathbb{C}$ cannot be distinguished by their real orbits. In the next section we will only use Lemma 3.9, whose proof does not make use of other results in this section.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Definition 3.1. A point $z \in \mathbb{C}$ is *mirrored* by w if $w \neq z$ and $\mathrm{Re}(z_n) = \mathrm{Re}(w_n)$ for all $n \geq 0$. We say that z is mirrored if there exists a w that mirrors z.

If P has real coefficients then every point $z = x + iy$ with $y \neq 0$ is mirrored by $\bar{z}$.

Definition 3.2. Consider points z and w that mirror each other. We say that the mirror *breaks* if there exists an $n \in \mathbb{N}$ so that $P^n(z) = P^n(w)$.

Recall that a polynomial $P = \sum_{n=0}^{d} a_n z^n$ is called exceptional if the equation $\mathrm{Re}(a_d i^d) = 0$ is satisfied.

LEMMA 3.3. *Let P be a non-exceptional polynomial. Suppose that there exists an open set U of points that are mirrored. Then the mirroring extends to an unbounded connected open set.*

Proof. We first write $P(x, y) = (A(x, y), B(x, y))$ as a map on $\mathbb{R}^2$. Here A, B are real polynomials of two real variables. We complexify this as a map $P'(u, v) = (A'(u, v), B'(u, v))$ from $\mathbb{C}^2$ to itself, where we replace x and y by independent complex variables. For $z \in U$, let $\phi(z)$ be a point that mirrors z. We may choose ϕ real analytic, and write V for the set of mirror images. The real analytic map $\phi(x, y)$ between U and V extends to a biholomorphic map $\psi(u, v) = (u, \Lambda(u, v))$ between open sets U', V' in $\mathbb{C}^2$. If we write the iterate $(P')^n(u, v) = (A'_n(u, v), B'_n(u, v))$ then we have, by analytic continuation, the equation $A'_n(u, v) = A'_n(\psi(u, v))$ on U'.

By our assumption that P is not exceptional, it follows that the polynomial $A(x, y)$ contains the term y^d. Then the level sets of $A'(u, v)$ are branched covers over the u axis. Inside such level sets, the map $\psi(u, v)$ maps points on a level set to the same level set without changing the u coordinate. This map extends by monodromy along any curve in the level set which avoids branch points. The map might be multiple valued. We show that, for large values of $|u|$, all the solutions of $A'(u, v) = 0$ are real values of v, when u is real, as follows.

Let $L_c = \{A(x, y) = c\}$ for real c, $|c| \leq C$. We want to show that if $A'(u, v) = A'(u, v') = c$ and $\|(u, v)\| \geq R$, then v' is real if u is real. This will ensure that the mirror exists on an open set near infinity in $\mathbb{C}$.

For large $\|(u, v)\|$, we have that $|v| \leq K|u|$ if $A'(u, v) = c$. Moreover, for fixed u, the equation $A'(u, v) = c$ has d complex roots v with multiplicity. However, we see that, for u real and large, $A(u, v) = c$ already has at least d real roots v: we see this by writing $x + iy = re^{i\theta}$. Then $P(x, y) = |a_d| r^d e^{id\theta + \psi} + \cdots$. The equation $\mathrm{Re}(P) = c$ will have at least d solutions.

By analytic continuation this mirroring extends for all n by the same monodromies. We next restrict back to the real coordinates. By the previous observation we have extended the mirroring to an unbounded connected open set. □

Let us denote by I_∞ the basin of attraction of infinity. Recall that near infinity a polynomial $P = a_d z^d + \cdots + a_0$ is conjugate to $z \to z^d$. The conjugation map ϕ_∞ is called the Böttcher map, and if $\phi_\infty(z) = re^{2\pi i\theta}$, then r and θ are called the Böttcher coordinates of z. A curve of the form $\mathrm{Arg}(\phi_\infty(z)) = 2\pi\theta$ is called an external ray with angle θ. Let G denote the Green function for the filled in Julia set of P.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

LEMMA 3.4. *Suppose that z is mirrored by w. If $G(w) > G(z)$ then, for sufficiently large $n \in \mathbb{N}$, the point w_n must lie on an external ray which at infinity is tangent to the imaginary axis.*

Proof. By our assumption $|w_n|/|z_n| \to \infty$, hence $|\mathrm{Arg}(w_n)|$ must converge to $\pi/2$. As $z \mapsto a_d z^d$ acts expansively on the external rays this convergence can only occur if, for all sufficiently large $n \in \mathbb{N}$, the point w_n lies on an external ray tangent to the imaginary axis. □

COROLLARY 3.5. *Except for points on a real one-dimensional set, $z \in I_\infty$ can only be mirrored by points on the same level curve of the Green function.*

As large sub-level sets of the Green function are strictly convex we obtain the following.

COROLLARY 3.6. *For sufficiently large $R > 0$, generic points $z \in \mathbb{C} \setminus D(R)$ can be mirrored by at most one point, which must lie on the same level curve of the Green function.*

LEMMA 3.7. *Suppose that there exists an unbounded open set on which all points are mirrored. Then a_d is real.*

Proof. Suppose a generic z, with $|z|$ large, is mirrored. As noted above, z can only be mirrored by the point w lying on the other intersection of the level curve of G with the vertical line through z. In particular, we notice that $\mathrm{Arg}(w) + \mathrm{Arg}(z) \to 0$ as $|z| \to \infty$. The same must hold for $P(w)$ and $P(z)$, unless $P(z) = P(w)$, which will not hold on an open set.

Denote the argument of a_d by θ. As $|z| \to \infty$ the arguments of $P(w)$ and $P(z)$ are approximately equal to the arguments of $a_d w^d$ and $a_d z^d$, which are $d\mathrm{Arg}(w) + \theta$ and $d\mathrm{Arg}(z) + \theta$. Hence, we see that

$$d\mathrm{Arg}(w) + \theta \sim -(d \cdot (-\mathrm{Arg}(w)) + \theta),$$

which means that the argument 2θ must be close to zero. As the error goes to zero as $|z| \to \infty$ while θ is fixed, it follows that a_d must be real. □

We will proceed to prove the following.

THEOREM 3.8. *Let P be a non-exceptional polynomial, and suppose that there exists a non-empty open set U so that every $z \in U$ is mirrored. Then P has real coefficients.*

Proof. By our assumption on U and by Lemmas 3.3 and 3.7, the leading coefficient of P is real. Hence, we may conjugate with a linear map $z \mapsto az + b$, with $a \neq 0$ real, such that P becomes of the form $z \mapsto z^d + a_{d-2} z^{d-2} + \cdots + a_0$. Such a conjugation maps mirrored points to mirrored points.

As noted above, the open set U extends to an unbounded open set on which mirroring occurs. Hence we may well assume that U lies in the neighborhood of infinity where the Böttcher coordinates are defined. Note that points in $P^n(U)$ must be mirrored for all $n \in \mathbb{N}$, unless mirrors break, which can only happen on one-dimensional subsets. Let $\phi_\infty(z) = z + l.o.t.$ be holomorphic in a neighborhood of infinity such that $\phi_\infty(P(z)) = \phi_\infty(z)^d$, for z sufficiently large. It follows that, for M sufficiently large, the set $\phi(P^M(U))$ contains an annulus centered at the origin.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

For the purpose of a contradiction we assume that the coefficients of P are not all real, and write $P = g + ih$, where g and h have real coefficients. We write

$$g(z) = z^d + a_{d-2}z^{d-2} + \cdots + a_0, \quad h(z) = b_k z^k + \cdots + b_0,$$

where $b_k \neq 0$, and without loss of generality we may assume that $b_k > 0$.

We consider the image under P and P^2 of an interval $I = \{R + iy\}$, with $R > 0$ large and $|y| < 1$. Note that

$$g(R+iy) = g(R) + idR^{d-1}y - \binom{d}{2}R^{d-2}y^2 + O(R^{d-4})O(y^2) + iO(R^{d-3})O(y),$$

and

$$ih(R+iy) = ih(R) - kb_kR^{k-1}y + iO(R^{k-2})O(y^2) + O(R^{k-2})O(y).$$

Hence

$$\begin{aligned} X(y) &:= \mathrm{Re}(P(R+iy)) \\ &= g(R) - \binom{d}{2}R^{d-2}y^2 + \mathcal{O}(R^{d-4})O(y^2) - kb_kR^{k-1}y + \mathcal{O}(R^{k-2})O(y). \end{aligned}$$

It follows that $X_{yy} < 0$ on the interval I for $R > 0$ sufficiently large.

Choose $C > 0$ sufficiently large so that, for $R > 0$ large enough and $y_1 = C \cdot R^{k+1-d}$, we have

$$\binom{d}{2}R^{d-2}|y_1|^2 >> k|b_k|R^{k-1}|y_1|.$$

Then we have that $X_y(y_1) < 0 < X_y(-y_1)$ and that $X(\pm 1) < X(y_1)$. For $R > 0$ sufficiently large the vertical line through $P(R + iy_1)$ intersects $P(I)$ in exactly one other point, say $P(R + iy_2)$.

Writing $y_2 = D \cdot R^{k+1-d}$ and solving for

$$\mathrm{Re}(P(R+iy_2)) = \mathrm{Re}(P(R+iy_1))$$

gives

$$D = -C - \frac{2k}{d(d-1)}b_k + O(R^{-1}),$$

or in other words

$$y_2 = -y_1 - \frac{2k}{d(d-1)}b_kR^{k+1-d} + O(R^{k-d}).$$

Since

$$\mathrm{Im}(P(R+iy)) = dyR^{d-1} + O(y)O(R^{d-3}) + b_kR^k + O(R^{k-1})$$

we have that

$$\begin{aligned} \mathrm{Im}(P(R+iy_1)) &= (dC + b_k)R^k + O(R^{k-1}) \quad \text{and} \\ \mathrm{Im}(P(R+iy_2)) &= \left(-dC - \left(\frac{2k}{d-1} - 1\right)b_k\right)R^k + O(R^{k-1}). \end{aligned}$$

Note that $k < d - 1$ and thus

$$-dC - \left(\frac{2k}{d-1} - 1\right)b_k > -(dC + b_k).$$

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Hence for R large enough we obtain

$$\mathrm{Re}(P^{\circ 2}(R+iy_2)) > \mathrm{Re}(P^{\circ 2}(R+iy_1)),$$

and therefore $R+iy_1$ and $R+iy_2$ are not mirrored. But $R+iy_2$ is the only point in the interval I that could mirror $R+iy_1$, and, for R large enough and $|y|>1$, one has that $G(R+iy) > G(R+iy_1)$. Hence Corollary 3.6 implies that, for R sufficiently large and generic y comparable to $C \cdot R^{k+1-d}$, the point $R+iy$ cannot be mirrored. However, since $\phi_\infty(P^M(U))$ contains a large annulus centered at the origin, it follows that $P^M(U)$ must contain such points $R+iy$, which contradicts our assumption that all points in U are mirrored. □

LEMMA 3.9. *Let $z \in \mathbb{C}$ and suppose that the vertical line containing z is not invariant. Then there are at most d^2-1 points w which mirror z.*

Proof. Let x_0, x_1, x_2 be the first three points in the orbit. Let L_0, L_1, L_2 denote the vertical lines through these points. If $P(L_0) \neq L_1$, then by Lemma 2.1, there are at most d points in L_0 mapped to L_1. So the mirror can have at most $d-1$ points. If $P(L_0) = L_1$, then by our hypothesis that P is not exceptional we have $L_1 \neq L_0$ and hence by Lemma 2.1, $P(L_1) \neq L_2$. But then at most d points in L_1 can be mapped to L_2 and each of those has at most d primages in L_0. So there are at most d^2-1 points in the mirror of z. □

We define the set

$$X := \{(z, w) \in \mathbb{C}^2 \mid z \text{ mirrors } w\}.$$

LEMMA 3.10. *The dimension of X is at most two.*

Proof. If we include the diagonal, the dimension is at least two. If there is an invariant vertical line L, then $L \times L$ consists of mirrored points. This set is two-dimensional. Except for this, there is, for every point z, at most a finite number of points w for which (z, w) are mirrored. This excludes components of dimension three or more. □

Let P again be a non-exceptional polynomial, and suppose that X contains a two-dimensional component, distinct from the diagonal and $L \times L$. Then there is a small piece of this which is an unbranched cover over an open set U in the z axis. Then by Theorem 3.8, P has real coefficients. Then X is a graph over U, with $w = \phi(z)$. Necessarily $\phi = x + i\lambda(x, y)$ and λ is real analytic. We can shrink U and make ϕ a diffeomorphism between U and $V = \phi(U)$: if the mirroring is not conjugation, then by continuation to infinity we have at least three mirrored points, which contradicts Corollary 3.6.

In conclusion we have the following.

LEMMA 3.11. *For non-exceptional polynomials the set of mirrored points is of dimension at most one, with the exception of conjugate points in the case where P has real coefficients.*

An easy way to obtain a curve of mirrored points is the following. Start with a generic z with two pre-images z_1, z_2 having the same real part. Then there is a biholomorphic

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

map $\phi : U(z_1) \to U(z_2)$ so that w and $\phi(w)$ have the same image. In a neighborhood of z_1, there is a zero set of the function $(\mathrm{Re})(\phi(w) - w)$ which is a real curve. This is a curve of mirrored points.

4. *Entropy*

In this section we will prove our main result, which compares both topological and metric entropy of the maps P and Q. As there are different equivalent definitions of entropy in the literature, let us recall the definitions that we will be using.

Let X be a metrizable compact topological space and let $F : X \to X$ be a continuous map. If U is an open neighborhood of the diagonal $\Delta := \{(x, x); x \in X\}$ we define (U, n) balls $B(x, U, n)$, centered at $x \in X$, by

$$B(x, U, n) = \{y \in X \mid \{(x, y), (F(x), F(y)), \ldots, (F^{n-1}(x), F^{n-1}(y))\} \subset U\}.$$

We set $N(U, n)$ to be the maximum of the number of pairwise disjoint $B(x, U, n)$ balls one can have in X. Next we define

$$H_{U,X} = H_U := \limsup \frac{1}{n} \log N(U, n).$$

If $V \subset U$ are two neighborhoods of the diagonal, then $H_V \geq H_U$. Let $U_1 \supset U_2 \cdots \supset U_n \cdots$ be any neighborhood basis of the diagonal. Then we can define the topological entropy

$$h_{\mathrm{top}}(F, X) = \lim_{n \to \infty} H_{U_n}.$$

If we let ρ denote any metric defining the topology of X, then we can set $U_\epsilon := \{(x, y) : \rho(x, y) < \epsilon\}$. We call the balls $B(x, \epsilon, n)$ and $N(U_\epsilon, n) =: N(\epsilon, n)$, $H_{U_\epsilon} =: H_\epsilon$. So then $h_{\mathrm{top}} = \lim_{\epsilon \to 0} H_\epsilon$.

Let λ be a probability measure on X:

$$h_\lambda(F, x, U) = \liminf_n -\frac{1}{n} \log(\lambda(B(x, U, n))).$$

We notice that $h_\lambda(F, x, U)$ increases when we replace U by a smaller set:

$$h_\lambda(F, x) = \sup_U h_\lambda(F, x, U).$$

If λ is invariant under F, then $h_\lambda(F, x) \geq h_\lambda(F, F(x))$. If λ is also ergodic, then this function is constant λ-almost everywhere. The metric entropy $h_\lambda(F)$ is defined to be this constant.

Let ρ be any metric on X defining the topology of X. We can then use the neighborhoods of the diagonal $U_\epsilon := \{(x, y); \rho(x, y) < \epsilon\}$. We call the balls $B(x, \epsilon, n)$. Also, we set $h_\lambda(F, x, \epsilon) = h_\lambda(F, x, U_\epsilon)$ and then

$$h_\lambda(F, x) = \sup_\epsilon h_\lambda(F, x, \epsilon).$$

Clearly, the metric entropy is independent of the metric. In fact, it is a topological invariant.

Let us recall a few standard facts that can be found in any book on ergodic theory, see for example [**8**] or [**6**]. If $F : (X, \lambda) \to (X, \lambda)$ and $G : (Y, \nu) \to (Y, \nu)$ are conjugate,

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

that is, there exists a homeomorphism $\Phi : X \to Y$ so that $\Phi \circ F = G \circ \Phi$, then $h_{\text{top}}(F) = h_{\text{top}}(G)$. If $\Phi_*(\lambda) = \nu$ then it also follows that $h_\lambda(F) = h_\nu(G)$.

If Φ is not necessarily invertible and F and G are merely semi-conjugate, then it follows that $h_{\text{top}}(F) \geq h_{\text{top}}(G)$ and $h_\lambda(F) \geq h_\nu(G)$.

From now on we consider the dynamics of a non-exceptional polynomial P and the induced maps $\Phi : \mathbb{C} \to \mathbb{R}^{N+1}$ and $Q : \mathbb{R}^{N+1} \to \mathbb{R}^{N+1}$. Since Φ gives a semi-conjugation between P and Q we have the following result.

COROLLARY 4.1. *If the non-exceptional polynomial P has degree d then the topological entropy of the map Q is at most* $\log(d)$.

Let λ be a probability measure on $\mathbb{P}^1$, and let $\nu = \Phi_*\lambda$ be the push forward to $\overline{S}$.

LEMMA 4.2. *If λ is invariant on $\mathbb{P}^1$, then the push forward ν is invariant on $\overline{S}$. If λ also is ergodic, then ν is ergodic as well.*

Proof. We prove first the invariance of ν. Let E be a Borel set in $\overline{S}$:

$$\begin{aligned}\nu(E) &= \lambda(\Phi^{-1}(E)) \\ &= \lambda(P^{-1}(\Phi^{-1}(E))) \\ &= \lambda((\Phi \circ P)^{-1}(E)) \\ &= \lambda((Q \circ \Phi)^{-1}(E)) \\ &= \lambda((\Phi^{-1}(Q^{-1}(E)))) \\ &= \nu(Q^{-1}(E)).\end{aligned}$$

Next we prove ergodicity. Let $E \subset \overline{S}$ be a Borel set. Assume that the set $Q^{-1}(E)$ is the same as E except for a set of ν-measure zero. So the sets $Q^{-1}(E)\backslash E$ and $E\backslash Q^{-1}(E)$ have ν-measure zero. The pull backs of E, $Q^{-1}(E)$ and the two difference sets to $\mathbb{P}^1$ have the same λ-measures. Consider the sets $F = \Phi^{-1}(E)$ and $P^{-1}(F)$. Since $P^{-1}(\Phi^{-1}(E)) = \Phi^{-1}(Q^{-1}(E))$ we see that F is invariant modulo sets of measure zero for P. Hence F has measure zero or one. Hence the ν-measure of E is also zero or one. □

By the semi-conjugacy we have the following.

LEMMA 4.3. $h_\lambda(P) \geq h_\nu(Q)$.

Recall from Lemma 3.9 that if P is a non-exceptional polynomial, then, for every $z \in \mathbb{P}^1$, there are at most $M = d^2 - 1$ points w with $\Phi(w) = \Phi(z)$. This is enough to guarantee that the entropy of Q actually equals the entropy of P. For the topological entropy this is a consequence of [**3**, Theorem 4.1.15]. We give a self-contained proof here for the convenience of the reader.

We denote by $d(\cdot\,,\cdot)$ the spherical metric on $\mathbb{P}^1$. For $k \in \mathbb{N}$ we define

$$U_k := \left\{(z, w) \;\middle|\; d(z, w) < \frac{1}{k}\right\}.$$

Note that the sets U_k form a neighborhood basis for the diagonal in $\mathbb{P}^1 \times \mathbb{P}^1$.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Let $\delta = \delta(k) > 0$ be sufficiently small such that if $z_0, w_0 \in \mathbb{P}^1$ are such that $d(z_0, w_0) < 2\delta$, then $d(z_j, w_j) < 1/k$ for $0 \le j \le k-1$. Let $\overline{x} \in \overline{S}$ and write $\Phi^{-1}(\overline{x}) = \{a_j\}_{j=1}^r$. By continuity of the map Φ and compactness of $\mathbb{P}^1$ there exists a neighborhood $\Omega = \Omega(\overline{x})$ whose lift to $\mathbb{P}^1$ satisfies

$$\Phi^{-1}(\Omega) \subset \bigcup_{j=1}^{r} D\left(a_j, \frac{\delta}{2}\right).$$

Then let $\eta = \eta(\overline{x}) > 0$ be such that

$$\bigcup_{j=1}^{r} D(a_j, \eta) \subset \Phi^{-1}(\Omega),$$

and define $\mathcal{N}(\overline{x})$ to be the maximal subset of $\bigcup_{j=1}^r D(a_j, \eta/3)$ for which

$$\Phi^{-1} \circ \Phi(\mathcal{N}(\overline{x})) = \mathcal{N}(\overline{x}).$$

Note that $\mathcal{N}(\overline{x})$ is obtained from $\bigcup_{j=1}^r D(a_j, \eta/3)$ by removing the relatively closed set of points whose mirrors do not all lie in $\bigcup_{j=1}^r D(z_j, \eta/3)$. We then define

$$V_k := \bigcup_{\overline{x} \in S} \mathcal{N}(\overline{x}) \times \mathcal{N}(\overline{x}).$$

LEMMA 4.4. *The sets $\Phi(V_k)$ form a neighborhood basis of the diagonal in $\overline{S} \times \overline{S}$, and $\Phi^{-1}\Phi(V_k) = V_k$ for each $k \in \mathbb{N}$.*

Let $z \in \mathbb{P}^1$ and consider all points $\overline{x} \in \overline{S}$ for which $z \in \mathcal{N}(\overline{x})$. Let μ be the supremum over all $\eta(\overline{x})$. Then there certainly exists an element in S with $z \in \mathcal{N}(\overline{x})$ for which $\eta(\overline{x}) \ge \frac{2}{3}\mu$. From now on let $\overline{x}$ be such an element.

Let $\overline{y}$ be any other element of $\overline{S}$ for which $z \in \mathcal{N}(\overline{y})$. Let us write $\Phi^{-1}(\overline{x}) = \{a_j\}_{j=1}^r$ and $\Phi^{-1}(\overline{y}) = \{b_j\}_{j=1}^s$.

LEMMA 4.5.

$$\mathcal{N}(\overline{y}) \subset \bigcup_{j=1}^{r} D(a_j, \delta).$$

Proof. Since $z \in \mathcal{N}(\overline{y})$, we may assume that $d(z, b_1) < (\eta(\overline{y}))/3$. Similarly, we have $d(z, a_1) < (\eta(\overline{x}))/3$. Since $\eta(\overline{y}) \le \mu$ and $\eta(\overline{x}) \ge \frac{2}{3}\mu$, it follows that

$$d(a_1, b_1) < \frac{\eta(\overline{x})}{3} + \frac{\eta(\overline{y})}{3} \le \left(\frac{1}{3} + \frac{1}{2}\right)\eta(\overline{x}) < \eta(\overline{x}).$$

Hence $b_1 \in \Phi^{-1}(\Omega(\overline{x}))$, and therefore $\Phi^{-1}(\overline{y}) \subset \Phi^{-1}(\Omega(\overline{x}))$. Since

$$\Phi^{-1}(\Omega(\overline{x})) \subset \bigcup D\left(a_j, \frac{\delta}{2}\right) \quad \text{and}$$

$$\mathcal{N}(\overline{y}) \subset \bigcup D\left(b_j, \frac{\delta}{2}\right),$$

it follows that

$$\mathcal{N}(\overline{y}) \subset \bigcup D(a_j, \delta). \qquad \square$$

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Let $z \in \mathbb{P}^1$ and define

$$W_z := \{w \in \mathbb{P}^1 \mid (z, w) \in V_k\}.$$

Recall from Lemma 3.9 that, for each $\overline{x} \in \overline{S}$, the set $\Phi^{-1}(\overline{x})$ contains at most $M = d^2 - 1$ elements.

COROLLARY 4.6. *The set W_z is contained in at most M disks of radius δ. Moreover, there exists an $\overline{x} \in \overline{S}$ so that the centers of these disks can be chosen to lie in the set $\Phi^{-1}(\overline{x})$.*

In order to estimate the topological entropy of Q from below we will work with maximally separated sets with respect to the neighborhood bases $\{U_k\}$ and $\{(\Phi \times \Phi)(V_k)\}$. Notice that a collection of points $\mathcal{A} \subset \mathbb{P}^1$ is (n, V_k)-separated if and only if $\Phi(\mathcal{A}) \subset S$ is $(n, (\Phi \times \Phi)(V_k))$-separated. This is useful as it allows us to only work in $\mathbb{P}^1$.

THEOREM 4.7. *Suppose that P is a non-exceptional polynomial. Then the topological entropy of Q is* $\log(d)$.

Proof. Let $k \in \mathbb{N}$, and define U_k, $V_k \subset \mathbb{P}^1 \times \mathbb{P}^1$ as above. Let $n \in \mathbb{N}$, and suppose that $\mathcal{A}$ is a collection of points in $\mathbb{P}^1$ that are (n, U_k)-separated. We would like to estimate from below the minimal number of points in $\mathcal{A}$ whose images under Φ are $(n, (\Phi \times \Phi)(V_k))$-separated, or equivalently, the minimal number of points in $\mathcal{A}$ that are (n, V_k)-separated. Let $z \in \mathcal{A}$, and let $Y \subset \mathcal{A}$ contain the points w^i that are not (n, V_k)-separated from z. We will estimate the size of the set Y from above.

By Corollary 4.6 the only possible way for (z_j, w_j^i) to lie in V_k but not in U_k is for the pair to lie in two distinct disks of radius $\delta = \delta(k)$ that contain the set W_z. Moreover, the centers of these disks must then be at least $(1/k) - 2\delta$ apart. If z_j and w_j^i instead lie in the same disk of radius δ, then by our assumption on δ it follows that (z_l, w_l^i) will lie in U_k for $l = j, \ldots, j + k - 1$.

It follows that we can represent the points in Y by unique words in the letters 1 through M. Let us be more precise. At time zero our set Y is covered by at most M disks. We assign to these disks the letters 1 through (at most) M. The first letter we assign to each element $w^i \in Y$ is naturally the letter of a disk that contains the point $w_0^i = w^i$.

Then consider the first time for which two points in Y with the same initial letter have drifted at least $1/k$ apart, say after j iterates. At that time we again assign the (at most) M disks a letter, and give each element w^i its second letter, namely the letter corresponding to a disk that contains w_j^i. Later letters are assigned similarly. By our assumption on $\delta(k)$ each word has at most $n/k + 1$ letters, and by our assumption that the points in Y are (n, U_k)-separated it follows that the words corresponding to these points are all unique. We can therefore estimate the number of elements in Y by

$$|Y| \leq M^{n/k+1}.$$

We denote by $N(n, U_k)$ and $N(n, V_k)$ the maximal number of respectively (n, U_k)- and (n, V_k)-separated points in $\mathbb{P}^1$. It follows that if $\mathcal{A}$ contains $N(n, U_k)$ points, then at least

$$N(n, U_k)/M^{n/k+1}$$

of those points are (n, V_k)-separated. In other words,

$$N(n, U_k) \leq M^{n/k+1} \cdot N(n, V_k).$$

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

It follows that

$$\frac{1}{n}\log N(n, U_k) \le \left(\frac{1}{k}+\frac{1}{n}\right)\log M + \frac{1}{n}\log N(n, V_k).$$

Hence the topological entropy of Q is at least $\log(d) - (1/k)\log M$, which holds for all $k \in \mathbb{N}$. Hence $h_{\text{top}}(Q) \ge \log(d)$. As we have already estimated the entropy from above, the proof is complete. □

We continue with the metric entropy. The following fact is standard; again we provide a proof for the convenience of the reader.

LEMMA 4.8. *Suppose that σ and τ are two invariant, non-zero, positive measures and let $\eta := \sigma + \tau$. Then η is invariant. If η is ergodic, then both σ and τ are ergodic. Moreover, they must both be multiples of η.*

Proof. Invariance of η is clear.

Next, assume that η is ergodic. We can suppose that η is a probability measure. Suppose that E is a set with positive σ-measure. Then $\tau(E) > 0$: let $H := \bigcap_k \bigcup_{n\ge k} F^{-n}(E)$. By invariance of τ this set has positive τ-measure, hence also positive η-measure. The set is also invariant, hence $\eta(H) = 1$. Therefore, also $\tau(H) > 0$. But this implies that also $\tau(E) > 0$.

Suppose next that σ is not ergodic. Then there exists a set E with $0 < \sigma(E) < \sigma(X)$ which is invariant, i.e. the sets $F^{-1}(E)\backslash E$ and $E\backslash F^{-1}(E)$ both have σ-measure 0. This implies that the same is true for τ. Hence E is also invariant for τ and therefore also for η. But $0 < \eta(E) < \eta(X)$, contradicting ergodicity of η.

It remains to show the last statement. Let $\sigma' = \sigma/\sigma(X)$, $\tau' = \tau/\tau(X)$, $\eta' = \eta/\eta(X)$. So these are invariant ergodic probability measures. Let E be a set. Then $\chi \circ F^n$ converge to a constant function, $\sigma'(E) = \tau'(E) = \eta'(E)$. Proportionality follows. □

LEMMA 4.9. *If ν is invariant ergodic on $\overline{S}$, then ν is the push forward of an invariant ergodic measure on $\mathbb{P}^1$.*

Proof. We define inductively a sequence (λ_n) of measures on $\mathbb{P}^1$, all of which have ν as push forward. We would like to find a measure which is also backwards invariant under P.

We divide $\mathbb{P}^1$ into finitely many sets A_j, where A_j consists of those points for which the fibers of Φ have exactly j points. So the mirrors consist of the points in $\bigcup_{j>1} A_j$. We define λ_1 by dividing the measure of $\Phi(A_j)$ into j equal measures.

We inductively define $\lambda_{n+1}(E) = \lambda_n(P^{-1}(E))$. Next we show that this procedure keeps the property of the push forward being ν. Let F be a subset of $\overline{S}$ and let E denote $\Phi^{-1}(F)$. Then

$$\begin{aligned}\lambda_{n+1}(E) &= \lambda_n(P^{-1}(E))\\ &= \lambda_n(P^{-1}(\Phi^{-1}(F)))\\ &= \lambda_n(\Phi^{-1}(Q^{-1}(F)))\\ &= \nu(Q^{-1}(F)) = \nu(F).\end{aligned}$$

Next, we use Cesaro means. Set $\lambda'_n = (1/n)\sum_{1\le j\le n}\lambda_j$. The push forwards of these measures are all equal to ν. The total mass of the signed measure $\lambda'_{n+1} - \lambda'_n$ is at

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

most $2/(n+1)$. We get

$$\begin{aligned}|\lambda'_{n+1}(E) - \lambda'_{n+1}(P^{-1}(E))| &\le |\lambda'_{n+1}(E) - \lambda'_n(P^{-1}(E))| + \frac{2}{n+1}\\ &= \left|\lambda'_{n+1}(E) - \frac{1}{n}\sum_{1\le j\le n}\lambda_j(P^{-1}(E))\right| + \frac{2}{n+1}\\ &= \left|\lambda'_{n+1}(E) - \frac{1}{n}\sum_{2\le j\le n+1}\lambda_j(E)\right| + \frac{2}{n+1},\end{aligned}$$

and, therefore,

$$\begin{aligned}|\lambda'_{n+1}(E) - \lambda'_{n+1}(P^{-1}(E))| &\le \left|\lambda'_{n+1}(E) - \frac{1}{n+1}\sum_{1\le j\le n+1}\lambda_j(E)\right| + \frac{4}{n+1}\\ &= \frac{4}{n+1}.\end{aligned}$$

It follows that any weak limit is invariant on $\mathbb{P}^1$, and also that ν is the push forward of any weak limit. Pick such a limit η. Suppose that η is not ergodic. Then we can write $\eta = \sigma + \tau$, where both measures are non-zero and invariant and their supports are disjoint. Let σ', τ' be their push forwards. These measures are invariant by Lemma 4.2 and their sum is ergodic. Hence by Lemma 4.8, they are both ergodic and in fact $\sigma' = a\nu$ and $\tau' = b\nu$ for $a + b = 1$ and $0 < a, b$. Hence we have that ν is the push forward of two measures, with disjoint supports. This implies that ν cannot charge the complement of the mirrors. We can repeat the argument if neither σ nor τ is ergodic. We then see that ν cannot charge any mirror with at most three points in the mirror. Repeating the procedure finitely many times and using that there is an upper bound on the number of mirrored points, we see that one of the measures obtained must be ergodic. Hence ν must be the push forward of an invariant, ergodic measure. □

Now we denote by μ the measure of maximal entropy on $\mathbb{P}^1$. We have the following.

LEMMA 4.10. *The metric entropy of the push forward of the measure μ is the same as the metric entropy of the measure μ.*

Proof. Let $\mathcal{X} \subset \overline{S}$ denote the finite set of points $\Phi(z)$ where z is either an isolated mirror point, a singular point of the one-dimensional mirror set, an isolated cluster point on the diagonal for mirrored points, or a critical point in either the Julia set or a parabolic basin. We include in $\mathcal{X}$ the full orbit for any periodic point in $\mathcal{X}$.

Let Y be the union of the vertical lines passing through the points in $\mathcal{X}$. Since P is not strongly exceptional, none of these vertical lines is forward invariant. It follows that each of these vertical lines has μ-measure 0, and hence $\mu(Y) = 0$.

Let $k >> 1$ be an integer. Since $\mu(Y) = 0$, there exists a $\delta_0 > 0$ so that $N_{\delta_0}(Y))$, the δ_0-neighborhood of Y, satisfies $\mu(N_{\delta_0}(Y)) < 1/k$. Let z now be a *generic point* in J, as defined in [**1**]. Then it follows that

$$\lim_{n\to\infty} \frac{\#\{0 \le j < n \mid P^j(z) \in N_{\delta_0}(Y)\}}{n} < \frac{1}{k}.$$

As Y has measure 0, we may also assume that the orbit of z never hits the set Y.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Write $\overline{x} = (x_0, \ldots, x_N) = \Phi(z)$, and let us estimate the entropy function for $\Phi_*(\nu)$ at $\overline{x}$.

For $\ell < n$ define balls in $\mathbb{P}^1$:

$$B'(n, \ell, \epsilon) := \bigcap_{r=0}^{n-\ell} \{w \in \mathbb{P}^1 \mid \rho(\Phi(P^r(w)), \overline{x}_{\ell+r}) < \epsilon\}.$$

Here we use the metric $\rho(\overline{x}, \overline{y}) = \max_{0 \le i \le N} |(x_i - y_i)|$ on a large ball in $\mathbb{R}^{N+1}$ containing the image of the filled-in Julia set.

The idea is to show that, for $\epsilon > 0$ small enough,

$$\mu(B'(n, \ell - 1, \epsilon)) \le \frac{\mu(B'(n, \ell, \epsilon))}{d}, \tag{4.1}$$

for all ℓ except at most a fraction of $1/k$, and for those we have

$$\mu(B'(n, \ell - 1, \epsilon)) \le \mu(B'(n, \ell, \epsilon)).$$

It follows that the μ-measure of $B'(n, 0, \epsilon)$ is at most $C(1/d)^{n-n/k}$. Therefore the metric entropy on $\overline{S}$ is at least $\log d$.

We need two additional lemmas to prove that equation (4.1) holds sufficiently often.

LEMMA 4.11. *There exists a $\lambda > 0$ so that if $x + iy_1$, $x + iy_2$ are distinct points in the filled-in Julia set with $\Phi(x + iy_1) = \Phi(x + iy_2)$, and the distance from $\Phi(x + iy_1)$ to $\mathcal{X}$ is at least δ_0, then $|y_1 - y_2| > \lambda$.*

Proof. Suppose to the contrary that there exist mirrored points $x_j + iy_j$, $x_j + iv_j$ with $x_j + iy_j$, $x_j + iv_j \to x + iy$. Then it follows that $\Phi(x + iy) \in \mathcal{X}$. □

LEMMA 4.12. *Let $\sigma < \lambda/2$. Then there exists $\epsilon, \tau > 0$ so that the following holds. If $(x_0, \ldots, x_N) \in \overline{S}$ with distance to $\mathcal{X}$ at least δ_0, then, for all $u + iv \in \hat{\mathbb{C}}$ with $\Phi(u + iv) \in B(\overline{x}, \epsilon)$ and $|u - x_0| < \tau$, there exists $x_1 + iy_1 \in \Phi^{-1}(\overline{x})$ for which $|v - y_1| < \sigma$.*

Proof. Follows immediately from the definition of $\mathcal{X}$. □

We continue the proof of Lemma 4.10. Since Y contains the critical points of P, there is a uniform $\delta < \delta_0$ such that P is one-to-one on the disk $D_\delta(w)$ for all $w \in \hat{\mathbb{C}} \backslash N_{\delta_0}(Y)$.

Consider now a set $B'(n, \ell, \epsilon)$ for ϵ sufficiently small. By the above lemmas, we can divide $B'(n, \ell, \epsilon)$ into *pieces*, where each subset is contained in a small ball centered at one of the mirror points of $P^\ell(z)$.

Let ℓ now be such that $P^{\ell-1}(z) \notin N_{\delta_0}(Y)$. Recall, from our choice of z, that $P^{\ell-1}(z) \notin N_{\delta_0}(Y)$ fails at most $1/k$ of the time for n sufficiently large.

The set $B'(n, \ell - 1, \epsilon)$ can be similarly divided into pieces, and it follows from the equidistribution property of the measure μ, plus our assumption that $P^{\ell-1}(z) \notin N_{\delta_0}(Y)$, that each of the components of $B'(n, \ell - 1, \epsilon)$ has mass at most $1/d$ times the component of $B'(n, \ell, \epsilon)$ that it is mapped into. While it is possible that several pieces of $B'(n, \ell - 1, \epsilon)$ are mapped into the same piece of $B'(n, \ell, \epsilon)$, the number of pieces is uniformly bounded and can never increase. This suffices to show that

$$\mu(B'(n, 0, \epsilon)) \le C\left(\frac{1}{d}\right)^{n-n/k}. \qquad \square$$

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press

Let us now restate and prove the main theorem of the paper.

THEOREM 1.4. *Let P be a non-exceptional complex polynomial of degree $d \geq 2$. Then the probability measure ν is invariant and ergodic. Moreover, it is the unique measure of maximal entropy,* $\log d$.

Proof. It follows from Lemma 4.2 that ν is invariant and ergodic. From Lemma 4.10 it follows that the metric entropy of ν is $\log d$. From Theorem 4.7 we see that P also has topological entropy $\log d$. If $\sigma \neq \nu$ is any other invariant ergodic probability measure on S, then by Lemma 4.9, σ is the push forward of an invariant ergodic probability measure τ on $\mathbb{P}^1$. Necessarily, $\tau \neq \mu$. Hence the metric entropy of τ is strictly less than $\log d$. It follows by Lemma 4.3 that the metric entropy of σ is strictly less than $\log d$. Therefore ν is the unique measure of maximal entropy. □

Acknowledgement. The second author was supported by a SP3-People Marie Curie Actions grant in the project Complex Dynamics (FP7-PEOPLE-2009-RG, 248443).

REFERENCES

[**1**] R. Bowen. Topological entropy for noncompact sets. *Trans. Amer. Math. Soc.* **184** (1973), 125–136.
[**2**] H. Brolin. Invariant sets under iteration of rational functions. *Ark. Mat.* **6** (1965), 103–144.
[**3**] T. Downarowicz. *Entropy in Dynamical Systems (New Mathematical Monographs, 18)*. Cambridge University Press, Cambridge, 2011.
[**4**] M. Ju. Lyubich. Entropy properties of rational endomorphisms of the Riemann sphere. *Ergod. Th. & Dynam. Sys.* **3** (1983), 351–385.
[**5**] R. Mañé. On the uniqueness of the maximizing measure for rational maps. *Bol. Soc. Brasil. Math.* **14** (1983), 27–43.
[**6**] K. Petersen. *Ergodic Theory (Cambridge Studies in Advanced Mathematics, 2)*. Cambridge University Press, Cambridge, 1983.
[**7**] F. Takens. Detecting strange attractors in turbulence. *Dynamical Systems and Turbulence (Warwick, 1980) (Lecture Notes in Mathematics, 898)*. Springer, New York, 1981, pp. 366–381.
[**8**] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York, 1982.

https://doi.org/10.1017/etds.2015.36 Published online by Cambridge University Press