ARKIV FÖR MATEMATIK Band 8 nr 4

1.68 15 1 Communicated 27 November 1968 by ÅKE PLEIJELL and LENNART CARLESON

Some connections between ergodic theory and the iteration of polynomials

By TOM S. PITCHER and JOHN R. KINNEY

I. Introduction

In a recent paper [1] Brolin has shown some connections between the theory of the iteration of polynomials in the complex plane and the ergodic transformations induced by the polynomials. [1] contains an exposition of the classical theory of iteration and a bibliography of the subject.

Consider a polynomial P of degree N and its iterates P_n given by $P_n(z) = P(P_{n-1}(z))$. The fixpoints of P, i.e., solutions of $P_n(z) = z$ are classified as repulsive if $|P_n'(z)| > 1$, indifferent if $|P_n'(z)| = 1$ and attractive if $|P_n'(z)| < 1$. Primary interest centers on the set F of points where (P_n) is not a normal family. F can also be characterized as the closure of the set of repulsive fixpoints. Replacing P by $L \circ P \circ L^{-1}$ with L a linear function only subjects the fixpoints to a linear transformation so we can assume that

$$P(z) = z^N + \sum_{i=0}^{N-2} a_i z^i.$$

It can be shown that F is compact, contains no open set and is completely invariant under P, i.e., $F = P(F) = P^{-1}(F)$.

II. The equilibrium measure for F

In [1] Brolin defines a natural probability measure on F as follows. Choose any point z_0 in the plane with at most two exceptions and let μ_n be the atomic measure assigning weight N^{-n} to each root of $P_n(z) = z_0$. The μ_n converge weakly to a probability measure μ supported on F, independent of the starting point z_0. μ is invariant under the transformation P and in fact, P is an ergodic transformation of F into itself under this measure.

It also turns out that μ is the equilibrium measure for F, that is, it minimizes the energy integral

$$I(\nu) = \iint \log \frac{1}{|z-w|} \, \nu(dz)\, \nu(dw)$$

among all Borel probability measures ν supported on F.

Let $c_1, \ldots, c_k$ be the critical points of the inverses of P and for each $0 \leqslant \theta < 2\pi$ let $l_i(\theta)$ be the half line $[c_i + \lambda e^{i\theta}, 0 \leqslant \lambda < \infty]$. We can find a θ_0 for which the half lines are all distinct and a $\delta > 0$ such that any two half lines $l_i(\theta_1)$, $l_j(\theta_2)$ with $\theta_0 - \delta < \theta_1, \theta_2 < \theta_0 + \delta$ intersect in a point outside F if at all. Thus the sets $A(\theta) = F \cap (l_1(\theta) \cup \ldots \cup l_k(\theta))$ are disjoint for θ in this interval so we can choose one, say $\bar{\theta}$ with $\mu(A(\bar{\theta})) = 0$. If we make the cuts $l_1(\bar{\theta}), \ldots, l_k(\bar{\theta})$ the inverses $g_1, \ldots, g_N$ of P are defined on $F - A(\bar{\theta})$. It is easily seen that

$$\frac{1}{N} \sum_{i=1}^{N} \int f(g_i(z))\, \mu_{n-1}(dz) = \int f(z)\, \mu_n(dz)$$

and hence that
$$\frac{1}{N} \sum_{i=1}^{N} \int f(g_i(z))\, \mu(dz) = \int f(z)\, \mu(dz).$$

It follows that $\mu(P_n(A(\bar{\theta})) = 0$ for all n and hence that

$$F_0 = F - \bigcup_{n=0}^{\infty} P_n(A(\bar{\theta}))$$

has μ-measure 1.

Now the g_i's are defined in a neighborhood of each point of F_0 and since each g_i takes F_0 into itself, all the inverses $g_{\alpha_1} \circ g_{\alpha_2} \circ \ldots \circ g_{\alpha_n}$ of P_n are defined in a neighborhood of each point of F_0. This does not imply that there are neighborhoods in which the inverses of all the P_n are defined.

We can now define the integer valued function $\alpha_n(z)$ for z in F_0 to be the solution of

$$g_{\alpha_n(z)}(P_n(z)) = P_{n-1}(z).$$

It is easily seen that

$$g_{\alpha_1(z)} \circ g_{\alpha_2(z)} \circ \ldots \circ g_{\alpha_n(z)}(P_n(z)) = z$$

and that $\alpha_n(P(z)) = \alpha_{n+1}(z)$. We will write

$$I_n(\beta_1, \ldots, \beta_n) = [z \mid \alpha_i(z) = \beta_i,\ i = 1, \ldots, n]$$

and
$$I_n(z) = I_n(\alpha_1(z), \ldots, \alpha_n(z)).$$

The transformation $z \to [\alpha_1(z), \alpha_2(z), \ldots]$ maps F into a sequence space and, as the following theorem shows, it takes μ into the "Bernoulli trial" measure.

Theorem 2.1. *Under μ the α_n are independent random variables with distribution*

$$\mu([z \mid \alpha_n(z) = k]) = \frac{1}{N} \quad (k = 1, \ldots, N).$$

Proof. The set $I_n(\beta_1, \ldots, \beta_n)$ contains all the points $g_{\beta_1} \circ \ldots \circ g_{\beta_n}(w)$ where $P_m(w) = z_0$ and no other solutions of $P_{n+m}(z) = z_0$. Hence, the set has μ_{n+m} measure N^{-n} and thus also μ measure N^{-n}.

In connection with the next theorem it should be remarked that in the case $P(z)=z^2$, F is the unit circle, μ is Lebesgue measure and the P_k are of course trigonometric functions and that in the case $P(z)=z^2-2$, $F=[-2,2]$, $\mu=C\,dx/\sqrt{4-x^2}$ and the P_k are a subsequence of the Chebycheff polynomials.

The functions $1, z, z^2, \dots$ are continuous and bounded on F, hence are square integrable with respect to μ. Let $Q_0=1, Q_1, \dots$ be the corresponding sequence of orthonormal polynomials having positive leading coefficients.

Theorem 2.2.

$$Q_{N^n}=\left[\int |z|^2\,\mu(dz)\right]^{-\frac{1}{2}} P_n \quad (n=0,1,2,\dots).$$

Proof. P_n has degree N^n and leading coefficient 1. Also

$$\int |P_n(z)|^2\,\mu(dz)=\int |z|^2\,\mu(dz).$$

For $n=0$, taking $P_0(z)=z$, we have

$$\int Q_0(z)\overline{P}_0(z)\,\mu(dz)=\int \bar{z}\mu(dz)=\frac{1}{N}\sum_{\alpha=1}^{N}\int \overline{g_\alpha(z)}\,\mu(dz)=0,$$

since $\sum_{\alpha=1}^{N} g_\alpha(z)$ is the coefficient of z^{N-1} in P which is 0. For $n>1$ and $k<N^n$ we have

$$\begin{aligned}\int z^k \overline{P}_n(z)\,\mu(dz) &= N^{-n}\sum_{\alpha_1\dots\alpha_n=1}^{N}\int (g_{\alpha_1}\circ\dots\circ g_{\alpha_n}(z))^k\overline{P}_n(g_{\alpha_1}\circ\dots\circ g_{\alpha_n}(z))\,\mu(dz)\\ &= N^{-n}\int \bar{z}\sum_{\alpha_1\dots\alpha_n=1}^{N}(g_{\alpha_1}\circ\dots\circ g_{\alpha_n}(z))^k\,\mu(dz).\end{aligned}$$

But the summation is $\sum w^k$ extended over the roots of $P_n(w)=z$ and this symmetric function depends only on the first k coefficients in

$$P_n(w)-z=w^{N^n}+c_1 w^{N^n-1}+\dots+c_{N^n}$$

and hence is a constant A independent of z. Thus,

$$\int z^k\,\overline{P_n(z)}\,\mu(dz)=AN^{-n}\int \bar{z}\mu(dz)=0.$$

III. The polynomials z^2-p for $p>2$

In this section we deal with a special class of P's. We assume that there exists a simply connected domain D containing F and containing none of the critical points of the functions P_n nor any limit points of them. It is known (see [1]) that in this case the set of inverses

$$[g_{\alpha_1} \circ \ldots \circ g_{\alpha_n} \mid 1 \leqslant \alpha_i \leqslant N,\ 1 \leqslant n < \infty]$$

forms a normal family in D having only constant limiting functions. We can extend the α_i to all of F in this case and we write

$$G_n(z, w) = g_{\alpha_1(z)} \circ g_{\alpha_2(z)} \circ \ldots \circ g_{\alpha_n(z)}(w).$$

Theorem 3.1. *For fixed z, $G_n(z, w)$ converges to z uniformly on compact subsets of D. The convergence is uniform on $F \times F$.*

Proof. To prove the first assertion we have only to show that the constant limit is z but this is obvious since $G_n(z, P_n(z)) = z$ for each n. For each $z \in F$ we can find an n such that $|G_m(z, w) - z| < \varepsilon/2$ for all $w \in F$ and $m \geqslant n$. Then for $z' \in I_n(z) \cap [z' \mid |z - z'| < \varepsilon/2]$ we have

$$|G_m(z', w) - z'| = |G_n(z, g_{\alpha_{n+1(z')}} \circ \ldots \circ g_{\alpha_{m(z')}}(w)) - z'| \leqslant \varepsilon/2 + |z - z'| < \varepsilon.$$

The $I_n(z)$ are open, in this special case, so this gives an open covering of F and the proof is now completed in the usual way.

We now choose a $w \in F$ which is not a fix point and set

$$\varrho_n(z) = G_n(z, w).$$

By Theorem 3.1

$$\varepsilon_n = \max_{z \in F} |\varrho_n(z) - z|$$

goes to zero as n goes to ∞. None of the numbers

$$g_{\alpha_j} \circ \ldots \circ g_{\alpha_{n+1}}(w) - g_{\alpha_j} \circ \ldots \circ g_{\alpha_n}(w)$$

vanishes since $P(w) \neq w$ so, setting $\alpha_k = \alpha_k(z)$, we can write

$$\frac{1}{n} \log |\varrho_{n+1}(z) - \varrho_n(z)|$$
$$= \frac{1}{n} \sum_{k=1}^{n-l} \log \left| \frac{g_{\alpha_k}(\varrho_{n+1-k}(P_k(z))) - g_{\alpha_k}(\varrho_{n-k}(P_k(z)))}{\varrho_{n+1-k}(P_k(z)) - \varrho_{n-k}(P_k(z))} \right|$$
$$+ \frac{1}{n} \log |\varrho_{l+1}(P_{n-l}(z)) - \varrho_l(P_{n-l}(z))|.$$

Using the facts that g_i, g_i', g_i'' and $(g_i')^{-1}$ are bounded on F we can easily show that

$$\log \left| \frac{g_{\alpha_k}(\varrho_{n+1-k}(P_k(z))) - g_{\alpha_k}(\varrho_{n-k}(P_k(z)))}{\varrho_{n+1-k}(P_k(z)) - \varrho_{n-k}(P_k(z))} \right| - \log |g'_{\alpha_k}(P_k(z))| \leqslant C\, \varepsilon_{n-k}.$$

Thus $$\frac{1}{n}\log|\varrho_{n+1}(z)-\varrho_n(z)|=\frac{1}{n}\sum_{k=1}^{n-}\log|g'_{\alpha_k}(P_k(z))|+0(\varepsilon_l)+A,$$

where $$|A|=\frac{1}{n}|\log|\varrho_{l+1}(P_{n-l}(z))-\varrho_l(P_{n-l}(z))||=0\left(\frac{1}{n}\right).$$

Theorem 3.2. *With μ-probability one*

$$\lim_{n\to\infty}\frac{1}{n}\log|\varrho_{n+1}(z)-\varrho_n(z)|=-H$$

and $$\limsup_{n\to\infty}\frac{1}{n}\log|\varrho_n(z)-z|=-H,$$

where $$H=\frac{-1}{N}\sum_{i=1}^{N}\int\log|g_i'(z)|\,\mu(dz).$$

Proof. The α_n form a stationary ergodic sequence and $\log|g'_{\alpha_k}(P_k(z))|$ is bounded so the ergodic theorem applies to

$$\frac{1}{n}\sum_{k=1}^{n-l}\log|g'_{\alpha_k}(P_k(z))|$$

and this plus the estimates above proves the first assertion. For any positive ε and large enough n,

$$|\varrho_n(z)-z|\leqslant\sum_{k=1}^{\infty}|\varrho_{n+k}(z)-\varrho_n(z)|\leqslant\sum_{k=1}^{\infty}e^{-(n+k)(H-\varepsilon)}=\frac{e^{-n(H-\varepsilon)}}{1-e^{-(H-\varepsilon)}},$$

so $$\limsup_{n\to\infty}\frac{1}{n}\log|\varrho_n(z)-z|\leqslant -H.$$

On the other hand

$$\max(|\varrho_n(z)-z|,|\varrho_{n+1}(z)-z|)\geqslant\tfrac{1}{2}|\varrho_{n+1}(z)-\varrho_n(z)|,$$

so the opposite inequality also obtains.

The polynomials $P(z)=z^2-p$ for $p>2$ satisfy the special requirements of this section. It can be shown [1] that in this case $F\subset[-\tfrac{1}{2}-\sqrt{\tfrac{1}{4}+p},\ \tfrac{1}{2}+\sqrt{\tfrac{1}{4}+p}]$ and the critical points are $-p$, $P(-p)$, $P_2(-p)$, etc. Computation shows that

$$-p<-\tfrac{1}{2}-\sqrt{\tfrac{1}{4}+p}\text{ and }\tfrac{1}{2}+\sqrt{\tfrac{1}{4}+p}<P(-p)<P_2(-p)<\dots$$

so we can take D to be the plane with the intervals $(-\infty,\ -p]$ and $[P(-p),\ \infty)$ removed.

Brolin [1] has given an upper bound for the Hausdorff dimension of F for $p \geqslant 2+\sqrt{2}$. We are now in a position to give a lower bound for $p>2$.

Theorem 3.3. *Let F_p be the F set for z^2-p, $p>2$ and μ_p the associated measure. Then*

$$\dim (F_p) \geqslant \frac{1}{1+\dfrac{\int \log (x+p)\, \mu_p(dx)}{2 \log 2}}.$$

Proof. In this case $g_i(x) = \pm\sqrt{x+p}$ and the right hand side is equal to $\log 2/H$. We are going to make use of Lemma 2 of [2] (or, more accurately, of the second half of the proof). It is proved there that if $D_n(x)$ is the dyadic interval of order n containing x and if A is a subset of

$$\left[x \,\middle|\, \limsup_{n\to\infty} \frac{1}{n} \log_2 \mu(D_n(x)) \leqslant -\alpha\right]$$

with $\mu(A)>0$ then $\dim (A) \geqslant \alpha$.

It is easily seen that the sets $I_n(x)$ are contained in disjoint intervals for this case (see [1], p. 126). If we write

$$|I| = \sup_{x,\, y\in I} |x-y|$$

and set $$A(n,\varepsilon) = [x|\, |I_m(x)| \geqslant e^{-m(H+\varepsilon)} \quad \text{for all} \quad m \geqslant n],$$

then $$[x|\, |\varrho_{m+1}(x) - \varrho_m(x)| \geqslant e^{-m(H+\varepsilon)} \quad \text{for all} \quad m \geqslant n] \subset A(n,\varepsilon),$$

so $\mu(A(n,\varepsilon)) \to 1$ as $n\to\infty$ for any positive ε.

Take n so large that $\mu(A(n,\varepsilon))>0$ and k so large that

$$\frac{-k \log 2}{H+\varepsilon} + 1 \leqslant -n.$$

If m_k is the largest integer such that

$$2^{-k} < e^{-m_k(H+\varepsilon)},$$

then $-(m_k+1)(H+\varepsilon) \leqslant -k \log 2$ so that

$$-m_k \leqslant \frac{-k \log 2}{H+\varepsilon} + 1 \leqslant -n.$$

At most two sets of the form $I_{m_k}(x)$ for $x \in A(n,\varepsilon)$ can intersect a dyadic interval of order k and $\mu(I_{m_k}(x)) = 2^{-m_k}$ so

$$\log_2(\mu(D_k(x)\cap A(n,\varepsilon)))\leqslant -m_k+1\leqslant \frac{-k\log 2}{H+\varepsilon}+2.$$

Replacing μ by μ_n, $$\mu_n(B)=\frac{\mu(B\cap A(n,\varepsilon))}{\mu(A(n,\varepsilon))}$$

in the result quoted above we see that $\dim(A(n,\varepsilon))\geqslant(\log 2)/(H+\varepsilon)$ for all n with $\mu(A(n,\varepsilon))>0$. Since $\bigcup_n A(n,\varepsilon)\subset F$

$$\dim(F)\geqslant(\log 2)/(H+\varepsilon)$$

and the proof is completed by letting $\varepsilon\to 0$.

We wish to estimate the integral in the above theorem.

$$A_p=\int\log(x+p)\,\mu_p(dx)=E\left(\log\left(p+\theta_1\sqrt{p+\theta_2\sqrt{p+\dots}}\right)\right),$$

when the θ_i are independent and are ± 1 with equal probability. Thus

$$\begin{aligned}A_p&=\tfrac{1}{2}\left[\log\left(p+\sqrt{p+\theta_2\sqrt{\dots}}\right)+\log\left(p-\sqrt{p+\theta_2\sqrt{\dots}}\right)\right]\\&=\tfrac{1}{2}\log\left(p^2-p-\theta_2\sqrt{p+\theta_3\sqrt{\dots}}\right)\\&=\tfrac{1}{4}\log\left((p^2-p)^2-p-\theta_3\sqrt{p+\theta_4\sqrt{\dots}}\right)\\&=2^{-n}E\left(\log\left(B_n(p)-\theta_{n+1}\sqrt{p+\theta_{n+2}\sqrt{\dots}}\right)\right),\end{aligned}$$

where $B_0(p)=p$ and $B_{n+1}(p)=B_n^2(p)-p$. Since $B_n(p)\uparrow\infty$ and $\theta_{n+1}\sqrt{p+\theta_{n+2}\sqrt{\dots}}$ is in F_p and hence is bounded, we have

$$A_p=\lim_{n\to\infty}2^{-n}\log B_n(p).$$

Now

$$\begin{aligned}2^{-(n+1)}\log B_{n+1}(p)&=2^{-(n+1)}\log(B_n^2(p)-p)\\&=2^{-n}\log B_n(p)+2^{-(n+1)}\log\left(1-\frac{p}{B_n^2(p)}\right)<2^{-n}\log B_n(p),\end{aligned}$$

so that $$A_p\leqslant\tfrac{1}{2}\log B_1(p)=\tfrac{1}{2}\log(p^2-p).$$

Combining this with Brolin's result we have

$$\left[1+\frac{\log\sqrt{p(p-1)}}{2\log 2}\right]^{-1}\leqslant\dim F_p\leqslant\left[1+\frac{\log(p-\frac{1}{2}-\sqrt{\frac{1}{4}+p})}{2\log 2}\right]^{-1},$$

where the left hand inequality holds for $p\geqslant 2$ and the right hand one for $p\geqslant 2+\sqrt{2}$.

ACKNOWLEDGEMENT

This work was partially supported by National Science Foundation grants GP-6216 and GP-6852.

Department of Mathematics, University of Southern California, Los Angeles, California 90007. U.S.A. (T. S. P.) and *Department of Mathematics, Michigan State University, East Lansing, Michigan, U.S.A.* (J. R. K.).

REFERENCES

1. Brolin, H., Invariant sets under iteration of rational functions. *Arkiv för Mat. 6*, nr 6 (1965).
2. Kinney, J. R., and Pitcher, T. S., Dimensional properties of a random distribution function on the square. *Ann. Math. Stat. 37*, no. 4 (1966).

Tryckt den 14 maj 1969

Uppsala 1969. Almqvist & Wiksells Boktryckeri AB