BACKWARD ITERATION ALGORITHMS FOR JULIA SETS OF MÖBIUS SEMIGROUPS

RICH STANKEWITZ AND HIROKI SUMI

ABSTRACT. We extend a result regarding the Random Backward Iteration algorithm for drawing Julia sets (known to work for certain rational semigroups containing a non-Möbius element) to a class of Möbius semigroups which includes certain settings not yet been dealt with in the literature, namely, when the Julia set is not a thick attractor in the sense given in [8].

1. Introduction

In this paper we consider the dynamics of rational semigroups (i.e., semigroups of rational maps on the Riemann sphere \( \hat{\mathbb{C}} \) where the semigroup operation is the composition of maps). In particular, we consider the algorithm to draw the figures of the Julia sets of finitely generated Möbius semigroups (i.e., semigroups of Möbius maps on \( \hat{\mathbb{C}} \)). In [20] two methods for generating graphical computer approximations of Julia sets of finitely generated rational semigroups were discussed, the full backward iteration algorithm and the random backward iteration algorithm. The former method was justified by the work of Boyd [5] and later generalized by Sumi [21], and the latter method was justified by the present authors in [20]. (See also [1, 2] where these methods were first explored. For the related topics in smooth ergodic theory, see [17].) However, because of complexities that do not exist when there is a map of degree two or more, neither method was verified for classes of Möbius semigroups in these papers. The goal of the present paper, however, is to verify both methods for certain classes of Möbius semigroups (see Theorems 2.2 and 3.1 for precise statements).

The full backward iteration algorithm creates successive approximations to the Julia set as follows. Starting with a seed value \( z_0 \in \mathbb{C} \), the set \( A_1 \) of all preimages of \( z_0 \) under all generating maps \( \{f_j : j = 1, \ldots, k\} \) of the semigroup is constructed. (Note, our use of the word preimage refers to a preimage of order 1 of \( z_0 \), i.e., a point \( y \) such that...
$f_j(y) = z_0$ for some $j$, as opposed to preimages of order $n \geq 2$, i.e., $y$ such that $f_{j_1} \circ \cdots \circ f_{j_n}(y) = z_0$.) Iteratively, the set $A_n$ is constructed to be the set of all preimages (of order 1) of all points in $A_{n-1}$ under all generating maps. Hence, $A_n$ is the set of all preimages of order $n$ of $z_0$. In general, for large $n$ the set $A_n$ approximates the Julia set of the semigroup (see precise statements in Section 2 where the $A_n$’s are exactly the supports of a corresponding convergent sequence of measures).

The random backward iteration algorithm (also known as the “ergodic method” or “chaos game” method) creates successive approximations to the Julia set as follows. Starting with a seed value $z_0 \in \mathbb{C}$ a random walk $\{z_n\}$ is generated by setting $z_n$ to be the outcome of randomly selecting one preimage of $z_{n-1}$ under a randomly selected generating map. In general, the plotted points of this random walk then will give a visual approximation to the Julia set of the semigroup (see precise statements in Section 3 given in terms of convergence of certain measures).

Analogous results, for both the “full” and “random” methods, in the context of attractor sets for contracting iterated function systems (see [13, 3, 7, 18]) and Julia sets of iterated single rational functions (see [10, 15, 16, 14]) are known. Both the full and random methods for both attractor sets of iterated function systems and Julia sets of rational semigroups are implemented using the freely available application Julia 2.0 [6].

We now introduce the basic terminology and notions needed to describe the full and random methods mentioned above.

For the entire paper we let $G = \langle f_1, \ldots, f_k \rangle$ be a rational semigroup generated by Möbius maps $f_j$ where the semigroup operation is the composition of maps, i.e., $G$ is the collections of all maps which can be expressed as a finite composition of maps from the generating set $\{f_j : j = 1, \ldots, k\}$. Also, since the results depend not on just the semigroup $G$ but also on the particular choice of generators, we assume each such semigroup comes with a particular fixed generating set of maps.

Research on the dynamics of rational semigroups was initiated by Hinkkanen and Martin in [11] (see also [12]), and this remains a primary source for background information. However, they studied only rational semigroups containing at least one element of degree at least two. For an in depth look at the dynamics of Möbius semigroups see [8].

We follow [11] in saying that the Fatou set $F(G)$ is the set of points in $\hat{\mathbb{C}}$ which have a neighborhood on which $G$ is normal, and its complement in $\hat{\mathbb{C}}$ is called the Julia set $J(G)$. 
We quote the following results from [11]. The Fatou set \( F(G) \) is *forward invariant* under each element of \( G \), i.e., \( g(F(G)) \subseteq F(G) \) for all \( g \in G \), and thus \( J(G) \) is *backward invariant* under each element of \( G \), i.e., \( g^{-1}(J(G)) \subseteq J(G) \) for all \( g \in G \). We note that the sets \( F(G) \) and \( J(G) \) are, however, not necessarily completely invariant under the elements of \( G \).

We define the *kernel Julia set* of \( G \) by \( J_{\text{ker}}(G) = \bigcap_{g \in G} g^{-1}(J(G)) \).

For a subset \( A \) of \( \hat{\mathbb{C}} \), we set \( G(A) = \bigcup_{g \in G} g(A) \). Also, we set \( G^{-1} := \{ g^{-1} : g \in G \} \) and this is called the inverse semigroup of \( G \). Note that \( G^{-1} \) is a Möbius semigroup generated by \( \{ f_1^{-1}, \ldots, f_k^{-1} \} \).

The *exceptional set* \( E(G) \) is defined to be the set of points \( z \) with a finite backward orbit \( G^{-1}(z) = \cup_{g \in G} g^{-1}(\{ z \}) \).

We record the following well-known facts for later use.

**Proposition 1.**

1. \( J_{\text{ker}}(G) \) is the largest forward invariant subset of \( J(G) \) under the action of \( G \).
2. For \( z \in \hat{\mathbb{C}} \setminus E(G) \) we have \( G^{-1}(z) \supseteq J(G) \). In particular, for \( z \in J(G) \setminus E(G) \) we have \( G^{-1}(z) = J(G) \).
3. \( G(E(G)) = E(G) \).
4. If \( \#J(G) \geq 3 \), then \( J(G) \) is both perfect (thus uncountable) and the closure of the set of points that are repelling fixed points under any map in \( G \).

**Proof.** (1) follows from the definition of \( J_{\text{ker}}(G) \) and the fact that \( J(G) \) is backward invariant under \( G \). (2) is shown in [11]. Because \( E(G) \) is clearly backward invariant, (3) follows from Proposition 2.17 in [8] when \( E(G) \) is finite. Now suppose \( E(G) \) is infinite. Since any three points in \( E(G) \) must have finite backward orbits under each \( g \in G \), we must then have each \( g \neq Id \) in \( G \) is elliptic of finite order. This implies then that \( g^{-1} \in G \) for each \( g \in G \) and thus \( G = G^{-1} \). Hence \( E(G) \) which is backward invariant under \( G \) must also be forward invariant under \( G \). (4) is shown in [19].

Our focus will be on Möbius semigroups and will require use of the following fundamental classification of maps in \( \mathcal{M} \), the group of all Möbius maps on \( \hat{\mathbb{C}} \). We recall the following definitions and results which can be found in [4]. For every complex matrix

\[
M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
\]
with \( ad - bc \neq 0 \) there is a corresponding \( m(z) = (az + b)/(cz + d) \) in \( \mathcal{M} \). We write \( \text{tr}(M) = a + d \) and \( \text{det}(M) = ad - bc \) for the trace and determinant of \( M \). Since \( m(z) \) determines \( M \) up to a scalar factor, we may unambiguously define \( \text{tr}^2[m] = \text{tr}^2(M)/\text{det}(M) \). This invariant of \( m(z) \) determines its dynamical properties, as follows.

**Theorem 1.1** (Classification of Möbius maps \( m \in \mathcal{M} \)). Let \( \text{Id} \) denote the identity map on \( \hat{\mathbb{C}} \) and let \( m \in \mathcal{M} \), \( m \neq \text{Id} \). Then

- (1) \( \text{tr}^2[m] = 4 \) if and only if \( m \) is parabolic, i.e., \( m \) has only one neutral fixed point and \( m \) is conjugate to the translation \( z \mapsto z + 1 \).
- (2) \( \text{tr}^2[m] \in [0, 4) \) if and only if \( m \) is elliptic, i.e., \( m \) has two neutral fixed points and \( m \) is conjugate to a rotation \( z \mapsto kz \) with \( |k| = 1 \).
- (3) \( \text{tr}^2[m] \in (4, +\infty) \) if and only if \( m \) is hyperbolic, i.e., \( m \) has an attracting and a repelling fixed point and \( m \) is conjugate to \( z \mapsto kz \) with \( k \in \mathbb{R} \) and \( |k| > 1 \).
- (4) \( \text{tr}^2[m] \notin [0, +\infty) \) if and only if \( m \) is strictly loxodromic, i.e., \( m \) has an attracting and a repelling fixed point and \( m \) is conjugate to \( z \mapsto kz \) with \( |k| > 1 \) and \( k \notin \mathbb{R} \).

Finally, a loxodromic map is any \( m \in \mathcal{M} \) which is either hyperbolic or strictly loxodromic, i.e., a map having an attracting and a repelling fixed point. Note, however, that this terminology is not universal.

2. **Full backward iteration algorithm**

Let \( a \in \hat{\mathbb{C}} \) be fixed. For each \( i_j \in \{1, \ldots, k\} \), we write \( z_{i_1,i_2,\ldots,i_n} = f_{i_1}^{-1} \circ \cdots \circ f_{i_n}^{-1}(a) \). Note that \( z_{i_1,i_2,\ldots,i_n} \) depends on the initial choice of \( a \) though we suppress this dependence in our notation.

Throughout we assume \( b = (b_1, \ldots, b_k) \) is a probability vector, i.e., each \( b_j > 0 \) and \( \sum_{j=1}^k b_j = 1 \).

Denoting by \( \delta_z \) the unit point mass measure at \( z \), we define probability measures \( \mu_{n}^{a,b} \) on \( \hat{\mathbb{C}} \) as follows:

\[
\mu_1^{a,b} = \sum_{i_1=1}^k b_1 \delta_{f_{i_1}^{-1}(a)}
\]
and, in general, for \( n > 1 \)

\[
\mu_{n,b} = \sum_{i_1,i_2,\ldots,i_n=1}^{k} b_{i_1} \cdots b_{i_n} \delta_{z_{i_1,i_2},\ldots,i_n}
\]

\[
= \sum_{i_1,i_2,\ldots,i_n=1}^{k} b_{i_1} \cdots b_{i_n} \delta_{f_{i_n}^{-1} \circ \cdots \circ f_{i_1}^{-1}(a)}.
\]

Following [5] and [21], we define a bounded linear operator \( T = T_{b}^G \) on the space \( C(\hat{C}) \) of continuous functions (endowed with the sup norm \( \| \cdot \|_{\infty} \)) on \( \hat{C} \) by

\[
(T\phi)(z) = \int_{\hat{C}} \phi(s) \, d\mu_{z,b}^{s}(s) = \sum_{i=1}^{k} b_{i} \phi(f_{i}^{-1}z),
\]

noting \( \| T \| = 1 \). Hence, \((T\phi)(z)\) is a weighted average of \( \phi \) evaluated at all \( k \) preimages of \( z \) under all generators \( f_{j} \).

Letting \( P(\hat{C}) \) denote the space of probability Borel measures on \( \hat{C} \), and noting that it is a compact metric space in the topology of weak* convergence, we have that the adjoint \( T_{b}^*: P(\hat{C}) \to P(\hat{C}) \) of \( T_{b}^G \) is given by

\[
(T_{b}^*)_{\rho}(A) = \int_{\hat{C}} \mu_{z,b}^{\rho}(A) \, d\rho(z)
\]

for all Borel sets \( A \subseteq \hat{C} \). Using the operator notation \( \langle \phi, \rho \rangle = \int \phi \, d\rho \) we express the action of the adjoint as \( \langle T\phi, \rho \rangle = \langle \phi, T_{b}^* \rho \rangle \). Note that the map \( \hat{C} \to P(\hat{C}) \) given by \( z \mapsto \mu_{z,b}^{\rho} \) is continuous since for a sequence \( z_{n} \to z_{0} \) in \( \hat{C} \), we have \( \langle \phi, \mu_{z_{n},b}^{\rho} \rangle = (T\phi)(z_{n}) \to (T\phi)(z_{0}) = \langle \phi, \mu_{z_{0},b}^{\rho} \rangle \) for any \( \phi \in C(\hat{C}) \), i.e., \( \mu_{z_{n},b} \to \mu_{z,b} \).

We give a claim which is obtained by using results from [22].

**Claim 1.** Let \( G = \langle f_{1},\ldots,f_{k} \rangle \) be a Möbius semigroup and let \( b = (b_{1},\ldots,b_{k}) \) be a probability vector. When \( J_{\ker}(G) = \emptyset \) and \( J(G) \neq \emptyset \), we have the following (a)(b)(c)(d)(e).

(a) \#J(G) \geq 3.

(b) There exists a unique minimal set \( L \) of \( G \), where we say that a non-empty compact subset \( L \) of \( \hat{C} \) is a minimal set of \( G \) if \( L = \bigcup_{g \in G} \{ g(z) \} \) for each \( z \in L \).

(c) \( L \cap F(G) \neq \emptyset \).

(d) \( L = \{ z \in L \mid \exists g \in G \text{ s.t. } z \text{ is an attracting fixed point of } g \} \).
(e) Let $b$ be a probability vector. Then there exists a unique Borel probability measure $\nu^b$ on $\hat{\mathbb{C}}$ such that $M^n(\phi)(z) \to \int \phi \, d\nu^b$ uniformly on $\hat{\mathbb{C}}$, where, $M = M^b_0$ is the transition operator with respect to $b$ given by $M(\phi)(z) = \sum_{j=1}^{k} b_j \phi(f_j(z))$, for any $\phi \in C(\hat{\mathbb{C}})$. Also, $\nu^b$ is the unique Borel probability measure on $\hat{\mathbb{C}}$ such that $M^* \nu^b = \nu^b$. Also, the support of $\nu^b$ is equal to $L$.

Proof. We now apply parts (2), (3), (8), (10), (13), (17), and (21) of Theorem 3.15 in [22] using the probability measure $\tau = \sum_{j=1}^{k} b_j \delta_{f_j}$ on the space of non-constant rational functions. Part (a) follows from (3). From (a) and Proposition 1(4) there exists a loxodromic map in $G$ and hence (b) follows from (21). If (c) did not hold, then $L \subseteq J(G)$ would be contained in $J_{ker}(G)$ contradicting our assumption. Part (d) follows from (17) noting that the $S_r$ in the reference is $L$ in this case. Lastly, (e) follows from (2), (8), (10), (13), and (21). □

We note that the complexities involved in the proof of Theorem 3.15 of [22] required a very delicate analysis based on the hyperbolic metric in this general setting (see Remark 6). Now, we prove the following.

Lemma 2.1. Let $G$ be a finitely generated Mobius semigroup such that $J_{ker}(G) = \emptyset$ and $J(G) \neq \emptyset$. Then, the unique minimal set $L$ of $G$ is equal to $J(G^{-1})$.

Proof. By (d) in Claim 1, we have $L \subseteq J(G^{-1})$ since any attracting fixed point of any $g \in G$ must be a repelling fixed point of $g^{-1}$.

We now consider the following two cases. Case (i) $\sharp L \geq 3$. Case (ii) $\sharp L \leq 2$.

Suppose we have case (i). Then $G(L) \subseteq L$ (thus $(G^{-1})^{-1}(L) \subseteq L)$, $\sharp L \geq 3$. Since $J(G^{-1})$ is a minimal backward invariant compact subset under $G^{-1}$ which has at least three elements, we obtain that $L \supseteq J(G^{-1})$. Hence $L = J(G^{-1})$.

Suppose we have case (ii). By (c) in Claim 1 we choose $z \in L \cap F(G)$. Since $\# L < +\infty$ and $F(G)$ is forward invariant under $G$, we see that $L = \overline{G(z)} = G(z) \subset F(G)$. By (a) in Claim 1, we can take hyperbolic distance on each connected component of $F(G)$ and for a small $\epsilon > 0$ let $V_\epsilon$ be the $\epsilon$-hyperbolic neighborhood of $L$ (we consider connected components $\{W_i\}$ of $F(G)$ which meet $L$ and we consider the $\epsilon$-hyperbolic neighborhood $A_{i,\epsilon}$ of $W_i \cap L$ in $W_i$ and let $V_\epsilon = \cup_i A_{i,\epsilon}$).

Then $(G^{-1})^{-1}(\overline{V_\epsilon}) = G(\overline{V_\epsilon}) \subseteq \overline{V_\epsilon}$ by Pick's Lemma and hence $J(G^{-1}) \subseteq \overline{V_\epsilon}$. Since $\epsilon$ is an arbitrary small number, it follows that $J(G^{-1}) \subseteq L$. Hence $J(G^{-1}) = L$. □
Let Theorem 2.2 provides the justification for the “full backward iteration algorithm” used to graphically approximate Remark 1. Theorem 2.2.

Let $G = \langle f_1, \ldots, f_k \rangle$ be a Möbius semigroup and let $b = (b_1, \ldots, b_k)$ be a probability vector. Suppose $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$. Then the measures $\mu_n^{a,b}$ converge weakly to a Borel probability measure $\mu^b = \mu^b_G$ on $\hat{C}$ independently of and uniform in $a \in \hat{C}$. Further, the support of $\mu^b$ is $J(G)$ and $T_1^b \mu^b = \mu^b$. Further, $\mu^b$ is the unique Borel probability measure on $\hat{C}$ such that $T_1^b \mu^b = \mu^b$.

Remark 1. Theorem 2.2 provides the justification for the “full backward iteration algorithm” used to graphically approximate $J(G)$. This method simply plots the $k^n$ (not necessarily distinct) points in the support of $\mu_n^{a,b}$. We note that this iterative process plots all $k$ inverses of each of the $k^{n-1}$ points in the support of $\mu_n^{a,b}$ to generate the support of $\mu_n^{a,b}$. Also, note that the support of $\mu_n^{a,b}$ is independent of $b$ (but not of $a$), which merely adjusts the weights on the point masses.

3. Random backward iteration algorithm

Let $\Sigma_k^+ = \prod_{n=1}^{\infty} \{1, \ldots, k\}$ denote the space of one-sided sequences on $k$ symbols, regarded with the usual topology and $\sigma$-algebra of Borel sets. Now let $P_b$ be the Bernoulli measure on $\Sigma_k^+$ which is then given on basis elements as follows: for fixed $j_n$ for $n = 1, \ldots, m$, we have

$$P_b(\{(i_1, i_2, \ldots) \in \Sigma_k^+ : i_n = j_n \text{ for all } n = 1, \ldots, m\}) = b_{j_1} \cdots b_{j_m}.$$ 

Let us define a random walk as follows. Starting at a point $z_0 = a \in \hat{C}$, we note that there are $k$ preimages of $a$ under the generators of the semigroup $G$. We randomly select $z_1$ to be $f_j^{-1}(a)$ with probability $b_j$. Likewise, $z_2$ is randomly selected to be one of the $k$ preimages of $z_1$. Continue in this fashion to generate what we call a random backward orbit $\{z_n\}$ of $z_0 = a$ under the semigroup $G$.

Utilizing the notation introduced in Section 2, we see that $\Sigma_k^+$ generates the entire set of backward orbits starting at $z_0 = a$ by letting the sequence $(i_1, i_2, \ldots)$ generate the backward orbit $\{g_{i_n} \circ \cdots \circ g_{i_1}(z_0)\} = \{z_{i_1,i_2,\ldots}\}$.

Formally, we define our random walk $\{z_n\}$ in terms of random variables $\{Z_n\}$ given as follows. For each $n \in \mathbb{N}$, we let $Z_n = Z_{n}^b$:
\((\Sigma^+_k, P_b) \to \hat{\mathbb{C}}\) by \(Z_n(i_1, i_2, \ldots) = z_{i_1, i_2, \ldots, i_n}\) and \(Z_0 = a\). Hence, \(Z_{n+1}(i_1, i_2, \ldots) = f_{i_n}^{-1} Z_n(i_1, i_2, \ldots)\).

The proof given for the analogous result in [20] shows the following.

**Claim 2.** Let \(G = \langle f_1, \ldots, f_k \rangle\) be a Möbius semigroup, let \(a \in \hat{\mathbb{C}}\), and let \(b = (b_1, \ldots, b_k)\) be a probability vector. Then the stochastic process \(\{Z_n : n = 0, 1, 2, \ldots\}\) forms a Markov process with transition probabilities \(\{\mu^{a,b}_n\}\), i.e., for each Borel set \(B \subseteq \hat{\mathbb{C}}\) we have \(P_b(\{Z_{n+1} \in B | Z_0, \ldots, Z_n\}) = \mu^{a,b}_n(B)\).

The main result can now be stated in terms of the probability measures, defined for each \(n \in \mathbb{N}\) and \((i_1, i_2, \ldots) \in \Sigma^+_k\), by

\[
\mu^{a}_{i_1, \ldots, i_n} = \frac{1}{n} \sum_{j=1}^{n} \delta_{i_1, i_2, \ldots, i_j}.
\]

**Theorem 3.1.** Let \(G = \langle f_1, \ldots, f_k \rangle\) be a Möbius semigroup, let \(a \in \hat{\mathbb{C}}\), and let \(b = (b_1, \ldots, b_k)\) be a probability vector. Suppose \(\ker(G^{-1}) = \emptyset\) and \(J(G^{-1}) \neq \emptyset\). Then, for \(P_b\) a.a. \((i_1, i_2, \ldots) \in \Sigma^+_k\), the probability measures \(\mu^{a}_{i_1, \ldots, i_n}\) converge weakly to \(\mu^b\) in \(\mathcal{P}(\hat{\mathbb{C}})\).

By using Theorem 2.2, the proof of Theorem 3.1 is now nearly identical with that given for the proof of the analogous result in [20], but is given in Appendix A of this paper for completeness. See [10] for the corresponding statement and proof in the case of iteration of one function of a rational map of degree two or more, which is the model for the analogous work here.

**Remark 2.** Theorem 3.1 provides the justification for the “random backward iteration algorithm” used to graphically approximate \(J(G)\). This method simply plots, for large \(n\), the \(n\) points in the support of \(\mu^{a}_{i_1, \ldots, i_n}\), i.e., the points of a random backward orbit, where \((i_1, i_2, \ldots) \in \Sigma^+_k\) is randomly selected according to \(P_b\). If \(a \notin J(G)\), then it is often appropriate to not plot the first hundred or so points in the random backward orbit since the earlier points in the orbit might not be very close to \(J(G)\).

By Claim 1, Lemma 2.1 and [22, Theorem 3.15], we can show the following result.

**Theorem 3.2.** Under the assumptions of Theorem 3.1, let \(a \in \hat{\mathbb{C}}\) and let \(b = (b_1, \ldots, b_k)\) be a probability vector. Then for \(P_b\) a.a. \((i_1, i_2, \ldots) \in \Sigma^+_k\), we have \(d(z_{i_1, \ldots, i_n}, J(G)) \to 0\) as \(n \to \infty\).
Proof. By Claim 1, there exists a unique minimal set $L$ of $G^{-1}$. By Lemma 2.1, $L = J(G)$. By [22, Theorem 3.15(7)], the statement of our theorem holds. \hfill \Box

**Remark 3.** In Theorems 2.2 and 3.1, both conditions $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$ are necessary. We see this in the following.

1. To see that the condition $J(G^{-1}) \neq \emptyset$ is necessary, let $k = 1, f_1(z) = e^{2\pi i \theta} z$ where $\theta \in \mathbb{Q}$ is a constant, and $G = \langle f_1 \rangle$. Then $J(G^{-1}) = \emptyset$ and $\mu_{1,1,\ldots,1}^n$ converges to a measure $\rho_a$ for each $a \in \hat{\mathbb{C}}$, and $\rho_a \neq \rho_{a'}$ if $|a| \neq |a'|$. Thus the condition $J(G^{-1}) \neq \emptyset$ is necessary.

2. To see that the condition $J_{\ker}(G^{-1}) = \emptyset$ is necessary, let $k = 2, f_1(z) = 2z, f_2(z) = \frac{1}{2}z$, and $G = \langle f_1, f_2 \rangle$. Then $\emptyset \neq \{0, \infty\} \subset J_{\ker}(G^{-1})$, and $\mu_{n}^{0,b} \to \delta_0, \mu_{\infty}^{0,b} \to \delta_{\infty}, \mu_{1,1,\ldots,i_n}^{0} \to \delta_0$, and $\mu_{1,1,\ldots,i_n}^{\infty} \to \delta_{\infty}$ as $n \to \infty$. Thus the condition $J_{\ker}(G^{-1}) = \emptyset$ is necessary.

Also, let $k \in \mathbb{N}$ be an arbitrary natural number and suppose a Möbius semigroup $G = \langle f_1, \ldots, f_k \rangle$ satisfies that $\emptyset \neq E(G) \cap J(G) \subsetneq J(G)$. Then we cannot have the conditions $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$ simultaneously. For, if we have both of them, then Lemma 2.1 implies that the unique minimal set $L$ of $G^{-1}$ is equal to $J(G)$. Since $G^{-1}(E(G) \cap J(G)) \subsetneq E(G) \cap J(G)$, we have a minimal set $L_0$ of $G^{-1}$ in $E(G) \cap J(G)$. Thus $L_0 = L = J(G)$. However, this contradicts the assumption $E(G) \cap J(G) \subsetneq J(G)$. Thus we cannot have the conditions $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$ simultaneously. Also, the conclusion $\text{supp} \mu^b = J(G)$ in Theorem 2.2 cannot hold for any $a \in E(G) \cap J(G)$. To construct an example of Möbius semigroup $G$ such that $\emptyset \neq E(G) \cap J(G) \subsetneq J(G)$, let $k = 2, f_1(z) = 2z$ and let $f_2$ be a Möbius transformation such that $0$ is an attracting fixed point of $f_2$ and $1$ is a repelling fixed point of $f_2$. Let $G = \langle f_1, f_2 \rangle$. Then $\emptyset \neq \{0\} = E(G) \cap J(G) \subsetneq J(G)$.

**Remark 4.** There are many examples of finitely generated Möbius semigroups $G$ such that $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$. In fact, these conditions are checkable and it is easy for us to construct examples of such Möbius semigroups. The class of such Möbius semigroups is very big. For various sufficient conditions for Möbius semigroups $G$ to satisfy $J_{\ker}(G^{-1}) = \emptyset$ and $J(G^{-1}) \neq \emptyset$ and examples of such Möbius semigroups $G$, see Section 5.

4. Consequences of the main result Theorem 3.1

Using Theorem 3.1 together with the fact that the support of $\mu^b$ is $J(G)$ no matter what probability vector $b$ is chosen, one can quickly
show the following results. Under the assumptions of Theorem 3.1, for each \( a \in \mathbb{C} \) and for each probability vector \( b \), let \( \Sigma_{a,b} := \{ (i_1, i_2, \ldots) \in \Sigma_+^n \mid \mu_{i_1,\ldots,i_n}^a \rightarrow \mu^b \text{ as } n \rightarrow \infty \} \). Note that by Theorem 3.1, we have \( P_b(\Sigma_{a,b}) = 1 \). Nearly identical details of the argument can be found in [10] and so we omit them here.

**Corollary 1.** Under the assumptions of Theorem 3.1, let \( a \in \mathbb{C} \) and \( (i_1, i_2, \ldots) \in \bigcup_b \Sigma_{a,b} \), where the union is taken over all probability vectors \( b \). Then

\[
J(G) \subseteq \bigcup_{j=0}^{\infty} \{z_{i_1, i_2, \ldots, i_j}\}.
\]

Furthermore, if \( a \in J(G) \), then

\[
J(G) = \bigcup_{j=0}^{\infty} \{z_{i_1, i_2, \ldots, i_j}\}.
\]

**Remark 5.** Meeting the conclusion of this corollary and Theorem 3.2 is generally what one is looking for when saying that the “random backward iteration” method works in drawing (an approximation of) \( J(G) \). Note that this conclusion holds for \( P_b \) a.a. \( (i_1, i_2, \ldots) \in \Sigma_+^n \), regardless of the choice of \( b \).

5. **Applications**

In this section, we give some examples to which we can apply Theorems 2.2, 3.1, 3.2 and Corollary 1. For Möbius semigroups \( S \), we give some sufficient conditions for \( J_{ker}(S) \) to be empty.

**Lemma 5.1.** Let \( S \) be a Möbius semigroup and \( S^{-1} \) its inverse semigroup. Assume \( J(S^{-1}) \setminus J(S) \neq \emptyset \) and \( E(S^{-1}) \cap J(S) = \emptyset \). Then \( J_{ker}(S) = \emptyset \).

**Proof.** Choose \( z_0 \in J(S^{-1}) \setminus J(S) \).

Then for all \( z \in \mathbb{C} \setminus E(S^{-1}) \) we see \( \overline{S(z)} = \overline{(S^{-1})^{-1}(z)} \) must contain \( z_0 \) by Proposition 1(2). Therefore for all \( z \in \mathbb{C} \setminus E(S^{-1}) \), we have \( S(z) \) meets \( F(S) \) since \( z_0 \in \mathbb{C} \setminus J(S) = F(S) \).

For \( z \in E(S^{-1}) \), since \( S(E(S^{-1})) = E(S^{-1}) \) by Proposition 1(3) we see that the assumption \( E(S^{-1}) \cap J(S) = \emptyset \) gives that \( S(z) \subseteq E(S^{-1}) \subset F(S) \).

So, for all \( z \in \mathbb{C} \) we have \( S(z) \) meets \( F(S) \). Thus \( J_{ker}(S) = \emptyset \) since \( J_{ker}(S) \) is a forward invariant subset of \( J(S) \). \( \square \)
Lemma 5.2. Let $S$ be a Möbius semigroup and suppose that $J(S^{-1}) \neq \hat{\mathbb{C}}$, $E(S^{-1}) \cap J(S) = \emptyset$, and int$(J(S)) \neq \emptyset$, where int denotes the set of interior points with respect to the topology in $\hat{\mathbb{C}}$. Then $J_{ker}(S) = \emptyset$.

Proof. Suppose $J(S^{-1}) \supset J(S)$. Since $S^{-1}(J(S)) \subset J(S)$, we have $S^{-1}(\text{int}(J(S))) \subset \text{int}(J(S))$. Since $J(S^{-1}) \neq \hat{\mathbb{C}}$, Montel’s theorem implies that int$(J(S)) \subset F(S^{-1})$. However, this is a contradiction. Thus we have $J(S^{-1}) \setminus J(S) \neq \emptyset$. By Lemma 5.1, we obtain that $J_{ker}(S) = \emptyset$. 

Lemma 5.3. Let $S$ be a Möbius semigroup and $S^{-1}$ its inverse semigroup. Assume $\#J(S) \geq 3$ and $J(S^{-1}) \cap J(S) = \emptyset$. Then $J_{ker}(S) = \emptyset$.

Proof. Suppose $w \in J_{ker}(S)$. Using Proposition 1(4) we may choose $g \in S$ such that $g$ has a repelling fixed point $p \neq w$. So $g$ has an attracting fixed point $q$ which must then be in $J(S^{-1})$. Hence $q \notin J(S)$. Since $g^n(w) \to q$ we contradict the fact that $J_{ker}(S)$ is both a closed forward invariant set under $S$ and contained in $J(S)$.

We apply the above results to the Caruso semigroups $S_\beta$ defined as follows. For a fixed nonzero complex number $\beta$ we let $f(z) = \beta + 1/z$ and $g(z) = -\beta + 1/z$. We define the Möbius semigroup $S_\beta = \{f, g\}$, and its inverse semigroup $S'_\beta = \{f^{-1}, g^{-1}\}$. See [8] for background, but note that there the notation is different in that Möbius groups are denoted by $\langle \ldots \rangle$ and semigroups are denoted $\langle \ldots \rangle$. We note that in Section 7 of [8] it is shown that neither $S_\beta$ nor $S'_\beta$ have finite orbits in $H^3 \cup \hat{\mathbb{C}}$ and so, in particular, $E(S'_\beta) = \emptyset$ and $E(S_\beta) = \emptyset$. Further it is shown that $\#J(S_\beta) \geq 3$ and $\#J(S'_\beta) \geq 3$.

Theorem 5.4. For each $\beta \in \mathbb{C} \setminus \{0\}$ we have either

(i) $J(S_\beta) \neq J(S'_\beta)$ and $J_{ker}(S_\beta) = \emptyset = J_{ker}(S'_\beta)$

or

(ii) $J(S_\beta) = J(S'_\beta)$ and $J_{ker}(S_\beta) = J_{ker}(S'_\beta)$.

Proof. From Proposition 7.14 of [8], $J(S_\beta) \neq J(S'_\beta)$ implies both $J(S_\beta) \setminus J(S'_\beta) \neq \emptyset$ and $J(S'_\beta) \setminus J(S_\beta) \neq \emptyset$. Hence by Lemma 5.1 the result follows (noting that $J(S_\beta) = J(S'_\beta)$ implies $J(S_\beta)$ is both forward and backward invariant under $S_\beta$ and thus $J_{ker}(S_\beta) = J(S_\beta) = J(S'_\beta) = J_{ker}(S'_\beta)$).

Hence whenever $J(S_\beta) \neq J(S'_\beta)$ we see by Theorem 3.1 (noting $\#J(S_\beta) \geq 3$ and $\#J(S'_\beta) \geq 3$) that both the random backward iteration algorithm and the full backward iteration algorithm work for both $S_\beta$ and $S'_\beta$, and these method can start with any seed value $a \in \hat{\mathbb{C}}$. 
This is an extension from what was shown in [8], and which we now describe.

Let \( K \subset \hat{\mathbb{C}} \) be a compact set that is forward invariant under each map in rational semigroup \( S \). Further suppose \( K \) is equipped with a metric \( d \), consistent with the topology on \( K \). If there exists a constant \( c, 0 < c < 1 \), such that for each \( s \in S \) we have \( d(s(z), s(w)) \leq cd(z, w) \) for all \( z, w \in K \), then we say that \( S \) is a CIFS (contracting iterated function system) on \( (K, d) \). We define the attractor \( A = A(K, S) \) as the closure of the set of all fixed points in \( K \) of the maps \( s \in S \). If \( \text{Int}(K) \neq \emptyset \) and \( A \subset \text{Int}(K) \), then \( A \) is called thick.

Theorem 5.7 in [8] shows that for any finitely generated M"obius semigroup \( S \) the following (i)-(iii) are equivalent: (i) \( S \) has a thick attractor \( J(S^{-1}) \), (ii) \( S^{-1} \) has a thick attractor \( J(S) \), and (iii) \( J(S^{-1}) \cap J(S) = \emptyset \) and the generators of \( S \) are loxodromic (having a repelling fixed point). Noted in [8], when a thick attractor exists both the (forward) random walk method and the (forward) full method work for drawing the corresponding Julia sets, however, the initial seed value must be chosen from the set \( K \) given in the definition of thick attractor.

**Remark 6.** Theorem 3.1 presented here is not so restrictive as Theorem 5.7 in [8] on the conditions that \( J(S^{-1}) \cap J(S) = \emptyset \) nor on the choice of starting seed value. This, however, required a more delicate analysis as conducted in the proof of Theorem 3.15 in [22] which was crucial in the proof of Theorem 3.1.

Concrete examples where Theorem 3.1 applies, but the result of [8] does not, are for \( \beta = \pm 1 \pm i \) where (see Theorem 8.2 in [8]), \( J(S_\beta) \) and \( J(S'_\beta) \) are unequal Cantor sets with \( J(S_\beta) \cap J(S'_\beta) = \{ \pm 1, \pm i \} \). Other examples of this type are for \( \beta = \pm 2 \) or \( \beta = \pm 2i \), where (see Theorem 8.1 in [8]) \( J(S_\beta) \) and \( J(S'_\beta) \) are unequal Cantor sets with \( J(S_\beta) \cap J(S'_\beta) = \{ \beta/2, -\beta/2 \} \).

We give further examples of M"obius semigroups \( G \) with \( J(G) \cap J(G^{-1}) \neq \emptyset \) to which we can apply Theorems 2.2, 3.1, 3.2 and Corollary 1.

**Example 1.** Denoting the open unit disk by \( \mathbb{D} \) with boundary \( S = \partial \mathbb{D} \), let \( f_1, \ldots, f_k \) be M"obius maps with \( f_j(\mathbb{D}) = \mathbb{D} \) for each \( j = 1, \ldots, k \) chosen such that for \( G = \langle f_1, \ldots, f_k \rangle \) we have \( J(G) \cap J(G^{-1}) \neq \emptyset \). (For example, any parabolic fixed points of any \( g \in G \) would necessarily lie in \( J(G) \cap J(G^{-1}) \). Also, if \( G = \langle f_1, \ldots, f_k \rangle \) is a Fuchsian group with \( J(G) \neq \emptyset \), then \( J(G) \cap J(G^{-1}) = J(G) \neq \emptyset \).) Note that \( J(G) \subseteq S \) and \( J(G^{-1}) \subseteq S \). Set \( f_{k+1}(z) = az \) with \( |a| > 1 \) and call \( G = \langle f_1, \ldots, f_k, f_{k+1} \rangle \). Since \( \hat{\mathbb{C}} \setminus \overline{\mathbb{D}} \) is forward invariant under \( G \) we must
have $J(\tilde{G}) \subseteq \tilde{D}$. Similarly, $J(\tilde{G}^{-1}) \subseteq \tilde{C} \setminus D$. In this way since $J(G) \subseteq J(\tilde{G})$ and $J(G^{-1}) \subseteq J(\tilde{G}^{-1})$, we have that $J(G) \cap J(G^{-1})$ contains $J(G) \cap J(G^{-1})$, which could be constructed to be a rather large subset of $S$. However, Theorems 2.2, 3.1, 3.2 and Corollary 1 still apply.

We conclude this paper with the following:

**Question 1.** Is $\{ \beta \in \mathbb{C} \setminus \{ 0 \} : J(S_\beta) \neq J(S'_\beta) \}$ open and dense in $\mathbb{C} \setminus \{ 0 \}$?

Acknowledgement. This work was partially supported by a grant from the Simons Foundation (#318239 to Rich Stankewitz). The research of the second author was partially supported by JSPS KAKENHI 24540211, 15K04899.

**Appendix A. Proof of Theorem 3.1**

In this section we introduce a key result of Furstenberg and Kifer to relate the invariance of the measure $\mu^b$ to the convergence of Cesaro averages of $\phi(z_n)$ where $\{z_n\}$ is a random backward orbit and $\phi \in C(\mathbb{C})$. As in [10, 20], this result is the key tool in the proof of our main theorem.

Let $Y$ be a compact metric space and let $\mathcal{P}(Y)$ be the space of Borel probability measures on $Y$, noting that $\mathcal{P}(Y)$ is a compact metric space in the topology of weak* convergence. Suppose there exists a continuous map $Y \rightarrow \mathcal{P}(Y)$ assigning to each $x \in Y$ a measure $\mu_x$. The corresponding Markov operator $H : C(Y) \rightarrow C(Y)$ is given by

$$Hf(x) = \int f(y) \, d\mu_x(y).$$

Suppose the stochastic process $\{ X_n : n = 0, 1, 2, \ldots \}$ is a Markov process with state space $Y$ corresponding to $H$, i.e.,

$$P(\{ X_{n+1} \in A | X_0, X_1, \ldots, X_n \}) = \mu_{X_n}(A).$$

Given these assumptions we then have the following, which is a weaker version (but sufficient for our purposes) of Theorem 1.4 in [9].

**Theorem A.1 ([9]).** Assume that there is a unique probability measure $\nu$ on $Y$ that is invariant under the adjoint operator $H^*$ on $\mathcal{P}(Y)$ and
let $\phi \in C(Y)$. Then with probability one

$$\frac{1}{N + 1} \sum_{n=0}^{N} \phi(X_n) \to \int \phi d\nu$$

as $N \to \infty$.

**Proof of Theorem 3.1.** Let $\phi \in C(\hat{C})$. We apply Theorem A.1 using $Y = \hat{C}$, $\mu_x = \mu^b_x$, $H = T$, $\nu = \mu^b$ and $X_n = Z_n$, noting that all the hypotheses have been met by Theorem 2.2 and Claim 2. Thus we obtain a set $\Sigma_{\phi} \subseteq \Sigma_d^+$ with $P_b(\Sigma_{\phi}) = 1$ such that for all $(i_1, i_2, \ldots) \in \Sigma_{\phi}$ we have

$$\langle \phi, \mu^a_{i_1, \ldots, i_n} \rangle = \frac{1}{n} \sum_{j=1}^{n} \phi(z_{i_1, i_2, \ldots, i_j}) \to \int \phi d\mu^b = \langle \phi, \mu^b \rangle.$$

Since the set $\Sigma_{\phi}$ depends on $\phi$, we require an extra step to achieve single such set to work for all maps in $C(\hat{C})$.

Since $\hat{C}$ is compact we know that $C(\hat{C})$ is separable. Let $\{\phi_j\}$ be dense in $C(\hat{C})$. Let $\Sigma_0 = \cap_{j=1}^{\infty} \Sigma_{\phi_j}$ and note that $P_b(\Sigma_0) = 1$. Let $\psi \in C(\hat{C})$. Let $\epsilon > 0$. Select $\phi_j$ such that $\|\psi - \phi_j\|_{\infty} < \epsilon$. Then for all $(i_1, i_2, \ldots) \in \Sigma_0$, we have

$$\left| \frac{1}{n} \sum_{j=1}^{n} \psi(z_{i_1, i_2, \ldots, i_j}) - \frac{1}{n} \sum_{j=1}^{n} \phi_j(z_{i_1, i_2, \ldots, i_j}) \right| \leq \frac{1}{n} \sum_{j=1}^{n} \phi_j(z_{i_1, i_2, \ldots, i_j}) - \frac{1}{n} \sum_{j=1}^{n} \phi_j(z_{i_1, i_2, \ldots, i_j})$$

$$+ \frac{1}{n} \sum_{j=1}^{n} \phi_j(z_{i_1, i_2, \ldots, i_j}) - \int_{\hat{C}} \phi_j \mu^b + \int_{\hat{C}} \psi \mu^b - \int_{\hat{C}} \psi \mu^b$$

$$< \epsilon + \frac{1}{n} \sum_{j=1}^{n} \phi_j(z_{i_1, i_2, \ldots, i_j}) - \int_{\hat{C}} \phi_j \mu^b + \int_{\hat{C}} \psi \mu^b + \epsilon < 3\epsilon$$

for large $n$.

Hence on $\Sigma_0$ we get the convergence we seek, for all $\psi \in C(\hat{C})$ and this completes the proof. \qed

**References**


BACKWARD ITERATION ALGORITHMS FOR MÖBIUS JULIA SETS 15


(Rich Stankewitz) DEPARTMENT OF MATHEMATICAL SCIENCES, BALL STATE UNIVERSITY, HTTP://RSTANKEWITZ.IWEB.BSU.EDU/
E-mail address: rstankewitz@bsu.edu

(Hiroki Sumi) DEPARTMENT OF MATHEMATICS, GRADUATE SCHOOL OF SCIENCE, OSAKA UNIVERSITY, 1-1, MACHIKANEYAMA, TOYONAKA, OSAKA, 560-0043, JAPAN, HTTP://WWW.MATH.SCI.OSAKA-U.AC.JP/~SUMI/WELCOMEOU-E.HTML
E-mail address: sumi@math.sci.osaka-u.ac.jp