NONLINEAR AUTOMORPHISMS OF PLANE DOMAINS

TIMO ERKAMA

Abstract. We prove that the number of holomorphic nonlinear polynomials mapping a plane domain one-to-one onto itself is at most countable.

1. Introduction

Several symmetric plane domains have a nondiscrete group of automorphisms defined by holomorphic polynomials of degree one. Examples of such domains are the whole complex plane $\mathbb{C}$, a disk, a strip and an annulus. A strip has a continuous group of automorphisms defined by monic polynomials whereas the unit disk $U = \{ z \in \mathbb{C} \mid |z| < 1 \}$ and the upper half plane $H = \{ z \in \mathbb{C} \mid \text{Im} z > 0 \}$ have a one-dimensional group of automorphisms defined by homogeneous polynomials of degree one. A complete characterization of domains with the above property appears in [4].

Some nonlinear polynomials can also map a planar domain one-to-one onto itself. For example, if a nonlinear polynomial is holomorphically conjugate to a rotation in a neighborhood of a neutral fixed point, then this point is the center of a Siegel disk $D$ and the polynomial generates a nondiscrete semigroup of holomorphic automorphisms of $D$. Another example of automorphisms defined by nonlinear polynomials is studied in Section 5.

For any subdomain $D$ of $\mathbb{C}$ let $\text{Pol}(D)$ be the set of all holomorphic polynomials mapping $D$ one-to-one onto itself. Then $\text{Pol}(D)$ is a semigroup with a topology induced by the Lie group $\text{Aut}(D)$ of all holomorphic automorphisms of $D$.

**Theorem 1.** For each $k \geq 2$ the set $\{ P \in \text{Pol}(D) \mid \deg P = k \}$ is discrete.

**Corollary 1.** $\text{Pol}(D)$ contains at most countably many nonlinear polynomials.

Corollary 1 is an immediate consequence of Theorem 1 because $\text{Aut}(D)$ is second countable. Hence $\text{Pol}(D)$ can be uncountable only if it has a nondiscrete subgroup of linear polynomials; then $D$ is one of the symmetric domains described in [4].

If $\text{Aut}(D)$ is not discrete, then $D$ is either simply or doubly connected. In both cases each nonlinear element of $\text{Pol}(D)$ is conformally conjugate to an elliptic, parabolic or hyperbolic Möbius transformation. In the elliptic case $D$ is an invariant subdomain of a Siegel disk; the example in Section 5 deals with the hyperbolic case.

The proof of Theorem 1 is based in the study of polymorphisms introduced in [6]. A polymorphism of a nonconstant holomorphic function $f : H \to \mathbb{C}$ is a pair $(\phi, P)$ of holomorphic polynomials such that $\phi \in \text{Pol}(H)$ and $f \circ \phi = P \circ f$. The...
set \( \Pi(f) \) of all polymorphisms of \( f \) is a topological semigroup with the topology induced by the map \((\phi, P) \mapsto \phi \) from \( \Pi(f) \) to \( \text{Pol}(H) \). In [6] we proved

**Theorem 2.** For each \( k \geq 2 \) the set \( \{(\phi, P) \in \Pi(f) \mid \deg P = k\} \) is discrete unless \( f \) is the composition of the exponential function \( e^z \) and two linear polynomials.

The property of Theorem 2 can be used to characterize the exponential function also in the whole complex plane \( C \) [5, Theorem 2].

We assume that the reader is familiar with complex analytic geometry; this will be needed in the study of some complex analytic subsets of \( \text{Aut}(C \cup \{\infty\}) \). For the terminology we refer to [2].

It is well known that all nonelementary groups of Möbius transformations contain loxodromic elements. We start with a corresponding result for semigroups which we shall need in Section 4.

---

### 2. Semigroups of \( SL(2, \mathbb{R}) \)

Let \( SL(2, \mathbb{R}) \) be the multiplicative Lie group of real \( 2 \times 2 \) matrices with determinant one. An element of \( SL(2, \mathbb{R}) \) is hyperbolic if it has two real distinct eigenvalues.

**Lemma 1.** Suppose that a subset \( \Gamma \subset SL(2, \mathbb{R}) \) is a multiplicative semigroup. Then either \( \Gamma \) is commutative or \( \Gamma \) contains a hyperbolic element.

**Proof.** Suppose that \( \Gamma \) contains two elements \( A_1 \) and \( A_2 \) which do not commute. We have to prove that \( \Gamma \) contains a hyperbolic element.

Recall that an element \( A \in SL(2, \mathbb{R}) \) is hyperbolic if and only if the trace of \( A \) satisfies \(|\text{tr } A| > 2\). If \( A \) is not hyperbolic, then either \( A \) or \(-A\) is conjugate to

\[
\begin{pmatrix}
1 & b \\
0 & 1
\end{pmatrix}
\quad \text{or} \quad
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}.
\]

Suppose first that \( A_1 \) has exactly one real eigenvalue and that \( A_2 \) is not hyperbolic. By conjugation (and if necessary replacing \( A_1 \) by \( A_1^2 \)) we may assume that \( A_1 = I + be_1e_2^T \) where \( e_1 \) and \( e_2 \) are the first and second columns of the identity matrix \( I \), respectively. Then for each positive integer \( k \) the binomial theorem implies that \( A_1^kA_2 = A_2 + kbe_1e_2^TA_2 \). Moreover, the trace of \( be_1e_2^TA_2 \) is nonzero, because \( A_1 \) and \( A_2 \) do not commute and \( A_2 \) is not hyperbolic. We conclude that \(|\text{tr } (A_1^kA_2)| > 2\) if \( k \) is large enough, so that \( A_1^kA_2 \) is hyperbolic. It remains to consider the case when \( A_1 \) and \( A_2 \) have no real eigenvalues. Then \( A_1 \) and \( A_2 \) are both conjugate to a matrix of the form

\[
\begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix},
\]

so that \( A_1^{-1} \) and \( A_2^{-1} \) can be approximated by positive powers of \( A_1 \) and \( A_2 \), respectively. More precisely, for \( i = 1, 2 \) every neighborhood of \( A_i^{-1} \) contains a power of \( A_i \). Since \( A_1 \) and \( A_2 \) do not commute, their commutator \( A_1A_2A_1^{-1}A_2^{-1} \) is hyperbolic [3, Lemma 3.2] and can be approximated by elements of \( \Gamma \) of the form \( A_1A_2A_1^kA_2^l \) for suitable positive integers \( k \) and \( l \). Since the set of hyperbolic elements is open in \( SL(2, \mathbb{R}) \), we conclude that \( A_1A_2A_1^kA_2^l \) is hyperbolic for some \( k, l \). This completes the proof of Lemma 1. \( \square \)
We finally mention that there is an epimorphism from $SL(2, \mathbb{R})$ to the complex analytic Lie group $Aut(C \cup \{\infty\})$ of holomorphic automorphisms of the Riemann sphere. The image of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ under this epimorphism will be denoted by $A_*$; then

$$A_*(z) = \frac{az+b}{cz+d} \quad \text{for each} \quad z \in C \cup \{\infty\}.$$ 

3. Holomorphic families of functional equations

For each subdomain $G \subset C$ let $H(G)$ denote the algebra of holomorphic functions of $G$.

Let $N$ be a complex analytic subset of $Aut(C \cup \{\infty\})$, and let $f \in H(U)$ be nonconstant. We say that a map $\beta: N \rightarrow H(f(U))$ is a holomorphic family of functional equations of $f$ if there is a subdomain $V \subset U$ such that $\phi(V) \subset U$ for each $\phi \in N$ and

$$(1) \quad f(\phi(z)) = \beta(\phi)(f(z))$$

for each $(\phi, z) \in N \times V$.

**Theorem 3.** Suppose that $\beta: N \rightarrow H(f(U))$ is a holomorphic family of functional equations of a nonconstant holomorphic function $f \in H(U)$, and suppose that $N$ has an accumulation point $\phi_0 \in N$ such that $\phi_0(U) = U$. Then $f$ has a holomorphic extension to all but at most two points of $\partial U$.

The function $f(z) = \log \frac{1+z}{1-z}$ provides an example of a situation where the hypotheses of Theorem 3 are satisfied but $f$ fails to have a holomorphic extension to two points of $\partial U$. In this case $N$ consists of Möbius transformations of the form $\phi_\alpha(z) = \frac{z-\alpha}{1-\alpha z}$ where $|\alpha| < \frac{1}{2}$, and $\beta(\phi_\alpha)(w) = w - f(\alpha)$ for each $w \in f(U)$.

**Proof.** For each $\phi \in N$ we can define a holomorphic extension of $f$ to $U \cup \phi(U)$ such that (1) holds for each $z \in U$. By replacing $N$ with a sufficiently small subspace of $N$ we may assume that for each $\phi \in N$ the closure of $\phi(U)$ contains the origin but does not contain the point at infinity. Then for each $\omega \in \partial U$ there is a holomorphic function $\Phi_\omega: N \rightarrow C$ such that

$$\Phi_\omega(\phi) = \phi(\omega) \quad (\phi \in N).$$

Let $V$ be the set of all points $\omega \in \partial U$ such that $\Phi_\omega$ is locally constant at $\phi_0$. Then $V$ contains at most two points, because elements of $N$ are uniquely determined by their values at any three distinct points.

Let $\omega_1$, $\omega_2$, and $\omega_3$ be three distinct points of $\partial U$ such that $\omega_1 \in \phi_0(V)$ if $V$ is not empty. It suffices to prove that $f$ has a holomorphic extension to either $\omega_2$ or $\omega_3$.

If $V$ is empty, choose $\omega \in \partial U$ such that $\phi_0(\omega) = \omega_2$. Since $\Phi_\omega$ is not locally constant at $\phi_0$, by the maximum principle [3] p. 234 the image of $\Phi_\omega$ contains an open neighborhood of $\phi(\omega_0) = \omega_2$. Thus there exists $\phi \in N$ and $\omega_0 \in \partial U$ such that $|\phi(\omega_0)| > 1$ and $\arg \phi(\omega_0) = \arg \omega_2$. Since the closure of $\phi(U)$ contains $\phi(\omega_0)$ and the origin, by convexity the point $\omega_2$ on the line segment joining $\phi(\omega_0)$ and 0 is contained in $\phi(U)$. Hence $f$ has a holomorphic extension to $\omega_2$.

If $V$ is not empty, we choose $h \in Aut(C \cup \{\infty\})$ such that $h(U) = H$, $h(\infty) = -i$ and $h(\omega_1) = \infty$. Let $x$ be any point between $h(\omega_2)$ and $h(\omega_3)$ on the real axis such
that \( h^{-1}(x) \not\in \phi_0(V) \), and let \( \omega = \phi_0^{-1}(h^{-1}(x)) \). Then \( \Phi_\omega \) is not locally constant at \( \phi_0 \), so that again by the maximum principle \( \Phi_\omega \) maps every neighborhood of \( \phi_0 \) to a neighborhood of \( h^{-1}(x) \). Thus there exists \( \phi \in N \) such that \( h(\phi(\omega)) \) is an interior point of the triangle \( T \) with vertices at \( h(\omega_2), h(\omega_3) \) and \( -i \), and by choosing \( \phi \) close to \( \phi_0 \) we may assume that \( \phi^{-1}(\omega_1) = \phi_0^{-1}(\omega_1) \), because \( \omega_1 \in \phi_0(V) \). Then \( h(\phi(U)) \) is a half plane containing at least one vertex of \( T \), because the point \( h(\phi(\omega)) \) on the boundary of \( h(\phi(U)) \) is an interior point of \( T \). On the other hand, the vertex \( -i \) of \( T \) is not contained in \( h(\phi(U)) \), so that either \( h(\omega_2) \) or \( h(\omega_3) \) is a point of \( h(\phi(U)) \). Thus either \( \omega_2 \) or \( \omega_3 \) is contained in \( \phi(U) \), so that \( f \) has a holomorphic extension to \( \omega_2 \) or \( \omega_3 \). This completes the proof of Theorem [3].


Suppose that the set \( \{ P \in \text{Pol}(D) \mid \deg P = k \} \) is not discrete for some \( k \); we have to prove that \( k = 1 \). Since \( \text{Pol}(D) \) is a subset of \( \text{Aut}(D) \), it follows that \( \text{Aut}(D) \) is not discrete. It is well known that \( \text{Aut}(D) \) can be nondiscrete only if \( D \) is either simply or doubly connected [10]. Thus we may assume that there exists a holomorphic isomorphism \( f : G \to D \) where \( G \) is either the whole plane, a punctured plane, a punctured disk, or an annulus.

Let \( N_0 \) be the set of all \( \phi \in \text{Aut}(G) \) such that \( f \circ \phi \circ f^{-1} \) is the restriction of a polynomial of degree \( \leq k \); then \( N_0 \) is not discrete in \( \text{Aut}(G) \). Let us first consider the case when the semigroup \( \langle N_0 \rangle \) generated by \( N_0 \) is commutative.

The number of polynomials of degree \( k \) commuting with a given nonlinear polynomial is finite [7]. Since \( \langle N_0 \rangle \) is commutative, it follows that \( f \circ \phi \circ f^{-1} \) can be the restriction of a nonlinear polynomial only for finitely many \( \phi \in N_0 \). Since \( N_0 \) is not discrete, we conclude that \( k = 1 \). Thus we may assume that \( \langle N_0 \rangle \) is not commutative.

If \( G \) is the whole complex plane \( \mathbb{C} \), the same is true of \( D \) and there is nothing to prove. If \( G \) is a punctured plane, a punctured disk or an annulus, then \( f \) can be chosen so that \( f \) linearizes each element of \( \text{Pol}(D) \), i.e. \( f^{-1} \circ P \circ f \) is the restriction of a linear polynomial for each \( P \in \text{Pol}(D) \). In this case the component of the identity of \( \text{Aut}(G) \) is abelian and contains \( N_0 \), so that \( \langle N_0 \rangle \) is commutative. Hence it remains to consider the case when \( G \) is a disk, and we may of course assume that \( G \) is the open unit disk \( U \). Then each element of \( N_0 \) is the restriction of an element of \( \text{Aut}(\mathbb{C} \cup \{ \infty \}) \), and from now on we shall identify \( N_0 \) with the corresponding subset of \( \text{Aut}(\mathbb{C} \cup \{ \infty \}) \).

Let \( V = \{ z \in U \mid |z| < \frac{1}{2} \} \) and let \( \mathcal{M} = \{ \phi \in \text{Aut}(\mathbb{C} \cup \{ \infty \}) \mid \phi(V) \subset U \} \); then \( \mathcal{M} \) is open in \( \text{Aut}(\mathbb{C} \cup \{ \infty \}) \), and the map

\[
(\phi, w) \mapsto (f \circ \phi \circ f^{-1})(w)
\]

is holomorphic in \( \mathcal{M} \times f(V) \). The partial derivatives of this map with respect to \( w \) are also holomorphic, and for each fixed \( w \in f(V) \) we can define a holomorphic function \( F_w : \mathcal{M} \to \mathbb{C} \) such that

\[
F_w(\phi) = \frac{\partial^{k+1}}{\partial w^{k+1}} (f \circ \phi \circ f^{-1})(w)
\]

for each \( \phi \in \mathcal{M} \). Let

\[
N = \{ \phi \in \mathcal{M} \mid F_w(\phi) = 0 \text{ for each } w \in f(V) \};
\]

then \( N \) is a complex analytic subset of \( \mathcal{M} \) [2].
From the definition of $N$ it follows that for each $\phi \in N$ there is a polynomial $P_\phi$ of degree $\leq k$ such that

$$P_\phi(w) = (f \circ \phi \circ f^{-1})(w) \text{ for each } w \in f(V).$$

Then

$$f \circ \phi(z) = P_\phi \circ f(z)$$

for each $z \in V$. It is also clear that $N_0 \subset N$, so that $N$ has an accumulation point $\phi_0 \in N$ such that $\phi_0(U) = U$. Moreover, the map $\phi \mapsto P_\phi$ is a holomorphic family of functional equations of $f$. From Theorem 3 it follows that $f$ has a holomorphic extension to all but at most two points of $\partial U$.

Let $\Omega$ be the set of points $\omega \in \partial U$ such that $f$ does not have a holomorphic extension to $\omega$. It is clear that $\phi^{-1}(\Omega) \subset \Omega$ for each $\phi \in N_0$; in fact, it follows from (2) that if $f$ does not have a holomorphic extension to $\omega$, then $f$ does not have a holomorphic extension to $\phi^{-1}(\omega)$ either. But $\phi^{-1}(\Omega) \subset \Omega$ implies that

$$\Omega \subset \phi(\Omega),$$

and since $\Omega$ and $\phi(\Omega)$ have the same cardinality, we conclude that

$$\Omega = \phi(\Omega)$$

for each $\phi \in N_0$.

Let us first consider the case when $\Omega$ is empty, so that $f$ has a holomorphic extension to a domain $U_1$ containing the closure of $U$. Let $\Gamma$ be the set of all $A \in SL(2, \mathbb{R})$ such that $A_* = (N_0)$ where $A_*$ is defined as in Section 2. Then $\Gamma$ is a multiplicative semigroup. Since $(N_0)$ is not commutative, the same is true of $\Gamma$, and by Lemma 1 $\Gamma$ contains a hyperbolic element $A$.

Since $A_* \in (N_0)$, there exist $\phi_1, \ldots, \phi_n \in N_0$ such that $A_* = \phi_1 \circ \cdots \circ \phi_n$, and by recalling the definition of $N_0$ we see that $f \circ A_* \circ f^{-1}$ is the restriction of a polynomial $P$. Therefore

$$f \circ A_* = P \circ f$$

in $U$. By analytic continuation this equation defines a holomorphic extension of $f$ to $A_*(U_1)$, $(A^2)_*(U_1)$, and by induction, to $(A^n)_*(U_1)$ for each positive integer $n$. But since $A$ is hyperbolic, these sets cover the whole Riemann sphere. However, this is not possible because all holomorphic functions of $C \cup \{\infty\}$ are constant. We conclude that $\Omega$ is not empty.

Since $\Omega$ contains at most two points, it follows from (3) that $\phi(\phi(\omega)) = \omega$ for each $\phi \in N_0$ and each $\omega \in \Omega$. Choose $h \in Aut(C \cup \{\infty\})$ such that $h(U) = H$ and $h^{-1}(\infty) \in \Omega$. Then $h \circ \phi \circ \phi \circ h^{-1} \in Pol(H)$ for each $\phi \in N_0$, and (2) implies that the set

$$(4) \quad \{(h \circ \phi \circ \phi \circ h^{-1}, P_\phi \circ P_\phi) \mid \phi \in N_0\}$$

is a subset of $\Pi(f \circ h^{-1})$. Moreover, $h \circ \phi \circ \phi \circ h^{-1}$ can be the identity of $Pol(H)$ only if $\deg P_\phi = 1$. It follows that either $k = 1$ or $(h \circ \phi_0 \circ \phi \circ h^{-1}, P_{\phi_0} \circ P_\phi)$ is an accumulation point of (4). Also, $f \circ h^{-1}$ cannot be the composite of $e^z$ and two linear polynomials, because $f \circ h^{-1}$ is univalent in $H$. In view of Theorem 2 we conclude that $k = 1$. The proof of Theorem 4 is now complete.
5. AN EXAMPLE

In this section we construct a plane domain where the quadratic polynomial

\[ P(z) = z^2 + \frac{1}{2}z \]

is holomorphically conjugate to a hyperbolic Möbius transformation. The example shows that a discrete group generated by a holomorphic polynomial need not be a Kleinian group.

Let \( g \) be an entire function linearizing \( P \) at the repelling fixed point \( z = \frac{1}{2} \). Then \( g(0) = \frac{1}{2} \),

\[ g\left(\frac{3}{2}z\right) = P(g(z)) \tag{5} \]

for each \( z \in \mathbb{C} \), and we may assume that \( g'(0) = -1 \). Such linearizing maps have been studied in detail by Myrberg [9].

Let \( T \) be the closed triangle with vertices at \( \frac{1}{2} \) and \( -\frac{1}{16}(1 \pm i) \). A straightforward computation shows that \( T \) is forward invariant under \( P \), i.e. \( P \) maps every point of \( T \) into a point of \( T \). A study of the branches of \( P^{-1} \) shows also that \( P \) is one-to-one in \( T \).

Since \( g \) is conformal at the origin and \( g'(0) = -1 \), there is \( \delta > 0 \) such that \( g(re^{i\theta}) \in T \) if \( 0 < r < \delta \) and \( |\theta| < \delta \). Since \( T \) is forward invariant, iteration of (5) shows that \( g \) maps the domain

\[ \{ \zeta \in \mathbb{C} \mid |\arg \zeta| < \delta \} \]

onto a subdomain \( D \) of \( T \). Moreover, \( P \) maps \( D \) onto itself, so that \( P \in \text{Pol}(D) \).

Figure 1 indicates the shape of \( D \) as well as some of the orbits and streamlines invariant under \( P \). Note that \( P \) generates a discrete group of automorphisms of \( D \). A result of Azarina [1] implies that for examples of this kind the boundary of \( D \) cannot be an analytic curve; accordingly the boundary in Figure 1 is not smooth at the fixed points of \( P \).

\[ \text{Figure 1.} \]
For this example it is not hard to prove that $\text{Pol}(D)$ is commutative. It is an open question whether for some other domain $\text{Pol}(D)$ could contain two nonlinear elements which do not commute.

REFERENCES


University of Joensuu, SF-80101 Joensuu, Finland
E-mail address: Timo.Erkama@joensuu.fi