CONFORMAL GEOMETRY AND DYNAMICS
An Electronic Journal of the American Mathematical Society
Volume 3, Pages 151–157 (November 30, 1999)
S 1088-4173(99)00051-X

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 $\mathbf{C}$, a disk, a strip and an annulus. A strip has a continuous group of automorphisms defined by monic polynomials whereas the unit disk $U = \{ z \in \mathbf{C} \mid |z| < 1 \}$ and the upper half plane $H = \{ z \in \mathbf{C} \mid \operatorname{Im} z > 0 \}$ 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 $\mathbf{C}$ let $\operatorname{Pol}(D)$ be the set of all holomorphic polynomials mapping D one-to-one onto itself. Then $\operatorname{Pol}(D)$ is a semigroup with a topology induced by the Lie group $\operatorname{Aut}(D)$ of all holomorphic automorphisms of D.

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

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

Corollary 1 is an immediate consequence of Theorem 1, because $\operatorname{Aut}(D)$ is second countable. Hence $\operatorname{Pol}(D)$ 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 $\operatorname{Aut}(D)$ is not discrete, then D is either simply or doubly connected. In both cases each nonlinear element of $\operatorname{Pol}(D)$ 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 $f : H \to \mathbf{C}$ is a pair $(\phi,\ P)$ of holomorphic polynomials such that $\phi \in \operatorname{Pol}(H)$ and $f \circ \phi\ = P \circ f$. The

Received by the editors June 23, 1999 and, in revised form, September 24, 1999.
1991 *Mathematics Subject Classification.* Primary 30D05; Secondary 58F23.

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 $\mathrm{Pol}(H)$. In [6] we proved

Theorem 2. *For each $k\ge 2$ 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 $\mathbf{C}$ [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 $\mathrm{Aut}(\mathbf{C}\cup\{\infty\})$. 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 $SL(2,\,\mathbf{R})$

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

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

Proof. Suppose that Γ contains two elements A_1 and A_2 which do not commute. We have to prove that Γ contains a hyperbolic element.

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

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

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 $A_1=I+be_1{e_2}^T$ 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 $A_1^kA_2=A_2+kbe_1{e_2}^TA_2$. Moreover, the trace of $be_1{e_2}^TA_2$ is nonzero, because A_1 and A_2 do not commute and A_2 is not hyperbolic. We conclude that $|\mathrm{tr}\ ({A_1}^kA_2)|>2$ 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 $A_1A_2{A_1}^{-1}{A_2}^{-1}$ is hyperbolic [3, Lemma 3.2] and can be approximated by elements of Γ of the form $A_1A_2{A_1}^k{A_2}^l$ for suitable positive integers k and l. Since the set of hyperbolic elements is open in $SL(2,\,\mathbf{R})$, we conclude that $A_1A_2{A_1}^k{A_2}^l$ is hyperbolic for some $k,\ l$. This completes the proof of Lemma 1. $\square$

We finally mention that there is an epimorphism from $SL(2, \mathbf{R})$ to the complex analytic Lie group $\mathrm{Aut}(\mathbf{C} \cup \{\infty\})$ 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

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

3. Holomorphic families of functional equations

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

Let N be a complex analytic subset of $\mathrm{Aut}(\mathbf{C} \cup \{\infty\})$, and let $f \in H(U)$ be nonconstant. We say that a map $\beta\colon N \to H(f(U))$ 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

$$f(\phi(z)) = \beta(\phi)(f(z)) \tag{1}$$

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

Theorem 3. *Suppose that $\beta\colon N \to H(f(U))$ 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 ∂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 ∂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 $\Phi_\omega : N \to \mathbf{C}$ such that

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

Let V be the set of all points $\omega \in \partial U$ such that Φ_ω is locally constant at ϕ_0. Then V contains at most two points, because elements of N are uniquely determined by their values at any three distinct points.

Let ω_1, ω_2 and ω_3 be three distinct points of ∂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 ω_2 or ω_3.

If V is empty, choose $\omega \in \partial U$ such that $\phi_0(\omega) = \omega_2$. Since Φ_ω is not locally constant at ϕ_0, by the maximum principle [8, p. 234] the image of Φ_ω contains an open neighborhood of $\phi_0(\omega) = \omega_2$. 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 ω_2 on the line segment joining $\phi(\omega_0)$ and 0 is contained in $\phi(U)$. Hence f has a holomorphic extension to ω_2.

If V is not empty, we choose $h \in \mathrm{Aut}(\mathbf{C} \cup \{\infty\})$ 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 $h^{-1}(x) \notin \phi_0(V)$, and let $\omega = {\phi_0}^{-1}(h^{-1}(x))$. Then Φ_ω is not locally constant at ϕ_0, so that again by the maximum principle Φ_ω maps every neighborhood of ϕ_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 ϕ close to ϕ_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 ω_2 or ω_3 is contained in $\phi(U)$, so that f has a holomorphic extension to ω_2 or ω_3. This completes the proof of Theorem 3. $\square$

4. Proof of Theorem 1

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

Let N_0 be the set of all $\phi \in \mathrm{Aut}(G)$ such that $f \circ \phi \circ f^{-1}$ is the restriction of a polynomial of degree $\le k$; then N_0 is not discrete in $\mathrm{Aut}(G)$. 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 $\mathbf{C}$, 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 $\mathrm{Pol}(D)$, i.e. $f^{-1} \circ P \circ f$ is the restriction of a linear polynomial for each $P \in \mathrm{Pol}(D)$. In this case the component of the identity of $\mathrm{Aut}(G)$ 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 $\mathrm{Aut}(\mathbf{C} \cup \{\infty\})$, and from now on we shall identify N_0 with the corresponding subset of $\mathrm{Aut}(\mathbf{C} \cup \{\infty\})$.

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

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

is holomorphic in $\mathfrak{M} \times f(V)$. 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 $F_w : \mathfrak{M} \to \mathbf{C}$ such that

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

for each $\phi \in \mathfrak{M}$. Let

$$N = \{\, \phi \in \mathfrak{M} \mid F_w(\phi) = 0 \text{ for each } w \in f(V) \,\};$$

then N is a complex analytic subset of $\mathfrak{M}$ [2].

From the definition of N it follows that for each $\phi \in N$ there is a polynomial P_ϕ of degree $\le k$ such that

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

Then

$$f \circ \phi(z) = P_\phi \circ f(z) \tag{2}$$

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 ∂U.

Let Ω be the set of points $\omega \in \partial U$ such that f does not have a holomorphic extension to ω. 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 ω, 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 Ω and $\phi(\Omega)$ have the same cardinality, we conclude that

$$\Omega = \phi(\Omega) \tag{3}$$

for each $\phi \in N_0$.

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

Since $A_* \in \langle N_0 \rangle$, there exist $\phi_1, \dots, \phi_n \in N_0$ 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 $\mathbf{C} \cup \{\infty\}$ are constant. We conclude that Ω is not empty.

Since Ω 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 $h \in \mathrm{Aut}(\mathbf{C} \cup \{\infty\})$ such that $h(U) = H$ and $h^{-1}(\infty) \in \Omega$. Then $h \circ \phi \circ \phi \circ h^{-1} \in \mathrm{Pol}(H)$ for each $\phi \in N_0$, and (2) implies that the set

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

is a subset of $\Pi(f \circ h^{-1})$. Moreover, $h \circ \phi \circ \phi \circ h^{-1}$ can be the identity of $\mathrm{Pol}(H)$ only if $\deg P_\phi = 1$. It follows that either $k = 1$ or $(h \circ \phi_0 \circ \phi_0 \circ h^{-1},\, P_{\phi_0} \circ P_{\phi_0})$ 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 1 is now complete.

5. An example

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

$$P(z) = z^2 + \tfrac{1}{2}z$$

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 $g(0) = \frac{1}{2}$,

$$g(\tfrac{3}{2}z) = P(g(z)) \tag{5}$$

for each $z \in \mathbf{C}$, 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

$$\{\zeta \in \mathbf{C} \mid |\arg \zeta| < \delta\}$$

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

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.

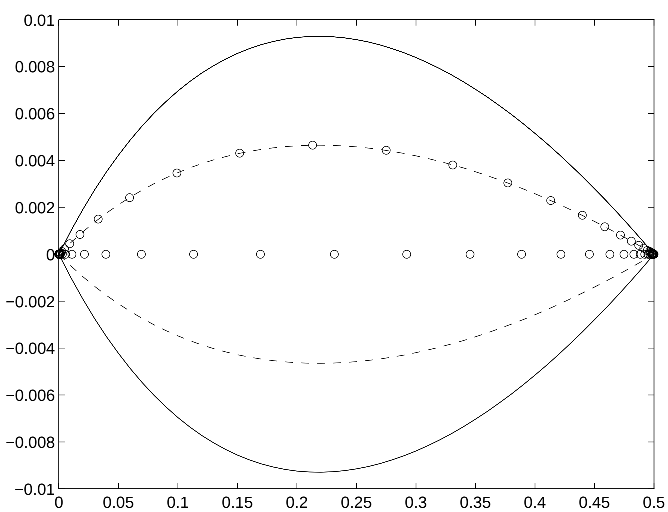

Figure 1.

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

References

1. Yu. V. Azarina, *Invariant analytic curves for entire functions*, Siberian Math. J. **30** (1989), no. 3, 349–353. MR **90j:**30038
2. E. M. Chirka, *Complex Analytic Sets*, Kluwer, 1989. MR **92b:**32016
3. T. Erkama, *Group actions and extension problems for maps of balls*, Ann. Acad. Sci. Fenn. Ser. A I Math. **556** (1973), 1–31. MR **57:**624
4. ———, *Möbius automorphisms of plane domains*, Ann. Acad. Sci. Fenn. Ser. A I Math. **10** (1985), 155–162. MR **87c:**30063
5. ———, *The exponential function and linearization of quadratic polynomials*, Analysis, Algebra and Computers in Mathematical Research, Lecture Notes in Pure and Applied Mathematics, Marcel Dekker, 1994, pp. 71–80. MR **95g:**39028
6. ———, *Polymorphisms and linearization of nonlinear polynomials*, Ann. Acad. Sci. Fenn. Ser. A I Math. **22** (1997), 113–121. MR **97j:**30008
7. E. Jacobsthal, *Über vertauschbare Polynome*, Math. Z. **63** (1955), 243–276. MR **17:**574a
8. S. Łojasiewicz, *Introduction to Complex Analytic Geometry*, Birkhäuser, 1991. MR **92g:**32002
9. P. J. Myrberg, *Über ganze Funktionen mit rationalem Multiplikationstheorem*, Ann. Acad. Sci. Fenn. Ser. A I Math. **414** (1968), 1–20. MR **37:**4256
10. R. Nevanlinna, *Uniformisierung*, Springer, 1953. MR **15:**208h

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