Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)
Vol. XVII (2017), 1417-1446

Polynomial semiconjugacies, decompositions of iterations, and invariant curves

FEDOR PAKOVICH

Abstract. We study the functional equation $A \circ X = X \circ B$, where A, B, and X are polynomials with complex coefficients. Using results of [13] about polynomials sharing preimages of compact sets in $\mathbb{C}$, we show that for given B its solutions may be described in terms of the filled-in Julia set of B. On this base, we prove a number of results describing a general structure of solutions. The results obtained imply in particular the result of Medvedev and Scanlon [10] about invariant curves of maps $F : \mathbb{C}^2 \to \mathbb{C}^2$ of the form $(x, y) \to (f(x), f(y))$, where f is a polynomial, and a version of the result of Zieve and Müller [22] about decompositions of iterations of a polynomial.

Mathematics Subject Classification (2010): 37F10 (primary); 14H99 (secondary).

1. Introduction

Let A and B be rational functions of degree at least two on the Riemann sphere. The functions A and B are called commuting if

$$A \circ B = B \circ A, \tag{1.1}$$

and conjugate if

$$A \circ X = X \circ B \tag{1.2}$$

for some rational function X of degree one.

If (1.2) is satisfied for some rational function X of degree at least two, the function B is called semiconjugate to A, and the function X is called a semiconjugacy from B to A. Unlike conjugation, semiconjugation is not an equivalency relation. We will use the notation $A \leq B$ if for given rational functions A and B there exists a non-constant rational function X such that (1.2) holds, and the notation $A \underset{X}{\leq} B$ if A,B, and X satisfy (1.2). The notation reflects the fact that the binary relation

Received October 30, 2015; accepted in revised form August 18, 2016.
Published online December 2017.

on the set of rational functions defined by equality (1.2) is a preorder. Indeed, if $A \underset{X}{\leq} B$ and $B \underset{Y}{\leq} C$ then $A \underset{X\circ Y}{\leq} C$.

Both equations (1.1) and (1.2) have "obvious" solutions. Namely, equation (1.1) has solutions of the form

$$A = R^{\circ m}, \quad B = R^{\circ n}, \tag{1.3}$$

where R is an arbitrary rational function and $m, n \geq 1$. Notice that such A and B have an iteration in common, that is

$$A^{\circ n} = B^{\circ m}. \tag{1.4}$$

In order to obtain solutions of equation (1.2) we can take arbitrary rational functions A_1, B_1 and set

$$F = A_1 \circ B_1, \quad G = B_1 \circ A_1.$$

Then the equality

$$(A_1 \circ B_1) \circ A_1 = A_1 \circ (B_1 \circ A_1) \tag{1.5}$$

implies that $F \underset{A_1}{\leq} G$. Similarly, $G \underset{B_1}{\leq} F$. Moreover, if now A_2, B_2 are rational functions such that the equality

$$G = A_2 \circ B_2 \tag{1.6}$$

holds, then the function $H = B_2 \circ A_2$ satisfies $G \underset{A_2}{\leq} H$ and $H \underset{B_2}{\leq} G$, implying that $F \underset{A_1\circ A_2}{\leq} H$ and $H \underset{B_2\circ B_1}{\leq} F$. This motivates the following definition of an equivalency relation on the set of rational functions: $F \sim G$ if there exist rational functions A_i, B_i, $1 \leq i \leq n$, such that

$$F = A_1 \circ B_1, \qquad G = B_n \circ A_n,$$

and

$$B_i \circ A_i = A_{i+1} \circ B_{i+1}, \quad 1 \leq i \leq n-1.$$

Clearly, $F \sim G$ implies that $F \leq G$ and $G \leq F$. Notice that, since for any rational function X of degree one the equality

$$A = (A \circ X) \circ X^{-1}$$

implies that $A \sim X^{-1} \circ A \circ X$, any equivalence class is a collection of conjugacy classes.

Functional equation (1.1) was first studied by Fatou, Julia, and Ritt in the papers [5,8], and [21]. In all these papers it was assumed that the considered commuting functions A and B have no iterate in common. Fatou and Julia described solutions of (1.1) under the additional assumption that the Julia set of A or B does not coincide with the whole complex plane, and Ritt investigated the general case. Briefly, the Ritt theorem states that if rational functions A and B commute and no iterate of A is equal to an iterate of B, then, up to a conjugacy, A and B are either powers, or Chebyshev polynomials, or Lattès functions. Another proof of the Ritt theorem was given by Eremenko in [4]. Notice however that a description of commuting A and B with a common iterate is known only in the polynomial case. Thus, in a certain sense the classification of commuting rational functions is not yet completed. On the other hand, it was shown by Ritt [19,21] that in the polynomial case equality (1.1) implies that, up to the change

$$A \to \lambda \circ A \circ \lambda^{-1}, \quad B \to \lambda \circ B \circ \lambda^{-1},$$

where λ is a polynomial of degree one, either

$$A = z^n, \quad B = \varepsilon z^m,$$

where $\varepsilon^n = \varepsilon$, or

$$A = \pm T_n, \quad B = \pm T_m,$$

or

$$A = \varepsilon_1 R^{\circ m}, \quad B = \varepsilon_2 R^{\circ n},$$

where $R = zS(z^\ell)$ for some polynomial S and ε_1, ε_2 are l-th roots of unity. In fact, this conclusion remains true if instead of (1.1) one were to assume only that A and B share a completely invariant compact set in $\mathbb{C}$ (see [13]).

Equation (1.2) was investigated in the recent paper [17]. The main result of [17] states that if a rational function B is semiconjugate to a rational function A, then either $A \sim B$, or A and B are "minimal holomorphic self-maps" between orbifolds of non-negative Euler characteristic on the Riemann sphere. The latter class of functions is a natural extension of the class of Lattès functions and admits a neat characterization. However, as with the description of commuting rational functions, the description of solutions of (1.2) given in [17] is not completely satisfactory, since it gives no information about equivalent rational functions. In particular, the results of [17] do not provide any bounds on the number of conjugacy classes in an equivalence class of a rational function B or more generally on the number of conjugacy classes of A such that $A \leq B$. Another related problem is the following: is it true that if conditions $A \leq B$ and $B \leq A$ hold simultaneously, then $A \sim B$? Finally, it would be desirable to obtain some handy structural descriptions of the totality of X satisfying (1.2) for given A and B, and of the totality of A satisfying $A \leq B$ for given B.

In this paper we study equation (1.2) with emphasis on the above questions in the case where all the functions involved are *polynomials*. Notice that in distinction with the general case, for polynomials there exists quite a comprehensive theory of functional decompositions developed by Ritt [20]. Nevertheless, questions regarding polynomial decompositions may be highly non-trivial, and a number of recent papers are devoted to such questions arising from different branches of mathematics. Let us mention for example the paper [22] with applications to algebraic dynamics [6], or the paper [16] with applications to differential equations [18]. Another example is the recent paper [10] about invariant varieties for dynamical systems defined by coordinatwise actions of polynomials, a considerable part of which concerns properties of polynomial solutions of (1.2).

The main distinction between this paper and the above mentioned papers is the systematical use of ideas and results from the paper [13] which relates polynomials sharing preimages of compact sets in $\mathbb{C}$ with the functional equation

$$A \circ C = D \circ B.$$

In particular, the main result of [13] leads to a characterization of polynomial solutions of (1.2) in terms of filled-in Julia sets. Recall that for a polynomial B the filled-in Julia set $K(B)$ is defined as the set of points in $\mathbb{C}$ whose orbits under iterations of B are bounded. Since equality (1.2) implies the equalities

$$A^{\circ n} \circ X = X \circ B^{\circ n}, \quad n \geq 1,$$

it it easy to see that if X is a semiconjugacy from B to A, then the preimage $X^{-1}(K(A))$ coincides with $K(B)$. We show that this property is in fact characteristic.

Theorem 1.1. *Let A, B and X be polynomials of degree at least two such that $A \underset{X}{\leq} B$. Then*

$$X^{-1}(K(A)) = K(B). \tag{1.7}$$

Conversely, if equality (1.7) *holds and* $\deg A = \deg B$, *then there exists a polynomial of degree one μ such that*

$$(\mu \circ A) \circ X = X \circ B$$

and $\mu(K(A)) = K(A)$. More generally, if for given B and X the condition

$$X^{-1}(K) = K(B) \tag{1.8}$$

holds for some compact set K in $\mathbb{C}$, then there exists a polynomial A such that $A \underset{X}{\leq} B$ and $K(A) = K$.

For a fixed polynomial B of degree at least two denote by $\mathcal{E}(B)$ the set of polynomials X of degree at least two such that $A \underset{X}{\leq} B$ for some polynomial A. An immediate corollary of Theorem 1.1 is that a polynomial X is contained in $\mathcal{E}(B)$ if and only if $K(B)$ is a union of fibers of X. Another corollary is that if $A \underset{X}{\leq} B$, then for any decomposition $X = X_1 \circ X_2$ there exists a polynomial C such that

$$A \underset{X_1}{\leq} C, \quad C \underset{X_2}{\leq} B.$$

Notice that in particular this casts the problem of the description of decompositions of iterations of a polynomial, first considered in the paper [22], into the context of equation (1.2). Indeed, since $B \circ B^{\circ d} = B^{\circ d} \circ B$, the polynomial $B^{\circ d}$ is contained in $\mathcal{E}(B)$ and hence for any decomposition $B^{\circ d} = Y \circ X$ the equalities

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B$$

hold for some polynomial A.

The following statement is another corollary of the main result of [13].

Theorem 1.2. *For any $X_1, X_2 \in \mathcal{E}(B)$ there exists $X \in \mathcal{E}(B)$ such that* $\deg X = \mathrm{LCM}(\deg X_1, \deg X_2)$ *and*

$$X = U_1 \circ X_1 = U_2 \circ X_2$$

for some polynomials U_1, U_2. Furthermore, there exists $W \in \mathcal{E}(B)$ such that $\deg W = \mathrm{GCD}(\deg X_1, \deg X_2)$ *and*

$$X_1 = V_1 \circ W, \quad X_2 = V_2 \circ W$$

for some polynomials V_1, V_2.

For fixed polynomials A, B denote by $\mathcal{E}(A, B)$ the subset of $\mathcal{E}(B)$ (possibly empty) consisting of polynomials X such that $A \underset{X}{\leq} B$. In particular, the set $\mathcal{E}(B, B)$ consists of polynomials of degree at least two commuting with B. We will call a polynomial P *special* if it is conjugate to z^n or $\pm T_n$, or equivalently if there exists a Möbius transformation μ which maps $K(P)$ to $\mathbb{D}$ or $[-1, 1]$. The following result describes a general structure of $\mathcal{E}(A, B)$ for non-special A, B.

Theorem 1.3. *Let A and B be fixed non-special polynomials of degree at least two such that the set $\mathcal{E}(A, B)$ is non-empty, and let X_0 be an element of $\mathcal{E}(A, B)$ of minimal degree. Then a polynomial X belongs to $\mathcal{E}(A, B)$ if and only if $X = \widetilde{A} \circ X_0$ for some polynomial $\widetilde{A}$ commuting with A.*

Notice that in a sense this result is a generalization of the result of Ritt about commuting polynomials. Indeed, applying Theorem 1.3 for $B = A$ and $X = B$, we obtain that if A is non-special and $B \in \mathcal{E}(A, A)$, then $B = \widetilde{A} \circ R$, where R is a polynomial of minimal degree in $\mathcal{E}(A, A)$. Now we can apply Theorem 1.3 again to the polynomial $\widetilde{A}$ and so on, eventually obtaining the representation $B = \mu_1 \circ R^{\circ m_1}$, where μ_1 is a polynomial of degree one commuting with A. In particular, since $A \in \mathcal{E}(A, A)$, the equality $A = \mu_2 \circ R^{\circ m_2}$ holds for some polynomial μ_2 of degree one commuting with A.

Another corollary of Theorem 1.3 is the following result obtained by Medvedev and Scanlon in [10]: if $\mathcal{C} \subset \mathbb{C}^2$ is an irreducible algebraic curve invariant under the map $F : (x, y) \to (f(x), f(y))$, where f is a non-special polynomial, then there exists a polynomial p which commutes with f such that $\mathcal{C}$ has the form $z_1 = p(z_2)$ or $z_2 = p(z_1)$. More generally, we prove the following statement which supplements the results of [10] about algebraic curves invariant under the map $F : (x, y) \to (f(x), g(y))$, where f and g are non-special polynomials.

Theorem 1.4. *Let f and g be non-special polynomials of degree at least two and $\mathcal{C}$ a curve in $\mathbb{C}^2$. Then $\mathcal{C}$ is an irreducible (f, g)-invariant curve if and only if $\mathcal{C}$ has the form $u(x) - v(y) = 0$, where u, v are polynomials of coprime degrees satisfying the equations*

$$t \circ u = u \circ f, \quad t \circ v = v \circ g \tag{1.9}$$

for some polynomial t.

Our next result describes the interrelations between the equivalence $\sim$, the preorder $\leq$, and decompositions of iterations.

Theorem 1.5. *Let A and B be polynomials of degree at least two. Then conditions $A \leq B$ and $B \leq A$ hold simultaneously if and only if $A \sim B$. Furthermore, $A \sim B$ if and only if there exist polynomials X, Y such that*

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B,$$

and $Y \circ X = B^{\circ d}$ for some $d \geq 0$.

For a fixed polynomial B of degree at least two denote by $\mathcal{F}(B)$ the set of polynomials A such that $A \leq B$. The following theorem gives a structural description of the set $\mathcal{F}(B)$.

Theorem 1.6. *Let B be a fixed non-special polynomial of degree $n \geq 2$. Then there exist $A \in \mathcal{F}(B)$ and a semiconjugacy X from B to A which are universal in the following sense: for any polynomial $C \in \mathcal{F}(B)$ there exist polynomials X_C, U_C such that $X = U_C \circ X_C$ and the diagram*

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow{B} & \mathbb{C} \\ \downarrow{\scriptstyle X_C} & & \downarrow{\scriptstyle X_C} \\ \mathbb{C} & \xrightarrow{C} & \mathbb{C} \\ \downarrow{\scriptstyle U_C} & & \downarrow{\scriptstyle U_C} \\ \mathbb{C} & \xrightarrow{A} & \mathbb{C} \end{array} \tag{1.10}$$

is commutative. Furthermore, the degree of X is bounded from above by a constant $c = c(n)$ which depends on n only.

We did not make special efforts to obtain an optimal estimation for $c(n)$, however our method of proof shows that

$$c(n) \leq (n-1)!\, n^{2\log_2 n + 3}.$$

Thus, Theorem 1.6 gives an effective bound on the number of conjugacy classes of polynomials A such that $A \leq B$.

The paper is organized as follows. In the second section we give a very brief overview of the Ritt theory. In the third section we recall basic results of [13] and prove Theorem 1.1 and Theorem 1.2. We also prove the corollaries of Theorem 1.1 mentioned above. In the fourth section we first show that if $A \leq B$ and one of polynomials A or B is special, then the other one also is special (Theorem 4.4). Then we prove Theorem 1.3 and deduce from it the result of Ritt about commuting polynomials. We also apply Theorem 1.3 to the problem of description of curves in $\mathbb{C}^2$ invariant under maps $F : (x, y) \to (f(x), g(y))$, where f and g are polynomials, and prove Theorem 1.4. Finally, we prove Theorem 1.5.

In the fifth section we first show (Theorem 5.2) that if B is a non-special polynomial of degree n, and $X \in \mathcal{E}(B)$, then the degree l of any special compositional factor of X satisfies the inequality $l \leq 2n$. On this base we prove that if $X \in \mathcal{E}(B)$ is not a polynomial in B, then $\deg X$ is bounded from above by a constant which depends on n only. In turn, from this result we deduce Theorem 1.6. As another corollary of the boundedness of $\deg X$ we obtain the following result of Zieve and Müller [22]: if B is a non-special polynomial of degree $n \geq 2$, and X and Y are polynomials such that $Y \circ X = B^{\circ s}$ for some $s \geq 1$, then there exist polynomials $\widetilde{X}$, $\widetilde{Y}$ and $i, j \geq 0$ such that

$$Y = B^{\circ i} \circ \widetilde{Y}, \quad X = \widetilde{X} \circ B^{\circ j}, \quad \text{and} \quad \widetilde{Y} \circ \widetilde{X} = B^{\circ \widetilde{s}},$$

where $\widetilde{s}$ is bounded from above by a constant which depends on n only.

ACKNOWLEDGEMENTS. The author is grateful to the Max-Planck-Institut für Mathematik for the hospitality and the support.

2. Overview of the Ritt theory

Let F be a polynomial with complex coefficients. The polynomial F is called *indecomposable* if the equality $F = F_2 \circ F_1$ implies that at least one of the polynomials F_1, F_2 is of degree one. Any representation of a polynomial F in the form $F = F_r \circ F_{r-1} \circ \cdots \circ F_1$, where $F_1, F_2, \ldots, F_r$ are polynomials, is called *a decomposition* of F. A decomposition is called *maximal* if all $F_1, F_2, \ldots, F_r$ are indecomposable and of degree greater than one. Two decompositions having an equal number of terms

$$F = F_r \circ F_{r-1} \circ \cdots \circ F_1 \quad \text{and} \quad F = G_r \circ G_{r-1} \circ \cdots \circ G_1$$

are called *equivalent* if either $r = 1$ and $F_1 = G_1$, or $r \geq 2$ and there exist polynomials μ_i, $1 \leq i \leq r-1$, of degree 1 such that

$$F_r = G_r \circ \mu_{r-1}, \quad F_i = \mu_i^{-1} \circ G_i \circ \mu_{i-1}, \quad 1 < i < r, \quad \text{and} \quad F_1 = \mu_1^{-1} \circ G_1.$$

The theory of polynomial decompositions established by Ritt can be summarized in the form of two theorems usually called the first and the second Ritt theorems (see [20]).

The first Ritt theorem states, roughly speaking, that any maximal decompositions of a polynomial may be obtained from any other by some iterative process involving the functional equation

$$A \circ C = D \circ B. \tag{2.1}$$

Theorem 2.1 ([20]). *Any two maximal decompositions $\mathcal{D}, \mathcal{E}$ of a polynomial P have an equal number of terms. Furthermore, there exists a chain of maximal decompositions $\mathcal{F}_i$, $1 \leq i \leq s$, of P such that $\mathcal{F}_1 = \mathcal{D}$, $\mathcal{F}_s \sim \mathcal{E}$, and $\mathcal{F}_{i+1}$ is obtained from $\mathcal{F}_i$ by a replacement of two successive polynomials $A \circ C$ in $\mathcal{F}_i$ by two other polynomials $D \circ B$ such that* (2.1) *holds.*

The second Ritt theorem in turn describes indecomposable polynomial solutions of (2.1). More precisely, it describes solutions satisfying the condition

$$\mathrm{GCD}(\deg A, \deg D) = 1, \quad \mathrm{GCD}(\deg C, \deg B) = 1, \tag{2.2}$$

which holds in particular if A, C, D, B are indecomposable (see Theorem 2.3 below).

Theorem 2.2 ([20]). *Let A, C, D, B be polynomials such that* (2.1) *and* (2.2) *hold. Then there exist polynomials $\sigma_1, \sigma_2, \mu, \nu$ of degree one such that, up to a possible replacement of A by D and of C by B, either*

$$A = \nu \circ z^s R^n(z) \circ \sigma_1^{-1}, \qquad C = \sigma_1 \circ z^n \circ \mu \tag{2.3}$$

$$D = \nu \circ z^n \circ \sigma_2^{-1}, \qquad B = \sigma_2 \circ z^s R(z^n) \circ \mu, \tag{2.4}$$

where R is a polynomial, $n \geq 1$, $s \geq 0$, and $\mathrm{GCD}(s, n) = 1$, *or*

$$A = \nu \circ T_m \circ \sigma_1^{-1}, \qquad C = \sigma_1 \circ T_n \circ \mu, \tag{2.5}$$

$$D = \nu \circ T_n \circ \sigma_2^{-1} \qquad B = \sigma_2 \circ T_m \circ \mu, \tag{2.6}$$

where T_n, T_m are the Chebyshev polynomials, $n, m \geq 1$, and $\mathrm{GCD}(n, m) = 1$.

Notice that the main difficulty in the practical use of Theorem 2.1 and Theorem 2.2 is the fact that classes of solutions appearing in Theorem 2.2 are not disjoint. Namely, any solution of the form (2.5), (2.6) with $n = 2$ can also be represented in the form (2.3), (2.4) (see, *e.g.*, [10,16,22] for further details).

The description of polynomial solutions of equation (2.1) in the general case in a certain sense reduces to the case where (2.2) holds by the following statement.

Theorem 2.3 ([3]). *Let A, C, D, B be polynomials such that* (2.1) *holds. Then there exist polynomials U, V, $\widetilde{A}$, $\widetilde{C}$, $\widetilde{D}$, $\widetilde{B}$, where*

$$\deg U = \mathrm{GCD}(\deg A, \deg D), \quad \deg V = \mathrm{GCD}(\deg C, \deg B),$$

such that

$$A = U \circ \widetilde{A}, \quad D = U \circ \widetilde{D}, \quad C = \widetilde{C} \circ V, \quad B = \widetilde{B} \circ V,$$

and

$$\widetilde{A} \circ \widetilde{C} = \widetilde{D} \circ \widetilde{B}.$$

In particular, if $\deg C = \deg B$, *then there exists a polynomial μ of degree one such that*

$$A = D \circ \mu^{-1}, \quad C = \mu \circ B.$$

Theorem 2.2 implies the following description of polynomial solutions of equation (1.2) under the condition

$$\mathrm{GCD}(\deg X, \deg B) = 1 \tag{2.7}$$

(see [7]).

Theorem 2.4 ([7]). *Let A, B, X be polynomials such that* (1.2) *and* (2.7) *hold. Then there exist polynomials μ, ν of degree one such that either*

$$A = \nu \circ z^s R^n(z) \circ \nu^{-1}, \quad X = \nu \circ z^n \circ \mu, \quad D = \mu^{-1} \circ z^s R(z^n) \circ \mu,$$

where R is a polynomial, $n \geq 1$, $s \geq 0$, and $\mathrm{GCD}(s, n) = 1$, *or*

$$A = \nu \circ \pm T_m \circ \nu^{-1}, \quad X = \nu \circ T_n \circ \mu, \quad D = \mu^{-1} \circ \pm T_m \circ \mu,$$

where T_n, T_m are the Chebyshev polynomials, $n, m \geq 1$, and $\mathrm{GCD}(n, m) = 1$.

Notice, however, that Theorem 2.2, even combined with Theorem 2.3, provides very little information about solutions of (1.2) if (2.7) is not satisfied. A possible way to investigate the general case is to analyze somehow the totality of all decompositions of a polynomial P, basing on Theorem 2.1 and Theorem 2.2, and then to apply this analysis to (1.2) using the fact that we can pass from the decomposition $P = A \circ X$ to the decomposition $P = X \circ B$. This idea was used in [10]. A similar technique was used in [22], where it was applied to the study of decompositions of iterations of a polynomial. In this paper we use another method completely bypassing Theorem 2.1. Notice by the way that Theorem 2.1 does not hold for arbitrary rational functions (see, *e.g.*, [12]).

3. Semiconjugacies and Julia sets

3.1. Polynomials sharing preimages of compact sets

Let $f_1(z)$, $f_2(z)$ be non-constant complex polynomials and $K_1, K_2 \subset \mathbb{C}$ compact sets. In the paper [13] we investigated the following problem. Under what conditions on the collection $f_1(z)$, $f_2(z)$, K_1, K_2 do the preimages $f_1^{-1}(K_1)$ and $f_2^{-1}(K_2)$ coincide, that is,

$$f_1^{-1}(K_1) = f_2^{-1}(K_2) = K \tag{3.1}$$

for some compact set $K \subset \mathbb{C}$?

Using ideas from approximation theory, we relate equation (3.1) to the functional equation

$$g_1(f_1(z)) = g_2(f_2(z)), \tag{3.2}$$

where $f_1(z)$, $f_2(z)$, $g_1(z)$, $g_2(z)$ are polynomials. It is easy to see that for any polynomial solution of (3.2) and any compact set $K_3 \subset \mathbb{C}$ we obtain a solution of (3.1) setting

$$K_1 = g_1^{-1}(K_3), \quad K_2 = g_2^{-1}(K_3). \tag{3.3}$$

Briefly, the main result of [13] states that, under a very mild condition on the cardinality of K, all solutions of (3.1) can be obtained in this way. Combined with Theorem 2.3 and Theorem 2.2 this leads to a very explicit description of solutions of (3.1).

Theorem 3.1 ([13]). *Let* $f_1(z)$, $f_2(z)$ *be polynomials,* $\deg f_1 = d_1$, $\deg f_2 = d_2$, $d_1 \leq d_2$, *and let* $K_1, K_2, K \subset \mathbb{C}$ *be compact sets such that* (3.1) *holds. Suppose that* $\mathrm{card}\{\mathrm{K}\} \geq \mathrm{LCM}(\mathrm{d}_1, \mathrm{d}_2)$. *Then, if* d_1 *divides* d_2, *there exists a polynomial* $g_1(z)$ *such that* $f_2(z) = g_1(f_1(z))$ *and* $K_1 = g_1^{-1}(K_2)$. *On the other hand, if* d_1 *does not divide* d_2, *then there exist polynomials* $g_1(z)$, $g_2(z)$, $\deg g_1 = d_2/d$, $\deg g_2 = d_1/d$, *where* $d = \mathrm{GCD}(d_1, d_2)$, *and a compact set* $K_3 \subset \mathbb{C}$ *such that* (3.2), (3.3) *hold. Furthermore, in this case there exist polynomials* $\widetilde{f}_1(z)$, $\widetilde{f}_2(z)$, $W(z)$, $\deg W(z) = d$, *such that*

$$f_1(z) = \widetilde{f}_1(W(z)), \quad f_2(z) = \widetilde{f}_2(W(z)) \tag{3.4}$$

and there exist linear functions $\sigma_1(z)$, $\sigma_2(z)$ *such that either*

$$\begin{aligned} g_1(z) &= z^c R^{d_1/d}(z) \circ \sigma_1^{-1}, \qquad & \widetilde{f}_1(z) &= \sigma_1 \circ z^{d_1/d}, \\ g_2(z) &= z^{d_1/d} \circ \sigma_2^{-1}, \qquad & \widetilde{f}_2(z) &= \sigma_2 \circ z^c R\big(z^{d_1/d}\big), \end{aligned} \tag{3.5}$$

for some polynomial $R(z)$ *and* c *equal to the remainder after division of* d_2/d *by* d_1/d, *or*

$$\begin{aligned} g_1(z) &= T_{d_2/d}(z) \circ \sigma_1^{-1}, \qquad & \widetilde{f}_1(z) &= \sigma_1 \circ T_{d_1/d}(z), \\ g_2(z) &= T_{d_1/d}(z) \circ \sigma_2^{-1}, \qquad & \widetilde{f}_2(z) &= \sigma_2 \circ T_{d_2/d}(z), \end{aligned} \tag{3.6}$$

for the Chebyshev polynomials $T_{d_1/d}(z)$, $T_{d_2/d}(z)$.

Theorem 3.1 may be used to prove many other results (see [13] for details), the most notable of which is the following description of solutions of (3.1) in the case where $K_1 = K_2$, first obtained by T. Dinh [1,2] by methods of complex dynamics.

Theorem 3.2 ([2,13]). *Let $f_1(z)$, $f_2(z)$ be polynomials such that*

$$f_1^{-1}(T) = f_2^{-1}(T) = K \tag{3.7}$$

holds for some infinite compact sets T, $K \subset \mathbb{C}$. Then, if d_1 divides d_2, there exists a polynomial $g_1(z)$ such that $f_2(z) = g_1(f_1(z))$ and $g_1^{-1}(T) = T$. On the other hand, if d_1 does not divide d_2, then there exist polynomials $\widetilde{f_1}(z)$, $\widetilde{f_2}(z)$, $W(z)$, $\deg W(z) = d$, *satisfying* (3.4). *Furthermore, in this case one of the following conditions holds:*

1) *T is a union of concentric circles and*

$$\widetilde{f_1}(z) = \sigma \circ z^{d_1/d}, \qquad \widetilde{f_2}(z) = \sigma \circ \gamma z^{d_2/d} \tag{3.8}$$

for some linear function $\sigma(z)$ and $\gamma \in \mathbb{C}$;

2) *T is a segment and*

$$\widetilde{f_1}(z) = \sigma \circ \pm T_{d_1/d}(z), \qquad \widetilde{f_2}(z) = \sigma \circ \pm T_{d_2/d}(z), \tag{3.9}$$

for some linear function $\sigma(z)$ and the Chebyshev polynomials $T_{d_1/d}(z)$, $T_{d_2/d}(z)$.

3.2. Proofs of Theorem 1.1 and Theorem 1.2

Proof of Theorem 1.1. If $A \underset{X}{\leq} B$, then for any $n \geq 1$ the equality

$$A^{\circ n} \circ X = X \circ B^{\circ n}$$

holds. Therefore, if $z_1 = X(z_0)$, then the sequence $A^{\circ n}(z_1)$ is bounded if and only if the sequence $X \circ B^{\circ n}(z_0)$ is bounded. In turn, the last sequence is bounded if and only if the sequence $B^{\circ n}(z_0)$ is bounded. Thus, $A \underset{X}{\leq} B$ implies

$$X^{-1}(K(A)) = K(B). \tag{3.10}$$

Conversely, if (3.10) holds, then it follows from $B^{-1}(K(B)) = K(B)$ that

$$(X \circ B)^{-1}(K(A)) = K(B).$$

Thus,

$$X^{-1}(K(A)) = (X \circ B)^{-1}(K(A)).$$

Since $\deg X \mid \deg(X \circ B)$, applying to the latter equality Theorem 3.1 we conclude that

$$\widetilde{A} \circ X = X \circ B$$

for some polynomial $\widetilde{A}$. Furthermore, since we proved that for such $\widetilde{A}$ the equality $X^{-1}(K(\widetilde{A})) = K(B)$ holds, we see that $X^{-1}(K(\widetilde{A})) = X^{-1}(K(A))$, implying that $K(\widetilde{A}) = K(A)$. Finally, it follows from Theorem 3.1 applied to

$$A^{-1}(K) = \widetilde{A}^{-1}(K) = K,$$

where $K = K(\widetilde{A}) = K(A)$, that there exists a polynomial of degree one μ such that $\widetilde{A} = \mu \circ A$ and $\mu(K(A)) = K(A)$.

More generally, if

$$X^{-1}(K) = K(B) \tag{3.11}$$

for some compact set $K \subset \mathbb{C}$, then

$$X^{-1}(K) = (X \circ B)^{-1}(K),$$

implying by Theorem 3.1 that (1.2) holds for some polynomial A. Furthermore, since for such a polynomial A equality (3.10) holds, we conclude that $X^{-1}(K) = X^{-1}(K(A))$ and $K = K(A)$. □

Corollary 3.3. *Let B be a polynomial of degree at least two. Then a polynomial X is contained in $\mathcal{E}(B)$ if and only $K(B)$ is a union of fibers of X. In particular, if B_1 and B_2 are polynomials such that $K(B_1) = K(B_2)$, then $\mathcal{E}(B_1) = \mathcal{E}(B_2)$.*

Proof. Clearly, condition (3.11) implies that $K(B)$ is a union of fibers of X. Conversely, if $K(B)$ is a union of fibers of X, then

$$K(B) = X^{-1}\big(X(K(B))\big),$$

implying that (3.11) holds for the compact set $K = X(K(B))$. □

Corollary 3.4. *Let A, B, and X be polynomials such that $A \underset{X}{\leq} B$. Then for any decomposition $X = X_1 \circ X_2$ there exists a polynomial C such that*

$$A \underset{X_1}{\leq} C, \quad C \underset{X_2}{\leq} B.$$

Proof. By Theorem 1.1, $K(B) = X^{-1}(K(A))$. Since $X = X_1 \circ X_2$, this implies that $K(B) = X_2^{-1}(\widetilde{K})$, where $\widetilde{K} = X_1^{-1}(K(A))$. Therefore, by Theorem 1.1, there exists a polynomial C such that

$$C \circ X_2 = X_2 \circ B. \tag{3.12}$$

Now we have:

$$A \circ X_1 \circ X_2 = X_1 \circ X_2 \circ B = X_1 \circ C \circ X_2,$$

implying that $A \circ X_1 = X_1 \circ C$. □

Remark 3.5. Corollary 3.4 may be proved without using Theorem 1.1. Indeed, if $X = X_1 \circ X_2$, then it follows from

$$A \circ (X_1 \circ X_2) = X_1 \circ (X_2 \circ B),$$

by Theorem 2.3, that

$$X_1 \circ X_2 = U \circ \widetilde{W}, \quad X_2 \circ B = V \circ \widetilde{W}, \tag{3.13}$$

where

$$\deg \widetilde{W} = \mathrm{GCD}(\deg(X_1 \circ X_2), \deg(X_2 \circ B)).$$

Since $\deg X_2 \mid \deg \widetilde{W}$, Theorem 2.3 applied to the first equality in (3.13) implies that $\widetilde{W} = S \circ X_2$ for some polynomial S. Therefore,

$$X_2 \circ B = V \circ \widetilde{W} = V \circ S \circ X_2$$

and hence (3.12) holds for $C = V \circ S$.

Proof of Theorem 1.2. By Theorem 1.1, the condition $X_1, X_2 \in \mathcal{E}(B)$ implies that there exist $K_1, K_2 \subset \mathbb{C}$ such that

$$X_1^{-1}(K_1) = K(B), \quad X_2^{-1}(K_2) = K(B).$$

It now follows from Theorem 3.1 that there exist polynomials X, W, U_1, U_2, V_1, V_2 such that

$$\deg X = \mathrm{LCM}(\deg X_1, \deg X_2), \quad \deg W = \mathrm{GCD}(\deg X_1, \deg X_2),$$

and that equalities

$$X = U_1 \circ X_1 = U_2 \circ X_2$$

and

$$X_1 = V_1 \circ W, \quad X_2 = V_2 \circ W \tag{3.14}$$

hold. Furthermore, there exists $K_3 \subset \mathbb{C}$ such that

$$K_1 = U_1^{-1}(K_3), \quad K_2 = U_2^{-1}(K_3).$$

Therefore, $X^{-1}(K_3) = K(B)$, implying by Theorem 1.1 that $X \in \mathcal{E}(B)$. Finally, any of equalities (3.14) implies that $W \in \mathcal{E}(B)$ by Corollary 3.4. □

4. Semiconjugacies between fixed *A* and *B*

4.1. Semiconjugacies between special polynomials

For a polynomial P and a finite set $K \subset \mathbb{C}$ denote by $P^{-1}_{\text{odd}}(K)$ the subset of $P^{-1}(K)$ consisting of points where the local multiplicity of P is odd. Notice that the chain rule implies that if $P = A \circ B$, then

$$P^{-1}_{\text{odd}}(K) = B^{-1}_{\text{odd}}\left(A^{-1}_{\text{odd}}(K)\right). \tag{4.1}$$

Lemma 4.1. *Let P be a polynomial of degree $n \geq 2$, and $K \subset \mathbb{C}$ a finite set containing at least two points. Assume that $P^{-1}_{\text{odd}}(K) = K$. Then K contains exactly two points, and P is conjugate to $\pm T_n$.*

Proof. Denote by e_z the multiplicity of P at $z \in \mathbb{C}$, and set $r = \text{card}(K)$. Since for any $y \in \mathbb{C}$ the set $P^{-1}(y)$ contains

$$n - \sum_{\substack{z\in\mathbb{C}\\ P(z)=y}} (e_z - 1)$$

points and

$$\sum_{z\in\mathbb{C}} (e_z - 1) = n - 1,$$

we have:

$$\text{card}\big(P^{-1}(K)\big) \geq rn - \sum_{z\in\mathbb{C}} (e_z - 1) = (r-1)n + 1 \tag{4.2}$$

(the minimum is attained if K contains all finite critical values of P). Therefore, if

$$\text{card}\left(P^{-1}_{\text{odd}}(K)\right) = \text{card}(K) = r,$$

then the set $P^{-1}(K)$ contains at least $(r-1)n + 1 - r$ points where the local multiplicity of P is greater than one, implying that

$$\sum_{z\in P^{-1}(K)} e_z \geq r + 2\left((r-1)n + 1 - r\right). \tag{4.3}$$

Since the sum in the left-hand side of (4.3) equals rn, this inequality implies that

$$(n-1)(r-2) \leq 0. \tag{4.4}$$

Thus, $r = 2$. Furthermore, since the equality in (4.4) is attained if and only if equality is attained in (4.3), we conclude that if $P^{-1}_{\text{odd}}(K) = K$, then $e_z = 2$ for each $z \in P^{-1}(K) \setminus K$, and the local multiplicity of P at each of the two points of K is equal to one.

Changing P to $\sigma^{-1} \circ P \circ \sigma$ for a convenient polynomial of degree one σ, we can assume that $K = \{-1, 1\}$. Then the condition on multiplicities of P implies that $P^2 - 1$ is divisible by $(P')^2$, and calculating the quotient we conclude that P satisfies the differential equation

$$n^2\left(1 - y^2\right) = \left(y'\right)^2\left(1 - z^2\right).$$

Since the general solution of the equation

$$\frac{y'}{\sqrt{1 - y^2}} = \pm\frac{n}{\sqrt{1 - z^2}}$$

is

$$\arccos y = \pm n \arccos z + c,$$

it follows now from $P(1) = \pm 1$ that

$$P = \pm\cos(n \arccos x) = \pm T_n(z).$$

□

Remark 4.2. Notice that the equality $T_n(-z) = (-1)^n T_n(z)$ implies that for even n the polynomials T_n and $-T_n$ are conjugate since $T_n = \alpha \circ (-T_n) \circ \alpha^{-1}$, where $\alpha(z) = -z$. For odd n however the polynomials T_n and $-T_n$ are not conjugate.

Lemma 4.3. *Let P be a polynomial and $a, b \in \mathbb{C}$. Then the set $P^{-1}_{\mathrm{odd}}\{a, b\}$ contains at least two points.*

Proof. It follows from the equality

$$2n = \sum_{\substack{z\in\mathbb{C}\\ P(z)=a}} e_z + \sum_{\substack{z\in\mathbb{C}\\ P(z)=b}} e_z$$

that the number

$$\sum_{z\in P^{-1}_{\mathrm{odd}}\{a,b\}} e_z$$

is even, implying that the number $\mathrm{card}(P^{-1}_{\mathrm{odd}}\{a, b\})$ also is even. On the other hand,

$$\mathrm{card}\left(P^{-1}_{\mathrm{odd}}\{a, b\}\right) \neq 0,$$

for otherwise $P^{-1}_{\mathrm{odd}}\{a, b\}$ contains at most $n/2 + n/2 = n$ points in contradiction with inequality (4.2). □

Theorem 4.4. *Let A and B be polynomials of degree at least two such that $A \le B$. Then A is conjugate to z^n if and only if B is conjugate to z^n. Similarly, A is conjugate to $\pm T_n$ if and only if B is conjugate to $\pm T_n$.*

Proof. Assume that B is conjugate to $\pm T_n$, and let X be a semiconjugacy from B to A. Changing B and X to $\sigma^{-1}\circ B\circ\sigma$ and $X\circ\sigma$, for a convenient polynomial σ of degree one, without loss of generality we can assume that $B=\pm T_n$. By Theorem 1.1, we have:

$$X^{-1}(K(A))=K(B)=[-1,1]. \tag{4.5}$$

Set $m=\deg X$. Since

$$T_m^{-1}([-1,1])=[-1,1], \tag{4.6}$$

equality (4.5) implies that

$$X^{-1}(K(A))=T_m^{-1}([-1,1]).$$

It now follows from Theorem 3.1 that there exists a polynomial δ of degree one such that $X=\delta\circ T_m$. Therefore, changing A and X to $\delta^{-1}\circ A\circ\delta$ and $\sigma^{-1}\circ X$, we can assume that $X=T_m$. Thus, we have:

$$A\circ T_m=T_m\circ\pm T_n=(-1)^m T_n\circ T_m, \tag{4.7}$$

implying that $A=\pm T_n$.

Similarly, if $B=z^n$, then the equalities

$$X^{-1}(K(A))=K(B)=\mathbb{D},$$

and $(z^m)^{-1}(\mathbb{D})=\mathbb{D}$ imply that $X=\delta\circ z^m$ for some polynomial δ of degree one, and arguing as above we conclude that A is conjugate to z^n.

Assume now that A is conjugate to $\pm T_n$. Without loss of generality we can assume that $A=\pm T_n$. Since $T_{n\,\mathrm{odd}}^{-1}\{-1,1\}=\{-1,1\}$, formula (4.1) implies that

$$(\pm T_n\circ X)_{\mathrm{odd}}^{-1}\{-1,1\}=X_{\mathrm{odd}}^{-1}\{-1,1\}.$$

It follows now from

$$\pm T_n\circ X=X\circ B \tag{4.8}$$

that

$$B_{\mathrm{odd}}^{-1}\left(X_{\mathrm{odd}}^{-1}\{-1,1\}\right)=X_{\mathrm{odd}}^{-1}\{-1,1\}. \tag{4.9}$$

Since by Lemma 4.3 the set $X_{\mathrm{odd}}^{-1}\{-1,1\}$ contains at least two points, this implies by Lemma 4.1 that the polynomial B is conjugate to $\pm T_n$.

Finally, if A is conjugate to z^n, we can assume that $A=z^n$, and considering zeroes of the left and the right parts of the equality

$$z^n\circ X=X\circ B,$$

we see that $B^{-1}(X^{-1}(0))=X^{-1}(0)$. It follows now from inequality (4.2) that $X^{-1}(0)$ consists of a single point, implying easily that the polynomial B is conjugate to z^n. □

Remark 4.5. Since for even n the polynomials T_n and $-T_n$ are conjugate (see Remark 4.2), Theorem 4.4 implies that if B is conjugate to $\pm T_n$ for even n, then A and B are conjugate. On the other hand, if B is conjugate to $-T_n$ for odd n, then A is not necessarily conjugate to $-T_n$, but only to $\pm T_n$. Still, it follows from (4.7) that if B is conjugate to T_n, then A is conjugate to T_n.

Notice that Theorem 4.4 combined with Remark 4.5 implies the following corollary.

Corollary 4.6. *Let A and B be polynomials such that the conditions $A \leq B$ and $B \leq A$ hold simultaneously, and at least one of A and B is special. Then A and B are conjugate.*

4.2. Proof of Theorem 1.3

The following lemma is a well-known fact from the complex dynamics. For the reader's convenience we give a short proof based on Theorem 3.1.

Lemma 4.7. *Let A be a polynomial of degree n such that $K(A)$ is a union of circles with a common center. Then $K(A)$ is a disk, and A is conjugate to z^n. Similarly, if $K(A)$ is a segment, then A is conjugate to $\pm T_n$.*

Proof. Since for a polynomial A the complement to $K(A)$ in $\mathbb{CP}^1$ is connected (see, *e.g.*, [11, Lemma 9.4]), if $K(A)$ is a union of circles with a common center, then $K(A)$ is a disk. Furthermore, changing if necessary A to a conjugate polynomial, we can assume that $K(A) = \mathbb{D}$. Thus, $A^{-1}(\mathbb{D}) = \mathbb{D}$. On the other hand, $(z^n)^{-1}(\mathbb{D}) = \mathbb{D}$, and applying to these equalities Theorem 3.1, we conclude that $A = \alpha z^n$, where $|\alpha| = 1$, implying that A is conjugate to z^n.

Similarly, if $K(A)$ is a segment, we can assume that $K(A) = [-1, 1]$, and to conclude in a similar way that A is conjugate to $\pm T_n$. □

Proof of Theorem 1.3. Set $d_0 = \deg X_0$, and let $X \in \mathcal{E}(A, B)$ be a polynomial of degree d. By Theorem 1.1, we have:

$$X_0^{-1}(K(A)) = K(B), \quad X^{-1}(K(A)) = K(B).$$

Applying to these equalities Theorem 3.2 and taking into account that, by Lemma 4.7, $K(A)$ is neither a union of concentric circles nor a segment, we conclude that $X = \widetilde{A} \circ X_0$ for some polynomial $\widetilde{A}$. Substituting now this expression in (1.2) and using that $X_0 \in \mathcal{E}(A, B)$ we have:

$$A \circ \widetilde{A} \circ X_0 = \widetilde{A} \circ X_0 \circ B = \widetilde{A} \circ A \circ X_0,$$

implying that $A \circ \widetilde{A} = A \circ \widetilde{A}$.

Conversely, if A commutes with $\widetilde{A}$, then

$$A \circ (\widetilde{A} \circ X_0) = \widetilde{A} \circ A \circ X_0 = (\widetilde{A} \circ X_0) \circ B.$$ □

Theorem 1.3 implies in particular the following classification of commuting polynomials obtained by Ritt.

Theorem 4.8 ([21]). *Let A and B be commuting polynomials of degree at least two. Then, up to the change*

$$A \to \lambda \circ A \circ \lambda^{-1}, \quad B \to \lambda \circ B \circ \lambda^{-1}, \tag{4.10}$$

where λ is a polynomial of degree one, either

$$A = z^n, \quad B = \varepsilon z^m, \tag{4.11}$$

where $\varepsilon^n = \varepsilon$, or

$$A = \pm T_n, \quad B = \pm T_m, \tag{4.12}$$

or

$$A = \varepsilon_1 R^{\circ m}, \quad B = \varepsilon_2 R^{\circ n}, \tag{4.13}$$

where $R = zS(z^\ell)$ for some polynomial S, and ε_1, ε_2 are l-th roots of unity.

Proof. Assume first that A is conjugate to z^n. Without loss of generality we may assume that $A = z^n$. Applying Theorem 1.1 for $B = A$ and $X = B$, we have:

$$B^{-1}(K(A)) = K(A).$$

Since $K(A) = \mathbb{D}$, arguing as in Lemma 4.7 we conclude that $B = \varepsilon z^m$, and it follows from $A \circ B = B \circ A$ that $\varepsilon^n = \varepsilon$. If A is conjugate to $\pm T_n$, the proof is similar.

On the other hand, if A is non-special, then Theorem 1.3 implies that any $B \in \mathcal{E}(A, A)$ has the form $B = \widetilde{A} \circ R$, where R is a polynomial of the minimum possible degree in $\mathcal{E}(A, A)$. Now we can apply Theorem 1.3 again to the polynomial $\widetilde{A}$ and so on, obtaining eventually the representation $B = \mu_1 \circ R^{\circ m_1}$, where μ_1 is a polynomial of degree one commuting with A. In particular, since $A \in \mathcal{E}(A, A)$, the equality $A = \mu_2 \circ R^{\circ m_2}$ holds for some polynomial μ_2 of degree one commuting with A. Furthermore, since R commutes with $A = \mu_2 \circ R^{\circ m_2}$, the polynomial μ_2 commutes with R. This implies easily that, up to a conjugacy, $R = zS(z^\ell)$ for some polynomial S, and $\mu_2 = \varepsilon_2 z$ for some lth root of unity ε_2. Finally, since μ_1 commutes with the polynomial A, and $A = \mu_2 \circ R^{\circ m_2}$ has the form $z\widetilde{S}(z^\ell)$ for some polynomial $\widetilde{S}$, we conclude that $\mu_1 = \varepsilon_1 z$ for some lth root of unity ε_1. □

4.3. Semiconjugacies and invariant curves

It was shown in the recent paper [10] that the problem of describing semiconjugate polynomials is closely related to the problem of describing algebraic curves $\mathcal{C}$ in $\mathbb{C}^2$ invariant under maps of the form $F : (x, y) \to (f(x), g(y))$, where f, g are polynomials of degree at least two. Briefly, this relation may be summarized as follows (see [10, Proposition 2.34] for more details).

If $\mathcal{C}$ is an irreducible (f, g)-invariant curve, then its projective closure $\overline{\mathcal{C}}$ in $\mathbb{CP}^1 \times \mathbb{CP}^1$ is also (f, g)-invariant. Denote by $\bar{h}$ the restriction of F on $\overline{\mathcal{C}}$. Let $\widetilde{\mathcal{C}}$ be the desingularization of $\mathcal{C}$ and $\beta : \widetilde{\mathcal{C}} \to \overline{\mathcal{C}}$ a map biholomorphic off a finite set. Clearly, $\bar{h}$ lifts to a holomorphic map $h : \widetilde{\mathcal{C}} \to \widetilde{\mathcal{C}}$. Consider now the commutative diagram

$$\begin{array}{ccc} \widetilde{\mathcal{C}} & \xrightarrow{h} & \widetilde{\mathcal{C}} \\ \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\beta} \\ \overline{\mathcal{C}} & \xrightarrow{\bar{h}} & \overline{\mathcal{C}} \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\alpha} \\ \mathbb{CP}^1 & \xrightarrow{f} & \mathbb{CP}^1, \end{array} \tag{4.14}$$

where $\alpha : \overline{\mathcal{C}} \to \mathbb{CP}^1$ is the projection map onto the first coordinate. Set $\pi = \alpha \circ \beta$. If π is a constant, then $\mathcal{C}$ is a line $z_1 = \xi$, where ξ is a fixed point of f, so assume that the degree of π is at least one. Observe that since $f^{-1}(\infty) = \infty$, the set $K = \pi^{-1}(\infty)$ and the map h satisfy the equality

$$h^{-1}(K) = K. \tag{4.15}$$

Since h is a holomorphic map between Riemann surfaces of the same genus and $\deg h = \deg f \geq 2$, it follows from the Riemann-Hurwitz formula that either $g(\widetilde{\mathcal{C}}) = 0$, or $g(\widetilde{\mathcal{C}}) = 1$ and h is unbranched. Since $\deg h \geq 2$, for unbranched h equality (4.15) is impossible. Therefore, $\widetilde{\mathcal{C}} = \mathbb{CP}^1$ and (4.15) implies easily that, up to the change $\alpha \circ h \circ \alpha^{-1}$, where α is a Möbius transformation, either $K = \infty$ and h is a polynomial, or $K = \{0, \infty\}$ and $h = z^{\pm \deg f}$. Thus,

$$f \circ \pi = \pi \circ h, \tag{4.16}$$

where either π and h are polynomials, or $h = z^{\pm \deg f}$ and π is a Laurent polynomial. The last case requires an additional investigation. The paper [10] refers (Fact 2.25) to a more general result of [9] (Theorem 10) implying that for a non-special polynomial f this possibility is excluded. Alternatively, one can use the results of [14] (*e.g.*, Theorem 6.4).

Considering in a similar way the projection onto the second coordinate, we obtain the equality

$$g \circ \rho = \rho \circ h. \tag{4.17}$$

Thus, for non-special f and g any irreducible (f,g)-invariant curve may be parametrized by some polynomials π, ρ satisfying a system given by equations (4.16), (4.17) for some polynomial h.

Notice that in a certain sense a description of (f, g)-invariant curves reduces to the case $f = g$ since the commutative diagram

$$\begin{array}{ccc} \mathbb{C}^2 & \xrightarrow{(h,h)} & \mathbb{C}^2 \\ \Big\downarrow{\scriptstyle (\pi,\rho)} & & \Big\downarrow{\scriptstyle (\pi,\rho)} \\ \mathbb{C}^2 & \xrightarrow{(f,g)} & \mathbb{C}^2 \end{array} \tag{4.18}$$

implies that any (f, g)-invariant curve is an image of an (h, h)-invariant curve under the map $(x, y) \to (\pi(x), \rho(y))$.

Theorem 1.3 allows to obtain easily the following description of (f, f)-invariant curves obtained in [10] (see Theorem 6.24 and the theorem on page 85).

Theorem 4.9. *Let f be a non-special polynomial of degree at least two, and $\mathcal{C}$ an irreducible (f, f)-invariant curve in $\mathbb{C}^2$. Then there exists a polynomial p which commutes with f such that $\mathcal{C}$ has either the form $z_1 = p(z_2)$ or $z_2 = p(z_1)$.*

Proof. If $\mathcal{C}$ is a line $z_1 = \xi$, then ξ is a fixed point of f, and the conclusion of the theorem holds for $p = \xi$. Similarly, the theorem holds if $\mathcal{C}$ is a line $z_2 = \xi$. Otherwise, as it was shown above, $\mathcal{C}$ may be parametrized by some non-constant polynomials π, ρ satisfying the system

$$f \circ \pi = \pi \circ h, \tag{4.19}$$

$$f \circ \rho = \rho \circ h \tag{4.20}$$

for some polynomial h. Furthermore, without loss of generality we may assume that there exists no polynomial w of degree greater than one such that

$$\pi = \widetilde{\pi} \circ w, \quad \rho = \widetilde{\rho} \circ w, \tag{4.21}$$

for some polynomials $\widetilde{\pi}$, $\widetilde{\rho}$. Indeed, if (4.21) holds, then applying Theorem 2.3 to the equality

$$(f \circ \widetilde{\pi}) \circ w = \widetilde{\pi} \circ (w \circ h),$$

we conclude that $w \circ h = \widetilde{h} \circ w$ for some polynomial $\widetilde{h}$, implying that we may change π to $\widetilde{\pi}$, ρ to $\widetilde{\rho}$, and h to $\widetilde{h}$.

Set $d = \mathrm{GCD}(\deg \rho, \deg \pi)$. Since f is not special, it follows from (4.19), (4.20) by Theorem 1.3 that if both ρ and π are of degree at least two, then $d > 1$, implying by Theorem 1.2 that (4.21) holds for some polynomials $\widetilde{\pi}$, $\widetilde{\rho}$ and w with $\deg w = d > 1$. Therefore, at least one of the polynomials ρ and τ is of degree one, say $\deg \rho = 1$. Then, $\mathcal{C}$ has the form $z_1 = p(z_2)$, where $p = \pi \circ \rho^{-1}$. Furthermore, equality (4.20) implies that $h = \rho^{-1} \circ f \circ \rho$, and substituting this expression into (4.19) we conclude that p commutes with f. □

Proof of Theorem 1.4. For any polynomials of coprime degrees u and v the curve $\mathcal{C}_{u,v}: u(x)-v(y)=0$ is irreducible (see [15, Proposition 3.1]). Furthermore, if (1.9) holds and (x_0, y_0) is a point on $\mathcal{C}_{u,v}$, then (1.9) yields the equality

$$u(f(x_0)) = t(u(x_0)) = t(v(y_0)) = v(g(y_0)),$$

implying that $(f(x_0), g(y_0))$ also is a point on $\mathcal{C}_{u,v}$.

Conversely, assume that $\mathcal{C}$ is an irreducible (f, g)-invariant curve which is not a line, and let π and ρ be polynomials parametrizing $\mathcal{C}$ and satisfying (4.16), (4.17) for some polynomial h. Then by Theorem 1.2, there exist polynomials u and v of coprime degrees such that

$$u \circ \pi = v \circ \rho.$$

Thus, any irreducible (f, g)-invariant curve $\mathcal{C}$ in $\mathbb{C}^2$ has the form $u(x)-v(y)=0$ for some polynomials u, v of coprime degrees. Furthermore, since the polynomial

$$s = u \circ \pi = v \circ \rho$$

belongs to $\mathcal{E}(h)$ we have:

$$t \circ u \circ \pi = u \circ \pi \circ h = u \circ f \circ \pi,$$
$$t \circ v \circ \rho = v \circ \rho \circ h = v \circ g \circ \rho,$$

implying (1.9). □

A further analysis of system (1.9) using Proposition 5.4 and Proposition 5.5 proved below leads to a more precise description of (f, g)-invariant curves apparently equivalent to the one given by [10, Theorem 6.2]. Notice however that in [10] a more general case of *skew-invariant* curves and *skew-twists* between polynomials is considered, and the methods of our paper involving Julia sets seem not to be extendable to this more general situation.

4.4. Semiconjugacies between equivalent A and B

For a natural number $n > 1$ with a prime decomposition $n = p_1^{a_1} p_2^{a_2} \dots p_s^{a_s}$ set $\operatorname{rad}(n) = p_1 p_2 \dots p_s$. The following two theorems in totality provide a proof of Theorem 1.5.

Theorem 4.10. *Let A and B be polynomials of degree at least two. Then conditions $A \leq B$ and $B \leq A$ hold simultaneously if and only if $A \sim B$.*

Proof. The "if" part follows from the definition of $\sim$ (see the introduction). Furthermore, if at least one of A and B is special, then conditions $A \leq B$ and $B \leq A$ imply by Corollary 4.6 that A and B are conjugate and hence equivalent. So, we may assume that A and B are non-special.

Let Y and X be polynomials such that

$$B \underset{Y}{\leq} A, \quad A \underset{X}{\leq} B. \tag{4.22}$$

Set $n = \deg A = \deg B$. We can assume that $\deg X > 1$, $\deg Y > 1$ since otherwise A and B are conjugate and hence $A \sim B$. Since (4.22) implies that $Y \circ X$ commutes with B, Theorem 4.8 implies that

$$\mathrm{rad}(\deg X) \mid \mathrm{rad}(n). \tag{4.23}$$

In particular,

$$\mathrm{GCD}(\deg X, n) > 1. \tag{4.24}$$

Applying Theorem 2.3 to the equality

$$A \circ X = X \circ B, \tag{4.25}$$

we conclude that there exist polynomials $\widetilde{X}$, $\widetilde{B}$, and W such that

$$B = \widetilde{B} \circ W, \quad X = \widetilde{X} \circ W, \tag{4.26}$$

and $\deg W = \mathrm{GCD}(\deg X, n)$. Clearly, $B \sim W \circ \widetilde{B}$, and equalities (4.25) and (4.26) imply that

$$A \circ \widetilde{X} = \widetilde{X} \circ (W \circ \widetilde{B}). \tag{4.27}$$

Furthermore, $\deg \widetilde{X} < \deg X$, since $\deg W > 1$ by (4.24). If $\deg \widetilde{X} = 1$, then $A \sim W \circ \widetilde{B}$ since A and $W \circ \widetilde{B}$ are conjugate; hence,

$$A \sim W \circ \widetilde{B} \sim B,$$

and we are done. Otherwise, we can apply Theorem 2.3 in a similar way to equality (4.27) and so on. Since condition (4.23) ensures that the degrees of corresponding semiconjugacies decrease, we obtain in this way a finite chain of equivalences from B to A. □

Theorem 4.11. *Let A and B be polynomials of degree at least two. Then $A \sim B$ if and only if there exist polynomials X and Y such that*

$$B \circ Y = Y \circ A, \quad A \circ X = X \circ B, \tag{4.28}$$

and $Y \circ X = B^{\circ d}$ for some $d \geq 0$.

Proof. Taking into account Theorem 4.10, we only need to show that if equalities (4.28) hold, then they hold for some $\widetilde{X}$, $\widetilde{Y}$ such that $\widetilde{Y} \circ \widetilde{X} = B^{\circ d}$, $d \geq 0$. Since (4.28) implies that $Y \circ X$ commutes with B, it follows from Theorem 4.8 that either B is special, or, up to a conjugacy,

$$Y \circ X = \varepsilon_1 R^{\circ m_1}, \quad B = \varepsilon_2 R^{\circ m_2},$$

where $R = zS(z^n)$ for some polynomial S, and ε_1, ε_2 are nth roots of unity. In the first case, Corollary 4.6 implies that A and B are conjugate. Therefore, in this case

(4.28) holds for some Möbius transformations $\widetilde{Y}$ and $\widetilde{X}$ such that $\widetilde{Y} \circ \widetilde{X} = B^0$. In the second case set

$$\widetilde{X} = X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)},$$

where $\varepsilon_3 = \varepsilon_2^{m_1}/\varepsilon_1$, and observe that the second of equalities (4.28) still holds for $\widetilde{X}$ since

$$A \circ \widetilde{X} = A \circ X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = X \circ B \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)}$$
$$= X \circ \varepsilon_2 R^{\circ m_2} \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = X \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} \circ \varepsilon_2 R^{\circ m_2} = \widetilde{X} \circ B.$$

On the other hand, we have:

$$Y \circ \widetilde{X} = \varepsilon_1 R^{\circ m_1} \circ \varepsilon_3 R^{\circ(m_2 m_1 - m_1)} = \varepsilon_1 \varepsilon_3 R^{\circ m_2 m_1} = \varepsilon_2^{m_1} R^{\circ m_2 m_1} = B^{\circ m_1}. \qquad \square$$

5. Semiconjugacies for fixed B

5.1. Special factors of semiconjugacies

Lemma 5.1. *Let A and B be polynomials of degree $n \geq 2$ such that*

$$A \circ T_\ell = T_\ell \circ B, \quad l \geq 2. \tag{5.1}$$

Then $l \leq 2n$, unless $A = \pm T_n$ and $B = \pm T_n$. Similarly, if

$$A \circ z^\ell = z^\ell \circ B, \quad l \geq 2, \tag{5.2}$$

then $l \leq n$, unless $A = \alpha z^n$, $\alpha \in \mathbb{C}$, and $B = \beta z^n$, $\beta \in \mathbb{C}$.

Proof. If

$$n \leq \frac{l-1}{2}, \tag{5.3}$$

then the set

$$(T_\ell \circ B)^{-1}_{\mathrm{odd}}\{-1, 1\} = B^{-1}_{\mathrm{odd}}\{-1, 1\}$$

contains at most $l - 1$ points. Therefore, if equality (5.1) holds, then the set

$$(A \circ T_\ell)^{-1}_{\mathrm{odd}}\{-1, 1\} \tag{5.4}$$

also contains at most $l - 1$ points. On the other hand, since -1 and 1 are the only finite critical values of T_n, if the set $A^{-1}_{\mathrm{odd}}\{-1, 1\}$ contains at least one point distinct from ± 1, then set (5.4) contains at least l points. Since by Lemma 4.3 the set $A^{-1}_{\mathrm{odd}}\{-1, 1\}$ contains at least two points, we conclude that if (5.3) holds, then

$$A^{-1}_{\mathrm{odd}}\{-1, 1\} = \{-1, 1\}. \tag{5.5}$$

Therefore, by Lemma 4.1, $A = \pm T_n$, It follows now from (5.1) that

$$\pm T_{nl} = T_\ell \circ B,$$

implying that

$$T_\ell \circ B = \pm T_\ell \circ T_n,$$

and applying to the last equality Theorem 2.3 we see that

$$T_\ell = \pm T_\ell \circ \mu, \quad B = \mu^{-1} \circ T_n, \tag{5.6}$$

for some polynomial μ of degree one. Finally, it is easy to see, using for example the explicit formula

$$T_n = \frac{n}{2} \sum_{k=0}^{[n/2]} (-1)^k \frac{(n-k-1)!}{k!(n-2k)!} (2x)^{n-2k}, \tag{5.7}$$

that T_n has non-zero coefficients of its terms of degree n and $n-2$, and the coefficient equal zero for its term of degree $n-1$. Thus, the first of equalities (5.6) implies the equality $\mu = \pm x$.

Assume now that equality (5.2) holds and $n \le l-1$. Then the polynomial in the right part of (5.2) has at most $l-1$ zeroes. On the other hand, since the unique finite critical value of z^ℓ is zero, it is easy to see that, unless

$$A = \alpha z^n, \quad \alpha \in \mathbb{C}, \tag{5.8}$$

the polynomial in the left part of (5.2) has at least l zeroes. Finally, (5.8) and (5.2) imply easily that $B = \beta z^n$, $\beta \in \mathbb{C}$. □

Theorem 5.2. *Let B be a non-special polynomial of degree $n \ge 2$, and X an element of $\mathcal{E}(B)$. Assume that $X = W_1 \circ z^\ell \circ W_2$ for some polynomials W_1, W_2 and $l \ge 1$. Then $l \le n$. Similarly, if $X = W_1 \circ \pm T_\ell \circ W_2$, then $l \le 2n$.*

Proof. If $X = W_1 \circ z^\ell \circ W_2$, then applying Corollary 3.4 twice we conclude that there exist polynomials $C_1$and C_2 such that the equalities

$$A \circ W_1 = W_1 \circ C_1, \quad C_1 \circ z^\ell = z^\ell \circ C_2, \quad C_2 \circ W_2 = W_2 \circ B \tag{5.9}$$

hold. Applying now Lemma 5.1 to the second equality in (5.9) we conclude that $l \le n$, unless C_1 and C_2 are conjugate to z^n. On the other hand, in the last case the third equality in (5.9) implies by Theorem 4.4 that B is conjugate to z^n. If $X = W_1 \circ \pm T_\ell \circ W_2$, the proof is similar. □

Corollary 5.3. *Let B be a non-special polynomial of degree $n \ge 2$. Assume that $B^{\circ d} = W_1 \circ z^\ell \circ W_2$ for some polynomials W_1, W_2, and $l \ge 1$, $d \ge 1$. Then $l \le n$. Similarly, if $B^{\circ d} = W_1 \circ \pm T_\ell \circ W_2$, then $l \le 2n$.*

Proof. Direct consequence of Theorem 5.2, since $B^{\circ d}$ is a semiconjugacy from B to B. □

5.2. Proof of Theorem 1.6

For natural numbers n and m define $l = l(n,m)$ as the maximum number coprime with n which divides m. Thus,

$$m = lb, \tag{5.10}$$

where $\text{rad}(b)|\text{rad}(n)$ and $\text{GCD}(n,l) = 1$. Define now $d = d(n,m)$ as the minimum number such that b in (5.10) satisfies $b \mid n^d$. The next proposition describes a general structure of elements of $\mathcal{E}(B)$ for non-special B.

Proposition 5.4. *Let B be a non-special polynomial of degree $n \geq 2$. Then any $X \in \mathcal{E}(B)$ has the form $X = \nu \circ z^{l(n,m)} \circ W$, where ν is a polynomial of degree one, and W is a compositional right factor of $B^{\circ d(n,m)}$. Furthermore, $l(n,m) < n$.*

Proof. Set $m = \deg X$, and let l, b, d be the numbers defined above. If A is a polynomial such that

$$A \circ X = X \circ B, \tag{5.11}$$

then the equality

$$A^{\circ d} \circ X = X \circ B^{\circ d}, \tag{5.12}$$

implies by Theorem 2.3 that

$$X = U \circ S, \quad B^{\circ d} = V \circ S, \tag{5.13}$$

for some polynomials U, V, S, where $\deg U = l$. Furthermore, equalities (5.11) and $X = U \circ S$ imply by Corollary 3.4 that

$$A \circ U = U \circ C \tag{5.14}$$

for some polynomial C. Since l is coprime with n, by Theorem 2.4 there exist polynomials μ, ν of degree one such that either

$$A = \nu \circ z^s R^{\ell}(z) \circ \nu^{-1}, \quad U = \nu \circ z^{\ell} \circ \mu, \quad C = \mu^{-1} \circ z^s R(z^{\ell}) \circ \mu,$$

where R is a polynomial, $n \geq 1$, $s \geq 0$, and $\text{GCD}(s,l) = 1$, or

$$A = \nu \circ \pm T_n \circ \nu^{-1}, \quad U = \nu \circ T_{\ell} \circ \mu, \quad C = \mu^{-1} \circ \pm T_n \circ \mu,$$

where $\text{GCD}(l,n) = 1$. In the last case however Theorem 4.4 applied to (5.11) implies that B is conjugate to T_n. Therefore, the first case must hold and hence $X = \nu \circ z^{\ell} \circ W$, where $W = \mu \circ S$ is a compositional right factor of $B^{\circ d}$. Moreover, since $n = rl + s$, where $r = \deg R$, the inequality $l < n$ holds whenever $r \neq 0$. On the other hand, if $r = 0$, then A is conjugate to z^n and hence B also is conjugate to z^n by Theorem 4.4. $\square$

For a natural number $n > 1$ with a prime decomposition $n = p_1^{a_1} p_2^{a_2} \dots p_s^{a_s}$ set $\text{ord}_p(n) = a_i$, if $p = p_i$ for some i, $1 \leq i \leq s$, and $\text{ord}_p\, n = 0$ otherwise.

Proposition 5.5. *If, under assumptions of Proposition* 5.4*, the polynomial X is not a polynomial in B, then $d(n,m) \le 2\log_2 n + 3$.*

Proof. Set

$$a = n^d/b. \tag{5.15}$$

Clearly, for any prime p,

$$\operatorname{ord}_p(b) + \operatorname{ord}_p(a) = \operatorname{ord}_p(n)d,$$

implying that

$$\operatorname{ord}_p b = \operatorname{ord}_p(n)(d-1) + \operatorname{ord}_p(n) - \operatorname{ord}_p(a). \tag{5.16}$$

Observe that the definition of $d(n,m)$ implies that a is not divisible by n. Moreover, the number b is not divisible by n either, since otherwise equality (5.11) implies by Theorem 2.3 that X is a polynomial in B. Observe also that by Theorem 4.4 any polynomial A such that (5.11) holds is not special.

It follows from Theorem 2.3 applied to equality (5.12) that there exist polynomials N, F and Y, Z, where

$$\deg Z = l, \quad \deg Y = a,$$

such that

$$A^{\circ d} = N \circ Y, \quad X = N \circ Z,$$

and

$$Y \circ X = Z \circ B^{\circ d}. \tag{5.17}$$

Applying now Theorem 2.3 and Theorem 2.2 to the equality

$$Y \circ X = \left(Z \circ B^{d-i}\right) \circ B^i$$

for each i, $1 \le i \le d-1$, we obtain a collection of polynomials Y_i, X_i, W_i U_i, K_i, L_i, $1 \le i \le d-1$, such that

$$Y = U_i \circ Y_i, \quad Z \circ B^{\circ d-i} = U_i \circ K_i, \quad X = X_i \circ W_i, \quad B^{\circ i} = L_i \circ W_i, \tag{5.18}$$

and

$$Y_i \circ X_i = K_i \circ L_i. \tag{5.19}$$

Furthermore,

$$\deg Y_i = a_i, \quad \deg X_i = lb_i,$$

where

$$a_i = \frac{a}{\operatorname{GCD}\left(a, n^{d-i}\right)}, \quad b_i = \frac{b}{\operatorname{GCD}\left(b, n^i\right)}, \tag{5.20}$$

and there exist polynomials of degree one ν_i, σ_i, μ_i $1 \leq i \leq d-1$, such that either

$$Y_i = \nu_i \circ z^{a_i} \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ z^c R(z^{a_i}) \circ \mu_i, \tag{5.21}$$

where $R \in \mathbb{C}[z]$ and $\mathrm{GCD}(c, a_i) = 1$, or

$$Y_i = \nu_i \circ z^c R^{lb_i}(z) \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ z^{lb_i} \circ \mu_i, \tag{5.22}$$

where $R \in \mathbb{C}[z]$ and $\mathrm{GCD}(c, lb_i) = 1$, or

$$Y_i = \nu_i \circ T_{a_i} \circ \sigma_i, \quad X_i = \sigma_i^{-1} \circ T_{lb_i} \circ \mu_i, \tag{5.23}$$

where $\mathrm{GCD}(a_i, lb_i) = 1$.

Observe first that

$$a_i \geq 2^i, \quad b_i \geq 2^{d-i}. \tag{5.24}$$

Indeed, since $n \nmid a$, there exists $p \in \mathrm{rad}(n)$ such that $\mathrm{ord}_p(n) - \mathrm{ord}_p(a) > 0$. Thus, $\mathrm{ord}_p(b) > \mathrm{ord}_p(n^{d-1})$ by (5.15), and hence for any i, $1 \leq i \leq d-1$, the equality

$$\mathrm{ord}_p\left(\mathrm{GCD}(b, n^i)\right) = \mathrm{ord}_p(n)i$$

holds. It follows now from (5.20) and (5.16) that

$$\mathrm{ord}_p(b_i) = \mathrm{ord}_p(b) - \mathrm{ord}_p\left(\mathrm{GCD}(b, n^i)\right) = \mathrm{ord}_p(n)(d-1-i) + \mathrm{ord}_p(n) - \mathrm{ord}_p(a),$$

implying that

$$b_i \geq p^{\mathrm{ord}_p(n)(d-1-i)+\mathrm{ord}_p(n)-\mathrm{ord}_p(a)} \geq p^{\mathrm{ord}_p(n)(d-1-i)+1} \geq p^{(d-1-i)+1} = p^{d-i}.$$

Similarly, since $n \nmid b$, there exists $q \in \mathrm{rad}(n)$ such that $\mathrm{ord}_q(n) - \mathrm{ord}_q(b) > 0$ implying by (5.20) and (5.16) that that for any i, $1 \leq i \leq d-1$, the inequality $a_i \geq q^i$ holds. Since $p \geq 2$, $q \geq 2$, this proves (5.24).

In order to establish the required bound, observe that since

$$A^{\circ d} = N \circ U_i \circ Y_i,$$

it follows from Corollary 5.3 that if (5.21) or (5.23) holds, then $a_i \leq 2n$. On the other hand, since $X = X_i \circ W_i$, if (5.22) or (5.23) holds, then $b_i \leq lb_i \leq 2n$, by Theorem 5.2. Thus, for any i, $1 \leq i \leq d-1$, the inequality

$$\min\{a_i, b_i\} \leq 2n$$

holds. On the other hand, it follows from (5.24) that for $i_0 = \lfloor d/2 \rfloor$ the inequality

$$\min\{a_i, b_i\} \geq 2^{\lfloor d/2 \rfloor}$$

holds. Therefore, $2^{\lfloor d/2 \rfloor} \leq 2n$, implying that $2^{d/2} \leq 2\sqrt{2}n$. Thus, $d/2 \leq \log_2 n + 3/2$ and $d \leq 2\log_2 n + 3$. □

Proof of Theorem 1.6. Observe first that if $X \in \mathcal{E}(B)$ is a semiconjugacy from B to A, then A is defined in a unique way since the equalities

$$A \circ X = X \circ B, \quad \widetilde{A} \circ X = X \circ B$$

imply $A \circ X = \widetilde{A} \circ X$, which in turn implies $A = \widetilde{A}$. In particular, this implies that for any $X_1, X_2 \in \mathcal{E}(B)$ such that $X_2 = \mu \circ X_1$ for some polynomial μ of degree one the corresponding polynomials $A_1, A_2 \in \mathcal{F}(B)$ are conjugate. Moreover, for any $A \in \mathcal{F}(B)$ there exists X such that

$$A \circ X = X \circ B \tag{5.25}$$

and X is not a polynomial in B, since (5.25) and $X = \widetilde{X} \circ B^{\circ s}$ imply that

$$A \circ \widetilde{X} = \widetilde{X} \circ B.$$

Finally, if $X_1, X_2 \in \mathcal{E}(B)$ and $\deg X_1 = \deg X_2$, then the corresponding polynomials in $A_1, A_2 \in \mathcal{F}(B)$ are conjugate, since Theorem 1.1 and Theorem 3.1 imply that there exists a polynomial μ of degree one such that $X_2 = \mu \circ X_1$.

Let X be an element of $\mathcal{E}(B)$ and $X = \nu \circ z^l \circ W$ its representation from Proposition 5.4. Then it follows from Proposition 5.5 that, unless X is a polynomial in B, the inequality $d \leq 2 \log_2 n + 3$ holds. Since, in addition, for the number l the inequality $l < n$ holds, this implies that up to the change $X \to \mu \circ X$, where μ is a polynomial of degree one, there exists at most a finite number of elements of $\mathcal{E}(B)$ which are not polynomials in B. Applying to these polynomials recursively Theorem 1.2 we obtain polynomials $X \in \mathcal{E}(B)$ and $A \in \mathcal{F}(B)$ which satisfy the conclusion of the theorem. □

Remark 5.6. Since the degree of the polynomial of X from Theorem 1.6 is equal to the least common multiple of degrees of all polynomials from $\mathcal{E}(B)$ which are not polynomials in B, it follows from Proposition 5.4 and Proposition 5.5 that $\deg X$ is bounded by the number $\psi(n) n^{2 \log_2 n + 3}$, where $\psi(n)$ denotes the least common multiple of all numbers less than n and coprime with n. In particular,

$$c(n) \leq (n-1)! n^{2 \log_2 n + 3}.$$

Corollary 5.7. *Let B be a polynomial of degree at least two. Then there exists at most a finite number of conjugacy classes of polynomials A such that $A \leq B$.*

Proof. If B is non-special, then the corollary follows from Theorem 1.6. For special B the corollary follows from Theorem 4.4. □

Corollary 5.8. *Each equivalence class of the relation $\sim$ contains at most a finite number of conjugacy classes.*

Proof. It follows from Corollary 5.7, since $A \sim B$ implies $A \leq B$. □

Corollary 5.9 ([22]). *Let B be a non-special polynomial of degree $n \geq 2$, and X and Y polynomials such that $Y \circ X = B^{\circ s}$ for some $s \geq 1$. Then there exist polynomials $\widetilde{X}$, $\widetilde{Y}$ and $i, j \geq 0$ such that*

$$Y = B^{\circ i} \circ \widetilde{Y}, \quad X = \widetilde{X} \circ B^{\circ j}, \quad \text{and} \quad \widetilde{Y} \circ \widetilde{X} = B^{\circ \widetilde{s}},$$

where $\widetilde{s}$ is bounded from above by a constant which depends on n only.

Proof. Clearly, without loss of generality we may assume that X is not a polynomial in B. Since $B \circ B^{\circ d} = B^{\circ d} \circ B$, the polynomial $B^{\circ d}$ is contained in $\mathcal{E}(B)$ and hence X is contained in $\mathcal{E}(B)$ by Corollary 3.4. Furthermore, since $\operatorname{rad}(\deg X) \mid \operatorname{rad}(n)$, it follows from Proposition 5.4 and Proposition 5.5 that there exists a polynomial $\widetilde{Y}$ such that $\widetilde{Y} \circ X = B^{\circ(2\log_2 n+3)}$. Therefore, if $s > 2\log_2 n + 3$, then

$$B^{\circ s} = B^{\circ(s-2\log_2 n-3)} \circ B^{\circ(2\log_2 n+3)} = B^{\circ(s-2\log_2 n-3)} \circ \widetilde{Y} \circ X = Y \circ X,$$

implying that $Y = B^{\circ(s-2\log_2 n-3)} \circ \widetilde{Y}$. This proves the corollary, and shows that $\widetilde{s} \leq 2\log_2 n + 3$. □

Remark 5.10. The bound $\widetilde{s} \leq 2\log_2 n + 3$ in Corollary 5.9 is not optimal. It was shown in [22] that in fact $\widetilde{s} \leq \log_2(n+2)$ and that this last bound cannot be improved. For more details we refer the reader to [22]. Notice however that for applications, similar to the ones given in [6], the actual form of the bound for $\widetilde{s}$ is not important.

References

[1] T. DINH, *Ensembles d'unicité pour les polynômes*, Ergodic Theory Dynam. Systems **22** (2002), 171–186.

[2] T. DINH, *Distribution des préimages et des points périodiques d'une correspondance polynomiale*, Bull. Soc. Math. France **133** (2005), 363–394.

[3] H. ENGSTROM, *Polynomial substitutions*, Amer. J. Math. **63** (1941), 249–255.

[4] A. EREMENKO, *Some functional equations connected with the iteration of rational functions* (Russian), Algebra i Analiz. **1** (1989), 102–116; translation in Leningrad Math. J. **1** (1990), 905–919.

[5] P. FATOU, *Sur l'iteration analytique et les substitutions permutables*, J. Math. Pures Appl. (9) **2** (1923), 343–384.

[6] D. GHIOCA, T. TUCKER and M. ZIEVE, *Linear relations between polynomial orbits*, Duke Math. J. **161** (2012), 1379–1410.

[7] H. INOU, *Extending local analytic conjugacies*, Trans. Amer. Math. Soc. **363** (2011), 331–343.

[8] G. JULIA, *Mémoire sur la permutabilité des fractions rationelles*, Ann. Sci. École Norm. Sup. **39** (1922), 131–215.

[9] A. MEDVEDEV, "Minimal sets in ACFA", Thesis (PhD) University of California, Berkeley. 2007. 96.

[10] A. MEDVEDEV and T. SCANLON, *Invariant varieties for polynomial dynamical systems*, Ann. of Math. **179** (2014), 81–177.

[11] J. MILNOR, "Dynamics in one Complex Variable", Princeton Annals in Mathematics, Vol. 160, Princeton, NJ: Princeton University Press, 2006.
[12] M. MUZYCHUK and F. PAKOVICH, *Jordan-Holder theorem for imprimitivity systems and maximal decompositions of rational functions*, Proc. Lond. Math. Soc. **102** (2011), 1–24.
[13] F. PAKOVICH, *On polynomials sharing preimages of compact sets, and related questions*, Geom. Funct. Anal. **18** (2008), 163–183.
[14] F. PAKOVICH, *Prime and composite Laurent polynomials*, Bull. Sci. Math. **133** (2009), 693–732.
[15] F. PAKOVICH, *Algebraic curves* $P(x) - Q(y) = 0$ *and functional equations*, Complex Var. Elliptic Equ. **56** (2011), 199–213.
[16] F. PAKOVICH, *Generalized "second Ritt theorem" and explicit solution of the polynomial moment problem*, Compos. Math. **149** (2013), 705–728.
[17] F. PAKOVICH, *On semiconjugate rational functions*, Geom. Funct. Anal. **26** (2016), 1217-1243.
[18] F. PAKOVICH, *Solution of the parametric Center Problem for Abel Equation*, J. Eur. Math. Soc. (JEMS) **19** (2017), 2343-2369.
[19] J. F. RITT, *On the iteration of rational functions*, Trans. Amer. Math. Soc. **21** (1920), 348–356.
[20] J. RITT, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.
[21] J. F. RITT, *Permutable rational functions*, Trans. Amer. Math. Soc. **25** (1923), 399–448.
[22] M. ZIEVE and P. MÜLLER, *On Ritt's polynomial decomposition theorem*, preprint, arXiv:0807.3578.

Department of Mathematics
Ben Gurion University of the Negev
P.O.B. 653 Beer Sheva
8410501 Israel
pakovich@math.bgu.ac.il