New York J. Math. 13 (2007) 89–96.

Commuting polynomials and self-similarity

Karl Zimmermann

ABSTRACT. Let F be an algebraically closed field of characteristic 0 and f(x) a polynomial of degree strictly greater than one in F[x]. We show that the number of degree k polynomials with coefficients in F that commute with f (under composition) is either zero or equal to the number of degree one polynomials with coefficients in F that commute with f. As a corollary, we obtain a theorem of E. A. Bertram characterizing those polynomials commuting with a Chebyshev polynomial.

Contents

1.	Introduction and summary of results	89
2.	Commuting polynomials	91
References		96

1. Introduction and summary of results

In the early 1920s, G. Julia [5] and J. F. Ritt [6] used topological and algebraic methods to study complex rational functions in one variable and in particular, those that commute under composition of functions. Even though their results are difficult to prove, the theorem classifying those *polynomials* that commute under composition is relatively easy to state and understand. We begin with a review of the ideas necessary to state the classification theorem. The review will provide context for our main results, most of which are related to the Julia–Ritt theorem but are obtained here using elementary techniques. Even though some of the results and definitions to follow make sense in a more general setting, for clarity of exposition, we make the assumption throughout the paper, that F is an algebraically closed field of characteristic 0, for example, the field of complex numbers.

Recall that the Chebyshev polynomials of the first kind, $T_n(x)$ where $n \ge 1$, are defined via

$$T_n(x) = \cos n(\cos^{-1}(x)).$$

In particular,

Received October 1, 2006.

ISSN 1076-9803/07

Mathematics Subject Classification. 12Y05.

Key words and phrases. Polynomial, commute, field, root of unity, Chebyshev polynomial.

$$T_{1}(x) = x$$

$$T_{2}(x) = 2x^{2} - 1$$

$$T_{3}(x) = 4x^{3} - 3x$$

$$\vdots$$

$$T_{n}(x) = 2xT_{n-1}(x) - T_{n-2}(x).$$

Using the defining equation above, it is easy to verify that $T_n \circ T_m = T_m \circ T_n$ for all $m, n \ge 1$. Note also that the recursion formula shows that T_n has degree n, leading coefficient 2^{n-1} (for $n \ge 1$), and that T_n is a sum of odd (respectively, even) powers of x when n is odd (respectively, even). This last clause can be restated: as a function on [-1, 1], T_n is an odd function when n is odd, an even function when n is even.

Conjugation by a degree one polynomial will prove to be a useful tool in studying commuting polynomials. Note that it is precisely the polynomials $\lambda(x) = ax + b \in F[x]$, $a \neq 0$ that are invertible in F[x] under composition.

Definition 1.1. Let $P(x), Q(x) \in F[x]$. Then P is similar to Q if and only if there exists $\lambda(x) = ax + b \in F[x], a \neq 0$ so that $Q(x) = \lambda^{-1}(P(\lambda(x)))$. If $f(x) \in F[X]$ is similar to itself via $\lambda(x) \in F[x]$ then λ is said to be a self-similarity of f. We let S_f denote the set of self-similarities of f.

Some easy consequences of the definition of similarity are:

- (1) Similarity is an equivalence relation
- (2) If P is similar to Q then $\deg(P) = \deg(Q)$
- (3) If f commutes with P, then $\lambda^{-1}(f(\lambda(x)))$ commutes with $\lambda^{-1}(P(\lambda(x)))$
- (4) S_f is a group under the operation of composition.

Like the set of Chebyshev polynomials, the set $\{P_n(x) = x^n\}_{n\geq 1}$ of monic mononomials, consists of pairwise commuting polynomials. These two sets play a major role in the classification theorem of Ritt and Julia.

Theorem 1.2 (Ritt–Julia). Let P(x) and Q(x) be nonlinear, nonconstant polynomials with coefficients in the complex numbers that commute under composition. Then one of the following holds:

- (1) P and Q are similar, via the same map λ , to Chebyshev polynomials.
- (2) P and Q are similar, via the same map λ , to monomials.
- (3) There exists a polynomial G(x) = xR(x^r) where R(x) is a polynomial, such that P and Q are similar, via the same λ, to ε₁G^a(x) and ε₂G^b(x) where ε₁ and ε₂ are rth roots of unity and G^a = G ∘ G ∘ · · · ∘ G, the a-fold iterate of G under composition.

Julia replaces (3) above with:

(3') There exist positive integers ν and μ such that $P^{\nu} = Q^{\mu}$.

For more about the statement of the above theorem and the relationship of commuting polynomials to Julia sets, please see [1].

A set of pairwise commuting polynomials containing one polynomial of each positive degree is called a chain. It is a consequence of Theorem 1.2 that up to similarity, there are exactly two chains, namely the Chebyshev polynomials and the monic monomials. Since the proof of the Ritt-Julia theorem is rather difficult, it is worth noting that in 1951, H. D. Block and H. P. Thielman [3] gave an elementary proof of this result for chains defined over an integral domain containing the integers. In 1955, E. J. Jacobsthal [4] also gave an elementary proof.

Given a polynomial, it is natural to ask about the set of polynomials with which it commutes. To be more precise, we make the following definition.

Definition 1.3. Let $k \ge 1$. For $f(x) \in F[x]$, we let $C_k(f)$ denote the set of degree k polynomials in F[x] that commute with f.

Our main result describes a relationship between $C_k(f)$, $k \ge 1$ and $C_1(f) = S_f$, the group of self-similarities of f.

Theorem 2.9. Let f(x) be a polynomial of degree strictly greater than one in F[x] and suppose that $C_k(f)$ is nonempty for some $k \ge 1$. Then the number of elements in $C_k(f)$ is the order of the group S_f .

Theorem 2.9 follows from a more general theorem relating commuting polynomials and self-similarities.

Theorem 2.7. Let f(x) be a polynomial in F[x] of degree n > 1 and assume $P, Q \in C_k(f)$ for some $k \ge 1$. Then $Q = \lambda_f \circ P$ where $\lambda_f \in S_f$.

Using Theorem 2.9, we also obtain an easy proof of the known result that a polynomial of degree n > 1 with coefficients in F can commute with at most n-1 polynomials of fixed degree $k \ge 1$ with coefficients in F (Corollary 2.10). To illustrate Corollary 2.10 and motivate our final result, note that the monomial $P_n(x)$ commutes with $R_{k,j}(x) = \rho^j x^k, j = 1, \ldots, n-1$, where ρ is a primitive $(n-1)^{\text{st}}$ root of unity. By Corollary 2.10, there can be no other degree k polynomials commuting with P_n . For the Chebyshev polynomials, it is clear that $T_n(x)$ commutes with $\pm T_k(x)$ if n is odd, and $T_k(x)$ if n is even, but it is not immediately obvious whether or not there are other polynomials of degree k with which T_n commutes. It turns out that there are no others, and Bertram [2] gives an elementary proof of this fact using a differential equation satisfied by the Chebyshev polynomials. Below, we give a different proof of Bertram's theorem (our Theorem 2.15), which follows as a corollary to Theorem 2.9 and Corollary 2.14.

Corollary 2.14. Let f(x) be a polynomial in F[x] of degree strictly greater than one and let m > 1 be an integer. Then S_f has m elements if and only if f is similar to a polynomial of the form $xg(x^m)$ with $g(x) \in F[x]$ but not similar to any polynomial of the form $xh(x^r)$ where $h(x) \in F[x]$ and r > m.

2. Commuting polynomials

We begin with the observation that if f is similar to g via the similarity λ , then for $k \geq 1$, there is a mapping, $\Phi_k : C_k(f) \to C_k(g)$ defined by $\Phi_k(P) = \lambda^{-1} \circ P \circ \lambda$, where $P \in C_k(f)$. For each $k \geq 1$, Φ_k is easily seen to be a bijection. Moreover, $\Phi_1 : C_1(f) \to C_1(g)$ is an isomorphism of groups; in other words, the self-similarity groups S_f and S_g are isomorphic.

This observation will be particularly useful when used in conjunction with the following definition and proposition.

Definition 2.1. The polynomial $f(x) = \sum_{k=0}^{n} a_k x^k$ is centered if $a_n \neq 0$ and $a_{n-1} = 0$.

In the proof below, we use the notation $O(x^k)$ to denote any polynomial with degree less than or equal to k.

Proposition 2.2. Let f(x) be a polynomial in F[x] of degree n > 1. Then f is similar to a monic, centered polynomial.

Proof. Let $f(x) = \sum_{j=0}^{n} a_j x^j$ and let g be similar to f via $\lambda(x) = ax + b$. A compute the second sec

tation shows

$$g(x) = a^{n-1}a_n x^n + a^{n-2}(na_n b + a_{n-1})x^{n-1} + O(x^{n-2}).$$

For g to be monic, let a be an $(n-1)^{st}$ root of a_n^{-1} . For g to be centered, let

$$b = -\frac{a_{n-1}}{na_n}.$$

The following proposition and corollary focus on polynomials that commute with a centered polynomial.

Proposition 2.3. Let f(x) be a polynomial in F[x] of degree n > 1 and P an element of $C_k(f)$ for some $k \ge 1$. Assume f is monic and centered and $P(x) = \sum_{i=0}^{k} b_i x^i$. Then:

- (1) For some j = 1, 2, ..., n-1, $b_k = \rho^j$ where ρ is a primitive $(n-1)^{\text{st}}$ root of unity.
- (2) $b_{k-1} = 0.$

Proof. The first result follows easily by comparing the degree kn coefficients in $f \circ P = P \circ f$. Using this result and comparing the degree kn - 1 coefficients yields the second.

Corollary 2.4. Let f(x) be a polynomial in F[x] of degree n > 1 and $P(x) \neq x$ an element of $C_k(f)$ for some $k \ge 1$. Then f is centered if and only if P is centered.

Proof. Assume f is centered. By the proof of Proposition 2.2, f is similar to a monic polynomial via $\lambda(x) = ax$. Such a similarity does not affect centering, so without loss of generality, we may assume f is centered and monic. Then P is centered by part (2) of Proposition 2.3.

A similar argument proves the converse unless $\deg(P) = 1$, that is, unless P(x) = bx, where $b \neq 0, 1$. However, if $f(x) = a_n x^n + a_{n-1} x^{n-1} + O(x^{n-2})$, and $a_n \neq 0$ and $a_{n-1} \neq 0$ then f(bx) = bf(x) yields $b^n = b = b^{n-1}$ which implies b = 0, 1. Thus, f must be centered.

Corollary 2.5. Let f(x) be a monic, centered polynomial in F[x] of degree n > 1. Then any self-similarity λ of f, is of the form $\lambda(x) = \rho^j x$ where ρ is a primitive $(n-1)^{\text{st}}$ root of unity.

Proof. Since f commutes with λ , this follows immediately from Proposition 2.3.

Corollary 2.6. Let f(x) be a polynomial in F[x] of degree n > 1. Then the order of S_f is less than or equal to n - 1.

Proof. By Proposition 2.2, f is similar to a monic, centered polynomial g. The result follows since the order of S_f equals the order of S_g which is less than or equal to n-1 by Corollary 2.5.

This sets the stage for the main results of the paper.

Theorem 2.7. Let f(x) be a polynomial in F[x] of degree n > 1 and assume $P, Q \in C_k(f)$ for some $k \ge 1$. Then $Q = \lambda_f \circ P$ where $\lambda_f \in S_f$.

Proof. By Proposition 2.2, f is similar to a monic, centered polynomial g via a similarity $\lambda(x) \in F[x]$. If $g(x) = x^n + \sum_{m=0}^{n-2} a_m x^m$ and $\hat{P} = \lambda^{-1} P \lambda$ and $\hat{Q} = \lambda^{-1} Q \lambda$, then, using Proposition 2.3, we see that

$$\hat{P}(x) = \rho^{i} x^{k} + \sum_{m=0}^{k-2} b_{m} x^{m}$$
 and
 $\hat{Q}(x) = \rho^{j} x^{k} + \sum_{m=0}^{k-2} c_{m} x^{m},$

where all coefficients of the polynomials are in F and ρ is a primitive $(n-1)^{\text{st}}$ root of unity. Let $\alpha = j - i$; we will show that $\hat{Q} = \rho^{\alpha} \hat{P}$ and that $\lambda_g(x) = \rho^{\alpha} x$ is a self-similarity of g.

To that end, let $r(x) = \rho^{\alpha} \hat{P}(x) - \hat{Q}(x)$, and note that $\deg(r(x)) = t < k$ or r(x) = 0. Then, setting $a_n = 1$ and $a_{n-1} = 0$, we have

$$\begin{aligned} r(g(x)) &= \rho^{\alpha} P(g(x)) - Q(g(x)) = \rho^{\alpha} g(P(x)) - g(Q(x)) \\ &= \rho^{\alpha} \sum_{m=0}^{n} a_{m}(\hat{P}(x))^{m} - \sum_{m=0}^{n} a_{m}(\hat{Q}(x))^{m} \\ &= a_{n}(\rho^{\alpha}(\hat{P}(x))^{n} - (\hat{Q}(x))^{n}) + O(x^{k(n-2)}) \\ &= (\rho^{\alpha} \hat{P}(x) - \hat{Q}(x)) \left(\sum_{s=0}^{n-1} (\rho^{\alpha} \hat{P}(x))^{n-1-s} (\hat{Q}(x))^{s} \right) + O(x^{k(n-2)}) \end{aligned}$$

Note that

$$\deg\left[\sum_{s=0}^{n-1} (\rho^{\alpha} \hat{P}(x))^{n-1-s} (\hat{Q}(x))^s\right] = k(n-1),$$

since the leading coefficient, that is, the degree k(n-1) coefficient, is

$$\sum_{s=0}^{n-1} (\rho^{\alpha})^{n-1-s} (\rho^{j})^{n-1-s} (\rho^{j})^{s} = \sum_{s=0}^{n-1} (\rho^{j})^{n-1-s} (\rho^{j})^{s} = \sum_{s=0}^{n-1} 1 = n.$$

Comparing degrees in the two sides of the equation for r(g(x)), we see that $tn = \deg(r(g(x))) = t + k(n-1)$, which implies that t = k. This is a contradiction since t < k, and therefore r(x) = 0 and $\hat{Q} = \rho^{\alpha} \hat{P}$.

To see that $\lambda_g(x) = \rho^{\alpha} x$ is a self-similarity of g, observe that $g(\hat{Q}(x)) = \hat{Q}(g(x))$ implies

$$g(\rho^{\alpha}\hat{P}(x)) = \rho^{\alpha}\hat{P}(g(x)) = \rho^{\alpha}g(\hat{P}(x)).$$

Letting $\hat{P}(x) = u$, we have $g(\rho^{\alpha}u) = \rho^{\alpha}g(u)$ as desired.

We define $\lambda_f = \lambda \circ \lambda_g \circ \lambda^{-1} = \Phi_1^{-1}(\lambda_g)$ where λ is the similarity chosen at the beginning of the proof and $\Phi_1 : S_f \to S_g$ is the isomorphism described at the

beginning of this section. Thus, $\lambda_f \in S_f$ and $\lambda_g = \Phi_1(\lambda_f) = \lambda^{-1} \circ \lambda_f \circ \lambda$. It follows that

$$\hat{Q} = \lambda_g \circ \hat{P}$$
$$\lambda \circ \hat{Q} \circ \lambda^{-1} = \lambda \circ \lambda_g \circ \hat{P} \circ \lambda^{-1}$$
$$\lambda \circ \lambda^{-1} \circ Q \circ \lambda \circ \lambda^{-1} = \lambda \circ \lambda^{-1} \circ \lambda_f \circ \lambda \circ \lambda^{-1} \circ P \circ \lambda \circ \lambda^{-1}$$
$$Q = \lambda_f \circ P$$

which completes the proof.

The idea of the proof above is similar to that used by Rivlin [7, Theorem 4.2, page 194] to prove that a degree 2 polynomial with coefficients in the real numbers can commute with at most one polynomial of a given degree. The generalization of this result to polynomials of arbitrary degree greater than one follows as an easy corollary to Theorem 2.9 below.

Lemma 2.8. Let f(x) and $\lambda(x)$ be polynomials in F[x] with $\deg(f) > 1$ and $\deg(\lambda) = 1$. Assume $P \in C_k(f)$ for some $k \ge 1$. Then $\lambda \circ P \in C_k(f)$ if and only if $\lambda \in S_f$.

Proof. Assume $\lambda \circ P \in C_k(f)$, that is, $\lambda \circ P \circ f = f \circ \lambda \circ P$. Composition with λ^{-1} yields $P \circ f = \lambda^{-1} \circ f \circ \lambda \circ P$ and so

$$f \circ P = P \circ f = \lambda^{-1} \circ f \circ \lambda \circ P.$$

Thus, $f = \lambda^{-1} \circ f \circ \lambda$, and $\lambda \in S_f$. Conversely, assume $\lambda \in S_f$. Then

$$f \circ \lambda \circ P = \lambda \circ \lambda^{-1} \circ f \circ \lambda \circ P$$
$$= \lambda \circ f \circ P$$
$$= \lambda \circ P \circ f$$

as desired.

Theorem 2.9. Let f(x) be a polynomial of degree strictly greater than one in F[x] and suppose that $C_k(f)$ is nonempty for some $k \ge 1$. Then the number of elements in $C_k(f)$ is the order of the group S_f .

Proof. Let $P \in C_k(f)$ and define $\Psi : S_f \to C_k(f)$ by $\Psi(\lambda) = \lambda \circ P$. By Lemma 2.8, $\Psi(\lambda) \in C_k(f)$, and since P is nonconstant, it follows easily that Ψ is injective. That Ψ is a surjection, follows immediately from Theorem 2.7.

Theorem 2.9 and Corollary 2.6 now give the following:

Corollary 2.10. Fix an integer $k \ge 1$ and let f(x) be a polynomial in F[x] of degree n > 1. Then there are at most n - 1 polynomials of degree k that commute with f.

It is natural to ask exactly which polynomials $f \in F[x]$ admit a nontrivial self-similarity group S_f . The answer is based on an obvious generalization of odd polynomials. Recall that a polynomial $f \in F[x]$ is odd if and only if f(-x) = -f(x); that is, if and only if $\lambda(x) = -x$ is a self-similarity of f.

Definition 2.11. Let $f(x) \in F[x]$ and m > 1 an integer. Then f is model if $f(\omega x) = \omega f(x)$ where ω is a primitive m^{th} root of unity.

94

Said another way, f is m-odd if and only if $\lambda(x) = \omega x$ is a self-similarity of f. Recall that f is an odd polynomial if and only if $f(x) = xg(x^2)$ where $g(x) \in F[x]$. Similarly, it is easy to show that f is m-odd if and only if $f(x) = xh(x^m)$ where $h(x) \in F[x]$. Thus, a polynomial is odd if and only if it is 2-odd. Moreover, any mk-odd polynomial is also m-odd.

Lemma 2.12. Let r > 1 and m > 1 be integers with least common multiple t. Assume $f \in F[x]$. Then f is r-odd and m-odd if and only if f is t-odd.

Proof. Assume f is both r-odd and m-odd. Then

$$\frac{f(x)}{x} = g(x^m) = h(x^r).$$

Thus, nonzero coefficients of f(x)/x may occur only in terms of degree a nonnegative multiple of both m and r, that is, in terms of degree a multiple of t. Therefore,

$$\frac{f(x)}{x} = p(x^t).$$

The converse follows immediately.

In order to characterize those polynomials with nontrivial self-similarity group, we first note that by the proof of Proposition 2.2, it follows that a similarity between centered polynomials must be of the form $\lambda(x) = ax$ where a is a nonzero element of F. Thus, if f and g are similar centered polynomials, the nonzero coefficients of f correspond precisely to the nonzero coefficients of g. But *m*-odd polynomials are centered. Therefore, a given centered polynomial h is similar to an *m*-odd polynomial if and only if h is *m*-odd.

Theorem 2.13. Let f(x) be a polynomial in F[x] of degree strictly greater than one and let m > 1 be an integer. Then f is similar to an m-odd polynomial but not an mk-odd polynomial for any k > 1 if and only if S_f has m elements.

Proof. By the remarks preceding the theorem, we may assume that f is monic, centered and m-odd (but not mk-odd). Therefore, $f(\omega x) = \omega f(x)$ where ω is a primitive m^{th} root of unity, which implies

$$S_f \supseteq \{\lambda_j(x) = \omega^j x \mid j = 1, 2, \dots, m\}.$$

Now let $\lambda(x) = \rho x \in S_f$ where ρ is an r^{th} root of unity. Then f is m-odd and r-odd and so by Lemma 2.12 it is t-odd, where t is the least common multiple of r and m. By assumption, we must have t = m whence r divides m and $\rho = \omega^j$ for some $j = 1, 2, \ldots, m$. Thus, S_f has m elements.

To prove the converse, let S_f have m elements and let g be a monic, centered polynomial similar to f. Then g has exactly m self-similarities, and all must be of the form $\lambda(x) = \rho x$ where ρ is a root of unity. Since the order of the group S_g equals the order of S_f which is m, we must have $\rho^m = 1$ for all elements of S_g . Thus g is m-odd. If g were mk-odd, where k > 1, then S_g and S_f would have more than m elements, a contradiction.

This result can be written in a slightly different form.

Corollary 2.14. Let f(x) be a polynomial in F[x] of degree strictly greater than one and let m > 1 be an integer. Then S_f has m elements if and only if f is

similar to a polynomial of the form $xg(x^m)$ with $g(x) \in F[x]$ but not similar to any polynomial of the form $xh(x^r)$ where $h(x) \in F[x]$ and r > m.

The question of which f have nontrivial self-similarity groups S_f becomes more interesting when f is a rational function and S_f is a subgroup of $\mathrm{PGL}_2(F) = \mathrm{Aut}(\mathbf{P}^1)$. In this situation, the groups that appear are the classical symmetry groups of the regular solids (see [8]).

We now apply the above results to the Chebyshev polynomials.

Theorem 2.15 (Bertram). Let $\{T_n\}_{n>1}$ be the sequence of nonlinear Chebyshev polynomials, and let $P(x) \in F[x]$ with $\deg(P) = k > 1$. If P commutes with at least one T_n , then $P = T_k$ if n is even, and $P = \pm T_k$ if n is odd.

Proof. Each polynomial in the sequence $\{T_n\}_{n>1}$ is centered. Observe that if n is even, then T_n is not m-odd for any m and by the remarks preceding Theorem 2.13, not similar to an m-odd polynomial for any m. Thus T_n admits no nontrivial self-similarities and $P = T_k$. If n is odd, T_n is 2-odd and again, by the remarks above, not similar to an m-odd polynomial for any m > 2. Thus $S_{T_n} = \{x, -x\}$ which implies $P = \pm T_k$.

Acknowledgements. The author extends sincere thanks to Andrew Browder, Joseph Silverman, and the referee for their very helpful comments.

References

- ATELA, P.; HU, J. Commuting polynomials with the same Julia set. International Journal of Bifurcation and Chaos 6 (1996), no. 12A, 2427–2432.
- [2] BERTRAM, E. A. Polynomials which commute with a Tchebycheff polynomial. Amer. Math. Monthly 78 (1971), no. 6, 650–653. MR0288104 (44 #5302), Zbl 0213.32103.
- [3] BLOCK, H. D.; THIELMAN, H. P. Commutative polynomials. Quart. J. Math. Oxford Ser. (2) 2 (1951), 241–243. MR0045250 (13,552f), Zbl 0044.00803.
- [4] JACOBSTHAL, ERNST. Über vertauschbare Polynome. Math. Z. 63 (1955) 243–276. MR0074373 (17,574a), Zbl 0065.24903.
- [5] JULIA, GASTON. Mémoire sur la permutabilité des fractions rationnelles. Ann. Sci. École Norm, Sup. (3) 39 (1922), 131–215. MR1509242, JFM 48.0364.02.
- [6] RITT, J. F. Permutable rational functions. Trans. Amer. Math. Soc. 25 (1923), no. 3, 399–488. MR1501252, JFM 49.0712.02. Errata Trans. Amer. Math. Soc. 26 (1924), no. 4, 494. MR1500495.
- [7] RIVLIN, THEODORE J. Chebyshev polynomials. From approximation theory to algebra and number theory. Second edition. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1990. MR1060735 (92a:41016), Zbl 0734.41029.
- [8] SILVERMAN, JOSEPH H. The field of definition for dynamical systems on P¹. Compositio Math. 98 (1995), no. 3, 269–304. MR1351830 (96j:11090), Zbl 0849.11090.

DEPARTMENT OF MATHEMATICS, UNION COLLEGE, SCHENECTADY, NY 12308, USA zimmermk@union.edu

This paper is available via http://nyjm.albany.edu/j/2007/13-5.html.