PERMUTABLE POLYNOMIALS AND RELATED TOPICS

H.Woracek G.Eigenthaler

1 Introduction

Let $k$ be a field and $k[x]$ the polynomial ring over $k$ in one indeterminate. Consider the semigroup $\langle k[x], \circ \rangle$ where $\circ$ denotes the composition of polynomials defined as $(f \circ g)(x) = f(g(x))$. The semigroup $\langle k[x], \circ \rangle$ is not commutative, i.e. the equality $f \circ g = g \circ f$ does not hold in general. Thus the question arises which specific polynomials $f$ and $g$ permute, i.e. satisfy $f \circ g = g \circ f$, and, in a more structural setting, which subsemigroups of $\langle k[x], \circ \rangle$ are commutative.

We consider mainly the case that $k = \mathbb{C}$, the field of complex numbers, as in this case a result of J.Ritt (see [38]) answers the first mentioned question. This leads us to an answer to the second question in form of a complete list of commutative subsemigroups of $\langle k[x], \circ \rangle$ for $k = \mathbb{C}$ (see Theorem 2).

In §2 we recall some basic facts concerning the semigroup $\langle k[x], \circ \rangle$. We introduce the notion of the gap-degree of a polynomial in §3. The gap-degree is an invariant under conjugation and provides a convenient tool for working with permutable polynomials. In §4 we give a classification of commutative subsemigroups of $\langle \mathbb{C}[x], \circ \rangle$ which contain no linear polynomials. Together with some well known results this leads us to a complete list of commutative subsemigroups of $\langle \mathbb{C}[x], \circ \rangle$ (see List IV in §5). Finally, in §6 we apply our results to some related questions, namely the classification of maximal commutative subsemigroups and the classification of generalized permutable chains.

We would like to point out that, as we refer to certain results of J.Ritt, in most parts of the paper $k$ equals $\mathbb{C}$. Apart from this some methods developed in this note depend only on the fact that $k$ is of characteristic zero and algebraically closed; in particular, this is the case for §3.

As a complete proof of the presented results employing only purely algebraic methods is still missing we have tried to collect literature concerning the subject of permutable polynomials and related topics, including various results in the case that $\text{char } k \neq 0$ and some papers which give connections to other areas.

2 Some basic facts about the semigroup $\langle k[x], \circ \rangle$

For $f \in k[x]$ denote by $[f]$ the (exact) degree of $f$. It is well known that $[.] : \langle k[x], \circ \rangle \rightarrow \langle \mathbb{N}_0, \cdot \rangle$ is a homomorphism (as $k$ denotes a field, $k$ contains no zero divisors). The polynomial $x$ is neutral with respect to composition and, again due to the fact that $k$ is a field, the units of $\langle k[x], \circ \rangle$ are exactly the linear polynomials. Furthermore the inverse of the polynomial $L(x) = ax + b$ is easily seen to be $\frac{1}{a}x - \frac{b}{a}$; we will denote the inverse of $L$ by $L^{-1}$. The $n$-th iterate of a polynomial $f$ will be abbreviated as $f^{(n)}$, i.e.

$$f^{(n)} = f \circ \cdots \circ f.$$
The operation of composition is right-distributive with respect to addition and multiplication of polynomials, i.e.

\[(f + g) \circ h = (f \circ h) + (g \circ h)\] and

\[(f \cdot g) \circ h = (f \circ h) \cdot (g \circ h)\]

holds. From this and the fact that \([\cdot]\) is a homomorphism it follows easily that every nonconstant polynomial is right-regular. On the other hand not every nonconstant polynomial is left-regular.

Given any invertible polynomial \(L(x) = ax + b\), we can define an (inner) automorphism \(\Phi_L\) of \(\langle k[x], \circ \rangle\) as \(\Phi_L(f) = L^{(-1)} \circ f \circ L\) for \(f \in k[x]\). Clearly the relations \(\Phi_L \circ \Phi_M = \Phi_{M \circ L}\) and \((\Phi_L)^{-1} = \Phi_{L^{-1}}\) are satisfied. Thus the set of all inner automorphisms forms a group acting on \(k[x]\), antiisomorphic to the group of linear polynomials.

We call two polynomials \(f\) and \(g\) conjugates of each other, if they belong to the same domain of transitivity with respect to this group, i.e. if there exists a linear polynomial \(L\), such that \(\Phi_L(f) = g\) holds. At some stage it is convenient to consider a coarser equivalence relation: We call two polynomials \(f\) and \(g\) weakly conjugates of each other, if there exist linear polynomials \(L_1\) and \(L_2\), such that \(L_1 \circ f \circ L_2 = g\) holds.

A polynomial \(f\) permutes with a constant polynomial \(a\) if and only if \(a\) is a fixed point of \(f\): \(f(a) = a\). Thus any commutative subsemigroup of \(\langle k[x], \circ \rangle\) contains at most one constant. On the other hand, given any commutative subsemigroup \(S\) we can extend \(S\) to a - still commutative - semigroup \(S' = S \cup \{a\}\) if and only if \(a\) is a common fixed point of all polynomials of \(S\). In view of this fact we can confine our attention to commutative semigroups which contain no constant polynomials. In the sequel the notion of commutative semigroup will always be understood in this manner.

A straightforward argument using Zorn’s Lemma shows that every commutative subsemigroup of \(\langle k[x], \circ \rangle\) is contained in some maximal commutative subsemigroup.

If \(f\) permutes with \(g\), then obviously \(f\) permutes also with every power \(g^{(n)}\) of \(g\). The converse statement is not true, as the following example shows: Let \(f(x) = \zeta x\) where \(\zeta\) is a primitive third root of unity and let \(g(x) = x^2\). Then

\[(x^2)^{(2)} \circ \zeta x = x^4 \circ \zeta x = \zeta x \circ x^4 = \zeta x \circ (x^2)^{(2)},\]

but

\[x^2 \circ \zeta x = \zeta^2 x \circ x^2 \neq \zeta x \circ x^2.\]

But at least the following result, which will be used in the subsequent sections, holds.

**Lemma 1** Let \(f\) permute with \(g^{(n_i)}\) for \(n_i \in \mathbb{N}, i = 1, \ldots, k\) and let \(d = \gcd(n_1, \ldots, n_k)\). Then \(f\) permutes with \(g^{(d)}\).

**Proof:** For \([g] = 0\) the statement is trivial, so let \([g] > 0\) and suppose w.l.o.g. \(d = \lambda_1 n_1 + \ldots + \lambda_i n_i + \ldots - \lambda_k n_k\) with \(\lambda_i \geq 0\). As \(f\) permutes with each \(g^{(n_i)}\) we find

\[f \circ g^{(\lambda_1 n_1 + \ldots + \lambda_i n_i)} = g^{(\lambda_1 n_1 + \ldots + \lambda_i n_i)} \circ f\]

and

\[f \circ g^{(\lambda_1 n_1 + \ldots + \lambda_k n_k)} = g^{(\lambda_1 n_1 + \ldots + \lambda_k n_k)} \circ f.\]

Thus

\[f \circ g^{(d)} \circ g^{(\lambda_1 n_1 + \ldots + \lambda_k n_k)} = f \circ g^{(\lambda_1 n_1 + \ldots + \lambda_i n_i)} = g^{(\lambda_1 n_1 + \ldots + \lambda_i n_i)} \circ f = \]
\[ = g^{(d)} \circ g^{(\lambda_{r+1}n_{r+1}+\ldots+\lambda_kn_k)} \circ f = g^{(d)} \circ f \circ g^{(\lambda_{r+1}n_{r+1}+\ldots+\lambda_kn_k)}, \]
which implies \( g^{(d)} \circ f = f \circ g^{(d)}. \)

\[ \square \]

### 3 The gap-degree of a polynomial

In this section we study an invariant of a polynomial under conjugation. The results proved in the sequel provide some tools for the study of permutable polynomials and give an insight to the behaviour of linear polynomials when composed with another (arbitrary) polynomial.

**Definition 1** Let \( f \in \mathbb{k}[x], \ |f| \geq 1. \) Consider the field extension \( \mathbb{k}(x) : \mathbb{k} \) and the subring \( \mathbb{k}[f]. \) Denote by \( \mathcal{G}(f) \) the group of all automorphisms of \( \mathbb{k}(x) : \mathbb{k} \) which leave \( \mathbb{k}[f] \) (as a whole) invariant. The number \( \mathcal{L}(f) = |\mathcal{G}(f)| \) will be called the gap-degree of \( f. \)

A well known theorem (see e.g. [43]) states that the automorphisms of \( \mathbb{k}(x) : \mathbb{k} \) are exactly those endomorphisms which map \( x \) to some linear fractional element \( L(x) = \frac{ax + b}{cx + d} \) of \( \mathbb{k}(x). \) Denote by \( \alpha_L \) the automorphism of \( \mathbb{k}(x) : \mathbb{k} \) with \( \alpha_L : x \mapsto L(x), \) i.e. \( \alpha_L(f) = f \circ L. \)

**Lemma 2** The group \( \mathcal{G}(f) \) consists exactly of those automorphisms \( \alpha_L, \) where \( L \) is a linear polynomial and satisfies an equation of the form

\[ L_1 \circ f = f \circ L \quad (1) \]

for some linear polynomial \( L_1. \)

**Proof:** The condition \( \alpha_L(\mathbb{k}[f]) \subseteq \mathbb{k}[f] \) implies that \( \alpha_L(f) = f \circ L \in \mathbb{k}[f]. \) Thus \( f \circ L \) is a polynomial (in the indeterminate \( x \), which implies that \( L \) is a polynomial. Furthermore \( f \circ L = L_1 \circ f \) where \( L_1 \) is a polynomial, which then must have degree 1.

Conversely suppose that (1) holds for some linear polynomial \( L_1. \) Then \( \alpha_L(\mathbb{k}[f]) = \mathbb{k}[f], \) as for arbitrary \( g \circ f \in \mathbb{k}[f] \) the formulas

\[ \alpha_L(g \circ f) = g \circ (f \circ L) = (g \circ L_1) \circ f \]

and

\[ \alpha_L(L_1^{-1} \circ f) = L_1^{-1} \circ f \circ L = L_1^{-1} \circ L_1 \circ f = f \]

hold.

\[ \square \]

We have \( \alpha_{L \circ M} = \alpha_M \circ \alpha_L, \) hence \( L \mapsto \alpha_L(-1) \) defines a group-isomorphism. In view of this fact and the above lemma we can consider \( \mathcal{G}(f) \) in the following as a group of linear polynomials.

Before we proceed to determine the group \( \mathcal{G}(f) \) for a given polynomial \( f, \) we introduce another notion.

**Definition 2** Let \( f \in \mathbb{k}[x], \) \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 \) with \( n = |f| \geq 2. \) Denote by \( l(f) \) the smallest number \( k \geq 1, \) such that \( a_{n-k} \neq 0. \) In the case \( f(x) = a_n x^n \) let \( l(f) = n. \) If \( l(f) > 1, \) i.e. \( a_{n-1} = 0, \) we say that \( f \) has gap-form. We call the number \( l(f) \) the gap of \( f. \)
From now on let us assume that \( \text{char} \, k = 0 \).

**Lemma 3** Let \( f \in k[x] \), \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 \) with \( [f] \geq 2 \), let \( L(x) = ax + b \) and consider the polynomial \( g = f \circ L \). Then \( l(g) > 1 \) if and only if

\[
b = \frac{a_{n-1}}{na_n}
\]

**(2)**

**Proof**: We compute the highest coefficients of \( g \):

\[
g(x) = f \circ L(x) = a_n(ax + b)^n + a_{n-1}(ax + b)^{n-1} + \ldots =
\]

\[
a_n a^n x^n + (a_n a^{n-1} b + a_{n-1} a^{n-1}) x^{n-1} + \ldots.
\]

Thus \( g \) has gap-form if and only if

\[
a_n a^{n-1} b + a_{n-1} a^{n-1} = 0,
\]

which is equivalent to (2), as \( a \neq 0 \).

\( \square \)

**Corollary 1** Let \( f \in k[x] \), \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 \) with \( [f] \geq 2 \). Then there is a conjugate \( g \) of \( f \) which has gap-form.

**Proof**: As multiplication from the left with linear polynomials does not change \( l(f) \), we can take

\[
g = (x + \frac{a_{n-1}}{na_n}) \circ f \circ (x - \frac{a_{n-1}}{na_n}).
\]

\( \square \)

From now on assume that \( k \) is not only of characteristic zero, but also algebraically closed.

The following proposition gives a method to compute \( L(f) \) from the coefficients of \( f \).

**Proposition 1** Let \( f \in k[x] \), \( f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_0 \), \( n = [f] \geq 2 \) and suppose that \( f \) has gap-form.

(i) In case that \( 1 < l(f) < n \) denote by \( n = k_1 > k_2 > \ldots > k_l > 0 \) those indices \( k_i \in \{1, \ldots, n\} \) for which \( a_{k_i} \neq 0 \). Then the formula

\[
L(f) = \gcd(k_1 - k_2, k_2 - k_3, \ldots, k_{l-1} - k_l)
\]

holds. Furthermore the gap-degree \( L(f) \) is the maximum number \( r \), such that \( f(x) - a_0 \) admits a representation of the form

\[
f(x) - a_0 = x^k p(x^r), \text{ for } k \geq 1 \text{ and } p \in k[x], \text{ } p(0) \neq 0,
\]

and we have \( G(f) = \{ \zeta x | \zeta L(f) = 1 \} \cong \mathbb{Z}_{L(f)} \).

(ii) In case that \( l(f) = n \), i.e. that \( f(x) = a_n x^n + a_0 \), we have \( L(f) = |k| \) and \( G(f) = \{ ax | a \in k, a \neq 0 \} \).
Case (ii) occurs if and only if $f$ is weakly conjugated to the power $x^n$.

**Proof:** Consider the case $1 < l(f) < n$. Denote by $c$ the maximum of all numbers $r$ such that a representation (3) is possible, and let $d = \gcd(k_1 - k_2, k_2 - k_3, \ldots, k_{l-1} - k_l)$. In first place we show $c = d$.

Let $r$ be any number such that (3) holds for some $k \in \mathbb{N}$ and $p \in \mathbb{k}[x]$. Then

$$\{k_1, \ldots, k_l\} \subseteq \{k, k + r, k + 2r, \ldots, k + [p]r\}$$

holds, which implies $r|(k_i - k_{i+1})$ for each $i = 1, \ldots, l - 1$ and therefore $r|d$.

To show the converse relation we establish a representation (3) for $r = d$:

$$f(x) - a_0 = x^{k_1}(a_{k_1}x^{k_1-k_1} + a_{k_2}x^{k_2-k_1} + \ldots + a_{k_l})$$

As $d|(k_i - k_{i+1})$ for each $i$, $d$ also divides $k_i - k_l$. Thus $f(x) - a_0 = x^kp(x^d)$ with $k = k_l$ and

$$p(x) = a_{k_1}x^{d_1} + a_{k_2}x^{d_2} + \ldots + a_{k_l}.$$ 

We proceed to determine $G(f)$. Let $\zeta$ be a $d$-th root of unity, then

$$f \circ \zeta x = (f - a_0) \circ \zeta x + a_0 = x^kp(x^d) \circ \zeta x + a_0 = \zeta^k x \circ x^kp(x^d) + a_0 =$$

$$= \zeta^k x \circ (f - a_0) + a_0 = \zeta^k x \circ f + (1 - \zeta^k)a_0 = L_1 \circ f,$$

where $L_1(x) = \zeta^k x + (1 - \zeta^k)a_0$. Thus $\zeta x \in G(f)$ by Lemma 2. Let conversely $L(x) = ax + b$ be an element of $G(f)$. Then there exists, again by Lemma 2, a linear polynomial $L'(x) = a'x + b'$, such that $L' \circ f = f \circ L$. Thus $l(f \circ L) = l(L' \circ f) = l(f) > 1$, which implies $b = 0$ by Lemma 3. Let $p(x) = b_mx^m + b_{m-1}x^{m-1} + \ldots + b_0$, then

$$f \circ L = (x^d p(x^d) + a_0) \circ (ax) = a^k x^k p(a^d x^d) + a_0 =$$

$$a^k (b_m a^{dm} x^{dm+k} + b_{m-1} a^{d(m-1)} x^{d(m-1)+k} + \ldots + b_0 x^k) + a_0$$

and

$$L' \circ f = a' (b_m x^{dm+k} + b_{m-1} x^{d(m-1)+k} + \ldots + b_0 x^k) + b'.$$

Comparing coefficients we find

$$a' = a^{dm+k} \text{ and } b' = a_0.$$

Let $b_i (i \neq m)$ be any nonzero coefficient of $p$ and compare the coefficients of $x^{di+k}$:

$$b_i a^{di+k} = a^{dm+k} b_i,$$

that is $a^{d(m-i)} = 1$. The numbers $k_1, \ldots, k_l$ are exactly those numbers $di+k$ with $b_i \neq 0$ and $k_1 = dm + k$. Thus $d = \gcd(k_1 - k_2, k_2 - k_3, \ldots, k_1 - k_l) = \gcd(\{d(m-i)\} | b_i \neq 0) \text{ and therefore } a^d = 1$. We found $G(f) = \{\zeta x | \zeta^d = 1\}$ and thus $L(f) = |G(f)| = d$.

Consider now the case $l(f) = n$. Similar to the first case we find for each $L(x) = ax + b \in G(f)$ that $l(f \circ L) = l(L' \circ f) = l(f) > 1$, hence $b = 0$. On the other hand

$$f \circ ax = (a_n x^n + a_0) \circ ax = (a_n x + (1 - a^n)a_0) \circ f$$

holds, and therefore $ax \in G(f)$ for arbitrary $a \in k, a \neq 0$. Thus $G(f) = \{ax | a \in k, a \neq 0\}$.

If $f(x) = a_n x^n + a_0$, we find $f = (a_n x + a_0) \circ x^n$. Thus $f$ is weakly conjugated to $x^n$. Conversely, let $f$ be weakly conjugated to $x^n$, i.e. $f = (a'x + b') \circ x^n \circ (ax + b)$. As $f$ has gap-form Lemma 3 implies $b = 0$ and we find $f(x) = a' a^n x^n + b'$.

$\square$
Proposition 2  The gap-degree is constant on weak conjugacy classes. Let \( f \in \mathbb{k}[x] \), \( f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0 \), \( n = \lceil f \rceil \geq 2 \) and let \( f' = \Phi_L(f) \) with \( L(x) = x - \frac{a_{n-1}}{na_n} \). Then \( \mathcal{L}(f) = \mathcal{L}(f') \) and the latter can be computed via Proposition 1.

(i) In case that \( 1 < l(f') < n \) the group \( \mathcal{G}(f) \) is given by

\[
\mathcal{G}(f) = \{ \zeta x + (\zeta - 1)\frac{a_{n-1}}{na_n} | \zeta^\mathcal{L}(f) = 1 \} \cong \mathbb{Z}\mathcal{L}(f).
\]

(ii) In case that \( l(f') = n \) we have

\[
\mathcal{G}(f) = \{ ax + (a - 1)\frac{a_{n-1}}{na_n} | a \in \mathbb{k}, a \neq 0 \}.
\]

Again case (ii) occurs if and only if \( f \) is weakly conjugated to the power \( x^n \).

Proof : Since \( \mathbb{k}[f] = \mathbb{k}[L \circ f] \) for any linear polynomial \( L \), we have \( \mathcal{G}(f) = \mathcal{G}(L \circ f) \). Suppose \( g = L_1 \circ f \circ L_2 \), then \( \mathcal{G}(g) = \mathcal{G}(L_2^{-1} \circ L_1^{-1} \circ g) = \mathcal{G}(L_2^{-1} \circ f \circ L_2) \). In the following we establish an isomorphism between \( \mathcal{G}(f) \) and \( \mathcal{G}(L_2^{-1} \circ f \circ L_2) \). This will prove the first assertion.

Let \( L \) be a linear polynomial and consider the conjugation \( \Phi_L \). If \( H \in \mathcal{G}(f) \) we have

\[
H_1 \circ f = f \circ H \text{ for some } H_1,
\]

and therefore also

\[
\Phi_L(H_1) \circ \Phi_L(f) = \Phi_L(f) \circ \Phi_L(H).
\]

Thus \( \Phi_L(H) \in \mathcal{G}(\Phi_L(f)) \), i.e.

\[
\Phi_L(\mathcal{G}(f)) \subseteq \mathcal{G}(\Phi_L(f)).
\]

Since a similar argument shows

\[
\Phi_L^{-1}(\mathcal{G}(\Phi_L(f))) \subseteq \mathcal{G}(f),
\]

we have \( \mathcal{G}(\Phi_L(f)) \subseteq \Phi_L(\mathcal{G}(f)) \), thus the conjugation \( \Phi_L \) yields an isomorphism between \( \mathcal{G}(f) \) and \( \mathcal{G}(\Phi_L(f)) \).

In case that \( 1 < l(f') < n \) Proposition 1 yields \( \mathcal{G}(f') = \{ \zeta x | \zeta^\mathcal{L}(f') = 1 \} \) and therefore

\[
\mathcal{G}(f) = \Phi_L^{-1}(\mathcal{G}(f')) = \{ \zeta x + (\zeta - 1)\frac{a_{n-1}}{na_n} | \zeta^\mathcal{L}(f) = 1 \}.
\]

In the case \( l(f') = n \) the result follows by a similar argument. \( \square \)

For future reference we state a result of H.Engström (see [16]), and give some formulas involving gap and gap-degree.

Proposition 3 (H.Engström 1941) Let \( f, g \in \mathbb{k}[x] \). If \( \mathbb{k}[f] \cap \mathbb{k}[g] \neq \mathbb{k} \) then there exists a polynomial \( h \) of degree \( \lceil h \rceil = \gcd([f],[g]) \), such that

\[
f = f_1 \circ h \text{ and } g = g_1 \circ h
\]

holds for appropriate \( f_1 \) and \( g_1 \).
Let us remark that Proposition 3 also follows from a result in [28] (ch.6, Theorem 5.84) by using the well known fact that \( \mathbb{k}[f] = \mathbb{k}(f) \cap \mathbb{k}[x] \) for any \( f \in \mathbb{k}[x] \).

Lemma 4 Let \( f, g \in \mathbb{k}[x] \), \( n = [f], m = [g] \geq 2 \), \( f(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0 \), \( g(x) = b_mx^m + b_{m-1}x^{m-1} + \ldots + b_0 \) and suppose \( f(x) \neq a_nx^n \) and \( g(x) \neq b_mx^m \). Then the gap of the polynomial \( f \circ g \) is given by the following formulas:

(i) If \( l(g) \neq [g]l(f) \) then \( l(f \circ g) = l(g) \),

(ii) if \( l(g) = [g]l(f) \) then \( l(f \circ g) \geq l(g) \).

In the last case equality holds if and only if

\[
a_nmb_m^{l(f)-1} b_{m-l(g)} + a_{n-l(f)} \neq 0. \tag{4}
\]

The gap of the polynomial \( f \cdot g \) is given by the formulas:

(iii) If \( l(f) \neq l(g) \) then \( l(f \cdot g) = \min(l(f), l(g)) \),

(iv) If \( l(f) = l(g) \) then \( l(f \cdot g) \geq l(f) = l(g) \).

In the last case equality holds if and only if

\[
a_nb_{m-l(g)} + a_{n-l(f)}b_m \neq 0. \tag{5}
\]

Furthermore we have:

(v) \( l(a_nx^n \circ g) = l(g) \),

(vi) \( l(f \circ b_mx^m) = ml(f) \),

(vii) \( l(a_nx^n \cdot g) = l(g) \),

(viii) \( l(g^{(r)}) = l(g) \) for \( r \in \mathbb{N} \).

Proof: Let \( k = l(f) \) and \( l = l(g) \), then \( f(x) = a_nx^n + a_{n-k}x^{n-k} + \ldots + a_0 \) and \( g(x) = b_mx^m + b_{m-l}x^{m-l} + \ldots + b_0 \) where \( a_{n-k}, b_{m-l} \neq 0 \).

We compute the highest coefficients of \( f \circ g \):

\[
f \circ g(x) = a_n(b_mx^m + b_{m-l}x^{m-l} + \ldots + b_0)^n + \\
+ a_{n-k}(b_mx^m + b_{m-l}x^{m-l} + \ldots + b_0)^{n-k} + \ldots = \\
a_nb_m^{n-m} + a_nb_m^{n-1}b_{m-l}x^{n-m-l} + a_{n-k}b_m^{n-k}x^{n-m-k} + \ldots. \tag{6}
\]

From (6) the assertions (i), (ii) and (4) follow.

We compute the highest coefficients of \( f \cdot g \):

\[
f \cdot g(x) = (a_nx^n + a_{n-k}x^{n-k} + \ldots + a_0) \cdot (b_mx^m + b_{m-l}x^{m-l} + \ldots + b_0) = \\
= a_nb_mx^{n+m} + a_{n-k}b_nb_m^{n-k} + a_nb_{m-l}x^{n+m-l} + \ldots. \tag{7}
\]

From (7) we obtain (iii), (iv) and (5).
To prove (v) we compute
\[ a_n x^n \circ g = a_n (b_m x^m + b_{m-1} x^{m-1} + \ldots + b_0)^n = a_n (b_m^n x^{nm} + n b_m^{n-1} b_{m-1} x^{nm-1} + \ldots ). \]

The relations (vi) and (vii) are obvious. To establish (viii) we use induction on \( r \). For \( r = 1 \) (viii) is clearly true. Let \( r \geq 2 \) and suppose that \( l(g^{(r-1)}) = l(g) \). As \( [g] \geq 2 \) holds. Thus (i) applies to \( g^{(r)} = g^{(r-1)} \circ g \) and we obtain \( l(g^{(r)}) = l(g) \).

\[ \square \]

**Lemma 5** Let \( f, g \in \mathbb{k}[x] \) be as in Lemma 4 and assume furthermore that \( f \circ g = g \circ f \), then \( l(f) = l(g) \).

**Proof:** By (i) and (ii) of Lemma 4 we obtain
\[ l(f \circ g) = l(g \circ f) \geq l(f). \]

In case \( l(g) \neq [g] l(f) \) we have \( l(f \circ g) = l(g) \) and therefore \( l(g) \geq l(f) \). If \( l(g) = [g] l(f) \) we also have \( l(g) \geq l(f) \). By symmetry \( l(f) \geq l(g) \) and thus \( l(f) = l(g) \).

\[ \square \]

Before we can state the next result we have to introduce another notation. Denote by \( G^*(f) \) the group of all linear polynomials \( L_1 \) such that a relation of the form \( L_1 \circ f = f \circ L_2 \) holds. Thus \( f \circ L_1 = (L_2 \circ L_3^{(-1)}) \circ f \circ L_2 \), i.e. \( L \in G(f^{(m)}) \).

**Lemma 6** Let \( f \in \mathbb{k}[x] \), \( [f] \geq 2 \). Then
\[ G(f) \supseteq G(f^{(2)}) \supseteq G(f^{(3)}) \supseteq \ldots. \] (8)

We have \( G(f) = G(f^{(2)}) \) if and only if \( G^*(f) \subseteq G(f) \). In this case equality holds throughout the chain (8).

The condition \( G^*(f) \subseteq G(f) \) is satisfied for instance if \( 1 < l(f) < [f] \) and \( f(0) = 0 \).

**Proof:** Let \( L \in G(f^{(m)}) \) and let \( m < n \). Then
\[ f^{(n-m)} \circ (f^{(m)} \circ L) = (L_1 \circ f^{(n-m)}) \circ f^{(m)}, \]
and Proposition 3 implies that there exists a polynomial \( h \) of degree \([h] = [f^{(m)}]\), such that \( f^{(m)} \circ L = L_2 \circ h \) and \( f^{(m)} = L_3 \circ h \) holds. Thus \( f^{(m)} \circ L = (L_2 \circ L_3^{(-1)}) \circ f^{(m)} \), i.e. \( L \in G(f^{(m)}) \).
Suppose $G(f) = G(f^{(2)})$ and let $L_1 \in G^*(f)$. Then $L_1 \circ f = f \circ L$ for some $L \in G(f) = G(f^{(2)})$. Let $L_2 \circ f^{(2)} = f^{(2)} \circ L$, then

$$L_2 \circ f^{(2)} = f \circ f \circ L = f \circ L_1 \circ f.$$  

Thus $L_2 \circ f = f \circ L_1$, which shows $L_1 \in G(f)$. Conversely suppose that $G^*(f) \subseteq G(f)$ and let $L \in G(f)$. Then

$$f^{(2)} \circ L = f \circ L_1 \circ f = L_2 \circ f^{(2)}$$

holds since $L_1 \in G^*(f) \subseteq G(f)$, and therefore $L \in G(f^{(2)})$. Repeating the above argument $(m-1)$ times we obtain

$$f^{(m)} \circ L = L_m \circ f^{(m)}$$

which shows $L \in G(f^{(m)})$.

Let $f$ satisfy $1 < l(f) < [f]$ and $f(0) = 0$. Then we have a representation

$$f(x) = x^k p(x^r)$$

with $k \geq 1$ and $r = L(f)$, and

$$G(f) = \{ \zeta x | \zeta^r = 1 \}.$$  

As $f \circ \zeta x = \zeta^k x \circ f$ we find

$$G^*(f) = \{ \zeta^k x | \zeta^r = 1 \} \subseteq G(f).$$  

Another sufficient condition in order that $G^*(f) \subseteq G(f)$ holds is given in the next lemma.

**Lemma 7** Let $f \in k[x]$, $[f] \geq 2$. Suppose that $1 < l(f)$ and that for some $n \in \mathbb{N}$ and a linear polynomial $\epsilon x$ the relation

$$f^{(n)} \circ \epsilon x = \delta x \circ f^{(n)}$$  

holds, where $\delta \neq 1$. If furthermore $l(f) < [f]$, then $G^*(f) \subseteq G(f)$ and in particular $L(f^{(k)}) = L(f)$ for $k \in \mathbb{N}$.

**Proof:** The relation (9) shows that $\epsilon x \in G(f^{(n)})$. Thus, by Lemma 6 we also have $\epsilon x \in G(f^{(n-1)})$, i.e.

$$f^{(n-1)} \circ \epsilon x = L \circ f^{(n-1)}$$

holds for some linear polynomial $L$. We compute

$$f \circ L \circ f^{(n-1)} = f \circ f^{(n-1)} \circ \epsilon x = f^{(n)} \circ \epsilon x = \delta x \circ f^{(n)},$$

and therefore

$$f \circ L = \delta x \circ f.$$  

(10)

As $l(f) > 1$ the right hand side of (10) has gap-form and, according to Lemma 3, the polynomial $L$ must be a multiplication $L(x) = \gamma x$. Due to this fact we may compute

$$f(0) = (f \circ \gamma x)(0) = (\delta x \circ f)(0) = \delta f(0).$$

As $\delta \neq 1$ we find $f(0) = 0$ which implies together with Lemma 6 the assertion.

In the following definition we consider another group of linear polynomials. This concept has been studied in [13].
Definition 3 Let $f \in k[x]$, $[f] \geq 2$. Denote by $\mathcal{G}^p(f)$ the group of all linear polynomials $L$ which permute with $f$:

$$L \circ f = f \circ L.$$ 

Furthermore let $\mathcal{L}^p(f) = |\mathcal{G}^p(f)|$.

Some results on $\mathcal{G}^p(f)$ will be recalled in §5.

For future reference let us state the following rather obvious facts.

Lemma 8 Let $f \in k[x]$, $[f] \geq 2$. Then $\mathcal{G}^p(f) \subseteq \mathcal{G}^p(f^{(n)})$ holds for each $n \in \mathbb{N}$. In case $\mathcal{L}(f)$ is finite we also have $\mathcal{L}^p(f^{(n)})|\mathcal{L}(f)$.

Proof: The first assertion is obvious. To prove the remaining part of the lemma, note that

$$\mathcal{G}^p(f^{(n)}) \subseteq \mathcal{G}(f^{(n)}) \subseteq \mathcal{G}(f).$$

$\square$

4 Commutative semigroups which do not contain a linear polynomial

In this section we give a complete list of all commutative subsemigroups of $\langle C[x], \circ \rangle$ which contain no linear polynomials (except possibly the trivial one, namely $x$).

To start with let us recall two well known results. Here $k$ again is a field of characteristic zero.

Proposition 4 Let $k$ be algebraically closed and let $\mathcal{S}$ be a commutative subsemigroup of $\langle k[x], \circ \rangle$. If $\mathcal{S}$ contains an element which is conjugated to a power $ax^n$ ($a \in k$, $a \neq 0$, $n \geq 2$), then

$$\Phi_L(\mathcal{S}) \subseteq P_r = \{ \zeta x^k | k \equiv 1 \ mod \ r, \zeta^r = 1 \}$$

for some $r$ and an appropriate conjugation $\Phi_L$.

Proposition 5 Let $\mathcal{S}$ be a commutative subsemigroup of $\langle k[x], \circ \rangle$. If $\mathcal{S}$ contains an element which is conjugated to a Chebyshev polynomial $t_n$ or $-t_n$, respectively ($n \geq 2$), then either (i) or (ii) holds for some conjugation $\Phi_L$:

(i) $\Phi_L(\mathcal{S}) \subseteq D_1 = \{ t_n | n \in \mathbb{N} \}$,

(ii) $\Phi_L(\mathcal{S}) \subseteq D_2 = \{ \pm t_k | k \ odd \}$.

A proof of these results can be found e.g. in [14]. Another interesting proof of Proposition 5 is given in [3]. All these proofs employ only purely algebraic methods.

We proceed treating the remaining case of a commutative subsemigroup $\mathcal{S}$ of $\langle C[x], \circ \rangle$ with no element conjugated to a power $ax^n$ or a Chebyshev polynomial $\pm t_n$. To do so we use two results of J.Ritt (see [36] and [38]) which have been proved using topological and analytic methods.
Proposition 6 (J.Ritt 1923) Let \( f, g \in \mathbb{C}[x] \), \([f], [g] \geq 2\) be permutable and suppose that neither \( f \) nor \( g \) is conjugated to a power or a Chebyshev polynomial. Then there exists a conjugation \( \Phi_L \) and a polynomial \( h(x) = x^p(x^r) \), such that
\[
\Phi_L(f) = \epsilon_1 h^{(m)} \quad \text{and} \quad \Phi_L(g) = \epsilon_2 h^{(n)}.
\]
Here \( \epsilon_1^* = \epsilon_2^* = 1 \) if \( r \) is chosen maximal. If furthermore \( r > 1 \), we have \( r = \mathcal{L}^p(h) \).

Proposition 7 (J.Ritt 1920) Let \( f, g \in \mathbb{C}[x] \), \([f], [g] \geq 2\) and let \( \epsilon \) and \( \delta \) be roots of unity with
\[
f \circ \epsilon x = \epsilon' x \circ f \quad \text{and} \quad g \circ \delta x = \delta' x \circ g\]
for some \( l, l' \in \mathbb{N} \). If \( \epsilon f^{(m)} = \delta g^{(n)} \) for some numbers \( m, n \in \mathbb{N} \), then there exists a conjugation \( \Phi_L \) and a polynomial \( h(x) = x^k p(x^r) \) (\( k \geq 0, p(0) \neq 0 \)), such that
\[
\Phi_L(f) = \epsilon_1 h^{(s)} \quad \text{and} \quad \Phi_L(g) = \epsilon_2 h^{(t)}.
\]
Here \( \epsilon_1^* = \epsilon_2^* = 1 \) if \( r \) is chosen maximal. If furthermore \( k \geq 1 \) and \( r > 1 \), we have \( r = \mathcal{L}(h) \).

Proposition 7 actually is proved in [36] only in the case \( \epsilon = \delta = 1 \), but it is easy to see that the proof given there also works in the more general situation.

Before we proceed to generalize Proposition 6 we have to state a corollary of Proposition 6 and a lemma.

Corollary 2 The formulation "there exists a conjugation \( \Phi_L \) and a polynomial \( h(x) \)" in Proposition 6 and Proposition 7 can be strengthened to "for each conjugation \( \Phi_L \) which conjugates \( f \) (and thus \( g \)) to gap-form there exists a polynomial \( h(x) \)."

Proof: If we have \( \epsilon_1 = \epsilon_2 = 1 \) we can apply any conjugation \( \Phi_M \) to the representations given in Proposition 6 and Proposition 7, respectively and obtain a representation of the same type. If \( \epsilon_1 \neq 1 \) or \( \epsilon_2 \neq 1 \) the polynomial \( h \) has gap-form. Any conjugation \( \Phi_M \) which conjugates \( f \) (and thus \( g \)) to gap-form differs only by a conjugation \( \Phi_{\gamma x} \) from \( \Phi_L \). Applying \( \Phi_{\gamma x} \) to the representations of Proposition 6 and Proposition 7, respectively does not change the type of representation given.

Lemma 9 Let \( f_1, f_2, \ldots \) be a (possibly finite) sequence of permutable polynomials which admit the representation
\[
f_i(x) = \epsilon_i[x^k p(x^r)]^{(k_i)}
\]
where \( \epsilon_i^* = 1 \), i.e. \( f_i \circ \epsilon_i x = \epsilon_i^* x \circ f_i \) for convenient numbers \( l_i \). Then (i) or (ii) holds:

(i) \( f_i = h^{(k_i)} \) for some polynomial \( h \).

(ii) For some \( i \) we have \( \epsilon_i^k \neq 1 \). In this case we have in particular \( k \geq 1 \) and \( r > 1 \).
Proof: Assume that (ii) does not hold, i.e. \( \epsilon_i^k = 1 \) for all \( i \). Then

\[
x^k p(x^r) \circ \epsilon_i x = x^k p(x^r)
\]

and thus also

\[
x^k p(x^r) \circ \epsilon_i^{-1} x = x^k p(x^r).
\]

This implies

\[
f_i(x) = [\epsilon_i x \circ x^k p(x^r) \circ \epsilon_i^{-1} x]^{(k_i)} \text{ for } i = 1, 2, \ldots.
\]

As the polynomials \( f_i \) permute we have that all numbers \( \epsilon_i \) are equal and therefore obtain a representation of the desired form.

\[\square\]

The following two lemmata generalize Proposition 6.

Lemma 10 Let \( n \in \mathbb{N} \), let \( f_1, \ldots, f_n \in \mathbb{C}[x] \), \([f_i] \geq 2\) be pairwise permutable polynomials with gap-form and suppose that no polynomial \( f_i \) is conjugated to a power or a Chebyshev polynomial. Then (i) or (ii) holds:

(i) There exist a polynomial \( g \) and numbers \( m_1, \ldots, m_n \in \mathbb{N} \), such that

\[
f_i = g^{(m_i)} \text{ for } i = 1, \ldots, n.
\]

(ii) There exist a polynomial \( g(x) = x^k g_1(x^r) \) where \( k \geq 1 \), \( g_1(0) \neq 0 \) and \( r = \mathcal{L}(g) > 1 \), numbers \( m_1, \ldots, m_n \in \mathbb{N} \) and \( r \)-th roots of unity \( \epsilon_1, \ldots, \epsilon_n \), such that

\[
f_i = \epsilon_i g^{(m_i)} \text{ for } i = 1, \ldots, n.
\]

Here

\[
\epsilon_i^{k_i m_j - 1} = \epsilon_j^{k_j m_i - 1} \text{ for } i, j = 1, \ldots, n.
\]

and for at least one \( i \in \{1, \ldots, n\} \) we have \( \epsilon_i^k \neq 1 \).

Proof: First of all we clarify condition (13). To do so assume \( f_i = \epsilon_i g^{(m_i)} \) with \( g \) as above and compute \( f_i \circ f_j \) and \( f_j \circ f_i \):

\[
f_i \circ f_j = \epsilon_i x \circ g^{(m_i)} \circ \epsilon_j x \circ g^{(m_j)} = \epsilon_i \epsilon_j^{k_i m_j} x \circ g^{(m_i + m_j)},
\]

\[
f_j \circ f_i = \epsilon_j x \circ g^{(m_j)} \circ \epsilon_i x \circ g^{(m_i)} = \epsilon_j \epsilon_i^{k_j m_i} x \circ g^{(m_i + m_j)}.
\]

Thus \( f_i \) and \( f_j \) permute if and only if (13) holds.

To prove the lemma use induction on \( n \). Consider first the case that \( n = 2 \). Then Proposition 6 implies together with Lemma 9 the assertion of the lemma.

We proceed to treat the case \( n > 2 \). The inductive hypothesis implies that (i) or (ii) holds for the polynomials \( f_1, \ldots, f_{n-1} \).

Case 1: Assume (i) holds for \( f_1, \ldots, f_{n-1} \), i.e. we have a representation

\[
f_i = g^{(m_i)} \text{ for } i = 1, \ldots, n - 1.
\]
Let \( \gcd(m_1, \ldots, m_{n-1}) = d \). As \( f_n \) permutes with each polynomial \( f_i \), Lemma 1 implies that \( f_n \) permutes with \( g^{(d)} \). Note that by Lemma 4 \( l(g^{(d)}) = l(g) = l(f_1) > 1 \). Thus Proposition 6 implies together with Corollary 2 that

\[
\begin{align*}
g^{(d)}(x) &= \epsilon_1[x^p(x^s)]^{(k_1)} \quad \text{and} \\
f_n(x) &= \epsilon_2[x^p(x^s)]^{(k_2)} \quad \text{with } \epsilon_1^u = \epsilon_2^u = 1.
\end{align*}
\]

Proposition 7 shows that

\[
\begin{align*}
g(x) &= \delta_1[x^t q(x^s)]^{(l_1)} \quad \text{and} \\
x^p(x^s) &= \delta_2[x^t q(x^s)]^{(l_2)} \quad \text{with } \delta_1^s = \delta_2^s = 1.
\end{align*}
\]

From this we find

\[
\begin{align*}
g^{(d)}(x) &= \epsilon_1\delta_2^{s_i}[x^t q(x^s)]^{(l_2k_1)} \quad \text{and} \\
f_n(x) &= \epsilon_2\delta_2^{s_i}[x^t q(x^s)]^{(l_2k_2)}
\end{align*}
\]

for some \( n_1, n_2 \in \mathbb{N} \). In order to apply Lemma 9 we have to show that \( \epsilon_1^s = 1 \), as then

\[
f_i(x) = \gamma_i[x^t q(x^s)]^{(l_2k_1m_i)} \quad \text{for } i = 1, \ldots, n - 1
\]

with some \( s \)-th roots of unity \( \gamma_i \).

If \( \epsilon_1 = \epsilon_2 = 1 \) we are done. So suppose that \( \epsilon_1 \neq 1 \) or \( \epsilon_2 \neq 1 \). We have

\[
[x^t q(x^s)]^{(l_2)} \circ \epsilon_i x = \epsilon_i x \circ [x^t q(x^s)]^{(l_2)} \quad \text{for } i = 1, 2.
\]

As all occurring polynomials have gap-form Lemma 7 is applicable and we find \( l \geq 1 \) and \( s = \mathcal{L}(x^t q(x^s)) \). Furthermore

\[
u|\mathcal{L}([x^t q(x^s)]^{(l_2)}) = \mathcal{L}(x^t q(x^s)) = s,
\]

thus \( \epsilon_1^s = 1 \), and Lemma 9 yields a representation of type (i) or (ii).

\textbf{Case 2:} Assume (ii) holds for \( f_1, \ldots, f_{n-1} \), i.e. we have a representation

\[
f_i = \epsilon_i [g^{(m_i)}] \quad \text{for } i = 1, \ldots, n - 1.
\]

Here \( g(x) = x^k g_1(x^r) \) with \( r = \mathcal{L}(g) > 1 \) and (w.l.o.g.) \( \epsilon_1^i \neq 1 \). As \( f_1 \) permutes with \( f_n \) Proposition 6 together with Corollary 2 implies that

\[
\begin{align*}
f_n(x) &= \delta_1[xq(x^s)]^{(s_1)} \quad \text{and} \\
f_1(x) &= \delta_2[xq(x^s)]^{(s_2)} = \epsilon_1[x^k g_1(x^r)]^{(m_1)}
\end{align*}
\]

with \( \delta_1^s = \delta_2^s = 1 \). As \( [xq(x^s)] > l(xq(x^s)) = l(f_1) > 1 \) we have by Lemma 6

\[
s|\mathcal{L}(xq(x^s)) = \mathcal{L}(f_1),
\]

and similarly

\[
\mathcal{L}(f_1) = \mathcal{L}(x^k g_1(x^r)) = r.
\]

Applying Proposition 7 we find

\[
x^k g_1(x^r) = \gamma_1[x^t q(x^s)]^{(l_1)} \quad \text{and}
\]

(15)
for some numbers \( l \) and \( r \) for some numbers \( i, \) or of type (ii).

The splitting of the assertion of Lemma 10 into types (i) and (ii) is justified by the following observation: If type (ii) occurs the polynomials \( f \) with gap-form may occur.

Remark 1 The splitting of the assertion of Lemma 10 into types (i) and (ii) is justified by the following observation: If type (ii) occurs the polynomials \( f \) have gap-form and satisfy \( f_i(0) = 0. \)

On the other hand, in case of type (i), arbitrary polynomials with gap-form may occur.

From (14) and (15) we obtain the following fact.

Remark 2 In each inductive step the degree of the representing polynomial \( g \) does not increase.

The next lemma shows that the assertion of Lemma 10 remains true in the case that \( n \to \infty. \)

Lemma 11 Let \( f_1, f_2, \ldots \in \mathbb{C}[x], \) \( |f_i| \geq 2 \) be a sequence of pairwise permutable polynomials with gap-form, and suppose that no polynomial \( f_i \) is conjugated to a power or a Chebyshev polynomial. Then a representation of type (i) or type (ii) of Lemma 10 holds.

Proof : We apply Lemma 10 to \( \{f_1, f_2\} \) and obtain a polynomial \( g_2 \) and a representation of the form (11) or (12) employing \( g_2. \) Proceeding from this representation we apply the inductive step of the proof of Lemma 10 to \( \{f_1, f_2, f_3\} \) and obtain a polynomial \( g_3 \) and a representation of the form (11) or (12) employing \( g_3. \) Carrying on in this way we get a sequence of polynomials \( g_n, \) such that for each \( n \) a representation of the form (11) or (12) employing \( g_n \) holds for \( f_i, i \) up to \( n. \) By Remark 2 we have \( |g_2| \geq |g_3| \geq \ldots. \)

Let \( g_2 \) be a polynomial of minimal possible degree, such that \( f_1 \) and \( f_2 \) admit a representation of the form (11) or (12) involving \( g_2. \) Starting the above procedure with this polynomial we get a sequence \( g_2, g_3, \ldots \) with \( |g_n| \geq |g_2|, \) as each \( g_n \) yields in particular a representation of \( f_1 \) and \( f_2, \) by Lemma 9. Thus

\[
[g_2] = [g_3] = [g_4] = \ldots.
\]
We have representations
\[ f_i = \epsilon_{i,k} g_k^{(m_i)} \] for \( i = 1, \ldots, k \) and \( k \geq 2 \)

(independently of whether type (i) or (ii) of Lemma 10 occurs). In particular we have for each \( k \geq 3 \)
\[ f_1 = \epsilon_{1,2} g_2^{(m_1)} = \epsilon_{1,k} g_k^{(m_1)} \]
and thus due to Proposition 7
\[ g_2(x) = \delta_1 [x^1 p(x^r)]^{(s)} \text{ and } g_k(x) = \delta_2 [x^1 p(x^r)]^{(s)} \]
with \( \delta_1 = \delta_2 = 1 \). Here obviously \( \delta_i x \in G(f_i) \) for \( i = 1, 2 \). Thus the numbers \( \delta_i \) are \( \mathcal{L}(f_i) \)-th roots of unity, i.e. \( g_2 \) and \( g_k \) differ only by a \( \mathcal{L}(f_1) \)-th root of unity. Therefore one specific polynomial \( g = \zeta g_2 \) (where \( \zeta^{\mathcal{L}(f_1)} = 1 \)) must occur infinitely often among the polynomials \( g_k \) and we find
\[ f_i = \epsilon_i g^{(m_i)} \text{ for each } i \in \mathbb{N}. \]

Applying Lemma 9 yields the assertion.

\[ \square \]

Lemma 10 and Lemma 11 put us in position to give a complete list of commutative subsemigroups of \( \langle \mathbb{C}[x], \circ \rangle \) containing no linear polynomials.

**Theorem 1** Let \( S \) be a commutative subsemigroup of \( \langle \mathbb{C}[x], \circ \rangle \) which does not contain any linear polynomial (except possibly \( x \) itself). Then \( S \) is conjugated to a subsemigroup of one of the following commutative semigroups:

**List I**

(i) \( \mathcal{P}_r = \{ \zeta x^k | k \equiv 1 \pmod{r}, \zeta^r = 1 \} \),

(ii) \( \mathcal{D}_1 = \{ t_n | n \in \mathbb{N} \} \),

(iii) \( \mathcal{D}_2 = \{ \pm t_k | k \text{ odd} \} \),

(iv) \( \mathcal{R}_1 = \{ \epsilon_i g^{(m_i)} | i = 1, \ldots, n \}, \) where \( n \in \mathbb{N} \), or \( \mathcal{R}_1 = \{ \epsilon_i g^{(m_i)} | i = 1, 2, \ldots \}, \) with \( g(x) = x^k p(x^r), k \geq 0, \epsilon_i^r = 1 \) and
\[ \epsilon_i^{m_j-1} = \epsilon_j^{m_i-1} \text{ for all } i \text{ and } j. \]

**Proof** : Let \( S \) be a commutative semigroup which contains no linear polynomials. The cases (i)-(iii) are settled by Proposition 4 and Proposition 5. Thus assume that no element of \( S \) is conjugated to a power or a Chebyshev polynomial. Let \( f \in S, [f] = n(>1) \). A result of E. Jacobsthal (see [19]) states that for any given degree \( k(> 1) \) there are at most \( n-1 \) polynomials of degree \( k \) permuting with \( f \). Thus \( S \) is countable and has in particular an at most countable set of generators \( f_1, f_2, \ldots \). By Lemma 5 there is a conjugation \( \Phi_L \) such that all polynomials \( \Phi_L(f_i) \) have gap-form. Applying Lemma 10 or Lemma 11, respectively proves the assertion of the theorem.

\[ \square \]
5 A list of commutative subsemigroups of \(<\mathbb{C}[x], \circ>\)

Together with some well known results Theorem 1 will lead us to a complete list of arbitrary commutative subsemigroups of \(<\mathbb{C}[x], \circ>\).

First of all let us consider semigroups of linear polynomials. In [14] we find the following result, which holds for an arbitrary field \(k\).

**Proposition 8** Let \(S\) be a commutative semigroup consisting entirely of linear polynomials. Then \(S\) is conjugated to a subsemigroup of one of the two following commutative semigroups.

**List II**

(i) \(T = \{x + a | a \in k\}\),

(ii) \(S_0 = \{ax | a \in k, a \neq 0\}\).

The next result determines the set \(G^p(f)\), i.e. all linear polynomials permuting with a given nonlinear polynomial.

**Proposition 9** Let \(k\) be a field of characteristic zero, \(f \in \mathbb{k}[x]\), \([f] \geq 2\) and let \(\Phi_L\) be a conjugation, such that \(f_1 = \Phi_L(f)\) has gap-form. Then \(G^p(f) = \Phi_L^{-1}(G^p(f_1))\) and

**List III**

(i) \(G^p(f_1) = \{x\}\) if \(f_1(0) \neq 0\),

(ii) \(G^p(f_1) = \{\zeta x | \zeta^r = 1\}\) if \(f_1(0) = 0\) and

\[
\begin{align*}
  r &= \max\{u \in \mathbb{N} | f_1(x) = xg(x^u) \text{ for some } g \in \mathbb{k}[x]\}. \\
\end{align*}
\]

A proof of this result can be found in [13]. Proposition 9 yields the well known example of the commutative semigroup \(R = \langle \xi x, g \rangle\), where \(g(x) = xp(x^r)\) and \(\xi\) is a primitive \(r\)-th root of unity. This is a special case of the commutative semigroup \(R_1\) occurring in Theorem 1 (if we add the linear polynomials \(\xi x\): \(R_1 = \langle \epsilon_i g^{(m_i)} | i = 1, \ldots, n \rangle\), where \(n \in \mathbb{N}\), or \(R_1 = \langle \epsilon_i g^{(m_i)} | i = 1, 2, \ldots \rangle\) with \(g(x) = x^kp(x^r)\), \(k \geq 0\), \(\epsilon_1 = 1\) and

\[
\epsilon_i^{k^{m_j}-1} = \epsilon_j^{k^{m_i}-1} \text{ for all } i \text{ and } j. \quad (16)
\]

The following proposition shows that this seemingly more general type of commutative semigroups actually does not yield "new" semigroups.

**Proposition 10** Consider the commutative semigroup \(R = \langle \xi x, \gamma g \rangle\), where \(\gamma\) is a primitive \(L^p(g)-\)th root of unity, i.e. \(R\) is commutative. Furthermore \(L^p(g)\) computes as \(L^p(g) = \gcd(k - 1, r)\).
Before we can give a proof of Proposition 10, we have to state another lemma.

**Lemma 12** Let \( n \in \mathbb{N}, k, r \in \mathbb{N}, k, r \geq 2 \) and let \( m_1, \ldots, m_n \in \mathbb{N}, \gcd(m_1, \ldots, m_n) = 1 \). Then the set of solutions of the system of congruences

\[
x_i(k^{m_i} - 1) \equiv x_j(k^{m_j} - 1) \mod r \quad \text{for} \ i, j = 1, \ldots, n
\]

equals the set of all integral linear combinations of the vectors

\[
\left( \begin{array}{c} 1 + k + \ldots + k^{m_1-1} \\ 1 + k + \ldots + k^{m_2-1} \\ \vdots \\ 1 + k + \ldots + k^{m_n-1} \end{array} \right), \quad \left( \begin{array}{c} \frac{r}{\gcd(k-1,r)} \\ 0 \\ \vdots \\ 0 \end{array} \right), \quad \ldots, \quad \left( \begin{array}{c} 0 \\ \vdots \\ 0 \\ \frac{r}{\gcd(k-1,r)} \end{array} \right).
\]

**Proof**: Obviously every linear combination of vectors (18) gives a solution of the system (17). To show the converse let us first compute the number of modulo \( r \) incongruent linear combinations of the vectors (18). Note that

\[
\gcd(k^{m_1} - 1, \ldots, k^{m_n} - 1) = k - 1,
\]
i.e. there exist numbers \( \mu_1, \ldots, \mu_n \), such that

\[
k - 1 = \mu_1(k^{m_1} - 1) + \ldots + \mu_n(k^{m_n} - 1).
\]

Dividing by \( k - 1 \) we obtain

\[
1 = \mu_1(1 + k + \ldots + k^{m_1-1}) + \ldots + \mu_n(1 + k + \ldots + k^{m_n-1})
\]

which yields

\[
\gcd(1 + k + \ldots + k^{m_1-1}, \ldots, 1 + k + \ldots + k^{m_n-1}) = 1.
\]

Consider a linear combination

\[
\lambda_0 \left( \begin{array}{c} 1 + k + \ldots + k^{m_1-1} \\ 1 + k + \ldots + k^{m_2-1} \\ \vdots \\ 1 + k + \ldots + k^{m_n-1} \end{array} \right) + \lambda_1 \left( \begin{array}{c} \frac{r}{\gcd(k-1,r)} \\ 0 \\ \vdots \\ 0 \end{array} \right) + \ldots + \lambda_n \left( \begin{array}{c} 0 \\ \vdots \\ 0 \\ \frac{r}{\gcd(k-1,r)} \end{array} \right)
\]

and suppose that

\[
\lambda_0 \left( \begin{array}{c} 1 + k + \ldots + k^{m_1-1} \\ 1 + k + \ldots + k^{m_2-1} \\ \vdots \\ 1 + k + \ldots + k^{m_n-1} \end{array} \right) + \left( \lambda_1 \frac{r}{\gcd(k-1,r)} \right) + \ldots + \left( \lambda_n \frac{r}{\gcd(k-1,r)} \right) \equiv 0 \mod r.
\]

Then

\[
\lambda_0 \left( \begin{array}{c} 1 + k + \ldots + k^{m_1-1} \\ 1 + k + \ldots + k^{m_2-1} \\ \vdots \\ 1 + k + \ldots + k^{m_n-1} \end{array} \right) \equiv 0 \mod \frac{r}{\gcd(k-1,r)}
\]
and thus (19) implies that

$$\lambda_0 \equiv 0 \mod \frac{r}{\gcd(k - 1, r)}.$$

We found that linear combinations of type (20) with different coefficients \(\lambda_0\) in \(\{1, \ldots, \frac{r}{\gcd(k - 1, r)}\}\) yield incongruent solutions (modulo \(r\)). For fixed \(\lambda_0 \in \{1, \ldots, \frac{r}{\gcd(k - 1, r)}\}\) we obviously get \(\gcd(k - 1, r)^n\) incongruent solutions (modulo \(r\)). In fact we found \(\gcd(k - 1, r)^n = \gcd(k - 1, r)^{n-1} \cdot r\) incongruent solutions of (17), namely linear combinations of type (20) with \(\lambda_0 \in \{1, \ldots, \frac{r}{\gcd(k - 1, r)}\}\) and \(\lambda_1, \ldots, \lambda_n \in \{1, \ldots, \gcd(k - 1, r)\}\).

We use a method which can be found e.g. in [2] (Chapter 4 of Section 2) to compute the number of incongruent solutions of (17).

To do so consider the matrix \(A\) of coefficients of the system (17). To illustrate the form of this matrix consider for example the case \(n = 4\). Then we have

$$A = \begin{pmatrix}
  c_2 & -c_1 & 0 & 0 \\
  c_3 & 0 & -c_1 & 0 \\
  c_4 & 0 & 0 & -c_1 \\
  0 & c_3 & -c_2 & 0 \\
  0 & c_4 & 0 & -c_2 \\
  0 & 0 & c_4 & -c_3
\end{pmatrix}, \text{ with } c_i = k^{m_i} - 1.
$$

The rank of the matrix \(A\) equals \(n - 1\). We compute the greatest common divisors \(d_h\) of the \(h \times h\) subdeterminants of \(A\) for \(h = 1, \ldots, n - 1\). Obviously \((k - 1)^h\) divides \(d_h\). On the other hand we can find for each \(i = 1, \ldots, n\) an \(h \times h\) minor of \(A\) with determinant \(\pm (k^{m_i} - 1)^h\). As \(\gcd((k^{m_1} - 1)^h, \ldots, (k^{m_n} - 1)^h) = (k - 1)^h\) we have

$$d_h = (k - 1)^h \text{ for } h = 1, \ldots, n - 1,$$

where \(d_0 = 1\) by convention. Thus the so called elementary divisors of \(A\) are given as

$$e_h = \frac{d_h}{d_{h-1}} = k - 1 \text{ for } h = 1, \ldots, n - 1.$$

Then the number \(|\lambda, r|\) of incongruent solutions of (17) is given by the formula

$$|A, r| = \gcd(e_1, r) \cdot \ldots \cdot \gcd(e_{n-1}, r) \cdot r = \gcd(k - 1, r)^{n-1} \cdot r.$$

Thus each solution of the system (17) is a linear combination of the type (20) which proves the lemma.

\[\square\]

**Proof:** [of Proposition 10] If \(r = 1\), then no nontrivial roots of unity appear, i.e. \(\epsilon_i = 1\). Thus \(R_1 \subseteq \langle g \rangle\).

Suppose therefore \(r > 1\) and consider first of all the cases that \(k = 0\) or \(1\). If \(k = 0\) we have \(g(x) = p(x^r)\) and thus

$$g(x) \circ \epsilon x = g(x) \text{ whenever } \epsilon^r = 1.$$

As \(\epsilon_i g^{(m_i)}\) permutes with \(\epsilon_j g^{(m_j)}\) we find

$$\epsilon_i g^{(m_i + m_j)} = \epsilon_i g^{(m_i)} \circ \epsilon_j g^{(m_j)} = \epsilon_j g^{(m_j)} \circ \epsilon_i g^{(m_i)} = \epsilon_j g^{(m_j + m_i)}.$$

which implies \( \epsilon_i = \epsilon_j \). Put \( \delta = \epsilon_1 \) \((= \epsilon_2 = \epsilon_3 = \ldots)\), then
\[
(\delta g)^{(m_i)} = \delta g^{(m_i)}
\]
and therefore \( R_1 \subseteq \langle \delta g \rangle \).

In case that \( k = 1 \) we have
\[
g(x) \circ \epsilon x = \epsilon x \circ g(x) \text{ whenever } \epsilon^r = 1.
\]
Thus \( R_1 \subseteq \langle \zeta x, g \rangle \) where \( \zeta \) denotes a primitive \( r \)-th root of unity.

In the following we may assume that \( k \geq 2 \) and \( r \geq 2 \), which makes the preceding lemma applicable.

Assume for the moment that \( R_1 = \langle \epsilon_i g^{(m_i)} | i = 1, \ldots, n \rangle \) with \( n \in \mathbb{N} \). Let \( \zeta \) be a primitive \( r \)-th root of unity and let \( \epsilon_i = \zeta^{x_i} \). As the relations (16) hold, we find
\[
x_i(k^{m_i} - 1) \equiv x_j(k^{m_j} - 1) \text{ mod } r \text{ for } i, j = 1, \ldots, n.
\]
According to Lemma 12 we have
\[
x_i \equiv \lambda(1 + k + \ldots + k^{m_i - 1}) + \lambda_i r' \text{ mod } r \text{ for } i = 1, \ldots, n,
\]
where \( r' = \frac{r}{\gcd(k-1, r)} \).

Let \( \gamma = \zeta^\lambda \); then
\[
(\gamma g)^{(m_i)} = \gamma^{1+k+\ldots+k^{m_i-1}} g^{(m_i)} = \zeta^{\lambda(1+k+\ldots+k^{m_i-1})} g^{(m_i)} = \zeta^{-\lambda r'} \cdot \zeta^{x_i} g^{(m_i)}.
\]
As \( (\zeta^{-\lambda r'})^{\gcd(k-1, r)} = 1 \) we have
\[
\epsilon_i g^{(m_i)} = \delta_i (\gamma g)^{(m_i)} \text{ for } i = 1, \ldots, n,
\]
where \( \delta_i \) are \( \gcd(k-1, r) \)-th roots of unity. Thus
\[
\epsilon_i g^{(m_i)} \in R = \langle \xi x, \gamma g \rangle,
\]
where \( \xi \) is a primitive \( \gcd(k-1, r) \)-th root of unity.

Consider now the case that \( R_1 = \langle \epsilon_i g^{(m_i)} | i = 1, 2, \ldots \rangle \). The numbers
\[
f_h = \gcd(k^{m_1} - 1, \ldots, k^{m_h} - 1) \text{ for } h = 1, 2, \ldots
\]
satisfy \( f_{h+1} | f_h \) for each \( h \). Thus they must remain constant for sufficiently large \( h \), and for such \( h \) we have
\[
f_h = k - 1 = \gcd(k^{m_1} - 1, k^{m_2} - 1, \ldots)
\]
and \( \gcd(m_1, \ldots, m_h) = 1 \). For each such \( h \) we find
\[
\epsilon_i g^{(m_i)} \in \langle \xi x, \gamma_h g \rangle \text{ for } i = 1, \ldots, h,
\]
where \( \xi \) is a primitive \( \gcd(k-1, r) \)-th root of unity and \( \gamma_h \) is an \( r \)-th root of unity. So at least one of the \( \gamma_h \) must occur infinitely often as \( h \) increases. Thus we find eventually for each \( i \)
\[
\epsilon_i g^{(m_i)} \in \langle \xi x, g \rangle
\]
where \( \gamma \) is a fixed \( r \)-th root of unity, which proves the assertion of the proposition.

Putting together Proposition 8, Proposition 9, Proposition 10 and Theorem 1 we obtain a complete list of commutative semigroups:
Theorem 2 Let $S$ be a commutative subsemigroup of $(\mathbb{C}[x], \circ)$. Then $S$ is conjugated to a subsemigroup of one of the following commutative semigroups:

List IV

(i) (linear polynomials): $T$, $S_0$.

(ii) (nonlinear polynomials except the trivial linear polynomial $x$):

$$\mathcal{P}_1 = \{x^n | n \in \mathbb{N}\},$$
$$\mathcal{D}_1 = \{t_n | n \in \mathbb{N}\},$$
$$\mathcal{Q} = \{g^{(n)} | n \in \mathbb{N}_0\} \text{ with } g \in \mathbb{C}[x], [g] \geq 2,$$

where $g$ has gap-form and is not conjugated to a power or a Chebyshev polynomial.

(iii) (nonlinear polynomials and nontrivial linear polynomials):

$$\mathcal{P}_r = \{\zeta x^k | k \equiv 1 \mod r, \zeta^r = 1\} = \{\zeta x, x^k | k \equiv 1 \mod r\} \ (r \geq 2) \text{ where } \xi \text{ is a primitive } r\text{-th root of unity},$$
$$\mathcal{D}_2 = \{\pm t_k | k \text{ odd}\} = \{-x, t_k | k \text{ odd}\},$$
$$\mathcal{R} = \{\zeta x, g\}, \text{ where } [g] \geq 2$$

and $g$ is not conjugated to a power or a Chebyshev polynomial, $g(x) = xp(x^r)$, $r = \mathcal{L}^p(g) > 1$ and $\zeta$ is a primitive $r$-th root of unity.

Proof: Let $S$ be a commutative semigroup. If $S$ consists entirely of linear polynomials, consider Proposition 8.

If there are nonlinear polynomials in $S$, consider the subsemigroup $S'$ of $S$ which contains all nonlinear polynomials of $S$. Then, according to Theorem 1, a certain conjugate $\Phi_L(S')$ of $S'$ is contained in a semigroup occurring in List I.

Suppose first $\Phi_L(S') \subseteq \mathcal{P}_r$ and choose $r$ maximal. Then $\Phi_L(S) \subseteq \mathcal{P}_r$, as no other linear polynomials than $\zeta x$ ($\zeta^r = 1$) permute with $\Phi_L(S')$.

Suppose that $\Phi_L(S') \subseteq \mathcal{D}_1$ or $\Phi_L(S') \subseteq \mathcal{D}_2$. As $\mathcal{G}^p(t_n) = \{x\} \text{ or } \{\pm x\}$ if $n$ is even or odd respectively, we again have $\Phi_L(S) \subseteq \mathcal{D}_1$ or $\Phi_L(S) \subseteq \mathcal{D}_2$.

Finally consider the case $\Phi_L(S') = \mathcal{R}_1 = \{\varepsilon_i g^{(m_i)} | i = 1, \ldots, n\} \text{ or } \{\varepsilon_i g^{(m_i)} | i = 1, 2, \ldots\}$ (from the proof of Theorem 1 we see that in this case actually the equality sign holds). According to Proposition 10 we have $\Phi_L(S') \subseteq \langle \zeta x, g(d) \rangle$ where $d = \gcd(m_1, \ldots, m_n)$ ($\gcd(m_1, m_2, \ldots)$, respectively) and $\zeta$ is a primitive $\mathcal{L}^p(g^{(d)})$-th root of unity. Thus $\Phi_L(S) \subseteq \mathcal{Q}$ or $\Phi_L(S) \subseteq \mathcal{R}$ holds.

$\square$
6 Maximal commutative semigroups and permutable chains

In this section we give some applications of Theorem 2.

First it is easy to determine all maximal commutative subsemigroups of \( (\mathbb{C}[x], \circ) \). Note that a maximal commutative semigroup must be conjugated to one of the semigroups listed in List IV. Thus it suffices to give conditions whether a semigroup occurring in List IV is maximal or not. It turns out that, roughly speaking, this is the case whenever the contrary is not obvious.

**Theorem 3** The commutative semigroups \( T, S_0, P_r \ (r = 1, 2, 3, \ldots) \), \( D_1 \) and \( D_2 \) are maximal (among all commutative semigroups).

The semigroup \( Q = \{g^{(n)}|n \in \mathbb{N}\} \) is maximal if and only if \( L^p(g) = 1 \) and \( g \) cannot be written in the form \( g = f^{(k)} \) for some \( f \in \mathbb{C}[x] \) and \( k > 1 \).

The semigroup \( R = \langle \zeta x, g \rangle \) is maximal if and only if \( g \) cannot be written in the form \( g = \xi f^{(k)} \) where \( L^p(f) = L^p(g), k > 1 \) and \( \xi L^p(g) = 1 \).

**Proof:** The assertions involving \( T, S_0, P_r, D_1 \) and \( D_2 \) are proved in [14].

Consider the case of \( Q \): If \( L^p(g) \neq 1 \) then \( Q \subseteq \langle \zeta x, g \rangle \) where \( \zeta \) is a primitive \( L^p(g) \)-th root of unity. Suppose on the other hand that \( Q \subseteq S \) and \( L^p(g) = 1 \), then \( S \) cannot contain linear polynomials. A conjugate of \( S \) appears in List IV. So the only possibility is that \( S \) is itself of type \( Q \). Thus \( \{g^{(n)}|n \in \mathbb{N}\} \subseteq \{f^{(n)}|n \in \mathbb{N}\} \) which shows that \( g = f^{(k)} \) for some \( k > 1 \).

Consider the case of \( R \) and assume \( R \subseteq S \). Then a conjugate \( \Phi_L(S) \) must occur in List IV. The only possibility is that \( \Phi_L(S) \) is of type \( R \): \( \Phi_L(S) = \langle \xi x, f \rangle \). As \( g \) and \( f \) both have gap-form the linear polynomial \( L \) must be a multiplication. Thus \( S = \langle \xi x, f_1 \rangle \) with \( f_1 = \Phi_L(f) \).

Furthermore \( g = \xi^k x \circ f_1^{(k)} \) and therefore \( L^p(f_1)|L^p(g) \). On the other hand \( \zeta^t x \in R \subseteq S \) for any \( \zeta^t x \in G^p(g) \) and therefore \( L^p(g)|L^p(f_1) \). Thus \( L^p(g) = L^p(f_1) \) and as \( R \neq S \) we must have \( k > 1 \).

□

As another application we solve a problem of W.Nöbauer. To do so we have to recall some notation.

Let \( S \) be a commutative semigroup. Then the (multiplicative) semigroup \( M(S) = \{|f|f \in S\} \) of \( \mathbb{N} \) is called the degree of \( S \). Given any subsemigroup \( M \) of \( \mathbb{N} \) we obviously have the following commutative semigroups \( S \) with \( M(S) = M \):

**List V**

(i) \( S = \{x^n|n \in M\} \),

(ii) \( S = \{\zeta x^n|n \in M, \zeta^r = 1\} \) if \( M \) consists entirely of numbers \( \equiv 1 \mod r \),

(iii) \( S = \{t_n|n \in M\} \),

(iv) \( S = \{\pm t_n|n \in M\} \) if \( M \) consists entirely of odd numbers.

For which semigroups \( M \) are the above listed \( S \) the only commutative semigroups (apart of conjugates) with \( M(S) = M \)?
Theorem 4 Let $M$ be a (multiplicative) subsemigroup of $N$. Every commutative semigroup $S$ with $M(S) = M$ (apart of conjugates) occurs in List V above if and only if $M$ does not consist entirely of powers of one specific number.

Proof: Suppose that $M = \{k^n|n \in M_1\}$ for some $k \in N$ and $M_1 \subseteq N$. Let $g$ be a polynomial of degree $k$ which is neither conjugated to a power nor to a Chebyshev polynomial. Consider the commutative semigroup $S = \{g(n)|n \in M_1\}$, then $M(S) = M$ and no conjugate of $S$ can occur in List V.

On the other hand suppose $M(S) = M$ and no conjugate of $S$ occurs among the semigroups listed above. In the case $M = \{1\}$ we are done. If $M \neq \{1\}$ the only possibility for a conjugate of $S$ is to be a subsemigroup of either $Q$ or $R$ (of List IV). In both cases $M(S) \subseteq \{k^n|n \in N\}$ with $k = |g|$.

□

References


[35] O. Ore: On a special class of polynomials (+ Errata), Trans. AMS 35 (1933), 559-584; 36 (1934), 275.


