FORMAL POWER SERIES AND SOME THEOREMS OF J.F.RITT IN ARBITRARY CHARACTERISTIC

Gerhard Dorfer, Harald Woracek

1 Introduction

In [R1]-[R4] J.F.Ritt proved some theorems concerning composition of (complex) rational functions. These results, when considered for polynomials, expose some interesting facts of the composition semigroup $\langle \mathbb{C}[z], \circ \rangle$ of complex polynomials. For example the classification of commutative subsemigroups of $\langle \mathbb{C}[z], \circ \rangle$ relies heavily on some results of [R1]-[R4] (compare [EW]). J.F.Ritt used analytic methods in his proofs. At the present time there are different approaches and 'purely algebraic' proofs of various of the mentioned results known, see e.g. [B1], [B2], [DW], [E], [F], [FR1], [FR2], [Ja], [LN], [L], [S], [T], [Z]. At the same time generalizations to polynomials over fields $\mathbb{K}$ with certain properties are obtained. Mostly $\mathbb{K}$ is assumed to be algebraically closed and to have characteristic zero. The assumption of algebraic closedness is in some respects no essential restriction (compare f.i. [S], Lemma I.5.2). However, the consideration of fields with nonzero characteristic is more complicated and requires (naturally) additional assumptions, see e.g. [B2], [EN], [S], [Z]. In [B1] and [DW] there can be found some examples showing differences between zero and nonzero characteristic.

It is the aim of the present note to continue these efforts: We generalize some results of [R1]-[R4] to polynomials over algebraically closed fields of arbitrary characteristic. The theorems under consideration are: A characterization of polynomials which satisfy a certain rational functional equation (Theorem 4.1), a characterization of those pairs of polynomials which have, up to a multiplicative constant, common iterates (Theorem 4.3), and a characterization of pairs of permutable polynomials (Theorem 5.1). The additional assumption we impose is that the characteristic of the coefficient field $\mathbb{K}$ is not a divisor of the degree of any polynomial involved in the above statements. Some of the given proofs are even new in the case of characteristic zero. Also our method is fairly elementary, disregarding the application of a result of [Z] in the proof of Theorem 5.1.

The first part of this paper consists of two sections. In Section 2 we introduce certain fields of (one-sided finite) power series and study composition of such series. These expositions, while elementary, provide some basic tools for our further investigations. The subject of Section 3 is the study of the so-called Boettcher function of a power series. In complex analysis, in particular in iteration theory, the Boettcher function is commonly used (compare [Be], [Bo]). In the setting of formal power series Boettcher functions have been introduced in [K].

The second part is devoted to the proof of the three theorems of Ritt mentioned above. Section 4 is concerned with Theorem 4.1 and Theorem 4.3. In Section 5

we characterize permutable polynomials. Our proof relies on the characterization of so called standard solutions ('Ritt's second theorem'). An elementary proof of this result can be found in [S], we use a stronger form as given in [Z].

2 Semigroups of power series

Throughout this paper let $\mathbb{K}$ be an algebraically closed field of characteristic $\pi \geq 0$. We consider formal power series with coefficients in the field $\mathbb{K}$. As usual we write a (possibly double infinite) formal power series as

$$f = \sum_{n=-\infty}^{\infty} a_n z^n, \ a_n \in \mathbb{K}. \tag{2.1}$$

If $f \neq 0$ is given by (2.1), let

$$\operatorname{Ord}_0 f := \inf\{n \in \mathbb{Z} | a_n \neq 0\}, \ \operatorname{Ord}_\infty f := \sup\{n \in \mathbb{Z} | a_n \neq 0\},$$

and denote by $\mathcal{R}_0$ and $\mathcal{R}_\infty$ the sets

$$\mathcal{R}_0 := \{f | \operatorname{Ord}_0 f > -\infty\} \cup \{0\}, \ \mathcal{R}_\infty := \{f | \operatorname{Ord}_\infty f < \infty\} \cup \{0\}.$$

If f and g are two formal power series, the sum $f + g$ is defined as usual. If both f and g are contained in $\mathcal{R}_0$ ($\mathcal{R}_\infty$, respectively), we define the product $f \cdot g$ by means of the Cauchy product. Clearly, $\langle \mathcal{R}_0, +, \cdot \rangle$ and $\langle \mathcal{R}_\infty, +, \cdot \rangle$ are fields. Note that

$$\operatorname{Ord}_0 (f+g) \geq \min\{\operatorname{Ord}_0 f, \operatorname{Ord}_0 g\}, \ \operatorname{Ord}_\infty (f+g) \leq \max\{\operatorname{Ord}_\infty f, \operatorname{Ord}_\infty g\},$$

where equality holds if $\operatorname{Ord}_0 f \neq \operatorname{Ord}_0 g$ ($\operatorname{Ord}_\infty f \neq \operatorname{Ord}_\infty g$), that $\operatorname{Ord}_0 (f \cdot g) = \operatorname{Ord}_0 f + \operatorname{Ord}_0 g$ ($\operatorname{Ord}_\infty (f \cdot g) = \operatorname{Ord}_\infty f + \operatorname{Ord}_\infty g$) and that $\operatorname{Ord}_0 f^{-1} = -\operatorname{Ord}_0 f$ ($\operatorname{Ord}_\infty f^{-1} = -\operatorname{Ord}_\infty f$).

Remark 2.1. If $f \in \mathcal{R}_0 \cap \mathcal{R}_\infty$, the inverse f^{-1} may have a different form depending on whether it is computed in $\mathcal{R}_0$ or $\mathcal{R}_\infty$.

Many properties of $\mathcal{R}_\infty$ can be deduced from the corresponding properties of $\mathcal{R}_0$ (or vice versa) by making use of the following transformations. Define

$$\varphi : \sum_{n=-\infty}^{\infty} a_n z^n \mapsto \sum_{n=-\infty}^{\infty} a_{-n} z^n,$$

and

$$\Phi_{\frac{1}{z}} f := (\varphi f)^{-1} : \mathcal{R}_0 \setminus \{0\} \to \mathcal{R}_\infty \setminus \{0\}.$$

Note that $\Phi_{\frac{1}{z}}$ can also be considered as a mapping of $\mathcal{R}_\infty \setminus \{0\}$ onto $\mathcal{R}_0 \setminus \{0\}$ if the (multiplicative) inverse is computed in the field $\mathcal{R}_0$. For notational reasons we put $\Phi_{\frac{1}{z}} 0 := 0$.

In the following lemma we collect some properties of φ and $\Phi_{\frac{1}{z}}$. These are proved by a straightforward calculation.

Lemma 2.2. *The mapping φ satisfies $\varphi \circ \varphi = \mathrm{id}$ and $\mathrm{Ord}_0 f = -\mathrm{Ord}_\infty \varphi f$. The restriction of φ to $\mathcal{R}_0$ ($\mathcal{R}_\infty$) is a field isomorphism onto $\mathcal{R}_\infty$ ($\mathcal{R}_0$).*

The mapping $\Phi_{\frac{1}{z}}$ induces a bijection of $\mathcal{R}_0$ onto $\mathcal{R}_\infty$. It is an isomorphism with respect to $\cdot$. It satisfies $\Phi_{\frac{1}{z}} \circ \Phi_{\frac{1}{z}} = \mathrm{id}_{\mathcal{R}_0}$, where the right (left) factor is considered as a mapping of $\mathcal{R}_0$ to $\mathcal{R}_\infty$ ($\mathcal{R}_\infty$ to $\mathcal{R}_0$). If $f \in \mathcal{R}_0$, then $\mathrm{Ord}_\infty \Phi_{\frac{1}{z}} f = \mathrm{Ord}_0 f$.

Definition 2.3. Let $f, g \in \mathcal{R}_0$, $\mathrm{Ord}_0 g \geq 1$, be given and write

$$f = \sum_{n=-k}^{\infty} a_n z^n, \ g = \sum_{n=1}^{\infty} b_n z^n.$$

We define a composition $f \circ g$ as

$$f \circ g = \sum_{n=-k}^{\infty} a_n g^n. \tag{2.2}$$

In order to justify the definition (2.2), note that the n-th coefficient of $f \circ g$ is a finite sum of products of coefficients a_l and powers of b_m (compare [Je]).

Note that in certain cases the right hand side of the relation (2.2) makes sense even if $\mathrm{Ord}_0 g < 1$. For example if the left factor f is contained in $\mathcal{R}_0 \cap \mathcal{R}_\infty$, the composition $f \circ g$ is well defined by (2.2) for any $g \in \mathcal{R}_0$, $g \neq 0$, as an element of $\mathcal{R}_0$.

As usual the compositional power $\underbrace{f \circ \cdots \circ f}_{n \text{ times}}$ will be denoted by $f^{(n)}$ whenever it exists.

Let $f, g \in \mathcal{R}_\infty$, $\mathrm{Ord}_\infty g \geq 1$, be given. We define a composition $f \circ g$ as

$$f \circ g := \Phi_{\frac{1}{z}} \left((\Phi_{\frac{1}{z}} f) \circ (\Phi_{\frac{1}{z}} g) \right).$$

Clearly the mapping $\Phi_{\frac{1}{z}}$ is compatible with $\circ$. It is easily checked that this definition extends the usual definition for, say, rational functions.

In the following proposition we collect the basic properties of the introduced notions. Since these results are seen by fairly elementary arguments for $\mathcal{R}_0$ and follow on applying $\Phi_{\frac{1}{z}}$ for $\mathcal{R}_\infty$, we leave the proof to the reader.

Proposition 2.4. *If $f, g \in \mathcal{R}_0$, $\mathrm{Ord}_0 g \geq 1$, then $\mathrm{Ord}_0 (f \circ g) = \mathrm{Ord}_0 f \cdot \mathrm{Ord}_0 g$. The operations $+$ and $\cdot$ are left distributive with respect to $\circ$, i.e. for $f, g, h \in \mathcal{R}_0$, $\mathrm{Ord}_0 h \geq 1$, we have*

$$(f+g) \circ h = (f \circ h) + (g \circ h), \ (f \cdot g) \circ h = (f \circ h) \cdot (g \circ h), \ (f \circ h)^{-1} = f^{-1} \circ h.$$

In particular, $f \circ h = g \circ h$ implies $f = g$. The associative law for $\circ$ holds whenever all occurring products are defined. The element z of $\mathcal{R}_0$ is neutral with respect to $\circ$. The set

$$\mathcal{R}_0^* := \{f \in \mathcal{R}_0 | \operatorname{Ord}_0 f = 1\}$$

consists of units with respect to $\circ$. The same assertions hold if everywhere 0 is replaced by ∞.

The notion introduced next is a technical but important tool for our argumentations.

Definition 2.5. Let $f = \sum_{n=-\infty}^{\infty} a_n z^n$, $f \neq a_k z^k$. If $\operatorname{Ord}_0 f = m > -\infty$, put

$$\mathfrak{l}_0(f) := \operatorname{Ord}_0 (f - a_m z^m) - m.$$

If $\operatorname{Ord}_\infty f = m < \infty$, put

$$\mathfrak{l}_\infty(f) := m - \operatorname{Ord}_\infty (f - a_m z^m).$$

For notational reasons we put $\mathfrak{l}_0(a_k z^k) = \mathfrak{l}_\infty(a_k z^k) = \infty$.

In the case that $\mathfrak{l}_0(f) > 1$ ($\mathfrak{l}_\infty(f) > 1$) we say that f has gap form. In the next lemma we list some properties of $\mathfrak{l}_0$ and $\mathfrak{l}_\infty$. These results and those of the following corollary will be used extensively in the subsequent sections.

Lemma 2.6. *Let $f, g \in \mathcal{R}_0$. Then the following holds:*

(i) For all $k \in \mathbb{Z}$ we have $\mathfrak{l}_0(z^k f) = \mathfrak{l}_0(f)$.

(ii) If $k \in \mathbb{Z} \setminus \{0\}$ and the characteristic π of $\mathbb{K}$ does not divide k, we have $\mathfrak{l}_0(f^k) = \mathfrak{l}_0(f)$. Otherwise $\mathfrak{l}_0(f^k) > \mathfrak{l}_0(f)$, if these numbers are finite.

(iii) In any case $\mathfrak{l}_0(f \cdot g) \geq \min\{\mathfrak{l}_0(f), \mathfrak{l}_0(g)\}$. Equality holds in this relation if $\mathfrak{l}_0(f) \neq \mathfrak{l}_0(g)$.

(iv) For all $k \in \mathbb{N}$, we have $\mathfrak{l}_0(f \circ z^k) = k\mathfrak{l}_0(f)$.

In case that $\operatorname{Ord}_0 g \geq 1$ we have furthermore:

(v) If $m = \operatorname{Ord}_0 f \neq 0$, then

$$\mathfrak{l}_0(f \circ g) \geq \min\{\mathfrak{l}_0(g^m), \mathfrak{l}_0(f)\operatorname{Ord}_0 g\} \geq \min\{\mathfrak{l}_0(g), \mathfrak{l}_0(f)\operatorname{Ord}_0 g\}.$$

In the first relation equality holds if $\mathfrak{l}_0(g^m) \neq \mathfrak{l}_0(f)\operatorname{Ord}_0 g$. If additionally $\pi \nmid m$, then also in the second relation equality holds. In the case $\operatorname{Ord}_0 f = 0$, we have $\mathfrak{l}_0(f \circ g) = \mathfrak{l}_0(f)\operatorname{Ord}_0 g$.

(vi) If $\operatorname{Ord}_0 g > 1$ and $\pi \nmid \operatorname{Ord}_0 g$, then $\mathfrak{l}_0(g^{(k)}) = \mathfrak{l}_0(g)$ for all $k \in \mathbb{N}$.

(vii) If $g \in \mathcal{R}_0^*$, we have $\mathfrak{l}_0(g^{(-1)}) = \mathfrak{l}_0(g)$.

Moreover, we have

$$\mathfrak{l}_\infty(\Phi_{\frac{1}{z}} f) = \mathfrak{l}_0(f).$$

Hence the assertions (i)-(vii) hold for $f, g \in \mathcal{R}_\infty$, if $\mathfrak{l}_0$ (Ord_0) is replaced by $\mathfrak{l}_\infty$ (Ord_∞).

Proof : The relations *(i)* and *(iv)* are immediate.

Let $f = z^m \tilde{f}$, $\mathrm{Ord}_0 \tilde{f} = 0$ and $\tilde{f} = a_0 + \sum_{n=\mathfrak{l}_0(f)}^{\infty} a_n z^n$. Then

$$\tilde{f}^k = a_0^k + k a_0^{k-1} a_{\mathfrak{l}_0(f)} z^{\mathfrak{l}_0(f)} + \ldots,$$

which implies *(ii)*. In order to show $\mathfrak{l}_0(\tilde{f}^{-1}) = \mathfrak{l}_0(\tilde{f})$, compare coefficients in the relation $1 = \tilde{f} \cdot \tilde{f}^{-1}$. A similar argument shows that *(iii)* holds.

Using distributivity we obtain $f \circ g = g^m \cdot (\tilde{f} \circ g)$. Since

$$\tilde{f} \circ g = a_0 + a_{\mathfrak{l}_0(f)} g^{\mathfrak{l}_0(f)} + \ldots,$$

we find $\mathfrak{l}_0(\tilde{f} \circ g) = \mathfrak{l}_0(f)\mathrm{Ord}_0 g$. Applying *(ii)* and *(iii)* yields *(v)*.

The relation *(vi)* follows immediately from *(v)* and *(ii)*, *(vii)* can be checked easily by comparing coefficients in $g^{(-1)} \circ g = z$. The assertion concerning $\mathfrak{l}_\infty(\Phi_{\frac{1}{z}} f)$ follows by applying *(ii)*.

□

Corollary 2.7. *Let $f, g \in \mathcal{R}_0$ ($\mathcal{R}_\infty$) and assume that*

$$f \circ z^m = z^m \circ g,$$

for some $m \in \mathbb{N}$. If $\pi \nmid m$ and $\mathfrak{l}_0(g) \neq m\mathfrak{l}_0(f)$ ($\mathfrak{l}_\infty(g) \neq m\mathfrak{l}_\infty(f)$), then $f = az^n$ and $g = bz^n$ for some $n \in \mathbb{N}$ and $a, b \in \mathbb{K}$, $a = b^m$. If $\pi | m$, the same assertion holds under the hypothesis $\mathfrak{l}_0(g) \geq m\mathfrak{l}_0(f)$ ($\mathfrak{l}_\infty(g) \geq m\mathfrak{l}_\infty(f)$).

Proof : If f or g is a monomial, the assertion follows immediately. Otherwise, *(ii)* of Lemma 2.6 implies $\mathfrak{l}_0(z^m \circ g) = \mathfrak{l}_0(g)$ if $\pi \nmid m$ and $\mathfrak{l}_0(z^m \circ g) > \mathfrak{l}_0(g)$ if $\pi | m$. By *(iv)* we have $\mathfrak{l}_0(f \circ z^m) = m\mathfrak{l}_0(f)$.

□

For $f \in \mathcal{R}_0$, $f \neq a_k z^k$, denote by $k_1 < k_2 < \ldots$ those indices with $a_{k_i} \neq 0$. Then the gap degree of f is defined as the number

$$\mathfrak{L}(f) := \gcd\{k_2 - k_1, k_3 - k_2, \ldots\}.$$

For $f \in \mathcal{R}_\infty$ let $\mathfrak{L}(f) := \mathfrak{L}(\phi f)$. For notational convenience we put $\mathfrak{L}(a_k z^k) := 0$. A related notion has been considered in [T].

Lemma 2.8. *Let $f \in \mathcal{R}_0$ ($\mathcal{R}_\infty$) be given. An element $\varepsilon \in \mathbb{K}$ satisfies a relation of the form*

$$f \circ \varepsilon z = \delta z \circ f \tag{2.3}$$

for some $\delta \in \mathbb{K}$, if and only if $\varepsilon^{\mathfrak{L}(f)} = 1$.
Proof : In the case that f is a monomial or $\mathfrak{l}(f) = 1$, the assertion is obvious. If $f \neq a_k z^k$ we write $f = \sum_{i=1}^{\infty} a_{k_i} z^{k_i}$ and (2.3) for some $\varepsilon \neq 1$ implies

$$\varepsilon^{k_{i+1}-k_i} = 1,$$

for all i, hence $\varepsilon^{\mathfrak{L}(f)} = 1$. The converse is seen by the same argument.

□

3 The Boettcher function

In this section we study conjugation with respect to $\circ$.

Definition 3.1. Let $g \in \mathcal{R}_0^*$ be given. Denote by Φ_g the conjugation

$$\Phi_g f := g^{(-1)} \circ f \circ g, \ f \in \mathcal{R}_0, \ \operatorname{Ord}_0 f \geq 1.$$

Clearly then:

Lemma 3.2. *If $f \in \mathcal{R}_0$ and $g \in \mathcal{R}_0^*$, then $\operatorname{Ord}_0 f = \operatorname{Ord}_0 \Phi_g f$. The mapping Φ_g is a composition isomorphism of $\{f \in \mathcal{R}_0 | \operatorname{Ord}_0 f \geq 1\}$ onto itself.*

In Definition 3.1 and Lemma 3.2 we can replace $\mathcal{R}_0$ and $\operatorname{Ord}_0 f$ by $\mathcal{R}_\infty$ and $\operatorname{Ord}_\infty f$.

Next recall that certain elements of $\mathcal{R}_0$ ($\mathcal{R}_\infty$) can be conjugated to powers z^m. This fact was first observed by L.Boettcher (see [Bo]), and is commonly used in the framework of complex analysis (see e.g. [Be], [R1]). In an algebraic setting the following result has been proved in principle in [K] for power series f contained in $\mathcal{R}_0$. The assertion in the case $f \in \mathcal{R}_\infty$ follows by an application of $\Phi_{\frac{1}{z}}$. Also note that the uniqueness part of the following proposition appears implicitly in the proof of Hilfssatz 4 of [K].

Proposition 3.3. *Let $f \in \mathcal{R}_0$ ($f \in \mathcal{R}_\infty$), $\operatorname{Ord}_0 f = m \geq 2$ ($\operatorname{Ord}_\infty f = m \geq 2$), $\pi \nmid m$. Then there exists an element $\beta \in \mathcal{R}_0^*$ ($\beta \in \mathcal{R}_\infty^*$), such that*

$$\Phi_\beta f = \beta^{(-1)} \circ f \circ \beta = z^m. \tag{3.1}$$

With β also $\beta \circ \varepsilon z$ satisfies (3.1) if $\varepsilon^{m-1} = 1$. Conversely, any element of $\mathcal{R}_0^*$ ($\mathcal{R}_\infty^*$) satisfying (3.1) is of this form.

If β satisfies (3.1), we refer to β as a Boettcher function of f. In the sequel β_f will be the generic notation for a Boettcher function of f. In the following corollary we complete the answer to the question whether in the case $\pi | \operatorname{Ord}_0 f$ a Boettcher function exists or not.

Corollary 3.4. *Let f and m be as in Proposition 3.3, and assume that $m = d\pi^k$ with $k \geq 1$ and $d \geq 1$. Then there exists a Boettcher function β of f if and only if f can be written as $g \circ z^{\pi^k}$ for some g.*

Proof : Assume first that $f \circ \beta = \beta \circ z^m$ for some β. An application of the chain rule for differentiation shows that $f' = 0$, i.e. $f = f_1 \circ z^\pi$. Hence $f_1 \circ \beta_1 = \beta \circ z^{\frac{m}{\pi}}$, where β_1 is a solution of the equation $\beta_1 \circ z^\pi = z^\pi \circ \beta$, i.e. the coefficients of β_1 are exactly the π-th powers of the coefficients of β. If $\pi | \frac{m}{\pi}$ the chain rule again shows that $f_1' = 0$. Proceeding inductively we end up with a relation of the form

$$f_k \circ \beta_k = \beta \circ z^d \tag{3.2}$$

where the coefficients of β_k are exactly the π^k-th powers of those of β. Moreover, we have $f = f_k \circ z^{\pi^k}$.

Conversely assume that $f = g \circ z^{\pi^k}$ and consider first the case that $d \geq 2$. Then it is possible to solve the equation (3.2) with g in place of f_k. This is seen by the same argument as employed in the proof of Proposition 3.3, see [K]. Then, clearly, $f \circ \beta = \beta \circ z^m$.

If $d = 1$, a straightforward argument shows that there exists a choice of β in order to satisfy $f \circ \beta = \beta \circ z^m$.

□

Remark 3.5. As the following example shows, one cannot expect a uniqueness result like Proposition 3.3, if $\pi | m$: Let $\beta \in \mathcal{R}_0$, $\operatorname{Ord}_0 \beta \geq 1$, be such that every nonzero coefficient is a $(\pi^k - 1)$-th root of unity, then

$$z^{\pi^k} \circ \beta = \beta \circ z^{\pi^k}.$$

In the following we study some properties of Boettcher functions.

Corollary 3.6. *Let f be as in Proposition 3.3 and let $n \in \mathbb{N}$. Then each Boettcher function $\beta_{f^{(n)}}$ of $f^{(n)}$ is of the form*

$$\beta_{f^{(n)}} = \beta_f \circ \varepsilon z, \ \varepsilon^{m^n - 1} = 1.$$

Proof : Note that

$$\Phi_{\beta_f}(f^{(n)}) = (\Phi_{\beta_f} f)^{(n)} = z^{m^n}.$$

The assertion follows from the uniqueness part of Proposition 3.3.

□

In the sequel we will always consider polynomials as elements of $\mathcal{R}_\infty$, since then a nonlinear polynomial has a Boettcher function if π is not a divisor of its degree. Clearly, an element $f \in \mathcal{R}_\infty$ is a polynomial if and only if $\operatorname{Ord}_0 f \geq 0$. Hence

our main interest will be in studying Boettcher functions of elements of $\mathcal{R}_\infty$. As examples for Boettcher functions may serve:

Lemma 3.7. *A Boettcher function β_{z^m} is given by $\beta_{z^m} = z$. If t_n denotes the Dickson polynomial of degree n, a Boettcher function is given by*

$$\beta_{t_n} = z + \frac{1}{z}. \tag{3.3}$$

Moreover, a Boettcher function β_{-t_n} is $(z + \frac{1}{z}) \circ \alpha z$, where $\alpha^{n-1} = -1$.
Proof : The first assertion is self-evident. To prove (3.3) recall that (see f.i. Corollary I.5.2 in [S])

$$t_n \circ (z + \frac{1}{z}) = (z + \frac{1}{z}) \circ z^n.$$

□

In the following some properties of Boettcher functions are collected.

Proposition 3.8. *Let $f \in \mathcal{R}_\infty$, $\operatorname{Ord}_\infty f = m \geq 2$, $\pi \nmid m$, be given.*

(i) If $g \in \mathcal{R}_\infty^*$ then $\beta_{\Phi_g f} = g^{(-1)} \circ \beta_f$.

(ii) We have $\mathfrak{l}_\infty(\beta_f) = \mathfrak{l}_\infty(f)$.

(iii) Assume that $\varepsilon \in \mathbb{K}$ satisfies $\varepsilon^{\mathfrak{L}(f)} = 1$, then β_f commutes with εz. In fact $\mathfrak{L}(\beta_f) = \mathfrak{L}(f)$.

(iv) If $k | \mathfrak{L}(\beta_f)$, and $\varepsilon \in \mathbb{K}$, $\varepsilon^k = 1$, then for some $\delta \in \mathbb{K}$ with $\delta^{(m-1)k} = 1$ we have $\beta_{\varepsilon f} = \beta_f \circ \delta z$.

Proof : To prove the first assertion note that

$$(\beta_f^{(-1)} \circ g) \circ (g^{(-1)} \circ f \circ g) \circ (g^{(-1)} \circ \beta_f) = z^m.$$

Assume next that $f = \sum_{n=-\infty}^{m} a_n z^n$, $\beta_f = \sum_{n=-\infty}^{1} b_n z^n$. Since $f \circ \beta_f = \beta_f \circ z^m$, we have

$$f \circ \beta = a_m(b_1 z + b_0 + b_{-1}\frac{1}{z} + \ldots)^m + a_{m-1}(b_1 z + b_0 + b_{-1}\frac{1}{z} + \ldots)^{m-1} + \ldots =$$

$$= b_1 z^m + b_0 + b_{-1}\frac{1}{z^m} + \ldots. \tag{3.4}$$

If $\mathfrak{l}_\infty(f) = k > 1$, it follows from (3.4) by an inductive argument that $\mathfrak{l}_\infty(\beta_f) = k$. If on the other hand $\mathfrak{l}_\infty(f) = 1$, there exists a linear polynomial $L = z + c$, such that $\mathfrak{l}_\infty(\Phi_L f) > 1$, in fact $c = -\frac{a_{m-1}}{m a_m}$. By the already proved statement

$$\mathfrak{l}_\infty(L^{(-1)} \circ \beta_f) = \mathfrak{l}_\infty(\beta_{\Phi_L f}) > 1,$$

hence $\mathfrak{l}_\infty(\beta_f) = 1$.

The assertions *(iii)* and *(iv)* are obvious if f is not of gap form. Assume that $\mathfrak{l}_\infty(f) > 1$ and that ε is a primitive $\mathfrak{L}(f)$-th root of unity. By Lemma 2.8 there exists a number $l \in \mathbb{N}$, such that $f \circ \varepsilon z = \varepsilon^l z \circ f$, hence for all $k \in \mathbb{N}$ and certain numbers l_k we have

$$z^{m^k} \circ \Phi_{\beta_f}(\varepsilon z) = \Phi_{\beta_f}(\varepsilon^{l_k} z) \circ z^{m^k}.$$

Since ε is a root of unity ε^{l_k} assumes only finitely many different values. For some k Corollary 2.7 is applicable and shows $\varepsilon z \circ \beta_f = \beta_f \circ az$, for some $a \in \mathbb{K}$. Since $\operatorname{Ord}_\infty \beta_f = 1$, we must have $a = \varepsilon$, i.e. εz commutes with β_f and we conclude that $\mathfrak{L}(f)|\mathfrak{L}(\beta_f)$.

Assume on the other hand that ε is a primitive $\mathfrak{L}(\beta_f)$-th root of unity. Then εz commutes with β_f and we find

$$\Phi_{\beta_f}(f \circ \varepsilon z) = z^m \circ \varepsilon z = \varepsilon^m z \circ z^m.$$

Hence $f \circ \varepsilon z = \varepsilon^m z \circ f$ and we conclude that $\mathfrak{L}(\beta_f)|\mathfrak{L}(f)$.

To prove the remaining part let $k|\mathfrak{L}(\beta_f)$ and $\varepsilon^k = 1$ be given. Then

$$\beta_f^{(-1)} \circ \varepsilon f \circ \beta_f = \varepsilon z \circ \beta_f^{(-1)} \circ f \circ \beta_f = \varepsilon z \circ z^m = \delta z \circ z^m \circ \delta^{-1} z.$$

□

Lemma 3.9. *Let $f, g \in \mathcal{R}_\infty$, $\operatorname{Ord}_\infty g \geq 1$, be given and assume that $\operatorname{Ord}_0(f \circ g)$, $\operatorname{Ord}_0 g \geq 0$. Then also $\operatorname{Ord}_0 f \geq 0$.*

Proof : Write $f = f_+ + f_-$ with $\operatorname{Ord}_0 f_+ \geq 0$ and $\operatorname{Ord}_\infty f_- < 0$. Then $\operatorname{Ord}_\infty(f_- \circ g) < 0$ and $\operatorname{Ord}_0(f_+ \circ g) \geq 0$. Since $f \circ g = (f_+ \circ g) + (f_- \circ g)$ the assumption implies that $f_- \circ g = 0$, hence $f_- = 0$.

□

Lemma 3.10. *Let $\beta \in \mathcal{R}_\infty^*$, $\mathfrak{l}_\infty(\beta) > 1$, $n \in \mathbb{N}$, $n \geq 2$, and $\varepsilon \in \mathbb{K} \setminus \{0\}$. Assume that*

$$f = \beta \circ z^n \circ \beta^{(-1)},\ g = \beta \circ \varepsilon z^n \circ \beta^{(-1)}$$

are both polynomials. Then β commutes with εz.

Proof : We have

$$\beta \circ \varepsilon z \circ \beta^{(-1)} \circ f = \beta \circ \varepsilon z \circ \beta^{(-1)} \circ \beta \circ z^n \circ \beta^{(-1)} = g.$$

By Lemma 3.9, $\beta^{(-1)} \circ \varepsilon z \circ \beta = L$ where L is a linear polynomial. Since $\mathfrak{l}_\infty(\beta^{(-1)} \circ \varepsilon z \circ \beta) > 1$, L must be equal to εz.

□

The following proposition characterizes permutable power series by means of their Boettcher function. A partial result in this direction has been obtained in [K].

Proposition 3.11. *Let $f, g \in \mathcal{R}_\infty$, $\operatorname{Ord}_\infty f = m_f \geq 2$, $\operatorname{Ord}_\infty g = m_g \geq 2$, $\pi \nmid m_f, m_g$, be given. Then f commutes with g if and only if*

$$\beta_g = \beta_f \circ \varepsilon z,$$

for some ε with $\varepsilon^{(m_f-1)(m_g-1)} = 1$.
Proof : Assume that $f \circ g = g \circ f$ holds, then

$$z^{m_f} \circ \Phi_{\beta_f} g = \Phi_{\beta_f} g \circ z^{m_f}. \tag{3.5}$$

Corollary 2.7 implies that $\Phi_{\beta_f} g = \delta z^{m_g}$. Now (3.5) shows that $\delta^{m_f-1} = 1$. Choose ε such that $\varepsilon^{m_g-1} = \delta^{-1}$. Then

$$\Phi_{\beta_f \circ \varepsilon z}(g) = \Phi_{\varepsilon z}(\Phi_{\beta_f}(g)) = \varepsilon^{-1} z \circ \delta z^{m_g} \circ \varepsilon z = z^{m_g},$$

hence $\beta_f \circ \varepsilon z$ is a Boettcher function for g. Clearly $\varepsilon^{(m_g-1)(m_f-1)} = 1$.

Conversely, assume that $\beta_g = \beta_f \circ \varepsilon z$ for some ε with $\varepsilon^{(m_g-1)(m_f-1)} = 1$. Then

$$f \circ g = \beta_f \circ z^{m_f} \circ \beta_f^{(-1)} \circ \beta_f \circ \varepsilon z \circ z^{m_g} \circ \varepsilon^{-1} z \circ \beta_f^{(-1)} =$$

$$= \beta_f \circ z^{m_f} \circ \varepsilon^{1-m_g} z \circ z^{m_g} \circ \beta_f^{(-1)} = \beta_f \circ \varepsilon^{m_f(1-m_g)} \circ z^{m_f m_g} \circ \beta_f^{(-1)} =$$

$$= \beta_f \circ \varepsilon^{1-m_g} z \circ z^{m_g m_f} \circ \beta_f^{(-1)} = \beta_f \circ \varepsilon z \circ z^{m_g} \circ \varepsilon^{-1} z \circ \beta_f^{(-1)} \circ \beta_f \circ z^{m_f} \circ \beta_f^{(-1)} = g \circ f.$$

□

4 Polynomials with a rational functional equation and polynomials with a common iterate

The following result characterizes those polynomials f whose Boettcher function is rational, i.e. $\operatorname{Ord}_0 \beta_f > -\infty$. Note that this condition means that f satisfies the functional equation

$$f(\beta(z)) = \beta(z^n),\ n = \operatorname{Ord}_\infty f,$$

for the rational function β (compare [R2]).

Theorem 4.1. *Let f be a polynomial, $\mathrm{Ord}_\infty f = n \geq 2$, $n \neq \pi^k$, and assume that $f \circ \beta = \beta \circ z^n$ for some $\beta \in \mathcal{R}^*_\infty$. Then $\mathrm{Ord}_0 \beta > -\infty$, if and only if there exists a linear polynomial L, such that $\Phi_L(f)$ either is a power or up to the sign a Dickson polynomial.*

Proof : Let $\mathrm{Ord}_0 \beta = -m > -\infty$, i.e.

$$\beta = bz + b_0 + b_1 \frac{1}{z} + \ldots + b_m \frac{1}{z^m},$$

and $f \circ \beta = \beta \circ z^n$ for the polynomial f. We assume that $m \geq 2$ and deduce a contradiction. Note that with β also $L \circ \beta$ satisfies these conditions (replacing f by $L \circ f \circ L^{-1}$). Hence we may assume that $b_0 \neq 0$.

Write $f = a_n z^n + \ldots + a_0$. We compute some of the powers of z which occur in $f \circ \beta$ or $\beta \circ z^n$. Let

$$i_* := \min\{i \geq 1 | \pi \nmid \binom{n}{i}\},$$

and note that by our assumption $n \neq \pi^k$ we have $i_* \leq \left[\frac{n}{2}\right]$. The highest (lowest) powers occuring in β^n are z^n, z^{n-i_*} (z^{-mn}, $z^{-mn+i_* \mathfrak{l}_0(\beta)}$). If follows that $a_{n-1} = \ldots = a_{n-i_*+1} = 0$ and $a_{n-i_*} \neq 0$. The lowest power which occurs in β^{n-i_*} is $z^{-m(n-i_*)}$. It follows that

$$-m(n-i_*) = -mn + i_* \mathfrak{l}_0(\beta),$$

hence $\mathfrak{l}_0(\beta) = m$, i.e. β is of the form $bz + b_0 + b_m \frac{1}{z^m}$.

Again we use the freedom of composing β with a linear polynomial from the left and assume that $b_0 = 0$. The highest (lowest) powers occuring in β^n are then z^n, $z^{n-i_*-mi^*}$ (z^{-mn}, $z^{-m(n-i_*)+i_*}$). If $n - i_* - mi_* \geq 0$, we obtain with similar arguments as above that $m(n-i_*+mi_*) = m(n-i_*)+i_*$, a contradiction. Otherwise it follows that f is a monomial, which obviously is a contradiction again.

If $\mathrm{Ord}_0 \beta_f = 1$ we immediately have that f is conjugated to the power z^n. If $\mathrm{Ord}_0 \beta_f = -1$, $\beta_f = bz + b_1 \frac{1}{z}$, we can write

$$\beta_f = \frac{b}{\alpha} z \circ (z + \frac{1}{z}) \circ \alpha z,$$

where $\alpha^2 = \frac{b}{b_1}$. It follows that

$$(z + \frac{1}{z}) \circ \alpha^{1-n} z \circ z^n \circ (z + \frac{1}{z})^{(-1)},$$

is a polynomial conjugated to f. Lemma 3.10 implies $\alpha^{1-n} = \pm 1$, and the assertion follows.

If $\Phi_L(f)$ is a power or a Dickson polynomial for some linear polynomial L, then by Lemma 3.7 and Proposition 3.8 we have $\mathrm{Ord}_0 \beta_f \geq -1$.

□

Remark 4.2. If $\mathrm{Ord}_\infty f = \pi^k$, then Theorem 4.1 fails as Remark 3.5 shows. However, by Corollary 3.4, the existence of β implies that f is conjugated to z^{π^k}.

In the remaining part of this section we consider polynomials with a common iterate (compare [R1] for the case $\varepsilon = 1$).

Theorem 4.3. *Let f, g be polynomials of degree at least two, $\pi \nmid \mathrm{Ord}_\infty f, \mathrm{Ord}_\infty g$, let $\varepsilon \in \mathbb{K}$ be such that $\varepsilon^{\mathfrak{L}(f)} = \varepsilon^{\mathfrak{L}(g)} = 1$ and let L be a linear polynomial such that $\Phi_L(f)$ has gap form. If for some numbers $n, m \in \mathbb{N}$*

$$\varepsilon f^{(n)} = g^{(m)},$$

then $\mathfrak{L}(f) = \mathfrak{L}(g)$ and there exists a polynomial h, $\mathfrak{l}_\infty(h) > 1$, $\mathfrak{L}(f) | \mathfrak{L}(h)$, elements $\chi_1, \chi_2 \in \mathbb{K}$, $\chi_i^{\mathfrak{L}(h)} = 1$ and numbers $l, k \in \mathbb{N}$, such that

$$f = \Phi_{L^{-1}}(\chi_1 h^{(l)}),\ g = \Phi_{L^{-1}}(\chi_2 h^{(k)}).$$

If already $\mathfrak{l}_\infty(f) > 1$, we have in fact $\chi_i^{\mathfrak{L}(f)} = 1$.

Before we can prove Theorem 4.3, we need another lemma.

Lemma 4.4. *Let $g_1, g_2 \in \mathcal{R}_\infty$, $g \in \mathcal{R}_\infty^*$ and $n, m \in \mathbb{N}$ be given, such that $m | n$ and*

$$\mathrm{Ord}_0 (g_1 \circ z^n \circ g_2),\ \mathrm{Ord}_0 (g \circ z^m \circ g_2) \geq 0.$$

Then also $\mathrm{Ord}_0 (g_1 \circ z^{\frac{n}{m}} \circ g^{(-1)}) \geq 0$.

Proof : We have

$$(g_1 \circ z^{\frac{n}{m}} \circ g^{(-1)}) \circ (g \circ z^m \circ g_2) = g_1 \circ z^n \circ g_2,$$

and the assertion follows by Lemma 3.9.

□

Proof (of Theorem 4.3): We consider first the case that $\mathfrak{l}_\infty(f) > 1$. Note that this implies $\mathfrak{l}_\infty(g) > 1$. Assume that $\varepsilon f^{(n)} = g^{(m)}$ and let β_f and β_g be Boettcher functions of f and g. By Corollary 3.6 and Proposition 3.8 we find

$$\beta_g = \beta_f \circ \tilde{\delta} z, \tag{4.1}$$

for some $\tilde{\delta} \in \mathbb{K}$. Note that (4.1) implies that $\mathfrak{L}(\beta_f) = \mathfrak{L}(\beta_g)$. Clearly our assumption implies $(\mathrm{Ord}_\infty f)^n = (\mathrm{Ord}_\infty g)^m$, hence there exist numbers $r, s, t \in \mathbb{N}$ such that $\mathrm{Ord}_\infty f = r^s$ and $\mathrm{Ord}_\infty g = r^t$. Write $u := \gcd\{s, t\} = bt - as$ with $a, b \in \mathbb{N}$. Then

$$f^{(a)} = \beta_f \circ z^{r^{as}} \circ \beta_f^{(-1)},\ g^{(b)} = \beta_g \circ z^{r^{bt}} \circ \beta_g^{(-1)} = (\beta_f \circ \tilde{\delta}^{1-r^{bt}} z) \circ z^{r^{bt}} \circ \beta_f^{(-1)},$$

are both polynomials. By Lemma 4.4 also

$$h := (\beta_f \circ \tilde{\delta}^{1-r^{bt}} z) \circ z^{r^u} \circ \beta_f^{(-1)} \tag{4.2}$$

is a polynomial. Note that (4.2) implies $\mathfrak{l}_\infty(h) = \mathfrak{l}_\infty(\beta_f)$ and $\mathfrak{L}(\beta_f)|\mathfrak{L}(h)$.

Consider the polynomials

$$h^{(\frac{s}{u})} = (\beta_f \circ \gamma z) \circ z^{r^s} \circ \beta_f^{(-1)},\ f = \beta_f \circ z^{r^s} \circ \beta_f^{(-1)}.$$

By Lemma 4.4 also $(\beta_f \circ \gamma z) \circ z \circ \beta_f^{(-1)}$ is a polynomial, and hence must equal γz. We find that

$$f = \gamma^{-1} h^{(\frac{s}{u})},\ \gamma^{\mathfrak{L}(f)} = 1.$$

A similar argument applies to g and yields the desired result.

If $\mathfrak{l}_\infty(f) = 1$, also $\varepsilon = 1$, hence our assumption reads as $f^{(n)} = g^{(m)}$. The assertion in this case follows from the already proved applied to $\Phi_L(f)$ and $\Phi_L(g)$.

□

5 Permutable polynomials

In this section we prove a characterization of permutable polynomials (compare [R4], [T]).

Theorem 5.1. *Let $f, g \in \mathbb{K}[z]$, $\tilde{n} = \mathrm{Ord}_\infty f, \tilde{m} = \mathrm{Ord}_\infty g \geq 2$, $\pi \nmid \tilde{n}, \tilde{m}$, be given and assume that $f \circ g = g \circ f$. Then there exists a linear polynomial L, such that one of the following cases occurs:*

(i) $\Phi_L(f) = \alpha_f z^{\tilde{n}}$, $\Phi_L(g) = \alpha_g z^{\tilde{m}}$, *with* $\alpha_g^{\tilde{n}-1} = \alpha_f^{\tilde{m}-1}$.

(ii) $\Phi_L(f) = \alpha_f t_{\tilde{n}}$, $\Phi_L(g) = \alpha_g t_{\tilde{m}}$, *with* $\alpha_f^2 = \alpha_g^2 = 1$ *and* $\alpha_g^{\tilde{n}-1} = \alpha_f^{\tilde{m}-1}$.

(iii) $\Phi_L(f) = \alpha_f h^{(s)}$, $\Phi_L(g) = \alpha_g h^{(t)}$, *for some* $h \in \mathbb{K}[z]$ *with* $\alpha_f z \circ h = h \circ \alpha_f z$, $\alpha_g z \circ h = h \circ \alpha_g z$.

Remark 5.2. If we drop the assumption $\pi \nmid \tilde{n}$ and $\pi \nmid \tilde{m}$ the conclusion of Theorem 5.1 is not true as the following example shows: Let $\mathbb{K}$ be an algebraically closed field of characteristic $\pi \neq 0$ and consider the polynomial $g := z^\pi$. Obviously g permutes with a polynomial f if and only if every coefficient of f is a $(\pi - 1)$-th root of unity (compare Remark 3.5).

It is easily seen that for every $k \geq 4$ there exists a choice of f, $\mathrm{Ord}_\infty f = k$, such that f and g cannot be represented as in *(i)*, *(ii)* or *(iii)* of Theorem 5.1. In fact, choose

$$f(z) = z^k + z^{k-2} + z^{k-3} + \ldots + z + 1.$$

If f could be represented as in *(i)*, we would obtain a relation of the form $f = L_1 \circ z^k \circ L_2$ for some linear polynomials L_1, L_2, a contradiction since all coefficients

of f with exception of the $(k-1)$-th are nonzero. A similar argument shows that *(ii)* cannot occur. Now assume *(iii)*. Since π is prime, we must have $t = 1$. Hence $h = L_1 \circ z^\pi \circ L_2$ for some linear polynomials L_1, L_2 and since $\alpha_g z$ permutes with h, we conclude that $f = M_1 \circ z^{\pi^s} \circ M_2$ for some linear polynomials M_1, M_2, a contradiction as we have seen above.

To prepare the proof of Theorem 5.1 we provide another lemma:

Lemma 5.3. *Let p, q be linear polynomials, $m \geq 3$ and $r \geq 2$ be given. Assume that $\pi \nmid m, (r+m)$ and that there exist linear polynomials L_1 and L_2 such that*

$$L_1 \circ z^r p(z)^m = z^r q(z^m) \circ L_2, \tag{5.1}$$

then p and q are multiplications, i.e. $p = \varepsilon z$, $q = \delta z$.

Proof : Without loss of generality we may assume that $L_1 = z + \mu$, $L_2 = z + \nu$. Write $p = \varepsilon z + \alpha$, $q = \delta z + \gamma$, then (5.1) becomes

$$z^r(\varepsilon z + \alpha)^m + \mu = (z + \nu)^r(\delta(z + \nu)^m + \gamma). \tag{5.2}$$

Comparing the coefficients of $z^{r+m}, z^{r+m-1}, z^{r+m-2}$ in (5.2) yields, since $m \geq 3$,

$$\varepsilon^m = \delta,\ m\varepsilon^{m-1}\alpha = \delta(r+m)\nu,\ \binom{m}{2}\varepsilon^{m-2}\alpha^2 = \delta\binom{r+m}{2}\nu^2. \tag{5.3}$$

Consider first the case $\pi \neq 2$. If $\pi \nmid (m-1)$, $\pi \nmid (r+m-1)$ and $\alpha \neq 0$, the above relations imply

$$\frac{m}{r+m} = \frac{\varepsilon\nu}{\alpha} = \frac{m-1}{r+m-1},$$

a contradiction since $r \neq 0$. We conclude that $\alpha = 0$ and clearly then also $\nu = 0$. If $\pi \nmid (m-1)$ and $\pi | (r+m-1)$ ($\pi | (m-1)$ and $\pi \nmid (r+m-1)$) we obtain from the last relation in (5.3) that $\alpha = 0$ ($\nu = 0$). The second relation then implies $\nu = 0$ ($\alpha = 0$).

Now assume that $\pi | (m-1)$ and $\pi | (r+m-1)$, which is in particular the case if $\pi = 2$. Then $\pi | r$ and a comparison of the coefficients of z^1 in (5.2) yields $\delta(r+m)\nu^{r+m-1} = 0$. Hence $\nu = 0$ and thus also $\alpha = 0$.

Finally, comparing the coefficient of z^0 (z^r) in (5.2) shows that $\mu = 0$ and $\gamma = 0$.

□

Now we come to the proof of Theorem 5.1, which is done in several steps.

Choose L such that $\Phi_L(f)$ has gap form. An elementary consideration using Lemma 2.6 shows that $\mathfrak{l}_\infty(\Phi_L(f)) = \mathfrak{l}_\infty(\Phi_L(g))$. Hence we may assume throughout that $\mathfrak{l}_\infty(f) = \mathfrak{l}_\infty(g) > 1$ which will imply that we may choose $L = z$. Also, without loss of generality, let $\tilde{n} \geq \tilde{m}$.

Step 1: First we reduce the proof of Theorem 5.1 to the case that $\mathrm{Ord}_\infty f \nmid \mathrm{Ord}_\infty g$ and $\mathrm{Ord}_\infty g \nmid \mathrm{Ord}_\infty f$.

We construct inductively a (finite) sequence of polynomials f_i, g_i: Put $f_0 = f$, $g_0 = g$, and if f_i and g_i have already been determined, $\mathrm{Ord}_\infty f_i, \mathrm{Ord}_\infty g_i \neq 1$, and satisfy $\mathrm{Ord}_\infty f_i | \mathrm{Ord}_\infty g_i$ (or $\mathrm{Ord}_\infty g_i | \mathrm{Ord}_\infty f_i$, respectively), then define f_{i+1} and g_{i+1} by

$$f_{i+1} = f_i,\ g_i = g_{i+1} \circ f_i,$$

or $g_{i+1} = g_i,\ f_i = f_{i+1} \circ g_i$, respectively. This construction is possible due to Proposition 1 of [T] since we have $f_i \circ g_i = g_i \circ f_i$ for all i (Proposition 1 in [T] is the only result proved there for arbitrary characteristic). The constructed sequence is finite since in each step at least one of the degrees of f_i and g_i strictly decreases. It terminates if either at some point $\mathrm{Ord}_\infty f_i \nmid \mathrm{Ord}_\infty g_i$ and $\mathrm{Ord}_\infty g_i \nmid \mathrm{Ord}_\infty f_i$, or if one of the degrees of f_i and g_i is 1. In order to visualize this construction consider the following example:

$$\begin{array}{ccccccc} f = f_0 & \leftarrow & f_1 & = & f_2 & = & f_3 \\ & \nwarrow & & \swarrow & & \swarrow & \\ g = g_0 & = & g_1 & \leftarrow & g_2 & \leftarrow & g_3 \end{array}$$

In particular, since we assume that f and g are of gap form and π does not divide the degree of any of the occuring polynomials, all members of the constructed sequences are of gap form.

Assume first that the last polynomial (in the example g_3) is linear, i.e. of the form γz. Going backwards in the above construction, we find that f and g are of the form *(iii)* (in the example we have $f = \gamma f_3^{(3)}$, $g = \gamma f_3^{(2)}$).

Now consider the case that the last polynomials satisfy $\mathrm{Ord}_\infty f_i \nmid \mathrm{Ord}_\infty g_i$ and $\mathrm{Ord}_\infty g_i \nmid \mathrm{Ord}_\infty f_i$. Note that this assumption automatically excludes case *(iii)*. If the assertion of the theorem holds for f_i and g_i, i.e. f_i and g_i are of the form *(i)* or *(ii)*, so are f and g. This is again seen by going backwards the above construction.

During the remainder of our argumentation we assume that $\mathrm{Ord}_\infty f \nmid \mathrm{Ord}_\infty g$ and $\mathrm{Ord}_\infty g \nmid \mathrm{Ord}_\infty f$, and hence our efforts will have as their aim to show that *(i)* or *(ii)* of Theorem 5.1 holds.

Step 2: We show that if $f^{(k)}$ and $g^{(l)}$ are of the form *(i)* or *(ii)* for some $k, l \in \mathbb{N}$, then f and g also are.

Assume first that $f^{(k)} = \alpha_1 z^{\tilde{n}^k}$, $g^{(l)} = \alpha_2 z^{\tilde{m}^l}$. By *(vi)* of Lemma 2.6, we conclude that $f = \alpha_f z^{\tilde{n}}$, $g = \alpha_g z^{\tilde{m}}$. Since f commutes with g we must have $\alpha_g^{\tilde{n}-1} = \alpha_f^{\tilde{m}-1}$.

Next we assume that $f^{(k)} = \alpha_1 t_{\tilde{n}^k}$ and $g^{(l)} = \alpha_2 t_{\tilde{m}^l}$. By Proposition 3.3 and Lemma 3.7 a Boettcher function of f is given by

$$\beta = (z + \frac{1}{z}) \circ \varepsilon z,$$

for some ε. Lemma 3.10 shows that $\varepsilon^{1-\tilde{n}} = \pm 1$, hence $f = \pm t_{\tilde{n}}$. A similar argument shows that $g = \pm t_{\tilde{m}}$.

Step 3: Let $\tilde{n} = \mathrm{Ord}_\infty f = nd$, $\tilde{m} = \mathrm{Ord}_\infty g = md$, with $d = \gcd\{\tilde{n}, \tilde{m}\}$. Then m

and n are relatively prime and at least two. We show, using Step 2, that without loss of generality

$$m, n \geq 4,\ |m-n| \geq 3,\ 1 < \frac{n}{m} < 2, \tag{5.4}$$

can be assumed.

The numbers $\log \operatorname{Ord}_\infty f$ and $\log \operatorname{Ord}_\infty g$ are linearly independent over $\mathbb{Z}$, hence there exist numbers $k, l \in \mathbb{N}$, such that

$$0 < k \log \operatorname{Ord}_\infty f - l \log \operatorname{Ord}_\infty g < \log \sqrt{2}.$$

This implies $\operatorname{Ord}_\infty f^{(k)} \nmid \operatorname{Ord}_\infty g^{(l)}$, $\operatorname{Ord}_\infty g^{(l)} \nmid \operatorname{Ord}_\infty f^{(k)}$, and $1 < \frac{\operatorname{Ord}_\infty f^{(2k)}}{\operatorname{Ord}_\infty g^{(2l)}} < 2$, hence the polynomials $f^{(2k)}$ and $g^{(2l)}$ meet our requirements (5.4).

Step 4: In this step we define three sequences of polynomials f_i, g_i, h_i. This construction is analogous to that in [R4]. By Proposition 1 of [T] there exist polynomials f_0, g_0, h_0, such that

$$f = f_0 \circ h_0,\ g = g_0 \circ h_0,$$

and $\operatorname{Ord}_\infty h_0 = \gcd\{\operatorname{Ord}_\infty f, \operatorname{Ord}_\infty g\}$. It follows that $h_0 \circ f_0$ and $h_0 \circ g_0$ are permutable, and by the same argument as above we find polynomials f_1, g_1, h_1, such that

$$h_0 \circ f_0 = f_1 \circ h_1,\ h_0 \circ g_0 = g_1 \circ h_1, \tag{5.5}$$

$\operatorname{Ord}_\infty h_1 = \operatorname{Ord}_\infty h_0$. Since $f_0 \circ h_0 \circ g_0 = g_0 \circ h_0 \circ f_0$ we have $f_0 \circ g_1 = g_0 \circ f_1$. Proceeding inductively we obtain sequences f_i, g_i, h_i, which satisfy

$$h_i \circ f_i = f_{i+1} \circ h_{i+1},\ h_i \circ g_i = g_{i+1} \circ h_{i+1}, \tag{5.6}$$

$$f_i \circ h_i \circ g_i = g_i \circ h_i \circ f_i,\ f_i \circ g_{i+1} = g_i \circ f_{i+1}. \tag{5.7}$$

Since $\mathfrak{l}_\infty(f) = \mathfrak{l}_\infty(g) > 1$, the polynomials f_i, g_i, h_i may be chosen such that all of them have gap form. Then $\mathfrak{l}_\infty(h_0) = \mathfrak{l}_\infty(f) = \mathfrak{l}_\infty(g)$ and $\mathfrak{l}_\infty(h_{i+1}) = \mathfrak{l}_\infty(f_i) = \mathfrak{l}_\infty(g_i)$ for $i \geq 0$.

We claim that if at some stage $f_i = L_1 \circ z^n$, $g_i = L_2 \circ z^m$ and $h_i = \gamma z^d$, for some $\gamma \in \mathbb{K}$ and linear polynomials L_1, L_2, then f and g are of the form *(i)*. In fact, by comparing the gaps in the first equation of (5.7), L_1 and L_2 must be multiplications. By (5.6) also f_{i-1}, g_{i-1} and h_{i-1} are monomials. Proceeding inductively yields the claim.

Recall that $\operatorname{Ord}_\infty f_i = n$, $\operatorname{Ord}_\infty g_i = m$, that these numbers are relatively prime, $\pi \nmid n, m$ and that $\operatorname{Ord}_\infty h_i = d$.

Step 5: Now the results of [Z] on standard solutions are applied to the second equation of (5.7). We assume that m and n are chosen as in Step 3 and that $m < n$. By the Main Theorem of [Z], there exist for each i linear polynomials $L_{i,j}$, $j = 1, \ldots, 4$, such that one of the following cases occurs:

(i) for some polynomials p_i, q_i which are linear by our choice of m and n, we have

$$\begin{array}{l} L_{i,1} \circ f_i \circ L_{i,3}^{(-1)} = z^r p_i(z)^m \\ L_{i,3} \circ g_{i+1} \circ L_{i,2} = z^m \\ L_{i,1} \circ g_i \circ L_{i,4}^{(-1)} = z^m \\ L_{i,4} \circ f_{i+1} \circ L_{i,2} = z^r p_i(z^m) \end{array}, \tag{5.8}$$

(ii) for some $\lambda_i \in \mathbb{K} \setminus \{0\}$ and the Dickson polynomials t_n and t_m, we have

$$\begin{array}{l} L_{i,1} \circ f_i \circ L_{i,3}^{(-1)} = \lambda_i^{-mn} z \circ t_n \circ \lambda_i^m z \\ L_{i,3} \circ g_{i+1} \circ L_{i,2} = \lambda_i^{-m} z \circ t_m \circ \lambda_i z \\ L_{i,1} \circ g_i \circ L_{i,4}^{(-1)} = \lambda_i^{-mn} z \circ t_m \circ \lambda_i^n z \\ L_{i,4} \circ f_{i+1} \circ L_{i,2} = \lambda_i^{-n} z \circ t_n \circ \lambda_i z \end{array}, \tag{5.9}$$

Note that the cases - power solution or Dickson solution - of [Z] when solving $f_i \circ g_{i+1} = g_i \circ f_{i+1}$ for different values of i cannot mix. This is seen as follows: since we assume $m \geq 4$ and $\pi \nmid m$, we have $\mathfrak{L}(t_m) = 2$ and therefore t_m is not of the form $M_1 \circ z^m \circ M_2$ with linear M_1, M_2.

Step 6: Consider case *(i)* of Step 5. By our choice of n and m due to Step 3 and Lemma 5.3, we find comparing the representations of f_{i+1} in the equations (5.8) for i and $i+1$ that $p_i = \varepsilon_i z$. From the representations of f_i, g_i and g_{i+1} in (5.8) we conclude that $L_{i,2}, L_{i,3}$, $L_{i,4}$ have no constant term. Hence f_{i+1} and g_{i+1} are monomials for all $i \geq 0$. If for some $i \geq 0$ also h_{i+1} is a monomial we are done. Otherwise (5.6) shows that $\mathfrak{l}_\infty(h_{i+1}) = (r+m)\mathfrak{l}_\infty(h_i)$ for all $i \geq 1$, a contradiction.

Step 7: Consider case *(ii)* of Step 5. Again $L_{i,2}$, $L_{i,3}$ and $L_{i,4}$ have no constant term. Comparing the representations of f_1 and g_1 in the equations (5.9) for $i = 0$ and $i = 1$, we obtain

$$L_{1,1}^{(-1)} \circ \lambda_1^{-mn} z \circ t_n \circ \lambda_1^m z \circ L_{1,3} = L_{0,4}^{(-1)} \circ \lambda_0^{-n} z \circ t_n \circ \lambda_0 z \circ L_{0,2}^{(-1)},$$

$$L_{1,1}^{(-1)} \circ \lambda_1^{-mn} z \circ t_m \circ \lambda_1^n z \circ L_{1,4} = L_{0,3}^{(-1)} \circ \lambda_0^{-m} z \circ t_m \circ \lambda_0 z \circ L_{0,2}^{(-1)}.$$

Since t_n (t_m) has gap form and the gap degree $\mathfrak{L}(t_n)$ ($\mathfrak{L}(t_m)$) equals two, we conclude by Lemma 2.8 that

$$\lambda_1^m z \circ L_{1,3} \circ L_{0,2} \circ \lambda_0^{-1} z = \pm z,\ \lambda_0^n z \circ L_{0,4} \circ L_{1,1}^{(-1)} \circ \lambda_1^{-mn} z = \pm z, \tag{5.10}$$

and

$$\lambda_1^n z \circ L_{1,4} \circ L_{0,2} \circ \lambda_0^{-1} z = \pm z,\ \lambda_0^m z \circ L_{0,3} \circ L_{1,1}^{(-1)} \circ \lambda_1^{-mn} z = \pm z. \tag{5.11}$$

Comparing these relations, we find

$$\lambda_0^n z \circ L_{0,4} = \pm \lambda_0^m z \circ L_{0,3},\ \lambda_1^m z \circ L_{1,3} = \pm \lambda_1^n z \circ L_{1,4}.$$

If we put

$$f_0^* = f_0 \circ L_{0,3}^{(-1)} \circ \lambda_0^{-m} z,\ g_0^* = g_0 \circ L_{0,3}^{(-1)} \circ \lambda_0^{-m} z,\ h_0^* = \lambda_0^m z \circ L_{0,3} \circ h_0,$$

$$f_1^* = \lambda_0^m z \circ L_{0,3} \circ f_1,\ g_1^* = \lambda_0^m z \circ L_{0,3} \circ g_1,\ h_1^* = h_1 \circ L_{0,3}^{(-1)} \circ \lambda_0^{-m} z,$$

we get from (5.5)

$$h_0^* \circ f_0^* = f_1^* \circ h_1^*,\ h_0^* \circ g_0^* = g_1^* \circ h_1^*.$$

Substituting from the equations (5.9) and using (5.10) leeds to

$$\tilde{h}_0 \circ t_n = \pm(t_n \circ \tilde{h}_1),\ \tilde{h}_0 \circ t_m = \pm(t_m \circ \tilde{h}_1),$$

where $\tilde{h}_0 = h_0^* \circ L_{0,1}^{(-1)} \circ \lambda_0^{-mn} z$ and $\tilde{h}_1 = \lambda_0 z \circ L_{0,2}^{(-1)} \circ h_1^*$. Now conjugate these relations with a Boettcher function $\beta = z + \frac{1}{z}$ of t_n:

$$\pm\Phi_\beta \tilde{h}_0 \circ z^n = z^n \circ \Phi_\beta \tilde{h}_1,\ \pm\Phi_\beta \tilde{h}_0 \circ z^m = z^m \circ \Phi_\beta \tilde{h}_1.$$

If $\Phi_\beta \tilde{h}_0$ ($\Phi_\beta \tilde{h}_1$) is not of the form εz^d (δz^d), then by Corollary 2.7 both relations $n\mathfrak{l}_\infty(\Phi_\beta \tilde{h}_0) = \mathfrak{l}_\infty(\Phi_\beta \tilde{h}_1)$ and $m\mathfrak{l}_\infty(\Phi_\beta \tilde{h}_0) = \mathfrak{l}_\infty(\Phi_\beta \tilde{h}_1)$ hold, a contradiction. We arrive at the conclusion that

$$\Phi_\beta \tilde{h}_0 = \varepsilon z^d.$$

Lemma 3.10 implies that $\varepsilon = \pm 1$ and $\tilde{h}_0 = \varepsilon t_d$. Substituting this relation and the first and third relation of (5.9) into $f = f_0 \circ h_0$, $g = g_0 \circ h_0$, and using the definition of $\tilde{h}_0$ shows that f and g are of the form *(ii)* of Theorem 5.1.

All assertions of Theorem 5.1 are proved.

Acknowledgement: The authors wish to thank the referee for pointing out to their attention the papers of P.Tortrat and U.Zannier, and for suggesting the consideration of fields with nonzero characteristic.

References

[Ba] I.N.Baker: *Permutable power series and regular iteration,* J. Austral. Math. Soc. 2 (1961-62), 265-294.

[Be] A.F.Beardon: *Iteration of rational functions,* Graduate Texts in Mathematics, Springer 1991.

[B1] F.Binder: *Polynomial Decomposition,* Master Thesis, Universität Linz 1994.

[B2] F.Binder: *Characterizations of polynomial prime bidecompositions: A simplified proof,* Contributions to General Algebra 9 (1995), 61-72, Hölder-Pichler-Tempsky und B.G.Teubner, Wien-Stuttgart.

[Bo] L.Böttcher: *Beiträge zur Theorie der Iterationsrechnung (russian),* Bull. Kasan Math. Soc. 14 (1905), 176.

[DW] F.Dorey, G.Whaples: *Prime and composite polynomials,*
J. Algebra 28 (1974), 88-101.

[EN] G.Eigenthaler, W.Nöbauer: *Über die mit einem Polynom vertauschbaren linearen Polynome,*
Sb. d. Österr. Akad. d. Wiss., math.-nat. Klasse, Abt.II 196 (1990), 143-153.

[EW] G.Eigenthaler, H.Woracek: *Permutable polynomials and related topics,*
Contributions to General Algebra 9 (1995), 163-182, Hölder-Pichler-Tempsky und B.G.Teubner, Wien-Stuttgart.

[E] H.T.Engstrøm: *Polynomial substitutions,*
Amer. J. Math. 63 (1941), 249-255.

[F] M.Fried: *On a theorem of Ritt and related Diophantine problems,*
J. reine angew. Math. 264 (1973), 40-55.

[FR1] M.Fried, R.MacRae: *On the Invariance of chains of fields,*
Illinois J. Math. 13 (1969), 165-171.

[FR2] M.Fried, R.MacRae: *On curves with separated variables,*
Math. Ann. 180 (1969), 220-226.

[Ja] E.Jacobsthal: *Über vertauschbare Polynome,*
Math. Z. 63 (1955), 243-276.

[Je] S.A.Jennings: *Substitution groups of formal power series,*
Canad. J. Math. 6 (1954), 325-340.

[Jo] D.L.Johnson: *The group of formal power series under substitution,*
J. Austral. Math. Soc. 45 (1988), 296-302.

[K] H.Kautschitsch: *Kommutative Teilhalbgruppen der Kompositionshalbgruppe von Polynomen und formalen Potenzreihen,*
Monatsh. Math. 74 (1970), 421-436.

[LN] H.Lausch & W.Nöbauer: *Algebra of polynomials,*
North-Holland, Amsterdam 1973.

[L] H.Levi: *Composite polynomials with coefficients in an arbitrary field of characteristic zero,*
Amer. J. Math. 64 (1942), 389-400.

[R1] J.F.Ritt: *On the iteration of rational functions,*
Trans. Amer. Math. Soc. 21 (1920), 348-356.

[R2] J.F.Ritt: *Periodic functions with a multiplication theorem,*
Trans. Amer. Math. Soc. 23 (1922), 16-25.

[R3] J.F.Ritt: *Prime and composite polynomials,*
Trans. Amer. Math. Soc. 23 (1922), 51-66.

[R4] J.F.Ritt: *Permutable rational functions,*
Trans. Amer. Math. Soc. 25 (1923), 399-448.

[S] A.Schinzel: *Selected topics on polynomials,*
University of Michigan Press, Ann Arbor 1982.

[T] P.Tortrat: *Sur la composition des polynomes,*
Colloq. Math. 55 (1988), 329-353.

[Z] U.Zannier: *Ritt's second theorem in arbitrary characteristic,*
J. reine angew. Math. 445 (1993), 175-203.

Gerhard Dorfer
Institut für Algebra und
Diskrete Mathematik
Technische Universität Wien
Wiedner Hauptstraße 8-10/118.3
A-1040 Wien
AUSTRIA
email: g.dorfer@tuwien.ac.at

Harald Woracek
Institut für Analysis und Technische
Mathematik
Technische Universität Wien
Wiedner Hauptstraße 8-10/114.1
A-1040 Wien
AUSTRIA
email: hworacek@pop.tuwien.ac.at

AMS Classification numbers: 12E05, 13F25