CANONICAL REPRESENTATIONS OF HYPERGEOMETRIC TERMS

S. A. ABRAMOV AND M. PETKOVŠEK

Abstract. We describe a multiplicative normal form for rational functions which exhibits the shift structure of the factors. Using this normal form we construct a minimal representation of hypergeometric terms in one variable. Our main result concerns hypergeometric terms in two variables: every such term is the product of a Z-term and a rational sequence.

Résumé. Nous décrivons une forme normale multiplicative pour les fractions rationnelles montrant la structure décalage des facteurs. Cette forme nous permet de construire une représentation minimale de termes hypergéométriques en une variable. Notre résultat principal concerne les termes hypergéométriques en deux variables: tous tels termes sont obtenus comme produits d'un Z-terme avec une fraction rationelle.

1. Introduction

A d-dimensional sequence $T(n_1, n_2, \ldots, n_d)$ is a *hypergeometric term* if all the quotients $T(n_1, \ldots, n_i + 1, \ldots, n_d)/T(n_1, \ldots, n_d)$ are rational functions of $n_1, n_2, \ldots, n_d$. The ubiquity of hypergeometric terms in combinatorial enumeration is widely recognized. While for $d = 1$ hypergeometric terms have simple structure and are well understood, this is not the case when $d \geq 2$. Especially important among hypergeometric terms are the holonomic ones because they can be summed by Zeilberger's Creative Telescoping algorithm [11]. A well-known class of holonomic hypergeometric terms are the proper terms which can be expressed as a product of a polynomial with several factorials of linear forms with integer coefficients, their reciprocals, and exponential functions. In this paper we give a structure theorem for hypergeometric terms which shows that every hypergeometric term can be expressed similarly if the polynomial factor is replaced by a rational function. This opens the way to a proof, to be given in a subsequent paper [3], of a conjecture of Wilf and Zeilberger [10, p. 585] which states that every holonomic hypergeometric term is proper.

In Section 2 we introduce a normal form for rational functions which reveals the shift structure of its factors, and is symmetric with respect to its numerator and denominator. In the rest of the paper we use this form to construct canonical representations of univariate and bivariate hypergeometric terms. Section 3 is devoted to minimal representation of univariate hypergeometric terms as products with rational factor whose numerator and denominator are simultaneously of the lowest possible degrees. In Section 4 we prepare another tool, a sufficient condition for a bivariate polynomial to factor into integer-linear factors. In Section 5 we prove our main result: every bivariate hypergeometric term is the product of a Z-term and a rational sequence. By a Z-term we mean a hypergeometric term whose defining rational functions factor into integer-linear factors.

Throughout the paper, K is a field of characteristic zero, and $\mathbb{N}$ denotes the set of nonnegative integers. We write $p \perp q$ to indicate that polynomials $p, q \in K[x, y]$ are relatively prime, and $p \sim q$ if there is a $u \in K \setminus \{0\}$ such that $p(x, y) = u\, q(x, y)$. The *leading coefficient* of a rational function is the quotient of the leading coefficients of its numerator and denominator. A rational function is *monic* if its leading coefficient is 1.

Partially supported by the French-Russian Lyapunov Institute under grant 98-03, and by MZT RS under grant J2-8549.

2. Normal forms for rational functions

In this section, E denotes the automorphism of $K(x)$ defined by $Ef(x) = f(x+1)$. Following [8] we introduce the notion of shift-equivalence among polynomials.

Definition 1. *Irreducible polynomials $p, q \in K[x]$ are* shift-equivalent *if $p = E^k q$ for some $k \in \mathbb{Z}$. In this case we write $p \overset{\text{sh}}{\sim} q$. A rational function is* shift-homogeneous *if all nonconstant irreducible monic factors of its numerator and denominator belong to the same shift-equivalence class.*

It is clear that by grouping together shift-equivalent irreducible monic factors of its numerator and denominator every rational function can be written in the form

$$R(x) = z\, R_1(x) R_2(x) \cdots R_k(x) \tag{1}$$

where $z \in K$, $k \geq 0$, each R_i is a monic shift-homogeneous rational function, and any two nonconstant monic irreducible factors p of R_i and q of R_j are pairwise shift-inequivalent whenever $i \neq j$. We call (1) a *shift-homogeneous factorization* of R.

Lemma 1. *Let $R(x) = z\, R_1(x)R_2(x)\cdots R_k(x) = w\, S_1(x)S_2(x)\cdots S_k(x)$ be two shift-homogeneous factorizations of R such that the nonconstant monic irreducible factors of R_i and S_i belong to the same shift-equivalence class. Then $z = w$ and $R_i = S_i$ for all i.*

Proof. Clearly $z = w$ because they are both equal to the leading coefficient of R. Therefore

$$\frac{R_i(x)}{S_i(x)} = \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{S_j(x)}{R_j(x)} \quad (i = 1, 2, \ldots, k).$$

As every nonconstant irreducible monic factor of the left-hand fraction is shift-inequivalent with every such factor of the right-hand fraction, $R_i(x)/S_i(x) = 1$, and $R_i = S_i$, for all i. ☐

Multiplicative normal forms of rational functions which exhibit the shift structure of the factors are useful in design of algorithms for summation and solution of difference equations in closed form. The following well-known form is used in algorithms for hypergeometric summation [4], finding hypergeometric solutions of difference equations [6], and rational summation [8].

Definition 2. *Let $R \in K(x) \setminus \{0\}$ be a rational function. If $z \in K$ and monic polynomials $a, b, c \in K[x]$ satisfy*

- (i) $R = z \cdot \frac{a}{b} \cdot \frac{Ec}{c}$,
- (ii) $a \perp E^k b$ *for all integer* $k \geq 0$,
- (iii) $a \perp c$ *and* $b \perp Ec$,

then (z, a, b, c) is the polynomial normal form (PNF) *of R.*

Definition 3. *Let $R \in K(x) \setminus \{0\}$ be a rational function. If $z \in K$ and monic polynomials $r, s, u, v \in K[x]$ satisfy*

- (i) $R = z \cdot \frac{r}{s} \cdot \frac{EF}{F}$ *where* $F = \frac{u}{v}$ *and* $u \perp v$,
- (ii) $r \perp E^k s$ *for all* $k \in \mathbb{Z}$,

then (z, r, s, u, v) is a rational normal form (RNF) *of R. If, in addition,*

- (iii) $r \perp u \cdot Ev$ *and* $s \perp Eu \cdot v$,

then (z, r, s, u, v) is a strict RNF *of R.*

While PNF is unique [6, Cor. 3.1], a rational function can have several different RNFs, even strict ones.

Example 1. *Let $R(x) = x(x+2)/((x-1)(x+1)^2(x+3))$. Then we can write*

$$\begin{aligned}
R(x) &= \frac{1}{(x-1)(x+3)} \cdot \frac{x(x+2)}{(x+1)^2} = \frac{1}{(x-1)(x+3)} \cdot \frac{EF}{F}, \quad F = \frac{x+1}{x}, \\
R(x) &= \frac{1}{(x+1)^2} \cdot \frac{x(x+2)}{(x-1)(x+3)} = \frac{1}{(x+1)^2} \cdot \frac{EF}{F}, \quad F = \frac{x-1}{x+2}, \\
R(x) &= \frac{1}{(x+1)(x+3)} \cdot \frac{x(x+2)}{(x-1)(x+1)} = \frac{1}{(x+1)(x+3)} \cdot \frac{EF}{F}, \quad F = (x-1)(x+1), \\
R(x) &= \frac{1}{(x-1)(x+1)} \cdot \frac{x(x+2)}{(x+1)(x+3)} = \frac{1}{(x-1)(x+1)} \cdot \frac{EF}{F}, \quad F = \frac{1}{x(x+2)},
\end{aligned}$$

so $R(x)$ has four different strict RNFs.

Proposition 1. *Let $\varphi = (z, r, s, u, v)$ be an RNF of $R = p/q$ where $p, q \in K[x]$.*

(i) *If φ is strict then $r \mid p$ and $s \mid q$.*
(ii) *$\varphi^{-1} = (1/z, s, r, v, u)$ is an RNF of $1/R$. If φ is strict then so is φ^{-1}.*
(iii) *If (z, r, s, u', v') is another RNF of R then $u' = u$ and $v' = v$ (i.e., RNF is uniquely determined by r and s).*

Proof. (i) As $p\, s\, u\, Ev = z\, q\, r\, v\, Eu$, and $r \perp s\, u\, Ev$, it follows that $r \mid p$. Similarly we see that $s \mid q$.

(ii) Clearly,

$$\frac{1}{z} \cdot \frac{s}{r} \cdot \frac{E(v/u)}{v/u} = \left(z \cdot \frac{r}{s} \cdot \frac{E(u/v)}{u/v} \right)^{-1} = \frac{1}{R}.$$

Properties (ii) and (iii) of RNF are invariant on exchanging r with s and u with v, so they remain satisfied for the new form.

(iii) Write $F' = u'/v'$. Then

$$u\, Ev\, EF' - v\, Eu\, F' = 0,$$

which is a first-order homogeneous linear recurrence with polynomial coefficients. Its general solution is $F' = Cu/v$ where C is an arbitrary constant. As u, v, u', v' are monic, $u \perp v$, and $u' \perp v'$, the equation $u'/v' = Cu/v$ implies that $u = u'$ and $v = v'$. □

Does every rational function have an RNF, and how do we find one? It turns out that any algorithm for computing the PNF (see [6], also [8] and [7]) can be used to construct an RNF by the following lemma.

Lemma 2. *Let (z, a, b, c) be the PNF of R, $(1, a_1, b_1, c_1)$ the PNF of b/a, and $c/c_1 = d/d_1$ where $d, d_1 \in K[x]$ are monic and $d \perp d_1$. Then (z, b_1, a_1, d, d_1) is a strict RNF of R. In particular, every nonzero rational function has a strict RNF.*

Proof. As

$$R = z \cdot \frac{a}{b} \cdot \frac{Ec}{c} \quad \text{and} \quad \frac{b}{a} = \frac{a_1}{b_1} \cdot \frac{Ec_1}{c_1},$$

it follows that

$$R = z \cdot \frac{b_1}{a_1} \cdot \frac{E(c/c_1)}{c/c_1} = z \cdot \frac{b_1}{a_1} \cdot \frac{E(d/d_1)}{d/d_1},$$

proving property (i) of RNF. Because $a_1 \perp E^k b_1$ for $k \geq 0$, we have $b_1 \perp E^k a_1$ for $k \leq 0$. From $a_1 \mid b$, $b_1 \mid a$ and $a \perp E^k b$ for $k \geq 0$, it follows that $b_1 \perp E^k a_1$ for $k \geq 0$ as well, proving property (ii). To prove (iii), note that $d \mid c$ and $d_1 \mid c_1$. Because $(1, a_1, b_1, c_1)$ is a PNF, we have $a_1 \perp d_1$ and $b_1 \perp Ed_1$. Because $a_1 \mid b$ and $b_1 \mid a$ and (z, a, b, c) is a PNF, we have $b_1 \perp d$ and $a_1 \perp Ed$. Finally, $d \perp d_1$ by construction. □

Example 2. *Take $R(x) = (x^2-1)/(x^2+2x)$. As*
$$R(x) = \frac{x-1}{x+2} \cdot \frac{x+1}{x},$$
we have $z=1$, $a=x-1$, $b=x+2$, $c=x$. Next,
$$\frac{x+2}{x-1} = \frac{x(x+1)(x+2)}{(x-1)x(x+1)},$$
so $a_1=b_1=1$, $c_1=x(x^2-1)$, $d=1$, and $d_1=x^2-1$. Thus $(1,1,1,1,x^2-1)$ is an RNF for $R(x)$. Incidentally, we have discovered that $R=EF/F$ where $F(x)=1/(x^2-1)$ (cf. [6, Lemma 5.1]*).*

Even though RNF is not unique, the RNFs representing the same rational function are closely related to each other. To describe this relationship, we use localization to shift-equivalence classes.

Lemma 3. *Let $R \in K(x)$ be shift-homogeneous. If $(1,r,s,u,v)$ and $(1,r_1,s_1,u_1,v_1)$ are two RNFs of R then $r=r_1=1$ and $\deg s = \deg s_1$, or $s=s_1=1$ and $\deg r = \deg r_1$.*

Proof. From
$$\frac{r}{s} \cdot \frac{E(u/v)}{u/v} = \frac{r_1}{s_1} \cdot \frac{E(u_1/v_1)}{u_1/v_1}$$
it follows that $r\, s_1\, Eu\, u_1\, v\, Ev_1 = r_1\, s\, u\, Eu_1\, Ev\, v_1$, so $\deg r - \deg r_1 = \deg s - \deg s_1$. Because R is shift-homogeneous, property (ii) of the RNF implies that $r=1$ or $s=1$. In the same way, $r_1=1$ or $s_1=1$. We distinguish four cases: If $r=r_1=1$ it follows that $\deg s = \deg s_1$. If $s=s_1=1$ it follows that $\deg r = \deg r_1$. If $r=s_1=1$, then $\deg s + \deg r_1 = 0$, so $s=r_1=1$. If $r_1=s=1$, then $\deg r + \deg s_1 = 0$, so $r=s_1=1$. In all four cases, the assertion is true. □

Theorem 1. *Let (z,r,s,u,v) and (z',r',s',u',v') be two RNFs of $R \in K(x)$. Then*
- (i) *$z=z'$,*
- (ii) *$\deg r = \deg r'$ and $\deg s = \deg s'$,*
- (iii) *there is a one-to-one correspondence f between the multisets of nonconstant irreducible monic factors of r and r' such that $p \overset{\mathrm{sh}}{\sim} f(p)$ for all $p \mid r$,*
- (iv) *there is a one-to-one correspondence g between the multisets of nonconstant irreducible monic factors of s and s' such that $q \overset{\mathrm{sh}}{\sim} g(q)$ for all $q \mid s$.*

Proof. Obviously $z=z'$ because they are both equal to the leading coefficient of R. Let $r = r_1 r_2 \cdots r_k$, $s = s_1 s_2 \cdots s_k$, $u = u_1 u_2 \cdots u_k$, $v = v_1 v_2 \cdots v_k$, and likewise for r', s', u', v', be shift-homogeneous factorizations where nonconstant irreducible monic factors of polynomials with the same subscript belong to the same shift-equivalence class. For $i=1,2,\ldots,k$ write
$$R_i = \frac{r_i}{s_i} \cdot \frac{E(u_i/v_i)}{u_i/v_i}, \quad R_i' = \frac{r_i'}{s_i'} \cdot \frac{E(u_i'/v_i')}{u_i'/v_i'}.$$
Then, clearly, $(1,r_i,s_i,u_i,v_i)$ is an RNF of R_i, and $(1,r_i',s_i',u_i',v_i')$ is an RNF of R_i'. As $R = R_1 R_2 \cdots R_k = R_1' R_2' \cdots R_k'$, Lemma 1 implies that $R_i = R_i'$, for all i. By Lemma 3, $\deg r_i = \deg r_i'$ and $\deg s_i = \deg s_i'$. It follows that $\deg r = \deg r'$ and $\deg s = \deg s'$. To obtain the desired correspondences f resp. g, let the nonconstant irreducible monic factors of r_i (resp. s_i) correspond to the nonconstant irreducible monic factors of r_i' (resp. s_i'). □

Corollary 1. *Let (z,r,s,u,v) be an RNF of $R \in K(x)$. If*
$$R(x) = \frac{p(x)}{q(x)} \frac{EF(x)}{F(x)}$$
where $p,q \in K[x]$ and $F \in K(x)$ then $\deg r \le \deg p$ and $\deg s \le \deg q$.

Proof. Write $F = g/h$ where $g,h \in K[x]$ are relatively prime. Let (z',r',s',u',v') be a strict RNF of p/q. Then $(z',r',s',u'g,v'h)$ is an RNF of R. Hence by Theorem 1, $\deg r = \deg r'$ and $\deg s = \deg s'$. By Proposition 1 (i), $r' \mid p$ and $s' \mid q$. Consequently $\deg r \le \deg p$ and $\deg s \le \deg q$. □

3. Hypergeometric terms in one variable

Definition 4. *A sequence $h : \mathbb{N} \to K$ is a (univariate)* hypergeometric term *if there are polynomials $p, q \in K[x] \setminus \{0\}$ such that*

$$q(n)h_{n+1} = p(n)h(n)$$

for all $n \in \mathbb{N}$.

To obtain an explicit representation of a hypergeometric term h_n, write $R = p/q \in K(x)$, and let $n_0 \in \mathbb{N}$ be such that $p(n)q(n) \neq 0$ whenever $n \geq n_0$. Then we have $h_{n+1} = R(n)h(n)$ for $n \geq n_0$, and so

$$h_n = h_{n_0} \prod_{k=n_0}^{n-1} R(k) \qquad (n \geq n_0).$$

Let (z, r, s, u, v) be an RNF of R, $F = zr/s$ and $G = u/v$. Then

$$h_n = h_{n_0} \prod_{k=n_0}^{n-1} \left(F(k) \cdot \frac{G(k+1)}{G(k)} \right) = h_{n_0} \frac{G(n)}{G(n_0)} \prod_{k=n_0}^{n-1} F(k),$$

and we claim that the numerator and denominator of F are *simultaneously of the minimal possible degrees* among all representations of h_n in the form

$$h_n = U(n) \prod_{k=n_0}^{n-1} V(k) \tag{2}$$

where $U, V \in K(x)$ are rational functions. Indeed, from (2) we have

$$R(n) = \frac{h_{n+1}}{h_n} = V(n) \frac{U(n+1)}{U(n)},$$

so it follows from Corollary 1 that the degrees of the numerator and denominator of F do not exceed those of the numerator and denominator of V.

Example 3. *Consider the hypergeometric term h_n defined by*

$$\begin{aligned} h_0 &= 1, \\ \frac{h_{n+1}}{h_n} &= \frac{(4n^2 + 13n + 3)(6n^2 + 17n + 5)}{(n^2 + 5n + 4)(6n^2 + 11n + 4)} \qquad (n \geq 0). \end{aligned}$$

We can express this hypergeometric term explicitly as

$$h_n = \prod_{k=0}^{n-1} \frac{(4k^2 + 13k + 3)(6k^2 + 17k + 5)}{(k^2 + 5k + 4)(6k^2 + 11k + 4)}.$$

An RNF *of h_{n+1}/h_n is*

$$\left(4, n + \frac{1}{4}, n + 4, (n+1)(n+2)\left(n + \frac{1}{2}\right)\left(n + \frac{3}{2}\right), n + \frac{1}{3}\right),$$

so we can also write

$$h_n = \frac{(n+1)(n+2)(2n+1)(2n+3)}{6(3n+1)} \prod_{k=0}^{n-1} \frac{4k+1}{k+4}$$

where the factor in the product has numerators and denominators of minimal possible degrees.

4. Products of integer-linear polynomials

Lemma 4. *Let $f \in K(x)$, $a \in K$, $a \neq 0$. If $f(x+a) = f(x)$ then $f(x) = c \in K$.*

Proof. Write $f(x) = p(x)/q(x)$ where $p, q \in K[x]$. Let $x_0 \in K$ be such that $q(x_0 + ka) \neq 0$ for all $k \in \mathbb{N}$, and $c = f(x_0)$. Then $r(x) = p(x) - c\,q(x) \in K[x]$ vanishes on $\{x_0 + ka;\ k \in \mathbb{N}\}$. In characteristic zero this is an infinite set, hence r is the zero polynomial, and $f(x) = c$ as claimed. □

Lemma 5. *Let $P \in K[x,y]$, $a, b \in K$, $a \neq 0$. If $P(x+a, y+b) = P(x,y)$ then there is a univariate polynomial $f \in K[x]$ such that $P(x,y) = f(bx - ay)$.*

Proof. By induction on $\deg_y P$.

- $\deg_y P = 0$: Write $P(x,y) = p(x)$. Then $p(x+a) = p(x)$, so $p(x) = c \in K$ by Lemma 4. Take $f(x) = c$.
- $\deg_y P > 0$: By induction on k it follows that $P(x+ka, y+kb) = P(x,y)$ for all $k \in \mathbb{N}$. In particular, $P(ka, kb) = P(0,0)$ for all $k \in \mathbb{N}$. Denote $w = P(0,0)$. Dividing $P(x,y) - w$ by $bx - ay$ in $K[x,y]$ (w.r.t. lexicographic order on monomials with $x \prec y$), we have

$$P(x,y) - w = Q(x,y)(bx - ay) + R(x) \tag{3}$$

  where $Q \in K[x,y]$ and $R \in K[x]$. Also, $\deg_y Q = \deg_y P - 1$. Set $x = ka$ and $y = kb$ in (3) to see that $R(ka) = 0$ for all $k \in \mathbb{N}$, hence $R(x) = 0$ and $P(x,y) - w = Q(x,y)(bx - ay)$. Substituting $x + a$ for x and $y + b$ for y yields $P(x,y) - w = Q(x+a, y+b)(bx - ay)$, so $Q(x,y)(bx-ay) = Q(x+a, y+b)(bx-ay)$. As $K[x,y]$ is an integral domain, $Q(x+a, y+b) = Q(x,y)$. By the induction hypothesis there is a $g \in K[x]$ such that $Q(x,y) = g(bx - ay)$. We conclude that $P(x,y) = f(bx - ay)$ where $f(x) = xg(x) + w$.

□

Corollary 2. *Assume that K is algebraically closed. Let $P \in K[x,y] \setminus K$ be irreducible, $a, b \in K$, $a \neq 0$. If $P(x+a, y+b) = P(x,y)$ then there are $u \in K \setminus \{0\}$, $v \in K$ such that $P(x,y) = u(bx - ay + v)$.*

Proof. By Lemma 5, there is an $f \in K[x]$ such that $P(x,y) = f(bx - ay)$. As P is irreducible, so is f, hence $\deg f \leq 1$. Thus there are $u, w \in K$ such that $f(x) = ux + w$ and, consequently, $P(x,y) = u(bx - ay) + w$. As P is nonconstant, we have $u \neq 0$. Taking $v = w/u$ proves the claim. □

Definition 5. *A polynomial $p \in K[x,y]$ is* integer-linear *if $p(x,y) = ax + by + c$ where $a, b \in \mathbb{Z}$ and $c \in K$.*

Theorem 2. *Let K be algebraically closed, and $P \in K[x,y]$. If for every irreducible factor p of P there are $a, b \in \mathbb{Z}$, $a > 0$, such that $p(x+a, y+b)$ also divides $P(x,y)$ then P factors into integer-linear factors.*

Proof. By induction on the total degree of P.

- $\deg P = 0$: If P is constant then it is integer-linear.
- $\deg P > 0$: Pick a nonconstant irreducible factor p_0 of P. Construct a sequence of nonconstant irreducible factors p_i of P such that $p_{i+1}(x,y) = p_i(x + a_i, y + b_i)$ where $a_i, b_i \in \mathbb{Z}$ and $a_i > 0$, for $i \geq 0$. As $K[x,y]$ is a unique factorization domain, there are indices $i_0 < j_0$ such that $p_{i_0} = c\,p_{j_0}$ for some $c \in K \setminus \{0\}$. By definition of p_i, it follows that

$$p_{i_0}(x,y) = c\,p_{i_0}(x + A, y + B) \tag{4}$$

  where $A = a_{i_0} + a_{i_0+1} + \cdots + a_{j_0-1} > 0$ and $B = b_{i_0} + b_{i_0+1} + \cdots + b_{j_0-1}$ are integers. Order monomials lexicographically and compare the leading coefficients in (4) to determine that $c = 1$. By Lemma 2, there are $u, v \in K$, $u \neq 0$, such that $p_{i_0}(x,y) = u(Bx - Ay + v)$. Let

$$P(x,y) = q_1(x,y)q_2(x,y)\cdots q_r(x,y)$$

  be a factorization of P into irreducibles, and $S = \{i;\ q_i(x,y) \sim Bx - Ay + w,\ \text{for some } w \in K\}$. Clearly $i_0 \in S$. Write $P = QR$ where $Q(x,y) = \prod_{i \in S} q_i(x,y)$. Then Q factors into integer-linear factors and $Q \perp R$. Let p be an irreducible factor of R. Then there are $a, b \in \mathbb{Z}$, $a > 0$, such that $p(x+a, y+b)$ also divides $P(x, y)$. If $p(x+a, y+b)$ divides $Q(x, y)$ then $p(x+a, y+b) \sim Bx - Ay + w$, for some $w \in K$, hence $p(x, y) \sim Bx - Ay + (w - aB + Ab)$. But then $p(x, y)$ divides $Q(x, y)$ which is impossible. Thus $p(x+a, y+b)$ must divide $R(x, y)$. Also, $\deg R < \deg P$. By induction hypothesis, R factors into integer-linear factors, hence so does $P = QR$.

□

A different proof using algebraic functions is given in [2, Lemma 3].

5. A structure theorem for hypergeometric terms in two variables

In this section, a *sequence* $T(n, k)$ is a function $T : \mathbb{N} \times \mathbb{N} \to K$. We call a set $A \subseteq \mathbb{N} \times \mathbb{N}$ *algebraic* if there is a polynomial $p \in K[x, y] \setminus \{0\}$ which vanishes on A. Clearly, if A is algebraic and B is not, then $B \setminus A$ is not algebraic. Also, a finite union of algebraic sets is algebraic.

Proposition 2. *Let $F, G \in K(x, y)$ be rational functions which agree on a non-algebraic set $B \subseteq \mathbb{N} \times \mathbb{N}$. Then $F = G$.*

Proof. Let $F = p/q$, $G = u/v$, where $p, q, u, v \in K[x, y]$. The polynomial $pv - qu$ vanishes on the non-algebraic set B, hence it is the zero polynomial, and so $F = G$. □

Definition 6. *We write $T =_a T'$ if the sequences $T(n, k)$ and $T'(n, k)$ agree outside some algebraic set. A sequence $T(n, k)$ is* trivial *if $T =_a 0$.* (equality modulo an algebraic set)

Equality modulo an algebraic set is clearly an equivalence relation. It is also a congruence because $T_1 =_a T_2$ and $T_1' =_a T_2'$ imply $T_1 + T_1' =_a T_2 + T_2'$ and $T_1 T_1' =_a T_2 T_2'$. Trivial sequences can be described as those with algebraic support. Note however that a nontrivial sequence can vanish on a non-algebraic set.

Example 4. The sequence $T(n, k) = \binom{n}{k}$ is nontrivial because $\operatorname{supp} T = \{(n, k) \in \mathbb{N} \times \mathbb{N};\ n \geq k\}$ is not algebraic. But neither is its complement $\{(n, k) \in \mathbb{N} \times \mathbb{N};\ n < k\}$.

Definition 7. *A sequence $T(n, k)$ is a (bivariate)* hypergeometric term *if there are polynomials $f_0, f_1, g_0, g_1 \in K[x, y] \setminus \{0\}$ such that*

$$f_1(n, k) T(n+1, k) = f_0(n, k) T(n, k), \tag{5}$$

$$g_1(n, k) T(n, k+1) = g_0(n, k) T(n, k) \tag{6}$$

for all $n, k \in \mathbb{N}$.

Clearly the product of two hypergeometric terms is again a hypergeometric term.

Proposition 3. *If T is a hypergeometric term and $T' =_a T$ then T' is a hypergeometric term.*

Proof. Let T satisfy (5), (6) and let $p(n, k) T'(n, k) = p(n, k) T(n, k)$ for all $n, k \in \mathbb{N}$. Then

$$p(n, k) p(n+1, k) f_1(n, k) T'(n+1, k) = p(n, k) p(n+1, k) f_0(n, k) T'(n, k), \tag{7}$$

$$p(n, k) p(n, k+1) g_1(n, k) T'(n, k+1) = p(n, k) p(n, k+1) g_0(n, k) T'(n, k) \tag{8}$$

for all $n, k \in \mathbb{N}$, hence T' is a hypergeometric term. □

Identification of bivariate sequences which agree outside an algebraic set is consistent with identification of univariate sequences which agree outside a finite set (cf. [9]). Such identification enables us to regard every rational function $R \in K(x, y)$ as a sequence $R(n, k)$, without actually having to specify its values at the singular points of R. Therefore, if T is a hypergeometric term satisfying (5), (6), we can write

$$T(n+1, k) =_a F(n, k) T(n, k), \quad T(n, k+1) =_a G(n, k) T(n, k), \tag{9}$$

where $F = f_0/f_1$ and $G = g_0/g_1$.

Definition 8. *Two rational functions $F, G \in K(x,y)$ are* compatible *if they satisfy*

$$F(x,y)\,G(x+1,y) = F(x,y+1)\,G(x,y). \tag{10}$$

Proposition 4. *Let $T(n,k)$ be a hypergeometric term which satisfies* (9). *If $T \neq_a 0$ then*

(i) *F and G are compatible,*
(ii) *F and G are unique.*

Proof. (i) By (9) we have

$$\begin{aligned} T(n+1,k+1) &= F(n,k+1)T(n,k+1) = F(n,k+1)G(n,k)T(n,k) \\ &= G(n+1,k)T(n+1,k) = G(n+1,k)F(n,k)T(n,k) \end{aligned}$$

for (n,k) outside some algebraic set A. Hence $F(n,k+1)G(n,k) = G(n+1,k)F(n,k)$ on $\operatorname{supp} T \setminus A$. As this is a non-algebraic set, Proposition 2 implies that $F(x,y+1)G(x,y) = G(x+1,y)F(x,y)$.

(ii) Assume that in addition to (9), $T(n+1,k) =_a F_1(n,k)T(n,k)$ and $T(n,k+1) =_a G_1(n,k)T(n,k)$. Then $F(n,k) = F_1(n,k)$ and $G(n,k) = G_1(n,k)$ on $\operatorname{supp} T \setminus A$, for some algebraic set A. By Proposition 2, $F = F_1$ and $G = G_1$. □

The following property of divisibility in $K[x,y]$ will be used without mention.

Proposition 5. *Let $p, q \in K[x,y]$, p irreducible. Then $p \mid q$ in $K[x,y]$ if and only if $p \mid q$ in $K(x)[y]$.*

Proof. Let $p \in K[x,y]$ be irreducible. The only units of $K[x,y]$ are the nonzero constants, hence p is primitive when considered as an element of $K[x][y]$. Thus divisibility by p is not affected when the coefficient ring $K[x]$ is replaced by its field of fractions $K(x)$. □

Lemma 6. (cf. [1]) *Let $a, b, u, v \in K[x] \setminus \{0\}$, $u \perp v$, $r = u/v$, p an irreducible factor of v, and*

$$a(x)r(x+1) = b(x)r(x). \tag{11}$$

Then there are $m, n \in \mathbb{N}$, $m \geq 1$, $n \geq 0$, such that $p(x+m)$ divides $a(x)$ and $p(x-n)$ divides $b(x)$.

Proof. Rewrite (11) as

$$a(x)u(x+1)v(x) = b(x)u(x)v(x+1). \tag{12}$$

Let $m \in \mathbb{N}$, $m \geq 1$, be such that $p(x+m-1)$ divides $v(x)$ but $p(x+m)$ does not. Then (12) implies that $p(x+m) \mid a(x)u(x+1)v(x)$. As $p(x+m) \perp u(x+1)v(x)$, it follows that $p(x+m) \mid a(x)$.

Let $n \in \mathbb{N}$, $n \geq 0$, be such that $p(x-n)$ divides $v(x)$ but $p(x-n-1)$ does not. Then (12) implies that $p(x-n) \mid b(x)u(x)v(x+1)$. As $p(x-n) \perp u(x)v(x+1)$, it follows that $p(x-n) \mid b(x)$. □

Theorem 3. *Let $F, G \in K(x,y)$ be compatible rational functions over an algebraically closed field K. Then there are compatible rational functions $F', G' \in K(x,y)$ which factor into integer-linear factors, and a rational function $R \in K(x,y)$ such that*

$$F(x,y) = F'(x,y)\frac{R(x+1,y)}{R(x,y)}, \tag{13}$$

$$G(x,y) = G'(x,y)\frac{R(x,y+1)}{R(x,y)}. \tag{14}$$

Proof. Take an RNF (G', R) of $G(x,y)$, considered as a rational function of y over $K(x)$. Then $G'(x,y)$ is shift-reduced w.r.t. y and satisfies (14). Define $F'(x,y)$ so that it satisfies (13). The compatibility condition (10) for F, G implies that

$$F'(x,y)\,G'(x+1,y) = F'(x,y+1)\,G'(x,y), \tag{15}$$

so F', G' are compatible. It remains to show that they factor into integer-linear factors. Write

$$F'(x,y) = \frac{s(x,y)}{t(x,y)}, \quad G'(x,y) = \frac{u(x,y)}{v(x,y)} \tag{16}$$

where $s, t, u, v \in K[x, y]$, $s(x, y) \perp t(x, y)$, and $u(x, y) \perp v(x, y + m)$ for all $m \in \mathbb{Z}$. We use two lemmas.

Lemma 7. *Let F', G', s, t, u, v be as in (15), (16). If $p \in K[x, y]$ is an irreducible factor of uv then there are $A, B \in \mathbb{Z}$, $A > 0$, such that $p(x + A, y + B)$ divides st.*

Proof. a) If $p \mid v$ rewrite (15) as

$$s(x, y)t(x, y + 1)G'(x + 1, y) = s(x, y + 1)t(x, y)G'(x, y).$$

By Lemma 6, there is $m \in \mathbb{Z}$, $m \geq 1$, such that

$$p(x + m, y) \mid s(x, y)t(x, y + 1).$$

Then $p(x + m, y) \mid s(x, y)$ or $p(x + m, y - 1) \mid t(x, y)$. Take $(A, B) = (m, 0)$ in the former case, $(A, B) = (m, -1)$ in the latter.

b) If $p \mid u$ rewrite (15) as

$$s(x, y + 1)t(x, y)\frac{1}{G'(x + 1, y)} = s(x, y)t(x, y + 1)\frac{1}{G'(x, y)}.$$

By Lemma 6, there is $m \in \mathbb{Z}$, $m \geq 1$, such that

$$p(x + m, y) \mid s(x, y + 1)t(x, y).$$

Then $p(x + m, y - 1) \mid s(x, y)$ or $p(x + m, y) \mid t(x, y)$. Take $(A, B) = (m, -1)$ in the former case, $(A, B) = (m, 0)$ in the latter. □

Lemma 8. *Let F', G', s, t, u, v be as in (15), (16) where $G'(x, y)$ is shift-reduced w.r.t. y. If $q \in K[x, y]$ is an irreducible factor of st then there is $C \in \mathbb{Z}$ such that $q(x, y + C)$ divides uv.*

Proof. a) If $q \mid t$ rewrite (15) as

$$u(x, y)v(x + 1, y)F'(x, y + 1) = u(x + 1, y)v(x, y)F'(x, y).$$

By Lemma 6, there are $m, n \in \mathbb{Z}$ such that

$$q(x, y + m) \mid u(x, y)v(x + 1, y) \quad \text{and} \quad q(x, y - n) \mid u(x + 1, y)v(x, y).$$

Since u/v is shift-reduced w.r.t. y it follows that $q(x, y + m) \mid u(x, y)$ or $q(x, y - n) \mid v(x, y)$. Take $C = m$ in the former case, $C = -n$ in the latter.

b) If $q \mid s$ rewrite (15) as

$$u(x + 1, y)v(x, y)\frac{1}{F'(x, y + 1)} = u(x, y)v(x + 1, y)\frac{1}{F'(x, y)}.$$

By Lemma 6, there are $m, n \in \mathbb{Z}$ such that

$$q(x, y + m) \mid u(x + 1, y)v(x, y) \quad \text{and} \quad q(x, y - n) \mid u(x, y)v(x + 1, y).$$

Since u/v is shift-reduced w.r.t. y it follows that $q(x, y + m) \mid v(x, y)$ or $q(x, y - n) \mid u(x, y)$. Take $C = m$ in the former case, $C = -n$ in the latter. □

Proof of Thm. 3 (cont'd): If p is an irreducible factor of uv then by Lemma 7, there are $A, B \in \mathbb{Z}$, $A > 0$, such that $p(x+A, y+B)$ divides st. By Lemma 8, there is $C \in \mathbb{Z}$ such that $p(x+A, y+B+C)$ divides uv. Hence by Theorem 2, uv factors into integer-linear factors.

If q is an irreducible factor of st then by Lemma 8, there is $C \in \mathbb{Z}$ such that $q(x, y + C)$ divides uv. By Lemma 7, there are $A, B \in \mathbb{Z}$, $A > 0$, such that $q(x + A, y + B + C)$ divides st. Hence by Theorem 2, st factors into integer-linear factors.

This shows that u, v, s, t, and hence F' and G', factor into integer-linear factors. □

Definition 9. *A hypergeometric term T over K is a* Z-term over K *if the polynomials f_0, f_1, g_0, g_1 in (5) factor into integer-linear factors.*

Corollary 3. *Let $T(n,k)$ be a hypergeometric term over an algebraically closed field K. Then there is a rational function $R \in K(x,y)$ and a Z-term $T'(n,k)$ such that $T(n,k) =_a R(n,k)T'(n,k)$.*

Proof. Let $F, G \in K(x,y)$ be such that $T(n+1,k) =_a F(n,k)T(n,k)$ and $T(n,k+1) =_a G(n,k)T(n,k)$, and let $R, F', G' \in K(x,y)$ be the rational functions associated with F, G by Theorem 3. Define

$$T'(n,k) =_a \frac{T(n,k)}{R(n,k)}.$$

Then $T(n,k) =_a R(n,k)T'(n,k)$ and

$$T'(n+1,k) =_a \frac{T(n+1,k)}{R(n+1,k)} =_a F(n,k)\frac{R(n,k)}{R(n+1,k)} \cdot \frac{T(n,k)}{R(n,k)} =_a F'(n,k)T'(n,k),$$

$$T'(n,k+1) =_a \frac{T(n,k+1)}{R(n,k+1)} =_a G(n,k)\frac{R(n,k)}{R(n,k+1)} \cdot \frac{T(n,k)}{R(n,k)} =_a G'(n,k)T'(n,k).$$

As F', G' factor into integer-linear factors, $T(n,k)$ is a Z-term. □

Acknowledgement. The authors wish to express their thanks to Ha Q. Le for his careful reading of a preliminary version of this paper.

References

[1] S. A. Abramov, Rational solutions of linear difference and differential equations with polynomial coefficients *USSR Comput. Math. Phys.* 29, 7–12. Transl. from *Zh. vychisl. mat. mat. fyz.* **29** (1989), 1611–1620.
[2] S. A. Abramov, H. Q. Le, Applicability of Zeilberger's algorithm to rational functions, *Proc. FPSAC '00*, LNCS, Springer-Verlag 2000, 91–102.
[3] S. A. Abramov, M. Petkovšek, Proof of a conjecture of Wilf and Zeilberger, in preparation.
[4] R. W. Gosper, Jr., Decision procedure for indefinite hypergeometric summation, *Proc. Natl. Acad. Sci. USA* **75** (1978) 40–42.
[5] W. Koepf, *Hypergeometric Summation: An Algorithmic Approach to Summation and Special Function Identities*, Vieweg, 1998.
[6] M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients, *J. Symb. Comp.* **14** (1992) 243–264.
[7] M. Petkovšek, H. S. Wilf and D. Zeilberger, $A = B$, A K Peters, 1996.
[8] R. Pirastu and V. Strehl, Rational summation and Gosper-Petkovšek representation, *J. Symb. Comp.* **19** (1995).
[9] R. P. Stanley, Differentially finite power series, *European J. Combin.* **1** (1980) 175–188.
[10] H. S. Wilf and D. Zeilberger, An algorithmic proof theory for hypergeometric (ordinary and "q") multi-sum/integral identities, *Invent. Math.* **108** (1992) 575–633.
[11] D. Zeilberger, The method of creative telescoping, *J. Symb. Comput.* **11** (1991) 195–204.

E-mail address: abramov@ccas.ru, marko.petkovsek@uni-lj.si

Computer Center of the Russian Academy of Science, Vavilova 40, Moscow 117967, Russia, and Department of Mathematics and Mechanics, University of Ljubljana, Jadranska 19, SI-1000 Ljubljana, Slovenia