J. Symbolic Computation (1998) **11**, 1–22

Hypergeometric Solutions of Linear Recurrences with Polynomial Coefficients

MARKO PETKOVŠEK

Department of Mathematics, University of Ljubljana, Slovenia

(Received 21 September 1990)

We describe algorithm Hyper which can be used to find all hypergeometric solutions of linear recurrences with polynomial coefficients.

1. Introduction

Let K be a field of characteristic zero. We assume that K is *computable*, meaning that the elements of K can be finitely represented and that there exist algorithms for carrying out the field operations.

Let $K^{\mathbb{N}}$ denote the ring of all sequences over K, with addition and multiplication defined term-wise. Following Stanley (1980) we identify two sequences if they agree from some point on. Formally, we are working in the quotient ring $\mathcal{S}_K := K^{\mathbb{N}}/J$ where J is the ideal of sequences with at most finitely many non-zero terms. Hence all equalities of the form $L(n) = R(n)$ are to be interpreted as valid for all large enough n. Accordingly, a sequence is considered non-zero if and only if it has infinitely many non-zero terms.

Definition 1.1. A sequence $a \in \mathcal{S}_K$ will be called *polynomial* (respectively *rational*) over K if there is a polynomial $f(x) \in K[x]$ (respectively a rational function $f(x) \in K(x)$) such that $a(n) = f(n)$ for all large enough $n \in \mathbb{N}$.

A non-zero sequence $a \in \mathcal{S}_K$ is a *hypergeometric term* (or simply *hypergeometric*) over K if there is a rational function $r(x) \in K(x)$ such that $a(n+1) = r(n)a(n)$ for all large enough $n \in \mathbb{N}$. Two hypergeometric terms are *similar* if their ratio is rational.

We will denote the sets of polynomial, rational, and hypergeometric sequences over K by $\mathcal{P}_K$, $\mathcal{R}_K$, and $\mathcal{H}_K$, respectively. By $\mathcal{L}(\mathcal{H}_K)$ we will denote the linear subspace spanned by $\mathcal{H}_K$ in the K-linear space $\mathcal{S}_K$.

Since two polynomials which agree on infinitely many arguments are identical, the polynomial defining a polynomial sequence is unique, and we will not distinguish between the two. The same goes for rational functions and sequences. In fact, $\mathcal{P}_K$ is a ring isomorphic to $K[x]$ and $\mathcal{R}_K$ is a field isomorphic to $K(x)$. It is easy to see that $\mathcal{H}_K$ is a multiplicative group. For any field F, let F^* denote its multiplicative group. Then $f : \mathcal{H}_K \to {\mathcal{R}_K}^*$ defined by $f : h(n) \mapsto h(n+1)/h(n)$ is an epimorphism which induces an

0747–7171/90/000000 + 00 \$03.00/0 © 1998 Academic Press Limited

isomorphism between $\mathcal{H}_K/K^*$ and $\mathcal{R}_K{}^*$. Obviously, similarity is a congruence relation in $\mathcal{H}_K$ induced by the normal subgroup $\mathcal{R}_K{}^*$.

EXAMPLE 1.1. Let $a(n) = 0, b(n) = 1, c(n) = n^2-1, d(n) = 2^n, e(n) = 2^n-1, f(n) = n!$. Then sequences a, b, c are polynomial over K, sequences b, c, d, f are hypergeometric over K, and sequences a, e belong to $\mathcal{L}(\mathcal{H}_K)$ but are not hypergeometric over K.

Let S denote the *shift operator* in the ring $K^{\mathbb{N}}$, defined by $Sa(n) = a(n+1)$. There is a unique endomorphism E of $\mathcal{S}_K$ such that $E(a+J) = Sa+J$, for every $a \in K^{\mathbb{N}}$. It is easy to see that E is in fact an automorphism of $\mathcal{S}_K$, and thus $(\mathcal{S}_K, E)$ is an *inversive difference ring*, according to the definition in Cohn (1965).

Let d be a positive integer and $r_0(n)$, $r_1(n)$, ..., $r_d(n)$ sequences from $\mathcal{R}_K$ such that $r_d \neq 0$. Then $L := \sum_{i=0}^{d} r_i E^i$ is a *linear recurrence operator with rational coefficients* over K. We will be interested in algorithms for finding hypergeometric solutions of equations of the form $La = f$ where L is as above and $f \in \mathcal{S}_K$.

Clearly, for every linear recurrence operator with rational coefficients there exists a linear recurrence operator with polynomial coefficients having the same kernel. Furthermore, every linear recurrence operator L can be written in the form $L = L_0 E^m$ where L_0 is an operator with non-zero trailing coefficient. Thus the affine solution space $L^{-1}f$ equals $E^{-m}L_0^{-1}f$, for any $f \in \mathcal{S}_K$. Therefore we will from now on assume that L is of the form

$$L := \sum_{i=0}^{d} p_i E^i \tag{1.1}$$

where $p_0(n)$, $p_1(n)$, ..., $p_d(n)$ are sequences from $\mathcal{P}_K$ such that $p_0, p_d \neq 0$. The *order* of L as defined in (1.1) is d, and its *degree* is $m := \max_{0 \le i \le d} \deg p_i(n)$. With L as in (1.1), the equation

$$La = 0 \tag{1.2}$$

is a *homogeneous linear recurrence with polynomial coefficients* for the unknown sequence $a \in \mathcal{S}_K$. We will also consider the corresponding *non-homogeneous recurrence*

$$La = f \tag{1.3}$$

where L is as in (1.1) and f is a given sequence from $\mathcal{S}_K$. The *order* and *degree* of recurrences (1.2) and (1.3) are defined to be those of L.

Abramov (1989a) gives an algorithm for finding all polynomial solutions of (1.3). Abramov (1989b) does the same for rational solutions. In this paper we present algorithm Hyper which can be used to find all hypergeometric solutions of (1.3). To give some motivation, we describe first an application of algorithm Hyper to definite hypergeometric summation.

The problem of indefinite hypergeometric summation was solved by Gosper (1978) who discovered an algorithm for finding hypergeometric solutions of the non-homogeneous first-order recurrence

$$a(n+1) - a(n) = h(n)$$

where $h(n)$ is a hypergeometric term. In other words, Gosper's algorithm decides whether the indefinite sum of a hypergeometric term is hypergeometric (apart from an additive constant).

One can ask the same question about definite sums of the form

$$a(n) = \sum_k F(n,k) \tag{1.4}$$

where summation ranges over all integers, and the two quotients $F(n+1,k)/F(n,k)$ and $F(n,k+1)/F(n,k)$ are rational functions of n and k. Using Bernstein's theory of holonomic functions, Zeilberger (1990) proved that every a of the form (1.4) satisfies a recurrence of the form (1.2). Zeilberger (1991) gives an algorithm which constructs such a recurrence. Wilf and Zeilberger (1992) generalize this to multiple and/or q-hypergeometric sums. Since algorithm Hyper can be used to obtain a basis for the space of solutions of (1.2) belonging to $\mathcal{L}(\mathcal{H}_K)$, the combination of Zeilberger's algorithm with Hyper, followed by solving a system of linear algebraic equations gives an algorithm for deciding whether (1.4) is a linear combination of hypergeometric terms.

The overview of the paper is as follows. In order to make the description of Hyper self-contained, we start in Section 2 by deriving algorithm Poly which finds a basis for the space of polynomial solutions of (1.2). A similar algorithm (but expecting the recurrence to be given in terms of the difference operator Δ) has been developed independently by Abramov (1989a).

Section 3 describes a normal form for rational functions which plays an important role both in Gosper's algorithm and in algorithm Hyper.

In Section 4 we develop algorithm Hyper which decides existence of hypergeometric solutions of (1.2) over some extension of the coefficient field of (1.2). The algorithm works by constructing a finite set of auxiliary recurrences of order d such that (1.2) has a hypergeometric solution if and only if some auxiliary recurrence has a non-zero polynomial solution. Algorithm Poly is then used on each auxiliary recurrence.

In Section 5 we show how to solve recurrences (1.2) and (1.3) over $\mathcal{L}(\mathcal{H}_K)$. Abramov (1989b) presents an algorithm for finding rational solutions of (1.3). With a simple modification, this algorithm can be used to find hypergeometric solutions of (1.3) when $f \neq 0$. Finally, we discuss the necessity to compute in algebraic extensions of the coefficient field of the recurrence in order to find all its hypergeometric solutions.

2. Polynomial solutions

In this section we assume that there exists an algorithm for finding integer roots of polynomials over K. Such a field is called *adequate* in Abramov (1989a, b). We follow the convention that $0^0 = 1$ and that $\binom{n}{k} = 0$ when $k < 0$ or $0 \leq n < k$.

The question of existence of non-zero polynomial solutions of (1.2) reduces to the problem of finding an upper bound N for the possible degrees of such solutions. Once we have N, we can substitute in the recurrence a generic polynomial of degree N for $a(n)$ and solve the resulting system of linear algebraic equations. The following example shows that it is not possible to bound the degree of polynomial solutions of (1.2) in terms of its order and degree.

Example 2.1. For any fixed non-negative integer k, the solutions of

$$na(n+1) - (n+k)a(n) = 0$$

are constant multiples of $\binom{n+k-1}{k}$ which is a polynomial in n of degree k, while the recurrence is of order and degree 1. Similar examples can be constructed for higher-order recurrences.

We consider first the second-order recurrence

$$p(n)a(n+2) + q(n)a(n+1) + r(n)a(n) = 0 \tag{2.1}$$

where p, q, r are polynomials. Assume that $a(n)$ is a non-zero polynomial solution of (2.1), of unknown degree N. Let m be the degree of (2.1) and

$$\begin{aligned} p(n) &= u_0 n^m + u_1 n^{m-1} + u_2 n^{m-2} + O(n^{m-3})\,, \\ q(n) &= v_0 n^m + v_1 n^{m-1} + v_2 n^{m-2} + O(n^{m-3})\,, \\ r(n) &= w_0 n^m + w_1 n^{m-1} + w_2 n^{m-2} + O(n^{m-3})\,. \end{aligned}$$

By definition of m, at least one of u_0, v_0, w_0 is non-zero. Further, let

$$a(n) = \alpha_N n^N + \alpha_{N-1} n^{N-1} + \alpha_{N-2} n^{N-2} + O(n^{N-3})$$

with $\alpha_N \neq 0$. Then, by the Binomial Theorem,

$$\begin{aligned} a(n+1) &= \alpha_N \left(n^N + N n^{N-1} + \binom{N}{2} n^{N-2} \right) \\ &+ \alpha_{N-1} \left(n^{N-1} + (N-1) n^{N-2} \right) + \alpha_{N-2} n^{N-2} + O(n^{N-3}) \end{aligned}$$

and

$$\begin{aligned} a(n+2) &= \alpha_N \left(n^N + 2N n^{N-1} + 4\binom{N}{2} n^{N-2} \right) \\ &+ \alpha_{N-1} \left(n^{N-1} + 2(N-1) n^{N-2} \right) + \alpha_{N-2} n^{N-2} + O(n^{N-3})\,. \end{aligned}$$

We substitute these expansions into (2.1) and equate coefficients of like powers of n. The coefficient of n^{N+m} yields, after cancelling α_N,

$$u_0 + v_0 + w_0 = 0\,. \tag{2.2}$$

If this condition is not fulfilled then (2.1) has no non-zero polynomial solution. Otherwise we proceed with the coefficient of n^{N+m-1}, use (2.2), cancel α_N, and find that

$$(2u_0 + v_0)N + u_1 + v_1 + w_1 = 0\,. \tag{2.3}$$

If $2u_0 + v_0 \neq 0$ then (2.3) determines a single possible value for N, and we are done. On the other hand, if

$$2u_0 + v_0 = 0 \tag{2.4}$$

then we must also have

$$u_1 + v_1 + w_1 = 0 \tag{2.5}$$

or else (2.1) has no non-zero polynomial solution. If both (2.4) and (2.5) hold we continue with the coefficient of n^{N+m-2}. Here we find after using (2.2), (2.4), and (2.5), and after cancelling α_N, that

$$u_0 N^2 + (2u_1 - u_0 + v_1)N + u_2 + v_2 + w_2 = 0\,. \tag{2.6}$$

Now $u_0 \neq 0$ or else it would follow from (2.2) and (2.4) that $u_0 = v_0 = w_0 = 0$, contrary to our assumption. Therefore (2.6) determines at most two possible values for N.

Next we summarize this algorithm and, after giving an example, generalize it to recurrences of arbitrary order.

Algorithm Poly for $d = 2$

INPUT: Polynomials

$$p(n) = \sum_{j=0}^{m} u_j n^{m-j},\ q(n) = \sum_{j=0}^{m} v_j n^{m-j},\ r(n) = \sum_{j=0}^{m} w_j n^{m-j}$$

with coefficients in K such that at least one of u_0, v_0, w_0 is non-zero.

OUTPUT: A basis $\mathcal{B}$ for the space of polynomial solutions over K of (2.1).

[1] If $u_0 + v_0 + w_0 \neq 0$ then $\mathcal{D} := \emptyset$
else if $u_0 \neq w_0$ then $\mathcal{D} := \{N;\ (u_0 - w_0)N + u_1 + v_1 + w_1 = 0\} \cap \mathbb{N}$
else if $u_1 + v_1 + w_1 \neq 0$ then $\mathcal{D} := \emptyset$
else $\mathcal{D} := \{N;\ u_0 N^2 + (u_1 - u_0 - w_1)N + u_2 + v_2 + w_2 = 0\} \cap \mathbb{N}$.

[2] If $\mathcal{D} = \emptyset$ then $\mathcal{B} := \emptyset$
else
$k := \max \mathcal{D}$;
find a basis $\mathcal{B}$ for the space of polynomial solutions of (2.1) over K of degree at most k, using the method of undetermined coefficients.

[3] Return $\mathcal{B}$ and stop.

Example 2.2. The recurrence

$$n(n+1)a(n+2) - 2n(n+100)a(n+1) + (n+99)(n+100)a(n) = 0$$

has $m = 2$, $u_0 = w_0 = u_1 = 1$, $v_0 = -2$, $v_1 = -200$, $w_1 = 199$, $u_2 = v_2 = 0$, and $w_2 = 9900$. Therefore $u_0 + v_0 + w_0 = u_0 - w_0 = u_1 + v_1 + w_1 = 0$, and the possible degrees are among the roots of

$$N^2 - 199N + 9900 = 0\,,$$

which are $N = 99$ and $N = 100$. In fact, there exist polynomial solutions both of degree 99 and of degree 100, e.g., $\binom{n+98}{99}$ and $\binom{n+98}{100}$.

Algorithm Poly for arbitrary d

INPUT: Polynomials

$$p_i(n) = \sum_{j=0}^{m} c_{i,j} n^{m-j}, \quad \text{for } i = 0, 1, \ldots, d\,, \tag{2.7}$$

with coefficients in K such that at least one of $c_{i,0}$, $0 \leq i \leq d$, is non-zero. We assume that $c_{i,j} = 0$ when $j < 0$ or $j > m$.

OUTPUT: A basis $\mathcal{B}$ for the space of polynomial solutions over K of (1.2).

[1] Initialize $s := -1$.

[2] Repeat
increment s by 1;
for $j = 0, 1, \ldots, s$ compute

$$b_j^{(s)} := \sum_{i=0}^{d} i^j c_{i,s-j} \tag{2.8}$$

until $\exists j \in \{0, 1, \ldots, s\}$ such that $b_j^{(s)} \neq 0$.

[3] Let $\mathcal{D}$ be the set of non-negative integer roots N of the polynomial

$$D(N) := \sum_{j=0}^{s} \binom{N}{j} b_j^{(s)} \ . \tag{2.9}$$

[4] If $\mathcal{D} = \emptyset$ then $\mathcal{B} := \emptyset$
else
$k := \max \mathcal{D}$;
find a basis $\mathcal{B}$ for the space of polynomial solutions of (1.2) over K of degree at most k, using the method of undetermined coefficients.

[5] Return $\mathcal{B}$ and stop.

We will call $D(N)$ the *degree polynomial* of (1.2).

To prove correctness of algorithm Poly we must show that the loop in step [2] eventually terminates, that the set $\mathcal{D}$ is finite, and that the degree of any non-zero polynomial solution of recurrence (1.2) belongs to $\mathcal{D}$. These facts are established by the following three lemmas.

LEMMA 2.1. *In algorithm* Poly*, $s \leq d$ at all times.*

PROOF. Assume that at some point $s = s_0$. Then $b_j^{(s)} = 0$ for $0 \leq j \leq s < s_0$. In particular, $b_s^{(s)} = 0$ for $0 \leq s < s_0$. If $s_0 > d$ this implies that

$$b_s^{(s)} = \sum_{i=0}^{d} i^s c_{i,0} = 0 \,, \quad \text{for } 0 \leq s \leq d \,.$$

Since $\det\,(i^s)_{i,s=0}^{d} = V(0, 1, \ldots, d)$ is a Vandermonde determinant it follows that $c_{i,0} = 0$ for $i = 0, 1, \ldots, d$. But these are the coefficients of n^m in $p_i(n)$, and by our assumption at least one of them is non-zero, a contradiction. □

LEMMA 2.2. *In algorithm* Poly*, $\mathcal{D}$ is a finite set.*

PROOF. By Lemma 2.1, the loop in step [2] terminates after at most $d+1$ iterations. When this happens at least one of the $b_j^{(s)}$ is non-zero, hence the degree polynomial (2.9) does not vanish identically and has a finite set of roots. It follows that $\mathcal{D}$ is finite. □

LEMMA 2.3. *Let* s_0 *be the final value of* s *in algorithm* Poly. *Let*

$$a(n) = \sum_{k=0}^{N_0} \alpha_k \, n^k, \quad \textit{where } \alpha_{N_0} \neq 0 \tag{2.10}$$

be a polynomial solution of recurrence (1.2). *Then*

$$\sum_{j=0}^{s_0} \binom{N_0}{j} b_j^{(s_0)} = 0 \tag{2.11}$$

where $b_j^{(s)}$ *is as in* (2.8).

PROOF. By the Binomial Theorem,

$$a(n+i) = \sum_{k=0}^{N_0} \alpha_k \sum_{l=0}^{k} \binom{k}{l} i^{k-l} n^l \,. \tag{2.12}$$

We adopt the convention that $c_{i,j} = 0$ and $\alpha_k = 0$ whenever any of i, j, k are out of bounds indicated in (2.7) and (2.10). By substituting (2.7) and (2.12) into (1.1) we obtain

$$La = \sum_{i,j,k,l} c_{i,j} \alpha_k \binom{k}{l} i^{k-l} n^{m+l-j}$$

where each summation index ranges over all integers and the order of summation can be chosen arbitrarily. Replacing j by $r = m + l - j$ yields

$$La = \sum_r n^r \sum_k \alpha_k \sum_l \binom{k}{l} \sum_i i^{k-l} c_{i,m+l-r} \,.$$

The coefficient of n^r can be written as

$$\sum_k \alpha_k \sum_l \binom{k}{l} b_j^{(s)} \tag{2.13}$$

where $j = k-l$ and $s = k+m-r$. By definition of s_0, we have $b_j^{(s)} = 0$ for $0 \leq j \leq s < s_0$. By our convention, this also holds when $j > s$ since then $c_{i,s-j} = 0$. If $j < 0$ the binomial coefficient in (2.13) is zero; if $j \geq 0$ and $s < s_0$ then $b_j^{(s)} = 0$. So the inner sum in (2.13) vanishes whenever $s < s_0$, i.e., when $k < s_0 - m + r$. On the other hand, $\alpha_k = 0$ for $k > N_0$, hence the terms of the outer sum in (2.13) are zero unless $s_0 - m + r \leq k \leq N_0$. As a consequence, (2.13) is zero unless $r \leq m + N_0 - s_0$, so

$$\deg La \leq m + N_0 - s_0 \,.$$

Now replace r by $t = s_0 - m + r$. Then (2.13) can be rewritten as

$$\sum_{k=t}^{N_0} \alpha_k \sum_l \binom{k}{l} b_{k-l}^{(k-t+s_0)} \,. \tag{2.14}$$

If $La = 0$ then (2.14) is zero for all values of t. In particular, for $t = N_0$ we have

$$\alpha_{N_0} \sum_l \binom{N_0}{l} b_{N_0-l}^{(s_0)} = 0 \,.$$

After cancelling α_{N_0} and replacing l by $j = N_0 - l$, this turns into (2.11). $\square$

Remark 2.1. To implement Poly efficiently, we note that equating (2.14) to zero for $t = N_0, N_0 - 1, \dots, 0$ gives a system of linear equations for the unknown coefficients $\alpha_{N_0}, \alpha_{N_0-1}, \dots, \alpha_0$ which is already in triangular form.

Remark 2.2. We can modify Poly so that it will find a basis for the affine space of polynomial solutions of the non-homogeneous recurrence (1.3). Let $f = f(n) = \sum_{i=0}^{r} f_i n^i$ in (1.3) where $r = \deg f(n)$. In the proof of Lemma 2.3 we saw that either $\deg La = m + N_0 - s_0$, or $\deg La < m + N_0 - s_0$ and (2.11) holds. Hence either $N_0 = r + s_0 - m$, or N_0 is a non-negative integer root of the degree polynomial (2.9). Choosing the largest of these values as N_0, we find the unknown coefficients α_i from

$$\sum_{k=t}^{N_0} \alpha_k \sum_l \binom{k}{l} b_{k-l}^{(k-t+s_0)} = f_{t+m-s_0}\,, \quad \text{for } t = N_0, N_0 - 1, \dots, 0\,,$$

which is a triangular system of linear equations.

3. A normal form for rational functions

The following lemma appears in Gosper (1978), though without mentioning properties (AC) and (BC).

Lemma 3.1. *Let K be a field of characteristic zero and $r(x)$ a non-zero rational function over K. Then there exists a non-zero constant $Z \in K$ and monic polynomials $A(x)$, $B(x)$, and $C(x)$ over K such that*

$$r(x) = Z \frac{A(x)}{B(x)} \frac{C(x+1)}{C(x)} \tag{3.1}$$

where

(i) $\gcd(A(x), B(x+k)) = 1$ *for every non-negative integer k,* (AB)
(ii) $\gcd(A(x), C(x)) = 1$, (AC)
(iii) $\gcd(B(x), C(x+1)) = 1$. (BC)

Proof. Let $r(x) = p(x)/q(x)$ where p, q are polynomials over K with leading coefficients α and β, respectively. Let $Z := \alpha/\beta$. Then $r(x) \neq 0$ implies $p(x) \neq 0$ which implies $\alpha \neq 0$ which in turn implies $Z \neq 0$.

Denote by $R(k)$ the resultant of $p(x)$ and $q(x+k)$ with respect to x. Let N be the largest non-negative integer root of $R(k)$, or -1 if there is none. Construct polynomials $p_i(x)$, $q_i(x)$ for $-1 \leq i \leq N$ inductively by

$$p_{-1}(x) = \frac{p(x)}{\alpha}, \quad q_{-1}(x) = \frac{q(x)}{\beta},$$

and for $0 \leq i \leq N$,

$$\begin{aligned} s_i(x) &= \gcd(p_{i-1}(x), q_{i-1}(x+i))\,, \\ p_i(x) &= \frac{p_{i-1}(x)}{s_i(x)}\,, \\ q_i(x) &= \frac{q_{i-1}(x)}{s_i(x-i)}\,. \end{aligned}$$

Let

$$\begin{aligned} A(x) &:= p_N(x)\,, \\ B(x) &:= q_N(x)\,, \\ C(x) &:= \prod_{i=1}^{N}\prod_{k=1}^{i} s_i(x-k)\,. \end{aligned}$$

Then

$$\begin{aligned} Z\frac{A(x)}{B(x)}\frac{C(x+1)}{C(x)} &= \frac{\alpha}{\beta}\frac{p_N(x)}{q_N(x)}\prod_{i=1}^{N}\prod_{k=1}^{i}\frac{s_i(x-k+1)}{s_i(x-k)} \\ &= \frac{\alpha}{\beta}\frac{p_{-1}(x)}{\prod_{i=0}^{N} s_i(x)}\frac{\prod_{i=0}^{N} s_i(x-i)}{q_{-1}(x)}\prod_{i=1}^{N}\frac{s_i(x)}{s_i(x-i)} \\ &= \frac{\alpha p_{-1}(x)}{\beta q_{-1}(x)} = \frac{p(x)}{q(x)} = r(x)\,. \end{aligned}$$

Polynomials $A(x), B(x), C(x)$ are monic by construction. Now we verify properties (AB), (AC), and (BC).

By definition of p_k, q_k, and s_k,

$$\gcd(p_k(x), q_k(x+k)) = \gcd(\frac{p_{k-1}(x)}{s_k(x)}, \frac{q_{k-1}(x+k)}{s_k(x)}) = 1$$

for all k such that $0 \le k \le N$. Let $0 \le k \le i, j \le N$. Then $p_i, p_j \mid p_k$ and $q_i, q_j \mid q_k$, therefore $\gcd(p_i(x), q_j(x+k)) \mid \gcd(p_k(x), q_k(x+k))$. Hence

$$\gcd(p_i(x), q_j(x+k)) = 1\,. \tag{3.2}$$

(AB): Setting $i = j = N$ in (3.2), we see that $A(x)$ and $B(x+k)$ are relatively prime for all k such that $0 \le k \le N$. By definition of N this implies that they are relatively prime for all $k \ge 0$.

(AC): If $A(x)$ and $C(x)$ have a common factor then so do $p_N(x)$ and $s_i(x-k)$, for some i and k such that $1 \le k \le i \le N$. Then $p_N(x)$ and $q_{i-1}(x+i-k)$ have a common factor, too, by definition of q_i. But as $0 \le i-k \le i-1 < N$, this contradicts (3.2).

(BC): If $B(x)$ and $C(x+1)$ have a common factor then so do $q_N(x)$ and $s_i(x-k)$, for some i and k such that $0 \le k \le i-1 \le N-1$. Then $q_N(x)$ and $p_{i-1}(x-k)$ have a common factor, too, by definition of p_i. Hence $q_N(x+k)$ and $p_{i-1}(x)$ have a common factor. But as $0 \le k \le i-1 < N$, this contradicts (3.2). □

Note that the proof of Lemma 3.1 actually gives an algorithm to compute (3.1). This algorithm utilizes gcd and resultant computations in $K[x]$ and $K[x,k]$, respectively.

Lemma 3.2. *Let K be a field of characteristic zero. Let $a(x)$, $b(x)$, and $c(x)$ be polynomials over K such that* $\gcd(a(x), c(x)) = \gcd(b(x), c(x+1)) = 1$. *If*

$$r(x) = \frac{a(x)}{b(x)}\frac{c(x+1)}{c(x)} \tag{3.3}$$

and $C(x)$ is related to $r(x)$ as in Lemma 3.1, then $c(x)$ divides $C(x)$.

PROOF. Use (3.1) to rewrite (3.3) as

$$ZA(x)b(x)c(x)C(x+1) = a(x)B(x)C(x)c(x+1)\,. \tag{3.4}$$

Let

$$\begin{aligned} g(x) &:= \gcd(c(x), C(x))\,, && (3.5)\\ d(x) &:= c(x)/g(x)\,, && (3.6)\\ D(x) &:= C(x)/g(x)\,. && (3.7) \end{aligned}$$

Then $\gcd(d(x), D(x)) = \gcd(a(x), d(x)) = \gcd(b(x), d(x+1)) = 1$. Cancelling $g(x)g(x+1)$ from (3.4) gives

$$ZA(x)b(x)d(x)D(x+1) = a(x)B(x)D(x)d(x+1)\,.$$

It follows that

$$\begin{aligned} d(x) &\mid B(x)d(x+1)\,,\\ d(x+1) &\mid A(x)d(x)\,. \end{aligned}$$

Using this repeatedly we obtain

$$\begin{aligned} d(x) &\mid B(x)B(x+1)\cdots B(x+k-1)d(x+k)\,,\\ d(x) &\mid A(x-1)A(x-2)\cdots A(x-k)d(x-k)\,, \end{aligned}$$

for all $k \in \mathbb{N}$. Since F has characteristic zero $\gcd(d(x), d(x+k)) = \gcd(d(x), d(x-k)) = 1$, for all large enough k. It follows that $d(x)$ divides both $B(x)B(x+1)\cdots B(x+k-1)$ and $A(x-1)A(x-2)\cdots A(x-k)$ for all large enough k. But these two polynomials are relatively prime by the assumptions of the Lemma, so $d(x)$ is constant. Hence $c(x)$ divides $C(x)$, by (3.6) and (3.5). □

COROLLARY 3.1. *Factorization* (3.1) *described in Lemma 3.1 is unique.*

PROOF. Assume that

$$r(x) = Z\frac{A(x)}{B(x)}\frac{C(x+1)}{C(x)} = z\frac{a(x)}{b(x)}\frac{c(x+1)}{c(x)} \tag{3.8}$$

where A, B, C as well as a, b, c are monic polynomials satisfying properties (AB), (AC), and (BC) of Lemma 3.1. Then $Z = z$. By Lemma 3.2, $c(x)$ divides $C(x)$ and vice versa. As they are both monic, $c(x) = C(x)$. Therefore $A(x)b(x) = a(x)B(x)$. By property (AB) of Lemma 3.1, $a(x)$ divides $A(x)$ and vice versa, so $a(x) = A(x)$ and $b(x) = B(x)$. □

4. Hypergeometric solutions

Let K_0 be a field of characteristic zero and K an extension field of K_0. Given a linear recurrence operator L with polynomial coefficients over K_0, we seek solutions of (1.2) which are hypergeometric over K. We will call K_0 the *coefficient field* of the recurrence. We assume that there exist algorithms for finding integer roots of polynomials over K and for factoring polynomials over K into factors irreducible over K.

As with polynomial solutions, we consider first the second-order recurrence (2.1). Assume that $a(n)$ is a hypergeometric solution of (2.1). Then there is a rational sequence $S(n)$ such that $a(n+1) = S(n)a(n)$. Substituting this into (2.1) and cancelling $a(n)$ gives

$$p(n)S(n+1)S(n) + q(n)S(n) + r(n) = 0\,.$$

Let

$$S(n) = Z\frac{A(n)}{B(n)}\frac{C(n+1)}{C(n)}$$

be the normal form for $S(n)$ described in Lemma 3.1. Then

$$Z^2 p(n)A(n+1)A(n)C(n+2)+Zq(n)B(n+1)A(n)C(n+1)+r(n)B(n+1)B(n)C(n) = 0\,. \tag{4.1}$$

The first two terms contain $A(n)$ as a factor, so $A(n)$ divides $r(n)B(n+1)B(n)C(n)$. By properties (AB) and (AC) of the normal form, $A(n)$ is relatively prime with $C(n)$, $B(n)$, and $B(n+1)$, so $A(n)$ divides $r(n)$. Similarly we find that $B(n+1)$ divides $p(n)A(n+1)A(n)C(n+2)$, therefore by properties (AB) and (BC) of the normal form, $B(n+1)$ divides $p(n)$. This leaves us with a finite set of candidates for $A(n)$ and $B(n)$, namely, the monic factors of $r(n)$ and $p(n-1)$, respectively. We can cancel $A(n)B(n+1)$ from the coefficients of (4.1) to obtain

$$Z^2\frac{p(n)}{B(n+1)}A(n+1)C(n+2) + Zq(n)C(n+1) + \frac{r(n)}{A(n)}B(n)C(n) = 0\,. \tag{4.2}$$

To determine the value of Z, we consider the leading coefficient of the left-hand side in (4.2) and find out that Z satisfies a quadratic equation with known coefficients. So given the choice of $A(n)$ and $B(n)$, there are at most two choices for Z.

For a fixed choice of $A(n)$, $B(n)$, and Z, we can use algorithm Poly to determine if (4.2) has any non-zero polynomial solution $C(n)$. If yes, we have found a hypergeometric solution of (2.1). If (4.2) has no non-zero polynomial solution whatever the choice of $A(n)$, $B(n)$, and Z, then (2.1) has no hypergeometric solution.

Algorithm Hyper for $d = 2$

INPUT: Polynomials $p(n)$, $q(n)$, and $r(n)$ over K_0; an extension field K of K_0.

OUTPUT: A hypergeometric solution of (2.1) over K if it exists; 0 otherwise.

[1] For all monic factors $A(n)$ of $r(n)$ and $B(n)$ of $p(n-1)$ over K do:

> $P(n) := (p(n)/B(n+1))A(n+1)$;
> $Q(n) := q(n)$;
> $R(n) := (r(n)/A(n))B(n)$;
> $m := \max\{\deg P, \deg Q, \deg R\}$;
> let α, β, γ be the coefficients of n^m in P, Q, R, respectively;
>
> for all non-zero $Z \in K$ such that $\alpha Z^2 + \beta Z + \gamma = 0$ do:
>
> > If the recurrence
> >
> > $$Z^2 P(n)C(n+2) + ZQ(n)C(n+1) + R(n)C(n) = 0$$
> >
> > has a non-zero polynomial solution $C(n)$ over K then
> >
> > > $S(n) := Z(A(n)/B(n))(C(n+1)/C(n))$;
> > > return a non-zero solution $a(n)$ of $a(n+1) = S(n)a(n)$ and stop.

[2] Return 0 and stop.

EXAMPLE 4.1. The recurrence

$$(n+4)a(n+2) + a(n+1) - (n+1)a(n) = 0$$

appears in Bender and Orszag (1978), p. 43, Example 2.3.5. The monic factors of $r(n)$ are 1 and $n+1$ and those of $p(n-1)$ are 1 and $n+3$. Taking $A(n) = B(n) = 1$ leads to $Z = \pm 1$; however, algorithm Poly finds no non-zero polynomial solutions of the auxiliary recurrences. If exactly one of $A(n), B(n)$ is equal to 1, the equation for Z is either $Z = 0$ or contradictory. Finally, the choice $A(n) = n+1$, $B(n) = n+3$ leads to $Z = \pm 1$ again. For $Z = 1$, the auxiliary recurrence is

$$(n+2)C(n+2) + C(n+1) - (n+3)C(n) = 0$$

with, up to a constant factor, unique polynomial solution $C(n) = 1$. This gives $S(n) = (n+1)/(n+3)$ and

$$a(n) = \frac{1}{(n+1)(n+2)}\,.$$

For $Z = -1$, the auxiliary recurrence is

$$(n+2)C(n+2) - C(n+1) - (n+3)C(n) = 0$$

with, up to a constant factor, unique polynomial solution $C(n) = 2n+3$. This gives $S(n) = -((n+1)/(n+3))((2n+5)/(2n+3))$ and

$$a(n) = \frac{(-1)^n(2n+3)}{(n+1)(n+2)}\,.$$

EXAMPLE 4.2. The number $d(n)$ of derangements of a set with n elements satisfies the recurrence

$$a(n) = (n-1)a(n-1) + (n-1)a(n-2)\,. \tag{4.3}$$

Taking $A(n) = B(n) = 1$ yields $Z = -1$, but the auxiliary recurrence has no non-zero polynomial solution. The remaining choice $A(n) = n+1$, $B(n) = 1$ leads to $Z^2 - Z = 0$, so $Z = 1$. The auxiliary recurrence

$$(n+2)C(n+2) - (n+1)C(n+1) - C(n) = 0$$

has, up to a constant factor, the only polynomial solution $C(n) = 1$. Then $S(n) = n+1$ and, up to a constant factor,

$$a(n) = n!$$

is the only hypergeometric solution of (4.3). Using the well-known method of reduction of order, we find another fundamental solution (which happens to be precisely the number of derangements):

$$d(n) = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}\,. \tag{4.4}$$

Since this is not a constant multiple of $n!$ (although it comes very close, being equal to the integer nearest to $n!/e$), we have proved that (4.4) is not a hypergeometric term. As the summand in (4.4) does not depend on n this can also be shown by Gosper's algorithm.

EXAMPLE 4.3. In van der Poorten (1979), it is shown that the numbers

$$a(n) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \tag{4.5}$$

satisfy the recurrence

$$(n+2)^3 a(n+2) - (2n+3)(17n^2+51n+39)a(n+1) + (n+1)^3 a(n) = 0\,. \tag{4.6}$$

Here all the coefficients are of the same degree, therefore the equation for Z will have no non-zero solution unless $A(n)$ and $B(n)$ are of the same degree as well. But they are both monic factors of $(n+1)^3$, so they must be equal. Then $P(n) = p(n)$, $Q(n) = q(n)$, $R(n) = r(n)$, and the equation for Z is $Z^2 - 34Z + 1 = 0$ with solutions $Z = 17 \pm 12\sqrt{2}$. In both cases the auxiliary recurrence has no non-zero polynomial solutions, proving that (4.6) has no hypergeometric solution. As a consequence, (4.5) is not hypergeometric.

Algorithm Hyper for arbitrary d

INPUT: Polynomials $p_i(n)$ over K_0, for $i = 0, 1, \ldots, d$; an extension field K of K_0.

OUTPUT: A hypergeometric solution of (1.2) over K if it exists; 0 otherwise.

[1] For all monic factors $A(n)$ of $p_0(n)$ and $B(n)$ of $p_d(n-d+1)$ over K do:

$P_i(n) := p_i(n) \prod_{j=0}^{i-1} A(n+j) \prod_{j=i}^{d-1} B(n+j)$, for $i = 0, 1, \ldots, d$;
$m := \max_{0 \le i \le d} \deg P_i(n)$;
let α_i be the coefficient of n^m in $P_i(n)$, for $i = 0, 1, \ldots, d$;

for all non-zero $Z \in K$ such that

$$\sum_{i=0}^{d} \alpha_i Z^i = 0 \tag{4.7}$$

do:

If the recurrence

$$\sum_{i=0}^{d} Z^i P_i(n) C(n+i) = 0 \tag{4.8}$$

has a non-zero polynomial solution $C(n)$ over K then

$S(n) := Z(A(n)/B(n))(C(n+1)/C(n))$;
return a non-zero solution $a(n)$ of $a(n+1) = S(n)a(n)$ and stop.

[2] Return 0 and stop.

THEOREM 4.1. *Let $a(n)$ be a non-zero solution of* (1.2) *such that $a(n+1) = S(n)a(n)$ where $S(n)$ is a rational sequence. Let*

$$S(n) = Z \frac{A(n)}{B(n)} \frac{C(n+1)}{C(n)} \tag{4.9}$$

be the normal form described in Lemma 3.1. Let $P_i(n)$ *and* α_i*, for* $i = 0, 1, \ldots, d$*, be defined as in algorithm* Hyper. *Then* $\sum_{i=0}^{d} \alpha_i Z^i = 0$*,* $A(n)$ *divides* $p_0(n)$*,* $B(n)$ *divides* $p_d(n-d+1)$*, and* $C(n)$ *satisfies* (4.8).

Conversely, if Z *is an arbitrary constant,* $A(n)$ *and* $B(n)$ *arbitrary sequences,* $C(n)$ *satisfies* (4.8) *where* $P_i(n)$*, for* $i = 0, 1, \ldots, d$*, is defined as in algorithm* Hyper, *and* $a(n+1) = S(n)a(n)$ *where* $S(n)$ *is as in* (4.9)*, then* $a(n)$ *satisfies* (1.2).

Proof. From (1.2) and $a(n+1) = S(n)a(n)$ it follows that

$$\sum_{i=0}^{d} p_i(n) \left(\prod_{j=0}^{i-1} S(n+j) \right) a(n) = 0\,, \tag{4.10}$$

hence after cancelling $a(n)$ and using (4.9) we have

$$\sum_{i=0}^{d} p_i(n) Z^i \left(\prod_{j=0}^{i-1} \frac{A(n+j)}{B(n+j)} \right) \frac{C(n+i)}{C(n)} = 0\,. \tag{4.11}$$

Multiplication by $C(n) \prod_{j=0}^{d-1} B(n+j)$ now gives (4.8). All terms of the sum in (4.8) with $i > 0$ contain the factor $A(n)$, thus $A(n)$ divides the term with $i = 0$ which is $p_0(n)C(n) \prod_{j=0}^{d-1} B(n+j)$. By properties (AB) and (AC) of the normal form, it follows that $A(n)$ divides $p_0(n)$. Similarly, $B(n+d-1)$ divides $Z^d p_d(n) C(n+d) \prod_{j=0}^{d-1} A(n+j)$, hence by properties (AB) and (BC) of the normal form, $B(n+d-1)$ divides $p_d(n)$, so $B(n)$ divides $p_d(n-d+1)$. Finally, a look at the leading coefficient of the left-hand side of (4.8) shows that $\sum_{i=0}^{d} \alpha_i Z^i = 0$.

To prove the converse, we retrace the steps which led from (1.2) through (4.10) and (4.11) to (4.8), in reverse order. □

Remark 4.1. Although the worst-case time complexity of Hyper is exponential in the degree of (1.2), a careful implementation can speed it up in several places. Here we give a few suggestions.

We can reduce the degree of recurrence (4.8) by cancelling the factor $A(n)B(n+d-1)$, as we did in (4.2) for the case $d = 2$.

Observe that the coefficients of equation (4.7) which determines Z depend only on the difference $D(A, B) = \deg B(n) - \deg A(n)$ and not on $A(n)$ or $B(n)$ themselves. Therefore it is advantageous to test pairs of factors $A(n)$, $B(n)$ in order of the value of $D(A, B)$.

We can skip those values of $D(A, B)$ for which (4.7) has a single non-zero term and thus no non-zero solution. For example, this happens when $\deg p_d(n) = \deg p_0(n) \geq \deg p_i(n)$ for $0 \leq i \leq d$, and $D(A, B) \neq 0$. Hence in this case it suffices to test pairs $A(n)$, $B(n)$ of equal degree (cf. Examples 4.1 and 4.3).

We can skip all pairs $A(n)$, $B(n)$ which do not satisfy property (AB) of Lemma 3.1.

Example 4.4. The number $i(n)$ of involutions of a set with n elements satisfies the recurrence

$$a(n) = a(n-1) + (n-1)a(n-2)\,.$$

More generally, let $r \geq 2$. The number $i_r(n)$ of permutations which contain no cycles longer than r satisfies the recurrence

$$\begin{aligned} a(n) &= a(n-1) + (n-1)a(n-2) + (n-1)(n-2)a(n-3) + \ldots \\ &+ (n-1)\cdots(n-r+2)(n-r+1)a(n-r) \end{aligned} \tag{4.12}$$

(see, e.g., Wimp and Zeilberger (1985)). In Hyper, the degrees of the coefficients of auxiliary recurrences are obtained by adding to the degree sequence of the coefficients of the original recurrence (starting with the leading coefficient) an arithmetic progression with increment $D = D(A, B)$. In case of (4.12), the degree sequence is $0, 0, 1, 2, \ldots, r-1$. Adding to this sequence any arithmetic progression with integer increment D will produce a sequence with a single term of maximum value (the first one if $D < 0$; the last one if $D \geq 0$), implying that (4.7) has a single non-zero term for all choices of $A(n)$ and $B(n)$. Therefore (4.12) has no hypergeometric solution. This example shows that for any $d \geq 2$, there exist recurrences of order d without hypergeometric solutions. In particular, for $r = 2$ this means that the sum

$$i(n) = \sum_k \frac{n!}{(n-2k)!\, 2^k\, k!}, \tag{4.13}$$

(see, e.g., Comtet (1974)), is not a hypergeometric term.

Example 4.5. (Gessel *et al.*, 1991) encountered the following sums when evaluating inversion polynomials of plane trees at $q = -1$ (see their paper for the relevant definitions):

$$f(n) = \sum_{k=0}^{n} \binom{3k}{k}\binom{3n-3k}{n-k}, \tag{4.14}$$

$$g(n) = \sum_{k=0}^{n-1} \binom{3k}{k}\binom{3n-3k-2}{n-k-1}. \tag{4.15}$$

As shown by S.B. Ekhad using the algorithm of Zeilberger (1991), $Lf = 0$ and $Mg = 0$ where

$$\begin{aligned} L &= 8(n+2)(2n+3)E^2 - 6(36n^2+99n+70)E + 81(3n+2)(3n+4), \\ M &= 16(n+2)(2n+3)(2n+5)E^3 - 12(2n+3)(54n^2+153n+130)E^2 \\ &+ 324(27n^3+72n^2+76n+30)E - 2187n(3n+1)(3n+2), \end{aligned}$$

and E denotes the shift operator. Using Hyper on $La = 0$ we find that $\dim(\mathrm{Ker}L \cap \mathcal{L}(\mathcal{H}_K)) = 1$. A fundamental hypergeometric solution is

$$x(n) = \left(\frac{27}{4}\right)^n$$

(corresponding to $A(n) = B(n) = C(n) = 1$, $Z = 27/4$). Therefore $L = L_1 L_2$ where $L_1 = 2(n+2)(2n+3)E - 3(3n+2)(3n+4)$, and

$$L_2 = 4E - 27. \tag{4.16}$$

From this we obtain a second fundamental solution

$$y(n) = x(n) \sum_{k=0}^{n} \frac{\binom{3k}{k}}{3k-1} \left(\frac{27}{4}\right)^{-k}.$$

From $f(0) = 1$, $f(1) = 6$ we determine that

$$f(n) = (x(n) - y(n))/2\,. \tag{4.17}$$

Also, $f(n)$ satisfies the non-homogeneous first-order recurrence

$$4f(n+1) - 27f(n) = -3\frac{\binom{3n+1}{n+1}}{2n+1}\,.$$

Using Hyper on $Ma = 0$ we find that $\dim(\mathrm{Ker}M \cap \mathcal{L}(\mathcal{H}_K)) = 2$. Two fundamental hypergeometric solutions are $x(n)$ and

$$u(n) = \frac{27^n}{n\binom{2n}{n}}$$

(corresponding to $A(n) = n$, $B(n) = n + 1/2$, $C(n) = 1$, $Z = 27/4$). By means of these solutions we find that $M = M_1 M_2 L_2$ where L_2 is as in (4.16), $M_1 = 2(n+2)(2n+3)E - 3(3n+1)(3n+2)$, and $M_2 = 2(2n+3)E - 27n$. From this we obtain a third fundamental solution

$$v(n) = x(n)\sum_{k=0}^{n-1}\frac{\binom{3k}{k}}{2k+1}\left(\frac{27}{4}\right)^{-k}\,.$$

From $g(1) = 1$, $g(2) = 7$, $g(3) = 48$ we determine that for $n \geq 1$,

$$g(n) = (3x(n) + v(n))/27\,. \tag{4.18}$$

The fact that $u(n)$ does not appear in (4.18) hints at existence of a second-order recurrence with polynomial coefficients satisfied by $g(n)$. Indeed, we find that $Tg = 0$ where $T = 8(n+1)(2n+3)E^2 - 6(36n^2 + 63n + 31)E + 81(3n+1)(3n+2) = T_1 L_2$, with $T_1 = 2(n+1)(2n+3)E - 3(3n+1)(3n+2)$, and L_2 as in (4.16). Also, $g(n)$ satisfies the non-homogeneous first-order recurrence

$$4g(n+1) - 27g(n) = \frac{\binom{3n}{n}}{2n+1}\,.$$

Observe that (4.17) and (4.18) are identities which express $f(n)$ and $g(n)$ in terms of sums whose summand does not depend on n, in contrast to their definitions in (4.14) and (4.15).

We conclude that neither $f(n)$ nor $g(n)$ belongs to $\mathcal{L}(\mathcal{H}_K)$.

Example 4.6. Consider the recurrence

$$n(n+1)a(n+2) - 2n(n+k+1)a(n+1) + (n+k)(n+k+1)a(n) = 0 \tag{4.19}$$

over the field $\mathrm{Q}(k)$ where k is transcendental over Q. Since its degree polynomial $D(N) = (N-k)(N-k-1)$ has no integer roots, (4.19) has no polynomial solutions. However, Hyper succeeds with $Z = 1$, $A(n) = n + k$, $B(n) = n$, and $C(n)$ any linear polynomial. Thus we find two fundamental hypergeometric solutions, say, $(k+1)^{\overline{n-1}}/(n-1)!$ and $(k+2)^{\overline{n-2}}/(n-2)!$ where

$$k^{\overline{n}} = \prod_{i=0}^{n-1}(k+i)$$

is the rising power. If we substitute a non-negative integer value for k these solutions can be written as $\binom{n+k-1}{k}$ and $\binom{n+k-1}{k+1}$, respectively. It follows that for every specific non-negative integer k all solutions of (4.19) are polynomial over $\mathbb{Q}$; cf. Example 2.2 which is a special case with $k = 99$.

5. Ramifications of Hyper

5.1. Linear combinations of hypergeometric terms

Sometimes hypergeometric terms are used to define the notion of "closed form", but this leaves out such nice sequences as $2^n - 1$. The class $\mathcal{L}(\mathcal{H}_K)$ is larger and has more closure properties than $\mathcal{H}_K$, while still being acceptable as a paradigm for "closed form". Algorithm Hyper can be modified to find a basis for the space of all those solutions of (1.2) which belong to $\mathcal{L}(\mathcal{H}_K)$.

PROPOSITION 5.1. *Let L be as in* (1.1)*, and let a be a hypergeometric term such that $La \neq 0$. Then La is hypergeometric and similar to a.*

PROOF. Let $r := Ea/a$. Then $E^i a = E^{i-1}(ra) = (E^{i-1}r)(E^{i-1}a) = \left(\prod_{j=0}^{i-1} E^j r\right) a$, so

$$La = \sum_{i=0}^{d} p_i E^i a = \left(\sum_{i=0}^{d} p_i \prod_{j=0}^{i-1} E^j r\right) a$$

is a non-zero rational multiple of a. $\square$

PROPOSITION 5.2. *Every sequence from $\mathcal{L}(\mathcal{H}_K)$ can be written as a sum of pairwise dissimilar hypergeometric terms.*

PROOF. Since the sum of two similar hypergeometric terms is either hypergeometric or zero, the desired representation can be achieved by grouping together similar terms. $\square$

THEOREM 5.1. *Let S be a non-empty set of hypergeometric terms. If S is linearly dependent over $\mathcal{R}_K$ then S contains two similar hypergeometric terms.*

PROOF. Let $S \subseteq \mathcal{H}_K$ be linearly dependent over $\mathcal{R}_K$. We prove the assertion by induction on $|S|$.

If $|S| = 1$ then S is not linearly dependent over $\mathcal{R}_K$. If $|S| > 1$, let $S' = \{a_1, a_2, \ldots, a_m\}$ be a minimal subset of S which is linearly dependent over $\mathcal{R}_K$. Then there exist non-zero elements $r_1, r_2, \ldots, r_m \in \mathcal{R}_K$ such that $\sum_{i=1}^m r_i a_i = 0$. Let $b_i := r_i a_i$, for $i = 1, 2, \ldots, m$. Then $b_i \in \mathcal{H}_K$, and

$$\sum_{i=1}^{m} b_i = 0\,. \tag{5.1}$$

For $i = 1, 2, \ldots, m$, let $s_i := Eb_i/b_i$. Apply E to (5.1), multiply (5.1) by s_m, then subtract the resulting equations to find

$$\sum_{i=1}^{m-1} (s_i - s_m) b_i = 0\,. \tag{5.2}$$

If $s_i = s_m$ for some i then b_i and b_m are constant multiples of each other, hence a_i and a_m are similar. Otherwise all terms in (5.2) are hypergeometric. By induction hypothesis there exist j and k, $j \neq k$, such that $(s_j - s_m)b_j$ and $(s_k - s_m)b_k$ are similar. But then so are a_j and a_k. $\square$

COROLLARY 5.1. *Let L be a linear recurrence operator with polynomial coefficients, and $a \in \mathcal{L}(\mathcal{H}_K)$ such that $La = 0$. If $a = \sum_{i=1}^k a_i$ where a_i are pairwise dissimilar hypergeometric terms then*

$$La_i = 0, \quad \textit{for} \quad i = 1, 2, \ldots, k\,.$$

PROOF. By Proposition 5.1, for each i there exists a rational sequence r_i such that $La_i = r_i a_i$. Therefore

$$0 = La = \sum_{i=1}^k La_i = \sum_{i=1}^k r_i a_i\,.$$

Since the a_i are pairwise dissimilar Theorem 5.1 implies that $r_i = 0$, for all i. $\square$

COROLLARY 5.2. *Let L be a linear recurrence operator with polynomial coefficients. Then the space $KerL \cap \mathcal{L}(\mathcal{H}_K)$ has a basis in $\mathcal{H}_K$.*

PROOF. Let $a \in \mathcal{L}(\mathcal{H}_K)$ satisfy $La = 0$. By Proposition 5.2, we can write $a = \sum_{i=1}^k a_i$ where a_i are pairwise dissimilar hypergeometric terms. By Corollary 5.1, each a_i satisfies $La_i = 0$. It follows that hypergeometric solutions of (1.2) span the space of solutions from $\mathcal{L}(\mathcal{H}_K)$. To obtain a basis for $\mathrm{Ker}L \cap \mathcal{L}(\mathcal{H}_K)$, select a maximal linearly independent set of hypergeometric solutions of (1.2). $\square$

PROPOSITION 5.3. *Let r, h, h_i, C, C_i be non-zero sequences from $\mathcal{S}_K$ such that*

$$\frac{h(n+1)}{h(n)} = r(n)\frac{C(n+1)}{C(n)}\,, \tag{5.3}$$

$$\frac{h_i(n+1)}{h_i(n)} = r(n)\frac{C_i(n+1)}{C_i(n)}\,, \quad \textit{for} \;\; i = 1, 2, \ldots, k\,, \tag{5.4}$$

and C is a K-linear combination of C_i. Then h is a K-linear combination of h_i.

PROOF. From (5.3) and (5.4) it follows that there are constants $\alpha, \alpha_i \in K$ and $j_0 \in \mathbb{N}$ such that

$$h(n) = \alpha \left(\prod_{j=j_0}^{n-1} r(j) \right) C(n)$$

and

$$h_i(n) = \alpha_i \left(\prod_{j=j_0}^{n-1} r(j) \right) C_i(n)\,, \quad \text{for} \;\; i = 1, 2, \ldots, k\,.$$

Let $C = \sum_{i=1}^{k} \lambda_i C_i$ where $\lambda_i \in K$. Then

$$h(n) = \sum_{i=1}^{k} \alpha_i \left(\prod_{j=j_0}^{n-1} r(j) \right) \lambda_i C_i(n) = \sum_{i=1}^{k} \alpha \frac{\lambda_i}{\alpha_i} h_i(n) \, .$$

□

Modify Hyper so that it will not stop after finding one hypergeometric solution but will rather test all possible choices of $A(n)$, $B(n)$, and Z. Return all hypergeometric solutions h which satisfy $h(n+1) = S(n)h(n)$ where $S(n) = Z\frac{A(n)}{B(n)}\frac{C(n+1)}{C(n)}$ and C is a basic polynomial solution found by Poly on the corresponding auxiliary recurrence. From Theorem 4.1, Corollary 5.2, and Proposition 5.3 it follows that this modified algorithm will produce a generating set for the space of solutions which belong to $\mathcal{L}(\mathcal{H}_K)$. In general, this will not be a basis because the same hypergeometric solution may be obtained from several auxiliary recurrences. One can prune this set until it becomes a basis, testing for linear dependence by means of the Casoratian determinant.

Another way to obtain such a basis is to find one hypergeometric solution with Hyper, then reduce the order of the recurrence, recursively find the corresponding basis for the reduced recurrence, and use Gosper's algorithm to put the antidifferences of these solutions into closed form if possible. This method will actually yield a larger class of solutions (cf. Examples 4.2 and 4.5).

5.2. Non-homogeneous recurrences

In this subsection we show how to solve (1.3) over $\mathcal{L}(\mathcal{H}_K)$ when $f \neq 0$.

Proposition 5.4. *Up to the order of the terms, the representation of sequences from $\mathcal{L}(\mathcal{H}_K)$ as sums of pairwise dissimilar hypergeometric terms is unique.*

Proof. Assume that $a_1, a_2, \ldots, a_k$ and $b_1, b_2, \ldots, b_m$ are pairwise dissimilar hypergeometric terms with

$$\sum_{i=1}^{k} a_i = \sum_{j=1}^{m} b_j \, . \tag{5.5}$$

Using induction on $k+m$ we prove that $k = m$ and that each a_i equals some b_j. If $k+m = 0$ this holds trivially. Let $k+m > 0$. Then by Theorem 5.1 it follows that $k > 0$, $m > 0$, and some a_i is similar to some b_j. Relabel the terms so that a_k is similar to b_m, and let $h := a_k - b_m$. If $h \neq 0$ then we can use induction hypothesis both on $\sum_{i=1}^{k-1} a_i + h = \sum_{j=1}^{m-1} b_j$ and $\sum_{i=1}^{k-1} a_i = \sum_{j=1}^{m-1} b_j - h$, to find that $k = m-1$ and $k-1 = m$. This contradiction shows that $h = 0$, so $a_k = b_m$ and $\sum_{i=1}^{k-1} a_i = \sum_{j=1}^{m-1} b_j$. By induction hypothesis, $k = m$ and each a_i with $1 \leq i \leq k-1$ equals some b_j with $1 \leq j \leq m-1$. □

If $a \in \mathcal{L}(\mathcal{H}_K)$ and $La = f$ then $f \in \mathcal{L}(\mathcal{H}_K)$, by Proposition 5.1. Let $a = \sum_{j=1}^{m} a_j$ and $f = \sum_{j=1}^{k} f_j$ where a_j and f_j are pairwise dissimilar hypergeometric terms. Without loss of generality assume that there is an $l \leq m$ such that $La_j \neq 0$ if and only if $j \leq l$. Then by Proposition 5.4, $l = k$ and we can relabel the f_j so that

$$La_j = f_j\,, \quad \text{for} \ \ j = 1, 2, \ldots, k\,. \tag{5.6}$$

By Proposition 5.1, there are non-zero rational sequences r_j such that $a_j = r_j f_j$, for $j = 1, 2, \ldots, k$. Let $s_j := Ef_j/f_j$. With L as in (1.1), it follows from (5.6) that r_j satisfies

$$L_j r_j = 1\,, \quad \text{for} \ \ j = 1, 2, \ldots, k \tag{5.7}$$

where

$$L_j = \sum_{i=0}^{d} p_i \left(\prod_{l=0}^{i-1} E^l s_j \right) E^i\,, \quad \text{for} \ \ j = 1, 2, \ldots, k\,.$$

This gives the following algorithm for solving (1.3) over $\mathcal{L}(\mathcal{H}_K)$:

1. Write $f = \sum_{j=1}^{k} f_j$ where f_j are pairwise dissimilar hypergeometric terms.
2. For $j = 1, 2, \ldots, k$, find a non-zero rational solution r_j of (5.7) using the algorithm of Abramov (1989b). If none exists for some j then (1.3) has no solution in $\mathcal{L}(\mathcal{H}_K)$.
3. Use Hyper to find a basis $a_1, a_2, \ldots, a_m$ for the space $\mathrm{Ker} L \cap \mathcal{L}(\mathcal{H}_K)$.
4. Return $\sum_{j=1}^{m} C_j a_j + \sum_{j=1}^{k} r_j f_j$ where C_j are arbitrary constants.

Note that step 2 could also be carried out by means of Hyper, since any solution of a non-homogeneous recurrence with hypergeometric right-hand side satisfies a homogeneous recurrence of one-higher order. However, we prefer Abramov's algorithm for this task in order to avoid exhaustive search for suitable monic factors of the two sentinel coefficients of the recurrence.

It remains to explain how to perform step 1. We need to group together similar hypergeometric terms, so we need to decide if a given hypergeometric term is rational. The following lemma gives an algorithm for this.

Lemma 5.1. *Let $h \in \mathcal{H}_K$ and $r := Eh/h$. Let*

$$r(n) = Z \frac{A(n)}{B(n)} \frac{C(n+1)}{C(n)}$$

and

$$\frac{B(n)}{A(n)} = \frac{a(n)}{b(n)} \frac{c(n+1)}{c(n)} \tag{5.8}$$

be the normal forms for $r(n)$ and $B(n)/A(n)$, as described in Lemma 3.1. Then $h \in \mathcal{R}_K$ if and only if $Z = a(n) = b(n) = 1$.

Proof. If $Z = a(n) = b(n) = 1$ then $r(n) = c(n)C(n+1)/(c(n+1)C(n))$, so $h(n) = \alpha C(n)/c(n)$ where $\alpha \in K$ is some constant. Hence h is rational.

Conversely, assume that $h(n) = p(n)/q(n)$ where p, q are relatively prime polynomials. Let

$$\frac{h(n+1)}{h(n)} = \frac{q(n)}{q(n+1)} \frac{p(n+1)}{p(n)} = Z \frac{A(n)}{B(n)} \frac{C(n+1)}{C(n)} \tag{5.9}$$

as in Lemma 3.1. Obviously, $Z = 1$. By Lemma 3.2, $p(n)$ divides $C(n)$. Write $C(n) =$

$p(n)s(n)$ where $s(n)$ is a polynomial. Then from (5.9)

$$\frac{B(n)}{A(n)} = \frac{q(n+1)}{q(n)} \frac{s(n+1)}{s(n)} .$$

By Corollary 3.1, normal form (5.8) is unique. So $c(n) = q(n)s(n)$ and $a(n) = b(n) = 1$. □

5.3. Computing in algebraic extensions

Observe that all computation in algorithm Poly takes place in the coefficient field K_0 of the recurrence. Therefore the basic polynomial solutions found by Poly constitute a basis for the space of polynomial solutions over any extension field K of K_0.

On the other hand, algorithm Hyper used on (1.2) requires factorization of the leading and trailing coefficients of (1.2), as well as of polynomials encountered in (4.7). This factorization has to be performed in the field K over which we seek hypergeometric solutions.

Example 5.1. The Fibonacci recurrence

$$a(n+2) - a(n+1) - a(n) = 0 \tag{5.10}$$

has no hypergeometric solution over Q but it has two basic hypergeometric solutions, $(\frac{1+\sqrt{5}}{2})^n$ and $(\frac{1-\sqrt{5}}{2})^n$, over $Q(\sqrt{5})$. This is the splitting field of polynomial (4.7) of the recurrence, namely $Z^2 - Z - 1$.

Example 5.2. Similarly, the recurrence

$$a(n+2) - (2n+1)a(n+1) + (n^2-5)a(n) = 0$$

has no hypergeometric solution over Q while it has two basic hypergeometric solutions, $(+\sqrt{5})^{\overline{n}}$ and $(-\sqrt{5})^{\overline{n}}$, over $Q(\sqrt{5})$. This time, $Q(\sqrt{5})$ is the splitting field of the trailing coefficient of the recurrence, $n^2 - 5$.

The proof of correctness of Hyper shows that if (1.2) has a hypergeometric solution in some extension K of the coefficient field K_0, then a constant multiple of it belongs to the splitting field of the sentinel coefficients of (1.2) and of polynomials (4.7). Thus computing in the algebraic closure of K_0 will produce all hypergeometric solutions of recurrences over K_0.

However, suppose that a is a sequence over K_0. The next theorem shows that to determine if a is hypergeometric over some extension K of K_0 it is sufficient to answer this question for $K = K_0$.

Theorem 5.2. *Let K_0 be a field of characteristic zero and K an extension field of K_0. Let a be a sequence over K_0. If a is rational (respectively hypergeometric) over K then a is rational (respectively hypergeometric) over K_0.*

Proof. Let $a(n) = r(n)$ for $n \geq n_0$ where r is rational over K. Let $r = p/q$ where p and q are polynomials over K of respective degrees k and m, and q is monic. Then

$$p(n) - a(n)q(n) = 0, \quad \text{for} \quad n = n_0, n_0+1, \ldots, n_0+k+m \tag{5.11}$$

is a system of linear equations over K_0 for the coefficients of p and q. Since (5.11) has a non-trivial solution in K, its determinant is zero. It follows that (5.11) has a non-trivial solution in K_0. The number of equations in (5.11) exceeds the degrees of both p or q, therefore any non-trivial solution must have both p and q non-zero. So there are non-zero polynomials p' and q' over K_0 of respective degrees $k' \leq k$ and $m' \leq m$ such that

$$p'(n)q(n) - p(n)q'(n) = 0, \quad \text{for} \quad n = n_0, n_0 + 1, \ldots, n_0 + k + m\,.$$

The left-hand side is a polynomial of degree not exceeding $k + m$ which has $k + m + 1$ distinct zeros. It follows that $p'q = pq'$, so $r = p/q = p'/q'$ is rational over K_0.

If a is hypergeometric over K then Ea/a is rational over K. By the preceding paragraph, Ea/a is rational over K_0, hence a is hypergeometric over K_0. $\square$

For example, to prove that the sequence of Fibonacci numbers is not hypergeometric over any field of characteristic zero, it suffices to show that (5.10) has no hypergeometric solution over Q.

It is interesting to observe that in recurrences arising in combinatorics the leading and trailing coefficients often factor completely over Q; cf. Examples 4.3, 4.4, and 4.5.

Acknowledgements

This work is part of author's Ph.D. Thesis, Petkovšek (1990). The author wishes to thank his advisor, Dana Scott, as well as Herb Wilf for their invaluable encouragement and support. He is also indebted to Bruce Sagan who corrected several errors in a preliminary version, and to István Nemes for many fruitful discussions.

References

Abramov, S.A. (1989a). Problems in computer algebra that are connected with a search for polynomial solutions of linear differential and difference equations. *Moscow Univ. Comput. Math. Cybernet.* no. 3, 63–68. Transl. from *Vestn. Moskov. univ. Ser. XV. Vychisl. mat. kibernet.* no. 3, 56–60.

Abramov, S.A. (1989b). Rational solutions of linear differential and difference equations with polynomial coefficients. *U.S.S.R. Comput. Maths. Math. Phys.* **29** 7–12. Transl. from *Zh. vychisl. mat. mat. fiz.* **29** 1611–1620.

Bender, C.M., Orszag, S.A. (1978). *Advanced Mathematical Methods for Scientists and Engineers*. New York: McGraw-Hill.

Cohn, R.M. (1965). *Difference Algebra*. New York: Interscience Publishers.

Comtet, L. (1974). *Advanced Combinatorics: The Art of Finite and Infinite Expansions*. Dordrecht-Holland: D. Reidel Publ. Co. Transl. of *Analyse Combinatoire, Tomes I et II*, Presses Universitaires de France, Paris 1970.

Gessel, I.M., Sagan, B.E., Yeh, Y.-N. (1991). Enumeration of trees by inversions. Submitted.

Gosper, R.W., Jr. (1978). Decision procedure for indefinite hypergeometric summation. *Proc. Natl. Acad. Sci. USA* **75**, 40–42.

Petkovšek, M. (1991). *Finding Closed-Form Solutions of Difference Equations by Symbolic Methods*. Ph.D. Thesis. Pittsburgh PA: Carnegie Mellon University, CMU-CS-91-103.

van der Poorten, A. (1979). A proof that Euler missed... Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer* **1**, 195–203.

Stanley, R.P. (1980). Differentiably finite power series. *European J. Combin.* **1**, 175–188.

Wilf, H.S., Zeilberger, D. (1992). An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities. To appear.

Wimp, J., Zeilberger, D. (1985). Resurrecting the asymptotics of linear recurrences. *J. Math. Anal. Appl.* **111**, 162–176.

Zeilberger, D. (1990). A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.* **32**, 321–368.

Zeilberger, D. (1991). The method of creative telescoping. *J. Symb. Comput.* **11**, 195–204.