Order-Degree Curves for Hypergeometric Creative Telescoping

Shaoshi Chen*
Department of Mathematics
NCSU
Raleigh, NC 27695, USA
schen21@ncsu.edu

Manuel Kauers†
RISC
Johannes Kepler University
4040 Linz, Austria
mkauers@risc.jku.at

ABSTRACT

Creative telescoping applied to a bivariate proper hypergeometric term produces linear recurrence operators with polynomial coefficients, called telescopers. We provide bounds for the degrees of the polynomials appearing in these operators. Our bounds are expressed as curves in the (r, d)-plane which assign to every order r a bound on the degree d of the telescopers. These curves are hyperbolas, which reflect the phenomenon that higher order telescopers tend to have lower degree, and vice versa.

Categories and Subject Descriptors

I.1.2 [**Computing Methodologies**]: Symbolic and Algebraic Manipulation—*Algorithms*

General Terms

Algorithms

Keywords

Symbolic Summation, Creative Telescoping, Degree Bounds

1. INTRODUCTION

We consider the problem of finding linear recurrence equations with polynomial coefficients satisfied by a given definite single sum over a proper hypergeometric term in two variables. This is one of the classical problems in symbolic summation. Zeilberger [17] showed that such a recurrence always exists, and proposed the algorithm now named after him for computing one [16, 18]. Also explicit bounds are known for the order of the recurrence satisfied by a given sum [15, 10, 4]. Little is known however about the degrees of the polynomials appearing in the recurrence. These are investigated in the present paper.

*Supported by the NSF grant CCF-1017217.

†Supported by the FWF grant Y464-N18.

Ideally, we would like to be able to determine for a given sum all the pairs (r, d) such that the sum satisfies a linear recurrence of order r with polynomial coefficients of degree at most d. This is a hard question which we do not expect to have a simple answer. The results given below can be viewed as answers to simplified variants of the problem. One simplification is that we restrict the attention to the recurrences found by creative telescoping [18], called "*telescopers*" in the symbolic summation community (see Section 2 below for a definition). The second simplification is that instead of trying to characterize all the pairs (r, d), we confine ourselves to sufficient conditions.

Our main results are thus formulas which provide bounds on the degree d of the polynomial coefficients in a telescoper, depending on its order r. The formulas describe curves in the (r, d)-plane with the property that for every integer point (r, d) above the curve, there is a telescoper of order r with polynomial coefficients of degree at most d. As the curves are hyperbolas, they reflect the phenomenon that higher order recurrence equations may have lower degree coefficients. This feature can be used to derive a complexity estimate according to which, at least in theory, computing the minimum order recurrence is more expensive than computing a recurrence with slightly higher order (but drastically smaller polynomial coefficients). This phenomenon is analogous to the situation in the differential case, which was first analyzed by Bostan et al. [5] for algebraic functions, and recently for integrals of hyperexponential terms by the authors [6].

Our analysis for non-rational proper hypergeometric input (Section 3) follows closely our analysis for the differential case [6]. It turns out that the summation case considered here is slightly easier than the differential case in that it requires fewer cases to distinguish and in that the resulting degree estimation formula is much simpler than its differential analogue. For rational input (Section 4), we derive a degree estimation formula following Le's algorithm for computing telescopers of rational functions [3, 9].

2. PROPER HYPERGEOMETRIC TERMS AND CREATIVE TELESCOPING

Let $\mathbb{K}$ be a field of characteristic zero and let $\mathbb{K}(n, k)$ be the field of rational functions in n and k. We will be considering extension fields E of $\mathbb{K}(n, k)$ on which two isomorphisms S_n and S_k are defined which commute with each other, leave every element of $\mathbb{K}$ fixed, and act on n and k via $S_n(n) = n + 1$, $S_k(n) = n$, $S_n(k) = k$, $S_k(k) =$

$k+1$. A *hypergeometric term* is an element h of such an extension field E with $S_n(h)/h \in \mathbb{K}(n,k)$ and $S_k(h)/h \in \mathbb{K}(n,k)$. Proper hypergeometric terms are hypergeometric terms which can be written in the form

$$h = p\,x^n y^k \prod_{m=1}^{M} \frac{\Gamma(a_m n + a'_m k + a''_m)\Gamma(b_m n - b'_m k + b''_m)}{\Gamma(u_m n + u'_m k + u''_m)\Gamma(v_m n - v'_m k + v''_m)}, \tag{1}$$

where $p \in \mathbb{K}[n,k]$, $x, y \in \mathbb{K}$, $M \in \mathbb{N}$ is fixed, a_m, a'_m, b_m, b'_m, u_m, u'_m, v_m, v'_m are nonnegative integers, a''_m, b''_m, u''_m, $v''_m \in \mathbb{K}$ and the expressions x^n, y^k, and $\Gamma(\dots)$ refer to elements of E on which S_n and S_k act as suggested by the notation, e.g.,

$$S_n(\Gamma(2n-k+1)) = (2n-k+1)(2n-k+2)\Gamma(2n-k+1),$$
$$S_k(\Gamma(2n-k+1)) = \frac{1}{2n-k}\Gamma(2n-k+1).$$

Throughout the paper the symbols p, x, y, M, a_m, a'_m, a''_m, ... will be used with the meaning they have in (1). We also use the abbreviations

$$A_m := a_m n + a'_m k + a''_m, \quad B_m := b_m n - b'_m k + b''_m,$$
$$U_m := u_m n + u'_m k + u''_m, \quad V_m := v_m n - v'_m k + v''_m.$$

The hidden assumption that the number of Γ-terms is the same for all four kinds is without loss of generality, since we can always put further terms $\Gamma(0n+0k+1)$ if necessary.

We follow the paradigm of creative telescoping. For a given hypergeometric term h as above, we want to determine polynomials $\ell_0, \dots, \ell_r \in \mathbb{K}[n]$ (free of k, not all zero), and a rational function $C \in \mathbb{K}(n,k)$ (possibly involving k, possibly zero), such that

$$\ell_0 h + \ell_1 S_n(h) + \dots + \ell_r S_n^r(h) = S_k(Ch) - Ch.$$

In this case, the operator $L := \ell_0 + \ell_1 S_n + \dots + \ell_r S_n^r \in \mathbb{K}[n][S_n]$ is called a *telescoper* for h, and the rational function $C \in \mathbb{K}(n,k)$ is called a *certificate* for L (and h). The number r is called the *order* of L, and $d := \max_{i=0}^r \deg_n \ell_i$ is called its *degree*. If h represents an actual sequence $f(n,k)$, then a recurrence for the definite sum $\sum_{k=-\infty}^{\infty} f(n,k)$ can be obtained from such a pair (L,C) as explained in the literature on symbolic summation [12]. We shall not embark on the technical subtleties of this correspondence here but restrict ourselves to analyzing of the set of all pairs (r,d) for which there exists a telescoper of order r and degree d.

The following notation will be used.

- For $p \in \mathbb{K}[n,k]$ and $m \in \mathbb{N}$, let
$$p^{\overline{m}} := p(p+1)(p+2)\cdots(p+m-1)$$
with the conventions $p^{\overline{0}} = 1$ and $p^{\overline{1}} = p$.
- For $p \in \mathbb{K}[n,k]$, $\deg_n p$ and $\deg_k p$ denote the degree of p with respect to n or k, respectively. $\deg p$ without any subscript denotes the total degree of p.
- For $z \in \mathbb{R}$, let $z^+ := \max\{0, z\}$.

With this notation, we have

$$\frac{S_n(h)}{h} = x\frac{S_n(p)}{p} \prod_{m=1}^{M} \frac{A_m^{\overline{a_m}} B_m^{\overline{b_m}}}{U_m^{\overline{u_m}} V_m^{\overline{v_m}}},$$
$$\frac{S_k(h)}{h} = y\frac{S_k(p)}{p} \prod_{m=1}^{M} \frac{A_m^{\overline{a'_m}} (V_m - v'_m)^{\overline{v'_m}}}{U_m^{\overline{u'_m}} (B_m - b'_m)^{\overline{b'_m}}}.$$

3. THE NON-RATIONAL CASE

We consider in this section the case where h cannot be split into $h = qh_0$ for $q \in \mathbb{K}(n,k)$ and another hypergeometric term h_0 with $S_k(h_0)/h_0 = 1$. Informally, this means that we exclude terms h where $y = 1$ and every Γ-term involving k can be cancelled against another one to some rational function. Those terms are treated separately in Section 4 below. If h cannot be split as indicated, then also Ch cannot be split in this way, for any nonzero rational function $C \in \mathbb{K}(n,k)$. In particular, we then have $S_k(Ch)/(Ch) \neq 1$, hence $S_k(Ch) \neq Ch$, hence $S_k(Ch) - Ch \neq 0$, and hence whenever (L,C) is such that $L(h) = S_k(Ch) - Ch$, we can be sure that L is not the zero operator, and we need not worry about this requirement any further.

The analysis in the present case is similar to the one of Apagodu and Zeilberger [10], who used it for deriving a bound on the order r of L, and similar to our analysis [6] of the differential case. The main idea is to follow step by step the execution of Zeilberger's algorithm when applied to h. This leads to a homogeneous linear system of equations with coefficients in $\mathbb{K}(n)$ which will have a solution whenever it is underdetermined. The condition of having more variables than equations in this linear system is the source of the estimate for choices (r,d) that lead to a solution.

3.1 Zeilberger's Algorithm

Recall the steps of Zeilberger's algorithm: for some choice of r, it makes an ansatz $L = \ell_0 + \ell_1 S_n + \dots + \ell_r S_n^r$ with undetermined coefficients ℓ_i, and then calls Gosper's algorithm on $L(h)$. Gosper's algorithm [7] proceeds by writing

$$\frac{S_k(L(h))}{L(h)} = \frac{S_k(P)}{P}\frac{Q}{S_k(R)}$$

for some polynomials P, Q, R such that $\gcd(Q, S_k^i(R)) = 1$ for all $i \in \mathbb{N}$. It turns out that the undetermined coefficients $\ell_0, \dots, \ell_r$ appear linearly in P and not at all in Q or R. Next, the algorithm searches for a polynomial solution Y of the *Gosper equation*

$$P = Q\,S_k(Y) - RY$$

by making an ansatz $Y = y_0 + y_1 k + y_2 k^2 + \dots + y_s k^s$ for some suitably chosen degree s, substituting the ansatz into the equation, and comparing powers of k on both sides. This leads to a linear system in the variables $\ell_0, \dots, \ell_r, y_0, \dots, y_s$ with coefficients in $\mathbb{K}(n)$. Any solution of this system gives rise to a telescoper L with the corresponding certificate $C = RY/P$. If no solution exists, the procedure is repeated with a greater value of r.

For a hypergeometric term h and an operator $L = \ell_0 + \ell_1 S_n + \dots + \ell_r S_n^r$, we have

$$L(h) = \sum_{i=0}^{r} \ell_i x^i \frac{S_n^i(p)}{p} \prod_{m=1}^{M} \frac{A_m^{\overline{ia_m}} B_m^{\overline{ib_m}}}{U_m^{\overline{iu_m}} V_m^{\overline{iv_m}}} h$$
$$= \frac{\sum_{i=0}^{r} \ell_i x^i S_n^i(p) \prod_{m=1}^{M} P_{i,m}}{p \prod_{m=1}^{M} U_m^{\overline{ru_m}} V_m^{\overline{rv_m}}} h$$
$$= \left(\sum_{i=0}^{r} \ell_i x^i S_n^i(p) \prod_{m=1}^{M} P_{i,m}\right)$$

$$\times\, x^n y^k \prod_{m=1}^{M} \frac{\Gamma(A_m)\Gamma(B_m)}{\Gamma(U_m + r u_m)\Gamma(V_m + r v_m)},$$

where

$$P_{i,m} = A_m^{\overline{i a_m}} B_m^{\overline{i b_m}} (U_m + i u_m)^{\overline{(r-i) u_m}} (V_m + i v_m)^{\overline{(r-i) v_m}}.$$

We can write

$$\frac{S_k(L(h))}{L(h)} = \frac{S_k(P)}{P} \frac{Q}{S_k(R)},$$

where

$$P = \sum_{i=0}^{r} \ell_i x^i S_n^i(p) \prod_{m=1}^{M} P_{i,m},$$

$$Q = y \prod_{m=1}^{M} A_m^{\overline{a'_m}} (V_m + r v_m - v'_m)^{\overline{v'_m}},$$

$$R = \prod_{m=1}^{M} (U_m + r u_m - u'_m)^{\overline{u'_m}} B_m^{\overline{b'_m}}.$$

Depending on the actual values of the coefficients appearing in h, this decomposition may or may not satisfy the requirement $\gcd(Q, S_k^i(R)) = 1$ for all $i \in \mathbb{N}$. But even if it does not, it only means that we may overlook some solutions, but every solution we find still gives rise to a correct telescoper and certificate. Since we are interested only in bounding the size of the telescopers of h, it is sufficient to study under which circumstances the Gosper equation

$$P = Q\, S_k(Y) - R\, Y$$

with the above choice of P, Q, R has a solution.

3.2 Counting Variables and Equations

Apagodu and Zeilberger [10] proceed from here by analyzing the linear system over $\mathbb{K}(n)$ resulting from the Gosper equation for a suitable choice of the degree of Y. They derive a bound on r but give no information on the degree d. General bounds for the degrees of solutions of linear systems with polynomial coefficients could be applied, but as these degree bounds grow with the size of the matrix, it seems difficult to capture the phenomenon that increasing r may allow for decreasing d using such general bounds.

We proceed differently. Instead of a coefficient comparison with respect to powers of k leading to a linear system over $\mathbb{K}(n)$, we consider a coefficient comparison with respect to powers of n and k leading to a linear system over $\mathbb{K}$. This requires us to make a choice not only for the degree of Y in k but also for the degree of Y in n as well as for the degrees of the ℓ_i $(i = 0, \dots, r)$ in n. For expressing the number of variables and equations in this system, it is helpful to adopt the following definition.

Definition 1. *For a proper hypergeometric term h as in (1), let*

$$\delta = \deg p,$$

$$\vartheta = \max\Big\{\sum_{m=1}^{M}(a_m + b_m), \sum_{m=1}^{M}(u_m + v_m)\Big\},$$

$$\lambda = \sum_{m=1}^{M}(u_m + v_m),$$

$$\mu = \sum_{m=1}^{M}(a_m + b_m - u_m - v_m),$$

$$\nu = \max\Big\{\sum_{m=1}^{M}(a'_m + v'_m), \sum_{m=1}^{M}(u'_m + b'_m)\Big\}.$$

Note that these parameters are integers which only depend on h but not on r or d. Except for μ, they are all nonnegative. Note also that we have $\lambda + \mu \geq 0$ and $\vartheta = \lambda + \mu^+ \geq |\mu|$.

Lemma 2. *Let $d_i := \deg_n \ell_i$ $(i = 0, \dots, r)$. Then*

$$\deg P \leq \delta + \lambda r + \max_{i=0}^{r}(d_i + i\mu).$$

Furthermore, $\deg_k P \leq \delta + \vartheta r$.

Proof. It suffices to observe that

$$\deg P_{i,m} \leq i a_m + i b_m + (r-i) u_m + (r-i) v_m$$

for all $m = 1, \dots, M$ and all $i = 0, \dots, r$. For the degree with respect to k, observe also that $\deg_k \ell_i = 0$ for all i. ∎

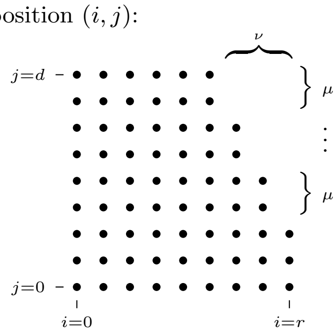
We have some freedom in choosing the d_i. The choice influences the number of variables in the ansatz

$$L = \sum_{i=0}^{r} \sum_{j=0}^{d_i} \ell_{i,j} n^j S_n^i$$

as well as the number of equations. We prefer to have many variables and few equations. For a fixed target degree d, the maximum possible number of variables is $(d+1)(r+1)$ by choosing $d_0 = d_1 = \dots = d_r = d$. But this choice also leads to many equations. A better balance between number of variables and number of equations is obtained by lowering some of the d_i with indices close to zero (if μ is negative) or with indices close to r (if μ is positive). Specifically, we choose

$$d_i := d - \begin{cases} (\nu + i - r)^+ |\mu| & \text{if } \mu \geq 0 \\ (\nu - i)^+ |\mu| & \text{if } \mu < 0. \end{cases}$$

The motivation for this choice goes back to Verbaeten [14]. See [6, Ex. 11, Ex. 15.5 and the remarks after Thm. 14] for an explanation of the corresponding choice in the differential case. The support of the ansatz for L looks as in the following diagram, where every term $n^j S_n^i$ is represented by a bullet at position (i, j):



With this choice for the degrees d_i, the number of variables in the ansatz for L is

$$\sum_{i=0}^{r}(d_i + 1) = (d+1)(r+1) - \tfrac{1}{2}|\mu|\nu(\nu+1),$$

provided that $d \geq |\mu|\nu$. The number of resulting equations is as follows.

Lemma 3. *If the d_i are chosen as above, then P contains at most*

$$\tfrac{1}{2}\big(\delta + \vartheta r + 1\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2\big)$$

terms $n^i k^j$.

Proof. If $\mu \geq 0$, we have

$$d_i + i\mu = d - (\nu + i - r)^+\mu + i\mu \leq d - \nu|\mu| + r\mu$$

for all $i = 0, \dots, r$. Likewise, when $\mu < 0$, we have

$$d_i + i\mu = d - (\nu - i)^+|\mu| + i\mu \leq d - \nu|\mu|$$

for all $i = 0, \dots, r$. Together with Lemma 2, it follows that

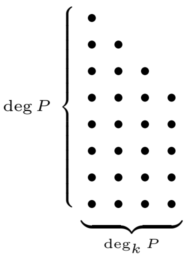
$$\deg P \leq \delta + (\lambda + \mu^+)r + d - \nu|\mu| = \delta + \vartheta r + d - \nu|\mu|$$

regardless of the sign of μ. We also have $\deg_k P \leq \delta + \vartheta r$ from Lemma 2. For the number of terms $n^i k^j$ in P we have

$$\sum_{i=0}^{\deg_k P} (1 + \deg P - i) = \tfrac{1}{2}(\deg_k P + 1)(2\deg P + 2 - \deg_k P).$$

Plugging the estimates for $\deg P$ and $\deg_k P$ into the right hand side gives the expression claimed in the Lemma. ∎

The support of P has a trapezoidal shape which is determined by the total degree and the degree with respect to k:



The next step is to choose the degrees for Y in n and k. This is done in such a way that $Q\,S_k(Y) - RY$ only contains terms which are already expected to occur in P, so that no additional equations will appear.

Lemma 4. *Let the d_i be chosen as before and suppose that $Y \in \mathbb{K}[n,k]$ is such that $\deg Y \leq \deg P - \nu$ and $\deg_k Y \leq \deg_k P - \nu$. Then $P - (Q\,S_k(Y) - RY)$ contains at most*

$$\tfrac{1}{2}\big(\delta + \vartheta r + 1\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2\big)$$

terms $n^i k^j$.

Proof. As for Lemma 3, using also $\max\{\deg Q, \deg R\} = \max\{\deg_k Q, \deg_k R\} = \nu$. ∎

Lemma 4 suggests the ansatz

$$Y = \sum_{i=0}^{s_1} \sum_{j=0}^{s_2 - i} y_{i,j} k^i n^j$$

with $s_1 = \deg_k P - \nu$ and $s_2 = \deg P - \nu$, which provides us with

$$\tfrac{1}{2}\big(\delta + \vartheta r + 1 - \nu\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2 - \nu\big)$$

variables. We are now ready to formulate the main result of this section. Note that the inequality for d is a considerably simpler formula than the corresponding result in the differential case (Thm. 14 in [6]).

Theorem 5. *Let h be a proper hypergeometric term which cannot be written $h = qh_0$ for some $q \in \mathbb{K}(n,k)$ and a hypergeometric term h_0 with $S_k(h_0)/h_0 = 1$. Let $\delta, \lambda, \mu, \nu$ be as in Definition 1, let $r \geq \nu$ and*

$$d > \frac{(\vartheta\nu - 1)r + \frac{1}{2}\nu\big(2\delta + |\mu| + 3 - (1 + |\mu|)\nu\big) - 1}{r - \nu + 1}.$$

Then there exists a telescoper L for h of order at most r and degree at most d.

Proof. A sufficient condition for the existence of a telescoper of order r and degree d is that for some particular ansatz, the equation

$$P = Q\,S_k(Y) - RY$$

has a nontrivial solution. A sufficient condition for the existence of a solution is that the linear system resulting from coefficient comparison has more variables than equations. For all d in question, we have $d > \vartheta\nu \geq |\mu|\nu$. Therefore, with the ansatz described above, we have

$$(d+1)(r+1) - \tfrac{1}{2}|\mu|\nu(\nu+1)$$

variables $\ell_{i,j}$ in P,

$$\tfrac{1}{2}\big(\delta + \vartheta r + 1 - \nu\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2 - \nu\big)$$

variables $y_{i,j}$ in Y, and

$$\tfrac{1}{2}\big(\delta + \vartheta r + 1\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2\big)$$

equations. Solving the inequality

$$\begin{aligned}&(d+1)(r+1) - \tfrac{1}{2}|\mu|\nu(\nu+1)\\ &\quad + \tfrac{1}{2}\big(\delta + \vartheta r + 1 - \nu\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2 - \nu\big)\\ &> \tfrac{1}{2}\big(\delta + \vartheta r + 1\big)\big(\delta + 2d + \vartheta r - 2|\mu|\nu + 2\big)\end{aligned}$$

under the assumption $r \geq \nu$ for d gives the claimed degree estimate. ∎

3.3 Examples and Consequences

Example 6. *1. For*

$$h = (n^2 + k^2 + 1)\frac{\Gamma(2n + 3k)}{\Gamma(2n - k)}$$

we have $\delta = 2$, $\vartheta = 2$, $\mu = 0$, $\nu = 4$. Theorem 5 predicts a telescoper of order r and degree d whenever $r \geq 4$ and

$$d > \frac{7r + 5}{r - 3}.$$

The left figure below shows the curve defined by the right hand side (black) together with the region of all points (r,d) for which we found telescopers of h with order r and degree d by direct calculation (gray). In this example, the estimate overshoots by very little only.

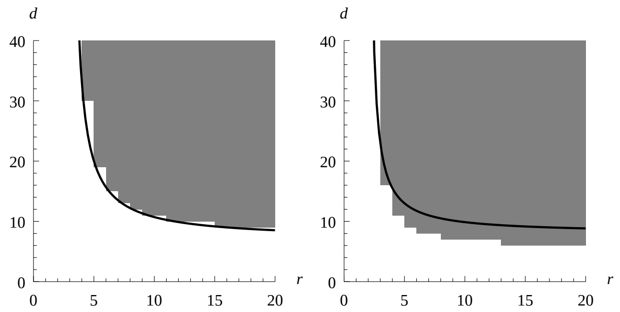
2. The corresponding picture for

$$h = \frac{\Gamma(2n + k)\Gamma(n - k + 2)}{\Gamma(2n - k)\Gamma(n + 2k)}$$

is shown below on the right. Here, $\delta = 0$, $\vartheta = 3$, $\mu = 0$, $\nu = 3$ and Theorem 5 predicts a telescoper of order r and degree d whenever $r \geq 3$ and

$$d > \frac{8r - 1}{r - 2}.$$

In this example, the estimate is less tight.



The gray region consists of all the points (r, d) for which the ansatz as described in the proof above results in a linear system (over $\mathbb{K}$) which has a nontrivial solution. The points (r, d) in the portion of the gray region which is below the black curve represent telescopers where this linear system is overdetermined but, for some strange reason, nevertheless nontrivially solvable. The small portions of white space which lie above the curves are not in contradiction with our theorem because they do not contain any points with integer coordinates. (The theorem says that every point $(r, d) \in \mathbb{Z}^2$ above the curve belongs to the gray region.)

Theorem 5 supplements the bound given in [10] on the order of telescopers for a hypergeometric term by an estimate for the degree that these operators may have. In addition, it provides lower degree bounds for higher orders and admits a bound on the least possible degree for a telescoper.

Corollary 7. *With the notation of Theorem 5, h admits a telescoper of order $r = \nu$ and degree*

$$d = \left\lceil \tfrac{1}{2}\nu(2\delta + 2\nu\vartheta + |\mu| - \nu|\mu|)\right\rceil$$

as well as a telescoper of order

$$r = \left\lceil \tfrac{1}{2}\nu(1 + 2\delta + 2(\nu - 1)(\vartheta - |\mu|))\right\rceil$$

and degree $d = \vartheta\nu$.

Proof. Immediate by checking that the two choices for r and d satisfy the conditions stated in Theorem 5. ∎

An accurate prediction for the degrees of the telescopers can also be used for improving the efficiency of creative telescoping algorithms. Although most implementations today compute the telescoper with minimum order, it may be less costly to compute a telescoper of slightly higher order. If we know in advance the degrees d of the telescopers for every order r, we can select before the computation the order r which minimizes the computational cost. Of course, the cost depends on the algorithm which is used. It is not necessary (and not advisable) to follow the steps in the derivation of Theorem 5 and do a coefficient comparison over $\mathbb{K}$. Instead, it is more efficient to follow the common practice [8] of comparing coefficients only with respect to powers of k and solve a linear system over $\mathbb{K}(n)$. For nonminimal choices of r, this system will have a nullspace of dimension greater than one, of which we need not compute a complete basis, but only a single vector with components of low degree. There are algorithms known for computing such a vector using $\mathrm{O}^\sim(m^3 t)$ field operations when the system has at most m variables and equations and the solution has degree at most t [13]. In the situation at hand, we have $m = (r+1)+(\delta+\vartheta r+1)$ variables and a solution of degree $t = \delta + \vartheta r + d - (|\mu| + 1)\nu + 1$.

Therefore, in order to compute a telescoper and its certificate most efficiently, we should minimize the cost function

$$C(r, d) := \big((\vartheta + 1)r + \delta + 2\big)^3 \big(\delta + \vartheta r + d - (|\mu| + 1)\nu + 1\big).$$

According to the following theorem, for asymptotically large input it is significantly better to choose r slightly larger than the minimal possible value.

Theorem 8. *Let h and λ, μ, ν be as in Theorem 5, $\tau \geq \max\{\vartheta, \nu\}$, let κ be an increasing sublinear function with the property that degree t solutions of a linear system with m variables and at most m equations over $\mathbb{K}(n)$ can be computed with $\kappa(\max\{m, t\})m^3 t$ operations in $\mathbb{K}$. Then:*

1. *A telescoper of order $r = \tau$ along with a corresponding certificate can be computed using*
$$\kappa(\tau^3)\tau^9 + \tfrac{1}{2}(7 - |\mu|)\kappa(\tau^3)\tau^8 + \mathrm{O}^\sim(\tau^7)$$
operations in $\mathbb{K}$.
2. *If $\alpha > 1$ is some constant and r is chosen such that $r = \alpha\tau + \mathrm{O}(1)$, then a telescoper of order r and a corresponding certificate can be computed using*
$$\frac{\alpha^5}{\alpha - 1}\kappa(\tau^3)\tau^8 + \mathrm{O}^\sim(\tau^7)$$
operations in $\mathbb{K}$.

Proof. According to Theorem 5, for every $r \geq \tau$ there exists a telescoper of order r and degree d for any

$$d > f(r) := \frac{(\tau^2 - 1)r + \mathrm{O}(\tau^2)}{r - \tau - 1}.$$

By assumption, such a telescoper can be computed using no more than

$$C(r, d) = \kappa\big((\tau + 1)r + \delta + 2\big)^3 \big(\delta + \tau r + d - (|\mu| + 1)\tau + 1\big)$$

operations in $\mathbb{K}$. The claim now follows from the asymptotic expansions of $C(\tau, f(\tau)+1)$ and $C(\alpha\tau, f(\alpha\tau)+1)$ for $\tau \to \infty$, respectively. ∎

The leading coefficient in part 2 is minimized for $\alpha = 5/4$. This suggests that when ϑ and ν are large and approximately equal, the order of the cheapest telescoper is about 20% larger than the minimal expected order.

4. THE RATIONAL CASE

This case is considered separately for two reasons. First, because this allowed us to easily escape from the difficulty of ensuring a nonzero telescoper in the previous section. Secondly, because the bound we obtain below for the rational case is much sharper than the bound we would obtain by treating the rational case as in the previous section with a modification like in [6] to handle the problem of avoiding a trivial telescoper.

We assume now that h can be written as $h = qh_0$ for some hypergeometric term h_0 with $S_k(h_0)/h_0 = 1$. By the following transformation, we may assume without loss of generality $h_0 = 1$.

Lemma 9. *Let h be a hypergeometric term and suppose that $h = qh_0$ for some $q \in \mathbb{K}(n, k)$ and a hypergeometric term h_0 with $S_k(h_0)/h_0 = 1$. Let $a, b \in \mathbb{K}[n, k]$ be such that $S_n(h_0)/h_0 = a/b$. Let L be a telescoper for q of order r and degree d. Then there exists a telescoper for h of order r and degree at most $d + r\max\{\deg_n a, \deg_n b\}$.*

Proof. Write $L = \ell_0 + \ell_1 S_n + \cdots + \ell_r S_n^r$ and let $C \in \mathbb{K}(n,k)$ be a certificate for L and q, so $L(q) = S_k(Cq) - Cq$. For $i = 0, \dots, r$, let

$$\tilde{\ell}_i := \ell_i \frac{b}{a} S_n\Big(\frac{b}{a}\Big) \cdots S_n^{i-1}\Big(\frac{b}{a}\Big)$$

and $\tilde{L} := \tilde{\ell}_0 + \tilde{\ell}_1 S_n + \cdots + \tilde{\ell}_r S_n^r$. Then

$$\tilde{L}(qh_0) = L(q)h_0 = (S_k(Cq) - Cq)h_0 = S_k(Cqh_0) - Cqh_0.$$

Because of

$$S_k\Big(\frac{a}{b}\Big) = \frac{S_k(S_n(h_0))}{S_k(h_0)} = \frac{S_n(S_k(h_0))}{S_k(h_0)} = \frac{S_n(h_0)}{h_0} = \frac{a}{b},$$

the operator $\tilde{L}$ is free of k. Thus, after clearing denominators, $\tilde{L}$ is a telescoper for h with coefficients of degree at most $d + r \max\{\deg_n a, \deg_n b\}$. ∎

From now on, we assume that h is at the same time a proper hypergeometric term and a rational function, or equivalently, that h is a rational function whose denominator factors into integer-linear factors. Le [9] gives a precise description of the structure of telescopers in this case, and he proposes an algorithm different from Zeilberger's for computing them. Our degree estimate is derived following the steps of his algorithm, so we start by briefly summarizing the main steps of Le's approach.

4.1 Le's Algorithm

Given a rational proper hypergeometric term h, Le's algorithm computes a telescoper L for h as follows.

1. Compute $g \in \mathbb{K}(n,k)$ and polynomials $p, q \in \mathbb{K}[n,k]$ with $\gcd(q, S_k^i(q)) = 1$ for all $i \in \mathbb{Z} \setminus \{0\}$ such that
$$h = S_k(g) - g + \frac{p}{q}.$$
Then an operator L is a telescoper for h if and only if L is a telescoper for $\frac{p}{q}$. Abramov [1, 2] and Paule [11] explain how to compute such a decomposition.

2. Compute a polynomial $u \in \mathbb{K}[n]$, operators $V_1, \dots, V_s$ in $\mathbb{K}[n][S_n]$, and rational functions $f_1, \dots, f_s$ of the form $f_i = (a_i n + a_i' k + a_i'')^{-e_i}$ $(i = 1, \dots, s)$ such that
$$\frac{p}{q} = \frac{1}{u} \sum_{i=0}^{s} V_i(f_i).$$
Such data always exists according to Lemma 5 in [9] in combination with the assumption $\gcd(q, S_k^i(q)) = 1$ $(i \in \mathbb{Z} \setminus \{0\})$. It can be further assumed that the f_i are chosen such that $a_i' > 0$, $e_i > 0$, $\gcd(a_i, a_i') = 1$ for all i, and
$$\Big(\frac{a_i}{a_i'} - \frac{a_j}{a_j'}\Big)n + \Big(\frac{a_i''}{a_i'} - \frac{a_j''}{a_j'}\Big) \notin \mathbb{Z}$$
for all $i \neq j$ with $e_i = e_j$.

3. For $i = 1, \dots, s$, compute an operator $L_i \in \mathbb{K}(n)[S_n]$ such that $S_n^{a_i'} - 1$ is a right divisor of $L_i(\frac{1}{u} V_i)$. It follows from Le's Lemma 4 that the operators L_i with this property are precisely the telescopers of the rational functions $V_i(f_i)$.

4. Compute a common left multiple $L \in \mathbb{K}[n][S_n]$ of the operators $L_1, \dots, L_s$. Then L is a telescoper for h.

The main part of the computational work happens in the last two steps. It therefore appears sensible to assume in the following degree analysis that we already know the data $u, V_1, \dots, V_s, f_1, \dots, f_s$ computed in step 2, and to express the degree bounds in terms of their degrees and coefficients rather than in terms of the degrees of numerator and denominator of h, say.

4.2 Counting Variables and Equations

Also in the present case, the degree estimate is obtained by balancing the number of variables and equations of a certain linear system over $\mathbb{K}$. The linear system we consider originates from a particular way of executing steps 3 and 4 of the algorithm outlined above.

Theorem 10. *Let $u \in \mathbb{K}[n]$ and let $V_1, \dots, V_s \in \mathbb{K}[n][S_n]$ be operators of degree δ_i $(i = 1 \dots, s)$. Let $f_i = (a_i n + a_i' k + a_i'')^{-e_i}$ for some $a_i'' \in \mathbb{K}$, $a_i, a_i' \in \mathbb{Z}$ with $a_i' > 0$ and $\gcd(a_i, a_i') = 1$, $e_i > 0$, suppose*

$$\Big(\frac{a_i}{a_i'} - \frac{a_j}{a_j'}\Big)n + \Big(\frac{a_i''}{a_i'} - \frac{a_j''}{a_j'}\Big) \notin \mathbb{Z}$$

for all $i \neq j$ with $e_i = e_j$. Let $h = \frac{1}{u}\sum_{i=1}^s V_i(f_i)$. Then for every $r \geq \sum_{i=1}^s a_i'$ and every

$$d > \frac{-r - 1 + \sum_{i=1}^{s} a_i' \delta_i}{r + 1 - \sum_{i=1}^{s} a_i'} + \deg_n u$$

there exists a telescoper L for h of order at most r and degree at most d.

Proof. According to Le's algorithm, it suffices to find some $L \in \mathbb{K}[n][S_n]$ and operators $R_i \in \mathbb{K}(n)[S_n]$ with the property that $L(\frac{1}{u} V_i) = R_i(S_n^{a_i'} - 1)$ for all i.

Denote by ρ_i the order of V_i. Writing $\tilde{d} := d - \deg_n u$, we make an ansatz $L = \tilde{L} u$ with

$$\tilde{L} = \sum_{i=0}^{r} \sum_{j=0}^{\tilde{d}} \ell_{i,j} n^j S_n^i$$

so that L has degree d and $L \frac{1}{u} V_i = \tilde{L} V_i$ $(i = 1, \dots, s)$. It thus remains to construct operators $R_i \in \mathbb{K}[n][S_n]$ with $\tilde{L} V_i = R_i(S_n^{a_i'} - 1)$. Since $L\tilde{V}_i$ has order $r + \rho_i$ and degree $\tilde{d} + \delta_i$, we consider ansatzes for the R_i of order $r + \rho_i - a_i'$ and degree $\tilde{d} + \delta_i$, respectively, because $S_n^{a_i'} - 1$ has order a_i' and degree 0. Then we have altogether

$$(r+1)(\tilde{d}+1) + \sum_{i=1}^{s} (r + \rho_i - a_i' + 1)(\tilde{d} + \delta_i + 1)$$

variables in $\tilde{L}$ and the R_i, and comparing coefficients with respect to n and S_n in all the required identities $\tilde{L} V_i = R_i(S_n^{a_i'} - 1)$ leads to a linear system with

$$\sum_{i=1}^{s} (r + \rho_i + 1)(\tilde{d} + \delta_i + 1)$$

equations. This system will have a nontrivial solution whenever the number of variables exceeds the number of equations. Under the assumption $r \geq \sum_{i=1}^{s} a'_i$, the inequality

$$(r+1)(\tilde{d}+1) + \sum_{i=1}^{s}(r+\rho_i - a'_i + 1)(\tilde{d}+\delta_i+1)$$
$$> \sum_{i=1}^{s}(r+\rho_i+1)(\tilde{d}+\delta_i+1)$$

is equivalent to

$$\tilde{d} > \frac{-r-1+\sum_{i=1}^{s} a'_i\delta_i}{r+1-\sum_{i=1}^{s} a'_i}.$$

This completes the proof. ∎

4.3 Examples and Consequences

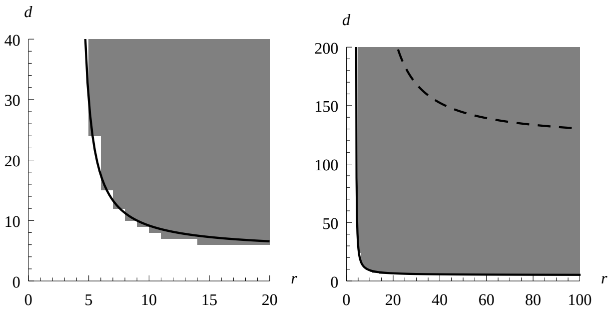
Example 11. *1. The rational function*

$$h = \frac{(2n-3k)(3n-2k)^2}{(n+k+2)(n+2k+1)(2n+k+1)(3n+k+1)}$$

can be written in the form $h = \frac{1}{u}\sum_{i=1}^{4} V_i(f_i)$ *where* $u = (n-1)n(n+3)(2n-1)(3n+1)(5n+1)$, *the* f_i *are such that* $a'_1 = a'_2 = a'_3 = 1$, $a'_4 = 2$, *and the* V_i *are such that* $\delta_1 = \cdots = \delta_4 = 6$. *Therefore, Theorem 10 predicts a telescoper of order* r *and degree* d *whenever* $r \geq 5$ *and*

$$d > \frac{29-r}{r-4} + 6.$$

This curve together with the region of all points (r,d) *for which a telescoper of order* r *and degree* d *exists is shown solid in the figure below. The dashed curve on the right hand side indicates the curve which would be obtained using the reasoning from Section 3.*



2. The corresponding picture for the rational function

$$h = \frac{(n-k+1)^2(2n-3k+5)}{(n+k+3)(n+k+5)(n+2k+1)(2n+k+1)^2}$$

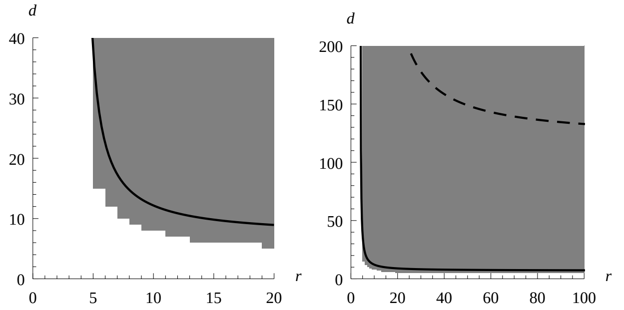
is shown in the figure below on the right. This input can be written in the form

$$h = S_k(g) - g + \frac{1}{u}\sum_{i=1}^{4} V_i(f_i)$$

with $g = \frac{10(n+3)^2(n+4)(2n+2k+7)}{(n-4)^2(n+9)(n+k+3)(n+k+4)}$, $u = (3n+1)^2(n-4)^2(n-2)^2(n+5)(n+9)$, *the* f_i *such that* $a'_1 = a'_2 = a'_3 = 1$, $a'_4 = 2$, *and the* V_i *such that* $\delta_1 = 8$, $\delta_2 = \delta_3 = \delta_4 = 7$. *According to Theorem 10, we therefore expect a telescoper for* h *of order* r *and degree* d *whenever* $r \geq 5$ *and*

$$d > \frac{35-r}{r-4} + 8.$$

In this example, the estimate is not as tight as in the previous one, but still much better than the approach from Section 3.



Again, it is an easy matter to specialize the general degree bound to a degree estimate for a low order telescoper, or to an order estimate for a low degree telescoper.

Corollary 12. *With the notation of Theorem 10,* h *admits a telescoper of order* $r = \sum_{i=1}^{s} a'_i$ *and degree* $d = \deg_n u + \sum_{i=1}^{s}(\delta_i - 1)a'_i$ *as well as a telescoper of order* $r = \sum_{i=1}^{s} a_i\delta_i$ *and degree* $d = \deg_n u$.

Proof. Clear by checking that the proposed choices for r and d are consistent with the bounds in Theorem 10. ∎

Also as in the non-rational case, the bounds for the degrees of the telescopers can be used for deriving bounds on the computational complexity. Let us assume for simplicity that the cost of steps 1 and 2 of Le's algorithm is negligible, or equivalently, that the input h is of the form $\frac{1}{u}\sum_{i=0}^{s} V_i(f_i)$ with $V_i \in \mathbb{K}[n][S_n]$. We shall analyze the algorithm which carries out steps 3 and 4 of Section 4.1 in one stroke by making an ansatz over $\mathbb{K}(n)$ for an operator $L = \ell_0 + \ell_1 S_n + \cdots + \ell_r S_n^r$, computing the right reminders of $L\frac{1}{u}V_i$ with respect to $S_n^{a'_i} - 1$ and equating their coefficients to zero. We assume, as before, that the resulting linear system is solved using an algorithm whose runtime is linear in the output degree and cubic in the matrix size. Then the algorithm requires $\mathrm{O}^{\sim}(r^3 d)$ operations in $\mathbb{K}$.

Theorem 13. *Let* $u \in \mathbb{K}[n]$, $V_1, \ldots, V_s \in \mathbb{K}[n][S_n]$, *and* $f_1, \ldots, f_s \in \mathbb{K}(n,k)$ *be as in Theorem 10 and consider* $h = \frac{1}{u}\sum_{i=1}^{s} V_i(f_i)$. *Suppose that* κ *is an increasing sublinear function such that degree* t *solutions of a linear system with* m *variables and at most* m *equations over* $\mathbb{K}(n)$ *can be computed with* $\kappa(\max\{m,t\})m^3 t$ *operations in* $\mathbb{K}$. *Assume* $\delta_1 = \cdots = \delta_s =: \delta > 0$ *and* $a'_1 = a'_2 = \cdots = a'_s =: a'$ *are fixed. Then:*

1. *A telescoper of order* $r = a's$ *can be computed using*

$$a'^4(\delta-1)\kappa(a's\delta)s^4 + \mathrm{O}^{\sim}(s^3)$$

operations in $\mathbb{K}$.

2. *If* $\alpha > 1$ *is some constant and* r *is chosen such that* $r = \alpha a's + \mathrm{O}(1)$ *then a telescoper of order* r *can be computed using*

$$\frac{\alpha^3}{\alpha-1}a'^3(\delta - 1 + (\alpha-1)\deg_n u)\kappa(a's\delta)s^3 + \mathrm{O}^{\sim}(s^2)$$

operations in $\mathbb{K}$.

Proof. According to the first estimate stated in Theorem 10, for every $r \geq a's$ there exists a telescoper of order r and degree d for any

$$d > f(r) := \frac{sa'\delta - r - 1}{r + 1 - sa'} + \deg_n u.$$

By assumption, such a telescoper can be computed using no more than $C(r,d) := \kappa r^3 d$ operations in $\mathbb{K}$. The claim now follows from the asymptotic expansions of $C(a's, f(a's)+1)$ and $C(\alpha a's, f(\alpha a's)+1)$ for $s \to \infty$, respectively. ∎

When $\deg_n u = 0$, the leading coefficient in part 2 is minimized for $\alpha = 3/2$. This suggests that when s is large and all the δ_i, and a_i' are approximately equal, the order of the cheapest operator exceeds the minimal expected order by around 50%.

It must not be concluded from a literal comparison of the exponents in Theorems 5 and 10 that Le's algorithm is faster than Zeilberger's, because τ in Theorem 5 and s in Theorem 10 measure the size of the input differently. Nevertheless, it is plausible to expect that Le's algorithm is faster, because it finds the telescopers without also computing a (potentially big) corresponding certificate. Our main point here is not a comparison of the two approaches, but rather the observation that both of them admit a degree analysis which fits to the general paradigm that increasing the order can cause a degree drop which is significant enough to leave a trace in the computational complexity.

It can also be argued that the situations considered in Theorems 8 and 13 are chosen somewhat arbitrarily (ϑ and ν growing while μ remains fixed; resp. s growing while all the δ_i and a_i' remain fixed). Indeed, it would be wrong to take these theorems as an advice which telescopers are most easily computed for a particular input at hand. Instead, in order to speed up an actual implementation, one should let the program calculate the optimal choice for r from the degree estimates given Theorems 5 and 10 with the particular parameters of the input.

Unfortunately, we are not able to illustrate the speedup obtained in this way by an actual runtime comparison for a concrete example, because for examples which can be handled on currently available hardware, the computational cost turns out to be minimized for the least order operator. But already for examples which are only slightly beyond the capacity of current machines, the degree predictions in Theorems 5 and 10 indicate that computing the telescoper of order one more than minimal will start to give an advantage. We therefore expect that the results presented in this paper will contribute to the improvement of creative telescoping implementations in the very near future.

5. REFERENCES

[1] S. A. Abramov. The rational component of the solution of a first-order linear recurrence relation with rational right hand side. *U.S.S.R. Comput. Math. Math. Phys.*, 15(4):216–221, 1975.

[2] S. A. Abramov. Indefinite sums of rational functions. In *Proc. ISSAC'95*, pp 303–308, 1995.

[3] S. A. Abramov and H. Q. Le. A criterion for the applicability of Zeilberger's algorithm to rational functions. *Discr. Math.*, 259(1–3):1–17, 2002.

[4] A. Bostan, S. Chen, F. Chyzak, and Z. Li. Complexity of creative telescoping for bivariate rational functions. In *Proc. ISSAC'10*, pp 203–210, 2010.

[5] A. Bostan, F. Chyzak, B. Salvy, G. Lecerf, and É. Schost. Differential equations for algebraic functions. In *Proc. ISSAC'07*, pp 25–32, 2007.

[6] S. Chen and M. Kauers. Trading order for degree in creative telescoping. *J. Symbol. Comput.* 47(8):968–995, 2012.

[7] W. Gosper. Decision procedure for indefinite hypergeometric summation. *Proc. Nation. Acad. Sci.*, 75:40–42, 1978.

[8] C. Koutschan. A fast approach to creative telescoping. *Math. Comput. Sci.*, 4(2–3):259–266, 2010.

[9] H. Q. Le. A direct algorithm to construct the minimal Z-pairs for rational functions. *Adv. Appl. Math.*, 30(1–2):137–159, 2003.

[10] M. Mohammed and D. Zeilberger. Sharp upper bounds for the orders of the recurrences output by the Zeilberger and q-Zeilberger algorithms. *J. Symbol. Comput.*, 39(2):201–207, 2005.

[11] P. Paule. Greatest factorial factorization and symbolic summation. *J. Symbol. Comput.*, 20:235–268, 1995.

[12] M. Petkovšek, H. Wilf, and D. Zeilberger. *A = B*. AK Peters, Ltd., 1997.

[13] A. Storjohann and G. Villard. Computing the rank and a small nullspace basis of a polynomial matrix. In *Proc. ISSAC'05*, pages 309–316, 2005.

[14] P. Verbaeten. The automatic construction of pure recurrence relations. *ACM Sigsam Bulletin*, 8, 1974.

[15] H. S. Wilf and D. Zeilberger. An algorithmic proof theory for hypergeometric (ordinary and q) multisum/integral identities. *Invent. Math.*, 108:575–633, 1992.

[16] D. Zeilberger. A fast algorithm for proving terminating hypergeometric identities. *Discr. Math.*, 80:207–211, 1990.

[17] D. Zeilberger. A holonomic systems approach to special function identities. *J. Comput. Appl. Math.*, 32:321–368, 1990.

[18] D. Zeilberger. The method of creative telescoping. *J. Symbol. Comput.*, 11:195–204, 1991.