

On the length of integers in telescopers for proper hypergeometric terms

Manuel Kauers*

Research Institute for Symbolic Computation
Johannes Kepler University
Linz, Austria

Lily Yen

Capilano University and
Simon Fraser University
Vancouver, Canada

November 18, 2013

Abstract

We show that the number of digits in the integers of a creative telescoping relation of expected minimal order for a bivariate proper hypergeometric term has essentially cubic growth with the problem size. For telescopers of higher order but lower degree we obtain a quintic bound. Experiments suggest that these bounds are tight. As applications of our results, we give an improved bound on the maximal possible integer root of the leading coefficient of a telescoper, and the first discussion of the bit complexity of creative telescoping.

1 Introduction

Creative telescoping is a backbone of symbolic summation. It permits the construction of recurrence equations for definite sums. In its classical version,

*Supported by FWF project Y464-N18.

it is applied to sums whose summands are hypergeometric terms. This situation was intensively studied during the 1990s (see Petkovšek et al. (1997) and the references given there for an overview on the classical results). While during the first decade of this century most research in the area focussed on generalizing creative telescoping to sums whose summands are more complicated (see, for instance, the survey articles of Koutschan (2013) and Schneider (2013) and the references given there), the hypergeometric case is recently getting back into the focus. There is now a general interest in getting a better understanding of the sizes of the output of summation algorithms, and of the amount of time spent on the computation. This paper contributes to these investigations. We work out bounds for the length of the integers that may appear in the output of creative telescoping algorithms, complementing earlier results given for the order and the degree of creative telescoping relations. As corollaries of our bounds, we obtain a new bound on the maximal integer root of the leading coefficient as well as a first bound on the bit complexity of creative telescoping.

Throughout this article, we consider a proper hypergeometric term

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

where $p \in \mathbb{Z}[n, k]$, $M \in \mathbb{N}$ is fixed, $x, y, a_m, a'_m, a''_m, b_m, b'_m, b''_m, u_m, u'_m, u''_m, v_m, v'_m, v''_m$ are fixed nonnegative integers, and n and k are variables. To avoid discussion of degenerate cases, we assume throughout that h is not a rational function. The assumption that there are exactly M Gamma-terms for each of the four types is without loss of generality, because we can always add further terms $\Gamma(0n + 0k + 0)$ without changing h .

A creative telescoping relation for h is a pair (L, C) , where $L = \ell_0 + \ell_1 S_n + \dots + \ell_r S_n^r \in \mathbb{Z}[n][S_n] \setminus \{0\}$ is a nonzero recurrence operator in n , free of k , and $C \in \mathbb{Q}(n, k)$ a bivariate rational function in n and k (which may well be zero), with the property

$$L(h) = (S_k - 1)(Ch).$$

The symbols S_n and S_k refer to the usual shift operators $n \rightsquigarrow n+1, k \rightsquigarrow k+1$, respectively. The operator L is called a *telescoper* for h , and C is called a *certificate* for L and h .

If h has finite support, i.e., $\lim_{|k| \rightarrow \infty} h(n, k) = 0$ for every fixed $n \in \mathbb{N}$, then a telescoper L annihilates the definite hypergeometric sum $H(n) :=$

$\sum_k h(n, k)$. If not, a creative telescoping relation still gives rise to an inhomogeneous recurrence for finite definite sums such as $\sum_{k=0}^n h(n, k)$ or $\sum_{k=n}^{2n} h(n, k)$. See Petkovšek et al. (1997) for details.

The classical Zeilberger algorithm (Zeilberger, 1990, 1991; Petkovšek et al., 1997) finds a creative telescoping relation for any given proper hypergeometric term h . This algorithm is based on Gosper's algorithm (Gosper, 1978) for indefinite hypergeometric summation and delivers a creative telescoping relation (L, C) for which the order r of L is minimal. An alternative algorithm proposed by Apagodu and Zeilberger (2005) does not use Gosper's algorithm during the computation but only in its correctness proof. The key observation behind this algorithm is that $L = \ell_0 + \ell_1 S_n + \cdots + \ell_r S_n^r \in \mathbb{Q}[n][S_n]$ is a telescoper for h if there exists some polynomial $Y \in \mathbb{Q}[n, k]$ with the property

$$\ell_0 P_0 + \cdots + \ell_r P_r = Q S_k(Y) - R Y, \quad (2)$$

where

$$P_i = x^i S_n^i(p) \prod_{m=1}^M \left((a_m n + a'_m k + a''_m)^{\overline{ia_m}} (b_m n - b'_m k + b''_m)^{\overline{ib_m}} \right. \\ \left. \times (u_m n + u'_m k + u''_m + i u_m)^{\overline{(r-i)u_m}} (v_m n - v'_m k + v''_m + i v_m)^{\overline{(r-i)v_m}} \right) \\ (i = 0, \dots, r),$$

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

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

For a justification, see the article of Apagodu and Zeilberger (2005). The following definition contains certain quantities in terms of which bounds on the size of the telescoper of h can be formulated.

Definition 1. For a proper hypergeometric term h as above, define

$$\nu = \max \left\{ \sum_{m=1}^M (a'_m + v'_m), \sum_{m=1}^M (u'_m + b'_m) \right\}, \quad \delta = \deg(p), \\ \vartheta = \max \left\{ \sum_{m=1}^M (a_m + b_m), \sum_{m=1}^M (u_m + v_m) \right\}, \quad \lambda = \sum_{m=1}^M (u_m + v_m),$$

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

Furthermore, we let

$$\Omega := \max_{m=1}^M \max\{|a_m|, |a'_m|, |a''_m|, |b_m|, |b'_m|, |b''_m|, |u_m|, |u'_m|, |u''_m|, |v_m|, |v'_m|, |v''_m|\}$$

be a bound on the integers appearing in the arguments of the Γ terms of h .

Apagodu and Zeilberger show that h admits a telescoper of order r for every $r \geq \nu$, or in other words, that if r is the order of the minimal telescoper, then $r \leq \nu$. Generically this bound is tight. Chen and Kauers (2012) supplement this result with information about the degrees of the coefficients of the telescoper. They show that for every $r \geq \nu$ and every d satisfying

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

there exists a telescoper $L = \ell_0 + \dots + \ell_r S_n^r$ with $\max_{i=0}^r |\ell_i| \leq d$. The purpose of the present article is to refine the analysis one step further by giving bounds on the length of the integers appearing in the coefficients ℓ_i of a telescoper L of h . In Theorem 7 in Section 4, we show that hypergeometric terms h have a telescoper of order $r = \nu$ whose integer coefficients have no more than $O(\Omega^3 \log(\Omega))$ digits. In Theorem 10 in Section 4, we show furthermore that there are telescopers of order $r = O(\Omega)$ and degree $d = O(\Omega^2)$ whose integer coefficients have no more than $O(\Omega^5 \log(\Omega))$ digits. For both estimates, we provide experimental data that indicate that our bounds are sharp.

2 Bounding Tools

In order to bound the integers arising in the coefficients of a telescoper, we need to know by how much the size of the integers can grow during the various steps of the computation. In particular, we need to know how adding, multiplying, and shifting of polynomials may affect the length of their coefficients, and how long the integer coefficients can become in the solution of a system of linear equations with polynomial coefficients. In this section we provide a collection of results in this direction.

The coefficient length of a polynomial depends on the basis with respect to which the polynomial is expressed. We are mostly interested in the coefficient

length with respect to the standard monomial basis $1, x, x^2, x^3, \dots$, but we will also have occasion to use alternative bases. In the following definition we introduce the notational distinction which will be used below.

- Definition 2.** 1. For $p = \sum_{i=0}^d p_i n^i \in \mathbb{Q}[n]$, we call $|p| := |p|_s := \max_{i=0}^d |p_i|$ the (standard) height or the (standard) norm of p .
2. For $p = \sum_{i=0}^d p_i \binom{n}{i} \in \mathbb{Q}[n]$, we call $|p|_b := \max_{i=0}^d |p_i|$ the binomial height or the binomial norm of p .
3. For $p = \sum_{i=0}^d \sum_{j=0}^e p_{i,j} n^i k^j \in \mathbb{Q}[n, k]$, we define $\|p\|_{s,s} := \max_{i=0}^d \max_{j=0}^e |p_{i,j}|$.
4. For $p = \sum_{i=0}^d \sum_{j=0}^e p_{i,j} n^i \binom{k}{j} \in \mathbb{Q}[n, k]$, we define $\|p\|_{s,b} := \max_{i=0}^d \max_{j=0}^e |p_{i,j}|$.

Note that $|\cdot|_s$, $|\cdot|_b$, $\|\cdot\|_{s,s}$, and $\|\cdot\|_{s,b}$ are indeed norms, i.e., they satisfy absolute scalability, triangle inequality, and they are zero only when the argument is zero. The following lemmas give bounds for shifted polynomials, for products of polynomials, and, to begin with, a connection between the standard norm and the binomial norm.

Lemma 3 (Conversion). For all $p \in \mathbb{Q}[n, k]$, we have $\|p\|_{s,b} \leq \deg_k(p)!^2 \|p\|_{s,s}$.

Proof. Recall from Equation (6.10) on page 262 of Graham et al. (1994):

$$k^m = \sum_{i \geq 0} S_2(m, i) k^i = \sum_{i \geq 0} S_2(m, i) i! \frac{k^i}{i!} = \sum_{i \geq 0} S_2(m, i) i! \binom{k}{i},$$

where $S_2(m, i)$ is the Stirling number of the second kind. Write $p = p_0 + p_1 k + \dots + p_d k^d$ with $p_0, \dots, p_d \in \mathbb{Q}[n]$. Then

$$p = \sum_{j=0}^d p_j k^j = \sum_{j=0}^d \left(p_j \sum_{i=0}^j S_2(j, i) i! \binom{k}{i} \right) = \sum_{i=0}^d \left(\sum_{j=i}^d p_j S_2(j, i) i! \right) \binom{k}{i}.$$

Thus, for the binomial height of p , we find

$$\begin{aligned} \|p\|_{s,b} &= \max_{i=0}^d \left| \sum_{j=i}^d p_j S_2(j, i) i! \right| \leq \max_{i=0}^d \sum_{j=i}^d |p_j| S_2(j, i) i! \leq \max_{i=0}^d \sum_{j=i}^d \|p\|_{s,s} S_2(j, i) d! \\ &\leq \|p\|_{s,s} d! \sum_{j=0}^d S_2(d, j) \leq \|p\|_{s,s} d! \mathbf{B}_d \leq \|p\|_{s,s} d!^2, \end{aligned}$$

where \mathbf{B}_d denotes the d th Bell number. □

Lemma 4 (Shift). *For $q \in \mathbb{Q}[n, k]$ and $r \in \mathbb{N}$, we have $\|S_n^r(q)\|_{s,s} \leq (1+r)^{\deg_n(q)} \|q\|_{s,s}$ and $\|S_n^r(q)\|_{s,b} \leq (1+r)^{\deg_n(q)} \|q\|_{s,b}$.*

Proof. For $\|\cdot\|_{s,s}$, this is Lemma 3.4 of Yen (1996). It then also holds for $|\cdot|_s$ and polynomials in $\mathbb{Q}[n] \subseteq \mathbb{Q}[n, k]$. If finally $q = \sum_{i=0}^d q_i \binom{k}{i} \in \mathbb{Q}[n, k]$ for certain $q_i \in \mathbb{Q}[n]$, then $\|S_n^r(q)\|_{s,b} = \max_{i=0}^d |S_n^r(q_i)|_s \leq (1+r)^d \max_{i=0}^d |q_i|_s = (1+r)^d \|q\|_{s,b}$, so it also holds for the norm $\|\cdot\|_{s,b}$. \square

Lemma 5 (Product). 1. *For $p_1, \dots, p_m \in \mathbb{Q}[n]$, we have*

$$\left| \prod_{i=1}^m p_i \right| \leq \left(\max_{i=1}^m \deg(p_i) + 1 \right)^{m-1} \prod_{i=1}^m |p_i|.$$

2. *Let $p_1, p_2, \dots, p_m \in \mathbb{Q}[n, k]$ be polynomials of total degree 1, and $M \in \mathbb{N}$ be such that $\|p_i\|_{s,s} \leq M$ for $i = 1, \dots, m$. Then for every $q \in \mathbb{Q}[n, k]$, we have*

$$\|p_1 p_2 \cdots p_m q\|_{s,b} \leq (2M)^m (\deg_k(q) + 2)^{\overline{m}} \|q\|_{s,b}.$$

Proof. 1. It suffices to prove the case $m = 2$. The general case then follows immediately by induction. To show the case $m = 2$, consider two polynomials $p = \sum_{i=0}^d p_i n^i$ and $q = \sum_{i=0}^e q_i n^i$. The coefficient of n^j in pq is $\sum_{i=0}^{d+e} p_i q_{j-i}$, where we understand coefficients as being zero if their index is out of range. For every j , the sum can have at most $\min\{\deg(p), \deg(q)\} + 1$ nonzero terms, and as each term is bounded by $|p_i q_{j-i}| \leq |p| |q|$, the claim follows.

2. It suffices to prove the case $m = 1$. The general case then follows immediately by induction on m . Consider $p = a + bk + cn \in \mathbb{Q}[n, k]$ and write $q = \sum_{i=0}^d q_i \binom{k}{i}$ with $q_0, \dots, q_d \in \mathbb{Q}[n]$. Observe that

$$(uk + v) \binom{k}{i} = (ui + v) \binom{k}{i} + u(i+1) \binom{k}{i+1}.$$

Therefore

$$\begin{aligned} pq &= (an + bk + c) \sum_{i=0}^d q_i \binom{k}{i} \\ &= \sum_{i=0}^d (bq_i k + (an + c)q_i) \binom{k}{i} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=0}^d (an + bi + c)q_i \binom{k}{i} + b(i+1)q_i \binom{k}{i+1} \\
&= \sum_{i=0}^{d+1} ((an + bi + c)q_i + biq_{i-1}) \binom{k}{i}.
\end{aligned}$$

Because of

$$|(an + bi + c)q_i + biq_{i-1}| \leq 2(i+1)M \max\{|q_i|, |q_{i-1}|\},$$

and $|q_i| \leq \|q\|_{s,b}$ for all i , it follows that

$$\|pq\|_{s,b} \leq \max_{i=0}^{d+1} 2(i+1)M \max\{|q_i|, |q_{i-1}|\} \leq 2(d+2)M \|q\|_{s,b}$$

as claimed. \square

Finally, we need a bound on the length of the integers which may appear in the basis vectors of the nullspace of a matrix with univariate polynomial entries. The result below takes into account that the columns of the matrix may be split into two groups for which different bounds on the degrees and height are known. Although matrices and vectors all have coefficients in $\mathbb{Z}[x]$, all linear algebra notions (rank, kernel, linear independence, etc.) are understood with respect to the ground field $\mathbb{Q}(x)$.

Lemma 6. *Let $A = (A_0, A_1) \in \mathbb{Z}[x]^{n \times (m_0 + m_1)}$ be a matrix of rank ρ . For $i = 0, 1$, let d_i and M_i be bounds on the degrees and standard heights of the entries of $A_i \in \mathbb{Z}[x]^{n \times m_i}$. Assume that A_0 has full rank. Then $\ker A$ has a basis of vectors from $\mathbb{Z}[x]^{m_0 + m_1}$ that are bounded in degree by $(m_0 - 1)d_0 + (\rho - m_0)d_1 + \max\{d_0, d_1\}$ and in height by*

$$\rho! (\max\{d_0, d_1\} + 1)^{\rho-1} M_0^{m_0-1} M_1^{\rho-m_0} \max\{M_0, M_1\}.$$

Proof. By selecting a maximal linearly independent set of rows from A , we may assume without loss of generality that $n = \rho$. Furthermore, because A_0 has full rank, we have $\rho \geq m_0$, and by exchanging columns within A_1 if necessary, we may assume that $A_1 = (W, V)$ where $W \in \mathbb{Z}[x]^{\rho \times (\rho - m_0)}$, $V \in \mathbb{Z}[x]^{\rho \times (m_1 - (\rho - m_0))}$ and $U := (A_0, W) \in \mathbb{Z}[x]^{\rho \times \rho}$ has full rank.

A basis of $\ker A$ is given by the vectors $(v_i, -e_i) \in \mathbb{Q}(x)^{\rho + (m_0 + m_1 - \rho)}$ where $e_i \in \mathbb{Q}(x)^{m_0 + m_1 - \rho}$ is the i th unit vector and $v_i \in \mathbb{Q}(x)^\rho$ is the unique solution

of the inhomogeneous linear system $Uv_i = Ve_i$. The right hand side is of course just the i th column of V . According to Cramer's rule, the j th component of v_i is given by $\frac{\det U'}{\det U}$ where U' is the matrix obtained from U by replacing its j th column by the i th column of V . Multiplying all the basis vectors by $\det U$ gives a basis of polynomial entries with integer coefficients. By Lemma 5.(1), and from the definition of the determinant,

$$\det((a_{i,j}))_{i,j=1}^n = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{i,\pi(i)},$$

the heights of the determinants $\det U'$ corresponding to columns $j \leq m_0$ are bounded by $\rho!(\max\{d_0, d_1\} + 1)^{\rho-1} M_0^{m_0-1} M_1^{\rho-m_0+1}$; and by $\rho!(\max\{d_0, d_1\} + 1)^{\rho-1} M_0^{m_0} M_1^{\rho-m_0}$ for $j > m_0$. Combining both cases gives the claimed bound. The degree estimate follows from the defining formula for the determinant by the same reasoning. \square

3 Bounds for P_0, \dots, P_r, Q , and R

We will obtain our bounds on the height of the telescope by making an ansatz for ℓ_0, \dots, ℓ_r and the coefficients of the polynomial Y in equation (2), comparing coefficients, and applying Lemma 6 to the linear system obtained from comparing coefficients in (2). For doing so, we need to determine the heights and degrees of the polynomials in this equation.

For the degrees, we have $\deg(P_i) \leq \delta + r\vartheta$ and $\deg(Q), \deg(R) \leq \nu$ by Lemmas 2 and 4 of Chen and Kauers (2012), where \deg refers to the total degree.

For the heights, we apply the lemmas of the previous section. Noting that the products over the rising factorials consist of linear factors all of which have heights bounded by $(r+2)\Omega - 1$, it follows that

$$\begin{aligned} \|P_i\|_{s,b} &\leq (2(r+2)\Omega - 2)^{r\lambda+i\mu} (\delta + 2)^{\overline{r\lambda+i\mu}} \|x^i S_n^i(p)\|_{s,b} && \text{by Lemma 5.(2)} \\ &\leq (2(r+2)\Omega - 2)^{\vartheta r} (\delta + 2)^{\overline{\vartheta r}} \|x^i S_n^i(p)\|_{s,b} && \text{because } r\lambda + i\mu \leq \vartheta r \\ &\leq |x|^i (\delta + \vartheta r + 1)! (2(r+2)\Omega - 2)^{\vartheta r} (1+i)^{\deg_n(p)} \|p\|_{s,b} && \text{by Lemma 4} \\ &\leq \|p\|_{s,s} \delta!^2 (1+r)^\delta |x|^r (\delta + \vartheta r + 1)! (2(r+2)\Omega - 2)^{\vartheta r} && \text{by Lemma 3} \end{aligned}$$

for every $i = 0, \dots, r$. Note that the right hand side does not depend on i but only on r and quantities that are determined by the hypergeometric term h .

For $Y_j = \binom{k}{j}$, we have $S_k(Y_j) = \binom{k+1}{j} = \binom{k}{j} + \binom{k}{j+1}$; therefore, $\|S_k(Y_j)\|_{s,b} = \|Y_j\|_{s,b} = 1$. Hence, since also the linear factors in the rising factorials in Q and R are all bounded in height by $(r+2)\Omega - 1$, we obtain, again by using Lemma 5.(2),

$$\begin{aligned} \|Q S_k(Y_j)\|_{s,b} &\leq |y|(2(r+2)\Omega - 2)^{\sum_{m=1}^M (a'_m + v'_m)} (j+2)^{\overline{\sum_{m=1}^M (a'_m + v'_m)}} \|S_k(Y_j)\|_{s,b} \\ &\leq |y|(j + \nu + 1)^\nu (2(r+2)\Omega - 2)^\nu, \end{aligned}$$

and likewise

$$\|R Y_j\|_{s,b} \leq (j + \nu + 1)^\nu (2(r+2)\Omega - 2)^\nu$$

for every $j \in \mathbb{N}$.

4 The minimal telescoper

Choose $r = \nu$ and $s = \delta + (\vartheta - 1)\nu$, and make an ansatz

$$Y = y_0 + y_1 \binom{k}{1} + \cdots + y_s \binom{k}{s}$$

with undetermined coefficients y_0, \dots, y_s . Then, comparing like coefficients of $\binom{k}{j}$ in the equation

$$\ell_0 P_0 + \cdots + \ell_r P_r = Q S_k(Y) - R Y$$

leads to a system of homogeneous linear equations with $(r+1) + (s+1) = \delta + \vartheta\nu + 2$ variables $\ell_0, \dots, \ell_r, y_0, \dots, y_s$ and no more than

$$\max\left\{1 + \max_{i=0}^r \deg_k(P_i), 1 + \deg_k(Q) + s, 1 + \deg_k(R) + s\right\} \leq \delta + \vartheta\nu + 1$$

equations. This system obviously has a nontrivial solution.

The matrix $A \in \mathbb{Z}[n]^{(\delta + \vartheta\nu + 1) \times (\delta + \vartheta\nu + 2)}$ encoding this system has the form $A = (A_L, A_C)$ where $A_L \in \mathbb{Z}[n]^{(\delta + \vartheta\nu + 1) \times (\nu + 1)}$ consists of the columns corresponding to the variables ℓ_j in the telescoper part, and $A_C \in \mathbb{Z}[n]^{(\delta + \vartheta\nu + 1) \times (\delta + (\vartheta - 1)\nu + 1)}$ consists of the columns corresponding to the variables y_j in the certificate part. More precisely, the entry of A_L in row i and column j is the coefficient of $\binom{k}{i-1}$ in P_{j-1} ($i = 1, \dots, \delta + \vartheta\nu + 1$; $j = 1, \dots, \nu + 1$), and the entry of A_C in row i and column j is the coefficient of $\binom{k}{i-1}$ in $Q S_k\left(\binom{k}{j-1}\right) - R\left(\binom{k}{j-1}\right)$ ($i = 1, \dots, \delta + \vartheta\nu + 1$; $j = 1, \dots, \delta + (\vartheta - 1)\nu + 1$).

By the results of the previous section, the entries of A_L have degree at most $\delta + \vartheta\nu$ and height at most $\|p\|_{s,s}\delta!^2(\nu+1)^\delta|x|^\nu(\delta+\vartheta\nu+1)!(2(\nu+2)\Omega-2)^{\vartheta\nu}$, and the entries of A_C have degree at most $\delta + \vartheta\nu$ and height at most $(|y|+1)(\delta+\vartheta\nu+1)^\nu(2(\nu+2)\Omega-2)^\nu$.

We want to determine the height of the polynomials in the solution vectors of A . There are two cases to distinguish. If A_L has full rank, then we can apply Lemma 6 with $A_0 = A_L$, $A_1 = A_C$, $\rho \leq n = \delta + \vartheta\nu + 1$, $m_0 = \nu + 1$, $m_1 = \delta + (\vartheta - 1)\nu + 1$. It implies the existence of a solution $(\ell_0, \dots, \ell_\nu, y_0, \dots, y_{\delta+(\vartheta-1)\nu}) \in \mathbb{Z}[n]^{(\nu+1)+(\delta+(\vartheta-1)\nu+1)}$ with

$$\begin{aligned}
|\ell_i| &\leq (\delta + \vartheta\nu + 1)!(\max\{\delta + \vartheta\nu, \delta + \vartheta\nu\} + 1)^{\delta+\vartheta\nu} \\
&\quad \times \left(\|p\|_{s,s}\delta!^2(\nu+1)^\delta|x|^\nu(\delta+\vartheta\nu+1)!(2(\nu+2)\Omega-2)^{\vartheta\nu} \right)^\nu \\
&\quad \times \left((|y|+1)(\delta+\vartheta\nu+1)^\nu(2(\nu+2)\Omega-2)^\nu \right)^{\delta+\vartheta\nu+1-\nu} \\
&\quad \times \max\left\{ \|p\|_{s,s}\delta!^2(\nu+1)^\delta|x|^\nu(\delta+\vartheta\nu+1)!(2(\nu+2)\Omega-2)^{\vartheta\nu}, \right. \\
&\quad \left. (|y|+1)(\delta+\vartheta\nu+1)^\nu(2(\nu+2)\Omega-2)^\nu \right\} \\
&\leq ((\delta + \vartheta\nu + 1)!)^2(\delta + \vartheta\nu + 1)^{\delta+\vartheta\nu} \\
&\quad \times \left(\|p\|_{s,s}\delta!^2(\nu+1)^\delta|x|^\nu(\delta+\vartheta\nu+1)!(2(\nu+2)\Omega-2)^{\vartheta\nu} \right)^\nu \\
&\quad \times \left((|y|+1)(\delta+\vartheta\nu+1)^\nu(2(\nu+2)\Omega-2)^\nu \right)^{\delta+(\vartheta-1)\nu+1} \\
&\quad \times \|p\|_{s,s}\delta!^2(\nu+1)^\delta(\delta+\vartheta\nu+1)^\nu(2(\nu+2)\Omega-2)^{\vartheta\nu} \max\{|x|^\nu, |y|+1\} \\
&\leq \max\{|x|^\nu, |y|+1\} \|p\|_{s,s}^{\nu+1}(\delta+\vartheta\nu+1)!^{\nu+2}(\nu+1)^{\delta(\nu+1)}(|y|+1)^{\delta+(\vartheta-1)\nu+1} \\
&\quad \times \delta!^{2(\nu+1)}|x|^{\nu^2}(\delta+\vartheta\nu+1)^{\delta+(\vartheta+\delta+2)\nu+(\vartheta-1)\nu^2}(2(\nu+2)\Omega-2)^{(\delta+\vartheta+1)\nu+(2\vartheta-1)\nu^2}
\end{aligned}$$

for $i = 0, \dots, \nu$.

If A_L does not have full rank, then it has itself a nonempty kernel. In this case, if $(\ell_0, \dots, \ell_\nu)$ is a nontrivial kernel element of A_L , then $(\ell_0, \dots, \ell_\nu, 0, \dots, 0)$ is a nontrivial kernel element of $A = (A_L, A_C)$. Therefore, in this case it suffices to estimate the height of the polynomial entries in the kernel of A_L . To this end, we use again Lemma 6, this time taking A_0 to be some nonzero column (w.l.o.g. the first), A_1 the remaining columns, $n = \delta + \vartheta\nu + 1$, $m_0 = 1$,

$m_1 = \nu$, $\rho \leq \nu - 1$. Using for both A_0 and A_1 the degree and height estimates stated above for A_L , we get the bound

$$|\ell_i| \leq (\nu - 1)! (\delta + \vartheta\nu + 1)^{\nu-2} \left(\|p\|_{s,s} \delta!^2 (\nu + 1)^\delta |x|^\nu (\delta + \vartheta\nu + 1)! (2(\nu + 2)\Omega - 2)^{\vartheta\nu} \right)^{\nu-1}$$

for $i = 0, \dots, r$. As this is always less than or equal to the bound obtained before for the case when A_L has full rank, we have completed the proof of the following theorem.

Theorem 7. *Let h be a non-rational proper hypergeometric term as in (1), and let $\delta, \vartheta, \nu, \Omega$ be as in Definition 1. Then there exists a telescoper for h of order $r = \nu$ whose polynomial coefficients are bounded in height by*

$$\begin{aligned} & \max\{|x|^\nu, |y| + 1\} \|p\|_{s,s}^{\nu+1} (\delta + \vartheta\nu + 1)^{\nu+1} (\nu + 1)^{\delta(\nu+1)} (|y| + 1)^{\delta+(\vartheta-1)\nu+1} \\ & \times \delta!^{2(\nu+1)} |x|^{\nu^2} (\delta + \vartheta\nu + 1)^{\delta+(\vartheta+\delta+2)\nu+(\vartheta-1)\nu^2} (2(\nu + 2)\Omega - 2)^{(\delta+\vartheta+1)\nu+(2\vartheta-1)\nu^2}. \end{aligned}$$

- Remarks 8.**
1. *In general, a hypergeometric term h does not have any telescoper of order smaller than ν , so the theorem makes a statement about the integers appearing in the minimal order telescoper of a “generic” hypergeometric term h . For hypergeometric terms which do possess a smaller telescoper, the theorem remains true as it stands, but does not say anything about the size of the integers in the minimal telescoper.*
 2. *Lemma 6 also yields the degree bound $(\delta + \nu\vartheta)(\delta + \nu\vartheta + 1) = O(\Omega^4)$, which is worse than the degree bound $O(\Omega^3)$ given by Chen and Kauers (2012). In the generic case, when the minimal telescoper order is ν , the solution space of the linear system discussed above has dimension 1, so that at least in this case there is a telescoper of degree $O(\Omega^3)$ and height as stated above. We do not know if this also applies to the degenerate case.*
 3. *Considering $\|p\|_{s,s}$, δ , $|x|$, $|y|$, and M as fixed, and noting that ν and ϑ are bounded by $2M\Omega$, the bound of Theorem 7 is equal to $e^{64(M\Omega)^3 \log(\Omega) + O(\Omega^3)}$ as Ω tends to infinity. Combined with the degree bound $O(\Omega^3)$ (when ν is minimal) or $O(\Omega^4)$ (when it’s not), it follows that there is a telescoper of order $r = \nu = O(\Omega)$ of bit size $O(\Omega^7 \log(\Omega))$ or $O(\Omega^8 \log(\Omega))$, respectively.*
 4. *The choice of the binomial basis in the ansatz for Y is motivated by the fact that with respect to this basis the shift does not increase the norm.*

In the standard basis we have $S_k(k^j) = (k+1)^j = \sum_i \binom{j}{i} k^i$, whose standard norm is $\binom{j}{\lfloor j/2 \rfloor} \leq 2^j$. Using this (almost tight) bound in the argument above leads to a suboptimal bound of the form $e^{O(\Omega^4 \log(\Omega))}$. Of course, the choice of the basis with respect to k used in the ansatz for Y does not have any effect on the output telescoper L , which is free of k by construction.

We conclude the section by a family of hypergeometric terms which gives evidence that the bound of Theorem 7 seems to be asymptotically accurate.

Example 9. For $\Omega = 1, 2, 3, \dots$ consider the proper hypergeometric term $h_\Omega = \frac{\Gamma(\Omega k)}{\Gamma(\Omega n - k)}$. We have computed the minimal telescoper L_Ω of h_Ω for $\Omega = 1, \dots, 23$ and determined the length of the integers appearing in them. Let H_Ω be the logarithm of the maximum over the absolute values of all integers appearing in L_Ω . In Figure 1, we plot the normalized values $\frac{H_\Omega}{\Omega^3}$ (bullets, •) against the following least square fits, testing the four hypotheses $H_\Omega = \Theta(\Omega^3 \log(\Omega))$, $\Theta(\Omega^3)$, $\Theta(\Omega^2 \log(\Omega))$, or $\Theta(\Omega^2)$, respectively:

1. $\log(\Omega) \left(1.43 + \frac{3.30}{\Omega} - \frac{1.66}{\Omega^2} \right)$ (solid line, ———)
2. $1 \left(5.06 - \frac{9.22}{\Omega} + \frac{4.23}{\Omega^2} \right)$ (densely dashed, -----)
3. $\frac{\log(\Omega)}{\Omega} \left(34.8 - \frac{167}{\Omega} + \frac{221}{\Omega^2} \right)$ (loosely dashed, - - - -)
4. $\frac{1}{\Omega} \left(58.9 - \frac{182}{\Omega} + \frac{123}{\Omega^2} \right)$ (dotted,)

The best fit is given by the first hypothesis, suggesting that the bound proven above is asymptotically accurate.

5 Nonminimal telescopers

As shown by Chen and Kauers (2012), telescopers of order $r > \nu$ may have much smaller degrees than the (generically) minimal telescoper of order $r = \nu$. More precisely, the arithmetic size, i.e., the number of monomials $n^i S_n^j$ with a nonzero coefficient appearing in a telescoper, which is bounded by

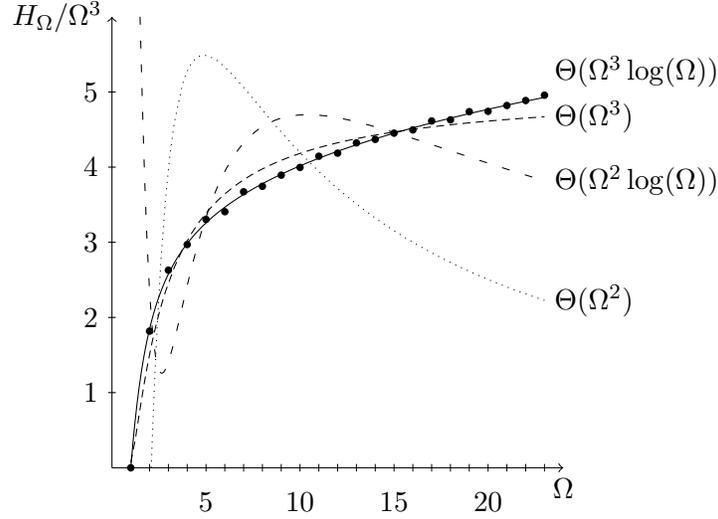


Figure 1: Heights of minimal telescopers

$(r + 1)(d + 1)$, is asymptotically smaller by one order of magnitude when $r = \alpha\nu$ for any fixed constant $\alpha > 1$. It is therefore also interesting to bound the length of the integers appearing in telescopers of nonminimal order.

In this section, we derive such a bound. Following Chen and Kauers (2012), we proceed by analyzing the linear system of equations obtained from the parameterized Gosper equation (2) by comparing coefficients with respect to both n and k . The corresponding matrix is much larger but its entries are integers instead of integer polynomials.

As the resulting bound turns out to be much larger than the bound obtained in the previous section for the height of the telescoper of order ν , we confine ourselves to giving only an asymptotic estimate rather than an exact formula. This makes the expressions in the calculations a little simpler.

Choose $r = 2\nu = O(\Omega)$, $s = \delta + r\vartheta - \nu$, $d = 4\nu\vartheta = O(\Omega^2)$, and make an ansatz

$$L = \sum_{j=0}^r \sum_{i=0}^d \ell_{i,j} n^i S_n^j, \quad Y = \sum_{j=0}^{s+d} \sum_{i=0}^s y_{i,j} n^i \binom{k}{j},$$

with undetermined coefficients $\ell_{i,j}$ and $y_{i,j}$. Then, comparing like coefficients

of $n^i \binom{k}{j}$ in the equation

$$\sum_{j=0}^r \sum_{i=0}^d \ell_{i,j} n^i P_j = Q S_k(Y) - RY$$

leads to a system of homogeneous linear equations with

$$(r+1)(d+1) + (s+d+1)(s+1) = 12\nu^2\vartheta^2 + (12+8\delta)\nu\vartheta + \nu^2 + O(\Omega) = O(\Omega^4)$$

variables $\ell_{i,j}$ and $y_{i,j}$ and no more than

$$\begin{aligned} & \max \left\{ (\delta + r\vartheta + d + 1)(\delta + r\vartheta + 1), (\nu + s + d + 1)(\nu + s + 1) \right\} \\ & = (\delta + r\vartheta + d + 1)(\delta + r\vartheta + 1) = 12\nu^2\vartheta^2 + (8 + 8\delta)\nu\vartheta + O(\Omega) = O(\Omega^4) \end{aligned}$$

equations. As $12 > 8$, this system has a nontrivial solution if $\nu\vartheta \rightarrow \infty$, as $\Omega \rightarrow \infty$.

Let $A = (A_L, A_C)$ be the matrix encoding this linear system, with A_L the submatrix consisting of the columns corresponding to the variables $\ell_{i,j}$ and A_C the part consisting of the columns corresponding to the variables $y_{i,j}$, respectively. As the coefficients of P_i , Q , or R do not change when these polynomials are multiplied by some term n^j (only the exponents change), we can use the same bounds for the heights of the matrix entries as before. Hence A_L is an integer matrix with $(r+1)(d+1) = O(\Omega^3)$ columns and $O(\Omega^4)$ rows whose entries are bounded in absolute value by $e^{O(\Omega^2 \log(\Omega))}$, and A_C is an integer matrix with $O(\Omega^4)$ rows and columns whose entries are bounded in absolute value by $e^{O(\Omega \log(\Omega))}$.

If A_L happens to have full rank, we can apply Lemma 6 to A , interpreting its entries as integer polynomials of degree zero. It follows that the solution space has a basis whose components are bounded by

$$O(\Omega^4)! (e^{O(\Omega^2 \log(\Omega))})^{O(\Omega^3)} (e^{O(\Omega \log(\Omega))})^{O(\Omega^4)} = e^{O(\Omega^5 \log(\Omega))}.$$

If A_L does not have full rank, then, as before, any nontrivial solution of A_L gives rise to a nontrivial solution of A by padding the solution vectors with zeros. Applying Lemma 6 to an arbitrary decomposition of A_L into a block of full rank and the rest gives the bound

$$O(\Omega^3)! (e^{O(\Omega^2 \log(\Omega))})^{O(\Omega^3)} = e^{O(\Omega^5 \log(\Omega))}$$

for the size of the integers in a basis of the solution space of A_L . We have thus completed the proof of the following theorem.

Theorem 10. *For every $\Omega \in \mathbb{N}$, let h_Ω be a proper hypergeometric term as in (1) for which the integer coefficients appearing in the Γ terms are bounded in absolute value by Ω , for which p , x and y are fixed, and for which $\nu\vartheta \rightarrow \infty$ as $\Omega \rightarrow \infty$. Then, as Ω approaches infinity, each term h_Ω admits a telescoper L_Ω of order $O(\Omega)$ and polynomial degree $O(\Omega^2)$ with integer coefficients bounded in absolute value by $e^{O(\Omega^5 \log(\Omega))}$.*

There is nothing special about the choice $r = 2\nu$ in the above derivation. The argument works more generally for any choice $r = \alpha\nu$ where $\alpha > 1$ is a constant (assumed to remain fixed as Ω grows). Choosing $d = \frac{1+2\alpha}{\alpha-1}\nu\vartheta$ also leads to the bound $e^{O(\Omega^5 \log(\Omega))}$.

For the (generically) minimal order $r = \nu$, the approach of this section only delivers the height bound $e^{O(\Omega^6 \log(\Omega))}$ for a telescoper of degree $O(\Omega^3)$, which is much worse than the height bound $e^{O(\Omega^3 \log(\Omega))}$ obtained in Theorem 7 for a telescoper of degree at most $O(\Omega^4)$.

To conclude the section, we again compare the theoretical bound with the actual heights found on a particular example.

Example 11. *For $\Omega = 1, 2, 3, \dots$ consider the same proper hypergeometric term $h_\Omega = \frac{\Gamma(\Omega k)}{\Gamma(\Omega n - k)}$ as in Example 9. From the minimal telescopers L_Ω of order $\nu = \Omega + 1$, we constructed nonminimal telescopers of order 2Ω of small degree and height.*

For each L_Ω , we computed many terms of a randomly chosen sequence solution, and used these to construct a candidate operator M_Ω of order 2Ω and minimal degree by guessing. Checking that the M_Ω are left-multiples of the L_Ω proves that they are indeed telescopers. Unlike the minimal order operators L_Ω , the minimal degree operators of order 2Ω are typically not unique but form a vector space over \mathbb{Q} of dimension greater than 1. For example, for $\Omega = 6$, the telescopers of order 12 and degree 53 form a vector space of dimension 3 and there are no telescopers of order 12 and degree 52 or less. Using LLL, we determined an element of these vector spaces with small (but not necessarily smallest possible) integer coefficients. Let H_Ω be the logarithm of the maximum of the absolute values of the coefficients of the vector computed in this way.

In Figure 2, we plot the values of $\frac{H_\Omega}{\Omega^5}$ (bullets, \bullet) against the least square fits

1. $\log(\Omega) \left(0.269 + \frac{0.599}{\Omega} \right)$ (solid, ---)

$$2. \frac{\log(\Omega)}{\Omega} \left(2.73 - \frac{3.39}{\Omega} \right) \text{ (dashed, - - - -)}$$

for comparing the hypotheses $H_\Omega = \Theta(\Omega^5 \log(\Omega))$ or $H_\Omega = \Theta(\Omega^4 \log(\Omega))$. Unfortunately, because of the high computational cost of computing H_Ω , we were not able to produce more data points. However, despite being less convincing than the test in the previous example, also here the solid curve seems to catch the trend better than the dashed curve, suggesting that the (quasi-)quintic bound can probably not be improved to a (quasi-)quartic bound in general.

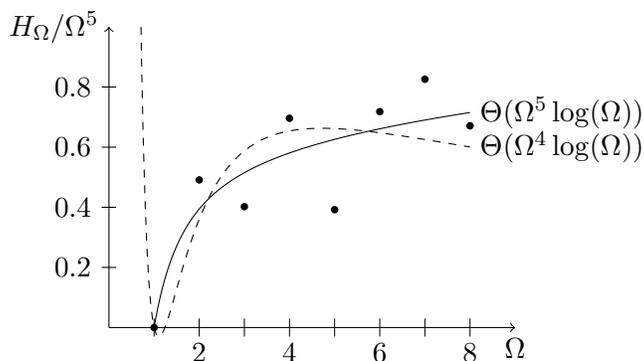


Figure 2: Heights of nonminimal telescopers

6 Consequences

Theorems 7 and 10 are primarily interesting for two reasons. First, they give rise to a significant improvement of Yen’s “two-line algorithm” for proving hypergeometric summation identities (Yen, 1993, 1996), and second, they imply a bound on the bit complexity of creative telescoping. No such bound was known before.

The two-line algorithm rests on the following observations.

Proposition 12 (Yen 1993, 1996). *Let $L \in \mathbb{Z}[n][S_n]$ be an operator of order r and degree d , and let $\ell_r \in \mathbb{Z}[n] \setminus \{0\}$ be the coefficient of S_n^r in L .*

1. *Suppose there is a sequence $(a_n)_{n=0}^\infty$ which is annihilated by L and contains a run of at least $r + d + 1$ consecutive 1’s (i.e., there exists an index $n_0 \in \mathbb{N}$ with $a_{n_0} = a_{n_0+1} = \dots = a_{n_0+r+d+1} = 1$). Then L also annihilates 1.*

2. Let $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ be sequences which are annihilated by L . If $a_n = b_n$ for all $n \leq r + n_0$, where n_0 is the greatest nonnegative integer root of ℓ_r (or $n_0 = 0$ if ℓ_r has no nonnegative integer roots). Then $a_n = b_n$ for all $n \in \mathbb{N}$.
3. If n_0 is an integer root of ℓ_r , then $n_0 \leq d|\ell_r|$.

In view of these facts, in order to prove a hypergeometric summation identity

$$\sum_k h(n, k) = 1,$$

for a given proper hypergeometric term $h(n, k)$ which has finite support and no singularities in $\mathbb{N} \times \mathbb{Z}$, it suffices to proceed as follows:

1. Determine bounds on the order r , the degree d , and the height H , of some telescoper of the summand h .
2. Check the identity for $n = 0, \dots, r + dH$. It holds for all $n \in \mathbb{N}$ iff it holds for all these points.

For step 1, Yen gives an explicit formula for a bound with asymptotic growth $e^{O(\Omega^6 \log(\Omega))}$ ($\Omega \rightarrow \infty$). Our bound from Theorem 7 is significantly better, albeit still exponential. Although, as illustrated in the example, our bound seems to be tight in general, it turns out that in virtually all examples the integer roots of the leading coefficient ℓ_r are much smaller than they could be. In these cases, it remains much more efficient to compute a telescoper for the summand and inspect the linear factors of ℓ_r .

For the cost of computing a telescoper, Theorem 8 of Chen and Kauers (2012) says that a telescoper of order $r = \nu$ [resp. $r = O(\Omega)$] and degree $d = O(\Omega^3)$ [resp. $d = O(\Omega^2)$] can be computed using $O^\sim(\Omega^9)$ [resp. $O^\sim(\Omega^8)$] arithmetic operations, where the soft-O notation $O^\sim(\cdot)$ suppresses possible logarithmic terms. If we use these algorithms to compute telescopers modulo various primes and then use Chinese remaindering and rational reconstruction to combine the results of the modular computations into a telescoper with integer coefficients, this will take time proportional to the length of the integers appearing in the output times the number of arithmetic operations spent for a single prime. We thus obtain a bound $O^\sim(\Omega^3) \times O^\sim(\Omega^9) = O^\sim(\Omega^{12})$ for the time to compute a telescoper of order $r = \nu$ if no lower order telescoper exists, and a bound of

$O\sim(\Omega^5) \times O\sim(\Omega^8) = O\sim(\Omega^{13})$ for the time to compute a nonminimal telescoper of order $r = O(\Omega)$.

There is another, somewhat more heuristic algorithm which makes use of the fact that all the telescopers of a given term h form a left ideal in the operator algebra $\mathbb{Q}(n)[S_n]$ (see Bronstein and Petkovšek (1996) for a tutorial on arithmetic in such algebras). The algorithm proceeds as follows. Choose a prime $p \in \mathbb{Z}$ and compute several nonminimal telescopers, then take their greatest common right divisor in $\mathbb{Z}_p(n)[S_n]$, and hope that this is the modular image of the minimal telescoper. With high probability, this will be the case. Repeat the computation for various primes and use Chinese remaindering and rational reconstruction to recover an operator in $\mathbb{Q}(n)[S_n]$ from all the modular greatest common right divisors. If we assume that the cost of computing the greatest common right divisor can be neglected, then this algorithm spends $O\sim(\Omega^8)$ operations in \mathbb{Z}_p for every prime p , and if we further assume that possible issues related to unlucky primes can be neglected as well, we expect to need $O\sim(\Omega^3)$ primes of size $O\sim(1)$. The resulting bit complexity is thus $O\sim(\Omega^3) \times O\sim(\Omega^8) = O\sim(\Omega^{11})$ for terms h whose minimal telescoper has order $r = \nu$.

As pointed out above, for proving a hypergeometric identity it is not necessary to explicitly compute a telescoper for the summand. Yen's algorithm gets away without computing any information about the telescoper. It is however very expensive. On the other hand, explicitly computing a complete telescoper is more than we need, even though it is cheaper. The algorithm proposed by Guo et al. (2008) is an attempt to compromise between these two extremes: it actually sets up the linear system for computing a telescoper, but then, instead of solving it, it determines a bound on the height of the solution, taking into account special features of the particular matrix at hand, such as sparsity, in a more careful way than it would be easily possible to do in a general analysis. Unfortunately, Guo et al. do not make any statement about the complexity of their algorithm. It would be interesting to know whether their improvement can be translated into better bounds on either the height of a telescoper or, more generally, on the bit complexity of creative telescoping.

7 Acknowledgement

We would like to thank Mogens Lemvig Hansen for producing the graphs for both examples.

References

- Manuel Bronstein and Marko Petkovšek. An introduction to pseudo-linear algebra. *Theoretical Computer Science*, 157(1):3–33, 1996.
- Shaoshi Chen and Manuel Kauers. Order-degree curves for hypergeometric creative telescoping. In *Proceedings of ISSAC'12*, pages 122–129, 2012.
- William Gosper. Decision procedure for indefinite hypergeometric summation. *Proceedings of the National Academy of Sciences of the United States of America*, 75:40–42, 1978.
- Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, second edition, 1994.
- Qiang-Hui Guo, Qing-Hu Hou, and Lisa H. Sum. Proving hypergeometric identities by numerical verifications. *Journal of Symbolic Computation*, 43(12):895–907, 2008.
- Christoph Koutschan. Creative telescoping for holonomic functions. In Johannes Blümlein and Carsten Schneider, editors, *Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions*, Texts and Monographs in Symbolic Computation, pages 171–194. Springer, 2013.
- Mohamud Mohammed and Doron Zeilberger. Sharp upper bounds for the orders of the recurrences output by the Zeilberger and q -Zeilberger algorithms. *Journal of Symbolic Computation*, 39(2):201–207, 2005.
- Marko Petkovšek, Herbert Wilf, and Doron Zeilberger. *A = B*. AK Peters, Ltd., 1997.
- Carsten Schneider. Simplifying multiple sums in difference fields. In Johannes Blümlein and Carsten Schneider, editors, *Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions*, Texts and Monographs in Symbolic Computation, pages 325–360. Springer, 2013.

Lily Yen. *Contributions to the Proof Theory of Hypergeometric Identities*. PhD thesis, University of Pennsylvania, 1993. URL <http://courses.capilanou.ca/faculty/lyen/thesis/>.

Lily Yen. A two-line algorithm for proving terminating hypergeometric identities. *Journal of Mathematical Analysis and Applications*, 198(3):856–878, 1996. ISSN 0022-247X. doi: <http://dx.doi.org/10.1006/jmaa.1996.0118>.

Doron Zeilberger. A fast algorithm for proving terminating hypergeometric identities. *Discrete Mathematics*, 80:207–211, 1990.

Doron Zeilberger. The method of creative telescoping. *Journal of Symbolic Computation*, 11:195–204, 1991.