ABSTRACT

The long-term goal initiated in this work is to obtain fast algorithms and implementations for definite integration in Almkvist and Zeilberger’s framework of (differential) creative telescoping. Our complexity-driven approach is to obtain tight degree bounds on the various expressions involved in the method. To make the problem more tractable, we restrict to bivariate rational functions. By considering this constrained class of inputs, we are able to blend the general method of creative telescoping with the well-known Hermite reduction. We then use our new method to compute diagonals of rational power series arising from combinatorics.

Categories and Subject Descriptors:
I.1.2 [Computing Methodologies]: Symbolic and Algebraic Manipulations — Algebraic Algorithms

General Terms: Algorithms, Theory.

Keywords: Algorithms, Theory.

1. INTRODUCTION

The long-term goal of the research initiated in the present work is to obtain fast algorithms and implementations for the definite integration of general special functions, in a complexity-driven perspective. As most special-function integrals cannot be expressed in closed form, their evaluation cannot be based on table lookups only, and even when closed forms are available, they may prove to be intractable in further manipulations. In both cases, the difficulty can be mitigated by representing functions by annihilating differential operators. This motivated Zeilberger to introduce a method now known as creative telescoping [20], which applies to a large class of special functions: the D-finite functions [14] defined by sets of linear differential equations of any order, with polynomial coefficients. Zeilberger’s method applies in general to multiple integrals and sums.

A sketch of Zeilberger’s method is as follows. Given a D-finite function \( f \) of the variables \( x \) and \( y \), the definite integral \( F(x) = \int_0^y f(x, y) \, dy \) is D-finite, and a linear differential equation satisfied by \( F \) can be constructed [20]. To explain this, let \( k \) be a field of characteristic zero, \( D_x \) and \( D_y \) be the usual derivations on the rational-function field \( k(x, y) \), both restricting to zero on \( k \), and let \( k(x, y)(D_x, D_y) \) be the ring of linear differential operators over \( k(x, y) \). The heart of the method is to solve the differential telescoping equation (1) below for \( L \in \mathbb{K}[x][D_x] \setminus \{0\} \) and \( g = R(f) \) for some \( R \in k(x)(D_x, D_y) \). The operator \( L \) is called a telescoper for \( f \), and \( g \) a certificate of \( L \) for \( f \). Under the assumption

\[
\lim_{y \to \alpha} g(x, y) = \lim_{y \to \beta} g(x, y)
\]

for \( x \) in some domain,

\( L(x, D_x) \) is then proved to be an annihilator of \( F \).

The main emphasis in works since the 1990’s has been on finding telescopers of order minimal over all telescopers for \( f \), which are called minimal telescopers. (Two minimal telescopers differ by a multiplicative factor in \( k(x) \).) In view of the computational difficulty of solving (1), there has been special attention to subclasses of inputs. Of particular importance is the case of hyperexponential functions, defined by first-order differential equations, studied by Almkvist and Zeilberger in [1]. Their method is a direct differential analogue of Zeilberger’s algorithm for the recurrence case [21].

On the other hand, very little is known about the complexity of creative telescoping: the only related result seems to be an analysis in [9] of an algorithm for hyperexponential indefinite integration. In order to get complexity estimates, we simplify the problem by restricting to a smaller class of inputs, namely that of bivariate rational functions. Although restricted, this class already has many applications, for instance in combinatorics, where many nontrivial problems are encoded as diagonals of rational formal power series, themselves expressible as integrals. Our goal thus reads as follows.

Problem. Given \( f = P/Q \in k(x, y) \setminus \{0\} \), find a pair \((L, g)\) with \( L = \sum_{\eta \neq 0} \eta(x)D_x^\eta \) in \( k[x][D_x] \setminus \{0\} \) and \( g \) in \( k(x, y) \) such that

\[
L(x, D_x)(f) = D_y(g).
\] (1)

By considering this more constrained class of inputs, we are indeed able to blend the general method of creative telescoping with the well-known Hermite reduction [10].

Essentially two algorithms for minimal telescopers can be found in the literature: The classical way [1] is to apply a
differential analogue of Gosper’s indefinite summation algorithm, which reduces the problem to solving an auxiliary linear differential equation for polynomial solutions. An algorithm developed later in [7] (see also [12]) performs Hermite reduction on $f$ to get an additive decomposition of the form $f = D_y(a) + \sum_{i=1}^{m} u_i/v_i$, where the $u_i$ and $v_i$ are in $k[x][y]$ and the $v_i$ are squarefree. Then, the algorithm in [1] is applied to each $u_i/v_i$ to get a telescoper $L_i$ minimal for it. The least common multiple of the $L_i$‘s is then proved to be a minimal telescope for $f$. This algorithm performs well only for specific inputs (both in practice and from the complexity viewpoint), but it inspired our Lemma 22 via [12].

As a first contribution in this article, we present a new, provably faster algorithm for computing minimal telescopes for bivariate rational functions. Instead of a single use of Hermite reduction as in [12], we apply Hermite reduction to the $D_y(f)$‘s, iteratively for $i = 0, 1, \ldots$, which yields

$$D_y(n)(f) = D_y(g) + \sum_{i=0}^{\nu_i} \frac{u_i}{w_i}$$

for some factor $w$ of the squarefree part of the denominator of $f$. If $\eta_0, \ldots, \eta_{\nu_i} \in k(x)$ are not all zero and such that $\sum_{i=0}^{\nu_i} \eta_i w_i = 0$, then the operator $\sum_{i=0}^{\nu_i} \eta_i D_y^{(i)}$ is a telescoper for $f$, and more specifically, the first nontrivial linear relation obtained in this way yields a minimal telescope for $f$.

As a second contribution, we give the first proof of a polynomial complexity for creative telescoping on a specific class of inputs, namely on bivariate rational functions. For minimal telescopes, only a polynomial bound on $d_y$ (but none on $d_x$) was given for special inputs in [7]; more specifically, we derive complexity estimates for all mentioned methods (see Fig. 1), showing that our approach is faster. Furthermore, we analyse the bidegrees of non minimal telescopes generated by other approaches: Lipshitz’s work [13] can be rephrased into an existence theorem for telescopes with polynomial size; the approach followed in the recent work on algebraic functions [3] leads to telescopes of smaller degree sizes. These are new instances of the philosophy, promoted in [3], that relaxing minimality can produce smaller outputs.

A third contribution is a fast Maple implementation [22], incorporating a careful implementation of the original Hermite reduction algorithm, making use of the special form of $u_i/w_i$ in (2) and of usual modular techniques (probabilistic rank estimate) to determine when to invoke the solver for linear algebraic equations. Experimental results indicate that our implementation outperforms Maple’s core routine.

Note that for the fastest method we propose, denoted by $H^1$ in Tables 1–3, we chose to output the certificate as a mere sum of (small) rational functions, without any form of normalisation. This choice seems to be uncommon for creative-telescoping algorithms, but a motivation is how the certificate is used in practice. Very often, like for applications to diagonals in §5, the certificate is actually not needed. In other applications, the next step of the method of creative telescoping is to integrate (1) between $\alpha$ and $\beta$, leading to $L(F)(x) = g(x, \alpha) - g(x, \beta)$. Therefore, only evaluations of the certificate are really needed, and normalisation can be postponed to after specialising at $\alpha$ and $\beta$.

The end of this section, §1.1, provides classical complexity results, notation, and hypotheses that will be used throughout. We then study Hermite reduction over $k(x)$ in §2, proving output degree bounds and a low-complexity algorithm. This is then applied in §3 to derive our new algorithm for creative telescoping, and to compare its complexity with that of Almkvist and Zeilberger’s approach. For non-minimal telescopes, we show the existence of some of lower arithmetic size in §4: cubic for nonminimal order instead of quartic for minimal order. See the summary in Figure 1, where the low complexity of algorithms for minimal telescopes relies on Storjohann and Villard’s algorithms [19], thus inducing a certified probabilistic feature. We apply our results to the calculation of diagonals in §5, and describe our implementation and comment on execution timings in §5.

1.1 Background on complexity — Notation

We recall basic notation and complexity facts for later use. Let $k$ be again a field of characteristic zero. Unless otherwise specified, all complexity estimates are given in terms of arithmetical operations in $k$, which we denote by “ops”. Let $k[x]_{\leq d}^{m \times n}$ be the set of $m \times n$ matrices with coefficients in $k[x]$ of degree at most $d$. Let $\omega \in [2,3]$ be a feasible exponent of matrix multiplication, so that two matrices from $k^{n \times n}$ can be multiplied using $O(n^\omega)$ ops. Facts 1 and 2 below show the complexity of multipoint evaluation, rational interpolation, and algebraic operations on polynomial matrices using fast arithmetic, where the notation $\tilde{O}(\cdot)$ indicates cost estimates with hidden logarithmic factors [6, Def. 25.8].

Fact 1 For $p \in k[x]$ of degree less than $n$, pairwise distinct $u_0, \ldots, u_{n-1} \in k$, and $v_0, \ldots, v_{n-1} \in k$, we have:

(i) Evaluating $p$ at the $u_i$‘s takes $O(n)$ ops.

(ii) For $m \in \{1, \ldots, n\}$, constructing $f = s/t \in k(x)$ with

$$\deg_x(s) < m \quad \text{and} \quad \deg_y(t) \leq n - m \quad \text{such that} \quad t(u_1) \neq 0$$

and $f(u_i) = v_i$ for $0 \leq i \leq n - 1$ takes $\tilde{O}(n)$ ops.

Fact 2 For $M$ in $k[x]_{\leq d}^{m \times n}$, $d > 0$, we have:

(i) If $M = (M_1 \ M_2)$ is an invertible $n \times n$ matrix with $M_i \in k[x]_{\leq d}^{m \times n}$, where $i = 1, 2$ and $n_1 + n_2 = n$, then the degree of $\det(M)$ is at most $n_1d_1 + n_2d_2$.

(ii) If $M = (M_1 \ M_2)$ is not of full rank and with $M_i \in k[x]^{m \times n}_{\leq d}$, where $i = 1, 2$ and $n_1 + n_2 = n$, then there exists a nonzero $u \in k[x]^n$ with coefficients of degree at most $n_1d_1 + n_2d_2$ such that $Mu = 0$.

(iii) The rank $r$ and a basis of the null space of $M$ can be computed using $\widetilde{O}(nmr^{\omega - 2}d)$ ops.
In contrast, the method of Horowitz and Ostrogradsky takes $O(n^2)$ operations in $K$ [6, §22.2]. Thus, Hermite’s method is quasi-optimal and asymptotically faster than the former.

From now on, we fix $K = k(x)$ and analyse the complexity of Hermite reduction over $k(x)$ in terms of operations in $k$.

To this end, we use an evaluation-interpolation approach.

2.1 Output size estimates

We derive an upper bound on the bidegrees of $g$ and $r$ satisfying (3) by studying the linear system in (11).

Analysing Hermite reduction (under (H)) shows the existence of $A, a \in k(x)[y]$ with $\deg_y(A) < d_y^*$, $\deg_y(a) < d_y^*$ and

$$\frac{P}{Q} = D_y \left( \frac{A}{Q} \right) + \frac{a}{Q}.\tag{4}$$

In order to bound the bidegrees of $A$ and $a$, we reformulate (4) into the equivalent form

$$P = Q^* D_y(A) - \left( \frac{Q^* D_y(Q^{-1})}{Q^{-1}} \right) A + Q^{-1} a,\tag{5}$$

where $Q^* D_y(Q^{-1})/Q^{-1}$ is a polynomial in $k(x)[y]$ of bidegree at most $(d_y^*, d_y^* - 1)$ - Fact 3. Viewing $A$ and $a$ as polynomials in $k(x)[y]$ with undetermined coefficients, we form the following linear system, equivalent to (5),

$$(H_1 \, H_2) \left( \begin{array}{c} \hat{A} \\ \hat{a} \end{array} \right) = \hat{P},\tag{6}$$

where $H_1 \in k[x]^{d_x \times d_y^*}$, $H_2 \in k[x]^{d_x \times d_y^*}$, and $\hat{A}$, $\hat{a}$, and $\hat{P}$ are the coefficient vectors of $A$, $a$, and $P$ with sizes $d_x^*$, $d_y^*$, and $d_y$, respectively. Under the constraint of properness of $A/Q^*$ and $a/Q^*$, $(A, a)$ is unique by Lemma 4. Then (6) has a unique solution, which leads to the following lemma.

Lemma 6 The matrix $(H_1 \, H_2)$ is invertible over $k(x)$.

As the matrix $(H_1 \, H_2)$ is uniquely defined by $Q$, we call it the matrix associated with $Q$, denoted by $H(Q)$. Let $\delta$ be its determinant, so that $\deg_y(\delta) \leq \mu:=d_x^* d_y^* + d_x^* d_y^*$ by Fact 2(i). For later use, we also define $\bar{\delta}$ as the determinant of $H(Q^{∗2})$, so that $\deg_y(\bar{\delta}) \leq \mu’:=2d_x^* d_y^*$ by Fact 2(i) and since $(Q^{∗2})^{-1} = Q^{-1}$.

Lemma 7 There exist $B, b \in k[x,y]$ with $\deg_y(B) < d_y^*$ and $\deg_y(b) < d_y^*$, and such that:

(i) $\frac{P}{Q} = D_y \left( \frac{B}{Q^{∗2}} \right) + \frac{b}{Q^{∗2}}$;

(ii) $\deg_y(B) \leq \mu - d_x^* + \deg_y(P)$ and $\deg_y(b) \leq \mu - d_x^* + \deg_y(P)$.

Proof. Applying Cramer’s rule to (6) leads to (i). Assertion (ii) next follows by determinant expansions.

In what follows, we shall encounter proper rational functions with denominator $Q$ satisfying $Q = Q^{∗2}$. The following lemma is an easy corollary of Lemma 7 for such functions.

Corollary 8 Assuming $Q = Q^{∗2}$ in addition to Hypothesis (H), there exist $B, b \in k[x,y]$ with $\deg_y(B) < d_y^*$ and $\deg_y(b) < d_y^*$, and such that

(i) $\frac{P}{Q^{∗2}} = D_y \left( \frac{B}{Q^{∗2}} \right) + \frac{b}{Q^{∗2}}$;

(ii) $\deg_y(B)$ and $\deg_y(b)$ are bounded by $\mu’ - d_x^* + \deg_y(P)$. 

(For proofs, see [6, Cor. 10.8, 5.18, 11.6] and [19, Th. 7.3]).

We call squarefree factorisation of $Q \in k[x,y][k[x,y]$ w.r.t. $y$ the unique product $Q_1 Q_2 \cdots Q_m$ equal to $Q$ for $g \in k[x]$ and $Q_i \in k[x,y]$ satisfying $\deg_y(Q_i) > 0$ and such that the $Q_i$’s are primitive, squarefree, and pairwise coprime. The squarefree part $Q^*$ of $Q$ w.r.t. $y$ is the product $Q_1 Q_2 \cdots Q_m$. Let $Q^*$ denote the polynomial $Q_1 Q_2 \cdots Q_m$. and $\deg_y(Q^*)$ the leading coefficient of $Q$ w.r.t. $y$. The following two formulas about $Q$, $Q^*$, and $Q^*$ can be proved by mere calculations.

Fact 3 Let $\hat{Q}$ denote $Q^*/Q$. Then we have

(i) $Q^* D_y(Q^*)/Q^* = \sum_{i=1}^m (i - 1) \hat{Q}_i D_y(\hat{Q}_i) \in k[x,y]$;

(ii) $D_y(Q)/Q^* = \sum_{i=1}^m \hat{Q}_i D_y(\hat{Q}_i) \in k[x,y]$.

Let $f = P/Q$ be a nonzero element in $k(x,y)$, where $P, Q$ are two coprime polynomials in $k[x,y]$. The degree of $f$ in $x$ is defined to be $\max \{\deg_x(P), \deg_x(Q)\}$, and denoted by $\deg_x(f)$. The degree of $f$ in $y$ is defined similarly. The bidegree of $f$ is the pair $(\deg_x(f), \deg_y(f))$, which is denoted by $\text{bideg}(f)$. The bidegree of $f$ is said to be bounded (above) by $(\alpha, \beta)$, written $\text{bideg}(f) \leq (\alpha, \beta)$, when $\deg_x(f) \leq \alpha$ and $\deg_y(f) \leq \beta$.

We say that $f = P/Q$ is proper if the degree of $P$ in $y$ is less than that of $Q$. For creative teasing, we may always assume w.l.o.g. that $f = P/Q$ is proper. If not, rewrite $f = D_y(p) + \hat{f}$ with $p \in k(x)[y]$ and $\hat{f}$ proper. A telescope $L$ for $f$ with certificate $\hat{g}$ is a telescope for $f$ with certificate $L(p) + \hat{g}$.

Hypothesis (H) From now on, $P$ and $Q$ are assumed to be nonzero polynomials in $k[x,y]$ such that $\deg_y(P) < \deg_y(Q)$, $\gcd(P, Q) = 1$, and $Q$ is primitive w.r.t. $y$.

Notation From now on, we write $(d_x, d_y)$, $(d_x^*, d_y^*)$, and $(d_x^*, d_y^*)$ for the bidegrees of $Q$, $Q^*$, and $Q^*$, respectively.

The following notation makes our estimates concise.

Hypothesis (H’) Occasionally, we shall require the extended hypothesis: Hypothesis (H) and $\deg_y(P) \leq d_x^*$.

2. HERMITE REDUCTION

Let $K$ be a field of characteristic zero, either $k$ or $k(x)$ in what follows. Let $K(y)$ be the field of rational functions in $y$ over $K$, and $D_y$ be the usual derivation on it. For a rational function $f \in K(y)$, Hermite reduction [10] computes rational functions $g$ and $r = a/b$ in $K(y)$ satisfying

$f = D_y(g) + r, \quad \deg_y(a) < \deg_y(b), \quad b$ is squarefree. (3)

Horowitz and Ostrogradsky’s method [15, 11] computes the same decomposition as in (3) by solving a linear system. For the details of those methods, see [4, Chapter 2].

Lemma 4 If $f$ is proper, a pair $(g, r)$ satisfying (3) for proper $g, r$ is unique.

Proof. This is a consequence of [11, Theorem 2.10] after writing $r$ as a sum $\sum_{i=1}^m a_i/(x - b_i)$ and integrating. ☐

Lemma 5 Let $f$ be a nonzero rational function in $K(y)$ of degree at most $n$ in $y$, then Hermite reduction on $f$ can be performed using $O(n)$ operations in $K$.

Proof. See [6, Theorem 22.7]. ☐
2.2 Algorithm by evaluation and interpolation

We observe that an asymptotically optimal complexity can be achieved by evaluation and interpolation at each step of Hermite reduction over k(x). This inspires us to adapt Gerhard’s modular method \[8, 9\] to k(x, y). Recall that, by Hyp. (H), Q \in k[x, y] is nonzero and primitive over k[x].

**Definition** An element \( x_0 \in k \) is lucky if \( \text{lcm}_y(Q)(x_0) \neq 0 \) and \( \text{deg}_y(\text{gcd}(Q(x_0, y), D_y(Q(x_0, y)))) = d_y \).

**Lemma 9** There are at most \( d_y(2d_y^3 - 1) \) unlucky points.

**Proof.** Let \( \sigma \in k[x] \) be the \( d_y \)th subresultant w.r.t. y of Q and \( D_y(Q) \). By \[9, \text{Corollary 5.5}\], all unlucky points are in the set \( U = \{ x_0 \in k \mid \sigma(x_0) = 0 \} \). By \[9, \text{Corollary 3.2(ii)}\], \( \text{deg}_y(\sigma) \leq d_y(2d_y^3 - 1) \).

**Lemma 10** Let B, b, and \( \delta \) be the same as in Lemma 7, and let \( x_0 \in k \) be lucky. Then \( \delta(x_0) \neq 0 \) and \( (B(x_0, y), b(x_0, y)) \) is the unique pair such that

\[
P(x_0, y) = D_y \left( \frac{B(x_0, y)}{\delta(x_0)Q^*(x_0, y)} \right) + \frac{b(x_0, y)}{\delta(x_0)Q^*(x_0, y)}. \tag{7}
\]

**Proof.** By the luckiness of \( x_0 \), \( \text{deg}_y(\text{gcd}(Q(x_0, y), \delta)) = d_y \) and \( Q(x_0, y)^+ = Q^*(x_0, y) \), so \( Q(x_0, y)^* = Q^*(x_0, y) \). This implies \( \mathcal{H}(Q)(x_0, y) = \mathcal{H}(Q(x_0, y)) \), which, by Lemma 6, is invertible over k(x). Hence \( \delta(x_0) \neq 0 \), and the evaluation at \( x = x_0 \) of the equality in Lemma 7(ii) is well-defined. Thus, \( (B(x_0, y), b(x_0, y)) \) is a solution of (7). Uniqueness follows from Lemma 4.

**Theorem 11** Algorithm HermiteEvalInterp in Figure 2 is correct and takes \( \mathcal{O}(d_x d_y^3 + \text{deg}_y(P) d_y) \) ops.

**Proof.** Set \( \nu \) to \( d_y(2d_y^3 - 1) \). Lemma 9 implies that the \( \lambda + \nu \) lucky points found in Step 3 are all less than \( \lambda + \nu + 1 \). By Lemmas 4 and 7(ii), \( A = B/\delta \) and \( a = b/\delta \). By Lemma 10, \( A_0 = B(x_0, y)/\delta(x_0) \) and \( a_0 = b(x_0, y)/\delta(x_0) \). By Lemma 7(ii) and since \( \text{deg}_y(\delta) \leq \mu \), it suffices to rationally interpolate \( A \) and \( a \) from values at \( \lambda + 1 \) lucky points. This shows the correctness. The dominant computation in Step 1 is the gcd, which takes \( \mathcal{O}(d_x d_y) \) ops by \[6, \text{Cor. 11.9}\]. For each integer \( 1 \leq \lambda + \nu \), testing luckiness amounts to evaluating at \( x_0 \) and computing \( \text{gcd}(Q(x_0, y), D_y(Q(x_0, y))) \), which takes \( \mathcal{O}(d_y) \) ops by Fact 1(ii) and \[6, \text{Cor. 11.6}\]. Then, generating \( S \) in Step 3 costs \( \mathcal{O}((\lambda + \nu + 1)d_y) \) ops. By Fact 1(i), evaluations in Step 4 take \( \mathcal{O}(\lambda + 1 + d_y) \) ops. For each \( x_0 \in S \), the cost of the Hermite reduction in Step 4 is \( \mathcal{O}(d_y) \) ops by Lemma 5. Thus, the total cost of Step 4 is \( \mathcal{O}(\lambda + 1 + d_y) \) ops. By Fact 1(ii), Step 5 takes \( \mathcal{O}(\lambda + 1 + d_y) \) ops. Since \( \lambda \leq 2d_x d_y + \text{deg}_y(P) \) and \( \nu \leq 2d_x d_y \), the total cost is as announced.

As the generic output size of Hermite reduction is proportional to \( \lambda d_y \), which is \( \mathcal{O}(d_x d_y + \text{deg}_y(P) d_y) \), Algorithm HermiteEvalInterp has quasi-optimal complexity.

3. MINIMAL TELESCOPERS

We analyse two algorithms for constructing minimal telescopers for bivariate rational functions and their certificates.

![Figure 2: Hermite reduction over k(x) via evaluation and interpolation.](image)

**Algorithm HermiteEvalInterp**

**Input:** \( P, Q \in k[x, y] \) satisfying Hypothesis (H).

**Output:** \((A, a) \in k(x)[y]^2\) solving (4).

1. Compute \( Q^- := \text{gcd}(Q, D_y(Q)) \) and \( Q^+ := Q/Q^- \);
2. Set \( \lambda := 2(d_x^2 d_y^3 + d_x^3 d_y) + \text{deg}_y(P) - \min(d_x^2, d_y^2) \);
3. Set \( S \) to the set of \( \lambda + 1 \) smallest nonnegative integers that are lucky for \( Q \);
4. For each \( x_0 \in S \), compute \((A_0, a_0) \in k[y]^2\) such that
   \[
   \frac{P(x_0, y)}{Q(x_0, y)} = D_y \left( \frac{A_0}{Q^-} \right) + \frac{a_0}{Q^+} \]
   using Hermite reduction over \( k \);
5. Compute \((A, a) \in k(x)[y]\) by rational interpolation and return this pair.

3.1 Hermite reduction approach

We design a new algorithm, presented in Figure 3, to compute minimal telescopers for rational functions by basing on Hermite reduction. For \( f = P/Q \in k(x, y) \) and \( n \in \mathbb{N} \), Hermite reduction decomposes \( D_x(f) \) into

\[
D_x(f) = D_y(g_i) + r_i, \tag{8}
\]

where \( g_i, r_i \in k(x, y) \) are proper. Since the squarefree part of the denominator of \( D_x(f) \) divides \( Q^+ \), so does the denominator of \( r_i \). The following lemma shows that (8) recombines into telescopers and certificates; next, Lemma 13 implies that the first pair obtained in this way by Algorithm HermiteTelescopin in Figure 3 yields a minimal telescopers.

**Lemma 12** The rational functions \( r_0, \ldots, r_{d_y} \) are linearly dependent over \( k(x) \).

**Proof.** The constraints on \( r_i \) imply \( \text{deg}_y(r_i Q^+) < d_y \) for all \( i \in \mathbb{N} \), from which follows the existence of a nontrivial linear dependence among the \( r_i \)'s over \( k(x) \).

**Lemma 13** An integer \( p \) is minimal such that \( \sum_{i=0}^p \eta_i r_i = 0 \) for \( \eta_0, \ldots, \eta_p \in k(x, y) \) not all zero if and only if \( \sum_{i=0}^p \eta_i D_x^i \) is a minimal telescopers for \( f \) with certificate \( \sum_{i=0}^p \eta_i g_i \).

**Proof.** Multiplying (8) by \( \eta_i \) before summing yields

\[
L(f) = D_y \left( \sum_{i=0}^p \eta_i g_i \right) + \sum_{i=0}^p \eta_i r_i \quad \text{for} \quad L := \sum_{i=0}^p \eta_i D_x^i,
\]

where the first two sums are proper. Thus, by Lemma 4, \( L \) is a telescopers of order \( p \) for \( f \) with certificate \( \sum_{i=0}^p \eta_i g_i \) if and only if \( \sum_{i=0}^p \eta_i r_i = 0 \) with \( \eta_0 \neq 0 \). The lemma follows.

3.1.1 Order bounds for minimal telescopers

Lemmas 12 and 13 combine into an upper bound on the order of minimal telescopers for \( f \).

**Corollary 14** Minimal telescopers have order at most \( d_y^3 \).

The bound \( 6d_y \) is shown in [3] for rational functions of the form \( yD_y(Q)/Q \in k[x, y] \). Apagodu and Zeilberger [2] obtain a similar bound for a class of nonrational
hyperexponential functions, but their proof does not seem to apply to rational functions, as it heavily relies on the presence of a nontrivial exponential part.

We also derive a lower bound on the order of the minimal telescoper, to be used as an optimisation at the end of §3.1.3: choosing a lucky \( x_0 \in k \), next applying Hermite reduction in \( k(y) \) to \( D_y^i(f)(x_0,y) \), yields

\[
D_y^i(f)(x_0,y) = D_y(g_{0,i}) + r_0,1,
\]

where \( g_{0,i} \) are proper and the denominator of \( r_0,1 \) divides \( Q^i(x_0,y) \). Let \( r_0 \) be the smallest integer such that \( r_0,0, \ldots , r_0,\rho \) are linearly dependent over \( k \).

**Lemma 15** A minimal telescoper has order at least \( r_0 \).

Proof. We first claim that \( r_0,1 = r_1(x_0,y) \) for \( r_1 \) as in (8). Note that the squarefree part w.r.t. \( y \) of the denominator of \( D_y^i(f) \) divides \( Q^i \) for all \( i \in \mathbb{N} \). By [9, Cor. 5.5], \( x_0 \) is lucky for the denominator of \( D_y^i(f) \) for all \( i \in \mathbb{N} \). Then, the claim on \( r_0,1 \) follows from Lemma 10 applied to \( D_y^i(f) \). Let \( \rho \) be the minimal order of a telescoper, then \( r_0, \ldots , r_\rho \) are linearly dependent over \( k(x) \) by Lemma 13. Thus \( r_0,0, \ldots , r_0,\rho \) are linearly dependent over \( k \), which implies \( \rho_0 \leq \rho \).

3.1.2 Degree bounds for minimal telescopers

To derive degree bounds for \( g_i \) and \( r_i \) in (8), let \( \delta, \delta', \mu \) and \( \mu' \) be defined as before Lemma 7, and set \( \mu'' = \mu + \mu' - 1 \).

**Lemma 16** Let \( W \mid k[x,y] \) with \( \deg_x(W) < d_y^* \). Then, for all \( \delta \in \mathbb{N} \), \( B, b \in k[x,y] \) with both \( \deg(B) \) and \( \deg(b) \) bounded by \( \deg_x(W) + \mu'' \) and \( \mu'' \), bounded by \( \deg_x(W) + \mu'' - 1 \), \( d_y^* - 1 \), such that

\[
D_x \left( \frac{W}{\delta + 1}\right) = D_y \left( \frac{B}{\delta + 1}\right) + \frac{b}{\delta + 1},
\]

Proof. A straightforward calculation leads to

\[
D_x \left( \frac{W}{\delta + 1}\right) = \frac{\bar{W}}{\delta + 2\delta' + 1} - \frac{1}{\delta + 1} \cdot \frac{W D_x(Q)}{Q^2},
\]

where \( \deg(\bar{W}) \leq \deg_x(W) + \mu'' - 1 \). By Corollary 8, there exist \( \bar{B}, \bar{b} \in k[x,y] \) such that

\[
D_x \left( \frac{W}{\delta + 1}\right) = \frac{1}{\delta + 1} \cdot \frac{W D_x(Q)}{Q^2} = \frac{1}{\delta + 1} \cdot \frac{D_y \left( \frac{\bar{B}}{Q} \right) + \bar{b}}{\delta + 1},
\]

with \( \deg(\bar{B}) \) and \( \deg(\bar{b}) \) bounded by \( \deg_x(W) + \mu'' - 1 \) and \( d_y^* - 1 \). Setting \( (B, b) = (-\bar{B}, \bar{W} - \bar{b}) \) ends the proof.

**Lemma 17** For \( i \in \mathbb{N} \), there exist \( B_i, b_i \in k[x,y] \) such that

\[
D_y^i(f) = D_y \left( \frac{B_i}{\delta+i+1}\right) + \frac{b_i}{\delta+i+1},
\]

Moreover, \( \deg(\bar{B}_i) \leq \deg_x(P) + \mu + i\mu'' + (i-1)d_x + d_y^* - 1 \) and \( \deg(b_i) \leq \deg_x(P) + \mu + i\mu'' - d_x, d_y^* - 1 \).

Proof. We proceed by induction on \( i \). For \( i = 0 \), the claim follows from Lemma 7. Assume that \( i > 0 \) and that the claim holds for the values less than \( i \). For brevity, we set \( \gamma = \deg_x(P) + \mu \), \( F_i = B_i-1(\delta^{i+1}Q^{-i+1}Q^{-1}) \), and \( G_i = b_i-1(\delta^{i+1}Q^{-i+1}Q^{-1}) \).

The induction hypothesis implies

\[
D_y^i(f) = D_y D_y^i(F_{i-1}) + D_y G_{i-1},
\]

with bidegree bounds on \( B_{i-1} \) and \( b_{i-1} \). Fact (3) implies that \( \bar{Q} = Q^i D_y(Q^-)/Q^- \) is in \( k[x,y] \), with \( \deg(\bar{Q}) \leq \]

**Algorithm HermiteTelescoping(f)**

**Input:** \( f = P/Q \in k(x,y) \) satisfying Hypothesis (H).

**Output:** A minimal telescoper \( L \in k[x](D_x) \) with certificate \( g \in k(x,y) \).

1. Apply HermiteEvalInterp to \( f \) to get \( (g_0, a_0) \) such that \( f = D_y(g_0) + a_0/Q^* \). If \( a_0 = 0 \), return \( (1, g_0) \).
2. For \( i \) from 1 to \( \deg_y(Q^*) \) do
   (a) Apply HermiteEvalInterp to \( -a_{i-1} D_x(Q^*)/Q^2 \) to express it as \( D_y(g_i) + a_i/Q^* \).
   (b) Set \( g_i = D_x(g_{i-1}) + g_i \) and \( a_i = D_x(a_{i-1}) + a_i \).
   (c) Solve \( \sum_{j=0}^i \eta_j a_j = 0 \) for \( \eta_j \in k(x) \) using [19]. If there exists a nontrivial solution, then set \( (L, g) := (\sum_{j=0}^i \eta_j D_x^j, \sum_{j=0}^i \eta_j g_j) \), and break.
3. Compute the content \( c \) of \( L \) and return \((c^{-1}L, c^{-1}g)\).

**Figure 3:** Creative telescoping by Hermite reduction \((d_y - 1, d_y^*)\). Hence \( D_y(1/Q^-) = -Q/Q \). This observation and an easy calculation imply that

\[
D_y(F_{i-1}) = \frac{\bar{B}_{i-1}}{\delta+i+1}\bar{Q}/Q^2 - \frac{\bar{B}_{i-1}}{\delta+i+1}\bar{Q}/Q^2.
\]

Setting \( B_i = \bar{B}_{i-1} + \bar{B}_i Q^{-1} \) and \( b_i = \bar{b}_{i-1} \), we arrive at (10). It remains to verify the degree bounds. The induction hypothesis implies that both \( \deg_x(\bar{B}_i) \) and \( \deg(b_i) \) are bounded by \( \gamma + \mu'' - d_y^* \). It follows that \( \deg_x(\bar{B}_i Q^{-1} \bar{Q}) \) is bounded by \( \gamma + \mu'' + (i-1)d_x \). Similarly, \( \deg_x(\bar{B}_{i-1}) \) is bounded by \( \gamma + \mu'' + (i-1)d_x \), and so is \( \deg_x(B_i) \). The bounds on degrees in \( y \) are obvious.

We next derive degree bounds for the minimal telescopers obtained at an intermediate stage of HermiteTelescoping; refined bounds on the output will be given by Theorem 25.

**Lemma 18** Under (H'), Step 2(c) of Algorithm HermiteTelescoping computes a minimal telescoper \( L \in k[x](D_x) \) with order \( \rho \) and a certificate \( g \in k(x,y) \) for \( P/Q \) with \( \deg_x(L) \in \mathcal{O}(d_x d_y \rho^2) \) and \( \deg(g) \in \mathcal{O}(d_x d_y \rho^2) \times \mathcal{O}(d_y \rho) \).

Proof. By Lemma 13, we exhibit a minimal telescoper by considering the first nontrivial linear dependence among the \( a_i \)'s in (10). Let \( M \) be the coefficient matrix of the system in (\( \eta_i \)) obtained from \( \sum_{i=0}^\infty a_i = 0 \). By Lemma 17, \( M \) is of size at most \( (\rho + 1) \times d_y^* \) and coefficients of degree at most \( \sigma := d_x + \mu + \mu'' - d_y^* + d_y^* \). This introduces a \( \mathcal{O}(d_x d_y \rho^2) \times \mathcal{O}(d_y \rho) \).

3.1.3 Complexity estimates

We proceed to analyse the complexity of the algorithm in Figure 3 and of an optimisation.
Theorem 19 Under Hyp. (H'), Algorithm HermiteTelescoping in Figure 3 is correct and takes $O(\rho^{r+1} d x d y^2)$ ops, where $\rho$ is the order of the minimal telescoper.

Proof. The formulas in Step 2(a) create the loop invariant $D_1(f) = D_y(g_i) + a_i/\rho^r$. Correctness then follows from Lemmas 12 and 20. Step 1 takes $O(dx d y^2)$ ops by Theorem 11 under (H'). By Lemma 17, deg$_x(-a_i-1 D_x Q^*) \in O(dx d y)$. So the cost for performing Hermite reduction on $-a_i-1 D_x Q^*/\rho^r$ in Step 2(a) is $O(dx d y^2)$ ops by Theorem 11. The bigrades of $g_i$ and $a_i$ in Step 2(b) are in $O(dx d y) \times O(dx d y)$ by Lemma 17. Since adding and differentiating have linear complexity, Step 2(c) takes $O(dx d y^2)$ ops. For each $i$, the coefficient matrix of $\sum_{j=0}^i \eta_j a_j = 0$ in Step 2(c) is of size at most $(i+1) \times d y^2$ and with coefficients of degree at most deg$_x(a_i) \in O(dx d y)$. Moreover, the rank of this matrix is either $i$ or $i+1$. Then, Step 2(c) takes $O(dx d y^2)$ ops by Fact 2(iii). Computing the content and divisions in Step 3 has complexity $O(dx d y^3 \rho^3)$. If the algorithm returns when $\rho = \rho_0$, then the total cost is in

$$\sum_{i=0}^\rho O(i^2 d x d y) + \sum_{i=1}^\rho O(\rho^i d x d y^2) \subset O(\rho^{r+1} d x d y^2) \text{ ops,} \quad (11)$$

which is as announced. □

An optimisation, based on Lemma 15, consists in guessing the order $\rho$ so as to perform Step 2(c) a few times only: as a preprocessing step, choose $x_0 \in k$ lucky for $Q$, then detect linear dependence of $(r_0,0,\ldots, r_0)$ in (9). The minimal $j$ for dependence is a lower bound $\rho_0$ on $\rho$. So Step 2(c) is then performed only when $i \geq \rho_0$. In practice, the lower bound $\rho_0$ computed in this way almost always coincides with the actual order $\rho$. So normalising the $g_i$'s becomes the dominant step, as observed in experiments. We analyse this optimisation by first estimating the cost for computing $\rho_0$.

Lemma 20 Under Hypothesis (H'), computing a lower order bound $\rho_0$ for minimal telescopers takes $O(dx d y \rho_0^3)$ ops.

Proof. Since differentiating has linear complexity, the derivative $D_x^j(f)$ takes $O(i^2 d x d y)$ ops. By Fact 1(i), the evaluation $D_x^j(f)(x_0,y)$ takes as much. The cost of Hermite reduction on $D_x^j(f)(x_0,y)$ is $O(dx d y)$ ops by Lemma 5. By Fact 2(iii) with $d = 1$, computing the rank of the coefficient matrix of $\sum_{j=0}^\rho \eta_j r_0 j, r_0$, with $r_0, j$ as in (9), takes $O(dx d y^{\rho_0-1})$ ops. Thus, the total cost for computing a lower bound on $\rho_0$ is $O(dx d y^{\rho_0-1}) \subset O(dx d y \rho_0^3)$ ops. □

Corollary 21 For runs such that $\rho_0 = \rho = O(1)$, the previous optimisation of Hermite Telescoping takes $O(dx d y^2)$ ops.

Proof. In view of Lemma 20, the estimate (11) becomes $O(dx d y^2) + \sum_{i=1}^\rho O(i^2 d x d y^2) + \sum_{i=0}^\rho \sum_{j=0}^i O(i^2 d x d y^2)$, which is $O(dx d y^2) + O((\rho-\rho_0) \rho^r d x d y^2)$ ops, whence the result. □

3.2 Almkvist and Zeilberger's approach

We analyse the complexity of Almkvist and Zeilberger's algorithm [1] when restricted to bivariate rational functions. In order to get a telescoper whose order $\rho$ is minimal, the resulting algorithm, denoted RatAZ, solves (1) for increasing, prescribed values of $\rho$ until it gets a solution $(\eta_1, \ldots, \eta_r, g) \in k(x)^{r+1} \times k(x,y)$ with the $\eta_i$'s not all zero. For the analysis, we start by studying the parameterisation of the differential Goser algorithm of [1] under the same restriction to $k(x,y)$.

Definition (99) Let $K$ be a field and $a,b \in K[y]$ be non-zero polynomials. A triple $(p, q, r) \in K[y]^3$ is said to be a differential Goser form of the rational function $a/b$ if

$$a = b Q^r p + \frac{a}{r} \text{ and gcd}(p, q - \tau Q^r(r)) = 1 \text{ for all } \tau \in N.$$  

For hyperexponential $f$, a key step in [1] is to compute a differential Goser form of the logarithmic derivative of $F = \sum_{i=0}^\rho \eta_i D_x^i(f)$, where the $\eta_i$'s are undetermined from $k(x)$. In the analogue RatAZ, this form is predicted by Lemma 22 below, which is a technical generalisation of a result by Le [12] on $F$ when $f$ has a squarefree denominator.

Write $Q = t(x,y)$, splitting content and primitive part w.r.t. $x$. By an easy induction, $D_x^i(f) = N_1/(Q^{iT})$ for $N_1 \in k[x,y]$. For this section, set $F = \sum_{i=0}^\rho \eta_i D_x^i(f)$, $N = \sum_{i=0}^\rho \eta_i N_i T^{i-\rho}$, and $H = -D_y(Q)/Q - \rho^r Q^r D_y(T^r)$.

Lemma 22 If $F$ is nonzero, the triple $(N, H, Q^r)$ is a differential Goser form of $D_y(F)/F$.

Proof. First, observe $F = N/(Q^{iT})$ and $Q^r = t^r T^r$. Next, $D_y(F)/F = D_y(N)/(N - D_y(Q)/\rho - \rho^r D_y(T^r)) = D_y(N)/(N + H/Q^r)$, which remains to prove $gcd(Q^r, H - \tau D_y(Q^r)) = 1$, for any $\tau \in N$. Recall that the squarefree part $Q^r$ of $Q$ is the product $Q_1 Q_2 \cdots Q_m$ and that $Q_1$ denotes $Q^r/Q_1$. By Fact 2(iii),

$$Z := H - \tau D_y(Q^r) = -\rho^r D_y(T^r) - \sum_{i=1}^\rho (i + \tau) Q_i D_y(Q_i).$$

If $Q_i$ divides $t^r$, $Z$ reduces to $-(j + \tau) Q_i D_y(Q_i)$ modulo $Q_i$. If not, it reduces to $-(j + \tau) Q_i D_y(Q_i) - \rho^r (Q_i D_y(T^r)/Q_i)$, which rewrites to $-(j + \tau + \rho) Q_i D_y(Q_i)$ modulo $Q_i$. In both cases, $Z$ is coprime with $Q^r$, as $j > 0$, $\tau > 0$, and $\rho > 0$.

By another induction, we observe $bideg(N_i) \leq (deg_x(T^r) - i d_y + i \deg_y(T^r) - 1)$, so that $bideg(N) \leq (deg_x(T^r) + \rho d_x(T^r) - \rho d_y + \rho \deg_y(T^r) - 1)$. The next step in RatAZ is, for fixed $\rho$, to reduce (1) by the change of unknown $g = z/(Q^{iT})$, so as to determine all $(\eta_i) \in k(x)^{r+1}$ for which the differential equation in $z$

$$\sum_{i=0}^\rho \eta_i N_i T^{i-\rho} = Q^r D_y(z) + (D_y(Q^r)/Q + H) z$$  

has a polynomial solution in $k(x)[y]$. For later use, we recall the following consequence of [9, Corollary 9.6].

Lemma 23 Let $a,b \in K[y]$ be such that $b = -lc_y(b)/lc_y(a)$ is a nonnegative integer and $deg_y(b) = deg_y(a) - 1$. Let $c \in K[y]$ be such that $b \geq deg_y(c) - deg_y(a) + 1$. If $u$ is a polynomial solution of $a D_y(z) + bzcz = c$, then $deg_y(u) \leq b$.

The following lemma generalises [12, Lemma 2) to present a degree bound for $z$.

Lemma 24 If $u \in k(x[y]$ is a solution of (12) for $(\eta_i) \in k(x)^{r+1}$, then $deg_y(u)$ is bounded by $\beta = d_x + d_y + \rho \deg_y(T^r)$.

Proof. Let $a = Q^r$ and $b = D_y(Q^r) + H$. By the definition of $H$, $b = -Q^r D_y(Q^r)/Q - \rho^r D_y(T^r)$. Fact 3(i) implies that $lc_y(b) = -(d_x + \rho \deg_y(T^r)) lc_y(a)$. Therefore, $b = -lc_y(b)/lc_y(a) = d_y + \rho \deg_y(T^r)$. As deg$_x(N) < d_x + \rho \deg_y(T^r)$ and $d_x = d_x + d_y + \beta \geq deg_y(N) - d_y + 1$. The lemma holds by Lemma 23. □
Algorithm RatAZ($f$)

**INPUT:** $f = P/Q \in k(x,y)$ satisfying Hypothesis (H).

**OUTPUT:** A minimal telescoper $L \in k[x](D_x)$ with certificate $g \in k(x,y)$.

1. Compute $Q^- = \gcd(Q, D_y(Q))$. $Q^* = Q/Q^-$, and $T, T^*$ primitive parts of $Q, Q^*$ w.r.t. $x$, respectively;

2. Set $(\hat{N}, N, \beta, H)$ to $(P, P_d, -Q^*D_y(Q))/Q$;

3. For $\ell = 0, 1, \ldots$ do
   (a) Set $z$ to $\sum_{i=0}^{\beta} z_i y^i$, extract the linear system $\mathcal{M}(\eta z)^T = 0$ from (12) for $\ell = \ell$ and compute a basis $S$ of the null space of $\mathcal{M}$ by [19].
   (b) If $S$ contains a solution $(\eta_0, \ldots, \eta_s, s)$ such that $\eta_0, \ldots, \eta_s$ are not all nonzero, then set $(L, g) := (\sum_{i=0}^{\beta} \eta_i D_y, s/(Q^-T^{-\ell}))$, and go to Step 4;
   (c) Update $\hat{N} := D_y(\hat{N})T^* - N(T^*D_y(T)/T + iD_x(T^*))$, $N := NT^* + \eta_{s+1}\hat{N}$, $\beta := \beta + \deg_y(T^*)$, and $H := H - tD_y(T^*)$.

4. Compute the content $c$ of $L$ and return $(c^{-1}L, c^{-1}g)$.

**Figure 4:** Improved Almkvist–Zeilberger algorithm

We end the present section using the approach of Almkvist and Zeilberger to provide tight degree bounds on the outputs from Algorithms HermiteTelescoping and RatAZ.

**Theorem 25** Under Hypothesis (H'), there exists a minimal telescoper $L \in k[x](D_x)$ with certificate $g(x,y)$ with $\deg_y(L) \in \mathcal{O}(d_x d_y)$ and $\deg(g) \in \mathcal{O}(d_x d_y)$.

**Proof.** By Corollary 14, there exists a smallest $\rho \in \mathbb{N}$ at most $d_y$, for which (1) is not all zero. For this $\rho$, we estimate the size of the polynomial matrix $\mathcal{M}$ derived from (12) by undetermined coefficients. By the remark on $N$ after Lemma 22, we have $\deg(N) \leq (n_x, n_y)$ where $n_x := d_x + \rho \deg_y(T^*) - 1 \in \mathcal{O}(d_y, d_y)$. The matrix $\mathcal{M}$ contains two blocks $\mathcal{M}_1 \in k[x]_{\leq n_x}$ with coefficients of degree at most $n_x (\rho + 1) + d_x (\beta + 1) \in \mathcal{O}(d_y, d_y)$. The degree bound in $y$ for $\rho$ is obvious.

We now analyse the complexity of the algorithm in Fig. 4.

**Theorem 26** Under Hypothesis (H'), Algorithm RatAZ in Figure 4 is correct and takes $\mathcal{O}(d_x d_y \rho^{\nu+2})$ ops, where $\rho$ is the order of the minimal telescoper.

**Proof.** By the existence of a telescoper, Corollary 14, and Lemma 24, the algorithm always terminates and returns a minimal telescoper $L$, of order $\rho$ at most $d_y$. Gcd computations dominate the cost of Steps 1 and 2, which take $\mathcal{O}(d_y)$ ops. For each $\ell \in \mathbb{N}$, the dominating cost in Step 3 is computing the null space of $\mathcal{M}$. Let $n_x = d_x + \ell \deg_y(T^*) - 1 \in \mathcal{O}(d_y)$ and $n_y = d_y + \ell \deg_y(T^*) \in \mathcal{O}(d_y)$. By the same argument as in the proof of Theorem 25, the matrix $\mathcal{M}$ is of size at most $(n_y + 1) \times (\ell + \beta + 2)$ and with coefficients of degree at most $n_y$. Let $\tau$ be the rank of $\mathcal{M}$, which is either $\ell + \beta + 2$ or $\ell + \beta + 1$ by construction. Thus, a basis of the null space of $\mathcal{M}$ can be computed within $\mathcal{O}(n_x (n_y + 1) (\ell + \beta + 2) \rho^{-\nu-2})$ ops by Fact 2(iii). Since $\beta \in \mathcal{O}(d_y)$, $\mathcal{O}(n_x (n_y + 1) (\ell + \beta + 1) \rho^{-\nu-2})$ is included in $\mathcal{O}(d_x d_y \rho^{\nu+2})$. Since Step 3 terminates at $\ell = \rho$, the total cost of the algorithm is $\sum_{\ell=0}^{\rho} d_x d_y \rho^{\nu+1}$ ops. This is within the announced complexity, $\mathcal{O}(d_x d_y \rho^{\nu+2})$ ops.

**Corollary 27** Algorithms HermiteTelescoping and RatAZ in Figs. 3 and 4 both output the primitive minimal telescoper $L$ together with its certificate $g$, which satisfy $\deg_y(L) \leq d_y$, $\deg_y(L) \leq \mathcal{O}(d_x d_y \rho^{\nu+2})$, and $\deg(g) \in \mathcal{O}(d_y)$.

**Proof.** Both algorithms output the primitive minimal telescoper, as they compute a minimal telescoper at an intermediate step, and owing to their last step of content removal. Bounds follow from Corollary 14 and Theorem 25.

4. NONMINIMAL TELESCOPERS

Here, we discard Hypothesis (H) and trade the minimality of telescopers for smaller total output sizes. To this end, we adapt and slightly extend the arguments in [15] and [3, §3].

Given $f = P/Q \in k(x,y)$ of bidegree $(d_x, d_y)$, our goal is to find a (possibly nonminimal) telescoper for $f$. It is sufficient to find a nonzero differential operator $A(x, D_x, D_y)$ that annihilates $f$. Indeed, any $A \in k[x](D_x, D_y) \setminus \{0\}$ such that $A(f) = 0$ can be written $A = D_y(L + D_y R)$, where $L$ is nonzero in $k[x](D_x, D_y)$. If $r = 0$, then clearly $L$ is a telescoper for $f$; otherwise, $A(f) = 0$ yields $L(f) = D_y(-R(f) - \sum_{i=0}^{\rho} \beta_i f^{i+1})$ for some $a_i \in k(x)$, which implies that $L$ is again a telescoper for $f$. Moreover, in both cases, $\deg_y(L) \leq \deg_y(A)$ and $\deg_y(L) \leq \deg_y(A)$. Furthermore, for any $(i, j, \ell) \in \mathbb{N}^3$, a direct calculation yields

\[ x^i D_y^j D_x^k(f) = \frac{H_{i,j,k}}{Q^{i+j+k+1}}. \]

where $H_{i,j,k} \in k[x,y]$ and $\deg_y(H_{i,j,k}) \leq (j + \ell + 1)d_x + i - j$ and $\deg_x(H_{i,j,k}) \leq (j + \ell + 1)d_y - \ell$. From these inequalities, we derive the size and complexity estimates in Figure 1 (bottom half), using two different filtrations of $k[x](D_x, D_y)$.

Lipshitz’s filtration ([13]). Let $F_\nu$ be the $k$-vector space of dimension $f_{\nu} := \binom{n_x}{\nu+1}$ spanned by $\{x^i D_y^j D_x^k | i + j + \ell \leq \nu \}$. By (13), $F_\nu(f)$ is contained in the vector space of dimension $f_{\nu} := \binom{n_x}{\nu+1} + \binom{n_y}{\nu+1} + 1$ spanned by $\{x^i D_y^j D_x^k | i \leq (\nu + 1)d_x + \nu, j \leq (\nu + 1)d_y \}$. Choosing $\nu = 6(d_x + 1)(d_y + 1)$ yields $f_{\nu} > g_{\nu}$; therefore, there exists $A \in k(x, D_x, D_y) \setminus \{0\}$ with total degree at most $6(d_x + 1)(d_y + 1)$ in $x, D_x, D_y$ that annihilates $f$. Moreover, $A$ is found by linear algebra in dimension $\mathcal{O}(d_x d_y)$.

A better filtration ([3]). Instead of taking total degree, set $F_{\kappa,\nu}$ to the $k$-vector space of dimension $f_{\kappa,\nu} := (\kappa + 1) \binom{\nu+2}{\nu+1}$ generated by $\{x^i D_y^j D_x^k | i \leq \kappa, j + \ell \leq \nu \}$. By (13), $F_{\kappa,\nu}(f)$ is contained in the vector space of dimension $g_{\kappa,\nu} := (\nu + 1)(d_x + \kappa + 1)(d_y + 1)$ spanned by $\{x^i D_y^j D_x^k | i \leq (\nu + 1)d_x + \kappa, j \leq (\nu + 1)d_y \}$. Choosing $\kappa = 3d_x d_y$ and $\nu = 6d_y$ results in $f_{\kappa,\nu} > g_{\kappa,\nu}$. This implies the existence of $A \in k(x, D_x, D_y) \setminus \{0\}$ with total degree at most $6d_y$ in $D_x$ and $D_y$ and degree at most $3d_x d_y$ in $x$ that annihilates $f$. Again, $A$ is found by linear algebra over $k$, but in smaller dimension $\mathcal{O}(d_x d_y)$. 
Table 1: Creative telescoping on random instances
Timings in ms for algorithms in Table 3 (stopped after 30 min).

| No. | AZ | Abr | RAZ | H1 | R2 | HD | EI | MG |
|-----|----|-----|-----|----|----|----|----|----|
| 29  | 44 | 72  | 32  | 28 | 36 | 20 | 608| 528|
| 45  | 52 | 76  | 36  | 20 | 24 | 32 | 652| 584|
| 46  | 428| 1436| 784 | 492| 1288| 752| 34433| 18945|
| 49  | 474269 | 34694 | 20977 | 10336 | 36254 | 22417 | ∞ | 652968 |

Table 2: Computation of the diagonals of (14)
Timings in ms by creative telescoping of $f(y, x/y)/y$ (upper half) or $f(y/x)/x$ (second half). Algorithms listed in Table 3.

| No. | AZ | Abr | RAZ | H1 | R2 | HD | RR | GP |
|-----|----|-----|-----|----|----|----|----|----|
| 4   | 176 | 136 | 110 | 128 | 108 | 220 | 956 |
| 8   | 5362 | 4244 | 4380 | 1976 | 5344 | 4396 | 10336 | 154409 |
| 10  | 11740 | 12816 | 7108 | 7445 | 24565 | 7076 | 46882 | 1118313 |
| 4   | 184 | 168 | 120 | 120 | 220 | 116 | 224 | 1340 |
| 8   | 3540 | 3704 | 2540 | 2092 | 6976 | 2516 | 10348 | 271480 |
| 10  | 16817 | 17013 | 9200 | 8068 | 32218 | 9092 | 46750 | 412309 |

Table 3: List of the algorithms for the experiments

5. IMPLEMENTATION AND TIMINGS

We implemented in Maple 13 all the algorithms described; as we used Maple’s generic solver SolveTools:-Linear, all of our implementations are deterministic.

The evaluation-interpolation algorithm Algorithm HermiteEvalInterp for Hermite reduction (Fig. 2) does not perform well, mainly because Maple’s rational interpolation routines are far too slow. We thus implemented Algorithm HermiteReduce (original version) in [4, §2.2] (carefully avoiding redundant extended gcd calculations), and noted that it performs better.

We then implemented a variant of Algorithm HermiteTelescoping in Figure 3, using HermiteReduce in place of HermiteEvalInterp, and including the optimisation at the end of §3.1.3, refined by additional modular calculations.

For a rational function, Algorithm HermiteTelescoping returns the minimal telescoper $L$ and the certificate $g$. The algorithm separates the computation for $L$ from that for $g$. Indeed, $g$ is formed by the coefficients of $L$, $g_0$, the $\tilde{g}_i$ and their derivatives given in Figure 3. This feature enables us to either return the certificate $g$ as a sum of unnormalised rational functions, or a normalised rational function.

A selection of timings by this implementation and others are given in Table 1; our code, the full table, as well as the random inputs are given in [22]. For our experiments, we exhaustively considered all 49 degree patterns in factorisations of denominators $Q_1 \cdots Q_m$ $(m \leq 5)$ that add up to bidegree (5,5), and generated corresponding random denominators, imposing the integers of the expanded forms to have around 26 digits. Numerators were generated as random bidegree-(5,5) polynomials with coefficients of 26 digits.

Application to diagonals. The diagonal of a formal power series $s = \sum_{i,j \geq 0} f_{i,j} x^i y^j$ in $k[x,y]$ is defined to be the power series $\Delta(f) := \sum_{i,j \geq 0} f_{i,j} x^i$. For a $D$-finite power series $s$, it is known to be $D$-finite [13], and it is even algebraic for a bivariate rational function $f \in k(x,y) \cap k[x,y]$ [18, §6.3]. A linear differentiator operator $L(x,y)(\Delta(f))$ that annihilates $\Delta(f)$ can then be computed via rational-function telescoping, owing to the following classical lemma from [13].

Lemma 28 Any telescoper for $f(y, x/y)/y$ annihilates $\Delta(f)$.

By this lemma, it suffices to compute a telescoper without its certificate to get an annihilator. Algorithm HermiteTelescoping is suitable for this task, since it separates computation of telescopers and certificates. Alternatively, for $f = P/Q$, we can compute an annihilator of $\Delta(f)$ either as the differential resolvent of the resultant $\text{Res}_y(Q,P - \tau D_x Q)$, or simply guess it from the first terms of the series expansion of $\Delta(f)$.

We compare the various algorithms on an example borrowed from [5] (timings of execution are given in Table 2):

$$f = \frac{1}{1 - x - y - xy(1 - x^2)}, \quad \text{where } d \in \mathbb{N}. \quad (14)$$

All computer calculations have been performed on a Quad-Core Intel Xeon X5482 processor at 3.20GHz, with 3GB of RAM, using up to 6.5GB of memory allocated by Maple.

6. REFERENCES

[1] G. Almkvist and D. Zeilberger. The method of differentiation under the integral sign. J. Symb. Comput., 10:571–591, 1990.
[2] M. Águgodu and D. Zeilberger. Multi-variable Zeilberger and Almkvist-Zeilberger algorithms and the sharpening of Wilf-Zeilberger theory. Adv. in Appl. Math., 37(2):139–152, 2006.
[3] A. Bostan, F. Chyzak, B. Salvy, G. Lecerf, and É. Schost. Differential equations for algebraic functions. In ISSAC’07, pages 25–32. ACM, New York, 2007.
[4] M. Bronstein. Symbolic Integration I: Transcendental functions, volume 1 of Algorithms and Computation in Mathematics. Springer-Verlag, Berlin, second edition, 2005.
[5] A. Flaxman, A. W. Harrow, and G. B. Sorkin. Strings with maximally many distinct subsequences and substrings. Electron. J. Combin., 11(1):R8, 10 pp., 2004.
[6] J. von zur Gathen and J. Gerhard. Modern Computer Algebra. Cambridge University Press, Cambridge, second edition, 2003.
[7] K. O. Geddes and H. Q. Le. An algorithm to compute the minimal telescopes for rational functions (differential-integral case). In Mathematical Software, pages 453–463. WSP, 2002.
[8] J. Gerhard. Fast modular algorithms for squarefree factorization and Hermite integration. Appl. Algebra Engrg. Comm. Comput., 11(3):203–226, 2001.
[9] J. Gerhard. Modular Algorithms in Symbolic Summation and Symbolic Integration (LNCS). Springer-Verlag, 2004.
[10] C. Hermite. Sur l’intégration des fractions rationnelles. Ann. Sci. École Norm. (2), 1:215–218, 1872.
[11] E. Horowitz. Algorithms for partial fraction decomposition and rational function integration. In SYMSAC’71, pages 441–457, New York, USA, 1971. ACM.
[12] H. Q. Le. On the differential-integral analogue of Zeilberger’s algorithm to rational functions. In Proc. of the 2000 Asian Symposium on Computer Mathematics, pages 204–213, 2000.
[13] L. Lipshitz. The diagonal of a $D$-finite power series is $D$-finite. J. Algebra, 113(2):373–378, 1989.
[14] L. Lipshitz. $D$-finite power series. J. Algebra, 122(2):353–373, 1989.
[15] M. Ostrogradsky. De l’intégration des fractions rationnelles. Bull. de la classe physico-mathématique de l’Acad. Impériale des Sciences de Saint-Pétersbourg, 4:145–167, 286–300, 1845.
[16] R. H. Risch. The problem of integration in finite terms. Trans. Amer. Math. Soc., 139:167–189, 1969.
[17] R. H. Risch. The solution of the problem of integration in finite terms. Bull. Amer. Math. Soc., 76:605–608, 1970.
[18] R. P. Stanley. Enumerative Combinatorics. Vol. 2, volume 62 of Cambridge Studies in Advanced Mathematics. CUP, 1999.
[19] A. Storjohann and G. Villard. Computing the rank and a small nullspace basis of a polynomial matrix. In ISSAC’05, pages 399–406. ACM, New York, 2005.
[20] D. Zeilberger. A holonomic systems approach to special functions identities. J. Comput. Appl. Math., 32:321–368, 1990.
[21] D. Zeilberger. The method of creative telescoping. J. Symbolic Comput., 11(3):195–204, 1991.
[22] http://algo.inria.fr/chs/ByvaKAT/. 2010.