Efficient Rational Creative Telescoping

Mark Giesbrecht

Symbolic Computation Group, Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Hui Huang

Symbolic Computation Group, Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

George Labahn

Symbolic Computation Group, Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada

Eugene Zima

Physics and Computer Science, Wilfrid Laurier University, Waterloo, ON, N2L 3C5, Canada

Abstract

We present a new algorithm to compute minimal telescopers for rational functions in two discrete variables. As with recent reduction-based approach, our algorithm has the nice feature that the computation of a telescoper is independent of its certificate. Moreover, our algorithm uses a sparse representation of the certificate, which allows it to be easily manipulated and analyzed without knowing the precise expanded form. This representation hides potential expression swell until the final (and optional) expansion, which can be accomplished in time polynomial in the size of the expanded certificate. A complexity analysis, along with a Maple implementation, suggests that our algorithm has better theoretical and practical performance than the reduction-based approach in the rational case.

Keywords: Rational function, GGSZ reduction, Left division with remainder, Telescoper

1. Introduction

Creative telescoping is a powerful method pioneered by Zeilberger (1990a,b, 1991) in the 1990s and has now become the cornerstone for finding closed forms for definite sums and definite integrals in computer algebra. The method mainly constructs a recurrence (resp. differential) equation admitting the prescribed definite sum (resp. integral) as a solution. Employing other

Email addresses: mwg@uwaterloo.ca (Mark Giesbrecht), h2huang@uwaterloo.ca (Hui Huang), glabahn@uwaterloo.ca (George Labahn), ezima@wlu.ca (Eugene Zima)

algorithms applicable to the resulting recurrence or differential equation, it is then possible to find closed form solutions or prove that there is no such solution. In the latter case, one can still make use of creative telescoping for such operations as determining asymptotic expansions about the sum or integral under investigation.

In the case of summation, in order to compute a sum of the form $\sum_{y=a}^{b} f(x,y)$, the main task of creative telescoping consists of constructing polynomials $c_0, \ldots, c_\rho$ in x, not all zero, and another function g in the same domain as f such that

$$c_0(x)f(x,y) + c_1(x)f(x+1,y) + \cdots + c_\rho(x)f(x+\rho,y) = g(x,y+1) - g(x,y). \tag{1.1}$$

The number ρ may or may not be part of the input. If $c_0, \ldots, c_\rho$ and g are as above, then we say that $L = c_0 + c_1 \mathrm{S}_x + \cdots + c_\rho \mathrm{S}_x^\rho$ with S_x the recurrence operator in x is a *telescoper* for f and g is a *certificate* for L. If $c_\rho \neq 0$ then the integer ρ is the *order* of L. Finally, the maximum degree in x among the polynomials c_ℓ is the *degree* of L.

The technique of creative telescoping has seen various algorithmic generalizations and improvements over the past two decades. As outlined in the introduction of (Chen et al., 2015), we can distinguish four generations amongst these algorithms based on the techniques they employed. Details about the first two generations can be found in (Zeilberger, 1990b, 1991; Petkovšek et al., 1996) and the third generation was initiated in (Mohammed and Zeilberger, 2005; Apagodu and Zeilberger, 2006). It is worthwhile mentioning that many of the best output size estimates known so far are obtained from the analysis of algorithms from the third generation (Chen and Kauers, 2012a,b; Chen et al., 2014).

The fourth generation originating from (Bostan et al., 2010), is the so-called reduction-based approach, and has drawn the most attention recently. The main idea of this approach is to iteratively apply a chosen reduction method a priori to bring each function $f(x+\ell, y)$ on the left-hand side of (1.1) into some kind of normal form, referred to as a *shift-remainder* for brevity, modulo summable rational functions of the forms as the right-hand side of (1.1). Then finding the c_ℓ amounts to seeking a linear dependency amongst these shift-remainders. This idea equips the approach with the useful feature (compared with earlier generations) that it allows one to find a telescoper without necessarily also computing the corresponding certificate. In other words, the computation of the c_ℓ in (1.1) is separated from the computation of g. In a typical situation where the size of the c_ℓ is much smaller than the size of g and the right-hand side of (1.1) collapses to zero when summing over the defining interval, this approach enables one to merely compute the c_ℓ avoiding the costly yet unnecessary computation of the certificate g. In applications where a certificate is required, the approach also allows one to express the certificate as an unnormalized sum so that the summands are concatenated symbolically without actually calculating the sum. These summands are often of much smaller sizes than the original certificate. So far, the reduction-based approach has been worked out for many special functions, for example, in the differential case (Bostan et al., 2010, 2013b,a; Chen et al., 2016, 2018; van der Hoeven, 2017; Bostan et al., 2018) and in the shift case (Chen et al., 2015; Huang, 2016; Bostan et al., 2017; Chen et al., 2019). These algorithms turn out to be more efficient in practice than those of the previous generations.

However, it is also the case that the unnormalized expression for the certificate returned by the reduction-based approach can introduce superfluous terms which eventually cancel out when normalized. These terms will not contribute to the final output but will increase sizes of intermediate results and thus deteriorate the performance of the approach in these applications.

In order to illustrate this issue, let us consider a simple discrete rational function of the form

$$f(x,y) = \frac{x}{x+3y+3m} - \frac{x}{x+3y+3} + \frac{x}{x+3y}, \tag{1.2}$$

where m is an integer greater than one. Applying a reduction method, for example, in (Abramov, 1975) or (Chen et al., 2015), to the given rational function f yields

$$f(x,y) = g_0(x,y+1) - g_0(x,y) + r_0 \quad \text{with } g_0(x,y) = \sum_{k=1}^{m-1} \frac{x}{x+3y+3k} \quad \text{and} \quad r_0 = \frac{x}{x+3y}. \tag{1.3}$$

Based on the form (1.3), iteratively applying the reduction method to each $f(x+\ell,y)$ for $\ell \geq 0$ gives

$$f(x+\ell,y) = g_\ell(x,y+1) - g_\ell(x,y) + r_\ell \quad \text{with a shift-remainder } r_\ell = \frac{x+\ell}{x+3y+\bar{\ell}},$$

where $\bar{\ell} \in \{0,1,2\}$ is ℓ reduced modulo 3 and

$$g_\ell(x,y) = g_0(x+\ell,y) + \sum_{k=1}^{\lfloor \ell/3 \rfloor} \frac{x+\ell}{x+3y+3(k-1)+\bar{\ell}}.$$

Finding the first linear dependency amongst the shift-remainders r_ℓ reduces to solving the linear system

$$\begin{pmatrix} 9x & 9x+9 & 9x+18 & 9x+27 \\ 6x^2+9x & 6x^2+12x+6 & 6x^2+15x+6 & 6x^2+27x+27 \\ x^3+3x^2+2x & x^3+3x^2+2x & x^3+3x^2+2x & x^3+6x^2+11x+6 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{1.4}$$

The polynomial solution $(c_0, c_1, c_2, c_3) = (-(x+3), 0, 0, x)$ then gives

$$L = x\, S_x^3 - (x+3), \tag{1.5}$$

a telescoper for f of minimal order with the corresponding certificate

$$g(x,y) = x \cdot g_3(x,y) - (x+3) \cdot g_0(x,y) = \frac{x(x+3)}{x+3y+3m} - \frac{x(x+3)}{x+3y+3} + \frac{x(x+3)}{x+3y} \tag{1.6}$$

obtained by canceling out the common $m-2$ terms in the summation. As the m increases, the size of each g_ℓ grows rapidly, whereas the expanded certificate g may still be small. In this particular example, it is actually more reasonable to use the decomposition

$$f(x,y) = g_0(x,y+1) - g_0(x,y) + r_0, \quad \text{with } g_0(x,y) = -\frac{x}{x+3y} \text{ and } r_0 = \frac{x}{x+3y+3m},$$

instead of (1.3). This leads to an alternate choice of shift-remainders r_ℓ, with the resulting g_ℓ of smaller sizes, for the $f(x+\ell,y)$. With this choice one gets the same telescoper L and certificate g as before, but this time there is no cancellation happening in (1.6). That is, the unnormalized sum gives the final size of the certificate. This suggests a solution to the above issue. Namely, find an initial decomposition (1.3) with the smallest possible g_0 using the method proposed in

(Polyakov, 2011; Zima, 2011) to initiate the iterative process of the reduction-based approach. However this process requires a full irreducible factorization of polynomials.

Separate from the previously mentioned work, there is an alternate method developed by Le (2003) which constructs telescopers in a direct fashion. This method was later used by Chen and Kauers (2012a) to obtain a much sharper order-degree curve for telescopers in the case of bivariate rational functions, compared with the curve obtained by using a third-generation algorithm. Currently, the method is only worked out for bivariate rational functions in the (q-) shift case. Nevertheless, the method is still interesting because it also has the feature that the computation of a telescoper does not depend on its certificate. In order to demonstrate its main idea, consider again the rational function f given in (1.2). As with the reduction-based approach, this method first decomposes f as in (1.3). The difference is that it later decomposes the shift-remainder r_0 as the sum of several simple fractions of numerators in x only, which in our example is merely $x \cdot \frac{1}{x+3y}$. By viewing $x = x\,\mathrm{S}_x^0$ as a recurrence operator of order zero and using the fact that $\mathrm{S}_x^3 - 1$ is a minimal telescoper for $\frac{1}{x+3y}$ with a corresponding certificate $\frac{1}{x+3y}$, Le's method then computes the least common left multiple of x and $\mathrm{S}_x^3 - 1$ with the left cofactor of x (resp. $\mathrm{S}_x^3 - 1$), giving rise to the same telescoper L as in (1.5) (resp. its certificate $\frac{x(x+3)}{x+3y}$) for the simple fraction $x \cdot \frac{1}{x+3y} = r_0$. In the more general case where there is more than one simple fraction in r_0, one finds a telescoper of minimal order for r_0 by calculating the least common left multiple of all resulting telescopers for simple fractions. Together with (1.3), the method yields a telescoper of minimal order for f, namely L, as well as its (optional) certificate of the form

$$g = L \cdot g_0 + \frac{x(x+3)}{x+3y}.$$

Rather than leaving the certificate as a (potentially large) unnormalized sum as done by the reduction-based approach, this method represents the certificate by recurrence operators. This representation enables one to more easily manipulate the certificate or analyze its various properties such as the singularities without knowing its expanded form. However, the intermediate expression swell which happens in the certificate is still unavoidable due to (1.3). A second disadvantage is that this method requires the numerator of each simple fraction appearing in the decomposition of the shift-remainder to be independent of y, often requiring one to work in algebraic extensions of the base field.

1.1. Proposed new approach

Our new algorithm constructs a telescoper for a rational function in a similar fashion as the reduction-based approach, but incorporating the idea from the method of Le (2003). As a result, our algorithm completely avoids algebraic extensions of the base field and intermediate expression swell in the certificate. In order to describe the main idea of the new algorithm, let us continue the example (1.2). Unlike the reduction-based approach and the method of Le, we first find a recurrence operator M allowing us to rewrite f in the form

$$f = \underbrace{(x\,\mathrm{S}_x^{3m} - x\,\mathrm{S}_x^3 + x)}_{M} \cdot \frac{1}{x+3y}.$$

Using the relation $\mathrm{S}_x^3(x+3y) = \mathrm{S}_y(x+3y)$ with S_y the recurrence operator in y, we find the left remainder $R = x\,\mathrm{S}_x^0$ from the (special) left division of M by $\mathrm{S}_y - 1$. This implies that $\frac{x}{x+3y}$ is a shift-remainder of f. By fixing a positive integer ρ, say $\rho = 3$ in this example, we make an ansatz $L = c_3 \mathrm{S}_x^3 + c_2 \mathrm{S}_x^2 + c_1 \mathrm{S}_x + c_0$ with the c_ℓ to be determined. Note that upper bounds for the order ρ are available, provided that a telescoper exists (see (Huang, 2016, Theorem 5.5) or a refined version in Lemma 4.4). As before, we calculate the left remainder

$$\tilde{R} = (x+2)c_2 \mathrm{S}_x^2 + (x+1)c_1 \mathrm{S}_x + ((x+3)c_3 + xc_0)$$

from the left division of $L \cdot R$ by $\mathrm{S}_y - 1$. We show that L is a telescoper if and only if $\tilde{R} = 0$, or equivalently,

$$\begin{pmatrix} x & 0 & 0 & x+3 \\ 0 & 0 & x+2 & 0 \\ 0 & x+1 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{1.7}$$

One immediately reads a polynomial solution $(c_0, c_1, c_2, c_3) = (-(x+3), 0, 0, x)$, yielding a telescoper for f. In terms of certificates, we either, follow the idea from (Gerhard et al., 2003) and use the sparse representation

$$g = \mathrm{LeftQuot}(L \cdot M, \mathrm{S}_y - 1) \cdot \frac{1}{x+3y}, \quad \text{where LeftQuot denotes the left quotient,}$$

or expand it as (1.6) by noticing $\mathrm{LeftQuot}(L \cdot M, \mathrm{S}_y - 1) = x(x+3)\,\mathrm{S}_x^{3m} - x(x+3)\,\mathrm{S}_x^3 + x(x+3)$.

As with the reduction-based approach in the shift case, the termination of the new algorithm is guaranteed by the existence criterion for telescopers given in (Abramov, 2003). This essentially says that the given rational function f has a telescoper if and only if the denominator of its shift-remainder is integer-linear, that is, each of its (nontrivial) irreducible factors has the form $P(\lambda x + \mu y)$ for two integers λ, μ and a univariate polynomial P. The particular pair (λ, μ) with $\gcd(\lambda, \mu) = 1$ and $\mu \geq 0$ is called an *integer-linear type*. In the general case, the operator S_x in M is replaced by a special recurrence operator acting particularly on integer-linear rational functions of one type, and the given rational function is initially separated into several simple fractions according to integer-linear types.

When a telescoper is guaranteed and an upper bound ρ has been determined, then the above process can be executed once by constructing a telescoper of minimal order from the nontrivial polynomial solution of the resulting linear system with the last nonzero entry having the smallest possible index. A second, iterative approach, lets $\rho = 0, 1, 2, \ldots$ until a nontrivial polynomial solution of the resulting linear system is found. When the chosen bound is equal to the actual order of minimal telescopers, our complexity analysis suggests that the upper-bound version is faster than the iterative version by a factor of the actual order. However, the iterative version often performs much better in practice when the upper bound is not sufficiently sharp.

In summary, our main contribution is a new algorithm for computing minimal telescopers for rational functions. As with the reduction-based approach and the method of Le, our algorithm separates the computation of the telescoper from that of the certificate. When the certificate is needed our algorithm computes it in a sparse form, hiding potential expression swell until a final, optional expansion. Unlike Le's method our algorithm avoids the need for algebraic extensions. In addition, if an expanded form for the certificate is desired then this can be computed easily by a left division in time polynomial in the size of the expanded certificate. Moreover, comparing (1.7) with (1.4) suggests that our algorithm also has better control for the size of intermediate expression involved in the computation of the telescoper.

Arithmetic costs of both the upper-bound and iterative versions of the new algorithm, as well as that of the reduction-based approach in the rational case, are analyzed in this paper. We note that, until recently, most complexity analyses were done for the differential case (Bostan et al., 2010, 2013b; van der Hoeven, 2017; Bostan et al., 2018) whereas little has been known for the shift case. The analysis result shows that our new algorithm is at least one order of magnitude faster than the reduction-based approach in the rational case when the certificate is not expanded. A Maple implementation further confirms that (the iterative version of) our approach outperforms the reduction-based approach when restricted to the rational case. In addition, the new algorithm is easy to analyze and leads to a tight order-degree curve for telescopers, a property shared with the method of Le.

The remainder of the paper proceeds as follows. Some basic notions and results are recalled in the next section for later use. In particular, we review the notion of shiftless decomposition and the GGSZ reduction in the context of bivariate rational functions. A new recurrence operator ring particularly working on integer-linear rational functions of one type is introduced in Section 3. Based on left division with remainder in this operator ring, Section 4 describes a new algorithm to construct a telescoper of minimal order for bivariate rational functions along with its several variants. Section 5 provides the cost analysis of our new algorithm, followed in Section 6 by a brief summary and the cost analysis of the reduction-based approach in the rational case. Section 7 contains some experimental comparison among all above-mentioned approaches. The paper ends with some topics for future research.

2. Preliminaries

Throughout the paper C denotes a field of characteristic zero with $C(x, y)$ the field of rational functions in x, y over C. We denote by σ_x and σ_y the *shift operators* over $C(x, y)$, which, for any $f \in C(x, y)$, are defined by

$$\sigma_x(f(x, y)) = f(x + 1, y) \quad \text{and} \quad \sigma_y(f(x, y)) = f(x, y + 1).$$

Recall that a rational function $f \in C(x, y)$ is called *summable* with respect to y (or *σ_y-summable* for short) if $f = \sigma_y(g) - g$ for some $g \in C(x, y)$. Two polynomials $f, g \in C[x, y]$ are called *shift-coprime* with respect to y (or *σ_y-coprime* for short) if $\deg_y(\gcd(f, \sigma_y^\ell(g))) = 0$ for all $\ell \in \mathbb{Z}$, and called *shift-equivalent* with respect to y (or *σ_y-equivalent* for short), denoted by $f \sim_y g$, if $f = \sigma_y^m(g)$ for some $m \in \mathbb{Z}$. Clearly, two σ_y-coprime polynomials are coprime to each other in $C[x, y]$, and $\sim_y$ is an equivalence relation. A nonzero polynomial $f \in C[x, y]$ is called *shift-free* with respect to y (or *σ_y-free* for short) if $\deg_y(\gcd(f, \sigma_y^\ell(f))) = 0$ for all $\ell \in \mathbb{Z} \setminus \{0\}$. All these notions can be straightforwardly generalized to polynomials in $C(x)[y]$ and have analogous definitions in terms of the variable x (or other variables). Nevertheless, we note that these notions invoked later are defaulted to be defined with respect to the variable y unless otherwise stated.

Let f be a polynomial in $C[x, y]$. Throughout this paper, we will order terms using a pure lexicographic order with $x \prec y$. For this order, we let $\mathrm{lc}_{x,y}(f)$ and $\deg_{x,y}(f)$ denote the leading coefficient and total degree, respectively, of f over C with respect to x, y. We follow the convention that $\deg_{x,y}(0) = -\infty$ and say that the polynomial f is *monic* if $\mathrm{lc}_{x,y}(f) = 1$. The *content*, denoted by $\mathrm{cont}_{x,y}(f)$, of f with respect to x, y is the greatest common divisor (GCD) over C of the coefficients of f with respect to x, y with f being *primitive* if $\mathrm{cont}_{x,y}(f) = 1$. The *primitive part* $\mathrm{prim}_{x,y}(f)$ of f with respect to x, y is then defined to be $f / \mathrm{cont}_{x,y}(f)$. In certain instances, we also need to consider the above notions with respect to a particular variable, say y. In these cases, we will instead write $\mathrm{lc}_y(f)$, $\deg_y(f)$, $\mathrm{cont}_y(f)$ and $\mathrm{prim}_y(f)$ by viewing f as a polynomial in y over the domain $C[x]$.

Let $C(x,y)[\mathrm{S}_x, \mathrm{S}_y]$ be the ring of linear recurrence operators in x, y over $C(x,y)$. Here $\mathrm{S}_x, \mathrm{S}_y$ commute with each other, and $\mathrm{S}_x(f) = \sigma_x(f)\,\mathrm{S}_x$, $\mathrm{S}_y(f) = \sigma_y(f)\,\mathrm{S}_y$ for any $f \in C(x,y)$. The application of an operator $M = \sum_{i,j\geq 0} a_{ij}\,\mathrm{S}_x^i\,\mathrm{S}_y^j$ in $C(x,y)[\mathrm{S}_x, \mathrm{S}_y]$ to a rational function $f \in C(x,y)$ is then defined as

$$M(f) = \sum_{i,j\geq 0} a_{ij}\sigma_x^i\sigma_y^j(f).$$

Definition 2.1. *Let f be a rational function in $C(x,y)$. A nonzero operator $L \in C[x][\mathrm{S}_x]$ is called a* telescoper *for f if $L(f)$ is σ_y-summable, or equivalently, there exists a rational function $g \in C(x,y)$ such that*

$$L(f) = (\mathrm{S}_y - 1)(g),$$

where 1 denotes the identity map of $C(x,y)$. We call g a corresponding certificate *for L. The* order *and* degree *of L are defined to be its degree in S_x and the maximum degree in x of its coefficients with respect to S_x, respectively. A telescoper of minimal order is also called a* minimal telescoper.

In the rest of this section, we recall from (Gerhard et al., 2003) the notion of shiftless decomposition, and the GGSZ reduction in the context of bivariate rational functions, as follows. Both of these notions will play important roles in our later algorithms.

Definition 2.2. *Let $g \in C[x,y]$ be a polynomial admitting the decomposition*

$$c\prod_{i=1}^{m}\prod_{j=1}^{n_i}\sigma_y^{\ell_{ij}}(g_i)^{e_{ij}}, \tag{2.1}$$

where $c \in C[x]$, $g_i \in C[x,y]$, $m, n_i, \ell_{ij}, e_{ij} \in \mathbb{N}$ and $e_{ij} > 0$. Then (2.1) is called a shiftless decomposition *of g (with respect to y) if*

- *g_i is monic, squarefree and of positive degree in y;*
- *g_i is primitive and shift-free with respect to y;*
- *the g_i are pairwise σ_y-coprime;*
- $0 = \ell_{i1} < \ell_{i2} < \cdots < \ell_{in_i}$.

The above definition coincides with (Gerhard et al., 2003, Definition 1) in the case of univariate polynomials.

A polynomial in $C[x,y]$ may have more than one shiftless decomposition. Depending on the size of m, we can distinguish different types of shiftless decompositions. The most refined type, corresponding to the maximum m, has the form (2.1) with all g_i being irreducible, and is obtained by full factorization in $C[x,y]$. Using the auto-dispersion set of the given polynomial, that is, the set of all integers i with the property that the given polynomial has a nontrivial common divisor with its ith shift, a quartic-time algorithm over $C(x)$ for computing a shiftless decomposition was first developed in (Gerhard et al., 2003, §3). This algorithm returns the coarsest shiftless decomposition, namely the type corresponding to the minimum m, which has the form (2.1) satisfying the property that for all $1 \leq i < i' < m$ at least one of the conditions

$$(\ell_{i1}, \ldots, \ell_{in_i}) \neq (\ell_{i'1}, \ldots, \ell_{i'n_{i'}}) \quad \text{or} \quad (e_{i1}, \ldots, e_{in_i}) \neq (e_{i'1}, \ldots, e_{i'n_{i'}})$$

is satisfied. Note that the auto-dispersion set of a polynomial can be obtained by using the algorithm of Man and Wright (1994) based on full factorization, or more efficiently, in the case of integer polynomials using the modular procedure **pDispersionSet** from (Gerhard et al., 2003).

Based on the shiftless decomposition, a reduction algorithm in the case of univariate rational functions, named **RatSum**, was also developed in the same paper (Gerhard et al., 2003). This algorithm can be carried over to the bivariate case in a straightforward manner, to which we will refer as the GGSZ reduction later for convenience, named after the authors. The input and output of the GGSZ reduction are given below.

GGSZReduction. Given two coprime polynomials $f, g \in C[x, y]$ with $g \neq 0$, compute rational functions h, r in $C(x, y)$ with $r = a/b$, $a, b \in C[x, y]$, $\deg_y(a) < \deg_y(b)$ and b being σ_y-free such that

$$\frac{f}{g} = (\mathrm{S}_y - 1)(h) + r. \tag{2.2}$$

Such a reduction algorithm is vital for many creative telescoping approaches, including the reduction-based one in (Chen et al., 2015), the method of Le (2003) and the algorithm introduced in this paper. Unlike previous reduction algorithms as given in (Abramov, 1971; Paule, 1995; Abramov and Petkovšek, 2001; Gerhard et al., 2003), the GGSZ reduction uses a sparse representation of h in the output in terms of left quotients, and hence works in polynomial-time of the size of the input without the final expansion.

Example 2.3. *Let g be a polynomial of the form*

$$g = g_0 \sigma_y(g_0) \sigma_y^{29}(g_0) \sigma_y^{30}(g_0) \cdot b, \tag{2.3}$$

where $g_0 = xy + 1$ and $b = ((-5x + 2y)^3 + 1)((-5x + 2y + 1)^3 + 1)(3x + 10y + 1)^2$. Up to making b monic, the above equation gives a shiftless decomposition of g with respect to y. Consider now the rational function f/g admitting the partial fraction decomposition

$$\frac{f}{g} = \frac{2x+3}{\sigma_y^{30}(g_0)} - \frac{2x+3}{\sigma_y^{29}(g_0)} - \frac{1}{\sigma_y(g_0)} + \frac{1}{g_0} + \frac{2x^2+1}{(-5x+2y)^3+1} + \frac{x-1}{(-5x+2y+1)^3+1} + \frac{xy+1}{(3x+10y+1)^2}.$$

We remark that all the decomposed forms given in our examples are for readability only. Applying the GGSZ reduction to f/g then yields (2.2) with

$$h = \mathit{LeftQuot}((2x+3)\,\mathrm{S}_y^{30} - (2x+3)\,\mathrm{S}_y^{29} - \mathrm{S}_y + 1, \mathrm{S}_y - 1)\left(\frac{1}{g_0}\right) = ((2x+3)\,\mathrm{S}_y^{29} - 1)\left(\frac{1}{g_0}\right)$$

$$\text{and} \quad r = \frac{2x^2+1}{(-5x+2y)^3+1} + \frac{x-1}{(-5x+2y+1)^3+1} + \frac{xy+1}{(3x+10y+1)^2}, \tag{2.4}$$

where LeftQuot denotes the left quotient in the ring $\mathbb{Q}(x)[y][\mathrm{S}_y]$. In this example, the only left quotient in h is a sparse operator although it is of relatively high order 29. Hence the expanded form of h is small. Since $r \neq 0$, the given rational function f/g is not σ_y-summable by (Gerhard et al., 2003, Theorem 12). We will use f/g as a running example in this paper.

3. Integer-linear operator ring

One of the key ideas of our new algorithm described in the next section is to convert operations among rational functions to arithmetic operations for recurrence operators (such as multiplication and left division with remainder). In order to achieve this goal, in this section we introduce a new recurrence operator ring acting on integer-linear rational functions of a given type.

Recall that an irreducible polynomial $g \in C[x,y]$ is called *integer-linear* (over C) if it is of the form $P(\lambda x + \mu y)$ for some univariate polynomial $P(z) \in C[z]$ and two integers λ, μ. By pulling out a common factor, one may assume without loss of generality that the integers λ, μ are coprime if $g \notin C$ or $(\lambda, \mu) = (0,0)$ otherwise, and that $\mu \geq 0$. Such a pair (λ, μ) is unique and is called the *integer-linear type* of g. A polynomial in $C[x,y]$ is called *integer-linear* (over C) if all its irreducible factors are integer-linear, possibly with different integer-linear types. A rational function in $C(x,y)$ is called *integer-linear* (over C) if its denominator and numerator are both integer-linear.

Rather than merely giving an affirmative answer to the case where an input polynomial is integer-linear, we look for the integer-linear decomposition of the polynomial, with the intent to make the integer-linear part more manipulatable.

Definition 3.1. *Let $g \in C[x,y]$ be a polynomial admitting the decomposition*

$$P_0(x,y) \prod_{i=1}^{\tilde{m}} P_i(\tilde{\lambda}_i x + \tilde{\mu}_i y), \tag{3.1}$$

where $\tilde{m} \in \mathbb{N}$, $\tilde{\lambda}_i, \tilde{\mu}_i \in \mathbb{Z}$, $P_0 \in C[x,y]$ and $P_i(z) \in C[z]$. Then (3.1) is called the integer-linear decomposition *of g if*

(i) none of non-constant irreducible factors of P_0 is integer-linear;

(ii) each $P_i(z)$ is monic with respect to z and of positive degree in z.

(iii) each $(\tilde{\lambda}_i, \tilde{\mu}_i)$ satisfies $\gcd(\tilde{\lambda}_i, \tilde{\mu}_i) = 1$ and $\tilde{\mu}_i \geq 0$;

(iv) any two pairs of the $(\tilde{\lambda}_i, \tilde{\mu}_i)$ are distinct.

The $(\tilde{\lambda}_i, \tilde{\mu}_i)$ are called integer-linear types *of g. If g is clear from the context, we will simply say that the $(\tilde{\lambda}_i, \tilde{\mu}_i)$ are integer-linear types. In addition, we call the product of the $P_i(\tilde{\lambda}_i x + \tilde{\mu}_i y)$ the* integer-linear part *of g.*

Clearly, g is integer-linear if and only if $P_0 \in C$ in (3.1). By using full factorization, we see that every polynomial admits an integer-linear decomposition. Moreover, the decomposition is unique up to the order of the factors, according to the uniqueness of integer-linear types and full factorization.

In terms of computation, an efficient algorithm for finding the integer-linear decomposition for general multivariate polynomials was recently proposed by the authors (Giesbrecht et al., 2019). Compared with the previous known algorithms (Abramov and Le, 2002; Li and Zhang, 2013), this procedure performs better both in theory and in practice.

Consider now integer-linear rational functions of a single type, (λ, μ), with λ, μ a coprime integer pair with $\mu \geq 0$. By Bézout's relation, there exist unique integers α, β with $\alpha\lambda + \beta\mu = 1$ and

$$\begin{cases} 0 \leq \alpha < \mu \text{ and } |\beta| \leq |\lambda|, & \text{if } \lambda\mu \neq 0, \\ \alpha = 1 \text{ and } \beta = 0, & \text{if } (\lambda, \mu) = (1,0), \\ \alpha = 0 \text{ and } \beta = 1, & \text{if } (\lambda, \mu) = (0,1). \end{cases}$$

Define

$$\sigma_{(\lambda,\mu)} = \sigma_x^{\alpha}\sigma_y^{\beta} \quad \text{and} \quad \mathrm{S}_{(\lambda,\mu)} = \mathrm{S}_x^{\alpha}\,\sigma_y^{\beta}.$$

Then $\sigma_{(\lambda,\mu)}$ is an automorphism of $C(x, y)$ mapping $f(x, y) \in C(x, y)$ to $f(x + \alpha, y + \beta)$ and $\mathrm{S}_{(\lambda,\mu)} \in C(x, y)[\mathrm{S}_x, \mathrm{S}_y, \mathrm{S}_y^{-1}]$ with $\mathrm{S}_{(\lambda,\mu)}(f) = \sigma_{(\lambda,\mu)}(f)\,\mathrm{S}_{(\lambda,\mu)}$. In particular, $\sigma_{(\lambda,\mu)}(P(z)) = P(z + 1)$ for $P(z) \in C(z)$ with $z = \lambda x + \mu y$. This allows us to treat integer-linear rational functions of one type as univariate rational functions.

For the pair (λ, μ) given above, consider the operator ring $C(x)[y][\mathrm{S}_{(\lambda,\mu)}, \mathrm{S}_{(\lambda,\mu)}^{-1}]$, in which the addition is defined termwise and the multiplication is defined by the rule

$$\sum_i a_i\,\mathrm{S}_{(\lambda,\mu)}^i \cdot \sum_j b_j\,\mathrm{S}_{(\lambda,\mu)}^j = \sum_{i,j} a_i\,\sigma_x^{\alpha i}\sigma_y^{\beta i}(b_i)\,\mathrm{S}_{(\lambda,\mu)}^{i+j}.$$

We are particularly interested in applying the operators from the ring $C(x)[y][\mathrm{S}_{(\lambda,\mu)}, \mathrm{S}_{(\lambda,\mu)}^{-1}]$ to the elements of the field $C(z)$ with $z = \lambda x + \mu y$. In this case, the operators $\mathrm{S}_x, \mathrm{S}_y$ can be regarded as elements in the ring $C(x)[y][\mathrm{S}_{(\lambda,\mu)}, \mathrm{S}_{(\lambda,\mu)}^{-1}]$ since

$$\mathrm{S}_x(P(z)) = \mathrm{S}_{(\lambda,\mu)}^{\lambda}(P(z)) \;\text{ and }\; \mathrm{S}_y(P(z)) = \mathrm{S}_{(\lambda,\mu)}^{\mu}(P(z)) \quad \text{for } P(z) \in C(z) \text{ with } z = \lambda x + \mu y.$$

Therefore, for $M = \sum_i c_i\,\mathrm{S}_{(\lambda,\mu)}^i \in C(x)[y][\mathrm{S}_{(\lambda,\mu)}, \mathrm{S}_{(\lambda,\mu)}^{-1}]$, the left multiplication by

$$M_1 = \sum_{i\geq 0} a_i\,\mathrm{S}_x^i \in C(x)[y][\mathrm{S}_x] \quad \text{or} \quad M_2 = \sum_{i\geq 0} b_i\,\mathrm{S}_y^i \in C(x)[y][\mathrm{S}_y]$$

in the ring $C(x)[y][\mathrm{S}_{(\lambda,\mu)}, \mathrm{S}_{(\lambda,\mu)}^{-1}]$ is well-defined in the sense that

$$M_1 \cdot M(P(z)) = \sum_{i\geq 0}\sum_j a_i\,\sigma_x^i(c_j)\sigma_{(\lambda,\mu)}^{\lambda i+j}(P(z)) \text{ and } M_2 \cdot M(P(z)) = \sum_{i\geq 0}\sum_j b_i\,\sigma_y^i(c_j)\sigma_{(\lambda,\mu)}^{\mu i+j}(P(z)), \tag{3.2}$$

where $P(z) \in C(z)$ with $z = \lambda x + \mu y$.

3.1. *Refinement for the integer-linear decomposition*

In this subsection we refine the integer-linear part of a given polynomial in $C[x, y]$, so as to better describe the arithmetic in our operator ring.

Recall that two polynomials $f, g \in C[x, y]$ are called *shift-equivalent* with respect to x, y (or (σ_x, σ_y)*-equivalent* for short), denoted by $f \sim_{x,y} g$, if there exist $\ell, m \in \mathbb{Z}$ such that $f = \sigma_x^{\ell}\sigma_y^{m}(g)$. Clearly, $\sim_{x,y}$ is an equivalence relation and contains the relation $\sim_y$. Suppose that f, g are integer-linear of the forms $f(x, y) = P_1(\lambda_1 x + \mu_1 y)$ and $g(x, y) = P_2(\lambda_2 x + \mu_2 y)$ for $P_i(z) \in C[z]$ and $\lambda_i, \mu_i \in \mathbb{Z}$ with $\mu_i \geq 0$ and $\gcd(\lambda_i, \mu_i) = 1$. Then $f \sim_{x,y} g$ implies that $(\lambda_1, \mu_1) = (\lambda_2, \mu_2)$ and $P_1(z) = P_2(z + \ell)$ for some $\ell \in \mathbb{Z}$, namely $f = \mathrm{S}_{(\lambda_1,\mu_1)}^{\ell}(g)$, and conversely.

Suppose a polynomial $b \in C[x, y]$ has the integer-linear decomposition of the form (3.1) with $P_0 = 1$. Then we wish to compute integers $m, n_i, \mu_i, \nu_{ij}, e_{ij} \in \mathbb{N}$, $\lambda_i \in \mathbb{Z}$, and univariate polynomials $c_b(x) \in C[x]$, $p_i(z) \in C[z]$ such that

$$b = c_b(x)\prod_{i=1}^{m}\prod_{j=1}^{n_i} p_i(\lambda_i x + \mu_i y + \nu_{ij})^{e_{ij}}, \tag{3.3}$$

where

(i) each $p_i(z)$ is monic, squarefree, shift-free with respect to z and of positive degree in z;

(ii) each (λ_i, μ_i) is an integer-linear type with $\mu_i > 0$;

(iii) for two distinct integers i, i' with $1 \le i, i' \le m$, every non-constant irreducible factor of $p_i(\lambda_i x + \mu_i y)$ is (σ_x, σ_y)-inequivalent with any factor of $p_{i'}(\lambda_{i'} x + \mu_{i'} y)$, or equivalently, either $(\lambda_i, \mu_i) \neq (\lambda_{i'}, \mu_{i'})$ or $p_i(z)$ is shift-coprime with $p_{i'}(z)$ with respect to z;

(iv) $0 = \nu_{i1} < \cdots < \nu_{in_i}$ and $e_{ij} > 0$.

We refer to (3.3) as the *refined integer-linear decomposition* of the given polynomial b. The following algorithm provides a way to compute such a decomposition.

ILDRefinement. Given a nonzero polynomial $b \in C[x, y]$ admitting the integer-linear decomposition of the form (3.1) with $P_0 = 1$, compute the refined integer-linear decomposition of b. The steps are:

1. set $m = 0$ and $c_b(x) = 1$.
   for $k = 1, \ldots, \tilde{m}$ do

   if $\tilde{\mu}_k = 0$ then update $c_b(x) = P_k(x)$ else

   1.1 compute a shiftless decomposition $\prod_{i=1}^{m_k} \prod_{j=1}^{n_{ik}} \tilde{g}_{ik}(z + \tilde{\nu}_{ijk})^{\tilde{e}_{ijk}}$ of $P_k(z)$ with respect to z.

   1.2 for $i = 1, \ldots, m_k$ do
   update m to $m + 1$; set $n_m = n_{ik}$, $p_m(z) = \tilde{g}_{ik}(z)$ and $(\lambda_m, \mu_m) = (\tilde{\lambda}_k, \tilde{\mu}_k)$;
   for $j = 1, \ldots, n_m$ set $\nu_{m,j} = \tilde{\nu}_{ijk}$ and $e_{m,j} = \tilde{e}_{ijk}$.

2. return $c_b(x)$, $(\lambda_i, \mu_i)_{1 \le i \le m}$, $(\nu_{ij}, e_{ij})_{1 \le i \le m, 1 \le j \le n_i}$ and $(p_i(z))_{1 \le i \le m}$.

The correctness of the refinement procedure follows directly from the definition of the shiftless decomposition.

Example 3.2. *Consider the polynomial g defined by (2.3). By definition, it is easy to see that g admits the integer-linear decomposition*

$$g = P_0(x, y) \cdot P_1(-5x + 2y) \cdot P_2(3x + 10y), \tag{3.4}$$

with $P_0 = g_0 \sigma_x(g_0) \sigma_y^{29}(g_0) \sigma_y^{30}(g_0)$, $P_1(z) = (z^3 + 1)((z + 1)^3 + 1)$ and $P_2(z) = (z + 1)^2$. Here $g_0 = xy + 1$. Furthermore refining the integer-linear part of g (the polynomial b in Example 2.3) yields

$$b = p_1(-5x+2y) \cdot p_1(-5x+2y+1) \cdot p_2(3x+10y) \quad \text{with } p_1(z) = z^3+1 \text{ and } p_2(z) = (z+1)^2. \tag{3.5}$$

4. Telescoping via operator

In this section, we will demonstrate how to construct a telescoper and (optionally) its certificate using left division with remainder among recurrence operators from integer-linear operator rings introduced in the preceding section, in a similar spirit of the GGSZ reduction.

To this end, let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$. Assume that g admits the integer-linear decomposition (3.1) and that b is the integer-linear part of g. Then there exist $f_0, a \in C(x)[y]$ with $\deg_y(a) < \deg_y(b)$ such that

$$\frac{f}{g} = \frac{f_0}{P_0} + \frac{a}{b}. \tag{4.1}$$

By Abramov's criterion (Abramov, 2003, Theorem 10), f_0/P_0 is σ_y-summable provided that a telescoper for f/g exists. As b is integer-linear it admits a refined version (3.3) of its integer-linear decomposition. The unique partial fraction decomposition of a/b with respect to (3.3) over $C(x)[y]$ is then given by

$$\frac{a}{b} = \frac{1}{\delta(x)} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \sum_{k=1}^{e_{ij}} \frac{a_{ijk}(x,y)}{p_i(\lambda_i x + \mu_i y + \nu_{ij})^k},$$

where $\delta \in C[x]$, $a_{ijk} \in C[x,y]$ with $\deg_y(a_{ijk}) < \deg_z(p_i)$. Let $d_i = \max_{1\le j\le n_i}\{e_{ij}\}$ and assume that $a_{ijk} = 0$ in case $k > e_{ij}$. Interchanging the order of summations and introducing the operator $M_{ik} = \frac{1}{\delta}\sum_{j=1}^{n_i} a_{ijk}\, \mathrm{S}_{(\lambda_i,\mu_i)}^{\nu_{ij}} \in C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}]$ then yields

$$\frac{a}{b} = \sum_{i=1}^{m}\sum_{k=1}^{d_i}\left(\frac{1}{\delta(x)}\sum_{j=1}^{n_i}\frac{a_{ijk}(x,y)}{p_i(\lambda_i x + \mu_i y + \nu_{ij})^k}\right) = \sum_{i=1}^{m}\sum_{k=1}^{d_i} M_{ik}\left(\frac{1}{p_i(\lambda_i x + \mu_i y)^k}\right). \tag{4.2}$$

Note that for each fixed i with $1 \le i \le m$, we have

$$\mathrm{S}_x(p_i(z)) = \mathrm{S}_{(\lambda_i,\mu_i)}^{\lambda_i}(p_i(z)) \quad \text{and} \quad \mathrm{S}_y(p_i(z)) = \mathrm{S}_{(\lambda_i,\mu_i)}^{\mu_i}(p_i(z)) \quad \text{with } z = \lambda_i x + \mu_i y.$$

By the equation (4.1) and the definition of telescopers, a nonzero operator $L \in C[x][\mathrm{S}_x]$ is a telescoper for f/g if and only if there exists $h \in C(x,y)$ such that

$$L\left(\frac{a}{b}\right) = (\mathrm{S}_y - 1)(h). \tag{4.3}$$

Since $\deg_y(a) < \deg_y(b)$ and because the denominators of the reduced fractions h and $(\mathrm{S}_y - 1)(h)$ have the same (σ_x,σ_y)-equivalence classes with the same multiplicities, the rational function h can be written in a form analogous to (4.2). That is,

$$h = \sum_{i=1}^{m}\sum_{k=1}^{d_i} Q_{ik}\left(\frac{1}{p_i(\lambda_i x + \mu_i y)^k}\right)$$

for some $Q_{ik} \in C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ with coefficients of degrees in y less than $\deg_z(p_i)$. The above expression is unique because all the $p_i(z)$ are shift-free with respect to z and for any two distinct integers i, i' with $1 \le i, i' \le m$, every non-constant factor of $p_i(\lambda_i x + \mu_i y)$ is (σ_x,σ_y)-inequivalent to all factors of $p_{i'}(\lambda_{i'} x + \mu_{i'} y)$. Thus (4.3) implies that

$$LM_{ik} = (\mathrm{S}_y - 1)Q_{ik} \quad \text{for all } i = 1,\ldots,m \text{ and } k = 1,\ldots,d_i, \tag{4.4}$$

where the multiplications are defined by the rules (3.2) in the ring $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$. With the left quotient in the ring $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ denoted by LeftQuot, the equation (4.4) implies that $Q_{ik} = \mathrm{LeftQuot}(LM_{ik}, \mathrm{S}_y - 1)$. Notice that L and $\mathrm{S}_y - 1$ commute with each other. Let R_{ik} be the left remainder from the left division of M_{ik} by $\mathrm{S}_y - 1$ in $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}]$. Then (4.4) is equivalent to say that the left remainder from the left division of LR_{ik} by $\mathrm{S}_y - 1$ in $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ is equal to zero. This then provides us a rational telescoping criterion, in analogy to the rational summation criterion given by (Gerhard et al., 2003, Theorem 12).

Theorem 4.1. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$. Assume that f/g can be decomposed as (4.1), in which a/b is further assumed to admit the decomposition (4.2). Then a necessary and sufficient condition for a nonzero operator $L \in C[x][\mathrm{S}_x]$ to be a telescoper for $f/g \in C(x, y)$ is that there are operators $Q_{ik} \in C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ such that (4.4) holds. This is also equivalent to the condition that for any $1 \le i \le m$ and $1 \le k \le d_i$, the left remainder from the left division of the operator LR_{ik} by $\mathrm{S}_y - 1$ in $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ is equal to zero, where R_{ik} is the left remainder from the left division of M_{ik} by $\mathrm{S}_y - 1$.*

4.1. Left remainders and left quotient formulas

In what follows, we discuss the concrete formulas for left remainders and left quotients from the left divisions by $\mathrm{S}_y - 1$, inducing a linear system containing the information for telescopers.

Observe that each M_{ik} is equal to $\frac{1}{\delta} \sum_{j=1}^{n_i} a_{ijk} \mathrm{S}_{(\lambda_i,\mu_i)}^{\nu_{ij}}$. A direct calculation shows that the left remainder R_{ik} from the left division of M_{ik} by $\mathrm{S}_y - 1$ in the ring $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}]$ is given by

$$R_{ik} = \frac{1}{\delta} \sum_{r=0}^{\mu_i - 1} \tilde{a}_{ikr} \mathrm{S}_{(\lambda_i,\mu_i)}^{r} \quad \text{with } \tilde{a}_{ikr} = \sum_{j_r} \sigma_y^{-q_{j_r}}(a_{ij_r k}), \tag{4.5}$$

where the summation in $\tilde{a}_{ikr}$ runs over all integers j_r in $\{1, \ldots, n_i\}$ satisfying the property that there exists an integer $q_{j_r} \in \mathbb{Z}$ such that $\nu_{ij_r} = \mu_i q_{j_r} + r$. Inspired by the proof of (Chen and Kauers, 2012a, Theorem 10), we look for a telescoper L of the form $\sum_{\ell=0}^{\rho} c_\ell \sigma_x^\ell(\delta) \mathrm{S}_x^\ell$ for $\rho \in \mathbb{N}$ and $c_\ell \in C[x]$. This in fact does not lose any generality since, for a given telescoper $\tilde{L} = \sum_{\ell=0}^{\rho} \tilde{c}_\ell \mathrm{S}_x^\ell \in C[x][\mathrm{S}_x]$, multiplying with the least common multiple δ_ρ of the $\sigma_x^\ell(\delta)$ gives

$$L = \delta_\rho \cdot \tilde{L} = \sum_{\ell=0}^{\rho} c_\ell \sigma_x^\ell(\delta) \mathrm{S}_x^\ell \quad \text{with } c_\ell = \tilde{c}_\ell \cdot \frac{\delta_\rho}{\sigma_x^\ell(\delta)},$$

yielding a telescoper with the same order as $\tilde{L}$ and of the required form.

It follows from (4.5) that

$$LR_{ik} = \sum_{\ell=0}^{\rho} \sum_{r=0}^{\mu_i - 1} c_\ell \sigma_x^\ell(\tilde{a}_{ikr}) \mathrm{S}_{(\lambda_i,\mu_i)}^{r+\ell\lambda_i} \in C[x, y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}].$$

The left remainder from the left division of LR_{ik} by $\mathrm{S}_y - 1$ in the ring $C(x)[y][\mathrm{S}_{(\lambda_i,\mu_i)}, \mathrm{S}_{(\lambda_i,\mu_i)}^{-1}]$ is simply of the form

$$\tilde{R}_{ik} = \sum_{\tilde{r}=0}^{\mu_i - 1} \left(\sum_{\ell_{\tilde{r}}} c_{\ell_{\tilde{r}}} \sigma_y^{-q_{r\ell_{\tilde{r}}}} \sigma_x^{\ell_{\tilde{r}}}(\tilde{a}_{ikr}) \right) \mathrm{S}_{(\lambda_i,\mu_i)}^{\tilde{r}}, \tag{4.6}$$

where the inner summation runs over all integers $\ell_{\tilde{r}}$ in $\{0, \ldots, \rho\}$ satisfying the property that there exist integers $r, q_{r\ell_{\tilde{r}}}$ such that $0 \le r \le \mu_i - 1$ and $r + \ell_{\tilde{r}}\lambda_i = \mu_i q_{r\ell_{\tilde{r}}} + \tilde{r}$. Notice that for any fixed integer ℓ, the $r + \ell\lambda_i$ with $0 \le r \le \mu_i - 1$ have distinct images modulo μ_i. Thus for each $0 \le \tilde{r} \le \mu_i - 1$ and $0 \le \ell_{\tilde{r}} \le \rho$, the integer pair $(r, q_{r\ell_{\tilde{r}}})$ is unique as long as it exists. This means that each c_ℓ appears at most once in every coefficient of $\tilde{R}_{ik}$ with respect to $\mathrm{S}_{(\lambda_i,\mu_i)}$.

For the left quotient, we will only be interested in the one from the left division of LM_{ik} by $\mathrm{S}_y - 1$ and its precise formula is summarized in the following remark. Note that formulas for other left quotients are also available if desired.

Remark 4.2. *For each $1 \le i \le m$ and $1 \le k \le d_i$, if $\lambda_i \ge 0$ then*

$$\mathit{LeftQuot}(LM_{ik}, \mathrm{S}_y - 1) = \sum_{\alpha=0}^{\nu_{in_i}+\rho_0\lambda_i-\mu_i} \left(\sum_{(\ell_\alpha, j_\alpha)} c_{\ell_\alpha} \sigma_y^{-q_{\ell_\alpha, j_\alpha}} \sigma_x^{\ell_\alpha}(a_{ij_\alpha k}) \right) \mathrm{S}^{\alpha}_{(\lambda_i,\mu_i)},$$

where the inner summation runs over all integer pairs (ℓ_α, j_α) in $\{0, \ldots, \rho\} \times \{1, \ldots, n_i\}$ satisfying the property that there exists a positive integer $q_{\ell_\alpha, j_\alpha}$ such that $\nu_{ij_\alpha} + \ell_\alpha \lambda_i = \mu_i q_{\ell_\alpha, j_\alpha} + \alpha$. Otherwise,

$$\begin{aligned} &\mathit{LeftQuot}(LM_{ik}, \mathrm{S}_y - 1) \\ &= -\sum_{\alpha=-1}^{\nu_{i1}+\rho_0\lambda_i} \left(\sum_{(\ell_\alpha, j_\alpha)} c_{\ell_\alpha} \sigma_y^{-q_{\ell_\alpha, j_\alpha}} \sigma_x^{\ell_\alpha}(a_{ij_\alpha k}) \right) \mathrm{S}^{\alpha}_{(\lambda_i,\mu_i)} + \sum_{\alpha=0}^{\nu_{in_i}-\mu_i} \left(\sum_{(\ell_\alpha, j_\alpha)} c_{\ell_\alpha} \sigma_y^{-q_{\ell_\alpha, j_\alpha}} \sigma_x^{\ell_\alpha}(a_{ij_\alpha k}) \right) \mathrm{S}^{\alpha}_{(\lambda_i,\mu_i)}, \end{aligned}$$

where the first inner summation runs over all integer pairs (ℓ_α, j_α) in $\{0, \ldots, \rho\} \times \{1, \ldots, n_i\}$ satisfying the property that there exists a nonpositive integer $q_{\ell_\alpha, j_\alpha}$ so that $\nu_{ij_\alpha} + \ell_\alpha \lambda_i = \mu_i q_{\ell_\alpha, j_\alpha} + \alpha$, while the second inner summation runs over all integer pairs (ℓ_α, j_α) in $\{0, \ldots, \rho\} \times \{1, \ldots, n_i\}$ satisfying the property that there exists a positive integer $q_{\ell_\alpha, j_\alpha}$ so that $\nu_{ij_\alpha} + \ell_\alpha \lambda_i = \mu_i q_{\ell_\alpha, j_\alpha} + \alpha$. Note that the left quotients may become dense although the LM_{ik} are sparse operators.

The following corollary is an immediate result of Theorem 4.1 and the equation (4.6).

Corollary 4.3. *With the assumptions and notations of Theorem 4.1, further assume that (4.5) holds. Then a nonzero operator $L = \sum_{\ell=0}^{\rho} c_\ell\, \sigma_x^\ell(\delta)\, \mathrm{S}_x^\ell \in C[x][\mathrm{S}_x]$ is a telescoper for $f/g \in C(x, y)$ if and only if*

$$\sum_{\ell_{\tilde r}} c_{\ell_{\tilde r}}\, \sigma_y^{-q_{r\ell_{\tilde r}}} \sigma_x^{\ell_{\tilde r}}(\tilde a_{ikr}) = 0 \quad \text{for all } i = 1, \ldots, m,\ k = 1, \ldots, d_i \text{ and } \tilde r = 0, \ldots, \mu_i - 1, \tag{4.7}$$

where the summation, for each $(i, k, \tilde r)$, runs over all integers $\ell_{\tilde r}$ in $\{0, \ldots, \rho\}$ satisfying the property that there exist integers $r, q_{r\ell_{\tilde r}}$ such that $0 \le r \le \mu_i - 1$ and $r + \ell_{\tilde r}\lambda_i = \mu_i q_{r\ell_{\tilde r}} + \tilde r$.

Using the well-known fact that any linear system with more variables than equations admits a nontrivial solution, one obtains an upper bound for the order of minimal telescopers, which coincides with the known bound given in (Huang, 2016, Theorem 5.5) for "generic" rational functions since the known bound is already generically sharp. However, there are instances in which our bound is better than the known one (e.g. Example 4.9 versus Example 6.3).

Lemma 4.4. *With the assumptions and notations of Theorem 4.1, further assume that (4.5) holds. Then the order of a minimal telescoper for f/g is at most*

$$\rho_0 = \sum_{i=1}^{m} \sum_{k=1}^{d_i} \mu_i \cdot \left(\max_{0 \le r \le \mu_i - 1} \{-1, \deg_y(\tilde a_{ikr})\} + 1 \right). \tag{4.8}$$

Proof. Let $L \in C[x][\mathrm{S}_x]$ be a minimal telescoper for f/g of order $\rho \in \mathbb{N}$. As before, we may assume without loss of generality that L is of the form $\sum_{\ell=0}^{\rho} c_\ell \sigma_x^\ell(\delta)\, \mathrm{S}_x^\ell$. We then conclude from Corollary 4.3 that (4.7) holds, which in turn can be viewed as a linear system over $C[x]$ with $\rho + 1$ unknowns $c_0, \ldots, c_\rho$ by equating coefficients of like powers of y to zero. Meanwhile, this linear system contains at most ρ_0 equations over $C[x]$ with ρ_0 given by (4.8). Hence this system has a nontrivial solution over $C[x]$ whenever $\rho + 1$ beats ρ_0, which gives rise to a telescoper for f/g by Corollary 4.3. This thus yields that $\rho \le \rho_0$, concluding the proof. ∎

4.2. A new creative telescoping algorithm

Putting things together, we obtain a new creative telescoping algorithm for rational functions.

RationalOperatorCT (upper-bound version)**.** Given two coprime polynomials $f, g \in C[x, y]$ with $g \neq 0$, compute a minimal telescoper $L \in C[x][\mathrm{S}_x]$ for f/g and a corresponding certificate $h \in C(x, y)$ if telescopers exist. The steps are:

1. compute the integer-linear decomposition (3.1) of g.
2. compute the decomposition (4.1) of f/g with b being the integer-linear part of g.
3. apply the GGSZ reduction to f_0/P_0 to find $h, r \in C(x, y)$ with h of a sparse form such that

$$\frac{f_0}{P_0} = (\mathrm{S}_y - 1)(h) + r. \tag{4.9}$$

4. if $r \neq 0$ then return "No telescoper exists!".
5. compute the refined version (3.3) of the integer-linear decomposition of b.
6. compute the partial fraction decomposition (4.2) of a/b with respect to (3.3) over $C(x)[y]$.
7. for $i = 1, \ldots, m$ and $k = 1, \ldots, d_i$ do

   7.1 set $\tilde{a}_{ik0} = \tilde{a}_{ik1} = \cdots = \tilde{a}_{ik,\mu_i - 1} = 0$.

   7.2 for $j = 1, \ldots, n_i$ compute $q, r \in \mathbb{Z}$ with $0 \leq r \leq \mu_i - 1$ such that $\nu_{ij} = \mu_i q + r$ and update $\tilde{a}_{ikr} = \tilde{a}_{ikr} + \sigma_y^{-q}(a_{ijk})$.
8. if all $\tilde{a}_{ikr} = 0$ then set $L = 1$ and update

$$h = h + \sum_{i=1}^{m} \sum_{k=1}^{d_i} \mathrm{LeftQuot}(M_{ik}, \mathrm{S}_y - 1)\left(\frac{1}{p_i(\lambda_i x + \mu_i y)^k}\right),$$

   and return L, h.
9. set ρ_0 by (4.8).
10. make an ansatz $L = \sum_{\ell=0}^{\rho_0} c_\ell \sigma_x^\ell(\delta)\, \mathrm{S}_x^\ell$ with the c_ℓ indeterminates.

    for $i = 1, \ldots, m$ and $k = 1, \ldots, d_i$ do

    10.1 set $R_{ik}^{(0)} = R_{ik}^{(1)} = \cdots = R_{ik}^{(\mu_i - 1)} = 0$.

    10.2 for $\ell = 0, \ldots, \rho_0$ and $r = 0, \ldots, \mu_i - 1$ compute $q, \tilde{r} \in \mathbb{Z}$ such that $r + \ell\lambda_i = \mu_i q + \tilde{r}$ and $0 \leq \tilde{r} \leq \mu_i - 1$, and update

$$R_{ik}^{(\tilde{r})} = R_{ik}^{(\tilde{r})} + c_\ell \sigma_y^{-q} \sigma_x^\ell(\tilde{a}_{ikr}).$$

11. find a nontrivial solution $(c_0, \ldots, c_{\rho_0}) \in C[x]^{\rho_0 + 1}$ with the last nonzero entry having the smallest possible index so that $R_{ik}^{(\tilde{r})} = 0$ for $i = 1, \ldots, m$, $k = 1, \ldots, d_i$ and $\tilde{r} = 0, \ldots, \mu_i - 1$.
12. update $h = L(h) + \sum_{i=1}^{m} \sum_{k=1}^{d_i} \mathrm{LeftQuot}(LM_{ik}, \mathrm{S}_y - 1)\left(\frac{1}{p_i(\lambda_i x + \mu_i y)^k}\right)$, and return L, h.

Theorem 4.5. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$. Then the above algorithm* **RationalOperatorCT** *correctly finds a minimal telescoper for f/g and a corresponding certificate when such telescopers exist.*

Proof. By (Gerhard et al., 2003, Theorem 12) and (Abramov, 2003, Theorem 10), steps 1-4 are correct. For the case where all $\tilde{a}_{ikr} = 0$ in step 8, Theorem 4.1 implies that one is a minimal telescoper for a/b (and thus for f/g). If this is not the case, then by Theorem 4.1, Corollary 4.3 and Lemma 4.4, together with the discussions in between, the algorithm eventually returns a telescoper for f/g of order ρ with $0 < \rho \leq \rho_0$ and also a corresponding certificate. It remains to show that ρ is indeed the minimal order. Let $\tilde{\rho}$ be the order of a minimal telescoper for f/g and suppose that $\tilde{\rho}$ is strictly less than ρ. Notice that there is a minimal telescoper for f/g of the form $\sum_{i=0}^{\tilde{\rho}} \tilde{c}_\ell \, \sigma_x^\ell(\delta) \, \mathrm{S}_x^\ell$ with $\tilde{c}_\ell \in C[x]$ and $\tilde{c}_{\tilde{\rho}} \neq 0$. It follows from Corollary 4.3 that $(\tilde{c}_0, \ldots, \tilde{c}_{\tilde{\rho}}, 0, \ldots, 0) \in C[x]^{\rho_0+1}$ is also a nontrivial solution of the linear system in step 11, whose last nonzero entry having index $\tilde{\rho}$ since $\tilde{c}_{\tilde{\rho}} \neq 0$. This contradicts with the minimum of ρ stated in step 11 and concludes the proof. $\square$

Remark 4.6. *In step 3, by letting $s \in C[x, y]$ and $t \in C[x]$ be the numerator and denominator of f_0, respectively, it is sufficient to apply the GGSZ reduction to s/P_0 (or even $s/\operatorname{prim}_y(P_0)$) since $s/P_0 = (\mathrm{S}_y - 1)(ht) + rt$ and $rt = 0$ if and only if $r = 0$. This will reduce the total cost of this step.*

Remark 4.7. *Roughly speaking, by replacing step 9 by "for $\rho_0 = 1, 2, \ldots$ do" and iteratively repeating steps 10-11 until a nontrivial solution $(c_0, \ldots, c_{\rho_0}) \in C[x]^{\rho_0+1}$ is found, one obtains an iterative analog of the above algorithm, which will be referred to as the* iterative version.

Remark 4.8. *Under the assumptions of Theorem 4.1, we can also follow an alternate method, suggested in (Le, 2003), for computing a minimal telescoper for f/g. Namely, after obtaining the decomposition (4.2), we separately take each simple fraction*

$$M_{ik}\left(\frac{1}{p_i(\lambda_i x + \mu_i y)^k}\right)$$

and individually compute their minimal telescopers $L_{ik} \in C[x][\mathrm{S}_x]$ and then take the least common left multiple of the L_{ik}. One can show that this least common left multiple gives a minimal telescoper for f/g.

Note that, for each integer pair (i, k) with $1 \leq i \leq m$ and $1 \leq k \leq d_i$, we can compute the minimal telescoper L_{ik} by letting

$$\rho_{ik} = \mu_i \cdot \left(\max_{0 \leq r \leq \mu_i - 1} \{-1, \deg_y(\tilde{a}_{ikr})\} + 1\right) \quad \text{with } \tilde{a}_{ikr} \text{ given in (4.5)}$$

and solving a linear system of ρ_{ik} rows and $\rho_{ik} + 1$ columns. This compares to the algorithm **RationalOperatorCT** *which solves a linear system of $\sum_{i=1}^m \sum_{k=1}^{d_i} \rho_{ik}$ rows and $\sum_{i=1}^m \sum_{k=1}^{d_i} \rho_{ik} + 1$ columns to obtain the final minimal telescoper. However, the alternative method also requires time to compute the least common left multiple of the L_{ik}. Preliminary experiments suggest that in practice such a method does not outperform the algorithm* **RationalOperatorCT** *and it is often less efficient than the iterative version mentioned in Remark 4.7.*

In order to demonstrate our new algorithm, let us work on some examples.

Example 4.9. *Consider the rational function r_1 of the form*

$$r_1 = \frac{2x^2 + 1}{(-5x + 2y)^3 + 1} + \frac{x - 1}{(-5x + 2y + 1)^3 + 1}.$$

By definition, the denominator of r_1 is integer-linear of a single type $(-5, 2)$. Thus our algorithm proceeds to step 5, calculating the refined integer-linear decomposition of the denominator

$$p_1(-5x + 2y) \cdot p_1(-5x + 2y + 1) \quad \text{with } p_1(z) = z^3 + 1,$$

with respect to which, step 6 shows that r_1 admits the partial fraction decomposition

$$r_1 = \underbrace{((x-1)\,\mathrm{S}_{(-5,2)} + 2x^2 + 1)}_{M_1}\left(\frac{1}{p_1(-5x+2y)}\right) \quad \text{with } \mathrm{S}_{(-5,2)} = \mathrm{S}_x\,\mathrm{S}_y^3.$$

By left division with remainder in $\mathbb{Q}(x)[y][\mathrm{S}_{(-5,2)}]$, we know that M_1 is already the left remainder from its left division by $\mathrm{S}_y - 1$. Hence r_1 has a telescoper of order at most $\rho_0 = 2$ by Lemma 4.4. By making an ansatz $L_1 = \sum_{\ell=0}^{2} c_\ell\,\mathrm{S}_x^\ell$ with c_ℓ indeterminates, steps 10-11 then build a linear system

$$\begin{pmatrix} 2x^2+1 & \sigma_y^2\sigma_x(x-1) & \sigma_y^5\sigma_x^2(2x^2+1) \\ x-1 & \sigma_y^3\sigma_x(2x^2+1) & \sigma_y^5\sigma_x^2(x-1) \end{pmatrix}\begin{pmatrix} c_0 \\ c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{4.10}$$

of which the nullspace has only dimension one. Solving the linear system gives a minimal telescoper

$$L_1 = (4x^4 + 8x^3 + 7x^2 + 5x + 3)\,\mathrm{S}_x^2 + 2(2x^2 - 5)\,\mathrm{S}_x - (4x^4 + 24x^3 + 55x^2 + 59x + 27), \tag{4.11}$$

and step 12 returns the corresponding certificate in the sparse expression

$$h_1 = \mathit{LeftQuot}(L_1 M_1, \mathrm{S}_y - 1)\left(\frac{1}{p_1(-5x+2y)}\right). \tag{4.12}$$

We note that $L_1 M_1$ here is equal to

$$\begin{aligned} &c_2\,\sigma_x^2(2x^2+1)\,\mathrm{S}_{(-5,2)}^{-10} + c_2\,\sigma_x^2(x-1)\,\mathrm{S}_{(-5,2)}^{-9} + c_1\,\sigma_x(2x^2+1)\,\mathrm{S}_{(-5,2)}^{-5} \\ &+ c_1\,\sigma_x(x-1)\,\mathrm{S}_{(-5,2)}^{-4} + c_0\,(2x^2+1) + c_0(x-1)\,\mathrm{S}_{(-5,2)} \end{aligned}$$

and hence is a sparse operator. The left quotient in (4.12), however, is a dense operator of exponents in $\mathrm{S}_{(-5,2)}$ ranging consecutively from -10 to -1.

Example 4.10. *Consider the same rational function f/g as Example 2.3. By Example 3.2, the polynomial g admits the integer-linear decomposition (3.4) with the integer-linear part b satisfying the refined decomposition (3.5). Then step 2 computes the decomposition (4.1) with*

$$\frac{f_0}{P_0} = \frac{2x+3}{\sigma_y^{30}(g_0)} - \frac{2x+3}{\sigma_y^{29}(g_0)} - \frac{1}{\sigma_y(g_0)} + \frac{1}{g_0},$$

to which applying the GGSZ reduction yields (4.9) with h of the sparse form given in (2.4) and $r = 0$. In step 6, one gets that

$$\frac{a}{b} = \underbrace{((x-1)\,\mathrm{S}_{(-5,2)} + 2x^2 + 1)}_{M_1}\left(\frac{1}{p_1(-5x+2y)}\right) + \underbrace{(xy+1)\,\mathrm{S}_{(3,10)}^0}_{M_2}\left(\frac{1}{p_2(3x+10y)}\right),$$

where $S_{(-5,2)} = S_x S_y^3$ *and* $S_{(3,10)} = S_x^7 S_y^{-2}$. *Again, by left division with remainder in* $\mathbb{Q}(x)[y][S_{(-5,2)}]$ *and* $\mathbb{Q}(x)[y][S_{(3,10)}]$, *we know that both* M_1 *and* M_2 *are left remainders from the left divisions by* $S_y - 1$. *It then follows from Lemma 4.4 that* f/g *has a telescoper of order at most* $\rho_0 = 22$. *With the ansatz* $L = \sum_{\ell=0}^{22} c_\ell S_x^\ell$ *for indeterminates* c_ℓ, *steps 10-11 set up a system of* ρ_0 *linear equations in unknowns* c_ℓ, *in which each linear equation is of degree in x at most 2. The only basis to the nullspace of this linear system over* $\mathbb{Q}(x)$ *yields a minimal telescoper*

$$L = (3x^2 + 42x + 82) S_x^{22} - (3x^2 + 30x + 10) S_x^{20} - 2(3x^2 + 72x + 142) S_x^{12} + 2(3x^2 + 60x + 10) S_x^{10} + (3x^2 + 102x + 802) S_x^2 - (3x^2 + 90x + 610), \quad (4.13)$$

and then step 12 returns a corresponding certificate in the sparse expression

$$L(h) + \mathit{LeftQuot}(LM_1, S_y - 1)\left(\frac{1}{p_1(-5x + 2y)}\right) + \mathit{LeftQuot}(LM_2, S_y - 1)\left(\frac{1}{p_2(3x + 10y)}\right). \quad (4.14)$$

5. Arithmetic cost for the new algorithm

In this section, we give a complexity analysis of the new algorithm described in the preceding section. For this purpose, we first collect some classical complexity notations and facts needed in this paper. More background on these can be found in (von zur Gathen and Gerhard, 2013).

5.1. Complexity background

In this paper, the cost of algorithms will be counted by the number of arithmetic operations in the field C. All costs are analyzed in terms of O-estimates for classical arithmetic and $O^\sim$-estimates for fast arithmetic, where the *soft-Oh notation* "$O^\sim$" is basically "O" but suppressing logarithmic factors (see (von zur Gathen and Gerhard, 2013, Definition 25.8) for a precise definition).

We summarize the facts needed for our analysis below and will freely use them in our theorems. For proofs, we refer to (von zur Gathen and Gerhard, 2013), (Gerhard, 2004, §3 and §5) and (Zhou et al., 2012, Theorem 4.1).

The first fact gives sharp degree bounds for two basic arithmetic operations – division with remainder and partial fraction decomposition. This turns out to be very useful in estimating degree sizes. The proofs are mainly based on Cramer's rule and determinant expansions and will be skipped.

Fact 5.1 (Degree bounds)**.** *Let* f, g *be two nonzero polynomials in* $C[x, y]$.

(i) Assume that $\deg_y(f) \geq \deg_y(g)$. *Then there exist unique* $q, r \in C[x, y]$ *with*

$$(\deg_x(q), \deg_y(q)) \leq ((\deg_y(f) - \deg_y(g)) \deg_x(g) + \deg_x(f), \deg_y(f) - \deg_y(g))$$

$$\text{and} \quad (\deg_x(r), \deg_y(r)) \leq ((\deg_y(f) - \deg_y(g) + 1) \deg_x(g) + \deg_x(f), \deg_y(g) - 1)$$

such that $\mathrm{lc}_y(g)^{\deg_y(f) - \deg_y(g) + 1} f = qg + r$.

(ii) *Assume that* $\deg_y(f) < \deg_y(g)$ *and* $g = g_1^{e_1} \ldots g_m^{e_m}$ *with* $e_i \in \mathbb{N} \setminus \{0\}$ *and* $g_i \in C[x, y]$ *being pairwise coprime. Then there exists* $\delta \in C[x]$ *and* $\{f_{ij}\}_{1 \le i \le m, 1 \le j \le e_i} \subseteq C[x, y]$ *with*

$$\deg_x(\delta) \le \deg_x(g)\deg_y(g) - \sum_{i=1}^{m} \frac{e_i(1 + e_i)}{2} \deg_x(g_i)\deg_y(g_i)$$

and $\quad (\deg_x(f_{ij}), \deg_y(f_{ij})) \le (\deg_x(g)\deg_y(g) + \deg_x(f) - \deg_x(g) + j\deg_x(g_i), \deg_y(g_i) - 1)$

such that

$$\frac{f}{g} = \frac{1}{\delta}\left(\frac{f_{11}}{g_1} + \cdots + \frac{f_{1e_1}}{g_1^{e_1}} + \cdots + \frac{f_{m1}}{g_m} + \cdots + \frac{f_{me_m}}{g_m^{e_m}}\right).$$

The next fact contains the cost of some basic arithmetics for univariate polynomials.

Fact 5.2 (Arithmetic of univariate polynomials)**.** *Let* $f, g \in C[x]$ *with* $\deg_x(f), \deg_x(g) \le d_x$. *Then the following operations can be performed at most in* $\mathrm{O}(d_x^2)$ *arithmetic operations in* C *with classical arithmetic and* $\mathrm{O}^{\sim}(d_x)$ *with fast arithmetic.*

(i) *Addition, multiplication, division with remainder, GCD computation of* f *and* g*;*

(ii) *Evaluation* f *at* $d_x + 1$ *distinct points in* C *or interpolation in* $C[x]$ *at these points;*

(iii) *Squarefree decomposition of* f *over* $C[x]$*;*

(iv) *Partial fraction decomposition of* f/g *with respect to a given factorization of* g*, provided that* f, g *are nonzero coprime polynomials with* $\deg_x(f) < \deg_x(g)$*.*

In order to analyze the cost for operations on bivariate polynomials, a general (although not optimal) technique is to use evaluation and interpolation on polynomials and to perform operations on univariate polynomials based on the above fact. We will frequently use this technique without explicitly pointing it out.

As mentioned the introduction, most of recent creative telescoping algorithms, including our new one presented in Section 4, eventually reduce the problem of finding telescopers to the problem of solving linear systems, which can be accomplished efficiently.

Fact 5.3 (Solving linear systems)**.** *Let* M *be a polynomial matrix in* $C[x]^{m \times n}$ *with entries being polynomials in* $C[x]$ *of degree in* x *less than* d_x*. Assume that* $n \in \mathrm{O}(m)$*. Then a basis of the null space of* M *in* $C[x]$ *can be computed using* $\mathrm{O}(m^3 d_x^2)$ *arithmetic operations in* C *with classical arithmetic (Gaussian elimination) and* $\mathrm{O}^{\sim}(m^{\omega-1} n d_x)$ *with fast arithmetic, where* $\omega \in \mathbb{R}$ *with* $2 < \omega \le 3$ *is the exponent of matrix multiplication over* C*.*

5.2. Output size estimates

In order to analyze the complexity of our new algorithm, we first need to estimate sizes of intermediate results.

Using Fact 5.1, one gets size estimates for the output of the GGSZ reduction. It turns out that the sizes of the summable part depend on the dispersion of the denominator of the input rational function. Recall that the *dispersion* of a polynomial $g \in C[x, y]$ with respect to y, denoted by $\mathrm{dis}_y(g)$, is the maximal integer ℓ with the property that g and $\sigma_y^{\ell}(g)$ have a common divisor of positive degree in y. For simplicity, we confine ourselves to O-estimates and omit the simple yet tedious proof.

Lemma 5.4. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$ and $\deg_y(f), \deg_y(g) \leq d_y$. Assume that applying the GGSZ reduction to f/g yields*

$$\frac{f}{g} = (\mathrm{S}_y - 1)\left(\frac{s}{t}\right) + \frac{a}{\delta b},$$

where $\delta \in C[x]$, $s, t, a, b \in C[x, y]$ with $\deg_y(s) < \deg_y(t)$, $\deg_y(a) < \deg_y(b)$ and b being primitive and shift-free with respect to y. Then

$$\begin{aligned}
&\deg_x(\delta) \in \mathrm{O}(\deg_x(g) \cdot d_y),\\
&(\deg_x(b), \deg_y(b)) \in \mathrm{O}(\deg_x(g)) \times \mathrm{O}(d_y),\\
&(\deg_x(a), \deg_y(a)) \in \mathrm{O}(\deg_x(g) \cdot d_y + \deg_x(f)) \times \mathrm{O}(d_y),\\
&(\deg_x(s), \deg_y(s)) \in \mathrm{O}(\deg_x(g) \cdot d_y + \deg_x(f) + \mathrm{dis}_y(g) \deg_x(g)) \times \mathrm{O}(\mathrm{dis}_y(g) \cdot d_y),\\
\text{and} \quad &(\deg_x(t), \deg_y(t)) \in \mathrm{O}(\deg_x(g) \cdot d_y + \mathrm{dis}_y(g) \deg_x(g)) \times \mathrm{O}(\mathrm{dis}_y(g) \cdot d_y).
\end{aligned}$$

Lemma 5.5. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$ and $\deg_y(f), \deg_y(g) \leq d_y$. Assume that f/g can be decomposed as (4.1), in which a/b further admits the decomposition (4.2) and (4.5) holds. Let $s \in C[x, y]$ and $t \in C[x]$ be the numerator and denominator of f_0, respectively. Then*

$$\begin{aligned}
&\deg_x(\delta) \in \mathrm{O}(\deg_x(g) \cdot d_y),\\
&(\deg_x(\tilde{a}_{ijk}), \deg_y(\tilde{a}_{ijk})) \leq (\deg_x(a_{ijk}), \deg_y(a_{ijk})) \in \mathrm{O}(\deg_x(g) \cdot d_y + \deg_x(f)) \times \mathrm{O}(\deg_z(p_i)),\\
&(\deg_x(s), \deg_y(s)) \in \mathrm{O}(\deg_x(g) \cdot d_y + \deg_x(f)) \times \mathrm{O}(d_y),\\
&\deg_x(t) \in \mathrm{O}(\deg_x(g) \cdot d_y) \quad \text{and} \quad (\deg_x(P_0), \deg_y(P_0)) \in \mathrm{O}(\deg_x(g)) \times \mathrm{O}(d_y).
\end{aligned}$$

Proof. It is readily seen from (4.1)-(4.2) that

$$\frac{f}{g} = \frac{f_0}{P_0} + \frac{1}{\delta(x)} \sum_{i=1}^{m} \sum_{j=1}^{n_i} \sum_{k=1}^{e_{ij}} \frac{a_{ijk}(x, y)}{p_i(\lambda_i x + \mu_i y + \nu_{ij})^k}.$$

Since P_0 and the $p_i(\lambda_i x + \mu_i y + \nu_{ij})$ are pairwise coprime, the degree bounds then follows straightforwardly from Fact 5.1. □

The following depicts an order-degree curve of telescopers for bivariate rational functions.

Theorem 5.6. *With the assumptions and notations of Lemma 5.5, let ρ_0 be defined by (4.8). For each $1 \leq i \leq m$ and $1 \leq k \leq d_i$, let $\alpha_{ik} = \max_{0 \leq r \leq \mu_i - 1}\{-1, \deg_{x,y}(\tilde{a}_{ikr})\}$ and $\beta_{ik} = \max_{0 \leq r \leq \mu_i - 1}\{-1, \deg_y(\tilde{a}_{ikr})\}$. Then for any nonnegative integer pair (ρ, τ) with $\rho \geq \rho_0$ and*

$$\tau > \deg_x(\delta) - 1 + \frac{\sum_{i=1}^{m} \sum_{k=1}^{d_i} \mu_i(\alpha_{ik} - \frac{1}{2}\beta_{ik})(\beta_{ik} + 1)}{\rho + 1 - \rho_0}, \tag{5.1}$$

there exists a telescoper for f/g of order at most ρ and degree at most τ. In particular, letting $\deg_x(f), \deg_x(g) \leq d_x$ and $\deg_y(f), \deg_y(g) \leq d_y$, if $\rho = \rho_0$ then the right hand side of (5.1) is in $\mathrm{O}(\rho_0 d_x d_y)$ while if $\rho = 2\rho_0$ then it belongs to $\mathrm{O}(d_x d_y)$.

Proof. Let $\rho, \tau \in \mathbb{N}$ with $\rho \geq \rho_0$ and τ satisfying (5.1). To prove the lemma, it is sufficient to show that there exist $c_0, \ldots, c_\rho \in C[x]$, not all zero, with $\deg_x(c_\ell) \leq \tau - \deg_x(\delta)$ such that (4.7) holds, because then Corollary 4.3 asserts that $\sum_{\ell=0}^{\rho} c_\ell \, \sigma_x^\ell(\delta) \, \mathrm{S}_x^\ell$ gives a desired telescoper for f/g. Now we consider the linear system over C (instead of $C[x]$) obtained by vanishing coefficients of like powers of x and y in (4.7). In other words, we view the coefficients of the c_ℓ with respect to x, not the c_ℓ themselves, as unknowns. This then gives us $(\tau - \deg_x(\delta) + 1)(\rho + 1)$ unknowns in total. On the other hand, we see that each equation in (4.7) has total degree in x, y at most $\tau - \deg_x(\delta) + \alpha_{ik}$ and degree in y at most β_{ik}. It follows that the resulting linear system contains at most

$$(\tau - \deg_x(\delta) + 1)\rho_0 + \sum_{i=1}^{m} \sum_{k=1}^{d_i} \mu_i(\alpha_{ik} - \frac{1}{2}\beta_{ik})(\beta_{ik} + 1)$$

equations over C. Since $\rho \geq \rho_0$, one concludes from (5.1) that the linear system over C resulting from (4.7) have more unknowns than equations, assuring such a nontrivial solution. $\square$

Remark 5.7. *Under the assumptions of the above theorem, in the context of (Chen and Kauers, 2012a, §4), all $\tilde{a}_{ikr}$ are actually in $C[x]$, yielding $\alpha_{ik} = \max_{0 \leq r \leq \mu_i - 1}\{-1, \deg_x(\tilde{a}_{ik})\}$ and $\beta_{ik} = 0$. Thus $\rho_0 = \sum_{i=1}^{m} \sum_{k=1}^{d_i} \mu_i$ and (5.1) becomes*

$$\tau > \deg_x(\delta) - 1 + \frac{\sum_{i=1}^{m} \sum_{k=1}^{d_i} \mu_i \alpha_{ik}}{\rho + 1 - \rho_0},$$

which in fact coincides with the sharp order-degree curve given in (Chen and Kauers, 2012a, Theorem 10) (after correcting the typos in the formula of the lower bound for d there).

5.3. Cost analysis of algorithm

Lemma 5.8. *A shiftless decomposition of a polynomial $g \in C[x, y]$ with $\deg_x(g) = d_x$ and $\deg_y(g) = d_y$ can be computed using*

(i) $\mathrm{O}(d_y^4)$ *arithmetic operations in $C(x)$ with classical arithmetic and* $\mathrm{O}^\sim(d_y^3)$ *with fast arithmetic, or*

(ii) $\mathrm{O}(d_x d_y^4 + d_x^2 d_y)$ *arithmetic operations in C with classical arithmetic and* $\mathrm{O}^\sim(d_x d_y^3)$ *with fast arithmetic,*

plus the cost of computing the auto-dispersion set of g.

Proof. (i) See (Gerhard et al., 2003, Theorem 10) for the proof.
(ii) Notice that the c and all g_i are factors of g and thus $\deg_x(c), \deg_x(g_i) \leq d_x$. The claimed cost then follows by making use of evaluation-interpolation techniques. $\square$

With the help of the above lemma, in analogy to (Gerhard et al., 2003, Theorem 13), we derive the cost of the GGSZ reduction recalled in Section 2.

Theorem 5.9. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$ and $\deg_y(f), \deg_y(g) \leq d_y$. Then the GGSZ reduction computes $h, r \in C(x, y)$ with h of a sparse form such that (2.2) holds, using $\mathrm{O}(\deg_x(g) d_y^4 + \deg_x(g)^2 d_y^3 + \deg_x(f) \deg_x(g) d_y^2 + \deg_x(f)^2 d_y)$ arithmetic operations in C with classical arithmetic and $\mathrm{O}^\sim(\deg_x(g) d_y^3 + \deg_x(f) d_y)$ with fast arithmetic, plus the cost of computing the auto-dispersion set of g.*

Proof. By (Gerhard et al., 2003, Theorem 13), the cost of the GGSZ reduction is dominated by computing the shiftless decomposition of g and one subsequent partial fraction decomposition of f/g (in terms of the variable y). Note that the latter operation takes $O(\deg_x(g)^2 d_y^3 + \deg_x(f)\deg_x(g)d_y^2 + \deg_x(f)^2 d_y)$ with classical arithmetic and $\tilde{O}(\deg_x(g)d_y^2 + \deg_x(f)d_y)$ with fast arithmetic. This, together with Lemma 5.8 (ii), concludes the claimed cost. □

If the expanded form of the h in (2.2) is expected, then a direct yet cumbersome calculation based on Lemma 5.4 shows that the whole algorithm can be accomplished in time polynomial in the size of the final output.

Corollary 5.10. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$ and $\deg_y(f), \deg_y(g) \leq d_y$. Then the GGSZ reduction returns expanded output using*

$$\begin{aligned}O(&\operatorname{dis}_y(g)^3 \deg_x(g)^2 d_y^2 + \operatorname{dis}_y(g)^2 \deg_x(f)\deg_x(g)d_y^2 + \operatorname{dis}_y(g)^2 \deg_x(g)^2 d_y^3\\ &+ \operatorname{dis}_y(g)\deg_x(f)\deg_x(g)d_y^3 + \operatorname{dis}_y(g)\deg_x(f)^2 d_y^2 + \operatorname{dis}_y(g)\deg_x(g)^2 d_y^4)\end{aligned}$$

arithmetic operations in C with classical arithmetic and

$$\tilde{O}(\operatorname{dis}_y(g)^2 \deg_x(g)d_y^2 + \operatorname{dis}_y(g)\deg_x(f)d_y^2 + \operatorname{dis}_y(g)\deg_x(g)d_y^3)$$

with fast arithmetic, plus the cost of computing the auto-dispersion set of g.

Now we are ready to study the cost of the algorithm **RationalOperatorCT**.

Theorem 5.11. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$, $\deg_x(f), \deg_x(g) \leq d_x$ and $\deg_y(f), \deg_y(g) \leq d_y$. Assume that f/g has a telescoper and let ρ be the actual order of its minimal telescopers. Further define ρ_0 by (4.8). Then the algorithm* **RationalOperatorCT** *finds a minimal telescoper for f/g and a corresponding certificate of a sparse form using $O(d_x d_y^4 + \rho_0 d_x^2 d_y^3 + \rho_0^3 d_x^2 d_y^2)$ arithmetic operations in C with classical arithmetic and $\tilde{O}(d_x d_y^3 + \rho_0 d_x d_y^2 + \rho_0^\omega d_x d_y)$ with fast arithmetic, plus the cost of computing auto-dispersion sets and finding rational roots.*

Proof. According to (Giesbrecht et al., 2019, Theorem 3.5), step 1 takes $O(d_x^2 d_y + d_x d_y^3 + d_y^4)$ arithmetic operations with classical arithmetic and $\tilde{O}(d_x d_y^2 + d_y^3)$ with fast arithmetic, plus the cost of finding rational roots. Steps 2 and 6 essentially compute the partial fraction decomposition of f/g with respect to the decomposition $g = P_0 \prod_{i=1}^m \prod_{j=1}^{n_i} p_i(\lambda_i x + \mu_i y + \nu_{ij})^{e_{ij}}$ over $C(x)[y]$, taking $O(d_x^2 d_y^3)$ with classical arithmetic and $\tilde{O}(d_x d_y^2)$ with fast arithmetic in total. Regardless of the cost of computing auto-dispersion sets, one concludes from Remark 4.6, Lemma 5.5 and Theorem 5.9 that step 3 uses $O(d_x d_y^4 + d_x^2 d_y^3)$ with classical arithmetic and $\tilde{O}(d_x d_y^3)$ with fast arithmetic, dominating the cost of step 5 by Lemma 5.8 (i). Note that by assumption, $r = 0$ and thus the algorithm continues after step 4.

In step 7, the computation of the $\tilde{a}_{ikr}$ dominates the cost of other steps. By Lemma 5.5, each nonzero polynomial a_{ijk} has degree in x in $O(d_x d_y)$ and degree in y in $O(\deg_z(p_i))$. This implies that computing each nonzero term $\sigma_y^{-q}(a_{ijk})$ requires $O(d_x^2 d_y^2 \deg_z(p_i) + d_x d_y \deg_z(p_i)^2)$ with classical arithmetic and $\tilde{O}(d_x d_y \deg_z(p_i))$ with fast arithmetic. Observe that $a_{ijk} = 0$ for $k > e_{ij}$. Therefore, step 7 in total takes $O(\rho_0 d_x^2 d_y^3)$ with classical arithmetic and $\tilde{O}(\rho_0 d_x d_y^2)$ with fast arithmetic. By postponing the computation of the $\sigma_x^\ell(\delta)$ to step 12, one obtains the same cost for step 10 along similar lines as above.

Step 8 deals with the trivial case and takes no arithmetic operations in C. In step 9, computing the number ρ_0 takes linear time $O(d_y)$ since $m \le \deg_y(b) \le d_y$. In step 11, the coefficient matrix with respect to y produced by the linear system has at most ρ_0 rows and $\rho_0 + 1$ columns, whose each nonzero entry is of degree in x in $O(d_x d_y)$ according to Lemma 5.5 and rank is exactly the actual order ρ of minimal telescopers for f/g. Thus Fact 5.3 implies that step 11 needs $O(\rho_0^3 d_x^2 d_y^2)$ with classical arithmetic and $O^\sim(\rho_0^\omega d_x d_y)$ with fast arithmetic. In the final step 12, it virtually takes no arithmetic operations for returning h in this sparse representation, while computing the $\sigma_x^\ell(\delta)$ in the minimal telescoping L requires $O(\rho d_x^2 d_y^2)$ with classical arithmetic and $O^\sim(\rho d_x d_y)$ with fast arithmetic. In addition, since each c_ℓ is of degree in x in $O(\rho_0 d_x d_y)$ by Theorem 5.6, further expanding the coefficients of the telescoper L takes $O(\rho_0 d_x^2 d_y^2)$ with classical arithmetic and $O^\sim(\rho_0 d_x d_y)$ with fast arithmetic. The announced cost follows. □

Corollary 5.12. *With the assumptions of Lemma 5.5, further assume that* $\deg_x(f), \deg_x(g) \le d_x$*. Let* ρ_0 *be defined by* (4.8) *and let* $\mu = \max\{\mu_1, \dots, \mu_m\}$*. Then* $\rho_0 \in O(\mu d_y)$*, and the algorithm* **RationalOperatorCT** *takes* $O(\mu^3 d_x^2 d_y^5)$ *arithmetic operations in* C *with classical arithmetic and* $O^\sim(\mu^\omega d_x d_y^{\omega+1})$ *with fast arithmetic, plus the cost of computing auto-dispersion sets and finding rational roots.*

Proof. By assumption, the integer-linear part b of g satisfies the refined decomposition (3.3). Thus $\deg_y(b) = \sum_{i=1}^m \sum_{j=1}^{n_i} e_{ij} \deg_z(p_i) \le d_y$. It follows from the equation (4.5) that $\deg_y(\tilde{a}_{ikr}) \le \max_{1\le j\le n_i}\{\deg_y(a_{ijk})\} \le \deg_z(p_i) - 1$ and thus $\rho_0 \in O(\mu d_y)$. Now let ρ be the actual order of minimal telescopers for f/g. Then we conclude from Lemma 4.4 that $\rho \le \rho_0 \in O(\mu d_y)$. Thus the announced cost is evident by Theorem 5.11. □

Remark 5.13. *Under the assumptions of Corollary 5.12, as a direct corollary of Theorem 5.6, the size of a minimal telescoper for* f/g *is in* $O(\mu^2 d_x d_y^3)$*, which is relatively sharp by Remark 5.7. With this size bound, one sees from the above corollary that the algorithm* **RationalOperatorCT** *is almost optimal when the exponent of matrix multiplication* ω *is close to 2.*

Remark 5.14. *Under the assumptions of Corollary 5.12, following from similar arguments as in the proofs of Theorem 5.11 and Corollary 5.12, the iterative version of our algorithm presented in Remark 4.7 takes* $O(\mu^4 d_x^2 d_y^6)$ *arithmetic operations in* C *with classical arithmetic and* $O^\sim(\mu^{\omega+1} d_x d_y^{\omega+2})$ *with fast arithmetic, plus the cost of computing auto-dispersion sets and finding rational roots.*

Remark 5.15. *Under the assumptions of Corollary 5.12, in the particular case where* $C = \mathbb{Q}$*, one can assume without loss of generality that* $f, g \in \mathbb{Z}[x, y]$*. Then by incorporating the cost of computing the auto-dispersion set of an integer polynomial (cf. (Gerhard et al., 2003, Theorem 14)) and the cost of finding rational roots of an integer polynomial (cf. (von zur Gathen and Gerhard, 2013, Theorem 15.21)), one sees from the above corollary that the algorithm* **RationalOperatorCT***, as well as its iterative version, has the total running time bounded by* $(\mu + d_x + d_y + \log \|f\|_\infty + \log \|g\|_\infty)^{O(1)}$ *word operations, where the max-norm of* $f = \sum_{i,j\ge 0} f_{ij} x^i y^j \in \mathbb{Z}[x, y]$ *is defined as* $\|f\|_\infty = \max_{i,j\le 0} |f_{ij}|$*. See (Gerhard, 2004; von zur Gathen and Gerhard, 2013) for more information on word operations.*

Similar to Corollary 5.10, by a straightforward and cumbersome calculation, one confirms that the expanded form of the certificate can be obtained in time polynomial in the size of the final result.

Corollary 5.16. *With the assumptions of Lemma 5.5, further assume that* $\deg_x(f), \deg_x(g) \le d_x$. *Let* $L \in C[x][S_x]$ *be a minimal telescoper for* f/g *of order* ρ *with a corresponding certificate* $h \in C(x,y)$. *Define* ρ_0 *by* (4.8)*, and let* $\mu = \max\{\mu_1, \ldots, \mu_m\}$*,* $\xi = \max_{1\le i\le m}\{\nu_{in_i} + \rho|\lambda_i| - \mu_i\}$ *and* $\xi_0 = \mathrm{dis}_y(P_0)$.

(i) Let $h_n, h_d \in C[x,y]$ *be the numerator and denominator of* h*. Then*

$$\begin{aligned} &(\deg_x(h_n), \deg_y(h_n)) \in \mathrm{O}(\xi d_x + \rho\xi_0 d_x + \rho_0 d_x d_y) \times \mathrm{O}(\xi d_y + \rho\xi_0 d_y) \\ \textit{and}\quad &(\deg_x(h_d), \deg_y(h_d)) \in \mathrm{O}(\xi d_x + \rho\xi_0 d_x + \rho d_x d_y) \times \mathrm{O}(\xi d_y + \rho\xi_0 d_y). \end{aligned}$$

(ii) The telescoper L *along with the expanded certificate* h *can be found using*

$$\mathrm{O}(\xi_0^3\mu^3 d_x^2 d_y^4 + \xi_0^3\mu^3 d_x d_y^5 + \xi_0^2\mu^3 d_x^2 d_y^5 + \xi_0^2\mu^3 d_x d_y^6 + \xi_0\mu^3 d_x^2 d_y^6 + \xi^3 d_x^2 d_y^2 + \xi^2\mu d_x^2 d_y^4 + \xi\mu^2 d_x^2 d_y^6)$$

arithmetic operations in C *with classical arithmetic and*

$$\mathrm{O}^{\sim}(\xi_0^2\mu^2 d_x d_y^3 + \xi_0\mu^2 d_x d_y^4 + \xi^2 d_x d_y^2 + \xi\mu^2 d_x d_y^4)$$

with fast arithmetic, plus the cost of computing auto-dispersion sets and finding rational roots.

6. Arithmetic cost for the reduction-based approach

In this section, we review the reduction-based creative telescoping algorithm from (Chen et al., 2015) in the case of bivariate rational functions and further analyze its cost in this setting. As indicated by the name of the algorithm, a reduction method plays a crucial role. The original reduction method used by (Chen et al., 2015) in the rational case was first developed by Abramov (1971). Due to its complexity, we instead employ the GGSZ reduction given in Section 2 to carry out all the reduction steps in the algorithm, so as to highlight the more significant discrepancy between this algorithm and the one developed in Section 4.

Before discussing the concrete algorithm, let us first recall some notions. As a generalization of auto-dispersion sets, the *dispersion set* of a polynomial $f \in C[x,y]$ with respect to another polynomial $g \in C[x,y]$ is defined to be the integer set

$$\mathrm{DS}_y(f,g) = \{\ell \in \mathbb{Z} \mid \deg_y(\gcd(f, \sigma_y^\ell(g))) > 0\}.$$

Using modular techniques (cf. (von zur Gathen and Gerhard, 2013, Theorem 6.26)), computing (a superset of) the dispersion set of two bivariate polynomials amounts to finding that of two univariate polynomials, which in turn can be achieved by the algorithm of Man and Wright (1994) or by the procedure **pDispersionSet** from (Gerhard et al., 2003) in the particular case where $C = \mathbb{Q}$.

For a rational function $f \in C(x,y)$, another rational function $r \in C(x,y)$ is called a *shift-remainder* with respect to y (or σ_y*-remainder* for short) of f if $f - r$ is σ_y-summable and r can be written as $r = a/b$ with $a, b \in C[x,y]$, $\deg_y(a) < \deg_y(b)$ and b being σ_y-free. For brevity, we just say that r is a σ_y-remainder if f is clear from the context. Then the GGSZ reduction reduces a rational function to a σ_y-remainder modulo σ_y-summable rational functions. Clearly, any integer shift of a σ_y-remainder with respect to x is again a σ_y-remainder. A rational function in $C(x,y)$ usually has more than one σ_y-remainder and any two of them differ by a σ_y-summable rational function. The following theorem implies that zero is the only σ_y-remainder in the case of a σ_y-summable rational function.

Theorem 6.1 (Chen et al. 2015, Proposition 4.5)**.** *A rational function in $C(x, y)$ is σ_y-summable if and only if any of its σ_y-remainders is zero.*

We summarize below the main idea of the reduction-based algorithm in (Chen et al., 2015). Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$, and assume that applying the GGSZ reduction to f/g yields (2.2). When the existence of telescopers for f/g is guaranteed, that is, when the denominator of r in (2.2) is integer-linear, then the reduction-based algorithm proceeds in an iterative fashion. It begins by fixing the order ρ of a telescoper for f/g, say starting from $\rho = 0$, and then looks for a telescoper of that order; if none exists, it looks for one of the next higher order. This pattern continues until one telescoper is found, with termination assured by the existence.

The key task is then to find a telescoper of the fixed order ρ. In this respect, we make an ansatz

$$L = c_0 + c_1 \mathrm{S}_x + \cdots + c_\rho \mathrm{S}_x^\rho,$$

with $c_\ell \in C[x]$ to be determined. Write $r_0 = r$ and then for $\ell = 1, \ldots, \rho$ iteratively adjusting the σ_y-remainder $\sigma_x(r_{\ell-1})$ with respect to $\sum_{i=0}^{\ell-1} c_i r_i$ by (Chen et al., 2015, Theorem 5.6), along with a subsequent normalization, leads to

$$\sigma_x^\ell\left(\frac{f}{g}\right) = (\mathrm{S}_y - 1)(h_\ell) + r_\ell \quad \text{for } \ell = 0, \ldots, \rho, \tag{6.1}$$

where $h_\ell, r_\ell \in C(x, y)$ with $h_0 = h$ and r_ℓ being a σ_y-remainder. Note that the adjustment of σ_y-remainders is used to make sure that any $C[x]$-linear combination of r_ℓ, particularly $\sum_{\ell=0}^{\rho} c_\ell r_\ell$, is again a σ_y-remainder. A direct calculation shows that

$$L\left(\frac{f}{g}\right) = (\mathrm{S}_y - 1)\left(\sum_{\ell=0}^{\rho} c_\ell h_\ell\right) + \sum_{\ell=0}^{\rho} c_\ell r_\ell.$$

From Theorem 6.1 we see that $\sum_{\ell=0}^{\rho} c_\ell r_\ell = 0$, reducing the problem of telescopers to the simple task of solving a linear system over $C[x]$. More specifically, we set up the system of linear equations equivalent to $\sum_{\ell=0}^{\rho} c_\ell r_\ell = 0$ and solve it for the unknowns c_ℓ over $C[x]$. Any solution with $c_\rho \neq 0$ will give rise to a telescoper for f/g of the desired order ρ. Failing to find such a solution implies that no telescopers of order ρ exist. In this case, we update ρ to $\rho + 1$ and repeat the above process.

In order to complete the description of the reduction-based algorithm, we still need to provide the details about the adjustment of σ_y-remainders, which we will fulfil now. The core algorithm hidden in (Chen et al., 2015, Theorem 5.6) addresses the following problem.

ShiftRemainderAdjustment. Given two nonzero σ_y-remainders $r, r_0 \in C(x, y)$, compute a rational function $\tilde{h} \in C(x, y)$ and another σ_y-remainder $\tilde{r} \in C(x, y)$ such that

$$r = (\mathrm{S}_y - 1)(\tilde{h}) + \tilde{r} \quad \text{and} \quad c_0 r_0 + c_1 \tilde{r} \text{ is a } \sigma_y\text{-remainder for any } c_0, c_1 \in C[x]. \tag{6.2}$$

With the input of two nonzero σ_y-remainders $r, r_0 \in C(x, y)$, let $a, b \in C[x, y]$ be the numerator and denominator of r, respectively, and let $b_0 \in C[x, y]$ be the denominator of r_0. By definition, $\deg_y(a) < \deg_y(b)$ and b, b_0 are both σ_y-free. Using polynomial factorization and the dispersion set, one can uniquely decompose b as

$$c \cdot \sigma_y^{\ell_1}(p_1^{e_1}) \cdots \sigma_y^{\ell_m}(p_m^{e_m}), \tag{6.3}$$

where $c \in C[x]$, $p_i \in C[x,y]$ are distinct monic irreducible polynomials of positive degrees in y, $e_i \in \mathbb{N} \setminus \{0\}$ are multiplicities of $\sigma_y^{\ell_i}(p_i)$ in b, and $\ell_i \in \mathbb{Z}$ satisfy the conditions that (i) $\ell_i = 0$ if and only if $\sigma_y^{\ell}(p_i) \nmid b_0$ for any nonzero integer ℓ; (ii) if $\ell_i \neq 0$ then p_i is a factor of b_0. We refer to (6.3) as the *shift-coprime decomposition* of b with respect to b_0. Note that $\{\ell_1, \ldots, \ell_m\} \setminus \{0\}$ is equal to the integer set $\mathrm{DS}_y(b, b_0) \setminus \{0\}$.

Since b, b_0 are σ_y-free, we know that the polynomials $\sigma_y^{\ell_1}(p_1), \ldots, \sigma_y^{\ell_m}(p_m)$, as well as those $p_1, \ldots, p_m$, in (6.3) are pairwise coprime. By partial fraction decomposition, there exist unique $f_1, \ldots, f_m \in C(x)[y]$ with $\deg_y(f_i) < e_i \deg_y(p_i)$ such that

$$r = \frac{a}{b} = \frac{f_1}{\sigma_y^{\ell_1}(p_1^{e_1})} + \cdots + \frac{f_m}{\sigma_y^{\ell_m}(p_m^{e_m})}.$$

From (Chen et al., 2015, Theorem 5.6) we then have that

$$\tilde{r} = \frac{\sigma_y^{-\ell_1}(f_1)}{p_1^{e_1}} + \cdots + \frac{\sigma_y^{-\ell_m}(f_m)}{p_m^{e_m}}, \tag{6.4}$$

with the corresponding $\tilde{h}$ in (6.2) given by

$$\tilde{h} = \sum_{i=1, \ell_i<0}^{m} \sum_{j=0}^{-\ell_i-1} \sigma_y^{j}\left(-\frac{f_i}{\sigma_y^{\ell_i}(p_i^{e_i})}\right) + \sum_{i=1, \ell_i>0}^{m} \sum_{j=1}^{\ell_i} \sigma_y^{-j}\left(\frac{f_i}{\sigma_y^{\ell_i}(p_i^{e_i})}\right). \tag{6.5}$$

We remark that one may even force the ℓ_i in (6.3) to be distinct by grouping together those factors $p_i^{e_i}$ of the same integer shift order with respect to y and then the process described in the preceding paragraph can still be carried out almost literally. In this way, the computation of the new decomposition (6.3) requires, instead of full polynomial factorization, GCD computation only, provided that the dispersion set is known.

With the adjusted σ_y-remainders at hand, the reduction-based algorithm now works smoothly in the iterative manner as mentioned before.

Remark 6.2. *As already pointed out in (Chen et al., 2019, §5.2), it is actually sufficient to adjust the $\sigma_x(r_{\ell-1})$ for $1 \leq \ell \leq \rho$ with respect to r_0 only (rather than $\sum_{i=0}^{\ell-1} c_i r_i$) so as to insure the $C[x]$-linearity of the σ_y-remainders r_ℓ. This may further reduce the total cost for adjusting σ_y-remainders.*

Let us return to the two examples from Section 4. We will use the above reduction-based algorithm in order to illustrate the difference between the two approaches.

Example 6.3. *Let r_1 be the rational function given in Example 4.9. Then we know that r_1 remains unchanged after applying the GGSZ reduction and it has a telescoper. By Theorem 6.6, we get the upper bound $\rho_0 = 6$ for the order of minimal telescopers for r_1, which exceeds the actual bound since Example 4.9 implies that a minimal telescoper for r_1 has only order 2. Now in the iteration step $\rho = 2$, the reduction-based algorithm finds the adjusted additive decompositions*

$$\sigma_x^{\ell}(r_1) = (\mathrm{S}_y - 1)(h_\ell) + \frac{a_\ell}{b_\ell} \quad \text{for } \ell = 0, 1, 2,$$

where $h_\ell \in \mathbb{Q}(x,y)$, a_ℓ is an integer polynomial with $\deg_x(a_\ell) = 1$ and $\deg_y(a_\ell) = 3$, and $b_\ell = ((-5x+2y)^3 - 1)((-5x+2y)^3 - 1)$. Note that the h_ℓ and a_ℓ are not displayed here for space *reasons. In order to find a $\mathbb{Q}[x]$-linear dependency among the a_ℓ/b_ℓ, we set up a linear system attached by the coefficient matrix (with a scale applied so as to fit in one line)*

$$\begin{pmatrix} 16x^2+8x & 16x^2+40x+24 & 16x^2+72x+80 \\ -120x^3-36x^2+12 & -120x^3-300x^2-168x & -120x^3-516x^2-504x+108 \\ 300x^4+30x^3+12x^2-60x+6 & 300x^4+750x^3+390x^2+6x & 300x^4+1230x^3+1032x^2-492x+54 \\ -250x^5+25x^4-30x^3+79x^2-14x+1 & -250x^5-625x^4-300x^3-13x^2+6x+3 & -250x^5-975x^4-680x^3+559x^2-118x+19 \end{pmatrix}.$$

This linear system admits the same solutions as (4.10), in other words, it leads to the same minimal telescoper as Example 4.9. The corresponding certificate is left as an unnormalized dense sum.

Example 6.4. *Consider the same rational function f/g as Example 2.3. From the same example, we see that f/g satisfies (2.2) with h, r given in (2.4). Moreover, there exist telescopers for f/g since the denominator of r is integer-linear. Again, Theorem 6.6 gives an exceeded upper bound $\rho_0 = 26$ and by Example 4.10, the actual order of a minimal telescoper for f/g is merely 22. Then in the iteration step $\rho = 22$ of the reduction-based algorithm, we obtain the adjusted additive decompositions (6.1), where $h_0 = h$, $r_0 = r$ and other quantities h_ℓ, r_ℓ are not displayed here to save spaces. Finding a $\mathbb{Q}[x]$-linear dependency among the r_ℓ yields a linear system with the coefficient matrix of 26 rows and 23 columns and having entries of degrees in x at most 25. This linear system confirms the same minimal telescoper as Example 4.10, leaving the corresponding certificate as a large, unnormalized dense sum.*

6.1. *Output size estimates*

As before, we use Fact 5.1 to estimate degree sizes of intermediate results. We start by those of the adjusted σ_y-remainders in reduction steps.

Lemma 6.5. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$ and $\deg_y(f), \deg_y(g) \leq d_y$. Let $\rho \in \mathbb{N}$ and assume that f/g admits the decomposition (6.1), where r_0 is the σ_y-remainder obtained by applying the GGSZ reduction to f/g and each r_ℓ for $1 \leq \ell \leq \rho$ is the σ_y-remainder obtained by applying the shift-remainder adjustment to $\sigma_x(r_{\ell-1})$ with respect to r_0. For $0 \leq \ell \leq \rho$, write $r_\ell = a_\ell/(\delta_\ell b_\ell)$ with $\delta_\ell \in C[x]$, $a_\ell, b_\ell \in C[x, y]$, $\deg_y(a_\ell) < \deg_y(b_\ell)$, $\gcd(a_\ell, \delta_\ell b_\ell) = 1$ and b_ℓ being primitive and shift-free with respect to y. Then for $0 \leq \ell \leq \rho$, the following bounds hold:*

$$\deg_x(\delta_\ell) \leq \deg_x(\delta_0) + \deg_x(b_0)\deg_y(b_0), \quad (\deg_x(b_\ell), \deg_y(b_\ell)) = (\deg_x(b_0), \deg_y(b_0)),$$

$$\text{and} \quad (\deg_x(a_\ell), \deg_y(a_\ell)) \leq \Big(\deg_x(a_0) + \deg_x(b_0)\deg_y(b_0), \deg_y(b_0) - 1\Big).$$

Moreover,

$$\deg_x(\delta_\ell) \in \mathrm{O}(\deg_x(g)\cdot d_y), \quad (\deg_x(b_\ell), \deg_y(b_\ell)) \in \mathrm{O}(\deg_x(g)) \times \mathrm{O}(d_y),$$

$$\text{and} \quad (\deg_x(a_\ell), \deg_y(a_\ell)) \in \mathrm{O}(\deg_x(g)\cdot d_y + \deg_x(f)) \times \mathrm{O}(d_y).$$

Proof. The assertions for $\ell = 0$ are evident by assumption and Lemma 5.4. Now let ℓ be an arbitrary but fixed integer with $1 \leq \ell \leq \rho$. Assume that $\sigma_x(b_{\ell-1})$ admits the shift-coprime decomposition of the form (6.3) with respect to b_0. Then $c \in C$ in (6.3) since $b_{\ell-1}$ is primitive with respect to y. By (Huang, 2016, Proposition 5.2), there is a one-to-one correspondence between the multisets of monic irreducible factors of positive degrees in y of $b_{\ell-1}$ and $\sigma_x^{\ell-1}(b_0)$

such that for any such a factor q of $b_{\ell-1}$, there exists a unique factor q' of $\sigma_x^{\ell-1}(b_0)$ with the same multiplicity as q in $b_{\ell-1}$ satisfying $q \sim_y q'$, and vice versa. It then follows that

$$\sigma_x^{\ell}(b_0) = c' \cdot \sigma_y^{\ell'_1}(p_1^{e_1}) \cdots \sigma_y^{\ell'_m}(p_m^{e_m}) \quad \text{for some } c' \in C \text{ and } \ell'_i \in \mathbb{Z}. \tag{6.6}$$

Because both b_0 and $b_{\ell-1}$ are primitive and shift-free with respect to y, the $\sigma_y^{\ell'_i}(p_i)$ are pairwise coprime and also each p_i is primitive with respect to y. Thus there exist unique $f'_i \in C(x)[y]$ with $\deg_y(f'_i) < e_i \deg_y(p_i)$ such that

$$\sigma_x^{\ell}(r_0) = \sigma_x^{\ell}\left(\frac{a_0}{b_0}\right) = \sum_{i=1}^{m} \frac{f'_i}{\sigma_y^{\ell'_i}(p_i^{e_i})}. \tag{6.7}$$

By assumption, r_ℓ is the adjusted σ_y-remainder of $\sigma_x(r_{\ell-1})$ with respect to r_0. Then one sees from (6.4) that b_ℓ is equal to $\prod_{i=1}^m p_i^{e_i}$ up to a nonzero constant in C, yielding $(\deg_x(b_\ell), \deg_y(b_\ell)) = (\deg_x(b_0), \deg_y(b_0))$ by (6.6). Notice that each $f'_i/\sigma_y^{\ell'_i}(p_i^{e_i})$ differs from $\sigma_y^{-\ell'_i}(f'_i)/p_i^{e_i}$ by a σ_y-summable rational function. We know from (6.1) and (6.7) that

$$r_\ell = (\mathrm{S}_y - 1)(h') + \sum_{i=1}^{m} \frac{\sigma_y^{-\ell'_i}(f'_i)}{p_i^{e_i}} \quad \text{for some } h' \in C(x, y).$$

Since $r_\ell = a_\ell/(\delta_\ell b_\ell)$ with $\deg_y(a_\ell) < \deg_y(b_\ell)$ and $\gcd(a_\ell, \delta_\ell b_\ell) = 1$, and $b_\ell / \prod_{i=1}^m p_i^{e_i} \in C$, the above equation implies that h' actually belongs to $C(x)$. Thus $r_\ell = \sum_{i=1}^m \sigma_y^{-\ell'_i}(f'_i)/p_i^{e_i}$. The first assertion now follows straightforwardly by applying Fact 5.1 to (6.7). The second assertion is just one application of Lemma 5.4. □

We note that sharper degree bounds can be obtained in the special case where b_0 (and thus every b_ℓ) in the above lemma is integer-linear. These bounds, however, will not affect the orders of magnitude for these intermediate results. Thus we do not pursue more refined accuracy here.

The reduction-based approach also provides us an order-degree curve of telescopers for bivariate rational functions.

Theorem 6.6. *With the assumptions of Lemma 6.5, further assume that the polynomial b_0 admits the decomposition (3.3) and let $\rho_0 = \sum_{i=1}^m \mu_i d_i \deg_z(p_i)$ with $d_i = \max_{1\le j\le n_i}\{e_{ij}\}$. Then for any nonnegative integer pair (ρ, τ) with $\rho \ge \rho_0$ and*

$$\tau > \frac{\big((\rho+1)\deg_x(b_0)\deg_y(b_0) + \rho\deg_x(\delta_0) + \deg_x(a_0) + \rho_0\big)\rho_0 - \frac{1}{2}\rho_0(\rho_0 - 1) - (\rho+1)}{\rho + 1 - \rho_0}, \tag{6.8}$$

there exists a telescoper for f/g of order at most ρ and degree at most τ. In particular, letting $\deg_x(f), \deg_x(g) \le d_x$ and $\deg_y(f), \deg_y(g) \le d_y$, if $\rho = \rho_0$ then the right hand side of (6.8) is in $\mathrm{O}(\rho_0^2 d_x d_y)$ while if $\rho = 2\rho_0$ then it belongs to $\mathrm{O}(\rho_0 d_x d_y)$.

Proof. Let $\rho, \tau \in \mathbb{N}$ with $\rho \ge \rho_0$ and τ satisfying (6.8). Let $c_0, \ldots, c_\rho$ be indeterminates. By Theorem 6.1 and Remark 6.2, the operator $\sum_{\ell=0}^{\rho} c_\ell \mathrm{S}_x^\ell$ gives a desired telescoper for f/g if and only if the equation $\sum_{\ell=0}^{\rho} c_\ell r_\ell = \sum_{\ell=0}^{\rho} c_\ell a_\ell/(\delta_\ell b_\ell) = 0$ holds for $c_\ell \in C[x]$, not all zero, with $\deg_x(c_\ell) \le \tau$. Thus it amounts to verifying that, for the linear system over C induced by that equation, the number of unknowns, namely $(\tau+1)(\rho+1)$ in this case, is greater than the number of equations over C. According to (Huang, 2016, Theorem 5.5), the least common multiple of the b_ℓ has total degree in x, y at most ρ_0. Hence, based on Lemma 6.5, one calculates that there are at most

$$\Big(\tau + (\rho + 1)\deg_x(b)\deg_y(b) + \rho\deg_x(\delta) + \deg_x(a) + \rho_0\Big)\rho_0 - \frac{1}{2}\rho_0(\rho_0 - 1)$$

equations over C. Since $\rho \geq \rho_0$ and (6.8) holds, a direct comparison between the number of unknowns and the above number completes the proof. □

Compared with the above theorem, the order-degree curve of telescopers for bivariate rational functions depicted by Theorem 5.6 is much sharper..

6.2. Cost analysis of algorithm

Based on Fact 5.2, one easily obtains the following cost for adjusting σ_y-remainders.

Lemma 6.7. *Let $r, r_0 \in C(x, y)$ be valid inputs of the algorithm* **ShiftRemainderAdjustment**. *Write $r = a/(\delta b)$ with $\delta \in C[x]$, $a, b \in C[x, y]$, $\deg_y(a) < \deg_y(b)$, $\gcd(a, \delta b) = 1$ and b being primitive and shift-free with respect to y. Let $b_0 \in C[x, y]$ be the primitive part of the denominator of r_0 with respect to y. Assume that $\deg_x(b), \deg_x(b_0) \leq d_x$ and $\deg_y(b), \deg_y(b_0) \leq d_y$. Then the algorithm finds $\tilde{h} \in C(x, y)$ of the unnormalized form (6.5) and a σ_y-remainder $\tilde{r} \in C(x, y)$ such that (6.2) holds, using $O(\deg_x(a)^2 d_y + d_x^2 d_y^3 + \deg_x(a) d_x d_y^2 + \deg_x(\delta)^2 + \deg_x(\delta) d_x d_y)$ arithmetic operations in C with classical arithmetic and $\tilde{O}(\deg_x(a) d_y + d_x d_y^2 + \deg_x(\delta))$ with fast arithmetic, plus the cost of computing the dispersion set of b with respect to b_0.*

Now we are ready to analyze the cost of the reduction-based creative telescoping algorithm for bivariate rational functions.

Theorem 6.8. *Let $f, g \in C[x, y]$ be two coprime polynomials with $g \neq 0$, $\deg_x(f), \deg_x(g) \leq d_x$ and $\deg_y(f), \deg_y(g) \leq d_y$. Assume that f/g has a telescoper and let ρ be the actual order of its minimal telescopers. Further define ρ_0 as in Theorem 6.6. Then the reduction-based algorithm in (Chen et al., 2015) finds a minimal telescoper for f/g and an unnormalized certificate, using $O(d_x d_y^4 + \rho d_x^2 d_y^3 + \rho_0^3 \rho^3 d_x^2 d_y^2 + \rho_0^4 \rho^2 d_x d_y + \rho_0^5 \rho)$ arithmetic operations in C with classical arithmetic and $\tilde{O}(d_x d_y^3 + \rho d_x d_y^2 + \rho_0^{\omega-1} \rho^3 d_x d_y + \rho_0^{\omega} \rho^2)$ with fast arithmetic, plus the cost of computing (auto-)dispersion sets and finding rational roots.*

Proof. By Theorem 5.9, the GGSZ reduction step takes $O(d_x d_y^4 + d_x^2 d_y^3)$ arithmetic operations with classical arithmetic and $\tilde{O}(d_x d_y^3)$ with fast arithmetic, plus the cost of computing the auto-dispersion set. In addition to the cost of finding rational roots in the integer-linearity detection, the cost of the remaining algorithm is dominated by adjusting σ_y-remainders and solving linear systems in iteration steps. For the ℓth iteration with $0 \leq \ell \leq \rho$, by Lemmas 6.5 and 6.7, adjusting the ℓth σ_y-remainder takes $O(d_x^2 d_y^3)$ with classical arithmetic and $\tilde{O}(d_x d_y^2)$ with fast arithmetic, plus the cost of computing the dispersion sets in the shift-remainder adjustment. After the adjustment, we need to solve a linear system with the coefficient matrix having at most ρ_0 rows and $\ell + 1$ columns and of rank either ℓ or $\ell + 1$. Moreover, the entries of the matrix are of degrees in x in $O(\ell d_x d_y + \rho_0)$. By Fact 5.3, finding a solution requires $O(\rho_0^3 \ell^2 d_x^2 d_y^2 + \rho_0^4 \ell d_x d_y + \rho_0^5)$ with classical arithmetic and $\tilde{O}(\rho_0^{\omega-1} \ell^2 d_x d_y + \rho_0^{\omega} \ell)$ with fast arithmetic. Since there are ρ iterations, this step in total takes $O(\rho d_x^2 d_y^3 + \rho_0^3 \rho^3 d_x^2 d_y^2 + \rho_0^4 \rho^2 d_x d_y + \rho_0^5 \rho)$ with classical arithmetic and $\tilde{O}(\rho d_x d_y^2 + \rho_0^{\omega-1} \rho^3 d_x d_y + \rho_0^{\omega} \rho^2)$ with fast arithmetic, yielding the announced cost. □

The above theorem, together with Theorem 6.6, immediately yields the following.

Corollary 6.9. *With the assumptions of Theorem 6.6, further assume that* $\deg_x(f), \deg_x(g) \le d_x$ *and* $\deg_y(f), \deg_y(g) \le d_y$. *Let* $\mu = \max\{\mu_1, \ldots, \mu_m\}$. *Then, without normalizing the certificate, the reduction-based algorithm takes* $\mathrm{O}(\mu^6 d_x^2 d_y^8)$ *arithmetic operations in* C *with classical arithmetic and* $\tilde{\mathrm{O}}(\mu^{\omega+2} d_x d_y^{\omega+3})$ *with fast arithmetic, plus the cost of computing (auto-)dispersion sets and finding rational roots.*

Due to intermediate expression swell in the unnormalized expression of the certificate part as mentioned in the introduction, we do not expect to gain much for normalizing the certificate in the reduction-based algorithm and thus do not investigate this aspect further.

In addition, for a polynomial $b \in C[x, y]$, computing its auto-dispersion set and computing the dispersion set $\mathrm{DS}_y(\sigma_x(b), b)$ take almost the same cost. Hence the extra cost for the algorithm **RationalOperatorCT** (cf. Theorem 5.11) is no more than that for the reduction-based algorithm.

7. Implementation and timings

We have implemented our algorithms in the computer algebra system Maple 2018. In order to get an idea about the efficiency, we compared their runtime, as well as the memory requirements, to the performance of two known algorithms – the one developed by Le (2003) and the reduction-based one reviewed in Section 6. The implementation for the former uses the built-in Maple procedure **SumTools[Hypergeometric][ZpairDirect]**, while the implementation for the latter was done in accordance with the description of the algorithm **ReductionCT** from (Chen et al., 2015) restricted to the rational case, by embracing the GGSZ reduction and Remark 6.2. All timings are measured in seconds on a Linux computer with 128GB RAM and fifteen 1.2GHz Dual core processors. The computations for the experiments did not use any parallelism.

We take examples of the expanded form of

$$r(x, y) = (\mathrm{S}_y - 1)\left(\frac{f_0(x, y)}{g_0(x, y)}\right) + \frac{f(x, y)}{g_1(x, y) \cdot g_2(x, y)}, \tag{7.1}$$

where

- $f_0, f \in \mathbb{Z}[x, y]$ of total degree $m \ge 0$ and max-norm $\|f_0\|_\infty, \|f\|_\infty \le 20$;
- $g_0 \in \mathbb{Z}[x, y]$ of total degree $n \ge 0$ and max-norm $\|g_0\|_\infty \le 20$;
- $g_i = p_i \cdot \sigma_x(p_i) \cdot \sigma_x^{\mu}(p_i) \cdot \sigma_x^{\mu+1}(p_i)$ with $p_i = P_i((-1)^i \lambda x + \mu y)$ for positive integers λ, μ and integer polynomials $P_i(z) \in \mathbb{Z}[z]$ of total degree $n > 0$ and max-norm $\|P_i\|_\infty \le 20$.

For a selection of random rational functions of this type for different choices of (m, n, λ, μ), Table 1 collects the timings, without expanding the certificate, of the algorithm of Le (DCT), the reduction-based algorithm (RCT) and two variants of our algorithm from Section 4: for the columns OCT_1 and OCT_2, we both compute a minimal telescoper and a sparse certificate, but the difference is that the first one constructs the telescoper by exactly following the steps of the algorithm **RationalOperatorCT** while the second one proceeds in an iterative fashion as described in Remark 4.7. The columns *order* and *upper-order* are respectively used to record the actual order of the output minimal telescoper and the upper bound given in Lemma 4.4 for the order of minimal telescopers for the input.

(m,n,λ,μ)	DCT	RCT	OCT_1	OCT_2	order	upper-order
(1, 1, 1, 1)	0.18	0.17	0.16	0.16	2	4
(1, 1, 4, 1)	0.18	0.20	0.16	0.16	2	4
(1, 1, 16, 1)	0.19	0.21	0.17	0.17	2	4
(5, 1, 4, 1)	0.22	0.23	0.19	0.19	3	4
(10, 1, 4, 1)	0.26	0.27	0.21	0.21	3	4
(15, 1, 4, 1)	0.46	0.40	0.26	0.27	4	4
(15, 1, 4, 5)	10.43	14.63	0.88	0.90	10	10
(15, 1, 4, 7)	46.39	69.64	1.87	1.92	14	14
(15, 1, 4, 9)	181.34	283.65	3.46	3.58	18	18
(15, 1, 4, 11)	456.69	851.72	6.49	7.49	22	22
(15, 1, 4, 13)	892.44	2436.57	9.48	13.59	26	26
(1, 2, 4, 1)	–	15.24	6.81	2.48	7	8
(1, 3, 4, 1)	–	1220.58	1107.94	49.19	11	12
(1, 4, 4, 1)	–	30599.21	76681.82	935.41	15	16
(10, 2, 4, 1)	–	21.00	21.25	3.96	7	8
(20, 2, 4, 1)	–	27.27	66.43	5.92	7	8
(30, 2, 4, 1)	–	51.82	13.83	14.55	8	8
(30, 2, 4, 3)	–	504.78	52.67	51.93	12	12
(30, 2, 4, 5)	–	6437.51	387.80	436.25	20	20
(30, 2, 4, 7)	–	47763.39	1464.22	1283.01	28	28

Table 1: Comparison of four algorithms for a collection of rational functions of the form (7.1).

From the finding we see that both of our creative telescoping algorithms have comparable timings for random problems of small size. In particular none of the four algorithms have significant set up costs. As m increases our algorithms show significant improvement over both the direct and reduction-based methods. In the two cases $(1,4,4,1)$ and $(20,2,4,1)$, our algorithm OCT_1 is dramatically worse than the RCT. This is because the upper bound used in the algorithm is not sufficiently sharp. The dash in the column DCT indicates that the current built-in procedure for DCT in Maple 2018 is not applicable for random inputs with this choice of (m,n,λ,μ). The issue in these cases is that the denominator of the input rational function has irreducible factors of degrees greater than one, and then the algorithm of Le (2003) requires recurrence operators with coefficients being polynomials over algebraic numbers, something not yet included in the current implementation of DCT in Maple.

8. Conclusion and future work

A new algorithm of creative telescoping for bivariate rational functions was developed in this paper. Our algorithm is based on left division with remainder in the ring of recurrence operators and expresses the certificate part by a sparse representation, which can be expanded in time polynomial in the size of the final result if desired. In terms of complexity, our algorithm outperforms the reduction-based approach in the case of bivariate rational functions by at least one order of magnitude ignoring the certificate part. In practice, (the iterative version of) our algorithm is also more efficient according to the experiments.

With the rational case being settled, it is natural to wonder about a possible analogous algorithm for hypergeometric terms. Recall that a bivariate function $f(x,y)$ is called a *hypergeometric* *term* if its two shift-quotients $f(x+1,y)/f(x,y)$ and $f(x,y+1)/f(x,y)$ are both rational functions in x, y. The hypergeometric term is a basic and ubiquitous class of special functions appearing in combinatorics (Petkovšek et al., 1996). It is more interesting and also more challenging than the rational case.

In the hypergeometric case, there exists no direct analog of the partial fraction decomposition of rational functions. Thus the method described in this paper will not work directly for this setting. One possible way to proceed is to first compute a multiplicative decomposition of the given hypergeometic term and then reduce the problem to a rational one (cf. (Abramov and Petkovšek, 2001, 2002; Chen et al., 2015)). This way, however, may introduce left division with remainder on recurrence operators over $C(x,y)$ instead of $C(x)[y]$, and thus makes it more difficult to derive a hypergeometric telescoping criterion, namely an analog of Theorem 4.1. In the future, we hope to explore this topic further and aim at generalizing our results to the class of hypergeometric terms and beyond.

References

Abramov, S. A., 1971. On the summation of rational functions. USSR Comput. Math. Math. Phys. 11 (4), 324–330. URL https://doi.org/10.1016/0041-5553(71)90028-0

Abramov, S. A., 1975. The rational component of the solution of a first-order linear recurrence relation with a rational right side. USSR Comput. Math. Math. Phys. 15 (4), 216–221. URL https://doi.org/10.1016/0041-5553(75)90181-0

Abramov, S. A., 2003. When does Zeilberger's algorithm succeed? Adv. in Appl. Math. 30 (3), 424–441. URL http://dx.doi.org/10.1016/S0196-8858(02)00504-3

Abramov, S. A., Le, H. Q., 2002. A criterion for the applicability of Zeilberger's algorithm to rational functions. Discrete Math. 259 (1-3), 1–17. URL http://dx.doi.org/10.1016/S0012-365X(02)00442-9

Abramov, S. A., Petkovšek, M., 2001. Minimal decomposition of indefinite hypergeometric sums. In: Proceedings of ISSAC'01. ACM, New York, pp. 7–14. URL http://dx.doi.org/10.1145/384101.384103

Abramov, S. A., Petkovšek, M., 2002. Rational normal forms and minimal decompositions of hypergeometric terms. J. Symbolic Comput. 33 (5), 521–543. URL http://dx.doi.org/10.1006/jsco.2002.0522

Apagodu, M., Zeilberger, D., 2006. Multi-variable Zeilberger and Almkvist-Zeilberger algorithms and the sharpening of Wilf-Zeilberger theory. Adv. in Appl. Math. 37 (2), 139–152. URL http://dx.doi.org/10.1016/j.aam.2005.09.003

Bostan, A., Chen, S., Chyzak, F., Li, Z., 2010. Complexity of creative telescoping for bivariate rational functions. In: Proceedings of ISSAC'10. ACM, New York, pp. 203–210. URL http://dx.doi.org/10.1145/1837934.1837975

Bostan, A., Chen, S., Chyzak, F., Li, Z., Xin, G., 2013a. Hermite reduction and creative telescoping for hyperexponential functions. In: Proceedings of ISSAC'13. ACM, New York, pp. 77–84. URL http://dx.doi.org/10.1145/2465506.2465946

Bostan, A., Chyzak, F., Lairez, P., Salvy, B., 2018. Generalized Hermite reduction, creative telescoping and definite integration of d-finite functions. In: Proceedings of ISSAC'18. ACM, New York, pp. 95–102. URL http://doi.acm.org/10.1145/3208976.3208992

Bostan, A., Lairez, P., Salvy, B., 2013b. Creative telescoping for rational functions using the Griffiths-Dwork method. In: Proceedings of ISSAC'13. ACM, New York, pp. 93–100. URL http://dx.doi.org/10.1145/2465506.2465935

Bostan, A., Lairez, P., Salvy, B., 2017. Multiple binomial sums. J. Symbolic Comput. 80 (part 2), 351–386. URL https://doi.org/10.1016/j.jsc.2016.04.002

Chen, S., Hou, Q.-H., Huang, H., Labahn, G., Wang, R.-H., 2019. Constructing minimal telescopers for rational functions in three discrete variables. Preprint: arXiv:1904.11614.

Chen, S., Huang, H., Kauers, M., Li, Z., 2015. A modified Abramov-Petkovšek reduction and creative telescoping for hypergeometric terms. In: Proceedings of ISSAC'15. ACM, New York, pp. 117–124. URL http://doi.acm.org/10.1145/2755996.2756648

Chen, S., Kauers, M., 2012a. Order-degree curves for hypergeometric creative telescoping. In: Proceedings of ISSAC'12. ACM, New York, pp. 122–129.
URL http://dx.doi.org/10.1145/2442829.2442850

Chen, S., Kauers, M., 2012b. Trading order for degree in creative telescoping. J. Symbolic Comput. 47 (8), 968–995.
URL http://dx.doi.org/10.1016/j.jsc.2012.02.002

Chen, S., Kauers, M., Koutschan, C., 2014. A generalized Apagodu-Zeilberger algorithm. In: Proceedings of ISSAC'14. ACM, New York, pp. 107–114.
URL http://dx.doi.org/10.1145/2608628.2608641

Chen, S., Kauers, M., Koutschan, C., 2016. Reduction-based creative telescoping for algebraic functions. In: Proceedings of ISSAC'16. ACM, New York, pp. 175–182.
URL http://doi.acm.org/10.1145/2930889.2930901

Chen, S., van Hoeij, M., Kauers, M., Koutschan, C., 2018. Reduction-based creative telescoping for fuchsian D-finite functions. J. Symbolic Comput. 85, 108–127.
URL https://doi.org/10.1016/j.jsc.2017.07.005

von zur Gathen, J., Gerhard, J., 2013. Modern Computer Algebra, 3rd Edition. Cambridge University Press, Cambridge.
URL http://dx.doi.org/10.1017/CBO9781139856065

Gerhard, J., 2004. Modular Algorithms in Symbolic Summation and Symbolic Integration (Lecture Notes in Computer Science). Springer-Verlag.

Gerhard, J., Giesbrecht, M., Storjohann, A., Zima, E. V., 2003. Shiftless decomposition and polynomial-time rational summation. In: Proceedings of ISSAC'03. ACM, New York, pp. 119–126.
URL http://dx.doi.org/10.1145/860854.860887

Giesbrecht, M., Huang, H., Labahn, G., Zima, E., 2019. Efficient integer-linear decomposition of multivariate polynomials. In: Proceedings of ISSAC'19. ACM, New York, pp. 171–178.
URL http://doi.acm.org/10.1145/3326229.3326261

van der Hoeven, J., 2017. Constructing reductions for creative telescoping. Preprint: hal-01435877.

Huang, H., 2016. New bounds for hypergeometric creative telescoping. In: Proceedings of ISSAC'16. ACM, New York, pp. 279–286.
URL http://doi.acm.org/10.1145/2930889.2930893

Le, H. Q., 2003. A direct algorithm to construct the minimal Z-pairs for rational functions. Adv. in Appl. Math. 30 (1-2), 137–159, Formal power series and algebraic combinatorics (Scottsdale, AZ, 2001).
URL http://dx.doi.org/10.1016/S0196-8858(02)00529-8

Li, Z., Zhang, Y., 2013. An algorithm for decomposing multivariate hypergeometric terms. A contributed talk in CM'13.

Man, Y.-K., Wright, F. J., 1994. Fast polynomial dispersion computation and its application to indefinite summation. In: Proceedings of ISSAC'94. ACM, New York, pp. 175–180.
URL http://doi.acm.org/10.1145/190347.190413

Mohammed, M., Zeilberger, D., 2005. Sharp upper bounds for the orders of the recurrences output by the Zeilberger and q-Zeilberger algorithms. J. Symbolic Comput. 39 (2), 201–207.
URL http://dx.doi.org/10.1016/j.jsc.2004.10.002

Paule, P., 1995. Greatest factorial factorization and symbolic summation. J. Symbolic Comput. 20 (3), 235–268.
URL http://dx.doi.org/10.1006/jsco.1995.1049

Petkovšek, M., Wilf, H. S., Zeilberger, D., 1996. $A = B$. A K Peters, Ltd., Wellesley, MA.

Polyakov, S. P., 2011. Indefinite summation of rational functions with factorization of denominators. Program. Comput. Softw. 37 (6), 322–325, translated from Programmirovanie **37** (2011), no. 4.
URL https://doi.org/10.1134/S0361768811060077

Zeilberger, D., 1990a. A fast algorithm for proving terminating hypergeometric identities. Discrete Math. 80 (2), 207–211.
URL http://dx.doi.org/10.1016/0012-365X(90)90120-7

Zeilberger, D., 1990b. A holonomic systems approach to special functions identities. J. Comput. Appl. Math. 32 (3), 321–368.
URL http://dx.doi.org/10.1016/0377-0427(90)90042-X

Zeilberger, D., 1991. The method of creative telescoping. J. Symbolic Comput. 11 (3), 195–204.
URL http://dx.doi.org/10.1016/S0747-7171(08)80044-2

Zhou, W., Labahn, G., Storjohann, A., 2012. Computing minimal nullspace bases. In: Proceedings of ISSAC'12. ACM, New York, pp. 366–373.
URL https://doi.org/10.1145/2442829.2442881

Zima, E., 2011. Synthetic division in the context of indefinite summation. In: Proceedings of SNC'11. ACM, New York, pp. 151–152.
URL http://doi.acm.org/10.1145/2331684.2331708