AN EXTENSION OF ZEILBERGER'S FAST ALGORITHM TO GENERAL HOLONOMIC FUNCTIONS

FRÉDÉRIC CHYZAK

Abstract. We extend Zeilberger's fast algorithm for definite hypergeometric summation to non-hypergeometric holonomic sequences. The algorithm generalizes to differential and q-cases as well. Its theoretical justification is based on a description by linear operators and on the theory of holonomy.

Résumé. Nous étendons l'algorithme rapide de Zeilberger pour la sommation hypergéométrique définie aux suites holonomes non-hypergéométriques. L'algorithme se généralise aussi au cas différentiel et du q-calcul. Sa justification théorique se fonde sur une description par opérateurs linéaires et sur la théorie de l'holonomie.

Introduction

In [28], D. Zeilberger initiated an algorithmic treatment of special functions that led to efficient algorithms for summation and integration [21]. In this approach, he considered a large class of functions and sequences that enjoys numerous closure properties, the class of *holonomic functions*. Simple definitions of holonomy in the continuous and discrete cases are as follows: a function $f(x_1, \dots, x_n)$ is called holonomic when its derivatives span a finite-dimensional vector space over the field of rational functions in the x_i's; a sequence is then defined to be holonomic when its multivariate generating function is holonomic. We use *holonomic function* to refer to either case.

Algorithms for sums of holonomic sequences rely on the method of *creative telescoping* [29]. Given a bivariate sequence $(u_{n,k})$, this method computes a linear recurrence satisfied by the definite sum $U_n = \sum_{k\in\mathbb{Z}} u_{n,k}$. The calculation is as follows: assume that another sequence $(v_{n,k})$ and rational functions η_i *in n only* satisfy the identity

$$\sum_{i=0}^{L} \eta_i(n) u_{n+i,k} = v_{n,k+1} - v_{n,k}; \tag{1}$$

summing over k and considering technical assumptions on v then yields a linear recurrence satisfied by U_n. The method extends to differential and q-cases [5, 19, 20, 22].

A univariate sequence (u_n) such that u_{n+1}/u_n is a rational function in n is called *hypergeometric*. Similarly in the multivariate case, a hypergeometric sequence is a sequence $(u_{n_1,\dots,n_r})$ such that each $u_{n_1,\dots,n_i+1,\dots,n_r}/u_{n_1,\dots,n_r}$ is a rational functions in the n_i's. Equivalently, hypergeometric sequences are defined by linear first order equations. Hypergeometry does not imply holonomy, as exemplified by the sequence u given by $u_{n,k} = 1/(n^2 + k^2)$ (see [25]).

To solve the *elimination problem* of determining an equation like (1), Zeilberger first gave a general but theoretical algorithm based on a skew Euclidean algorithm [28]. He

This work was supported in part by the Long Term Research Project Alcom-IT (#20244) of the European Union.

himself called this algorithm the *slow algorithm*, and proposed his *fast algorithm* [27] for a restricted class of sequences: this algorithm is guaranteed to terminate on sequences which are simultaneously hypergeometric and holonomic. Such sequences are called *holonomic hypergeometric*. Zeilberger's theory extends to multiple summations of holonomic hypergeometric sequences, with counterparts for (possibly multiple) integrals and their q-analogues [25, 26]. As an example of application, Zeilberger's algorithm computes the following sum in closed form

$$\sum_{k=0}^{2n}(-1)^k\binom{2n}{k}\binom{2k}{k}\binom{4n-2k}{2n-k}=\binom{2n}{n}^2.$$

In [12], we described unified but rather slow algorithms based on skew Gröbner basis calculations to perform creative telescoping in general classes of functions and sequences, including the class of holonomic functions. This can be viewed as a generalization of Zeilberger's *slow* algorithm. Our main contribution in the present article is to extend Zeilberger's *fast* algorithm to a class of *∂-finite functions*, i.e., functions defined by linear equations of any order, in the unified setting of Ore operators.

For instance, our algorithm rediscovers identities like

$$\sum_{n=0}^{\infty}P_n(x)y^n=\frac{1}{1-2xy+y^2},\qquad \sum_{n=0}^{\infty}J_{2n+1/2}(x)=\int_0^x\frac{\cos t}{\sqrt{2\pi t}}dt,$$

where the $P_n(x)$ are the Legendre polynomials and the $J_\nu(x)$ are the Bessel functions of the first kind. In each case, we start from a description of the summand in the left-hand side in terms of linear operators at which it vanishes and we compute the right-hand side by summation. Note that in each case, the summand is not a hypergeometric term.

Zeilberger's fast algorithm for *definite* hypergeometric summation is based on an algorithm for *indefinite* hypergeometric summation due to R. W. Gosper [15, 16]. For sequences (u_k) and (U_k) such that $U_{k+1}-U_k=u_k$, U is called an *indefinite sum* of u. Gosper's algorithm recognizes whether there exists a hypergeometric indefinite sum U of a hypergeometric sequence u, and if so computes such a U. When a solution is found, the sum $\sum_{j=0}^{k-1}u_j$ is U_k-U_0. The sequences u and U are related by an equation of the form $U_k=\theta(k)u_k$ with θ a rational function, so that the summation problem reduces to computing θ. It turns out that θ satisfies a linear recurrence with polynomial coefficients, which can be solved for rational solutions θ by S. A. Abramov's algorithm [1]. Alternatively, Gosper's clever remark is that it suffices to solve a derived equation for *polynomial* solutions, which is done by a method of undetermined coefficients. (See [3] for a refinement.) As an example of application, Gosper's algorithm proves

$$\sum_{j=0}^{k}\frac{4^j}{\binom{2j}{j}}=\frac{2(k+1)4^k}{\binom{2k}{k}}+\frac{1}{3}.$$

If a positive integer L and rational functions η_i were known to be such that the left-hand side of Eq. (1) admits a hypergeometric indefinite sum, Gosper's algorithm would apply to solve (1) for it. Based on this observation, Zeilberger's fast algorithm introduces undetermined coefficients for the η_i's and uses an extension of Gosper's algorithm to solve for a hypergeometric indefinite sum (v_k) together with rational η_i's. This process in run with increased values of L until the indefinite summation problem becomes solvable. When u is a holonomic hypergeometric sequence, the termination of the algorithm is guaranteed by holonomy. The algorithm then yields Eq. (1) from which creative telescoping computes a linear recurrence satisfied by the definite sum U_n.

In this article, we generalize Zeilberger's algorithm to the case when the linear equations satisfied by $(u_{n,k})$ have orders larger than 1, and are not necessarily recurrences. The definition of ∂-finite functions [12] is recalled in the next section. Next, we extend Abramov's alternative approach to Gosper's algorithm, then Zeilberger's algorithm to ∂-finite functions. We then detail how the normal forms for ∂-finite functions used in those algorithms are obtained by methods of *Gröbner bases*. We finally define certificates and companion identities in the context of ∂-finite identities.

1. Algebras of operators and ∂-finite functions

A differential counterpart to Zeilberger's slow algorithm for sequences is available in the case of functions and both versions extend to q-analogues [25]. All these algorithms are very similar in their structures and behaviours, so that a unified description is in terms of linear operators. To this end, we introduced [12] a large class of operator algebras which are well suited to accommodate linear differential and difference operators, their q-analogues and numerous other generalized differential operators. Let $\mathbb{A}$ be a ring endowed with a ring endomorphism σ. Following [13], a *σ-derivation* δ on $\mathbb{A}$ is an additive endomorphism such that $(ab)^\delta = a^\sigma b^\delta + a^\delta b$ for all $a, b \in \mathbb{A}$. (We denote the application of σ's and δ's by powers, referring to the prime notation for derivatives.) Since the corresponding generalized differential operators are those of interest to our study, we often call a σ-derivation a derivation.

Definition. Let $\mathbb{K}$ be a (possibly skew) field and $\boldsymbol{\partial} = (\partial_1, \dots, \partial_r)$ be a tuple of indeterminates that commute pairwise. We assume that the field $\mathbb{K}$ is endowed with injective field endomorphisms σ_i's and additive endomorphisms δ_i's, one pair for each $i = 1, \dots, r$, such that each δ_i is a σ_i-derivation. We assume further that σ_i and δ_j, σ_i and σ_j, δ_i and δ_j commute for $i \neq j$. The *Ore algebra* $\mathbb{K}[\partial_1; \sigma_1, \delta_1] \dots [\partial_r; \sigma_r, \delta_r]$, which we also denote $\mathbb{K}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$, is the ring of polynomials in $\boldsymbol{\partial}$ with coefficients in $\mathbb{K}$, with usual addition and a product defined by associativity from the commutation rules

$$\partial_i a = a^{\sigma_i} \partial_i + a^{\delta_i}$$

between the ∂_i's and elements $a \in \mathbb{K}$.

An Ore algebra $\mathbb{O}$ is clearly a $\mathbb{K}$-algebra. In order to view it as an algebra of *linear operators*, we assume that we are given a commutative $\mathbb{K}$-algebra $\mathcal{F}$ whose elements we call *functions*, and we require $\mathcal{F}$ to be a *left $\mathbb{O}$-module* containing $\mathbb{K}$. For instance, in the case of the Ore algebra $\mathbb{O} = \mathbb{K}(z)[\partial; 1, d/dz]$ of linear differential operators, the algebra of Laurent formal power series $\mathbb{K}((z))$ is a left $\mathbb{O}$-module for the action $(\partial \cdot f)(z) = f'(z)$ and $(z \cdot f)(z) = zf(z)$; in the case of the Ore algebra $\mathbb{O} = \mathbb{K}(n)[\partial; S_n, 0]$ of linear recurrence operators, the algebra $\mathbb{K}^{\mathbb{N}}$ of sequences for term-wise addition and product is a left $\mathbb{O}$-module for the action $(\partial \cdot u)(n) = u_{n+1}$ and $(n \cdot u)(n) = nu_n$.

When viewed as operators, elements of Ore algebras are called *Ore operators*. By a *derivative* of a function $f \in \mathcal{F}$, we mean the result of the action of ∂_i on f, which we denote $\partial_i \cdot f$. More generally, any $\boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f$ is also called a derivative. For a function $f \in \mathcal{F}$, the left ideal $\operatorname{Ann} f = \{P \in \mathbb{O} \mid P \cdot f = 0\}$ describes much of the structure of the derivatives of f. It is called the *annihilating ideal* of f and satisfies $\mathbb{O}/\operatorname{Ann} f \simeq \mathbb{O} \cdot f$.

INPUT: a basis B for the annihilating ideal of a ∂-finite function f.
OUTPUT: a basis for all operators Q such that $Q \cdot f = \partial^{-1} \cdot f$, or $\perp$.

1. from B, compute a Gröbner basis G and get the finite basis $(\boldsymbol{\partial}^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in I}$ of $\mathbb{O}/\operatorname{Ann} f$ canonically associated to G;
2. introduce undetermined coefficients $\phi_{\boldsymbol{\alpha}}$ and rewrite $\partial \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} - 1$ in this basis by reduction by G;
3. solve the corresponding system of first order linear equations for solutions $\phi_{\boldsymbol{\alpha}} \in \mathbb{K}$;
4. if solvable, return $Q = \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}$; otherwise return $\perp$.

ALGORITHM 1. indefinite ∂-finite summation

Of particular interest are *∂-finite functions*, which correspond in applications to functions and sequences defined by a finite number of equations and initial conditions.

Definition. Let $\mathbb{O} = \mathbb{K}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ be an Ore algebra. A function f in a left $\mathbb{O}$-module is called *∂-finite* when its derivatives span a finite-dimensional vector space $\mathbb{O} \cdot f$ over $\mathbb{K}$. In this case, the left ideal $\operatorname{Ann} f = \{P \in \mathbb{O} \mid P \cdot f = 0\}$ is also called a *∂-finite ideal.*

In the case of the Ore algebra $\mathbb{O} = \mathbb{C}(x_1, \ldots, x_n)[\partial_1; 1, d/dx_1] \ldots [\partial_n; 1, d/dx_n]$ built on differential operators ∂_i's, we recover the definition of holonomy [28], so that ∂-finiteness extends holonomy of (continuous) functions.

2. INDEFINITE ∂-FINITE ∂^{-1}

For an Ore algebra $\mathbb{O} = \mathbb{K}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$, let ∂ be any of the ∂_i's and $\mathcal{F}$ be a left $\mathbb{O}$-module of functions. We call a function $F \in \mathcal{F}$ an *anti-derivative* of $f \in \mathcal{F}$ when $\partial \cdot F = f$. Alternatively, we write $\partial^{-1} \cdot f$ to denote *any* of those anti-derivatives. We develop an algorithm to compute the anti-derivatives $F = \partial^{-1} \cdot f$ of a ∂-finite function f, when there exists such an F in $\mathbb{O} \cdot f$. Moreover, the algorithm always terminates, detecting when no such $\partial^{-1} \cdot f$ exists in $\mathbb{O} \cdot f$ and returning the special symbol $\perp$ in this case. In the case of hypergeometric sequences (and Ore algebras built on shift or difference operators), we recover the variant of Gosper's algorithm that solves the linear recurrence for rational solutions by Abramov's algorithm.

2.1. **Algorithm.** We proceed to establish the following theorem.

Theorem. *Assume that $\mathbb{K}$ admits a decision algorithm to solve linear equations $L \cdot f = 0$ where $L \in \mathbb{K}[\partial; \sigma, \delta]$ for solutions in $\mathbb{K}$. Then Algorithm 1 is a decision algorithm to compute a basis of all the anti-derivatives of a ∂-finite function f in $\mathbb{O} \cdot f$.*

Note that the requirement that the input be the whole annihilating ideal of a ∂-finite function can be weakened: the algorithm terminates on any ∂-finite subideal of the annihilating ideal of a ∂-finite function; however, it may fail to find anti-derivatives with such an incomplete input. This change of ideal corresponds to a change of ∂-finite function f by introducing parasitic solutions.

The algorithm reduces the problem to that of solving a system of linear Ore operators for *rational* function solutions. Those rational functions are then viewed as the coefficients of the operator Q such that $\partial^{-1} \cdot f = Q \cdot f$.

The key point is to make the action of the differentiation operator ∂ on the finite-dimensional vector space $\mathbb{O} \cdot f$ explicit. Let F be any function in $\mathbb{O} \cdot f$. We fix a $\mathbb{K}$-basis of $\mathbb{O} \cdot f$ of the form $(\boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f)_{\boldsymbol{\alpha} \in I}$ for a finite set I of indices. Then $F = Q \cdot f$ where $Q \in \mathbb{O}/\operatorname{Ann} f$ can be written $Q = \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}}(\mathbf{x}) \boldsymbol{\partial}^{\boldsymbol{\alpha}}$. With the assumption $F = \partial^{-1} \cdot f$, i.e., $\partial \cdot F = f$, we have $\partial Q = 1 \mod \operatorname{Ann} f$. In other words:

$$\partial Q = \sum_{\boldsymbol{\alpha} \in I} \phi^{\sigma}_{\boldsymbol{\alpha}}(\mathbf{x}) \boldsymbol{\partial}^{\boldsymbol{\alpha}} \partial + \sum_{\boldsymbol{\alpha} \in I} \phi^{\delta}_{\boldsymbol{\alpha}}(\mathbf{x}) \boldsymbol{\partial}^{\boldsymbol{\alpha}} = 1. \tag{2}$$

Now, 1 and each $\boldsymbol{\partial}^{\boldsymbol{\alpha}} \partial$ in this equation can be rewritten in the basis $(\boldsymbol{\partial}^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in I}$. From the computational point of view, this rewriting is performed by methods of Gröbner basis and for a particular choice of basis of $\mathbb{O} \cdot f$. For the sake of clarity, we postpone the description of these two ingredients to Section 4.

Next, for each $\boldsymbol{\alpha} \in I$, extracting the coefficients in $\boldsymbol{\partial}^{\boldsymbol{\alpha}}$ yields an equation

$$\sum_{\boldsymbol{\beta} \in I} \lambda_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{x}) \phi^{\sigma}_{\boldsymbol{\beta}}(\mathbf{x}) + \phi^{\delta}_{\boldsymbol{\alpha}}(\mathbf{x}) = \mu_{\boldsymbol{\alpha}}(\mathbf{x}), \tag{3}$$

where the $\lambda_{\boldsymbol{\alpha},\boldsymbol{\beta}}$ and $\mu_{\boldsymbol{\alpha}}$ are rational functions in $\mathbf{x}$. Denoting vectors and matrices by capital letters, we get the following linear differential system

$$\Lambda(\mathbf{x}) \Phi^{\sigma}(\mathbf{x}) + \Phi^{\delta}(\mathbf{x}) = M(\mathbf{x}). \tag{4}$$

We next solve this system in a way which depends on the algebra of operators under consideration. Either the system is solvable, and each Q yields an anti-derivative $Q \cdot f$ in $\mathbb{O} \cdot f$; or it is not solvable, and no anti-derivative can be found in $\mathbb{O} \cdot f$.

Let us detail how to solve Eq. (4). Each equation of the system may involve several unknown functions. We do not know of algorithms to solve this kind of linear system directly; the first step is therefore to uncouple the system so as to get equations in a single unknown function. This can be achieved for any Ore operator ∂ by appealing to Abramov's and Zima's algorithm [4]. Indeed, introduce the new Ore algebra $\mathbb{K}[\partial'; \sigma', \delta']$ where $\sigma' = \sigma^{-1}$ and $\partial' = \delta' = -\sigma^{-1}\delta$ on $\mathbb{K}$. Applying σ^{-1} to Eq. (4) yields the system

$$\Lambda^{\sigma^{-1}}(\mathbf{x}) \Phi(\mathbf{x}) - \partial' \cdot \Phi(\mathbf{x}) = M^{\sigma^{-1}}(\mathbf{x}),$$

where $\Lambda^{\sigma^{-1}}(\mathbf{x})$ and $M^{\sigma^{-1}}(\mathbf{x})$ are known and $\Phi(\mathbf{x})$ is the unknown. This is exactly the input form of the algorithm in [4]. Once the system has been uncoupled, we have to solve several linear equations in a single unknown function for rational solutions $\phi_{\boldsymbol{\alpha}}$. This resolution in turn depends on the operator ∂'.

The case of (ordinary or q-) recurrences. This is an instance of the more general case when $\partial = \delta = \sigma - 1$ (where 1 is the identity). We then usually work with the σ operator of (ordinary or q-) shift instead of the δ operator of (ordinary or q-) difference, because both operator algebras $\mathbb{K}[\delta; \sigma, \delta]$ and $\mathbb{K}[\sigma; \sigma, 0]$ are equal when $\delta = \sigma - 1$. After the uncoupling step described above, we are led to linear equations in the shift or q-shift operator. In each case, an algorithm of Abramov's applies [2, 1].

The case of (ordinary) differential equations. In this case, σ is the identity, so that the change of Ore operators in the uncoupling step above is trivial ($\partial' = \partial$). We next solve each uncoupled differential equation by Abramov's algorithm [1].

Finally, note that the value 1 in the right-hand side of Eq. (2) was inessential. Changing (2) into the more general equation

$$\partial Q = \sum_{\boldsymbol{\alpha}\in I} \phi^{\sigma}_{\boldsymbol{\alpha}}(\mathbf{x})\boldsymbol{\partial}^{\boldsymbol{\alpha}}\partial + \sum_{\boldsymbol{\alpha}\in I} \phi^{\delta}_{\boldsymbol{\alpha}}(\mathbf{x})\boldsymbol{\partial}^{\boldsymbol{\alpha}} = H, \tag{5}$$

where H is any element of $\mathbb{O}/\operatorname{Ann} f$ makes it possible to detect if $H \cdot f$ has an antiderivative in $\mathbb{O} \cdot f$. This only affects the vector M in Eq. (4) in a linear way. This fact will be used in our fast algorithm for creative telescoping in the next section.

2.2. **Example: Harmonic summation.** Harmonic summation identities like

$$\sum_{k=1}^{n} \binom{k}{m} H_k = \binom{n+1}{m+1}\left(H_{n+1} - \frac{1}{m+1}\right),$$

where H_n denotes the harmonic number $\sum_{k=1}^{n} k^{-1}$, can be proved using our algorithm. Identities of this kind are classically proved by summation by parts or by techniques of generating functions. (See also M. Karr's general algorithm [17, 18].) Introducing $f_n = \binom{n}{m} H_n$, we show the equivalent form

$$\sum_{k=1}^{n} f_k = \frac{(n+1)^2}{(m+1)^2} f_n - \frac{(n-m)(n-m+1)}{(m+1)^2} f_{n+1}. \tag{6}$$

First, f satisfies the following linear recurrence:

$$(n-m+1)(n-m+2)f_{n+2} - (2n+3)(n-m+1)f_{n+1} + (n+1)^2 f_n = 0.$$

Such an equation is obtained by simplifying f_{n+1} and f_{n+2} by the relation

$$(n+1-k)f_{n+1} = (n+1)f_n + 1$$

and searching for a linear dependency. Thus, the sequence f is a ∂-finite function with respect to the Ore algebra $\mathbb{O} = \mathbb{Q}(n,m)[S_n; S_n, 0]$, where S_n is the shift operator with respect to n. Since $\mathbb{O} \cdot f$ is a two-dimensional vector space with basis $(f, S_n \cdot f)$, we introduce a generic operator $Q = \alpha_n + \beta_n S_n$ and compute $\partial Q - 1$. Then, Eq. (4) takes the form

$$\begin{cases} -(n-m+1)(n-m+2)\alpha_n + (n+1)(n-m+2)\alpha_{n+1} \\ \qquad\qquad - (n+1)(n-m+2)\beta_n + (n+1)(n+2)\beta_{n+1} \\ \qquad\qquad = (n+1)(n-m+2), \\ -(n-m+1)(n-m+2)\alpha_n + (2n+1)(n+2-m)\alpha_{n+1} \\ \qquad\qquad - (2n+1)(n-m+2)\beta_n + (3n^2+6n+2)\beta_{n+1} \\ \qquad\qquad = (2n+1)(n-m+2). \end{cases}$$

Uncoupling this system so as to get rid of α yields the recurrence

$$\begin{aligned} -(n+2)^2\beta_{n+2} &+ (2n+3)(n-m+3)\beta_{n+1} \\ &- (n-m+2)(n-m+3)\beta_n = (n-m+3)(n-m+2), \end{aligned}$$

which is solved for rational solutions by Abramov's algorithm. Replacing in the system and eliminating α_{n+1} between both equations, we find

$$\alpha_n = \frac{(n+1)^2}{(m+1)^2} \qquad \text{and} \qquad \beta_n = -\frac{(n-m)(n-m+1)}{(m+1)^2},$$

which yields the right-hand side of Eq. (6).

INPUT: a basis B for the annihilating ideal of a ∂-finite function f.
OUTPUT: a pair of operators (P, Q) satisfying (7).

1. from B, compute a Gröbner basis G and get the finite basis $(\boldsymbol{\partial}^{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in I}$ of $\mathbb{O}/\operatorname{Ann} f$ canonically associated to G;
2. for $L = 0, 1, \ldots$:
   (a) introduce undetermined coefficients η_i and $\phi_{\boldsymbol{\alpha}}$ and rewrite $\partial' \sum_{\boldsymbol{\alpha}\in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} - \sum_{i=0}^{L} \eta_i \partial^i$ in this basis by reduction by G;
   (b) solve the corresponding system of first order linear equations for solutions $\eta_i \in \mathbb{K}$ and $\phi_{\boldsymbol{\alpha}} \in \mathbb{K}(\mathbf{u})$;
   (c) if solvable, return the solution $\left(\sum_{i=0}^{L} \eta_i \partial^i, \sum_{\boldsymbol{\alpha}\in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}\right)$; otherwise loop.

ALGORITHM 2. definite ∂-finite summation

3. FAST DEFINITE ∂-FINITE $\partial^{-1}|_\Omega$

For an Ore algebra $\mathbb{O} = \mathbb{K}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$, let ∂ be any of the ∂_i's and $\mathcal{F}$ be a left $\mathbb{O}$-module of functions. To extend the case of definite summation and integration operators like $\sum_{k=a}^{b}$ and $\int_a^b dx$, we assume there is an operator $\partial^{-1}|_\Omega$ defined on $\mathcal{F}$ such that $\partial\partial^{-1}|_\Omega = 0$. (In [12], we used a less general definition for $\partial^{-1}|_\Omega$, requiring that $\partial^{-1}|_\Omega\partial$ also be 0. This corresponds to analytical assumptions on $\mathcal{F}$ which are irrelevant here.) In this section, we build on Algorithm 1 to perform the elimination step of creative telescoping on ∂-finite functions. In other terms, we solve Eq. (1). This in turn allows us to perform definite (q-)summation or (q-)integration of a (q-)holonomic function, or more generally the problem of computing a definite anti-derivative of a ∂-finite function, as described in [12].

Zeilberger's fast algorithm is guaranteed to terminate on holonomic hypergeometric sequences only. In the case of differential and difference operators, we similarly call a simultaneously ∂-finite and holonomic function *holonomic ∂-finite*. Our algorithm inputs a description of the annihilating ideal of a ∂-finite function and we prove its termination for holonomic ∂-finite functions.

3.1. Algorithm.

A (continuous) holonomic function $f(x, y)$ is a ∂-finite function with respect to an Ore algebra $\mathbb{O} = \mathbb{K}(x, y)[\partial_x; 1, d/dx][\partial_y; 1, d/dy]$ built on (ordinary) differential operators. (Here, $\partial_x = \delta_x = d/dx$ and $\partial_y = \delta_y = d/dy$.) The original definition of holonomy in the framework of $\mathcal{D}$-modules [7, 8] implies that there exists a non-zero operator in $\operatorname{Ann} f \cap \mathbb{K}(x)[\boldsymbol{\partial}; \mathbf{1}, \boldsymbol{\delta}]$ [28, Lemma 4.1]. We refer the reader to [9, 14] for textbooks on holonomy. As a result, there is a non-trivial identity of the form

$$\sum_{i=0}^{L} \eta_i(x)\partial_x^i \cdot f = \partial_y \cdot (Q(x, y, \partial_x, \partial_y) \cdot f)$$

mimicking (1) for $Q \in \mathbb{O}$. This existence property transfers to the discrete case by generating functions and similar results hold for q-analogues [23].

More generally, in the case of a ∂-finite function f with respect to an Ore algebra $\mathbb{O} = \mathbb{K}(u_1, \ldots, u_s)[\partial; \sigma, \delta][\partial'; \sigma', \delta']$ such that ∂' commutes with elements of $\mathbb{K}$ but not with the u_i's, we look for solutions of

$$P(\partial) \cdot f = \sum_{i=0}^{L} \eta_i \partial^i \cdot f = \partial' \cdot (Q(\mathbf{u}, \partial, \partial') \cdot f)\,, \tag{7}$$

where $P \neq 0$ and the η_i's *do not depend on* $\mathbf{u}$.

We now summarize the result of this section in the following theorem.

Theorem. *Assume that $\mathbb{K}(\mathbf{u})$ admits a decision algorithm to solve linear equations $L \cdot f = 0$ where $L \in \mathbb{K}(\mathbf{u})[\partial'; \sigma', \delta']$ for solutions in $\mathbb{K}(\mathbf{u})$. When there exists a pair (P, Q) that satisfies (7), Algorithm 2 terminates and returns such a pair. This happens in particular as soon as f is a holonomic ∂-finite function.*

As soon as we know an operator P that makes Eq. (7) solvable in Q, we can use our indefinite summation algorithm to get Q. Indeed, it was noted that the value of H in Eq. (5) is inessential; letting $H = P$ makes it possible (after reduction modulo $\operatorname{Ann} f$) to apply our indefinite summation algorithm, the vector M in Eq. (4) depending linearly on the η_i's. However, we do not want to solve for Q uniformly in the parameters η_i's; we need to find for which values of the η_i's the equation is solvable in Q. Therefore, we use a variant of our indefinite summation algorithm so that it solves Eq. (4) in Φ and M simultaneously. This corresponds to classical refinements of Abramov's algorithms described in [29].

Thus, our algorithm proceeds like Zeilberger's fast algorithm: we make a choice for L, introduce undetermined coefficients η_i's and apply our indefinite summation algorithm; if Eq. (4) is solvable, we have finished, otherwise we increase L.

3.2. **Example: Neumann's addition theorem.** We illustrate the previous algorithm with Neumann's addition theorem

$$1 = J_0(z)^2 + 2 \sum_{k=1}^{\infty} J_k(z)^2$$

for the Bessel functions of the first kind $J_k(z)$. The latter are defined as ∂-finite functions by the following operators

$$z^2 D_z^2 + z D_z + z^2 - k^2, \qquad z D_z S_k + (k+1) S_k - z, \qquad z D_z + z S_k - k,$$

in the Ore algebra $\mathbb{O} = \mathbb{K}(k, z)[S_k; S_k, 0][D_z; 1, D_z]$. It follows from an algorithm described in [12] that the squares $J_k(z)^2$ are also ∂-finite and defined by the system

$$\left\{ \begin{array}{r} z D_z^2 + (-2k+1) D_z - 2 S_k z + 2z, \\ z D_z S_k + z D_z + (2k+2) S_k - 2k, \\ z^2 S_k^2 - 4(k+1)^2 S_k - 2z(k+1) D_z + 4k(k+1) - z^2. \end{array} \right.$$

This system generates the ideal $\operatorname{Ann} J_k(z)^2$ in $\mathbb{O}$. Thus, $\mathbb{O}/\operatorname{Ann} J_k(z)^2$ is a three-dimensional vector space, with basis $(1, D_z, S_k)$, and we introduce a generic $Q = u_k + v_k S_k + w_k D_z$. We let $L = 1$ and introduce two parameters $\eta_0(z)$ and $\eta_1(z)$ in Eq. (7) to get a solution. Then, we get the following equations for the system (3)

$$u_k = \frac{k}{z} \eta_1(z), \qquad v_k = 0, \qquad w_k = \frac{1}{2} \eta_1(z),$$

together with the constraint that $\eta_0 = 0$ ($\eta_1(z)$ is any rational function in z). We set $\eta_1(z)$ to 1, so that $P = D_z$ and $Q = k/z + D_z/2$. With these values for P and Q, we have after creative telescoping:

$$P \cdot \left(\sum_{k=0}^{\infty} J_k(z)^2 \right) + \left[Q \cdot J_k(z)^2 \right]_{k=0}^{k=\infty} = 0,$$

from which follows by linearity that

$$D_z \cdot \left(2 \sum_{k=0}^{\infty} J_k(z)^2 - J_0(z)^2 - 1 \right) = -D_z \cdot (J_0(z)^2 + 1) - 2 \left[Q \cdot J_k(z)^2 \right]_{k=0}^{k=\infty} = 0,$$

since $\lim_{k\to+\infty} J_k(z) = \lim_{k\to+\infty} J_k'(z) = 0$. Thus $2\sum_{k=0}^{\infty} J_k(z)^2 - J_0(z)^2 - 1$ is a constant, checked to be 0 when $z = 0$. This proves Neumann's theorem.

4. Effective calculations with ∂-finite ideals

In the algorithms for hypergeometric summation, an important role is played by the relation of *similarity*. Two hypergeometric terms t_n and t_n' are called *similar* when t_n/t_n' is a non-zero rational function in n. When summing a hypergeometric term t_n, Gosper's algorithm therefore searches for an indefinite sum similar to the summand; the algorithm works in the *one-dimensional* vector space $\mathbb{K}(n) \cdot t_n$, so that each sequence under consideration can be represented by a single rational function.

In our extension to the case of ∂-finite functions with respect to an Ore algebra $\mathbb{O} = \mathbb{K}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$, the role of $\mathbb{K}(n) \cdot t_n$ is undertaken by the *finite-dimensional* vector space $\mathbb{O} \cdot f = \bigoplus_{\boldsymbol{\alpha}\in I} \mathbb{K}\, \boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f$ for a finite set I. Each function under consideration in the algorithm can be represented by its rational coordinates $\phi_{\boldsymbol{\alpha}} \in \mathbb{K}$ on the basis of the $\boldsymbol{\partial}^{\boldsymbol{\alpha}}$'s. Two problems arise naturally: one is to compute a set I which determines a basis; another is to compute normal forms in $\mathbb{O} \cdot f$. In particular, when an operator $P \in \mathbb{O}$ is applied on a function $\sum_{\boldsymbol{\alpha}\in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f \in \mathbb{O} \cdot f$, we need to normalize the result $\left(P \sum_{\boldsymbol{\alpha}\in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}\right) \cdot f$ in a form $\sum_{\boldsymbol{\alpha}\in I} \psi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f$.

Both problems are solved using methods of Gröbner bases that are described in [12]. Any Gröbner basis $\{G_1, \dots, G_\ell\}$ of the left ideal $\operatorname{Ann} f \subset \mathbb{O}$ with respect to a term order $\preceq$ (see definitions in [12]) determines a suitable set I in the following way. Call $h_i = \boldsymbol{\partial}^{\boldsymbol{\alpha}_i}$ the head term of G_i with respect to $\preceq$. Then, consider the set of those terms $\boldsymbol{\partial}^{\boldsymbol{\alpha}}$ less than all the h_i's and let $I = \{\boldsymbol{\alpha} \mid \forall i \quad \boldsymbol{\partial}^{\boldsymbol{\alpha}} \prec h_i\}$. This set defines a basis $(\boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f)_{\boldsymbol{\alpha}\in I}$ of $\mathbb{O} \cdot f$. We call it *canonically associated to* $\{G_1, \dots, G_\ell\}$ in Algorithms 1 and 2. Moreover, the procedure of reduction of operators in $\mathbb{O}$ with respect to $\preceq$ by the Gröbner basis provides us with a procedure of normal form in $\mathbb{O}/\operatorname{Ann} f \simeq \mathbb{O} \cdot f$.

5. Holonomic certificates and companion identities

In the case of definite hypergeometric summation, the *certificate* of an identity

$$\sum_{i=0}^{L} \eta_i(n) U_{n+i} = 0 \qquad \text{where} \qquad U_n = \sum_{k\in\mathbb{Z}} u_{n,k},$$

is defined [24, 26] as the tuple $(R_{n,k}, \eta_0(n), \dots, \eta_L(n))$, where $R_{n,k} = v_{n,k}/u_{n,k}$ for a hypergeometric v in Eq. (1). In the case of an Ore algebra $\mathbb{O} = \mathbb{K}(\mathbf{u})[\partial; \sigma, \delta][\partial'; \sigma', \delta']$, we define the *certificate* of an identity

$$P \cdot F = \sum_{i=0}^{L} \eta_i \partial^i \cdot F = 0 \qquad \text{where} \qquad F = \partial^{-1}|_\Omega \cdot f, \tag{8}$$

as the tuple $\left((\phi_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in I}, \eta_0, \dots, \eta_L\right)$, where the $\phi_{\boldsymbol{\alpha}}$'s are defined to satisfy Eq. (5) for $H = P$. As in the hypergeometric case, this certificate alone allows the *verification* of Eq. (8), and a multivariate extension is possible.

Companion identities [24] are also found in our generalized setting. Starting from Eq. (7), we write $P = R + \partial S$ and apply $\partial^{-1}|_\Omega$ to get the *companion identity*

$$\partial' \partial^{-1}|_\Omega Q \cdot f + \partial^{-1}|_\Omega R \cdot f + \partial^{-1}|_\Omega \partial S \cdot f = 0.$$

Very often in applications, $R = 0$ or $\partial^{-1}|_\Omega \partial = 0$, which simplifies the identity. (The second case happens for instance when summing over natural boundaries.)

As an example, we develop a companion identity obtained from a generating function for the Bessel functions $J_n(z)$. We have

$$\sum_{n\in\mathbb{Z}} J_n(z) u^n = e^{\frac{uz}{2}\left(1-\frac{1}{u^2}\right)}, \tag{9}$$

which can be proved using the algorithms of the previous sections. More precisely, proving the identity obtained after dividing by the right-hand side with our algorithms, we get operators $P = 2uD_z$ and $Q = 2uD_z + S_n + u^2$ in the Ore algebra $\mathbb{K}(z,u,n)[D_z;1,D_z][S_n;S_n,0]$, that satisfy Eq. (7). A certificate for the identity (9) could be derived from the pair (P,Q). Writing

$$f_n = J_n(z) u^n e^{-\frac{uz}{2}\left(1-\frac{1}{u^2}\right)},$$

we have $P \cdot f + (S_n - 1) Q \cdot f = 0$. *Summation* of this equality over $\mathbb{Z}$ yields (9); *integration* over $(0, +\infty)$ yields

$$[2uf]_0^{+\infty} + (S_n - 1) \cdot \int_0^{+\infty} (Q \cdot f) dz = 0.$$

The left-hand term of the sum is zero when $n \geq 1$, so that the integral is constant for $n \geq 1$. Evaluating it at $n = 1$, the companion identity takes the form

$$\int_0^\infty u^n e^{-\frac{uz}{2}\left(1-\frac{1}{u^2}\right)} \left((1 + 2nuz^{-1}) J_n(z) - u J_{n+1}(z)\right) dz = 2u.$$

Conclusions

The value of the left factor ∂ in Eq. (2) and Eq. (5) does not play an important role in Algorithm 1, and can in fact be changed by any $L \in \mathbb{K}[\partial;\sigma,\delta]$. As an application, this yields an algorithm to compute particular solutions y_0 of a non-homogeneous linear equation $L \cdot y = f$ for a ∂-finite function f when a particular solution exists in $\mathbb{O} \cdot f$: solve $LQ = 1 \mod \operatorname{Ann} f$ by a clear extension of Algorithm 1 and set $y_0 = Q \cdot f$. This particular solution often has a nicer expression than that computed by the method of variation of the constant. More generally, a problem solved by Algorithm 1 is that of determining if the sum of a left ideal and a principal right ideal $L\mathbb{O}$ for $L \in \mathbb{K}[\partial;\sigma,\delta]$ contains a given element of an Ore algebra. This problem of solving a *mixed equation* is also close to questions related to the factorization of operators.

The crucial step of Algorithm 2 for definite summation and integration is the resolution of the linear system (4), which we perform by first uncoupling the system using an algorithm in [4], before appealing to specialized algorithms [1, 2] to solve equations in a *single* unknown function. Other uncoupling algorithms are available [6, 10], but we emphasize the desire for an algorithm that works directly at the level of *systems* of Ore operators. Indeed, from our first experiments, the uncoupling step is the computational bottleneck of Algorithm 2, in relation to the dimension of the vector space $\mathbb{O} \cdot f$; we hope that avoiding it could allow calculations in vector spaces of higher dimensions.

In the case of a sequence $(u_{n,k})$ with finite support for each n, the operator Q in (7) need not be computed to perform creative telescoping, since summing the right-hand side of (7) clearly yields 0. More generally, we call definite $\partial^{-1}|_\Omega$ over *natural boundaries* the case of definite $\partial^{-1}|_\Omega$ when the right-hand side of

$$P(\partial)\partial'^{-1}|_\Omega \cdot f = \partial'^{-1}|_\Omega \partial' \cdot (Q(\mathbf{u}, \partial, \partial') \cdot f)$$

can be predicted to be 0. In [12], we built on ideas of N. Takayama's to develop an algorithm which takes advantage of this situation to achieve efficiency. When both sides of Eq. (7) are needed, this algorithm from [12] used in conjunction to Algorithm 1 is an alternative to the fast algorithm presented above: after computing P by our algorithm from [12], the application of Algorithm 1 with $H = P$ in Eq. (5) makes it possible to compute Q from P. However, note that Algorithm 2 is more robust than this method in the sense that it does not need more than a ∂-finite description of the input to find a solution (see [12] for further details).

Finally, we point out that our algorithms allowed us to prove the following identity due to N. Calkin [11]

$$\sum_{k=0}^{n} \left(\sum_{j=0}^{k} \binom{n}{j} \right)^3 = n2^{3n-1} + 2^{3n} - 3n2^{n-2}\binom{2n}{n}$$

in only a few minutes of calculations. Using the multivariate extension of Zeilberger's algorithm [26] would require a not so easy four-fold summation.

References

[1] Abramov, S. A. Rational solutions of linear differential and difference equations with polynomial coefficients. *USSR Computational Mathematics and Mathematical Physics 29*, 11 (1989), 1611–1620. Translation of the Zhurnal vychislitel'noi matematiki i matematicheskoi fiziki.

[2] Abramov, S. A. Rational solutions of linear difference and q-difference equations with polynomial coefficients. In *Symbolic and algebraic computation* (New York, 1995), A. Levelt, Ed., ACM Press, pp. 285–289. Proceedings ISSAC'95, Montreal, Canada.

[3] Abramov, S. A., Bronstein, M., and Petkovšek, M. On polynomial solutions of linear operator equations. In *Symbolic and algebraic computation* (New York, 1995), A. Levelt, Ed., ACM Press, pp. 290–296.

[4] Abramov, S. A., and Zima, E. V. A universal program to uncouple linear systems, 1996. Preprint.

[5] Almkvist, G., and Zeilberger, D. The method of differentiating under the integral sign. *Journal of Symbolic Computation 10* (1990), 571–591.

[6] Barkatou, M. A. An algorithm for computing a companion block diagonal form for a system of linear differential equations. *Applicable Algebra in Engineering, Communication and Computing 4* (1993), 185–195.

[7] Bernstein, I. N. Modules over a ring of differential operators, study of the fundamental solutions of equations with constant coefficients. *Functional Analysis and Applications 5*, 2 (1971), 1–16 (Russian); 89–101 (English translation).

[8] Bernstein, I. N. The analytic continuation of generalized functions with respect to a parameter. *Functional Analysis and Applications 6*, 4 (1972), 26–40 (Russian); 273–285 (English translation).

[9] Björk, J. E. *Rings of Differential Operators*. North Holland P. C., Amsterdam, 1979.

[10] Bronstein, M., and Petkovšek, M. An introduction to pseudo-linear algebra. *Theoretical Computer Science 157*, 1 (1996).

[11] Calkin, N. J. A curious binomial identity. *Discrete Mathematics 131*, 1-3 (1994), 335–337.

[12] Chyzak, F., and Salvy, B. Non-commutative elimination in Ore algebras proves multivariate holonomic identities. To appear. Preliminary version available as INRIA Research Report #2799, `ftp://ftp.inria.fr/INRIA/publication/publi-ps-gz/RR/RR-2799.ps.gz`.

[13] Cohn, P. M. *Free Rings and Their Relations*. No. 2 in London Mathematical Society Monographs. Academic Press, 1971.

[14] Coutinho, S. C. *A Primer of Algebraic D-modules*. No. 33 in London Mathematical Society Student Texts. Cambridge University Press, 1995.

[15] Gosper, R. W. Decision procedure for indefinite hypergeometric summation. *Proceedings of the National Academy of Sciences USA 75*, 1 (Jan. 1978), 40–42.

[16] Graham, R. L., Knuth, D. E., and Patashnik, O. *Concrete Mathematics*. Addison-Wesley, 1989. A Foundation for Computer Science.

[17] Karr, M. Summation in finite terms. *Journal of the ACM 28*, 2 (1981), 305–350.

[18] Karr, M. Theory of summation in finite terms. *Journal of Symbolic Computation 1* (1985), 303–315.

[19] Koornwinder, T. H. On Zeilberger's algorithm and its q-analogue. *Journal of Computational and Applied Mathematics 48* (1993), 91–111.

[20] Paule, P., and Riese, A. A Mathematica q-analogue of Zeilberger's algorithm based on an algebraically motivated approach to q-hypergeometric telescoping. In *Fields Proceedings of the Workshop "Special Functions, q-Series and Related Topics", 12–23 June 1995* (Toronto, Ontario, 1996), Fields Institute for Research in Mathematical Sciences at University College. To appear.

[21] Petkovšek, M., Wilf, H., and Zeilberger, D. *A=B*. A. K. Peters, Ltd., Wellesley, Massachusset, 1996. ISBN 1-56881-063-6.

[22] Riese, A. A generalization of Gosper's algorithm to bibasic hypergeometric summation. *The Electronic Journal of Combinatorics 3*, R19 (1996), 1–16.

[23] Sabbah, C. Systèmes holonomes d'équations aux q-différences. In *D-Modules and Microlocal Geometry* (Berlin, 1993), M. Kashiwara, T. Monteiro-Fernandes, and P. Schapira, Eds., Walter de Gruyter & Co., pp. 125–147. Proceedings of the Conference D-Modules and Microlocal Geometry, Lisbon, 1990.

[24] Wilf, H. S., and Zeilberger, D. Rational functions certify combinatorial identities. *Journal of the American Mathematical Society 3* (1990), 147–158.

[25] Wilf, H. S., and Zeilberger, D. An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities. *Inventiones Mathematicae 108* (1992), 575–633.

[26] Wilf, H. S., and Zeilberger, D. Rational function certification of multisum/integral/"q" identities. *Bulletin of the American Mathematical Society 27*, 1 (July 1992), 148–153.

[27] Zeilberger, D. A fast algorithm for proving terminating hypergeometric identities. *Discrete Mathematics 80* (1990), 207–211.

[28] Zeilberger, D. A holonomic systems approach to special functions identities. *Journal of Computational and Applied Mathematics 32* (1990), 321–368.

[29] Zeilberger, D. The method of creative telescoping. *Journal of Symbolic Computation 11* (1991), 195–204.

INRIA-Rocquencourt and École polytechnique (France)
E-mail address: `Frederic.Chyzak@inria.fr`