An Extension of Zeilberger's Fast Algorithm to General Holonomic Functions

Frédéric Chyzak

An Extension of Zeilberger's Fast Algorithm to General Holonomic Functions

Frédéric Chyzak [1]

INRIA-Rocquencourt and École polytechnique (France)
Frederic.Chyzak@inria.fr

Abstract

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

French: Nous étendons l'algorithme rapide de Zeilberger pour la sommation hypergéométrique définie au cas des 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.

Key words: ∂-finite functions, holonomic functions, symbolic integration, symbolic summation, Zeilberger's algorithm.

In [33], Zeilberger initiated an algorithmic treatment of special functions that led to efficient algorithms for summation and integration [24]. In this approach, he considered a large class of functions and sequences that enjoys numerous closure properties, the class of *holonomic functions*. He also suggested how special functions and sequences from q-calculus are amenable to a similar treatment. Simple definitions of holonomy in the classical continuous and discrete cases are as follows.

Definition. A function $f(x_1, \ldots, x_n)$ is *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 *holonomic* when its multivariate generating function is holonomic.

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

Complete definitions of the q-calculus counterparts are too involved to be recalled here. Indeed, holonomic and q-holonomic functions and sequences are classically introduced in a different way, adopting a module-theoretic approach. More specifically, Zeilberger's definition [33, Section 2.2.4] is based on Bernstein's theory of *holonomic modules* [8,9], i.e., $\mathcal{D}$-modules with a certain dimension property; this theory was generalized to a notion of holonomic modules in the q-case by Sabbah [26, Definition 2.3.1]. We refer the reader to [10,15] for textbooks on holonomy. In the case of (continuous) functions, both definitions are in fact equivalent owing to [33, Lemma 4.1, 27, Theorem 2.4 and Appendix 6]. In the case of sequences, they are equivalent owing to [33, Section 3.3.2]; a third alternative definition is provided by Lipshitz [22].

As to the q-case, we only remark at this point that for a constant q that is not a root of unity, the q-dilations $f(q^{\alpha_1}x_1, \dots q^{\alpha_n}x_n)$ for $\alpha_i \in \mathbb{N}$ of a *q-holonomic function* $f(x_1, \dots, x_n)$ span a finite-dimensional vector space over the field of rational functions in the x_i's, and that similarly the shifts $u_{k_1+\alpha_1,\dots,k_n+\alpha_n}$ for $\alpha_i \in \mathbb{N}$ of a *q-holonomic* sequence $u_{k_1,\dots,k_n}$ span a finite-dimensional vector space over the field of rational functions in the q^{k_i}'s. We generically use *holonomic function* to refer to either of the four cases above.

Algorithms for the summation of holonomic sequences rely on the method of *creative telescoping* [34]. Given a bivariate sequence $u = (u_{n,k})$, 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 $v = (v_{n,k})$ 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 the differential case and to q-calculus [5,21,23,25]. Note that, for it to work, the rational functions η_i must not depend on k. On the other hand, the sequence u to be summed is usually described in applications by linear recurrences whose coefficients do involve k. In this regard, Eq. (1) can be viewed as the result of a sort of *elimination* of k from the description of u. This could be made more precise in terms of polynomial elimination in skew algebras of operators (see Section 1).

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 sequence $(u_{n_1,\dots,n_r})$ is hypergeometric when each quotient $u_{n_1,\dots,n_i+1,\dots,n_r}/u_{n_1,\dots,n_r}$ is a rational function in the n_i's. Equivalently, hypergeometric sequences are defined by linear first-order recurrences. Hypergeometry does not imply holonomy, as exemplified by the sequence u given by $u_{n,k} = (n^2 + k^2)^{-1}$ (see [30]).

To perform the elimination problem of determining an equation like (1), Zeilberger first gave a general but theoretical algorithm based on the calculation of a skew resultant [33]. He himself called this algorithm the *slow algorithm*, and proposed his *fast algorithm* [32] 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 [30,31]. 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 [13], 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

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

where the $P_n(x)$'s are the Legendre orthogonal polynomials and the $J_\nu(x)$'s are the Bessel functions of the first kind. In each case, we start from a description of the summand s in the left-hand side in terms of linear operators which vanishes at s, and we obtain an operator that vanishes at the right-hand side. Note that in both cases, the summand is not a hypergeometric term, nor does it satisfy any first-order linear ordinary differential equation.

Zeilberger's fast algorithm for *definite* hypergeometric summation is based on an algorithm for *indefinite* hypergeometric summation due to R. W. Gosper [17,18]. For sequences $u=(u_k)$ and $U=(U_k)$ 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 computes

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

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 compute this sum. 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 is run with increasing 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.

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* [13] is recalled in the next section. In contrast with Zeilberger's algorithm which is based on Gosper's approach to hypergeometric indefinite summation, our extension of Zeilberger's algorithm relies on an alternative approach based on Abramov's algorithm. In Section 2, we modify Abramov's algorithm to obtain an algorithm for indefinite ∂-finite summation and integration. This first algorithm always terminates. Then, we extend Zeilberger's algorithm to ∂-finite functions in Section 3. More specifically, this second algorithm is guaranteed to terminate for the subclass of *holonomic ∂-finite* inputs only; although, it may also terminate for some non-holonomic ∂-finite inputs, it is to be viewed as a heuristic method in this case. In the same section, we show how the algorithm extends to the iterated calculation of multiple sums. We next detail in Section 4 how the normal forms for ∂-finite functions used in those algorithms are obtained by methods of *Gröbner bases*. In Section 5, 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 [27,33] and the method extends to q-analogues [30]. All these algorithms are very similar in their structures and behaviours, and a unified description is in terms of *linear operators*. To this end, we introduced [13] 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. In [13], we described various methods based on Gröbner basis calculations to solve the elimination problem of determining the relevant analogues to Eq. (1). In the following, we set up notation so as to consider linear operators with coefficients in a ring $\mathbb{A}$, and over a field of constants $\mathbb{K}$.

Let $\mathbb{A}$ be a ring endowed with a ring endomorphism σ. Following [14], a *σ-derivation* δ on $\mathbb{A}$ is an additive endomorphism that satisfies the skew Leibniz law $(ab)^\delta = a^\sigma b^\delta + a^\delta b$ for all $a, b \in \mathbb{A}$. (By analogy with the prime notation for derivatives, we denote the application of σ's and δ's by powers.) Since the corresponding generalized differential operators are those of interest to our study, we often simply call a σ-derivation a derivation.

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

$$\partial_i \partial_j = \partial_j \partial_i \qquad \text{and} \qquad \partial_i a = a^{\sigma_i} \partial_i + a^{\delta_i}$$

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

It follows from the commutation rules above that each element p of an Ore algebra has a unique representation in the form $p = \sum_{\alpha_1, \ldots, \alpha_r} c_{\alpha_1, \ldots, \alpha_r} \partial_1^{\alpha_1} \ldots \partial_r^{\alpha_r}$ for coefficients $c_{\boldsymbol{\alpha}} \in \mathbb{A}$. For each i, the *degree* of p in ∂_i is defined using this form as the largest α_i such that there exists a non-zero $c_{\boldsymbol{\alpha}}$, or as $-\infty$ when none exists. The injectivity of the σ_i's is crucial to recover the usual properties of the degree with respect to sums and products of polynomials.

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 a commutative $\mathbb{K}$-algebra $\mathcal{F}$ is given, whose elements we call *functions*, and we require $\mathcal{F}$ to be a *left $\mathbb{O}$-module* containing $\mathbb{K}$. Usually in applications, $\mathcal{F}$ even contains $\mathbb{A}$. In any case, this makes it possible to consider operators with coefficients in $\mathbb{A}$. For instance, in the case of the Ore algebra $\mathbb{O} = \mathbb{K}(z)[\partial; 1, d/dz]$ and 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 $\mathbb{O} = \mathbb{K}(n)[\partial; S_n, 0]$ and linear recurrence operators, the algebra $\mathbb{K}^{\mathbb{N}}$ of sequences for term-wise addition and term-wise product is a left $\mathbb{O}$-module for the action $(\partial \cdot u)(n) = u_{n+1}$ and $(n \cdot u)(n) = nu_n$. By an abuse of notation, in the applications we freely use the name of the operator instead of the indeterminate ∂. For example, both Ore algebras above are also denoted $\mathbb{K}(z)[d/dz; 1, d/dz]$ and $\mathbb{K}(n)[S_n; S_n, 0]$. This is justified by the fact that neither d/dz nor S_n satisfies any algebraic relation that holds globally on $\mathcal{F}$.

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 $\boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f$, where $\boldsymbol{\partial}^{\boldsymbol{\alpha}} = \partial_1^{\alpha_1} \dots \partial_r^{\alpha_r}$ with $\alpha_i \in \mathbb{N}$, 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 $\mathbb{O}/\operatorname{Ann} f \simeq \mathbb{O} \cdot f$ as left $\mathbb{O}$-modules.

Of particular interest are *∂-finite functions*, which correspond in applications to functions and sequences defined by a finite number of equations and initial conditions. To define ∂-finite functions, we focus on Ore algebras whose $\mathbb{K}$-algebra $\mathbb{A}$ of coefficients is in fact a field $\mathbb{F}$.
Definition. Let $\mathbb{O} = \mathbb{F}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ be an Ore algebra over a field $\mathbb{F}$. 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 $\mathbb{F}$. A left ideal $\mathfrak{I}$ such that $\mathbb{O}/\mathfrak{I}$ is a finite-dimensional $\mathbb{F}$-vector space is also called a *∂-finite ideal*.

Here, "∂" is a mere symbol which bears no relation to the indeterminates ∂_i on which the Ore algebra $\mathbb{O}$ is built. It follows from the definition that a function f is ∂-finite if and only if its annihilating ideal $\operatorname{Ann} f$ is ∂-finite. For the Ore algebra $\mathbb{O} = \mathbb{F}[\partial_1; 1, d/dx_1] \dots [\partial_n; 1, d/dx_n]$ built on differential operators ∂_i's over the field $\mathbb{F} = \mathbb{C}(x_1, \dots, x_n)$, we recover the definition of holonomy, so that ∂-finiteness extends holonomy of (continuous) functions. Similarly for Ore algebras built on q-dilation or q-shift operators, ∂-finiteness also extends q-holonomy of functions and sequences, respectively. In contrast, ∂-finite sequences (with respect to ordinary shifts) need not be holonomic: in particular, all hypergeometric sequences would otherwise be holonomic.

2 Indefinite ∂-Finite ∂^{-1}

For an Ore algebra $\mathbb{O} = \mathbb{F}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ over a field $\mathbb{F}$, 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 all the anti-derivatives of a ∂-finite function f that lie in $\mathbb{O} \cdot f$. The algorithm always terminates, detecting when no $\partial^{-1} \cdot f$ 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.

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 (see Section 4);
(2) introduce undetermined coefficients $\phi_{\boldsymbol{\alpha}}$ for $\boldsymbol{\alpha} \in I$ 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 all systems of solutions $\phi_{\boldsymbol{\alpha}} \in \mathbb{F}$;
(4) if solvable, return all the $Q = \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}$; otherwise return $\perp$.

Algorithm 1. Indefinite ∂-finite summation

Let us insist on the algorithm not requiring holonomy of the input function f, but merely ∂-finiteness; in contrast, the algorithm of Section 3 will require both. On the other hand, neither ∂-finiteness nor holonomy is sufficient to ensure the existence of anti-derivatives in the module $\mathbb{O} \cdot f$, as exemplified by the holonomic function $1/x$ and by the holonomic hypergeometric sequence $1/n$.

2.1 The Algorithm

Algorithm 1 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 operators Q such that $\partial^{-1} \cdot f = Q \cdot f$. We proceed to establish the following theorem.

Theorem. *Let $\mathbb{F}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ be an Ore algebra over the field $\mathbb{F}$ and ∂ be any of the ∂_i's. Assume that $\mathbb{F}$ admits a decision algorithm to solve linear equations $L \cdot \phi = 0$ where $L \in \mathbb{F}[\partial; \sigma, \delta]$ for all solutions ϕ in $\mathbb{F}$, and that σ is invertible. 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 in Algorithm 1 that the input be the whole annihilating ideal $\operatorname{Ann} f$ of a ∂-finite function f can be weakened: the algorithm also terminates on any ∂-finite subideal of $\operatorname{Ann} f$; however, it may fail to find some anti-derivatives with an incomplete input. This change of ideals corresponds to a change of ∂-finite functions by introducing parasitic solutions.

The key point is to make the action of the derivation operator ∂ on the finite-dimensional vector space $\mathbb{O} \cdot f$ explicit. Let F be any function in $\mathbb{O} \cdot f$. We fix an $\mathbb{F}$-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. Equivalently, this yields the $\mathbb{F}$-basis $\{\boldsymbol{\partial}^{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha} \in I}$ of $\mathbb{O}/\operatorname{Ann} f$. Then $F = Q \cdot f$ where $Q \in \mathbb{O}/\operatorname{Ann} f$ can be written $Q = \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}$. With the assumption $F = \partial^{-1} \cdot f$, i.e., $\partial \cdot F = f$, we have $\partial Q = 1 \mod \operatorname{Ann} f$, i.e.,

$$\partial Q = \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}}^{\sigma} \boldsymbol{\partial}^{\boldsymbol{\alpha}} \partial + \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}}^{\delta} \boldsymbol{\partial}^{\boldsymbol{\alpha}} = 1. \tag{2}$$

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 with a particular choice of basis of $\mathbb{O} \cdot f$. For the sake of clarity, we postpone the description of these ingredients to Section 4.

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

$$\sum_{\boldsymbol{\beta} \in I} \lambda_{\boldsymbol{\alpha},\boldsymbol{\beta}} \phi_{\boldsymbol{\beta}}^{\sigma} + \phi_{\boldsymbol{\alpha}}^{\delta} = \mu_{\boldsymbol{\alpha}},$$

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

$$\Lambda \Phi^{\sigma} + \Phi^{\delta} = M. \tag{3}$$

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, and no anti-derivative exists in $\mathbb{O} \cdot f$.

Let us detail how to solve (3). Each equation of the system may involve several unknown functions. Excluding ongoing research still to be further developed [7], we do not know of algorithms to solve this kind of linear system directly; the first step is therefore to "triangularize" the system, when possible, so as to obtain an equation in a single unknown function together with a system to be solved step by step. More precisely, the point is to put the system under the triangular, more generally trapezoidal, shape

$$\sum_{j=i}^{|I|} T_{i,j}(\partial) \cdot \psi_j = \nu_i \qquad i = 1, \ldots, d, \qquad d \leq |I|,$$

for *operators* $T_{i,j} \in \mathbb{F}[\partial; \sigma, \delta]$, rational functions $\nu_i \in \mathbb{F}$ and unknown functions ψ_j that are linear combinations of the $\phi_{\boldsymbol{\alpha}}$'s and such that the latter can be computed once the ψ_j's are known. This can be achieved for any Ore operator ∂, provided that σ be invertible, by appealing to an algorithm due to Abramov and Zima [4]. Indeed, introduce the new Ore algebra $\mathbb{F}[\partial^*; \sigma^*, \delta^*]$ where $\sigma^* = \sigma^{-1}$ (the inverse of σ) and ∂^* acts on $\mathbb{F}$ by $\delta^* = -\sigma^{-1}\delta$. Applying σ^{-1} to (3) yields the system

$$\Lambda^{\sigma^{-1}} \Phi - \partial^* \cdot \Phi = M^{\sigma^{-1}},$$

where $\Lambda^{\sigma^{-1}}$ and $M^{\sigma^{-1}}$ are known and Φ is the unknown. This is exactly the input form of the algorithm in [4]. Once the system has been "triangularized", we have to solve successive 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. Recurrences are an instance of the more general case when ∂ acts by $\delta = \sigma - 1$ (where 1 is the identity). We then usually work with the operator σ of (ordinary or q-) shift instead of the operator δ of (ordinary or q-) difference, because both operator algebras $\mathbb{F}[\delta; \sigma, \delta]$ and $\mathbb{F}[\sigma; \sigma, 0]$ are isomorphic when $\delta = \sigma - 1$. After the triangularization 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 [1,2].
The case of (ordinary) differential equations. In the differential case, the application σ is the identity, so that the change of Ore operators in the triangularization step above is trivial ($\partial^* = -\partial$). We next solve the successive differential equations by another algorithm of Abramov's [1].

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

$$\partial Q = \sum_{\boldsymbol{\alpha}\in I} \phi^{\sigma}_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} \partial + \sum_{\boldsymbol{\alpha}\in I} \phi^{\delta}_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}} = H \tag{4}$$

for $H \in \mathbb{O}/\operatorname{Ann} f$ makes it possible to detect if $H \cdot f$ has an anti-derivative in $\mathbb{O} \cdot f$. This affects the vector M in the system (3) in a linear way only, which will be used in the algorithm for creative telescoping of the next section.

2.2 Example: Harmonic Summation Identities

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 $\sum_{k=1}^{n} k^{-1}$, are classically obtained by summation by parts or by techniques of generating functions. (See also M. Karr's general algorithm [19,20].) One can alternatively find closed form evaluations for Harmonic sums using our algorithm. Introducing $f_n = \binom{n}{m} H_n$, we show the following equivalent form of the identity above:

$$F_n = \sum_{k=1}^{n} f_k = \frac{(n+1)^2}{(m+1)^2} f_n - \frac{(n-m)(n-m+1)}{(m+1)^2} f_{n+1}. \tag{5}$$

To this end, let us compute the first two shifts of f:

$$f_{n+1} = \frac{n+1}{n+1-m}\binom{n}{m}\left[H_n + \frac{1}{n+1}\right] = \frac{n+1}{n+1-m} f_n + \frac{1}{n+1-m}\binom{n}{m};$$
$$f_{n+2} = \frac{n+2}{n+2-m} f_{n+1} + \frac{n+1}{n+2-m}\left[\frac{1}{n+1-m}\binom{n}{m}\right].$$

Taking the appropriate linear combination of the above equations, one gets the following linear homogeneous recurrence with coefficients in $\mathbb{Q}(n,m)$

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

In the case of more complex expressions, one would appeal to the closure properties of ∂-finite functions under addition and product and use algorithms described in [13]. As a consequence of the above relation, 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 $(S_n - 1)Q - 1$. Then, the system (3) takes the form

$$\begin{cases} (2n+3)\beta_{n+1} + (n-m+2)(\alpha_{n+1} - \beta_n) = 0, \\ (n+1)^2\beta_{n+1} + (n-m+1)(n-m+2)(\alpha_n + 1) = 0. \end{cases}$$

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

$$\frac{(n+2)^2}{(n-m+2)(n-m+3)}\beta_{n+2} - \frac{2n+3}{n-m+2}\beta_{n+1} + \beta_n + 1 = 0,$$

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

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

The sum F satisfies $(S_n - 1)\cdot(F - f) = f = (S_n - 1)\cdot(Q \cdot f)$, whence $F_n - [(Q+1)\cdot f](n)$ is a constant seen to be 0 at $n = 1$. This proves Eq. (5).

The same algorithm would find evaluations of other harmonic sums like

$$\sum_{k=1}^{n} \frac{(-1)^k}{\binom{m}{k}} H_k = \frac{(-1)^n}{\binom{m}{n}} \left[\frac{n+1}{m+2} H_n + \frac{m+1-n}{(m+2)^2} \right] - \frac{m+1}{(m+2)^2},$$

$$\sum_{k=1}^{n} (2k+1) H_k^3 = (n+1)^2 H_n^3 - \frac{3}{2} n(n+1) H_n^2 + \frac{3n^2+3n+1}{2} H_n - \frac{3}{4} n(n+1),$$

$$\sum_{k=1}^{n} k^2 H_{n+k} = \frac{n(n+1)(2n+1)}{6} [2H_{2n} - H_n] - \frac{n(n+1)(10n-1)}{36}.$$

(For the second identity, note that the iterated application of the shift operator on H_n^3 spans a vector space with basis $\{1, H_n, H_n^2, H_n^3\}$; for the third identity, compute the indefinite sum $\sum_{k=1}^{n} k^2 H_{m+k}$, then set the parameter m to n.) However, the method cannot find

$$\sum_{k=1}^{n} H_k^3 = (n+1) H_n^3 - \frac{3}{2}(2n+1) H_n^2 + 3(2n+1) H_n + \frac{1}{2} H_n^{(2)} - 6n,$$

where $H_n^{(2)} = \sum_{k=1}^{n} k^{-2}$ denotes generalized harmonic numbers, unless the presence of the term in $H_n^{(2)}$ is *guessed*, together with its exponent, and the method slightly modified to perform Gröbner basis calculations and reductions in the module $\mathbb{O} \cdot H_n^3 + \mathbb{O} \cdot H_n^{(2)}$.

3 Fast Definite ∂-Finite ∂_Ω^{-1}

For an Ore algebra $\mathbb{O} = \mathbb{A}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$, let again ∂ be one 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 that there is a linear operator ∂_Ω^{-1} defined on $\mathcal{F}$ such that $\partial \partial_\Omega^{-1} = 0$. Here, the notation ∂_Ω^{-1} must be viewed as a single symbol, where Ω refers to no specific mathematical object, simply standing as a remembrance of the notation $\int_\Omega$ used to indicate an integration over a domain Ω. (In [13], we used a less general definition for ∂_Ω^{-1}, requiring that $\partial_\Omega^{-1} \partial$ 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; in the case of a ∂-finite function of a more general type, the algorithm can sometimes be fruitfully used in a heuristic way to compute a definite anti-derivative.

Zeilberger's fast algorithm is guaranteed to terminate on holonomic hypergeometric sequences only. Similarly in the case of (possibly mixed, possibly q-)

differential or difference operators, we call a simultaneously ∂-finite and holonomic function *holonomic ∂-finite*. We already described the connection between ∂-finiteness and holonomy: except in the case of sequences, both concepts are equivalent (up to a minor technical condition). Our phrasing may thus seem redundant; the analogy to the case of holonomic hypergeometric sequences and the fact that it refers to a restricted set of operator types make its raison d'être: our algorithm inputs a description of the annihilating ideal of any ∂-finite function; we prove its termination for holonomic ∂-finite functions.

3.1 The Algorithm

Let us first vindicate our algorithm by the case of holonomic functions in the differential case. A (continuous) holonomic function $f(x,y)$ is a ∂-finite function with respect to the Ore algebra $\mathbb{O} = \mathbb{C}(x,y)[D_x;1,d/dx][D_y;1,d/dy]$ built on (ordinary) differential operators. (Here, D_x and D_y act by $\delta_x = d/dx$ and $\delta_y = d/dy$, respectively.) The original description of holonomy in the framework of $\mathcal{D}$-modules [8,9] implies that there exists a non-zero operator in

$$\mathrm{Ann}_{\mathbb{O}}\, f \cap \mathbb{C}(x)[\boldsymbol{\partial};\mathbf{1},\boldsymbol{\delta}]$$

(see [33, Lemma 4.1]). As a result, there is a non-trivial identity of the form

$$\sum_{i=0}^{L} \eta_i(x) D_x^i \cdot f = D_y \cdot [Q(x,y,D_x,D_y)\cdot f]$$

mimicking Eq. (1) for $Q \in \mathbb{O}$. This existence property transfers to the discrete case through generating functions and similar results hold for q-analogues [26].

More generally, for a ∂-finite function f with respect to an Ore algebra

$$\mathbb{F}(x_1,\ldots,x_s)[\partial;\sigma,\delta][\partial';\sigma',\delta']$$

in *two* operators ∂ and ∂' and such that ∂ commutes with elements of $\mathbb{F}$ but not with the x_i's, we look for solutions of

$$P(\partial')\cdot f = \sum_{i=0}^{L} \eta_i {\partial'}^i \cdot f = \partial \cdot [Q(\mathbf{x},\partial,\partial')\cdot f]\,, \tag{6}$$

where $P \neq 0$ and the η_i's *do not depend on* $\mathbf{x}$. The existence of a non-trivial pair (P,Q) is not guaranteed in general, but it is thanks to [33, Lemma 4.1] in the (classical) holonomic setting and *mutatis mutandis* in the q-holonomic case [26]. We summarize the result of this section in the following theorem.
Theorem. *Let $\mathbb{F}(\mathbf{x})[\partial;\sigma,\delta][\partial';\sigma',\delta']$ be an Ore algebra and assume that $\mathbb{F}(\mathbf{x})$ admits a decision algorithm to solve linear equations $L\cdot\phi = 0$ where $L \in \mathbb{F}(\mathbf{x})[\partial;\sigma,\delta]$ for all solutions ϕ in $\mathbb{F}(\mathbf{x})$. Assume further that σ is invertible.*

INPUT: a basis B for the annihilating ideal of a ∂-finite function f.
OUTPUT: a pair of operators (P, Q) satisfying (6).

(1) from B, compute a Gröbner basis G and get the finite basis $\{\partial^\alpha \partial'^\beta\}_{(\alpha,\beta)\in I}$ of $\mathbb{O}/\operatorname{Ann} f$ canonically associated to G (see Section 4);
(2) for $L = 0, 1, \dots$:
   (a) introduce undetermined coefficients $\phi_{\alpha,\beta}$ for $(\alpha, \beta) \in I$ and η_i for $i = 0, \dots, L$ and rewrite $\partial \sum_{(\alpha,\beta)\in I} \phi_{\alpha,\beta} \partial^\alpha \partial'^\beta - \sum_{i=0}^L \eta_i \partial'^i$ in this basis by reduction by G;
   (b) solve the corresponding system of first order linear equations for all systems of solutions $\eta_i \in \mathbb{F}$ and $\phi_{\alpha,\beta} \in \mathbb{F}(\mathbf{x})$;
   (c) if solvable, return all the solutions (P, Q) for $P = \sum_{i=0}^L \eta_i \partial^i$ and $Q = \sum_{(\alpha,\beta)\in I} \phi_{\alpha,\beta} \partial^\alpha \partial'^\beta$; otherwise loop.

Algorithm 2. Definite ∂-finite summation

When there exists a pair (P, Q) that satisfies (6), Algorithm 2 terminates and returns such a pair. This is guaranteed to happen when f is holonomic ∂-finite.

Whenever we know an operator P that makes Eq. (6) solvable for Q, we can use Algorithm 1 to get Q. Indeed, it was noted that the value of H in Eq. (4) is inessential; letting $H = P$ (where ∂' replaces $\boldsymbol{\partial}$) makes it possible, after reduction modulo $\operatorname{Ann} f$, to apply Algorithm 1, the vector M in (3) 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 Eq. (6) is solvable for Q. Therefore, we use a variant of Algorithm 1 to solve the system (3) for Φ and M simultaneously. This corresponds to refinements of Abramov's algorithms that mimic Zeilberger's extension [34] of Gosper's algorithm.

Thus, our algorithm for the definite case 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 the system (3) is solvable, we have finished, otherwise we increase L. Note that even in the holonomic case, no bound on L is known except for the case of *(q-)proper-hypergeometric terms* designed by Wilf and Zeilberger [30] (see [28] for an improvement).

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 operators

$$z^2D_z^2 + zD_z + z^2 - k^2, \quad zD_zS_k + (k+1)S_k - z \quad \text{and} \quad zD_z + zS_k - k$$

in the Ore algebra $\mathbb{O} = \mathbb{Q}(k,z)[D_z;1,D_z][S_k;S_k,0]$. This relates to the previous section by setting $\partial = \delta = S_k - 1$, $\sigma = S_k$, $\partial' = \delta' = D_z$ and $\sigma' = 1$. An important fact whose significance will only become clear in Section 4 is that the three operators above constitute a Gröbner basis of $\operatorname{Ann} J_k(z)$ (with respect to a total degree order $\preceq$ satisfying $D_z \preceq S_k$). This Gröbner basis can be obtained from the classical pure differential equation and pure recurrence equation satisfied by $J_k(z)$. It follows from an algorithm from [13] that the squares $J_k(z)^2$ are also ∂-finite and are described by the Gröbner basis

$$\begin{cases} zD_z^2 + (-2k+1)D_z - 2S_kz + 2z, \\ zD_zS_k + zD_z + (2k+2)S_k - 2k, \\ z^2S_k^2 - 4(k+1)^2S_k - 2z(k+1)D_z + 4k(k+1) - z^2. \end{cases}$$

The module $\mathbb{O} \cdot J_k(z)^2$ has to contain $J_k(z)^2$ and its first two derivatives, or equivalently $J_k(z)^2$, $J_k(z)J_k'(z)$ and $J_k'(z)^2$, which are linearly independent. Thus, the system above generates the ideal $\operatorname{Ann} J_k(z)^2$ in $\mathbb{O}$ and the module $\mathbb{O}/\operatorname{Ann} J_k(z)^2$, which is isomorphic to $\mathbb{O} \cdot J_k(z)^2$, is a three-dimensional vector space, with basis $\{1, D_z, S_k\}$. Knowing the generating function of the $J_k(z)$, one could prove that $J_k(z)$ and $J_k(z)^2$ are holonomic in (k,z), so that Algorithm 1 has to terminate on $\operatorname{Ann} J_k(z)^2$; else, the algorithm has to be run in a heuristic way. To this end, we introduce a generic $Q = u_k + v_kS_k + w_kD_z$. Next, we let $L = 1$ and introduce two parameters $\eta_0(z)$ and $\eta_1(z)$ in Eq. (6) to get a solution. The system (3) then reads

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

and we obtain the constraint $\eta_0 = 0$ ($\eta_1(z)$ is any rational function in z). We set $\eta_1(z)$ to 1, so that

$$P = D_z \qquad \text{and} \qquad Q = -\left(\frac{k}{z} + \frac{D_z}{2}\right).$$

With these values for P and Q, we have after creative telescoping:

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

from which follows by linearity that

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

is identically zero 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, verified to be 0 when $z = 0$. This proves Neumann's theorem.

3.3 Extension to Multivariate Anti-Derivatives — Application to Iterated Multiple Sums

So far in this section, we have described the (q-)summation and the (q-)integration of *bivariate* functions f: starting from a system describing f in terms of *two* operators ∂ and ∂', we have used an algorithm to compute a *single* operator in ∂' for the definite anti-derivative with respect to ∂. In fact, the dependency of the sum or integral in a single variable is inessential and we can perform summations and integrations of *multivariate* ∂-finite functions f whose definite anti-derivatives with respect to ∂ are still multivariate ∂-finite. This stems from the fact that the requirement in Eq. (6) that P and Q be polynomials in a single ∂' (disregarding the dependency of Q in ∂) can be relaxed. For a tuple $\boldsymbol{\partial}$ of operators and *another* operator ∂ whose ∂_Ω^{-1} is to be computed, Eq. (6) then becomes

$$P(\boldsymbol{\partial}) \cdot f = \sum_{(\boldsymbol{\alpha},\beta)} \eta_{\boldsymbol{\alpha},\beta} \boldsymbol{\partial}^{\boldsymbol{\alpha}} \partial^{\beta} \cdot f = \partial \cdot \left[Q(\mathbf{x}, \boldsymbol{\partial}, \partial) \cdot f \right], \tag{7}$$

where $P \neq 0$ does not involve ∂ and the $\eta_{\boldsymbol{\alpha},\beta}$'s do not involve $\mathbf{x}$. Instead of running a loop over univariate polynomials of increasing degree L, like in step (2) of Algorithm 2, one runs a loop to allow polynomials P over more and more multivariate monomials in $\boldsymbol{\partial}$, in a way to be detailed below. Now, instead of stopping the loop in Algorithm 2 after the first solution found, one continues so as to obtain a system of operators $P - \partial Q$, until the P's span a ∂-finite ideal. Termination is proven in the case of holonomic ∂-finite functions in the same way as in the simpler univariate case, by appealing to [33, Lemma 4.1].

More specifically, let $\mathbb{F}(\mathbf{x})[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}][\partial; \sigma, \delta]$ be an Ore algebra in the operators ∂ and $\boldsymbol{\partial} = (\partial_1, \dots, \partial_r)$, ∂ being none of the ∂_i's. We still assume that $\mathbb{F}(\mathbf{x})$ admits a decision algorithm to solve linear equations $L \cdot f = 0$ where $L \in \mathbb{F}(\mathbf{x})[\partial; \sigma, \delta]$ for all solutions in $\mathbb{F}(\mathbf{x})$. In order to consider polynomials P over more and more multivariate monomials in $\boldsymbol{\partial}$ instead of the polynomial $\sum_{i=0}^{L} \eta_i \partial'^i$ in step (2) of Algorithm 2, we propose several options:

(1) Let $\{m_i\}_{i\in\mathbb{N}}$ be a sequence that runs over all monomials $\boldsymbol{\partial}^{\boldsymbol{\alpha}}$ in $\boldsymbol{\partial}$. Consider the polynomial $P = \sum_{i=0}^{L} \eta_i m_i$ with undetermined coefficients η_i .
(2) Consider the polynomial $P = \sum_{|\boldsymbol{\alpha}|\le L} \eta_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}$ with total degree L and undetermined coefficients $\eta_{\boldsymbol{\alpha}}$.

(3) Using a term order $\preceq$ on the monomials in $\boldsymbol{\partial}$ (see the definition in [13]), an extension of the FGLM algorithm [16] can be used to determine the successive sets of monomials involved in the P's. The key idea is that when a polynomial P has been obtained, any multiple of its leading monomial m need not be considered any longer in the next loops. In this way, we obtain an algorithm for definite anti-derivative that is very close in spirit to the algorithms for addition and product of ∂-finite functions which are described in [13].

We now illustrate this algorithm by proving the double summation identity

$$\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}(-1)^{n+r+s}\binom{n}{r}\binom{n}{s}\binom{n+r}{r}\binom{n+s}{s}\binom{2n-(r+s)}{n}=\sum_{k=0}^{\infty}\binom{n}{k}^4, \quad (8)$$

which can be found in [24]. The difficult part of the proof is to compute a recurrence operator for the left-hand side. Such an operator was given by K. Wegschaider by a direct evaluation of the double sum [28]. Here, we compute the left-hand side using Algorithm 2 in an iterated way:

(1) Starting from the system

$$\begin{cases} (n+1)(n+1-r)(n+1-s)(n+1-r-s)S_n \\ \quad + (n+s+1)(n+r+1)(2n+1-r-s)(2n+2-r-s), \\ (r+1)^2(2n-r-s)S_r + (n+r+1)(n-r)(n-r-s), \\ (s+1)^2(2n-r-s)S_s + (n+s+1)(n-s)(n-r-s), \end{cases} \quad (9)$$

which defines the summand in the double sum as a ∂-finite function with respect to the Ore algebra $\mathbb{Q}(n,r,s)[S_n;S_n,0][S_r;S_r,0][S_s;S_s,0]$, we first compute the simple sum over s. To this end, we apply the multivariate extension of Algorithm 2, setting $\boldsymbol{\partial}=(S_n,S_r)$ and $\partial=S_s-1$, so as to perform the inner summation (with respect to s) first. We select option (2); for $L=2$, we introduce the polynomials

$$\begin{cases} P=\eta_{0,0}+\eta_{1,0}S_n+\eta_{0,1}S_r+\eta_{2,0}S_n^2+\eta_{1,1}S_nS_r+\eta_{0,2}S_r^2, \\ Q=\phi(s) \end{cases}$$

under undetermined form. Reducing $Z=P-(S_s-1)Q$ by the system (9) and solving the reduced system for rational functions ϕ and $\eta_{i,j}$ yields two operators P, one with monomials S_n^2, S_r^2, S_n, S_r and 1, the other with monomials S_r^2, S_n, S_r and 1, both with large polynomial coefficients in n and r. Solving the corresponding Eq. (7), it turns out that the right-hand side vanishes so that both P's in fact annihilate the sum over s.

(2) Moreover, these two P's span a ∂-finite ideal with respect to the Ore algebra $\mathbb{Q}(n,r)[S_n;S_n,0][S_r;S_r,0]$. Thus, we can apply the simple case of Algorithm 2 so as to perform the summation over r. This yields the

following operator R that annihilates the double sum:

$$(n+2)^3 S_n^2 - 2(2n+3)(3n^2+9n+7)S_n - 4(4n+5)(4n+3)(n+1).$$

So far, we have *computed* an operator R which annihilates the left-hand side of Eq. (8). Applying Algorithm 2 again to the right-hand side, we get the *same* second-order operator. Since both sides of Eq. (8) agree at $n=0$ and $n=1$, where they are 1 and 2, respectively, and since the leading coefficient $(n+2)^3$ of R does not vanish for $n \in \mathbb{N}$, we get Eq. (8) by induction on $n \in \mathbb{N}$.

4 Effective Calculations with ∂-Finite Ideals

In the algorithms for hypergeometric summation, an important rôle 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 $\mathbb{Q}(n) \cdot t_n$, so that each sequence t' under consideration can be represented by the single rational function r such that $t' = rt$.

In our extension to the case of ∂-finite functions with respect to an Ore algebra $\mathbb{O} = \mathbb{F}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ over a field $\mathbb{F}$, the rôle of $\mathbb{F}(n) \cdot t_n$ is undertaken by the *finite-dimensional* vector space $\mathbb{O} \cdot f = \bigoplus_{\boldsymbol{\alpha} \in I} \mathbb{F}\, \boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f$ for a finite set I. Each function under consideration in our algorithms are represented by its rational coordinates $\phi_{\boldsymbol{\alpha}} \in \mathbb{F}$ in 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 to 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 $(P \sum_{\boldsymbol{\alpha} \in I} \phi_{\boldsymbol{\alpha}} \boldsymbol{\partial}^{\boldsymbol{\alpha}}) \cdot f$ 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 [13]. 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 [13]) determines a suitable set I in the following way. Call $h_i = \boldsymbol{\partial}^{\boldsymbol{\alpha}_i}$ the leading term of G_i with respect to $\preceq$. Then, consider the set $I = \{\boldsymbol{\alpha} \mid \forall i \quad h_i \not\preceq \boldsymbol{\partial}^{\boldsymbol{\alpha}}\}$ of those terms $\boldsymbol{\partial}^{\boldsymbol{\alpha}}$ greater than none of the h_i's. This set defines a basis $\{\boldsymbol{\partial}^{\boldsymbol{\alpha}} \cdot f\}_{\boldsymbol{\alpha} \in I}$ of $\mathbb{O} \cdot f$ which we have called *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 $\mathbb{O}/\operatorname{Ann} f \simeq \mathbb{O} \cdot f$. Finally, note that (skew) Gröbner bases can be computed from any basis by a variant of Buchberger's algorithm [13].

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 [29,31] as the tuple $(R_{n,k}, \eta_0(n), \ldots, \eta_L(n))$, where $R_{n,k} = v_{n,k}/u_{n,k}$ for a hypergeometric v in Eq. (1). In the case of functions specified by operators in the Ore algebra $\mathbb{K}(\mathbf{x})[\partial;\sigma,\delta][\partial';\sigma',\delta']$, we define the *certificate* of an identity

$$P \cdot F = \sum_{i=0}^{L} \eta_i \partial'^i \cdot F = 0 \qquad \text{where} \qquad F = \partial_\Omega^{-1} \cdot f, \tag{10}$$

as the tuple $\left((\phi_{\boldsymbol\alpha})_{\boldsymbol\alpha\in I}, \eta_0, \ldots, \eta_L\right)$, where the $\phi_{\boldsymbol\alpha}$'s are defined to satisfy Eq. (4) for $H = P$ (where ∂' replaces $\boldsymbol\partial$). As in the hypergeometric case, this certificate alone allows the *verification* of Eq. (10), and a multivariate extension is possible along the lines of Section 3.3.

An extension of the *companion identities* described by Wilf and Zeilberger in the hypergeometric case [29] is available in the generalized setting of ∂-finite functions. Starting from Eq. (6), we write the Euclidean division $P = R + \partial' S(\partial')$ of P by ∂', and we apply ∂'^{-1}_Ω to get the following new form of a *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

$$\sum_{n\in\mathbb{Z}} J_n(z) u^n = e^{\frac{uz}{2}\left(1-\frac{1}{u^2}\right)}, \tag{11}$$

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 $\mathbb{K}(u,z,n)[D_z;1,D_z][S_n;S_n,0]$, that satisfy Eq. (6) with $\partial = S_n - 1$ and $\partial' = D_z$. A certificate for the identity (11) 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 with respect to n over $\mathbb{Z}$ yields (11); *integration* with respect to z over $(0, +\infty)$ yields

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

when $u > 1$ or $-1 < u \leq 0$. 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 $n = 1$ and after some rewriting, the companion identity for $n \geq 1$, $u > 1$ or $-1 < u \leq 0$ reads

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

Conclusions

The value of the left factor ∂ in Eq. (2) and Eq. (4) plays no important rôle in Algorithm 1, and can in fact be changed to any $L \in \mathbb{F}[\partial; \sigma, \delta]$. As an application, this yields an algorithm to compute particular solutions y_0 of a non-homogeneous linear equation $L \cdot y = H \cdot f$ for a ∂-finite function f and $H \in \mathbb{F}[\boldsymbol{\partial}; \boldsymbol{\sigma}, \boldsymbol{\delta}]$ when a particular solution exists in $\mathbb{O} \cdot f$: solve $LQ = H \mod \operatorname{Ann} f$ by an obvious 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 constants. 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 $L \in \mathbb{F}[\partial; \sigma, \delta]$ 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 (3), which we perform by first uncoupling the system using an algorithm from [4], before appealing to specialized algorithms [1,2] to solve equations in a *single* unknown function. Other uncoupling algorithms are available [6,11], 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.

Besides, the theory of holonomy is restricted to three types of operators (derivation, shift and q-shift). A challenging problem is to develop a holonomic theory for other types of operators, which would extend the scope of Algorithm 2 by enlarging the class of holonomic ∂-finite functions. In the same vein, designing new classes of (non-holonomic) functions for which Eq. (6) is *a priori* guaranteed to be solvable for non-trivial pairs (P, Q) would turn the so far heuristic use of the algorithm into a guaranteed method.

An extension of Algorithm 2 has been presented to compute a system of operators that annihilate a multivariate sum or integral known to be ∂-finite. We noted that there is some freedom in the way the outer loop is run in step (2) of this algorithm. In particular, the version based on the FGLM algorithm [16] seems the most interesting in practice, because it refrains from introducing useless monomials. Elaborating on works by P. Verbaeten, K. Wegschaider recently obtained a clever algorithm for hypergeometric multiple summation [28]. The method dramatically reduces the number of terms to be considered in recurrences in order to obtain an annihilating operator for the sum. Trying to combine this approach with our algorithm for the ∂-finite definite case is a promising direction for research.

In the case of a sequence $(u_{n,k})$ with finite support for each n, the operator Q in (6) need not be computed to perform creative telescoping, since summing the right-hand side of (6) clearly yields 0. More generally, the case of definite ∂_Ω^{-1} when the right-hand side of

$$P(\partial')\partial_\Omega^{-1} \cdot f = \partial_\Omega^{-1}\partial \cdot [Q(\mathbf{x}, \partial, \partial') \cdot f]$$

can be predicted to be 0 is called definite ∂_Ω^{-1} over *natural boundaries*. In [13], 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. (6) are needed, this algorithm from [13] used in conjunction with Algorithm 1 is an alternative to the fast algorithm presented above: after computing P by our algorithm from [13], the application of Algorithm 1 with $H = P$ in Eq. (4) 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 [13] for further details).

As a last example, we point out that our algorithms allowed us to prove the following identity due to N. Calkin [12]

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

in only a few minutes of calculations. Using the algorithm for multivariate summation that was developed by H. Wilf and D. Zeilberger basing on Sister Celine's technique [31] would require a not-so-easy four-fold summation.

We finish with a few words about the programs used and the timings obtained so as to demonstrate the efficiency of our algorithms. No integrated implementation of them is available yet, so that we ran them step by step, using our package MGFUN to compute the Gröbner bases needed. (This package is available from the URL `http://algo.inria.fr/libraries/` and by anonymous ftp from `ftp.inria.fr:/INRIA/Projects/algo/programs`; the part of the package concerning Gröbner basis calculations has been integrated into MAPLE V RELEASE 5.) The calculations of the examples in the previous sections were performed with the system MAPLE on a DecStation 3000 300X (Alpha); they required between a few seconds and a few minutes each: a matter of seconds for each harmonic identity in Section 2, for Neumann's addition theorem in Section 3 and for the generating function of the Bessel functions and its companion identity in Section 5; 195 seconds (and 12 MB) for Calkin's identity above, and 390 seconds (and 15 MB) only for the double sum in Eq. (8). As a comparison, the latter two identities could previously neither be obtained performing the elimination by Gröbner bases only, due to too long computations (over a month), nor even with our extension of Takayama's algorithm from [13], due to a run out of memory (over 300 MB used).

Acknowledgement. I thank Peter Paule, Bruno Salvy and anonymous referees for their very accurate and constructive remarks on preliminary versions.

References

[1] S. A. Abramov, Rational solutions of linear differential and difference equations with polynomial coefficients, *USSR Computational Mathematics and Mathematical Physics* **29**:11 (1989) 1611–1620, transl. from: Zh. vychisl. Mat. i mat. Fiz.

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

[3] S. A. Abramov, M. Bronstein and M. Petkovšek, On polynomial solutions of linear operator equations, in: A. Levelt, ed., *Symbolic and algebraic computation*, Proceedings of ISSAC'95, Montreal, Canada (ACM Press, New York, 1995) 290–296.

[4] S. A. Abramov and E. V. Zima, A universal program to uncouple linear systems, in: *Proceedings of CMCP'96* **7**, Proceedings of the International Conference on Computational Modeling and Computing in Physics, Dubna, Russia, Sept. 16–21, 1996, (1996) 16–26.

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

[6] M. A. Barkatou, 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] M. A. Barkatou, private communication, (1997).

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

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

[10] J. E. Björk, *Rings of Differential Operators* (North Holland, Amsterdam, 1979).

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

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

[13] F. Chyzak and B. Salvy, Non-commutative elimination in Ore algebras proves multivariate holonomic identities, *Journal of Symbolic Computation* **26**:2 (1998) 187–227.

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

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

[16] J. C. Faugère, P. Gianni, D. Lazard and T. Mora, Efficient computation of zero-dimensional Gröbner bases by change of ordering, *Journal of Symbolic Computation* **16** (1993) 329–344.

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

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

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

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

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

[22] L. Lipshitz, D-Finite Power Series, *Journal of Algebra* **122** (1989) 353–373.

[23] P. Paule and A. Riese, A Mathematica q-analogue of Zeilberger's algorithm based on an algebraically motivated approach to q-hypergeometric telescoping, in: *Special functions, q-series and related topics*, Toronto, Ontario, 1995 (Fields Inst. Commun. **14**, Amer. Math. Soc., Providence, 1997) 179–210.

[24] M. Petkovšek, H. Wilf and D. Zeilberger, $A{=}B$ (A. K. Peters, Ltd., Wellesley, Massachussets, 1996) ISBN 1-56881-063-6.

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

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

[27] N. Takayama, An approach to the zero recognition problem by Buchberger algorithm, *Journal of Symbolic Computation* **14** (1992) 265–282.

[28] K. Wegschaider, Computer generated proofs of binomial multi-sum identities, diploma dissertation, RISC, Univ. Linz (1997). Available at the URL `http://www.risc.uni-linz.ac.at/research/combinat/risc/`

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

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

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

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

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

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