SYMBOLIC SUMMATION — SOME RECENT DEVELOPMENTS

PETER PAULE
Research Institute for Symbolic Computation (RISC), Johannes-Kepler-Universität
A-4040 Linz, Austria
email: Peter.Paule@risc.uni-linz.ac.at

and

VOLKER STREHL
IMMD-Informatik I, Friedrich-Alexander-Universität
D-91058 Erlangen, Germany
email: strehl@informatik.uni-erlangen.de

ABSTRACT

In recent years, the problem of symbolic summation has received much attention due to the exciting applications of Zeilberger's method for definite hypergeometric summation. This lead to renewed interest in the central part of the algorithmic machinery, Gosper's "classical" method for indefinite hypergeometric summation. We review some of the recent (partly unpublished as yet) work done in this field, with a particular emphasis on the unifying and guiding rôle of normal forms for polynomials and rational functions, especially adapted to summation algorithms.

1. Introduction

The algorithmic problem of symbolic summation can be most easily introduced by presenting it as the discrete analogue of the well-known problem of "symbolic integration" or "integration in finite terms". Without being too formal, we can simply say that the rôle of the differential operator in the latter problem is taken over by the difference operator Δ:

$$(\Delta\, g)\,(x) := g(x+1) - g(x) \quad ,$$

where g is a function belonging to some appropriate domain. It is also convenient to introduce the shift operator E (in the variable x):

$$(E\, g)\,(x) := g(x+1) \quad ,$$

so that $\Delta = E - Id$. Positive and negative powers of E denote the corresponding shifts in x.

The problem of *indefinite* summation essentially asks for solving the first-order difference equation

$$\Delta\, g = f \quad ,$$

where f belongs to some "nice" domain of functions,* and g is searched for in the same class or some suitable extension of it. For such a g, by telescoping,

$$\sum_{x=a}^{b} f(x) = g(b+1) - g(a) \quad , \quad \text{if } b-a \in \mathbf{N},$$

i.e. the "antidifference" g of f provides an expression for sums involving f. It is usually no serious restriction if we consider functions f, g etc. on integer arguments only, and thus speak of a summation problem with respect to "sequences" instead of "functions".

It must be admitted that the summation problem, despite its deep roots in classical difference calculus, has never attracted the same attention as the related, and very prestigious, problem of symbolic integration. A first survey of summation methods in the context of computer algebra has been given by Lafon[16]. The main themes treated there are

- summation of polynomials and rational functions;
- hypergeometric summation and Gosper's algorithm;
- summation using extensions of function domains.

In all cases, only *indefinite* summation was treated. Thus it may come as a surprise that most of the present interest in these questions was stimulated by the ingenious use Zeilberger made of Gosper's method for *indefinite* hypergeometric summation for the purpose of *definite* hypergeometric summation. Thus "Summation" has become a very active field of research, and what we present below is a glance into some of the most recent work under the unifying perspective of *normal forms*, motivated and stimulated by Gosper's and Zeilberger's algorithms. We will not be complete, and, in particular, we did not intend to write an "update" of Lafon's survey. Especially, we do not touch the third theme, "summation by extensions" at all, an area which is marked by the outstanding work of Karr[13,14]. On the other hand, we do come back to the interesting part of the first theme, the summation rational functions, for two good reasons.

Before giving a short outline of the sections to follow, let us mention that Graham et al.[10] give a gentle introduction into summation with many examples. In particular, Gosper's algorithm is treated in detail, and the applications made by Zeilberger and Wilf are discussed and illustrated.

Here we first briefly recall Gosper's algorithm and fix the notation for later reference. In Sec. 3 we present the main ideas and concepts of Zeilberger's method for definite hypergeometric sums. In the fourth section we look back to the most elementary case of hypergeometric summation, i.e. the summation of rational

---

*For computer algebra this means, in particular, that functions to be dealt with must be specifiable with a finite amount of information, as is the case for "polynomials", "rational function", or "hypergeometric terms".

functions. We motivate the concept of *greatest factorial factorization* (GFF) of polynomials, which, in the context of rational summation, is the appropriate analogue of the well-known *squarefree factorization*, used in rational integration. Looking (in Sec. 5) at GOSPER's algorithm with the GFF in mind, the "mystery" of this method disappears and, more importantly, a certain normal form for rational functions, which we call the GOSPER-PETKOVŠEK form, appears naturally. These strongly normalizing representations, GFF and GOSPER-PETKOVŠEK form, are then reliable guides for formulating the analogues of GOSPER's and ZEILBERGER's method in the world of q-hypergeometric series. We follow this track in Sec. 6 up to the point where the analogies become clearly visible, and demonstrate the possibility of effective implementation by examples. To conclude, we come back to the rational summation problem and present an "optimal" summation algorithm, which is essentially based on the GOSPER-PETKOVŠEK form for rational functions.

We conclude this introduction with a technical remark: throughout this article $\mathcal{K}$ denotes a field of characteristic 0. Polynomials $A \in \mathcal{K}[x]$ are usually assumed to be monic (unless stated otherwise) — this is no restriction for the problems under consideration.

2. GOSPER's algorithm for indefinite hypergeometric summation

In order to prepare the stage for hypergeometric summation, let us briefly recall the classical notion of hypergeometric series (or "functions"), cf. GRAHAM ET AL.[10]: these are infinite series

$$ {}_pF_q\left[\begin{matrix} a_1\, a_2\, \ldots\, a_p \\ b_1\, b_2\, \ldots\, b_q \end{matrix}; z\right] = \sum_{k=0}^{\infty} \frac{a_1^{\overline{k}}\, a_2^{\overline{k}}\, \ldots\, a_p^{\overline{k}}}{b_1^{\overline{k}}\, b_2^{\overline{k}}\, \ldots\, b_q^{\overline{k}}} \cdot \frac{z^k}{k!} \quad , \tag{1} $$

where the a_i, b_j are (complex) parameters, z is a (complex) variable, and where, for $k \geq 0$,

$$ a^{\overline{k}} := a \cdot (a+1) \cdot \ldots \cdot (a+k-1) $$

denotes the *rising factorial* of length k. Most of the classically known "special functions" (exponential, logarithmic, trigonometric functions, orthogonal polynomials, Bessel functions etc.) can actually be written as hypergeometric series by an appropriate choice of parameters. As a rule, binomial sums, such as compiled by GOULD[9], are nothing but terminating hypergeometric series, which is, loosely speaking, the case if one of the numerator parameters in Eq. (1) is a negative integer. These sums, in particular, belong to the objects to which ZEILBERGER's method can be applied — see the references given in Sec. 3 for numerous examples.

The particular feature of the summation in (1) is the fact that, if we write f_k for the k-th term of the series, then the quotient of successive terms

$$ \frac{f_{k+1}}{f_k} = \frac{(a_1+k)(a_2+k)\ldots(a_p+k)}{(b_1+k)(b_2+k)\ldots(b_q+k)} \cdot \frac{z}{(1+k)} $$

is a rational function of k. Conversely, up to normalization, any rational function of k can be written in this form. This leads us to the notion of a hypergeometric sequence.

A sequence $(f_k)_{k\geq 0}$ of elements in $\mathcal{K}$ is *hypergeometric* if there exist polynomials $A, B \in \mathcal{K}[x]$ such that

$$\frac{f_{k+1}}{f_k} = \frac{A(k)}{B(k)} \quad (k \geq 0) \ .$$

The rational function $A/B \in \mathcal{K}(x)$ is called a *rational representation* of $(f_k)_{k\geq 0}$. The problem of *indefinite hypergeometric summation* asks the following:

> *given* a hypergeometric sequence $(f_k)_{k\geq 0}$ over $\mathcal{K}$, e.g. specified by its hypergeometric representation $A/B \in \mathcal{K}(x)$,
> *decide* whether there exists a hypergeometric sequence $(g_k)_{k\geq 0}$ such that
>
> $$g_{k+1} - g_k = f_k \quad (k \geq 0) \tag{2}$$
>
> and if this is the case, *determine* its rational representation.

As mentioned before, a solution $(g_k)_{k\geq 0}$ of the first-order difference equation (2) leads, by telescoping, to

$$\sum_{k=a}^{b-1} f_k = g_b - g_a \quad (0 \leq a < b) \ ,$$

which justifies to speak about an indefinite summation problem.

The following classical algorithm, presented by Gosper[8] in 1978, provides a complete solution to the indefinite hypergeometric summation problem:

> Let $r \in \mathcal{K}(x)$ be a rational function representation of some hypergeometric sequence $(f_k)_{k\geq 0}$. Then
>
> 1. *determine* polynomials $P, Q, R \in \mathcal{K}[x]$ such that
>
> $$r = \frac{E\,P}{P} \cdot \frac{Q}{E\,R} \ , \tag{3}$$
>
> where
>
> $$\gcd(Q, E^j R) = 1 \quad \text{for all} \quad j \geq 1 \ .$$
>
> 2. *try to solve* the "key equation"
>
> $$P = Q \cdot E\,Y - R \cdot Y \tag{4}$$
>
> for $Y \in \mathcal{K}[x]$ (!). Now:

- If such a polynomial solution $Y \in \mathcal{K}[x]$ exists, then $(g_k)_{k\geq 0}$, where
$$g_k = \frac{Y(k)\,R(k)}{P(k)} \cdot f_k \quad (k \geq 0)\ , \tag{5}$$
is a hypergeometric solution of (2) and
$$s = \frac{Q}{R} \cdot \frac{EY}{Y} \tag{6}$$
is a rational representation of it.
- If no polynomial solution $Y \in \mathcal{K}[x]$ of (4) exists, then no hypergeometric solution $(g_k)_{k\geq 0}$ of (2) can exist.

It follows, in particular, that a hypergeometric solution $(g_k)_{k\geq 0}$ of (2), if it exists, is necessarily a rational multiple of the input sequence $(f_k)_{k\geq 0}$. As a consequence, we can note: if the input sequence $(f_k)_{k\geq 0}$ is *rational*, then a potential hypergeometric "antidifference" $(g_k)_{k\geq 0}$ must itself be rational.

The feasibility of GOSPER's algorithm, what really turns it into a decision method, depends on the fact that one can give an *a priori* degree bound on potential polynomial solutions $Y \in \mathcal{K}[x]$ of the key equation (4). This is discussed in the presentations given by GOSPER[8], LAFON[16], GRAHAM ET AL.[10], and, with particular attention paid to uniqueness questions and the behaviour on rational inputs, by LISONĚK ET AL.[17].

In Sec. 5 we will look again on GOSPER's algorithm. It turns out that the concept of *greatest factorial factorization*, introduced in Sec. 4 in the context of the rational summation problem, leads to a strengthening of the particular representation of a rational function, as contained in Eq. (3). We will show that this "normal form" for rational functions, which we call the GOSPER-PETKOVŠEK representation, makes the machinery behind the "key equation" (4) and the statements (5) and (6) better understandable.

GOSPER's algorithm is at the computational heart of ZEILBERGER's fast algorithm for definite hypergeometric summation, which is discussed in the next section. We restrict ourselves here to the presentation of two very simple illustrating examples.

– Let $f_k = (k+1)(k+1)!$ for $k \geq 0$, then
$$\frac{f_{k+1}}{f_k} = \frac{k+2}{k+1} \cdot \frac{k+2}{1} \quad .$$
Hence $P(x) = x+1, Q(x) = x+2$, and $R(x) = 1$. The constant polynomial $Y(x) = 1$ satisfies the key equation
$$x + 1 = (x+2)Y(x+1) - Y(x) \quad .$$
Consequently, according to Eq. (5), the hypergeometric solution of the difference equation $f_k = g_{k+1} - g_k$ is
$$g_k = \frac{1}{k+1} \cdot f_k = (k+1)! \quad ,$$

with rational representration

$$s = \cdot\frac{x+2}{1}\cdot\frac{1}{1} = x+2 \quad .$$

and thus, for $n \geq 0$,

$$\sum_{k=0}^{n}(k+1)(k+1)! = g_{n+1} - g_0 = (n+2)! - 1 \quad .$$

– Let now $f_k = \binom{n}{k}$ for $k \geq 0$ and $n \geq 0$ fixed. Then

$$\frac{f_{k+1}}{f_k} = \frac{n-k}{k+1} \quad .$$

Hence $P(x) = 1, Q(x) = n - x$ and $R(x) = x$. The key equation reads

$$1 = (n-x)\cdot Y(x+1) - x \cdot Y(x) \quad ,$$

and it is easily checked that this equation does not admit a polynomial solution $Y \in \mathcal{K}[x]$. Consequently, $\sum_{k=0}^{m}\binom{n}{k}$ is not hypergeometric as a function of m.

3. Gosper's algorithm and Zeilberger's summation method

It was observed by Zeilberger[35,36,38] that the indefinite summation algorithm of Gosper can be used in a non-obvious and nontrivial way for *definite* hypergeometric summation. The methodology which evolved from that approach has important applications in verifying or finding binomial identities "automatically", finding annihilating recurrence operators for hypergeometric sums etc.. Expositions of what is now known as "Zeilberger's method", together with examples and applications, have been given e.g. by Zeilberger[39], Wilf[33], Cartier[4], Koornwinder[15], Strehl[31,32], Graham et al.[10]. A generalization of that method to multisum identities, has been presented by Wilf and Zeilberger[34]. This fundamental article is also an excellent source for examples and further references. Maple versions of Zeilberger's algorithm have been written and published by Zeilberger[37] and Koornwinder[15]. A Mathematica implementation by Paule and Schorn[22] is available by email request from the first author.

The following short description of the basic mechanism of Zeilberger's method is adapted from Paule[21], where the q-analogue (see section 6 of this paper) is spelled out.

Let $f := (f_{n,k})$ be a double-indexed sequence with values in the ground-field $\mathcal{K}$. We shall consider only sequences where n runs through the nonnegative integers, whereas the second parameter k may run through all integers.

The sequence f is called *hypergeometric* in both parameters if both quotients

$$\frac{f_{n+1,k}}{f_{n,k}} \quad \text{and} \quad \frac{f_{n,k+1}}{f_{n,k}}$$

are rational functions in n and k over the field $\mathcal{K}$ (for all n and k for which the quotients are well-defined).

For instance, for the binomial coefficient sequence $f_{n,k} = \binom{n}{k}$ we have

$$\binom{n+1}{k+1} / \binom{n}{k+1} = \frac{n+1}{n-k} \quad \text{and} \quad \binom{n+1}{k+1} / \binom{n+1}{k} = \frac{n-k+1}{k+1} \quad . \tag{7}$$

We assume that the hypergeometric sequence f has finite support with respect to k, i.e. for each fixed n one has $f_{n,k} = 0$ for all k outside a finite integer interval I_n. In this case, the sum $S_n := \sum_k f_{n,k} = \sum_{k \in I_n} f_{n,k}$ is finite, and we may use the convention that the summation parameter k runs through all the integers, in case the summation range is not specified explicitly.

It can then be shown that, under mild conditions, the sequence $S := (S_n)_{n \geq 0}$ is *holonomic*, i.e. it satisfies a homogeneous linear recursion with rational function (or polynomial) coefficients:

$$C_0(n)S_n + C_1(n)S_{n-1} + \cdots + C_d(n)S_{n-d} = 0 \tag{8}$$

for $n \geq d$, with $C_i \in \mathcal{K}(x)$, $C_0 \neq 0$. If we denote by N the shift operator w.r.t. n, then we can state that S is uniquely specified by such an annihilating operator

$$A_S(N) := \sum_{l=0}^{d} C_l(n) N^{-l}, \tag{9}$$

together with the d initial values $S_0, \ldots, S_{d-1}$ (provided that $C_0(n) \neq 0$ for $n \geq d$).

What is the rôle of GOSPER's algorithm in *constructing* such an operator $A_S(N)$? Note that from

$$N^{-1} f_{n,k} = \frac{f_{n-1,k}}{f_{n,k}} f_{n,k}, \quad N^{-2} f_{n,k} = \frac{f_{n-2,k}}{f_{n-1,k}} \frac{f_{n-1,k}}{f_{n,k}} f_{n,k}, \text{ a.s.o.}$$

it follows that $A_S(N)\, f_{n,k}$ is a rational multiple of $f_{n,k}$, i.e.

$$A_S(N)\, f_{n,k} = r(n,k)\, f_{n,k} \quad , \tag{10}$$

where $r(n,k)$ is a rational function in n, k over $\mathcal{K}$. Hence, for each fixed n, the sequence $(A_S(N)\, f_{n,k})_{k \geq 0}$ is hypergeometric (w.r.t. k). We can use GOSPER's algorithm in order to solve the hypergeometric summation problem

$$A_S(N)\, f_{n,k} = g_{n,k} - g_{n,k-1} \quad , \tag{11}$$

where $g_{n,k}$ is hypergeometric in n, k, and indeed, if it exists, $g_{n,k}$ will be rational multiple of $f_{n,k}$:

$$g_{n,k} = cert_S(n,k)\, f_{n,k} \quad . \tag{12}$$

In this situation, it follows by telescoping over Eq. (11) and using the finite support property of f that the operator $A_S(N)$ has the desired property:

$$A_S(N)\, S_, = A_S(N) \sum_k f_{n,k} = \sum_k g_{n,k} - g_{n,k-1} = 0 \quad .$$

Now the following remarks are crucial for the understanding of the method:

- Up to now we have argued as if $A_S(N)$ were known — this is not the case, of course. But if start by writing $A_S(N)$ with undetermined coefficients $C_j \in \mathcal{K}(x)$, then we can use the second phase of GOSPER's algorithm (solving the "key equation") to solve for these unknowns as well, since they enter linearly into this equation. This approach will work if the assumed degree d of $A_S(N)$ is sufficiently high.

  For the class of "proper" hypergeometric inputs f, for which the algorithm is known to work, one has *a priori* bounds for the order d. In practice, one tries orders $d = 1, 2, 3, \ldots$ until GOSPER's algorithms admits a solution.

- If successful, ZEILBERGER's algorithm will return, as a result from applying GOSPER's algorithm, a difference operator $A_S(N)$, as above, and a (rational) "certificate function" $cert_S(n,k)$ such that Eq. (11) and Eq. (12) hold. This fact can be checked *independently*, because this amounts to checking the *rational function identity*

  $$r(n,k) = cert_S(n,k) - cert_S(n,k-1)\frac{f_{n,k-1}}{f_{n,k}} \quad , \tag{13}$$

  where $r(n,k)$, rational in n and k, comes from Eq. (10). This justifies the notion "certificate" for the rational function $cert_S(n,k)$.

As a very simple illustrating example we use Zeilberger's algorithm in order to obtain a computer proof for the binomial theorem.

Let $f_{n,k} = \binom{n}{k} x^k$, $(x \in \mathcal{K})$. Then the output of Zeilberger's algorithm is

$$A_S(N) = N^0 - (1+x)N^{-1} \quad , \quad cert_S(n,k) = (k-n)x/n \quad .$$

The first part of the result asserts that $S_n = \sum_k f_{n,k}$ satisfies the same (first-order) recursion as $(1+x)^n$. To actually *verify* this claim, we compute

$$A_S(N) f_{n,k} = \binom{n}{k} x^k - (1+x)\binom{n-1}{k} x^k = \left(1 - (1+x)\frac{n-k}{k}\right) f_{n,k} \quad ,$$

i.e.

$$r(n,k) = \frac{(x+1)k - xn}{n} \quad ,$$

and equate it with

$$cert_S(n,k) - cert_S(n,k-1)\frac{f_{n,k-1}}{f_{n,k}} = \frac{(k-n)x}{n} - \frac{(k-n-1)x}{n} \cdot \frac{k}{(n-k+1)x} \ .$$

This simple example does not at all demonstrate the full power of the method. For instance, about 80% of the more than 500 binomial coefficient identities listed by GOULD[9] can be proved automatically this way. Numerous further interesting and nontrivial applications of ZEILBERGER's algorithm can be found in the references given above.

We conclude this section by an example in which we use the PAULE/SCHORN[22] implementation of ZEILBERGER's algorithm. We shall prove

$$\sum_{k=0}^{n} \binom{n}{k}^2 \binom{2k}{n} \binom{2k}{k+a} \binom{2k}{k}^{-1} = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n}{k+a} \quad , \tag{14}$$

an identity which is a generalization of a key identity (case $a = 0$) used in a study by STREHL[31] on a problem arising from number theory.

Invoking the program one gets:

```
zeder!8> math
Mathematica 2.1 for HP Apollo Domain/OS
Copyright 1988-92 Wolfram Research, Inc.

In[1]:= <<zb_alg.m

Out[1]= Peter Paule and Markus Schorn's implementation loaded.

In[2]:= Zb[Binomial[n,k]^2 Binomial[2k,n] Binomial[2k,k+a] *
        Binomial[2k,k]^(-1), k,n,3]
                      2        2
Out[2]= {-8 (1 + n)  (2 + n)  SUM[n] -

                 2           2               2
>       (2 + n)  (88 - 4 a  + 69 n + 15 n ) SUM[1 + n] -

                                        2           2
>       (4 + n) (5 + 2 n) (19 - a  + 15 n + 3 n ) SUM[2 + n] +

                     2            2
>       (3 - a + n)  (3 + a + n)  SUM[3 + n] == 0}
```

```
In[3]:= Zb[Binomial[n,k]^2 Binomial[n,k+a], k,n, 3]
                     2         2
Out[3]= {8 (1 + n)  (2 + n)  SUM[n] +

                 2          2                  2
>       (2 + n)  (88 - 4 a  + 69 n + 15 n ) SUM[1 + n] +

                                        2            2
>       (4 + n) (5 + 2 n) (19 - a  + 15 n + 3 n ) SUM[2 + n] -

                     2            2
>       (3 - a + n)  (3 + a + n)  SUM[3 + n] == 0}
```

This shows that both sides of Eq. (14) satisfy the same third-order recurrence in n, thus the validity of Eq. (14) follows from checking the the first three initial values of both sides.

4. Summation of rational functions

Let us start with a general *proviso* for this section and the last one, both dealing with the summation problem for rational functions: Since indefinite integration and summation of *polynomials* can be considered as trivial tasks, we will assume that all rational functions appearing in these sections are *proper*, i.e. degree of the numerator polynomial < degree of the denominator polynomial.

The problem of indefinite *integration* of rational functions asks for the following:

> *given* a rational function $r = A/B \in \mathcal{K}(x)$,
> *determine* rational functions $s = C/D \in \mathcal{K}(x)$ and $t = F/G \in \mathcal{K}(x)$ such that
>
> $$r = \frac{d}{dx} s + t \quad \text{ i.e. } \int \frac{A}{B}\, dx = \frac{C}{D} + \int \frac{F}{G}\, dx \quad , \tag{15}$$
>
> where the "logarithmic" part t is as "small" as possible, so that no further rational parts of the integral can be "extracted" from it.

A solution, in this sense, is obtained once polynomials C, D, F, G have been found such that (15) holds and such that the denominator polynomial G of the logarithmic part t is a product of pairwise relatively prime squarefree polynomials. A classical algorithm, due to HERMITE[11] (see also GEDDES ET AL.[7] or SUBRAMANIAM AND MALM[30]), shows how this goal can always be reached iteratively via partial fraction decomposition and partial integration. An alternative method, due to HOROWITZ[12] (see also GEDDES ET AL.[7]), proceeds as follows:

Let $B = B_1^1 \cdot B_2^2 \cdot \ldots \cdot B_k^k$ denote the squarefree factorization of the denominator polynomial B. Then put

$$\begin{aligned} D &:= \gcd(B, \frac{d}{dx}B) = B_1^0 \cdot B_2^1 \cdot \ldots \cdot B_k^{k-1} \quad , \\ G &:= B/D = B_1 \cdot B_2 \cdot \ldots \cdot B_k \quad , \end{aligned} \tag{16}$$

and obtain polynomials C, F, such that Eq. (15) holds, by the method of undetermined coefficients, applied to the polynomial equation (the "HERMITE-OSTROGRADSKI formula")

$$A = G \cdot \left(\frac{d}{dx}C\right) - \left(\frac{d}{dx}D\right)\frac{G}{D} \cdot C + D \cdot F \quad . \tag{17}$$

Let us remark the following:

— Eq. (17) is indeed a polynomial equation, since D divides $G \cdot \frac{d}{dx}D$, which can be seen from

$$G \cdot \frac{d}{dx}D = \frac{d}{dx}B - \frac{d}{dx}G \cdot D \quad .$$

— The feasibility of the approach with undetermined coefficients follows from the *a priori* degree bounds $\deg C < \deg D$ and $\deg F < \deg G$.

— The denominator polynomial G is indeed a product of pairwise relatively prime squarefree polynomials.

— The squarefree factorization, which plays the essential rôle in this approach, can be computed by gcd-calculations only, i.e. no factoring is required (see GEDDES ET AL.[7]).

We now turn to the parallel problem of indefinite *summation* of rational functions. Here our problem is

given a rational function $r = A/B \in \mathcal{K}(x)$,
determine rational functions $s = C/D \in \mathcal{K}(x)$ and $t = F/G \in \mathcal{K}(x)$ such that

$$r = \Delta\, s + t \quad \text{i.e.}^\dagger \sum \frac{A}{B}\,\delta x = \frac{C}{D} + \sum \frac{F}{G}\,\delta x \tag{18}$$

where the "transcendental" part t is as "small" as possible, so that no further rational parts of the sum can be "extracted" from it.

In this situation, the notion of "smallness" can be made precise by requiring that the degree of the denominator polynomial G should be as small as possible for all polynomials C, D, F, G satisfying (18). It is not too difficult to see (cf. ABRAMOV[1],

†We use the suggestive δ-notation proposed in GRAHAM ET AL.[10].

Paule[19], Pirastu[24]) that an equivalent condition can be given in terms of the *dispersion* of a (nonconstant) polynomial: this is defined as the maximum integer distance occurring between any two roots of the polynomial (in the algebraic closure of $\mathcal{K}$) — see the example below. A proper solution of the rational summation problem for $A/B \in \mathcal{K}[x]$ is then given by polynomials C, D, F, G such that (18) holds, where the denominator polynomial G has zero dispersion. Note that this condition alone is too weak to ensure uniqueness in a strong sense. For a study of the uniqueness question see Paule[19], Pirastu[24,26], and Pirastu and Strehl[27].

The two standard approaches for the rational integration problem do have their counterparts in the summation case. A Hermite-type approach was presented by Moenck[18,16], but the theoretical base for the version he gives is not sufficiently precise, and this has lead to at least one defective implementation in a computer algebra system. See Pirastu[24,25] for details and for a correct adaptation of Hermite's idea to the summation situation.

For a summation analog of the Horowitz approach, as outlined above, one needs first an analog of the squarefree decomposition of a polynomial. The corresponding notion, the *greatest factorial factorization* (GFF) of a polynomial, has been introduced only recently and extensively studied by Paule[19]. To discuss it, we need a bit of notation.

If $G \in \mathcal{K}[x]$ is a polynomial, then, for $k \geq 0$, the k-th *(falling) factorial* of G is defined by

$$[G]^{\underline{k}} := \begin{cases} G \cdot \left(E^{-1}G\right) \cdot \left(E^{-2}G\right) \cdot \ldots \cdot \left(E^{-k+1}G\right) & \text{if } k > 0 \quad , \\ 1 & \text{if } k = 0 \quad . \end{cases}$$

Using this notion, one can show that every non-constant polynomial $G \in \mathcal{K}[x]$ has a representation of the form

$$G = [G_1]^{\underline{1}} \cdot [G_2]^{\underline{2}} \cdot \ldots \cdot [G_k]^{\underline{k}} \quad , \tag{19}$$

where $G_1, \ldots, G_k \in \mathcal{K}[x]$ are polynomials such that G_k is non-constant, and that such a representation is *unique* if the following condition is satisfied:

$$\forall\, 1 \leq i, j \leq k \;:\; \max(j-i, 0) < h \leq j \quad \Rightarrow \quad \gcd(E^h G_i, G_j) = 1 \quad . \tag{20}$$

It is this particular situation which gives rise to the GFF. The parameter k in this representation is named *factorial exponent* of G.

The "geometrical" meaning behind this uniqueness condition is the following: we call an *interval* of G any factorial polynomial $[H]^{\underline{h}}$ which divides G. Naturally, a polynomial can, in general, be written as a product of intervals in many ways. A *unique normal form* can be achieved if it is required that no two such factors "overlap" or "touch".

Let us illustrate this concept by an example (over $\mathcal{K} = \mathbf{Q}$):

$$\begin{aligned} G \;=\; & x^{18} + 11\,x^{17} + 30\,x^{16} - 38\,x^{15} - 148\,x^{14} + 316\,x^{13} \\ & + 378\,x^{12} - 1250\,x^{11} + 415\,x^{10} + 1365\,x^{9} - 1736\,x^{8} + 584\,x^{7} \\ & + 500\,x^{6} - 1692\,x^{5} + 1328\,x^{4} + 704\,x^{3} - 768\,x^{2} \end{aligned}$$

¿From its factored form

$$G = (x^2 - 2\,x + 2)^2\,(x^2 + 1)\,(x - 1)^3\,x^2\,(x + 1)^3\,(x + 3)\,(x + 4)^3$$

one can easily read off the squarefree factorization

$$\begin{aligned} G &= \big((x^2+1)(x+3)\big)^1 \big((x^2 - 2\,x + 2)(x)\big)^2 \big((x-1)(x+1)(x+4)\big)^3 \\ &= \big(x^3 + 3\,x^2 + x + 3\big)^1 \big(x^3 - 2\,x^2 + 2\,x\big)^2 \big(x^3 + 4\,x^2 - x - 4\big)^3 \quad , \end{aligned}$$

i.e.

$$G_1 = x^3 + 3\,x^2 + x + 3\ ,\ G_2 = x^3 - 2\,x^2 + 2\,x\ ,\ G_3 = x^3 + 4\,x^2 - x - 4 \quad .$$

For the GFF $G = [G_1]^{\underline{1}} \cdot [G_2]^{\underline{2}} \cdot \ldots \cdot [G_k]^{\underline{k}}$ we find, again by inspecting the factored form,

$$G = [(x+4)^2(x+1)\,(x-1)\,(x^2 - 2\,x + 2)]^{\underline{1}}\ [(x+4)\,(x^2+1)]^{\underline{2}}\ [(x+1)^2]^{\underline{3}}\ ,$$

so that

$$\begin{aligned} G_1 &= (x+4)^2\,(x+1)\,(x-1)\,(x^2 - 2\,x + 2) \\ &= x^6 + 6\,x^5 + x^4 - 22\,x^3 + 30\,x^2 + 16\,x - 32 \\ G_2 &= (x+4)\,(x^2+1) = x^3 + 4\,x^2 + x + 4 \\ G_3 &= (x+1)^2 = x^2 + 2\,x + 1 \quad . \end{aligned}$$

Here the factorial exponent is $k = 3$, whereas the dispersion of G has value 5, since -4 and 1 are zeros of G, and no other zeros of G have integer distance > 5.

As with the squarefree factorization, it is important to note that the GFF of a polynomial can be calculated by gcd-computations alone, i.e. without factoring. To demonstrate this, we outline one possible method of computation, which can easily be extended to a proof of the uniqueness property. In order to avoid heavy notation, the meet-sign $\wedge$ will occasionally be used to denote the gcd-operation for polynomials.

Let $G \in \mathcal{K}[x]$ be a non-constant polynomial, and let $k \geq 1$, be the *factorial exponent* of G, i.e. the maximal positive integer j such that

$$G \wedge E\,G \wedge \ldots \wedge E^{j-1}G \neq 1 \quad ,$$

and define

$$G_k := G \wedge E\,G \wedge \ldots \wedge E^{k-1}G \quad .$$

By definition,

$$G_k \neq 1 \quad \text{and} \quad G_k \wedge E^k G = 1 \quad ,$$

and thus

$$G_k \wedge E^j G_k = 1 \quad (0 < j \leq k) \quad .$$

It follows that the polynomials

$$G_k\,,\ E^{-1}G_k\,,\ \ldots\,,\ E^{-k+1}G_k$$

are pairwise relatively prime, and each of them divides G. Hence

$$[G_k]^{\underline{k}} = G_k \cdot \left(E^{-1}G_k\right) \cdot \ldots \cdot \left(E^{-k+1}G_k\right) \ \mid G \quad .$$

Now put $H := G/[G_k]^{\underline{k}}$, then either $H = 1$, or else H is a non-constant polynomial with factorial exponent $< k$, so that the same procedure just outlined may be applied iteratively to H.

We note that the GFF has properties similar to the squarefree factorization, in particular, if (19) is the GFF of $G \in \mathcal{K}[x]$, then

$$\gcd(G, \Delta G) = \gcd(G, E\,G) = G \wedge E\,G = [G_1]^{\underline{0}} \cdot [G_2]^{\underline{1}} \cdot \ldots \cdot [G_k]^{\underline{k-1}} \quad ,$$

so that

$$\frac{G}{E^{-1}\gcd(G, \Delta G)} = G_1 \cdot G_2 \cdot \ldots \cdot G_k \quad .$$

Using the concept of GFF, a summation analog of Horowitz' method can be formulated. Some care about relative primeness has to be taken, however, the analogy is not as straightforward as one might wish.

Let $A/B \in \mathcal{K}(x)$ be a proper rational function. It will now be necessary to "blow up" the fraction A/B by multiplying numerator and denominator by the same polynomial. The new fraction, $\widetilde{A}/\widetilde{B}$ say, which denotes the same element of $\mathcal{K}(x)$ as A/B, has to be "shift-saturated", in a sense to be explained below.

Under this condition, one can indeed show that a solution of the rational summation problem (18) can be obtained from

$$\frac{A}{B} = \frac{\widetilde{A}}{\widetilde{B}} = \Delta\,\frac{C}{E^{-1}\Big([\widetilde{B}_1]^{\underline{0}}\,[\widetilde{B}_2]^{\underline{1}}\,\cdots\,[\widetilde{B}_k]^{\underline{k-1}}\Big)} + \frac{F}{\widetilde{B}_1\,\widetilde{B}_2\,\cdots\,\widetilde{B}_k} \quad .$$

Here

$$D = E^{-1}\Big([\widetilde{B}_1]^{\underline{0}}\,[\widetilde{B}_2]^{\underline{1}}\,\cdots\,[\widetilde{B}_k]^{\underline{k-1}}\Big) = \gcd(\widetilde{B}, E^{-1}\widetilde{B})$$

and

$$G = \widetilde{B}_1 \cdot \widetilde{B}_2 \cdot \ldots \cdot \widetilde{B}_k = \frac{\widetilde{B}}{D}$$

come from the GFF of $\widetilde{B}$:

$$\widetilde{B} = [\widetilde{B}_1]^{\underline{1}} \cdot [\widetilde{B}_2]^{\underline{2}} \cdot \ldots \cdot [\widetilde{B}_k]^{\underline{k}} \quad ,$$

and $C, F \in \mathcal{K}[x]$ are polynomials with $\deg(C) < \deg(D)$ and $\deg(F) < \deg(G)$.

Again, the solution polynomials C, F can be obtained via undetermined coefficients from the polynomial equation

$$\tilde{A} = \frac{\tilde{B}}{(E\,D)} \cdot (E\,C) - G \cdot C + D \cdot F$$

using the degree bounds. Note, in particular, that the "blowing up" procedure mentioned above has to be performed in such a way that the polynomial $G = \tilde{B}_1\,\tilde{B}_2\,\cdots\,\tilde{B}_k$ has dispersion 0, i.e. $\gcd(G, E^h G) = 1$ for all $h \neq 0$. A method for doing this has been presented by Paule[19]. His way of solving the rational summation problem *à la* Horowitz as an application of the GFF has the disadvantage that the "blow up" may lead to degree bounds which are much bigger that necessary. In Section 7, we will present an alternative approach which avoids this inefficiency — indeed, the method presented there is "optimal" with respect to the parameter "degree bound".

To conclude this section, let us mention that another iterative method, different from Moenck's approach *à la* Hermite and not based on partial fraction decomposition, has been proposed by Abramov[2] — see also Pirastu[24,25] for a discussion and an improvement of this approach.

5. Greatest factorial factorization and Gosper's algorithm

With the concept of GFF in mind, we now look back at Gosper's algorithm from Sec. 2, in particular at the crucial representation (3) of rational functions. Following Paule[20], let us consider the homogeneous first-order difference equation

$$B \cdot E\,Y - A \cdot Y = 0 \quad , \tag{21}$$

where $A, B \in \mathcal{K}[x]$ with $\gcd(A, B) = 1$. Let $Y \in \mathcal{K}[x]$ be a non-constant *polynomial* solution of Eq. (21), and let

$$Y = [Y_1]^{\underline{1}} \cdot [Y_2]^{\underline{2}} \cdot \ldots \cdot [Y_k]^{\underline{k}}$$

denote its GFF. Dividing Eq. (21) by

$$\gcd(Y, E\,Y) = [Y_2]^{\underline{1}} \cdot [Y_3]^{\underline{2}} \cdot \ldots \cdot [Y_k]^{\underline{k-1}}$$

gives

$$B \cdot (E\,Y_1) \cdot \ldots \cdot (E\,Y_k) = A \cdot Y_1 \left(E^{-1} Y_2\right) \cdot \ldots \cdot \left(E^{-k+1} Y_k\right) \quad ,$$

and from the coprimeness condition (20) we get

$$A = (E\,Y_1) \cdot (E\,Y_2) \cdot \ldots \cdot (E\,Y_k) \quad , \tag{22}$$

$$B = Y_1 \cdot \left(E^{-1} Y_2\right) \cdot \ldots \cdot \left(E^{-k+1} Y_k\right) \quad . \tag{23}$$

This gives us a way of computing the solution Y by constructing iteratively its GFF:

— put $A^{(0)} = A$ and $B^{(0)} = B$ and let k denote the maximum integer i such that $\gcd(A, E^i B) \neq 1$;

— for j from 1 to k compute

$$\begin{aligned} Y_j &:= \gcd(E^{-1}A^{(j-1)}, B^{(j-1)}) \\ A^{(j)} &:= A^{(j-1)}/E\,Y_j \\ B^{(j)} &:= B^{(j-1)}/E^{-j+1}Y_j \end{aligned}$$

— if a polynomial solution $Y \in \mathcal{K}[x]$ of Eq. (21) exists, then one ends up with $(A^{(k)}, B^{(k)}) = (1,1)$ and $Y = [Y_1]^{\underline{1}} \cdot \ldots \cdot [Y_k]^{\underline{k}}$ is this solution;

— if no such solution exists, then one ends up with $(A^{(k)}, B^{(k)}) \neq (1,1)$, which means that

$$\begin{aligned} A &= (E\,Y_1) \cdot \ldots \cdot (E\,Y_k) \cdot A^{(k)} \quad , \\ B &= Y_1 \cdot \ldots \cdot \left(E^{-k+1}Y_k\right) \cdot B^{(k)} \quad , \end{aligned}$$

where $A^{(k)}, B^{(k)} \in \mathcal{K}[x]$ are polynomials such that

$$\gcd(A^{(k)}, E^j B^{(k)}) = 1 \quad \text{for all } j \geq 0 \quad . \tag{24}$$

In the "failure" situation just described we may nevertheless conclude that $Y = [Y_1]^{\underline{1}} \cdot \ldots \cdot [Y_k]^{\underline{k}}$ is the solution of the equation

$$\frac{B}{B^{(k)}} \cdot E\,Y - \frac{A}{A^{(k)}} \cdot Y = 0 \quad ,$$

or equivalently,

$$\frac{A}{B} = \frac{E\,Y}{Y} \cdot \frac{A^{(k)}}{B^{(k)}} \quad ,$$

which, together with condition (24), brings us back to the situation of Eq. (3). But even more can be said: it turns out that the output of the algorithm described above satisfies

$$\gcd(A^{(k)}, Y) = 1 = \gcd(B^{(k)}, E\,Y) \quad .$$

In his work on hypergeometric solutions of linear difference equations with polynomial coefficients, Petkovšek[23] showed that under these additional assumptions the Gosper representation (3) is *unique*. We state this result, the Gosper-Petkovšek representation for rational functions, for later reference:

> For any (nonzero) rational function $r \in \mathcal{K}(x)$ there exist unique monic polynomials $P, Q, R \in \mathcal{K}[x]$ and a unique constant‡ $c \in \mathcal{K}$ such that
>
> $$r = c \cdot \frac{E\,P}{P} \cdot \frac{Q}{E\,R} \quad , \tag{25}$$
>
> with $\gcd(P,Q) = 1 = \gcd(P,R)$ and $\gcd(Q, E^j R) = 1$ for all $j \geq 1$.

‡For the following, it is convenient to assume that the constant c, if $\neq 1$, is associated with the polynomial Q, so that we admit a polynomial which is not necessarily monic in this position.

The above algorithm shows how this representation can be computed, again using (essentially) only gcd-calculations. Note that the value k of the factorial exponent of Y can be obtained by a resultant computation, since this is the maximum integer value of the resultant $Res_x(A, E^i B)$, seen as a polynomial in i. As a technical remark, let us note for later use that in the situation of Eq. (25) the gdc-conditions ensure in particular that $Q \mid A$ and $(E\,R) \mid B$, where $r = A/B$ is written in reduced form.

With the GOSPER-PETKOVŠEK (in short: GP) representation in hands, GOSPER's ingenious, but somewhat mysterious, algorithm can be given the following "natural" explanation, which is due to PAULE[20].

Let $A/B \in \mathcal{K}(x)$ with $\gcd(A,B) = 1$ be the *reduced* rational representation of the hypergeometric sequence $(f_k)_{k\geq 0}$. Suppose a hypergeometric solution $(g_k)_{k\geq 0}$ of

$$g_{k+1} - g_k = f_k \quad (k \geq 0) \tag{26}$$

exists, and let $C/D \in \mathcal{K}(x)$ with $gcd(C,D) = 1$ be the reduced rational representation of $(g_k)_{k\geq 0}$. Then $(g_k)_{k\geq 0}$ is a rational multiple of the input, or written explicitly,

$$g_k = \frac{D(k)}{C(k) - D(k)} \cdot f_k \quad (k \geq 0)\ . \tag{27}$$

Conversely, using this representation as an "Ansatz" for g_k we find

$$\frac{A}{B} = \frac{E(C-D)}{C-D} \cdot \frac{C}{ED} \quad . \tag{28}$$

Hence, solving (26) is equivalent to finding relatively prime polynomials C, D satisfying (28). If this can be done, $C/D \in \mathcal{K}(x)$ is the (reduced) rational representation of the solution $(g_k)_{k\geq 0}$.

Note that $\gcd(C, C-D) = 1 = \gcd(D, C-D)$, thus the right hand side of (28) is "very close" to a GP representation. If it actually were that representation, then from writing the "true" GP representation of A/B as

$$\frac{A}{B} = \frac{EP}{P} \cdot \frac{Q}{ER} \quad , \tag{29}$$

for $P, Q, R \in \mathcal{K}[x]$, say, we could conclude that $C = Q$, $D = R$ (and $P = C - D$) by uniqueness. But in general there is no guarantee to have $\gcd(C, E^j D) = 1$ for all $j \geq 1$. As a way out, consider the GP representation for C/ED,

$$\frac{C}{ED} = \frac{E\tilde{P}}{\tilde{P}} \cdot \frac{\tilde{Q}}{E\tilde{R}} \tag{30}$$

say, for polynomials $\tilde{P}, \tilde{Q}, \tilde{R} \in \mathcal{K}[x]$, which turns the right hand side of Eq. (28) into a true GP representation:

$$\frac{E((C-D)\tilde{P})}{(C-D)\tilde{P}} \cdot \frac{\tilde{Q}}{E\tilde{R}} \quad . \tag{31}$$

Now the comparison with the GP representation (29) results in

$$\tilde{Q} = Q \quad , \quad \tilde{R} = R \quad , \tag{32}$$

and

$$P = (C - D)\tilde{P} \quad . \tag{33}$$

Eq. (33) can be rewritten as

$$P = Q \cdot E(\frac{D}{R}\tilde{P}) - R \cdot (\frac{D}{R}\tilde{P}) \quad , \tag{34}$$

which shows that $Y = \tilde{P}\, D/R$ is a solution of the "key equation" (4). Note that Y is a polynomial, since R divides D by the properties of the GP representation, applied to Eq. (30). It is now an easy matter to check that this Y actually leads to Eqs. (5) and (6).

6. q-analogues of Gosper's and Zeilberger's algorithms

In this section we present a sketch of how the concepts introduced above, greatest factorial factorization and Gosper-Petkovšek representation, extend quite naturally into the realm of q-analogues. As direct consequences, q-analogues of the Gosper and Zeilberger algorithms can be put into action. Proofs, further details, and applications will appear in a forthcoming article by Paule[20].

For this section we adjoin to the ground-field $\mathcal{K}$ the indeterminate q. In applications q might be specialized to a nonzero complex number, e.g. with $|q| < 1$ for limit considerations. We shall make use of the standard notation

$$\begin{aligned} (a;q)_n = (a)_n &= (1-a)(1-aq)\cdots(1-aq^{n-1}) \text{ for } n \geq 1, \\ (a;q)_0 = (a)_0 &= 1 \quad , \quad \text{and } 1/(q)_n = 0 \text{ for } n < 0. \end{aligned}$$

Following Gasper and Rahman[6], let us define a q-hypergeometric sequence $(f_k)_{k\geq 0}$ with m (≥ 0) numerator and n (≥ 0) denominator parameters $a_i, b_j \in \mathcal{K}(q)$, $(1 \leq i \leq m, 1 \leq j \leq n)$ as

$$f_k = \frac{(a_1)_k \ldots (a_m)_k}{(b_1)_k \ldots (b_n)_k} \cdot \frac{z^k}{(q)_k} \cdot q^{\alpha\binom{k}{2}+\beta k} \quad ,$$

with argument[§] $z \in \mathcal{K}$, and where α, β are integers.

Writing $x = q^k$, we can view the rational representation of f_{k+1}/f_k, namely

$$\frac{f_{k+1}}{f_k} = \frac{(1-a_1x)\ldots(1-a_mx)}{(1-b_1x)\ldots(1-b_nx)(1-qx)} \cdot x^\alpha \cdot q^\beta \cdot z \tag{35}$$

[§] For sake of simplicity we restrict ourselves to $z \in \mathcal{K}$. All works the same way when taking $z = z(q) \in \mathcal{K}(q)$ with $z(0) \neq 0$.

as an element of $\mathcal{K}(q, x)$.

In this section we will look at polynomials $G \in \mathcal{K}(q)[x]$, and the q-shift operator ϵ

$$(\epsilon\, G)(x) := G(qx)$$

will play an analogous rôle to that of the shift operator E. This motivates the following definitions. A polynomial $G \in \mathcal{K}(q)[x]$ is said to be *q-monic* if $G(0) = 1$. Note that this property is invariant with respect to the q-shift operator ϵ, whereas the property *monic*, defined as usual, is invariant with respect to the shift operator $(EG)(x) = G(x+1)$. For a polynomial $G \in \mathcal{K}(q)[x]$, the k-th falling *q-factorial* of G is defined by

$$[G]_{q}^{\underline{k}} := \begin{cases} G \cdot \left(\epsilon^{-1} G\right) \cdot \left(\epsilon^{-2} G\right) \cdot \ldots \cdot \left(\epsilon^{-k+1} G\right) & \text{if } k > 0 \quad , \\ 1 & \text{if } k = 0 \quad . \end{cases}$$

For example, $(x)_n = [1 - xq^{n-1}]_q^{\underline{n}}$. By separating the q-monic part and considering the q-shift operator ϵ instead of E, we are able to define the unique *q-greatest factorial factorization* (qGFF) of a polynomial $G \in \mathcal{K}(q)[x]$:

$$G = c \cdot x^j \cdot [G_1(x)]_q^{\underline{1}} \cdot [G_2(x)]_q^{\underline{2}} \cdot \ldots \cdot [G_k(x)]_q^{\underline{k}} \quad ,$$

where $c \in \mathcal{K}(q)$, and $G_i \in \mathcal{K}(q)[x]$ are q-monic. Note that the monomial x^j $(j \geq 0)$, corresponds to the root 0 of G with multiplicity j. One can prove that for the qGFF of $G \in \mathcal{K}(q)[x]$ (as above) the following analogue of the gcd-property of the GFF holds:

$$\gcd(G, \epsilon G) = x^j \cdot [G_1(x)]_q^{\underline{0}} \cdot [G_2(x)]_q^{\underline{1}} \cdot \ldots \cdot [G_k(x)]_q^{\underline{k-1}} \quad .$$

By an argument analogous to that in Sec. 5, it is possible to derive a unique q-analogue of the Gosper-Petkovšek representation, in short the qGP form, of the rational representation $r(x)$ of a q-hypergeometric sequence. Let us introduce some notation first. We consider rational functions

$$r(x) = \frac{A(x)}{B(x)} = \frac{A_1(x)}{B_1(x)} \cdot x^{\alpha} \cdot q^{\beta} \cdot z \quad ,$$

where $A, B \in \mathcal{K}(q)[x]$ are such that $A_1, B_1 \in \mathcal{K}(q)[x]$ are q-monic, α and β are integers, and $z \in \mathcal{K}$. Let $\mu(x) := x^{\alpha} \in \mathcal{K}(x)$ and $\pi(q) := q^{\beta} \in \mathcal{K}(q)$. It is convenient to use $num(s)$ (resp. $den(s)$) for the numerator (resp. denominator) of a reduced rational function s. Now we can formulate the properties of the qGP form:

> For any (nonzero) $r \in \mathcal{K}(q)(x)$, as above, there exist unique q-monic polynomials $P, Q, R \in \mathcal{K}(q)[x]$ such that
>
> $$\frac{A_1}{B_1} = \frac{\epsilon\, P}{P} \cdot \frac{Q}{\epsilon\, R} \quad ,$$

the *q-monic part* of r, with $\gcd(P,Q)=1=\gcd(P,R)$ and $\gcd(Q,\epsilon^j R)=1$ for all $j\geq 1$, and

$$r=\frac{\epsilon\,\widetilde{P}}{\widetilde{P}}\cdot\frac{\widetilde{Q}}{\epsilon\,\widetilde{R}}\quad , \tag{36}$$

where

$$\begin{aligned}\widetilde{P} &= P\cdot num(\pi(x))\quad ,\\ \widetilde{Q} &= Q\cdot z\cdot num(\mu(x))/den(\pi(q))\quad ,\\ \epsilon\widetilde{R} &= (\epsilon R)\cdot den(\mu(x))\quad .\end{aligned}$$

With this qPG form in our hands, we are able to derive a q-analog of GOSPER's algorithm simply by following the steps made in Sec. 5:

1. Given the qGP form (36) of a rational representation $r\in\mathcal{K}(q)(x)$, as above, of a q-hypergeometric sequence $(f_k)_{k\geq 0}$, the q-key equation

   $$\widetilde{P}=\widetilde{Q}\cdot(\epsilon Y)-\widetilde{R}\cdot Y$$

   has to be solved for $Y\in\mathcal{K}(q)[x]$.

2. Then

   $$\frac{C}{D}=\frac{\epsilon Y}{Y}\cdot\frac{\widetilde{Q}}{\widetilde{R}}\in\mathcal{K}(q)(x)$$

   is the rational representation of the q-hypergeometric solution, i.e.

   $$g_k=\frac{D(q^k)}{C(q^k)-D(q^k)}\cdot f_k$$

   solves

   $$f_k=g_{k+1}-g_k\quad .$$

As an elementary illustrating example let us take, for any fixed $n\geq 0$ and for $k\geq 0$,

$$f_k=(-1)^k q^{\binom{k}{2}-(n-1)k}\begin{bmatrix}n\\k\end{bmatrix}\quad ,$$

where

$$\begin{bmatrix}n\\k\end{bmatrix}:=\begin{cases}\dfrac{(q)_n}{(q)_k(q)_{n-k}} & \text{, if } 0\leq k\leq n\quad ,\\ 0 & \text{, otherwise}\quad ,\end{cases}$$

is the Gaussian polynomial, which for $q=1$ turns into the ordinary binomial coefficient $\binom{n}{k}$. Writing $x=q^k$, the rational representation of f_{k+1}/f_k is

$$\frac{f_{k+1}}{f_k}=\frac{q\,(1-q^{-n}x)}{1-qx}=:r(x)\quad .$$

¿From the qGP form of $r(x)$, namely

$$r(x) = \frac{qx}{x} \cdot \frac{1-q^{-n}}{1-qx} \quad ,$$

we get that

$$\tilde{P}(x) = x \ , \ \tilde{Q}(x) = 1 - q^{-n}x \ , \quad \text{and} \ \tilde{R}(x) = 1 - x \quad .$$

Solving the q-key equation

$$x = (1 - q^{-n}x)\, Y(qx) - (1-x)\, Y(x)$$

yields

$$Y(x) = 1/(1-q^{-n}) \in \mathcal{K}(q,x) \quad .$$

Hence

$$C/D \in \mathcal{K}(q,x) \quad \text{with} \quad C(x) = 1 - q^{-n}x \ \ \text{and} \ \ D(x) = 1 - x$$

is the rational representation of the q-hypergeometric solution, i.e.

$$g_k = \frac{D(q^k)}{C(q^k) - D(q^k)} \cdot f_k = -q^{n-k}\frac{1-q^k}{1-q^n} \cdot f_k$$

solves $f_k = g_{k+1} - g_k$. Consequently, for $m \geq 0$,

$$\sum_{k=0}^{m} (-1)^k q^{\binom{k}{2}-(n-1)k} \begin{bmatrix} n \\ k \end{bmatrix} = g_{m+1} - g_0 = (-1)^m q^{\binom{m}{2}-(n-1)m} \begin{bmatrix} n-1 \\ m \end{bmatrix} \quad .$$

Note that for $q = 1$ this turns into the well-known identity

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

In section 2 we have seen that the sign-less variation of $\sum_{k=0}^{m} \binom{n}{k}$ is *not hypergeometric* as a function of m.

Being equipped with a q-analogue of GOSPER's algorithm, a q-analogue of ZEILBERGER's fast algorithm can be designed in a straightforward manner. Instead of considering hypergeometric sequences $(f_{n,k})$, as in section 3, we consider sequences which are *q-hypergeometric* in both parameters, i.e. $f_{n+1,k}/f_{n,k}$ and $f_{n,k+1}/f_{n,k}$ are rational functions in q^n and q^k over $\mathcal{K}(q)$. For instance, taking the sequence of Gaussian polynomials $f_{n,k} = \begin{bmatrix} n \\ k \end{bmatrix}$ we have (compare with Eq. (7))

$$\begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \Big/ \begin{bmatrix} n \\ k+1 \end{bmatrix} = \frac{1-q^{n+1}}{1-q^{n-k}} \quad \text{and} \quad \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \Big/ \begin{bmatrix} n+1 \\ k \end{bmatrix} = \frac{1-q^{n-k+1}}{1-q^{k+1}} \quad .$$

The annihilating operator $A_S(N) = \sum_{l=0}^{d} C_l(n) N^{-l}$ now comes with coefficient polynomials $C_l(x)$ from $\mathcal{K}(q)[x]$. Finally, the certificate $cert_S(n,k)$, as well as the corresponding $r(n,k)$, is a rational function over $\mathcal{K}(q)$, rational in q^n and q^k.

We give no further details, instead we conclude this section by an example for which we use RIESE's[28] implementation written in Mathematica. This package is not the first, but the most comprehensive implementation of a q-analogue of Zeilberger's fast algorithm. Maple programs for the q-case are due to ZEILBERGER[5] and KOORNWINDER[15].

We shall prove here, for any $n \geq 0$,

$$\sum_l q^{6l^2+l} \begin{bmatrix} 2n+2 \\ n-3l \end{bmatrix} \frac{1-q^{6l+2}}{1-q^{2n+2}} = 1 \quad , \tag{37}$$

an identity which is due to SCHUR[29]. Note that the sum actually is taken over a finite integer intervall.

Invoking the program one gets:

```
zeder!11> math
Mathematica 2.1 for HP Apollo Domain/OS
Copyright 1988-92 Wolfram Research, Inc.
 -- Display Manager graphics initialized --

In[1]:= <<qZeil.m
Out[1]= Axel Riese's q-Zeilberger implementation version 1.3 loaded

In[2]:= FullCert = True
Out[2]= True

In[3]:= qZeil[q^(6l^2+l) qBinomial[2n+2,n-3l,q] (1-q^(6l+2))/
        (1-q^(2n+2)), {l,-Infinity,Infinity},n,1]
Out[3]= SUM[n] == SUM[-1 + n]

In[4]:= Cert
Out[4]=
       1 + n  1 + 3 l    n    3 l    n    2 + 3 l    n
      q     (q        - q )(-q    + q )(-q        + q )
-------------------------------------------------------------
       1 + 3 l      1 + 3 l        n       n        1 + 2 n
(-1 + q       )(1 + q       )(-1 + q )(1 + q )(-1 + q       )
```

The expression `Cert` corresponds to the qWZ-certificate $cert_S(n,l)$. For $n \to \infty$ identity (37) turns into a classic identity, namely Euler's pentagonal number theorem,

$$\sum_{k=-\infty}^{\infty} (-1)^k q^{\frac{1}{2}k(3k+1)} = \prod_{m=1}^{\infty} (1-q^m),$$

see e.g. ANDREWS[3].

We also would like to remark that the specialization $q = 1$ of identity (37) results in the nontrivial binomial coefficient identity

$$\sum_l \binom{2n+2}{n-3l} \frac{3l+1}{n+1} = 1 \quad .$$

This relatively simple example does not reflect the full power of the method. For instance, with RIESE's implementation it is possible to give computer proofs of most of the terminating summation and transformation formulas listed in the appendix of the book by GASPER and RAHMAN[6]. The range of applications also includes the automatic treatment of various identities from additive number theory, such as the celebrated Rogers-Ramanujan identities, see e.g. EKHAD and TRE[5], or PAULE[21].

7. GOSPER-PETKOVŠEK representation and rational summation

In this last section we come back to the indefinite summation problem for rational functions, as already treated in Sec. 4. As remarked earlier, if a rational function has a hypergeometric "antidifference", then this antidifference is itself rational, and GOSPER's algorithm can be used to determine it. Here we look again at the summation problem as stated in (18). Surprisingly, the data provided by the GOSPER-PETKOVŠEK representation of rational functions can be put into action for a surprisingly simple solution of the summation problem *à la Horowitz*. This method is even "optimal" with respect to the degrees of the denominator polynomials involved. A detailed description of this method is given in an article by PIRASTU and STREHL[27].

We want to solve the equation

$$\frac{A}{B} = \Delta \frac{C}{D} + \frac{F}{G} \quad , \tag{38}$$

where $A/B \in \mathcal{K}(x)$ is a given proper rational function (in reduced form), and $C/D, F/G \in \mathcal{K}(x)$ are the proper rational functions to be determined such that G has dispersion zero, i.e. $\gcd(G, E^j G) = 1$ for all integers $j \neq 0$.

Eq. (38) can be solved in many different ways, and following HOROWITZ' idea we can do this by first appropriately fixing the denominator polynomials D and G, and the solving a linear system corresponding to Eq. (38) for the undetermined coefficients of C and F, respecting the degree bound for these polynomials. Note that we do not insist on finding C/D and F/G in reduced form. After determining the coefficients of C and F from the linear system, it may happen that in the fractions C/D and F/G a cancellation takes place.

Since in this approach we need to solve a linear system of size $\deg D + \deg G$, it is desirable to fix D and G such that

- a polynomial solution of Eq. (38) (for C and F) in the desired form is possible;

- $\deg D$ and $\deg G$ are as small as possible for any choice of denominator polynomials with this property.

It turns out that such an "optimal" way of fixing D and G *a priori* can be made in the following sense:

- the choice of D and G depends only on the denominator polynomial B of the input, and leads to a solution of Eq. (38) for *any* choice of A such that $\deg A < \deg B$;
- the polynomials D and G have minimum degree among all polynomials with this property.

In particular, this "optimal" method yields denominator polynomials D with (usually) much lower degree than the approach by PAULE mentioned in Sec. 4, whereas the degree of G is the same in both methods. We remark that, in general, even the degree of the "optimal" D is much larger that the degree of B — a notable difference to the situation with integration, see Eq. (16), where the degree of D is always less that the degree of B.

It is convenient to introduce the notion of *shift-equivalence* of irreducible polynomials $G, H \in \mathcal{K}[x]$:

$$G \sim_E H \quad \overset{def}{\Leftrightarrow} \quad \exists k \in \mathbf{Z} \; : \; E^k G = H$$

The shift-equivalence class containing G will be written as $\langle G \rangle_E$. From now on, Γ denotes an arbitrary system of representatives for the shift-equivalence classes of irreducible polynomials. If $F, G \in \mathcal{K}[x]$ are polynomials, with G irreducible, we define the *G-part of F* as

$$\Psi_G(F) := \prod_{j \in \mathbf{Z}} \left(E^j G \right)^{a_j} \quad ,$$

where $(E^j G)^{a_j}$ is the exact power of $E^j G$ dividing F — this is the largest factor of F that "belongs to $\langle G \rangle_E$". This factor may be represented by the corresponding sequence of exponents

$$\psi_G(F) := (a_j)_{j \in \mathbf{Z}} \quad ,$$

which belongs to the set of sequences

$$\Omega_0 = \left\{ \; \omega = (\omega_j)_{j \in \mathbf{Z}} \; ; \; \omega_j \in \mathbf{N}, \text{ finitely many } \neq 0 \; \right\} \quad ,$$

i.e. the free abelian monoid over $\mathbf{Z}$, where we freely use additional componentwise operations and relations (infimum, comparison) in the natural way.

Once a representative system Γ has been fixed, a polynomial $F \in \mathcal{K}[x]$ can be identified with a finitary mapping

$$F \; : \; \Gamma \to \Omega_0 \; : \; G \mapsto \psi_G(F) \quad ,$$

where "finitary" means that only finitely many sequences $\psi_G(F)$ are different from the zero sequence, the neutral element of the monoid Ω_0. Conversely, any such mapping belongs to a polynomial $\in \mathcal{K}[x]$. In this way, transformations acting on polynomials can be specified by describing the corresponding transformation acting on sequences $\in \Omega_0$.

Four simple transformations of Ω_0 will be used in the sequel. We first define, for any nonzero sequence $\omega = (\omega_j)_{j\in\mathbf{Z}} \in \Omega_0$, the index of the *first maximum*, $m(\omega)$, and the index of the *last nonzero component*, $l(\omega)$:

$$\begin{aligned} j < m(\omega) \quad &\Rightarrow \quad \omega_j < \omega_{m(\omega)} \quad , \quad j \geq m(\omega) \Rightarrow \omega_j \leq \omega_{m(\omega)} \quad , \\ \omega_{l(\omega)} \quad &\neq \quad 0 \ , \ j > l(\omega) \Rightarrow \omega_j = 0 \end{aligned}$$

We then put:

$\hat{\omega} = (\hat{\omega}_j)_{j\in\mathbf{Z}}$, the *cover* of ω : this is the least unimodal sequence $\eta \in \Omega_0$ s.th. $\eta \geq \omega$ (w.r.t. the componentwise partial ordering of Ω_0);

$\overline{\omega} = (\overline{\omega}_j)_{j\in\mathbf{Z}}$, the *reduced cover* of ω , defined by

$$\overline{\omega}_j = \begin{cases} \hat{\omega}_j & \text{if } j < m(\omega) (= m(\hat{\omega})) \quad , \\ \hat{\omega}_{j+1} & \text{if } j \geq m(\omega) \quad ; \end{cases}$$

$\omega^{max} = ({\omega^{max}}_j)_{j\in\mathbf{Z}}$, the *maximum indicator* of ω , defined by

$$\omega_j^{max} = \begin{cases} 0 & \text{if } j \neq m(\omega) \quad , \\ \omega_{m(\omega)} & \text{if } j = m(\omega) \quad ; \end{cases}$$

$\omega^+ = \left(\omega_j^+\right)_{j\in\mathbf{Z}}$, the *compression* of ω , where

$$\omega_j^+ = \begin{cases} 0 & \text{if } j \neq l(\omega) \quad , \\ \sum_{i\in\mathbf{Z}} \omega_i & \text{if } j = l(\omega) \quad . \end{cases}$$

We can now define the *covering transformation* $F \mapsto \hat{F}$ for polynomials by simply requiring that

$$\forall G \in \Gamma \ : \ \psi_G(\hat{F}) = \widehat{\psi_G(F)}$$

i.e. by executing the covering transformation $\omega \mapsto \hat{\omega}$ on all sequences $\psi_G(F)$ $(G \in \Gamma)$ "in parallel". The same can be done for the other three operations. Note that none of these transformations depends on the choice of Γ.

We can now state the crucial result about optimality, which follows from a detailed study of the action of the Δ-operator on shift equivalence classes:

- The optimal choice for the denominator polynomials D and G in the rational summation problem, given the input $A/B \in \mathcal{K}(x)$, is:

$$D = \overline{B} \quad , \quad G = B^{max}$$

The (perhaps surprising) link with the Gosper-Petkovšek representation is then:

- Let
$$\frac{B}{E\,B} = \frac{E\,P}{P} \cdot \frac{Q}{E\,R}$$
be the GOSPER-PETKOVŠEK representation of the rational function $B/(E\,B)$, then
$$\overline{B} = \frac{P \cdot B}{R} = \frac{\hat{B}}{R} \quad , \quad B^{max} = Q^{+} \quad .$$

We conclude that

> an "optimal" rational summation algorithm can be easily designed on the basis of any algorithm computing the GOSPER-PETKOVŠEK representation.

Note that the transformation $Q \mapsto Q^{+}$ can be implemented without factoring.

We illustrate the rational summation algorithm just outlined and the concepts introduced in this section by an example. Note that the factored form of the polynomials are given for the purpose of illustration only — computations are done using the expanded forms.

Consider the rational function $A/B \in \mathbf{Q}(x)$, where

$$\begin{array}{rcl} A & := & x^2 - 3\,x + 1 \\ B & := & x^{12} + 9\,x^{11} + 30\,x^{10} + 38\,x^9 - 12\,x^8 - 68\,x^7 - 46\,x^6 - 22\,x^5 - 5\,x^4 + 75\,x^3 \\ & = & (\,x-1\,)^2\,x^3\,(\,x+3\,)\,(\,x^2+1\,)\,(\,x^2+4\,x+5\,)^2 \end{array}$$

We have two shift classes, $\langle x \rangle$ and $\langle x^2+1 \rangle$, and

$$\begin{array}{rl} \Psi_x(B) = (x-1)^2 x^3 (x+3) & i.e. \quad \psi_x(B) = (\,\ldots 0\,2\,\underline{3}\,0\,0\,1\,0 \ldots\,) \\ \Psi_{x^2+1}(B) = (x^2+1)(x^2+4x+5)^2 & i.e. \quad \psi_{x^2+1}(B) = (\,\ldots 0\,\underline{1}\,0\,2\,0 \ldots\,) \end{array}$$

where underlining denotes the position of index zero. From

$$\begin{array}{ll} \psi_x(\hat{B}) = \widehat{\psi_x(B)} = (\,\ldots 0\,2\,\underline{3}\,1\,1\,1\,0 \ldots\,) & \psi_{x^2+1}(\hat{B}) = \widehat{\psi_{x^2+1}(B)} = (\,\ldots 0\,\underline{1}\,1\,2\,0 \ldots\,) \\ \psi_x(\overline{B}) = \overline{\psi_x(B)} = (\,\ldots 0\,2\,\underline{1}\,1\,1\,0 \ldots\,) & \psi_{x^2+1}(\overline{B}) = \overline{\psi_{x^2+1}(B)} = (\,\ldots 0\,\underline{1}\,1\,0 \ldots\,) \end{array}$$

and

$$\begin{array}{rcl} \psi_x(B^{max}) & = & \psi_x(B)^{max} = (\ldots\,0\,0\,\underline{3}\,0\,0\,\ldots) \\ \psi_{x^2+1}(B^{max}) & = & \psi_{x^2+1}(B)^{max} = (\ldots 0\,\underline{0}\,0\,2\,0\,0\,\ldots) \end{array}$$

we expect to find

$$C = (x-1)^2(x)(x+1)(x+2)(x^2+1)(x^2+2x+2) \quad , \quad G = x^3(x^2+4x+5)^2 \quad ,$$

Indeed, the computation of the GOSPER-PETKOVŠEK representation of $B/(E\,B)$ outputs the three polynomials

$$\begin{aligned}
P &= x^4+5\,x^3+10\,x^2+10\,x+4\\
&= (\,2+2\,x+x^2\,)\,(\,2+x\,)\,(\,1+x\,)\\
Q &= x^7+2\,x^6-x^5-4\,x^4+3\,x^3-6\,x^2+5\,x\\
&= (\,x-1\,)^2\,x\,(\,x^2+1\,)\,(\,x^2+4\,x+5\,)\\
R &= x^7+7\,x^6+20\,x^5+32\,x^4+28\,x^3+12\,x^2\\
&= x^2\,(\,3+x\,)\,(x^2+4\,x+5\,)^2 \quad .
\end{aligned}$$

Note that

$$\begin{aligned}
\psi_x(Q) = (\ldots\,0\,2\,\underline{1}\,0\,\ldots) \quad &, \quad \psi_{x^2+1}(Q) = (\ldots 0\,\underline{1}\,0\,1\,0\,\ldots)\,,\\
\psi_x(Q^+) = \psi_x(Q)^+ = (\ldots\,0\,\underline{3}\,0\,\ldots) \quad &, \quad \psi_{x^2+1}(Q^+) = \psi_{x^2+1}(Q)^+ = (\ldots\,0\,\underline{0}\,0\,2\,0\,\ldots).
\end{aligned}$$

Next we find

$$\begin{aligned}
D = P\cdot B/R &= x^9+3\,x^8+2\,x^7-2\,x^6-5\,x^5-3\,x^4-2\,x^3+2\,x^2+4\,x\\
&= (\,x^2+1\,)\,x\,(\,x-1\,)^2\,(\,2+2\,x+x^2\,)\,(\,2+x\,)\,(\,1+x\,)\quad ,\\
G = Q^+ &= x^7+8x^6+26x^5+40x^4+25x^3\\
&= x^3\,(\,x^2+4x+5\,)^2 \quad .
\end{aligned}$$

Solving now a polynomial equation equivalent to

$$\frac{A}{B} = \Delta\,\frac{C}{D} + \frac{F}{G} \tag{39}$$

for the coefficients of

$$C = c_0 + c_1 x + \cdots + c_8 x^8 \quad \text{and} \quad F = f_0 + f_1 x + \cdots + f_6 x^6$$

yields a solution of Eq. (39), where

$$\frac{C}{D} = \frac{-1900-24428\,x+25768\,x^2+\cdots-13947\,x^7+222\,x^8}{43200\;(x^2+1)\,x\,(x-1)^2\,(2+2\,x+x^2)\,(2+x)\,(1+x)}\quad ,$$

$$\frac{F}{G} = \frac{24000-34250\,x-56900\,x^2-27745\,x^3-4612\,x^4+37\,x^5}{72000\;(x^2+4\,x+5)^2\,x^3}\quad .$$

8. Acknowledgements

A first draft of this article was written during a stay of the second author at the RISC institute in October 1994, sponsored by the *Deutsch-Österreichisches Hochschullehrer-Austauschprogramm*. The hospitality and inspiring working atmosphere of RISC is gratefully acknowledged.

9. References

1. S. A. Abramov, On the summation of rational functions. *Zh. vychisl. mat. Fiz.* **11** (1971), pp. 1071–1075.
2. S. A. Abramov, Rational component of the solution of a first linear recurrence relation with a rational right hand side. *Zh. vychisl. mat. Fiz.* **14** (1975), pp. 1035–1039.
3. G.E. Andrews, *The Theory of Partitions*, Encyclopedia of Mathematics and its Applications, Vol.2, ed. G.-C. Rota (Addison-Wesley, Reading, 1976; Reissued: Cambridge University Press, London and New York, 1985).
4. P. Cartier, Démonstration "automatique" d'identités et fonctions hypergeométriques [d'après D. Zeilberger]. *Séminaire Bourbaki* **746** (1991).
5. S.B. Ekhad and S. Tre, A purely verification proof of the first Rogers-Ramanujan identity, *J. Combinatorial Theory (A)* **54** (1990), pp. 309–311.
6. G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and its Applications Vol.35, ed. G.-C. Rota (Cambridge University Press, London and New York, 1990).
7. K. O. Geddes, S. R. Czapor, and G. Labahn, *Algorithms for Computer Algebra.* (Kluwer, Boston, 1992).
8. R. W. Gosper jr., Decision procedure for indefinite hypergeometric summation. *Proc. Natl. Acad. Sci. USA*, **75** (1978), pp. 40–42.
9. H.W. Gould, *Combinatorial Identities - A Standardized Set of Tables Listing 500 Binomial Coefficient Summations.* Revised Edition (Morgantown, West Virginia University, 1972).
10. R.L. Graham, D.E. Knuth, and O. Patashnik, *Concrete Mathematics, A Foundation for Computer Science*, 2nd edition (Addison-Wesley, Reading, 1994).
11. E. Hermite, Sur l'intégration des fonctions rationnelles, *Nouv. Annal. Math.* (1872), pp.145-148.
12. E. Horowitz, Algorithms for partial fraction decomposition and rational integration. In *Proc. SYMSAM '71*, ed. S.R. Petrick, (ACM Press, 1971), pp.441-457.
13. M. Karr, Summation in finite terms. *Journal of the ACM* **28** (1081), 305–350.
14. M. Karr, Theory of summation in finite terms. *J. Symbolic Computation* **1** (1985), pp. 303–315.
15. T.H. Koornwinder, On Zeilberger's algorithm and its q-analogue: a rigorous description. *J. Comp. Appl. Math.* **48** (1993), pp. 91–111.
16. J. C. Lafon, Summation in finite terms. In *Computer Algebra : Symbolic and Algebraic Computation*, eds. B. Buchberger, G. E. Collins, and R. Loos (Springer-Verlag, Wien-New York, 1983), pp. 71–77.
17. P. Lisoněk, P. Paule and V. Strehl, An improvement of the degree setting in Gosper's algorithm. *J. Symbolic Computation* **16** (1993), pp. 243–258.
18. R. Moenck, On computing closed forms for summations. In *Proceedings of the*

1977 MACSYMA user's conference, Berkeley, California, 1977), pp. 225–236,

19. P. Paule, Greatest factorial factorization and symbolic summation I. RISC-Linz Report Series 93-02, J. Kepler University, Linz (1993).
20. P. Paule, Greatest factorial factorization and symbolic summation II (in preparation).
21. P. Paule, Short and Easy Computer Proofs of the Rogers-Ramanujan Identities and of Identities of Similar Type, *Electronic Journal of Combinatorics* **1** (1994) R10.
22. P. Paule and M. Schorn, A Mathematica version of Zeilberger's algorithm for proving binomial coefficient identities, RISC-Linz Report Series 93-36, J. Kepler University, Linz (1993).
23. M. Petkovšek, Hypergeometric solutions of linear recurrences with polynomial coefficients. *J. Symbolic Computation* **14** (1992), pp. 243–264.
24. R. Pirastu, *Algorithmen zur Summation rationaler Funktionen.* Diploma thesis, Univ. Erlangen-Nürnberg (1992), (in german).
25. R. Pirastu, Algorithms for indefinite summation of rational functions in Maple. Submitted to *The Maple Technical Newsletters* (1994).
26. R. Pirastu, A note on the minimality problem in indefinite summation of rational functions. *Actes 31e Séminaire Lotharingien de Combinatoire, Publications de l'I.R.M.A.* 1994/021, ed. J. Zeng, (Strasbourg, 1994), pp.95-101.
27. R. Pirastu and V. Strehl, Rational summation and Gosper-Petkovšek representation. IMMD-Erlangen, Technical Report 94-03, Univ. Erlangen-Nürnberg (1994).
28. A. Riese, *A Mathematica q-analogue of Zeilberger's algorithm for proving q-hypergeometric identities*, Diploma Thesis, Univ. Linz, in preparation.
29. I. Schur, Ein Beitrag zur additiven Zahlentheorie und zur Theorie der Kettenbrüche, *S.-B. Preuss. Akad. Wiss. Phys.-Math.Kl.* (1917), pp. 302–321. (Reprinted: Gesammelte Abhandlungen, Vol.2, Springer, Berlin (1973), 117–136.)
30. T.N. Subramaniam and D.E.G. Malm, How to integrate rational functions, *Amer. Math. Monthly* **99/8** (1992).
31. V. Strehl, Binomial identities — combinatorial and algorithmic aspects. IMMD-Erlangen, Technical Report 93-01, Univ. Erlangen-Nürnberg (1993). (To appear in *Discrete Mathematics*).
32. V. Strehl, Binomial sums and identities, *The Maple Technical Newsletter* **10** (1993), pp. 37–49.
33. H.S. Wilf, *Identities and their Computer Proofs.* "SPICE" lecture notes, available via anonymous `ftp` of the file `pub/wilf/lecnotes.ps` at the site `ftp.cis.upenn.edu` (1993).
34. H.S. Wilf and D. Zeilberger, An algorithmic proof theory for hypergeometric (ordinary and "q") multisum/integral identities. *Inventiones Math.* **108** (1992), pp. 575–633.
35. D. Zeilberger, A holonomic systems approach to special function identities. *J. Comp. Appl. Math.* **32** (1990) pp. 321–368.

36. D. Zeilberger, A fast algorithm for proving terminating hypergeometric identities. *Discrete Math.* **80** (1990), pp. 207–211.
37. D. Zeilberger, A Maple program for proving hypergeometric identities. *SIGSAM Bulletin, ACM Press* **25** (1991), pp. 4–13.
38. D. Zeilberger, The method of creative telescoping, *J. Symbolic Computation* **11** (1991), pp. 195–204.
39. D. Zeilberger, Three recitations on holonomic systems and hypergeometric series. *Proc. of the 24th Séminaire Lotharingien*, ed. D. Foata, (Publ. I.R.M.A. Strasbourg, 1993), pp. 5–37.
(Will be reprinted in the special issue "Symbolic Computation in Combinatorics Δ_1" of the *J. Symbolic Computation*)