RISC-Linz
Research Institute for Symbolic Computation
Johannes Kepler University
A-4040 Linz, Austria, Europe

Rational Summation and Gosper-Petkovšek Representation

Roberto PIRASTU, Volker STREHL

(February 9, 1995)

RISC-Linz Report Series No. 95-08

Supported by: Commission of the European Communities in the framework of the program "Human Capital and Mobility", contract nr. ERBCHBICT930501.

Published in: Journal of Symbolic Computation, Special Issue on *Symbolic Computation and Combinatorics* Δ_1, P. Paule, V. Strehl (eds.)

J. Symbolic Computation (1995) **11**, 1–000

Rational Summation and Gosper-Petkovšek Representation

Roberto Pirastu[§†] and Volker Strehl[¶]

[§]*RISC-Linz, J. Kepler University, Linz, Austria*
Roberto.Pirastu@risc.uni-linz.ac.at

[¶]*IMMD-1 (Informatik), University of Erlangen-Nürnberg, Germany*
strehl@informatik.uni-erlangen.de

(Received 22 September 1994)

Indefinite summation essentially deals with the problem of inverting the difference operator $\Delta : f(X) \to f(X+1) - f(X)$.

In the case of rational functions over a field k we consider the following version of the problem

> given $\alpha \in k(X)$, determine $\beta, \gamma \in k(X)$ such that $\alpha = \Delta\,\beta + \gamma$, where γ is as "small" as possible (in a suitable sense).

In particular, we address the question

> what can be said about the denominators of a solution (β, γ) by looking at the denominator of α only ?

An "optimal" answer to this question can be given in terms of the Gosper-Petkovšek representation for rational functions, which was originally invented for the purpose of indefinite hypergeometric summation. This information can be used to construct a simple new algorithm for the rational summation problem.

1. Introduction

The problem of indefinite summation is the discrete analog of the probleme of indefinite integration: it essentially asks for inverting the finite difference operator

$$\Delta : f(X) \mapsto f(X+1) - f(X)$$

on an appropriate space of functions.

In this article we consider Δ acting on the rational functions $k(X)$ over some field k of characteristic 0. (We may restrict our attention to *proper* rational functions, i.e. functions $f/g \in k(X)$, where the polynomials $f, g \in k[X]$ satisfy $\deg f < \deg g$, since inverting Δ on polynomials is trivial). Since Δ is not surjective on $k(X)$, e.g. $1/X \notin \Delta k(X)$, we

[†] Supported by the Commission of the European Communities in the framework of the program "Human Capital and Mobility", contract Nr. ERBCHBICT930501.

0747–7171/90/000000 + 00 \$03.00/0

consider the summation problem for (proper) rational functions in the following form (which was apparently first formulated in Abramov (1975))

> given $f/g \in k(X)$, find a/u and $b/v \in k(X)$ such that
> $$\frac{f}{g} = \Delta\left(\frac{a}{u}\right) + \frac{b}{v} \tag{1.1}$$
> where the "remainder" b/v is as "small" as possible (in some suitable sense).

The corresponding problem in the case of indefinite integration of rational functions, i.e. solving, for given $f, g \in k[X]$,

$$\frac{f}{g} = \frac{d}{dX}\left(\frac{a}{u}\right) + \frac{b}{v} \tag{1.2}$$

for $a, b, u, v \in k[X]$, is usually attacked in either of two ways (see ch. 11 in Geddes et al. (1992)):

- by iteratively building up the "rational part" a/u by extracting "integrable" constituents of the b/v-part; (Hermite's method)
- by first computing appropriate denominator polynomials u, v from g alone, and then solving the system (1.2) for a and b with undetermined coefficients (Horowitz' method).

Naturally, one is tempted to try to carry over these classical techniques to the situation of the difference operator. But this turns out to be less straightforward than one may expect. Indeed, the approach described in Moenck (1977), trying to simulate the Hermite method, suffers from theoretical deficiencies.

These problems were first noted by P. Paule who tries to establish (among others) a firm theoretical basis for the rational summation problem, introducing the concept of *greatest factorial factorization* in place of the *squarefree decomposition* used in the integration situation. Based on this concept, in Paule (1993) a Horowitz-type summation algorithm for rational functions is proposed, which has the disadvantage of producing denominator polynomials u, v in the situation (1.1) which are usually much too large (in degree), hence leading to much too large linear systems for the coefficients of a and b to solve.

Note, however, that iterative algorithms for rational summation, different from the above, have already been proposed in Abramov (1971,1975).

Pirastu (1992) gives an overview and a unified treatment of these algorithms and some improvements, together with implementations in the computer algebra System Maple (see also Pirastu (1994a)).

The present article has the following purposes:

- to establish an algebraic setup (decomposition w.r.t. shift-equivalence classes of irreducible polynomials) in order to deal with the summation problem for rational functions on a local-global basis, [Sections 2 and 3];
- to provide precise information about the "optimal" (in a sense to be specified) choice of the denominator polynomials u, v in (1.1) in a Horowitz-type approach, based on (local) information about the shift-classes associated with the denominator polynomial g, [Section 4] — this questions has been addressed earlier in Pirastu (1992) and Pirastu (1994b);

- to show (from local considerations) how this information can be obtained from the Gosper-Petkovšek representation of rational functions, which was originally invented in the context of indefinite hypergeometric summation, see Gosper (1978), Petkovšek (1992), and for which we give a purely combinatorial equivalent, [Section 5];
- to propose a new summation method for rational functions, based on (known) algorithms for producing the Gosper-Petkovšek representation of rational functions, [Section 6].

In Section 5.1 we give a detailed example, describing the combinatorial equivalent of the Gosper-Petkovšek representation. An example of complete application of the new summation method is presented in Section 7.

Let us underline the following aspect: although our investigations are mostly of *local* nature, i.e., referring to decompositions with respect to shift-equivalence classes of irreducible polynomials, and thus apparently referring to a factorization of the initial data into irreducible polynomials, the final algorithm operates *globally* and does not use any factorization procedure at all - it is based on gcd- and resultant-computation only.

2. Problem Description

Let k be a field of characteristic 0. As usual, $k(X)$ denotes the field of rational functions over k. Elements $\alpha = \alpha(X) \in k(X)$ are written as quotients $\alpha = f/g$, where f and $g \neq 0$ are polynomials in X over k. This representation is *normalized* if $\gcd(f, g) = 1$ and if g is monic. Usually, but not always, we will assume that elements of $k(X)$ are represented in this way. A *proper* rational function is an element $\alpha = f/g \in k(X)$ such that $\deg f < \deg g$. The constant 0 is the only proper rational function which is constant. The proper rational functions form a k-subalgebra of $k(X)$, denoted by $\mathcal{R}$. By polynomial division $k(X) \cong k[X] \oplus \mathcal{R}$.

The shift operator E and the (forward) difference operator Δ on $k(X)$ are defined as usual:

$$\begin{array}{rlll} E & : k(X) \longrightarrow k(X) & : \alpha(X) \mapsto \alpha(X+1) \\ \Delta = E - I & : k(X) \longrightarrow k(X) & : \alpha(X) \mapsto \alpha(X+1) - \alpha(X) \end{array}$$

Note that E is a k-algebra isomorphism, Δ is k-linear and has the constant functions as its kernel. If restricted to the subalgebra $\mathcal{R}$, E is still a k-algebra isomorphism, and Δ is now injective, since 0 is the only element in its kernel. *Indefinite summation* of rational function essentially asks for inverting the linear operator Δ. Since inverting Δ on polynomials is trivial, we can from now on restrict our attention to the algebra $\mathcal{R}$. This has the advantage that $\Delta^{-1}\alpha$ is uniquely determined - if it exists! As is well known, the latter is not always the case: Δ is not surjective on $\mathcal{R}$. E.g., for any $j > 0$, there is no $\beta \in k(X)$ such that $\Delta\beta = 1/X^j$. More generally: if $f/g^i \in \mathcal{R}$ is a rational function in reduced form, with g irreducible, then it does not belong to $\Delta\mathcal{R}$. In view of this phenomenon one has to make a choice between the following alternatives:

- Asking for a decision procedure for the existence of $\Delta^{-1}\alpha \in \mathcal{R}$, and giving an algorithm to construct such an element in the case of a positive answer. Different approaches and algorithms in this direction were presented in Abramov (1971), Gosper (1978), and Man (1993).

- Enlarging the domain of functions under consideration (e.g. by adding polygamma functions), so that at least every $\alpha \in \mathcal{R}$ has an inverse w.r.t. Δ. Moenck's approach mentioned above goes into this direction. A general method in analogy to Risch's integration method is described in Karr (1981, 1985).
- Considering a "refined" rational summation problem: given $\alpha \in \mathcal{R}$, construct $\beta \in \mathcal{R}$ which is as "close" as possible to what we expect $\Delta^{-1}\alpha$ to be: making $\gamma = \alpha - \Delta\beta \in \mathcal{R}$ as "small" as possible. In particular this requires: for $\alpha \in \Delta\mathcal{R}$ one should get the true inverse, i.e., $\gamma = 0$, and in general, if the same procedure is applied to the difference γ, this should not lead to any improvement. This problem has apparently been first stated in Abramov (1975).

In this note we will concentrate on the last alternative. As a reasonable measure of "smallness" we will take the degree of the denominator polynomial in the reduced presentation of γ. We define:

$$\text{for } \alpha = f/g \in \mathcal{R}, \text{ where } \gcd(f,g) = 1, \quad \| \alpha \| := \deg g$$

Note that $\| \; . \; \|$ induces a metric on $\mathcal{R}$.

The rational summation problem can the be stated as follows:

> Given $\alpha \in \mathcal{R}$, determine β and γ in $\mathcal{R}$ such that $\alpha = \Delta\beta + \gamma$, where $\| \gamma \|$ is minimal.

Thus one asks for an element of $\Delta\mathcal{R}$ which is closest to α w.r.t. the $\| \; . \; \|$-metric. Naturally, we will say that α is *summable* if $\alpha \in \Delta\mathcal{R}$, i.e., if there is such pair (β, γ) with $\gamma = 0$.

The *existence* part of the rational summation problem is no problem at all, but the answer to the question to what extent *uniqueness* holds is less obvious.

3. Localization

For any monic irreducible polynomial $g \in k[X]$ let $\mathcal{R}_g$ denote the subalgebra of rational functions $f/g^i \in \mathcal{R}$, where $i \geq 0$. By partial fraction representation

$$\mathcal{R} \cong \bigoplus_g \mathcal{R}_g$$

where $\bigoplus_g$ runs over all monic irreducible polynomials. Clearly

$$\text{if } \alpha = \sum_g \alpha_g \;, \text{ where } \; \alpha_g \in \mathcal{R}_g \;, \text{ then } \quad \| \alpha \| = \sum_g \| \alpha_g \|$$

When dealing with the shift operator E and the difference operator Δ, one has to consider E-orbits, i.e., shift-invariant subspaces of $\mathcal{R}$. For any monic irreducible polynomial $g \in \mathcal{R}$ we put

$$\mathcal{R}_{[g]} := \bigoplus_{i \in \mathbf{Z}} \mathcal{R}_{E^i g}$$

Here $[g]$ denotes the class of monic irreducible polynomials *shift-equivalent* to g, i.e., the polynomials $E^i g(X) = g(X+i)$ for $i \in \mathbf{Z}$. In the following the notation $\bigoplus_{[g]}$ and

$\sum_{[g]}$ will be used to indicate product and sums over a representative system for the shift-equivalence classes of monic irreducible polynomials. Thus

$$\mathcal{R} \cong \bigoplus_{[g]} \mathcal{R}_{[g]}$$

with

$$\alpha = \sum_{[g]} \alpha_{[g]} \quad \text{and} \quad \alpha_{[g]} = \sum_{i \in \mathbf{Z}} \alpha_{E^i g}$$

It is clear that any equation

$$\alpha = \Delta\beta + \gamma \quad (\alpha, \beta, \gamma \in \mathcal{R})$$

localizes to

$$\alpha_{[g]} = \Delta\beta_{[g]} + \gamma_{[g]}$$

for any monic irreducible polynomial g. And since

$$\| \alpha \| = \sum_{[g]} \| \alpha_{[g]} \|$$

we can state

> if $(\beta, \gamma) \in \mathcal{R}^2$ is a solution for the rational summation problem for $\alpha \in \mathcal{R}$, then for each monic irreducible polynomial g (i.e., for its shift-equivalence class) the pair $(\beta_{[g]}, \gamma_{[g]}) \in \mathcal{R}^2_{[g]}$ is a solution of the rational summation problem for $\alpha_{[g]}$ — and conversely.

This shows that for answering the uniqueness question it suffices to study the "local" situation in the components $\mathcal{R}_{[g]}$. And in particular:

> $\alpha \in \mathcal{R}$ is summable if and only if $\alpha_{[g]} \in \mathcal{R}_{[g]}$ is summable for each (shift-equivalence class of) monic irreducible polynomial(s)

4. The structure of the local solutions

4.1. Local transformations

We start with another simple observation:

> For each $d \in \mathbf{Z}$ the operator $\Delta_d = E^d - 1$ is divisible by $\Delta_1 = \Delta$:
>
> $$\Delta_d = E^d - I = \begin{cases} \Delta \cdot (I + E + E^2 + \cdots + E^{d-1}) & \text{if } d > 0 \\ -\Delta \cdot (E^{-1} + E^{-2} + \cdots + E^{d}) & \text{if } d < 0 \end{cases}$$

which means that

$$\Delta_d(\frac{f}{g^a}) = \Delta \left(\frac{\Delta_d}{\Delta} \frac{f}{g^a} \right) \in \Delta\mathcal{R}_{[g]}$$

for each $f/g^a \in \mathcal{R}_{[g]}$. Thus any two terms $E^i(f_1/g^a)$ and $E^j(f_2/g^b)$ with $i \neq j$, appearing in the canonical decomposition of some $\alpha \in \mathcal{R}_{[g]}$ may be transformed according to either of the two identities (assuming $d = j - i > 0$):

$$\begin{aligned} E^i \frac{f_1}{g^a} + E^j \frac{f_2}{g^b} &= E^i \left(\frac{f_1}{g^a} + \frac{f_2}{g^b} \right) + \Delta(E^{d-1} + \cdots + I) E^i \frac{f_2}{g^b} \\ E^i \frac{f_1}{g^a} + E^j \frac{f_2}{g^b} &= E^j \left(\frac{f_1}{g^a} + \frac{f_2}{g^b} \right) - \Delta(E^{-d} + \cdots + E^{-1}) E^j \frac{f_1}{g^a} \end{aligned}$$

Notice that up to a shift E^{j-i} (or E^{i-j}) the first terms on the right are the same, namely a shift of

$$\frac{f_1}{g^a} + \frac{f_2}{g^b} = \frac{f}{g^c} \quad \text{where } c \leq \max\{a, b\}.$$

If these rational functions are written in reduced form, then $c = \max\{a, b\}$ if $a \neq b$. Note that

$$\| \, E^i \frac{f_1}{g^a} + E^j \frac{f_2}{g^b} \, \| = (a + b) \cdot \deg g > c \cdot \deg g = \| \, \frac{f_1}{g^a} + \frac{f_2}{g^b} \, \|$$

4.2. The structure of the γ part

If we take an arbitrary element $\alpha \in \mathcal{R}_{[g]}$,

$$\alpha = E^{i_1} \frac{f_1}{g^{a_1}} + \cdots + E^{i_k} \frac{f_k}{g^{a_k}}$$

say, with pairwise distinct shifts i_j $(1 \leq j \leq k)$, then starting from $(\beta_0, \gamma_0) = (0, \alpha)$ one may produce through $k - 1$ of such "local transformations" a pair (β_k, γ_k) such that $\alpha = \Delta\beta_k + \gamma_k$, where $\beta_k, \gamma_k \in \mathcal{R}_{[g]}$, and where in particular γ_k is a shift of

$$\frac{f_1}{g^{a_1}} + \cdots + \frac{f_k}{g^{a_k}} = \frac{f}{g^a} \quad \text{with} \quad a \leq \max\{a_1, \ldots, a_k\}$$

(if all the rational function are written in reduced form).

The pair (β_k, γ_k) is a solution of the local summation problem, as the following consideration shows.

If γ_k vanishes, then α is summable. This can only happen if $\max\{a_1, \ldots, a_k\}$ is attained at least twice.

If γ_k does not vanish, then consider any two decompositions

$$\alpha = \Delta\beta^{(1)} + \gamma^{(1)} = \Delta\beta^{(2)} + \gamma^{(2)}$$

where $\alpha, \beta^{(i)}, \gamma^{(i)} \in \mathcal{R}_{[g]}$, $(i = 1, 2)$ and where the $\gamma^{(i)}$ cannot be reduced any further by local transformations. Solutions of the rational summation problem are of that kind. We have

$$\gamma^{(1)} = E^i \frac{f_1}{g^a} \quad , \quad \gamma^{(2)} = E^j \frac{f_2}{g^b}$$

for some i, j, f_1, f_2, a, b. Hence

$$\Delta(\beta^{(1)} - \beta^{(2)}) = \gamma^{(2)} - \gamma^{(1)} = E^j \frac{f_2}{g^b} - E^i \frac{f_1}{g^a}$$

If $i = j$, then the r.h.s is of the form $E^i \left(\frac{f_2}{g^b} - \frac{f_1}{g^a} \right) = E^i \frac{f}{g^c}$, which is possible only if the r.h.s. vanishes, i.e., if $a = b$ and $f_1 = f_2$.

If $i \neq j$, then one local transformation step leads to a similar situation, and by the same argument one concludes that $a = b$ and $f_1 = f_2$.

As a consequence:

> Let $\alpha \in \mathcal{R}_{[g]}$, then either $\alpha \in \Delta\mathcal{R}_{[g]}$ (i.e., α is summable), or else there exist unique $\beta \in \mathcal{R}_{[g]}, f$ and a such that
>
> $$\alpha = \Delta\beta + \frac{f}{g^a}$$
>
> where $\gamma = f/g^a$ is reduced.

Note that the γ part may be shifted freely, i.e., we could impose $\gamma = E^i(f/g^a)$ for any $i \in \mathbf{Z}$. This amounts to selecting a different representative g in $[g]$, but for a natural choice of g see $(*)$ below. The exponent a is independent of the shift i, of course, and $a \leq \max\{a_1, \ldots, a_k\}$. The uniqueness of the β part follows from the injectivity of Δ on $\mathcal{R}$.

4.3. The structure of the β part

Before considering the β part of the localized summation problem, let us introduce a bit of notation. In Section 5.1 we describe an example for the concepts introduced below.

Given $q \in k[x]$ and a monic irreducible polynomial g, we may consider the *spectrum* of q with respect to g, i.e., the doubly infinite sequence

$$\langle q, g\rangle = (a_i)_{i\in\mathbf{Z}} \quad , \text{ where } \quad E^i g^{a_i} \parallel q$$

i.e., a_i is the maximum integer, such that $E^i g^{a_i} \mid q$. For $\alpha = p/q \in \mathcal{R}$ in reduced form we define the spectrum of α with respect to g by $\langle \alpha, g\rangle := \langle q, g\rangle$. Note that $\langle \alpha, g\rangle = (a_i)_{i\in\mathbf{Z}}$ means

$$\alpha_{[g]} = \sum_{i\in\mathbf{Z}} \alpha_{E^i g} = \sum_{i\in\mathbf{Z}} E^i \frac{f_i}{g^{a_i}}$$

with respect to the canonical decomposition (in reduced form) of $\alpha_{[g]}$. Denote by $g^{\mathbf{a}}$ the polynomial $\prod_{i\in\mathbf{Z}} E^i g^{a_i}$ for any sequence $\mathbf{a}$ with only nonnegative, only finitely many nonzero components.

The sequences $\langle \alpha, g\rangle$ are integer sequences with finite support, but for reasons that become evident later on, we have to consider more general classes of sequences as well. In the following we will consider doubly infinite sequences $\mathbf{a} = (a_i)_{i\in\mathbf{Z}}$ and operators acting on them, but some care has to be taken in order to make sure that these operators are well defined. The arithmetic operations "$+$" and "$-$" on sequences, as well as the infimum "$\wedge$" and the comparison "$\leq$" are defined coordinatewise.

Naturally, we will have a shift and a delta operator on sequences:

$$\begin{aligned} \epsilon : (a_i)_{i\in\mathbf{Z}} \mapsto \epsilon\,\mathbf{a} &:= (a_{i-1})_{i\in\mathbf{Z}} \\ \delta : (a_i)_{i\in\mathbf{Z}} \mapsto \delta\,\mathbf{a} &:= (a_i - a_{i-1})_{i\in\mathbf{Z}} = (1-\epsilon)\,\mathbf{a} \end{aligned}$$

The inverse of the δ operator (on an appropriate subspace of sequences) is the summation operator σ given by

$$\sigma : (a_i)_{i\in\mathbf{Z}} \mapsto \sigma\,\mathbf{a} := \left(\textstyle\sum_{j\leq i} a_i\right)_{i\in\mathbf{Z}}$$

Two further operators, which are also only defined on appropriate subspaces, are the maximum operators "from the left" and "from the right".

$$\begin{aligned}\overrightarrow{\mu} : (a_i)_{i\in\mathbf{Z}} \mapsto \overrightarrow{\mu}\,\mathbf{a} &:= \left(\max_{j\le i} a_j\right)_{i\in\mathbf{Z}} \\ \overleftarrow{\mu} : (a_i)_{i\in\mathbf{Z}} \mapsto \overleftarrow{\mu}\,\mathbf{a} &:= \left(\max_{j\ge i} a_j\right)_{i\in\mathbf{Z}}\end{aligned}$$

Let g be monic irreducible polynomial, and let us consider $\alpha \in \mathcal{R}_{[g]}$ such that

$$\alpha = \sum_{i\in\mathbf{Z}} \alpha_{E^i g} = \sum_{i\in\mathbf{Z}} E^i \,\frac{f_i}{g^{a_i}}$$

Then

$$\begin{aligned}\alpha &= \sum_{i<0} E^i \,\frac{f_i}{g^{a_i}} + \frac{f_0}{g^{a_0}} + \sum_{j>0} E^j \,\frac{f_j}{g^{a_j}} \\ &= \sum_{i<0}\left(\frac{f_i}{g^{a_i}} - \Delta\sum_{s=i}^{-1} E^s \,\frac{f_i}{g^{a_i}}\right) + \frac{f_0}{g^{a_0}} + \sum_{j>0}\left(\frac{f_j}{g^{a_j}} + \Delta\sum_{t=0}^{j-1} E^t \,\frac{f_j}{g^{a_j}}\right) \\ &= \Delta\beta + \gamma\end{aligned}$$

where

$$\beta = -\sum_{s<0} E^s \sum_{i\le s} \frac{f_i}{g^{a_i}} + \sum_{t\ge 0} E^t \sum_{j>t} \frac{f_j}{g^{a_j}} \tag{4.1}$$

$$\gamma = \sum_{i\in\mathbf{Z}} \frac{f_i}{g^{a_i}} \tag{4.2}$$

If we write $\mathbf{a} = (a_i)_{i\in\mathbf{Z}} = \langle\alpha, g\rangle$ and $\mathbf{b} = (b_i)_{i\in\mathbf{Z}} = \langle\beta, g\rangle$ then it follows that

$$b_s \ \text{ is } \ \begin{cases} \le (\overrightarrow{\mu}\mathbf{a})_s & \text{for } s<0 \text{ in general} \\ = (\overrightarrow{\mu}\mathbf{a})_s & \text{for } s<0 \text{ with } (\delta\overrightarrow{\mu}\mathbf{a})_s > 0 \text{ in particular}\end{cases}$$

$$b_t \ \text{ is } \ \begin{cases} \le (\epsilon^{-1}\overleftarrow{\mu}\mathbf{a})_t & \text{for } t\ge 0 \text{ in general} \\ = (\epsilon^{-1}\overleftarrow{\mu}\mathbf{a})_t & \text{for } t\ge 0 \text{ such that } (\delta\epsilon^{-1}\overleftarrow{\mu}\mathbf{a})_t < 0 \text{ in particular}\end{cases}$$

This shows that in general (i. e., unless some cancellation takes place) the denominator polynomial of β is given by $g^{\hat{\mathbf{a}}}$, where

$$\hat{a}_s = \begin{cases} (\overrightarrow{\mu}\,\mathbf{a})_s & \text{for } s<0 \\ (\epsilon\,\overleftarrow{\mu}\,\mathbf{a})_s & \text{for } s\ge 0\end{cases}$$

Note that one has

$$\overrightarrow{\mu}\,\mathbf{a} \wedge \epsilon^{-1}\overleftarrow{\mu}\,\mathbf{a} \le \hat{\mathbf{a}}$$

with equality holding if and only if

$$\min_s \{\, a_s = \max\langle\alpha, g\rangle \,\} \le 0 \le \max_s \{\, a_s = \max\langle\alpha, g\rangle \,\} \tag{$*$}$$

This condition can always be satisfied by making a proper choice of the representative g from its shift-equivalence class $[g]$. A "natural" candidate for the denominator of β is thus the polynomial

$$g^{\overrightarrow{\mu}\,\mathbf{a} \wedge \epsilon^{-1} \overleftarrow{\mu}\,\mathbf{a}}$$

with g taken from $[g]$ such that $(*)$ holds. We will show that this candidate is "optimal" in the sense that for a suitable choice of the numerator polynomial of α no cancellation occurs and $g^{\overrightarrow{\mu}\,\mathbf{a} \wedge \epsilon^{-1} \overleftarrow{\mu}\,\mathbf{a}}$ is indeed the true denominator of β.

4.4. Optimality

Given $\alpha = p/q \in \mathcal{R}$, the values of the candidate denominator polynomials u, v of β, γ given above depend only on the denominator q of α, and on the choice of the proper representatives in each shift-equivalence class $[g]$. We now show that the "Ansatz"

$$u = \prod_g g^{\overrightarrow{\mu}\langle\alpha, g\rangle \wedge \epsilon^{-1} \overleftarrow{\mu}\langle\alpha, g\rangle}$$

is **optimal** in the sense that for any polynomial q there is always a choice of the polynomial p such that a solution β for the β-part of the summation problem for α has precisely this denominator u (and not a proper divisor of it). As usual, we may restrict our attention to a single shift-equivalence class $[g]$.

Let $\alpha \in \mathcal{R}_{[g]}, \alpha \neq 0$, written as

$$\alpha = \sum_{i \in \mathbf{Z}} E^i \, \frac{f_i}{g^{a_i}}$$

where the sequence $\mathbf{a} = \langle\alpha, g\rangle$ satisfies property $(*)$ above, i. e., a proper choice of the representative $g \in [g]$ has been made. Consider $\mathbf{a}$ as fixed, and the sequence $(f_i)_{i \in \mathbf{Z}}$ yet to be determined.

Let now $\mathbf{b} = \overrightarrow{\mu}\,\mathbf{a} \wedge \epsilon^{-1} \overleftarrow{\mu}\,\mathbf{a}$ (we assume $\mathbf{b} \neq \mathbf{0}$, otherwise the summation problem for α would have the trivial solution $\beta = 0, \gamma = \alpha$).

From property $(*)$ we know

$$\mathbf{b}_s = \begin{cases} (\overrightarrow{\mu}\,\mathbf{a})_s & \text{if } s < 0 \\ (\epsilon^{-1} \overleftarrow{\mu}\,\mathbf{a})_s & \text{if } s \geq 0 \end{cases}$$

Consider now the sequence of polynomials

$$f_i = \begin{cases} 1 & \text{if } i = k \\ 1 - g^{(\delta\,\mathbf{b})_i} & \text{if } k < i < 0 \end{cases}$$

where $k = \min_s \{ \, a_s \neq 0 \, \} = \min_s \{ \, b_s \neq 0 \, \}$.

It is easy to check that these polynomials satisfy the system of equations

$$\sum_{k \leq i \leq s} f_i \cdot g^{b_s - a_i} = 1 \qquad (k \leq s < 0)$$

Similarly, the family of polynomials

$$f_j = \begin{cases} 1 & \text{if } j = l \\ 1 - g^{-(\delta\,\mathbf{b})_j} & \text{if } 0 \leq j < l \end{cases}$$

where $l = \max_t \{ a_t \neq 0 \} = \max_t \{ b_t \neq 0 \} + 1$, satisfies the system of equations

$$\sum_{t<j\leq l} f_j \cdot g^{b_t - a_j} = 1 \qquad (0 \leq t < l)$$

We put $f_i = 0$ if $i < k$ or $i > l$. Note that for all $i \in \mathbf{Z}$ f_i is coprime g^{a_i} and $\deg f_i < \deg g^{a_i}$, so that

$$\alpha = -\sum_{i<0} \frac{f_i}{g^{a_i}} + \sum_{j\geq 0} \frac{f_j}{g^{a_j}} \in \mathcal{R}$$

and is written in reduced form. Comparison with (4.1) shows that the β-part of $\alpha = \Delta\beta + \gamma$ satisfies

$$\beta = \sum_{k\leq s<l} E^s \frac{1}{g^{b_s}}$$

and thus $\langle \beta, g \rangle = \mathbf{b}$. In other words: $g^{\mathbf{b}}$ is the true denominator of β.

4.5. Conclusion

The rational summation problem for α has a solution (β, γ) such that for each shift-equivalence class $[g]$ with $\alpha_{[g]} \neq 0$ we have:

- The denominator of $\beta_{[g]}$ divides $g^{\overrightarrow{\mu}\langle\alpha, g\rangle \wedge \epsilon^{-1}\overleftarrow{\mu}\langle\alpha, g\rangle}$
- The denominator of $\gamma_{[g]}$ divides $g^{\max\langle\alpha, g\rangle}$

provided appropriate representatives $g \in [g]$ have been choosen (i. e., representatives satisfying $(*)$).

An algorithm for computing (β, γ) may thus proceed as follows:

- Given $\alpha \in \mathcal{R}$, compute polynomials u, v:

$$\begin{aligned} u &= \prod_g g^{\overrightarrow{\mu}\langle\alpha, g\rangle \wedge \epsilon^{-1}\overleftarrow{\mu}\langle\alpha, g\rangle} \\ v &= \prod_g g^{\max\langle\alpha, g\rangle} \end{aligned}$$

  where the products run over an appropriately chosen representative system for the shift-equivalence classes of monic irreducible polynomials (respecting $(*)$).
- Put $\beta = a/u$, $\gamma = b/v$, where a, b are polynomials with $\deg a < \deg u$, $\deg b < \deg v$ with indeterminate coefficients.
- Determine the true a and b by solving

$$\alpha = \frac{Ea}{Eu} - \frac{a}{u} + \frac{b}{v}$$

An algorithm for computing u and v will be presented in Section 6. Note that the solution $(a/u, b/v)$ will not necessarily be reduced, but in general this "Ansatz" is the optimum one can do (in keeping u and v as "small" as possible) without taking properties of the numerator of α into account.

We finally remark, that our results about the denominator polynomial v can be rephrased using the notion of dispersion, as introduced in Abramov (1975): the *dispersion* $\mathrm{dis}(q)$ of a polynomial q is defined by

$$\mathrm{dis}(q) := \max\{h \in \mathbf{Z};\ Res_X(q, E^h\, q) = 0\}$$

From the definition it follows that $\mathrm{dis}(q)$ is the maximum integer root of $Res_X(q, E^h\, q)$ as a polynomial in the indeterminate h, i.e., it is the maximum integer h such that q and $E^h\, q$ have a common factor.

Clearly: v (as above) and all the variants, obtained by shifting the contributions from the shift-equivalence classes freely (see Section 4.2) are polynomials with dispersion zero.

As a consequence, as has been remarked earlier in Abramov (1975), solutions of the rational summation problem can be characterized as follows:

> Let $\alpha \in \mathcal{R}$, then $(\beta, \gamma) \in \mathcal{R}^2$ is a solution of the rational summation problem for α if and only if
>
> $$\alpha = \Delta\beta + \gamma \qquad \text{and} \qquad \mathrm{dis}(\mathrm{denom}(\gamma)) = 0$$

The existence of a solution (β, γ) follows from Section 4.2, where it was also shown how two distinct solutions are related. The latter question has first been answered in Paule (1993).

5. The Gosper-Petkovšek representation of rational functions

The following representation of rational functions is at the basis of Gosper's classical decision method for indefinite hypergeometric summation:

> For any rational function $\alpha \in k(X)$ there are polynomials $p, q, r \in k[X]$ such that
>
> $$\alpha = \frac{E\,p}{p} \cdot \frac{q}{E\,r} \quad \text{with} \quad \gcd(q, E^i r) = 1 \quad \text{for all } i \geq 1$$

Petkovšek showed in Petkovšek (1992) that a presentation of $\alpha \in k(X)$ as

$$\alpha = c \cdot \frac{E\,p}{p} \cdot \frac{q}{E\,r} \quad \text{with} \quad \gcd(q, E^i r) = 1 \quad \text{for all } i \geq 1 \quad , \quad \gcd(p, r) = 1 = \gcd(p, q)$$

with monic polynomials $p, q, r \in k[X]$ and $c \in k$ is *unique*.

Note that the usual algorithms for computing p, q, r (and c), as outlined in Gosper (1978) and Petkovšek (1992), are based on resultant- and gcd-computations, together with a search for integer zeros of polynomials. In the following, we will look at this representation from a different point of view, related to the decomposition of rational functions according to shift-equivalence classes of irreducible polynomials. This is not meant as an *algorithmic* approach, but it gives a *combinatorial* view of this classical result which turns out to be useful for the rational summation problem.

Note first that the Gosper-Petkovšek representation localizes perfectly (because it is a purely multiplicative statement). We may split α into a product $\alpha^{[g_1]} \cdots \alpha^{[g_k]}$, where the g_i are irreducible polynomials belonging to distinct shift-equivalence classes, and where each factor $\alpha^{[g_i]}$ accounts for the contribution of factors from the class of g_i to α. If we have a (unique) local representation

$$\alpha^{[g_i]} = \frac{E\,p_i}{p_i} \cdot \frac{q_i}{E\,r_i}$$

with the appropriate gcd-conditions satisfied, then the (unique) Gosper-Petkovšek-triple (p, q, r) for α results from multiplication†:

$$p = p_1 \cdots p_k \quad , \quad q = q_1 \cdots q_k \ , \ r = r_1 \cdots r_k$$

If we look now at the local situation for any irreducible polynomial g, then $\alpha^{[g]}$ may be represented by a doubly infinite sequence of integers

$$(a_i)_{i \in \mathbf{Z}} := \langle f, g \rangle - \langle h, g \rangle \quad \text{for } \alpha = f/h$$

where "−" is the componentwise difference of sequences.

Let us say that an integer sequence $(a_i)_{i \in \mathbf{Z}}$ is a *rational sequence* if there are only finitely many nonzero terms. If all terms are nonnegative, and again only finitely many different from 0, then it is a *polynomial sequence*. Adopting this terminology, the Gosper-Petkovšek representation boils down to the following combinatorial assertion, where δ and ϵ are operators on sequences as before and $\mathbf{0}$ is the all-zero sequence.

The combinatorial equivalent of the Gosper-Petkovšek representation is illustrated by an example in Section 5.1.

PROPOSITION 5.1. *(Combinatorial Gosper-Petkovšek representation) Let* $\mathbf{a}$ *be any rational sequence. Then there are unique polynomial sequences* $\mathbf{p}$, $\mathbf{q}$, *and* $\mathbf{r}$ *such that*

$$\mathbf{a} = -\delta\, \mathbf{p} + \mathbf{q} - \epsilon\, \mathbf{r}$$

where

$$\mathbf{p} \wedge \mathbf{q} = \mathbf{0} \ , \ \mathbf{p} \wedge \mathbf{r} = \mathbf{0} \ , \ \mathbf{q} \wedge \epsilon^j\, \mathbf{r} = \mathbf{0} \textit{ for all } j \geq 1$$

PROOF. We define $\mathbf{f} := \sigma\, \mathbf{a}$ and

$$\mathbf{q} := \delta\, \overrightarrow{\mu}\, \mathbf{f} \ , \ \epsilon\, \mathbf{r} := -\delta\, \overleftarrow{\mu}\, \mathbf{f}$$

We then have $\mathbf{q} \geq \mathbf{0}$, since $\overrightarrow{\mu}\, \mathbf{f}$ is non-decreasing, and $\mathbf{q}$ is a polynomial sequence, since $(\delta\, \overrightarrow{\mu}\, \mathbf{f})_i > 0$ implies $(\delta\, \mathbf{f})_i = a_i > 0$.
Similarly, $\mathbf{r} \geq \mathbf{0}$, since $\overleftarrow{\mu}\, \mathbf{f}$ in non-increasing, and $\mathbf{r}$ is a polynomial sequence since $(\delta\, \overleftarrow{\mu}\, \mathbf{f})_i > 0$ implies $(\delta\, \mathbf{f})_i = a_i < 0$.
Now obviously

$$(\delta\, \overrightarrow{\mu}\, \mathbf{f})_i > 0 \text{ and } (\delta\, \overleftarrow{\mu}\, \mathbf{f})_j < 0 \text{ implies } i \leq j$$

so that $\mathbf{q} \wedge \epsilon^j\, \mathbf{r} = \mathbf{0}$ holds for all $j \geq 1$.

Consider now

$$\mathbf{p} := \overrightarrow{\mu}\, \mathbf{f} + \overleftarrow{\mu}\, \mathbf{f} - \mathbf{f} - \mathbf{m_f} = (\overrightarrow{\mu} - id)\, \mathbf{f} \wedge (\overleftarrow{\mu} - id)\, \mathbf{f}$$

where $\mathbf{m_f}$ denotes the sequence which has value $m_{\mathbf{f}} := \max_{i \in \mathbf{Z}} f_i$ everywhere (note that this value is well-defined since $\delta\, \mathbf{f} = \mathbf{a}$ is a rational sequence). Again, it is easy to check that $\mathbf{p}$ is a polynomial sequence, and we have in addition

$$\delta\, \mathbf{p} = \delta\, \overrightarrow{\mu}\, \mathbf{f} + \delta\, \overleftarrow{\mu}\, \mathbf{f} - \delta\, \mathbf{f} = \mathbf{q} - \epsilon\, \mathbf{r} - \mathbf{a}$$

as desired.

We now have to show that $\mathbf{p} \wedge \mathbf{q} = \mathbf{0}$ and $\mathbf{p} \wedge \mathbf{r} = \mathbf{0}$ hold. Note

† The role of the scalar factor c in Petkovšek's assertion is completely irrelevant for our purpose.

- if $q_i = (\delta \overrightarrow{\mu}\, \mathbf{f})_i \neq 0$, then necessarily $f_i = (\overrightarrow{\mu}\, \mathbf{f})_i$ and thus $p_i = (\overleftarrow{\mu}\, \mathbf{f})_i - m_{\mathbf{f}} \leq 0$, which means $p_i = 0$, since $\mathbf{p}$ is a polynomial sequence.
- if $r_j = -(\delta \overleftarrow{\mu}\, \mathbf{f})_{j+1} \neq 0$, then necessarily $f_j = (\overleftarrow{\mu}\, \mathbf{f})_j$ and thus $p_i = (\overrightarrow{\mu}\, \mathbf{f})_j - m_{\mathbf{f}} \leq 0$, which means $p_i = 0$, since $\mathbf{p}$ is a polynomial sequence.

So far the *existence* part of the proposition has been established. For the uniqueness part, let $\mathbf{p}, \mathbf{q}, \mathbf{r}$ be any polynomial sequences such that the assertion of the lemma holds. We will show that these are identical with the corresponding sequences defined above.

Let $i_0 \in \mathbf{Z}$ be an index such that both

$$(\sigma\, \mathbf{q})_{i_0} = m_{\sigma\, \mathbf{q}} \quad \text{and} \quad (\epsilon\, \sigma\, \mathbf{r})_{i_0} = 0$$

hold. Note that the condition: "$\mathbf{q} \wedge \epsilon^j \mathbf{r} = \mathbf{0}$ for all $j \geq 1$" guarantees the existence of such an index. We then have

$$(\epsilon\, \sigma\, \mathbf{r})_i = 0 \;\; \text{for all } i \leq i_0 \quad , \quad (\sigma\, \mathbf{q})_j = m_{\sigma\, \mathbf{q}} \;\; \text{for all } j \geq i_0$$

For $i \leq i_0$ we now have

$$f_i = (\sigma\, \mathbf{a})_i = -p_i + (\sigma\, \mathbf{q})_i$$

hence $f_i \leq (\sigma\, \mathbf{q})_i$, and the "orthogonality" of $\mathbf{p}$ and $\mathbf{q}$ implies that $f_i = (\sigma\, \mathbf{q})_i$ whenever $q_i \neq 0$ holds. Both facts together imply

$$(\overrightarrow{\mu}\, \mathbf{f})_i = (\sigma\, \mathbf{q})_i \;\; \text{for all } i \leq i_0$$

Similarly, for $j \geq i_0$ we have

$$f_j = (\sigma\, \mathbf{a})_j = -p_j + m_{\sigma\, \mathbf{q}} - (\sigma\, \epsilon\, \mathbf{r})_j$$

hence $f_j \leq m_{\sigma\, \mathbf{q}} - (\sigma\, \epsilon\, \mathbf{r})_j$, and here the "orthogonality" of $\mathbf{p}$ and $\mathbf{r}$ implies $f_j = m_{\sigma\, \mathbf{q}} - (\sigma\, \epsilon\, \mathbf{r})_j$ whenever $r_j \neq 0$. Here these two facts imply

$$(\overleftarrow{\mu}\, \mathbf{f})_j = m_{\sigma\, \mathbf{q}} - (\sigma\, \epsilon\, \mathbf{r})_j \;\; \text{for all } j \geq i_0$$

We conclude, in particular, that $f_{i_0} = m_{\sigma\, \mathbf{q}}$, hence $m_{\mathbf{f}} = m_{\sigma\, \mathbf{q}}$, and consequently

$$\overrightarrow{\mu}\, \mathbf{f} = \sigma\, \mathbf{q} \;\; \text{and} \;\; \overleftarrow{\mu}\, \mathbf{f} = \mathbf{m_f} - \sigma\, \epsilon\, \mathbf{r}$$

which implies

$$\mathbf{q} = \delta\, \overrightarrow{\mu}\, \mathbf{f} \;\; \text{and} \;\; \epsilon\, \mathbf{r} = -\delta\, \overleftarrow{\mu}\, \mathbf{f}$$

Finally

$$\delta\, \mathbf{p} = -\delta\, \mathbf{f} + \mathbf{q} - \epsilon\, \mathbf{r} = \delta\, (-\mathbf{f} + \overrightarrow{\mu}\, \mathbf{f} + \overleftarrow{\mu}\, \mathbf{f})$$

and

$$\mathbf{p} = -\mathbf{f} + \overrightarrow{\mu}\, \mathbf{f} + \overleftarrow{\mu}\, \mathbf{f} - \mathbf{m_f}$$

follows. □

5.1. An Example

As an example, we determine the Gosper-Petkovšek representation of the rational function α given by

$$\alpha = \frac{(X-3)(X-2)^2(X+2)(X+5)^2}{(X-4)(X+1)^3(X+3)^2}$$

The rational sequence $\mathbf{a}$ associated to α with respect to the irreducible polynomial $g = X$ is

$$\mathbf{a} = \langle (X-3)(X-2)^2(X+2)(X+5)^2, X\rangle - \langle (X-4)(X+1)^3(X+3)^2, X\rangle$$

According to the proposition, we compute the following sequences, where the position of the index 0 is indicated by underlining.

$$\begin{array}{rcccccccccccccccc}
\mathbf{a} & = & \dots & 0 & -1 & 1 & 2 & 0 & \underline{0} & -3 & 1 & -2 & 0 & 2 & 0 & \dots \\
\mathbf{f} = \sigma\mathbf{a} & = & \dots & 0 & -1 & 0 & 2 & 2 & \underline{2} & -1 & 0 & -2 & -2 & 0 & 0 & \dots \\
\overrightarrow{\mu}\mathbf{f} & = & \dots & 0 & 0 & 0 & 2 & 2 & \underline{2} & 2 & 2 & 2 & 2 & 2 & 2 & \dots \\
\mathbf{q} = \delta\overrightarrow{\mu}\mathbf{f} & = & \dots & 0 & 0 & 0 & 2 & 0 & \underline{0} & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\overleftarrow{\mu}\mathbf{f} & = & \dots & 2 & 2 & 2 & 2 & 2 & \underline{2} & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
\epsilon\mathbf{r} = -\delta\overleftarrow{\mu}\mathbf{f} & = & \dots & 0 & 0 & 0 & 0 & 0 & \underline{0} & 2 & 0 & 0 & 0 & 0 & 0 & \dots \\
(\overrightarrow{\mu} - id)\mathbf{f} & = & \dots & 0 & 1 & 0 & 0 & 0 & \underline{0} & 3 & 2 & 4 & 4 & 2 & 2 & \dots \\
(\overleftarrow{\mu} - id)\mathbf{f} & = & \dots & 2 & 3 & 2 & 0 & 0 & \underline{0} & 1 & 0 & 2 & 2 & 0 & 0 & \dots \\
\mathbf{p} & = & \dots & 0 & 1 & 0 & 0 & 0 & \underline{0} & 1 & 0 & 2 & 2 & 0 & 0 & \dots
\end{array}$$

From this it follows that the Gosper-Petkovšek representation of α is given by

$$p = (X-4)(X+1)(X+3)^2(X+4)^2, \quad q = (X-2)^2, \quad r = X^2$$

One easily verifies that

$$\frac{Ep}{p}\cdot\frac{q}{Er} = \frac{(X-3)(X+2)(X+4)^2(X+5)^2}{(X-4)(X+1)(X+3)^2(X+4)^2}\cdot\frac{(X-2)^2}{(X+1)^2} = \alpha$$

and that $\gcd(q, E^i r) = 1$ for all $i \geq 1$ and $\gcd(p,r) = 1 = \gcd(q,r)$.

6. The Algorithm

The proof of the proposition shows that the Gosper-Petkovšek representation of rational functions provides an algorithmic way to compute the polynomials u and v from Section 4.5.

Let $s/t \in \mathcal{R}$ be in reduced form and g any irreducible polynomial. We apply the proposition to the rational function $\alpha_{[g]} = t_{[g]}/Et_{[g]}$. Let $\mathbf{t} := \langle t, g\rangle$, then we have $\mathbf{a} = \langle t_{[g]}, g\rangle - \langle Et_{[g]}, g\rangle = \mathbf{t} - \epsilon\,\mathbf{t}$ and this implies $\mathbf{f} = \sigma\,\mathbf{a} = \mathbf{t}$. For $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ from the Gosper-Petkovšek representation of α, the following holds

$$\begin{array}{rcll}
\langle p\cdot t, g\rangle - \langle r, g\rangle & = & \mathbf{p} + \mathbf{t} - \mathbf{r} & \\
& = & -\mathbf{t} + \overrightarrow{\mu}\,\mathbf{t} + \overleftarrow{\mu}\,\mathbf{t} - \mathbf{m_t} + \mathbf{t} + \epsilon^{-1}\,\delta\,\overleftarrow{\mu}\,\mathbf{t} & \\
& = & \overrightarrow{\mu}\,\mathbf{t} + \overleftarrow{\mu}\,\mathbf{t} - \mathbf{m_t} + \epsilon^{-1}\,\overleftarrow{\mu}\,\mathbf{t} - \overleftarrow{\mu}\,\mathbf{t} & \\
& = & \overrightarrow{\mu}\,\mathbf{t} + \epsilon^{-1}\,\overleftarrow{\mu}\,\mathbf{t} - \mathbf{m_t} & = \quad \overrightarrow{\mu}\,\mathbf{t} \wedge \epsilon^{-1}\,\overleftarrow{\mu}\,\mathbf{t}
\end{array}$$

Since this holds for any irreducible g, for the Gosper-Petkovšek representation (p,q,r) of $\alpha = t/Et$ we have that $u = pt/r$ is *globally* optimal, in the sense that condition $(*)$ is simultaneously satisfied for all shift equivalence classes. On an algorithmic level, this means that the optimum denominator polynomial u can be obtained directly from an algorithm computing the Gosper-Petkovšek representation - see Alg. 1 below. In particular, no factorization w.r.t. shift-equivalence classes is necessary.

Similarly, we show that also the denominator v can be obtained from the Gosper-Petkovšek representation of α. Consider again $\alpha_{[g]}$ as above and let i_0 be the smallest index such that $t_{i_0} = m_{\mathbf{t}}$, then from $\mathbf{q} = \delta\,\overrightarrow{\mu}\,\mathbf{t}$ it follows that i_0 is the biggest index such that $q_{i_0} \neq 0$. Let us define a sequence $\mathbf{q}^+ = (q_i^+)_{i\in\mathbf{Z}}$ by

$$q_i^+ = \begin{cases} 0 & \text{if } i \neq i_0 \\ \sum_{j\leq i_0} q_j & \text{if } i = i_0 \end{cases}$$

and the polynomial $q_{[g]}^+$ by $q_{[g]}^+ := g^{\mathbf{q}^+}$. Since $\sum_{i\leq i_0} q_i = m_{\mathbf{t}}$, the denominator $v_{[g]} = q_{[g]}^+$ is optimal with respect to the shift class $[g]$.

From this it follows that the denominator $v = q^+ := \prod_{[g]} q_{[g]}^+$ is optimal for s/t in the sense of condition $(*)$. In Alg. 2 we describe an algorithm for computing q^+ from the polynomial q of the Gosper-Petkovšek representation of t/Et, again using gcd and resultant computations only – no factorization is needed.

Summarizing, we have that

> for any $s/t \in \mathcal{R}$ an optimal choice of denominators u and v in the sense of condition $(*)$ is given by
>
> $$u = \frac{p\cdot t}{r} \quad\text{and}\quad v = q^+$$
>
> where (p, q, r) is the Gosper-Petkovšek representation of t/Et.

For the computation of the Gosper-Petkovšek representation of a rational function we follow the algorithm proposed in Petkovšek (1992), which we include for completeness as Alg. 1. Here $Res_X(p, q)$ denotes the resultant of $p(X)$ and $q(X)$ with respect to the indeterminate X.

$$(p, q, r) \longleftarrow \texttt{GP-Rep}\ (\alpha)$$

Inputs:

α : a rational function $\in \mathcal{R}$.

Outputs:

p, q, r : Gosper-Petkovšek representation of α, i. e.

$$\alpha = \frac{E\,p}{p}\cdot\frac{q}{E\,r} \text{ with } \gcd(q, E^i\,r) = 1 \text{ for all } i \geq 1 \text{ and } \gcd(p, r) = 1 = \gcd(p, q).$$

Begin

Step 1: *Initialization*

$p \longleftarrow 1$, $q \longleftarrow \texttt{numer}(\alpha)$, $r \longleftarrow \texttt{denom}(\alpha)$

Step 2:

for $h \in \{h' \in \mathbb{N};\ Res_X(q, E^{h'}\,r) = 0\}$ **do**
 $d \longleftarrow \texttt{gcd}(q, E^h\,r)$
 $q \longleftarrow q/d$
 $r \longleftarrow r/E^{-h}\,d$

$p \leftarrow p \cdot \prod_{i=1}^{h} E^{-i}\, d$
endfor
return (p, q, r)
End

Algorithm 1. `GP-Rep`

The computation of q^+ described in Alg. 2 refers to the concept of dispersion as mentioned at the end of Section 4.5. Several algorithms are known for computing the dispersion of a polynomial (see e.g. Abramov (1971), Pirastu (1992), Paule (1993), Man and Wright (1994)). In the context of computing the dispersion and, more generally, the positive integer roots of a resultant (as in step 2 of Alg. 1), one can make use of Loos' (1983) fast p-adic method for computing rational zeros of polynomials.

Furthermore, we say that v is the *part* of p in q for $v, p, q \in k[X]$ if $v \mid q$, $\gcd(p, q/v) = 1$ and only factors of p arise in v. In this case we write $v = \mathrm{part}(p, q)$. The computation of $\mathtt{part}(p, q)$ only needs repeated gcd-computations. Initialize $v_1 \leftarrow \mathtt{gcd}(p, q)$ and $q_1 \leftarrow q/v_1$, then compute $v_i \leftarrow \mathtt{gcd}(p, q_{i-1})$ and $q_i \leftarrow q_{i-1}/v_i$ for $i = 2, 3, \ldots$ until $v_n = 1$ for some n. Then all factors of p arising in q are isolated and we have $\mathtt{part}(p, q) \leftarrow v_1 v_2 \cdots v_n$.

With this notation the algorithm for computing the polynomial q^+ is described in Alg. 2.

$$q^+ \leftarrow \mathtt{plus}(q)$$

Inputs:
q : polynomial.

Outputs:
q^+ : polynomial with $\langle q^+, g\rangle = \langle q, g\rangle^+$ for all irreducible g.

Begin

Step 1: *Initialization*
$d \leftarrow \mathtt{dis}(q)$, $g \leftarrow q$, $q^+ \leftarrow 1$

Step 2: *Collect all factors at the end of each class*
for $j = d$ **downto** 0 **do**
$k \leftarrow \mathtt{part}(E^{-j}\, g, g)$
$g \leftarrow g/k$
$q^+ \leftarrow q^+ \cdot E^j\, k$
endfor
return (q^+)
End

Algorithm 2. `plus`

In Alg. 3 we show the algorithm for solving the rational summation problem with optimal bounds. Remark that Step 3 mainly reduces to the solution of a system of linear equation over the constant field k.

$$(\beta, \gamma) \leftarrow \texttt{Opt-Rat-Sum}(\alpha)$$

Inputs:
α : a rational function in $\mathcal{R}$.

Outputs:
β, γ : a solution of the rational summation problem for α

Begin

Step 1: *Initialization*

$t \leftarrow \texttt{denom}(\alpha)$

Step 2: *(Optimal) bounds for the denominators of β and γ*

$(p, q, r) \leftarrow \texttt{GP-Rep}(t/E\,t)$
$u \leftarrow \frac{p\,t}{r}$
$v \leftarrow \texttt{plus}(q)$

Step 3: *Computation of the numerators of β and γ*

$$a \leftarrow \sum_{i=0}^{\deg(u)-1} a_i X^i,\; b \leftarrow \sum_{j=0}^{\deg(v)-1} b_j X^j \text{ for indeterminates } a_i\text{'s and } b_j\text{'s}$$

determine the a_i's and b_j's by coefficient comparison from $\quad \alpha = E\,\dfrac{a}{u} - \dfrac{a}{u} + \dfrac{b}{v}$

return $(a/u, b/v)$
End

Algorithm 3. `Opt-Rat-Sum`

7. An Example

We want to compute a solution of the rational summation problem for the following rational function

$$\alpha = \frac{s}{t} = \frac{X^2 - 3X + 1}{(X-1)^2 X^3 (X+3)(X^2+1)(X^2+4X+5)^2}$$

For purpose of demonstration the denominator is written in factored form. Although this information is not required by the algorithm, we can read off the two shift-equivalence classes into which the denominator polynomial t splits, and we have

$$\begin{array}{rcccccccccc} \langle t, X\rangle & = & \ldots & 0 & 0 & 2 & \underline{3} & 0 & 0 & 1 & 0 & \ldots \\ \langle t, X^2+1\rangle & = & \ldots & 0 & 0 & 0 & \underline{1} & 0 & 2 & 0 & 0 & \ldots \end{array}$$

Step 2 in the algorithm first computes the Gosper-Petkovšek representation of t/Et.

We obtain

$$\begin{aligned}
\frac{t}{Et} &= \frac{(X-1)^2X(X+3)(X^2+1)(X^2+4X+5)^2}{(X+1)^3(X+4)(X^2+2X+2)(X^2+6X+10)^2} \\
&= \frac{(X+2)(X+3)(X^2+4X+5)}{(X+1)(X+2)(X^2+2X+2)} \cdot \frac{(X-1)^2X(X^2+1)(X^2+4X+5)}{(X+1)^2(X+4)(X^2+6X+10)^2} \\
&= \frac{Ep}{p} \cdot \frac{q}{Er}
\end{aligned}$$

and from this the optimal bounds

$$u = \frac{pt}{r} = (X-1)^2X(X+1)(X+2)(X^2+1)(X^2+2X+2) \text{ and } v = q^+ = X^3(X^2+4X+5)^2$$

Note that

$$\begin{array}{rcccccccccccc}
\langle u, X\rangle &=& \ldots & 0 & 0 & 2 & \underline{1} & 1 & 1 & 0 & 0 & \ldots \\
\langle u, X^2+1\rangle &=& \ldots & 0 & 0 & 0 & \underline{1} & 1 & 0 & 0 & 0 & \ldots \\
\langle v, X\rangle &=& \ldots & 0 & 0 & 0 & \underline{3} & 0 & 0 & 0 & 0 & \ldots \\
\langle v, X^2+1\rangle &=& \ldots & 0 & 0 & 0 & \underline{0} & 0 & 2 & 0 & 0 & \ldots
\end{array}$$

Step 3 defines $a = a_0 + a_1X + \cdots + a_8X^8$ and $b = b_0 + \cdots + b_6X^6$ and solves the polynomial equation equivalent to

$$\frac{s}{t} = E\frac{a}{u} - \frac{a}{u} + \frac{b}{v}$$

by coefficient comparison. This mainly consists in solving a system of linear equations in the indeterminates a_i's and b_i's.

This way we obtain the following result $\alpha = \Delta\beta + \gamma$,

$$\begin{aligned}
\beta &= \frac{a}{u} = \frac{-1900 - 24428\,X + 25768\,X^2 + \cdots - 13947\,X^7 + 222\,X^8}{43200\,(X^2+1)\,X\,(X-1)^2\,(2+2\,X+X^2)\,(2+X)\,(1+X)} \\
\gamma &= \frac{b}{v} = \frac{24000 - 34250\,X - 56900\,X^2 - 27745\,X^3 - 4612\,X^4 + 37\,X^5}{72000\,(X^2+4\,X+5)^2\,X^3}
\end{aligned}$$

The summation algorithm presented here has been implemented in the Maple language. Contact the second author for the code. Maple procedures for other summation algorithms are contained in Pirastu (1994a) and can be obtained from the first author.

References

Abramov, S. A. (1971), On the summation of rational functions. *Zh. vychisl. mat. Fiz.*, 11:1071 – 1075. English transl. in USSR Comput. Math. Phys.

Abramov, S. A. (1975), Rational component of the solution of a first linear recurrence relation with a rational right hand side. *Zh. vychisl. mat. Fiz.*, 14:1035–1039. English transl. in USSR Comput. Math. Phys.

Geddes, K. O., Czapor, S. R. and Labahn, G. (1992), *Algorithms for Computer Algebra*. Kluwer, Boston.

Gosper, R. W. (1978), Decision procedure for indefinite hypergeometric summation. *Proc. Natl. Acad. Sci. USA*, 75:40 –42.

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

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

Lafon, J. C. (1983), Summation in finite terms. In B. Buchberger, G. E. Collins, and R. Loos, editors, *Computer Algebra Symbolic and Algebraic Computation*, pages 71–77. Springer-Verlag.

Loos, R. (1983), Computing rational zeros of integral polynomials by p-adic expansion. *SIAM J. Computation*, 12(2):286-293.

Man, Y.-K. (1993), On computing closed forms for indefinite summation. *Journal of Symbolic Computation*, 16:355-376.

Man, Y.-K. and Wright, F. J. (1994), Fast polynomial dispersion computation and its application to indefinite summation, In *Proceedings of ISSAC '94*, pages 176–180, ACM press.

Moenck, R. (1977), On computing closed forms for summations. In *Proceedings of MACSYMA user's conference*, pages 225–236. Berkley, California.

Paule, P. (1993), Greatest factorial factorization and symbolic summation I. RISC-Linz Report Series 93-02, J. Kepler University, Linz.

Petkovšek, M. (1992), Hypergeometric solutions of linear recurrences with polynomial coefficients. *Journal of Symbolic Computation*, 14:243–264.

Pirastu, R. (1992), *Algorithmen zur Summation rationaler Funktionen.* Diploma thesis, Univ. Erlangen-Nürnberg. (in german).

Pirastu, R. (1994a), Algorithms for indefinite summation of rational functions in Maple. submitted to *The Maple Technical Newsletter*.

Pirastu, R. (1994b), A note on the minimality problem in indefinite summation of rational functions. Actes du Séminaire Lotharingien de Combinatoire, 31 session, Saint-Nabor, Ottrott, october 1993, J. Zeng, editeur, Publications de l'I.R.M.A. 1994/021, pages 95-101.