J. Symbolic Computation (1996) **11**, 1–000

Greatest Factorial Factorization and Symbolic Summation

PETER PAULE[†]

Research Institute for Symbolic Computation, J. Kepler University, A–4040 Linz, Austria

(Received 19 December 1992)

The greatest factorial factorization (GFF) of a polynomial provides an analogue to square-free factorization but with respect to integer shifts instead to multiplicities. We illustrate the fundamental role of that concept in the context of symbolic summation. Besides a detailed discussion of the basic GFF notions we present a new approach to the indefinite rational summation problem as well as to Gosper's algorithm for summing hypergeometric sequences.

1. Introduction

At present the most general algebraic and algorithmic frame for discussing the problem of indefinite summation is provided by the work of Karr (1981, 1985). His method, working over $\Pi\Sigma$-fields which are certain difference field extensions of a constant field $\mathbf{K}$, can be viewed as a summation analogue to Risch's integration method. A difference field simply is a field $\mathbf{F}$ together with an automorphism σ of $\mathbf{F}$.

Given a, f from a $\Pi\Sigma$-field $\mathbf{F}$, Karr's method constructively decides the existence of a solution $g \in \mathbf{E}$ of $\sigma g - a \cdot g = f$ where $\mathbf{E}$ is a fixed $\Pi\Sigma$-extension field of $\mathbf{F}$; f is called summable (with respect to $\mathbf{E}$) if the equation can be solved in the case $a = 1$. We distinguish two cases: The "telescoping problem", i.e., given $f \in \mathbf{F}$ find $g \in \mathbf{F}$ such that $\sigma g - g = f$, and the "general problem", i.e., given $f \in \mathbf{F}$ determine a $\Pi\Sigma$-extension field $\mathbf{E}$ of $\mathbf{F}$ such that $\sigma g - g = f$ for some $g \in \mathbf{E}$.

Despite the fact that Karr's method algorithmically decides whether a proposed extension $\mathbf{E}$ is a $\Pi\Sigma$-extension, the problem with applying Karr's method for the general case consists in finding an appropriate candidate for $\mathbf{E}$. But in view of his analogue to Liouville's theorem on elementary integrals (Karr, 1985, *RESULT* p. 314) one has the following: If $f \in \mathbf{F}$ is summable in $\mathbf{E}$, then the "interesting" part of it already is summable in $\mathbf{F}$, and the remainder consists of formal sums that have been adjoined to $\mathbf{F}$ in the construction of $\mathbf{E}$. This justifies to consider the telescoping problem separately.

Pointing to rational and hypergeometric summation techniques which do not require complete factorization, Karr (1981, sect. 4.2) raises the question whether similar techniques can be "profitably applied" in his $\Pi\Sigma$-field theory. Despite the fact that the present

[†] Supported in part by grant P7220 of the Austrian FWF. Email: `peter.paule@risc.uni-linz.ac.at`

0747–7171/90/000000 + 00 $03.00/0

paper focuses on rational and hypergeometric summation only, it can be seen as a first step in this direction. This is made more explicit as follows.

As a basic tool for a unified treatment of rational and hypergeometric summation, "greatest factorial factorization" (GFF) of a polynomial in analogy to square-free factorization is introduced. Instead of collecting irreducibles according to their multiplicities, the GFF is obtained by extracting divisors of factorial type $p(x)p(x-1)\ldots p(x-k+1)$ of greatest length. As with square-free factorization the computation of the GFF-form of a polynomial does not require complete factorization. With Karr's theory as a guiding principle in the background, here the following problems are treated:

(i) rational telescoping (Section 4), i.e., $\mathbf{F} = \mathbf{K}(x)$ with $\sigma c = c$ for all $c \in \mathbf{K}$ and $\sigma x = x + 1$;
(ii) hypergeometric telescoping (Section 5), i.e., $\mathbf{F} = \mathbf{K}(x, f)$ with σ acting on $\mathbf{K}(x)$ as before and $\sigma f = rf$ for some fixed $r \in \mathbf{K}(x)$;
(iii) general rational summation (Section 6), i.e., where $\mathbf{F}$ as in (i) and $\mathbf{E}$ has to be determined.

In all of these applications a certain type of polynomial gcd plays a basic role, namely $\gcd(p, \sigma p)$ for $p \in \mathbf{K}[x]$. From the GFF-form of a polynomial p the GFF-form of $\gcd(p, \sigma p)$ can be read off directly (Section 2, Fundamental Lemma), just as the square-free factorization of $\gcd(p, Dp)$, D the derivation operator, from the square-free factorization of p. This fundamental property is used throughout the paper. For treating rational telescoping a new canonical "S-form" representation of rational functions is introduced (Section 3), i.e., a representation as a quotient of two polynomials where the denominator has an especially nice GFF-form. Gosper's algorithm for hypergeometric telescoping finds a new explanation using only basic GFF notions from Section 2, in particular the Fundamental Lemma. It is well known (see Abramov, 1975) that the general rational summation problem can be solved in full generality (in the sense of Karr); in our approach both the GFF and the S-form play a crucial role.

Concerning "profitable applications" in $\Pi\Sigma$-field theory, the GFF approach is flexible enough to carry over to the "q-case" as well; see Paule & Strehl (1995). This corresponds to q-rational and q-hypergeometric summation which is treated in Karr's theory by choosing $\mathbf{F} = \mathbf{K}(x)$ with $\sigma c = c$ for all $c \in \mathbf{K}$ and $\sigma x = qx$ for a fixed $q \in \mathbf{K}$, and $\mathbf{F} = \mathbf{K}(x, f)$ with σ acting on $\mathbf{K}(x)$ as before and $\sigma f = rf$ for some fixed $r \in \mathbf{K}(x)$, respectively. From this fact one might expect that GFF or some suitable generalization could be of some use also for more general aspects of Karr's theory.

This paper is self-contained, no difference field knowledge but only basic facts from algebra are required. In the following we briefly review its sections. Section 2 presents the basic GFF notions, in particular the Fundamental Lemma and an algorithm for computing the GFF-form of a polynomial. In Section 3 we investigate the relation to the dispersion function (Abramov, 1971) and discuss "shift-saturated" polynomials which are polynomials with sufficiently nice GFF-form. Due to lattice properties of $\mathbf{K}[x]$ with respect to gcd, a minimal shift-saturated polynomial $\mathrm{sat}(p)$ can be assigned to each $p \in \mathbf{K}[x]$. The canonical S-form of a rational function is introduced as the quotient of two polynomials with denominator of type $\mathrm{sat}(p)$. In Section 4 rational telescoping is treated; based on S-form representation, Theorem 4.1 explains why factorials rather than powers play the essential role in summation. Section 5 presents a new and algebraically motivated approach to Gosper's algorithm; together with the basic notions of

Section 2 this section can be read independently from the rest of the paper. In Section 6 we consider the general rational summation problem from GFF point of view. Two new algorithms are given. The first one works iteratively similar to the approach sketched by Moenck (1977). His approach is implemented in the computer algebra system Maple to sum rational functions, but due to several gaps in Moenck's original description the Maple algorithm fails on certain rational function inputs as observed by the author of this paper; see Example 6.6. The second algorithm provides an analogue to what is called "Horowitz' Method" or "Hermite-Ostrogradsky Formula" for rational function integration. In addition, discussing minimal-degree answers to the general rational summation problem we present a new Theorem 6.3 which explicitly tells in which way two "minimal solutions" differ.

2. Greatest Factorial Factorization

In this section "greatest factorial factorization" (GFF) of a polynomial is introduced. It is a new canonical form representation which can be viewed as an analogue to square-free factorization. One of the crucial features of GFF is that, analogous to square-free factorization, it can be computed in an iterative manner essentially involving only gcd computations.

2.1. Basic Definitions

By $\mathbf{N}$ we understand the set of all nonnegative integers. We assume all rings or fields to be of *characteristic zero*. It will be convenient to assume $\mathbf{K}$ to be a *field*, especially in the context of indefinite rational or hypergeometric summation. But for a large part of the theory it would suffice to take for $\mathbf{K}$ a suitable ring such that the polynomial ring $\mathbf{K}[x]$ is a unique factorization domain. As usual we shall assume the result of any gcd (greatest common divisor) computation in $\mathbf{K}[x]$ as being normalized to a *monic* polynomial.

By E we denote the shift operator on $\mathbf{K}[x]$, i.e., $(Ep)(x) = p(x+1)$ for any $p \in \mathbf{K}[x]$. The extension of this shift operator to the rational function field $\mathbf{K}(x)$, the quotient field of $\mathbf{K}[x]$, will be also denoted by E.

The polynomial degree of any $p \in \mathbf{K}[x]$, $p \neq 0$, is denoted by $\deg(p)$. We define $\deg(0) := -\infty$.

Definition. For any monic polynomial $p \in K[x]$ and $k \in \mathbf{N}$ the *k-th falling factorial* $[p]^{\underline{k}}$ of p is defined as

$$[p]^{\underline{k}} := \prod_{i=0}^{k-1} E^{-i} p.$$

Note that by the null convention $\prod_{i \in \emptyset} p_i = 1$ we have $[p]^{\underline{0}} = 1$.

This factorial notion, introduced by Moenck (1977), is crucial in the context of a certain polynomial factorization, in the following called *greatest factorial factorization*. It can be viewed as a polynomial extension of the falling factorial notion, introduced usually in the form $(x)^{\underline{k}} = x(x-1)\dots(x-k+1)$; for the notation see, e.g., (Graham *et al.*, 1989).

In the following we often make use of the elementary fact that an integer-shift $E^n\, t$ of an irreducible polynomial $t \in \mathbf{K}[x]$ again is irreducible over $\mathbf{K}$. This fact corresponds to the multiplicative property of the shift operator, i.e., $E^n\,(p \cdot q) = (E^n\, p) \cdot (E^n\, q)$.

Let n be a positive integer and $p \in \mathbf{K}[x]$. Then $\gcd(p, E^n p) \neq 1$ is equivalent to the existence of an irreducible polynomial $t \in \mathbf{K}[x]$ such that $t\ (E^{-n} t) \mid p$. Also equivalent to that is the existence of two roots of p in the splitting field of p over $\mathbf{K}$ at integer distance n.

Similarly, $\gcd(p, Ep, \ldots, E^n p) \neq 1$ is equivalent to the existence of an irreducible polynomial $t \in \mathbf{K}[x]$ such that $[t]^{\underline{n+1}} \mid p$. In this case there exist $n+1$ roots of p in the splitting field of p forming a sequence $\langle \alpha, \alpha+1, \ldots, \alpha+n \rangle$, i.e., $p(\alpha + i) = 0$ for all $i \in \{0, 1, \ldots, n\}$.

If a polynomial has many roots at integer distance, there are many possibilities to rewrite it using factorials.

Example 2.1. *Consider* $p(x) = x^5 + 2\,x^4 - x^3 - 2\,x^2 = (x+2)(x+1)x^2(x-1) \in \mathbf{Q}[x]$, $\mathbf{Q}$ *the rational number field, then* $p(x) = [x]^{\underline{1}}\,[(x+2)x]^{\underline{2}} = [x(x-1)]^{\underline{1}}\,[x+2]^{\underline{3}} = [(x+2)x]^{\underline{1}}\,[x+1]^{\underline{3}} = [x]^{\underline{2}}\,[x+2]^{\underline{3}} = [x]^{\underline{1}}\,[x+2]^{\underline{4}}$, *etc.* □

From all these possibilities the last one which takes care of maximal chains is of particular importance. Intuitively, it can be obtained as follows: One selects irreducible factors of p in such a way that their product, say $q_1(x)\, q_1(x-1) \ldots q_1(x-k+1)$, forms a falling factorial $[q_1]^{\underline{k}}$ of maximal length k. For the remaining irreducible factors of p this procedure is iterated in order to find all k-th falling factorial divisors $[q_1]^{\underline{k}}, [q_2]^{\underline{k}}$, etc., of that type. Then $[q_1 \cdot q_2 \cdot \ldots]^{\underline{k}}$ forms the factorial factor of p of maximal length k. Iterating this procedure one gets a factorization of p in terms of “greatest” factorials factors.

Definition. We say that $\langle p_1, \ldots, p_k \rangle$, $p_i \in \mathbf{K}[x]$, is a *GFF-form* of a monic polynomial $p \in \mathbf{K}[x]$ if the following conditions hold:
(GFF1) $p = [p_1]^{\underline{1}} \cdots [p_k]^{\underline{k}}$,
(GFF2) each p_i is monic, and $k > 0$ implies $\deg(p_k) > 0$,
(GFF3) $i \leq j \Rightarrow \gcd([p_i]^{\underline{i}}, Ep_j) = 1 = \gcd([p_i]^{\underline{i}}, E^{-j} p_j)$.

Note that, due to the null convention, $\langle\rangle$ is the GFF-form of $1 \in \mathbf{K}[x]$. Condition (GFF3) intuitively can be understood as prohibiting “overlaps” of chains that violate length maximality.

The following theorem explicitly states the fact that the GFF-form provides a canonical form. For instance, $\langle x, 1, 1, x+2 \rangle$ is *the* GFF-form of the polynomial p from the example above.

Theorem 2.1. *If* $\langle p_1, \ldots, p_k \rangle$ *and* $\langle q_1, \ldots, q_l \rangle$ *are* GFF-*forms of a monic* $p \in \mathbf{K}[x]$ *then* $k = l$ *and* $p_i = q_i$ *for all* $i \in \{1, \ldots, k\}$.

Proof. The proof proceeds by induction on $\deg(p)$. The case for 0 is obvious. Assume $\deg(p) > 0$. Let $t \in \mathbf{K}[x]$ be an irreducible factor of p_k, which exists by (GFF2), then $t | [q_i]^{\underline{i}}$ for some i by (GFF1). Equivalently, $t | E^{-h} q_i$ with $0 \leq h < i$, and choose i and h so that $i - h$ is maximal. All what we need is to show that $h + k \leq i$, because then $k \leq h + k \leq i \leq l \leq k$, the last inequality by symmetry. This implies $h = 0$, $k = i = l$, $t | q_k$, and the proof is completed by using induction hypothesis on $\langle p_1, \ldots, p_k/t \rangle$ and $\langle q_1, \ldots, q_k/t \rangle$, each with tailing ones removed, which are GFF-forms of $p/[t]^{\underline{k}}$. Now, assume $h+k > i$. Since $[t]^{\underline{k}} | p$ we have $E^{h-i} t | [q_j]^{\underline{j}}$ for some j, or equivalently, $E^{h-i} t | E^{-g} q_j$ with $0 \leq g < j$. If $i \geq j$ then $E^{h-i} t | \gcd([q_j]^{\underline{j}}, E^{-i} q_i)$, violating (GFF3). If $i < j$ and $0 > g - i + 1$, then $E^{h-i+g+1} | \gcd(E^{g-i+1} q_i, E q_j)$, violating the other part of (GFF3) because of $E^{g-i+1}q_i | [q_i]^{\underline{i}}$. Finally, if $i < j$ and $0 \leq g - i + 1$, then $t | E^{i-g-h} q_j$ with $i - h < j + i - g - h$, a contradiction to the maximal choice of $i - h$. □

If $\langle p_1, \ldots, p_k \rangle$ is the GFF-form of a monic $p \in \mathbf{K}[x]$ we sometimes express this fact for short by $\mathrm{GFF}(p) = \langle p_1, \ldots, p_k \rangle$.

2.2. The Fundamental Lemma

As pointed out in the introduction the "gcd-shift", i.e., the gcd of a polynomial p and its shift Ep, plays a basic role in rational and hypergeometric summation. The GFF-concept takes special care of that observation, as made explicit by the following simple but crucial lemma which is a perfect analogue to what one has for square-free factorization.

Lemma 2.1. *("Fundamental Lemma") Given a monic polynomial $p \in \mathbf{K}[x]$ with* GFF-*form $\langle p_1, \ldots, p_k \rangle$. Then*

$$\gcd(p, Ep) = [p_1]^{\underline{0}} \cdots [p_k]^{\underline{k-1}}.$$

Proof. Proceeding by induction on k the case $k = 0$ is trivial. For $k > 0$,

$$\begin{aligned}
&\gcd(p, Ep) = \\
&[p_k]^{\underline{k-1}} \cdot \gcd([p_1]^{\underline{1}} \cdots [p_{k-1}]^{\underline{k-1}} \cdot E^{-k+1} p_k, E([p_1]^{\underline{1}} \cdots [p_{k-1}]^{\underline{k-1}} \cdot p_k)) = \\
&[p_k]^{\underline{k-1}} \cdot \gcd([p_1]^{\underline{1}} \cdots [p_{k-1}]^{\underline{k-1}}, E([p_1]^{\underline{1}} \cdots [p_{k-1}]^{\underline{k-1}})).
\end{aligned}$$

The first equality is obvious, the second is a consequence of (GFF3) because for $i < k$ we have $\gcd([p_i]^{\underline{i}}, Ep_k) = \gcd(E^{-k+1}p_k, E[p_i]^{\underline{i}}) = E(\gcd(E^{-k}p_k, [p_i]^{\underline{i}})) = 1$ together with $\gcd(E^{-k+1}p_k, Ep_k) | \gcd([p_k]^{\underline{k}}, Ep_k) = 1$. The rest follows from applying the induction hypothesis. □

In other words, from the GFF-form of p, i.e., $\mathrm{GFF}(p) = \langle p_1, \ldots, p_k \rangle$ one directly can extract the GFF-form of its "gcd-shift", i.e., $\mathrm{GFF}(\gcd(p, Ep)) = \langle p_2, \ldots, p_k \rangle$.

Example 2.2. *From* $\mathrm{GFF}(p) = \langle x, 1, 1, x + 2 \rangle$ *one immediately gets by Lemma 2.1 that* $\gcd(p, Ep) = [x + 2]^{\underline{3}}$ *and* $\mathrm{GFF}(\gcd(p, Ep)) = \langle 1, 1, x + 2 \rangle$. □

It will be convenient to introduce the following abbreviation for the "gcd-shift":

Definition. Given a monic polynomial $p \in \mathbf{K}[x]$: $\mathrm{gcdE}(p) := \gcd(p, Ep)$.

For various applications that will follow it is useful to keep in mind that dividing p with $\mathrm{GFF}(p) = \langle p_1, \ldots, p_k \rangle$ by $E^{-1}\,\mathrm{gcdE}(p)$ or $\mathrm{gcdE}(p)$ results in separating the product of the first, respectively last, falling factorial entries:

$$\frac{p}{E^{-1}\,\mathrm{gcdE}(p)} = p_1 p_2 \ldots p_k \quad \text{and} \quad \frac{p}{\mathrm{gcdE}(p)} = p_1 (E^{-1} p_2) \ldots (E^{-k+1} p_k).$$

Remark. The analogous lemma used in standard square-free factorization algorithms reads as follows. Let $q = q_1^1\, q_2^2 \ldots q_k^k$ be the square-free factorization of $q \in \mathbf{K}[x]$, i.e., each irreducible factor of q_i arises exactly with multiplicity i in the complete factorization of q in $\mathbf{K}[x]$. Then for the derivation operator D on $\mathbf{K}[x]$:

$$\gcd(q, Dq) = q_1^0\, q_2^1 \cdots q_k^{k-1}.$$

The analogy to the proposition above is made fully transparent by the elementary fact that

$$\gcd(p, Ep) = \gcd(p, \Delta p).$$

One is tempted to view these two different types of representations, related by the operator analogue above, as somewhat "orthogonal" to each other. In a concrete example this statement becomes more transparent. Let $p = x^{14} - 2x^{13} + 4x^{12} - 2x^{11} - 2x^{10} + 10x^9 - 16x^8 + 2x^7 + 5x^6 - 16x^5 + 20x^4 + 8x^3 - 12x^2 \in \mathbf{Q}[x]$ with $p = (x^2 - 2x + 3)(x^3 + 2x)^2(x^2 - 1)^3$ as its square-free factorization. The representation of p according to its GFF-form is

$$p = [(x^2+2)(x^2-1)]^{\underline{1}}\ [x^2+2]^{\underline{2}}\ [(x+1)^2]^{\underline{3}}.$$

Comparing both representations, one observes that the constituents q_i of the square-free factorization violate several GFF-properties, for instance, (GFF3) by $\gcd(x^2 - 1, E^{-2}(x^2 - 1)) = x - 1$. Vice versa, the constituents of the GFF-form need not be relatively prime nor square-free. More information on square-free factorization, for instance, can be found in the book (Geddes *et al.*, 1992). □

2.3. Computing the GFF-Form

One crucial feature of greatest factorial factorization is that, analogous to square-free factorization, it can be obtained *without any factorization* by an iterative procedure essentially involving only gcd computations. As with square-free factorization this goal can be achieved in several ways. Nevertheless, most of these algorithms rely on the Fundamental Lemma. That one we give below uses Lemma 2.1 together with the trivial fact $p = p/\operatorname{gcdE}(p) \cdot \operatorname{gcdE}(p)$. It is especially simple in structure and also verified easily.

Algorithm GFF INPUT: a monic polynomial $p \in \mathbf{K}[x]$; OUTPUT: the GFF-form $\mathrm{GFF}(p)$ of p.

If $p = 1$ then $\mathrm{GFF}(p) := \langle\rangle$.
Otherwise, let $\langle p_2, \ldots, p_k\rangle := \mathrm{GFF}(\operatorname{gcdE}(p))$. Then:

$$\mathrm{GFF}(p) := \langle p/(\operatorname{gcdE}(p)(E^{-1}p_2)\cdots(E^{-k+1}p_k)), p_2, \ldots, p_k\rangle.$$

Remark. (i) To present this method for computing the GFF-form was suggested by one of the referees; another variant, "Algorithm 2" proposed in Paule (1993), requires one more gcd-operation, but only $O(k)$ polynomial operations in comparison to $O(k^2)$ as in Algorithm GFF. Nevertheless empirical tests suggest that still Algorithm GFF is more efficient. As the referee points out, the heuristic explanation is that it is often better to have more operations on smaller polynomials than to have fewer operations on larger ones.

(ii) Another alternative to compute the GFF-form can be derived from the fact that the algorithm of Petkovšek (1992) for computing the Gosper-Petkovšek representation ("GP-form") for rational functions, a normalized version of the G-form representation also described in Section 5, contains the GFF-form computation as a special case; see Lemma 5.2. This also was briefly described in Paule & Strehl (1995). □

3. Shift-Equivalence Classes and Saturation

In this section we first investigate how Abramov's dispersion function is related to GFF. Then we discuss "saturated" polynomials; these are polynomials with sufficiently nice GFF-form. Due to lattice properties one can assign to any monic polynomial p the minimal saturated multiple $\mathrm{sat}(p)$, called "shift-saturation" of p. This gives rise to a new canonical "S-form" representation of rational functions, i.e., as a quotient of polynomials where the denominator is of type $\mathrm{sat}(p)$. As worked out in the following sections, the advantage of using S-forms for rational summation is due to the simple GFF-structure of their denominators. We would like to mention that Strehl (1992) was the first who pointed out the lattice aspects of the GFF.

3.1. Dispersion

As a basic notion for the algorithmic treatment of the rational summation problem, Abramov (1971) defined the *dispersion* $\mathrm{dis}(p)$ of a polynomial $p \in \mathbf{K}[x]$ with $\deg(p) \geq 1$ as $\mathrm{dis}(p) := max\{k \in \mathbf{N} : \gcd(p, E^k p) \neq 1\}$. We find it convenient to extend this definition to nonzero constant polynomials by defining $\mathrm{dis}(p) := 0$ for $p \in \mathbf{K}[x]$ with $\deg(p) = 0$.

Thus $\mathrm{dis}(p) = n$ is equivalent to saying that the maximal integer root-distance $|\alpha - \beta|$, α and β being roots of p in its splitting field over $\mathbf{K}$, is equal to n. For instance, $\mathrm{dis}([x]^{\underline{n}}) = n-1$, or $\mathrm{dis}((x+2)x(x-1)(x-2)) = 4$ where $\mathrm{GFF}((x+2)x(x-1)(x-2)) = \langle x+2, 1, x\rangle$. More precisely, dispersion and greatest factorial factorization are related as follows: Let $p \in \mathbf{K}[x]$ be monic with GFF-form $\langle p_1, \ldots, p_k\rangle$. Suppose that $\mathrm{dis}(p) = n$, then there exists an irreducible polynomial $t \in \mathbf{K}[x]$ such that $t|[p_i]^{\underline{i}}$ and $E^{-n}t|[p_j]^{\underline{j}}$ for some i and j. From the maximality property of n it follows that $t|p_i$ and $E^{-n}t|E^{-j+1}p_j$. Thus only the factor

$$p_1(x)\ p_2(x)\ p_2(x-1)\ p_3(x)\ p_3(x-2) \ldots p_k(x)\ p_k(x-k+1) \tag{3.1}$$

of p contributes to $\mathrm{dis}(p)$.

As a by-product we also obtain that multiplying p by

$$p_1(x+1)\ p_2(x+1) \ldots p_k(x+1)\ \left(= \frac{Ep}{\mathrm{gcdE}(p)}\right) \tag{3.2}$$

increases the dis-function exactly by one. Since the least common multiple $\mathrm{lcm}(p, Ep) = p(Ep)/\,\mathrm{gcdE}(p)$, this means that $\mathrm{dis}(\mathrm{lcm}(p, Ep)) = \mathrm{dis}(p) + 1$.

At this place we introduce an obvious but useful lemma.

Lemma 3.1. *Let $p \in \mathbf{K}[x]$ be a monic polynomial with* GFF*-form $\langle p_1, \ldots, p_k\rangle$ then $\mathrm{lcm}(p, Ep)$ has* GFF*-form $\langle 1, Ep_1, \ldots, Ep_k\rangle$.*

Proof. From (3.2) it is immediate that $\mathrm{lcm}(p, Ep) = [Ep_1]^{\underline{2}} \ldots [Ep_k]^{\underline{k+1}}$. The easy check of (GFF2) and (GFF3) completes the proof. □

As with gcdE, it will be convenient to introduce the corresponding abbreviation with respect to the "lcm-shift":

Definition. Given a monic $p \in \mathbf{K}[x]$: $\mathrm{lcmE}(p) := \mathrm{lcm}(p, Ep)$.

The dispersion statistics can be extended to rational functions as follows. For *relatively prime* polynomials $a, b \in \mathbf{K}[x]$: $\mathrm{dis}(a/b) := \mathrm{dis}(b)$.

This extension, only depending on the denominator b, is justified by the following proposition which is due to Abramov (1971):

Proposition 3.1. *For relatively prime polynomials $a, b \in \mathbf{K}[x]$ with $\deg(b) \geq 1$:*

$$\operatorname{dis}(\Delta\ \frac{a}{b}) = \operatorname{dis}(\frac{a}{b}) + 1.$$

Proof. We give a proof, different from Abramov's original one, using the GFF concept. In view of $\Delta(a/b) = ((Ea) \cdot b - a \cdot (Eb))/(b \cdot (Eb))$ we define $B := \operatorname{gcdE}(b)$. Then

$$\begin{aligned} &\gcd((Ea) \cdot b - a \cdot (Eb), b \cdot (Eb)) = \\ &B \cdot \gcd((Ea) \cdot b/B - a \cdot (Eb)/B, \operatorname{lcmE}(b)) = \\ &B \cdot \gcd((Ea) \cdot b/B - a \cdot (Eb)/B, \operatorname{gcdE}(b)), \end{aligned} \tag{3.3}$$

where the last line follows from $\gcd(a, b) = 1$. Denote the gcd on the right hand side of (3.3) by B_0, then

$$\operatorname{dis}(\Delta\frac{a}{b}) = \operatorname{dis}(\frac{b}{B_0} \cdot \frac{Eb}{B}) = \operatorname{dis}(\frac{b}{B} \cdot \frac{Eb}{B} \cdot \frac{B}{B_0}) = \operatorname{dis}(\operatorname{lcmE}(b)) = 1 + \operatorname{dis}(b)$$

where the equation before last follows from the observations related to (3.1) and (3.2). □

We want to note that the dis-function on rational functions works "opposite" to the deg-function on polynomials in connection with the Δ operator. The same applies if deg is extended to rational functions as, for instance, in (Karr, 1981): For $a, b \in \mathbf{K}[x]$, $b \neq 0$, define $\deg(a/b) := \deg(a) - \deg(b)$. Evidently deg is well-defined on $\mathbf{K}(x)$, i.e., if $a/b = c/d$ for $a, b, c, d \in \mathbf{K}[x]$ then $\deg(a/b) = \deg(c/d)$.

Proposition 3.2. *For nonzero $a, b \in \mathbf{K}[x]$ with $\deg(a/b) \neq 0$:*

$$\deg(\Delta\frac{a}{b}) = \deg(\frac{a}{b}) - 1.$$

Proof. In $\Delta(a/b) \cdot b(Eb) = (Ea)b - a(Eb)$ rewrite the right hand side as $(\Delta a)b - a(\Delta b)$. For any nonzero $p \in \mathbf{K}[x]$ we have $\operatorname{lcf}(\Delta p) = \deg(p) \cdot \operatorname{lcf}(p)$, thus the leading coefficients $\operatorname{lcf}((\Delta a)b)$ and $\operatorname{lcf}(a(\Delta b))$ are equal iff $\deg(a) = \deg(b)$. Hence, if $\deg(a/b) \neq 0$ then $\deg((\Delta a)b - a(\Delta b)) = \deg(a) + \deg(b) - 1$, and comparison to the degree of $\Delta(a/b) \cdot b(Eb)$ completes the proof. □

Proposition 3.1, for instance, gives a simple criterion, due to Abramov (1971), whether a given rational function is rational summable:

Proposition 3.3. *Let $a, b \in \mathbf{K}[x]$ be relatively prime with $\deg(b) \geq 1$ and $\operatorname{dis}(b) = 0$. Then there exists no rational function solution $s \in \mathbf{K}(x)$ of the equation $\Delta s = a/b$.*

Proof. For any rational function $s \in \mathbf{K}(x)$ we have $\operatorname{dis}(\Delta s) = 1 + \operatorname{dis}(s) \geq 1$, by Proposition 3.1. This contradicts $\operatorname{dis}(a/b) = \operatorname{dis}(b) = 0$. □

Example 3.1. *It is well-known that the sequence of* harmonic *numbers of order α, $H_n^{(\alpha)} := \sum_{k=1}^{n} 1/k^\alpha$, $\alpha \in \mathbf{N}\backslash\{0\}$, is not rational. By Proposition 3.3 this can be seen quickly as follows. Suppose $H_n^{(\alpha)} = r(n)$ for some $r \in \mathbf{K}(x)$. Thus, $r(n+1) - r(n) = 1/(n+1)^\alpha$ for all integers $n \geq 1$, and hence, as an identity in $\mathbf{K}(x)$: $r(x+1) - r(x) = 1/(x+1)^\alpha$.*

But $\mathrm{dis}(1/(x+1)^\alpha) = 0$, *a contradiction to* $r \in \mathbf{K}(x)$. *- Note that* $(H_n^{(1)})_{n\geq 1} = (H_n)_{n\geq 1}$.
□

3.2. Saturated Polynomials

A certain type of polynomials which play a basic role in rational summation has a sufficiently nice GFF-form, i.e, the GFF-constituents are relatively prime and their factors do not differ by any integer-shift. For studying these polynomials, which will be called "saturated", the following equivalence relation on $\mathbf{K}[x]$ plays a fundamental role. In fact, it is a special case of definition 13 of (Karr, 1981) for monic polynomials.

Definition. Two polynomials $p_1, p_2 \in \mathbf{K}[x]$ are said to be *shift-equivalent* if there exists an integer k such that $p_2(x) = p_1(x+k)$.

It is easily checked that this indeed defines an equivalence relation on $\mathbf{K}[x]$.

Example 3.2. *For* $p(x) = x^2(x-1/3)(x-1/2)(x-1)(x-7/3)^3(x-3)^2 \in \mathbf{Q}[x]$ *the set of monic irreducible factors splits into three equivalence classes, namely* $F_1 = \{x - 1/2\}, F_2 = \{x-1/3, x-7/3\}, F_3 = \{x, x-1, x-3\}$. □

Let p_1, p_2 be shift-equivalent irreducible factors of p, i.e., $p_2(x) = p_1(x-k)$ for some integer k. Defining

$$p_1 > p_2 \;:\Leftrightarrow\; k > 0$$

imposes a total order, which we shall call the *shift-order*, on the elements of each shift-equivalence class.

Example 3.3. *In the example above, according to shift-order for the elements of* F_3 *we have:* $x > x-1 > x-3$. □

We introduce a canonical choice of representatives $\mathrm{ShiftEq}(p)$ of the shift-equivalence classes of the monic irreducible factors of p by choosing from each class the *maximal* element with respect to the shift-order.

For monic $p \in \mathbf{K}[x]$ let $\mathrm{ShiftEq}(p) = \{p_1, p_2, \ldots, p_n\}$ be this uniquely determined set of representatives. For each p_i-class let q_i denote the *minimal* element. Filling up "shift-gaps" by multiplying extra irreducibles $t \in \mathbf{K}[x]$ with $q_i \leq t \leq p_i$, $i \in \{1, \ldots, n\}$, amounts to gluing together factorial chains.

Example 3.4. *The* GFF*-form of* p *from Example 3.2 is*

$$\langle x(x-1/3)(x-1/2)(x-7/3)^3(x-3)^2, x\rangle$$

Let $p_1(x) := p(x)\cdot(x-4/3)(x-2)$ *and* $p_2(x) := p_1(x)\cdot(x-1/3)^2(x-4/3)^2(x-1)(x-2)$, *then for the* GFF*-forms we have,*

$$\mathrm{GFF}(p_1) = \langle x(x-1/2)(x-7/3)^2(x-3), 1, x-1/3, x\rangle,$$
$$\mathrm{GFF}(p_2) = \langle x-1/2, 1, (x-1/3)^3, x^2\rangle.$$

□

In the example above the constituents of $\mathrm{GFF}(p)$ and $\mathrm{GFF}(p_1)$ neither are relatively prime nor belong their factors to different shift-equivalence classes. By multiplication of further extra factors this property, being crucial for rational summation, is achieved for GFF(p_2). This gives rise to the following definition:

Definition. Let $p \in \mathbf{K}[x]$ be monic with GFF-form $\langle p_1, \dots, p_k \rangle$. Then p is called *shift-saturated* (for short: *saturated*) if $\gcd(p_i, E^h p_j) \neq 1$ implies $i = j$ and $h = 0$.

Example 3.5. *The polynomial $p_2(x)$ from Example 3.4 is saturated. It is a divisor of $p_3(x) := [(x(x-1/3)(x-1/2))^3]^{\underline{4}}$ which is also saturated.* □

Saturatedness is invariant under the gcd operation:

Proposition 3.4. *The* gcd *of two saturated polynomials is saturated.*

Proof. Immediate consequence of Lemma 3.2 below. □

There are several proofs of Proposition 3.4. For instance, it is an immediate consequence of the following lemma, which we shall also use later, describing *minimal* reduction steps.

Lemma 3.2. *Given non-constant, monic and saturated polynomials $p, q \in \mathbf{K}[x]$ with* GFF*-forms $\langle p_1, \dots, p_k \rangle$ and $\langle q_1, \dots, q_l \rangle$, respectively, where $k \leq l$. If $k = l$ assume $p_k \neq q_k$ and $\deg(p_k) \leq deg(q_k)$. Then there exists a monic and saturated $q' \in \mathbf{K}[x]$ such that*

$$\gcd(p, q') = \gcd(p, q) \quad \textit{and} \quad \deg(q') < \deg(q). \tag{3.4}$$

In addition, there exists a monic divisor r of q_l with $\deg(r) > 0$ which determines q' in one of the following ways:
(a) $q' = q/[r]^{\underline{l}}$,
(b) $\gcd(r, q_l/r) = 1$ and $q' = q/r$,
(c) $\gcd(r, q_l/r) = 1$ and $q' = q/E^{-l+1}r$.

Proof. If $\gcd([p_i]^{\underline{i}}, [q_l]^{\underline{l}}) = 1$ for all $i \in \{1, \dots, k\}$, then $q' := q/[q_l]^{\underline{l}}$ with GFF-form $\langle q_1, \dots, q_{l-1} \rangle$ is monic and saturated, and also (3.4) holds.

Suppose $\gcd([p_i]^{\underline{i}}, [q_l]^{\underline{l}}) \neq 1$. This is equivalent to $d := \gcd(E^{-\alpha} p_i, q_l) \neq 1$ where $-l < \alpha < i$. Choose r as the maximal divisor of q_l containing only irreducible factors t of d, a choice implying $\gcd(r, q_l/r) = 1$. *(i)* If $\alpha \in \{-l+1, \dots, -1\}$ these irreducibles t cannot be in $\gcd(p, q)$, because then $t | E^{-\alpha} p_i$ and $t | E^{-\beta} p_j$ for some $\beta \geq 0$ and some j violates saturatedness of p. In this case $q' := q/r$ with GFF-form $\langle q_1, \dots, q_{l-1} \cdot Er, q_l/r \rangle$, with trailing 1's dropped, is monic, saturated and satisfies (3.4). *(ii)* If $\alpha \in \{0, \dots, i-1\}$ the irreducibles $E^{-l+1}t$ cannot be in $\gcd(p, q)$, because then $E^{-l+1}t | E^{-\alpha-l+1} p_i$ and $E^{-l+1} t | E^{-\beta} p_j$ with $0 \leq \beta < j \leq k$. *Assuming $k < l$*, this violates saturatedness of p, since then $1 \leq (\alpha + l - 1) - \beta$. In this case $q' := q/E^{-l+1}r$ with GFF-form $\langle q_1, \dots, q_{l-1} \cdot r, q_l/r \rangle$, with trailing 1's dropped, is monic, saturated and satisfies (3.4).

As we have seen, argument *(ii)* works only if we assume $k < l$. *If $k = l$* this argument only would fail if $\beta = k - 1$, which means $j = k$ and thus $t | p_k$ and $t | \gcd(E^{-\alpha} p_i, q_k)$. Because of saturatedness of p this implies $\alpha = 0$ and $i = k$, and therefore $t | \gcd(p_k, q_k)$. Consequently, if $k = l$ we assume that $s_0 := \gcd(p_k, q_k) \neq 1$, otherwise we are done as in (i) or (ii) above. Let s be the maximal divisor of q_k containing only irreducibles of p_k. *(i')*

If $s = q_k$ then for $r := q_k/s_0$ we have $\deg(r) > 0$ because of the degree assumption on p_k and q_k. In this case $q' := q/r$ with GFF-form $\langle q_1, \ldots, q_{l-1} \cdot Er, q_k/r\rangle$ is monic, saturated and satisfies (3.4). *(ii')* In the case $s \neq q_k$ two possibilities arise. If $\gcd([p_i]^{\underline{i}}, q_k/s) = 1$ for all $i \in \{1, \ldots, k-1\}$ then $r := q_k/s$ and $q' := q/[r]^{\underline{k}}$, etc. If $\gcd([p_i]^{\underline{i}}, q_k/s) \neq 1$ for some $i \in \{1, \ldots, k-1\}$ then compute r as the maximal divisor of q_k/s as in the cases (i) and (ii) above. □

3.3. Shift-Saturation

Evidently, the subset of monic and saturated multiples of p of the lattice of polynomials from $\mathbf{K}[x]$, ordered by divisibility, has exactly one minimal element. This allows to assign to any monic polynomial a unique multiple which is saturated and thus equipped with "nice" GFF-form:

Definition. Given monic $p \in \mathbf{K}[x]$ then the *shift-saturation* $\mathrm{sat}(p)$ of p is defined as the monic, saturated polynomial from $\mathbf{K}[x]$ of lowest degree that is divisible by p.

Using maximal and minimal elements of shift-equivalence classes allows a more explicit description of $\mathrm{sat}(p)$. Let $\mathrm{ShiftEq}(p) = \{p_1, \ldots, p_n\}$ with q_i the minimal elements of the p_i-classes, as above. Then $p_i(x) = q_i(x + k_i)$ for some $k_i \in \mathbf{N}$. For each $i \in \{1, \ldots, n\}$ we define the *length* of the p_i-class as $l(p_i) := k_i + 1$. Let $\mathrm{mult}(p_i)$ be the maximum of the multiplicities of all irreducibles contained in the p_i-class. Now it is easily checked, for instance, using Lemma 3.2 that the shift-saturation of p is the polynomial

$$\mathrm{sat}(p) = [p_1^{\mathrm{mult}(p_1)}]^{\underline{l(p_1)}} \cdots [p_n^{\mathrm{mult}(p_n)}]^{\underline{l(p_n)}}. \tag{3.5}$$

From this representation the GFF-form of $\mathrm{sat}(p)$ almost directly can be read off. One just has to merge factorials according to

$$[p_i^{\mathrm{mult}(p_i)}]^{\underline{l}}\ [p_j^{\mathrm{mult}(p_j)}]^{\underline{l}} = [p_i^{\mathrm{mult}(p_i)} p_j^{\mathrm{mult}(p_j)}]^{\underline{l}}$$

in case $l = l(p_i) = l(p_j)$ for $i \neq j$, to reorder the factorials involved, and to insert 1's corresponding to trivial factorials of the form $[1]^{\underline{l}}$.

Example 3.6. *Again we take the polynomial p from Example 3.2, then:*
$\mathrm{ShiftEq}(p) = \{x - 1/2, x - 1/3, x\}$, $\mathrm{mult}(x - 1/2) = 1$, $\mathrm{mult}(x - 1/3) = 3$, $\mathrm{mult}(x) = 2$, *and* $l(x - 1/2) = 1$, $l(x - 1/3) = 3$, $l(x) = 4$. *Therefore,*

$$\mathrm{sat}(p) = [x - 1/2]^{\underline{1}}\ [(x - 1/3)^3]^{\underline{3}}\ [x^2]^{\underline{4}} = p_2.$$

with GFF-*form* $\langle x - 1/2, 1, (x - 1/3)^3, x^2\rangle$. □

Besides the maximal and minimal elements, the *multiplicity element* m_i of a shift-equivalence class plays a distinguished role, a fact which was pointed out by Pirastu (1992). It is defined as the smallest, with respect to shift-order, irreducible m_i in the p_i-class such that $m_i^{\mathrm{mult}(p_i)} | p$.

As a lemma we state a gcd property used in connection with Theorem 4.1, Section 4.

Lemma 3.3. *Let* $p \in \mathbf{K}[x]$ *be monic and* m *the multiplicity element from some shift-equivalence class of* p*, then* $\gcd(m, p/\gcd\mathrm{E}(p)) \neq 1$.

Proof. Assume $m^\mu | \operatorname{gcdE}(p)$ with $\mu = mult(p_i)$ if m belongs to the p_i-shift-equivalence class of p. Then $E^{-1}m^\mu | p$, a contradiction to the definition of m. □

The proof of the following lemma is left to the reader:

Lemma 3.4. *Given monic $p, t \in \mathbf{K}[x]$ such that $t|p$ and t does not cancel the maximal, minimal or multiplicity element in any shift-equivalence class of p. Then $\operatorname{sat}(p/t) = \operatorname{sat}(p)$.*

Proof. Obvious from description (3.5). □

From description (3.5) also the following is obvious:

Lemma 3.5. *Let $p \in \mathbf{K}[x]$ be monic with $\deg(p) \geq 1$ and let $\langle s_1, \ldots, s_k \rangle$ be the GFF-form of $\operatorname{sat}(p)$, then:*
(a) $\operatorname{dis}(p) = 0 \Rightarrow k = 1$ *and* $s_1 = p = \operatorname{sat}(p)$,
(b) $\operatorname{dis}(p) \geq 1 \Rightarrow \deg(s_1 \cdots s_k) < \deg(p)$.

Proof. Obvious from description (3.5). □

Shift-saturation commutes with the shift operator:

Lemma 3.6. *For all monic $p \in \mathbf{K}[x]$: $\operatorname{sat}(Ep) = E \operatorname{sat}(p)$.*

Proof. Let $\langle p_1, \ldots, p_k \rangle$ be the GFF-form of $\operatorname{sat}(p)$. Then $\langle Ep_1, \ldots, Ep_k \rangle$ ist the GFF-form of $E \operatorname{sat}(p)$ which evidently is monic and saturated. By definition, $\operatorname{sat}(p)$ is the saturated polynomial of lowest degree divisible by p, thus $E \operatorname{sat}(p)$ is the saturated polynomial of lowest degree divisible by Ep. Hence, $E \operatorname{sat}(p) = \operatorname{sat}(Ep)$. □

Shift-Saturation also commutes with the "lcm-shift":

Lemma 3.7. *For all monic $p \in \mathbf{K}[x]$: $\operatorname{sat}(\operatorname{lcmE}(p)) = \operatorname{lcmE}(\operatorname{sat}(p))$.*

Proof. Clearly $\operatorname{sat}(p)$ and $\operatorname{sat}(Ep)$ divide $\operatorname{sat}(\operatorname{lcmE}(p))$, thus by Lemma 3.6 we have $\operatorname{lcmE}(\operatorname{sat}(p)) = \operatorname{lcm}(\operatorname{sat}(p), \operatorname{sat}(Ep)) | \operatorname{sat}(\operatorname{lcmE}(p))$. On the other hand, let $\langle p_1, \ldots, p_k \rangle$ be the GFF-form of $\operatorname{sat}(p)$, then $\langle 1, Ep_1, \ldots, Ep_k \rangle$ by Lemma 3.1 is the GFF-form of $\operatorname{lcmE}(\operatorname{sat}(p))$ which is monic, saturated and divisible by p and Ep. Hence, $\operatorname{sat}(\operatorname{lcmE}(p))$ divides $\operatorname{lcmE}(\operatorname{sat}(p))$. □

It is important to note that $\operatorname{sat}(p)$, like the GFF-form, can be computed using *only* gcd computations, i.e., a procedure for a complete factorization of p is not required. Such an algorithm is given in (Pirastu, 1992) or in (Paule, 1993). The first algorithm is more efficient because the latter unnecessarily uses square-free factorization.

3.4. S-forms of Rational Functions

A typical feature of symbolic summation is that for different purposes different representations of rational functions are more appropriate. In order to avoid the repetition of lengthy specifications we define different types of representations as certain *forms*. First we consider the usual canonical form, i.e., the quotient of two relatively prime polynomials.

DEFINITION. The pair $\langle c, d\rangle$, $c, d \in \mathbf{K}[x]$, is called the *reduced form* of $s \in \mathbf{K}(x)$ if $s = c/d$, d monic, and $\gcd(c, d) = 1$.

Shift-saturation gives rise to another type of rational function representation.

DEFINITION. The pair $\langle \gamma, \delta\rangle$, $\gamma, \delta \in \mathbf{K}[x]$, is called a *saturated representation* for $s \in \mathbf{K}(x)$ if $s = \gamma/\delta$ and δ is monic and saturated.

DEFINITION. A saturated representation $\langle \gamma, \delta\rangle$ for $s \in \mathbf{K}(x)$ is called a *saturated form* (for short: *S-form*) for s if δ has minimal degree among all saturated representations of s.

The next proposition states that S-forms are canonical forms for rational functions. Because of the nice properties of the GFF-constituents of the denominators, S-form representation is tailored for rational summation application.

PROPOSITION 3.5. *For any $s \in \mathbf{K}(x)$ with reduced form $\langle c, d\rangle$ there exists a unique S-form which is $\langle c \cdot \mathrm{sat}(d)/d, \mathrm{sat}(d)\rangle$.*

PROOF. $\langle c \cdot \mathrm{sat}(d)/d, \mathrm{sat}(d)\rangle$ certainly is a saturated representation for s. Let $\langle \gamma, \delta\rangle$ be another saturated representation for s with $\deg(\delta) \leq \deg(\mathrm{sat}(d))$. From $c \cdot \delta = \gamma \cdot d$ we have $d|\delta$, hence $\mathrm{sat}(d)|\delta$ by the definition of shift-saturation, and therefore $\delta = \mathrm{sat}(d)$. □

The following proposition, which we shall use later, implicitly tells how a saturated representation can be reduced to S-form.

PROPOSITION 3.6. *Let $\langle \gamma, \delta\rangle$, $\gamma, \delta \in \mathbf{K}[x]$, be a saturated representation (for γ/δ) with $\mathrm{GFF}(\delta) = \langle p_1, \ldots, p_k\rangle$. Then $\langle \gamma, \delta\rangle$ is in S-form if and only if there exist no i and $r \in \mathbf{K}[x]$ with $\deg(r) > 0$ such that (i) $r|p_i$, and one of the following:*
(iia) $[r]^{\underline{i}}|\gamma$,
(iib) $\gcd(r, p_i/r) = 1$ *and* $r|\gamma$,
(iic) $\gcd(r, p_i/r) = 1$ *and* $E^{-i+1}r|\gamma$.

PROOF. If δ is not minimal, then from $\gamma/\delta = (c \cdot \mathrm{sat}(d)/d)/\mathrm{sat}(d)$ as in the proof above we have $\mathrm{sat}(d) = \gcd(\mathrm{sat}(d), \delta)$ with $\deg(\mathrm{sat}(d)) < \deg(\delta)$. Thus δ can be reduced by one of the factors listed in Lemma 3.2, and also γ must be divisible by the same factor. □

4. Rational Telescoping

In this section we discuss the difference equation

$$s(x + 1) - s(x) = r(x) \tag{4.1}$$

over the rational function field $\mathbf{K}(x)$. We call (4.1) the *telescoping equation* for $r \in \mathbf{K}(x)$.

Example 4.1. *If* $r = 1/(x^2 - x - 3/4)$ *then* $s = (-x+1)/(x^2 - 2x + 3/4)$ *is a solution of (4.1) and due to telescoping:*

$$\sum_{k=1}^{n} r(k) = \sum_{k=1}^{n} (s(k+1) - s(k)) = s(n+1) - s(1) = -\frac{n}{n^2 - 1/4}.$$

□

It is well-known that the telescoping equation for any $p \in \mathbf{K}[x]$ finds a polynomial solution. Thus, by splitting off the polynomial part, one can restrict solving (4.1) to given $r \in \mathbf{K}(x)$ with $\deg(r) < 0$. Such rational functions usually are called *proper*.

Assume that for proper $r \in \mathbf{K}(x)$ there exists a rational solution $s \in \mathbf{K}(x)$ of $\Delta s = r$. Such an r will be called *rational summable*. If $\deg(s) \neq 0$ then, by Proposition 3.2, $\deg(s) = deg(r) + 1 \leq 0$. If $\deg(s) = 0$ then $s = \bar{s} + c$ where $\bar{s} \in \mathbf{K}(x)$ is proper and $c \in \mathbf{K} \setminus \{0\}$. In this case $r = \Delta s = \Delta \bar{s}$. This implies the well-known fact that given a proper rational summable $r \in \mathbf{K}(x)$ the telescoping equation $\Delta s = r$ always has a proper solution $s \in \mathbf{K}(x)$. It is obvious that two rational solutions must differ by adding a constant, thus all other rational solutions are improper of degree 0.

In this context several questions naturally arise, for example:

(a) Are there simple criteria to decide whether a given proper rational function is rational summable?
(b) How to treat situations, as $r(x) = 1/x$, where no rational solution of $\Delta s = r$ exists.
(c) How to compute $s \in \mathbf{K}(x)$ such that $\Delta s = r$ for a given rational summable $r \in \mathbf{K}(x)$?

Concerning question (a), certainly all algorithms treating rational summation give a practical answer to this question. We especially point to the difference field approach by Karr (1981) which provides the most general theoretic and algorithmic setting. A very elementary, well-known and useful criterion follows directly from the degree reasoning above:

Lemma 4.1. *If* $r \in \mathbf{K}(x)$ *is proper with* $\deg(r) = -1$*, then* r *is not rational summable.*

Another practical criterion, for instance, is Proposition 3.3 due to Abramov. In the following we shall add some more criteria, one (Proposition 4.2) includes Abramov's result as a special case.

To our knowledge the first answer to question (b) has been given by Abramov (1975) in an algorithmic way. The general machinery of Karr (1981) in principle can deal with the problem but, as explained in the introduction, before running the algorithm one has to supply appropriate information about the difference field extension in which the solution is expected. Moenck (1977) sketched an algorithm working analogously to that called Hermite-iteration for rational function integration. His approach is taken by the computer algebra system Maple to sum rational functions. Due to several gaps in Moenck's paper, observed by the author of this article, the Maple algorithm is unable to treat *arbitrary* rational function input. An example for that is given below in Example 6.6. The entire problem, viewed in the light of shift-saturation, will be discussed in Section 6. That section also contains two new algorithms solving $\Delta s = r$ in general, i.e., also for given non-rational summable $r \in \mathbf{K}(x)$. One of those can be considered as an analogue to what is called Horowitz' method for rational function integration. Pirastu (1992) closes the gaps in Moenck's paper and discusses the relation of the above mentioned algorithms according to implementations carried out by himself in Maple.

All algorithms mentioned above include an answer to question (c) as a special case. Due to the fact that any rational sequence is *hypergeometric*, also Gosper's algorithm could be applied in order to answer (c); see Section 5. In this section we present a new approach (Theorem 4.1) which provides an algorithmic solution as well as an algebraic explanation along the concept of shift-saturation. In Section 5 we briefly discuss its relation to Gosper's algorithm working on rational function inputs from GFF point of view.

4.1. Telescoping via S-forms

One of the crucial observations is a simple explicit connection between the S-form of a rational function $s \in \mathbf{K}(x)$ and the S-form of its difference Δs. Before making this explicit in Proposition 4.1, we first state a lemma for technical reasons.

Lemma 4.2. *Given $a, b, c, p \in \mathbf{K}[x]$ such that $a \cdot Ep + b \cdot p = c$. Let $t \in \mathbf{K}[x]$ be an irreducible divisor of Ep, then for $k \geq 1$:*

$$[t]^{\underline{k}}|c \quad \textit{and} \quad \gcd([t]^{\underline{k}}, a) = 1 \Rightarrow [t]^{\underline{k}}|Ep.$$

Proof. Using induction on k the case $k = 1$ is trivial. Thus assume that $[t]^{\underline{k+1}}|c$ and $\gcd([t]^{\underline{k+1}}, a) = 1$, then $[t]^{\underline{k}}|c$ and $\gcd([t]^{\underline{k}}, a) = 1$, hence $[t]^{\underline{k}}|Ep$ by the induction hypothesis. From $E^{-k}t|c$, $E^{-k}t|p$, and $\gcd(E^{-k}t, a) = 1$ we have $E^{-k}t|Ep$ which completes the proof of $[t]^{\underline{k+1}}|Ep$. □

Now we are ready for the proposition relating the S-form of s to that of Δs.

Proposition 4.1. *For $\gamma, \delta \in \mathbf{K}[x]$ let $\langle \gamma, \delta \rangle$ be in S-form, then the S-form of $\Delta(\gamma/\delta)$ is $\langle \alpha, \beta \rangle$ where*

$$\alpha = \frac{\delta}{\mathrm{gcdE}(\delta)} \cdot E\gamma - \frac{E\delta}{\mathrm{gcdE}(\delta)} \cdot \gamma \quad \textit{and} \quad \beta = \mathrm{lcmE}(\delta).$$

Proof. It will be convenient to define $\delta_i := (E^i \delta)/\operatorname{gcdE}(\delta)$ for $i \in \{0, 1\}$. Let $\langle p_1, \ldots, p_k \rangle$ be the GFF-form of δ then, by Lemma 3.1, $\langle 1, Ep_1, \ldots, Ep_k \rangle$ is the GFF-form of β which is monic and saturated. Thus we have $\Delta(\gamma/\delta) = \alpha/\beta$ where $\langle \alpha, \beta \rangle$ is a saturated representation. Assume that $\langle \alpha, \beta \rangle$ is not in S-form. Then there exist i and irreducible $r \in \mathbf{K}[x]$ such that $r|Ep_i$ and satisfying one of the possibilities of Proposition 3.6. *(i)* Assume $[r]^{\underline{i+1}}|\alpha$. Clearly, $[r]^{\underline{i}}|\alpha$ and $\gcd([r]^{\underline{i}}, \delta_0) = 1$ because of $[r]^{\underline{i}}|[Ep_i]^{\underline{i}}$ and saturatedness of δ. From $r|\alpha$ and $r|\delta_1$ we have $r|(E\gamma) \cdot \delta_0$ and thus $r|E\gamma$. Applying Lemma 4.2 implies $[E^{-1}r]^{\underline{i}}|\gamma$, together with $E^{-1}r|p_i$ according to Proposition 3.6 a contradiction to $\langle \gamma, \delta \rangle$ in S-form. *(ii)* Assume $\gcd(r, Ep_i/r) = 1$ and $r|\alpha$. As in (i) we have $E^{-1}r|\gamma$ which together with $E^{-1}r|p_i$ and $\gcd(E^{-1}r, p_i/E^{-1}r) = 1$ contradicts the S-form of $\langle \gamma, \delta \rangle$. *(iii)* Assume $\gcd(r, Ep_i/r) = 1$ and $E^{-i}r|\alpha$. Because of $E^{-i}r|E^{-i+1}p_i|\delta_0$ and $\gcd(E^{-i}r, \delta_1) = 1$ we have $E^{-i+1}(E^{-1}r)|\gamma$ which together with $E^{-1}r|p_i$ and $\gcd(E^{-1}r, p_i/E^{-1}r) = 1$ contradicts the S-form of $\langle \gamma, \delta \rangle$. □

For the proof of the main result of this section, Theorem 4.1, we need the following inverse relation lemma:

LEMMA 4.3. *(a) For monic* $p \in \mathbf{K}[x]$ *with* GFF*-form* $\langle 1, p_2, \ldots, p_n \rangle$*:*

$$q = E^{-1}\,\mathrm{gcdE}(p) \Rightarrow p = \mathrm{lcmE}(q).$$

(b) For monic $q \in K[x]$*:*

$$p = \mathrm{lcmE}(q) \Rightarrow q = E^{-1}\,\mathrm{gcdE}(p).$$

PROOF. (a) By the Fundamental Lemma $q = [E^{-1}p_2]^{\underline{1}}\,[E^{-1}p_3]^{\underline{2}} \ldots [E^{-1}p_n]^{\underline{n-1}}$ which gives, by Lemma 3.1, $\mathrm{lcmE}(q) = [p_2]^{\underline{2}}\,[p_3]^{\underline{3}} \ldots [p_n]^{\underline{n}} = p$.
(b) Let $q = [q_1]^{\underline{1}}\,[q_2]^{\underline{2}} \ldots [q_m]^{\underline{m}}$ be the GFF-form of q. By Lemma 3.1 and the Fundamental Lemma, $p = [Eq_1]^{\underline{2}}\,[Eq_2]^{\underline{3}} \ldots [Eq_m]^{\underline{m+1}}$ and thus $E^{-1}\,\mathrm{gcdE}(p) = [q_1]^{\underline{1}}\,[q_2]^{\underline{2}} \ldots [q_m]^{\underline{m}}$ which is q. □

Now we are ready for the announced theorem which in the context of shift-saturation describes the solution of the difference equation $\Delta s = r$ over $\mathbf{K}(x)$ in terms of a solution of a difference equation over $\mathbf{K}[x]$. In addition, it provides an algorithm for deciding the existence of a rational solution s as well as for the computation of s if the answer is positive. How the theorem is related to the rational instance of Gosper's algorithm will be discussed in the next section.

THEOREM 4.1. *Let* $r, s \in \mathbf{K}(x)$ *both be proper and with S-forms* $\langle \alpha, \beta \rangle$ *and* $\langle \gamma, \delta \rangle$*, respectively. If* $\Delta s = r$*, then*

$$\delta = E^{-1}\,\mathrm{gcdE}(\beta). \tag{4.2}$$

and γ *is a polynomial solution of*

$$\alpha = \frac{\beta}{\mathrm{gcdE}(\beta)} \cdot E\gamma - \frac{\beta}{E^{-1}\,\mathrm{gcdE}(\beta)} \cdot \gamma. \tag{4.3}$$

In addition, among all polynomial solutions of (4.3) γ *is uniquely determined by the degree condition*

$$\deg(\gamma) = \deg(\alpha) - \deg(\beta) + \deg(\delta) + 1. \tag{4.4}$$

PROOF. The essential part follows directly from Proposition 4.1. In order to express the solution denominator δ in terms of β, i.e., to get (4.2) apply Lemma 4.3 (b). The first equation of Proposition 4.1 is equivalent to (4.3) because $\mathrm{gcdE}(\mathrm{lcmE}(\delta)) = E\delta$ by Lemma 4.3 (b). The degree estimate (4.4) for γ is immediate from Proposition 3.2 applied to proper γ/δ. Any other polynomial solution $\bar{\gamma}$ of (4.3) gives rise to another telescoping solution, i.e., $\Delta(\bar{\gamma}/\delta) = \alpha/\beta$. Hence $\deg(\bar{\gamma}/\delta) = 0$, which means $\deg(\bar{\gamma}) = \deg(\delta) > \deg(\alpha) - \deg(\beta) + \deg(\delta) + 1$ where the inequality holds by Lemma 4.1. □

One basic application is the practical computation of a rational function solution s of $\Delta s = r$. If such a solution exists, then its S-form $\langle \gamma, \delta \rangle$ can be determined first by computing $\delta = E^{-1}\,\mathrm{gcdE}(\beta)$ and then by solving the *polynomial* difference equation (4.3) for γ using the degree estimate (4.4). A concrete elementary example is given below.

EXAMPLE 4.2. *Given $r \in \mathbf{Q}(x)$ as in Example 4.1. Then the S-form $\langle \alpha, \beta \rangle$ of r is constituted by $\alpha(x) = x - 1/2$ and $\beta(x) = [x + 1/2]^{\underline{3}}$. If $\Delta s = r$ for proper $s \in \mathbf{Q}(x)$ with S-form $\langle \gamma, \delta \rangle$, then by Theorem 4.1: $\delta(x) = \gcd(\beta(x), \beta(x-1)) = [x - 1/2]^{\underline{2}}$ (Fundamental Lemma) and $\gamma \in \mathbf{Q}[x]$ is the solution with $\deg(\gamma) = 1$ (eq. (4.4)) of $x - 1/2 = (x-3/2) \cdot \gamma(x+1) - (x+1/2) \cdot \gamma(x)$. It is easy to determine γ as $\gamma = -x+1$. Hence $s(x) = (-x+1)/[x-1/2]^{\underline{2}}$.* □

As a by-product we get another necessary condition for r being rational summable. We will also use this proposition later, in Section 6, discussing the problem of rational summation in full generality.

PROPOSITION 4.2. *Let the rational function $r \in \mathbf{K}(x)$ be proper with S-form $\langle \alpha, \beta \rangle$ and $\mathrm{GFF}(\beta) = \langle p_1, \ldots, p_k \rangle$, then:*

$$p_1 \neq 1 \Rightarrow r \text{ is not rational summable.}$$

PROOF. By Lemma 4.3, (4.2) is equivalent to $\beta = \mathrm{lcmE}(\delta) = [Eq_1]^{\underline{2}} \cdots [Eq_n]^{\underline{n+1}}$ where $\langle q_1, \ldots, q_n \rangle$ is the GFF-form of δ as in Theorem 4.1. Uniqueness of GFF-form implies $p_1 = 1$. □

REMARK. This contains Abramov's criterion, Proposition 3.3, as a special case because $\mathrm{dis}(b) = 0$ by Lemma 3.5 implies $b = \mathrm{sat}(b) = \beta$ with GFF-form $\langle b \rangle$. □

We give an application where $r \in \mathbf{Q}(x)$ is a rational function with $\mathrm{dis}(r) = 2$:

EXAMPLE 4.3. *By Proposition 4.2, $r(x) = 1/((x+1/2)x(x-3/2)) \in \mathbf{Q}(x)$ is not rational summable because its S-form $\langle \alpha, \beta \rangle = \langle x - 1/2, x\,[x+1/2]^{\underline{3}} \rangle$ and $\mathrm{GFF}(\beta) = \langle x, 1, x+1/2 \rangle$ with $x \neq 1$. - According to Theorem 4.1 the crucial difference equation (4.3) in this case (cf. Example 4.2) reads as*

$$x - \frac{1}{2} = x\ (x - \frac{3}{2}) \cdot \gamma(x+1) - x\ (x + \frac{1}{2}) \cdot \gamma(x),$$

and the non-existence of a rational function solution is explained by the non-existence of a polynomial solution γ which is clear by observing x being a factor of the right and not of the left hand side.

□

4.2. S-form versus Reduced Form

In order to work out more explicitly what one gains in this context by changing from the usual representation of a rational function in reduced form to S-form representation, we give an alternative proof of the essence of Theorem 4.1, i.e., of Proposition 4.1. In this proof we will not use Proposition 3.6, but only fundamental properties of shift-saturation.

ALTERNATIVE PROOF OF PROPOSITION 4.1. Let $r, s \in \mathbf{K}(x)$ be both proper with reduced forms $\langle a, b \rangle$ and $\langle c, d \rangle$, and with S-forms $\langle \alpha, \beta \rangle$ and $\langle \gamma, \delta \rangle$, respectively. Assume $r = \Delta s$. The first equation of Proposition 4.1 is obvious from $\alpha/\beta = \Delta(\gamma/\delta)$. For the

S-form of r we have $\langle \alpha, \beta \rangle = \langle a \cdot \mathrm{sat}(b)/b, \mathrm{sat}(b) \rangle$ by Proposition 3.5. If $d_i := E^i d / \gcd\mathrm{E}(p)$ then $r = \Delta s$ is equivalent to

$$a = \frac{d_0 \cdot Ec - d_1 \cdot c}{t} \quad \text{and} \quad b = \frac{\mathrm{lcmE}(d)}{t}$$

where

$$t := \gcd(d_0 \cdot Ec - d_1 \cdot c, \mathrm{lcmE}(d)) = \gcd(d_0 \cdot Ec - d_1 \cdot c, \gcd\mathrm{E}(d)). \tag{4.5}$$

The last equality is by the same reason as (3.3). If we could prove that

$$\mathrm{sat}(\frac{\mathrm{lcmE}(d)}{t}) = \mathrm{sat}(\mathrm{lcmE}(d)), \tag{4.6}$$

then by Lemma 3.7, $\beta = \mathrm{sat}(b) = \mathrm{sat}(\mathrm{lcmE}(d)/t) = \mathrm{sat}(\mathrm{lcmE}(d)) = \mathrm{lcmE}(\mathrm{sat}(d)) = \mathrm{lcmE}(\delta)$, and the proof of Proposition 4.1 would be completed.

Equation (4.6) immediately follows from Lemma 3.4 if we can guarantee that t does not cancel maximal, minimal, or multiplicity elements of any shift-equivalence class of $p := \mathrm{lcmE}(d)$.

The first two cases are obvious from (4.5), i.e., $t | \gcd\mathrm{E}(d)$, together with

$$\frac{\mathrm{lcmE}(d)}{\gcd\mathrm{E}(d)} = (Eq_1)q_1 \cdot (Eq_2)(E^{-1}q_2) \cdot \ldots \cdot (Eq_k)(E^{-k+1}q_k)$$

where $\langle q_1, \ldots, q_k \rangle$ is the GFF-form of d.

Finally, assume $\gcd(t, m) \neq 1$ for some multiplicity element m, i.e., an irreducible from some p_i-shift-equivalence class of p such that $m^{\mathrm{mult}(p_i)} | p$. Hence $\gcd(t, m) \neq 1$ is equivalent to $m | t$. We have $m | p / \gcd\mathrm{E}(p)$ by Lemma 3.3 and $p / \gcd\mathrm{E}(p) = d_0$ by Lemma 4.3. On the other hand, $m | d_0 \cdot Ec - d_1 \cdot c$. Hence $m | d_1 \cdot c$ which reduces to $m | d_1$ because of $\gcd(c, d) = 1$. Thus $m | d_0$ and $m | d_1$, a contradiction to $\gcd(d_0, d_1) = 1$. □

5. Hypergeometric Telescoping

In this section we consider hypergeometric telescoping. Equipped with the GFF concept we present a new and algebraically motivated approach to the problem. It leads to essentially the same algorithm Gosper came up with in 1978, but in a new setting where its underlying mechanism finds a more transparent explanation than in the descriptions given so far. Of course, to a certain extent this judgement is subjective, so we invite the interested reader to form his/her own by comparison to Gosper's original argumentation as described, for instance, in Gosper (1978) or in the book (Graham *et al.*, 1989). In a subsection we briefly relate rational telescoping, as a special case of Gosper's algorithm, to Theorem 4.1.

A sequence $\langle f_k \rangle_{k \geq 0}$ is called *hypergeometric* over $\mathbf{K}$ if there exists a rational function $\rho \in \mathbf{K}(x)$ such that $f_{k+1}/f_k = \rho(k)$ for all $k \in \mathbf{N}$. Given hypergeometric $\langle f_k \rangle_{k \geq 0}$, the problem of *hypergeometric telescoping* is to find a hypergeometric solution $\langle g_k \rangle_{k \geq 0}$ of

$$g_{k+1} - g_k = f_k. \tag{5.1}$$

Rational telescoping, Section 4, is a special case, because for any $r \in \mathbf{K}(x)$ the sequence $\langle f_k \rangle_{k \geq l}$, where $f_k := r(k)$ and l a sufficiently large integer, evidently is hypergeometric. For the sake of simplicity we will restrict to consider (5.1) for $k \geq 0$.

5.1. Gosper's Algorithm Revisited

Assume that a hypergeometric solution $\langle g_k\rangle_{k\geq 0}$ of (5.1) exists. Let $\sigma \in \mathbf{K}(x)$ be such that $g_{k+1}/g_k = \sigma(k)$ for all $k \in \mathbf{N}$, then evidently

$$g_k = \tau(k) \cdot f_k \tag{5.2}$$

where $\tau(x) = 1/(\sigma(x) - 1) \in \mathbf{K}(x)$. By this relation, eq. (5.1) is equivalent to

$$a \cdot E\tau - b \cdot \tau = b, \tag{5.3}$$

where $\langle a, b\rangle$ is the reduced form of ρ. Vice versa, any rational solution $\tau \in \mathbf{K}(x)$ of (5.3) gives rise to a hypergeometric solution $g_k := \tau(k)\cdot f_k$ of (5.1). This means, hypergeometric telescoping is equivalent to finding a *rational* solution τ of (5.3).

In case such a solution $\tau \in \mathbf{K}(x)$ with reduced form $\langle u, v\rangle$ exists, assume we know v or a multiple $V \in \mathbf{K}[x]$ of v. Then by clearing denominators in $a \cdot EU/EV - b \cdot U/V = b$ the problem reduces further to finding a polynomial solution $U \in \mathbf{K}[x]$ of the resulting difference equation with polynomial coefficients,

$$a \cdot V \cdot EU - b \cdot (EV) \cdot U = b \cdot V \cdot EV. \tag{5.4}$$

(Note that at least one polynomial solution, namely $U = u \cdot V/v$, exists.) Furthermore, equations of that type simplify by canceling gcdE's. For instance, in order to get more information about the denominator v, let $v_i := E^i v/\operatorname{gcdE}(v)$, $i \in \{0, 1\}$. Then (5.3) is equivalent to

$$a \cdot v_0 \cdot Eu - b \cdot v_1 \cdot u = b \cdot v_0 \cdot v_1 \cdot \operatorname{gcdE}(v). \tag{5.5}$$

Now, if $\langle p_1, \ldots, p_m\rangle$, $m \geq 0$, is the GFF-form of v, it follows from $\gcd(u, v) = 1 = \gcd(v_0, v_1)$ and the Fundamental Lemma that

$$v_0 = (E^0 p_1)\cdots(E^{-m+1}p_m)|b \quad\text{and}\quad v_1 = (Ep_1)\cdots(Ep_m)|a. \tag{5.6}$$

This observation gives rise to a simple and straightforward algorithm for computing a multiple $V := [P_1]^{\underline{1}} \ldots [P_n]^{\underline{n}}$ of v. For instance, if $P_1 := \gcd(E^{-1}a, b)$ then obviously $p_1|P_1$. Indeed, we shall see below that by exploiting GFF-properties one can extract iteratively p_i-multiples P_i such that $EP_i|a$ and $E^{-i+1}P_i|b$:

Algorithm VMULT INPUT: the reduced form $\langle a, b\rangle$ of $\rho \in \mathbf{K}(x)$; OUTPUT: polynomials $\langle P_1, \ldots, P_n\rangle$ such that $V := [P_1]^{\underline{1}} \cdots [P_n]^{\underline{n}}$ is a multiple of the reduced denominator v of $\tau \in \mathbf{K}(x)$.
(i) Compute $n = \min\{j \in \mathbf{N}|\gcd(E^{-1}a, E^{k-1}b) = 1 \text{ for all integers } k > j\}$.
(ii) Set $a_0 = a$, $b_0 = b$, and compute for i from 1 to n:

$$P_i = \gcd(E^{-1}a_{i-1}, E^{i-1}b_{i-1}), \tag{5.7}$$

$$a_i = a_{i-1}/EP_i, \tag{5.8}$$

$$b_i = b_{i-1}/E^{-i+1}P_i. \tag{5.9}$$

The lemma tells that the polynomials P_i indeed are multiples of the p_i's:

Lemma 5.1. *Let $\langle u, v\rangle$ with $u, v \in \mathbf{K}[x]$ be the reduced form of a rational function solution τ of eq. (5.3) and $\operatorname{GFF}(v) = \langle p_1, \ldots, p_m\rangle$. Let n and $\langle P_1, \ldots, P_n\rangle$ be computed as in Algorithm VMULT. Then:*

$$n \geq m \text{ and } p_i|P_i \text{ for all } i \in \{1, \ldots, m\}.$$

Proof. The first part, $n \geq m$, is obvious from (5.6). For $n = 0$ the lemma is trivial, hence we assume $n \geq 1$. In view of (5.6) define $\alpha_0, \beta_0 \in \mathbf{K}[x]$ such that

$$a = (Ep_1)\cdots(Ep_m)\cdot\alpha_0 \text{ and } b = (E^0 p_1)\cdots(E^{-m+1}p_m)\cdot\beta_0.$$

For i from 1 to n define

$$g_i := \gcd(E^{-1}\alpha_{i-1}, E^{i-1}\beta_{i-1})$$

and

$$\alpha_i := \alpha_{i-1}/Eg_i \text{ and } \beta_i := \beta_{i-1}/E^{-i+1}g_i.$$

Note that $\alpha_i, \beta_i \in \mathbf{K}[x]$ such that $\alpha_i|\alpha_{i-1}$ and $\beta_i|\beta_{i-1}$, and $\gcd(E^{-1}\alpha_n, E^{k-1}\beta_n) = 1$ for all $k \in \mathbf{N}$. Besides using the null convention $\prod_{j\in\emptyset} q_j = 1$, it will be convenient to define $p_i := 1$ for $i \in \{m+1, \ldots, n\}$. In order to prove $p_i|P_i$, we prove more generally by induction on i that for $i \in \{1, \ldots, n\}$:

$$P_i = p_i \cdot g_i, \tag{5.10}$$
$$a_i = (Ep_{i+1})\cdots(Ep_m)\cdot\alpha_i, \tag{5.11}$$
$$b_i = (E^{-i}p_{i+1})\cdots(E^{-m+1}p_m)\cdot\beta_i. \tag{5.12}$$

The induction relies on the following facts: For any $i \in \{0, \ldots, m-2\}$,
(I) $\forall l, j \in \{i+2, \ldots, m\}$: $\gcd(p_l, E^{-j+i+1}p_j) = 1$,
(II) $\forall l \in \{i+2, \ldots, m\}$: $\gcd(p_l, E^i\beta_i) = 1$,
(III) $\forall j \in \{i+2, \ldots, m\}$: $\gcd(E^{-j+i+1}p_j, E^{-1}\alpha_i) = 1$.

Fact (I) is immediate from the definition of GFF-form; see especially (GFF3) in section 2. The other facts are consequences of the Fundamental Lemma and eq. (5.5) rewritten in the form

$$\alpha_0 \cdot Eu - \beta_0 \cdot u = b \cdot \mathrm{gcdE}(v). \tag{5.13}$$

To prove fact (II) assume that $t|\gcd(p_l, E^i\beta_i)$ for a monic $t \in \mathbf{K}[x]$. This means, $E^{-i}t|E^{-i}p_l$ and $E^{-i}t|\beta_i$ which, by the Fundamental Lemma and $\beta_i|\beta_0$, is equivalent to $E^{-i}t|\mathrm{gcdE}(v)$ and $E^{-i}t|\beta_0$. Thus $E^{-i}t|\alpha_0 \cdot Eu$ by (5.13), hence $E^{-i}t = t = 1$.

To prove fact (III) assume that $t|\gcd(E^{-j+i+1}p_j, E^{-1}\alpha_i)$ for monic $t \in \mathbf{K}[x]$. This means, $Et|E^{-j+i+2}p_j$ and $Et|\alpha_i$ which, by the Fundamental Lemma and $\alpha_i|\alpha_0$, is equivalent to $Et|\mathrm{gcdE}(v)$ and $Et|\alpha_0$. Thus $Et|\beta_0 \cdot u$ by (5.13), hence $Et = t = 1$.

Now the base case $i = 1$ is shown as follows:

$$\begin{aligned} P_1 &= \gcd(E^{-1}a_0, b_0) \\ &= \gcd(p_1\cdots p_m \cdot E^{-1}\alpha_0, (E^0p_1)\cdots(E^{-m+1}p_m)\cdot\beta_0) \\ &= p_1 \cdot \gcd(p_2\cdots p_m \cdot E^{-1}\alpha_0, (E^{-1}p_2)\cdots(E^{-m+1}p_m)\cdot\beta_0) \\ &= p_1 \cdot g_1, \end{aligned}$$

where the last equality follows by the facts (I),(II), and (III). In addition,

$$a_1 = \frac{a_0}{EP_1} = (Ep_2)\cdots(Ep_m)\cdot\frac{\alpha_0}{Eg_1} = (Ep_2)\cdots(Ep_m)\cdot\alpha_1,$$

and

$$b_1 = \frac{b_0}{P_1} = (E^{-1}p_2)\cdots(E^{-m+1}p_m)\cdot\frac{\beta_0}{g_1} = (E^{-1}p_2)\cdots(E^{-m+1}p_m)\cdot\beta_1.$$

The induction step $i \to i+1$ works analogously, and is left to the reader. □

Summarizing, hypergeometric telescoping can be decided constructively as follows: Given the reduced form $\langle a, b\rangle$ of $\rho \in \mathbf{K}(x)$ for which $f_{k+1}/f_k = \rho(k)$, compute polynomials $\langle P_1, \dots, P_n\rangle$ by Algorithm VMULT and take $V := [P_1]^{\underline{1}} \dots [P_n]^{\underline{n}}$; if eq. (5.4) can be solved for $U \in \mathbf{K}[x]$ then $g_k := f_k \cdot U(k)/V(k)$ solves (5.1), if eq. (5.4) admits no polynomial solution then no hypergeometric solution of (5.1) exists.

How this approach relates to Gosper's original one, and to work of Petkovšek, is described in the next subsection. We also want to remark that in practice, before solving for $U \in \mathbf{K}[x]$, equation (5.4) is simplified; see also the next subsection.

5.2. Gosper's Original Approach

Also in this section, let $\langle a, b\rangle$ with $a, b, \in \mathbf{K}[x]$ be the reduced form of $\rho \in \mathbf{K}(x)$ for which $\rho(k) = f_{k+1}/f_k$.

In Gosper's original approach the following type of rational function representation plays a crucial role:

Definition. The triple $\langle p, q, r\rangle$ with polynomials $p, q, r \in \mathbf{K}[x]$ is called a *G-form* for the rational function $a/b \in \mathbf{K}(x)$, if

$$\frac{a}{b} = \frac{Ep}{p} \cdot \frac{q}{r} \quad \text{and} \quad \gcd(q, E^k r) = 1 \text{ for all } k \geq 0.$$

In the previous section we used Algorithm VMULT to compute multiples P_i of the GFF-constituents p_i of the rational solution denominator v. Petkovšek (1992) used exactly the same algorithm in order to compute a *canonical* G-form representation. That canonical form, called *GP-form*, serves as a key ingredient for his algorithm “Hyper”; it is defined as the unique G-form where additionally p and r are supposed to be monic, and $\gcd(p, q) = \gcd(Ep, r) = 1$.

Lemma 5.1 focuses on the $p_i | P_i$ property; the following lemma lists additional facts about Algorithm VMULT which can be proved in a similar fashion:

Lemma 5.2. *Let n, a_n, b_n and $\langle P_1, \dots, P_n\rangle$ be computed as in Algorithm VMULT, then:*
(i) $a = (EP_1) \cdots (EP_n) \cdot a_n$,
(ii) $b = (E^0 P_1) \cdots (E^{-n+1} P_n) \cdot b_n$,
(iii) $\forall k \in \mathbf{N}$: $\gcd(a_n, E^k b_n) = 1$,
(iv) $\forall i \in \{1, \dots, n\}$: $\gcd([P_i]^{\underline{i}}, a_n) = 1$,
(v) $\forall i \in \{1, \dots, n\}$: $\gcd([P_i]^{\underline{i}}, E^{-1} b_n) = 1$,
(vi) $\mathrm{GFF}([P_1]^{\underline{1}} \cdots [P_n]^{\underline{n}}) = \langle P_1, \dots, P_n\rangle$.

Proof. For statements (i)-(v) cf. Petkovšek (1992); the verification of (vi) is left to the reader. □

As a by-product of Lemma 5.2 we obtain that VMULT can be used as an alternative to Algorithm GFF to compute $\mathrm{GFF}(P) = \langle P_1, \dots, P_n\rangle$ for given monic $P \in \mathbf{K}[x]$: simply set $a = EP/\,\mathrm{gcdE}(P)$ and $b = P/\,\mathrm{gcdE}(P)$; cf. Remark (ii) of Section 2.3.

In the previous section the essential part of hypergeometric telescoping was solved by Algorithm VMULT which computes from the reduced form $\langle a, b\rangle$ a multiple V of the “a priori” unknown denominator v of the reduced solution $\tau \in \mathbf{K}(x)$ of eq. (5.3). In the light of Lemma 5.2, the multiple $V = [P_1]^{\underline{1}} \cdots [P_n]^{\underline{n}}$ is nothing but the first part of the

GP-form for a/b; this is true because evidently

$$\langle V, a_n, b_n \rangle \text{ is the GP-form for } a/b.$$

We want to note that this fact was conjectured independently by Schorn (1995).

The algorithmic elegance of Gosper's original approach, which attacked the problem from a different point of view, relies on the crucial observation that not only *one* specific G-form, but *any* G-form for a/b provides a suitable multiple of v. This is made explicit as follows.

First we state an elementary lemma.

LEMMA 5.3. *If $\langle V, q, r \rangle$ with $V, q, r \in \mathbf{K}[x]$ is a G-form for a/b, then for $U \in \mathbf{K}[x]$:*

$$a \cdot \frac{EU}{EV} - b \cdot \frac{U}{V} = b \Leftrightarrow q \cdot EU - r \cdot U = r \cdot V. \tag{5.14}$$

PROOF. [The easy verification is left to the reader.] □

In case that V arises from a G-form for a/b, this lemma rewrites the corresponding difference equation (5.4) for U in more convenient form. For the sake of abbreviation we define

$$G(U, V, q, r) := q \cdot EU - r \cdot U - r \cdot V.$$

The following crucial lemma of Gosper finds a simple GFF proof.

LEMMA 5.4. *Given a G-form $\langle V, q, r \rangle$ for a/b with $V, q, r \in \mathbf{K}[x]$, then:*

$$G(U, V, q, r) = 0 \text{ for } U \in \mathbf{K}(x) \Rightarrow U \in \mathbf{K}[x].$$

PROOF. Assume $U = C/D$, i.e., $\langle C, D \rangle$ is the reduced form of the rational function U. Then $G(U, V, q, r) = 0$ is equivalent to $q \cdot EC/ED - r \cdot C/D = r \cdot V$, and for $\mathrm{GFF}(D) = [d_1]^{\underline{1}} \cdots [d_j]^{\underline{j}}$ we obtain analogously as in the situation of eq. (5.5): $(Ed_1) \cdots (Ed_j) | q$ and $(E^0 d_1) \cdots (E^{-j+1} d_j) | r$. Hence $Ed_j \,|\, \gcd(q, E^j r)$ and $D = 1$. □

Now we are in the position to prove the crucial fact the general mechanism of Gosper's original approach is based on:

PROPOSITION 5.1. *If there exist a G-form $\langle V, q, r \rangle$ for a/b with $V, q, r \in \mathbf{K}[x]$ and $U \in K[x]$ such that*

$$G(U, V, q, r) = 0,$$

then for any G-form $\langle \bar{V}, \bar{q}, \bar{r} \rangle$ for a/b with $\bar{V}, \bar{q}, \bar{r} \in \mathbf{K}[x]$ there exists $\bar{U} \in K[x]$ such that

$$G(\bar{U}, \bar{V}, \bar{q}, \bar{r}) = 0.$$

PROOF. By the assumption and Lemma 5.3 we have $a \cdot EU/EV - b \cdot U/V = b$. Define $\bar{U} := \bar{V} \cdot U/V \in \mathbf{K}(x)$, then

$$\begin{aligned} G(\bar{U}, \bar{V}, \bar{q}, \bar{r}) &= \bar{q} \cdot E\bar{V} \cdot \frac{EU}{EV} - \bar{r} \cdot \bar{V} \cdot \frac{U}{V} - \bar{r} \cdot \bar{V} \\ &= \bar{r} \cdot \bar{V} \, (\frac{\bar{q}}{\bar{r}} \cdot \frac{E\bar{V}}{\bar{V}} \cdot \frac{EU}{EV} - \frac{U}{V} - 1) = 0; \end{aligned}$$

hence $\bar{U} \in \mathbf{K}[x]$ by Lemma 5.4. □

We want to add that in practice the polynomial solution $U \in \mathbf{K}[x]$ of $G(U,V,q,r)=0$ is computed from the following straight-forward variation of the problem:

LEMMA 5.5. *Let $\langle V,q,r\rangle$ be a G-form for a/b where $G(U,V,q,r)=0$ for $U \in \mathbf{K}[x]$ then $U=(E^{-1}r)\cdot W$, where $W \in \mathbf{K}[x]$ solves*

$$q \cdot EW - (E^{-1}r)\cdot W = V. \tag{5.15}$$

PROOF. Because of $\gcd(q,r)=1$ we have $r|EU$. Hence there exists $W \in \mathbf{K}[x]$ such that $U=(E^{-1}r)\cdot W$ for which $G(U,V,q,r)=0$ reduces to (5.15). □

It is the form (5.15) in which the difference equation associated to a G-form is to find in Gosper (1978) or in the book (Graham *et al.*, 1989).

Consider the problem of deciding constructively over $\mathbf{K}(x)$ the general, first-order linear difference equation $a\cdot E\tau - b\cdot\tau = c$ with nonzero polynomials $a,b,c \in \mathbf{K}[x]$. We conclude this section by the remark that following the derivation above, one can easily see how each step of this approach can be modified for solving also this more general problem. We present the result of this modification in form of a proposition. Its proof is entirely analogous to what we did above and is left to the reader. For alternative methods and the general n-th order case see, for instance, Abramov (1995).

PROPOSITION 5.2. *Given nonzero $a,b,c \in \mathbf{K}[x]$, the problem of solving*

$$a\cdot E\tau - b\cdot\tau = c \tag{5.16}$$

for $\tau \in \mathbf{K}(x)$ can be decided constructively as follows:
(i) Compute a G-form $\langle V,q,r\rangle$ for a/b with $V,q,r \in \mathbf{K}[x]$ such that $b|(r\cdot V)$.
(ii) If

$$b\cdot q\cdot EU - b\cdot r\cdot U = c\cdot r\cdot V \tag{5.17}$$

has a solution $U \in \mathbf{K}[x]$ then $\tau = U/V$ solves (5.16), otherwise (5.16) has no rational solution $\tau \in \mathbf{K}(x)$.

PROOF. [Left to the reader.] □

We want to note that $a,b \in \mathbf{K}[x]$ need not be relatively prime; furthermore it is easy to see how any G-form for a/b can be modified to achieve $b|r\cdot V$ in step (i).

EXAMPLE 5.1. *Equation (5.16) with $a = x(x-1)^2$, $b=(x-1)^2$ and $c = x$ has no rational solution $\tau \in \mathbf{Q}(x)$; cf. Example 20 from Karr (1981). Evidently, $\langle V,q,r\rangle = \langle (x-1)^2,(x-1)^2,x\rangle$ is a G-form for a/b such that $b|r\cdot V$, and $(x-1)^2\cdot EU - x\cdot U = x^2$ has no solution $U \in \mathbf{Q}[x]$. The latter can be seen most easily by the observation $x|EU$, i.e., $U=(x-1)\cdot W$ for $W \in \mathbf{K}[x]$, which reduces the equation to $(x-1)^2\cdot EW-(x-1)\cdot W = x$.*
□

5.3. Rational Telescoping as a Special Case

In this section we briefly relate rational telescoping, as a special case of Gosper's algorithm, to Theorem 4.1.

Assume for the telescoping problem (5.1) that $\langle f_k \rangle_{k\geq 0}$ is a rational sequence, i.e., $f_k = \alpha(k)/\beta(k)$ for some $\alpha, \beta \in \mathbf{K}[x]$. As above let $\langle a, b\rangle$ be the reduced form of $\rho \in \mathbf{K}(x)$ for which $f_{k+1}/f_k = \rho(k)$. Then

$$\frac{a}{b} = \frac{E\alpha}{\alpha} \cdot \frac{\beta}{E\beta} = \frac{E\alpha}{\alpha} \cdot \frac{\beta_0}{\beta_1}$$

where $\beta_i := E^i\beta/\operatorname{gcdE}(\beta)$ for $i \in \{0, 1\}$.

Proposition 5.3. *Let $\langle a, b\rangle$, $\langle \alpha, \beta\rangle$ and b_i be as above, then:*

$$\beta \textit{ is saturated} \Rightarrow \langle \alpha, \beta_0, \beta_1\rangle \textit{ a G-form for } a/b.$$

Proof. It remains to show that $\gcd(\beta_0, E^k\beta_1) = 1$ for all $k \in \mathbf{N}$. If $\langle p_1, \ldots, p_n\rangle$ is the GFF-form of β then we have $\beta_0 = (E^0p_1)\cdots(E^{-n+1}p_n)$ and $\beta_1 = (Ep_1)\cdots(Ep_n)$, hence the gcd condition is obvious from the saturatedness of β. □

According to Gosper's algorithm the denominator of τ is $V = \alpha$ and the numerator $U = (E^{-1}\beta_1)\cdot W$, where $W \in \mathbf{K}[x]$ solves

$$\beta_0 \cdot EW - (E^{-1}\beta_1)\cdot W = \alpha, \tag{5.18}$$

the associated equation in simplified form (5.15). Thus,

$$g_k = \frac{\beta_1(k-1)\cdot W(k)}{\alpha(k)} \cdot f_k = \frac{W(k)}{\gcd(\beta(k-1), \beta(k))}$$

because of $(E^{-1}\beta_1)/\beta = 1/E^{-1}\operatorname{gcdE}(\beta)$.

Summarized, given the rational input $\langle f_k\rangle_{k\geq 0}$ by a *saturated representation* $\langle \alpha, \beta\rangle$ then $\langle \alpha, \beta_0, \beta_1\rangle$ is an appropriate G-form for which Gosper's algorithm outputs a *saturated representation* of the rational solution $\langle g_k\rangle_{k\geq 0}$ of (5.1), in case it exists, in the same form as spelled out by Theorem 4.1. This means, the denominator $\gcd(\beta(k-1), \beta(k))$ of g_k is determined as in (4.2), the numerator of g_k as a polynomial solution of (5.18) which is equivalent to (4.3). In the special case where $\langle f_k\rangle_{k\geq 0}$ is given in *S-form* representation $\langle \alpha, \beta\rangle$, Theorem 4.1 says that Gosper's algorithm delivers the proper rational output $\langle g_k\rangle_{k\geq 0}$ also in *S-form* if $\langle \alpha, \beta_0, \beta_1\rangle$ is used as G-form.

Finally we want to remark that a careful analysis of the possible degrees of solutions $W \in \mathbf{K}[x]$ of (5.18) is given in (Lisonek *et al.*, 1993), or in Paule (1993) using the GFF concept.

6. Rational Summation

In this section we apply the GFF concept to the situation where for given $r \in \mathbf{K}(x)$ the telescoping equation $\Delta s = r$ finds no rational solution $s \in \mathbf{K}(x)$.

Considering indefinite rational *integration* the analogous problem for given $r \in \mathbf{K}(x)$ consists in finding the *rational part* $s \in \mathbf{K}(x)$ and the *transcendental part* $t \in \mathbf{K}(x)$ of the decomposition $\int r = s + \int t$. It is well-known, for instance, (Davenport *et al.*, 1988) or (Geddes *et al.*, 1992) that the non-rational part t is determined uniquely under slight side-conditions.

EXAMPLE 6.1. *A summation analogue of*

$$\int \frac{x+1}{x} = x + \int \frac{1}{x} = x + \log(x),$$

for instance, is

$$\sum_{k=1}^{n} \frac{k+1}{k} = n + \sum_{k=1}^{n} \frac{1}{k} = n + H_n,$$

i.e., the harmonic number sequence $\langle H_n \rangle$ is taking the part of the logarithm function. □

Discussing the indefinite *summation* analogue, $\sum r = s + \sum t$, we shall treat this question, referred to as the *decomposition problem*, in the equivalent form

$$r = \Delta s + t. \tag{6.1}$$

For obvious reasons throughout this section we restrict to *proper* rational functions.

EXAMPLE 6.2. *As we have seen in Example 4.1, the pair*

$$\langle s, t \rangle = \langle \frac{(-1)(x-1)}{(x-1/2)(x-3/2)}, 0 \rangle$$

solves the decomposition problem for $r = 1/((x+1/2)(x-3/2))$. The pairs

$$\langle s_1, t_1 \rangle = \langle \frac{(-1/3)(4x-5)}{(x-1/2)(x-3/2)}, \frac{(-1/3)(4x-1)}{(x+1/2)x(x-1/2)} \rangle$$

and

$$\langle s_2, t_2 \rangle = \langle \frac{(-2/3)(x-1)}{(x-1/2)(x-3/2)}, \frac{-2/3}{x(x+1/2)} \rangle$$

solve the decomposition problem for $r = 1/((x+1/2)x(x-3/2))$; cf. Example 4.3. □

Concerning uniqueness of the non-rational part the situation for rational summation is a little bit more subtle than for rational integration. Analogous to integration what one intuitively expects is uniqueness with respect to degree:

DEFINITION. The pair $\langle s, t \rangle$ of proper $s, t \in \mathbf{K}(x)$ is called a *minimal decomposition* of proper $r \in \mathbf{K}(x)$ if $r = \Delta s + t$ such that $\deg(f)$ is minimal where $\langle e, f \rangle$ is the reduced form of t.

But the following lemma shows that a minimal decomposition is not uniquely determined.

LEMMA 6.1. *Given $p, q \in \mathbf{K}[x]$ with q monic and $q = q_1^{m_1} \cdots q_n^{m_n}$ the complete factorization of q over $\mathbf{K}[x]$. Then for any tupel $\langle k_1, \ldots, k_n \rangle$ over the integers there exist $s \in \mathbf{K}(x)$ and $\bar{p} \in \mathbf{K}[x]$ such that*

$$\frac{p}{q_1^{m_1} \cdots q_n^{m_n}} = \Delta s + \frac{\bar{p}}{(E^{k_1} q_1)^{m_1} \cdots (E^{k_n} q_n)^{m_n}}.$$

PROOF. Because of partial fraction decomposition it is sufficient to show the statement for the case $n = 1$, i.e., $q = q_1^{m_1}$. The case $k_1 = 0$ is trivial. Assume $k_1 > 0$, then

$$\frac{p}{q} - \frac{E^{k_1} p}{E^{k_1} q} = (I - E^{k_1})\frac{p}{q} = -\Delta \frac{I - E^{k_1}}{I - E}\frac{p}{q},$$

hence $s = -\sum_{j=0}^{k_1-1} E^j(p/q)$ and $\bar{p} = E^{k_1}p$. By exchanging the roles of p and $\bar{p}$ the statement is proved also for $k_1 < 0$. □

This lemma explains the existence of arbitrary integer-shift variations of a decomposition keeping the denominator degree of the non-rational part invariant.

EXAMPLE 6.3. *The pairs*

$$\langle s_1, t_1 \rangle = \langle \frac{8}{x-2}, \frac{1}{x} \rangle \text{ and } \langle s_2, t_2 \rangle = \langle \frac{7x+2}{x(x-2)}, \frac{1}{x+1} \rangle$$

are decompositions of $r = (x^2 - 11x + 2)/[x]^{\underline{3}}$. *We have*

$$\deg(x) = \deg(x+1) = 1 \text{ and } \operatorname{dis}(t_1) = \operatorname{dis}(t_2) = 0,$$

hence both decompositions are minimal because otherwise there exists a decomposition of r *of the form* $\langle s, 0 \rangle$, *i.e.,* $r = \Delta s = \Delta s_1 + t_1 = \Delta s_2 + t_2$, *hence* $\Delta(s - s_1) = t_1$ *and* $\Delta(s - s_2) = t_2$ *violating Proposition 3.3.* □

A general criterion for deciding whether $\langle s, t \rangle$ is a minimal decomposition was derived by Abramov (1975) who was the first to observe that minimality is guaranteed by requiring the dispersion of t to be zero.

THEOREM 6.1. *Given proper* $r, s, t \in \mathbf{K}(x)$ *such that* $r = \Delta s + t$. *Then* $\langle s, t \rangle$ *is a minimal decomposition of* r *if and only if* $\operatorname{dis}(t) = 0$.

PROOF. See Section 6.3. □

Using GFF and shift-saturation, in Section 6.3 we give a proof of this theorem together with an explicit description in which manner two minimal decompositions differ.

As already pointed out, Abramov (1975) was the first who computed minimal decompositions in an algorithmic way. Viewing the problem in the light of GFF and shift-saturation we present two new algorithms. The first one, as described in Section 6.1, works iteratively similar to the approach sketched by Moenck (1977). The second one, Theorem 6.2 of Section 6.2, provides an analogue to what is called "Horowitz' Method" for rational function integration described, for instance, in (Davenport *et al.*, 1988) or (Geddes *et al.*, 1992). Some authors, for instance, Subramaniam & Malm (1992) refer to that method as the "Hermite-Ostrogradsky Formula".

Based on implementations in Maple a detailed comparison of Abramov's algorithm to the "Horowitz analogue" given in Theorem 6.2 can be found in Pirastu (1995*a*). Pirastu & Siegl (1995*b*) discuss various aspects of these algorithms according a parallel implementation in ‖MAPLE‖ on a workstation network.

6.1. Minimal Decomposition by Iteration

In order to compute a minimal decomposition $\langle s,t\rangle$ of r, in view of Proposition 4.2 and Theorem 4.1, a natural first step is to take the S-form $\langle \alpha,\beta\rangle$ of r and to split off the nontrivial first GFF-constituent of β which possibly exists. Before doing so, for sake of abbreviation we introduce a definition.

Definition. A saturated representation $\langle \gamma,\delta\rangle$ is called *proper* if $\deg(\gamma) < \deg(\delta)$.

Now we are ready for the first basic decomposition step.

Proposition 6.1. *Given a proper S-form $\langle \sigma,\tau\rangle$ with polynomials $\sigma,\tau \in \mathbf{K}[x]$ such that $\mathrm{GFF}(\tau) = \langle p_1,\ldots,p_n\rangle$. Then there exist unique proper S-forms $\langle \alpha,\beta\rangle$ and $\langle \gamma,\delta\rangle$ with polynomials from $\mathbf{K}[x]$ such that*

$$\frac{\sigma}{\tau} = \frac{\alpha}{\beta} + \frac{\gamma}{\delta} \tag{6.2}$$

where

$$\mathrm{GFF}(\beta) = \langle 1,p_2,\ldots,p_n\rangle \quad \textit{and} \quad \mathrm{GFF}(\delta) = \langle p_1\rangle. \tag{6.3}$$

Proof. Let $\beta := [p_2]^{\underline{2}}\ldots[p_n]^{\underline{n}}$ and $\delta := p_1 = \tau/\beta$. Because of $\gcd(\beta,\delta) = 1$ there exist unique polynomials $\alpha,\gamma \in \mathbf{K}[x]$, which can be computed by Extended Euclidean Algorithm (e.g., Geddes *et al.*, 1992), such that $\deg(\alpha) < \deg(\beta)$, $\deg(\gamma) < \deg(\delta)$, and $\sigma = \alpha\cdot\delta + \gamma\cdot\beta$ which is equivalent to (6.2). Clearly (6.3) holds and both $\langle \alpha,\beta\rangle$ and $\langle \gamma,\delta\rangle$ are proper saturated representations. Assume that $\langle \gamma,\delta\rangle$ is not in S-form, then according to Proposition 3.6 there exists r with $r|p_1 = \delta$, $\deg(r) > 0$, and $r|\gamma$. Hence $r|\sigma$, by Proposition 3.6 a violation of the S-form of $\langle \sigma,\tau\rangle$. Analogously it is proved that $\langle \alpha,\beta\rangle$ is in S-form. □

Remark. The proof of Proposition 6.1 shows how the numerators α and γ can be obtained constructively. □

Now the second basic decomposition step is as follows:

Proposition 6.2. *Given a proper S-form $\langle \alpha,\beta\rangle$ with polynomials $\alpha,\beta \in \mathbf{K}[x]$ such that $\mathrm{GFF}(\beta) = \langle 1,p_2,\ldots,p_n\rangle$. Then there exist proper saturated representations $\langle \bar\alpha,\bar\beta\rangle$ and $\langle \gamma,\delta\rangle$ with polynomials from $\mathbf{K}[x]$ such that*

$$\frac{\alpha}{\beta} = \Delta\frac{\gamma}{\delta} + \frac{\bar\alpha}{\bar\beta} \tag{6.4}$$

where

$$\mathrm{GFF}(\bar\beta) = \langle p_2,\ldots,p_n\rangle \quad \textit{and} \quad \mathrm{GFF}(\delta) = \langle E^{-1}p_2,\ldots,E^{-1}p_n\rangle. \tag{6.5}$$

Proof. If α/β is rational summable then, by Theorem 4.1, $\delta = E^{-1}\,\mathrm{gcdE}(\beta)$, and γ a polynomial solution of (4.3). In this case we take $\langle \bar\alpha,\bar\beta\rangle := \langle 0, E\delta\rangle$. If α/β is not rational summable then also define $\delta := E^{-1}\,\mathrm{gcdE}(\beta)$ and $\bar\beta := \mathrm{gcdE}(\beta) = E\delta$. In this case we modify the inhomogeneous part α of (4.3) by adding a polynomial such that the resulting equation admits a polynomial solution. Because of $\gcd(\beta/E\delta,\beta/\delta) = 1$, for instance, by Extended Euclidean Algorithm one can find $\gamma,\bar\gamma \in \mathbf{K}[x]$ such that $\alpha = \beta/E\delta\cdot\bar\gamma - \beta/\delta\cdot\gamma$

with $\deg(\gamma) < \deg(\beta/E\delta) \leq \deg(\delta)$. Defining $\bar{\alpha} \in \mathbf{K}[x]$ by $\bar{\gamma} = E\gamma + \bar{\alpha}$ this equation is equivalent to (6.4). In any case $\langle \gamma, \delta \rangle$ and $\langle \bar{\alpha}, \bar{\beta} \rangle$ are proper saturated representations and also (6.5) holds by definition of $\bar{\beta}$ and δ. □

Remark. The proof of Proposition 6.2 shows how the numerators $\bar{\alpha}$ and γ can be obtained constructively. □

One should note that generally the numerators $\bar{\alpha}$ and γ in Proposition 6.2 are *not* uniquely determined. In addition, neither $\langle \bar{\alpha}, \bar{\beta} \rangle$ nor $\langle \gamma, \delta \rangle$ need to be in S-form. Both facts are made explicit by the following example.

Example 6.4. *Different decompositions of type as in Proposition 6.2 of the same rational function* $r(x) = (x^2 - 11x + 2)/[x]^{\underline{3}}$*:*

$$r(x) = \Delta \frac{-x+10}{(x-1)(x-2)} + \frac{x-10}{x(x-1)} = \Delta \frac{8x-8}{(x-1)(x-2)} + \frac{x-1}{x(x-1)}.$$

□

Despite this lack of uniqueness, Proposition 6.2 together with Proposition 6.1 provide the basic reduction steps to solve the decomposition problem in an iterative manner. The algorithm works by iterated reduction of shift-saturated representations until one arrives at a non-rational part with dispersion zero. The mechanism of the algorithm will be clear from the example below.

Remark. For a discussion how this method relates to Moenk (1977) see the diploma thesis of Pirastu (1992), and also Pirastu (1995*a*). □

Example 6.5. **("Minimal Decomposition by Iteration")** *Consider* $r \in \mathbf{Q}(x)$ *from Example 4.3 with S-form* $\langle \sigma/\tau \rangle = \langle x - 1/2, \tau = [x]^{\underline{1}}\ [1]^{\underline{2}}\ [x+1/2]^{\underline{3}} \rangle$*. For the first* GFF-*constituent of* τ *we have* $x \neq 1$*, hence by Proposition 6.1* r *decomposes as*

$$\frac{\sigma}{\tau} = -\frac{4/3}{[x]^{\underline{1}}} + \frac{4/3(x-1/2)(x-1)}{[x+1/2]^{\underline{3}}}$$

Now, according to Proposition 6.2 one computes the decomposition

$$\frac{4/3(x-1/2)(x-1)}{[x+1/2]^{\underline{3}}} = \Delta \frac{(-4/3)x + 5/3}{[x-1/2]^{\underline{2}}} + \frac{4/3(x-1)}{[x+1/2]^{\underline{2}}}$$

($\delta = [x-1/2]^{\underline{2}}$*,* $\bar{\gamma} = -1$*,* $\gamma = -(4/3)x + 5/3$*, and* $\bar{\alpha} = \bar{\gamma} - E\gamma = (4/3)(x-1)$*).*

Again applying Proposition 6.2, which is possible because $\bar{\alpha}/(E\delta)$ *is in S-form with first* GFF-*constituent of the denominator equal to* 1*, yields*

$$\frac{4/3(x-1)}{[x+1/2]^{\underline{2}}} = \Delta \frac{2/3}{x-1/2} + \frac{4/3}{x+1/2}.$$

Collecting all parts together one obtains a minimal decomposition

$$r = \frac{\sigma}{\tau} = \Delta s + t$$

with

$$s = \frac{(-4/3)x + 5/3}{(x-1/2)(x-3/2)} + \frac{2/3}{x-1/2} = -\frac{2/3(x-1)}{(x-1/2)(x-3/2)}$$

as the rational and

$$t = -\frac{4/3}{x} + \frac{4/3}{x+1/2} = -\frac{2/3}{x(x+1/2)}$$

as the non-rational part for which $\mathrm{dis}(t) = 0$*. Equivalently, splitting off the harmonic number expression by partial fraction decomposition we get for* $n \geq 1$*:*

$$\sum_{k=1}^{n} \frac{1}{(k+1/2)k(k-3/2)} = -\frac{8}{3}\frac{n}{(2n+1)(2n-1)} - \frac{4}{3}H_n + \frac{4}{3}\sum_{k=1}^{n}\frac{1}{k+1/2}.$$

According to Theorem 6.1 no further essential simplification is possible. □

6.2. Minimal Decomposition by Horowitz Analogue

For indefinite integration $\int r$ of a rational function, besides others there is a method usually called Horowitz' method since it was studied in detail by Horowitz (1971). The method relies on the following fact:

Given proper $r \in \mathbf{K}(x)$ with reduced form $\langle a, b\rangle$ where $b = b_1^1\ b_2^2 \ldots b_n^n$ is the square-free factorization of b. Then there exist uniquely determined polynomials $c, e \in \mathbf{K}[x]$ such that

$$\frac{a}{b} = D\frac{c}{b_2^1\ b_3^2 \ldots b_n^{n-1}} + \frac{e}{b_1\ b_2 \ldots b_n} \tag{6.6}$$

and

$$\deg(c) < \deg(d), \quad \deg(e) < \deg(\frac{b}{d})$$

where $d = \gcd(b, Db) = b_2^1\ b_3^2 \ldots b_n^{n-1}$. (D is the derivation operator.)

Observing that $b_1\ b_2 \ldots b_n = b/d$, Horowitz' method reduces the problem of finding c and e to solving a system of linear equations. This reduction is done by rewriting eq. (6.6) as a polynomial differential equation

$$a = \frac{b}{d}\cdot(Dc) - \frac{(Dd)b}{d^2}\cdot c + d\cdot e. \tag{6.7}$$

(Note that $(Dd)b/d^2$ indeed is a polynomial.) Coefficient comparison after replacing c and e by sums $\sum c_k x^k$ and $\sum e_k x^k$ with undetermined coefficients leads to the linear system to solve.

As we shall see, in the case of indefinite rational summation there is an analogue to Horowitz' method in which GFF plays the part of square-free factorization. Another substantial difference consists in the fact that one has to take the S-form of the rational function r instead of the usual reduced form. This analogue of the Horowitz decomposition solves the decomposition problem and reads as follows:

Theorem 6.2. *Given a proper S-form* $\langle \alpha, \beta\rangle$*,* $\alpha, \beta \in \mathbf{K}[x]$*, with* $\mathrm{GFF}(\beta) = \langle p_1, \ldots, p_n\rangle$*. Then there exist unique proper saturated representations* $\langle \alpha_0, \beta_0\rangle$ *and* $\langle \gamma, \delta\rangle$ *with polynomials from* $\mathbf{K}[x]$ *such that*

$$\frac{\alpha}{\beta} = \Delta\frac{\gamma}{\delta} + \frac{\alpha_0}{\beta_0} \tag{6.8}$$

where

$$\mathrm{GFF}(\beta_0) = \langle p_1 \cdots p_n\rangle \quad \textit{and} \quad \mathrm{GFF}(\delta) = \langle E^{-1}p_2, \ldots, E^{-1}p_n\rangle. \tag{6.9}$$

Moreover, the polynomials α_0 and γ can be obtained constructively as the uniquely determined solutions of the polynomial difference equation

$$\alpha = \frac{\beta}{\operatorname{gcdE}(\beta)} \cdot E\gamma - \frac{\beta}{E^{-1}\operatorname{gcdE}(\beta)} \cdot \gamma + E^{-1}\operatorname{gcdE}(\beta) \cdot \alpha_0. \tag{6.10}$$

PROOF. The existence part is an easy induction exercise. The case $n = 0$ is trivial. Assuming $n \geq 1$ one distinguishes the cases $p_1 \neq 1$ and $p_1 = 1$, and uses Propositions 6.1 and 6.2. The details are left to the reader. Concerning uniqueness, by (6.9) we have $\beta_0 = \beta/E^{-1}\operatorname{gcdE}(\beta)$ and $\delta = E^{-1}\operatorname{gcdE}(\beta)$, thus (6.10) is equivalent to (6.8), and it remains to show that α_0 and γ are the uniquely determined polynomial solutions of (6.10). Let $l := \deg(\delta)$, $k := \deg(\beta_0) = \deg(\beta/\delta) = \deg(\beta/E\delta)$ and $\alpha_0 := \sum_{i=0}^{k-1} a_i x^i$, $\gamma := \sum_{j=0}^{l-1} g_j x^j$ where a_i, g_j are undetermined coefficients over $\mathbf{K}$. Then for the left hand side of (6.10): $\deg(\alpha) < \deg(\beta) = \deg(\delta \cdot \beta/\delta) = l + k$, and $\deg(\beta/E\delta \cdot E\gamma - \beta/\delta \cdot \gamma) < \deg(\beta/\delta) + \deg(\gamma)$ implies for the right hand side: $\deg(\beta/E\delta \cdot E\gamma - \beta/\delta \cdot \gamma + \delta \cdot \alpha_0) = \deg(\delta) + \deg(\alpha_0) \leq k + l - 1$. Hence by coefficient comparison of the left and right hand sides, solving (6.10) for $\alpha_0, \gamma \in \mathbf{K}[x]$ is equivalent to solving $k+l$ linear equations in $k+l$ unknowns. □

Since $\operatorname{dis}(\alpha_0/\beta_0) = \operatorname{dis}(p_1 \cdots p_n) = 0$, Theorem 6.2 delivers a minimal decomposition because of Theorem 6.1. Thus the algorithm essentially consists of two parts only: computing the S-form of the input function r and finding polynomial solutions α_0, γ of (6.10). The efficiency of this algorithm is discussed in (Pirastu, 1995*a*) and (Pirastu & Siegl, 1995*b*). Below we illustrate the algorithm along elementary Maple steps, a careful implementation in Maple was done by Pirastu (1995*a*).

EXAMPLE 6.6. **("Minimal Decomposition by Horowitz Analogue")**
In MAPLE V.3 its not possible to evaluate

```
> sum((k+2)/( k*(k^2+2)^2 * (k^2-1)^2 ), k);

                      -----
                       \            k + 2
                        )    ---------------------
                       /         2      2   2     2
                      ----- k (k  + 2)  (k  - 1)
                        k
```

But using MAPLE in connection with Theorem 6.2 needs only a few lines to decompose the summand. Note that in simple cases, like this one, the S-form is computed easily by inspection.

```
> alpha:=(x+2)*x;
                          alpha := (x + 2) x

> beta:=(x^2+2)^2 * ((x+1)*x*(x-1))^2;
                          2     2         2  2         2
                beta := (x  + 2)  (x + 1)  x  (x - 1)
```

```
> delta:=factor(gcd(beta,subs(x=x-1,beta)));
                                          2        2
                              delta := x  (x - 1)

> gam:=sum(g[i]*x^i,i=0..3);
                                                   2          3
                    gam := g[0] + g[1] x + g[2] x  + g[3] x

> alphaz:=sum(a[i]*x^i,i=0..5);
                             2          3          4          5
alphaz := a[0] + a[1] x + a[2] x  + a[3] x  + a[4] x  + a[5] x

> match(alpha=beta/subs(x=x+1,delta)*subs(x=x+1,gam) -
> (beta/delta)*gam + delta*alphaz, x, 'coeffs');
                                      true

> factor(subs(coeffs,gam));
                                                     2
                           - 1/12 - 1/12 x + 1/12 x

> factor(subs(coeffs,alphaz));
bytes used=400056, alloc=196572, time=2.916
                                         2         3
                          1 + 2/3 x + 1/3 x  + 1/6 x
```

Hence we get the minimal decomposition

$$\frac{x+2}{x(x^2+2)^2(x^2-1)^2} = \Delta \frac{1}{12}\frac{x^2-x-1}{x^2(x-1)^2} + \frac{1}{6}\frac{x^3+2x^2+4x+6}{(x^2+2)^2(x+1)^2}$$

or, equivalently, for $n \geq 2$*:*

$$\sum_{k=2}^{n} \frac{k+2}{k(k^2+2)^2(k^2-1)^2} = -\frac{1}{48}\frac{(n+2)^2(n-1)^2}{(n+1)^2n^2} + \frac{1}{6}\sum_{k=2}^{n}\frac{k^3+2k^2+4k+6}{(k^2+2)^2(k+1)^2}.$$

The sum expression on the right hand side constitutes the non-rational part of the original sum. □

Remark. According to Theorem 6.1 no further essential reduction is possible. But in general the non-rational part by partial-fraction decomposition over **K** can be decomposed into smaller subparts, for instance, in the example above over **Q** as

$$\frac{1}{18}\sum_{k=2}^{n}\frac{1}{(k+1)^2} + \frac{7}{54}\sum_{k=2}^{n}\frac{1}{k+1} - \frac{1}{54}\sum_{k=2}^{n}\frac{7k-4}{k^2+2} - \frac{1}{9}\sum_{k=2}^{n}\frac{k-1}{(k^2+2)^2}.$$

One should note that the denominators already have been delivered by the Horowitz analogue. Extending the ground field **Q** to the splitting field of the irreducible denominator x^2+2 further refinement is possible. This raises the following question: How would an analogue to what is called Rothstein/Trager-method for rational integration (see Rothstein, 1976, or Trager, 1976, or Geddes *et al.*, 1992) look like in the case of rational summation? □

6.3. Uniqueness of Minimal Decomposition

In this section we give a proof of Abramov's $\mathrm{dis}(t) = 0$ criterion, Theorem 6.1, for minimal decompositions $\langle s, t\rangle$ of r. A key ingredient of this proof is the following theorem which tells how two minimal decompositions differ. Together with Theorem 6.1 its essential message is:

- A minimal decomposition $\langle s, t\rangle$ is unique up to variations induced by arbitrary integer-shifts of the irreducibles of the reduced denominator of t.

THEOREM 6.3. *Given proper $s, \bar{s}, t, \bar{t} \in \mathbf{K}(x)$ such that*

$$\Delta s + t = \Delta \bar{s} + \bar{t}, \tag{6.11}$$

let $\langle e, f\rangle$ and $\langle \bar{e}, \bar{f}\rangle$ be the reduced forms for t and $\bar{t}$, respectively, with

$$f = p_1^{\alpha_1} \cdots p_m^{\alpha_m} \text{ and } \bar{f} = q_1^{\beta_1} \cdots q_n^{\beta_n}$$

the complete factorization of f and $\bar{f}$ over $\mathbf{K}[x]$. If

$$\mathrm{dis}(t) = \mathrm{dis}(\bar{t}) = 0$$

then $m = n$ and for all $i \in \{1, \dots, n\}$: $\alpha_i = \beta_i$ and $q_i = E^{k_i} p_i$ for some integer k_i.

PROOF. Let $g := \gcd(f, \bar{f})$, $\varphi := f/g$, $\bar{\varphi} := \bar{f}/g$, $d := \gcd(\bar{e}\cdot\varphi - e\cdot\bar{\varphi}, g\cdot\varphi\cdot\bar{\varphi})$, and $\psi := g/d$ which is in $\mathbf{K}[x]$, then $\langle(\bar{e}\cdot\varphi - e\cdot\bar{\varphi})/d, \psi\cdot\varphi\cdot\bar{\varphi}\rangle$ is the reduced form of $t - \bar{t}$ which is rational summable by (6.11). Hence by Proposition 4.2 any shift-equivalence class of irreducibles of $\psi\cdot\varphi\cdot\bar{\varphi}$ has at least two elements, i.e.,

$$t \text{ irreducible and } t|\psi\cdot\varphi\cdot\bar{\varphi} \Rightarrow \exists k \neq 0 \text{ such that } E^k t|\psi\cdot\varphi\cdot\bar{\varphi}. \tag{6.12}$$

The crucial observation is that $\psi = 1$, because if an irreducible $t|\psi$ then $E^k t|\psi\cdot\varphi\cdot\bar{\varphi}$ for some $k \neq 0$, which causes a contradiction: $E^k t$ cannot divide ψ, φ, or $\bar{\varphi}$, otherwise one of the dis conditions $\mathrm{dis}(\psi) = \mathrm{dis}(\varphi) = \mathrm{dis}(\bar{\varphi}) = 0$ would be violated. Because of (6.12), $\psi = 1$, and $\mathrm{dis}(\varphi) = \mathrm{dis}(\bar{\varphi}) = 0$ each shift-equivalence class of irreducible factors of $\varphi\cdot\bar{\varphi}$ contains exactly two elements: one belonging to φ and one to $\bar{\varphi}$. Thus $m = n$ and $q_i = E^{k_i} p_i$ with integer k_i for all $i \in \{1, \dots, n\}$. It remains to show that $\alpha_i = \beta_i$. By Lemma 6.1 there exist $\tilde{s} \in \mathbf{K}(x)$ and $\tilde{e} \in \mathbf{K}[x]$ such that $\bar{e}/\bar{f} = \Delta\tilde{s} + \tilde{e}/(p_1^{\beta_1} \cdots p_n^{\beta_n})$, hence $\Delta(s - \bar{s} - \tilde{s}) = \tilde{e}/(p_1^{\beta_1} \cdots p_n^{\beta_n}) - e/(p_1^{\alpha_1} \cdots p_n^{\alpha_n})$. Let $\langle a, b\rangle$ be the reduced form of the right hand side of the last equation, then evidently $\mathrm{dis}(b) = 0$, a contradiction to Proposition 3.3. Thus a/b must be 0 which implies $\alpha_i = \beta_i$ for all $i \in \{1, \dots, n\}$. □

Now we are ready for the

PROOF OF THEOREM 6.1. Let $\langle e, f\rangle$ be the reduced form of t. Assume that $\langle s, t\rangle$ is a minimal decomposition of r with $\mathrm{dis}(t) > 0$. If $\langle\alpha, \beta\rangle$ is the S-form of t with $\mathrm{GFF}(\beta) = \langle p_1, \dots, p_n\rangle$, then also $\mathrm{dis}(\beta) > 0$. By Theorem 6.2 there exist $\bar{s} \in \mathbf{K}(x)$, $\alpha_0 \in \mathbf{K}[x]$ such that $t = \Delta\bar{s} + \alpha_0/(p_1 \cdots p_n)$. Lemma 3.5 implies $\deg(p_1 \cdots p_n) < \deg(f)$, a contradiction to minimality of $\langle s, t\rangle$ because of $r = \Delta(s + \bar{s}) + \alpha_0/(p_1 \cdots p_n)$. For the other direction, assume $\mathrm{dis}(t) = 0$. For any minimal decomposition $\langle \bar{s}, \bar{t}\rangle$ of r with reduced form $\langle \bar{e}, \bar{f}\rangle$ of $\bar{t}$ by the part proved above we also have $\mathrm{dis}(\bar{t}) = 0$. Now Theorem 6.3 implies $\deg(f) = \deg(\bar{f})$, hence $\langle s, t\rangle$ must be minimal. □

7. Conclusion

Besides the applications discussed in this paper, the GFF concept can be used, for instance, also for dealing with q-hypergeometric summation. In this context, instead of the shift operator $(Ep)(x) := p(x + 1)$ the q-shift operator $(\epsilon p)(x) := p(qx)$ plays the fundamental role. A brief description of using q GFF for deriving a q-analogue of Gosper's algorithm is given in (Paule & Strehl, 1995). Both types of shift operators are special cases of difference field extensions considered in Karr's general summation theory (Karr, 1981 and 1985). Thus, as pointed out in the introduction, one might expect that GFF or a suitable generalization could also play some role there; cf. the question raised in section 4.2 of (Karr, 1981).

Concerning rational summation we want to point out that Malm and Subramaniam (1995) came up with another Horowitz analogue. A definite answer concerning the question of optimality of such analogues has been given by Pirastu & Strehl (1994) in the following sense: given the proper rational function r with a generic numerator they are able to solve the decomposition problem $r = \Delta s + t$ optimally in the sense that also the degree of the reduced denominator of s is minimal.

With respect to computer algebra software, for the Maple system it is planned to replace the implementation of Moenck's algorithm, which was used so far for rational summation, by Pirastu's optimization (Pirastu, 1995*a*) of Abramov's algorithm. — C. Mallinger implemented most of the forms and algorithms presented in this paper in Mathematica; the programs are available via email request to the author.

Acknowledgements: I would like to thank Volker Strehl, Roberto Pirastu and Markus Schorn for critical comments and helpful discussions. Special thanks go to an anonymous referee whose detailed, instructive and penetrating comments were extremely helpful for the revision of the paper. One of his valuable suggestions for streamlining the presentation was to introduce the reduction Lemma 3.2.

References

Abramov, S.A. (1971). On the summation of rational functions. *Zh. vychisl. Mat. mat. Fiz. , 4* **11**, 1071–1075. English transl. in USSR Comput. Math. Phys.

Abramov, S.A. (1975). The rational component of the solution of a first-order linear recurrence relation with a rational right side. *Zh. vychisl. Mat. mat. Fiz. , 4* **15**, 1035–1039. English transl. in USSR Comput. Math. Phys.

Abramov, S.A. (1995). Rational solutions of linear difference and q-difference equations with polynomial coefficients. In: (Levelt, T., ed) *Proc. ISSAC '95*, pp. 285–289, New York: ACM Press.

Davenport, J.H., Siret Y., Tournier, E. (1988). *Computer Algebra - Systems and Algorithms for Algebraic Computation*. London: Academic Press.

Geddes, K.O., Czapor, S.R., Labahn, G. (1992). *Algorithms for Computer Algebra*. Boston/Dordrecht/London: Kluwer Acad. Publ.

Gosper, R.W. (1978). Decision procedures for indefinite hypergeometric summation. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 40–42.

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

Horowitz, E. (1971) Algorithms for Partial Fraction Decomposition and Rational Integration. In: *Proc. of SYMSAM '71*, pp. 441-457, Los Angeles.

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

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

Lisonek, P., Paule, P., Strehl, V. (1993). Improvement of the Degree Setting in Gosper's Algorithm. *J. Symbolic Computation* **16**, 243–258.

Malm, D.E.G., Subramaniam, T.N. (1995). The summation of rational functions by an extended Gosper algorithm. *J. Symbolic Computation* **19**, 293–304.

Moenck, R. (1977). On Computing Closed Forms for Summations. In: *Proc. 77 MACSYMA Users Conf.*, pp. 225-236, Berkeley.

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

Paule, P., Strehl, V. (1995). Symbolic summation - some recent developments. *RISC-Linz Report Series 95-11*, Johannes Kepler University, Linz. To appear in: "Computeralgebra in science and engineering - algorithms, systems, applications", J. Fleischer, J. Grabmeier, F. Hehl, and W. Kuechlin (eds.), World Scientific, Singapore.

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

Pirastu, R. (1992). *Algorithmen zur Summation rationaler Funktionen*, (Algorithms for Summation of Rational Functions, in German), Diploma Thesis, University Erlangen-Nürnberg.

Pirastu, R., Strehl, V. (1994). Rational summation and Gosper-Petkovšek representation. *IMMD-Erlangen, Technical Report 94-03* , Univ. Erlangen-Nürnberg. To appear in: *J. Symbolic Computation*.

Pirastu, R. (1995*a*). Algorithms for indefinite summation of rational functions in Maple. *MapleTech* **2**, 29–38.

Pirastu, R., Siegl, K. (1995*b*). Parallel computation and indefinite summation: a ‖MAPLE‖ application for the rational case. To appear 1995 in: *J. Symbolic Computation*.

Rothstein, M. (1976). *Aspects of Symbolic Integration and Simplification of Exponential and Primitive Functions*. Ph.D. Thesis, Univ. of Wisconsin, Madison.

Schorn, M. (1995). Contributions to Symbolic Summation. *RISC-Linz Diploma Thesis*, J. Kepler University Linz.

Strehl, V. (1992). A calculus of intervals - preliminary notes. Unpublished manuscript (29. 8. 1992).

Subramaniam, T.N., Malm, D.E.G. (1992). How to integrate rational functions. *Amer. Math. Monthly* **99**, 762–772.

Trager, B. M. (1976). Algebraic Factoring and Rational Function Integration. In: *Proc. of 1976 ACM Symposium on Symbolic and Algebraic Computation*, pp. 219-226, New York: ACM Inc.