An Algorithm for Deciding Zero Equivalence of Nested Polynomially Recurrent Sequences

Manuel Kauers [1]

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

Abstract

We introduce the class of nested polynomially recurrent sequences which includes a large number of sequences that are of combinatorial interest. We present an algorithm for deciding zero equivalence of these sequences, thereby providing a new algorithm for proving identities among combinatorial sequences. This algorithm is able to treat mathematical objects which are not covered by any other known symbolic method for proving combinatorial identities. Despite its theoretical flavor and its high complexity, an implementation of the algorithm can be successfully applied to nontrivial examples.

Key words: symbolic computation, combinatorial sequences, nested polynomially recurrent sequences, zero equivalence, decision procedure

1 Introduction

Computer proofs of combinatorial identities came up in the early nineties when Zeilberger [1] presented algorithms for deciding whether a representation of a *holonomic function* represents the zero function. The study of holonomic and ∂-finite functions [2,3] was motivated by the observation that many special functions are of such type and by the fact that the defining differential-difference system of such functions provides a convenient representation of the mathematical object which can be used in computations.

Email address: `manuel.kauers@risc.uni-linz.ac.at` (Manuel Kauers).

[1] This work was financially supported by the Austrian Science Foundation, FWF

However, little is known about the algorithmic treatment of functions which are not holonomic. In this paper, we restrict our attention to univariate sequences, i.e., functions with domain $\mathbb{N}$. We note that in this case, holonomic sequences are also called P-finite [4]. We present an algorithm for proving zero equivalence (and hence, for proving identities) of *nested polynomially recurrent sequences,* to be defined in Section 3. The class of these sequences contains all holonomic sequences, but it contains in addition also plenty of interesting objects which are not holonomic and which, to our knowledge, could not be handled so far by symbolic methods. Section 6 contains some examples.

We will employ the notions of *difference algebra* [5], which can be seen as a discrete analogue to differential algebra [6]. Sequences will be defined by annihilating *difference polynomials* from a polynomial *difference ring* whose variables represent the sequences under consideration, or shifts of these sequences (see Section 2). Compared to the definition of sequences by annihilating linear operators, as it is done in algorithms for holonomic objects [1–3], the use of difference polynomials allows the definition of sequences of a more general type, as we will see in Section 3. However, we employ the notion of difference algebra only as a convenient language, but we will not need any deep results from the theory of difference algebra [5]. The basic definitions we need will be stated in Section 2. For our arguments we assume familiarity with elementary concepts of commutative algebra only, as it is presented, e.g., in [7].

On one hand, our algorithm is of theoretical interest. It provides a decision procedure for deciding zero equivalence of nested polynomially recurrent sequences. On the other hand, our algorithm is of practical relevance, for instance, to prove entries from mathematical tables like [8] (see also Example 11). Despite its very high worst case complexity, we succeeded in proving nontrivial identities using a straightforward implementation.

This paper is organized as follows. Section 2 introduces some basic notions and convenient notation. In Section 3 we introduce the class of nested polynomially recurrent sequence. We will give some examples and some closure properties of the class of these sequences. Section 4 presents how nested polynomially recurrent sequences are translated into the language of difference rings by means of *defining relations.* Section 5 presents the algorithm for proving zero equivalence, along with proofs of its correctness and termination. Section 6 concludes the paper by some illustrative examples.

2 Difference Rings and Difference Ideals

Let K be a computable algebraically closed field of characteristic zero. By computable, we mean that every element $a \in K$ should have a finite representation $\bar{a}$ and for any representation $\bar{a}$ of an element in $a \in K$ it should be decidable if a is the zero element of K. Natural choices for K are the algebraic closure of $\mathbb{Q}$, or transcendental extensions hereof.

We recall the basic definitions of differential algebra [5]. A *difference ring* R is a commutative ring with unity, equipped with an injective ring homomorphism $s\colon R \to R$. The homomorphism s is called the *shift operator* of R.

An important example of a difference ring is the m-fold *polynomial difference ring,* to be denoted by $R^{(m)}$. Let $\{t_{i,j} : i = 1, \ldots, m, j \in \mathbb{N}_0\}$ be algebraically independent over K and let s be canonically defined by $s(c) = c$ $(c \in K)$ and $s(t_{i,j}) := t_{i,j+1}$. Then

$$R^{(m)} := K[t_{1,0}, \ldots, t_{m,0}, t_{2,1}, \ldots, t_{m,1}, t_{3,1}, \ldots\ldots],$$

equipped with this s forms a difference ring. We view $R^{(m)}$ as a polynomial ring with infinitely many indeterminates. The elements of $R^{(m)}$ are called *difference polynomials.* Readers familiar with differential algebra will note the similarity with differential polynomials. However, it is worth noting that $s(ab) = s(a)s(b)$ whereas the derivative in a differential ring obeys the more complicated Leibniz rule $D(ab) = D(a)b + aD(b)$.

Writing $s^n := s \circ s^{n-1}$, $s^0 = \mathrm{id}$, defining $t_i := t_{i,0}$ $(i = 1, \ldots, m)$ and omitting parentheses we will often write $s^j t_i$ in place of $t_{i,j}$. This shall remind us that the index j corresponds to the jth shift of the object represented by the variable t_i. We will use similar shortcuts not only for the indeterminates, but also for polynomials, sets of polynomials, etc. As an example, $sp = s^3 t_2 + s^2 t_3 + st_1$ if $p = s^2 t_2 + st_3 + t_1$.

As the polynomial difference ring $R^{(m)}$ is not Noetherian [5, Chapter 2, Ex. 3], we would have difficulties to compute in $R^{(m)}$ directly. Therefore we introduce restrictions $R_r^{(m)}$ of $R^{(m)}$ where only shifts up to some finite order $r \in \mathbb{N}$ appear. We define

$$R_r^{(m)} := K[t_{1,0}, t_{2,0}, \ldots, t_{m,0}, \ldots\ldots t_{1,r}, \ldots, t_{m,r}] \subseteq R^{(m)} \qquad (r \in \mathbb{N}).$$

We view $R_r^{(m)}$ as a polynomial ring over K in $m(r+1)$ indeterminates. Note that the shift operator s on $R^{(m)}$ does not induce a shift operator of $R_r^{(m)}$ because $s(t_{i,r}) \notin R_r^{(m)}$ $(i = 1, \ldots, m)$. We call r the *order* and m the *depth* of $R_r^{(m)}$. As m will be fixed in all our considerations, we can safely write $R_r := R_r^{(m)}$ for short.

A *difference ideal* I in some difference ring R is an ideal in R such that $sI \subseteq I$. If $S \subseteq R$ is any subset of R and I is the intersection of all difference ideals in R containing S, then we say I is *generated* by S and S is called a *basis* of I. We write $I = \langle\!\langle S \rangle\!\rangle$. Note that if S itself has the property that $sS \subseteq S$ then I is the usual ring ideal generation of S, denoted by $I = \langle S \rangle$.

For $S \subseteq R_r$, we write $\langle S \rangle_r$ for the polynomial ideal generated by S in R_r. Note that $\langle S \rangle_r \neq \langle S \rangle_{r+1}$ unless S is trivial. The notation $I \trianglelefteq R_r$ expresses that the set $I \subseteq R_r$ is a polynomial ideal in R_r, i.e., $I = \langle I \rangle_r$. If $I \trianglelefteq R_r$ is some ideal, then $\operatorname{Rad} I$ denotes the radical of I.

3 Nested Polynomially Recurrent Sequences

We have already fixed a computable algebraically closed field K of characteristic zero in the previous section. Let the class $\mathcal{C}$ of sequences $\mathbb{N} \to K$ be defined by structural induction as follows. Let $f_1, \dots, f_s \in \mathcal{C}$ and $p \in K[Y_1, \dots, Y_s, X_1, \dots, X_{r-1}]$ ($s, r \geq 0$ fixed). A sequence $f = (f(n))_{n=1}^{\infty}$ in K belongs to $\mathcal{C}$ if it satisfies (1) the recurrence

$$f(n+r) = p\Big(f_1(n), \dots, f_s(n),\ f(n), f(n+1), \dots, f(n+r-1)\Big)$$

or (2) it satisfies the recurrence

$$f(n+r) = \frac{1}{p\Big(f_1(n), \dots, f_s(n),\ f(n), f(n+1), \dots, f(n+r-1)\Big)}.$$

Note that the base of the induction is given by the case $s = 0$.

We call $\mathcal{C}$ the class of *nested polynomially recurrent sequences.* Its elements are "polynomially recurrent" in the sense that the polynomial p is not limited to polynomials that are linear in $f(n+i)$. Nonlinear recurrences are allowed as well. The term "nested" reflects the fact that the definition of f may involve other nested polynomially recurrent functions.

The number r in the definition is called the *order* of f. The sequences $f_1, \dots, f_s$ from the definition are called *subexpressions* of f, and the notion of subexpression is understood transitively. The total number m of subexpressions of f is called the *depth* of f.

Note that the notions of depth and order depend on the definition of a sequence rather than on the sequence itself, e.g., $(1/n - 1/n)_{n=1}^{\infty}$ has order 0 and depth 2, $(f(n))_{n=1}^{\infty}$ with $f(n+2) := 3f(n)$, $f(1) = f(2) = 0$ has order 2 and depth 1, and $(0)_{n=1}^{\infty}$ has depth 0 and order 0, yet all three sequences are equal, only their representations differ.

The class $\mathcal{C}$ contains a large variety of sequences which appear frequently in practice. It is immediate that all holonomic sequences are contained in $\mathcal{C}$. In addition, $\mathcal{C}$ contains sequences like $(\alpha^{\beta^n})_{n=1}^{\infty}$ ($\alpha \in K, \beta \in \mathbb{Z}$; by $f(n+1) = f(n)^{\beta}$) or $(\alpha^{\mathrm{F}(n)})_{n=1}^{\infty}$ ($\alpha \in K$, $\mathrm{F}(n)$ the nth Fibonacci number; by $f(n+2) =$

$f(n+1)f(n))$ which are easily seen not to be holonomic. A class of nonlinear recurrent sequences which arise in combinatorial and number theoretic considerations is studied in [9,10]. These sequences satisfy recurrences of the form $f(n+1) = f(n)^2 + \alpha f(n) + \beta$ for certain $\alpha, \beta \in \mathbb{Q}$, thus they are members of $\mathcal{C}$. An example is Sylvester's sequence defined by $s(n+1) = s(n)^2 - s(n) + 1$, $s(0) = 0$ (cf. [11, M0865]).

The Handbook of Mathematical Functions [8] contains a lot of families $f_n(x)$ of special functions which, for fixed x, admit sequences in n which belong to $\mathcal{C}$. Examples include

$$\begin{aligned} E_n(x) &:= \int_1^\infty t^{-n} \exp(-xt)\, dt \\ \text{by} \qquad E_{n+1}(x) &= \frac{1}{n}(\exp(-x) - xE_n(x)), \\ \Gamma(n,x) &:= \int_x^\infty t^{n-1} \exp(-t)\, dt \\ \text{by} \qquad \Gamma(n,x) &= \Gamma(n-1,x)(n-1) + x^{n-1}\exp(-x) \\ Q(\chi^2|n) &:= 1 - \frac{1}{2^{n/2}\Gamma(n/2)} \int_0^{\chi^2} t^{n/2-1} \exp(-t/2)\, dt \\ \text{by} \quad Q(\chi^2|n+2) &= Q(\chi^2|n) + \frac{(\chi^2/2)^{n/2}\exp(-\chi^2/2)}{\Gamma(n/2+1)}, \end{aligned}$$

as well as Bessel functions $J_n(x)$, their integrals, Struve functions, orthogonal polynomials, etc. If these functions can be handled for every fixed $n \in \mathbb{N}$, then our algorithm provides a tool to prove relations among them for general n.

It is clear that $\mathcal{C}$ is closed under field operations provided that they are meaningful, i.e., denominators must not vanish anywhere on the natural numbers. It is also quite clear that $\mathcal{C}$ is closed under taking indefinite sums and products, for $F(n) = \sum_{k=1}^n f(k)$ satisfies $F(n+1) = F(n) + f(n+1)$ and $F(n) = \prod_{k=1}^n f(k)$ satisfies $F(n+1) = f(n+1)F(n)$. It follows that $\mathcal{C}$ contains all $\Pi\Sigma$-sequences [12]. Many definite sums $F(n) = \sum_{k=1}^n f(k,n)$ obey linear recurrences and therefore also belong to $\mathcal{C}$. For large classes of summands $f(k,n)$, suitable recurrences for $F(n)$ can be computed by the methods of Schneider [13] or Zimmermann [14], and a recurrence computed by one of these methods can be used as a definition of F in the present context.

It may be remarked that $\mathcal{C}$ is also closed under taking indefinite continued fractions. Given $(f(n))_{n=1}^\infty \in \mathcal{C}$ with $f(n) \neq 0$ $(n \in \mathbb{N})$ we introduce the notation

$$F(n) := \mathop{\mathrm{K}}_{k=1}^{n} f(k) := f(1) + {}^{1}\!\big/ f(2) + {}^{1}\!\big/ \cdots + {}^{1}\!\big/ f(n).$$

It is at the heart of the theory of continued fractions [15] that $F(n)$ can be written as a quotient $\varkappa_1(n)/\varkappa_2(n)$ where

$$\begin{aligned}\varkappa_1(n+2) &= f(n+2)\varkappa_1(n+1)+\varkappa_1(n),\\ \varkappa_1(1) &= f(1), \quad \varkappa_1(2) = 1+f(1)f(2),\\ \varkappa_2(n+2) &= f(n+2)\varkappa_2(n+1)+\varkappa_2(n),\\ \varkappa_2(1) &= 1, \quad \varkappa_2(2) = f(2).\end{aligned}$$

Obviously, $(\varkappa_1(n))_{n=1}^\infty, (\varkappa_2(n))_{n=1}^\infty \in \mathcal{C}$, and as $\mathcal{C}$ is closed under arithmetic operations, it follows $(F(n))_{n=1}^\infty \in \mathcal{C}$.

4 Defining Relations

Let $f \in \mathcal{C}$ be given. The goal of this section is the construction of a finite set $D \subseteq R_r$ of *defining relations* for f, according to the following definition.

Definition 1 *Let $f_1, \dots, f_m \in \mathcal{C}$. A finite set $D \subseteq R_r$ is called a* set of defining relations *for $f_1, \dots, f_m$ if*

(1) For all $n \in \mathbb{N}$, the ideal

$$\left\langle D \cup \{s^i t_j - f_j(n+i) : i = 0, \dots, r-1, j = 1, \dots, m\}\right\rangle_r$$

has a unique solution, and in this solution $s^r t_j = f_j(n+r)$ $(j = 1, \dots, m)$.
(2) For every $j \in \{1, \dots, m\}$, there exists exactly one polynomial $p \in D$ of the form $p = s^r t_j + q$ or $p = qs^r t_j - 1$ where q depends only on $s^{r'} t_{j'}$ with $j' \leq j$ and $r' \leq r$, but not on $s^r t_j$. This p is called the defining polynomial *or the* defining relation *of $s^r t_j$.*
(3) For all $p \in R_{r-1}$, we have $p \in D \iff sp \in D$.

The number r is called the order *of D. We say that the sequence f_i* corresponds *to the variable t_i.*

If $f_1, \dots, f_{m-1}$ are the subexpressions of f_m, then we also say D is a set of defining relations for f_m.

We next collect some important properties of sets of defining relations. If we say that the set $A \cup sA \subseteq R_{r+1}$ is obtained on shift of the set $A \subseteq R$, then the essence of the following lemma is that the property of being a set of defining relations is preserved under shift.

Lemma 2 *Let $f \in \mathcal{C}$ be a nested polynomially recurrent sequence of depth m and $D \subseteq R_r$ be a set of defining relations for f. Let $f_1, \dots, f_m$ be the sequences corresponding to the variables $t_1, \dots, t_m$, respectively. Then $D \cup sD \subseteq R_{r+1}$ is also a set of defining relations for f.*

PROOF. It is clear that the conditions (2) and (3) of Def. 1 are satisfied for $D \cup sD$. As for (1), take an arbitrary $n \in \mathbb{N}$. Then,

$$\Big\langle D \cup \{s^i t_j - f_j(n+i) : i = 0, \ldots, r-1,\ j = 1, \ldots, m\}\Big\rangle_r$$

has a unique solution with $s^r t_j = f_j(n+r)$ $(j = 1, \ldots, m)$ and

$$\Big\langle sD \cup \{s^i t_j - f_j(n+i) : i = 0, \ldots, r,\ j = 1, \ldots, m\}\Big\rangle_{r+1}$$

has a unique solution with $s^{r+1} t_j = f_j(n+r+1)$ $(j = 1, \ldots, m)$. It follows that

$$\Big\langle (D \cup sD) \cup \{s^i t_j - f_j(n+i) : i = 0, \ldots, r,\ j = 1, \ldots, m\}\Big\rangle_{r+1}$$

also as a unique solution with $s^{r+1} t_j = f_j(n+r+1)$ $(j = 1, \ldots, m)$. □

The following lemma states that the polynomial ideal generated by a set of defining relations in some R_r coincides with the intersection of R_r with the difference ideal it generates in $R^{(m)}$. The proof of the lemma proceeds by considering lexicographic Gröbner bases and using the elimination property in polynomial rings with finitely many variables. As the lemma is not needed in the sequel, we omit the details of the proof.

Lemma 3 *If D is a set of defining relations of order r and $\langle\langle D\rangle\rangle$ is the difference ideal generated by D in $R^{(m)}$, then $\langle\langle D\rangle\rangle \cap R_r^{(m)} = \langle D\rangle_r$.*

We now turn to the construction of sets of defining relations for the elements of $\mathcal{C}$. Given $f \in \mathcal{C}$ with subexpressions $f_1, \ldots, f_{m-1}$, a set of defining relations can easily be obtained using the recurrences fulfilled by the f_i. Suppose the f_i are numbered such that all subexpressions of f_i are among the $f_1, \ldots, f_{i-1}$, let r_i be the order of f_i and put $r := \max_i r_i$. We distinguish two cases, according to the two cases in the definition of $\mathcal{C}$.

(1) $f_i(n+r_i) = p\big(f_1(n), \ldots, f_{i-1}(n), f_i(n), \ldots, f_i(n+r_i-1)\big)$ for some polynomial $p \in K[X_1, \ldots, X_{i-1}, Y_0, \ldots, Y_{r_i-1}]$. For each such f_i, let

$$d_i := s^{r_i} t_i - p(t_1, \ldots, t_{i-1}, t_i, \ldots, s^{r_i-1} t_i) \in R_{r_i}.$$

(2) $f_i(n+r_i) = 1/p\big(f_1(n), \ldots, f_{i-1}(n), f_i(n), \ldots, f_i(n+r_i-1)\big)$ for some polynomial $p \in K[X_1, \ldots, X_{i-1}, Y_0, \ldots, Y_{r_i-1}]$. For each such f_i, let

$$d_i := p(t_1, \ldots, t_{i-1}, t_i, \ldots, s^{r_i-1} t_i) s^{r_i} t_i - 1 \in R_{r_i}.$$

Using this notation, define $D \subseteq R_r$ as

$$D := \{d_1, sd_1, \ldots, s^{r-r_1} d_1,\ d_2, sd_2, \ldots, s^{r-r_2} d_2,\ \ldots\ldots,\ d_m, sd_m, \ldots, s^{r-r_m} d_m\}.$$

It is immediate by construction that D is a set of defining relations for f. In practice, we will of course represent common subexpressions by a single variable queue $t_i, st_i, sst_i, \ldots$ rather than by separate ones, and more subtle optimizations for reducing the number of variables are thinkable as well.

Example 4 *Consider the sequence* $(f(n))_{n=1}^{\infty} \in \mathcal{C}$ *defined by*

$$f(n) := \frac{\mathrm{F}(n)}{\mathrm{F}(n+1)} + \sum_{k=1}^{n} \frac{(-1)^k}{\mathrm{F}(k)\,\mathrm{F}(k+1)},$$

where $\mathrm{F}(n)$ *denotes the* n*th Fibonacci number. An appropriate set of defining relations for* f *is*

$$\begin{aligned}
D = \{&st_1 + t_1, sst_1 + st_1, && (\, t_1 \sim (-1)^n \,) \\
&sst_2 - st_2 - t_2, && (\, t_2 \sim \mathrm{F}(n) \,) \\
&sst_3 - st_3 - t_3, && (\, t_3 \sim \mathrm{F}(n+1) \,) \\
&t_4 t_3 - 1, st_4 st_3 - 1, sst_4 sst_3 - 1, && (\, t_4 \sim 1/\,\mathrm{F}(n+1) \,) \\
&t_5 t_2 t_3 - 1, st_5 st_2 st_3 - 1, sst_5 sst_2 sst_3 - 1, && (\, t_5 \sim 1/\,\mathrm{F}(n)\,\mathrm{F}(n+1) \,) \\
&st_6 - t_6 - st_1 st_5, sst_6 - st_6 - sst_1 sst_5, && (\, t_6 \sim \Sigma_{k=1}^{n} \ldots) \\
&t_7 - t_2 t_4 - t_6, st_7 - st_2 st_4 - st_6, && (\, t_7 \sim f(n) \,) \\
&\qquad sst_7 - sst_2 sst_4 - sst_6 \,\}
\end{aligned}$$

This representation has been obtained by mechanically applying the definitions of the various subexpressions, and representing identical subexpressions by the same variable. However, yet a "better" set of defining relations for f *can be obtained by exploiting that* $t_2 \sim \mathrm{F}(n)$ *implies* $st_2 \sim \mathrm{F}(n+1)$.

$$\begin{aligned}
D' = \{&t_1 + st_1, st_1 + sst_1, && (\, t_1 \sim (-1)^n \,) \\
&sst_2 - st_2 - t_2, && (\, t_2 \sim \mathrm{F}(n) \,) \\
&t_3 t_2 - 1, st_3 st_2 - 1, sst_3 sst_2 - 1, && (\, t_3 \sim 1/\,\mathrm{F}(n) \,) \\
&st_4 - t_4 - t_1 t_2 st_3, sst_4 - st_4 - st_1 st_2 sst_3, && (\, t_4 \sim \Sigma_{k=1}^{n-1} \ldots) \\
&st_5 - t_2 st_3 - st_4, sst_5 - st_2 sst_3 - sst_4 \,\} && (\, t_5 \sim f(n-1) \,).
\end{aligned}$$

5 Proving Zero Equivalence

We now turn to the algorithm for deciding $f \overset{?}{\equiv} 0$ for elements $f \in \mathcal{C}$ given by a set D of defining relations.

The key idea is an induction argument. The algorithm computes a number $k \in \mathbb{N}$ such that $f(n) = \cdots = f(n+k-1) = 0$ implies $f(n+k) = 0$ for arbitrary $n \in \mathbb{N}$. After that, f is evaluated at k consecutive points, and either there is a counterexample among these values, or there is no counterexample at all.

The algorithm reads as follows.

Algorithm 1

Input: f — *a nested polynomially recurrent sequence*

D — *a set of defining relations for* f

Output: *true or false, depending on whether* $f \equiv 0$ *or not*

Assumptions: t_m *corresponds to* $(f(n))_{n=1}^{\infty}$, D *is of order* r

```
1  function isZeroEquivalent($f, D$)
2     $k \leftarrow 0$
3     $I_0 \leftarrow \langle D \rangle_r + \langle t_m, st_m, \ldots, s^{r-1}t_m \rangle_r$
4     while $s^{k+r}t_m \notin \mathrm{Rad}\, I_k$ do
5        $k \leftarrow k+1$
6        $I_k \leftarrow \langle I_{k-1} \rangle_{k+r} + \langle s^{r+k-1}t_m \rangle_{k+r} + \langle s^k D \rangle_{k+r}$
7     end do
8     for $n$ from 1 to $k+r$ do
9        if $f(n) \neq 0$ then
10          return false, counterexample $= n$
11    return true
```

The rest of this section consists of the proofs for correctness and termination of Algorithm 1.

Theorem 5 *Algorithm 1 is correct.*

PROOF. It is clear that $f \not\equiv 0$ whenever the algorithm returns "false" because this does only happen when a counterexample has been found. Now suppose the algorithm returns "true". We will prove $f(n) = 0$ $(n \in \mathbb{N})$ by induction on n.

First, according to lines 8–10, we have $f(1) = \cdots = f(k+r) = 0$ as base of the induction. Now let $n \in \mathbb{N}$ be arbitrary such that $f(n) = f(n+1) = \cdots = f(n+k+r-1) = 0$. Prove $f(n+k+r) = 0$.

By repeated application of Lemma 2, $D \cup sD \cup \cdots \cup s^k D$ is a set of defining relations for f because D is. Let f_j $(j = 1, \ldots, m)$ be the sequences corresponding to the variables t_j $(j = 1, \ldots, m)$, respectively. By assumption of the algorithm, $f_m = f$. Condition (1) of Def. 1 asserts that the ideal

$$J = \left\langle D \cup \cdots \cup s^k D \cup \{s^i t_j - f_j(n+i) : i = 0, \ldots, k+r-1, j = 1, \ldots, m\} \right\rangle_{k+r}$$

has a unique solution, and this solution satisfies $s^{k+r}t_j = f_j(n+k+r)$. Now, by induction hypothesis, $I_k \subseteq J$. Every solution of J must be a solution of I_k as well. But by the termination condition in line 4, I_k has only solutions with $s^{k+r}t_m = 0$. This implies $f(n+k+r) = 0$. □

The next theorem will assert the termination of Algorithm 1. For its proof, we will need two technical lemmas.

Lemma 6 *Let $\mathfrak{p} \trianglelefteq K[X] =: K[x_1, \ldots, x_n]$ be a prime ideal. Then*

(1) For all $q \in K[X]$, the ideal $\mathfrak{p}' := \langle \mathfrak{p} \cup \{p\} \rangle \trianglelefteq K[X, y]$ with $p = y - q$ is prime and the quotient fields of the coordinate rings are isomorphic, $Q(K[X]/\mathfrak{p}) \cong Q(K[X,y]/\mathfrak{p}')$.

(2) For all $q \in K[X] \setminus \mathfrak{p}$, the ideal $\mathfrak{p}' := \langle \mathfrak{p} \cup \{p\} \rangle \trianglelefteq K[X, y]$ with $p = qy - 1$ is prime and $Q(K[X]/\mathfrak{p}) \cong Q(K[X,y]/\mathfrak{p}')$.

PROOF. Let $R = K[X]/\mathfrak{p}$ and $R' = K[X,y]/\mathfrak{p}'$.

(1) $p = y - q$ for $q \in K[X]$. Consider the homomorphisms ϕ, ψ defined by

$$\begin{array}{lll} \phi\colon K[X] \to R' & x_i \mapsto x_i \ (i = 1, \ldots, n), & \\ \psi\colon K[X,y] \to R & x_i \mapsto x_i \ (i = 1, \ldots, n), & y \mapsto q(x_1, \ldots, x_n). \end{array}$$

As $\mathfrak{p} \subseteq \ker \phi$ and $\mathfrak{p}' \subseteq \ker \psi$, these homomorphisms induce homomorphisms $\bar{\phi}\colon R \to R'$ and $\bar{\psi}\colon R' \to R$. $\bar{\phi}$ and $\bar{\psi}$ are inverses of each other because $\bar{\psi}(\bar{\phi}(x_i)) = \bar{\phi}(\bar{\psi}(x_i)) = x_i$ for $i = 1, \ldots, n$ and $\bar{\phi}(\bar{\psi}(y)) = \bar{\phi}(q) = q \equiv_{\mathfrak{p}'} y$. It follows that $R \cong R'$ and consequently $Q(R) \cong Q(R')$.

(2) $p = qy - 1$ for $q \in K[X] \setminus \mathfrak{p}$. Note that, by $\mathfrak{p}' = \langle \mathfrak{p} \rangle + \langle p \rangle$, $R' \cong R[y]/\langle p \rangle$. As $q \notin \mathfrak{p}$, there is some element $1/q \in Q(R)$. It suffices to show the existence of an embedding $R' \hookrightarrow Q(R)$, for then $R \hookrightarrow R' \hookrightarrow Q(R)$, and hence $Q(R) \hookrightarrow Q(R') \hookrightarrow Q(R)$, and hence $Q(R') \cong Q(R)$.

The evaluation homomorphism $\phi\colon R[y] \to Q(R)$, $\phi(y) = 1/q$ induces a homomorphism $\bar{\phi}\colon R' \to Q(R)$ because $\langle p \rangle \subseteq \ker \phi$. If furthermore $\ker \phi \subseteq \langle p \rangle$ then $\bar{\phi}$ is injective, and we are done. Indeed, $\ker \phi \subseteq \langle p \rangle$: Let $a = \sum_{i=0}^{n} a_i y^i \notin \langle qy - 1 \rangle$ be in canonical form, i.e., fully reduced wrt. $qy - 1$. Then $q \nmid a_n$ in $K[X]$. Suppose $a \in \ker \phi$, i.e.,

$$0 = \phi(a) = \sum_{i=0}^{n} a_i \phi(y)^i = \sum_{i=0}^{n} \frac{a_i}{q^i} = \frac{1}{q^n} \sum_{i=0}^{n} a_i q^{n-i}.$$

As $1/q^n \neq 0$, it follows that

$$0 = \sum_{i=0}^{n} a_i q^{n-i} = a_n + q \underbrace{\sum_{i=0}^{n-1} a_i q^{(n-1)-i}}_{\in K[X]},$$

and hence $q \mid a_n$, a contradiction. □

For the present context, the previous lemma provides an invariant property for certain ideals in R_k on extending them to ideals in R_{k+1}. This observation is crucial for the termination proof. Recall that the dimension of a primary ideal $\mathfrak{a} \trianglelefteq K[X]$ is defined as the transcendence degree of $Q(K[X]/\operatorname{Rad}\mathfrak{a})$ over K.

Lemma 7 *Let $D \subseteq R_r$ be a set of defining relations for some $f \in \mathcal{C}$, where r is the order of D. For all $n \geq 0$, define $I_n := \langle D \cup sD \cup \cdots \cup s^n D\rangle_{n+r}$. Furthermore, let $I_k \subseteq \mathfrak{a} \trianglelefteq R_{k+r}$ and $\mathfrak{a}' := \langle \mathfrak{a}\rangle_{k+r+1} + I_{k+1} \trianglelefteq R_{k+r+1}$ for some fixed $k \in \mathbb{N}$. Let $\mathfrak{a} = \bigcap_{i=1}^{s} \mathfrak{p}_i$ be a primary decomposition of $\mathfrak{a}$. Then*

(1) $\mathfrak{a}' = \bigcap_{i=1}^{s} \Big(\mathfrak{p}_i + I_{k+1}\Big)$ is a primary decomposition of $\mathfrak{a}'$ and

(2) for $i = 1, \ldots, s$ we have $\dim_{k+r} \mathfrak{p}_i = \dim_{k+r+1} \Big(\mathfrak{p}_i + I_{k+1}\Big)$.

PROOF. As $I_k \subseteq \mathfrak{a}$, we have $\mathfrak{a}' = \langle \mathfrak{a}\rangle_{k+r+1} + \langle d_1, \ldots, d_m\rangle_{k+r+1}$ where d_i is the defining relation of $s^{k+r+1}t_i$ $(i = 1, \ldots, m)$.

Consider the special case $\mathfrak{a}' = \langle \mathfrak{a}\rangle + \langle d\rangle \trianglelefteq R_{k+r}[s^{k+r+1}t_1]$. Applying Lemma 6 to the associated prime ideals of $\mathfrak{p}_i$, we obtain that $\operatorname{Rad}\mathfrak{p}_i + \langle d\rangle$ is prime and its function field is isomorphic to that of $\operatorname{Rad}\mathfrak{p}_i$, so in particular $\mathfrak{p}_i + \langle d\rangle$ is primary and $\dim \mathfrak{p}_i = \dim(\mathfrak{p}_i + \langle d\rangle)$.

The general case $\mathfrak{a}' = \langle \mathfrak{a}\rangle_{k+r+1} + \langle d_1, \ldots, d_m\rangle_{k+r+1}$ is proven by repeating the argument m times. Note that this is possible because d_i does not depend on variables $s^{k+r+1}t_j$ with $j > i$ by Def. 1.(2). □

Theorem 8 *Algorithm 1 terminates.*

PROOF. The only critical part is the loop in lines 4–7. We define an ordering $\prec$ on ideals as follows. Let $\mathfrak{a} \trianglelefteq A$, $\mathfrak{b} \trianglelefteq B$ be two ideals in some rings A, B. By a_d, b_d denote the respective number of primary components of dimension d in the ideals $\mathfrak{a}, \mathfrak{b}$. Let d_0 be the greatest integer such that $a_{d_0} \neq b_{d_0}$, or $d_0 = 0$ if $a_d = b_d$ for all d. Then we write $\mathfrak{a} \prec \mathfrak{b}$ iff $a_{d_0} < b_{d_0}$.

It suffices to show that $I_{k+1} \prec I_k$, because then the ideal sequence $I_1, I_2, \ldots$ computed by the algorithm is strictly decreasing wrt. $\prec$, and hence, by Dixon's lemma, it must be finite. Eventually, there will be a k with $I_k = \langle 1 \rangle$ and at least then the loop is left.

Suppose $s^{k+r}t_m \notin \operatorname{Rad} I_k$ at the end of the loop body, otherwise we are done. Let $I_k = \bigcap_i \mathfrak{p}_i$ be a primary decomposition of I_k. As $s^{k+r}t_m \notin \operatorname{Rad} I_k$, there must be some component $\mathfrak{p}_i$ with $s^{k+r}t_m \notin \operatorname{Rad} \mathfrak{p}_i$, and so $\dim(\mathfrak{p}_i + \langle s^{k+r}t_m \rangle_{k+r}) < \dim \mathfrak{p}_i$. It follows I_k has at least one component which is replaced by components of strictly lower dimension in $I_k + \langle s^{k+r}t_m \rangle_{k+r}$. Lemma 7 ensures that on passing from $I_k + \langle s^{k+r}t_m \rangle_{k+r}$ to the ideal $I_{k+1} = \langle I_k \rangle_{k+1+r} + \langle s^{k+r}t_m \rangle_{k+1+r} + \langle s^{k+1}D \rangle_{k+1+r}$, there will not appear new components, and the dimension of no component will increase. Therefore $I_{k+1} \prec I_k$ as claimed. □

6 Examples

Example 9 *(Example 4 continued) We apply Algorithm 1 to show*

$$f(n) = \frac{\mathrm{F}(n)}{\mathrm{F}(n+1)} + \sum_{k=1}^{n} \frac{(-1)^k}{\mathrm{F}(k)\,\mathrm{F}(k+1)} = 0 \qquad (1)$$

for all $n \in \mathbb{N}$. Using D from page 8 as set of defining relations, we find $sst_7 \notin \operatorname{Rad}\langle D \cup \{t_7, st_7\}\rangle_2$, $s^3t_7 \in \operatorname{Rad}\langle D \cup sD \cup \{t_7, st_7, sst_7\}\rangle_3$. We conclude $k = 1$ and we have to check $k + r = 3$ initial values. It is easily verified $f(1) = f(2) = f(3) = 0$, and this implies $f(n) = 0$ for all $n \in \mathbb{N}$. Using D' instead of D leads to the same result, but fewer variables may speed up the computations.

A careful inspection of the proofs in Section 5 shows that condition (2) of Def. 1 is only used in the termination proof, but not needed for the correctness. If we apply Algorithm 1 to a set $D \subseteq R_r$ which satisfies conditions (1) and (3) of Def. 1 and we obtain an answer, then this result is correct — we may, however, obtain no answer at all. The next example provides an application of this observation.

Example 10 *(from [16], Exercise 5.93) We want to show for all functions f and all $\alpha \neq 0$ the identity*

$$\sum_{k=1}^{n} \frac{\prod_{i=1}^{k-1} \big(f(i) + \alpha\big)}{\prod_{i=1}^{k} f(i)} = \frac{1}{\alpha}\Big(\prod_{k=1}^{n} \frac{f(k) + \alpha}{f(k)} - 1\Big). \qquad (2)$$

The idea is to omit the defining relation for the variable corresponding to $f(n)$.

There are two possible ways to treat the constant α: Either we regard it as a transcendental element and compute in $\mathbb{Q}(\alpha)$, or we regard it as a constant sequence and represent it by one of the variables t_i. We will follow the second approach, as it is more rigorous with respect to analytical correctness.

We use the following set as a set of defining relations. The variable t_3 will correspond to $f(n)$.

$$\begin{aligned}
D = \{&st_1 - t_1, && (\, t_1 \sim \alpha\,)\\
&t_2t_1 - 1, st_2st_1 - 1, && (\, t_2 \sim 1/\alpha\,)\\
&st_4 - t_4st_3, && (\, t_4 \sim \Pi f(n)\,)\\
&t_5t_3 - 1, st_5st_3 - 1, && (\, t_5 \sim 1/f(n)\,)\\
&st_6 - t_6(st_3 + st_1)st_5, && (\, t_6 \sim \Pi((f(n)+\alpha)/f(n))\,)\\
&st_7 - t_7(t_3 + t_1), && (\, t_7 \sim \Pi(f(n)+\alpha)\,)\\
&t_8t_4 - 1, st_8st_4 - 1, && (\, t_8 \sim 1/\Pi(f(n))\,)\\
&st_9 - t_9 - st_8st_7, && (\, t_9 \sim \Sigma(\Pi/\Pi)\,)\\
&t_{10} - t_9 + t_2(t_6 - 1), && (\, t_{10} \sim \textit{identity candidate})\\
&\quad st_{10} - st_9 + st_2(st_6 - 1)\,\}
\end{aligned}$$

It is easily checked that $st_{10} \notin \mathrm{Rad}\langle D \cup \{t_{10}\}\rangle_1$ and $sst_{10} \in \mathrm{Rad}\langle D \cup sD \cup \{t_{10}, st_{10}\}\rangle_2$. The loop terminates with $k = 1$ and we have to check $k + r = 2$ initial values. For $n = 1$, the left hand side evaluates to $1/f(1)$, and the right hand side evaluates to

$$\frac{1}{\alpha}\Big(\prod_{k=1}^{1}\frac{f(k)+\alpha}{f(k)} - 1\Big) = \frac{1}{\alpha}\Big(\frac{f(1)+\alpha}{f(1)} - 1\Big) = \frac{1}{f(1)}.$$

For $n = 2$, the left hand side evaluates to

$$\sum_{k=1}^{2}\frac{\prod_{i=1}^{k-1}\big(f(i)+\alpha\big)}{\prod_{i=1}^{k} f(i)} = \frac{1}{f(1)} + \frac{f(1)+\alpha}{f(1)f(2)} = \frac{\alpha+f(1)+f(2)}{f(1)f(2)}.$$

The right hand side evaluates to

$$\frac{1}{\alpha}\Big(\prod_{k=1}^{2}\frac{f(k)+\alpha}{f(k)} - 1\Big) = \frac{1}{\alpha}\Big(\frac{(f(1)+\alpha)(f(2)+\alpha)}{f(1)f(2)} - 1\Big) = \frac{\alpha+f(1)+f(2)}{f(1)f(2)}.$$

This completes the proof.

Despite the tremendous number of variables needed, an implementation completes the above examples virtually instantaneously — at least if special purpose software is used for the Gröbner basis computations in the radical membership test. We implemented the algorithm in the Maple system and used Faugère's Gb system [17] for Gröbner basis computations.

The examples above were selected in order to illustrate the computations of the algorithm in detail, the next example just lists some identities which we were able to check automatically using our algorithm. Note that these identities were up to now out of the scope of algorithmic computer proofs.

Example 11 *(1) Exercise 6.61 in [16]. If* $\mathrm{F}(n)$ *denotes the nth Fibonacci number then*

$$\sum_{k=0}^{n} \frac{1}{\mathrm{F}(2^k)} = 3 - \frac{\mathrm{F}(2^n - 1)}{\mathrm{F}(2^n)}. \tag{3}$$

(2) (5.1.45) in [8]. Let $E_n(x)$ *denote the nth exponential integral and* $\Gamma(n, x)$ *be the incomplete Gamma function (See page 3 for definitions and defining relations). Then*

$$E_n(x) = x^{n-1}\Gamma(1 - n, x) \tag{4}$$

(3) (26.4.5) in [8]. Let $Q(\chi^2|n)$ *be the quantile of the Chi Square distribution (cf. page 3), and let* $n!! = 2 \cdot 4 \cdot 6 \cdots n$ *for* $n \in \mathbb{N}$ *even. Then*

$$Q(\chi^2|n) = \exp(-\chi^2/2)\Big(1 + \sum_{r=1}^{n/2-1} \frac{(\chi^2)^r}{(2r)!!}\Big). \tag{5}$$

(4) Recall the notation $\mathrm{K}_{k=1}^{n} a_k$ *introduced in Section 3 for continued fractions, and let*

$$h(n) = \frac{(n+1)((-1)^n(\pi+2) + \pi - 2)\Gamma(\frac{k}{2})}{4\sqrt{\pi}\,\Gamma(\frac{n+1}{2})} \qquad (n \in \mathbb{N}).$$

Then

$$\sum_{k=0}^{n} \frac{1}{k!} = \mathop{\mathrm{K}}_{k=1}^{n} \Big(h(k) - \prod_{i=1}^{k} \big(-\tfrac{7}{4}i^2 + 9i - \tfrac{45}{4}\big)\Big). \tag{6}$$

Though our main interest is not in efficiency, we want to point out some timings as evidence that Algorithm 1 is of practical relevance. Our timings are taken on a 2.4GHz machine with 1Gb of memory. In the table below, m, r, d denote the depth, order and maximum total degree of the used set of defining relations, respectively. The number k is as in Algorithm 1, and t is the approximate CPU time in seconds which was required to compute k.

Eq.	(1)	(2)	(3)	(4)	(5)	(6)
m	7	10	7	9	11	11
r	2	1	1	1	2	2
d	3	2	2	3	3	4
k	1	1	1	3	3	2
t	0	0	0	6	67	8*)

*) with slight human support

7 Conclusion

We have defined the class $\mathcal{C}$ of nested polynomially recurrent sequences and we have shown that zero equivalence is decidable on this class. This result in particular admits a new algorithm for proving combinatorial identities on a class of sequences that were formerly out of the scope of computer proofs. Although it is likely that there are more efficient ways to implement the algorithm, an unoptimized ad-hoc implementation of the algorithm is already capable of doing nontrivial examples in reasonable time, which underlines its practical relevance.

Acknowledgements

I am grateful to my advisor Peter Paule and my colleagues in his group at RISC for fruitful discussions on earlier drafts of this paper. I am also indebted to Gábor Bodnár for continuous support concerning commutative algebra, and to Tobias Beck for considerable simplifications of the original proofs.

References

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

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

[3] F. Chyzak, An extension of Zeilberger's fast algorithm to general holonomic functions, Discrete Mathematics 217 (2000) 115–134.

[4] R. P. Stanley, Enumerative Combinatorics, Volume 2, Cambridge Studies in Advanced Mathematics 62, Cambridge University Press, 1999.

[5] R. M. Cohn, Difference Algebra, Interscience Publishers, John Wiley & Sons, 1965.

[6] J. F. Ritt, Differential Algebra, American Mathematical Society, Colloquium Publications, 1950.

[7] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms, Springer, 1992.

[8] M. Abramowitz, I. A. Stegun, Handbook of Mathematical Functions, 9th Edition, Dover Publications, Inc., 1972.

[9] A. V. Aho, N. J. A. Sloane, Some doubly exponential sequences, Fibonacci Quarterly 11 (1973) 429–437.

[10] S. W. Golomb, On certain nonlinear recurring sequences, American Mathematical Monthly 70 (1963) 403–405.

[11] N. J. A. Sloane, S. Plouffe, The Encyclopedia of Integer Sequences, Academic Press, 1995, http://research.att.com/~njas/sequences/.

[12] M. Karr, Theory of summation in finite terms, Journal of Symbolic Computation 1 (3) (1985) 303–315.

[13] C. Scheider, Symbolic summation in difference fields, Ph.D. thesis, RISC-Linz, Johannes Kepler Universität Linz (2001).

[14] B. Zimmermann, Definite summation and integration of special functions, Ph.D. thesis, RISC-Linz, Johannes Kepler Universität Linz (to appear).

[15] O. Perron, Die Lehre von den Kettenbrüchen, 2nd Edition, Chelsea Publishing Company, 1929.

[16] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics, 2nd Edition, Addison-Wesley, 1989.

[17] J.-C. Faugère, Gb tutorial, http://www-calfor.lip6.fr/~jcf/ (2002).