Computing the Algebraic Relations of C-finite Sequences and Multisequences

Manuel Kauers [1] and Burkhard Zimmermann [2]

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

Abstract

We present an algorithm for computing generators for the ideal of algebraic relations among sequences which are given by homogeneous linear recurrence equations with constant coefficients. Knowing these generators makes it possible to use Gröbner basis methods for carrying out certain basic operations in the ring of such sequences effectively. In particular, one can answer the question whether a given sequence can be represented in terms of other given sequences. A collection of examples, which were done with an implementation of our algorithm, is included.

1 Introduction

A *C-finite sequence* over a field k is a function $a\colon \mathbb{Z} \to k$ which satisfies a linear homogeneous recurrence with constant coefficients $c_0, c_1, \dots, c_r \in k$ with $c_0 \neq 0$ and $c_r \neq 0$,

$$c_0 a(n) + c_1 a(n+1) + \cdots + c_r a(n+r) = 0 \qquad (n \in \mathbb{Z});$$

a *P-finite sequence* over k satisfies a recurrence of the same type, but with polynomial coefficients $c_i(n) \in k[n]$ (Zeilberger, 1990). Clearly, every C-finite sequence is P-finite. C-finite sequences are well studied in the literature (Everest et al., 2003). The most famous C-finite sequence is the sequence of Fibonacci numbers satisfying $F_{n+2} = F_{n+1} + F_n$ and $F_0 = 0, F_1 = 1$.

Email addresses: `manuel.kauers@risc.uni-linz.ac.at` (Manuel Kauers), `burkhard.zimmermann@risc.uni-linz.ac.at` (Burkhard Zimmermann).

[1] Partially supported by the Austrian Science Foundation (FWF) grants F1305 and P16613-N12.
[2] Supported by the Austrian Science Foundation (FWF) grant F1301.

An *algebraic relation over* k among r sequences $a_1, \dots, a_r \colon \mathbb{Z} \to k$ is a polynomial $f \in k[x_1, \dots, x_r]$ such that $f(a_1(n), \dots, a_r(n)) = 0$ for all $n \in \mathbb{Z}$. For instance, the polynomial $x_1x_2 - x_3^2 - x_4$ is an algebraic relation over $\mathbb{Q}$ among the four sequences F_{n-1}, F_{n+1}, F_n and $(-1)^n$ by Cassini's identity $F_{n-1}F_{n+1} - F_n^2 = (-1)^n$.

It is sometimes of interest to decide whether or not a given polynomial is an algebraic relation of given sequences. This is trivial for the case of C-finite (Nemes and Petkovšek, 1995) sequences and, nowadays, routine for P-finite sequences (Salvy and Zimmermann, 1994) and many other classes of sequences. However, finding the algebraic relations among given sequences in the first place is a completely different task. Note that the set of algebraic relations among sequences $a_1, \dots, a_r$ forms an ideal of $k[x_1, \dots, x_r]$. The aim of this paper is to give algorithms for computing generators for this ideal in the case of C-finite sequences (Section 4) and C-finite multisequences (Section 7).

Let $k[a_1, \dots, a_r]$ be the smallest subring of $k^{\mathbb{Z}}$ containing the sequences $a_1, \dots, a_r$ and all constant sequences, and let I be the ideal of all algebraic relations among $a_1, \dots, a_r$. A Gröbner basis (Buchberger, 1965; Adams and Loustaunau, 1994) of I allows us to compute in $k[a_1, \dots, a_r]$ via the presentation by generators and relations

$$k[a_1, \dots, a_r] \simeq k[x_1, \dots, x_r]/I.$$

In particular, we can carry out addition, multiplication and canonical simplification effectively. Moreover, the question of whether a given C-finite sequence is representable in terms of other given C-finite sequences can be answered. The following is a typical example.
Example 1. (Graham et al., 1994, Exercise 7.26).

The second-order Fibonacci numbers $\mathfrak{F}_n$ are defined by the recurrence

$$\mathfrak{F}_n = \mathfrak{F}_{n-1} + \mathfrak{F}_{n-2} + F_n \quad (n \geq 2), \qquad \mathfrak{F}_0 = 0, \mathfrak{F}_1 = 1.$$

Express $\mathfrak{F}_n$ in terms of the usual Fibonacci numbers F_n and F_{n+1}.

It is an easy matter to compute the recurrence

$$\mathfrak{F}_{n+4} = 2\mathfrak{F}_{n+3} + \mathfrak{F}_{n+2} - 2\mathfrak{F}_{n+1} - \mathfrak{F}_n \quad (n \geq 0);$$

we use this recurrence as the "C-finite definition" of the second order Fibonacci numbers $\mathfrak{F}_n$. Using Algorithm RatRep, it is a matter of less than a second to prove that $\mathfrak{F}_n$ cannot be represented as a rational function in F_n and F_{n+1} alone; and Algorithm AlgRep tells us that $\mathfrak{F}_n$ cannot even be represented by an algebraic function in F_n and F_{n+1}. However, $\mathfrak{F}_n$ can be expressed as a polynomial in F_n, F_{n+1} and n, and algorithm PolyRep finds the representation $\mathfrak{F}_n = \frac{1}{5}(2(n+1)F_n + nF_{n+1})$; see Section 8 for details. No other algorithm is known to us which provides both the negative and the positive answers. $\square$

Countless identities in the literature on Fibonacci numbers (Hoggatt, 1979) are algebraic relations among C-finite sequences of several arguments; Catalan's identity

$$F_n^2 - F_{n+m}F_{n-m} = (-1)^{n-m}F_m^2, \tag{1}$$

a typical example. With Algorithm 3 (Section 7) all such identities can be found – and proved – automatically.

2 Problem Specification

In this section, we give a concrete description of the problem that we are dealing with. The *shift operator* E is defined on univariate sequences $a\colon \mathbb{Z} \to k$ by

$$(E \cdot a)(n) = a(n+1) \qquad (n \in \mathbb{Z}).$$

Polynomials in $k[E]$ represent linear constant coefficient recurrence operators. For instance, $(E^2 - E - 1) \cdot F = 0$ is the recurrence $F_{n+2} - F_{n+1} - F_n = 0$ in operator notation. The *i-th partial shift operator* E_i is defined on multisequences $a\colon \mathbb{Z}^d \to k$ by

$$(E_i \cdot a)(n_1, \ldots, n_i, \ldots, n_d) := a(n_1, \ldots, n_i + 1, \ldots, n_d) \qquad (n_1, \ldots, n_d \in \mathbb{Z}).$$

Following Zeilberger (1990), we define:

Definition 1 (C-finite sequences and multisequences). A sequence $a\colon \mathbb{Z} \to k$ is *C-finite over* k iff it is annihilated by some nonzero operator $P \in k[E]$:

$$P \cdot a = 0, \quad P \in k[E], \quad P \neq 0.$$

A multisequence $a\colon \mathbb{Z}^d \to k$ is *C-finite over* k iff for each i with $1 \leq i \leq d$ there is a nonzero operator P_i in $k[E_i]$ such that

$$P_i \cdot a = 0.$$

If $a : \mathbb{Z} \to k$ is a C-finite sequence and $\alpha_1, \ldots, \alpha_d$ are integers, then

$$b(n_1, \ldots, n_d) = a(\alpha_1 n_1 + \cdots + \alpha_d n_d)$$

is a C-finite multisequence.

Definition 2 (Algebraic Relations). Let $k \subseteq K$ be fields and let S be a set. The *ideal of algebraic relations over* k among functions $a_1, \ldots, a_r\colon S \to K$ is the kernel of the ring map $\varphi\colon k[x_1, \ldots, x_r] \to k^S$ which maps x_i to a_i for $1 \leq i \leq r$ and which maps elements of k to corresponding constant functions.

We denote it by $I(a_1, \ldots, a_r; k)$. Algebraic relations among sequences and multisequences are defined by taking $S = \mathbb{Z}$ and $S = \mathbb{Z}^d$ respectively.

By Hilbert's basis theorem, $I(a_1, \ldots, a_r; k)$ is finitely generated. The aim of this paper is to give an algorithm for computing generators for $I(a_1, \ldots, a_r; \mathbb{Q})$ in the case where $a_1, \ldots, a_r : \mathbb{Z}^d \to \mathbb{Q}$ are C-finite multisequences:

Problem MCRels.
Input: *C-finite multisequences $a_1, \ldots, a_r : \mathbb{Z}^d \to \mathbb{Q}$, where each sequence is given by d recurrences – one for each argument – and sufficiently many initial values.*
Output: *A set $\{g_1, \ldots, g_m\} \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ such that*

$$I(a_1, \ldots, a_r; \mathbb{Q}) = \langle g_1, \ldots, g_m \rangle$$

To solve Problem MCRels in full generality, we solve special cases of it first: The algorithm for the C-finite multisequences calls an algorithm for C-finite univariate sequences. That algorithm, in turn, calls an algorithm for the case of univariate geometric sequences. In summary, the problem reductions are:

GeoRels (Section 3) $\longleftarrow$ CRels (Section 4) $\longleftarrow$ MCRels (Section 7)

3 Relations among Geometric Sequences

Let $\bar{\mathbb{Q}}$ be the algebraic closure of $\mathbb{Q}$ and $\bar{\mathbb{Q}}^\times = \bar{\mathbb{Q}} \setminus \{0\}$. It is well-known that any C-finite sequence over $\mathbb{Q}$ can be represented in terms of various geometric sequences $n \mapsto \zeta^n$ with $\zeta \in \bar{\mathbb{Q}}^\times$ and the sequence $n \mapsto n$. (For the Fibonacci numbers, Binet's formula (7) gives such a representation.) We study the algebraic relations among such sequences.

Problem GeoRels.
Input: $\zeta_1, \ldots, \zeta_r \in \bar{\mathbb{Q}}^\times$.
Output: *A set $\{g_1, \ldots, g_m\} \subseteq \bar{\mathbb{Q}}[x_0, x_1, \ldots, x_r]$ such that*

$$I(n, \zeta_1^n, \ldots, \zeta_r^n; \bar{\mathbb{Q}}) = \langle g_1, \ldots, g_m \rangle$$

where x_0 corresponds to the arithmetic sequence $n \mapsto n$, and that x_i corresponds to the geometric sequence $n \mapsto \zeta_i^n$, for $i = 1, \ldots, r$.

Multiplicative relations among the numbers $\zeta_1, \ldots, \zeta_r$ immediately imply corresponding relations among the geometric sequences $\zeta_1^n, \ldots, \zeta_r^n$: A trivial calculation shows that

$$\prod_{i=1}^{r}(\zeta_i^n)^{a_i} - \prod_{i=1}^{r}(\zeta_i^n)^{b_i} = 0 \qquad (n \in \mathbb{Z}), \tag{2}$$

for any integers $a_1, \ldots, a_r$ and $b_1, \ldots, b_r$ satisfying

$$\prod_{i=1}^{r} \zeta_i^{a_i - b_i} = 1. \tag{3}$$

We recall the following usual definitions:
Definition 3. A *lattice* is a submodule of the $\mathbb{Z}$-module $\mathbb{Z}^r$. The *exponent lattice* of nonzero elements $\zeta_1, \ldots, \zeta_r$ of a field is given by

$$L(\zeta_1, \ldots, \zeta_r) := \left\{ (m_1, \ldots, m_r) \in \mathbb{Z}^r : \prod_{i=1}^{r} \zeta_i^{m_i} = 1 \right\}.$$

The *lattice ideal* $I(L)$ of a lattice $L \subseteq \mathbb{Z}^r$ is the ideal

$$I(L) := \left\langle \left\{ \prod_{i=1}^{r} x_i^{a_i} - \prod_{i=1}^{r} x_i^{b_i} : a \in \mathbb{N}^r, b \in \mathbb{N}^r, \text{ and } a - b \in L \right\} \right\rangle$$

of $\bar{\mathbb{Q}}[x_1, \ldots, x_r]$.

These definitions allow us to state (2)–(3) concisely as

$$I(\zeta_1^n, \ldots, \zeta_r^n; \bar{\mathbb{Q}}) \supseteq I(L(\zeta_1, \ldots, \zeta_r)). \tag{4}$$

In fact, equality holds true in (4), and throwing in the linear sequence $n \mapsto n$ does not introduce any new relations:
Proposition 1. *The relations among the $r+1$ sequences $n, \zeta_1^n, \ldots, \zeta_r^n$ over $\bar{\mathbb{Q}}$ form the ideal of $R := \bar{\mathbb{Q}}[x_0, x_1, \ldots, x_r]$ generated by the lattice ideal of the exponent lattice of $\zeta_1, \ldots, \zeta_r$:*

$$I(n, \zeta_1^n, \ldots, \zeta_r^n; \bar{\mathbb{Q}}) = R\, I(L(\zeta_1, \ldots, \zeta_r))$$

Proof. Let $I := I(n, \zeta_1^n, \ldots, \zeta_r^n; \bar{\mathbb{Q}})$ and $J := R\ I(L(\zeta_1, \ldots, \zeta_r))$. We already know that $I \supseteq J$ by (2)–(3). It remains to show $I \subseteq J$. Let G be a Gröbner basis of J with respect to some fixed term order $\prec$. We show that we can reduce any $f \in I$ to 0 by G. Let $f \in I$ be arbitrary. Assume that f is totally reduced by G. We have to show that $f = 0$. Write f as

$$f = \sum_{a \in S} f_a(x_0) \prod_{i=1}^{r} x_i^{a_i}$$

with a minimal $S \subseteq \mathbb{Z}^r$, i.e., with $f_a \neq 0$ for $a \in S$. Since $f \in I$,

$$\sum_{a \in S} f_a(n) \left(\prod_{i=1}^{r} \zeta_i^{a_i} \right)^n = 0 \tag{5}$$

for all integers n. In (5), the bases $\prod_{i=1}^r \zeta_i^{a_i}$ of the geometric sequences are pairwise distinct. (Suppose, to the contrary, that $\prod_{i=1}^r \zeta_i^{a_i} = \prod_{i=1}^r \zeta_i^{b_i}$ for $a \neq b$ with $a \in S$ and $b \in S$. Then f would involve monomials $x_0^{a_0} \prod_{i=1}^r x_i^{a_i}$ and $x_0^{b_0} \prod_{i=1}^r x_i^{b_i}$ with $\prod_{i=1}^r x_i^{a_i} - \prod_{i=1}^r x_i^{b_i} \in J$, contradicting the assumption that f is totally reduced with respect to G.) Geometric sequences over a field k with pairwise distinct bases are linearly independent over $k[n]$ (for a proof of this basic fact, see, for instance, Milne-Thomson, 1933, Section 13.0). Therefore, (5) implies that $f_a = 0$ for all $a \in S$. But we assumed $f_a \neq 0$ for all $a \in S$. So $S = \emptyset$, which means that $f = 0$. $\square$

Algorithm 1 (GeoRels) is a straightforward implementation of Proposition 1. It builds on two procedures LatticeIdeal and ExponentLattice, which

Algorithm 1 Algebraic Relations among Geometric Sequences

Input: $\zeta_1, \ldots, \zeta_r \in \bar{\mathbb{Q}}^\times$.
Output: A set $\{g_1, \ldots, g_m\} \subseteq \bar{\mathbb{Q}}[x_0, x_1, \ldots, x_r]$ such that

$$I(n, \zeta_1^n, \ldots, \zeta_r^n; \bar{\mathbb{Q}}) = \langle g_1, \ldots, g_m \rangle .$$

```
function GeoRels(ζ_1, ..., ζ_r)
    L := ExponentLattice(ζ_1, ..., ζ_r)
    I := LatticeIdeal(L)
    return I
end function
```

solve the following problems:

Problem ExponentLattice.
Input: *A tuple $(\zeta_1, \ldots, \zeta_r)$ of algebraic numbers, none of them zero.*
Technically, the input consists of polynomials $p_1, \ldots, p_r$ in $\mathbb{Q}[x]$ and an irreducible polynomial q in $\mathbb{Q}[x]$. The algebraic numbers $\zeta_1, \ldots, \zeta_r$ are defined by $\zeta_i = p_i(\alpha)$ where α is a root of q. They are in $\mathbb{Q}(\alpha)$.
Output: *A set $\{v_1, \ldots, v_n\} \subseteq \mathbb{Z}^r$ such that*

$$L(\zeta_1, \ldots, \zeta_r) = \mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n.$$

Problem LatticeIdeal.
Input: *A finite set $\{v_1, \ldots, v_n\}$ of vectors from $\mathbb{Z}^r$.*
Output: *A set $\{g_1, \ldots, g_m\} \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ such that*

$$I(\mathbb{Z}v_1 + \cdots + \mathbb{Z}v_n) = \langle g_1, \ldots, g_m \rangle .$$

Ge (1993) gives an efficient algorithm for solving Problem ExponentLattice. Algorithms for Problem LatticeIdeal can be found, for instance, in Hemmecke and Malkin (2005).

Example 2. What are the algebraic relations among n, ζ_+^n, ζ_-^n, and $(-1)^n$ over $\bar{\mathbb{Q}}$, where $\zeta_+ = (1+\sqrt{5})/2$ and $\zeta_- = (1-\sqrt{5})/2$? Ge's algorithm ExponentLattice delivers

$$L(\zeta_+, \zeta_-, -1) = (1,1,1)\mathbb{Z} + (0,0,2)\mathbb{Z}$$

corresponding to $\zeta_+\zeta_- = -1$ and $(-1)^2 = 1$. Calling LatticeIdeal on that lattice gives

$$I(n, \zeta_+^n, \zeta_-^n, (-1)^n; \bar{\mathbb{Q}}) = \langle y_1 y_2 - y_3, y_3^2 - 1\rangle$$

which means that all algebraic relations among n, ζ_+^n, ζ_-^n and $(-1)^n$ are consequences of $\zeta_+^n\zeta_-^n - (-1)^n = 0$ and $((-1)^n)^2 - 1 = 0$. □

4 Relations among C-finite Sequences over $\mathbb{Q}$

A fundamental and well known fact is that every C-finite sequence $a\colon \mathbb{Z} \to k$ can be written as a linear combination of geometric sequences with polynomial coefficients. If a satisfies the recurrence

$$c_0 a(n) + c_1 a(n+1) + \cdots + c_{r-1} a(n+r-1) + a(n+r) = 0 \quad (n \in \mathbb{Z})$$

then it has a representation of the form

$$a(n) = p_1(n)\zeta_1^n + \cdots + p_s(n)\zeta_s^n \quad (n \in \mathbb{Z}) \tag{6}$$

where $\zeta_1, \ldots, \zeta_s$ are the disctinct roots of the *characteristic polynomial*

$$c(z) = c_0 + c_1 z + \cdots + c_{r-1} z^{r-1} + z^r$$

and $p_i(n)$ is a polynomial in n whose degree is less than the multiplicity of the root ζ_i $(i = 1, \ldots, m)$. As we may assume $c_0 \neq 0$ without loss of generality, we can assume that all roots ζ_i be different from 0. Representation (6) allows us to reduce the problem of finding all relations among C-finite sequences (Problem CRels) to the problem of finding all relations among geometric sequences $\zeta_1^n, \ldots, \zeta_s^n$ and the arithmetic sequence n (Problem GeoRels).

Algorithm 2 (CRels) receives recurrences for $a_1, \ldots, a_r$ as input, and starts by solving them in terms of suitable geometric sequences ζ_i^n and the arithmetic sequence n (line 2). Next, it computes a set A of generators for the ideal $J := I(n, \zeta_1^n, \ldots, \zeta_s^n; \bar{\mathbb{Q}}) \subseteq \bar{\mathbb{Q}}[y_0, y_1, \ldots, y_s]$ of relations among these helper sequences (line 3) by calling Algorithm 1 (GeoRels). Since $a_j(n) = \sum_{i=1}^s p_{ij}(n)\zeta_i^n$, the ideal $I(a_1, \ldots, a_r; \bar{\mathbb{Q}})$ is the kernel of the ring map $\psi : \bar{\mathbb{Q}}[x_1, \ldots, x_r] \to \bar{\mathbb{Q}}[y_0, y_1, \ldots, y_s]/J$ given by

$$\psi(x_j) := \sum_{i=1}^s p_{ij}(y_0) y_i + J, \qquad \psi(c) = c + J \text{ for } c \in \bar{\mathbb{Q}}.$$

A set G of generators for this kernel is computed by elimination using a Gröbner basis (line 4 – line 7) with respect to a suitable elimination ordering; the technique used is based on (Adams and Loustaunau, 1994, Theorem 2.4.2).

Algorithm 2 Algebraic Relations among C-finite Sequences over $\mathbb{Q}$.

Input: A tuple of C-finite sequences $(a_1, \ldots, a_r)$ over $\mathbb{Q}$. Each sequence is given by a recurrence and initial values.
Output: A set $\{g_1, \ldots, g_m\} \subseteq \mathbb{Q}[x_1, \ldots, x_r]$,
such that

$$I(a_1, \ldots, a_r; \mathbb{Q}) = \langle g_1, \ldots, g_m \rangle .$$

1: **function** $\textsc{CRels}(a_1, \ldots, a_r)$
2: Compute $\zeta_i \in \bar{\mathbb{Q}}$ and $p_{ij} \in \bar{\mathbb{Q}}[y_0]$ for $i = 1, \ldots, s$ and $j = 1, \ldots, r$ such that $a_j(n) = \sum_{i=1}^{s} p_{ij}(n)\zeta_i^n$ for $j = 1, \ldots, r$ and every $n \in \mathbb{Z}$.
3: $A := \textsc{GeoRels}(\zeta_1, \ldots, \zeta_s)$, as an ideal of $\bar{\mathbb{Q}}[y_0, \ldots, y_s]$
4: $B := \{x_j - \sum_{i=1}^{s} p_{ij}(y_0)\, y_i : j = 1, \ldots, r\}$
5: Endow $R := \bar{\mathbb{Q}}[y_0, y_1, \ldots, y_s, x_1, \ldots, x_r]$ with an elimination order $\prec$ that has $y_0, y_1, \ldots, y_s$ higher than $x_1, \ldots, x_r$.
6: $G := \textsc{MonicReducedGröbnerBasis}(A \cup B)$ in R with respect to $\prec$
7: **return** $G \cap \mathbb{Q}[x_1, \ldots, x_r]$
8: **end function**

Example 3. What are the algebraic relations among F_n, F_{n+1}, and $(-1)^n$ over $\mathbb{Q}$, where F_n is the sequence of Fibonacci numbers?

Factorization of the characteristic polynomial $z^2 - z - 1$ and consideration of initial values gives Binet's formula

$$F_n = \frac{1}{\sqrt{5}}\zeta_+^n - \frac{1}{\sqrt{5}}\zeta_-^n, \quad F_{n+1} = \frac{1+\sqrt{5}}{2\sqrt{5}}\zeta_+^n - \frac{1-\sqrt{5}}{2\sqrt{5}}\zeta_-^n \qquad (n \in \mathbb{Z}), \tag{7}$$

where $\zeta_\pm = (1 \pm \sqrt{5})/2$ as in Example 2. There we got the result

$$I(n, \zeta_+^n, \zeta_-^n, (-1)^n; \bar{\mathbb{Q}}) = \langle y_1 y_2 - y_3, y_3^2 - 1 \rangle.$$

By elimination via Buchberger's algorithm,

$$\begin{aligned} &I(F_n, F_{n+1}, (-1)^n; \bar{\mathbb{Q}}) \\ &= \langle x_1 - \frac{1}{\sqrt{5}}y_1 + \frac{1}{\sqrt{5}}y_2,\ x_2 - \frac{1+\sqrt{5}}{2\sqrt{5}}y_1 + \frac{1-\sqrt{5}}{2\sqrt{5}}y_2,\ x_3 - y_3, \\ &\qquad y_1y_2 - y_3,\ y_3^2 - 1 \rangle \cap \bar{\mathbb{Q}}[x_1, x_2, x_3] \\ &= \langle x_1^2 + x_1x_2 - x_2^2 + x_3,\ x_3^2 - 1 \rangle. \end{aligned}$$

The generators of this ideal correspond to the identities

$$F_n^2 + F_nF_{n+1} - F_{n+1}^2 + (-1)^n = 0 \text{ and } ((-1)^n)^2 - 1 = 0;$$

all other polynomial identities among F_n, F_{n+1}, and $(-1)^n$ are consequences of those two. $\square$

By construction, Algorithm 2 (CRels) returns a set of generators $G \subseteq \bar{\mathbb{Q}}[x_1, \ldots, x_r]$ for the ideal $I(a_1, \ldots, a_r; \bar{\mathbb{Q}})$ of $\bar{\mathbb{Q}}[x_1, \ldots, x_r]$. However, Problem CRels asks for generators $G \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ for the ideal $I(a_1, \ldots, a_r; \mathbb{Q})$ of $\mathbb{Q}[x_1, \ldots, x_r]$. For proving Algorithm 2 (CRels) correct in that sense (Theorem 1 below), we need two lemmata.

Lemma 1. *Let $f \in K[x_1, \ldots, x_r]$ be an algebraic relation of sequences $a_1, \ldots, a_r : \mathbb{Z} \to k$ where K is an extension field of k. Then f is a linear combination of algebraic relations whose coefficients are in k.*

Proof. As K is an extension field of k, we can write f as

$$f = \alpha_1 f_1 + \cdots + \alpha_m f_m \tag{8}$$

with $f_1, \ldots, f_m \in k[x_1, \ldots, x_r]$ and coefficients $\alpha_1, \ldots, \alpha_m \in K$ which are linearly independent over k. We show that $f_1, \ldots, f_m$ are algebraic relations of $a_1, \ldots, a_r$. Fix an arbitrary $n \in \mathbb{Z}$. As f is an algebraic relation, it follows by (8) that

$$\alpha_1 f_1(a_1(n), \ldots, a_r(n)) + \cdots + \alpha_k f_k(a_1(n), \ldots, a_r(n)) = 0.$$

Note that $f_i(a_1(n), \ldots, a_r(n)) \in k$ for $i = 1, \ldots, m$. As $\alpha_1, \ldots, \alpha_m$ are linearly independent over k, it follows that $f_i(a_1(n), \ldots, a_r(n)) = 0$ for $i = 1, \ldots, m$. Therefore, $f_1, \ldots, f_m$ are algebraic relations of $a_1, \ldots, a_r$. $\square$

Lemma 2. *Let $I \subseteq K[x_1, \ldots, x_r]$ be the ideal of algebraic relations over K among sequences $a_1, \ldots, a_r$ that take values in a subfield k of K. Then I has a finite set of generators in $k[x_1, \ldots, x_r]$, i.e., I is defined over k.*

Proof. By Hilbert's Basis Theorem, I is generated by finitely many elements of $K[x_1, \ldots, x_r]$.

In that ideal basis, we can replace each element $f \in K[x_1, \ldots, x_r]$ by elements $f_1, \ldots, f_m \in I \cap k[x_1, \ldots, x_r]$ according to Lemma 1. $\square$

Theorem 1. *Algorithm 2 (*CRels*) is correct. Its output G satisfies*

(1) $G \subseteq \mathbb{Q}[x_1, \ldots, x_r]$,
(2) G generates the ideal $I(a_1, \ldots, a_r; \mathbb{Q})$ of $\mathbb{Q}[x_1, \ldots, x_r]$.

Proof. 1. By Lemma 2 with $k = \mathbb{Q}$, $K = \bar{\mathbb{Q}}$ there is an $A \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ that generates $I(a_1, \ldots, a_r; \bar{\mathbb{Q}})$ over $\bar{\mathbb{Q}}$. Let B be the monic reduced Gröbner basis of A. As computing a Gröbner basis involves only field operations on the coefficient level, $B \subseteq \mathbb{Q}[x_1, \ldots, x_r]$, too. By construction, both G and B

are monic reduced Gröbner bases of $I(a_1, \ldots, a_r; \bar{\mathbb{Q}})$. Since the monic reduced Gröbner basis of an ideal is unique, $G = B$, and $G \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ follows.

2. Let $f \in I(a_1, \ldots, a_r; \mathbb{Q})$ be arbitrary. As $G = \{g_1, \ldots, g_m\}$ generates $I(a_1, \ldots, a_r; \bar{\mathbb{Q}})$ over $\bar{\mathbb{Q}}$, we can find, by reduction, cofactors $u_1, \ldots, u_m$ in $\bar{\mathbb{Q}}[x_1, \ldots, x_r]$ such that

$$f = u_1 g_1 + \cdots + u_m g_m. \tag{9}$$

But, in fact, $u_1, \ldots, u_m \in \mathbb{Q}[x_1, \ldots, x_r]$: Both f and $g_1, \ldots, g_m$ have coefficients in $\mathbb{Q}$, and reduction involves only rational operations on the coefficient level. By way of (9), G generates $I(a_1, \ldots, a_r; \mathbb{Q})$ over $\mathbb{Q}$. □

5 Separation of C-finite Multisequences

We say that a multisequence $a : \mathbb{Z}^d \to k$ is *quasiunivariate* if $a(n_1, \ldots, n_d)$ depends only on one of its d arguments, i.e., if there is an index i and a sequence $b : \mathbb{Z} \to k$ such that $a(n_1, \ldots, n_d) = b(n_i)$ for all $n_1, \ldots, n_d \in \mathbb{Z}$. In this section we show that any C-finite multisequence can be expressed as a polynomial in quasiunivariate C-finite multisequences (Theorem 2). We call such a representation *separated*. While this result is almost trivial, it is the key for reducing Problem MCRels to Problem CRels in Section 7. Note that separated representations are particular to C-finite multisequences; P-finite multisequences in general do not admit them.

Example 4. The well-know addition theorem for the Fibonacci numbers

$$F_{m+n} = F_{m+1}F_n + F_m F_{n+1} - F_m F_n$$

gives a separated representation for F_{m+n}. □

The C-finite sequences annihilated by a fixed recurrence operator $P \in k[E]$ of order r form an r-dimensional vector space over k. The sequences $e_{P,0}, \ldots, e_{P,r-1} : \mathbb{Z} \to k$ defined by the recurrence $P \cdot e_{P,i} = 0$ and the "canonical" initial values

$$e_{P,i}(n) = \begin{cases} 1 & \text{if } n = i \\ 0 & \text{if } n \neq i \end{cases} \qquad \text{for } 0 \leq n < r.$$

form a basis of this vector space. Indeed, any solution $a : \mathbb{Z} \to k$ of $P \cdot a = 0$ can be written as

$$a(n) = \sum_{0 \leq i < r} a(i) e_{P,i}(n) \qquad (n \in \mathbb{Z}). \tag{10}$$

(Equation (10) is true by induction on n. For the induction step, note that both sides of it satisfy the same order r recurrence given by P; for the induction base, note that both sides agree for $n = 0, 1, \ldots, r-1$.)

Lemma 3. *Let $a : \mathbb{Z}^d \to k$ be a C-finite multisequence satisfying the system of recurrences $P_1 \cdot a = 0, \ldots, P_d \cdot a = 0$ with $P_i \in k[E_i] \setminus \{0\}$ for $i = 1, \ldots, d$. Then*

$$a(n_1, \ldots, n_d) := \sum_{0 \le i_1 < r_1} \cdots \sum_{0 \le i_d < r_d} a(i_1, \ldots, i_d)\, e_{P_1, i_1}(n_1) \, \cdots \, e_{P_d, i_d}(n_d). \tag{11}$$

where $r_i = \deg P_i$ for $i = 1, \ldots, d$.

Proof. By induction on d. The induction base $d = 1$ is Equation (10). Let $(n_1, \ldots, n_{d-1}) \in \mathbb{Z}^{d-1}$ be arbitrary but fixed and consider $a(n_1, \ldots, n_{d-1}, n_d)$ as a univariate sequence in n_d. According to Equation (10), it has the representation

$$a(n_1, \ldots, n_{d-1}, n_d) = \sum_{0 \le i_d < r_d} a(n_1, \ldots, n_{d-1}, i_d) e_{P_d, i_d}(n_d). \tag{12}$$

As $(n_1, \ldots, n_{d-1})$ were arbitrary, (12) holds for all $(n_1, \ldots, n_d) \in \mathbb{Z}^d$. Consider the term $a(n_1, \ldots, n_{d-1}, i_d)$ appearing under the sum as a C-finite multisequence of $d-1$ arguments. By the induction hypothesis, it can be written as a $(d-1)$-fold sum of the shape (11). $\square$

Theorem 2. *Any C-finite multisequence can be separated: For any C-finite multisequence $a : \mathbb{Z}^d \to k$ there exists an $m \in \mathbb{N}$, C-finite sequences $b_1, \ldots, b_m : \mathbb{Z} \to k$ and a polynomial $f \in k[x_{11}, \ldots, x_{dm}]$ such that*

$$\begin{aligned} a(n_1, \ldots, n_d) = f(\; & b_1(n_1), \;\ldots,\; b_m(n_1), \\ & \quad\vdots \qquad\qquad\quad \vdots \\ & b_1(n_d), \;\ldots,\; b_m(n_d) \;) \end{aligned}$$

for all $(n_1, \ldots, n_d) \in \mathbb{Z}^d$.

Proof. Equation (11) in Lemma 3 gives a suitable representation. $\square$

Theorem 2 states that the set of quasiunivariate multisequences generates the ring of all C-finite multisequences. Note that Equation (11) shows how to compute quasiunivariate representations effectively.

6 Separation and Algebraic Relations

Separation leaves us with the problem of computing the ideal I_* of relations among quasiunivariate multisequences

$$\begin{array}{ccc} b_1(n_1), & \ldots, & b_m(n_1), \\ \vdots & & \vdots \\ b_1(n_d), & \ldots, & b_m(n_d) \end{array} \tag{13}$$

where $b_1, \ldots, b_m$ are C-finite. Computing the algebraic relations among the entries of a fixed row in this table is, essentially, a univariate problem; Algorithm CRels applies. Is I_* already generated by the union (taken over all the rows) of the relations among the entries in one row? In this section we prove that this is indeed the case. The next Lemma proves it in the special case $d = 2$.

Lemma 4. *Assume that the functions $a_1, \ldots, a_r : U \times V \to k$ depend only on their first argument, i.e., the one in U, while the functions $b_1, \ldots, b_s : U \times V \to k$ depend only on their second argument, i.e., the one in V. Let us write their algebraic relations in the ring $R = k[x_1, \ldots, x_r, y_1, \ldots, y_s]$ where x_i corresponds to a_i and y_j to b_j, for $i = 1, \ldots, r$ and $j = 1, \ldots, s$.*

(1) Let F be a Gröbner basis for $I(a_1, \ldots, a_r; k)$ and let G be a Gröbner basis for $I(b_1, \ldots, b_s; k)$ with respect to some fixed term order. Then $F \cup G$ is a Gröbner basis for $I(a_1, \ldots, a_r, b_1, \ldots, b_s; k)$.
(2) The relations among $a_1, \ldots, a_r, b_1, \ldots, b_s$ are generated by the relations among $a_1, \ldots, a_r$ together with the relations among $b_1, \ldots, b_s$:

$$I(a_1, \ldots, a_r, b_1, \ldots, b_s; k) = RI(a_1, \ldots, a_r; k) + RI(b_1, \ldots, b_s; k).$$

Proof. Part 2 immediately follows from Part 1; we prove Part 1.

Let $I_* = I(a_1, \ldots, a_r, b_1, \ldots, b_s; k)$. To show that $F \cup G$ is a Gröbner basis for $I_* := I(a_1, \ldots, a_r, b_1, \ldots, b_s; k)$, it suffices to show (a) that $F \cup G \subseteq I_*$ and (b) that any element of I_* reduces to 0 by $F \cup G$.

(a): $F \cup G \subseteq I_*$ since $F \subseteq I(a_1, \ldots, a_r; k) \subseteq I_*$ and $G \subseteq I(b_1, \ldots, b_s; k) \subseteq I_*$.

(b): Let $f \in I_*$ be fully reduced with respect to $F \cup G$. We have to show that $f = 0$. Fix an arbitrary $u \in U$. Define a ring map

$$\phi_u : k[x_1, \ldots, x_r, y_1, \ldots, y_s] \to k[y_1, \ldots, y_s]$$

fixing k by $\phi_u(x_i) = a_i(u)$ for $i = 1, \ldots, r$ and $\phi_u(y_i) = y_i$ for $i = 1, \ldots, s$. Note that $f \in I_*$ implies $\phi_u(f) \in I(b_1, \ldots, b_s; k)$. By assumption, f is fully reduced with respect to G. Since the head terms of elements of G involve only $y_1, \dots, y_s$ while they are free of $x_1, \dots, x_r$, this implies that also $\phi_u(f)$ is fully reduced with respect to G. As $\phi_u(f) \in I(b_1, \dots, b_s; k)$ is fully reduced by a Gröbner basis of $I(b_1, \dots, b_s; k)$, we know that, in fact, $\phi_u(f) = 0$.

Let us write the polynomial $f \in k[x_1, \dots, x_r, y_1, \dots, y_s]$ as a finite sum

$$f = \sum_{m \in \mathbb{N}^s} f_m \, y_1^{m_1} \dots y_s^{m_s} \tag{14}$$

with coefficient polynomials $f_m \in k[x_1, \dots, x_r]$. Since $\phi_u(f) = 0$, we have $\phi_u(f_m) = 0$ for all $m \in \mathbb{N}^s$. To show that $f = 0$, it remains to show that all coefficient polynomials f_m vanish. Fix an arbitrary m. As we have shown $\phi_u(f_m) = 0$ for an arbitrary $u \in U$, we know that $f_m \in I(a_1, \dots, a_r; k)$. Since, by assumption, f is fully reduced with respect to F, and since $F \subseteq k[x_1, \dots, x_r]$, we know by (14) that also f_m is fully reduced with respect to F. We have shown that $f_m \in I(a_1, \dots, a_r; k)$ is fully reduced with respect to a Gröbner basis of $I(a_1, \dots, a_r; k)$. Therefore, $f_m = 0$. $\square$

Generalizing Lemma 4 from functions of 2 to functions of d arguments is a simple matter of induction. The result is:

Theorem 3. *Consider an array*

$$\begin{array}{ccc} b_{11}(n_1), & \dots, & b_{1m}(n_1) \\ \vdots & & \vdots \\ b_{d1}(n_d), & \dots, & b_{dm}(n_d) \end{array}$$

of $d \times m$ quasiunivariate multisequences $b_{ij} : \mathbb{Z}^d \to k$, in which multisequences in the i-th row depend only on their i-th argument n_i. Let $I_i = I(b_{i1}, \dots, b_{im}; k) \subseteq k[y_{i1}, \dots, y_{im}]$ be the ideal of relations of the entries in the i-th row, and let $I_* = I(b_{11}, \dots, b_{dm}; k) \subseteq k[y_{11}, \dots, y_{dm}]$ be the ideal of relations of all the entries in the array. Then I_* is generated by $I_1, \dots, I_d$:

$$I_* = \sum_{i=1}^{d} k[y_{11}, \dots, y_{dm}] I_i.$$

Proof. By induction on d. For $d = 1$, there is nothing to prove. In the induction step from d to $d+1$, use Lemma 4 Part 2 with $U = \mathbb{Z}^d$, $V = \mathbb{Z}$, $(a_1, \dots, a_r) = (b_{1,1}, \dots, b_{d,m})$, and $(b_1, \dots, b_s) = (b_{d+1,1}, \dots, b_{d+1,m})$. $\square$

Example 5. Determine the ideal $I_* := I(F_m, F_{m+1}, (-1)^m, F_n, F_{n+1}, (-1)^n; \mathbb{Q}) \subseteq R := \mathbb{Q}[x_1, x_2, x_3, y_1, y_2, y_3]$. (Notation: F_m stands for the multisequence $(m, n) \mapsto F_m$ etc.)

By Example 3 (twice), both $I_1 := I(F_m, F_{m+1}, (-1)^m; \mathbb{Q}) \subseteq \mathbb{Q}[x_1, x_2, x_3]$ and $I_2 := I(F_n, F_{n+1}, (-1)^n; \mathbb{Q}) \subseteq \mathbb{Q}[y_1, y_2, y_3]$ are known. Clearly, I_* contains $RI_1 + RI_2$. The question is whether or not I_* contains anything beyond that. As F_m, F_{m+1} and $(-1)^m$ depend only on m while F_n, F_{n+1} and $(-1)^n$ depend only on n, this is not the case, by Lemma 4. Therefore,

$$I_* = \langle x_1^2 + x_1x_2 - x_2^2 + x_3,\ x_3^2 - 1,\ y_1^2 + y_1y_2 - y_2^2 + y_3,\ y_3^2 - 1\rangle.$$

□

7 Relations among C-finite Multisequences

Now we have all the tools for solving Problem MCRels. All we need to do is to combine separation (Section 5, Theorem 2) with Theorem 3 and Algorithm 2 (CRels); the result is Algorithm 3 below.

Algorithm 3 Algebraic Relations among C-finite Multisequences over $\mathbb{Q}$.

Input: C-finite multisequences $a_1, \ldots, a_r : \mathbb{Z}^d \to \mathbb{Q}$, where each sequence is given by d recurrences (one for each argument) and sufficiently many initial values.

Output: A finite set $G \subseteq \mathbb{Q}[x_1, \ldots, x_r]$ generating $I(a_1, \ldots, a_r; \mathbb{Q})$.

```
function MCRels(a_1, ..., a_r)
    Compute a separated representation for a_1, ..., a_r. It consists of polynomials p_1, ..., p_r ∈ Q[y_11, ..., y_dm] and univariate C-finite sequences b_1, ..., b_m : Z → Q such that
```

$$a_k(n_1, \ldots, n_d) = p_k\begin{pmatrix} b_1(n_1), & \ldots, & b_m(n_1), \\ \vdots & & \vdots \\ b_1(n_d), & \ldots, & b_m(n_d) \end{pmatrix}$$

```
    for k = 1, ..., r and all (n_1, ..., n_d) ∈ Z^d.
    F := CRels(b_1, ..., b_m), as an ideal of Q[z_1, ..., z_m].
    A := ⋃_{i=1}^{d} {f(y_i1, ..., y_im) : f ∈ F}
    B := {x_k − p_k : k = 1, ..., r}
    Endow R := Q[y_11, ..., y_dm; x_1, ..., x_r] with a term order ≺ for eliminating y_11, ..., y_dm.
    G := MonicReducedGröbnerBasis(A ∪ B) in R with respect to ≺
    return G ∩ Q[x_1, ..., x_r]
end function
```

Theorem 4. *Algorithm 3 is correct: Its output G generates $I(a_1, \dots, a_r; \mathbb{Q})$.*

Proof. By the correctness of Algorithm CRels and renaming of variables, the set $\{f(y_{i1}, \dots, y_{im}) : f \in F\}$ generates the ideal $I_i := I(b_1(n_i), \dots, b_m(n_i); \mathbb{Q}) \subseteq k[y_{i1}, \dots, y_{im}]$ for $i = 1, \dots, d$. By Theorem 3, this implies that A generates $I_* := I(b_1(n_1), \dots, b_m(n_d); \mathbb{Q})$. From the representation of $a_1, \dots, a_r$ in terms of $b_1(n_1), \dots, b_m(n_d)$ computed in step 2, it follows that $I(a_1, \dots, a_r; \mathbb{Q})$ is the kernel of the ring map $\psi : \bar{\mathbb{Q}}[x_1, \dots, x_r] \to \mathbb{Q}[y_{11}, \dots, y_{dm}]$ given by $\psi(x_j) := p_k + I_*$ for $j = 1, \dots, r$ and $\psi(c) = c + I_*$ for $c \in \mathbb{Q}$. By (Adams and Loustaunau, 1994, Theorem 2.4.2), the set G computed in Step 5 – Step 8 generates the kernel of ψ. $\square$

8 Finding Representations

It is sometimes of interest to know whether a given C-finite sequence can be represented in terms of other given C-finite sequences.

Problem Rep (**variants:** PolyRep/RatRep/AlgRep)**.**
Input: *A C-finite (multi-)sequence a and C-finite (multi-)sequences $b_1, \dots, b_r$.*
Output: *Either a polynomial (resp. a rational function, resp. an algebraic function) f in r variables such that*

$$a(n_1, \dots, n_d) = f(b_1(n_1, \dots, n_d), \dots, b_r(n_1, \dots, n_d)) \tag{15}$$

for all $(n_1, \dots, n_d) \in \mathbb{Z}^d$ or the string "no such representation exists."

All three variants of the problem can be easily solved by looking at a Gröbner basis of

$$I(a(n), b_1(n), \dots, b_r(n); k) \subseteq k[x_0, x_1, \dots, x_r]$$

with respect to an elimination ordering for the variable x_0 corresponding to $a(n)$:

(1) A polynomial $f \in \mathbb{Q}[x_1, \dots, x_r]$ such that (15) holds exists if and only if the reduced Gröbner basis contains a polynomial of the form $x_0 + q$ for some polynomial $q \in k[x_1, \dots, x_m]$; in this case, $f = -q$.
(2) A rational function $f \in \mathbb{Q}(x_1, \dots, x_r)$ such that (15) holds exists if and only if the Gröbner basis contains a polynomial of the form $px_0 + q$ for some polynomials $p, q \in k[x_1, \dots, x_m]$, $p \neq 0$; in this case, $f = -q/p$.
(3) An algebraic function $f(x_1, \dots, x_r)$ such that (15) holds exists if and only if the Gröbner basis contains a polynomial in which x_0 appears.

From another point of view, Problem Rep is about solving recurrences: We solve the defining recurrence of a in terms of the sequences $b_1, \dots, b_r$.

Example 1 *(continued from page 2).* A lexicographic Gröbner basis of $I(\mathfrak{F}(n), F_n, F_{n+1}; \mathbb{Q})$ with respect to $x_0 \succ x_1 \succ x_2$ is $\{-1 + x_1^4 + 2x_1^3x_2 - x_1^2x_2^2 - 2x_1x_2^3 + x_2^4\}$. As the generator of this ideal is free of x_0, we can conclude that there does not exist any algebraic function A with $\mathfrak{F}_n = A(F_n, F_{n+1})$.

Taking the arithmetic sequence $n \mapsto n$ into account, we find that a lexicographic Gröbner basis of $I(\mathfrak{F}(n), F_n, F_{n+1}, n; \mathbb{Q})$ with respect to $x_0 \succ x_1 \succ x_2 \succ x_3$ is $\{-5x_0+2x_1+2x_1x_3+x_2x_3, -1+x_1^4+2x_1^3x_2-x_1^2x_2^2-2x_1x_2^3+x_2^4, 16-40x_0x_1^3-60x_0x_1^2x_2-8x_1^3x_2+70x_0x_1x_2^2-12x_1^2x_2^2+45x_0x_2^3+14x_1x_2^3-16x_2^4+16x_3-25x_2^4x_3\}$, the first generator of which implies $\mathfrak{F}_n = \frac{1}{5}(2(n+1)F_n + nF_{n+1})$. $\square$

9 C-finite Sequences over $\mathbb{Q}(z_1, \ldots, z_n)$

So far, our algorithms deal with C-finite sequences over the field $\mathbb{Q}$ of rational numbers. In fact, they work also for C-finite sequences over the algebraic numbers $\bar{\mathbb{Q}}$ without any modification. In this section, we briefly sketch how to extend them to C-finite sequences over a field of rational functions $\mathbb{Q}(z_1, \ldots, z_n)$.

It turns out that the only problem with generalizing the algorithms from $\mathbb{Q}$ to $\mathbb{Q}(z_1, \ldots, z_n)$ is that Ge's algorithm ExponentLattice works for algebraic numbers $\zeta_1, \ldots, \zeta_r \in \mathbb{Q}[\alpha]^\times$ with $\alpha \in \bar{\mathbb{Q}}$, while for our present generalization we would need it for algebraic functions $\zeta_1, \ldots, \zeta_r \in \mathbb{Q}(z_1, \ldots, z_n)[\alpha]^\times$ with $\alpha \in \overline{\mathbb{Q}(z_1, \ldots, z_n)}$. There is a pragmatic approach for extending Ge's algorithm to the latter case: To get rid of the indeterminates $z_1, \ldots, z_n$, substitute randomly chosen rational numbers $z_1^{(1)}, \ldots, z_n^{(1)}$ for them in the defining relations of $\zeta_1, \ldots, \zeta_r$ and α. That way we obtain images $\zeta_1^{(1)}, \ldots, \zeta_r^{(1)} \in \mathbb{Q}[\alpha^{(1)}]^\times$, with $\alpha^{(1)} \in \bar{\mathbb{Q}}$, of $\zeta_1, \ldots, \zeta_r$, unless we run into a degenerate case, which we reject. Note that any multiplicative relation $\zeta_1^{m_1} \ldots \zeta_r^{m_r} = 1$ among $\zeta_1, \ldots, \zeta_r$ implies a corresponding relation $(\zeta_1^{(1)})^{m_1} \ldots (\zeta_r^{(1)})^{m_r} = 1$ among their images $\zeta_1^{(1)}, \ldots, \zeta_r^{(1)}$. Therefore, the lattice $L = L(\zeta_1, \ldots, \zeta_r)$ is contained in the lattice $L^{(1)} = L(\zeta_1^{(1)}, \ldots, \zeta_r^{(1)})$. Generators for $L^{(1)}$ can be computed by Ge's algorithm. In unlucky cases, the images $\zeta_1^{(1)}, \ldots, \zeta_r^{(1)}$ may satisfy additional multiplicative relations, and so we cannot conclude at this point that $L = L^{(1)}$. To make sure that we did not run into an unlucky case, all we have to do is to check membership in L for each generator $m \in \mathbb{Z}^r$ of $L^{(1)}$, i.e., to check that indeed $\zeta_1^{m_1} \ldots \zeta_r^{m_r} = 1$. This can be done, for instance, by an ideal membership test using Gröbner basis methods. If this check succeeds, ExponentLattice$(\zeta_1, \ldots, \zeta_r)$ finishes by returning the generators of $L = L^{(1)}$. Otherwise, in the unlucky case, the algorithm repeats the same steps with different values for $z_1, \ldots, z_n$, and so on. It seems that unlucky cases can be made unlikely by drawing $z_1, \ldots, z_n$ from a large enough (finite) subset of

$\mathbb{Q}^m$ with uniform probability. It would be interesting to find bounds for the probability of running into an unlucky case, or, better, to give a deterministic – but still efficient – algorithm.

In case we use N different images of $\zeta_1, \ldots, \zeta_r$, leading to N superlattices $L^{(1)}, \ldots, L^{(N)}$ of L, an optimization is possible: As a candidate for L, use their intersection $L^{(1)} \cap \cdots \cap L^{(N)}$, as it is, in general, smaller than each of them; Cohen (1993) describes how to intersect integer lattices.

Example 6. The Chebyshev polynomials of the first kind $T_n(z)$ are C-finite over $\mathbb{Q}(z)$:

$$T_{n+2}(z) - 2zT_{n+1}(z) + T_n(z) = 0 \qquad (n \in \mathbb{Z}).$$

With Algorithm MCREL we can compute

$$I(T_{n-m}(z), T_n(z), T_{m+n}(z), T_m(z); \mathbb{Q}(z)) =$$
$$\langle -x_1 - x_3 + 2x_2x_4, x_2^2 + x_4^2 - x_1x_3 - 1, -2x_4^3 + 2x_1x_3x_4 + 2x_4 - x_1x_2 - x_2x_3 \rangle.$$

The second generator gives the identity

$$T_m(z)^2 + T_n(z)^2 - T_{n-m}(z)T_{m+n}(z) - 1 = 0$$

which is a well-known analog of Catalan's identity (1) for the Chebyshev polynomials. □

10 Examples and Applications

If the ideal of algebraic relations of some C-finite sequences is explicitly known, then a lot of information about these sequences can be computed algorithmically.

Proving and Finding Identities. In order to decide whether a conjectured algebraic relation of some given C-finite multisequences holds, it suffices to compute the ideal of the algebraic relations of these sequences by Algorithm MCRELS and to check whether the polynomial corresponding to the conjectured identity belongs to that ideal. For instance, Catalan's identity (1) can be proved in that way. Textbooks on Fibonacci numbers (Hoggatt, 1979, e.g.) list dozens of such identities. More interesting might be that such identities can also be *found* in an automated way, provided that it is specified where to search. In order to find, for instance, an identity that relates F_n, F_m, F_{n+m}, F_{n-m}, $(-1)^n$ and $(-1)^m$, it is sufficient to compute

$$I(F_n, F_m, F_{n+m}, F_{n-m}, (-1)^n, (-1)^m; \mathbb{Q}).$$

The ideal basis returned by Algorithm MCRELS contains a polynomial corresponding to (1).

We are by no means restricted to the Fibonacci numbers. Many other combinatorial sequences also obey C-finite recurrences, and Algorithm 2 can be used to study their algebraic relations.
Example 7. The sequence $f(n)$ defined via

$$f(n) = 5f(n-1) - 7f(n-2) + 4f(n-3) \quad (n \geq 3), \qquad f(0) = \tfrac{5}{16}, f(1) = \tfrac{3}{4}, f(2) = 2$$

describes the number of HC-polyominoes for $n \geq 2$ (Stanley, 1997, Example 4.7.18). With Algorithm 2 (CRELS), we find that $f(n), f(n+1), f(n+2)$ are algebraically dependent with 2^n via

$$\begin{aligned}2^{2n} = {} & 256f(n)^3 - 896f(n)^2f(n+1) + 1104f(n)f(n+1)^2 - 496f(n+1)^3 \\ & + 320f(n)^2f(n+2) - 752f(n)f(n+1)f(n+2) + 512f(n+1)^2f(n+2) \\ & + 112f(n)f(n+2)^2 - 160f(n+1)f(n+2)^2 + 16f(n+2)^3 \quad (n \geq 0).\end{aligned}$$

This identity might not have been known before, and it seems hard to prove it in a combinatorial way.

With Algorithm ALGREP, we prove that $f(n)$ cannot be represented as an algebraic function in terms of $F_n, F_{n+1}, (-1)^n$ and n. We do not know of any other method – combinatorially or not – for proving the absence of such representations. □
Example 8. The "Tribonacci" numbers T_n (Sloane and Plouffe, 1995, A000073), defined via

$$T_{n+3} = T_n + T_{n+1} + T_{n+2} \quad T_0 = 0, T_1 = T_2 = 1,$$

satisfy the identity

$$T_{2n}^3 + T_n^2T_{4n} + 2T_{3n}T_{4n}T_{5n} + T_{2n}T_{4n}T_{6n} = 2T_nT_{2n}T_{3n} + T_{4n}^3 + T_{2n}T_{5n}^2 + T_{3n}^2T_{6n}.$$

This identity was discovered by Algorithm 2 (CRELS). It appeared, together with some further polynomials, as basis element of $I(T_n, T_{2n}, \ldots, T_{6n}; \mathbb{Q})$. □
Example 9. For the Perrin numbers P_n (Sloane and Plouffe, 1995, A001608), defined via

$$P_{n+3} = P_n + P_{n+1} \quad P_0 = 3, P_1 = 0, P_2 = 2,$$

we find

$$I(P_n, P_{2n}, P_{3n}; \mathbb{Q}) = \langle x_1^3 - 3x_1x_2 + 2x_3 - 6\rangle,$$

and hence the identity $P_n^3 - 3P_nP_{2n} + 2P_{3n} = 6$. □

Solving Recurrences

Example 10. It is easy to see that the sum

$$a(n) = \sum_{k=0}^{n} \binom{n}{k} F_k$$

satisfies

$$a(n+2) = 3a(n+1) - a(n), \quad a(0) = 0, \quad a(1) = 1.$$

Using Algorithm PolyRep, we can solve this recurrence in terms of Fibonacci numbers, i.e., $b_1(n) = F_n$ and $b_2(n) = F_{n+1}$, getting

$$a(n) = F_n(2F_{n+1} - F_n)$$

which is well-known. □

Example 11. The sum

$$a(n) = \sum_{k=0}^{n} \binom{n}{k} F_{n+k}$$

satisfies the recurrence

$$a(n+2) = 4a(n+1) + a(n) \quad a(0) = 0, \quad a(1) = 2.$$

Using Algorithm PolyRep, we find the representation

$$a(n) = F_n(2F_n^2 - 3F_nF_{n+1} + 3F_{n+1}^2).$$

□

Example 12. The sum

$$a(n) = \sum_{k=0}^{n} \binom{n}{k} F_{2k}$$

satisfies the recurrence

$$a(n+2) = 5a(n+1) - 5a(n) \quad a(0) = 0, \quad a(1) = 1.$$

Algorithm AlgRep proves that $a(n)$ cannot be written as an algebraic function in n, F_n, and F_{n+1}. □

Proving Divisibility Relations.

Example 13. In order to prove the divisibility property

$$L_n \mid L_{n+2m}^4 - (L_{2m}^2 - 4)^2 \qquad (n, m \geq 0) \tag{16}$$

for the Lucas numbers L_n defined by $L_{n+2} = L_{n+1} + L_n$, $L_0 = 2$, $L_1 = 1$, it suffices to find an identity of the form

$$L_{n+2m}^4 - (L_{2m}^2 - 4)^2 = q(n, m)L_n \qquad (n, m \geq 0)$$

for some integer sequence $q(n,m)$. If $q(n,m)$ can itself be expressed in terms of L_n, L_{2m}, and L_{n+2m}, then it can be computed. For, if

$$\mathfrak{a} := I(L_n, L_{2m}, L_{n+2m}; \mathbb{Q}) = \langle a_1, \ldots, a_\ell \rangle \trianglelefteq \mathbb{Q}[x_1, x_2, x_3],$$

then, by an extended Gröbner basis computation (Becker et al., 1993, Section 5.6) we can find polynomials $g_0, \ldots, g_\ell$ such that

$$x_3^4 - (x_2^2 - 4)^2 = x_1 g_0 + g_1 a_1 + \cdots + g_\ell a_\ell.$$

In this way, we have found that

$$q(n,m) = (L_n - 2L_{n+2m}L_{2m})(L_n^2 + 2L_{n+2m}^2 - L_n L_{2m+n} L_{2m})$$

does the job. (Observe that $q(n,m) \neq 0$ for all $n, m \geq 0$.)

In fact, the present example is even simpler: (16) follows by inspection from

$$\mathfrak{a} = \langle -16 + x_1^4 + 8x_2^2 - x_2^4 - 2x_1^3 x_2 x_3 + 2x_1^2 x_3^2 + x_1^2 x_2^2 x_3^2 - 2x_1 x_2 x_3^3 + x_3^4 \rangle.$$

□

Example 14. The problem proposed by Furdui (2002) can be treated in a similar way: Prove that $\gcd(L_n, F_{n+1}) = 1$ for all $n \geq 1$.

Using Algorithm CRels, we find that

$$I(L_n, F_{n+1}; \mathbb{Q}) = \langle x_1^4 - 10x_1^3 x_2 + 35x_1^2 x_2^2 - 50x_1 x_2^3 + 25x_2^4 - 1 \rangle.$$

Let us denote the generator of this ideal by g. An extended Gröbner basis computation shows that

$$1 = (x_1)^3 \cdot (x_1) + (-10x_1^3 + 35x_1^2 x_2 - 50x_1 x_2^2 + 25x_2^3) \cdot (x_2) + (-1) \cdot g.$$

Hence there are integer sequences $p(n), q(n)$ such that

$$1 = p(n)L_n + q(n)F_{n+1} + 0 \qquad (n \geq 1).$$

The claim follows. □

Example 15. For the sequence $a(n)$ defined via

$$a(n+2) = 5a(n+1) - a(n) \quad (n \geq 0), \qquad a(0) = a(1) = 1$$

we have

$$I(a(n), a(n+1); \mathbb{Q}) = \langle x_2^2 + x_1^2 + 3 - 5x_1 x_2 \rangle.$$

An immediate consequence is that $a(n)a(n+1) \mid a(n+1)^2 + a(n)^2 + 3$ for all $n \in \mathbb{N}$. Friendman (1995) has asked for a proof of this divisibility property. Such problems can easily be generated using our algorithm. □

11 An Implementation

A package for the computer algebra system Mathematica 5 implementing Algorithm MCRels is available for download at

http://www.risc.uni-linz.ac.at/research/combinat/software/

It provides a function "Dependencies" which computes the ideal of algebraic relations among a given list of C-finite multisequences over $\mathbb{Q}$. We illustrate the usage of this package by a short example, and refer to the user manual (Kauers and Zimmermann, 2005) for further information.

Example 1 *(continued).* In order to compute the algebraic relations among

$\mathfrak{F}(n)$, F_n, F_{n+1} and n, we type

$$\text{In[1]:=}\ \text{Dependencies}[\{\mathfrak{F}[n], \text{Fibonacci}[n], \text{Fibonacci}[n+1], n\}, x,\\ \text{Where} \to \{\mathfrak{F}[n+2] == \mathfrak{F}[n+1] + \mathfrak{F}[n] + \text{Fibonacci}[n+2],\\ \mathfrak{F}[0] == 0, \mathfrak{F}[1] == 1\}]$$

and obtain in less than a second the following basis:

$$\text{Out[1]=}\ \{-5x_1 + 2x_2 + 2x_2x_4 + x_3x_4, -1 + x_2^4 + 2x_2^3x_3 - x_2^2x_3^2 - 2x_2x_3^3 + x_3^4, 16 - 40x_1x_2^3 - 60x_1x_2^2x_3 - 8x_2^3x_3 + 70x_1x_2x_3^2 - 12x_2^2x_3^2 + 45x_1x_3^3 + 14x_2x_3^3 - 16x_3^4 + 16x_4 - 25x_3^4x_4\}$$

□

12 Further Work

Our algorithm depends heavily on the fact that linear recurrence equations (or differential equations) with constant coefficients admit closed form solutions in terms of exponentials and polynomials. In general, this is no longer true if the coefficients $c_i(n)$ in a recurrence equation

$$c_0(n)a(n) + c_1(n)a(n+1) + \cdots + c_r(n)a(n+r) = 0 \qquad (n \in \mathbb{Z})$$

can be polynomials in n. Solutions $a(n)$ of such recurrence equations are called P-finite. It would be very interesting to have an algorithm for computing the algebraic relations among given P-finite sequences. Such an algorithm would be extremely useful in the field of symbolic summation and integration of special functions.

Our initial motivation for studying algebraic relations came from symbolic summation. A known limitation of the celebrated Karr summation algorithm (Karr, 1981, 1985) is that the sequences appearing in the summand expression have to be algebraically independent, i.e., the Karr algorithm is limited to the transcendental case. In particular, sums like $\sum_{k=1}^{n}(-1)^k \sum_{i=1}^{k} 1/i$ involving $(-1)^k$ cannot be dealt with, due to the algebraic relation $((-1)^k)^2 = 1$. Using that particular relation, Schneider's extension (Schneider, 2001) of Karr's algorithm is able to deal with sums involving alternating signs $(-1)^k$, but no further progress has been made with respect to the non-transcendental case. We believe that this is partly due to the lack of an algorithm for computing the algebraic relations of the sequences involved. The algorithm described in the present paper could thus be useful for extending the Karr algorithm to summand expressions involving arbitrary C-finite sequences.

We did not analyze the complexity of our algorithms. The computation of a primitive element α for $\mathbb{Q}(\zeta_1, \ldots, \zeta_r)$, as required for Ge's algorithm, is costly and dominates the runtime in may cases. Experiments suggest that it is the runtime bottleneck if the degrees of the minimal polynomials for $\zeta_1, \ldots, \zeta_r$ exceeds approximately 15. Less frequently, the runtime bottleneck is the Gröbner basis computation in Algorithm 2.

Acknowledgement. We wish to thank Peter Paule for helpful discussion.

References

Adams, W. W., Loustaunau, P., 1994. An Introduction to Gröbner Bases. No. 3 in Graduate Studies in Math. Amer. Math. Soc., New York.

Becker, T., Weispfenning, V., Kredel, H., 1993. Gröbner Bases. Springer.

Buchberger, B., 1965. Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal. Ph.D. thesis, Universität Innsbruck.

Cohen, H., 1993. A Course in Computational Algebraic Number Theory. Springer.

Everest, G., van der Poorten, A., Shparlinski, I., Ward, T., 2003. Recurrence Sequences. Vol. 104 of Mathematical Surveys and Monographs. American Mathematical Society.

Friendman, J., 1995. Problem B-785. The Fibonacci Quarterly 33 (2).

Furdui, O., 2002. Problem B-931. The Fibonacci Quarterly 40 (1), 85.

Ge, G., 1993. Algorithms related to multiplicative representations of algebraic numbers. Ph.D. thesis, U.C. Berkeley.

Graham, R. L., Knuth, D. E., Patashnik, O., 1994. Concrete Mathematics, 2nd Edition. Addison-Wesley.

Hemmecke, R., Malkin, P., 2005. Computing generating sets of lattice ideal-sArXiv:math.CO/0508359.

Hoggatt, V. E., 1979. Fibonacci and Lucas numbers. The Fibonacci Association.

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

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

Kauers, M., Zimmermann, B., 2005. Dependencies — a Mathematica package for computing algebraic dependencies of C-finite sequences. Tech. rep., SFB F013, Johannes Kepler Universität, (in preparation).

Milne-Thomson, L. M., 1933. The Calculus of Finite Differences. Macmillan and Co., ltd.

Nemes, I., Petkovšek, M., 1995. Rcomp: A mathematica package for computing with recursive sequences. Journal of Symbolic Computation 20 (5–6), 745–753.

Salvy, B., Zimmermann, P., 1994. Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. ACM Transactions on Mathematical Software 20 (2), 163–177.

Schneider, C., 2001. Symbolic summation in difference fields. Ph.D. thesis, RISC-Linz, Johannes Kepler Universität Linz.

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

Stanley, R. P., 1997. Enumerative Combinatorics, Volume 1. Cambridge Studies in Advanced Mathematics 62. Cambridge University Press.

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