Order bounds for C^2-finite sequences

M. Kauers, P. Nuspl, V. Pillwein

February 2023

RISC Report Series No. 23-03

ISSN: 2791-4267 (online)

Available at https://doi.org/10.35011/risc.23-03

**JOHANNES KEPLER
UNIVERSITY LINZ**
Altenberger Str. 69
4040 Linz, Austria
www.jku.at
DVR 0093696

Order bounds for C^2-finite sequences*

Manuel Kauers[1], Philipp Nuspl[2], and Veronika Pillwein[2]

manuel.kauers@jku.at, philipp.nuspl@jku.at, veronika.pillwein@risc.jku.at

[1] *Johannes Kepler University Linz, Institute for Algebra*
[2] *Johannes Kepler University Linz, Research Institute for Symbolic Computation*

Abstract

A sequence is called C-finite if it satisfies a linear recurrence with constant coefficients. We study sequences which satisfy a linear recurrence with C-finite coefficients. Recently, it was shown that such C^2-finite sequences satisfy similar closure properties as C-finite sequences. In particular, they form a difference ring.

In this paper we present new techniques for performing these closure properties of C^2-finite sequences. These methods also allow us to derive order bounds which were not known before. Additionally, they provide more insight in the effectiveness of these computations.

The results are based on the exponent lattice of algebraic numbers. We present an iterative algorithm which can be used to compute bases of such lattices.

1 Introduction

Infinite objects that can be represented by a finite amount of information and that can be effectively computed with, e.g., by means of closure properties, are natural objects of study in symbolic computation. This includes in particular sequences that can be defined by linear recurrences with coefficients that, in turn, have a finite description. If these coefficients are polynomials, the sequences are called holonomic or D-finite and the special case of constant coefficients is referred to as C-finite sequences.

It is well known [13, 17] that these form classes that are closed under several operations such as addition, multiplication, interlacing, taking subsequences, etc. These closure properties are algorithmic, have been implemented in several computer algebra systems, and contribute to the "holonomic toolkit" [11] for automatically proving and deriving identities.

It has been shown [8, 9, 24, 22] that many of these closure properties also hold and can be implemented for sequences that are defined by linear recurrences with C-finite coefficients, also called C^2-finite. To our knowledge, C^2-finite sequences have first been introduced formally by Kotek and Makowsky [14] in the context of graph polynomials. Thanatipanonda and Zhang [29] give an overview on different properties of polynomial, C-finite and holonomic sequences and consider the extension under the name X-recursive sequences. A survey on closure properties of linear recurrence sequences including C^2-finite sequences is given by Krityakierne and Thanatipanonda [15].

The main computational issue when dealing with this new class of C^2-finite sequences is the presence of zero divisors. Even though it was shown that C^2-finite sequences form a difference ring, so far it was not clear whether their closure properties can be effectively computed, see also the discussion in Section 2.2 below.

In this paper, we introduce a new method for executing closure properties that, in particular, comes with order bounds. A key ingredient for this technique is the computation of a basis for the exponent lattice for the eigenvalues of the coefficient recurrences. For the computation, we

*The research was funded by the Austrian Science Fund (FWF) under the grant W1214-N15, project DK15 and under the grants P33530, P31571, and I6130.

introduce an iterative version of Ge's algorithm [7] described in Section 3. The resulting order bounds for the ring operations, interlacing, and taking subsequences are presented in Section 5.

2 Preliminaries

In this section we introduce some basic notation and definitions which are used throughout the paper. We denote the set of natural numbers by $\mathbb{N} = \{0, 1, 2, \dots\}$. Furthermore, $\mathbb{K} \supseteq \mathbb{Q}$ denotes an algebraic number field. The $\mathbb{K}$-algebra of sequences under termwise addition and termwise multiplication is denoted by $\mathbb{K}^{\mathbb{N}}$. For the sake of a cleaner notation, $c(n)$ can denote both a sequence $(c(n))_{n\in\mathbb{N}}$ and the term at index n. The meaning is always clear from the context. The *shift operator* σ acts as $\sigma((a(n))_{n\in\mathbb{N}}) = (a(n+1))_{n\in\mathbb{N}}$ on a sequence $(a(n))_{n\in\mathbb{N}}$. A difference subring $R \subseteq \mathbb{K}^{\mathbb{N}}$ is a subring which is additionally closed under taking shifts, i.e., $\sigma(a) \in R$ for all $a \in R$. The ring of recurrence operators $R[\sigma]$ is, in general, non-commutative and an element $\mathcal{A} := \sum_{i=0}^{r} c_i \sigma^i \in R[\sigma]$ with $c_i \in R$ acts on a sequence $a = (a(n))_{n\in\mathbb{N}}$ as $\mathcal{A}a = \sum_{i=0}^{r} c_i(n) a(n+i)$. If $\mathcal{A}a = 0$, we say that the operator $\mathcal{A}$ annihilates the sequence a. If $c_r \neq 0$, then r is called the *order* of the operator $\mathcal{A}$. The minimal order of an operator which annihilates a is called the order of the sequence a and is denoted by $\operatorname{ord}(a)$.

2.1 C-finite sequences

Sequences $c \in \mathbb{K}^{\mathbb{N}}$ which are annihilated by an operator $\mathcal{C} = \sum_{i=0}^{r} \gamma_i \sigma^i \in \mathbb{K}[\sigma]$ are called *C-finite*. Equivalently, these are sequences that satisfy a linear recurrence with constant coefficients

$$\gamma_0 c(n) + \cdots + \gamma_r c(n+r) = 0 \quad \text{for all } n \in \mathbb{N}.$$

The set of C-finite sequences over $\mathbb{K}$ forms a $\mathbb{K}$-algebra which we denote by $\mathcal{R}_C$. Suppose $c, d, c_1, \dots, c_m \in \mathcal{R}_C$. Then, the following *closure properties* are well known (e.g., [13]):

1. $c + d \in \mathcal{R}_C$ with $\operatorname{ord}(c+d) \leq \operatorname{ord}(c) + \operatorname{ord}(d)$,
2. $cd \in \mathcal{R}_C$ with $\operatorname{ord}(cd) \leq \operatorname{ord}(c)\operatorname{ord}(d)$,
3. $c_{\ell,k} := (c(\ell n + k))_{n\in\mathbb{N}} \in \mathcal{R}_C$ with $\operatorname{ord}(c_{\ell,k}) \leq \operatorname{ord}(c)$ for all $\ell, k \in \mathbb{N}$.
4. Let e be the interlacing of $c_1, \dots, c_m$, i.e., $e(n) = c_r(q)$ for all $n = qm + r$ with $0 \leq r < m$. Then, $e \in \mathcal{R}_C$ and $\operatorname{ord}(e) \leq m \sum_{j=1}^{m} \operatorname{ord}(c_j)$.

The same closure properties and order bounds hold for D-finite sequences, i.e., sequences which are annihilated by an operator $\mathcal{A} \in \mathbb{K}[x][\sigma]$ [17, 12].

Let $\mathcal{C} := \sum_{i=0}^{r-1} \gamma_i \sigma^i + \sigma^r$ be the unique monic minimal annihilating operator of $c \in \mathcal{R}_C$. The polynomial $\sum_{i=0}^{r-1} \gamma_i x^i + x^r \in \mathbb{K}[x]$ is called the *characteristic polynomial* of c. Over the splitting field $\mathbb{L}$ the polynomial completely factors as $x^{n_0} \prod_{i=1}^{m} (x - \lambda_i)^{d_i}$ with pairwise different $\lambda_1, \dots, \lambda_m \in \mathbb{L}$ and $n_0, d_1, \dots, d_m \in \mathbb{N}$. We call these λ_i the *eigenvalues* of the sequence c. The sequence can also be written as polynomial-linear combination of exponential sequences λ_i^n: In particular, there are polynomials $p_1, \dots, p_m \in \mathbb{L}[x]$ with $\deg(p_i) = d_i - 1$ for $i = 1, \dots, m$ such that

$$c(n) = \sum_{i=1}^{m} p_i(n) \lambda_i^n \text{ for all } n \geq n_0. \tag{1}$$

This is called the *closed form* of c [13, 19].

A C-finite sequence c is called *degenerate* if it has eigenvalues $\lambda \neq \mu$ such that $\frac{\lambda}{\mu}$ is a root of unity. Otherwise, the sequence is called *non-degenerate*.

Theorem 1 ([1, 5])**.** Let c be a non-degenerate C-finite sequence. Then, c is either the zero sequence or it only has finitely many zeros, i.e., there is an $n_0 \in \mathbb{N}$ such that $c(n) \neq 0$ for all $n \geq n_0$.

Suppose c, d are C-finite sequences with eigenvalues $\lambda_1, \dots, \lambda_{m_1}$ and $\mu_1, \dots, \mu_{m_2}$, respectively. From the closed form of c and d, it is clear, that $c + d$ has eigenvalues $\lambda_1, \dots, \lambda_{m_1}, \mu_1, \dots, \mu_{m_2}$ and cd has eigenvalues $\lambda_i \mu_j$ with $1 \leq i \leq m_1$ and $1 \leq j \leq m_2$. The sequence $(c(\ell n + k))_{n \in \mathbb{N}}$ has eigenvalues $\lambda_1^\ell, \dots, \lambda_m^\ell$.

2.2 C^2-finite sequences

A generalization of C-finite sequences are C^2-finite sequences. These extend C-finite and D-finite sequences and include many more sequences which appear in combinatorics.

Definition 2. A sequence $a \in \mathbb{K}^{\mathbb{N}}$ is called C^2-finite over $\mathbb{K}$ if there are C-finite sequences $c_0, \dots, c_r$ over $\mathbb{K}$ with $c_r(n) \neq 0$ for all $n \geq n_0$ for some $n_0 \in \mathbb{N}$ such that

$$c_0(n)a(n) + \dots + c_r(n)a(n + r) = 0, \qquad \text{for all } n \in \mathbb{N}. \tag{2}$$

Several examples for C^2-finite sequences are given in [29, 9]. Throughout this article, additional examples are given.

A C-finite sequence c can be uniquely described by a minimal recurrence and $\operatorname{ord}(c)$ many initial values. Similarly, a C^2-finite sequence can be described uniquely by its recurrence and by finitely many initial values. The number of initial values which is needed to uniquely determine the sequence depends on the zeros of the leading coefficient c_r of the recurrence. It can be decided whether the leading coefficient only has finitely many zeros [1]. However, it is not known if these finitely many zeros can be computed. This is known as the *Skolem problem* [26].

Previously, it was shown that C^2-finite sequences, analogously to C-finite sequences, form a $\mathbb{K}$-algebra and they are furthermore closed under taking subsequences at arithmetic progressions and interlacing [8, 9]. So far, it was not known whether these closure properties can be computed effectively. In this article we show a method how these closure properties can be performed effectively. As a caveat, these computations might introduce finitely many zeros in the leading coefficient which can yield to problems when one has to decide how many initial values are needed to uniquely define the sequence. In practice we have, however, observed that even though the Skolem problem is very difficult in general it can usually be solved for most examples that appear in practice [23].

Note that in [8, 9] it is assumed that the leading coefficient c_r in (2) has no zero terms at all. The two definitions are equivalent. If a sequence satisfies a C^2-finite recurrence as in Definition 2, then shifting the recurrence yields a recurrence where the leading coefficient has no zero terms. The definition here allows us to derive bounds for the orders of closure properties similar to the C-finite case that cannot be derived otherwise (see Example 17).

If $c_r(n) \neq 0$, then the recurrence (2) can be used to compute the term $a(n + r)$ provided that the previous terms $a(n), \dots, a(n + r - 1)$ are known:

$$a(n + r) = -\tfrac{c_0(n)}{c_r(n)} a(n) - \dots - \tfrac{c_{r-1}(n)}{c_r(n)} a(n + r - 1).$$

This is also captured by the companion matrix M_a of a which is defined as

$$M_a \coloneqq \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0/c_r \\ 1 & 0 & \dots & 0 & -c_1/c_r \\ 0 & 1 & \dots & 0 & -c_2/c_r \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{r-1}/c_r \end{pmatrix}.$$

If $c_r(n) \neq 0$ for all $n \in \mathbb{N}$, then

$$(\sigma a, \sigma^2 a, \dots, \sigma^r a) = (a, \sigma a, \dots, \sigma^{r-1} a)\, M_a.$$

In the special case that a is a C-finite sequence, we have $M_a \in \mathbb{K}^{r \times r}$. Since the recurrence of a sequence is not unique, neither is the companion matrix.

In the recurrence (2) we can assume that $n_0 = 0$ holds in the closed form (1) for all coefficients $c_0, \dots, c_r$. This can be achieved by extending the closed form representation to all $n \in \mathbb{N}$ and introducing polynomial factors $n(n-1)\cdots(n-n_0)$ in all coefficients c_i. This only increases the order of the coefficients c_i and leaves the order of the overall recurrence intact.

2.3 Lattices

A $\mathbb{Z}$-submodule L of $\mathbb{Z}^m$ is called a lattice. Every lattice L admits a finite basis $v_1, \dots, v_\ell \in \mathbb{Z}^m$, i.e., a set of linearly independent generators of the module L. We call ℓ the *rank* of the lattice L.

The LLL algorithm can be used to compute a basis of "short" vectors for the lattice L [16, 3]. Such a basis is called a *reduced* basis. Let $L = \langle v_1, \dots, v_\ell \rangle \subseteq \mathbb{Z}^m$, i.e., L is the lattice generated by $v_1, \dots, v_\ell \in \mathbb{Z}^m$. Let $b_1, \dots, b_r$ be a reduced basis of L and $\bar{b}_1, \dots, \bar{b}_r$ the corresponding Gram-Schmidt vectors. The reduced basis is "short" in the sense that (cf. (1.7) in [16] or Theorem 2.6.2 in [3])

$$\|b_j\|_2^2 \leq 2^{k-1} \left\|\bar{b}_k\right\|_2^2 \text{ for all } 1 \leq j \leq k \leq r. \tag{3}$$

Suppose $V \in \mathbb{Z}^{m \times \ell}$ and $r = \min(m, \ell)$. Then, we can compute unimodular (i.e., invertible) matrices $P \in \mathbb{Z}^{m \times m}, Q \in \mathbb{Z}^{\ell \times \ell}$ and a diagonal matrix $D = \operatorname{diag}(d_1, \dots, d_r) \in \mathbb{Z}^{m \times \ell}$ with $d_i \mid d_{i+1}$ for all $i = 1, \dots, r-1$ such that $PVQ = D$. The unique matrix D is called the *Smith normal form* of V and the largest diagonal entry d_r is called the *invariant factor* of V. If e_i denotes the i-th determinantal divisor of V, i.e., the greatest common divisor of all i-by-i minors of V, then $d_r = \frac{e_r}{e_{r-1}}$ [21, 20].

3 The exponent lattice of algebraic numbers

Let $\lambda_1, \dots, \lambda_m \in \overline{\mathbb{Q}}$. The set of relations of these algebraic numbers

$$L \coloneqq L(\lambda_1, \dots, \lambda_m) \coloneqq \{(e_1, \dots, e_m) \in \mathbb{Z}^m \mid \lambda_1^{e_1} \cdots \lambda_m^{e_m} = 1\}$$

forms a lattice. In his PhD thesis [7] Ge gave an algorithm for computing a basis of L. It is a combination of LLL with a bound on the size of the basis vectors and the fact that membership of L is easy to decide. Variants of Ge's algorithm were given in [10, 6, 35, 33, 34]. Here we present another variant. Our version is inspired by how LLL is applied in van Hoeij's algorithm for polynomial factorization [30, 32, 31]. One feature of this version is that it uses approximations that are only as good as necessary for the particular input, rather than approximations whose accuracy is determined by the worst case behavior. Another advantage of our version is that its correctness admits a very concise proof.

Like Ge, we start by observing that

$$(e_1, \dots, e_m) \in L(\lambda_1, \dots, \lambda_m) \iff \lambda_1^{e_1} \cdots \lambda_m^{e_m} = 1 \iff \sum_{i=1}^m e_i \log(\lambda_i) \in 2\pi \mathrm{i} \mathbb{Z}.$$

Hence, instead of finding a basis for $L = L(\lambda_1, \dots, \lambda_m)$, we can compute a basis of the lattice

$$L_+ = \left\{ (e_1, \dots, e_{m+1}) \in \mathbb{Z}^m \mid \sum_{i=1}^m e_i \log(\lambda_i) + e_{m+1} 2\pi \mathrm{i} = 0 \right\}$$

and drop the last coordinates to find a basis of the original lattice L. If we agree to always choose the standard branch of the logarithm, the last coordinate will be bounded by md, where d is the degree of the field extension of $\lambda_1, \dots, \lambda_m$. By a result by Masser [18], we can compute a constant $M \geq md$ such that L and therefore L_+ have a basis of vectors b with $\|b\|_\infty \leq M$.

It remains to provide an algorithm which can compute a basis of

$$L_+ = \{(e_1, \dots, e_n) \in \ \mathbb{Z}^n \mid e_1 x_1 + \cdots + e_n x_n = 0\}$$

where $x_1, \ldots, x_n \in \mathbb{C} \setminus \{0\}$. Due to the special shape of the x_i in our case, we can compute rational approximations $\xi_i \in \mathbb{Q}(\mathrm{i})$ of arbitrary precision [2]. In particular, for every $\epsilon > 0$ we can compute $\xi_1, \ldots, \xi_n \in \mathbb{Q}(\mathrm{i})$ such that $|\Re(x_i) - \Re(\xi_i)| < \epsilon$ and $|\Im(x_i) - \Im(\xi_i)| < \epsilon$ for all $i = 1, \ldots, n$. Furthermore, we can use the fact that membership $(e_1, \ldots, e_n) \in L_+$ can be checked and that we know that a basis with vectors bounded by M exists.

Algorithm 3: Computing a basis for L_+

```
Input:  Computable numbers $x_1, \ldots, x_n \in \mathbb{C} \setminus \{0\}$ and $M \in \mathbb{Q}$ such that the lattice
            $$L_+ = \left\{ (e_i)_{i=1,\ldots,n} \in \mathbb{Z}^n \mid \sum_{i=1}^{n} e_i x_i = 0 \right\}$$
        has a basis of vectors $b \in \mathbb{Z}^n$ with $\|b\|_\infty \leq M$
Needs:  One can decide whether $b \in L_+$ for any $b \in \mathbb{Z}^n$
Output: A basis of $L_+$
$w \leftarrow 1$
$B \leftarrow \{(1, 0, \ldots, 0, 0, 0), \ldots, (0, \ldots, 0, 1, 0, 0)\} \subseteq \mathbb{Z}^{n+2}$
while $\exists\ (e_1, \ldots, e_n, *, *) \in B$: $(e_1, \ldots, e_n) \notin L_+$ do
    $w \leftarrow 2w$
    find $\xi_1, \ldots, \xi_n \in \mathbb{Q}(\mathrm{i})$ with $|\Re(\xi_i) - \Re(x_i)| < \frac{1}{nw}$ and $|\Im(\xi_i) - \Im(x_i)| < \frac{1}{nw}$ for
      all $i = 1, \ldots, n$
    replace every vector $(e_1, \ldots, e_n, *, *) \in B$ by
            $$\left( e_1, \ldots, e_n, w \sum_{i=1}^{n} e_i \Re(\xi_i), w \sum_{i=1}^{n} e_i \Im(\xi_i) \right)$$
    apply LLL to $B$, call the output vectors $b_1, \ldots, b_r$ and the corresponding
      Gram-Schmidt vectors $\bar{b}_1, \ldots, \bar{b}_r$
    while $r > 0$ and $\|\bar{b}_r\|_2 > \sqrt{n+2}M$ do $r \leftarrow r - 1$
    $B \leftarrow \{b_1, \ldots, b_r\}$
end
return $\{(e_1, \ldots, e_n) : (e_1, \ldots, e_n, *, *) \in B\}$
```

For proving the correctness of Algorithm 3 we will employ the following lemma:

Lemma 4. [32, Lemma 2] If $b_1, \ldots, b_r$ is a lattice basis and $\bar{b}_1, \ldots, \bar{b}_r$ is the corresponding Gram-Schmidt basis, then for every $v \in \langle b_1, \ldots, b_r \rangle$ with $\|v\|_2 < \|\bar{b}_r\|_2$ we have in fact $v \in \langle b_1, \ldots, b_{r-1} \rangle$.

Proof. Let $v \in \langle b_1, \ldots, b_r \rangle$ be such that $\|v\|_2 < \|\bar{b}_r\|_2$, say $v = \alpha_1 b_1 + \cdots + \alpha_r b_r$ for certain $\alpha_1, \ldots, \alpha_r \in \mathbb{Z}$. We have to show that $\alpha_r = 0$. Let $\mu_{i,j}$ be such that $b_i = \sum_{j \leq i} \mu_{i,j} \bar{b}_j$ for all i, j;

note that $\mu_{i,i} = 1$. Now

$$\begin{aligned}
\left\|\bar{b}_r\right\|_2^2 > \|v\|_2^2 &= \left\|\sum_{i=1}^{r} \alpha_i b_i\right\|_2^2 = \left\|\sum_{i=1}^{r}\sum_{j=1}^{i} \alpha_i \mu_{i,j} \bar{b}_j\right\|_2^2 \\
&= \left\|\sum_{j=1}^{r}\left(\sum_{i=j}^{r} \alpha_i \mu_{i,j}\right)\bar{b}_j\right\|_2^2 \overset{\text{Pythagoras}}{=} \sum_{j=1}^{r}\left|\sum_{i=j}^{r} \alpha_i \mu_{i,j}\right|^2 \left\|\bar{b}_j\right\|_2^2 \\
&= \underbrace{\sum_{j=1}^{r-1}\left|\sum_{i=j}^{r} \alpha_i \mu_{i,j}\right|^2 \left\|\bar{b}_j\right\|_2^2}_{\geq 0} + |\alpha_r|^2 \left\|\bar{b}_r\right\|_2^2 \geq |\alpha_r|^2 \left\|\bar{b}_r\right\|_2^2
\end{aligned}$$

together with $\alpha_r \in \mathbb{Z}$ forces $\alpha_r = 0$, as claimed. □

Theorem 5. Algorithm 3 is correct and terminates.

Proof. It is clear that every output vector is an element of L_+. To see that the output vectors generate L_+, we need to justify the removals in line 9. By assumption, we know that L_+ has a basis whose elements have components bounded by M. For every vector $(e_1, \dots, e_n) \in L_+$ with $|e_i| < M$ for all i we have

$$\begin{aligned}
w\left|\sum_{i=1}^{n} e_i \Re(\xi_i)\right| &= w\left|\sum_{i=1}^{n} e_i \Re(\xi_i) - e_i \Re(x_i)\right| \leq w \sum_{i=1}^{n} |e_i| \left|\Re(\xi_i) - \Re(x_i)\right| \\
&\leq w \sum_{i=1}^{n} |e_i| /(nw) < M
\end{aligned}$$

and likewise for the imaginary parts. Therefore, we are only interested in vectors $b = (e_1, \dots, e_n, *, *)$ in the lattice generated by B with

$$\|b\|_2 \leq \sqrt{M^2 + \cdots + M^2 + M^2 + M^2} = \sqrt{n+2}M.$$

By Lemma 4, these vectors are still in the lattice after the removals in line 9.

It remains to show that the algorithm terminates. Suppose it does not terminate, i.e., the set B eventually contains r vectors in every iteration which are not all in the lattice L_+. We show that from some point on in the algorithm (i.e., for big enough w), this cannot be the case because vectors which are not in L_+ are too long and are therefore removed in line 9 of the algorithm.

There are only finitely many vectors $(e_1, \dots, e_n) \in \mathbb{Z}^n$ with $|e_i| \leq \sqrt{n+2}M$ for all $i = 1, \dots, n$. Therefore, there exists an $\epsilon > 0$ such that

$$\max(|e_1 \Re(x_1) + \cdots + e_n \Re(x_n)|, |e_1 \Im(x_1) + \cdots + e_n \Im(x_n)|) > \epsilon$$

for all $(e_1, \dots, e_n) \in \mathbb{Z}^n \setminus L_+$ with $|e_i| \leq \sqrt{n+2}M$ for all i. Choose such an ϵ. Suppose we are in line 3 of the algorithm with $w \geq \frac{\sqrt{n+2}M(1+2^{(r-1)/2})}{\epsilon}$ and $b_j \in B \setminus L$ with $j \in \{1, \dots, r\}$. Let

$$b_j = \left(e_1, \dots, e_n, w\sum_{i=1}^{n} e_i \Re(\xi_i), w\sum_{i=1}^{n} e_i \Im(\xi_i)\right).$$

Since b_j has not been removed in line 9 in the previous iteration, we have $|e_i| \leq \sqrt{n+2}M$ for all $i = 1, \dots, n$. By the choice of ϵ for either $f = \Re$ or $f = \Im$ we have

$$\begin{aligned}
w\left|\sum_{i=1}^{n} e_i f(\xi_i)\right| &= w\left|\sum_{i=1}^{n} e_i (f(\xi_i) - f(x_i) + f(x_i))\right| = w\left|\sum_{i=1}^{n} e_i (f(\xi_i) - f(x_i)) + \sum_{i=1}^{n} e_i f(x_i)\right| \\
&\geq w\left(\left|\sum_{i=1}^{n} e_i f(x_i)\right| - \left|\sum_{i=1}^{n} e_i (f(\xi_i) - f(x_i))\right|\right) > w\left(\epsilon - \left|\sum_{i=1}^{n} e_i (f(\xi_i) - f(x_i))\right|\right).
\end{aligned}$$

Furthermore, we have

$$\left|\sum_{i=1}^{n} e_i(f(\xi_i) - f(x_i))\right| \leq \sum_{i=1}^{n} |e_i| \, |f(\xi_i) - f(x_i)| < \sum_{i=1}^{n} \sqrt{n+2}M \tfrac{1}{nw} = \tfrac{\sqrt{n+2}M}{w}.$$

Using this and the condition on w in the inequality above yields

$$w\left|\sum_{i=1}^{n} e_i f(\xi_i)\right| > we - \sqrt{n+2}M \geq 2^{(r-1)/2}\sqrt{n+2}M.$$

Therefore, $\|b_j\|_2 > 2^{(r-1)/2}\sqrt{n+2}M$. In particular, using (3),

$$\left\|\bar{b}_k\right\|_2 \geq 2^{-(k-1)/2} \|b_j\|_2 > 2^{(r-k)/2}\sqrt{n+2}M \geq \sqrt{n+2}M$$

for all $k = j, \ldots, r$. Hence, $b_j, \ldots, b_r$ would have been removed already in the past iteration and cannot be in the set B anymore, a contradiction. □

An implementation of the algorithm is part of the `rec_sequences` package[1] and is publicly available [22].

4 Torsion number

For proving the order bounds for C^2-finite sequences, we will heavily rely on the fact that a C-finite sequence c can be written as interlacing of non-degenerate sequences $c(dn), \ldots, c(dn+d-1)$ [5, Theorem 1.2]. More generally, if we have a finitely generated difference algebra of C-finite sequences, we will determine a number $d \in \mathbb{N}$ (which we will call the *torsion number*) such that every sequence in the algebra can be written as the interlacing of d non-degenerate subsequences.

Let $c_0, \ldots, c_r \in \mathcal{R}_C$ with eigenvalues $\lambda_1, \ldots, \lambda_m$ and let $R_d := \mathbb{K}_\sigma[c_0(dn), \ldots, c_r(dn)]$ be the smallest difference algebra which contains the sequences $c_0(dn), \ldots, c_r(dn)$. Suppose $c \in R_d$. Then, every eigenvalue λ of c is of the form $\lambda = \lambda_1^{e_1} \cdots \lambda_m^{e_m}$ for some $e_1, \ldots, e_m \in \mathbb{N}$. We want to find a d such that every sequence $c \in R_d$ is non-degenerate. Equivalently, we want to find a d such that

$$\left(\frac{\lambda_1^{de_1} \cdots \lambda_m^{de_m}}{\lambda_1^{df_1} \cdots \lambda_m^{df_m}}\right)^k = 1 \implies \lambda_1^{de_1} \cdots \lambda_m^{de_m} = \lambda_1^{df_1} \cdots \lambda_m^{df_m}$$

for all $k, e_1, \ldots, e_m, f_1, \ldots, f_m \in \mathbb{N}$. In order to write this more concisely we define the multiplicative group $G := \langle \lambda_1, \ldots, \lambda_m \rangle \leq (\mathbb{C}^\times, \cdot)$. Then, this condition reads as

$$\forall k \in \mathbb{N}_{\geq 1} \forall \lambda \in G\colon \lambda^{kd} = 1 \implies \lambda^d = 1.$$

The following lemma shows that this number d also has a purely group-theoretical and a purely lattice-theoretical description.

Lemma 6. Let $G := \langle \lambda_1, \ldots, \lambda_m \rangle \leq (\mathbb{C}^\times, \cdot)$. The following conditions on $d \in \mathbb{N}_{\geq 1}$ are equivalent:

1. The number d satisfies

$$\forall k \in \mathbb{N}_{\geq 1} \forall \lambda \in G\colon \lambda^{kd} = 1 \implies \lambda^d = 1.$$

2. Let $T(G) := \{\lambda \in G \mid \operatorname{ord}(\lambda) < \infty\}$ be the torsion subgroup of G. Then, d satisfies

$$\operatorname{ord}(\lambda) \mid d \text{ for all } \lambda \in T(G).$$

[1] The code is available at `https://github.com/PhilippNuspl/rec_sequences` in the `IntegerRelations` class.

3. Let

$$L \coloneqq L(\lambda_1, \dots, \lambda_m) \coloneqq \{(e_1, \dots, e_m) \in \mathbb{Z}^m \mid \lambda_1^{e_1} \cdots \lambda_m^{e_m} = 1\}$$

be the lattice of integer relations among $\lambda_1, \dots, \lambda_m$. Then, d satisfies

$$\forall k \in \mathbb{N}_{\geq 1}\, \forall v \in \mathbb{Z}^m \colon kdv \in L \implies dv \in L. \tag{4}$$

Proof. $1 \implies 2$: Let $\lambda \in T(G)$ and let $m \in \mathbb{N}_{\geq 1}$ be minimal with $\lambda^m = 1$. Then, clearly $\lambda^{md} = 1$. By assumption, $\lambda^d = 1$. As m was chosen minimal, we have $m \mid d$.

$2 \implies 3$: Let $k \in \mathbb{N}_{\geq 1}, v = (e_1, \dots, e_m) \in \mathbb{Z}^m$ and $kdv \in L$. Let $\lambda = \lambda_1^{e_1} \cdots \lambda_m^{e_m}$. By definition of L,

$$\lambda^{kd} = \lambda_1^{kde_1} \cdots \lambda_m^{kde_m} = 1.$$

Hence, $\lambda \in T(G)$. Therefore, by assumption, $\operatorname{ord}(\lambda) \mid d$, so $\lambda^d = 1$ and $dv \in L$.

$3 \implies 1$: Let $k \in \mathbb{N}_{\geq 1}, \lambda = \lambda_1^{e_1} \cdots \lambda_m^{e_m} \in G$ and $\lambda^{kd} = 1$. Defining $v \coloneqq (e_1, \dots, e_m) \in \mathbb{Z}^m$ yields $kdv \in L$. By assumption, $dv \in L$, i.e., $\lambda^d = 1$. □

Considering condition 2 of Lemma 6, we can see that the smallest d which satisfies the condition is the exponent of the torsion group.

Definition 7. The *torsion number* $d \in \mathbb{N}_{\geq 1}$ of $\lambda_1, \dots, \lambda_m \in \overline{\mathbb{Q}}$ is defined as

$$d \coloneqq \exp(T(G)) \coloneqq \operatorname{lcm}(\operatorname{ord}(\lambda) \mid \lambda \in T(G))$$

where $G \coloneqq \langle \lambda_1, \dots, \lambda_m \rangle \leq (\mathbb{C}^\times, \cdot)$.

We also call d the torsion number of the lattice L if it is the smallest number satisfying (4). Using the terminology of pure modules, (4) is equivalent to $\{v \in \mathbb{Z}^m \mid dv \in L\}$ being a pure lattice [4, III.16].

Lemma 8. *Let d be the torsion number of the lattice L. Then, d' satisfies the conditions from Lemma 6 if and only if $d \mid d'$.*

Proof. $\implies$: By definition, d is the smallest number satisfying the conditions from Lemma 6. Hence, $d' > d$ and we can write $d' = dq + r$ with $0 \leq r < d$. Let $G \coloneqq \langle \lambda_1, \dots, \lambda_m \rangle$ and $\lambda \in T(G)$. Then, $\operatorname{ord}(\lambda) \mid d$ and $\operatorname{ord}(\lambda) \mid d'$. Hence,

$$\operatorname{ord}(\lambda) \mid (d' - dq) = r.$$

Since $d > r$ is the smallest number with this property we have $r = 0$, so $d \mid d'$.

$\impliedby$: Clear from condition 2 of Lemma 6. □

Now, we want to show that the torsion number of algebraic numbers $\lambda_1, \dots, \lambda_m \in \overline{\mathbb{Q}}$ can actually be computed. First, we have devised an algorithm in Section 3 which computes a basis $v_1, \dots, v_\ell \in \mathbb{Z}^m$ for the lattice $L \coloneqq L(\lambda_1, \dots, \lambda_m)$. Then, the invariant factor of the matrix built by the basis is precisely the torsion number of the lattice:

Theorem 9. *Let $v_1, \dots, v_\ell \in \mathbb{Z}^m$ be a basis of the lattice $L = \langle v_1, \dots, v_\ell \rangle \subseteq \mathbb{Z}^m$. Let $V \coloneqq (v_1, \dots, v_\ell) \in \mathbb{Z}^{m \times \ell}$. Then, the invariant factor of V is the torsion number of L.*

Proof. We write

$$PVQ = \begin{pmatrix} D \\ 0 \end{pmatrix} \eqqcolon \overline{D} \in \mathbb{Z}^{m \times \ell}$$

where $D = \operatorname{diag}(d_1, \dots, d_{\ell-1}, d)$ is the Smith normal form and $P \in \mathbb{Z}^{m \times m}, Q \in \mathbb{Z}^{\ell \times \ell}$ are unimodular matrices.

First, we show that the invariant factor d of the matrix V satisfies (4): Let $k \in \mathbb{N}_{\geq 1}$ and $v \in \mathbb{Z}^m$ with $kdv \in L$. Then, there is a $w = (w_1, \dots, w_\ell) \in \mathbb{Z}^\ell$ such that

$$kdv = w_1 v_1 + \cdots + w_\ell v_\ell = Vw.$$

Therefore,

$$kdPv = PVw = PVQQ^{-1}w = \overline{D}\overline{w}$$

with $\overline{w} = Q^{-1}w$. The ℓ-th row yields $kd(Pv)_\ell = d\overline{w}_\ell$ where $(Pv)_\ell$ denotes the ℓ-th entry of the vector Pv. Hence, $k \mid \overline{w}_\ell$. For the i-th row with $i < \ell$ we have $kd(Pv)_i = d_i\overline{w}_i$. By the property of the Smith normal form, we have $d_i \mid d$ and therefore $k \mid \overline{w}_i$. As Q is unimodular, [20, Corollary 158] yields

$$k \mid \gcd(\overline{w}) = \gcd(Q^{-1}w) = \gcd(w).$$

Hence, $\frac{w}{k} \in \mathbb{Z}^\ell$ and

$$dv = \tfrac{1}{k}kdv = V\tfrac{w}{k} \in L.$$

Secondly, we show that the invariant factor d is the smallest number: Suppose d' is the smallest number which satisfies (4). By Lemma 8 there is a k such that $d = d'k$. Let $P^{-1} = (p_1, \dots, p_m) \in \mathbb{Z}^{m\times m}$. As the columns of V are a basis of L and Q is unimodular we have

$$L = V\mathbb{Z}^\ell = P^{-1}\overline{D}Q^{-1}\mathbb{Z}^\ell = P^{-1}\overline{D}\mathbb{Z}^\ell.$$

Therefore,

$$\{d_1p_1, \dots, d_{\ell-1}p_{\ell-1}, dp_\ell\}$$

is also a basis of L. Let $v := dp_\ell = d'kp_\ell \in L$. By assumption, $d'p_\ell \in L$. Hence,

$$\{d_1p_1, \dots, d_{\ell-1}p_{\ell-1}, d'p_\ell\}$$

is a basis of L as well. Therefore, there is a unimodular change-of-basis matrix $U \in \mathbb{Z}^{m\times m}$ with

$$U(d_1p_1, \dots, d_{\ell-1}p_{\ell-1}, dp_\ell) = (d_1p_1, \dots, d_{\ell-1}p_{\ell-1}, d'p_\ell).$$

In particular, the last column yields $Udp_\ell = d'p_\ell$. As U is unimodular, we have

$$d\gcd(p_\ell) = \gcd(Udp_\ell) = \gcd(d'p_\ell) = d'\gcd(p_\ell).$$

As $\gcd(p_\ell) \neq 0$, we have $d = d'$. □

Example 10. Let

$$\lambda_1 = 2^{1/2}, \lambda_2 = (-2)^{1/3}, \lambda_3 = \mathrm{i}, \lambda_4 = -\mathrm{i}.$$

The columns of

$$V := \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 3 \\ 1 & 2 & -1 \\ 1 & -2 & 1 \end{pmatrix} = P^{-1}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}Q^{-1}$$

are a basis of $L(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. Hence, $d = 4$ is the torsion number of $\lambda_1, \dots, \lambda_4$.

Let $c_0, \dots, c_r \in \mathcal{R}_C$ with eigenvalues $\lambda_1, \dots, \lambda_m$. Then, we have seen that we can compute a number $d \in \mathbb{N}_{\geq 1}$ (namely the torsion number) such that the algebra

$$R := \mathbb{K}_\sigma[c_0(dn), \dots, c_r(dn)]$$

only contains sequences which are non-degenerate, i.e., sequences which contain only finitely many zeros. A non-degenerate sequence might still be a zero divisor in the ring $\mathbb{K}^\mathbb{N}$. However, we can still define the localization $Q(R) := \{\frac{c}{d} \mid c \in R, d \in R \setminus \{0\}\}$. This localization $Q(R)$ is a field. An element of $Q(R)$ can, however, only be interpreted as a sequence in $\mathbb{K}^\mathbb{N}$ from some term on (cf. the discussion in Section 8.2 in [27] or [28]). For instance, the sequence $\frac{3^n}{2^n-1}$ cannot be evaluated at the term $n = 0$. This is not a problem for our applications as we will see in Section 5. We summarize the discussions of the section in the following theorem:

Theorem 11. Let $c_0, \dots, c_r \in \mathcal{R}_C$ with eigenvalues $\lambda_1, \dots, \lambda_m$. Then, we can compute a number $d \in \mathbb{N}_{\geq 1}$ (namely the torsion number) such that the localization $Q(R)$ of the algebra

$$R \coloneqq \mathbb{K}_\sigma[c_0(dn), \dots, c_r(dn)]$$

is a field. The elements of the field $Q(R)$ can be considered as sequences which are non-zero from some term on.

From the closed form of C-finite sequences it is clear that these sequences can be seen as special cases of sums of single nested product expressions. The torsion number can be used to find a certain algebraic independent basis of these sequences [28].

5 Order bounds

In this section we will derive order bounds for the ring operations and additional closure properties of C^2-finite sequences.

The computation of closure properties of C^2-finite sequences can be reduced to solving linear systems of equations [8, 9]. A C^2-finite recurrence

$$x_0(n) + x_1(n)\sigma + \cdots + x_s(n)\sigma^s$$

with $x_i \in R$ for some suitable ring R of sequences is obtained by computing an element $(x_0, \dots, x_s)$ in the kernel of a matrix

$$\big(w_0, w_1, \dots, w_s\big) \in Q(R)^{r \times (s+1)}. \tag{5}$$

The w_i can be computed iteratively using $w_{i+1} = M\sigma(w_i)$ for a suitable matrix $M \in Q(R)^{r\times r}$ (where the shift operator σ is applied componentwise).

- In the case a recurrence for $a+b$ is computed, we use $w_0 = e_0 \oplus \tilde{e}_0$ and $M = M_a \oplus M_b$ where M_a, M_b are the companion matrices of a, b and $e_0, \tilde{e}_0$ are the first unit vectors of appropriate sizes.
- In the case a recurrence for ab is computed, we use $w_0 = e_0 \otimes \tilde{e}_0$ and $M = M_a \otimes M_b$.
- In the case a recurrence for $a(\ell n)$ is computed, we use $w_0 = e_0$ and $M = M_a(\ell n) \cdots M_a(\ell n + \ell - 1)$.
- In the case a C^2-finite recurrence for $c(jn^2 + kn + \ell)$ with $j, k, \ell \in \mathbb{N}$ and a C-finite sequence c (which does not have 0 as an eigenvalue) is computed, we use

  $$w_0 = M_c^{kn+\ell-r+1} e_{r-1} \text{ and } M = M_c^{j(2n+1)} \tag{6}$$

  where M_c is the companion matrix of c and e_{r-1} the last unit vector.

The underlying ring R is the difference algebra $\mathbb{K}_\sigma[c_0, \dots, c_r]$ generated by the C-finite sequences $c_0, \dots, c_r$ appearing in w_0 and M.

5.1 Interlacing and subsequence

Theorem 12. Let $a_1(n), \dots, a_d(n)$ be C^2-finite sequences of maximal order r. Let b be the interlacing of these sequences. We can compute a C^2-finite recurrence of order at most dr for b.

Proof. By shifting the recurrences of the a_s appropriately, we can assume that they all satisfy a C^2-finite recurrence of order r of the form

$$c_{s,0}(n)a_s(n) + \cdots + c_{s,r}(n)a_s(n+r) = 0$$

for $s = 1, \dots, d$ for C-finite sequences $c_{s,i}$ where the $c_{s,r}$ only have finitely many zeros. Let e_{di} be the interlacing of $c_{1,i}, \dots, c_{d,i}$ for $i = 0, \dots, r$. These e_{di} are then C-finite and e_{dr} only has finitely many zeros. Then, b satisfies the recurrence

$$e_0(n)b(n) + e_d(n)b(n+d) + \cdots + e_{dr}(n)b(n+dr) = 0.$$

□

As seen in the proof of Theorem 12, computing the interlacing of C^2-finite sequences is simpler than in the case of C-finite and D-finite sequences. This is because the coefficients of the recurrence, namely C-finite sequences, are closed under interlacing themselves.

Example 13. Let c be C-finite satisfying

$$c(n) - c(n+r) = 0, \quad c(0) = 1, c(1) = \cdots = c(r-1) = 0.$$

Furthermore, let a be the interlacing of c and $d-1$ times the zero sequence. Theorem 12 shows that a is C^2-finite of order at most dr. The sequence a is cyclic and has $rd-1$ consecutive zeros. Hence, the sequence a also has to have order at least rd as otherwise, the sequence would be constantly zero. The bound in Theorem 12 is therefore tight in general.

Lemma 14. Let a be C^2-finite of order r and let d be the torsion number of the eigenvalues appearing in the recurrence of a. Let $\ell \in \mathbb{N}$. We can compute a C^2-finite recurrence of order at most r which is satisfied by all sequences $a(d\ell n + i)$ for $i = 0, \dots, d\ell - 1$.

Proof. The sequences $a(n+i)$ for $i = 0, \dots, d-1$ all satisfy the same recurrence. By the choice of d, all sequences in the ring R generated by the sequences appearing in

$$M = M_a(d\ell n) \cdots M_a(d\ell n + d\ell - 1)$$

are non-degenerate. By Theorem 11, $Q(R)$ is a field. Therefore, if $s = r$, then the linear system (5) is underdetermined and we can compute an element (after clearing denominators) $(x_0, \dots, x_r) \in R^{r+1}$ in the kernel with $x_t \neq 0$ and $x_{t+1} = \cdots = x_r = 0$ for some $t \leq r$. This gives rise to a C^2-finite recurrence

$$x_0(n) + x_1(n)\sigma + \cdots + x_t(n)\sigma^t$$

as x_t only has finitely many zeros by the choice of d. □

To extend Lemma 14 to subsequences at arbitrary arithmetic progressions we write such an arbitrary subsequence as the interlacing of certain subsequences for which Lemma 14 can be applied.

Theorem 15. Let a be C^2-finite of order r and let d be the torsion number of the eigenvalues appearing in the recurrence of a. Let $\ell \in \mathbb{N}$. We can compute a C^2-finite recurrence of order at most dr which is satisfied by the sequence $a(\ell n)$.

Proof. By Lemma 14 we can compute a recurrence of order at most r satisfied by $a(d\ell n + i)$ for $i = 0, \dots, d\ell - 1$. Let b be the interlacing of the d sequences

$$a(d\ell n), a(d\ell n + \ell), \dots, a(d\ell n + (d-1)\ell).$$

By Theorem 12, b has order at most dr. We show that $b(n) = a(\ell n)$: Let $n = qd + s$ with $0 \leq s < d$. Then, by the definition of b

$$b(n) = b(qd+s) = a(d\ell q + sl) = a(\ell(dq+s)) = a(\ell n).$$

□

5.2 Ring operations

Theorem 16. Let a, b be C^2-finite of order r_1, r_2, respectively and let d be the torsion number of the eigenvalues appearing in the recurrences of a, b. Then,

1. the sequence $a + b$ is C^2-finite of order at most $d(r_1 + r_2)$ and
2. the sequence ab is C^2-finite of order at most $dr_1 r_2$.

Furthermore, such recurrences can be computed.

Proof. We can compute C^2-finite recurrences of maximal order r_1, r_2 for $a(dn + i), b(dn + i)$ by Lemma 14. The closure properties $a(dn + i) + b(dn + i)$ and $a(dn + i)b(dn + i)$ can be computed again by solving a linear system of equations over the field $Q(R)$. Then, the same order bounds as in the C-finite and D-finite case apply, so the sequences $a(dn+i)+b(dn+i), a(dn+i)b(dn+i)$ have maximal orders $r_1 + r_2, r_1r_2$, respectively. By Theorem 12, we can interlace these sequence and obtain a recurrence of order $d(r_1 + r_2), dr_1r_2$ for $a + b$ and ab, respectively. □

In the special case that both C^2-finite sequences are C-finite or D-finite, the torsion number is 1 and the bounds simplify to the known order bounds for these rings.

Theorem 16 does not imply that the ring of C^2-finite sequences is computable. We can compute C^2-finite recurrences for the sum and the product. These recurrences, however, have leading coefficients which can have finitely many zeros. To uniquely determine the sequences $a + b, ab$ we might need to define additional initial values at these singularities. However, by the Skolem problem, we do not know whether these singularities can be computed. This is also illustrated in the next example.

Example 17. Let $a(n) = 2^{\binom{n+1}{2}}$ (A006125 in the OEIS [25]) and $b(n) = 4^{\binom{n}{2}}$ (A053763). Both sequences are C^2-finite satisfying the recurrences

$$2^{n+1}\, a(n) - a(n + 1) = 0, \quad 4^n\, b(n) - b(n + 1) = 0.$$

The torsion number of $L(1, 2, 4)$ is $d = 1$. The coefficients for a recurrence of $c = a + b$ are given by an element in the kernel of

$$\begin{pmatrix} 1 & 2^{n+1} & 2^{2n+3} \\ 1 & 2^{2n} & 2^{4n+2} \end{pmatrix}.$$

A recurrence is therefore, for instance, given by

$$2^{3n+3}(2^n - 1)c(n) - 2^{n+2}(2^{2n} - 2)c(n + 1) + (2^n - 2)c(n + 2) = 0.$$

The recurrence has order $\operatorname{ord}(a) + \operatorname{ord}(b) = 2$ as expected but the leading coefficient has a zero term at $n = 1$. Shifting the recurrence yields a recurrence of higher order with a leading coefficient which does not have any zero terms anymore.

Example 18. Let c be C-finite of order 2 satisfying

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

Let a, b be C^2-finite satisfying

$$a(n) = 1 \quad c(n)b(n) - b(n + 1) = 0, \quad b(0) = 1.$$

The eigenvalues that appear are 1 and -1. The torsion number is therefore $d = 2$. Let $a_i(n) = a(2n + i)$ and $b_i(n) = b(2n + i)$ for $i = 0, 1$. These are even C-finite of order 1 satisfying

$$a_i(n) - a_i(n + 1) = 0, \quad b_i(n) + b_i(n + 1) = 0.$$

Let $s_i = a_i + b_i$. These s_i are C-finite of order 2 satisfying

$$s_i(n) - s_i(n + 2) = 0.$$

The interlacing $s = a + b$ of s_0, s_1 satisfies the C-finite recurrence of order $4 = d(\mathrm{ord}(a) + \mathrm{ord}(b))$

$$s(n) - s(n+4) = 0.$$

However, s also satisfies a C^2-finite recurrence of order 3, namely

$$c_0(n)s(n) + c_2(n)s(n+2) + s(n+3) = 0$$

with

$$\begin{aligned} &c_0(n) - c_0(n+2) = 0, && c_0(0) = -1, && c_0(1) = 0, \\ &c_2(n) - c_2(n+2) = 0, && c_2(0) = 0, && c_2(1) = -1. \end{aligned}$$

There cannot be a shorter recurrence for $s(n)$ as it contains 2 consecutive zeros.

5.3 Sparse subsequences

Theorem 19. *Let c be C-finite of order r and $\lambda_1, \dots, \lambda_m$ its eigenvalues and $\lambda_i \neq 0$ for all $i = 1, \dots, m$. Let d be the torsion number of the eigenvalues. Then, we can compute a C^2-finite recurrence of*

$$c(jn^2 + kn + \ell)$$

of maximal order dr for all $j, k, \ell \in \mathbb{N}$.

Proof. In a first step, we show how we can find a recurrence of order r for the sequence

$$a(n) = c(d(jn^2 + kn) + \ell).$$

Lemma 11 in [14] shows that M^{pn+q} for $p, q \in \mathbb{Z}$ is a matrix of C-finite sequences. The proof shows that the characteristic polynomials of the sequences is the characteristic polynomial of M^p. Let M_c be the companion matrix of c. Suppose

$$(x - \lambda_1)^{d_1} \cdots (x - \lambda_m)^{d_m}$$

is the characteristic polynomial of c which, by definition of the companion matrix, is also equal to the characteristic polynomial of M_c. Then, by the closed form of C-finite sequences, the characteristic polynomial of $c(pn)$ is given by

$$(x - \lambda_1^p)^{d_1} \cdots (x - \lambda_m^p)^{d_m}$$

which, in turn, is equal to the characteristic polynomial of M_c^p. By (6), the sequences that generate the underlying ring R used for computing a recurrence for $a(n)$ all have characteristic polynomial equal to the characteristic polynomials of M_c^{dk} and M_c^{2dj}. An element in the kernel of the linear system over the field $Q(R)$ can easily be computed if $s = r$. This gives rise to a C^2-finite recurrence of order r for a.

An arbitrary sequence

$$b(n) = c(jn^2 + kn + \ell)$$

can be written as interlacing of sequences

$$a_r(n) = c(d(djn^2 + (2jr + k)n) + jr^2 + kr + \ell)$$

for $r = 0, \dots, d-1$ as the term at index $n = qd + r$ of the interlacing is precisely given by

$$\begin{aligned} a_r(q) &= c(d(djq^2 + (2jr + k)q) + jr^2 + kr + \ell) \\ &= c(j(d^2q^2 + 2rq + r^2) + k(dq + r) + \ell) = c(jn^2 + kn + \ell). \end{aligned}$$

We can compute C^2-finite recurrences of order r for these sequences a_r by the first part of the proof (choosing $j = dj, k = 2jr + k, \ell = jr^2 + kr + \ell$). By Theorem 12 we can therefore compute a C^2-finite recurrence of order dr for b. □

Example 20. Let c be the C-finite sequence (A006131 in the OEIS) satisfying

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

The sequence has eigenvalues $\frac{1\pm\sqrt{17}}{2}$ and their torsion number is 1. The sparse subsequence $a(n) = c(n^2)$ is C^2-finite of order 2 satisfying

$$c_0(n)a(n) - c(4n+3)a(n+1) + c(2n)a(n+2) = 0$$

where c_0 is C-finite of order 2 satisfying

$$4096\,c_0(n) - 144\,c_0(n+1) + c_0(n+2) = 0, \quad c_0(0) = -20, c_0(1) = -1856.$$

6 Outlook

Recently, the class of *simple C^2-finite sequences* has been introduced [24] that satisfies the same computational properties as C^2-finite sequences, but does not share the same technical issues. In particular, it is possible to derive bounds for the asymptotic behavior, there is a characterization through the generating function and closure properties can be computed more efficiently.

It is, however, not clear whether it is possible to derive order bounds for simple C^2-finite sequences as we have presented here for C^2-finite sequence. In that case, one is dealing with an inhomogeneous linear system and the underlying ring is not a principal ideal domain. Hence, one cannot simply bound the rank of modules.

Typically, given a defining recurrence for a C^2-finite sequence, it is difficult to argue that it does not satisfy a shorter recurrence. For D-finite sequences, it is a common strategy to use Guess-and-prove to derive a shorter recurrence (or to find evidence that it is holonomic in the first place). It would be desirable to have a guessing routine for C^2-finite sequences. As a naive approach leads to a non-linear system (see also [29]), it needs to be investigated how this can be solved efficiently.

Acknowledgment

We warmly thank Thierry Combot and Carsten Schneider for their advice.

References

[1] Jean Berstel and Maurice Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bulletin de la Société Mathématique de France*, 104:175–184, 1976.

[2] Richard P. Brent. Fast multiple-precision evaluation of elementary functions. *J. ACM*, 23(2):242–251, 1976.

[3] Henri Cohen. *A Course in Computational Algebraic Number Theory.* Graduate Texts in Mathematics. Springer Berlin Heidelberg, 2013.

[4] Charles W. Curtis and Irving Reiner. *Representation Theory of Finite Groups and Associative Algebras.* AMS Chelsea Publishing Series. Interscience Publishers, 1966.

[5] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence Sequences.* Mathematical Surveys and Monographs. American Mathematical Society, 2015.

[6] Paolo Faccin. *Computational problems in algebra: units in group rings and subalgebras of real simple Lie algebras.* PhD thesis, University of Trento, 2014.

[7] Guoqiang Ge. *Algorithms Related to Multiplicative Representations of Algebraic Numbers.* PhD thesis, U.C. Berkeley, 1993.

[8] Antonio Jiménez-Pastor, Philipp Nuspl, and Veronika Pillwein. On C^2-finite sequences. In *Proceedings of ISSAC 2021, Virtual Event Russian Federation, July 18–23, 2021*, pages 217–224, 2021.

[9] Antonio Jiménez-Pastor, Philipp Nuspl, and Veronika Pillwein. An extension of holonomic sequences: C^2-finite sequences. *Journal of Symbolic Computation*, 116:400–424, 2023.

[10] Manuel Kauers. *Algorithms for Nonlinear Higher Order Difference Equations.* PhD thesis, Johannes Kepler University Linz, 2005.

[11] Manuel Kauers. The Holonomic Toolkit. In *Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions*, Texts and Monographs in Symbolic Computation, pages 119–144. Springer, 2013.

[12] Manuel Kauers. Bounds for D-Finite Closure Properties. In *Proceedings of ISSAC 2014, Kobe, Japan*, pages 288–295, New York, NY, USA, 2014. Association for Computing Machinery.

[13] Manuel Kauers and Peter Paule. *The Concrete Tetrahedron.* Texts and Monographs in Symbolic Computation. Springer, 2011.

[14] Tomer Kotek and Johann A. Makowsky. Recurrence relations for graph polynomials on bi-iterative families of graphs. *Eur. J. Comb.*, 41:47–67, 2014.

[15] Tipaluck Krityakierne and Thotsaporn Aek Thanatipanonda. Ansatz in a Nutshell: A comprehensive step-by-step guide to polynomial, C-finite, holonomic, and C^2-finite sequences, 2022. `https://arxiv.org/abs/2201.08035`.

[16] Arjen K. Lenstra, Hendrik W. jun. Lenstra, and László Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261:515–534, 1982.

[17] Christian Mallinger. Algorithmic Manipulations and Transformations of Univariate Holonomic Functions and Sequences. Diplomarbeit, Johannes Kepler University Linz, 1996.

[18] David W. Masser. Linear relations on algebraic groups. *New Advances in Transcendence Theory*, 1988.

[19] Stephen Melczer. *An Invitation to Analytic Combinatorics.* Texts and Monographs in Symbolic Computation. Springer, 2021.

[20] Johannes Middeke. Symbolic linear algebra, 2019. `https://www3.risc.jku.at/education/courses/ss2019/sla/script_sla2019.pdf`.

[21] Morris Newman. *Integral Matrices.* ISSN. Elsevier Science, 1972.

[22] Philipp Nuspl. C-finite and C^2-finite Sequences in SageMath. RISC Report Series 22-06, Research Institute for Symbolic Computation (RISC), Johannes Kepler University Linz, 2022.

[23] Philipp Nuspl and Veronika Pillwein. A Comparison of Algorithms for Proving Positivity of Linearly Recurrent Sequences. In *Computer Algebra in Scientific Computing*, volume 13366 of *LNCS*, pages 268–287. Springer International Publishing, 2022.

[24] Philipp Nuspl and Veronika Pillwein. Simple C^2-finite Sequences: a Computable Generalization of C-finite Sequences. In Marc Moreno Maza and Lihong Zhi, editors, *ISSAC '22: Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*, pages 45–53. Association for Computing Machinery, 2022.

[25] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2022. `http://www.oeis.org`.

[26] Joël Ouaknine and James Worrell. Decision Problems for Linear Recurrence Sequences. In *Lecture Notes in Computer Science*, pages 21–28. Springer, 2012.

[27] Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. *$A = B$*. CRC Press, 1996.

[28] Carsten Schneider. Minimal representations and algebraic relations for single nested products. *Programming and Computer Software*, 46:133–161, 2020.

[29] Thotsaporn Aek Thanatipanonda and Yi Zhang. Sequences: Polynomial, C-finite, Holonomic, ..., 2020. `https://arxiv.org/pdf/2004.01370`.

[30] Mark van Hoeij. Factoring polynomials and the knapsack problem. *Journal of Number Theory*, 95:167–189, 2002.

[31] Mark van Hoeij. The complexity of factoring univariate polynomials over the rationals. In *Proc. ISSAC'13*, pages 13–14, 2013.

[32] Mark van Hoeij and Andrew Novocin. Gradual sub-lattice reduction and a new complexity for factoring polynomials. *Algorithmica*, 63(3):616–633, 2012.

[33] Tao Zheng. Characterizing triviality of the exponent lattice of a polynomial through galois and galois-like groups. In François Boulier, Matthew England, Timur M. Sadykov, and Evgenii V. Vorozhtsov, editors, *Computer Algebra in Scientific Computing*, pages 621–641. Springer International Publishing, 2020.

[34] Tao Zheng. A fast algorithm for computing multiplicative relations between the roots of a generic polynomial. *Journal of Symbolic Computation*, 104:381–401, 2021.

[35] Tao Zheng and Bican Xia. An effective framework for constructing exponent lattice basis of nonzero algebraic numbers. In *Proceedings of the 2019 on International Symposium on Symbolic and Algebraic Computation*, page 371–378, 2019.