JOHANNES KEPLER
UNIVERSITY LINZ

Author
Philipp Nuspl
01556359

Submission
Research Institute for Symbolic Computation

Supervisor and First Evaluator
Assoc. Prof. Dr. **Veronika Pillwein**

Second Evaluator
Prof. Dr. **James Worrell**

May 2023

Algorithms for linear recurrence sequences

Doctoral Thesis

to obtain the academic degree of

Doktor der Technischen Wissenschaften

in the Doctoral Program

Technische Wissenschaften

JOHANNES KEPLER UNIVERSITY LINZ
Altenbergerstraße 69
4040 Linz, Austria
www.jku.at
DVR 0093696

Abstract

In the past few decades, numerous tools for automatically discovering and proving identities involving sequences and special functions were developed. These tools are often based on algorithms which manipulate sequences satisfying linear recurrences. If the recurrences have constant coefficients, these sequences are called C-finite and in the case of polynomial coefficients they are called D-finite.

We study sequences satisfying recurrences with coefficients which are C-finite themselves and call them C^2-finite. We investigate which properties and algorithms carry over from the classical C-finite and D-finite cases to this new setting. In particular, we show that most so-called closure-properties, which are known for the classical cases, also hold for C^2-finite sequences, i.e., they are closed under termwise addition, termwise multiplication, interlacing and taking subsequences at arithmetic progressions. In many cases these operations are effective and we present algorithms for performing them. In general, however, these algorithms are closely related to and limited by certain decision procedures of C-finite sequences. Deciding whether every term of a sequence is positive or nonzero is not known to be decidable in theory. Nevertheless, we show that it is often easy to decide these properties in practice.

Restricting the ring of C^2-finite sequence to sequences which satisfy a monic (i.e., having constant leading coefficient) linear recurrence with C-finite coefficients, we obtain a subring where all closure properties can be performed effectively. On the other hand, we can allow more general sequences as coefficients. This way we obtain increasingly larger rings where the operations are more difficult to perform.

Most of the theoretical results are also implemented in a package for the computer algebra system SageMath. The thesis contains a tutorial for this package. The tutorial shows how the examples given throughout the thesis can be performed automatically on the computer.

Kurzfassung

In den letzten Jahrzehnten wurden zahlreiche Werkzeuge für das automatische Entdecken und Beweisen von Identitäten von Folgen und speziellen Funktionen entwickelt. Diese Werkzeuge basieren oft auf Algorithmen, die Folgen, die lineare Rekursionen erfüllen, manipulieren. Falls die Rekursionen konstante Koeffizienten haben, nennt man die Folgen C-finit und im Falle von polynomiellen Koeffizienten nennt man sie D-finit.

Wir untersuchen Folgen, die Rekursionen mit Koeffizienten erfüllen, die selbst C-finit sind und nennen sie C^2-finit. Wir untersuchen, welche Eigenschaften und Algorithmen wir vom klassischen C-finiten und D-finiten Fall auf diese neue Klasse übertragen können. Insbesondere zeigen wir, dass die meisten sogenannten closure properties, die für die klassischen Fälle bekannt sind, auch für C^2-finite Folgen gelten. D.h., sie sind abgeschlossen bezüglich termweiser Addition, termweiser Multiplikation und Verflechtung. Außerdem ist die Teilfolge von C^2-finiten Folgen wieder C^2-finit. Diese Operationen sind häufig effektiv und wir stellen Algorithmen vor, die diese Berechnungen durchführen. Im Allgemeinen sind diese Algorithmen jedoch durch bestimmte schwierige Entscheidungsprobleme für C-finite Folgen eingeschränkt. Es ist zum Beispiel nicht bekannt, ob die Probleme, dass jeder Term einer Folge positiv oder ungleich null ist, entscheidbar sind. Wir zeigen jedoch, dass diese Probleme in der Praxis oft einfach zu entscheiden sind.

Schränkt man den Ring der C^2-finiten Folgen auf Folgen ein, die eine normierte (d.h. mit konstantem Leitkoeffizienten) lineare Rekursion mit C-finiten Koeffizienten erfüllen, so erhält man einen Unterring, in dem alle closure properties effektiv durchgeführt werden können. Andererseits können wir auch allgemeinere Folgen als Koeffizienten zulassen. Auf diese Weise erhält man größere Ringe, in denen die Operationen schwieriger durchzuführen sind.

Die meisten der theoretischen Ergebnisse sind auch in einem Softwarepaket für das Computeralgebrasystem SageMath implementiert. Die Dissertation enthält ein Tutorial für dieses Paket. Das Tutorial zeigt, wie die in der Dissertation gegebenen Beispiele automatisch auf dem Computer ausgeführt werden können.

Acknowledgments

First and foremost I would like to thank my advisor Veronika Pillwein. She gave me the freedom to investigate topics I had a strong interest in and at the same time guided me to fruitful and exciting problems in these areas. Thanks to her I learned a lot about mathematics and how research is and should be done. Furthermore, I thank James Worrell for agreeing to serve as the second evaluator of this thesis. It is an honor to have him as a leading expert for linear recurrence sequences and their computational problems in my committee.

Thanks to Antonio Jimenéz Pastor my start in the PhD program went very smoothly. Our discussions over the past years heavily influenced a lot of the research I have done. Likewise, I am very thankful for numerous discussions with Manuel Kauers who has shaped my understanding of computer algebra since the first lectures in linear algebra almost eight years ago. I am grateful for ample constructive feedback on my work during his algebra seminar. Furthermore, I greatly appreciate the comments I got during the combinatorics seminar organized by Peter Paule and Carsten Schneider as well as the DK seminar organized by Veronika Pillwein.

In the past three years I had the pleasure to be part of three different institutes, the DK, RISC and the Institute for Algebra. I enjoyed my time in all three very much and I am very thankful to my colleagues (many of them I now consider good friends) for giving me such a wonderful experience in my PhD program.

Finally, I am very grateful for the help and support I have received over the years from Katharina as well as from my family.

The research presented in this thesis was funded by the Austrian Science Fund (FWF) under the grants W1214-N15, project DK15, P33530 and P31571.

Contents

1 Introduction

Linear recurrence sequences have been a subject of significant interest in mathematics for a long time and the Fibonacci sequence is undoubtedly the most well-known and meticulously studied sequence of this kind. The sequence which was only later (from 1876 on, coined by Lucas [Kos11]) known as the Fibonacci sequence was apparently already studied several hundred years earlier in Indian poetry [Sin85]. Leonardo of Pisa (also called Fibonacci, around 1170–1240) introduced the sequence in his book *Liber Abaci* in the context of a combinatorial problem. He described the development of a rabbit population by the famous recurrence [Sig03]. Later, also Kepler (1571–1630) studied the Fibonacci numbers in his book *Harmonices Mundi* (1619). He observed that the ratio of two consecutive Fibonacci numbers tends to the golden ratio. Furthermore, Kepler found the identity which is known as the Cassini identity today [KADF97]. In the first half of the 18th century, de Moivre (1667–1754) and D. Bernoulli (1700–1782) were among the first ones to investigate similar sequences as the Fibonacci numbers. Both of them also found the closed form of the Fibonacci numbers (later known as Binet's formula) [Ber28, Kos11]. Moreover, de Moivre studied the generating functions of linear recurrence sequences [Moi22, Moi30]. According to [EPSW15], one of the key steps for developing our modern understanding of linear recurrence sequences and in particular their arithmetic properties are the works of Lucas (1842–1891). Today, we call sequences of this type, i.e., satisfying linear recurrences with constant coefficients, *C-finite* [Zei90].

The solutions of linear differential equations (with polynomial coefficients) were a topic of many articles in the 19th century. Deriving a linear recurrence (with polynomial coefficients) for the coefficients of the series solutions of such differential equations were routinely done by mathematicians such as Fuchs (1833–1902) and Frobenius (1849–1917) [Fuc66, Fro75]. In modern language such functions and sequences are called *D-finite* or *holonomic* (the precise correspondence between functions and their coefficient sequences is given in Theorem 2.7). Likewise, arithmetic operations using D-finite objects were already performed in the 19th century, for instance by Hurwitz (1859–1919) and Beke

(1862–1946) [Jun31, Bek94]. Nowadays, the operations which can be performed in the ring of D-finite sequences/functions are called closure properties (cf. Theorem 2.6).

These closure properties of D-finite (or C-finite) sequences are implemented in all major computer algebra systems and are used to automatically prove identities of sequences and special functions. For instance, the identity [Rao53]

$$\sum_{k=0}^{2n} f(k)f(k+1) = f(2n+1)^2 - 1,$$

where $f(n) = \langle 0, 1, 1, 2, 3, 5, \ldots \rangle$ denotes the Fibonacci numbers, can be shown routinely. Combining methods for finding closed form solutions of linear recurrences (cf. [Pet92, Hoe99]) together with a method called creative telescoping (cf. [Zei91, Chy14]) identities of the form

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$

can be discovered and proven automatically [PWZ96].

Considering the usefulness of D-finite and C-finite sequences it is natural to consider generalizations of these classes. Notable extensions from the past two decades include admissable sequences [Kau07a], sequences built by nested expressions [Sch07], coefficient sequences of DD-finite functions [JPP19] or poly-recursive sequences [CMP$^+$21]. In our approach we keep the aspect of considering linear recurrences but we allow more general coefficients. In particular, we consider sequences which satisfy a linear recurrence with coefficients which are C-finite themselves. Such sequences are called C^2-finite and numerous examples can be found in combinatorics and in other areas of research. For instance, the sequence $f(n^2)$ where f, again, denotes the Fibonacci sequence is C^2-finite satisfying the recurrence

$$f(2n+3)f(n^2) + f(4n+4)f((n+1)^2) - f(2n+1)f((n+2)^2) = 0.$$

In this thesis we investigate the computational properties of C^2-finite sequences. First, in Chapter 2 we give an overview of important results of C-finite and D-finite sequences which will be used throughout the thesis. Then, in Chapter 3, we introduce C^2-finite sequences and prove some important facts on their asymptotic behavior, the ring-structure and their generating functions. Chapter 4 discusses how (and which) closure properties of C^2-finite sequences can be performed. One of the differences to C-finite and D-finite sequences is that the same order bounds do not seem too hold. Therefore, in Chapter 5, we show a different way to compute with C^2-finite sequences. Using this approach we can also derive order bounds. Even though closure properties can often be performed in practice, in theory it seems that we are limited by the so-called Skolem Problem (which is the problem of deciding whether a given C-finite sequence contains a zero term, cf. Section 2.2). In Chapter 6 we show that if we fix the leading coefficient in the recurrence of a C^2-finite sequence to be constant, we get a subring where all closure properties can be computed effectively. On the other hand, by allowing other sequences as coefficients in the recurrence we can obtain a chain of increasingly larger difference rings as shown in Chapter 7. The Skolem Problem, which plays an important role in the computations of C^2-finite sequences, can be reduced to showing that a C-finite sequence is positive. In Chapter 8 we compare several, mostly well-known, algorithms for automatically proving positivity of certain sequences. To this effect, we use sequences from The On-Line Encyclopedia of Integer Sequences (OEIS, cf. [OEI23]) for testing the implementations. Furthermore, we study how many of these sequences from the OEIS are in fact C-finite or D-finite. Most of the algorithms discussed throughout the thesis are implemented in a software package for the computer algebra system SageMath [Sag23]. Finally, Chapter 9 can be seen as a tutorial and showcase of this software package.

2 Preliminaries

In this chapter we introduce the notions that we work with throughout the thesis. In particular, we discuss D-finite and C-finite sequences. These are sequences satisfying a linear recurrence with polynomial or constant coefficients, respectively. Furthermore, we introduce important properties of these sequences that we need later.

There are numerous expositions on sequences satisfying linear recurrences with constant or polynomial coefficients. Some of these, that were also used to prepare this thesis, include [Sta80, Zei90, Sta99, FS09, KP11, Kau13, EPSW15].

Throughout this thesis, if not further specified, $\mathbb{K}$ denotes a field of characteristic zero. Often we think about sequences arising from combinatorics, then $\mathbb{K}$ can usually be thought of as the field of rational numbers $\mathbb{Q}$. The natural numbers are denoted by $\mathbb{N} = \{0, 1, 2, \dots, \}$. The $\mathbb{K}$-algebra of sequences is denoted by $\mathbb{K}^{\mathbb{N}}$. For a sequence $a = (a(n))_{n \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$, we sometimes simply write $a(n)$. It is always clear from the context if $a(n)$ denotes the sequence $(a(n))_{n \in \mathbb{N}}$ or the specific term at index n.

2.1 D-finite sequences

The *shift-operator* $\sigma \colon \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$ acts on a sequence $a := (a(n))_{n \in \mathbb{N}}$ in the natural way as $\sigma(a) := (a(n+1))_{n \in \mathbb{N}}$. Let $R \subseteq \mathbb{K}^{\mathbb{N}}$ be a subring of the ring of sequences. The linear recurrence operators $R[\sigma]$ over R form an (in general) noncommutative ring under the usual (i.e., polynomial) addition and multiplication obeying the commutation rule $\sigma \cdot a = \sigma(a) \cdot \sigma$ for $a \in R$.

An element $\mathcal{A} := \sum_{i=0}^{r} c_i \sigma^i \in R[\sigma]$ acts on a sequence a as

$$\mathcal{A} a := \sum_{i=0}^{r} c_i(n) a(n+i).$$

If $c_r \neq 0$, then r is called the *order* of the operator $\mathcal{A}$. If $\mathcal{A}a = 0$, then we call $\mathcal{A}$ an *annihilator* of a and say that $\mathcal{A}$ annihilates a.

Definition 2.1. A sequence $a \in \mathbb{K}^{\mathbb{N}}$ is called *D-finite* (or *P-recursive* or *holonomic*) if there is a nonzero linear recurrence operator with polynomial coefficients $\mathcal{A} \in \mathbb{K}[n][\sigma]$ which annihilates a, i.e., $\mathcal{A}a = 0$.

A D-finite sequence a with annihilator $\mathcal{A}$ is also annihilated by $\mathcal{B}\mathcal{A}$ for any $\mathcal{B} \in \mathbb{K}[n][\sigma]$. Usually we are interested in an operator with minimal order. The order of such an operator is then called the *order* of the sequence a and is denoted by $\operatorname{ord}(a)$.

Let $\mathcal{A} = \sum_{i=0}^{r} p_i \sigma^i \in \mathbb{K}[n][\sigma]$. We can define the *characteristic polynomial* of $\mathcal{A}$ as

$$\chi(\mathcal{A}) = \operatorname{lc}_n \left(\sum_{i=0}^{r} p_i(n) y^i \right) \in \mathbb{K}[y]. \tag{2.1}$$

The roots of this polynomial are called the *eigenvalues* of $\mathcal{A}$. The eigenvalues of a minimal annihilating operator $\mathcal{A}$ of the sequence a are also called the eigenvalues of a. If an annihilating operator $\mathcal{A}$ of the sequence a is fixed, we denote the characteristic polynomial of the annihilator of a simply by $\chi(a)$, i.e., $\chi(a) := \chi(\mathcal{A})$.

Let a be a D-finite sequence annihilated by $\mathcal{A} = \sum_{i=0}^{r} p_i(n)\sigma^i$. Equivalently we can say that a satisfies a linear recurrence with polynomial coefficients

$$p_0(n)a(n) + p_1(n)a(n+1) + \cdots + p_r(n)a(n+r) = 0, \quad \text{for all } n \in \mathbb{N}.$$

A D-finite sequence can always be described by finite amount of data, namely by the coefficients of the recurrence $p_0, \ldots, p_r$ and finitely many initial values $a(0), \ldots, a(m)$. The number of initial values that are needed to uniquely determine the sequence depends on the order r of the recurrence and integer roots of the leading coefficient p_r.

D-finite sequences often appear in combinatorics counting certain objects like graphs, paths, permutations or tilings. Numerous example can be found in the OEIS, *The On-Line Encyclopedia of Integer Sequences* [OEI23]. This database, based on books by Neil Sloane and Simon Plouffe [Slo73, SP95], contains about 360 000 integer sequences at the time of writing (spring 2023), many with detailed information and references.

Example 2.2. Polynomial sequences $p(n)$ with $p \in \mathbb{K}[n]$, geometric sequences α^n with $\alpha \in \mathbb{K}$ and the factorial sequence $n!$ are all D-finite.

Example 2.3. Let $H(n) = \sum_{k=1}^{n} \frac{1}{k}$ be the sequence of harmonic numbers. The sequence is D-finite of order 2 and satisfies the recurrence

$$(n+1)H(n) - (2n+3)H(n+1) + (n+2)H(n+2) = 0, \quad \text{for all } n \in \mathbb{N}.$$

Example 2.4. The Catalan numbers (A000108 in the OEIS [OEI23]) defined as $C(n) = \frac{1}{n+1}\binom{2n}{n}$ are D-finite of order 1 satisfying the recurrence

$$2(2n+1)C(n) - (n+2)C(n+1) = 0.$$

The Catalan numbers have numerous combinatorial interpretations [Sta15].

For proving computational properties of D-finite sequences the following well-known equivalent characterization via vector spaces is often useful [Sta80, Zei90]:

Theorem 2.5. *The sequence $a \in \mathbb{K}^{\mathbb{N}}$ is D-finite if and only if the $\mathbb{K}(n)$-vector space*

$$\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{\mathbb{K}(n)}$$

has finite dimension.

In fact, the dimension of this vector space corresponds precisely to the order of the recurrence.

As an application of Theorem 2.5 the following so-called *closure properties* of D-finite sequences can be shown [Sta80, Zei90, Mal96]:

Theorem 2.6. *Let $a(n), b(n), a_0(n), \dots, a_{m-1}(n)$ be D-finite sequences. Then,*

1. $\sigma(a(n)) = a(n+1)$ *is D-finite of order at most* $\operatorname{ord}(a)$*,*
2. $a(n) + b(n)$ *is D-finite of order at most* $\operatorname{ord}(a) + \operatorname{ord}(b)$*,*
3. $a(n)b(n)$ *is D-finite of order at most* $\operatorname{ord}(a)\operatorname{ord}(b)$*,*

4. *$\sum_{k=0}^{n} a(k)$ is D-finite of order at most* $\operatorname{ord}(a)+1$,
5. *$\sum_{k=0}^{n} a(k)b(n-k)$ is D-finite,*
6. *$a(kn+\ell)$ is D-finite of order at most* $\operatorname{ord}(a)$ *for all $k, \ell \in \mathbb{N}$ and*
7. *the interlacing $e(n) = a_r(q)$ where $n = qm + r$ for $0 \leq r < m$ is D-finite of order at most $m \sum_{r=0}^{m-1} \operatorname{ord}(a_r)$.*

In order to give an idea how these properties are usually shown we prove part 2 on the addition of two sequences.

Proof. By Theorem 2.5 the vector spaces

$$V_a := \langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{\mathbb{K}(n)}, V_b := \langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{\mathbb{K}(n)}$$

both have finite dimension. In fact, V_a has dimension at most $\operatorname{ord}(a)$ and V_b has dimension at most $\operatorname{ord}(b)$. By linearity of σ we have

$$\langle \sigma^i(a+b) \mid i \in \mathbb{N} \rangle_{\mathbb{K}(n)} \subseteq V_a + V_b.$$

The vector space $V_a + V_b$ has dimension at most $\operatorname{ord}(a) + \operatorname{ord}(b)$ as does the subspace. Hence, by Theorem 2.5 the sequence $a + b$ is D-finite with order at most $\operatorname{ord}(a) + \operatorname{ord}(b)$. □

All the properties from Theorem 2.6 are algorithmic and implemented in various computer algebra systems (e.g., `GeneratingFunctions` [Mal96] and `HolonomicFunctions` [Kou10a, Kou10b] for Mathematica, `gfun` [SZ94] for Maple and `ore_algebra` [KJJ15] for SageMath).

D-finite sequences not only often appear in combinatorics but also as the coefficient sequences of functions satisfying linear differential equations:

Theorem 2.7 (Theorem 1.5 in [Sta80])**.** *The sequence $a(n)$ is D-finite if and only if the corresponding generating function $f(x) = \sum_{n\in\mathbb{N}} a(n)x^n \in \mathbb{K}[[x]]$ satisfies a linear differential equation of the form*

$$q_0(x)f(x) + q_1(x)f'(x) + \cdots + q_s(x)f^{(s)}(x) = 0$$

with $q_0, \dots, q_s \in \mathbb{K}[x]$ not all zero.

Using Theorem 2.7 we can see that the coefficient sequences of many special functions are D-finite.

Example 2.8. The trigonometric functions $\sin(x), \cos(x)$, the exponential function $\exp(x)$ and the Bessel functions all satisfy a linear differential equation with polynomial coefficients [DLMF21]. The corresponding coefficient sequences are therefore all D-finite.

Due to the recurrence satisfied by a D-finite sequence, the terms of such a sequence cannot grow arbitrarily. In fact, for every such sequence $a(n)$ there is a constant $\alpha \in \mathbb{Q}$ such that $|a(n)| \leq n!^{\alpha}$ for all $n \geq 2$ [Ger05, Proposition 1.2.1]. We can find even more precise asymptotics for D-finite sequences. For sequences $a(n), b(n) \in \mathbb{K}^{\mathbb{N}}$ we write $a(n) \sim b(n)$ if $\lim_{n\to\infty} \frac{a(n)}{b(n)} = 1$. Using this notion of asymptotic equivalence the following theorem can be proven [WZ85, FS09, MS10, Mel21]:

Theorem 2.9 (Theorem 2 in [Kau13])**.** *Let $a(n)$ be D-finite. Then, there are constants $c_1, \dots, c_m$, polynomials $p_1, \dots, p_m$, natural numbers $r_1, \dots, r_m$, constants $\gamma_1, \dots, \gamma_m, \varphi_1, \dots, \varphi_m, \alpha_1, \dots, \alpha_m$ and natural numbers $\beta_1, \dots, \beta_m$ such that*

$$a(n) \sim \sum_{k=1}^{m} c_k e^{p_k(n^{1/r_k})} n^{\gamma_k n} \varphi_k^n n^{\alpha_k} \log(n)^{\beta_k}.$$

These asymptotics can be used to determine that a sequence is not D-finite. Other techniques for proving that a sequence is not D-finite include the analysis of the corresponding generating function.

Example 2.10. As the terms grow too fast, the sequence 2^{n^2} is not D-finite. Furthermore, the sequences $n^n, \log(n)$ and $P(n)$ where $P(n)$ denotes the n-th prime number are not D-finite [Ger05, FGS05].

A technique that is commonly used for constructing D-finite recurrences is *guessing*. The idea is to guess an operator $\sum_{i=0}^{r} p_i(n)\sigma^i = \sum_{i=0}^{r}\sum_{k=0}^{d} p_{i,k} n^k \sigma^i$ of order r and degree d which annihilates the known terms $a(0), \dots, a(N-1)$ of a sequence. A straightforward approach is to set up a linear system for the $(r+1)(d+1)$ many unknown variables $p_{i,k}$ using the $N-r$ many equations $\sum_{i=0}^{r}\sum_{k=0}^{d} p_{i,k} n^k a(n+i) = 0$ for $n = 0, \dots, N-r-1$. If this system is overdetermined, i.e., if $N - r \geq (r+1)(d+1)$, then any solution of this system is a reasonable guess for an annihilating operator of the sequence a. More advanced techniques for guessing use, for instance, Hermite-Padé approximation, homomorphic images or methods from lattice theory [Kau13, Yur22, KK22].

2.2 C-finite sequences

A difference ring R is a subring of the ring of sequences $\mathbb{K}^{\mathbb{N}}$ which is closed under shifts, i.e., $\sigma(a) \in R$ for all $a \in R$. The closure properties in Theorem 2.6 show that the set of D-finite sequences forms a difference ring under termwise addition and termwise multiplication (also called the *Hadamard product*). Unless specified otherwise, these termwise operations are always our ring operations. A particularly interesting and well-studied subring of the ring of D-finite sequences is the ring of C-finite sequences.

Definition 2.11. A sequence $c \in \mathbb{K}^{\mathbb{N}}$ is called C-finite if there is a nonzero linear recurrence operator with constant coefficients $\mathcal{C} \in \mathbb{K}[\sigma]$ which annihilates c, i.e., $\mathcal{C}c = 0$.

Equivalently, c satisfies a linear recurrence with constant coefficients. The order $\operatorname{ord}(c)$ can again be defined as the order of the minimal nonzero operator which annihilates c. We denote the ring of C-finite sequences by $\mathcal{R}_C$.

Example 2.12. Polynomial sequences $p(n)$ with $p \in \mathbb{K}[n]$ and geometric sequences α^n with $\alpha \in \mathbb{K}$ are C-finite.

Example 2.13. The Fibonacci sequence $f(n) \in \mathbb{Q}^{\mathbb{N}}$ (A000045 in the OEIS) satisfying the recurrence

$$f(n) + f(n+1) - f(n+2) = 0, \quad \text{for all } n \in \mathbb{N},$$

with initial values $f(0) = 0, f(1) = 1$ is C-finite of order 2. The C-finite sequence $l(n)$ satisfying the same recurrence but having initial values $l(0) = 2, l(1) = 1$ is called the Lucas sequence (A000032 in the OEIS).

Example 2.14. The Perrin numbers (A001608 in the OEIS) $p(n) \in \mathbb{Q}^{\mathbb{N}}$ are C-finite of order 3 satisfying the recurrence

$$p(n) + p(n+1) - p(n+3) = 0, \quad \text{for all } n \in \mathbb{N},$$

with initial values $p(0) = 3, p(1) = 0, p(2) = 2$.

Analogous to Theorem 2.5 for D-finite sequences, a sequence c is C-finite if and only if the $\mathbb{K}$-vector space

$$\langle \sigma^i(c) \mid i \in \mathbb{N} \rangle_{\mathbb{K}}$$

has finite dimension. From this property, the same closure properties (including the same bounds for the orders) as for D-finite sequences, Theorem 2.6, can be derived. Further, a sequence $c(n)$ is C-finite if and only if its generating function $f(x) = \sum_{n \in \mathbb{N}} c(n)x^n \in \mathbb{K}[\![x]\!]$ is a rational function [Zei90, KP11].

In contrast to D-finite sequences, C-finite sequences have a nice closed form expression. Namely, they can be written as polynomial-linear combination of geometric sequences.

Theorem 2.15 (Theorem 4.1 in [KP11])**.** *Let c be a C-finite sequence over the field $\mathbb{K}$ with characteristic polynomial*

$$\sum_{i=0}^{r} \gamma_i y^i = y^{n_0} \prod_{i=1}^{m} (y - \lambda_i)^{d_i} \in \mathbb{K}[y]$$

where $\lambda_1, \ldots, \lambda_m \in \mathbb{L} \supseteq \mathbb{K}$ are the pairwise different nonzero eigenvalues of c and $d_1, \ldots, d_m$ their multiplicities. Then, there are $p_1, \ldots, p_m \in \mathbb{L}[n]$ with $\deg(p_i) = d_i - 1$ for $i = 1, \ldots, m$ such that

$$c(n + n_0) = \sum_{i=1}^{m} p_i(n) \lambda_i^n, \quad \text{for all } n \in \mathbb{N}.$$

Theorem 2.15 shows in particular that every C-finite sequence $c \in \mathbb{C}^{\mathbb{N}}$ can be bounded as $|c(n)| \leq \alpha^n$ for all $n \geq 1$ for some $\alpha \in \mathbb{Q}$. In fact, more precise asymptotics can be derived. Namely, the asymptotic behavior is completely governed be the eigenvalues of maximal modulus. Using the notions from Theorem 2.15, let

$$|\lambda_1| = \cdots = |\lambda_k| > |\lambda_{k+1}| \geq \cdots \geq |\lambda_m| \tag{2.2}$$

and $d = \max_{i=1,\dots,k} \deg(p_i)$. Then,

$$c(n) \sim n^d \sum_{i=1}^{k} \text{coeff}(p_i, d)\lambda_i^n \tag{2.3}$$

where $\text{coeff}(p_i, d)$ denotes the coefficient of n^d in $p_i \in \mathbb{K}[n]$ [KP11].

Example 2.16. Let f be the Fibonacci sequence as in Example 2.13. The closed form of f is given by the well known Binet's formula (which was already known to de Moivre and D. Bernoulli in the first half of the 18th century before it was rediscovered by Binet in 1843 [Kos11])

$$f(n) = \tfrac{1}{\sqrt{5}}\left(\tfrac{1+\sqrt{5}}{2}\right)^n - \tfrac{1}{\sqrt{5}}\left(\tfrac{1-\sqrt{5}}{2}\right)^n, \qquad \text{for all } n \in \mathbb{N}.$$

Clearly, $\frac{1+\sqrt{5}}{2}$ is the unique dominant eigenvalue, so $f(n) \sim \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$.

An important notation for C-finite sequences is degeneracy. A C-finite sequence with pairwise different eigenvalues $\lambda_1, \dots, \lambda_m$ is called *degenerate* if there is a quotient $\frac{\lambda_k}{\lambda_i}$ with $k \neq i$ which is a root of unity. Otherwise the sequence is called *nondegenerate*.

Theorem 2.17 ([BM76, EPSW15]). *1. Every C-finite sequence $c(n)$ can be written (effectively) as the interlacing of nondegenerate subsequences $c(dn), \dots, c(dn + d - 1)$ for some $d \in \mathbb{N}$.*

2. Let $c(n)$ be a nondegenerate C-finite sequence. The set

$$Z_c := \{n \in \mathbb{N} \mid c(n) = 0\}$$

is either equal to $\mathbb{N}$ (i.e., c is the zero sequence) or finite.

One of the most celebrated theorems on C-finite sequences is the Skolem-Mahler-Lech theorem which gives a description of the set of indices whose corresponding terms are zero. It was first proven by Skolem for sequences over the rational numbers [Sko33] and later extended by Mahler to number fields (finite algebraic extensions of the rational numbers) [Mah35] and by Lech to fields of characteristic zero [Lec53]. The theorem can be seen as a consequence of Theorem 2.17.

Theorem 2.18 (Skolem-Mahler-Lech)**.** *Let c be a C-finite sequence. The set*

$$Z_c := \{n \in \mathbb{N} \mid c(n) = 0\}$$

can be written as the union of a finite set S and a finite number of arithmetic progressions, i.e.,

$$Z_c = S \cup \{n_1 + pn \mid n \in \mathbb{N}\} \cup \cdots \cup \{n_k + pn \mid n \in \mathbb{N}\} \tag{2.4}$$

for some $n_1, \ldots, n_k, p \in \mathbb{N}$.

Equivalently, the Skolem-Mahler-Lech theorem states that the zeros of a C-finite sequence are cyclic from some term onwards.

Several computational problems are closely related to the Skolem-Mahler-Lech theorem [EPSW15]. First, since the arithmetic progressions in (2.4) can be computed, it can be decided whether or not a given C-finite sequence has infinitely many zeros [BM76]. More difficult is the problem of determining the finite set S [OW12, HHHK05]. The *Skolem Problem* asks whether a given C-finite sequence has any zeros (i.e., provided that the sequence only has finitely many zeros S, is this set S empty). Decidability of the Skolem Problem is only known for special cases, most notably for sequences of order at most 4 [MST84, Ver85]. For sequences of higher order it remains open whether the Skolem Problem is decidable. Some decidability results can be achieved by restricting the C-finite input sequences further or by restricting the set where zeros are sought [LLN+22, BLN+22, KLOW20, LOW21].

The situation for D-finite sequences is even more unclear. Even the Skolem-Mahler-Lech theorem is only known for special cases of D-finite sequences and remains open for general D-finite sequences [BBY12, BCH21]. Futhermore, decidability of checking whether a sequence has infinitely many zeros is only known for certain D-finite sequences of order 2 [NOW21].

3 C^2-finite sequences

D-finite sequences are a natural generalization of C-finite sequences by considering linear recurrences with polynomial instead of constant coefficients. In the same way, we can generalize D-finite sequences by considering linear recurrences with coefficients that are C-finite themselves. These are called C^2-finite sequences. To our knowledge they were described for the first time in the context of graph polynomials [KM14]. Their computational properties and their generating functions were also studied in [TZ20]. The approach we present in this thesis is more computational compared to [KM14]. Our goal is always to check how operations on C^2-finite sequences can be implemented and used in practice for studying concrete problems, e.g., in combinatorics.

Here, we give a basic introduction to C^2-finite sequences and prove some basic facts about them. This chapter is mostly based on [JPNP21, JPNP23, NP22b].

3.1 Definition and examples

Let $\mathbb{K}$ be some field of characteristic zero and $\mathcal{R}_C$ the ring of C-finite sequences over $\mathbb{K}$. For a ring $R \subseteq \mathbb{K}^{\mathbb{N}}$ we denote the set of sequences which are invertible (i.e., not zero divisors) as elements in $\mathbb{K}^{\mathbb{N}}$ by $R^\times$. In particular, $\mathcal{R}_C^\times$ denotes the multiplicatively closed set of all C-finite sequences which do not contain any zeros. The localization of R w.r.t. $R^\times$ is denoted by $Q(R)$ and known as the *total ring of fractions of* R. An element $c(n)/d(n) \in Q(R)$ can be interpreted naturally as a sequence in $\mathbb{K}^{\mathbb{N}}$.

Definition 3.1. A sequence $a \in \mathbb{K}^{\mathbb{N}}$ is called C^2-finite if there is a linear recurrence operator $\mathcal{A} \in \mathcal{R}_C[\sigma]$ with $\operatorname{lc}(\mathcal{A}) \in \mathcal{R}_C^\times$ which annihilates a, i.e., $\mathcal{A}a = 0$.

Equivalently, the sequence a is C^2-finite if there are C-finite sequences $c_0, \dots, c_r$ such that $c_r(n) \neq 0$ for all $n \in \mathbb{N}$ and

$$c_0(n)a(n) + \cdots + c_r(n)a(n+r) = 0, \quad \text{for all } n \in \mathbb{N}. \tag{3.1}$$

The order is again defined analogously to the C-finite and D-finite cases, namely as the order of the minimal operator which annihilates the sequence.

Example 3.2. As polynomials are C-finite, all D-finite sequences (and as such C-finite sequences) are C^2-finite. If the leading coefficient $p_r(n)$ of the linear recurrence has any roots $n \in \mathbb{N}$, the recurrence can be shifted such that all these roots disappear.

Example 3.3. Let $L := \mathbb{K}(q)$ with q transcendental. Then, $a \in \mathbb{L}^{\mathbb{N}}$ is called q-holonomic if it satisfies a linear recurrence

$$p_0(q^n)a(n) + \cdots + p_r(q^n)a(n+r) = 0, \quad \text{for all } n \in \mathbb{N},$$

with coefficients $p_0, \dots, p_r \in \mathbb{L}[x]$, not all zero [KK09]. As all coefficients $p_i(q^n)$ are C-finite over $\mathbb{L}$, such a sequence a is also C^2-finite over $\mathbb{L}$.

Example 3.4. Let $a(n)$ count the number of graphs on n labeled nodes (A006125 in the OEIS). Then, $a(n) = 2^{n(n-1)/2} = 2^{\binom{n}{2}}$ and a is C^2-finite as

$$2^n a(n) - a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

Similarly, it can be easily seen that all sequences α^{n^2} for $\alpha \in \mathbb{K}$ are C^2-finite. The sequence 2^{n^2} is not D-finite. Hence, the set of C^2-finite sequences is certainly a strict generalization of D-finite sequences.

Example 3.5. Let $c(n)$ be a C-finite sequence. Then, the sequence of partial products $a(n) = \prod_{k=0}^{n} c(k)$ is C^2-finite of order 1 satisfying the recurrence

$$c(n+1)a(n) - a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

If $c(n)$ denotes the Fibonacci sequence, the terms $a(n)$ are also known as *fibonorials* or *Fibonacci factorials* (A003266 in the OEIS). This sequence and similar sequences have been studied in several works, e.g., in [Bro72, Mar13, BCMS20].

Example 3.6. Let $c(n)$ be a C-finite sequence with $c(n) \neq 0$ for all $n \in \mathbb{N}$. The sequence $a(n) = \frac{1}{c(n)}$ is not C-finite in general (unless $c(n)$ is the interlacing of geometric sequences [LT90]). The sequence $a(n)$ is, however, C^2-finite satisfying

$$c(n)a(n) - c(n+1)a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

In particular, by the closure properties of C^2-finite sequences (Theorem 5.8), the sequence

$$a(n) = \frac{f(n)}{f(n+1)} + \sum_{k=1}^{n} \frac{(-1)^k}{f(k)f(k+1)}$$

is C^2-finite (where f denotes the Fibonacci numbers). In fact it was shown that $a(n) = 0$ for all $n \in \mathbb{N}$ [Kau05, Example 4.7].

Example 3.7. Let $f(n)$ denote the Fibonacci numbers. It was observed in [KM14] that

$$\begin{aligned} f(2n+3)(f(2n+1)f(2n+3) - f(2n+2)^2)f(n^2) \\ +f(2n+2)(f(2n+3)+f(2n+1))f((n+1)^2) \\ -f(2n+1)f((n+2)^2) = 0 \end{aligned}$$

holds for all $n \in \mathbb{N}$. Hence, $f(n^2)$ is C^2-finite (A054783 in the OEIS). In fact, the C-finite coefficients can be simplified and we obtain the recurrence:

$$f(2n+3)f(n^2) + f(4n+4)f((n+1)^2) - f(2n+1)f((n+2)^2) = 0.$$

Example 3.8. Let $f(n)$ denote the Fibonacci numbers and $l(n)$ the Lucas numbers (as in Example 2.13). Let $\mathrm{Fib}(n,k) = \prod_{i=1}^{k} f(n-i+1)/f(i)$ be the *fibonomial coefficient*. It has been shown [KAO12, Theorem 1] that

$$a(n) = \sum_{k=0}^{n} \mathrm{Fib}(2n+1,k) = \prod_{k=1}^{n} l(2k), \quad \text{for all } n \in \mathbb{N}. \tag{3.2}$$

Hence, the sequence a (one half of the sequence A294349 in the OEIS) is C^2-finite and satisfies the recurrence

$$l(2n+2)a(n) - a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

Numerous other identities of fibonomial coefficients can be found in [ST05, BR14].

Identities of the form of Example 3.8 can also be viewed as a multibasic hypergeometric identity. The difficulty comes from the relations among the bases [BP99]. The identity (3.2) can be proven fully automatically using the tools developed in [AS21] or using a form of multibasic creative telescoping.

Open Question 3.9. Can one develop a version of creative telescoping for C^2-finite sequences which can prove identities as in Example 3.8 automatically?

Such a tool could, for instance, be used to verify the following example.

Example 3.10. The sequence $a(n) = \sum_{k=0}^{n} \operatorname{Fib}(n,k)$ seems to be C^2-finite of order 4 satisfying the recurrence

$$-(l(2n+2)+2)a(n) - l(n+2)a(n+2) + a(n+4) = 0.$$

The recurrence was obtained by guessing.

Example 3.11. Let $c(n)$ be a sequence and let

$$a(n) = [c(0), c(1), \dots, c(n)] = c(0) + \cfrac{1}{c(1) + \cfrac{1}{\ddots + \cfrac{1}{c(n)}}}$$

be the n-th convergent of the continued fraction

$$[c(0), c(1), c(2), c(3), \dots] = c(0) + \cfrac{1}{c(1) + \cfrac{1}{c(2) + \cfrac{1}{c(3) + \dots}}}$$

The n-th convergents $a(n) = \frac{p(n)}{q(n)}$ can be described by the recurrences

$$\begin{aligned} p(n) + c(n+2)p(n+1) - p(n+2) &= 0, \quad & p(0) &= c(0), \quad & p(1) &= c(0)c(1) + 1, \\ q(n) + c(n+2)q(n+1) - q(n+2) &= 0, \quad & q(0) &= 1, \quad & q(1) &= c(1). \end{aligned}$$

In particular, if c is a C-finite sequence, these sequences p, q are C^2-finite. For instance, if c (A003417 in the OEIS) satisfies

$$c(n+1) - 2c(n+4) + c(n+7) = 0$$

with initial values $a(n) = (2, 1, 2, 1, 1, 4, 1, \dots)$, then $a(n) \to e$ [Eul44, Coh06].

From recurrence (3.1) it is clear that every C^2-finite sequence can be described by finite data: The coefficients of the recurrence can be described by finite amount of data and the initial values uniquely define the sequence. I.e., given the coefficients $c_0, \dots, c_r$ of the recurrence and the initial values $a(0), a(1), \dots, a(r-1)$, the term $a(n)$ of the sequence can be computed. Substantial amount of research has been done to see how was fast terms $c(n)$ for a C-finite, D-finite or q-holonomic sequence c can be evaluated [BGS07, Bos20, BM21].

Open Question 3.12. Let a be a C^2-finite sequence and $n \in \mathbb{N}$. How fast can we compute $a(n)$?

It is in general not known whether we can decide that the leading coefficient $c_r(n)$ in the recurrence has no zeros (cf. Skolem Problem on page 12). Alternatively, we could allow the leading coefficient $c_r(n)$ to have at most finitely many zeros (as in [KM14] and later in Chapter 5). This alternative definition is equivalent to the definition given here as the recurrence can be shifted in order to attain a recurrence with a leading coefficient without any zeros. The advantage in allowing finitely many zeros is that this property is decidable [BM76]. For computations in practice, we would still have to compute the finite set of zeros to know how many initial values we have to save (where we would be limited by the Skolem Problem again).

A sequence is called *X-recursive* if it satisfies a linear recurrence with C-finite coefficients [TZ20]. I.e., the leading coefficient can have infinitely many zeros in this case. The disadvantage when dealing with X-recursive sequences is that finite amount of data might not suffice to uniquely define a sequence. Also, sequences might grow arbitrarily fast.

Example 3.13. Let $c(n)$ be the interlacing of the constant 0 and the constant 1 sequence. Then, c is C-finite of order 2 satisfying $c(n) - c(n+2) = 0$ with $c(0) = 0, c(1) = 1$. Any sequence $a(n)$ satisfying

$$c(n)a(n) - c(n)a(n+1) = 0, \quad \text{for all } n \in \mathbb{N},$$

is X-recursive. The given recurrence only forces $a(n) = a(n+1)$ for odd $n \in \mathbb{N}$. Hence, there are uncountably many X-recursive sequences over $\mathbb{Q}$ whereas there are only countably many C^2-finite sequences over $\mathbb{Q}$.

3.2 Asymptotics and counterexamples

Proving that a sequence satisfies a recurrence of a certain type is usually significantly easier than proving that a recurrence does not satisfy any recurrence of a certain type.

Example 3.14. Let $a(n) = (-1)^{\lceil \log(n+1) \rceil} + 1$. Then, for every $r \in \mathbb{N}$ the sequence contains a run of at least r consecutive zeros. Hence, if the sequence would be C^2-finite of order r, then the sequence would be eventually zero. Therefore, the sequence cannot be C^2-finite.

Example 3.15. By definition, every term of a C^2-finite sequence can be obtained by field operations of finitely many elements from the ground field $\mathbb{K}$ (the initial values of the C-finite coefficients, the coefficients of the C-finite recurrences and the initial values of the C^2-finite recurrence). Hence, all terms of a C^2-finite sequence have to be in a finite extension field of $\mathbb{Q}$ (cf. [Ger05, Proposition 1.3.3]). Therefore, the sequence $a(n) = \sqrt{n}$ cannot be C^2-finite.

Another method for showing that a sequence is not C^2-finite is by showing that it grows too fast asymptotically. Lemma 5 in [KM14] states, without a proof, that every C^2-finite sequence $a(n)$ with leading coefficient $c_r(n) \in \mathbb{Z}$ for all $n \in \mathbb{N}$ can be bounded by $|a(n)| \leq \alpha^{n^2}$ for some $\alpha \in \mathbb{Q}$. In fact, such a bound can be shown for any C^2-finite sequence (our proof follows the proof for D-finite sequences [Ger05, Proposition 1.2.1]).

Lemma 3.16. *Let $a \in \mathbb{C}^{\mathbb{N}}$ be C^2-finite. Then, there is an $\alpha \in \mathbb{Q}$ such that $|a(n)| \leq \alpha^{n^2}$ for all $n \geq 1$.*

Proof. Suppose a satisfies the recurrence

$$c_0(n)a(n) + \cdots + c_{r-1}(n)a(n+r-1) + c_r(n)a(n+r) = 0$$

for all $n \in \mathbb{N}$ with $c_0, \ldots, c_r \in \mathcal{R}_C$ and $c_r(n) \neq 0$ for all $n \in \mathbb{N}$. Then, for all $c_i(n)$ with $i = 0, \ldots, r-1$ there exists an $\alpha_i \in \mathbb{Q}$ such that $|c_i(n)| \leq \alpha_i^n$ for all $n \geq 1$. Furthermore, there exists an $\alpha_r \in \mathbb{Q}$ such that $\frac{1}{|c_r(n)|} \leq \alpha_r^n$ for all $n \in \mathbb{N}$ (cf. [EPSW15, Theorem 2.3]). Let $1 \leq \alpha \in \mathbb{Q}$ be large enough such that

$$\sum_{i=0}^{r-1} \frac{\alpha_i^n}{\alpha_r^n} \leq r \left(\max_{i=0,\ldots,r-1} \frac{\alpha_i}{\alpha_r} \right)^n \leq \alpha^n$$

for $n \geq 1$ and large enough such that $|a(n)| \leq \alpha^{n^2}$ holds for $n = 1, \ldots, r-1$. We show $|a(n)| \leq \alpha^{n^2}$ by induction on n. Suppose the inequality holds for all $a(i)$ with $i \leq n+r-1$. In the induction step we have

$$\begin{aligned} |a(n+r)| &= \left| \sum_{i=0}^{r-1} \frac{c_i(n)}{c_r(n)} a(n+i) \right| \leq \sum_{i=0}^{r-1} \frac{|c_i(n)|}{|c_r(n)|} |a(n+i)| \leq \sum_{i=0}^{r-1} \frac{\alpha_i^n}{\alpha_r^n} \alpha^{(n+i)^2} \\ &\leq \alpha^{(n+r-1)^2} \sum_{i=0}^{r-1} \frac{\alpha_i^n}{\alpha_r^n} \leq \alpha^{(n+r-1)^2} \alpha^n \leq \alpha^{(n+r)^2}. \end{aligned}$$

□

Example 3.17. Lemma 3.16 shows that the sequences 2^{n^3} and $\prod_{i=0}^n i!$ are not C^2-finite.

For a given C^2-finite sequence a it is, in general, not clear whether such an α with $|a(n)| \leq \alpha^{n^2}$ can be computed. The obstacle is the computation of a number $\alpha_r \in \mathbb{Q}$ such that $\frac{1}{|c_r(n)|} \leq \alpha_r^n$ for all $n \in \mathbb{N}$ [EPSW15]. In special cases, however, such an α_r can be found, e.g., if $c_r(n)$ is a polynomial sequence.

The sequence α^{n^2} is C^2-finite. Hence, there are examples where the bound in Lemma 3.16 is tight. Determining more precise asymptotics for C^2-finite sequences could turn out useful to determine that certain sequences are not C^2-finite. For instance, the asymptotics of D-finite sequences (Theorem 2.9) can be used to show that the sequence of prime numbers is not D-finite [Mel21, Example 2.26].

Open Question 3.18. Determine the asymptotics of C^2-finite sequences (or a subclass) analogously to the asymptotics of D-finite sequences.

D-finite sequences in general exhibit more complex asymptotics than C-finite sequences. Let

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

denote Apéry's numbers (A005259 in the OEIS). The sequence is D-finite of order 2 but is not C-finite as can be seen by the asymptotics (cf. [Hir12])

$$a(n) \sim 2^{-9/4}\pi^{-3/2}(\sqrt{2}+1)^{4n+2}n^{-3/2}.$$

An analogous example for C^2-finite sequences could answer a question from Armin Straub (Open Question 3.18 might also shed some light on this problem):

Open Question 3.19. Find a C^2-finite integer sequence $a \in \mathbb{Z}^{\mathbb{N}}$ which does not grow faster than a D-finite (C-finite) sequence but is not D-finite (C-finite) itself.

The sequence n^n is another common counterexample for a sequence which is not D-finite [Ger04]. It turns out, this sequence is also not C^2-finite.

Example 3.20. The sequence n^n is neither polynomial nor rational recursive, i.e., it cannot be described by a certain system of polynomial or rational difference equations [CMP+21]. As C^2-finite sequences are examples of rational recursive sequences, the sequence n^n is not C^2-finite.

3.3 Algebraic characterization

In this section we derive an equivalent characterization of C^2-finite sequences in terms of finitely generated modules which is analogous to the characterization in Theorem 2.5 for D-finite sequences. This characterization can then be used to show that the set of C^2-finite sequences forms a difference ring.

For C-finite sequences $c_0, \dots, c_r \in \mathcal{R}_C$ we denote the smallest $\mathbb{K}$-difference algebra (i.e., a $\mathbb{K}$-algebra which is additionally closed under shifts) which contains the sequences $c_0, \dots, c_r$ by $\mathbb{K}_\sigma[c_0, \dots, c_r]$. For a sequence $a \in \mathbb{K}^\mathbb{N}$ and a subring $S \subseteq \mathcal{R}_C$, we consider the module of shifts over the ring $Q(S)$,

$$\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$$

where the scalar multiplication is given by the Hadamard product of sequences in $\mathbb{K}^\mathbb{N}$. In Theorem 3.23 below, we prove that this module (with $S = \mathcal{R}_C$) is finitely generated if and only if the sequence is C^2-finite. For this purpose, we show two auxiliary lemmas first.

Lemma 3.21. *Let a be C^2-finite with annihilating operator $\mathcal{A} = c_0 + \dots + c_r\sigma^r$ and let $R := \mathbb{K}_\sigma[c_0, \dots, c_r]$. If $S \supseteq R$ is a subring of the ring of sequences $\mathbb{K}^\mathbb{N}$, then $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ is finitely generated.*

Proof. By assumption we have $\operatorname{lc}(\mathcal{A}) = c_r \in \mathcal{R}_C^\times$ and $\mathcal{A}a = 0$. Let $i \in \mathbb{N}$. Then,

$$\sigma^i \mathcal{A} = \sigma^i(c_0)\sigma^i + \dots + \sigma^i(c_r)\sigma^{i+r}$$

and $\operatorname{lc}(\sigma^i\mathcal{A}) = \sigma^i(c_r) \in \mathcal{R}_C^\times$. Since $(\sigma^i\mathcal{A})a = \sigma^i(\mathcal{A}a) = 0$, we can write

$$\sigma^{i+r}(a) = -\frac{\sigma^i(c_0)}{\sigma^i(c_r)}\sigma^i(a) - \dots - \frac{\sigma^i(c_{r-1})}{\sigma^i(c_r)}\sigma^{i+r-1}(a).$$

Hence, for all $i \in \mathbb{N}$ the sequence $\sigma^{i+r}(a)$ is a $Q(R)$-linear combination of the sequences $\sigma^i(a), \dots, \sigma^{i+r-1}(a)$. By induction, $\sigma^{i+r}(a)$ is a $Q(R)$- and therefore a $Q(S)$-linear combination of $a, \sigma(a), \dots, \sigma^{r-1}(a)$. Thus, the module $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ is generated by $a, \sigma(a), \dots, \sigma^{r-1}(a)$. $\square$

Lemma 3.22. *Let $a \in \mathbb{K}^\mathbb{N}$ and $S \subseteq \mathcal{R}_C$ a subring. If $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ is finitely generated, then a is C^2-finite.*

Proof. As the module is finitely generated, we can write

$$\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)} = \langle g_1, \dots, g_m \rangle_{Q(S)}$$

for some m and some sequences $g_1, \dots, g_m$. There exists an $r \in \mathbb{N}$ such that the elements g_j can be written as $g_j = \sum_{i=0}^{r-1} c_{i,j}\sigma^i(a)$ for some $c_{i,j} \in Q(S)$. Then, $\sigma^r(a)$ is a $Q(S)$-linear combination of $g_1, \dots, g_m$, so in particular a linear combination of the sequences

$a, \sigma(a), \ldots, \sigma^{r-1}(a)$. Hence, there exist sequences $c_0, \ldots, c_{r-1} \in S$ and $d_0, \ldots, d_{r-1} \in S^\times$ with

$$\sigma^r(a) = \frac{c_0}{d_0}a + \frac{c_1}{d_1}\sigma(a) + \cdots + \frac{c_{r-1}}{d_{r-1}}\sigma^{r-1}(a).$$

Clearing denominators shows that a is C^2-finite of order at most r. □

Theorem 3.23. *The following are equivalent:*

1. *The sequence a is C^2-finite.*
2. *There exists $\mathcal{A} \in \mathcal{R}_C[\sigma]$ with $\operatorname{lc}(\mathcal{A}) \in \mathcal{R}_C^\times$ and a C^2-finite sequence b with $\mathcal{A}a = b$.*
3. *The module $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)}$ over the ring $Q(\mathcal{R}_C)$ is finitely generated.*

Proof. (1) ⇒ (2): We can choose the C^2-finite sequence $b = 0$.

(2) ⇒ (1): Since b is C^2-finite, there exists an operator $\mathcal{B} \in \mathcal{R}_C[\sigma]$ with $\operatorname{lc}(\mathcal{B}) \in \mathcal{R}_C^\times$ and $\mathcal{B}b = 0$. Then, $(\mathcal{B}\mathcal{A})a = \mathcal{B}(\mathcal{A}a) = \mathcal{B}b = 0$. Furthermore,

$$\operatorname{lc}(\mathcal{B}\mathcal{A}) = \operatorname{lc}(\mathcal{B})\sigma^{\operatorname{ord}(\mathcal{B})}(\operatorname{lc}(\mathcal{A})) \in \mathcal{R}_C^\times.$$

(1) ⇒ (3): Follows from Lemma 3.21 with $S = \mathcal{R}_C$.

(3) ⇒ (1): Follows from Lemma 3.22 with $S = \mathcal{R}_C$. □

In the case of D-finite and C-finite sequences, the base ring of the finitely generated module is a field. As we have seen in Theorem 2.6 the key step for proving that these sets form rings makes use of the fact that submodules of finitely generated modules over fields (i.e., finite dimensional vector spaces) are again finitely generated. This holds more generally for Noetherian rings. However, the rings $\mathcal{R}_C$ and $Q(\mathcal{R}_C)$ are not Noetherian.

Example 3.24. Let $c_k \in \mathcal{R}_C$ be defined by $c_k(n) - c_k(n+k) = 0$ for every $n \in \mathbb{N}$, and $c_k(0) = \cdots = c_k(k-2) = 1, c_k(k-1) = 0$ (i.e., c_k has a 0 at every k-th term and 1 else). Let $L_m = \langle c_2, \ldots, c_{2^m} \rangle$ be ideals in $\mathcal{R}_C$ for $m \in \mathbb{N}$. By construction, $c_{2^m} \notin L_{m-1}$ since $c(2^{m-1} - 1) \neq 0$ and for every sequence $d(n) \in L_{m-1}$ we have $d(2^{m-1} - 1) = 0$. Hence,

$$L_1 \subsetneq L_2 \subsetneq L_3 \subsetneq \cdots$$

is an infinitely properly ascending chain of ideals in $\mathcal{R}_C$. Therefore, $\mathcal{R}_C$ is not a Noetherian ring.

However, instead of considering the whole ring of C-finite sequences, we can instead limit the base ring to a ring of the form $\mathbb{K}_\sigma[c_0, \ldots, c_r]$. By the properties of C-finite sequences, these algebras are Noetherian rings.

Lemma 3.25. *Let $c_0, \ldots, c_r \in \mathcal{R}_C$. Then, $\mathbb{K}_\sigma[c_0, \ldots, c_r]$ is a Noetherian ring.*

Proof. All the $\mathbb{K}$-vector spaces $\langle \sigma^i(c_j) \mid i \in \mathbb{N} \rangle_{\mathbb{K}}$ are finitely generated. Hence, also the difference algebras $\mathbb{K}_\sigma[c_j]$ are finitely generated. Therefore, also $\mathbb{K}_\sigma[c_0, \ldots, c_r]$ is finitely generated and a Noetherian ring [AM69, Corollary 7.7]. $\square$

Theorem 3.26. *The set of C^2-finite sequences is a difference ring under termwise addition and termwise multiplication.*

Proof. Let a, b be C^2-finite sequences and $\mathcal{A} = c_0 + c_1\sigma + \cdots + c_{r_1}\sigma^{r_1}$ and $\mathcal{B} = d_0 + d_1\sigma + \cdots + d_{r_2}\sigma^{r_2}$ the corresponding annihilating operators with $c_0, \ldots, c_{r_1}, d_0, \ldots, d_{r_2} \in \mathcal{R}_C$.

Now, let

$$S = \mathbb{K}_\sigma[c_0, \ldots, c_{r_1}, d_0, \ldots, d_{r_2}].$$

By Lemma 3.25, this ring S is Noetherian. Therefore, also $Q(S)$ is a Noetherian ring [AM69, Proposition 7.3]. By Lemma 3.21, the modules $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ and $\langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{Q(S)}$ are both finitely generated $Q(S)$-modules. Hence, also the modules

$$\langle \sigma^i(a+b) \mid i \in \mathbb{N} \rangle_{Q(S)} \subseteq \langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)} + \langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{Q(S)}$$

and

$$\langle \sigma^i(ab) \mid i \in \mathbb{N} \rangle_{Q(S)} \subseteq \langle \sigma^i(a)\sigma^j(b) \mid i, j \in \mathbb{N} \rangle_{Q(S)}$$

are finitely generated as they are submodules of finitely generated modules over a Noetherian ring. By Lemma 3.22, the sequences $a + b$ and ab are C^2-finite. Therefore, the set of C^2-finite sequences is a ring.

The operator

$$\tilde{\mathcal{A}} = \sigma(c_0) + \sigma(c_1)\sigma + \cdots + \sigma(c_{r_1})\sigma^{r_1} \in \mathcal{R}_C[\sigma]$$

annihilates $\sigma(a)$ as

$$\tilde{\mathcal{A}}(\sigma(a)) = (\tilde{\mathcal{A}}\sigma)a = (\sigma\mathcal{A})a = \sigma(\mathcal{A}a) = 0.$$

Furthermore, we have $\operatorname{lc}(\tilde{\mathcal{A}}) = \sigma(c_{r_1}) \in \mathcal{R}_C^\times$. Hence, the ring of C^2-finite sequences is also closed under shifts. $\square$

The statement that C^2-finite sequences form a ring can also be found in [KM14, Corollary 15]. Our proof is different and closer resembles the proofs for the classical C-finite and D-finite cases. Furthermore, their proof of Lemma 14, on which their result builds, seems to contain a mistake. In the proof they choose a certain sequence s_n which is not guaranteed to exist.

3.4 Generating functions

Switching between the generating function representation $g(x) = \sum_{n\geq 0} a(n)x^n \in \mathbb{K}[\![x]\!]$ of a D-finite (or C-finite) sequence $a(n) \in \mathbb{K}^{\mathbb{N}}$ and back often turns out to be a useful technique when proving properties of sequences. In this section we examine the generating functions of C^2-finite sequences.

For a C^2-finite sequence a over the field $\mathbb{K}$ with annihilating operator $c_0 + \cdots + c_r\sigma^r$, the smallest field $\mathbb{L} \supseteq \mathbb{K}$ which contains all the splitting fields of the characteristic polynomials of $c_0, \ldots, c_r$ is called the *splitting field of a*.

For natural numbers $n, k \in \mathbb{N}$ we write

$$n^{\underline{k}} = n(n-1)\cdots(n-k+1)$$

for the *falling factorial*. Let $g(x) = \sum_{n\geq 0} a(n)x^n \in \mathbb{L}[\![x]\!]$. Then, for $\lambda \in \mathbb{L}$ we write $g^{(d)}(\lambda x)$ for the d-th derivative of the formal power series $g(\lambda x) \in \mathbb{L}[\![x]\!]$, i.e.,

$$g^{(d)}(\lambda x) = \sum_{n\geq d} n^{\underline{d}} \lambda^n a(n) x^{n-d}.$$

Theorem 3.27. *Let a be a C^2-finite sequence over $\mathbb{K}$ with splitting field $\mathbb{L}$. Let $g(x) = \sum_{n\geq 0} a(n)x^n$ be its generating function. Then, $g(x)$ satisfies a functional equation of the form*

$$\sum_{k=1}^{m} p_k(x) g^{(d_k)}(\lambda_k x) = p(x) \tag{3.3}$$

for $p, p_1, \ldots, p_m \in \mathbb{L}[x], d_1, \ldots, d_m \in \mathbb{N}$ and $\lambda_1, \ldots, \lambda_m \in \mathbb{L}$.

Proof. Consider the defining recurrence of a:

$$c_0(n)a(n) + \cdots + c_r(n)a(n+r) = 0, \quad \text{for all } n \in \mathbb{N}.$$

Multiplying by x^n and summing over all $n \in \mathbb{N}$ yields

$$\sum_{n\geq 0} c_0(n)a(n)x^n + \cdots + \sum_{n\geq 0} c_r(n)a(n+r)x^n = 0. \tag{3.4}$$

The coefficients $c_0, \ldots, c_r$ have some closed form for all $n \geq n_0$. Hence, the left-hand side of equation (3.4) is just an $\mathbb{L}$-linear combination of power series of the form $\tilde{h}(x) = \sum_{n\geq n_0} n^j \lambda^n a(n+i)x^n$ for $j \in \mathbb{N}, i \in \{0, \ldots, r\}, \lambda \in \mathbb{L}$. The first terms $n = 0, \ldots, n_0 - 1$ in (3.4) just yield some polynomial factors. Furthermore, it is sufficient to consider formal power series of the form $h(x) = \sum_{n\geq 0} n^j \lambda^n a(n+i)x^n$ as $h(x) - \tilde{h}(x)$ is again polynomial. Hence, also these factors $h(x) - \tilde{h}(x)$ contribute to the right-hand side of (3.3).

Let $S(k,l)$ denote the Stirling numbers of the second kind. Then, $n^k = \sum_{l=0}^{k} S(k,l) n^{\underline{l}}$. Therefore,

$$\begin{aligned}
h(x) &= \sum_{n \geq i} (n-i)^j \lambda^{n-i} a(n) x^{n-i} = \sum_{n \geq i} \left(\sum_{k=0}^{j} \binom{j}{k} n^k (-i)^{j-k} \right) \lambda^{n-i} a(n) x^{n-i} \\
&= \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{j}{k} (-i)^{j-k} S(k,l) \sum_{n \geq i} n^{\underline{l}} \lambda^{n-i} a(n) x^{n-i} \\
&= \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{j}{k} (-i)^{j-k} S(k,l) \frac{x^{l-i}}{\lambda^i} \sum_{n \geq i} n^{\underline{l}} \lambda^{n} a(n) x^{n-l} \\
&= \sum_{k=0}^{j} \sum_{l=0}^{k} \binom{j}{k} (-i)^{j-k} S(k,l) \frac{x^{l-i}}{\lambda^i} \left(g^{(l)}(\lambda x) + p_l(x) \right)
\end{aligned}$$

where $p_l(x) \in \mathbb{L}[x]$ is defined as

$$p_l(x) = \begin{cases} -\sum_{n=l}^{i-1} n^{\underline{l}} \lambda^n a(n) x^{n-l}, & \text{if } i > l, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $h(x) = \sum_{l=0}^{j} q_l(x) g^{(l)}(\lambda x) + q(x)$ with $q_0, \ldots, q_j, q \in \mathbb{L}(x)$. Using this in equation (3.4) and clearing the denominator x^r yields a functional equation of the desired form.

This functional equation is nontrivial, i.e., the left-hand side of (3.3) does not simplify to zero: Fix some λ and consider a term $n^j \lambda^n a(n+i)$ with j maximal and i minimal among these maximal j. This term yields a nonzero term $x^{j-i+r} g^{(j)}(\lambda x)$ in the functional equation which cannot be canceled because of the choice of j and i. □

The proof of Theorem 3.27 uses the closed form representation of the C-finite coefficients and generalizes the classical proof for D-finite sequences [Sta80, Mal96, KP11]. A close investigation of the proof shows the following bounds (in the special case of D-finite sequences we get precisely the known bounds):

1. We have $\deg(p_k) \leq r + \max_i(\operatorname{ord}(c_i)) - 1$.

2. The λ_k are exactly the eigenvalues of the C-finite coefficients c_i.

3. The derivatives d_k are each bounded by the multiplicity of the eigenvalue λ_k in any c_i. In particular, $\max_k(d_k) \leq \max_i(\operatorname{ord}(c_i)) - 1$.
4. Let n_0 be minimal such that all $c_0, \dots, c_r$ have closed forms from n_0 on. We have $\deg(p) < \max(r, n_0)$. If we differentiate the functional equation $\max(r, n_0)$ times, we get a homogeneous functional equation (i.e., $p = 0$). The functional equation then satisfies $\max_k(d_k) \leq \max_i(\operatorname{ord}(c_i)) - 1 + \max(r, n_0)$.

Theorem 3.27 also generalizes the result for q-holonomic sequences: Every generating function of a q-holonomic sequence satisfies a q-shift equation [KK09]. In this case we would have $\lambda_k = q^k$.

Example 3.28. Let $a(n) := f(n^2)$ be the sparse subsequence of the Fibonacci sequence f (cf. Example 3.7). The generating function g of a satisfies the functional equation

$$\left(\phi^3 x^2 - \phi^{-3}\right) g\left(\phi^2 x\right) - \left(\psi^3 x^2 - \psi^{-3}\right) g\left(\psi^2 x\right) + x g\left(\phi^4 x\right) - x g\left(\psi^4 x\right) = (\psi - \phi) x$$

where $\phi := \frac{1+\sqrt{5}}{2}$ denotes the golden ratio and $\psi := \frac{1-\sqrt{5}}{2}$ its conjugate.

Example 3.29. Since $\frac{1}{n!}$ is C^2-finite (as it is D-finite), the coefficient sequence of the exponential generating function $\sum_{n\geq 0} \frac{a(n)}{n!} x^n$ of a C^2-finite sequence a is again C^2-finite. Let b be the coefficient sequence of the exponential generating function of the fibonorial numbers $\prod_{i=1}^{n} f(i)$, where f denotes the Fibonacci numbers (cf. Example 3.5). Then, b satisfies

$$f(n+1)b(n) - (n+1)b(n+1) = 0, \qquad \text{for all } n \in \mathbb{N}.$$

Let $h(x) = \sum_{n\geq 0} b(n) x^n$ be the generating function of b. Then, h satisfies

$$\phi\, h(\phi x) - \psi\, h(\psi x) - (\phi - \psi)\, h'(x) = 0$$

where ϕ, ψ are as in Example 3.28.

Theorem 3.27 shows that the generating function $g(x)$ of a C^2-finite sequence satisfies a linear differential equation with polynomial coefficients which can also contain terms involving the generating function with scaled argument $g(\lambda x)$. On the contrary, we can ask whether the coefficient sequence of any function satisfying an equation of this type is C^2-finite. This is not necessarily the case. In general, the coefficient sequence only satisfies a linear recurrence with C-finite coefficients where the leading coefficient might have infinitely many zeros. I.e., such sequences are always X-recursive.

Theorem 3.30. *Let $g(x) = \sum_{n\geq 0} a(n)x^n$ satisfy a functional equation of the form*

$$\sum_{k=1}^{m} p_k(x) g^{(d_k)}(\lambda_k x) = p(x)$$

for $p, p_1, \ldots, p_m \in \mathbb{L}[x], d_1, \ldots, d_m \in \mathbb{N}$ and $\lambda_1, \ldots, \lambda_m \in \mathbb{L}$. Then, the coefficient sequence $a(n)$ satisfies a linear recurrence with C-finite coefficients over $\mathbb{L}$.

Proof. The functional equation is an $\mathbb{L}$-linear combination of functions

$$\begin{aligned} x^j g^{(d)}(\lambda x) &= x^j \sum_{n\geq d} n^{\underline{d}} \lambda^n a(n) x^{n-d} \\ &= \sum_{n\geq j} (n+d-j)^{\underline{d}} \lambda^{n+d-j} a(n+d-j) x^n. \end{aligned}$$

We can compute this for every factor appearing in the functional equation. Comparing the coefficients yields a linear recurrence with C-finite coefficients. □

Example 3.31. Let $g(x) = \sum_{n\geq 0} a(n)x^n$ satisfy the equation

$$xg(2x) + g(x) = 1.$$

Then, $a(0) = 1$ and

$$2^n a(n) + a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

I.e., $a(n)$ is the sequence from Example 3.4.

The equation satisfied by even and odd functions are of the form (3.3). A function $g(x) = \sum_{n\geq 0} a(n)x^n$ satisfies the equation $g(x) = g(-x)$ (i.e., is even) if and only if the coefficient sequence $a(n)$ satisfies $(1-(-1)^n)a(n) = 0$ for all $n \in \mathbb{N}$ (i.e., $a(n) = 0$ for all odd $n \in \mathbb{N}$). By construction, C^2-finite sequences are uniquely defined by finitely many elements $\alpha \in \mathbb{K}$. This means in particular that there are only countably many C^2-finite sequences if the underlying field $\mathbb{K}$ is countable. On the other hand, there are uncountably many even functions. Hence, the coefficient sequences of functions satisfying a functional equation of the form (3.3) are not C^2-finite in general.

D-finite and C-finite sequences a, b are closed under both termwise multiplication $(ab)(n) = a(n)b(n)$ and the Cauchy product $(a \odot b)(n) := \sum_{i=0}^{n} a(i)b(n-i)$. It is not clear whether C^2-finite sequences are closed under the Cauchy product.

Open Question 3.32. Is the Cauchy product of two C^2-finite sequences C^2-finite again?

The sequences $a(n) = 2^{n^2}, b(n) = 3^{n^2}$ are both C^2-finite. it is not known whether $a \odot b$ is C^2-finite. In the restricted setting that one of the two C^2-finite sequences is in fact C-finite, we know that the Cauchy product is C^2-finite again.

Lemma 3.33. *Let a be C^2-finite and b be C-finite over $\mathbb{K}$. Then, the Cauchy product $c := a \odot b$ is again C^2-finite over the splitting field $\mathbb{L}$ of the characteristic polynomial of b.*

Proof. First, let $b(n) = n^d \lambda^n$ for all $n \in \mathbb{N}$ for some $k \in \mathbb{N}, \lambda \in \mathbb{L}$ and $c = a \odot b$. Furthermore, we denote $a_l(n) := \sum_{i=0}^{n} a(i)(n-i)^l \lambda^{n-i}$ for $l = 0, \ldots, d$. Then, for all $j \in \mathbb{N}, n \in \mathbb{N}$ we have

$$\begin{aligned}\sigma^j(c(n)) &= \sum_{i=0}^{n+j} a(i)(n+j-i)^d \lambda^{n+j-i} = \sum_{l=0}^{d} \lambda^j \binom{d}{l} j^{d-l} \sum_{i=0}^{n+j} a(i)(n-i)^l \lambda^{n-i} \\ &= \sum_{l=0}^{d} \lambda^j \binom{d}{l} j^{d-l} a_l(n) + \sum_{l=0}^{d} \lambda^j \binom{d}{l} j^{d-l} \sum_{i=1}^{j} a(n+i)(-i)^l \lambda^{-i}.\end{aligned}$$

Let $\mathcal{A} = c_0 + c_1\sigma + \cdots + c_r\sigma^r$ be an annihilating operator of a. With Lemma 3.25, the ring $R := \mathbb{L}_\sigma[c_0, \ldots, c_r]$ is Noetherian. The computation above shows

$$\langle \sigma^j(c) \mid j \in \mathbb{N} \rangle_{Q(R)} \subseteq \langle a_0, \ldots, a_d \rangle_{Q(R)} + \langle \sigma^j(a) \mid j \in \mathbb{N} \rangle_{Q(R)}.$$

With Lemma 3.21, the module on the right-hand side is finitely generated, hence also the module on the left-hand side is finitely generated. Therefore, with Lemma 3.22, the sequence c is C^2-finite. As C^2-finite sequences are closed under termwise addition (Theorem 3.26) and every C-finite sequence is just a linear combination of such exponential sequences from some term n_0 on, the Cauchy product of a C^2-finite sequence with a C-finite sequence is again C^2-finite. $\square$

Example 3.34. Let $a(n) := 2^{n^2}, b(n) := 3^n, c := a \odot b$. Then, c is again C^2-finite and satisfies

$$4^n c(n) - (\tfrac{1}{3}4^n + \tfrac{1}{8})c(n+1) + \tfrac{1}{24}c(n+2) = 0, \quad \text{for all } n \in \mathbb{N}$$

and $c(0) = 1, c(1) = 5$.

C^2-finite sequences generalize the sequence definition of a D-finite/C-finite sequence. As these sequences have an equivalent characterization in terms of their generating functions (Theorem 2.7), one could also aim to generalize these. I.e., one could, for instance, study power series which satisfy a linear differential equation with coefficients that are D-finite themselves. Such functions are called DD-finite [JPP19, JPP18]. These functions are, however, not closed under termwise multiplication [BJP20]. Hence, there cannot be a one-to-one correspondence of a generalization of the sequence definition of D-finite sequences and of the function definition of D-finite functions. It might, however, still be interesting to investigate the relationship between these two variants.

Open Question 3.35. Study the relationship between C^2-finite (or D^2-finite sequences) and DD-finite functions.

4 Computations with C^2-finite sequences

The closure properties of D-finite and C-finite sequences are all entirely computable. The situation is more complicated for C^2-finite sequences. In this chapter we study how closure properties can, in principle, be computed for C^2-finite sequences. The chapter is mostly based on the papers [JPNP21, JPNP23].

Suppose we consider a C^2-finite sequence c for which no recurrence is known yet (think about c being the sum of two C^2-finite sequences, for instance). We can make an ansatz for a C^2-finite recurrence

$$y_0 c + y_1 \sigma(c) + \cdots + y_{s-1} \sigma^{s-1}(c) + y_s \sigma^s(c) = 0$$

with coefficients $y_0, \ldots, y_s \in \mathcal{R}_C$ and $y_s \in \mathcal{R}_C^\times$ that we have yet to determine. Dividing by the leading coefficient yields a recurrence of the form

$$x_0 c + x_1 \sigma(c) + \cdots + x_{s-1} \sigma^{s-1}(c) + \sigma^s(c) = 0 \tag{4.1}$$

with unknown $x_0, \ldots, x_{s-1} \in Q(\mathcal{R}_C)$.

By Theorem 3.23, since c is C^2-finite, there are finitely many sequences $g_1, \ldots, g_r$ such that

$$\langle \sigma^i(c) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)} = \langle g_1, \ldots, g_r \rangle_{Q(\mathcal{R}_C)}.$$

In particular, there are sequences $w_{i,k} \in Q(\mathcal{R}_C)$ such that $\sigma^i(c) = \sum_{k=1}^r w_{i,k} g_k$ for all $i \in \mathbb{N}$. Using these in equation 4.1 yields

$$\begin{aligned} 0 &= x_0 \sum_{k=1}^r w_{0,k} g_k + x_1 \sum_{k=1}^r w_{1,k} g_k + \cdots + x_{s-1} \sum_{k=1}^r w_{s-1,k} g_k + \sum_{k=1}^r w_{s,k} g_k \\ &= \sum_{k=1}^r g_k \left(\sum_{i=0}^{s-1} w_{i,k} x_i + w_{s,k} \right). \end{aligned}$$

Hence, if we find $x_0, \ldots, x_{s-1} \in Q(\mathcal{R}_C)$ such that

$$\sum_{i=0}^{s-1} w_{i,k} x_i = -w_{s,k}, \quad \text{for all } k = 1, \ldots, r, \tag{4.2}$$

then we have found a solution of (4.1) and therefore a C^2-finite recurrence for c. To write the linear system (4.2) more concisely we denote

$$w_i = \left(w_{i,1}, \ldots, w_{i,r} \right)^\top \in Q(\mathcal{R}_C)^r, \quad \text{for } i = 0, \ldots, s,$$

and $x = \left(x_0, \ldots, x_{s-1} \right)^\top$. Then, (4.2) reads as an inhomogeneous system of linear equations of size $r \times s$ over $Q(\mathcal{R}_C)$

$$\left(w_0, \ldots, w_{s-1} \right) x = -w_s. \tag{4.3}$$

In Section 4.1 we show how this linear system, in particular the vectors w_i, can be computed for different closure properties. The difference compared to the classical D-finite or C-finite case is that the order of the ansatz s is not known a priori. In Section 4.2 we show that for big enough s the linear system (4.3) has a solution, so a C^2-finite recurrence of order s can be found. Further, we show a method (which is limited by the Skolem Problem again) for solving such linear systems.

4.1 Setting up the linear system

Suppose we have a C^2-finite sequence c. By Theorem 3.23 there are finitely many sequences $G = (g_1, \dots, g_r)$ such that

$$\langle \sigma^i(c) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)} = \langle g_1, \dots, g_r \rangle_{Q(\mathcal{R}_C)}.$$

In this section we show how sequences $w_{i,k} \in Q(\mathcal{R}_C)$ with $\sigma^i(c) = \sum_{k=1}^r w_{i,k} g_k$ can be computed.

First, we study the case where a recurrence of order r for c is given and we want to find these coefficients with respect to the generators

$$G_c := \left(c, \sigma(c), \dots, \sigma^{r-1}(c)\right).$$

Based on this, we consider the case where $c = a + b$ and $c = ab$ where a, b are C^2-finite sequences of order r_1, r_2, respectively. The generating sets in these cases are given by

$$G_a \oplus G_b = \left(a, \sigma(a), \dots, \sigma^{r_1-1}(a), b, \sigma(b), \dots, \sigma^{r_2-1}(b)\right)$$

in the case of addition and

$$G_a \otimes G_b = \left(ab, a\sigma(b), \dots, a\sigma^{r_2-1}(b), \dots, \sigma^{r_1-1}(a)b, \sigma^{r_1-1}(a)\sigma(b), \dots, \sigma^{r_1-1}(a)\sigma^{r_2-1}(b)\right)$$

in the case of multiplication. Next, we consider the case $a(kn)$ where a is a C^2-finite sequence of order r and $k \in \mathbb{N}$. The generating set is given by $(a(kn), \dots, a(kn + r - 1))$. Finally, we consider the case $c(jn^2 + kn + \ell)$ for a C-finite sequence c where $j, k, \ell \in \mathbb{N}$. Here, the generating set is given by $(c(jn^2), \dots, c(jn^2 + r - 1))$.

4.1.1 Ring computations

Let a be C^2-finite of order r with annihilating operator $c_0 + \cdots + c_{r-1}\sigma^{r-1} + \sigma^r \in Q(\mathcal{R}_C)[\sigma]$. We write the components of a vector $u_i \in Q(\mathcal{R}_C)^r$ as $u_{i,k}$ for $k = 0, \dots, r-1$. The componentwise shift of a vector u_i is simply denoted by $\sigma(u_i)$, i.e., $(\sigma(u_i))_k = \sigma(u_{i,k})$. The i-th unit vector is denoted by $e_i^{(r)} \in Q(\mathcal{R}_C)^r$ for $i = 0, \dots, r-1$. Note that, e.g., $e_0^{(r)} = (1, 0, \dots, 0)^\top$.

Lemma 4.1. *Initialize $u_i := e_i^{(r)} \in Q(\mathcal{R}_C)^r$ with the unit vectors for $i = 0, \dots, r-1$ and define*

$$u_i := -\sum_{k=0}^{r-1} \sigma^{i-r}(c_k) u_{i+k-r} \tag{4.4}$$

inductively for $i \geq r$. These $u_i \in Q(\mathcal{R}_C)$ satisfy

$$\sigma^i(a) = \sum_{k=0}^{r-1} u_{i,k} \sigma^k(a) \tag{4.5}$$

for all $i \in \mathbb{N}$ or $\sigma^i(a) = G_a u_i$ written concisely.

Proof. We show equation (4.5) by induction on i. It clearly holds for $i = 0, \dots, r-1$ by the definition of the u_i. Shifting the defining recurrence of a yields

$$\sigma^i(a) = -\sum_{j=0}^{r-1} \sigma^{i-r}(c_j) \sigma^{i+j-r}(a),$$

for $i \geq r$. Let us assume that equation (4.5) holds for $a, \dots, \sigma^{i-1}(a)$. Then,

$$\sum_{k=0}^{r-1} u_{i,k} \sigma^k(a) = -\sum_{k=0}^{r-1} \sum_{j=0}^{r-1} \sigma^{i-r}(c_j) u_{i+j-r,k} \sigma^k(a) = -\sum_{j=0}^{r-1} \sigma^{i-r}(c_j) \sigma^{i+j-r}(a) = \sigma^i(a).$$

□

A different way to compute the vectors u_j is to use the companion matrix of a sequence. The companion matrix M_a of the sequence a with annihilator $c_0 + c_1\sigma + \cdots + c_{r-1}\sigma^{r-1} + \sigma^r \in Q(\mathcal{R}_C)[\sigma]$ is defined as

$$M_a := \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{r-1} \end{pmatrix} \in Q(\mathcal{R}_C)^{r \times r}.$$

Lemma 4.2. *Let M_a be the companion matrix of a. Let*

$$u_0 := e_0^{(r)} = (1, 0, \dots, 0)^\top$$

and define

$$u_i := M_a \sigma(u_{i-1})$$

inductively for $i \geq 1$.

1. *These u_i are identical to the vectors from Lemma 4.1.*
2. *The u_i satisfy equation* (4.5).

Proof. (1): Clearly $u_i = e_i^{(r)}$ for $i = 0, \dots, r-1$ by the definition of the companion matrix. For $i \geq r$ we show that equation (4.4) from Lemma 4.1 is satisfied using induction on i. For $i = r$ we have

$$u_r = \left(-c_0, \dots, -c_{r-1}\right)^\top$$

by the definition of the companion matrix. Therefore,

$$-\sum_{k=0}^{r-1} c_k u_k = -\sum_{k=0}^{r-1} c_k e_k^{(r)} = u_r.$$

Now, we assume that equation (4.4) from Lemma 4.1 holds for $i-1$, i.e.,

$$u_{i-1} = -\sum_{k=0}^{r-1} \sigma^{i-1-r}(c_k) u_{i-1+k-r}. \tag{4.6}$$

Using equation (4.6) and the definition of the u_i we have

$$u_i = M_a \sigma(u_{i-1}) = -M_a \sum_{k=0}^{r-1} \sigma^{i-r}(c_k) \sigma(u_{i-1+k-r}) = -\sum_{k=0}^{r-1} \sigma^{i-r}(c_k) u_{i+k-r}.$$

(2): Follows directly from part (1) and Lemma 4.1. $\square$

Consider two C^2-finite sequences a, b. To compute the vectors w_i in the linear system (4.3) for $a + b$ we can concatenate the vectors we get from Lemma 4.1. Alternatively, we can use a similar approach as in [JPP18]. Suppose a, b have orders r_1, r_2, respectively. Define

$$M := M_a \oplus M_b = \begin{pmatrix} M_a & 0 \\ 0 & M_b \end{pmatrix} \in Q(\mathcal{R}_C)^{r_1+r_2 \times r_1+r_2}$$

to be the direct sum of the companion matrices of a and b and

$$w_0 := e_0^{(r_1)} \oplus e_0^{(r_2)} = \begin{pmatrix} e_0^{(r_1)} \\ e_0^{(r_2)} \end{pmatrix} \in Q(\mathcal{R}_C)^{r_1+r_2}.$$

Let us define $w_i := M\sigma(w_{i-1})$ iteratively. If we denote the first r_1 components of w_i by u_i and the last r_2 components by v_i, these u_i, v_i clearly satisfy (4.5) by the construction of M and Lemma 4.2 for G_a, G_b, respectively. Therefore,

$$(G_a \oplus G_b)w_i = (G_a \oplus G_b)(u_i \oplus v_i) = G_a u_i + G_b v_i = \sigma^i(a) + \sigma^i(b) = \sigma^i(a + b).$$

Analogously, for the multiplication we can define M as the Kronecker product

$$M := M_a \otimes M_b \in Q(\mathcal{R}_C)^{r_1 r_2 \times r_1 r_2}$$

of the two companion matrices. Again, defining $w_0 := e_0^{(r_1)} \otimes e_0^{(r_2)}$ and $w_i := M\sigma(w_{i-1})$, we have $w_i = u_i \otimes v_i$ where u_i, v_i satisfy $G_a u_i = \sigma^i(a)$ and $G_b v_i = \sigma^i(b)$. Therefore,

$$(G_a \otimes G_b)w_i = (G_a \otimes G_b)(u_i \otimes v_i) = (G_a u_i)(G_b v_i) = \sigma^i(a)\sigma^i(b) = \sigma^i(ab).$$

Algorithm 1 summarizes the arguments from the introduction of Chapter 4 and this section. The algorithm computes a recurrence for the addition or multiplication of two C^2-finite sequences a, b of orders r_1, r_2 provided that we can solve linear systems of equations over $Q(\mathcal{R}_C)$. The termination of Algorithm 1 follows from Lemma 4.9 which we show in the next section.

Input : C^2-finite sequences a, b of order r_1, r_2, respectively
output: C^2-finite recurrence satisfied by $a + b$ (or ab, respectively)
$M \leftarrow M_a \oplus M_b$ (or $M_a \otimes M_b$ for the multiplication)
$A \leftarrow$ empty matrix
$w \leftarrow e_0^{(r_1)} \oplus e_0^{(r_2)}$ (or $e_0^{(r_1)} \otimes e_0^{(r_2)}$ for the multiplication)
for $s = 0, 1, 2, \ldots$ **do**
  if *solution* $x \in Q(\mathcal{R}_C)^s$ *of* $Ax = -w$ *exists* **then**
    return $\sum_{i=0}^{s-1} x_i \sigma^i + \sigma^s$
  else
    $A \leftarrow (A \mid w)$
    $w \leftarrow M\sigma(w)$
  end
end

Algorithm 1: Computing C^2-finite ring operations

Example 4.3. Let $b(n) := \frac{1}{f(n+1)}$ where f denotes the Fibonacci numbers (cf. Example 2.13 and Example 3.6). The sequence b is C^2-finite satisfying

$$f(n+1)b(n) - f(n+2)b(n+1) = 0.$$

We want to compute a recurrence for $c(n) := f(n)b(n)$. The companion matrices of f and b are given by

$$M_f = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, M_b = \begin{pmatrix} \frac{f(n+1)}{f(n+2)} \end{pmatrix}.$$

Therefore,

$$M = M_f \otimes M_b = \begin{pmatrix} 0 & \frac{f(n+1)}{f(n+2)} \\ \frac{f(n+1)}{f(n+2)} & \frac{f(n+1)}{f(n+2)} \end{pmatrix}, w = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Hence, the linear system corresponding to an ansatz of order 2 is

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{f(n+1)}{f(n+2)} \end{pmatrix} \begin{pmatrix} x_0(n) \\ x_1(n) \end{pmatrix} = \begin{pmatrix} -\frac{f(n+1)}{f(n+3)} \\ -\frac{f(n+1)}{f(n+3)} \end{pmatrix}.$$

The system clearly has the solution $x_0(n) = -\frac{f(n+1)}{f(n+3)}, x_1(n) = -\frac{f(n+2)}{f(n+3)}$ which gives rise to the recurrence

$$-f(n+1)c(n) - f(n+2)c(n+1) + f(n+3)c(n+2) = 0.$$

Analogously we can find that $d(n) := \sum_{k=1}^{n} \frac{(-1)^k}{f(k)f(k+1)}$ satisfies the recurrence

$$-f(n+2)d(n) - f(n+3)d(n+1) + f(n+4)d(n+2) = 0.$$

The sequence $c(n) + d(n)$ is then again C^2-finite and we can find a recurrence of order 3. As the first 3 initial values are 0, we have $c(n) + d(n) = 0$ for all $n \in \mathbb{N}$ as shown in [Kau05, Example 4.7].

4.1.2 Subsequences

Let a be C^2-finite of order r with annihilating operator $c_0 + \cdots + c_{r-1}\sigma^{r-1} + \sigma^r \in Q(\mathcal{R}_C)[\sigma]$. Let $b(n) = a(kn)$ for some $k \in \mathbb{N}$. Then,

$$\sigma^i(b(n)) = a(kn + ki).$$

Let

$$G := \Big(a(kn), \ldots, a(kn + r - 1)\Big).$$

The next lemma shows that

$$\langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)} = \langle G \rangle_{Q(\mathcal{R}_C)}.$$

In particular, we find vectors $w_i \in Q(\mathcal{R}_C)^r$ such that $Gw_i = \sigma^i(b)$.

Lemma 4.4. *Define* $w_0 := e_0^{(r)}$ *and*

$$w_i(n) := M_a(kn) \cdots M_a(kn + k - 1) w_{i-1}(n+1), \quad \text{for all } n \in \mathbb{N},$$

iteratively. Then, $Gw_i = \sigma^i(b)$ *for all* $i \in \mathbb{N}$.

Proof. We use induction on i. Clearly, $Gw_0 = Ge_0^{(r)} = b$. Suppose that $Gw_{i-1} = \sigma^{i-1}(b)$. By the definition of the companion matrix M_a, we have

$$\begin{aligned} G(n)w_i(n) &= G(n)M_a(kn)\cdots M_a(kn+k-1)w_{i-1}(n+1) \\ &= G(n+1)w_{i-1}(n+1) = b(n+i) \end{aligned}$$

where we use the shifted induction hypothesis in the last step. Hence, $Gw_i = \sigma^i(b)$. $\square$

Lemma 4.4 shows in particular that

$$\langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)}$$

is finitely generated. Hence, the sequence b is C^2-finite by Theorem 3.23.

Theorem 4.5. *Let a be C^2-finite and $k \in \mathbb{N}$. Then, the sequence $b(n) = a(kn)$ is C^2-finite.*

For computing a recurrence for $a(kn)$ we can adjust Algorithm 1. Choosing

$$M = M_a(kn)\cdots M_a(kn+k-1) \in Q(\mathcal{R}_C)^{r\times r}$$

and $w = e_0^{(r)} \in Q(\mathcal{R}_C)^r$ gives an explicit algorithm for computing a recurrence for the subsequence of a C^2-finite sequence.

4.1.3 Sparse subsequences of C-finite sequences

Let c be C-finite of order r with annihilating operator $c_0 + \cdots + c_{r-1}\sigma^{r-1} + \sigma^r \in \mathbb{K}[\sigma]$. Let $a(n) = c(jn^2 + kn + \ell)$ for some $j, k, \ell \in \mathbb{N}$ and

$$G \coloneqq \Big(c(jn^2), \ldots, c(jn^2 + r - 1)\Big).$$

The next lemma shows that

$$\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)} = \langle G \rangle_{Q(\mathcal{R}_C)}.$$

In particular, we find vectors $w_i \in Q(\mathcal{R}_C)^r$ such that $Gw_i = \sigma^i(a)$. The proof is based on ideas from [KM14, Theorem 1]. Since c is C-finite, $M_c \in \mathbb{K}^{r\times r}$. We use the fact that

$(M_c^{p(n)})_{n\in\mathbb{N}}$ for a linear polynomial $p \in \mathbb{N}[n]$ can also be viewed as a matrix of C-finite sequences [KM14, Lemma 11].

Lemma 4.6. *Let*

$$w_i(n) = M_c^{2jni+ji^2+kn+ki+\ell-r+1} e_{r-1}^{(r)}.$$

These $w_i \in Q(\mathcal{R}_C)^r$ satisfy $Gw_i = \sigma^i(a)$ for all $i \in \mathbb{N}$.

Proof. Let

$$G_c(n) = \Big(c(n), \dots, c(n+r-1)\Big).$$

Then, $G_c(jn^2) = G(n)$. By the definition of the companion matrix we have

$$G_c(n+1) = G_c(n)M_c \tag{4.7}$$

for all $n \in \mathbb{N}$. Using $n \to jn^2$ we have $G_c(jn^2+1) = G_c(jn^2)M_c$. Repeated application of equation (4.7) yields

$$G_c(j(n+i)^2 + k(n+i) + \ell - r + 1) = G_c(jn^2)M_c^{2jni+ji^2+kn+ki+\ell-r+1}.$$

Multiplying by $e_{r-1}^{(r)}$ and using the definition of the $G_c(n)$ yields $Gw_i = \sigma^i(a)$. $\square$

Lemma 4.6 shows in particular that

$$\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_C)}$$

is finitely generated. Hence, the sequence a is C^2-finite by Theorem 3.23.

Theorem 4.7. *Let c be C-finite and $j, k, \ell \in \mathbb{N}$. Then, the sequence $a(n) = c(jn^2 + kn + \ell)$ is C^2-finite.*

An alternative method for proving a variant of Theorem 4.7 using the closed form of a C-finite sequence is given in [JPNP21, Corollary 3.6] (and a more general version in Theorem 7.11 in this thesis). This alternative proof, however, only guarantees a C^2-finite recurrence for a sequence $c(jn^2 + kn + \ell) \in \mathbb{K}^{\mathbb{N}}$ over an extension field $\mathbb{L} \supseteq \mathbb{K}$. Suppose a sequence $a \in \mathbb{K}^{\mathbb{N}}$ is D-finite over some extension field $\mathbb{L} \supseteq \mathbb{K}$, then a is actually D-finite over $\mathbb{K}$ [Ger05, Lemma 1.3.2]. It would be interesting to see whether the same holds for C^2-finite sequences.

Open Question 4.8. Suppose $a \in \mathbb{K}^{\mathbb{N}}$ is C^2-finite with C-finite coefficients over $\mathbb{L} \supseteq \mathbb{K}$. Can we compute a C^2-finite recurrence with C-finite coefficients over $\mathbb{K}$ for a?

For computing a recurrence for $a(n)$ we can, again, adjust Algorithm 1. By Lemma 4.6 we have

$$w_0(n) = M_c^{kn+\ell-r+1} e_{r-1}^{(r)} \text{ and } w_i(n) = M_c^{j(2n+1)} w_{i-1}(n+1).$$

Hence, choosing $M = M_c^{j(2n+1)}$ and $w = w_0$ in Algorithm 1 gives an explicit algorithm for computing a recurrence for $a(n)$. By [KM14, Lemma 11], $w \in Q(\mathcal{R}_C)^r, M \in Q(\mathcal{R}_C)^{r \times r}$. These recurrences can be computed by using either the Cayley-Hamilton theorem (as suggested by the proof of the lemma) or using guessing.

In the very same way we could compute recurrences for $a(n) = c(j\binom{n}{2} + kn + \ell)$ for $j, k, \ell \in \mathbb{N}$.

4.2 Solving the linear system

So far, we have shown that we can set up a linear inhomogeneous system of equations over the ring $Q(\mathcal{R}_C)$ such that a solution to this system gives rise to a recurrence for the closure properties which we want to compute. First, we show that this algorithm terminates, i.e., if the linear system is big enough, then the system has a solution.

Lemma 4.9. *The order of the ansatz s in* (4.1) *can be chosen big enough such that the corresponding linear system* (4.3) *for the addition and multiplication of two C^2-finite sequences has a solution.*

Proof. Let $c_0 + \cdots + c_{r_1}\sigma^{r_1} \in Q(\mathcal{R}_C)[\sigma], d_0 + \cdots + d_{r_2}\sigma^{r_2} \in Q(\mathcal{R}_C)[\sigma]$ be annihilators of the C^2-finite sequences a, b. Writing $A_s = (w_0, \ldots, w_s)$, equation (4.3) reads as $A_s x = -w_s$. Let

$$R := \mathbb{K}_\sigma[c_0, \ldots, c_{r_1}, d_0, \ldots, d_{r_2}].$$

In Section 4.1.1 we have shown that all sequences in the linear system $A_s = -w_s$ are in the ring $Q(R)$. By Lemma 3.25 this ring is Noetherian. Hence, writing $\operatorname{Im} A$ for the image of a matrix A, the increasing chain of modules

$$\operatorname{Im} A_0 \subseteq \operatorname{Im} A_1 \subseteq \operatorname{Im} A_2 \subseteq \cdots$$

has to stabilize. In particular, there is an s such that $\operatorname{Im} A_s = \operatorname{Im} A_{s+1}$. Therefore, $w_s \in \operatorname{Im} A_{s+1} = \operatorname{Im} A_s$, so $A_s x = -w_s$ has a solution $x \in Q(R)^s$. □

The proof of Lemma 4.9 is not constructive as the properties of the Noetherian ring only gives us the existence of a number s. The same proof, however, can also be used to show that the ansatz for computing a subsequence of a C^2-finite sequence or a sparse subsequence of a C-finite sequence can be chosen big enough such that the corresponding linear system has a solution.

In the case of D-finite and C-finite sequences, the s in Lemma 4.9 can be chosen as at most $\operatorname{ord}(a) + \operatorname{ord}(b)$ in the case of addition $a + b$ and $\operatorname{ord}(a) \operatorname{ord}(b)$ in the case of multiplication ab. These results follow directly from a dimension argument of the corresponding vector spaces. For C^2-finite sequences these order bounds cannot be used anymore.

Example 4.10. Let $c \in \mathcal{R}_C$ be the cyclic sequence of order m defined by

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

and let a, b be C^2-finite sequence defined by

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

Suppose we make an ansatz of order $s < m+1$ for $a + b$. With the definition of c, the corresponding linear system at $n = m - s + 1$ is of the form

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_0(n) \\ \vdots \\ x_{s-1}(n) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Hence, the linear system has no solution. For $s = m+1$ we get a solution for every n and therefore a C^2-finite recurrence for $a + b$ of order $m + 1$.

Example 4.10 indicates that any order bounds for C^2-finite sequences should depend not only on the orders of the C^2-finite recurrences but also on the orders of the C-finite coefficients. In Chapter 5 we derive such order bounds.

Next, we study how the linear system (4.3) can be solved. The main obstacle here is, again, the Skolem Problem. If the zeros of all minors of the linear system can be computed, then we have an algorithm for computing a solution of the linear system. This algorithm is based on methods from [KM14].

Theorem 4.11. *Let $A \in Q(\mathcal{R}_C)^{r\times s}$ and $w \in Q(\mathcal{R}_C)^r$. Suppose the linear system $Ax = w$ has a solution $x \in Q(\mathcal{R}_C)^s$. Provided that the Skolem Problem is decidable, we can compute a solution $x \in Q(\mathcal{R}_C)^s$.*

Proof. All minors of A are sequences in $Q(\mathcal{R}_C)$. Consider the set of all these. By the Skolem-Mahler-Lech theorem, the zeros of these minors are cyclic. Let $p \in \mathbb{N}$ be such that all minors have cycle-length p from the term $n_0 \in \mathbb{N}$ on. These numbers p, n_0 can be computed if the Skolem Problem is decidable.

We write $A = (w_0, \dots, w_{s-1})$ for $w_0, \dots, w_{s-1} \in Q(\mathcal{R}_C)^r$. Now, for every $m \in \{n_0, \dots, n_0 + p - 1\}$ we can compute a subset $j_m \subseteq \{0, \dots, s-1\}$ such that the vectors

$$\{w_j(m) \mid j \in j_m\} \subseteq \mathbb{K}^r$$

are maximally linearly independent, i.e., they are linearly independent and generate the same subspace as $\{w_0(m), \dots, w_{s-1}(m)\}$. By the choice of n_0 and p this is also true for all $n = m + pk$ for $k \in \mathbb{N}$, i.e., the vectors $\{w_j(m + pk) \mid j \in j_m\} \subseteq \mathbb{K}^r$ are maximally linearly independent for every $k \in \mathbb{N}$. Let us denote by $A_m \in Q(\mathcal{R}_C)^{r\times|j_m|}$ the submatrix of A where we keep the columns w_j with $j \in j_m$.

For every m, we can solve the system

$$A_m(m + pk)x_m(k) = w(m + pk), \quad \text{for all } k \in \mathbb{N}, \tag{4.8}$$

using the Moore-Penrose-Inverse [BIG03]: By the choice of m, p, n_0, the matrix $A_m(m+pk)$ has linear independent columns for every $k \in \mathbb{N}$. Therefore, the Gramian matrix $G(k) = A_m(m+pk)^\top A_m(m+pk)$ is regular for every k and $(\det(G(k)))_{k\in\mathbb{N}} \in \mathcal{R}_C^\times$. Now, let

$$x_m(k) = \frac{1}{\det(G(k))} \operatorname{cof}(G(k))\, A_m(m+pk)^\top w(m+pk)$$

where $\operatorname{cof}(\cdot)$ denotes the transposed cofactor matrix. Then, since equation (4.8) has a termwise solution, $(x_m(k))_{k\in\mathbb{N}} \in Q(\mathcal{R}_C)^{|j_m|}$ satisfies equation (4.8) by the theory of Moore-Penrose-Inverses. Let $x'_m \in Q(\mathcal{R}_C)^s$ be the vector where we add $0 \in Q(\mathcal{R}_C)$ at the indices $j \in \{0, \dots, s-1\} \setminus \{j_m\}$.

Now, the solution x for the entire system can be computed as the interlacing of $x'_{n_0}, \dots, x'_{n_0+p-1}$ from n_0 on and the first n_0 values can be computed explicitly. Then, $x \in Q(\mathcal{R}_C)^s$ as $Q(\mathcal{R}_C)$ is closed under interlacing and specifying finitely many initial values. □

The proof of Theorem 4.11 in fact also shows that whenever the linear system $A(n)x(n) = w(n)$ with $A \in Q(\mathcal{R}_C)^{r\times s}$ and $w \in Q(\mathcal{R}_C)^r$ has a solution $x(n) \in \mathbb{K}^s$ for every $n \in \mathbb{N}$, then a solution $x \in Q(\mathcal{R}_C)^s$ exists. Of course, also vice versa, if a sequence solution exists, then we have a termwise solution.

In Algorithm 2 we give the algorithm suggested by Theorem 4.11 in pseudocode. For a sequence $c \in Q(\mathcal{R}_C)$ we denote the start and the length of the zero-cycle of the sequence c by `period_start`(c) and `period_length`(c), respectively, i.e., the numbers $n_0, p \in \mathbb{N}$ such that

$$(c(pk+n_0), \dots, c(pk+p-1+n_0))$$

has the same zero-pattern for every $k \in \mathbb{N}$.

Even though Theorem 4.11 heavily relies on the Skolem Problem in theory, in practice the algorithm can, in many cases, be used for solving linear systems over the C-finite sequence ring. Using the techniques from Chapter 8, the zeros of C-finite sequences can often be found. The problem of the algorithm is rather that it is computationally too expensive if the Moore-Penrose inverses are computed explicitly. However, also if a classical row reduction algorithm such as the fraction-free Bareiss algorithm (cf. [Bar68]) is used, the

Input : $A \in Q(\mathcal{R}_C)^{r \times s}, w \in Q(\mathcal{R}_C)^r$
output: $x \in Q(\mathcal{R}_C)^s$ with $Ax = w$ if it exists and `false` otherwise
$\Phi \leftarrow$ minors of A
$p \leftarrow \operatorname{lcm}(\texttt{period\_length}(\phi) \mid \phi \in \Phi)$
$n_0 \leftarrow \max(\texttt{period\_start}(\phi) \mid \phi \in \Phi)$
if $A(m)x(m) = w(m)$ *is not solvable for an* $m \in \{0, \dots, n_0 - 1\}$ **then**
  return `false`
end
for $m = n_0, \dots, n_0 + p - 1$ **do**
  $j_m \leftarrow$ indices of columns of $A(m)$ which are maximally linearly independent
  $A_m \leftarrow$ matrix built by columns j_m of A
  $A'_m \leftarrow (A_m(m + pk))_{k \in \mathbb{N}}$
  $w' \leftarrow (w(m + pk))_{k \in \mathbb{N}}$
  $G \leftarrow (A'_m)^\top A'_m$
  $x_m \leftarrow \frac{1}{\det(G)} \operatorname{cof}(G)(A'_m)^\top w' \in Q(\mathcal{R}_C)^{|j_m|}$
  $x'_m \leftarrow$ insert 0 in x_m at indices $j \in \{0, \dots, s-1\} \setminus \{j_m\}$
end
$x \leftarrow$ interlacing of sequences $x'_{n_0}, \dots, x'_{n_0+p-1}$ with prepended terms $x(0), \dots, x(n_0 - 1)$
if $Ax = w$ **then**
  return x
else
  return `false`
end

Algorithm 2: Solving linear systems over $Q(\mathcal{R}_C)$

computations are very expensive due to a blowup of the order of the C-finite sequences and their coefficients in the linear system. Furthermore, the solutions $x \in Q(\mathcal{R}_C)^s$ that are computed are often too big. Cancelling common factors is, however, again a difficult problem [KZ08, KZ18].

5 Order bounds for C^2-finite closure properties

In the previous chapters a C^2-finite sequence a was annihilated by an operator $\mathcal{A}$ with $\operatorname{lc}(\mathcal{A})(n) \neq 0$ for all $n \in \mathbb{N}$. As discussed on page 17 we can equivalently allow annihilating operators $\mathcal{A}$ with $\operatorname{lc}(\mathcal{A})(n) \neq 0$ for almost all $n \in \mathbb{N}$. The set of sequences annihilated by these operators are the same, but the order of a sequence might be different. For deriving order bounds for C^2-finite sequences, we need the latter definition, i.e., we need to allow finitely many zeros in the leading coefficient of the recurrence.

Example 5.1. Let $a(n) := 2^{\binom{n+1}{2}}$ (A006125 in the OEIS) 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 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 $\operatorname{ord}(a) + \operatorname{ord}(b) = 2$ 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.

Theorem 2.6 shows that the closure properties of D-finite sequences satisfy very simple order bounds. The same bounds hold for C-finite sequences. In Example 4.10 we have seen that these order bounds do not hold in the C^2-finite case any longer. However, in this chapter we derive similar bounds which also depend on the C-finite coefficients which appear in the recurrences of the C^2-finite sequences. The content of this chapter is based on [KNP23].

5.1 The exponent lattice

For proving the order bounds for C^2-finite sequences, we heavily rely on the fact that a C-finite sequence c can be written as interlacing of nondegenerate sequences (cf. [EPSW15, Theorem 1.2])

$$c(dn), \dots, c(dn+d-1).$$

More generally, if we have a finitely generated difference algebra of C-finite sequences, we determine a number $d \in \mathbb{N}$ (which we call the *torsion number*) such that every sequence in the algebra can be written as the interlacing of d nondegenerate subsequences.

Let $c_0, \dots, c_r \in \mathcal{R}_C$ with eigenvalues $\lambda_1, \dots, \lambda_m$ and let $R_d := \mathbb{K}_\sigma[c_0(dn), \dots, c_r(dn)]$ be the smallest difference algebra which contains the sequences $c_0(dn), \dots, c_r(dn)$. 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 $e_1, \dots, e_m \in \mathbb{N}$. We want to find a d such that every sequence $c \in R_d$ is nondegenerate. 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 $k, e_1, \dots, e_m, f_1, \dots, f_m \in \mathbb{N}$. In order to write this more concisely we define the multiplicative group $G := \langle \lambda_1, \dots, \lambda_m \rangle \leq (\mathbb{C}^\times, \cdot)$. 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 next lemma shows that this number d also has a purely group-theoretical and a purely lattice-theoretical description. A *lattice* is a $\mathbb{Z}$-submodule of $\mathbb{Z}^m$. 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.

Lemma 5.2. *Let $G := \langle \lambda_1, \dots, \lambda_m \rangle \leq (\mathbb{C}^\times, \cdot)$. 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*

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

3. *Let*

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

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

$$\forall k \in \mathbb{N}_{\geq 1} \forall v \in \mathbb{Z}^m \colon kdv \in L \implies dv \in L. \tag{5.1}$$

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 $v := (e_1, \dots, e_m) \in \mathbb{Z}^m$ yields $kdv \in L$. By assumption, $dv \in L$, i.e., $\lambda^d = 1$. $\square$

Considering condition 2 of Lemma 5.2, we can see that the smallest d which satisfies the condition is the exponent of the torsion group.

Definition 5.3. 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 (5.1). A useful tool in studying lattices is the Smith normal form of a matrix. 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 $d_r = \frac{e_r}{e_{r-1}}$ [New72, Mid19].

Let $v_1, \dots, v_\ell \in \mathbb{Z}^m$ be a basis of the lattice $L = \langle v_1, \dots, v_\ell \rangle \subseteq \mathbb{Z}^m$ and let $V \coloneqq (v_1, \dots, v_\ell) \in \mathbb{Z}^{m \times \ell}$ with Smith normal form $PVQ = D$ with nonzero invariant factors $d_1, \dots, d_r$. Since Q is unimodular we have

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

Let $p_1, \dots, p_m \in \mathbb{Z}^m$ be the columns of P^{-1}. Since P is unimodular, these columns form a basis of $\mathbb{Z}^m$ and $d_1 p_1, \dots, d_r p_r$ form a basis of L [Tao22].

A lattice L is called *pure* if for all $k \in \mathbb{Z}, v \in \mathbb{Z}^m$ the condition $kv \in L$ implies $v \in L$. Equivalently, L is pure if and only if L is a direct summand of $\mathbb{Z}^m$ [CR66, Chapter III.16A]. The *pure closure* $\overline{L}$ of L is the smallest pure lattice which contains L, i.e., the intersection of all pure lattices that contain L. We have (cf. [CFQ15])

$$\overline{L} = \{v \in \mathbb{Z}^m \mid \exists k \in \mathbb{Z} \setminus \{0\} \colon kv \in L\}.$$

Using the terminology of pure modules, (5.1) is equivalent to $L_d \coloneqq \{v \in \mathbb{Z}^m \mid dv \in L\} \supseteq L$ being a pure lattice. We show that for suitable d the lattice L_d is precisely the pure closure of L. The property of being pure is closely related to the invariant factors of the matrix built by a basis of a lattice. The following lemma is already given, without a proof, in [CMDS84] on page 80. For the sake of completeness we include a proof here.

Lemma 5.4. *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 $V := (v_1, \dots, v_\ell) \in \mathbb{Z}^{m \times \ell}$. Then, L is pure if and only if all invariant factors of V are* 1.

Proof. Let $PVQ = D$ be the Smith normal form of V with invariant factors $d_1, \dots, d_\ell$. Furthermore, let $p_1, \dots, p_m \in \mathbb{Z}^m$ denote the columns of the unimodular matrix P^{-1}.

$\Longrightarrow$: The set $\{d_1 p_1, \dots, d_\ell p_\ell\}$ forms a basis of L. As L is pure, $p_1, \dots, p_\ell$ also form a basis of L. Hence, there is a unimodular change-of-basis matrix U with $U d_\ell p_\ell = p_\ell$. In particular (cf. Corollary 158 in [Mid19]),

$$\gcd(U d_\ell p_\ell) = d_\ell \gcd(p_\ell) = \gcd(p_\ell).$$

Therefore, $d_\ell = 1$ and by the divisibility property of the invariant factors

$$d_1 = \cdots = d_\ell = 1.$$

$\Longleftarrow$: As $\{p_1, \dots, p_\ell\}$ form a basis of L and $\{p_1, \dots, p_m\}$ form a basis of $\mathbb{Z}^m$, L is a direct summand of $\mathbb{Z}^m$ and therefore pure. □

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, there are algorithms which compute a basis $v_1, \dots, v_\ell \in \mathbb{Z}^m$ for the lattice $L := L(\lambda_1, \dots, \lambda_m)$ [Ge93, Kau05, Fac14, ZX19, Zhe20, Zhe21, KNP23]. Then, the invariant factor of the matrix built by the basis is precisely the torsion number of the lattice:

Theorem 5.5. *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 $V := (v_1, \dots, v_\ell) \in \mathbb{Z}^{m \times \ell}$ with invariant factor d. Then, $L_d = \{v \in \mathbb{Z}^m \mid dv \in L\}$ is the pure closure of L. In particular, d is the torsion number of L.*

Proof. The lattices L and L_d have the same rank ℓ. Let

$$d_1, \dots, d_{\ell-1}, d_\ell = d$$

be the invariant factors of L and $\overline{d}_1, \dots, \overline{d}_\ell$ the invariant factors of L_d. The lattice L_d has a basis of the form

$$\overline{d}_1 p_1, \dots, \overline{d}_\ell p_\ell.$$

Let $\overline{V} := (\overline{d}_1 p_1, \dots, \overline{d}_\ell p_\ell)$. Then, $\overline{V} S = V$ for some matrix $S \in \mathbb{Z}^{\ell \times \ell}$ as $L \subseteq L_d$. Therefore, $\overline{d}_\ell \mid d_\ell = d$ [New72, Lemma II.2]. Hence, $dp_\ell \in L_d$, so $p_\ell \in L_d$. By the same argument as in Lemma 5.4, $d = \overline{d}_\ell = 1$, so L_d is pure.

As L_d is pure, the pure closure $\overline{L}$ of L is contained in L_d. Let $v \in L_d$. Then, $dv \in L$, so $v \in \overline{L}$. Therefore, $L_d = \overline{L}$. □

Example 5.6. 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 nondegenerate, i.e., sequences which contain only finitely many zeros (cf. Theorem 2.17). A nondegenerate sequence might 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. Note, that an element of $Q(R)$ can be interpreted only as a sequence in $\mathbb{K}^{\mathbb{N}}$ from some term on (cf. the discussion in Section 8.2 in [PWZ96] or [Sch20]). 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 see in Section 5.2. We summarize the discussions of the section in the following theorem:

Theorem 5.7. *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*

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

is a field. The nonzero elements of the field $Q(R)$ can be considered as sequences which are nonzero 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 [Sch20].

5.2 Order bounds

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

In Chapter 4 we have seen how computations of closure properties of C^2-finite sequences can be reduced to solving linear systems of equations. A C^2-finite recurrence (here we use a "homogeneous ansatz" compared to the "inhomogeneous ansatz" in (4.1) which yields an inhomogeneous linear system)

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

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

$$\left(w_0, w_1, \dots, w_s\right) \in Q(R)^{r \times (s+1)}. \tag{5.2}$$

The w_i can be computed iteratively using $w_{i+1} = M\sigma(w_i)$ for a suitable matrix $M \in Q(R)^{r \times r}$.

- In the case a recurrence for $a+b$ is computed, we use $w_0 = e_0^{(r_1)} \oplus e_0^{(r_2)}$ and $M = M_a \oplus M_b$ where $r_1 = \operatorname{ord}(a), r_2 = \operatorname{ord}(b)$.
- In the case a recurrence for ab is computed, we use $w_0 = e_0^{(r_1)} \otimes e_0^{(r_2)}$ and $M = M_a \otimes M_b$.
- In the case a recurrence for $a(kn)$ with $k \in \mathbb{N}$ is computed, we use $w_0 = e_0^{(r)}$ and $M = M_a(kn) \cdots M_a(kn+k-1)$ where $r = \operatorname{ord}(a)$.
- 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) of order r is computed, we use

$$w_0 = M_c^{kn+\ell-r+1} e_{r-1}^{(r)} \text{ and } M = M_c^{j(2n+1)}. \tag{5.3}$$

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

In the next sections we present order bounds for the closure properties that we discussed in Chapter 4. In particular, we prove the following theorem (cf. Theorem 2.6 for the D-finite order bounds):

Theorem 5.8. *Let $a(n), b(n), a_0(n), \ldots, a_{m-1}(n)$ be C^2-finite sequences. Let d_a be the torsion number of the eigenvalues appearing in the recurrence of a and $d_{a,b}$ the torsion number of the eigenvalues appearing in the recurrences of a, b. Then,*

1. *$\sigma(a(n)) = a(n+1)$ is C^2-finite of order at most $\operatorname{ord}(a)$,*
2. *$a(n)+b(n)$ is C^2-finite of order at most $d_{a,b}(\operatorname{ord}(a)+\operatorname{ord}(b))$,*
3. *$a(n)b(n)$ is C^2-finite of order at most $d_{a,b} \operatorname{ord}(a) \operatorname{ord}(b)$,*
4. *$\sum_{k=0}^{n} a(k)$ is C^2-finite of order at most $\operatorname{ord}(a)+1$,*
5. *$a(kn+\ell)$ is C^2-finite of order at most $d_a \operatorname{ord}(a)$ for all $k, \ell \in \mathbb{N}$ and*
6. *the interlacing $e(n) = a_r(q)$ where $n = qm+r$ for $0 \le r < m$ is C^2-finite of order at most $m \max_{r=0,\ldots,m-1} \operatorname{ord}(a_r)$.*

Proof. 1: Clear from the proof of Theorem 3.26.

2, 3: Theorem 5.14.

4: Let $b(n) := \sum_{k=0}^{n} a(k)$. Then, $\sigma(b) - b = \sigma(a)$. If $\mathcal{A}$ is an annihilator of $\sigma(a)$, then $\mathcal{A} \cdot (-1 + \sigma)$ is an annihilator of b of order $\operatorname{ord}(\mathcal{A}) + 1$.

5: Theorem 5.13.

6: Theorem 5.10. □

Example 5.9. Let p denote the Perrin sequence (cf. Example 2.14). With Theorem 5.8 and Theorem 5.19, the sequence

$$\sum_{k=0}^{5n+1} \left(p((2k+1)^2) + \frac{p(k^2)}{p(3k+2)} \right)$$

is C^2-finite of order at most 7.

5.2.1 Interlacing and subsequence

Theorem 5.10. *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 5.10, 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 5.11. 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 5.10 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 5.10 is therefore tight in general.

Lemma 5.12. *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 $k \in \mathbb{N}$. We can compute a C^2-finite recurrence of order at most r which is satisfied by all sequences $a(dkn+i)$ for $i = 0, \ldots, dk-1$.*

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

$$M = M_a(dkn) \cdots M_a(dkn + dk - 1)$$

are nondegenerate. By Theorem 5.7, $Q(R)$ is a field. Therefore, if $s = r$, then the linear system (5.2) is underdetermined and we can compute an element (after clearing denominators) $(x_0, \ldots, x_r) \in R^{r+1}$ 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. $\square$

To extend Lemma 5.12 to subsequences at arbitrary arithmetic progressions we write such an arbitrary subsequence as the interlacing of certain subsequences for which Lemma 5.12 can be applied.

Theorem 5.13. *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 $k \in \mathbb{N}$. We can compute a C^2-finite recurrence of order at most dr which is satisfied by the sequence $a(kn)$.*

Proof. By Lemma 5.12 we can compute a recurrence of order at most r satisfied by $a(dkn + i)$ for $i = 0, \ldots, dk-1$. Let b be the interlacing of the d sequences

$$a(dkn), a(dkn + k), \ldots, a(dkn + (d-1)k).$$

By Theorem 5.10, b has order at most dr. We show that $b(n) = a(kn)$: Let $n = qd + s$ with $0 \leq s < d$. Then, by the definition of b

$$b(n) = b(qd + s) = a(dkq + sk) = a(k(dq + s)) = a(kn).$$

□

5.2.2 Ring operations

Theorem 5.14. *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_1r_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 5.12. 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 5.10, 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.

Example 5.15. 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 $4 = d(\operatorname{ord}(a) + \operatorname{ord}(b))$

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

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

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

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

The order bound is not reached in the previous example. In fact, we could not find any example where the bounds from Theorem 5.14 are reached. This, of course, yields the obvious question whether the bounds we have found are, in general, sharp.

Open Question 5.16. Let $d, r_1, r_2 \in \mathbb{N}$. Can we find C^2-finite sequences a, b of orders r_1, r_2 such that the torsion number of the eigenvalues appearing in the recurrences is d and $\operatorname{ord}(a+b) = d(\operatorname{ord}(a) + \operatorname{ord}(b))$ (or $\operatorname{ord}(ab) = d \operatorname{ord}(a) \operatorname{ord}(b)$).

Even a single example with $d \neq 1$ where the bound is reached would already be interesting.

Theorem 5.14 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 by Example 5.1. Hence, the following question is still open.

Open Question 5.17. Is the ring of C^2-finite sequences computable? I.e., suppose we are given C^2-finite sequences a, b with their recurrences and enough initial values. Can we compute a C^2-finite recurrence for $c = a + b$ or $c = ab$ together with a number r such that the initial values $c(0), \dots, c(r)$ uniquely determine the sequence c?

Closely related is also the following problem: Suppose we have C^2-finite sequences a, b. For checking whether these sequences are identical we can compute a recurrence for $c = a - b$. Of course, $a = b$ if and only if $c = 0$. If we can compute the zeros of the leading coefficient of the recurrence of c, we can check whether $c = 0$ by checking sufficiently many initial values. Hence, identity checking of C^2-finite sequences can be reduced to the Skolem Problem. We do not know if the converse holds:

Open Question 5.18. Can the Skolem Problem be reduced to identity checking of C^2-finite sequences?

5.2.3 Sparse subsequences

Theorem 5.19. *Let c be C-finite of order r and $\lambda_1, \ldots, \lambda_m$ its eigenvalues and $\lambda_i \neq 0$ for all $i = 1, \ldots, m$. 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 [KM14] 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 (5.3), 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 $r = 0, \ldots, d-1$ 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 5.10 we can therefore compute a C^2-finite recurrence of order dr for b. $\square$

Example 5.20. 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

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

Computing a C^2-finite recurrence for $c(n^2)$ where c is a C-finite sequence of order 2 is usually possible as the corresponding linear system is small. However, if c has order 3 it can already be difficult.

Open Question 5.21. Compute a C^2-finite recurrence for $p(n^2)$ where $p(n)$ is the sequence of Perrin numbers (cf. Example 2.14).

In fact, setting up the linear system and solving it via guessing yields a recurrence for $p(n^2)$ of order 3 having coefficients with maximal order 32. Checking the first 1000 terms indicates that this recurrence is indeed correct. The recurrence is, however, much more complicated than the easy recurrence we have for the sparse Fibonacci numbers (cf. Example 3.7) and other sequences of order 2.

6 A computable subring: simple C^2-finite sequences

We have seen that many computations with C^2-finite sequences are limited by the Skolem Problem. The problem stems from possible zeros in the leading coefficient of the recurrence. This can be avoided by only considering sequences which satisfy a linear recurrence with C-finite coefficients and constant leading coefficient. This chapter is based on the article [NP22b].

6.1 Algebraic characterization

Our notion for simple C^2-finite sequences is based on the analogous notion of simple P-recursive sequences for D-finite sequences [Kot12].

Definition 6.1. A sequence $a \in \mathbb{K}^{\mathbb{N}}$ is called *simple C^2-finite* if there is a linear recurrence operator $\mathcal{A} \in \mathcal{R}_C[\sigma]$ with $\operatorname{lc}(\mathcal{A}) = 1$ which annihilates a, i.e., $\mathcal{A}a = 0$.

Equivalently, we could restrict $\operatorname{lc}(\mathcal{A}) \in \mathbb{K}$ in Definition 6.1. As C-finite sequences are closed under multiplication with a field element, multiplying the operator $\mathcal{A}$ by $\frac{1}{\operatorname{lc}(\mathcal{A})}$ yields an annihilating operator with leading coefficient 1.

Many of the C^2-finite sequences that we have considered earlier are in fact simple C^2-finite.

Example 6.2. Let f denote the Fibonacci sequence. In Example 3.7 we have seen that $a(n) := f(n^2)$ satisfies a C^2-finite recurrence of order 2 with coefficients having maximal order 2. The sequence a is even simple C^2-finite and satisfies a recurrence of order 3 with coefficients having order at most 4:

$$-f(6n+11)a(n) - c_1(n)a(n+1) + f(6n+9)a(n+2) + a(n+3) = 0$$

with

$$c_1(n) - 54c_1(n+1) + 331c_1(n+2) - 54c_1(n+3) + c_1(n+4) = 0$$

and initial values

$$c_1(0) = 136, c_1(1) = 6710, c_1(2) = 317434, c_1(3) = 14927768.$$

This recurrence can be found using guessing and fixing the coefficients of the recurrence to only involve C-finite sequences which have certain powers of the golden ratio (and its conjugate) as roots. The recurrence can then be verified using closure properties of C^2-finite sequences. Using an algorithm for computing algebraic relations of C-finite sequences due to Kauers and Zimmermann [KZ08], we can write c_1 in terms of the Fibonacci sequence as

$$c_1(n) = f(4n+6)(-1 - 2f(4n+4) + 3f(4n+6)).$$

In fact, for any C-finite sequence c and $j, k, \ell \in \mathbb{N}$ we can find a simple C^2-finite recurrence for the sequence $c(jn^2 + kn + \ell)$: In this section we show that simple C^2-finite sequences form a ring. Furthermore, this ring clearly contains all C-finite sequences. Hence, the proof of Corollary 3.6 in [JPNP21] can be adjusted to show that the subsequence $c(jn^2 + kn + \ell)$ is simple C^2-finite.

However, not all C^2-finite sequences are simple C^2-finite. In fact, not all D-finite sequences are simple C^2-finite.

Example 6.3. The Catalan numbers (Example 2.4) are D-finite but neither simple P-recursive [Kot12, Section 8.1.5] nor polynomial recursive [CMP+21, Corollary 8]. In particular, the Catalan numbers are not simple C^2-finite over $\mathbb{Q}$.

By Lemma 3.16, for every simple C^2-finite sequence $a \in \mathbb{C}^{\mathbb{N}}$ there is an $\alpha \in \mathbb{Q}$ such that $|a(n)| \leq \alpha^{n^2}$ for all $n \geq 1$. As discussed after the lemma, such an α can be computed explicitly for simple C^2-finite sequences.

Analogous to Theorem 3.23, we can find an equivalent characterization for simple C^2-finite sequences in terms of finitely generated modules.

Theorem 6.4. *The following are equivalent:*

1. *The sequence a is simple C^2-finite.*
2. *There exists $\mathcal{A} \in \mathcal{R}_C[\sigma]$ with $\operatorname{lc}(\mathcal{A}) = 1$ and a simple C^2-finite sequence b with $\mathcal{A}a = b$.*
3. *The module $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{\mathcal{R}_C}$ over the ring of C-finite sequences $\mathcal{R}_C$ is finitely generated.*

Based on this characterization, one can easily prove the following theorem analogous to Theorem 3.26:

Theorem 6.5. *The set of simple C^2-finite sequences is a difference ring under termwise addition and termwise multiplication.*

In Section 3.4 we have studied the generating functions of C^2-finite sequences. In the case of simple C^2-finite sequences we can find an equivalent characterization in terms of certain functional equations.

Theorem 6.6. *The sequence $a \in \overline{\mathbb{Q}}^{\mathbb{N}}$ is simple C^2-finite if and only if its generating function $g(x) := \sum_{n \geq 0} a(n) x^n$ satisfies a functional equation of the form*

$$\sum_{k=1}^{m} \alpha_k x^{j_k} g^{(d_k)}(\lambda_k x) = p(x) \tag{6.1}$$

for

1. *$\alpha_1, \ldots, \alpha_k, \lambda_1, \ldots, \lambda_k \in \overline{\mathbb{Q}} \setminus \{0\}$,*
2. *$j_1, \ldots, j_m, d_1, \ldots, d_m \in \mathbb{N}$,*
3. *$p \in \overline{\mathbb{Q}}[x]$ and*
4. *let $s := \max_{k=1,\ldots,m}(d_k - j_k)$, then for all $k = 1, \ldots, m$ with $d_k - j_k = s$ we have $d_k = 0$ and $\lambda_k = 1$.*

Proof. $\Longrightarrow$: With Theorem 3.27 we can clearly find a functional equation satisfying properties (1)–(3). The contribution from the leading term after clearing the common denominator x^r is given by $h(x) = \sum_{n \geq 0} a(n+r) x^n = g(x) - p_{0,0}(x)$ for some polynomial $p_{0,0}(x)$. In particular, $d_k = j_k = 0$ and $\lambda_k = 1$, so $d_k - j_k = 0$. The other terms $n^j \lambda^n a(n+i)$ with $i < r$ yield contributions with $j_k = l - i + r, d_k = l$ in (6.1). In particular, $d_k - j_k = i - r < 0$.

$\Longleftarrow$: According to the proof of Theorem 3.30, the leading coefficient in the recurrence is given by terms where $d_k - j_k$ is maximal. If $\lambda_k = 1$ and $d_k = 0$ in these cases, then this leading coefficient is just 1. $\square$

6.2 Computable ring

In Chapter 4 we have seen how closure properties for C^2-finite sequences can be reduced to solving systems of equations of the form

$$Ax = b \tag{6.2}$$

with $A \in Q(R)^{m \times s}, b \in Q(R)^m$ and

$$R := \mathbb{K}_\sigma[c_0, \ldots, c_r] \subsetneq \mathcal{R}_C \tag{6.3}$$

with $c_0, \ldots, c_r \in \mathcal{R}_C$. All the fractional sequences, in fact, originate from the leading coefficient of the C^2-finite recurrences. Therefore, if the sequences for which we perform closure properties are simple C^2-finite, the linear system (6.2) is of the form $A \in R^{m \times s}$ and $b \in R^m$. We show how such systems can be solved. Our method is based on the closed form of C-finite sequences. Therefore, we assume that the base field is always the field of algebraic numbers $\overline{\mathbb{Q}}$. Note that every C-finite sequence over $\overline{\mathbb{Q}}$ has again a closed form as $\overline{\mathbb{Q}}$ is algebraically closed itself. First, we consider the special case, where we compute a constant solution of such a system.

Lemma 6.7. *We can compute all constant solutions $x \in \overline{\mathbb{Q}}^s$ of the linear system $Ax = b$ where $A \in \mathcal{R}_C^{m \times s}$ and $b \in \mathcal{R}_C^m$. In particular, we can decide whether such a solution exists.*

Proof. It is sufficient to consider one equation, i.e., $A \in \mathcal{R}_C^{1 \times s}$. The set of constant solutions is an affine subspace of $\overline{\mathbb{Q}}^s$. For several equations we can compute the intersection of these affine subspaces to determine all solutions. Using the closed form of the sequences, we can rewrite the equation $Ax = b$ as

$$\sum_{k=1}^{l} \underbrace{\left(\sum_{i \in S_k} \epsilon_{k,i} x_i + \epsilon_k \right)}_{=:y_k} (n - n_0)^{d_k} \lambda_k^{n - n_0} = 0, \quad \text{for all } n \geq n_0 \tag{6.4}$$

with $n_0 \in \mathbb{N}$, and $\epsilon_{k,i}, \epsilon_k, \lambda_k \in \overline{\mathbb{Q}}, d_k \in \mathbb{N}$ and $S_k \subseteq \{1, \dots, s\}$ for all $k = 1, \dots, l$. Certainly, if $y_k = 0$ for all $k = 1, \dots, l$ we have a solution. On the other hand, evaluating this equation for $n = n_0, n_0 + 1, \dots,$ yields a linear system for the y_k. This linear system is a generalized Vandermonde matrix, in particular it is regular [LT08, Liu68]. Therefore, if equation (6.4) holds, then $y_k = 0$ for all $k = 1, \dots, l$. This yields a linear system over $\overline{\mathbb{Q}}$ which can be solved. For the initial terms $n = 0, 1, \dots, n_0 - 1$ the equation $Ax = b$ can simply be solved over $\overline{\mathbb{Q}}$. The affine space of all solutions of the single equation is now given as the intersection of the affine subspace arising from solving equation (6.4) and the affine subspaces arising from the initial terms. $\square$

Suppose the sequence c is C-finite over the algebraic numbers $\overline{\mathbb{Q}}$ with closed form

$$c(n + n_0) = \sum_{i=1}^{m} p_i(n)\lambda_i^n$$

as in Theorem 2.15, i.e., $\lambda_1, \dots, \lambda_m \in \overline{\mathbb{Q}}, p_1, \dots, p_m \in \overline{\mathbb{Q}}[n]$ and $\deg(p_1) = d_1, \dots, \deg(p_m) = d_m$. Let

$$B_c := \left\{ \left(n^j \lambda_i^n\right)_{n \in \mathbb{N}} \mid i \in \{1, \dots, m\}, j \in \{1, \dots, d_i - 1\} \right\}. \tag{6.5}$$

Then, the sequence c is an $\mathbb{L}$-linear combination of sequences in B_c from n_0 on. Suppose n_0 is the smallest index such that each c_i from (6.3) is a $\overline{\mathbb{Q}}$-linear combination of sequences in B_{c_i}, as defined in (6.5), from n_0 on. We write

$$B := B_{c_0} \cup \dots \cup B_{c_r} \cup \{1\}.$$

Then, for any sequence $c \in R$ there is an $N \in \mathbb{N}$ and coefficients $x_{d_1,\dots,d_N} \in \overline{\mathbb{Q}}$ such that

$$c(n) = \sum_{d_1,\dots,d_N \in B} x_{d_1,\dots,d_N} d_1(n) \cdots d_N(n), \quad \text{for all } n \geq n_0. \tag{6.6}$$

Furthermore, if $c \in R$ is given by a recurrence and initial values such a representation can be computed: We can compute the closed form of c. Now, every term in this closed form has to be a product of the finitely many (not necessarily distinct) sequences from B.

Lemma 6.8. *Let $A \in R^{m \times s}$ and $b \in R^m$. If $Ax = b$ has a solution $x \in R^s$, then we can compute a solution $x \in \mathcal{R}_C^s$.*

Proof. We can assume that every sequence in the linear system is given in the form (6.6). For $N = 1, 2, \ldots$ we write

$$x_i = \sum_{d_1,\ldots,d_N \in B} x_{i,d_1,\ldots,d_N} d_1 \cdots d_N$$

for unknown coefficients $x_{i,d_1,\ldots,d_N} \in \overline{\mathbb{Q}}$ for $i = 1, \ldots, s$. In particular, for fixed N, we can compute sequences $e_1, \ldots, e_l \in \mathcal{R}_C$ such that $x_i = \sum_{k=1}^{l} x_{i,k} e_k$ with $x_{i,k} \in \overline{\mathbb{Q}}$ unknown for $i = 1, \ldots, s$. Let

$$\hat{x} := \Big(x_{1,1}, \ldots, x_{s,1}, \ldots, x_{1,l}, \ldots, x_{s,l}\Big)^{\top}.$$

Then, $Ax = b$ has a solution for the $x_{i,d_1,\ldots,d_N}$ if and only if the $m \times ls$ linear system

$$\Big(e_1 A, \ldots, e_l A\Big)\, \hat{x} = b$$

has a solution for $\hat{x}$. With Lemma 6.7 we can check whether the linear system has a solution for $\hat{x}$. If we have found a solution we can easily compute the corresponding x_i. As we know that a solution x of this form exists, this algorithm has to terminate for large enough N. $\square$

Theorem 6.9. *The ring of simple C^2-finite sequences over $\overline{\mathbb{Q}}$ is computable.*

Proof. Suppose a, b are simple C^2-finite with annihilating operators $\sum_{i=0}^{r_1-1} c_i \sigma^i + \sigma^{r_1}$ and $\sum_{i=0}^{r_2-1} d_i \sigma^i + \sigma^{r_2}$, respectively. By Lemma 4.9 (which can be proven completely analogously for simple C^2-finite sequences), there is a linear system over the computable ring $R := \overline{\mathbb{Q}}_\sigma[c_0, \ldots, c_{r_1-1}, d_0, \ldots, d_{r_2-1}]$ which has a solution and whose solution gives rise to a recurrence for $a + b$ or ab. This linear system can be computed and a solution of the system can be obtained with Lemma 6.8. As we do not know a priori how big this order s from Lemma 4.9 is and how big the N in the proof of Lemma 6.8 has to be chosen, we can simultaneously increase s and N. Eventually, this algorithm terminates and any solution gives rise to a recurrence for $a + b$ or ab. $\square$

Because we are working with the closed form of C-finite sequences, the recurrences that we compute in Theorem 6.9 might be over a bigger field than we started with. E.g., it might be that the sequences a, b are simple C^2-finite over $\mathbb{Q}$, but the coefficients in the recurrence for $a + b$ are C-finite over $\mathbb{K} \supsetneq \mathbb{Q}$. However, we do know by Theorem 6.5 that a recurrence with coefficients over $\mathbb{Q}$ exist.

Open Question 6.10. Is the ring of simple C^2-finite sequences over a field $\mathbb{K}$ computable?

In fact, a positive answer to Question 4.8 would also give a positive answer to this question by combining Theorem 6.5 and Theorem 6.9.

C^2-finite sequences are also closed under taking differences, partial sums, subsequences at arithmetic progressions and interlacing. The same proofs carry over to simple C^2-finite sequences. Even more, as these operations can be reduced to solving linear systems, these closure properties can be computed effectively.

This method for computing closure properties can also yield nicer (i.e., smaller coefficients) recurrences than the method from Chapter 4 as the following example shows.

Example 6.11. Consider the sequences

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

Both are simple C^2-finite. We want to compute a recurrence for $c := a + b$. An ansatz of order 3 yields the linear system

$$\begin{pmatrix} 1 & -2^n & 2 \cdot 4^n \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 8 \cdot 8^n \\ 1 \end{pmatrix}.$$

This is the smallest system which has a solution. Using the generalized inverse method from Algorithm 2 to compute the solution we get the recurrence

$$\begin{aligned} &\left(-2^{5n+4} + 2^{4n+2} + 2^{3n+3} - 2^{2n+1}\right) c(n) \\ &\quad + \left(2^{5n+4} - 2^{3n+3} - 2^{2n+1} + 1\right) c(n+2) \\ &\qquad + \left(2^{4n+2} - 2^{2n+2} + 1\right) c(n+3) = 0 \end{aligned} \tag{6.7}$$

if we use columns 0 and 2 of the matrix. Using columns 1 and 2 we get

$$\begin{aligned}\left(2^{3n+4} - 3 \cdot 2^{2n+2} + 2^{n+1}\right) c(n+1) \\ + \left(2^{3n+4} - 2^{2n+3} - 2^{n+1} + 1\right) c(n+2) \\ + \left(2^{2n+2} - 2^{n+2} + 1\right) c(n+3) = 0.\end{aligned}$$

Both recurrences have coefficients with maximal order 4. By Theorem 3.26, we know that c also has to satisfy a recurrence with leading coefficient 1.

We make an ansatz $x_i = x_{i,1} + x_{i,2} 2^n$ and write

$$\hat{x} = (x_{0,1}, x_{1,1}, x_{2_1}, x_{0,2}, x_{1,2}, x_{2_2}).$$

The linear system for $\hat{x} \in \mathbb{Q}^6$ computed in Lemma 6.8 is given by

$$\begin{pmatrix} 1 & -2^n & 2 \cdot 4^n & 2^n & -4^n & 2 \cdot 8^n \\ 1 & -1 & 1 & 2^n & -2^n & 2^n \end{pmatrix} \hat{x} = \begin{pmatrix} 8 \cdot 8^n \\ 1 \end{pmatrix}.$$

Comparing the coefficients of $1, 2^n, 4^n, 8^n$ as in Lemma 6.7 yields the constant system

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \\ 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 \end{pmatrix} \hat{x} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 8 \\ 1 \\ 0 \end{pmatrix}.$$

This system has the unique solution

$$\hat{x} = \left(0, 2, 3, 2, 6, 4\right)$$

which gives rise to the recurrence

$$(2 \cdot 2^n)\, c(n) + (2 + 6 \cdot 2^n)\, c(n+1) + (3 + 4 \cdot 2^n)\, c(n+2) + c(n+3) = 0. \tag{6.8}$$

The ideas of Chapter 5 seem to not carry over easily to the case of simple C^2-finite sequences. It would be interesting to see whether similar order bounds can be proven for this case.

Open Question 6.12. Can we derive similar order bounds as in Theorem 5.14 for simple C^2-finite sequences?

Identities of C-finite sequences can often be proven fully automatically by checking a moderate number of initial values. For proving the identity in the introduction of the thesis, we can define the C-finite sequence

$$c(n) := \sum_{k=0}^{2n} f(k)f(k+1) - f(2n+1)^2 + 1$$

where f denotes the Fibonacci numbers. By the order bounds of C-finite sequences, this sequence c has order at most 10. Hence, if $c(0) = \cdots = c(9) = 0$, then c is the constant zero sequence which proves the identity.

Similar methods for D-finite sequences are more difficult because one has to take possible zeros in the leading coefficient of the recurrence into account. Therefore, for D-finite sequences this method is often not feasible in practice [Yen96, Yen97, GHS08]. However, deriving reasonable order bounds for sequences that are simple (P-recursive or C^2-finite), we might be able to extend the method used for C-finite sequences to larger classes.

7 Extension to C^k-finite and D^k-finite sequences

Instead of considering sequences satisfying linear recurrences with C-finite coefficients we can allow D-finite sequences. Analogous to D^n-finite functions (cf. [JPP19, JPPS20]) the construction can be iterated, i.e., we can, for instance, define C^3-finite sequences as sequences satisfying a linear recurrence with C^2-finite coefficients. This chapter shows that these C^k-finite and D^k-finite sequences form an increasing chain of difference rings. The results are based on [JPNP23].

7.1 Definition and examples

A sequence is called C^0-finite if it is constant and called D^0-finite if it is polynomial.

Definition 7.1. Let $k \geq 1$. A sequence $a \in \mathbb{K}^{\mathbb{N}}$ is called C^k-finite (or D^k-finite) over $\mathbb{K}$ if there are C^{k-1}-finite (or D^{k-1}-finite) sequences $c_0, \ldots, c_r$ over $\mathbb{K}$ with $c_r(n) \neq 0$ for all $n \in \mathbb{N}$ such that

$$c_0(n)a(n) + c_1(n)a(n+1) + \cdots + c_r(n)a(n+r) = 0$$

for all $n \in \mathbb{N}$.

Example 7.2. Let $a(n) := \prod_{k=1}^{n} k!$. The sequence a is D^2-finite satisfying the recurrence

$$(n+1)!\, a(n) - a(n+1) = 0, \quad \text{for all } n \in \mathbb{N}.$$

The sequence is called the superfactorial (A000178 in the OEIS).

Example 7.3. Let $\alpha \in \mathbb{K}$. Every sequence a with $a(n) := \alpha^{n^3}$ is C^3-finite satisfying the recurrence

$$c(n)a(n) - a(n+1) = 0, \quad \text{for all } n \in \mathbb{N},$$

where $c(n) = \alpha^{3n^2+3n+1}$ is C^2-finite (cf. Example 3.4). More generally, $a(n) = \alpha^{n^k}$ is C^k-finite for every $k \in \mathbb{N}$.

Example 7.4. Using the same argument as in [KM14] one can derive a C^3-finite recurrence for $f(n^3)$ where f denotes the Fibonacci numbers:

$$c_0(n)f(n^3) + c_1(n)f((n+1)^3) + c_2(n)f((n+2)^3) = 0, \quad \text{for all } n \in \mathbb{N},$$

with

$$\begin{aligned} c_0(n) &= f(3n^2+9n+7)f(3n^2+3n+3)f(3n^2+3n+1) \\ &\qquad - f(3n^2+9n+7)f(3n^2+3n+2)^2, \\ c_1(n) &= f(3n^2+9n+7)f(3n^2+3n+2) + f(3n^2+9n+6)f(3n^2+3n+1), \\ c_2(n) &= -f(3n^2+3n+1). \end{aligned}$$

These coefficients c_0, c_1, c_2 are C^2-finite with Theorem 3.26 and Theorem 4.7. Furthermore, clearly $c_2(n) \neq 0$ for all n.

7.2 Ring structure

Adapting Section 3.3 to this more general setting, we show that the sets of C^k-finite and D^k-finite sequences form difference rings. We denote the set of C^k-finite sequences by $\mathcal{R}_{C^k}$, the set of D-finite sequences by $\mathcal{R}_D$ and the set of D^k-finite sequences by $\mathcal{R}_{D^k}$.

Now, Lemma 3.21, Lemma 3.22 and Theorem 3.23 can be formulated completely analogously for C^k-finite and D^k-finite sequences:

Lemma 7.5. *Let a be C^k-finite (or D^k-finite) with annihilating operator $\mathcal{A} = c_0 + \cdots + c_r\sigma^r$ and let R be the difference ring generated by $c_0, \ldots, c_r$. If $S \supseteq R$ is a subring of the ring of sequences $\mathbb{K}^{\mathbb{N}}$, then $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ is finitely generated.*

Lemma 7.6. *Let $a \in \mathbb{K}^{\mathbb{N}}$ and S be a subring of the set of C^{k-1}-finite (or D^{k-1}-finite) sequences. If $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)}$ is finitely generated, then a is C^k-finite (or D^k-finite).*

The proofs of Lemma 7.5 and Lemma 7.6 are analogous to the proofs of the corresponding lemmas in Section 3.3. Using Lemma 7.5 and Lemma 7.6 one can again prove a characterization for C^k-finite and D^k-finite sequences.

Theorem 7.7. *Let $a \in \mathbb{K}^{\mathbb{N}}$.*

1. *The sequence a is C^k-finite if and only if $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_{C^{k-1}})}$ is finitely generated.*
2. *The sequence a is D^k-finite if and only if $\langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(\mathcal{R}_{D^{k-1}})}$ is finitely generated.*

Similarly as in the C^2-finite setting, we can use Theorem 7.7 to show that the sets of C^k-finite and D^k-finite sequences form difference rings. Example 3.24 shows that these rings are not Noetherian. Hence, the idea is, again, to restrict the underlying ring to a Noetherian subring.

Lemma 7.8.
1. *Let $\mathcal{A} = \sum_{i=0}^r c_i \sigma^i \in \mathcal{R}_{C^k}[\sigma]$. Then, the $\mathbb{K}$-difference-algebra $\mathbb{K}_\sigma[c_0, \dots, c_r]$ is contained in a Noetherian ring S.*
2. *Let $\mathcal{A} = \sum_{i=0}^r c_i \sigma^i \in \mathcal{R}_{D^k}[\sigma]$. Then, the $\mathbb{K}(n)$-difference-algebra $\mathbb{K}(n)_\sigma[c_0, \dots, c_r]$ is contained in a Noetherian ring S.*

Proof. We use induction on k. For $k = 0$ we have $\mathcal{R}_{C^0} = \mathbb{K}$ and $\mathcal{R}_{D^0} = \mathbb{K}[n]$ which are both Noetherian.

Now, let c be a coefficient of $\mathcal{A}$ and let $\mathcal{C}$ be its annihilator. By induction, the difference-algebra generated by the coefficients of $\mathcal{C}$ is contained in a Noetherian ring S_c. Then, also the localization $Q(S_c)$ is Noetherian [AM69, Proposition 7.3]. By Lemma 7.5, the module $\langle \sigma^i(c) \mid i \in \mathbb{N} \rangle_{Q(S_c)}$ is finitely generated. Hence, also the difference-algebra $A_c := Q(S_c)_\sigma[c]$ is finitely generated and is, in particular, a Noetherian ring containing $\mathbb{K}_\sigma[c]$ (or $\mathbb{K}(n)_\sigma[c]$ in the D-finite case). Then, S can be chosen as the smallest ring containing the Noetherian rings $A_{c_0}, \dots, A_{c_r}$. This ring S is again Noetherian [AM69, Corollary 7.7]. $\square$

Theorem 7.9. *The sets of C^k-finite (resp. D^k-finite) sequences are difference rings under termwise addition and termwise multiplication.*

Proof. Let a, b be C^k-finite (or D^k-finite) sequences and $\mathcal{A} = c_0 + c_1\sigma + \cdots + c_{r_1}\sigma^{r_1}$ and $\mathcal{B} = d_0 + d_1\sigma + \cdots + d_{r_2}\sigma^{r_2}$ the corresponding annihilating operators.

With Lemma 7.8, there is a Noetherian ring S which contains all difference rings generated by $c_0, \dots, c_{r_1}, d_0, \dots, d_{r_2}$. Hence, with Lemma 7.5, the modules

$$\langle \sigma^i(a+b) \mid i \in \mathbb{N} \rangle_{Q(S)} \subseteq \langle \sigma^i(a) \mid i \in \mathbb{N} \rangle_{Q(S)} + \langle \sigma^i(b) \mid i \in \mathbb{N} \rangle_{Q(S)}$$

and

$$\langle \sigma^i(ab) \mid i \in \mathbb{N} \rangle_{Q(S)} \subseteq \langle \sigma^i(a)\sigma^j(b) \mid i, j \in \mathbb{N} \rangle_{Q(S)}$$

are finitely generated as they are submodules of finitely generated modules over a Noetherian ring. By Lemma 7.6, the sequences $a + b$ and ab are C^k-finite (or D^k-finite).

The operator

$$\tilde{\mathcal{A}} := \sigma(c_0) + \sigma(c_1)\sigma + \cdots + \sigma(c_{r_1})\sigma^{r_1}$$

annihilates $\sigma(a)$. Hence, the ring is also closed under shifts. $\square$

Using the ansatz method described in Chapter 4 one can reduce the computation of ring operations to solving linear systems. For instance, for D^2-finite sequences, we need to solve linear systems over the D-finite sequence ring. The ideas from Theorem 4.5 and Theorem 5.10 can be used to show that C^k-finite and D^k-finite sequences are closed under taking subsequences at arithmetic progressions and interlacings.

Example 7.10. We define the D-finite sequences

$$\begin{aligned}
(n^2+1)c_0(n) + c_0(n+1) &= 0, & c_0(0) &= 2,\\
(n+7)c_1(n) + (-n-1)c_1(n+1) &= 0, & c_1(0) &= 2,\\
(n+1)d_0(n) - d_0(n+1) &= 0, & d_0(0) &= 1,\\
(n+2)d_1(n) + (-n^2-3)d_1(n+1) &= 0, & d_1(0) &= 4,
\end{aligned}$$

and the D^2-finite sequences

$$\begin{aligned} c_0(n)a(n) + c_1(n)a(n+1) &= 0, & a(0) &= 3, \\ d_0(n)b(n) + d_1(n)b(n+1) &= 0, & b(0) &= 5. \end{aligned}$$

By Theorem 7.9, the sequence $h := ab$ is D^2-finite. With the methods introduced in Chapter 4 we can compute the recurrence

$$e(n)h(n) + h(n+1) = 0, \quad h(0) = 15$$

with

$$(n^6 + 2n^5 + 5n^4 + 8n^3 + 7n^2 + 6n + 3)e(n) + (n^2 + 9n + 14)e(n+1) = 0$$

and $e(0) = -\frac{1}{4}$.

By induction, every C^k-finite sequence is D^k-finite and every D^k-finite sequence is C^{k+1}-finite. Therefore, we get the following chain of rings

$$\mathcal{R}_C \subseteq \mathcal{R}_D \subseteq \mathcal{R}_{C^2} \subseteq \mathcal{R}_{D^2} \subseteq \mathcal{R}_{C^3} \subseteq \cdots$$

Example 7.4 is true more generally and the following generalization of Theorem 4.7 can be shown:

Corollary 7.11. *Let c be a C-finite sequence over the field $\mathbb{K}$ and $p \in \mathbb{N}[n]$. Denote $k := \deg(p)$. Then, $c(p(n))$ is C^k-finite over the splitting field $\mathbb{L}$ of the characteristic polynomial of c.*

Proof. We use induction on k. For $k = 1$, this is precisely the fact that C-finite sequences are closed under taking subsequences at arithmetic progressions. Let $k \geq 2$. We can write c as an $\mathbb{L}$-linear combination of sequences $d(n) = n^i \alpha^n$ for $i \in \mathbb{N}$ and $\alpha \in \mathbb{L}$ from some term on. Let $p(n) = p_k n^k + q(n)$ with $\deg(q) \leq k - 1$. Then, we have

$$d(p(n)) = p(n)^i \alpha^{p(n)} = p(n)^i \left(\alpha^{p_k}\right)^{n^k} \alpha^{q(n)}.$$

The sequence $p(n)^i$ is polynomial and therefore C^k-finite. The sequence $\left(\alpha^{p_k}\right)^{n^k}$ is C^k-finite (as seen in Example 7.3). By induction $\alpha^{q(n)}$ is C^{k-1}-finite, so in particular C^k-finite.

Therefore, with Theorem 7.9, the sequence $d(p(n))$ is C^k-finite as it is the product of C^k-finite sequences. Since C^k-finite sequences are also closed under $\mathbb{L}$-linear combinations and shifts, $c(p(n))$ is C^k-finite. □

The same question from Corollary 7.11 can of course be asked for D-finite sequences. Neither the proof of Corollary 7.11 (as D-finite sequences do not have a nice closed form) nor the proof of Theorem 4.7 (as Lemma 11 in [KM14] does not hold for the D-finite case) carry over to this case.

Open Question 7.12. Let a be a D-finite sequence. Is $a(n^2)$ a D^2-finite sequence?

8 Positivity of C-finite sequences

In the previous chapters we have seen that the Skolem Problem, i.e., the problem of deciding whether a given C-finite sequence $c(n)$ has a zero, plays an important role when computing with C^2-finite sequences. Closely related and similarly difficult is the problem of deciding whether a sequence is positive, i.e., whether $c(n) > 0$ for all $n \in \mathbb{N}$. We call this the *Positivity Problem*. Methods for showing positivity of a sequence can usually be adjusted to show nonnegativity of a sequence. For the sake of a clear presentation, we focus on positivity here. Decidability for this problem is known for sequences of order at most 5 and for certain other classes of C-finite sequences of order at most 9 [HHH06, LT09, OW14b, OW14a].

In the D-finite case even less is known. Similar methods as in the D-finite case were applied to certain sequences of order 2 [NOW21]. For larger orders, procedures based on quantifier elimination (cf. [GK05, KP10, Pil13]) were successfully applied in practice [Kau07b, Pil08, Pil19]. Other techniques are based on writing a sequence as sum of squares or singularity analysis [Cha14, Hoe21, MM22].

Concerning implementation, only very few software packages are known which support proving inequalities of sequences automatically. Two implementations of the Gerhold-Kauers method for Mathematica and SageMath, respectively, are known [Kau06, Ura20]. These, however, do not implement any special procedures for C-finite sequences. An online tool for computing zeros of certain C-finite sequences based on SageMath is presented in [BLN$^+$22]. In the context of the thesis two more software packages were created for dealing with C-finite sequences specifically, the package `rec_sequences` for SageMath and the package `PositiveSequence` for Mathematica (the latter is part of the `RISCErgoSum` collection of packages).

The Positivity Problem is not only interesting by itself but also plays an important role because other problems can be reduced to it. A sequence $c(n)$ has no zeros if and only if the sequence $c(n)^2$ is positive. Hence, as C-finite sequences form a computable ring, the Skolem Problem can be reduced to the Positivity Problem. Further, an inequality problem of the form $c(n) > d(n)$ for all $n \in \mathbb{N}$ can be reduced to checking whether the sequence $c(n) - d(n)$ is positive.

As the Positivity Problem only makes sense over real valued sequences, we assume that the base field $\mathbb{K}$ is a real number field. We denote the field of real algebraic numbers by $\mathbb{A} := \overline{\mathbb{Q}} \cap \mathbb{R} \supsetneq \mathbb{K}$. Suppose a C-finite sequence c has a unique dominant eigenvalue $\lambda_1 \in \mathbb{A}$ (i.e., we have $k = 1$ in the setting of equation (2.2)). Then, (2.3) shows that $c \sim \gamma n^d \lambda_1^n$ for some $\gamma \in \mathbb{A}$. The sequence c can only be positive if $\gamma, \lambda_1 > 0$. Furthermore, c is positive if and only if $c(n)/\lambda_1^n$ is positive. Therefore, in the case of a unique dominant eigenvalue, it is sufficient to show positivity of a sequence

$$p(n) + \sum_{i=1}^{s} \left(o_i(n)\xi_i^n + \overline{o_i}(n)\overline{\xi_i}^n \right) + \sum_{i=1}^{l} q_i(n)\rho_i^n \tag{8.1}$$

with $p \in \mathbb{A}[x], o_1, \ldots, o_s \in \overline{\mathbb{Q}}[x], q_1, \ldots, q_l \in \mathbb{A}[x]$ and constants $\xi_1, \ldots, \xi_s \in \overline{\mathbb{Q}}, \rho_1, \ldots, \rho_l \in \mathbb{A}$ where the leading coefficient of p is positive [OW14b].

In Section 8.1 we discuss several different algorithms which can be used for proving positivity of certain C-finite sequences. These methods are all well-known or slight variations of algorithms which can be found in the literature. In Section 8.2 we compare these algorithms for C-finite sequences coming from the OEIS. Furthermore, we provide some statistics on how many sequences in the OEIS are C-finite or D-finite based on guessing procedures. This chapter is mostly based on [NP22a].

8.1 Algorithms

In this section we give an overview of some methods which can be used to prove positivity of a C-finite sequence. Algorithms 1 and 2 (as well as their adjusted versions Algorithm 1e and 2e) presented below in Sections 8.1.1 and 8.1.2 can be applied to D-finite sequences. As such they can be used to prove positivity of C-finite sequences. However, sometimes C-finite sequences satisfy a D-finite recurrence of lower order, which is better suited as input for these methods. In Section 8.1.3, we discuss when such a D-finite recurrence exists. A method based on the combination of Algorithms 1 and 2 as well as on the closed form of a C-finite sequence is introduced in Section 8.1.5. The methods described in Sections 8.1.4

and 8.1.6 also make use of the closed form of C-finite sequences. They are based on known results, but we believe that they had not been implemented so far.

8.1.1 Algorithm 1

In 2005 [GK05] a method based on quantifier elimination (in particular, cylindrical algebraic decomposition CAD [Col75, CH91, CJ98, BRPR03]) was introduced which can be used to show positivity of sequences that can be defined recursively along some discrete parameter. This procedure, however, is not guaranteed to terminate. For D-finite sequences of small order conditions which guarantee the termination of the algorithm are known [KP10, Pil13].

We give a short description of Algorithm 1 from [KP10]. Suppose c is D-finite of order r. By Theorem 2.5 there are rational functions $q_{\rho,0}(x), \ldots, q_{\rho,r-1}(x) \in \mathbb{K}(x)$ for all $\rho \in \mathbb{N}$ with $c(n+\rho) = \sum_{i=0}^{r-1} q_{\rho,i}(n)c(n+i)$. The idea of the Gerhold-Kauers method is to check with quantifier elimination whether $c(n), \ldots, c(n+r-1) > 0$ implies $c(n+r) > 0$ where $c(n+r)$ can be written in terms of the $c(n), \ldots, c(n+r-1)$. If this is true, then by induction it would be sufficient to check finitely many initial values to deduce positivity of the entire sequence. If, however, this cannot be shown, then we can add $c(n+r) > 0$ to the hypothesis and show $c(n+r+1) > 0$. This process is iterated. In the iteration step $\rho \geq r$ we try to show positivity of the formula

$$\begin{aligned} \Phi(\rho, c) := \forall y_0, \ldots, y_{r-1}, x \in \mathbb{R}\colon \ &\left(x \geq 0 \wedge \bigwedge_{j=0}^{\rho-1} \sum_{i=0}^{r-1} q_{j,i}(x) y_i > 0 \right) \\ &\implies \sum_{i=0}^{r-1} q_{\rho,i}(x) y_i > 0. \end{aligned}$$

The formula $\Phi(\rho, c)$ is a generalized induction formula over the reals. It is certainly sufficient to prove the induction step and has the advantage of being a valid input for CAD and other quantifier elimination methods. Here, we give a slightly adjusted version which searches for an index n_0 such that the sequence $\sigma^{n_0}(c)$ is positive, i.e., it checks whether the sequence is *eventually* positive (hence, we denote the algorithm by Algorithm 1e). If such an n_0 can be found by the algorithm, then it is sufficient to check the initial values $c(0), \ldots, c(n_0 - 1)$ of the sequence to prove positivity of c.

Input : D-finite sequence c of order r
output: n_0 such that $\sigma^{n_0}(c)$ is positive
$n, n_0 \leftarrow 0$
$d \leftarrow c$
while $n < r$ *or* $\neg\Phi(n, d)$ **do**
    if $d(n) > 0$ **then**
        $n \leftarrow n + 1$
    else
        $n_0 \leftarrow n_0 + n + 1$
        $d \leftarrow \sigma^{n+1}(d)$
        $n \leftarrow 0$
    end
end
return n_0

Algorithm 1e: Adjusted version of Algorithm 1 from [KP10]

Clearly, Algorithm 1e is not guaranteed to terminate. E.g., if the input sequence c is not eventually positive, then the algorithm never terminates. Suppose the sequence c is eventually positive, i.e., there exists an $n_0 \in \mathbb{N}$ such that $\sigma^{n_0}(c)$ is positive. As the characteristic polynomials agree, $\chi(c) = \chi(\sigma^{n_0}(c))$, the same termination conditions for Algorithm 1 in [KP10] now also apply to Algorithm 1e.

Example 8.1. The alternating sequence A000034 is C-finite of order 2 satisfying

$$c(n) - c(n+2) = 0$$

with initial values $c(0) = 1, c(1) = 2$. Algorithm 1e terminates for this sequence showing that c is positive.

Example 8.2. The sequence A005682 is C-finite of order 6 satisfying

$$c(n) + c(n+2) - 2c(n+5) + c(n+6) = 0$$

with initial values $c = \langle 1, 2, 4, 8, 15, 28, \dots \rangle$. Algorithm 1e cannot show positivity of c in 60 seconds.

8.1.2 Algorithm 2

Algorithm 2 in [KP10] again uses quantifier elimination to prove positivity of a D-finite sequence. The idea is to check whether there is a $\mu > 0$ such that $c(n+1) \geq \mu c(n)$ for all $n \in \mathbb{N}$. By induction, if there is a $\mu > 0$ such that

$$c(n+1) \geq \mu c(n) \wedge \cdots \wedge c(n+r-1) \geq \mu c(n+r-2) \implies c(n+r) \geq \mu c(n+r-1),$$

then it is again sufficient to check finitely many initial values to prove positivity of c. Hence, the important step in the algorithm is to use quantifier elimination to verify whether there exists a $\mu > 0$ such that the formula

$$\Psi(\xi, \mu, c) := \forall y_0, \ldots, y_{r-1} \in \mathbb{R}\, \forall x \in \mathbb{R}_{\geq \xi} \colon \left(y_0 > 0 \wedge \bigwedge_{i=0}^{r-2} y_{i+1} \geq \mu y_i \right) \implies \sum_{i=0}^{r-1} q_i(x) y_i \geq \mu y_{r-1}$$

is valid where $q_i \in \mathbb{K}(x)$ are again such that $c(n+r) = \sum_{i=0}^{r-1} q_i(n) c(n+i)$ for all $n \in \mathbb{N}$.

Again, we give a slightly adjusted version which searches for an index n_0 such that the sequence $\sigma^{n_0}(c)$ is positive. If the input sequence c is eventually positive, then the same termination conditions as for Algorithm 2 in [KP10] apply in this adjusted version.

Example 8.3. Algorithm 2e can show positivity of the sequence from Example 8.2 (which could not be done with Algorithm 1e).

Example 8.4. Algorithm 2e cannot show positivity of the alternating sequence from Example 8.1 (which could be proven with Algorithm 1e).

8.1.3 D-finite reduction

Clearly, every C-finite sequence is also D-finite. Sometimes, C-finite sequences satisfy D-finite recurrences of lower order. In these cases it can be helpful to use this shorter D-finite recurrence as the next example shows.

Example 8.5. Let c be the sequence defined by $c(n) = n^2 + 1$ for all $n \in \mathbb{N}$ (A002522). If c is considered as a C-finite sequence of order 3, then neither Algorithm 1e nor Algorithm 2e

Input : D-finite sequence c of order r
output: n_0 such that $\sigma^{n_0}(c)$ is positive
$n, n_0 \leftarrow 0$
$d \leftarrow c$
$\Psi(\xi, \mu) \leftarrow$ quantifier free formula equivalent to $\Psi(\xi, \mu, d)$
for $n = 0, 1, \ldots$ **do**
  if $d(n) \leq 0$ **then**
    $n_0 \leftarrow n_0 + n + 1$
    $d \leftarrow \sigma^{n+1}(d)$
    $\Psi(\xi, \mu) \leftarrow$ quantifier free formula equivalent to $\Psi(\xi, \mu, d)$
    $n \leftarrow 0$
  else if $\exists \mu > 0\colon \bigwedge_{i=0}^{r-2} d(n+i+1) \geq \mu d(n+i) \wedge \Psi(n, \mu)$ **then**
    return n_0
end

Algorithm 2e: Adjusted version of Algorithm 2 from [KP10]

terminate in 60 seconds. If c is, however, considered as a D-finite sequence of order 1 and degree 2, then both algorithms terminate and show that c is indeed positive.

The next lemma shows that we can find a shorter D-finite recurrence of a C-finite sequence c if and only if c has eigenvalues of higher multiplicities or equivalently the characteristic polynomial $\chi(c) \in \mathbb{K}[y]$ of c is not squarefree.

Lemma 8.6. *Let c be a C-finite sequence of order r with $y \nmid \chi(c)$. Then, c is D-finite of order $m < r$ if and only if $\chi(c)$ is not squarefree.*

Proof. Suppose c is given as in Theorem 2.15.

$\Longleftarrow$: The sequences $p_i(n)\lambda_i^n$ are D-finite of order 1 and degree d_i over the algebraic closure $\overline{\mathbb{K}}$. Hence, by Theorem 2.6, $c(n)$ is D-finite of order at most m over $\overline{\mathbb{K}}$. Lemma 2 in [Ger05] shows that the sequence is then also D-finite over $\mathbb{K}$ with the same order and degree. In particular, if $\chi(c)$ is not squarefree, then $r = \sum_{i=1}^{m} d_i > m$.

$\Longrightarrow$: Suppose that c satisfies a D-finite recurrence of order $m < r$ and degree d

$$\sum_{i=0}^{m} p_i(n) c(n+i) = 0 \quad \text{for all } n \in \mathbb{N} \tag{8.2}$$

with $p_i(n) = \sum_{k=0}^{d} p_{i,k} n^k$ where not all $p_{i,k}$ are zero. Furthermore, suppose that c is C-finite of order r with pairwise distinct eigenvalues $\lambda_1, \dots, \lambda_r \in \overline{\mathbb{K}}$, i.e., $c(n)$ can be written as $c(n) = \sum_{j=1}^{r} \gamma_j \lambda_j^n$ for some $\gamma_j \in \overline{\mathbb{K}}$. Using this closed form in (8.2) yields

$$\sum_{k=0}^{d} \left(\sum_{i=0}^{m} \sum_{j=1}^{r} p_{i,k} \gamma_j \lambda_j^{n+i} \right) n^k = 0. \tag{8.3}$$

Let $\gamma_{k,j} := \sum_{i=0}^{m} p_{i,k} \gamma_j \lambda_j^i$, then (8.3) is equivalent to $\sum_{k=0}^{d} \left(\sum_{j=1}^{r} \gamma_{k,j} \lambda_j^n \right) n^k = 0$. Evaluating $n = 0, \dots, r(d+1) - 1$ yields a homogeneous linear system for the $\gamma_{k,j}$. The corresponding matrix is regular (cf. Theorem 2.2.1 in [Li06] or Proposition 2.11 in [HHHK05]), so $\gamma_{k,j} = 0$ for all k, j. Let k be such that $p_{i,k} \neq 0$ for some i. Then,

$$0 = \sum_{j=1}^{r} \lambda_j^n \sum_{i=0}^{m} p_{i,k} \gamma_j \lambda_j^i = \sum_{i=0}^{m} \sum_{j=1}^{r} p_{i,k} \gamma_j \lambda_j^{n+i} = \sum_{i=0}^{m} p_{i,k} c(n+i).$$

Hence, c satisfies a C-finite recurrence of order $m < r$, a contradiction to c being C-finite of order r. $\square$

The proof of Lemma 8.6 shows that precisely the polynomial factors can be reduced in the D-finite recurrence, i.e., the m in the statement of Lemma 8.6 is the number of distinct eigenvalues of the sequence, which is also denoted by m in Theorem 2.15. The degree of the D-finite recurrence can be bounded by

$$(m(m+1) - m) \max_{i=1,\dots,m} d_i = m^2 \max_{i=1,\dots,m} d_i \leq r^3$$

using Theorem 2 in [Kau14].

In practice, we can easily check whether $\chi(c)$ is squarefree by checking whether $\chi(c)$ and its derivative are coprime. The shorter D-finite recurrence can then be either found by guessing or by computing it explicitly from the closed form of c.

8.1.4 Classical algorithm for sequences with unique dominant eigenvalue

If a C-finite sequence has a unique dominant eigenvalue, checking positivity of the sequence is known to be decidable. More details on the concrete time complexity are given in [OW14b]. Based on these results, we give a full description of such an algorithm in this section which is readily implemented. A similar description is given in [Kou05].

In the introduction of this chapter we have seen that a C-finite sequence c can be assumed to be in its closed form representation as

$$c(n) = p(n) + r(n) \tag{8.4}$$

where $p \in \mathbb{A}[x]$ with $\operatorname{lc}(p) > 0$ and $r(n) = \sum_{i=1}^{m} p_i(n)\lambda_i^n$ with $p_i \in \overline{\mathbb{Q}}[x], \lambda_i \in \overline{\mathbb{Q}}$ and $1 > |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_m|$. The idea is now to compute an $\varepsilon \in (0,1)$ and $n_0, n_1 \in \mathbb{N}$ such that $|r(n)| < (1-\varepsilon)^n$ for $n \geq n_0$ and $p(n) \geq (1-\varepsilon)^n$ for $n \geq n_1$. Then, clearly $c(n)$ is positive from $\max(n_0, n_1)$ on. The initial values can be checked separately again.

Input : C-finite sequence c of the form (8.4)
output: true if $c(n) > 0$ for all $n \in \mathbb{N}$ and false otherwise
$\varepsilon \leftarrow \frac{1-|\lambda_1|}{2}$
compute n_0 such that $|r(n)| < (1-\varepsilon)^n$ for all $n \geq n_0$
compute n_1 such that $p(n) \geq (1-\varepsilon)^n$ for all $n \geq n_1$
if $c(n) > 0$ *for* $n = 0, \ldots, \max(n_0, n_1)$ **then**
  return *true*
else
  return *false*
end

Algorithm C: Positivity for sequences with dominant eigenvalues [OW14b]

For a polynomial $p_i(x) = \sum_{j=0}^{d_i} \gamma_{i,j} x^j \in \mathbb{A}$ of degree d_i we can easily compute a constant $c_i \in \mathbb{A}$ such that $|p_i(n)| \leq c_i n^{d_i}$ for all $n \geq 1$. For example, we can choose $c_i := \sum_{j=0}^{d_i} |\gamma_{i,j}|$. Let $c := \sum_{i=1}^{m} c_i$ and $d := \max(d_1, \ldots, d_m)$, i.e., the maximal multiplicity of the eigenvalues $\lambda_1, \ldots, \lambda_m$. Furthermore, let $\varepsilon := \frac{1-|\lambda_1|}{2}$. Then, $1 - \varepsilon = |\lambda_1| + \varepsilon$.

First, we show how n_0 can be found such that $|r(n)| < (1-\varepsilon)^n$ for $n \geq n_0$. Let $\mu := \frac{|\lambda_1|+\varepsilon}{|\lambda_1|}$. If $d = 0$, then $|r(n)| \leq c\,|\lambda_1|^n$. Clearly,

$$c\,|\lambda_1|^n < (1-\varepsilon)^n = (|\lambda_1| + \varepsilon)^n \iff \frac{\log(c)}{\log(\mu)} < n.$$

Hence, we can choose $n_0 := \left\lceil \frac{\log(c)}{\log(\mu)} \right\rceil$ in this case. If $d > 0$, then $|r(n)| \leq c\,n^d\,|\lambda_1|^n$. Again,

$$c\,n^d\,|\lambda_1|^n < (1-\varepsilon)^n \iff \log(c^{1/d}) < \frac{n}{d}\log(\mu) - \log(n).$$

The derivative of the right-hand side of the latter inequality is positive if $n > \frac{d}{\log(\mu)}$, i.e., from $\left\lceil \frac{d}{\log(\mu)} \right\rceil$ on the sequence on the right-hand side is monotonously increasing. Hence, if the inequality is true for some $n_0 \geq \left\lceil \frac{d}{\log(\mu)} \right\rceil$, then it is true for all $n \geq n_0$. Checking these values one by one, we find a suitable n_0 eventually.

If the polynomial $p(x) = p_0$ is just constant, then $p(n) \geq (1-\varepsilon)^n$ if and only if $n \geq \frac{\log(p_0)}{\log(1-\varepsilon)}$. Otherwise, we can compute the largest real root x_1 of the derivative of $p(x)$. If $p(n_1) \geq (1-\varepsilon)^{n_1}$ for any $n_1 \geq \lceil x_1 \rceil$, then the inequality holds for all $n \geq n_1$.

Note that once we have established that a sequence has a unique dominant eigenvalue, all these computations can be done using arbitrary precision arithmetic. For our implementation in SageMath we make use of the `Arb` library [Joh17].

Example 8.7. The sequence from Example 8.2 has a unique dominant eigenvalue. Hence, Algorithm C shows positivity of the sequence after checking $\max(n_0, n_1) = 12$ initial terms.

Example 8.8. The sequence from Example 8.1 has dominant eigenvalues ± 1. Hence, as the sequence does not have a unique dominant eigenvalue, Algorithm C cannot establish positivity of the sequence.

8.1.5 Combination of Algorithm 1 and Algorithm 2

In the case that the C-finite sequence has a unique dominant eigenvalue, we can combine the closed form representation of the sequence together with Algorithm 1e and Algorithm 2e. As we know that the polynomial term $p(n)$ in (8.1) certainly dominates the exponential terms, we can find indices n_i using Algorithm 1e and Algorithm 2e from which on the exponential sequences are dominated by the polynomial term. These input sequences have very low order (maximum order 3). Therefore, the termination criteria in [KP10] show that these algorithms terminate in most instances.

Before we can prove termination criteria for Algorithm P, we prove some auxiliary results on the characteristic polynomial of D-finite sequences.

Input : C-finite sequence c of the form (8.1)
output: true if $c(n) > 0$ for all $n \in \mathbb{N}$ and false otherwise
for $i \leftarrow 1$ **to** s **do**
  $n_{i,\overline{\mathbb{Q}}} \leftarrow$ Algorithm 1e applied to $\frac{p(n)}{s+l} + o_i(n)\xi_i^n + \overline{o_i}(n)\overline{\xi_i}^n$
end
for $i \leftarrow 1$ **to** l **do**
  $n_{i,\mathbb{A}} \leftarrow$ Algorithm 2e applied to $\frac{p(n)}{s+l} + q_i(n)\rho_i^n$
end
$n_0 \leftarrow \max(n_{1,\overline{\mathbb{Q}}}, \dots, n_{s,\overline{\mathbb{Q}}}, n_{1,\mathbb{A}}, \dots, n_{l,\mathbb{A}})$
if $c(n) > 0$ *for* $n = 0, \dots, n_0$ **then**
  return *true*
else
  return *false*
end

Algorithm P: Positivity for sequences with dominant eigenvalues

First, we extend the notion of the characteristic polynomial from the ring $\mathbb{K}[n][\sigma]$ to the left Euclidean domain $\mathbb{K}(n)[\sigma]$. For a rational function $\frac{p(n)}{q(n)}$ with coprime $p, q \in \mathbb{K}[n]$ we define the *degree* as $\deg(p/q) := \deg(p) - \deg(q)$ and call

$$\operatorname{lc}(p/q) := \operatorname{coeff}(p/q, \deg(p/q)) := \operatorname{lc}(p)/\operatorname{lc}(q)$$

the leading coefficient of p/q. Now, for an operator $\mathcal{A} = \sum_{i=0}^{r} \frac{p_i(n)}{q_i(n)}\sigma^i \in \mathbb{K}(n)[\sigma]$ with $\deg(\mathcal{A}) := \max_{i=0,\dots,r} \deg(p_i/q_i)$ we define the characteristic polynomial as

$$\chi(\mathcal{A}) := \sum_{\substack{i=0 \\ \deg(p_i/q_i)=\deg(\mathcal{A})}}^{r} \operatorname{lc}(p_i/q_i)y^i \in \mathbb{K}[y].$$

If $\mathcal{A} \in \mathbb{K}[n][\sigma]$, i.e., if all q_i are constants, then this definition is identical to the original definition (2.1).

Next, in Lemma 8.9 and Lemma 8.10, we state some basic properties of the characteristic polynomial. Since we could not find references for those, we add the proofs for the sake of completeness.

Lemma 8.9. *Let $\mathcal{A}, \mathcal{B} \in \mathbb{K}(n)[\sigma]$. Then $\chi(\mathcal{AB}) = \chi(\mathcal{A})\chi(\mathcal{B})$.*

Proof. Let $\mathcal{A} := \sum_{i=0}^{r} p_i(n)\sigma^i \in \mathbb{K}(n)[\sigma]$ and $\mathcal{B} := \sum_{j=0}^{s} q_j(n)\sigma^j \in \mathbb{K}(n)[\sigma]$ and $d_{\mathcal{A}} := \max_{i=0,\dots,r} \deg(p_i), d_{\mathcal{B}} := \max_{j=0,\dots,s} \deg(q_j) \in \mathbb{Z}$ their respective degrees. We show that $\mathcal{AB}$ has degree $d_{\mathcal{A}} + d_{\mathcal{B}}$. By the definition of multiplication in $\mathbb{K}(n)[\sigma]$ and the properties of the degree of a rational function, the degree of $\mathcal{AB}$ is certainly bounded by $d_{\mathcal{A}} + d_{\mathcal{B}}$. Let i', j' be maximal such that $\deg(p_{i'}) = d_{\mathcal{A}}$ and $\deg(q_{j'}) = d_{\mathcal{B}}$. We show that the coefficient of $\sigma^{i'+j'}$ of $\mathcal{AB}$ has degree $d_{\mathcal{A}} + d_{\mathcal{B}}$. This coefficient is given by $\sum_{l=0}^{i'+j'} p_l(n) q_{i'+j'-l}(n+l)$. Because of the choices of i', j' we have

$$\deg(p_l(n)q_{i'+j'-l}(n)) = \deg(p_l(n)) + \deg(q_{i'+j'-l}(n+l)) < d_{\mathcal{A}} + d_{\mathcal{B}}$$

for all $l \neq i'$. For $l = i'$, we have $\deg(p_l(n)q_{i'+j'-l}(n)) = d_{\mathcal{A}} + d_{\mathcal{B}}$, so by the properties of the degree we have

$$\begin{aligned} \deg\left(\sum_{l=0}^{i'+j'} p_l(n)q_{i'+j'-l}(n+l)\right) &= \max_{l=0,\dots,i'+j'} \left(\deg\left(p_l(n)\right) + \deg\left(q_{i'+j'-l}(n+l)\right)\right) \\ &= d_{\mathcal{A}} + d_{\mathcal{B}}. \end{aligned}$$

Next, we show that all coefficients of $\chi(\mathcal{A})\chi(\mathcal{B})$ and $\chi(\mathcal{AB})$ agree. Let $i \in \{0, \dots, r+s\}$. Then,

$$\operatorname{coeff}\left(\chi(\mathcal{A}), i\right) = \operatorname{coeff}\left(p_i(n), d_{\mathcal{A}}\right), \quad \operatorname{coeff}\left(\chi(\mathcal{B}), i\right) = \operatorname{coeff}\left(q_i(n), d_{\mathcal{B}}\right)$$

and therefore

$$\operatorname{coeff}\left(\chi(\mathcal{A})\chi(\mathcal{B}), i\right) = \sum_{j=0}^{i} \operatorname{coeff}\left(p_j(n), d_{\mathcal{A}}\right) \operatorname{coeff}\left(q_{i-j}(n), d_{\mathcal{B}}\right).$$

In the first part of the proof we have shown that $\mathcal{A}\mathcal{B}$ has degree $d_{\mathcal{A}} + d_{\mathcal{B}}$. Therefore,

$$\begin{aligned}
\operatorname{coeff}\left(\chi(\mathcal{A}\mathcal{B}), i\right) &= \operatorname{coeff}\left(\sum_{j=0}^{i} p_j(n) q_{i-j}(n+j), d_{\mathcal{A}} + d_{\mathcal{B}}\right) \\
&= \sum_{j=0}^{i} \operatorname{coeff}\left(p_j(n) q_{i-j}(n+j), d_{\mathcal{A}} + d_{\mathcal{B}}\right) \\
&= \sum_{j=0}^{i} \operatorname{coeff}\left(p_j(n), d_{\mathcal{A}}\right) \operatorname{coeff}\left(q_{i-j}(n+j), d_{\mathcal{B}}\right) \\
&= \sum_{j=0}^{i} \operatorname{coeff}\left(p_j(n), d_{\mathcal{A}}\right) \operatorname{coeff}\left(q_{i-j}(n), d_{\mathcal{B}}\right).
\end{aligned}$$

□

Suppose $\mathcal{A}$ is an annihilator of a and $\mathcal{B}$ an annihilator of b. Then, the *least common left multiple* $\operatorname{lclm}(\mathcal{A}, \mathcal{B})$ is an annihilator of $a + b$ [Kau15].

Lemma 8.10. *Let $\mathcal{A}, \mathcal{B} \in \mathbb{K}[n][\sigma]$. Then*

$$\chi(\mathcal{A}) \mid \chi(\operatorname{lclm}(\mathcal{A}, \mathcal{B})) \text{ and } \chi(\mathcal{B}) \mid \chi(\operatorname{lclm}(\mathcal{A}, \mathcal{B})).$$

In particular, we have

$$\operatorname{lcm}(\chi(\mathcal{A}), \chi(\mathcal{B})) \mid \chi(\operatorname{lclm}(\mathcal{A}, \mathcal{B})).$$

Proof. Let $\mathcal{C} \in \mathbb{K}(n)[\sigma]$ be such that $\mathcal{C}\mathcal{A} = \operatorname{lclm}(\mathcal{A}, \mathcal{B})$. Then, with Lemma 8.9 we have

$$\chi(\operatorname{lclm}(\mathcal{A}, \mathcal{B})) = \chi(\mathcal{C}\mathcal{A}) = \chi(\mathcal{C})\chi(\mathcal{A}).$$

□

Example 8.11. In Lemma 8.10, divisibility cannot be replaced with equality. Consider $\mathcal{A} := 1 + \sigma$ and $\mathcal{B} := n + (n+1)\sigma$. Then,

$$\chi(\mathcal{A}) = \chi(\mathcal{B}) = 1 + y,$$

but

$$\chi(\mathrm{lclm}(\mathcal{A},\mathcal{B})) = \chi(n + (2n+2)\sigma + (n+2)\sigma^2) = 1 + 2y + y^2.$$

An operator $\mathcal{A} = \sum_{i=0}^{r} p_i\sigma^i \in \mathbb{K}[n][\sigma]$ is called *balanced* if

$$\deg p_0 = \deg p_r = \max_{i=0,\dots,r} \deg p_i.$$

Equivalently, $\mathcal{A}$ is balanced if and only if the degree of $\chi(\mathcal{A}) \in \mathbb{K}[y]$ equals the order of $\mathcal{A}$ and the trailing coefficient of $\chi(\mathcal{A})$ is nonzero, i.e., $y \nmid \chi(\mathcal{A})$.

As Algorithm 2e terminates for essentially all sequences of order 2, the real algebraic part of Algorithm P certainly terminates.

Theorem 8.12. *Algorithm P terminates if $s = 0$, i.e., if all eigenvalues of c are real algebraic.*

Proof. Each sequence $h(n) := \frac{p(n)}{s+l} + q_i(n)\rho_i^n$ is the sum of two balanced D-finite sequences g, f over $\mathbb{A}$ satisfying the recurrences

$$-p(n+1)g(n) + p(n)g(n+1) = 0, \quad -q_i(n+1)\rho_i f(n) + q_i(n)f(n+1) = 0$$

with characteristic polynomials

$$\chi(\mathcal{G}) = \mathrm{lc}(p)(y-1), \quad \chi(\mathcal{F}) = \mathrm{lc}(q_i)(y-\rho_i)$$

where $\mathcal{G}, \mathcal{F}$ denote the annihilating operators of g, f, respectively. As these characteristic polynomials are coprime, Lemma 8.10 yields

$$\chi(\mathcal{H}) = \chi(\mathcal{G})\chi(\mathcal{F}) = \gamma(y-1)(y-\rho_i)$$

for some constant γ where $\mathcal{H}$ denotes the annihilating operator of h. In particular, $\mathcal{H}$ is balanced. Furthermore, $h \sim p(n)$ by construction. With [KP10, Theorem 3], Algorithm 2e terminates with input h. $\square$

It is conjectured that Algorithm 1e terminates for sequences of order 3 if the eigenvalues are complex. This is the case if we apply Algorithm 1e. Hence, if the conjecture is true, Algorithm P terminates for all C-finite sequences with a unique dominant eigenvalue.

Theorem 8.13. *Assume Conjecture 1 from [KP10] is true. Then, Algorithm P terminates.*

Proof. The proof of Theorem 8.12 already shows that the algorithm terminates for the real algebraic eigenvalues. Analogously, in the complex case, the sequences $h(n) := \frac{p(n)}{s+l} + o_i(n)\xi_i^n + \overline{o_i}(n)\overline{\xi_i}^n$ are D-finite of order 3 with a balanced annihilating operator $\mathcal{H}$ with characteristic polynomial

$$\chi(\mathcal{H}) = \gamma(y-1)(y-\xi_i)(y-\overline{\xi_i})$$

for some constant γ. With Conjecture 1, Algorithm 1e terminates on this input. □

Example 8.14. The sequence A002248 is C-finite of order 4 satisfying the recurrence

$$4c(n) - 8c(n+1) + 7c(n+2) - 4c(n+3) + c(n+4) = 0$$

with initial values $c = \langle 2, 8, 14, 16, \dots \rangle$. The sequence has the unique dominant eigenvalue 2. Neither Algorithm 1e nor Algorithm 2e terminate in 60 seconds. However, both Algorithm C and Algorithm P terminate in negligible time.

8.1.6 Decomposition into nondegenerate sequences

Theorem 2.17 states that every C-finite sequence $c(n)$ can be written as the interlacing of nondegenerate sequences

$$c_1(n) := c(dn), \dots, c_d(n) := c(dn + d - 1).$$

Reducing the Positivity Problem for c to the Positivity Problem for the subsequences c_k for $k = 1, \dots, d$ often turned out useful [MST84, Ver85, OW14b]. For a given C-finite sequence c we can check whether they are degenerate by computing the ratio of all pairs of eigenvalues and checking whether they are a root of unity [Coh13]. Hence, we can compute the decomposition of c into nondegenerate sequences naively by computing the sequences $c_1, \dots, c_d$ and checking whether all these are nondegenerate. If they are not, we can increase d. Eventually, for large enough d, all subsequences are nondegenerate. This already works well in practice. A more efficient algorithm is given in [YLN95].

If decomposition into subsequences is used together with Algorithm C or Algorithm P, then it is more efficient to check whether every subsequence has a unique dominant root (which can be done numerically with arbitrary-precision arithmetic) instead for checking degeneracy. The main bottleneck is usually the computation of the subsequences. Hence, an efficient implementation should certainly aim to minimize these. Sequences of natural numbers which can be decomposed into subsequences with a unique dominant eigenvalue are $\mathbb{N}$-rational [Kou05, Theorem 2.5.12]. It was already observed that they seem to cover most C-finite sequences appearing in practical examples [Kou05].

Example 8.15. The sequence A000115 is C-finite of order 8 and satisfies the recurrence

$$\begin{aligned} c(n) - c(n+1) - c(n+2) &+ c(n+3) \\ -c(n+5) + c(n+6) + c(n+7) &- c(n+8) = 0. \end{aligned}$$

with initial values $c = \langle 1, 1, 2, 2, 3, 4, 5, 6, \dots \rangle$. It has 6 dominant eigenvalues and is degenerate. It can be decomposed into 10 nondegenerate sequences with unique dominant eigenvalues. For these subsequences Algorithm C and Algorithm P both have no problem showing positivity.

Example 8.16. The Berstel sequence A007420 is C-finite of order 3 satisfying

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

with $c(0) = c(1) = 0, c(2) = 1$. Checking the initial values of the C-finite sequence $d(n) := c(n+53)^2$ indicates that $d(n)$ is positive and that the only zeros of c are at the indices $n = 0, 1, 4, 6, 13, 52$ (this is, in fact, the maximal number of zeros a nondegenerate sequence of order 3 can have and therefore these already have to be all zeros [Beu91]). The sequence $d(n)$ is nondegenerate and does not have a unique dominant root. In fact, Ge's algorithm applied to the eigenvalues of c shows that there are no relations among them. Hence, we cannot expect this algorithm to work for the sequence $d(n)$.

8.2 Comparison

For comparing the different algorithms discussed in the previous section, we consider C-finite sequences from the OEIS. First, we discuss how we can use guessing to determine how many of the around 360 000 sequences are C-finite or D-finite. Next, we use 1000 of these C-finite sequences (where the first terms are positive) as a test set for comparing the various algorithms and testing the implementation.

8.2.1 Recurrence sequences in the OEIS

In a talk in 2003 Bruno Salvy estimated that about 25% of the 5488 sequences from the book *The Encyclopedia of Integer Sequences* (cf. [SP95]) are D-finite [Sal03] (in his master thesis 1991 Simon Plouffe, using guessing, estimated this number to be around 18% for a preliminary draft of the book). These 25% have been cited as an estimate for the ratio of D-finite sequences in the OEIS several times over the past two decades [Sal05, Kau13, Yur22]. To our knowledge, no estimate for the ratio of C-finite sequences in the OEIS is known. At the time of writing (spring 2023) the OEIS contains about 360 000 integer sequences. Due to this large number of sequences, only estimates for these ratios can be found. This can be done, for instance, by either inspecting a smaller subset of these sequences closer by hand or by using guessing routines on the terms saved in the database. We use the latter approach.

Guessing routines are limited by the number of terms which are known for a particular sequence. The more terms we have, the bigger recurrences we can guess or the more confident we can be that a guessed recurrence is indeed valid. The number of terms given in the OEIS vary widely. About 6% of the sequences have at most 10 terms, about 50% at most 100 terms and about 13% of the sequences have at least 10 000 terms given (note, however, that these terms are only given in the corresponding "B-files" and not displayed in the database itself). Figure 8.1 gives an overview of the number of terms of sequences which are given in the OEIS.

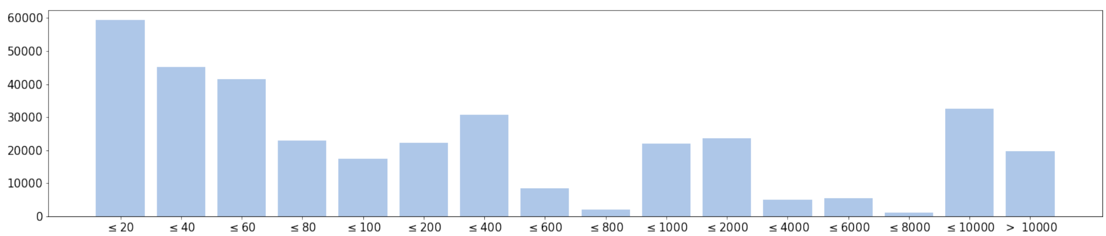
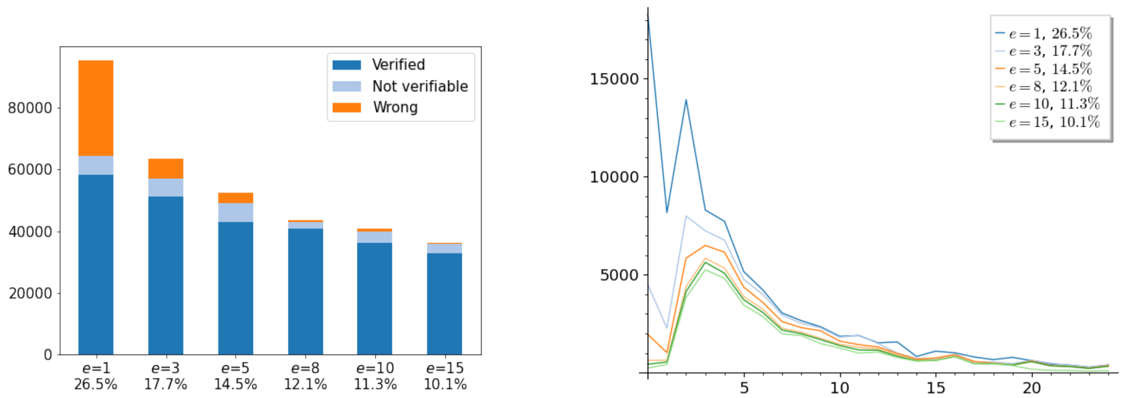


Figure 8.1: Number of sequences for which specific number of terms are given in the OEIS

For guessing C-finite recurrences, we fix a maximal order of 100, i.e., sequences which might satisfy a higher order recurrence are not considered C-finite. Guessing a D-finite recurrence of order r and degree d yields a linear system of equations with $(r+1)(d+1)$ many variables. In order to have some confidence in the guess, we have to make sure that the linear system is overdetermined. The more equations we have, the more confident we can be in our guess. We make sure that corresponding linear systems have

$$\lfloor (r+1)(d+1)(1+\tfrac{e}{2}) \rfloor + e$$

many variables for some $e \in \mathbb{N}$. I.e., the number of equations needed depends also relatively on the order/degree of the recurrence. The number e can be interpreted as the confidence level of our guess. However, if e is larger it might be that we do not have enough terms given to verify a recurrence that could be guessed with smaller e. This can also be seen in Figure 8.2a which shows the results for C-finite sequences. For instance, going from $e = 1$ to $e = 3$ we can see that many recurrences are recognized to be wrong (orange part in the figure). For larger e only a small number of recurrences are recognized as wrong, so we can assume that most of these recurrences are indeed correct. However, the figure also shows that by increasing e we might fail to verify a recurrence simply due to a lack of terms given in the OEIS (light blue parts in the figure). We can estimate that close to 15% of the sequences in the OEIS might be C-finite.



(a) (b)

Figure 8.2: (8.2a): Number of C-finite sequences in the OEIS according to different confidence levels. The percentages indicate the ratio of C-finite sequences in the OEIS according to the given confidence level
(8.2b): Number of C-finite sequences of specific orders in the OEIS

A sanity check can also be done using the OEIS Wiki page which gives an overview of the sequences which were already recognized to be C-finite.[1] For $e = 1$ we miss about 800 sequences from the Wiki page, about 33 600 sequences are considered C-finite both by the Wiki page as well as by guessing and about 61 600 additional (hypothetical) sequences were found. For $e = 5$ these numbers are about 5700, 28 800 and 23 600, respectively.

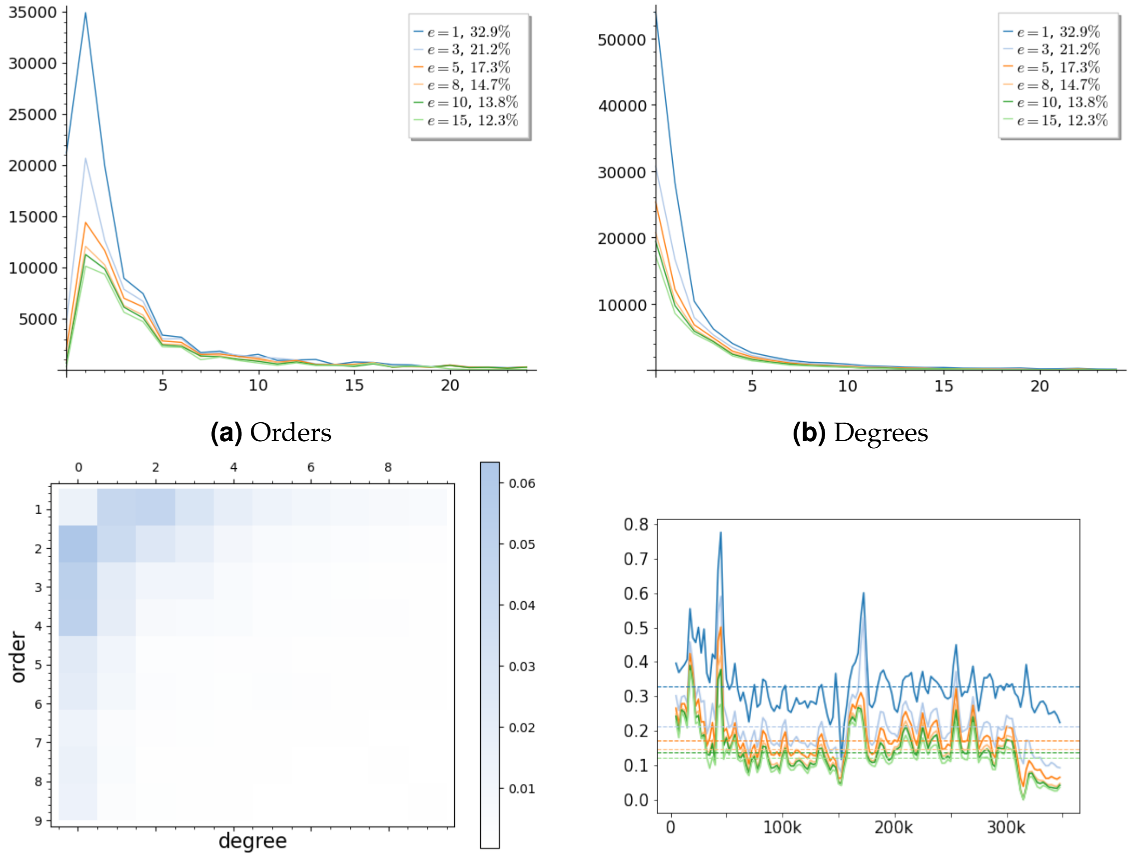
We can do a closer investigation of the recurrences that we guessed. For instance (for the confidence level $e = 5$), around one half of the C-finite sequences have eigenvalues of higher multiplicity. I.e., by Theorem 8.6 these are precisely the sequences which have a shorter (in terms of the order) D-finite recurrence. Almost 20% of the C-finite sequences are just polynomial sequences. The orders of the sequences are shown in Figure 8.2b. As can be expected, most sequences which are guessed have relatively small order (for $e = 5$ about 70% of the C-finite recurrences have orders at most 10 and only about 6% have order larger than 30).

For guessing D-finite recurrences we have to be a bit more careful as zero terms of a sequence might yield wrong guesses. We use the techniques from [KV19] to mitigate this problem. Recurrences for a D-finite sequence can be found on the so-called order-degree curve, i.e., the possible minimal values (r, d) for the order r and degree d of a recurrence lie on a hyperbola [Kau14]. Often, the minimal order recurrence, which is the one we are looking for, has large degree. For our guessing approach we search for the operator with $(r + 1)(d + 1) \leq 100$. The results can be observed in Figure 8.3. According to these, we can estimate that up to 20% of the sequences in the OEIS might be D-finite. For $e = 3$, about one third of the D-finite recurrences we guessed are hypergeometric, i.e., have order one. Close to 40% have degree zero, i.e., they are C-finite with eigenvalues having only multiplicities one. As we estimated before that around half of the C-finite sequences have eigenvalues of higher multiplicities we can guess that around 80% of the D-finite sequences in the OEIS are in fact C-finite. This agrees with our estimates that around 20% are D-finite and 15% are C-finite.

Plotting the ratio of D-finite sequences in the OEIS w.r.t. the OEIS identifier one can see that the ratio dropped slightly over the past twenty years of the database's existence (cf. Figure 8.3d). The first spike around sequence A042700 is due to a series of sequences related to continued fractions of certain algebraic numbers. The spike around the sequence A170000 is caused by a series of C-finite sequences counting the number of words in

[1] `https://oeis.org/wiki/Index_to_OEIS:_Section_Rec`

certain Coxeter groups (A168680 to A170731). For many of these only few terms are saved in the OEIS which explains the significant drop for larger confidence levels. Around the sequence A149000, the ratio drops significantly (especially for $e = 1$). These sequences describe lattice walks and most of them seem to not be D-finite [BK08].



(a) Orders

(b) Degrees

(c) Order-degrees density for $e = 5$

(d) Ratio of D-finite sequences in the OEIS

Figure 8.3: Number of D-finite sequences of specific orders and degrees in the OEIS in (8.3a) and (8.3b). Combined density plot of orders/degrees for $e = 5$ in (8.3c) and density of D-finite sequences in the OEIS w.r.t. their index in (8.3d).

8.2.2 Positive sequences in the OEIS

From the sequences where a C-finite recurrence could be guessed we take the first 1000 where the first 500 terms are strictly positive and are therefore highly likely to be positive altogether. [2]

The maximal order of these sequences is 42. The following table shows the number of sequences of each given order:

order	1	2	3	4	5	6	7	8	9	10	> 10
	73	134	117	139	120	80	87	36	47	27	140

More than half of these sequences, 567, have a unique dominant eigenvalue. There are 102, 40, 70, 32 sequences with $2, 3, 4, 5$ distinct dominant eigenvalues, respectively. Hence, there are 139 sequences with more than 6 distinct dominant eigenvalues. About half of the sequences, 513, have a characteristic polynomial which is not squarefree. By Lemma 8.6 these are the sequences which have a shorter D-finite recurrence.

We test the positivity methods implemented in the `rec_sequences` package on these sequences. SageMath provides an interface to `QEPCADB` which allows CAD computations [Sag23, Bro03]. This is used in the implementations of Algorithm 1 and Algorithm 2. For Algorithm C we rely on fast arbitrary-precision arithmetic using the library `Arb` which is included in SageMath [Joh17]. To decompose a sequence into subsequences with a unique dominant eigenvalue, we decompose the sequence into k subsequences and check, using arbitrary-precision arithmetic, whether all of these have a unique dominant eigenvalue. If they do not have a unique dominant eigenvalue, we increase k by one. The main bottleneck when decomposing is by far the computation of the subsequences. Checking whether a subsequence has a unique dominant eigenvalue or proving positivity of a sequence with a unique dominant eigenvalue using Algorithm C only takes negligible time in our examples.

[2] A table with these sequences and additional information is given on the website `https://www3.risc.jku.at/people/pnuspl/PositivityCFinite`. It also contains the detailed results of the SageMath (using `https://github.com/PhilippNuspl/rec_sequences/releases/tag/v0.1`) and Mathematica tests. The SageMath results in this thesis are using a more recent version of the package and are therefore slightly different.

We give a list of the methods that can be used on C-finite sequences to show positivity. Every method has a parameter `strict` which is `True` by default and indicates whether strict positivity or nonnegativity should be shown. The additional parameter `time` can be used to give an upper bound (in seconds) after which the algorithms should be terminated, the default value is -1, indicating that they should not stop prematurely.

- `is_positive_algo1` implements Algorithm 1 from [KP10]. As an additional parameter `bound` can be specified which gives an upper bound on the number of iterations.
- `is_positive_algo2` implements Algorithm 2 from [KP10]. Again, `bound` can be specified. This method is also implemented for general D-finite sequences and can be called using `is_positive` on D-finite sequences.
- `is_positive_dominant_root` implements Algorithm C for sequences with a unique dominant eigenvalue.
- `is_positive_dominant_root_decompose` first tries to decompose the sequence into sequences with a unique dominant eigenvalue and zero sequences and calls Algorithm C on each of those.

Using these methods, all of the 1000 sequences from the test set could be proven to be positive using a time limit of 60 seconds. The following table gives an overview of the number of sequences which could be proven to be positive by each method (a "D" indicates that decomposition of the sequence is used):

Algo. 1	D, Algo.1	Algo. 2	D, Algo.2	Algo. C	D, Algo. C
384	375	327	556	566	1000

It is clear that decomposing the sequences and using Algorithm C is the most powerful method and it can prove positivity of every single sequence in the test set. The implementation of Algorithm C is very fast and takes at most 0.3 seconds for every example we considered.

Similar experiments were done using a Mathematica implementation of some of the methods (more details in [NP22a]). Compared to the SageMath implementation, Algorithm C in the Mathematica implementation is significantly slower as it relies much more on exact computations. Due to the powerful quantifier elimination methods in Mathematica, Algorithm P can be tested as well. The results show that for the sequences in discussion the method is similarly powerful as Algorithm C.

Clearly, the implemented algorithms are already very powerful and can prove positivity of most C-finite sequences arising in practice. The situation for D-finite sequences is much bleaker. It would certainly be interesting to implement some of the algorithms mentioned in the introduction of this chapter and see how well they work on practical examples.

Open Question 8.17. How do other methods for proving positivity of C-finite and D-finite sequences compare to the algorithms presented here? Are they more efficient? Can they prove positivity of a wider range of sequences?

9 Implementation

As a proof of concept and for practical computations, most of the algorithms presented in the previous chapters are implemented in the computer algebra system SageMath [Sag23] in the package `rec_sequences` (developed under version SageMath 9.4). For most of the basic operations with C-finite or D-finite sequences the package relies on the `ore_algebra` package [KJJ15]. For proving inequalities of C-finite and D-finite sequences it relies on `Arb` (cf. [Joh17]) for efficient arbitrary precision computations and on `QEPCADB` for quantifier elimination using CAD (cf. [Bro03]). The latter needs to be installed separately if inequality methods based on the Gerhold-Kauers method should be used. The package `rec_sequences` is also described in the extended abstract [Nus22] on which this chapter is based.

9.1 Installation

The package is published under the GPL-3.0 license. The source code and extensive documentation can be found on Github.[1] Several different methods can be used for installation. Simplest, if SageMath was built from the sources, the command

```
sage --pip install git+https://github.com/PhilippNuspl/rec_sequences.git
```

installs the package together with the `ore_algebra` package (other methods can be found in the Github readme file). For using the functionality based on CAD, we can install `QEPCADB` by

```
sage -i qepcad
```

[1] https://github.com/PhilippNuspl/rec_sequences

9.2 C-finite sequences

After the installation of the package it can be loaded in any SageMath session. A C-finite sequence ring over a field of characteristic zero `K` can be created by `CFiniteSequenceRing(K)`. A sequence in this ring `C` can now be defined by a list of the coefficients of the recurrence and initial values:

$$\texttt{C([}\gamma_0,\dots,\gamma_r\texttt{], [}c_0,\dots,c_{r-1}\texttt{])} \leftrightarrow \begin{cases} \gamma_0 c(n) + \cdots + \gamma_r c(n+r) = 0, \\ c(0) = c_0, \dots, c(r-1) = c_{r-1}. \end{cases}$$

Alternatively, a symbolic expression in one variable or a list of initial terms can be used to define a C-finite sequence. In both cases guessing is used to find a recurrence.

```
sage: from rec_sequences.CFiniteSequenceRing import *
sage: C = CFiniteSequenceRing(QQ)
sage: fib = C([1,1,-1], [0,1], name="f") # Fibonacci numbers
sage: var("n");
sage: exp2 = C(2^n)
sage: alt = C(10*[1,-1])
sage: alt
C-finite sequence a(n): (1)*a(n) + (1)*a(n+1) = 0 and a(0)=1
```

Terms of a C-finite sequence can be obtained in the same way that elements of lists are obtained in Python.

```
sage: exp2[3], fib[:10]
(8, [0, 1, 1, 2, 3, 5, 8, 13, 21, 34])
```

Closure properties of C-finite sequences are computed using the `ore_algebra` package. These include difference ring operations (using `+`, `*` and `shift`), partial sums (using `sum`), Cauchy product (using `cauchy`), interlacing (using `interlace`) and subsequences (using `subsequence`). Equality of two C-finite sequences can be checked as well. The latter of the following examples proves the identity presented in the introduction of the thesis.

```
sage: fib.sum() == fib.shift(2)-1
True
sage: (fib*fib.shift()).sum().subsequence(2) == fib.subsequence(2,1)^2-1
True
```

Furthermore, one can obtain the recurrence coefficients, the initial values and the characteristic polynomial of a C-finite sequence (using `coefficients`, `initial_values` and `charpoly`, respectively) or compute the closed form:

```
sage: (fib^2-fib.shift()*fib.shift(-1)).closed_form() # Cassini identity
-(-1)^n
```

For proving positivity of a sequence, one can either use the methods presented in Section 8.2.2 directly or use the operators $>, <, >=, <=$. If these operators are used the computations abort after a set amount of time. Hence, for proving positivity of more complicated sequences, it can be useful to use the method `is_positive` explicitly as shown for the sequence A000115 from Example 8.15:

```
sage: fib < exp2, 10 > fib, alt >= -1
(True, False, True)
sage: c = C([1,-1,-1,1,0,-1,1,1,-1], [1, 1, 2, 2, 3, 4, 5, 6]) # A000115
sage: c.is_positive()
True
```

By the Skolem-Mahler-Lech theorem, Theorem 2.18, the set of zeros of a C-finite sequence is a finite set together with finitely many arithmetic progressions. In many cases, these zeros can be computed using the package:

```
sage: (alt+1).zeros()
Zero pattern with finite set {} and arithmetic progressions:
- Arithmetic progression (2*n+1)_n
```

Often, the sign pattern of a sequence is eventually cyclic (this is, however, not necessarily the case as Example 2.3 in [AKK$^+$21] shows):

```
sage: alt.sign_pattern()
Sign pattern: cycle <+->
```

More information on any of the methods can be obtained using `?`, e.g. `fib.interlace?`. In many cases, more detailed information on the algorithms which are performed can be viewed via the Python logging module. For instance, using the following command all subsequent methods display also intermediate results:

```
sage: logging.basicConfig(stream=sys.stdout, level=logging.DEBUG)
```

9.3 C^2-finite sequences

Analogous to C-finite sequences, C^2-finite sequences can again be defined by the coefficients of the recurrence and initial values.

```
sage: from rec_sequences.C2FiniteSequenceRing import *
sage: C2 = C2FiniteSequenceRing(QQ)
sage: fibonorial = C2([fib.shift(), -1], [1])
sage: fibonorial # A003266, fibonorial[n]==prod(fib[k] for k in range(1,n
    +1))
C^2-finite sequence of order 1 and degree 2 with coefficients:
 > c0 (n) : C-finite sequence c0(n): (1)*c0(n) + (1)*c0(n+1) + (-1)*c0(n+2)
     = 0 and c0(0)=1 , c0(1)=1
 > c1 (n) : C-finite sequence c1(n)=-1
and initial values a(0)=1
sage: fibonorial[:10]
[1, 1, 1, 2, 6, 30, 240, 3120, 65520, 2227680]
```

If a sequence $c(n)$ is C-finite, then the sparse subsequence $c(n^2)$ is C^2-finite (cf. Theorem 4.7) and we can compute the C-finite coefficients of the recurrence verifying the recurrence from Example 3.7:

```
sage: sparse_fib = fib.sparse_subsequence(C2) # A054783
sage: sparse_fib[:10] == [fib[n^2] for n in range(10)]
True
sage: coeffs = [-fib.subsequence(2, 3), -fib.subsequence(4, 4),
....:           fib.subsequence(2, 1)]
sage: sparse_fib.coefficients() == coeffs
True
```

Ring operations of C^2-finite sequences can be performed in the same way as for C-finite sequences. The operations are reduced to linear systems of equations over the C-finite sequence ring. Sometimes, these systems can be solved by computing a termwise solution first, using guessing to find a C-finite solution and verifying this solution. When defining a C^2-finite sequence ring it can be specified explicitly that this approach should be tried. This is demonstrated on Example 4.3:

```
sage: C2_guess = C2FiniteSequenceRing(QQ, guess=True)
sage: b = C2_guess([fib.shift(), -fib.shift(2)], [1/fib[0+1]])
sage: a = b*fib
sage: c = C2_guess([fib.shift()*fib.shift(2), -fib.shift(2)*fib.shift(3)],
```

```
....:                [1/(fib[0+1]*fib[0+2])])
sage: d = (c*alt.shift()).sum().prepend([0])
sage: c+d == 0
True
```

We can also compute a recurrence for $\sum_{k=0}^{\lfloor n/3 \rfloor} f((2k+1)^2)$ [JPNP23, Example 5.1]:

```
sage: h = fib.sparse_subsequence(C2).subsequence(2,1).sum().multiple(3)
sage: h.order(), h.degree(), h.leading_coefficient().order()
(9, 90, 84)
sage: 0 not in h.leading_coefficient()
True
```

Naive approaches for guessing a C^2-finite recurrence from given data yield polynomial systems of equations. However, if we fix the eigenvalues of the C-finite coefficients a potential recurrence can be obtained by solving a linear system. This way, we can, for instance, find and verify the simple C^2-finite recurrence from Example 6.2:

```
sage: K.<a> = NumberField(x^2-5)
sage: C2_K = C2FiniteSequenceRing(K)
sage: phi, psi = (1+a)/2, (1-a)/2
sage: eigenvalues = set([phi^4, psi^4, phi^6, psi^6, phi^8, psi^8])
sage: f2Data = [fib[n^2] for n in range(100)]
sage: sparse_fib_simp = C2_K.guess(f2Data, eigenvalues, order=3,
....:                                simple=True)
sage: sparse_fib_simp.degree()
4
sage: sparse_fib = C2_guess(sparse_fib)
sage: sparse_fib == sparse_fib_simp
True
```

Guessing is one of the most important tools that we have for C-finite and D-finite sequences. It would certainly be interesting to see if similar powerful algorithms can be developed for C^2-finite sequences.

Open Question 9.1. Can we find an efficient method for guessing a C^2-finite recurrence for the given terms of a sequence?

List of symbols

$\mathbb{N} = \{0, 1, 2, \dots\}$	Set of natural numbers
$\mathbb{N}_{\geq 1} = \{1, 2, \dots\}$	Set of positive natural numbers
$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}, \overline{\mathbb{Q}}$	Sets of integers, rational, real, complex and algebraic numbers
$\mathbb{K}[n], \mathbb{K}(n)$	The ring of polynomials and the field of rational functions over $\mathbb{K}$
$\deg(p)$	The degree of the polynomial p
$\mathrm{lc}(p), \mathrm{lc}_n(p)$	The leading coefficient of the polynomial p (w.r.t. n if specified)
$\mathrm{coeff}(p, i)$	The coefficient of n^i of the polynomial $p(n)$
$\mathbb{K}^{\mathbb{N}}$	The $\mathbb{K}$-algebra of sequences
$\langle a_1, a_2, \dots \rangle_R$	The R-module generated by the elements $a_1, a_2, \dots$
$\mathbb{K}[\![x]\!]$	The ring of formal power series over the field $\mathbb{K}$
$\sigma\colon \mathbb{K}^{\mathbb{N}} \to \mathbb{K}^{\mathbb{N}}$	The shift operator (page 4)
$\mathrm{ord}(\mathcal{A}), \mathrm{ord}(a)$	The order of an operator $\mathcal{A}$ or a sequence a (page 5)
$a(n) \sim b(n)$	Similarity of two sequences, i.e., $\lim_{n\to\infty} \frac{a(n)}{b(n)} = 1$ (page 8)
$\mathcal{R}_C$	Ring of C-finite sequences (page 9)
$R^{\times}$	Set of sequences of R which are units in $\mathbb{K}^{\mathbb{N}}$ (page 13)
$Q(R)$	The total ring of fractions of R (page 13)
$\mathbb{K}_{\sigma}[c_0, \dots, c_r]$	The smallest $\mathbb{K}$-difference algebra containing the sequences $c_0, \dots, c_r$ (page 21)
$n^{\underline{k}}$	Falling factorial, i.e., $n^{\underline{k}} = n(n-1)\cdots(n-k+1)$ (page 24)
$a \odot b$	The Cauchy product of the sequences a, b (page 29)
G_a	Suppose a is a sequence of order r, then $G_a = \left(a, \dots, \sigma^{r-1} a\right)$ (page 33)
$e_i^{(r)} \in R^r$	The i-th unit vector for $i \in \{0, \dots, r-1\}$ over the ring R (page 33)
M_a	Companion matrix of a sequence a (page 34)
$M \oplus N$	Direct sum of two matrices M, N (page 36)
$M \otimes N$	Kronecker product of two matrices M, N (page 36)

$\mathcal{R}_{C^k}, \mathcal{R}_{D^k}$	Ring of C^k-finite and D^k-finite sequences, respectively (page 73)
$\mathbb{A}$	Field of real algebraic numbers (page 79)
$\mathrm{lclm}(\mathcal{A}, \mathcal{B})$	Least common left multiple of the operators $\mathcal{A}, \mathcal{B}$ (page 89)

Bibliography

[AKK$^+$21] Shaull Almagor, Toghrul Karimov, Edon Kelmendi, Joël Ouaknine, and James Worrell. Deciding ω-regular properties on linear recurrence sequences. *Proc. ACM Program. Lang.*, 5(POPL), 2021.

[AM69] Michael F. Atiyah and Ian G. MacDonald. *Introduction to Commutative Algebra*. Addison-Wesley-Longman, 1969.

[AS21] Jakob Ablinger and Carsten Schneider. Solving linear difference equations with coefficients in rings with idempotent representations. In *Proceedings of the 2021 on International Symposium on Symbolic and Algebraic Computation*, ISSAC '21, page 27–34, New York, NY, USA, 2021. Association for Computing Machinery.

[Bar68] Erwin H. Bareiss. Sylvester's identity and multistep integer-preserving gaussian elimination. *Mathematics of Computation*, 22:565–578, 1968.

[BBY12] Jason P. Bell, Stanley N. Burris, and Karen Yeats. On the set of zero coefficients of a function satisfying a linear differential equation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 153(2):235–247, 2012.

[BCH21] Jason P. Bell, Shaoshi Chen, and Ehsaan Hossain. Rational Dynamical Systems, S-units, and D-finite Power Series. *Algebra and Number Theory*, 15(7):1699–1728, 2021.

[BCMS20] Curtis Bennett, Juan Carrillo, John Machacek, and Bruce Sagan. Combinatorial interpretations of lucas analogues of binomial coefficients and catalan numbers. *Annals of Combinatorics*, 24, 09 2020.

[Bek94] Emanuel Beke. Die Irreducibilität der homogenen linearen Differentialgleichungen. *Math. Ann.*, 45:278–294, 1894.

[Ber28] Daniel Bernoulli. Observationes de seriebus. *Acad. Sci. Petrop.*, 3:85–100, 1728.

[Beu91] Frits Beukers. The zero-multiplicity of ternary recurrences. *Compositio Mathematica*, 77(2):165–177, 1991.

[BGS07] Alin Bostan, Pierrick Gaudry, and Éric Schost. Linear Recurrences with Polynomial Coefficients and Application to Integer Factorization and Cartier-Manin Operator. *SIAM Journal on Computing*, 36(6):1777–1806, 2007.

[BIG03] Adi Ben-Israel and Thomas N.E. Greville. *Generalized Inverses: Theory and Applications*. CMS Books in Mathematics. Springer, 2003.

[BJP20] Alin Bostan and Antonio Jiménez-Pastor. On the exponential generating function of labelled trees. *Comptes Rendus. Mathématique*, 358(9-10):1005–1009, 2020.

[BK08] Alin Bostan and Manuel Kauers. Automatic classification of restricted lattice walks. 11 2008.

[BLN+22] Yuri Bilu, Florian Luca, Joris Nieuwveld, Joël Ouaknine, David Purser, and James Worrell. Skolem meets Schanuel. In *MFCS*, 2022.

[BM76] Jean Berstel and Maurice Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bulletin de la Société Mathématique de France*, 104:175–184, 1976.

[BM21] Alin Bostan and Ryuhei Mori. *A Simple and Fast Algorithm for Computing the N-th Term of a Linearly Recurrent Sequence*, pages 118–132. SIAM, 2021.

[Bos20] Alin Bostan. Computing the n-th term of a q-holonomic sequence. In *Proceedings of the 45th International Symposium on Symbolic and Algebraic Computation*, page 46–53, New York, NY, USA, 2020. Association for Computing Machinery.

[BP99] Andrej Bauer and Marko Petkovšek. Multibasic and mixed hypergeometric gosper-type algorithms. *Journal of Symbolic Computation*, 28(4):711–736, 1999.

[BR14] Arthur T. Benjamin and Elizabeth Reiland. Combinatorial Proofs of Fibonomial Identities. *The Fibonacci Quarterly*, 52(5), 2014.

[Bro72] Alfred Brousseau. *Fibonacci and Related Number Theoretic Tables: Fibonacci and related number theoretic tables*. Fibonacci and Related Number Theoretic Tables. Fibonacci Association, 1972.

[Bro03] Christopher W. Brown. QEPCAD B: A Program for Computing with Semi-Algebraic Sets Using CADs. *SIGSAM Bull.*, 37(4):97–108, 2003.

[BRPR03] Saugata Basu, S.B.R.P.M.F. Roy, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and computation in mathematics. Springer, 2003.

[CFQ15] Cristina Caldeira and João Filipe Queiró. Invariant factors of products over elementary divisor domains. *Linear Algebra and its Applications*, 485:345–358, 2015.

[CH91] George E. Collins and Hoon Hong. Partial Cylindrical Algebraic Decomposition for quantifier elimination. *Journal of Symbolic Computation*, 12(3):299–328, 1991.

[Cha14] Yongjae Cha. Closed form solutions of linear difference equations in terms of symmetric products. *Journal of Symbolic Computation*, 60:62–77, 2014.

[Chy14] Frédéric Chyzak. *The ABC of Creative Telescoping - Algorithms, Bounds, Complexity*. Habilitation à diriger des recherches, Ecole Polytechnique X, 2014.

[CJ98] Bob F. Caviness and Jeremy R. Johnson. *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Texts & Monographs in Symbolic Computation. Springer Vienna, 1998.

[CMDS84] David Carlson and Eduardo Marques De Sá. Generalized minimax and interlacing theorems. *Linear and Multilinear Algebra*, 15(1):77–103, 1984.

[CMP$^+$21] Michaël Cadilhac, Filip Mazowiecki, Charles Paperman, Michał Pilipczuk, and Géraud Sénizergues. On polynomial recursive sequences. *Theory of Computing Systems*, 2021.

[Coh06] Henry Cohn. A Short Proof of the Simple Continued Fraction Expansion of e. *The American Mathematical Monthly*, 113(1):57–62, 2006.

[Coh13] Henri Cohen. *A Course in Computational Algebraic Number Theory*. Graduate Texts in Mathematics. Springer Berlin Heidelberg, 2013.

[Col75] George E. Collins. Quantifier elimination for real closed fields by cylindrical algebraic decompostion. In *Automata Theory and Formal Languages*, pages 134–183. Springer Berlin Heidelberg, 1975.

[CR66] Charles W. Curtis and Irving Reiner. *Representation Theory of Finite Groups and Associative Algebras*. AMS Chelsea Publishing Series. Interscience Publishers, 1966.

[DLMF21] NIST Digital Library of Mathematical Functions. `http://dlmf.nist.gov`, Release 1.1.0 of 2020-12-15, 2021. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[EPSW15] Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence Sequences*. Mathematical Surveys and Monographs. American Mathematical Society, 2015.

[Eul44] Leonhard Euler. De fractionibus continuis dissertatio. *Commentarii academiae scientiarum Petropolitanae*, pages 98–137, 1744.

[Fac14] Paolo Faccin. *Computational problems in algebra: units in group rings and subalgebras of real simple Lie algebras*. PhD thesis, University of Trento, 2014.

[FGS05] Philippe Flajolet, Stefan Gerhold, and Bruno Salvy. On the non-holonomic character of logarithms, powers and the nth prime function. *Electronic Journal of Combinatorics*, 11(2):1–16, 2005.

[Fro75] Ferdinand G. Frobenius. über algebraisch integrirbare lineare differentialgleichungen. *J. Reine Angew. Math.*, 80:183–193, 1875.

[FS09] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[Fuc66] Lazarus Immanuel Fuchs. Zur theorie der linearen differentialgleichungen mit veränderlichen coefficienten. *J. Reine Angew. Math.*, 66:121–160, 1866.

[Ge93] Guoqiang Ge. *Algorithms Related to Multiplicative Representations of Algebraic Numbers*. PhD thesis, U.C. Berkeley, 1993.

[Ger04] Stefan Gerhold. On Some Non-Holonomic Sequences. *The Electronic Journal of Combinatorics*, 11(1), 2004.

[Ger05] Stefan Gerhold. *Combinatorial Sequences: Non-Holonomicity and Inequalities*. PhD thesis, Johannes Kepler University Linz, 2005.

[GHS08] Qiang-Hui Guo, Qing-Hu Hou, and Lisa H. Sun. Proving hypergeometric identities by numerical verifications. *Journal of Symbolic Computation*, 43(12):895–907, 2008.

[GK05] Stefan Gerhold and Manuel Kauers. A Procedure for Proving Special Function Inequalities Involving a Discrete Parameter. In *Proceedings of ISSAC 2005, Beijing, China, July 24–27, 2005*, pages 156–162, 2005.

[HHH06] Vesa Halava, Tero Harju, and Mika Hirvensalo. Positivity of second order linear recurrent sequences. *Discrete Applied Mathematics*, 154(3):447–451, 2006.

[HHHK05] Vesa Halava, Tero Harju, Mika Hirvensalo, and Juhani Karhumäki. Skolem's Problem: On the Border Between Decidability and Undecidability. Technical report, Turku Centre for Computer Science, 2005.

[Hir12] Michael D. Hirschhorn. Estimating the Apéry numbers. *The Fibonacci Quarterly*, pages 129–131, 2012.

[Hoe99] Mark van Hoeij. Finite singularities and hypergeometric solutions of linear recurrence equations. *Journal of Pure and Applied Algebra*, 139(1):109–131, 1999.

[Hoe21] Joris van der Hoeven. Fuchsian holonomic sequences, 2021. `https://hal.archives-ouvertes.fr/hal-03291372/`.

[Joh17] Fredrik Johansson. Arb: efficient arbitrary-precision midpoint-radius interval arithmetic. *IEEE Transactions on Computers*, 66:1281–1292, 2017.

[JPNP21] 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.

[JPNP23] 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.

[JPP18] Antonio Jiménez-Pastor and Veronika Pillwein. Algorithmic Arithmetics with DD-Finite Functions. In *Proceedings of ISSAC 2018, New York, NY, USA, July 16–19, 2018*, pages 231–237, 2018.

[JPP19] Antonio Jiménez-Pastor and Veronika Pillwein. A Computable Extension for Holonomic Functions: DD-Finite Functions. *J. Symb. Comput.*, 94:90–104, 2019.

[JPPS20] Antonio Jiménez-Pastor, Veronika Pillwein, and Michael F. Singer. Some structural results on d^n-finite functions. *Advances in Applied Mathematics*, 117, 2020.

[Jun31] Reinwald Jungen. *Sur les séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence*. PhD thesis, l'Ecole Polytechnique Fédérale, Zurich, 1931.

[KADF97] Johannes Kepler, Eric J. Aiton, Alistair M. Duncan, and Judith V. Field. *The Harmony of the World*. American Philosophical Society: Memoirs of the American Philosophical Society. American Philosophical Society, 1997.

[KAO12] Emrah Kilic, Ilker Akkus, and Hideyuki Ohtsuka. Some generalized fibonomial sums related with the gaussian q-binomial sums. *Bulletin mathématiques de la Société des sciences mathématiques de Roumanie*, 55, 2012.

[Kau05] Manuel Kauers. *Algorithms for Nonlinear Higher Order Difference Equations*. PhD thesis, Johannes Kepler University Linz, 2005.

[Kau06] Manuel Kauers. SumCracker: A package for manipulating symbolic sums and related objects. *Journal of Symbolic Computation*, 41(9):1039–1057, 2006.

[Kau07a] Manuel Kauers. An algorithm for deciding zero equivalence of nested polynomially recurrent sequences. *ACM Trans. Algorithms*, 3:18, 2007.

[Kau07b] Manuel Kauers. Computer Algebra and Power Series with Positive Coefficients. In *Proceedings of FPSAC'07*, pages 1–7, 2007.

[Kau13] 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.

[Kau14] 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.

[Kau15] Manuel Kauers. Algorithms for D-finite Functions. JNCF'15, Cluny, France, 2015.

[KJJ15] Manuel Kauers, Maximilian Jaroschek, and Fredrik Johansson. Ore Polynomials in Sage. In *Computer Algebra and Polynomials: Applications of Algebra and Number Theory*, pages 105–125. Springer International Publishing, 2015.

[KK09] Manuel Kauers and Christoph Koutschan. A mathematica package for q-holonomic sequences and power series. *The Ramanujan Journal*, 19(2):137–150, 2009.

[KK22] Manuel Kauers and Christoph Koutschan. Guessing with little data. In *Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*, ISSAC '22, page 83–90, 2022.

[KLOW20] George Kenison, Richard J. Lipton, Joël Ouaknine, and James Worrell. On the Skolem Problem and prime powers. In *ISSAC*, 2020.

[KM14] Tomer Kotek and Johann A. Makowsky. Recurrence relations for graph polynomials on bi-iterative families of graphs. *Eur. J. Comb.*, 41:47–67, 2014.

[KNP23] Manuel Kauers, Philipp Nuspl, and Veronika Pillwein. Order bounds for C^2-finite sequences, 2023.

[Kos11] Thomas Koshy. *Fibonacci and Lucas Numbers with Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2011.

[Kot12] Tomer Kotek. *Definability of Combinatorial Functions*. PhD thesis, Technion – Israel Institute of Technology, 2012.

[Kou05] Christoph Koutschan. Regular Languages and Their Generating Functions: The Inverse Problem. Diplomarbeit, Friedrich-Alexander-Universität Erlangen-Nürnberg, 2005.

[Kou10a] Christoph Koutschan. A fast approach to creative telescoping. *Mathematics in Computer Science*, 4(2-3):259–266, 2010.

[Kou10b] Christoph Koutschan. HolonomicFunctions (user's guide). Technical Report 10-01, RISC Report Series, Johannes Kepler University, Linz, Austria, 2010. https://risc.jku.at/sw/holonomicfunctions/.

[KP10] Manuel Kauers and Veronika Pillwein. When Can We Detect That a P-Finite Sequence is Positive? In *Proceedings of ISSAC 2010, Munich, Germany*, pages 195–201, New York, NY, USA, 2010. Association for Computing Machinery.

[KP11] Manuel Kauers and Peter Paule. *The Concrete Tetrahedron*. Texts and Monographs in Symbolic Computation. Springer, 2011.

[KV19] Manuel Kauers and Thibaut Verron. Why you should remove zeros from data before guessing. *ACM Commun. Comput. Algebra*, 53(3):126–129, 2019.

[KZ08] Manuel Kauers and Burkhard Zimmermann. Computing the Algebraic Relations of C-finite Sequences and Multisequences. *J. Symb. Comput*, 43(11):787–803, 2008.

[KZ18] Manuel Kauers and Doron Zeilberger. Factorization of c-finite sequences. In *Advances in Computer Algebra*, pages 131–147. Springer International Publishing, 2018.

[Lec53] Christer Lech. A note on recurring series. *Arkiv för Matematik*, 2:417–421, 1953.

[Li06] Hsuan-Chu Li. *Studies on Generalized Vandermonde Matrices: Their Determinants, Inverses, Explicit LU Factorizations, with Applications*. PhD thesis, National Chengchi University, 2006.

[Liu68] Chung L. Liu. *Introduction to combinatorial mathematics*. Computer Science Series. McGraw-Hill, first edition, 1968.

[LLN+22] Richard J. Lipton, Florian Luca, Joris Nieuwveld, Joël Ouaknine, David Purser, and James Worrell. On the Skolem problem and the Skolem conjecture. In *LICS*, 2022.

[LOW21] Florian Luca, Joël Ouaknine, and James Worrell. Universal Skolem sets. In *LICS*, 2021.

[LT90] Richard G. Larson and Earl J. Taft. The algebraic structure of linearly recursive sequences under hadamard product. *Israel Journal of Mathematics*, 72(1):118–132, 1990.

[LT08] Hsuan-Chu Li and Eng-Tjioe Tan. On a special generalized Vandermonde matrix and its LU factorization. *Taiwanese Journal of Mathematics*, 12:1651–1666, 2008.

[LT09] Vichian Laohakosol and Pinthira Tangsupphathawat. Positivity of third order linear recurrence sequences. *Discrete Applied Mathematics*, 157(15), 2009.

[Mah35] Kurt Mahler. Eine arithmetische Eigenschaft der Taylor-Koeffizienten rationaler Funktionen. *Akad. Wetensch. Amsterdam, Proc.*, 38:50–60, 1935.

[Mal96] Christian Mallinger. Algorithmic Manipulations and Transformations of Univariate Holonomic Functions and Sequences. Diplomarbeit, Johannes Kepler University Linz, 1996.

[Mar13] Diego Marques. The order of appearance of product of consecutive Lucas numbers. *The Fibonacci Quarterly*, 51, 2013.

[Mel21] Stephen Melczer. *An Invitation to Analytic Combinatorics*. Texts and Monographs in Symbolic Computation. Springer, 2021.

[Mid19] Johannes Middeke. Symbolic linear algebra, 2019. `https://www3.risc.jku.at/education/courses/ss2019/sla/script_sla2019.pdf`.

[MM22] Stephen Melczer and Marc Mezzarobba. Sequence Positivity Through Numeric Analytic Continuation: Uniqueness of the Canham Model for Biomembranes. *Accepted to Combinatorial Theory*, 2022.

[Moi22] Abraham de Moivre. Iii. de fractionibus algebraicis radicalitate immunibus ad fractiones simpliciores reducendis, deque summandis terminis quarumdam serierum æquali intervallo a se distantibus. *Philosophical Transactions*, 32, 1722.

[Moi30] Abraham de Moivre. *Miscellanea analytica de seriebus et quadraturis: accessere variae considerationes de methodis comparationum, combinationum & differentiarum ... itemque constructiones faciles orbium planetarum ...* J. Tonson & J. Watts, 1730.

[MS10] Marc Mezzarobba and Bruno Salvy. Effective bounds for P-recursive sequences. *Journal of Symbolic Computation*, 45(10):1075–1096, 2010.

[MST84] Maurice Mignotte, Tarlok Nath Shorey, and Robert Tijdeman. The distance between terms of an algebraic recurrence sequence. *Journal für die reine und angewandte Mathematik*, 349:63–76, 1984.

[New72] Morris Newman. *Integral Matrices*. ISSN. Elsevier Science, 1972.

[NOW21] Eike Neumann, Joël Ouaknine, and James Worrell. Decision problems for second-order holonomic sequences. In *LIPIcs: Leibniz International Proceedings in Informatics*, volume 198. Schloss Dagstuhl - Leibniz Center for Informatics, 2021.

[NP22a] 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.

[NP22b] 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.

[Nus22] 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.

[OEI23] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2023. `http://www.oeis.org`.

[OW12] Joël Ouaknine and James Worrell. Decision Problems for Linear Recurrence Sequences. In *Lecture Notes in Computer Science*, pages 21–28. Springer, 2012.

[OW14a] Joël Ouaknine and James Worrell. On the Positivity Problem for simple linear recurrence sequences. In *ICALP 14*, 2014.

[OW14b] Joël Ouaknine and James Worrell. Positivity problems for low-order linear recurrence sequences. In *SODA '14: Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, 2014.

[Pet92] Marko Petkovšek. Hypergeometric solutions of linear recurrences with polynomial coefficients. *Journal of Symbolic Computation*, 14(2):243–264, 1992.

[Pil08] Veronika Pillwein. Positivity of certain sums over Jacobi kernel polynomials. *Adv. in Appl. Math.*, 41(3):365–377, 2008.

[Pil13] Veronika Pillwein. Termination conditions for positivity proving procedures. In *Proceedings of ISSAC 2013, Boston USA, June 26–29, 2013*, pages 315–322, 2013.

[Pil19] Veronika Pillwein. On the positivity of the Gillis-Reznick-Zeilberger rational function. *Advances in Applied Mathematics*, 104:75–84, 2019.

[PWZ96] Marko Petkovšek, Herbert S. Wilf, and Doron Zeilberger. $A = B$. CRC Press, 1996.

[Rao53] K. Subba Rao. Some properties of fibonacci numbers. *The American Mathematical Monthly*, 60(10):680–684, 1953.

[Sag23] The Developers Sage. *SageMath, the Sage Mathematics Software System (Version 9.4)*, 2023. `https://www.sagemath.org`.

[Sal03] Bruno Salvy. Introduction aux s éries D-finies. Presented as the Alea'03, 2003. `https://perso.ens-lyon.fr/bruno.salvy/talks/Alea03/index.html`.

[Sal05] Bruno Salvy. D-finiteness: Algorithms and applications. In *Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation*, ISSAC '05, page 2–3, New York, NY, USA, 2005. Association for Computing Machinery.

[Sch07] Carsten Schneider. Symbolic summation assists combinatorics. *Seminaire Lotharingien de Combinatoire*, 56:1–36, 2007.

[Sch20] Carsten Schneider. Minimal representations and algebraic relations for single nested products. *Programming and Computer Software*, 46:133–161, 2020.

[Sig03] Laurence Sigler. *Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer New York, 2003.

[Sin85] Parmanand Singh. The so-called fibonacci numbers in ancient and medieval india. *Historia Mathematica*, 12(3):229–244, 1985.

[Sko33] Thoralf Skolem. Einige Sätze über gewisse Reihenentwicklungen und exponentiale Beziehungen mit Anwendung auf diophantische Gleichungen. *Oslo Vid. akad. Skrifter*, I(6), 1933.

[Slo73] Neil J.A. Sloane. *A Handbook of Integer Sequences*. Academic Press, 1973.

[SP95] Neil J.A. Sloane and Simon Plouffe. *The Encyclopedia of Integer Sequences*. Elsevier Science, 1995.

[ST05] Jaroslav Seibert and Pavel Trojovsky. On some identities for the Fibonomial coefficients. *Tatra Mountains Mathematical Publications*, 32, 2005.

[Sta80] Richard P. Stanley. Differentiably Finite Power Series. *European Journal of Combinatorics*, 1(2):175–188, 1980.

[Sta99] Richard P. Stanley. *Enumerative Combinatorics: Volume 2*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[Sta15] Richard P. Stanley. *Catalan Numbers*. Cambridge University Press, 2015.

[SZ94] Bruno Salvy and Paul Zimmermann. Gfun: a maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software (TOMS)*, 20(2):163–177, 1994.

[Tao22] Terrence Tao. Using the Smith normal form to manipulate lattice subgroups and closed torus subgroups, 2022. `https://terrytao.wordpress.com/2022/07/20/using-the-smith-normal-form-to-manipulate-lattice-subgroups-and-closed-torus-subgroups/`, Accessed: 24.4.2023.

[TZ20] Thotsaporn Aek Thanatipanonda and Yi Zhang. Sequences: Polynomial, C-finite, Holonomic, ..., 2020. `https://arxiv.org/pdf/2004.01370`.

[Ura20] Marcell János Uray. On proving inequalities by cylindrical algebraic decomposition. *Annales Univ. Sci. Budapest. Sect. Comp.*, pages 231–252, 2020.

[Ver85] Nikolai Konstantinovich Vereshchagin. Occurrence of zero in a linear recursive sequence. *Mat. Zametki*, 38(2):609–615, 1985.

[WZ85] Jet Wimp and Doron Zeilberger. Resurrecting the asymptotics of linear recurrences. *Journal of Mathematical Analysis and Applications*, 111(1):162–176, 1985.

[Yen96] Lily Yen. A two-line algorithm for proving terminating hypergeometric identities. *Journal of Mathematical Analysis and Applications*, 198(3):856–878, 1996.

[Yen97] Lily Yen. A two-line algorithm for proving q-hypergeometric identities. *Journal of Mathematical Analysis and Applications*, 213(1):1–14, 1997.

[YLN95] Kazuhiro Yokoyama, Ziming Li, and István Nemes. Finding roots of unity among quotients of the roots of an integral polynomial. In *Proceedings of ISSAC 1995, Montreal Quebec Canada July 10-12, 1995*, pages 85–89, 1995.

[Yur22] Sergey Yurkevich. The art of algorithmic guessing in gfun. In *Proceedings of the Maple Conference 2021*, volume 2, 2022.

[Zei90] Doron Zeilberger. A holonomic systems approach to special functions identities. *Journal of Computational and Applied Mathematics*, 32(3):321–368, 1990.

[Zei91] Doron Zeilberger. The method of creative telescoping. *Journal of Symbolic Computation*, 11(3):195–204, 1991.

[Zhe20] 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.

[Zhe21] Tao Zheng. A fast algorithm for computing multiplicative relations between the roots of a generic polynomial. *Journal of Symbolic Computation*, 104:381–401, 2021.

[ZX19] 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.

Index