université PARIS-SACLAY

universität wien

# DOCTORAL THESIS / THÈSE / DISSERTATION

Title of the doctoral thesis / Titre de la thèse / Titel der Dissertation

## “Integer sequences, algebraic series and differential operators”


submitted by / rédigée par / verfasst von

**Sergey Yurkevich, MMSc**


for the degree of / pour obtenir le grade de / angestrebter akademischer Grad

**Doctor of Philosophy (PhD)**

Thesis supervisors / directeurs de thèse / Dissertationsbetreuer:

**Dr. Alin Bostan** and

**Univ.-Prof. Mag. Dr. Herwig Hauser**

Composition of the thesis jury on July 6th 2023:

| | | |
|---|---|---|
| Andreas | Cap, | president, |
| Gilles | Villard, | |
| Wadim | Zudilin, | reviewers, |
| Alin | Bostan, | |
| Stéphane | Fischler, | |
| Herwig | Hauser, | examiners. |

Vienna, 2023 / Vienne, 2023 / Wien, 2023

# Acknowledgments

First and foremost I want to thank my supervisors Alin Bostan and Herwig Hauser without whom not only this thesis would have been impossible but also my personal academic career would have been doomed. I was extremely lucky to have both Alin and Herwig as my advisors for the thesis and also for the three years of my PhD studies. Alin taught me almost everything that I now know about algorithms and experimental mathematics, while my knowledge in algebraic geometry is solely due to Herwig. Thanks to Alin I learned how important good organization is, how to do good bibliography and how to write academic texts properly. Herwig encouraged and helped me to overcome academic difficulties and stressful situations, he precisely identified my strengths and guided me towards the completion of my goals. Keeping the necessary distance such that I could say that my achievements are well-deserved and realized by myself, I still always knew that I have his everlasting support. I extremely benefited from every single discussion with Alin which happened almost daily during the last 4 years. Apart from many philosophical conversations about weaknesses and strengths of present academia, he advised me on all my publications and presentations, as well as scientific applications, he helped me with many administrational struggles and important emails, prepared me for scientific discussions and showed to me the beauties and gems of academic life. After so many personal and online meetings and almost 200 thousand exchanged short messages with Alin, I am extraordinarily lucky being able to say that he has not only become my advisor, mentor and co-author but also a close friend. I am indebted to Herwig for the organization of four excellent workshops in Lisbon and many wonderful workshops across Austria where we could discuss very different kinds of mathematics and where I also had the chance to meet famous researchers and present my own contributions. On the same note, I also thank Alin for organizing the fantastic conference “Transient Transcendence In Transylvania”, as well as motivating me to participate at many other events that have become crucial for my academic career. Finally, I also am grateful for Herwig’s gracious financial support since he funded me as a master’s student, the first year of my PhD and also several trips to the aforementioned workshops.

I wish to thank all the other teachers from whom I learned mathematics. This includes on the one hand my school teachers and all professors at the Austrian Mathematical Olympiad who were the first ones showing me many wonderful mathematical gems, and on the other hand professors who taught me higher mathematics during lectures in my years at university. I especially thank those who left the greatest impression on me (listed in chronological order when we first met): Gerd Baron, Gerhard Kirchner, Clemens Heuberger, Andreas Cap, Christian Krattenthaler, Ilse Fischer, Kurt Hornik, Bruno Salvy, Gilles Villard, Frédéric Chyzak, Pierre Lairez and Vincent Neiger. Thank you for all your lectures!

I am very thankful to the MathExp group at Inria Saclay for the warm welcome in the team that I experienced. The permanent members Alin Bostan, Frédéric Chyzak, Philippe Dumas, Guy Fayolle and Pierre Lairez as well as the students Alexandre Goyer, Hadrien Notarantonio and Eric Pichon-Pharabod accepted me in their team and made me feel in Paris much more like at home. I am also indebted for Inria’s generous financial support that

crucially allowed me to visit many conferences, get a laptop and very valuable mathematical software.

As a PhD student I had the incredible opportunity to meet renowned and world-famous scientists and discuss with them about research. On this note I want to thank the Pol-Sys team, specifically Mohab Safey El Din for his support and many nice debates about academia, and Vincent Neiger for several pleasant scientific discussions while co-writing on a joint work. I thank Armin Straub for the joint productive week in Paris, and Wadim Zudilin for the interest in my work as well as so many brilliant comments and suggestions regarding it. I am indebted to Mireille Bousquet-Mélou for sharing her huge expertise about the combinatorics of discrete differential equations during eye-opening discussions in Oberwolfach and at the Institut Henri Poincaré. I thank Jacques-Arthur Weil for the amazing explanations of the current status of differential Galois theory and I am also thankful to Lucia Di Vizio for the kind invitation to the "Transcendance et combinatoire" seminar in the very beginning of my PhD. Finally, I am also indebted to all other outstanding researchers whom I met at different occasions and with whom I had the chance to exchange at least a few thoughts: Frits Beukers, Xavier Caruso, Frédéric Chyzak, Éric Delaygue, Javier Fresán, Charlotte Hardouin, Manuel Kauers, Hiraku Kawanoue, Christoph Koutschan, Maxim Kontsevich, Christian Krattenthaler, Pierre Lairez, Jean-Marie Maillard, Marc Mezzarobba, Karol Penson, Julien Roques, Bruno Salvy, Michael Singer, Duco van Straten, Michael Wibmer, Don Zagier.

I want to thank Alexandra Elbakyan for doing more for academia and research than possibly any living scientist.

I thank Andreas Cap and Stéphane Nonnenmacher for their great work not only as researchers but also as directors of the doctoral schools VSM and EDMH respectively. I am also very thankful to Gilles Villard and Wadim Zudilin for kindly agreeing to referee this thesis and I thank Stéphane Fischler for accepting to be an examiner.

I am grateful to the Austrian Academy of Sciences (ÖAW) for funding my PhD studies for two years via the DOC fellowship P-26101. I also thank the Austrian Science Fund (FWF) and the Agency for Education, Internationalisation (OeAD) as well as the Pro Scientia organization for various financial supports.

On the personal note, I want thank Florentina Stadlbauer for the everlasting support during the past 5 years we have been together. The moral aid and love I received from her was a crucial pillar for me that I value very much. I am also very thankful to all my friends who helped me a lot with both, the scientific difficulties and the mental hurdles in the last 3 years and also before. I am certain that without their support this dissertation would have not been possible. There are too many names to be mentioned here but I am immensely happy to know each and every one of them. I want to especially acknowledge the colleagues or co-authors who have became good friends: Giancarlo "John" Castellano, Johannes Droschl, Florian Fürnsinn, Hadrien Notarantonio, Hana Melánová, Chiara Novarini, Markus Reibnegger, Alex Stadler, Jakob Steininger.

I am enormously thankful to my family for their unconditional never-ending love and support. Most importantly, I want to express my deep gratitude and appreciation to my parents, Larisa and Andrey. Words cannot express how much they have done for me and how much I learned from them. I will always be astonished by their eternal optimism, strength and will to always go further, their ability to approach any problem rationally, their kindness and generosity, and their love to each other as well as their close ones. Dear parents, thank you for everything!

# Contents

# Abstract

This dissertation addresses mathematical and algorithmic problems and questions connected with integer sequences, algebraic series and differential operators. It is mainly composed of some of the articles the author (co-)wrote during his PhD studies. Explicitly, the thesis deals first with a family of hypergeometric sequences which can be represented as diagonals, the generating function of the Dubrovin-Yang-Zagier numbers, and a new formula for the reduced volume of any projection of the Clifford torus. Further, the dissertation presents three new algorithms solving the following problems more efficiently than previously possible: The computation of the $N$-th term of a $q$-holonomic sequence, the computation of the $N$-th power of a polynomial matrix, and the decision whether a given polyhedron has Rupert's property. Finally, the thesis also answers the following three explicitly stated but previously open questions: Is the Fibonacci sequence $(F_n)_{n\geq 0}$ a constant term sequence? (No), Does the $q$-analog of Pólya's Theorem hold? (Not in general but for some $q \in \mathbb{C}$), Does the Truncated icosidodecahedron have Rupert's property? (Yes). The last chapter contains a list of 60 open questions, problems and conjectures related to the topic of the dissertation.

# Résumé

Cette thèse aborde des problèmes et des questions mathématiques et algorithmiques liés aux suites d'entiers, aux séries algébriques et aux opérateurs différentiels. Elle est principalement composée de certains des articles que l'auteur a (co-)écrit pendant ses études de doctorat. Explicitement, la thèse traite d'abord d'une famille de suites hypergéométriques qui peuvent être représentées comme des diagonales, de la fonction génératrice des nombres de Dubrovin-Yang-Zagier, et d'une nouvelle formule pour le volume réduit de n'importe quelle projection du tore de Clifford. En outre, la thèse présente trois nouveaux algorithmes qui résolvent les problèmes suivants de manière plus efficace qu'auparavant : le calcul du $N$-ième terme d'une suite $q$-holonome, le calcul de la $N$-ième puissance d'une matrice polynomiale, et la décision si un polyèdre donné a la propriété de Rupert. Enfin, la thèse répond également aux trois questions suivantes, explicitement énoncées mais précédemment ouvertes : la suite de Fibonacci $(F_n)_{n\geq 0}$ est-elle une suite de termes constants ? (Non), Le $q$-analogue du théorème de Pólya est-il vrai ? (Pas en général mais pour certains $q \in \mathbb{C}$), L'icosidodécaèdre tronqué a-t-il la propriété de Rupert ? (Oui). Le dernier chapitre contient une liste de 60 questions ouvertes, problèmes et conjectures liés au sujet de la thèse.

# Zusammenfassung

Diese Dissertation befasst sich mit mathematischen und algorithmischen Problemen und Fragen im Zusammenhang mit ganzzahligen Folgen, algebraischen Reihen und Differentialoperatoren. Sie setzt sich hauptsächlich aus einigen Artikeln zusammen, die der Autor während seines Promotionsstudiums (mit)geschrieben hat. Explizit befasst sich die Dissertation zunächst mit einer Familie hypergeometrischer Folgen, die als Diagonalen dargestellt werden können, mit der Erzeugungsfunktion der Dubrovin-Yang-Zagier-Zahlen und mit einer neuen Formel für das reduzierte Volumen einer beliebigen Projektion des Clifford-Torus. Außerdem werden drei neue Algorithmen vorgestellt, die die folgenden Probleme effizienter lösen als bisher möglich: Die Berechnung des $N$-ten Terms einer $q$-holonomischen Folge, die Berechnung der $N$-ten Potenz einer Polynom-Matrix und die Entscheidung, ob ein gegebenes Polyeder die Rupert-Eigenschaft besitzt. Schließlich beantwortet die Arbeit auch die folgenden drei explizit formulierten, aber bisher offenen Fragen: Ist die Fibonacci-Folge $(F_n)_{n\geq 0}$ eine Folge mit konstantem Term? (Nein), Gilt die $q$-Analogie des Satzes von Pólya? (Nicht allgemein, aber für einige $q \in \mathbb{C}$), Hat das abgestumpfte Ikosidodekaeder die Rupertsche Eigenschaft? (Ja). Das letzte Kapitel enthält eine Liste von 60 offenen Fragen, Problemen und Vermutungen im Zusammenhang mit dem Thema der Dissertation.

# Chapter 1

# Introduction to the thesis

*“Мы все учились понемногу*
*Чему-нибудь и как-нибудь”*[1]
Александр С. Пушкин, *Евгений Онегин*, Глава первая

Instead of pursuing one specific mathematical or algorithmic problem, this dissertation is a product of a wonderful three-year journey of the author into the fantastic world of integer sequences, algebraic series and differential equations. During this time I learned many new concepts and ideas, discussed with prominent researchers in various areas and also had the opportunity to contribute himself to science and research.

The thesis consists of ten chapters: This Introduction, then eight Chapters 2–9 which are based on published or submitted research works by myself and co-authors, and the Conclusion (Chapter 10). More precisely, the chapters are built upon the following works:

- Chapter 2: “On a Class of Hypergeometric Diagonals” [82] with A. Bostan, published in *Proceedings of the American Mathematical Society* in 2022.
- Chapter 3: Work in preparation with A. Bostan and J.-A. Weil and part §3.1 in “The art of algorithmic guessing in `gfun`” [335] published in *Maple Transactions* in 2022.
- Chapter 4: “A hypergeometric proof that Iso is bijective” [81] with A. Bostan, published in *Proceedings of the American Mathematical Society* in 2022, part §3.2 in [335], and a work in preparation with A. Bostan and T. Yu.
- Chapter 5: “Fast Computation of the $N$-th Term of a $q$-Holonomic Sequence and Applications” [83] with A. Bostan, published in *Journal of Symbolic Computation* in 2023.
- Chapter 6: “Beating binary powering for polynomial matrices.” [75] with A. Bostan and V. Neiger, accepted at the conference ISSAC 23.
- Chapter 7: “On the $q$-analogue of Pólya’s Theorem” [84] with A. Bostan, published in *Electronic Journal of Combinatorics* in 2023.

[1] *Translation (A. S. Kline): “We’ve all acquired some education, A bit of this a bit of that”* Alexander S. Pushkin, *Eugene Onegin*, Chapter One.

- Chapter 8: "On the representability of sequences as constant terms" [79] with A. Bostan and A. Straub, submitted for publication in 2022.
- Chapter 9 "An algorithmic approach to Rupert's problem" [303] with J. Steininger, published in *Mathematics of Computation* in 2023 with extended abstract [302].

The Introduction summarizes and explains in the eight sections below each of the dissertation's chapters. Its purpose is to explain the initial motivation of each chapter, describe its main contribution and put it into a bigger perspective. I have the conviction that mathematical research is mainly led by open questions and conjectures. Therefore, for each of the chapters we formulate in its summary below an explicit question which was the main motivation for its creation. The answer to each of these questions, also formulated in the summaries, is usually the main contribution of each chapter. In order to stay compact in the introduction, I avoid giving all necessary references and details; they are, of course, present and explained in the corresponding chapters.

Following the devise that conjectures pave the way for mathematical research, I collected in the Conclusion a decent number of exceptionally interesting and still open problems, questions and conjectures related to the thesis. All of these problems I have encountered during the last three years; many of them are well-known and famous conjectures, others deal with explicit examples and are much less known. I hope that their variety ensures that everyone's taste is met at least once.

# Hypergeometric diagonals

The Chapter 2 is based on the joint work with A. Bostan [82]; the starting question that led to its creation reads as follows:

**Question 1.1.** *Can we generalize the main identities of Abdelaziz, Koutschan, Maillard [1]:*

$$ {}_3F_2\left(\left[\frac{2}{9},\frac{5}{9},\frac{8}{9}\right]\,;\left[1,\frac{2}{3}\right]\,;27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-y)^{1/3}}{1-x-y-z}\right) \quad \textit{and} $$

$$ {}_3F_2\left(\left[\frac{1}{9},\frac{4}{9},\frac{7}{9}\right]\,;\left[1,\frac{2}{3}\right]\,;27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-2y)^{2/3}}{1-x-y-z}\right)? $$

Before we hint on the interest in these identities and their importance, we first state this question properly and define the left- and right-hand sides of the equalities. The special functions on the left-hand sides are *generalized hypergeometric* functions defined as:

$$ {}_pF_q([a_1,\ldots,a_p]\,;[b_1,\ldots,b_q]\,;t) := \sum_{j\geq 0} \frac{(a_1)_j\cdots(a_p)_j}{(b_1)_j\cdots(b_q)_j}\,\frac{t^j}{j!}, $$

where $p,q\in\mathbb{N}$ (in the identities above $p=3$ and $q=2$), the numbers $a_1,\ldots,a_p,b_1,\ldots,b_q\in\mathbb{Q}$ are parameters, and $(x)_j := x(x+1)\cdots(x+j-1)$ denotes the rising factorial. For example,

$$ \begin{aligned} {}_3F_2\left(\left[\frac{2}{9},\frac{5}{9},\frac{8}{9}\right]\,;\left[1,\frac{2}{3}\right]\,;27\,t\right) &= 1+\frac{\frac{2}{9}\cdot\frac{5}{9}\cdot\frac{8}{9}}{\frac{2}{3}\cdot 1\cdot 1}27t+\frac{\frac{22}{81}\cdot\frac{70}{81}\cdot\frac{136}{81}}{\frac{10}{9}\cdot 2\cdot 2}27^2t^2+\cdots \\ &= 1+\frac{40}{9}t+\frac{5236}{81}t^2+\cdots. \end{aligned} \tag{1.1} $$

The *diagonal operator* of some $g(\mathbf{x}) = g(x_1, \ldots, x_n) \in \mathbb{Q}[\![x_1, \ldots, x_n]\!]$ on the right-hand side is defined as the univariate series

$$\mathrm{Diag}(g(\mathbf{x})) := \sum_{j \geq 0} g_{j,\ldots,j} t^j \in \mathbb{Q}[\![t]\!],$$

where $g(\mathbf{x})$ is a multivariate series given by

$$g(\mathbf{x}) = \sum_{(i_1,\ldots,i_n) \in \mathbb{N}^n} g_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n} \in \mathbb{Q}[\![\mathbf{x}]\!].$$

For example,

$$\frac{(1-x-y)^{1/3}}{1-x-y-z} = 1 + \frac{2}{3}x + \frac{2}{3}y + z + \frac{10}{9}xy + \frac{5}{3}xz + \cdots + \frac{40}{9}xyz + \cdots + \frac{5236}{81}x^2y^2z^2 + \cdots.$$

Notice that the coefficients of $x^n y^n z^n$ in this expansion (in red) and the coefficients (in blue) of $t^n$ in (1.1) agree for $n = 0, 1, 2$. It is precisely the statement of the first identity in Question 1.1 that they will agree for all $n \in \mathbb{N}$.

At first glance one might wonder why the identity in Question 1.1 is interesting apart of the *rigolo* numerological "coincidence". As very often the case with uncanny identities in mathematics, it turns out that these innocent-looking equalities lead the way to wonderful connections, algorithmic challenges and beautiful discoveries. These are explored in detail in Chapter 2 and the chapters afterwards and shall only be hinted on here.

The search for the answer to Question 1.1 is motivated by a fascinating conjecture due to Christol which connects arithmetics with geometry by trying to classify linear recurrences which yield (almost) integer sequences. This viewpoint leads to the theory of G-functions, relative periods and arithmetic gems such as the famous irrationality proof of $\zeta(3)$ by Apéry. A consequence of this conjecture states that

$${}_3F_2\left(\left[\frac{1}{9}, \frac{4}{9}, \frac{5}{9}\right]; \left[1, \frac{1}{3}\right]; 27\,t\right) = 1 + \frac{20}{9}t + \frac{2275}{81}t^2 + \frac{3124550}{6561}t^3 + \cdots$$

can be written as the diagonal of some algebraic function[2]. Even though the similarity of this function with those in the identities above is apparent, we remark that even this very particular case of the long-standing conjecture remains open. On the algorithmic side, one can show that the question of proving algorithmically identities like the ones in Question 1.1 algorithmically is decidable by a combination of two milestones in computer algebra: the method of *creative telescoping* for finding a differential equation for $\mathrm{Diag}(g)$ and *Petkovšek's algorithm* for deciding that a solution of this equation is indeed a hypergeometric function. An even more intriguing question is how to *find* such identities. This is where human intuition comes to play, in combination with an algorithmic guess-and-prove approach.

Without further ado we state our solution to Question 1.1:

[2]Recall that a (multivariate) function $g(\mathbf{x})$ is called *algebraic* if there exists a non-zero polynomial $P(\mathbf{x}, y)$ such that $P(\mathbf{x}, g(\mathbf{x})) = 0$; in other words $g(\mathbf{x})$ belongs to the algebraic closure of $\mathbb{C}(\mathbf{x})$ denoted by $\overline{\mathbb{C}(\mathbf{x})}$. A non-algebraic function is called *transcendental*.

**Answer 1.1.** *The diagonal of any finite product of algebraic functions of the form*

$$(1 - x_1 - \cdots - x_n)^R, \qquad R \in \mathbb{Q}, \tag{1.2}$$

*is a generalized hypergeometric function with explicitly determined parameters. More precisely, Theorem 2.3 generalizes the first identity in Question 1.1 and Theorem 2.4 generalizes the second one, in both cases for arbitrarily many variables and products.*

One advantage of our solution is that the proofs of the Theorems 2.3 and 2.4 are elementary and "human", i.e. they don't rely on previous results or algorithmic techniques. They show the *reason* for the existence of the identities in question and put them into a bigger context. Chapter 2 proceeds to classify diagonals obtained in the way (1.2). We prove that the Hadmard grade in each such case is finite, i.e. every function of the form $\mathrm{Diag}(R(\mathbf{x}))$ for $R(\mathbf{x})$ as in Theorem 2.3 or Theorem 2.4 can be written as a finite Hadamard (pointwise) product of algebraic functions. Moreover, we also manage to resolve 38 of previously unsolved particular cases of Christol's conjecture from a list of 116 potential counter-example constructed in 2011 by Bostan, Boukraa, Christol, Hassani, and Maillard.

To summarize, we could affirmatively answer Question 1.1, however Christol's conjecture stays open even in the particular case of hypergeometric functions. Still, we can say that Chapter 2 brought us one step closer to resolving this mystery.

# Dubrovin-Yang-Zagier numbers

The contents of the Chapter 3 are a combination of §3.1 in [335] as well as current work in progress with A. Bostan and J.-A. Weil. It was initially motivated by the question:

**Question 1.2.** *Consider the generating functions $F_a(t), F_b(t) \in \mathbb{Q}[\![t]\!]$ of the sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ in §10 of Zagier's amazing work [338]:*

$$F_a(t) = 1 - \frac{161}{2^8 3^4 5^2}t + \frac{26605753}{2^{18}3^9 5^6}t^2 + \cdots \quad \textit{and}$$
$$F_b(t) = 1 - \frac{161}{2^{10}3^3 5^2}t + \frac{3538565149}{2^{24}3^{10}5^6}t^2 + \cdots .$$

*In [338] it is only claimed on that $F_b(t)$ is algebraic. Can we also prove that $F_a(t) \in \overline{\mathbb{Q}(t)}$? Can we find analogous sequences to $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$? What about their generating functions?*

The precise definition (3.1), (3.2) of the numbers $a_n$ and $b_n$ looks scary at first glance. They correspond to Witten $r$-spin intersection numbers $\langle \tau_{s,m} \rangle$ for $r = 5$ and are defined as coefficients of integrals over some moduli spaces. It can be shown that $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ satisfy linear recurrences with polynomial coefficients, i.e. recursions of the form

$$p_r(n)u_{n+r} + \cdots + p_0(n)u_n = 0, \quad n \geq 0, \text{ for some } p_i(n) \in \mathbb{Q}[n]; p_0, p_r \neq 0. \tag{1.3}$$

Such sequences are called *P-recursive* (or *holonomic*) of order $r$. Equivalently, the generating functions are *D-finite* (or *holonomic*), i.e. they satisfy linear differential equations with polynomial coefficients of the form

$$q_s(t)y^{(s)}(t) + \cdots + q_0(t)y(t) = 0, \text{ for some } q_i(t) \in \mathbb{Q}[t]; q_s \neq 0. \tag{1.4}$$

For example, the numbers $(a_n)_{n\geq 0}$ satisfy the second-order recurrence

$$2^8 3^4 5^6 (5n+6)(n+2)(60n+43)a_{n+2} + 5^4(216000n^3 + 759600n^2 + 836940n + 290603)a_{n+1} \\ + (5n+4)(5n+3)(60n+103)a_n = 0,$$

and their generating function $F_a(t)$ satisfies the second-order differential equation

$$25t(7t-111600)(t^2 + 2^4 3^2 5^4 t + 2^8 3^4 5^5)F_a''(t) \quad + \quad (84t^2 - 3486325t - 2^4 3^2 5^6 4991)F_a(t) \\ +60(7t^3 + 51900t^2 - 2^6 3^3 5^5 1423t - 2^{10} 3^5 5^7 31)F_a'(t) \quad = 0.$$

It is not at all trivial to show that $a_n, b_n \in \mathbb{Z}[1/30]$. Computing terms using the recurrence means dividing in each step by $(5n+6)(n+2)(60n+43)$ which can have all possible prime factors. So from this perspective it looks like a miracle that all primes in the denominators of $a_n$'s except $2, 3, 5$ cancel out. One is, of course, immediately reminded of the famous Apéry numbers $(A_n)_{n\geq 0}$ which also follow a second-order recurrence

$$(n+1)^3 A_{n+1} - (2n+1)(17n^2 + 17n + 5)A_n + n^3 A_{n-1} = 0$$

with starting values $A_0 = 1, A_1 = 5$. The integrality of $A_n$ plays a major role in the irrationality proofs of $\zeta(3)$ and the fact that $A_n \in \mathbb{Z}$ follows from the following representation as a binomial sum:

$$A_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2. \tag{1.5}$$

Once found, this identity can be nowadays easily verified within a few milliseconds using the algorithmic technique of *creative telescoping*. Interestingly, this method, adapted for general use by Zeilberger, goes back to at least this precise example of the Apéry numbers and equality (1.5), see [319, §7].

Coming back to the numbers $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$, it is also possible to show a similar identity, for example:

$$a_n = 6^{-5n} \cdot \sum_{k=0}^{5n/2} \frac{(-9)^k}{10^{2k}} \cdot \frac{\left(\frac{3}{5}\right)_n \left(\frac{4}{5}\right)_n \left(\frac{1}{5}\right)_{3n-k}}{k!(5n-2k)!}. \tag{1.6}$$

It follows from this representation that $a_n \in \mathbb{Z}[1/30]$. There is, however, a major conceptual difference between the Apéry numbers and the Dubrovin-Yang-Zagier numbers $(a_n)_{n\geq 0}$. It is well-known (but not entirely trivial) that $\sum_{n\geq 0} A_n t^n$ is a transcendental function. On the other hand, we can show for $F_a(t)$ that:

**Answer 1.2.** *The generating function $F_a(t)$ is algebraic. Moreover, there are 7 other sequences of the same type as in Question 1.2. All of them belong to $\mathbb{Z}[1/30]$ and all the generating functions are most likely algebraic.*

The main purpose and contribution of Chapter 3 does not lie in Answer 1.2. This chapter is rather dedicated to explaining possible techniques for conjecturing and proving algebraicity of D-finite functions. This leads to a famous conjecture attributed to Grothendieck and its possible consequences, highly interesting numerical computations based on ideas of the Chudnovsky brothers and to differential Galois theory as well as invariant theory.

Moreover, it also touches the topic of algebraicity of hypergeometric functions which itself contains a whole world of fascinating theory. More precisely, we can prove that

$$F_a(t) = u_1(t) \cdot {}_2F_1\left[\begin{matrix}-1/60 & 11/60 \\ & 2/3\end{matrix}; q_1(t)\right] + u_2(t) \cdot {}_2F_1\left[\begin{matrix}19/60 & 31/60 \\ & 4/3\end{matrix}; q_2(t)\right],$$

for some explicit rational functions $q_1, q_2 \in \mathbb{Q}(t)$ and algebraic functions $u_1, u_2 \in \overline{\mathbb{Q}(t)}$.

# On Canham's problem

The Chapter 4 is based on joint work with A. Bostan [81], §3.2 in [335] and a current work in progress with A. Bostan and T. Yu. The question ([334, Thm. 1.1]) that led to this chapter reads as follows:

**Question 1.3.** *Is the stereographic projection of the Clifford torus to $\mathbb{R}^3$ uniquely determined by its isoperimetric ratio?*

As before, we shall first define the question properly by explaining the occurring technical terms before we proceed to describe its importance and answer.

The Clifford torus is a surface in $\mathbb{S}^3$, the four-dimensional 3-sphere, defined as

$$\mathcal{C} \coloneqq \left\{ [\cos u, \sin u, \cos v, \sin v]^T / \sqrt{2} : u, v \in [0, 2\pi) \right\}. \tag{1.7}$$

A stereographic projection of $\mathcal{C}$ to $\mathbb{R}^3$ is a perspective projection through a specific point on $\mathbb{S}^3$. The resulting surface, a so-called *Dupin cyclide*, is a Möbius transformation of the torus $T_{\sqrt{2}}$ which has major radius $R = \sqrt{2}$ and minor radius $r = 1$. The Figure 1.1 illustrates three such projections.

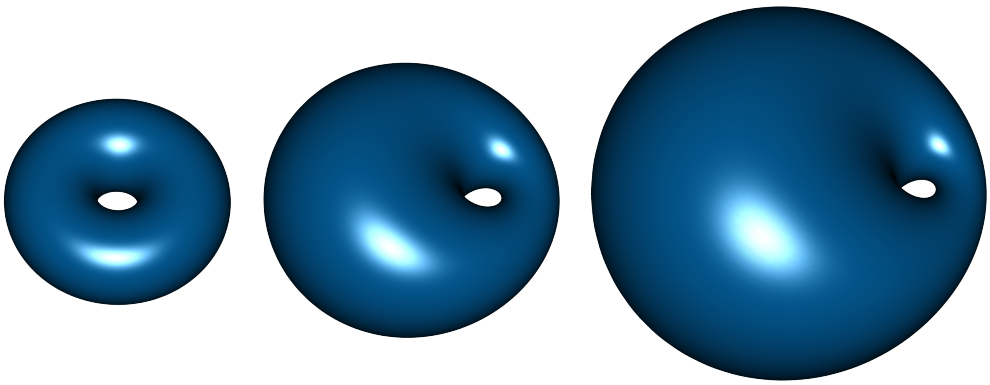


Figure 1.1: Three projections to $\mathbb{R}^3$ of the Clifford torus, equivalently three Möbius transformations of $T_{\sqrt{2}}$. The surface on the left is $T_{\sqrt{2}}$, and the two others are inversions of $T_{\sqrt{2}}$ at the unit sphere centered at $(0.2, 0, 0)$ and $(0.27, 0, 0)$.

The isoperimetric ratio $\iota(S)$ of a closed three-dimensional oriented surface $S$ is defined as

$$\iota(S) \coloneqq \pi^{1/6} \frac{\sqrt[3]{6V}}{\sqrt{A}},$$

where $A$ and $V$ are the surface area and volume of $S$. Clearly, $\iota(S)$ is scale-invariant. Moreover, by the isoperimetric inequality, $\iota(S) \leq 1$ with equality if and only if $S$ is the 2-sphere $\mathbb{S}^2$. The isoperimetric ratios of the three surfaces in Figure 1.1 are $\sqrt[3]{3/(2^{5/4}\pi^{1/2})} \approx 0.89$, $0.94$ and $0.96$ respectively.

Question 1.3 asks whether the shape of a Möbius transformed $T_{\sqrt{2}}$ is uniquely determined by $\iota$. This question appears naturally when studying the so-called *Canham model* in mathematical biology which asks for a surface $S$ of given isoperimetric ratio $\iota_0$ and genus $g$ such that the *Willmore energy* of $S$ is minimal. As a consequence of a recent theorem by Marques and Neves the set of all Willmore minimizers in genus $1$ without the constraint on $\iota$ is exactly given by stereographic projections of the Clifford torus. Hence, from this viewpoint, Question 1.3 asks whether the solution of Canham's problem is unique. For more details on the question's context we refer to Section 4.1 below.

One can prove that the set of all stereographic projections of $\mathcal{S}$ is a one-parameter family given by the $x$-coordinate of the center of the unit sphere at which $T_{\sqrt{2}}$ is inverted. It is the first contribution of Chapter 4 to show that the isoperimetric ratios of this family are given by

$$\sqrt[3]{\mathsf{Iso}(x)} = \sqrt[6]{\frac{9\sqrt{2}}{8\pi}} \cdot \frac{{}_2F_1\left[\begin{smallmatrix} -\frac{3}{2} & -\frac{3}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^{1/3}}{{}_2F_1\left[\begin{smallmatrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^{1/2}} \cdot \left(\frac{1-x^2}{1+x^2}\right)^{1/2}. \tag{1.8}$$

Then the answer to Question 1.3 follows:

**Answer 1.3.** *The function* $\mathsf{Iso}(x)$ *is strictly increasing on its domain and consequently Question 1.3 admits an affirmative answer.*

In Chapter 4 we proceed with proving an analogous formula for $\mathsf{Iso}_R(x)$, the isoperimetric ratio of a Möbius transformation of an arbitrary torus $T(R,r)$ and we prove that also this function is increasing in $x$ for any $R > 1$.

A crucial step in deriving (1.8) is finding linear differential equations for the surface area and volume of a torus $T_{\sqrt{2}}$ inverted at the unit sphere centered at $(x,0,0)$. As we show very explicitly in Section 4.3, this can be done by first finding the defining integrals and then applying *creative telescoping*. Then it remains to solve second-order linear differential equations in terms of hypergeometric functions which is another highly non-trivial task but for which we can also rely on modern computer algebra algorithms.

# Terms in $q$-holonomic sequences

The Chapter 5 consists of the joint work with A. Bostan [83]. Its initial motivation was to answer the following algorithmic question:

**Question 1.4.** *Given a field* $\mathbb{K}$*, some* $\alpha \in \mathbb{K}$ *and* $N \in \mathbb{N}$ *consider the polynomials*

$$F(x) = (\alpha - 1)(\alpha - x)\cdots(\alpha - x^{N-1}) \in \mathbb{K}[x] \quad \textit{and}$$
$$G(x) = 1 + \alpha x + \alpha^2 x^4 + \cdots + \alpha^N x^{N^2} \in \mathbb{K}[x].$$

*How fast can we evaluate* $F(q)$ *or* $G(q)$ *for some* $q \in \mathbb{K}$*?*

In this context "fast" means that we count the number of multiplications/divisions and additions/subtractions in $\mathbb{K}$ at unit cost and wish to keep this number low with respect to $N$. In other words, we are interested in an algorithm with good *arithmetic complexity*.

Clearly, if $\alpha = 0$ then $v(x)$ is trivially $1$ and $F(x) = \pm x^N$ in which case we can find $F(q)$ in $O(\log N)$ multiplications using *binary splitting*: Just repeat that $q^n = (q^{n/2})^2$ if $n$ is even and $q^n = q(q^{(n-1)/2})^2$ otherwise. The same idea works for $F(x)$ if $q = 1$ or any root of unity of small order compared to $N$. In general, however, for arbitrary $\alpha, q \in \mathbb{K}$, the question is very interesting for both $F(x)$ and $G(x)$ and, as we will see, for many other families of polynomials as well. The naive algorithm for $F(q)$ computes the powers $q^2, q^3, \ldots, q^{N-1}$ using $O(N)$ multiplications in $\mathbb{K}$, then computes the elements $\alpha-1$, $\alpha-q, \ldots, \alpha-q^{N-1}$ with $O(N)$ subtractions as well and finally multiplies these elements together using again $O(N)$ multiplications. The total arithmetic complexity is therefore $O(N)$. A similar reasoning shows that $G(q)$ can be computed in $O(N)$ operations as well.

Surprisingly, one can do better much and reduce $O(N)$ arithmetic operations to just $O(\mathbf{M}(\sqrt{N}))$. Here and later we write $\mathbf{M}(d)$ for the arithmetic complexity of the multiplication of two polynomials in $\mathbb{K}$ of degree at most $d$; using Fast Fourier Transform (FFT) techniques it is famously known that this task can be done in $O(d \log d \log\log d)$ operations and even $O(d \log d)$ under some assumptions on $\mathbb{K}$. We write $\tilde{O}(n)$ for hiding logarithmic factors inside the big-Oh notation, i.e. $\tilde{O}(n) \subseteq O(n^{1+\epsilon})$ for any $\epsilon > 0$. The main result of Chapter 5 reads as follows:

**Answer 1.4.** *Both polynomials in Question 1.4 can be evaluated using $\tilde{O}(\sqrt{N})$ operations in $\mathbb{K}$. Moreover, the same method applies for the class of $q$-holonomic sequences and has many applications, for example, in combinatorics, theory of orthogonal polynomials, $q$-difference equations and $q$-hypergeometric creative telescoping.*

The idea of $q$-deformations appears throughout mathematics. The analogous class of sequences to the P-recursive ones is the class of *$q$-holonomic sequences*. Often such sequences are defined for a variable $q$, however in view of Question 1.4 it makes sense for us to define them for $q \in \mathbb{K}$. In this sense a sequence $(u_n(q))_{\geq 0}$ of elements in $\mathbb{K}$ is called $q$-holonomic if it satisfies

$$c_r(q, q^n)u_{n+r}(q) + \cdots + c_0(q, q^n)u_n(q) = 0, \quad \text{for all } \ n \geq 0,$$

for some polynomials $c_0(x, y), \ldots, c_r(x, y)$ in $\mathbb{K}[x, y]$, with $c_r(x, y) \neq 0$ and some $q \in \mathbb{K}$. A typical example is the *$q$-factorial* which appears, for example, in combinatorics when counting more refined permutations:

$$[n]_q! := (1 + q)(1 + q + q^2) \cdots (1 + q + \cdots + q^{n-1}).$$

In view of this definition it is not difficult to see that both polynomials in Question 1.4 are $q$-holonomic sequences (if $N = n$ is seen as an index and $x$ is evaluated to $q$), for example $F_n(q) = \prod_{i=0}^{n-1}(\alpha - q^i)$ satisfies

$$F_{n+1}(q) - (\alpha - q^n)F_n(q) = 0.$$

It is an idea by Strassen that $n! \in \mathbb{K}$ can be computed using just $\tilde{O}(\sqrt{n})$ operations in $\mathbb{K}$. This fact is not only surprising but also carries very broad consequences, for example it

allows to factor $M \in \mathbb{Z}$ deterministically in $\tilde{O}(\sqrt[4]{M})$ bit operations. The trick is to use the so-called baby-step/giant-step technique which exploits the fact that multipoint evaluation and interpolation can be done quasi-optimally: given a polynomial $P(x) \in \mathbb{K}$ of degree $d$ and $d$ points $a_1, \ldots, a_d \in \mathbb{K}$, one can compute $P(a_1), \ldots, P(a_d)$ in $\tilde{O}(d)$ operations in $\mathbb{K}$.

The Chudnovsky brothers extended Strassen's idea to the class of P-recursive sequences. This works by observing that for computing the $N$-th term of a P-recursive sequence $(u_n)_{n\geq 0}$, one may first translate the recursion to a first-order matrix recurrence. More precisely, if $(u_n)_{n\geq 0}$ is defined by a recurrence of order $r$ (1.3), it is enough to perform the "matrix factorial" $A(n)A(n-1)\cdots A(1)$ for the companion matrix $A(x) \in \mathbb{K}[x]^{r\times r}$. This can be done with an adaptation of Strassen's method.

In order to prove the claim in Answer 1.4 we adapted Chudnovsky's work to the $q$-holonomic world. Surprisingly, the resulting algorithm is even somewhat simpler, essentially because in the mentioned multipoint evaluation we have to evaluate a polynomial $P(x) \in \mathbb{K}[x]$ at a geometric sequence instead an arithmetic one.

# Terms in C-finite sequences

This Chapter 6 consists of the joint work with A. Bostan and V. Neiger [75]. Its accompanying innocent-looking question sounds as follows:

**Question 1.5.** *How fast can we compute the $N$-th Fibonacci polynomial $F_N(x)$?*

By definition, for any $\mathbb{K}$ of characteristic zero, the Fibonacci polynomials $F_n(x) \in \mathbb{K}[x]$ are a straightforward polynomial analogue of the Fibonacci numbers:

$$F_{n+2}(x) \coloneqq xF_{n+1}(x) + F_n(x), \quad \text{for } n \geq 0, \tag{1.9}$$

with initial conditions $F_0(x) = 0$ and $F_1(x) = 1$. Clearly, $F_n(1)$ gives the $n$-th Fibonacci number and $\deg F(x) = n-1$ for $n \geq 1$. The polynomial sequence starts as

$$(F_n(x))_{n\geq 0} = (0, 1, x, x^2+1, x^3+2x, x^4+3x^2+1, \ldots),$$

and, completely analogously to the Fibonacci numbers, it holds that

$$\sum_{n\geq 0} F_n(x)y^n = \frac{y}{1-xy-y^2}.$$

More generally, C-finite sequences are defined over some ring $R$ (in case of Question 1.5 $R = \mathbb{K}[x]$) and satisfy a linear recurrence of the form

$$u_{n+r} = c_{r-1}u_{n+r-1} + \cdots + c_0 u_n \quad \text{for all } n \geq 0, \tag{1.10}$$

for some $r \geq 1$ and $c_0, \ldots, c_{r-1} \in R$. It is well known and easy to see that a sequence is C-finite if and only if its generating function $U(y) = \sum_{n\geq 0} c_n y^n \in R[\![y]\!]$ is rational.

If we are interested the number of operations in $R$ to compute the $N$-th term of $(u_n)_{n\geq 0}$ in (1.10) it is a good idea to consider the first-order matrix recurrence governed by the *companion matrix* of the recursion. Then it is easy to see that contrary to the matrix *factorial*

in Chapter 5 we just need to calculate the $N$-th *power* of a constant matrix in $R^{r\times r}$. This can be done efficiently using *binary powering* (or even better by an idea due to Fiduccia) and results in an algorithm with arithmetic complexity $O(\log N)$ steps in $R$.

When $R = \mathbb{K}[x]$, like for Fibonacci polynomials, we speak of *polynomial C-finite sequences*. Naturally in this case, we would like to count the number of operations in $\mathbb{K}$ instead in $\mathbb{K}[x]$. Since the $N$-th Fibonacci polynomial has degree $N-1$, the best potential algorithm we can hope for has complexity $O(N)$. Note that calculating $F_N(x)$ using the defining recurrence (1.9) results in an algorithm with complexity $O(N^2)$ since it computes all previous terms $F_k(x)$, with $k \leq N$ as well. The trick of powering the companion matrix is a big improvement compared to the naive method: this algorithm has complexity $O(\mathsf{M}(N))$, where, as before, $\mathsf{M}(n)$ stands for the arithmetic complexity of multiplication of two polynomials of degree at most $n$. Since FFT allows $\mathsf{M}(n) = O(n \log n \log\log n)$, we see that this algorithm is quasi-optimal with respect to the output size: The $N$-th Fibonacci polynomial over any field $\mathbb{K}$ can be computed in complexity $O(N \log N \log\log N)$ operations in $\mathbb{K}$. The remaining optimistic question is whether the factor $\log N$ can be dropped and whether it is possible to have a purely linear algorithm to compute $F_N(x)$. The surprisingly affirmative answer is given in Chapter 6:

**Answer 1.6.** *It is possible to compute the $N$-th term of any polynomial C-finite sequence in $O(N)$ operations in $\mathbb{K}$. In particular, $F_N(x) \in \mathbb{K}[x]$ can be computed in $O(N)$ field operations. Moreover, consequently, the $N$-th power any polynomial matrix $M(x) \in \mathbb{K}[x]^{r\times r}$ can be deduced in complexity*[3] *$O(N)$.*

At first glance, the most surprising part of this contribution is that the $N$-th power of a polynomial matrix can be computed faster than with $O(\mathsf{M}(N))$ multiplications. In the case $r = 1$ this means that $P(x)^N$ for some $P(x) \in \mathbb{K}[x]$ can be deduced faster than with binary powering – a striking fact that was first observed by Flajolet and Salvy in 1997. Indeed, we also compute $M(x)^N$ without ever multiplying polynomials!

Let us come back to the example of Fibonacci polynomials. By the trivial identity

$$\begin{pmatrix} F_{n+1}(x) & F_n(x) \\ F_n(x) & F_{n-1}(x) \end{pmatrix} = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix}^n,$$

it is clear that computing $F_N(x)$ in complexity $O(N)$ is equivalent to computing the $N$-th power of the matrix on the right-hand side in the same complexity. The key observation that allows to compute $F_N(x)$ faster than with binary powering is that writing $F_N(x) = \sum_{k=0}^{N-1} f_k x^k$ one can prove that $(f_n)_{n\geq 0}$ satisfies the recurrence relation

$$f_{k+2} = \frac{(N+k+1)(N-k-1)}{4(k+1)(k+2)} f_k \quad \text{for all } k \geq 0. \tag{1.11}$$

with initial conditions $(f_0, f_1) = (1, 0)$ for $N$ odd and $(f_0, f_1) = (0, N/2)$ otherwise. Simply unrolling this recurrence yields an $O(N)$ algorithm for all $f_0, \ldots, f_{N-1}$, consequently for $F_N(x)$ and therefore, repeating this also for $N+1$ and $N-1$, we find the $N$-th power of the corresponding matrix in arithmetic complexity $O(N)$ as well.

[3]Note that in this context we only consider the complexity with respect to $N$, i.e. we assume that $r$ and the degree of the entries of $M(x)$ are in $O(1)$.

Speaking generally, the main fact that leads to Answer 1.6 is that if $(u_n(x))_{n\geq 0}$ is a polynomial C-finite sequence then the sequence $(c_k)_{k\geq 0}$ given by $u_N(x) = \sum_{k\geq 0} c_k x^k$ is not only P-recursive but actually satisfies a recursion of order and degree independent of $N$. This can be seen either from the partial fraction decomposition of the rational generating function $U(x, y)$ and then application of the crucial Lemma 6.3 on algebraic substitution in D-finite functions, or using *reduction based creative telescoping* on the function $U(x, y)/y^{n+1}$ together with Lemma 6.5.

Inspired by the analogy between the arithmetic complexity model in $\mathbb{K}[x]$ and the bit complexity in $\mathbb{Z}$, a very intriguing question is whether it is possible to compute the $N$-th term of a C-finite sequence over $\mathbb{Z}$ in $O(N)$ binary operations instead $O(\mathbf{M}_{\mathbb{Z}}(N))$, where $\mathbf{M}_{\mathbb{Z}}(n)$ stands for the complexity of multiplying two numbers with bit-size at most $n$. To our knowledge, this question stays open even for the case $a^N$ for $a \in \mathbb{Z}$ since the digits of a power in a different basis seem to have no regularity one could exploit (see Question 10.27).

## On a $q$-analogue of Pólya's Theorem

The Chapter 7 consists of joint work with A. Bostan [84] and fully answers the following question asked by Aissen in [11]:

**Question 1.7.** *Let $n, k, a, b \in \mathbb{Z}$ be "admissible" integers. Is the bivariate series*

$$F(x, q) := \sum_{j\geq 0} \begin{bmatrix} n + aj \\ k + bj \end{bmatrix}_q x^j \in \mathbb{C}[q][\![x]\!] \tag{1.12}$$

*algebraic over $\mathbb{C}(x, q)$?*

Similarly to Chapter 5, the $q$-binomial coefficients in this question are defined as the analogues of usual binomial coefficients:

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{[n]_q!}{[k]_q![n-k]_q!} \in \mathbb{Z}[q],$$

where $[n]_q! := (1 + q) \cdots (1 + q + \cdots + q^{n-1})$. The integers $n, k, a, b$ are called admissible if $n \geq k \geq 0$, $a > b > 0$, $\gcd(a, b) = 1$, and either $n - k < a - b$ or $k < b$. This oddly-looking condition just means that for $j \geq 0$ the sequence $\begin{bmatrix} n+aj \\ k+bj \end{bmatrix}_q$ lies on a straight line in the $q$-Pascal triangle (that is the usual Pascal triangle but with binomials replaced with $q$-binomials) and that this line contains infinitely many points none of which are skipped.

The interest in the question above is motivated by a theorem of Pólya from 1921 that ensures that the analogous classical generating function $F(x)$ with usual binomial coefficients instead $q$-binomial coefficients is indeed algebraic over $\mathbb{C}(x)$. In this sense, Aissen asked in [11] whether the $q$-analogue of Pólya's theorem holds true.

The short Chapter 7 completely answers Aissen's question:

**Answer 1.8.** *The function $F(x, q)$ is transcendental over $\mathbb{C}(x, q)$. Moreover, if $q = \omega \in \mathbb{C}$ is specialized to some value then $F(x, \omega)$ is algebraic over $\mathbb{C}(x)$ if and only if $\omega$ is root of unity.*

The wider context of Question 1.7 is the domain of $q$-difference equations, that is (in case $q \in \mathbb{C}$) functional equations is of the form

$$f(q^m x) + b_{m-1}(x) f(q^{m-1} x) + \cdots + b_0(x) f(x) = 0,$$

for rational functions $b_0, \ldots, b_m \in \mathbb{C}(x)$ not all zero. Recall that this notion also appears in Chapter 6. For example, the generating function $f(x)$ of central $q$-binomial coefficients ($n = k = 0$ and $a = 2, b = 1$ in (1.12)) satisfies

$$q^3 x f(q^4 x) - (q(q+1)x + 1) f(q^2 x) + 2 f(qx) + (x-1) f(x) = 0.$$

A celebrated result by Ramis (1992) states that if a function $f(x)$ satisfies a linear differential equation and a $q$-difference equation for $q \in \mathbb{C}$ not a root of unity, then $f(x) \in \mathbb{C}(x)$ is a rational function. This fascinating dichotomy result is one of many in the area linear differential and difference equations.

Since algebraic functions are D-finite, Ramis' result allows to reduce the Question 1.7 to classifying rational functions $F(x, \omega)$ for $\omega \in \mathbb{C}$ not a root of unity. We remark, however, that this task is also not at all trivial and therefore in Chapter 7 we prove directly and in an elementary way that $F(x, \omega)$ is D-finite if and only if $\omega$ is a root of unity.

## Constant term sequences

The Chapter 8 incorporates joint work with A. Bostan and A. Straub [79]. The initial question that lead to it was asked by Straub in [309]:

**Question 1.9.** *Can the Fibonacci numbers $F_n$ be expressed as a constant term sequence?*

Recall that $(F_n)_{n \geq 0}$ is the coefficient sequence of $x/(1 - x - x^2)$ and satisfies for $n \geq 0$ the linear recurrence relation with constant coefficients:

$$F_{n+2} = F_{n+1} + F_n, n \geq 0, \quad \text{with } F_0 = 0 \text{ and } F_1 = 1.$$

A sequence $(A_n)_{n \geq 0}$ has a constant term representation if there exist multivariate Laurent polynomials $P(\mathbf{x}), Q(\mathbf{x}) \in \mathbb{Q}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ such that $A_n = \mathrm{ct}[P^n(\mathbf{x}) Q(\mathbf{x})]$. Here $\mathrm{ct}[F]$ stands for the constant term, i.e. for the coefficient of $x_1^0 \cdots x_n^0$ in $F$.

Constant term sequences are a very mysterious subclass of P-recursive sequences. Two such prominent examples are the Catalan numbers $C_n$ and the Apéry numbers $A_n$:

$$C_n = \mathrm{ct}\left[(x^{-1} + 2 + x)^n (1 - x)\right] \quad \text{and}$$
$$A_n = \mathrm{ct}\left[\left(\frac{(x+y)(z+1)(x+y+z)(y+x+1)}{xyz}\right)^n\right].$$

From the geometric series it follows that the generating function of a constant term sequence can be expressed as the diagonal (see Chapter 2) of a rational function. Moreover, a constant term representation obviously implies that the sequence is globally bounded but actually carries more arithmetic information than diagonals of rational functions. For

example, it is easy to see that if $(A_n)_{n\geq 0}$ is a constant term sequence with $Q(\mathbf{x}) = 1$ then $A_p \equiv A_1 \bmod p$ for all primes $p$.

For general constant term sequences we show in Chapter 8 that $A_p \equiv A_1 \bmod p$ for $p$ prime and *large enough*. Since the $p$-th Fibonacci number $F_p$ is $\pm 1 \bmod p$ dependent on $p \bmod 5$ it becomes evident that the answer to Question 1.9 is negative. The more ambitious and interesting task is therefore: Classify *all* constant term sequences that are C-finite, i.e. those whose generating function is rational (see also Chapter 6). We fully solve this task in Chapter 8:

**Answer 1.10.** *A C-finite sequence $(A_n)_{n\geq 0}$ is a constant term sequence if and only if it has one single characteristic root $\lambda$ and $\lambda \in \mathbb{Q}$.*

In the same chapter we proceed to classify C-finite sequences that are finite linear combinations of constant terms and discuss the same question for the much less understood class of hypergeometric sequences.

# On Rupert's problem

The Chapter 9 combines the joint works with J. Steininger [302, 303]. Its motivation is led by the following question from elementary geometry:

**Question 1.11.** *Does every polyhedron in $\mathbb{R}^3$ have Rupert's property?*

A polyhedron $\mathbf{P} \in \mathbb{R}^3$ is said to be *Rupert* or *have Rupert's property* if $\mathbf{Q}$, a copy of $\mathbf{P}$, can be moved through a straight tunnel ("hole") inside $\mathbf{P}$. For example, the three-dimensional cube is Rupert, because the unit square (the view from "above") strictly fits into the regular hexagon with side length $\sqrt{2/3}$ (viewed from the direction of a main diagonal), thus, in the left picture below, the black unit cube passes through a hole in the red one:

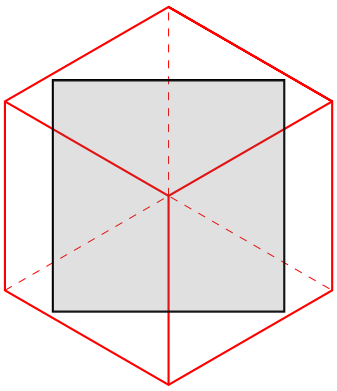


Figure 1.2: The cube is Rupert.

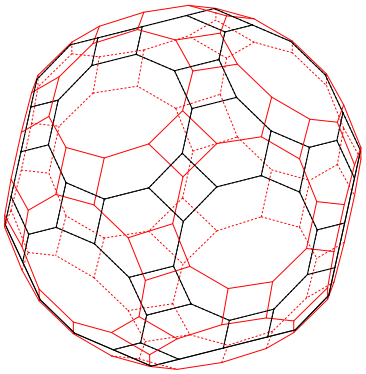


Figure 1.3: The TID is Rupert.

Given that the cube has this surprising property, the natural Question 1.11 is which other polyhedra admit it as well. During the last half century it was proven that all five Platonic solids are Rupert, and moreover, this property was confirmed 2018 for 8 out of the 13 Archimedean solids and in 2019 for a 9th Archimedean solid. The conjecture arose that possibly all polyhedra have Rupert's property.

While Question 1.11 stays open, Chapter 9 contains the following contribution:

**Answer 1.12.** *For a given* $\boldsymbol{P} \subseteq \overline{\mathbb{Q}}^3$ *the question whether* $\boldsymbol{P}$ *is Rupert is algorithmically decidable. Moreover, there exists a practical Las Vegas type algorithm*[4] *that proves that all 5 Platonic, 10 Archimedean and 82 Johnson solids have Rupert's property.*

In particular, a new Archimedean solid, the truncated icosidodecahedron (Figure 1.3), is resolved in Chapter 9. Moreover, we find improved *Nieuwland numbers* for most of the resolved solids, i.e. the obtained solutions are in some sense better than the previously best known. For the cube this was first investigated by Nieuwland who could prove that this solid with side length $3\sqrt{2}/4 \approx 1.06$ can be moved through a hole of the unit cube, and that this number is optimal. Conjecturally, the same number seems to be achieved by the optimal solution of the octahedron. Moreover, our recent findings suggest that the dodecahedron and icosahedron both have solutions (Figure 1.4 and Figure 1.5) with optimal "scaling factor" $\approx 1.0108$ which is a root of $P(x) = 2025x^8 - 11970x^6 + 17009x^4 - 9000x^2 + 2000$. The connection between duality and Rupert's property is evident, certainly very interesting but yet to be fully understood.

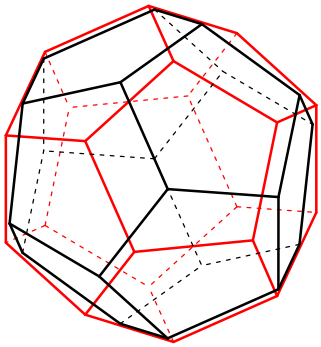


Figure 1.4: The dodecahedron is Rupert.

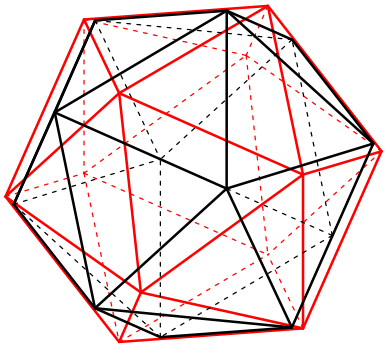
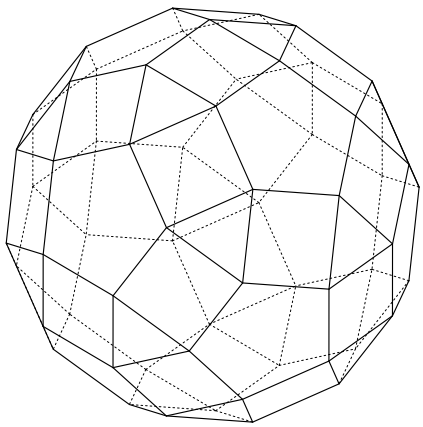


Figure 1.5: The icosahedron is Rupert.

In the same chapter we also investigate on the newly defined concept of *Rupertness* which describes how likely it is to find a solution to Rupert's problem for a given solid if searching for it "randomly". Based on our results, we conjecture that the Rhombiicosidodecahedron (RID in short, a point-symmetric Archimedean solid, depicted below) does not have Rupert's property.



[4]This means that this algorithm will find a (provable) solution *eventually* if it exists, but cannot prove the non-existence of a solution.

# Chapter 2

# Hypergeometric diagonals

*"Alles Gescheite ist schon gedacht worden.*
*Man muß nur versuchen, es noch einmal zu denken."*[1]
Johann Wolfgang von Goethe,
*Wilhelm Meisters Wanderjahre II*, Kunst, Ethisches, Natur

This chapter of the thesis is devoted to the study of diagonals of a family of multivariate algebraic functions. Explicitly, we prove that the diagonal of any finite product of algebraic functions of the form

$$(1 - x_1 - \cdots - x_n)^R, \qquad R \in \mathbb{Q},$$

is a generalized hypergeometric function, and we provide an explicit description of its parameters. The particular case $(1 - x - y)^R/(1 - x - y - z)$ corresponds to the main identity of Abdelaziz, Koutschan and Maillard in [1, §3.2]. Our result is useful in both directions: on the one hand it shows that Christol's conjecture holds true for a large class of hypergeometric functions, on the other hand it allows for a very explicit and general viewpoint on the diagonals of algebraic functions of the type above. Finally, in contrast to [1], our proof is completely elementary and does not require any algorithmic help.

This chapter is based on the joint work with A. Bostan [82].

## 2.1 Introduction

Let $\mathbb{K}$ be a field of characteristic zero and let $g \in \mathbb{K}[\![\mathbf{x}]\!]$ be a power series in $\mathbf{x} = (x_1, \ldots, x_n)$

$$g(\mathbf{x}) = \sum_{(i_1,\ldots,i_n)\in\mathbb{N}^n} g_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n} \in \mathbb{K}[\![\mathbf{x}]\!].$$

The *diagonal* $\mathrm{Diag}(g)$ of $g(\mathbf{x})$ is the univariate power series given by

$$\mathrm{Diag}(g) := \sum_{j\geq 0} g_{j,\ldots,j} t^j \in \mathbb{K}[\![t]\!].$$

[1] Translation: "Everything wise has already been thought. You just have to try to think it again."

A power series $h(\mathbf{x})$ in $\mathbb{K}[\![\mathbf{x}]\!]$ is called *algebraic* if there exists a non-zero polynomial $P(\mathbf{x}, T) \in \mathbb{K}[\mathbf{x}, T]$ such that $P(\mathbf{x}, h(\mathbf{x})) = 0$; otherwise, it is called *transcendental*.

If $g(\mathbf{x})$ is algebraic, then its diagonal $\mathrm{Diag}(g)$ is usually transcendental; however, by a classical result by Lipshitz [229], $\mathrm{Diag}(g)$ is D-finite, i.e., it satisfies a non-trivial linear differential equation with polynomial coefficients in $\mathbb{K}[t]$. Equivalently, the coefficients sequence $(g_{j,\ldots,j})_{j\geq 0}$ of $\mathrm{Diag}(g)$ is P-recursive, i.e., it satisfies a linear recurrence with polynomial coefficients (with respect to $j$).

When a P-recursive sequence satisfies a recurrence of order 1, we say that it is *hypergeometric*. An important class of power series, whose coefficients sequence is hypergeometric by design, is that of generalized hypergeometric functions. Let $p, q \in \mathbb{N}$ and $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$ be rational numbers such that $b_i + j \neq 0$ for any $i, j \in \mathbb{N}$. The *generalized hypergeometric function* ${}_pF_q$ with parameters $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$ is the univariate power series in $\mathbb{K}[\![t]\!]$ defined by

$$ {}_pF_q([a_1, \ldots, a_p]\,; [b_1, \ldots, b_q]\,; t) := \sum_{j\geq 0} \frac{(a_1)_j \cdots (a_p)_j}{(b_1)_j \cdots (b_q)_j} \frac{t^j}{j!}, $$

where $(x)_j := x(x+1)\cdots(x+j-1)$ is the rising factorial.

We are interested in this chapter in the following (dual) questions:

(i) What are the algebraic power series $g(\mathbf{x})$ whose diagonal $\mathrm{Diag}(g)$ is a generalized hypergeometric function ${}_pF_q$?[2]

(ii) What are the hypergeometric sequences $(a_j)_{j\geq 0}$ whose generating functions $\sum_{j\geq 0} a_j t^j$ can be written as diagonals of algebraic power series?

Note that from an algorithmic viewpoint, questions (i) and (ii) are very different in nature: while (i) is decidable (given an algebraic power series, one can decide if its diagonal is hypergeometric, for instance by combining the algorithms in [71] and [262]), the status of question (ii) is not known (does there exist an algorithm which takes as input a hypergeometric sequence and outputs an algebraic series whose diagonal is the generating function of the input sequence?).

Already for $n \in \{1, 2\}$ the above questions are non-trivial. The classes of diagonals of bivariate rational power series and of algebraic power series coincide [267, 145]. Hence, questions (i) and (ii) contain as a sub-question the characterization of algebraic hypergeometric functions. This problem was only recently solved in a famous paper by Beukers and Heckman [38].

Another motivation for studying questions (i) and (ii) comes from the well-known conjecture below, formulated in [99, 101] by Christol. Recall that $f \in \mathbb{Q}[\![t]\!]$ is called *globally bounded* if it has finite non-zero radius of convergence and $\beta \cdot f(\alpha \cdot t) \in \mathbb{Z}[\![t]\!]$ for some non-zero $\alpha, \beta \in \mathbb{Z}$.

> **Christol's conjecture.** If $f \in \mathbb{Q}[\![t]\!]$ is D-finite and globally bounded, then $f = \mathrm{Diag}(g)$ for some $n \in \mathbb{N}$ and some algebraic power series $g \in \mathbb{Q}[\![x_1, \ldots, x_n]\!]$.

[2] Note that a necessary condition is that $q = p - 1$, since the radius of convergence must be finite and non-zero.

Christol's conjecture is still largely open, even in the particular case when $f$ is a generalized hypergeometric function. In this case, it has been proved [99, 101] in two extreme subcases: when all the bottom parameters $b_i$ are integers (case of "minimal monodromy weight", in the terminology of [102]) and when they are all non-integers (case of "maximal monodromy weight"). In the first extremal case, the proof is based on the observation that

$$ {}_pF_q([a_1,\dots,a_p]\,;[1,\dots,1]\,;t) = (1-t)^{-a_1} \star \cdots \star (1-t)^{-a_p}, \tag{2.1}$$

where $\star$ denotes the Hadamard (term-wise) product, and on the fact that diagonals are closed under Hadamard product [100, Prop. 2.6]. In the second extremal case, it is based on the equivalence between being globally bounded and algebraic; this equivalence, proved by Christol [99, 101], is itself based on [38].

The other cases (of "intermediate monodromy weight") are widely open. A first explicit example of this kind, itself still open as of today, was given by Christol himself as soon as 1987 [99, §VII]:

Is $f(t) = {}_3F_2\left(\left[\tfrac{1}{9},\tfrac{4}{9},\tfrac{5}{9}\right]\,;\left[1,\tfrac{1}{3}\right]\,;t\right)$ the diagonal of an algebraic power series?

Two decades later, Bostan et al. [52, 53] produced a large list of about 100 similar ${}_3F_2$ (globally bounded) functions, which are potential counter-examples to Christol's conjecture (in the sense that, like ${}_3F_2([1/9,4/9,5/9]\,;[1,1/3]\,;t)$, they are not easily reducible to the two known extreme cases, via closure properties of diagonals, e.g., with respect to Hadamard products). In 2020, Abdelaziz, Koutschan and Maillard [1, §3] managed to show that two members of that list, namely

$$ {}_3F_2([1/9,4/9,7/9]\,;[1,1/3]\,;t) \quad \text{and} \quad {}_3F_2([2/9,5/9,8/9]\,;[1,2/3]\,;t)$$

are indeed diagonals. Precisely,

$$ {}_3F_2\left(\left[\frac{2}{9},\frac{5}{9},\frac{8}{9}\right]\,;\left[1,\frac{2}{3}\right]\,;27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-y)^{1/3}}{1-x-y-z}\right) \tag{2.2}$$

and

$$ {}_3F_2\left(\left[\frac{1}{9},\frac{4}{9},\frac{7}{9}\right]\,;\left[1,\frac{1}{3}\right]\,;27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-y)^{2/3}}{1-x-y-z}\right). \tag{2.3}$$

Section 3 of [1] also contains the following extension of (2.2) and (2.3), to any $R \in \mathbb{Q}$:

$$ {}_3F_2\left(\left[\frac{1-R}{3},\frac{2-R}{3},\frac{3-R}{3}\right]\,;[1,1-R]\,;27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-y)^{R}}{1-x-y-z}\right). \tag{2.4}$$

A common feature of identities (2.2) and (2.3) (and their extension (2.4)) is that the top parameters are in arithmetic progression, as opposed to Christol's initial example. However, they are the first known examples of generalized hypergeometric functions with intermediate monodromy weight, not trivially reducible to the two known extreme cases, and which are provably diagonals.

Our first result extends identity (2.4) to a much larger class of (transcendental) generalized hypergeometric functions.

**Theorem 2.1.** *Let $R, S \in \mathbb{Q}$ and $n, N \in \mathbb{N}$ such that $S \neq 0$ and $0 \leq n \leq N$. Set $s := N - n$ and $Q := S - R$. Then the generalized hypergeometric function*

$$ {}_{N+s}F_{N+s-1}\Big(\left[\frac{Q}{N}, \frac{Q+1}{N}, \ldots, \frac{Q+N-1}{N}, \frac{S}{s}, \ldots, \frac{S+s-1}{s}\right]; \\ \left[\frac{Q}{s}, \ldots, \frac{Q+s-1}{s}, 1, \ldots, 1\right]; N^N t\Big) $$

*is equal to the diagonal*

$$ \operatorname{Diag}\left(\frac{(1-x_1-\cdots-x_n)^R}{(1-x_1-\cdots-x_N)^S}\right). $$

Note that identity (2.4) corresponds to the particular case $(n, N, S) = (2, 3, 1)$ of Theorem 2.1. The proof of (2.4) given in [1, §3.2] relies on an algorithmic technique called *creative telescoping* [214], which works in principle[3] on any diagonal of an algebraic function, as long as the number $\max(n, N)$ of indeterminates is *fixed*. Our identity in Theorem 2.1 contains a number of indeterminates which is itself variable, hence it cannot be proved by creative telescoping in this generality. In §2.2.5 we offer instead a direct and elementary proof of a natural generalization. We note that Theorem 2.1 can also be proven by directly multiplying out the argument of the diagonal using the multinomial theorem, collecting needed terms and simplifying using the Chu-Vandermonde identity.

Note that in the theorem above, and similarly in later statements, the restriction on $R$ and $S$ to be rational numbers is actually superfluous. Indeed, from the proofs it is obvious that the identities hold for arbitrary (formal) parameters $R, S$; however, we include this condition because we wish $(1-x_1-\cdots-x_n)^R(1-x_1-\cdots-x_N)^{-S}$ to be an algebraic function.

In §2.2 we will further generalize Theorem 2.1 in two distinct directions. The first extension (Theorem 2.3) shows that the diagonal of the product of an arbitrary number of arbitrary powers of linear forms of the type $1-x_1-\cdots-x_m$ is again a generalized hypergeometric function. The second extension (Theorem 2.4) shows that under a condition on the exponents the same stays true if the product is multiplied with another factor of the form $(1-x_1-\cdots-x_{m-2}-2\,x_{m-1})^b$. For instance, when restricted to $m = 3$ variables, these results specialize as follows:

**Theorem 2.2.** *For any $R, S, T \in \mathbb{Q}$, we have:*

$$ \operatorname{Diag}\left((1-x)^R(1-x-y)^S(1-x-y-z)^T\right) = \tag{2.5} $$
$$ {}_6F_5\Big(\left[\frac{-(R+S+T)}{3}, \frac{1-(R+S+T)}{3}, \frac{2-(R+S+T)}{3}, \frac{-(S+T)}{2}, \frac{1-(S+T)}{2}, -T\right]; \\ \left[\frac{-(R+S+T)}{2}, \frac{1-(R+S+T)}{2}, -(S+T), 1, 1\right]; 27t\Big) $$

[3] Creative telescoping algorithms, such as the one in [71], compute a linear differential equation for $\operatorname{Diag}(g(\mathbf{x}))$. This equation is converted to a linear recurrence, whose hypergeometric solutions can be computed using Petkovšek's algorithm [262]. Note that the complexity (in time and space) of these algorithms increases with $n, N, R$ and $S$.

*and*

$$\operatorname{Diag}\left((1-x)^R(1-x-2\,y)^S(1-x-y-z)^{-1}\right) = \tag{2.6}$$
$${}_4F_3\left(\left[\frac{1-(R+S)}{3},\frac{2-(R+S)}{3},\frac{3-(R+S)}{3},\frac{1-S}{2}\right];\left[\frac{1-(R+S)}{2},\frac{2-(R+S)}{2},1\right];27t\right).$$

Note that (2.6) generalizes and explains the two identities observed in [1, Eq. (30)–(31)]

$${}_3F_2\left(\left[\frac{1}{9},\frac{4}{9},\frac{7}{9}\right];\left[1,\frac{2}{3}\right];27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-2\,y)^{2/3}}{1-x-y-z}\right) \tag{2.7}$$

and

$${}_3F_2\left(\left[\frac{2}{9},\frac{5}{9},\frac{8}{9}\right];\left[1,\frac{5}{6}\right];27\,t\right) = \operatorname{Diag}\left(\frac{(1-x-2\,y)^{1/3}}{1-x-y-z}\right). \tag{2.8}$$

Once again, our proofs of the (generalizations of) identities (2.5) and (2.6) are elementary, and do not rely on algorithmic tools.

One may wonder if other generalizations are possible, for instance whether the coefficient 2 can be replaced by a different one in (2.6). The following example shows that this is not the case. Let

$$U(t) = \operatorname{Diag}\left(\frac{\sqrt[3]{1-ax}}{1-x-y}\right),$$

then the coefficients sequence $(u_j)_{j\geq 0}$ of $U(t)$ satisfies the second-order recurrence relation

$$\begin{aligned}2a^2\,(6n+5)\,(3n+1)\,u_n - 3\,(n+1)\left(3\left(a^2+4a-4\right)n+2a^2+18a-18\right)u_{n+1}\\ = 9\,(1-a)\,(n+2)\,(n+1)\,u_{n+2}.\end{aligned}$$

When $a \in \{0,1,2\}$, the sequence $(u_j)_{j\geq 0}$ also satisfies a shorter recurrence, of order 1, as shown by our main results. In these cases, $U(t)$ is a hypergeometric function. When $a \notin \{0,1,2\}$, the second-order recurrence is the minimal-order satisfied by $(u_j)_{j\geq 0}$, hence $U(t)$ is not a hypergeometric function. This can be proved either using the explicit identity

$$U(t) = \sqrt[3]{\frac{a/2}{1-4\,t} + \frac{1-a/2}{(1-4\,t)^{\frac{3}{2}}}},$$

or by using the general approach in [85, §5].

An apparent weakness of our results is that they only provide examples with parameters in (unions of) arithmetic progressions. This is true, as long as identities are used alone. But symmetries may be broken by combining different identities and using for instance Hadamard products. As an illustration, by taking the Hadamard product in both sides of the following identities

$${}_3F_2\left(\left[\frac{Q}{3},\frac{Q+1}{3},\frac{Q+2}{3}\right];[1,Q];t\right) = \operatorname{Diag}\left(\frac{(1-\frac{x_1}{3}-\frac{x_2}{3})^{1-Q}}{1-\frac{x_1}{3}-\frac{x_2}{3}-\frac{x_3}{3}}\right)$$

and

$${}_2F_1\left(\left[\frac{Q}{6},\frac{Q+3}{6}\right];\left[\frac{Q}{3}\right];t\right) = \operatorname{Diag}\left(\frac{(1-\frac{x_4}{2})^{1-Q/3}}{1-\frac{x_4}{2}-\frac{x_5}{2}}\right),$$

both particular cases of Theorem 2.1, one deduces that the *non-symmetric* hypergeometric function

$$ {}_4F_3\left(\left[\frac{Q}{6}, \frac{Q+3}{6}, \frac{Q+1}{3}, \frac{Q+2}{3}\right]; [1,1,Q]; t\right) $$

is equal to the diagonal

$$ \mathrm{Diag}\left(\left(1-\frac{x_1}{3}-\frac{x_2}{3}\right)^{1-Q}\left(1-\frac{x_4}{2}\right)^{1-\frac{Q}{3}}\left(1-\frac{x_1}{3}-\frac{x_2}{3}-\frac{x_3}{3}\right)^{-1}\left(1-\frac{x_4}{2}-\frac{x_5}{2}\right)^{-1}\right). $$

Similarly, the *non-symmetric* hypergeometric function[4]

$$ {}_3F_2\left(\left[\frac{Q}{6}, \frac{Q+3}{6}, \frac{2Q+3}{6}\right]; \left[1, \frac{2\,Q}{3}\right]; t\right) $$

is equal to the diagonal

$$ \mathrm{Diag}\left(\left(1-\frac{x_1}{2}\right)^{1-\frac{Q}{3}}\left(1-\frac{x_3}{2}\right)^{1-\frac{2\,Q}{3}}\left(1-\frac{x_1}{2}-\frac{x_2}{2}\right)^{-1}\left(1-\frac{x_3}{2}-\frac{x_4}{2}\right)^{-1}\right). $$

A natural challenge is to prove (or disprove) that Christol's ${}_3F_2$ can be obtained in this way.

As a final remark, one should not think that every generalized hypergeometric function which is a diagonal needs to have a representation like in our Theorems 2.1 or 2.2. For instance, the diagonal $\mathrm{Diag}\left((1-(1+w)(x+y+z))^{-1}\right)$ is equal [54] to the generalized hypergeometric function

$$ {}_4F_3\left(\left[\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right]; \left[1, 1, \frac{1}{2}\right]; \frac{729}{4}\,t\right) = 1 + 18\,t + 1350\,t^2 + \cdots, $$

which is seemingly not of the form covered by any of our results.

## 2.2 General case

This section contains several parts: first we introduce in §2.2.1 and §2.2.2 some notation and state the two general Theorems 2.3 and 2.4. Then we explain them in §2.2.3 by means of four examples, showing that both Theorem 2.1 and Theorem 2.2 are special cases. Further, we continue in §2.2.4 with several lemmas and their proofs. Finally, the general theorems are proven in §2.2.5 and §2.2.6.

### 2.2.1 First Statement

Let $N \in \mathbb{N} \setminus \{0\}$ and $b_1, \ldots, b_N \in \mathbb{Q}$ with $b_N \neq 0$. We want to prove that the diagonal of

$$ R(x_1, \ldots, x_N) := (1+x_1)^{b_1}(1+x_1+x_2)^{b_2}\cdots(1+x_1+\cdots+x_N)^{b_N} \tag{2.9} $$

[4]Amusingly, the above ${}_3F_2$ is not only asymmetric, but it also shares another similarity with Christol's example: the sum of two of the three top parameters is equal to the third one. This pattern occurs in several other examples.

can be expressed as a hypergeometric function. For each $k = 1, \dots, N$ we define the tuple

$$u^k \coloneqq \left( \frac{B(k)}{N-k+1}, \frac{B(k)+1}{N-k+1}, \dots, \frac{B(k)+N-k}{N-k+1} \right),$$

where $B(k) \coloneqq -(b_k + \cdots + b_N)$. For $k = 1, \dots, N-1$ we set

$$v^k \coloneqq \left( \frac{B(k)}{N-k}, \frac{B(k)+1}{N-k}, \dots, \frac{B(k)+N-k-1}{N-k} \right).$$

Moreover set $v^N \coloneqq (1, 1, \dots, 1)$ with exactly $N-1$ ones. It follows by construction that the lengths of the tuples

$$u \coloneqq (u^1, \dots, u^N) \quad \text{and} \quad v \coloneqq (v^1, \dots, v^N)$$

are given by $M \coloneqq N + \cdots + 2 + 1 = N(N+1)/2$ and $M-1$ respectively. We have the following generalization of Theorem 2.1:

**Theorem 2.3.** *It holds that*

$$\mathrm{Diag}(R(x_1, \dots, x_N)) = {}_M F_{M-1}([u]\,;[v]\,;(-N)^N t).$$

### 2.2.2 Second Statement

Let $N \in \mathbb{N} \setminus \{0\}$ and $b_1, \dots, b_N \in \mathbb{Q}$. Assume that $b_N \neq 0$ and $b_{N-1} + b_N = -1$. We will prove that, for any $b \in \mathbb{Q}$, we can express

$$(1 + x_1 + \cdots + x_{N-2} + 2\,x_{N-1})^b \cdot R(x_1, \dots, x_N)$$

as a hypergeometric function as well. Again, let $B(k) \coloneqq -(b_k + \cdots + b_N)$. For each $k = 1, \dots, N-2$ we define the tuple

$$\tilde{u}^k \coloneqq \left( \frac{B(k)-b}{N-k+1}, \frac{B(k)-b+1}{N-k+1}, \dots, \frac{B(k)-b+N-k}{N-k+1} \right)$$

and set $\tilde{u}^{N-1} \coloneqq -(b_{N-1} + b_N + b)/2 = (1-b)/2$ and $\tilde{u}^N \coloneqq -b_N$. Moreover, for $k = 1, \dots, N-2$ we set

$$\tilde{v}^k \coloneqq \left( \frac{B(k)-b}{N-k}, \frac{B(k)-b+1}{N-k}, \dots, \frac{B(k)-b+N-k-1}{N-k} \right),$$

and $\tilde{v}^{N-1} \coloneqq (1, 1, \dots, 1)$ with exactly $N-1$ ones. It follows by construction that the lengths of the tuples

$$\tilde{u} \coloneqq (\tilde{u}^1, \dots, \tilde{u}^N) \quad \text{and} \quad \tilde{v} \coloneqq (\tilde{v}^1, \dots, \tilde{v}^{N-1})$$

are given by $M - 1 = N + \cdots + 4 + 3 + 1 + 1 = N(N+1)/2 - 1$ and $M-2$ respectively.

**Theorem 2.4.** *It holds that*

$$\mathrm{Diag}((1 + x_1 + \cdots + x_{N-2} + 2\,x_{N-1})^b \cdot R(x_1, \dots, x_N)) = {}_{M-1} F_{M-2}([\tilde{u}]\,;[\tilde{v}]\,;(-N)^N t).$$

### 2.2.3 Examples

Let us list some examples of the general theorems and draw the connection to previous statements.

1. First we emphasize that Theorem 2.1 follows promptly from the more general Theorem 2.3 by letting all $b_j = 0$ except $b_n = R$ and $b_N = -S$. Clearly, the change $\mathbf{x} \mapsto -\mathbf{x}$ in the algebraic function is reflected by the change $t \mapsto (-1)^N t$ in its diagonal.

2. Letting $N = 3$ in Theorem 2.3 we obtain immediately the first part of Theorem 2.2.

3. If moreover $T = -1$ in the case $N = 3$, we achieve a cancellation of the last parameter and are left with

$$\operatorname{Diag}\left(\frac{(1+x)^R(1+x+y)^S}{1+x+y+z}\right) = {}_5F_4\left(\left[\frac{1-(R+S)}{3}, \frac{2-(R+S)}{3}, \frac{3-(R+S)}{3}, \frac{1-S}{2}, \frac{2-S}{2}\right]; \left[\frac{1-(R+S)}{2}, \frac{2-(R+S)}{2}, 1-S, 1\right]; -27t\right).$$

4. Comparing with the similar situation of Theorem 2.4 in the case $N = 3$ and $b_{N-1} = -1 - b_N = 0$, we see that a family of ${}_4F_3$ functions remains and covers the second statement of Theorem 2.2.

### 2.2.4 Lemmas and Proofs

In this section we will state and prove necessary lemmas for the proofs of Theorems 2.3 and 2.4.

**Lemma 2.5.** *Let $N$ be a positive integer and $b_1, \dots, b_N \in \mathbb{Q}$ such that $b_N \neq 0$. It holds that*

$$[x_1^{k_1} \cdots x_N^{k_N}](1+x_1)^{b_1}(1+x_1+x_2)^{b_2} \cdots (1+x_1+\cdots+x_N)^{b_N} = \binom{b_N}{k_N}\binom{b_{N-1}+b_N-k_N}{k_{N-1}} \cdots \binom{b_1+\cdots+b_N-k_N\cdots-k_2}{k_1}.$$

This result contains the core identity of the present paper, since it enables the connection between the algebraic functions $R(\mathbf{x})$ of the form (2.9) and hypergeometric sequences. It can be proven in two ways: a direct approach works by multiplying the left-hand side out using the multinomial theorem, picking the needed coefficient and reducing the sum using the Chu-Vandermonde identity several times. This procedure is rather tedious and not instructive, therefore we present a combinatorially inspired proof.

*Proof.* Because $(1+x_1)^{b_1} \cdots (1+x_1+\cdots+x_{N-1})^{b_{N-1}}$ does not depend on $x_N$, we obtain

that the left-hand side of the equation is equal to

$$[x_1^{k_1}\cdots x_N^{k_N}]\prod_{i=1}^{N-1}\left(1+\sum_{j=1}^{i}x_j\right)^{b_i}(1+x_1+\cdots+x_{N-1})^{b_N}\left(1+\frac{x_N}{1+x_1+\cdots+x_{N-1}}\right)^{b_N}$$
$$=[x_1^{k_1}\cdots x_N^{k_N}]\prod_{i=1}^{N-1}\left(1+\sum_{j=1}^{i}x_j\right)^{b_i}(1+x_1+\cdots+x_{N-1})^{b_N}\sum_{k\geq 0}\binom{b_N}{k}\left(\frac{x_N}{1+x_1+\cdots+x_{N-1}}\right)^{k}$$
$$=\binom{b_N}{k_N}[x_1^{k_1}\cdots x_{N-1}^{k_{N-1}}](1+x_1)^{b_1}\cdots(1+x_1+\cdots+x_{N-2})^{b_{N-2}}(1+x_1+\cdots+x_{N-1})^{b_{N-1}+b_N-k_N}.$$

Now the claim follows by iteration. □

We remark that the same strategy as above can be used to prove an even more general statement. Let

$$R(\mathbf{x})=\prod_{i=1}^{N}\left(1+\sum_{j\in I_i}x_j\right)^{b_i},$$

for rational numbers $b_1,\ldots,b_N$, such that all variables $x_1,\ldots,x_N$ appear in $R(\mathbf{x})$, and sets $I_1,\ldots,I_N\subseteq\{1,\ldots,N\}$ with the property that $I_1\cup\cdots\cup I_{n-1}\subsetneq I_1\cup\cdots\cup I_{n-1}\cup I_n$ for all $n=1,\ldots,N$. Then, similarly to the statement in Lemma 2.5, the coefficient of $x^{k_1}\cdots x^{k_N}$ in $R(\mathbf{x})$ is a product of binomial coefficients and the diagonal of $R(\mathbf{x})$ is a generalized hypergeometric function. The notation, however, becomes quite cumbersome in this setting and no new ideas are needed; therefore we stick to the more insightful but less general case $I_n=\{1,\ldots,n\}$.

Note that Lemma 2.5 shares some similarities with Straub's Theorem 3.1 in [308], which provides explicit expressions of rational power series of the form

$$\left((1+x_1+\cdots+x_{\lambda_1})(1+x_{\lambda_1+1}+\cdots+x_{\lambda_1+\lambda_2})\cdots(1+x_{\lambda_1+\cdots+\lambda_{\ell-1}}+\cdots+x_N)-\alpha\cdot x_1x_2\cdots x_N\right)^{-1}.$$

In Lemma 2.5, we allow products of linear forms with arbitrary exponents however there is no term $\alpha\cdot x_1x_2\cdots x_N$, while in [308, Theorem 3.1] the linear forms have disjoint variables and appear with exponent 1. Setting $\alpha=0$ in Straub's formula also yields a product of binomial coefficients.

It is legitimate to wonder whether there is a common generalization of Lemma 2.5 and Thm. 3.1 in [308]. For instance, one may ask for which values of $\alpha$ is the diagonal

$$\operatorname{Diag}\left(\left(\sqrt{1-x}\,(1-y)-\alpha xy\right)^{-1}\right)=1+(\alpha+1/2)\,t+\left(\alpha^2+2\,\alpha+3/8\right)t^2+\cdots$$

hypergeometric? For a general $\alpha$, the minimal recurrence satisfied by the coefficients of the diagonal is of order 4, for $\alpha=\pm i/2$ it is of order 3, and it seems that the only rational value of $\alpha$ for which there exists a shorter recurrence is $\alpha=0$, in which case the diagonal is hypergeometric.

Now we want to verify a similar statement for the situation as in Theorem 2.4, so the case where we deal with the coefficient sequence of $(1+x_1+\cdots+x_{N-2}+2\,x_{N-1})^b\cdot R(x_1,\ldots,x_N)$. We lay the grounds for a lemma similar to Lemma 2.5, by starting with a rather surprising identity.

**Lemma 2.6.** *Let $k \in \mathbb{N}$ and $b \in \mathbb{Q}$ arbitrary. It holds that*

$$[x^k]\frac{(1+2x)^b}{(1+x)^{k+1}} = 4^k\binom{(b-1)/2}{k}.$$

*Proof.* First notice that for arbitrary $a, b$ we can compute

$$[x^k](1+2x)^b(1+x)^a = [x^k]\left(\sum_{i\geq 0} 2^i\binom{b}{i}x^i\right)\left(\sum_{j\geq 0}\binom{a}{j}x^j\right) = \sum_{j=0}^{k} 2^j\binom{b}{j}\binom{a}{k-j}.$$

So we set $a = -(k+1)$ and obtain

$$[x^k]\frac{(1+2x)^b}{(1+x)^{k+1}} = \sum_{j=0}^{k} 2^j\binom{b}{j}\binom{-k-1}{k-j} = 2^k\sum_{j=0}^{k}(-1)^j 2^{-j}\binom{b}{k-j}\binom{k+j}{k}.$$

It remains to prove the following identity[5]

$$\sum_{j=0}^{k}(-2)^{-j}\binom{b}{k-j}\binom{k+j}{k} = 2^k\binom{(b-1)/2}{k}. \tag{2.10}$$

To do this, we note that

$$\sum_{j=0}^{k}(-1)^j\binom{b}{k-j}\binom{k+j}{k}u^j = \binom{b}{k}{}_2F_1([-k, k+1]\,;[b+1-k]\,;u),$$

and

$${}_2F_1([-k, k+1]\,;[b+1-k]\,;1/2) = \frac{\Gamma((b+1-k)/2)\Gamma((b+2-k)/2)}{\Gamma((b+1-2k)/2)\Gamma((b+2)/2)} = 2^k\frac{\binom{(b-1)/2}{k}}{\binom{b}{k}},$$

by Kummer's identity [217, Eq. 3, p. 134]. □

The proof above explains the special role of the coefficient $a = 2$ mentioned in the introduction: it is one of the few values, along with $1$ and $-1$, for which there exists a closed form expression for the evaluation of a ${}_2F_1([\alpha, 1-\alpha]\,;[\gamma]\,;u)$ at $u = 1/a$.

Now we can proceed and prove the essential lemma for Theorem 2.4. Note that contrary to Lemma 2.5 the following statement is purely about diagonal coefficients and not for general exponents anymore. Except for the missing factor $\binom{b_{N-1}+b_N-k}{k}$ and the new two factors $4^k$ and $\binom{(b-1)/2}{k}$ the formulas are completely analogous.

**Lemma 2.7.** *Let $N$ be a positive integer and $b_1, \ldots, b_N \in \mathbb{Q}$ such that $b_{N-1} + b_N = -1$. For any $b \in \mathbb{Q}$ the coefficient of $x_1^k \cdots x_N^k$ in*

$$(1 + x_1 + \cdots + x_{N-2} + 2\,x_{N-1})^b \cdot (1+x_1)^{b_1}\cdots(1+x_1+\cdots+x_N)^{b_N}$$

*is given by*

$$4^k\binom{(b-1)/2}{k}\binom{b_N}{k}\cdot\binom{b_{N-2}+b_{N-1}+b_N+b-2k}{k}\cdots\binom{b_1+\cdots+b_N+b-(N-1)k}{k}.$$

[5]Note that identity (2.10) could alternatively be proven by using Zeilberger's creative telescoping algorithm [341], or derived from identity (3.42) in [163, p. 27] by setting $2n - x = b$, multiplying with $2^k$ and reverting the summation.

*Proof.* By the same argument as in the proof of Lemma 2.5, the left-hand side is equal to $\binom{b_N}{k}$ multiplied with the coefficient of $x_1^k \cdots x_{N-1}^k$ in

$$(1 + x_1 + \cdots + x_{N-2} + 2\,x_{N-1})^b \prod_{j=1}^{N-2} \left(1 + \sum_{i=1}^{j} x_i\right)^{b_j} (1 + x_1 + \cdots + x_{N-2} + x_{N-1})^{b_{N-1}+b_N-k}.$$

Because the product in the middle does not depend on $x_{N-1}$ and since we assumed that $b_{N-1} + b_N = -1$, we can first compute

$$[x_{N-1}^k] \left(1 + 2\,\frac{x_{N-1}}{1 + x_1 + \cdots + x_{N-2}}\right)^b \left(1 + \frac{x_{N-1}}{1 + x_1 + \cdots + x_{N-2}}\right)^{-1-k} = 4^k \binom{(b-1)/2}{k} (1 + x_1 + \cdots + x_{N-2})^{-k},$$

by Lemma 2.6. Therefore we are left with

$$4^k \binom{(b-1)/2}{k} \binom{b_N}{k} \cdot [x_1^k \cdots x_{N-2}^k] \prod_{j=1}^{N-3} \left(1 + \sum_{i=1}^{j} x_i\right)^{b_j} \cdot (1 + x_1 + \cdots + x_{N-2})^{b_{N-2}+b-2k-1},$$

which is easily computed using Lemma 2.5. □

Note that the requirement $b_{N-1} + b_N = -1$ comes from the $+1$ in the denominator of the left-hand side in Lemma 2.6. Since this identity is itself surprising and does not allow for obvious generalizations, the condition on the relationship of $b_{N-1}$ and $b_N$ is necessary.

### 2.2.5 Proof of Theorem 2.3

For the proof of Theorem 2.3 we will only use Lemma 2.5 and algebraic manipulations similar to the proof of Bober's Lemma 4.1 in [46].

By Lemma 2.5 we obtain the coefficient of $t^n$ for any $n \in \mathbb{N}$ on the left-hand side:

$$[t^n]\mathrm{Diag}((1 + x_1)^{b_1} \cdots (1 + x_1 + \cdots + x_N)^{b_N}) = \binom{b_N}{n} \cdots \binom{b_1 + \cdots + b_N - (N-1)n}{n}.$$

For the right-hand side we use the fact that for all $a, b$ and non-negative integers $n$ it holds

$$(a/b)_n((a+1)/b)_n \cdots ((a+b-1)/b)_n \cdot b^{bn} = (a)_{bn}.$$

Then

$$U_k := \prod_{i=1}^{N-k+1} (u_i^k)_n = \frac{(-b_k - \cdots - b_N)_{(N-k+1)n}}{(N-k+1)^{(N-k+1)n}},$$

for all $k = 1, \ldots, N$. Similarly,

$$V_k := \prod_{i=1}^{N-k} (v_i^k)_n = \frac{(-b_k - \cdots - b_N)_{(N-k)n}}{(N-k)^{(N-k)n}},$$

for all $k = 1, \dots, N-1$. Clearly $V_N \coloneqq \prod_{i=1}^{N-1}(v_i^N)_n = (n!)^{N-1}$. We deduce that

$$[t^n]_M F_{M-1}([u]\,;[v]\,;(-N)^N t) = \frac{(-N)^{nN}}{n!}\prod_{i=1}^{N}\frac{U_i}{V_i} = \frac{(-1)^{nN}}{(n!)^N}\prod_{i=1}^{N}\frac{(-b_i-\cdots-b_N)_{(N-i+1)n}}{(-b_i-\cdots-b_N)_{(N-i)n}}$$
$$= \frac{(-1)^{nN}}{(n!)^N}\prod_{i=1}^{N}(-b_i-\cdots-b_N+(N-i)n)_n.$$

The claim of Theorem 2.3 follows from the fact that

$$(-1)^n\frac{(-b_k-\cdots-b_N+(N-k)n)_n}{n!} = (-1)^n\binom{-b_k-\cdots-b_N+(N-k+1)n-1}{n}$$
$$= \binom{b_k+\cdots+b_N-(N-k)n}{n}. \qquad \square$$

### 2.2.6 Proof of Theorem 2.4

The proof of Theorem 2.4 is very similar: we will use Lemma 2.7 and the same reasoning as before. The only difference lies in the fact that because the hypergeometric function has one parameter less, we need to redefine $U_{N-1}, V_{N-1}$ and $V_N$. Recall that the denominator of $U_k$ was given by $(N-k+1)^{(N-k+1)n}$ and it cancelled with the denominator of $V_{k-1}$. In the present case, $\tilde{U}_{N-1}$ will have no denominator and therefore $2^{2n}$ from $\tilde{V}_{N-2}$ survives. This fits with the $4^k$ in the statement of Lemma 2.7 and is another indicator for the importance and essence of the constant $a = 2$.

Using Lemma 2.7 we obtain the coefficient of $t^n$ for any $n \in \mathbb{N}$ on the left-hand side:

$$[t^n]\mathrm{Diag}((1+x_1+\cdots+2\,x_{N-1}+x_N)^b R(x_1,\dots,x_N))$$
$$= 4^n\binom{(b-1)/2}{n}\binom{b_N}{n}\cdot\binom{b_{N-2}+b_{N-1}+b_N+b-2n}{n}\cdots\binom{b_1+\cdots+b_N+b-(N-1)n}{n}.$$

By the same reasoning as before, we have for all $k = 1, \dots, N-2, N$

$$\tilde{U}_k \coloneqq \prod_{i=1}^{N-k+1}(\tilde{u}_i^k)_n = \frac{(-b_k-\cdots-b_N-b)_{(N-k+1)n}}{(N-k+1)^{(N-k+1)n}},$$

and similarly

$$\tilde{V}_k \coloneqq \prod_{i=1}^{N-k}(\tilde{v}_i^k)_n = \frac{(-b_k-\cdots-b_N-b)_{(N-k)n}}{(N-k)^{(N-k)n}},$$

for $k = 1, \dots, N-2$. Clearly $\tilde{V}_{N-1} \coloneqq \prod_{i=1}^{N-1}(\tilde{v}_i^{N-1})_n = (n!)^{N-1}$ and we set $\tilde{V}_N \coloneqq 1$.

Moreover, this time we have $\tilde{U}_{N-1} \coloneqq (\tilde{u}^{N-1})_n = ((1-b)/2)_n$. Altogether, we find

$$\begin{aligned}
[t^n]_M F_{M-1}([\tilde{u}]\,;[\tilde{v}]\,;(-N)^N t) &= \frac{(-N)^{nN}}{n!}\prod_{i=1}^{N}\frac{\tilde{U}_i}{\tilde{V}_i}\\
&= \frac{(-1)^{nN}2^{2n}((1-b)/2)_n}{(n!)^N}\prod_{i=1}^{N-2}\frac{(-b_i-\cdots-b_N-b)_{(N-i+1)n}}{(-b_i-\cdots-b_N-b)_{(N-i)n}}\cdot\frac{(-b_N-b)_n}{1}\\
&= 4^n\frac{(-1)^n((1-b)/2)_n}{n!}\cdot\prod_{i=1}^{N-2}\frac{(-1)^n(-b_i-\cdots-b_N+(N-i)n)_n}{n!}\cdot\frac{(-1)^n(-b_N-b)_n}{n!}.
\end{aligned}$$

Using the same final observation as before we conclude the proof. □

## 2.3 Algebraicity and Hadamard grade

### 2.3.1 Algebraic cases

We address here the following question: given $b_1,\dots,b_N\in\mathbb{Q}$, $b_N\neq 0$, is the diagonal

$$\mathrm{Diag}(R(\mathbf{x})) = \mathrm{Diag}((1+x_1)^{b_1}\cdots(1+x_1+\cdots+x_N)^{b_N})$$

an algebraic function?

**Corollary 2.8.** $\mathrm{Diag}(R(\mathbf{x}))$ *is algebraic if and only if* $N=2$ *and* $b_2\in\mathbb{Z}$*, or* $N=1$.

In the proof below we will use several times the following useful fact [102, Thm. 33]: if a generalized hypergeometric function is algebraic, then its monodromy weight is zero, that is the number of integer bottom parameters is at most equal to the number of integer top parameters.

*Proof.* By Theorem 2.3 it is sufficient to study the algebraicity of the generalized hypergeometric function $H(t)$ defined by

$${}_{N(N+1)/2}F_{N(N+1)/2-1}\left([u^1,\dots,u^{N-1},-b_N]\,;[v^1,\dots,v^{N-1},1,1,\dots,1]\,;t\right),$$

where $u^k$ and $v^k$, $k=1,\dots,N-1$ are defined like in §2.2.1:

$$u^k \coloneqq \left(\frac{b}{\ell+1},\frac{b+1}{\ell+1},\dots,\frac{b+\ell}{\ell+1}\right)\quad\text{and}\quad v^k\coloneqq\left(\frac{b}{\ell},\frac{b+1}{\ell},\dots,\frac{b+\ell-1}{\ell}\right)$$

for $b=-(b_k+b_2+\cdots+b_N)$ and $\ell=N-k$.

By definition, $N-1$ of the bottom parameters are ones. We claim that each tuple $u^k$ contains at most one integer and if it does contain one, then $v^k$ does as well. From the definition of $u^k$ it follows that if some $u^k_i\in\mathbb{Z}$ then $b\in\mathbb{Z}$ and $b\equiv -i+1 \bmod \ell+1$. This shows that for any $k=1,\dots,N-1$ at most one $u^k_i\in\mathbb{Z}$. Because of the definition of $v^k$ we see that if $b\in\mathbb{Z}$ and $b\equiv i \bmod \ell$ for some $i\in\{1,\dots,\ell\}$, then $v^k_i\in\mathbb{Z}$. This proves the claim.

Henceforth, in order to introduce new integer parameters on the top, while not creating equally many on the bottom, it is only possible to choose $-b_N$ integer. Therefore, in order to achieve monodromy weight zero – a necessary condition for algebraicity of $H(t)$ – we need to have $N-1 \leq 1$. From the same argument it follows that in the case $N-1=1$, we need to have $-b_N \in \mathbb{Z}$.

Obviously for $N=1$ the diagonal is algebraic, so it remains to prove that, conversely, when $-b_2 =: S$ is an integer and $b_1 =: R \in \mathbb{Q}$ arbitrary, then the diagonal in

$$\operatorname{Diag}\left(\frac{(1-x_1)^R}{(1-x_1-x_2)^S}\right) = {}_3F_2\left(\left[\frac{S-R}{2}, \frac{S-R+1}{2}, S\right]; [1, S-R]; 4\,t\right). \tag{2.11}$$

is an algebraic function. If $R$ is an integer too, this follows from by [267, 145]. In the general case, one can rewrite the ${}_3F_2$ in (2.11) as the Hadamard product

$$\begin{aligned} {}_3F_2\left(\left[\frac{S-R}{2}, \frac{S-R+1}{2}, S\right]; [1, S-R]; t\right) = \\ {}_2F_1\left(\left[\frac{S-R}{2}, \frac{S-R+1}{2}\right]; [S-R]; t\right) \star (1-t)^{-S}. \end{aligned} \tag{2.12}$$

The ${}_2F_1$ is algebraic as it corresponds to Case I in Schwarz's table [288]. Since $S$ is an integer, $(1-t)^{-S}$ is a rational function. We conclude by applying Jungen's theorem [192, Thm. 8]: the Hadamard product of an algebraic and a rational function is algebraic, see also [300, Prop. 6.1.11]. □

### 2.3.2 Hadamard grade

Recall that the *Hadamard grade* [14] of a power series $S(t)$ is the least positive integer $h = h(S)$ such that $S(t)$ can be written as the Hadamard product of $h$ algebraic power series, or $\infty$ if no such product exists. Since algebraic power series are diagonals [145, §3], and diagonals are closed under Hadamard product [100, Prop. 2.6], any power series with finite Hadamard grade is a diagonal [14, Thm. 7]. Conversely, it is not clear whether diagonals always have finite Hadamard grade[6].

A natural question in relation with Corollary 2.8 is the following: given $b_1, \ldots, b_N \in \mathbb{Q}$, determine the Hadamard grade of $\operatorname{Diag}(R(\mathbf{x})) = \operatorname{Diag}((1+x_1)^{b_1} \cdots (1+x_1+\cdots+x_N)^{b_N})$, or at least decide if it is finite or not. We have have the following result as an application of the classification in [38] (in particular case 1 in Table 8.3) and our main theorem.

**Corollary 2.9.** *The Hadamard grade of* $\operatorname{Diag}(R(\mathbf{x}))$ *is at most* $N$.

*Proof.* Like in the proof of Corollary 2.8, we define

$$u^k := \left(\frac{b}{\ell+1}, \frac{b+1}{\ell+1}, \ldots, \frac{b+\ell}{\ell+1}\right) \quad \text{and} \quad v^k := \left(\frac{b}{\ell}, \frac{b+1}{\ell}, \ldots, \frac{b+\ell-1}{\ell}\right)$$

[6]There exist diagonals of any prescribed finite grade [273, Cor.1&2] assuming the Rohrlich–Lang conjecture [322, Conj.22]. If Christol's conjecture also holds, there exist diagonals of infinite grade [273, Prop. 1].

for $b = -(b_k + b_2 + \cdots + b_N)$, $\ell = N - k$ and $k = 1, \ldots, N-1$. Then by Theorem 2.3 it is sufficient to study the Hadamard grade of

$$H(t) = {}_{N(N+1)/2}F_{N(N+1)/2-1}\left([u^1, \ldots, u^{N-1}, -b_N]\,; [v^1, \ldots, v^{N-1}, 1, 1, \ldots, 1]\,; t\right).$$

Now notice that

$$H(t) = {}_NF_{N-1}([u^1]\,; [v^1]\,; t) \star \cdots \star {}_2F_1([u^{N-1}]\,; [v^{N-1}]\,; t) \star {}_1F_0([-b_N]\,; [\,]\,; t),$$

and that each hypergeometric function in the Hadamard product is algebraic [38, Thm. 7.1]. □

For instance, by Theorem 2.1, when $S = 1$, $R = 1/2$, the diagonal

$$\operatorname{Diag}\left(\frac{(1-x_1-x_2)^R}{(1-x_1-x_2-x_3)^S}\right)$$

is a transcendental ${}_2F_1$, which can be written as the Hadamard product of two algebraic functions

$${}_2F_1\left(\left[\frac{1}{6}, \frac{5}{6}\right]\,; \left[\frac{1}{2}\right]\,; t\right) \star (1-t)^{-1/2},$$

(the ${}_2F_1$ being algebraic by Schwarz's classification [288]) hence its Hadamard grade is 2.

Similarly, the diagonals from (2.2) and (2.3) have Hadamard grade 2 due to the identities

$${}_3F_2\left(\left[\frac{2}{9}, \frac{5}{9}, \frac{8}{9}\right]\,; \left[1, \frac{2}{3}\right]\,; t\right) = {}_3F_2\left(\left[\frac{2}{9}, \frac{5}{9}, \frac{8}{9}\right]\,; \left[\frac{1}{2}, \frac{2}{3}\right]\,; t\right) \star (1-t)^{-1/2}$$

and

$${}_3F_2\left(\left[\frac{1}{9}, \frac{4}{9}, \frac{7}{9}\right]\,; \left[1, \frac{1}{3}\right]\,; t\right) = {}_3F_2\left(\left[\frac{1}{9}, \frac{4}{9}, \frac{7}{9}\right]\,; \left[\frac{1}{2}, \frac{1}{3}\right]\,; t\right) \star (1-t)^{-1/2}$$

and to the fact that the two ${}_3F_2$'s on the right-hand side are algebraic by the interlacing criterion [38, Thm. 4.8]; see Figure 2.1 for a pictorial proof, where red points correspond to top parameters, and blue points to bottom parameters (and the additional parameter 1).

More generally, the diagonal from (2.4) has Hadamard grade 2 due to the identity

$${}_3F_2\left(\left[\frac{1-R}{3}, \frac{2-R}{3}, \frac{3-R}{3}\right]\,; [1, 1-R]\,; t\right) = {}_3F_2\left(\left[\frac{1-R}{3}, \frac{2-R}{3}, \frac{3-R}{3}\right]\,; \left[\frac{1}{2}, 1-R\right]\,; t\right) \star (1-t)^{-1/2},$$

since the ${}_3F_2$ on the right-hand side is an algebraic function for any $R \in \mathbb{Q}$ (with Fig. 2.1 replaced by a similar one, containing only interlacing blue right triangles and red equilateral triangles).

This observation provides an alternative (and probably the shortest) proof that the hypergeometric functions in (2.2), (2.3) and (2.4) are diagonals of algebraic functions. However, this viewpoint does not yield such a compact diagonal representation as found in [1] and in our main theorem.

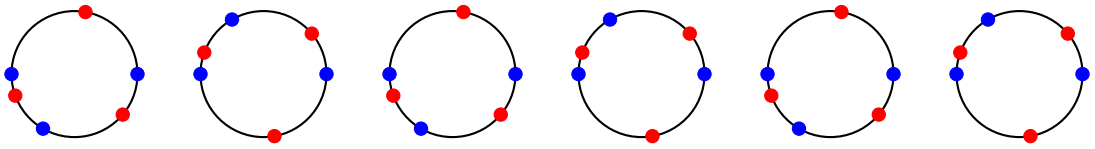


Figure 2.1: A pictorial proof of the algebraicity of ${}_3F_2\left(\left[\frac{1-R}{3}, \frac{2-R}{3}, \frac{3-R}{3}\right]\,;\left[\frac{1}{2}, 1-R\right]\,;t\right)$ for $R \in \{\frac{1}{3}, \frac{2}{3}\}$. There are $\varphi(18) = 6$ conditions to check, which lead to two distinct interlacing configurations.

The same observation also quickly solves two more cases amongst the 16 cases in the list [52, p. 58], namely those ${}_3F_2\left([N_1/9, N_2/9, N_3/9]\,;[1, M_1/3]\,;t\right)$ for which $(N_1, N_2, N_3; M_1)$ is $(1, 4, 7; 2)$ or $(2, 5, 8; 1)$.

Furthermore, using the interlacing criterion it is easy to see that the hypergeometric function ${}_3F_2([1/9, 4/9, 7/9]\,;[a, b]\,;t)$ is algebraic if $(a, b)$ or $(b, a)$ occurs in the set

$$\{(3/4, 1/4), (2/3, 1/3), (2/3, 1/6), (1/2, 1/3), (1/2, 1/6)\}.$$

Similarly, ${}_3F_2([2/9, 5/9, 8/9]\,;[a, b]\,;t)$ is algebraic if $(a, b)$ or $(b, a)$ is part of

$$\{(5/6, 1/2), (5/6, 1/3), (3/4, 1/4), (2/3, 1/2), (2/3, 1/3)\}.$$

Moreover, both ${}_3F_2\left(\left[\frac{1}{4}, \frac{3}{8}, \frac{7}{8}\right]\,;\left[\frac{2}{3}, \frac{1}{3}\right]\,;t\right)$ and ${}_3F_2\left(\left[\frac{1}{8}, \frac{3}{4}, \frac{5}{8}\right]\,;\left[\frac{2}{3}, \frac{1}{3}\right]\,;t\right)$ are algebraic.

The previous analysis proves the following corollary.

**Corollary 2.10.** *The hypergeometric function*

$${}_3F_2\left([A, B, C]\,;[1, D]\,;t\right)$$

*has Hadamard grade 2 (hence is a diagonal) for* $(A, B, C; D)$ *in the following set*

$$\Big\{(1/4, 3/8, 7/8; 1/3),\ 1/4, 3/8, 7/8; 2/3),\ (1/8, 5/8, 3/4; 1/3),\ (1/8, 5/8, 3/4; 2/3),$$
$$(1/9, 4/9, 7/9; 1/2),\ (1/9, 4/9, 7/9; 1/3),\ (1/9, 4/9, 7/9; 1/4),\ (1/9, 4/9, 7/9; 1/6),$$
$$(1/9, 4/9, 7/9; 2/3),\ (1/9, 4/9, 7/9; 3/4),\ (2/9, 5/9, 8/9; 1/2),\ (2/9, 5/9, 8/9; 1/3),$$
$$(2/9, 5/9, 8/9; 1/4),\ (2/9, 5/9, 8/9; 2/3),\ (2/9, 5/9, 8/9; 3/4),\ (2/9, 5/9, 8/9; 5/6)\Big\}.$$

Note that the authors of [52, 53] produced in 2011 a list of 116 potential counter-examples to Christol's conjecture; they displayed a sublist of 18 cases in the preprint [52, Appendix F], of which they selected 3 cases that were published in [53, §5.2]. As of today, to our knowledge, the 3 cases in [53] are still unsolved[7], while 2 of the 18 cases in [52] have been solved in [1] (in red, above) and 2 others in the current paper (in orange, above). From the list of 116 cases, only 2 were previously solved, in [1]. Corollary 2.10 solves 14 cases more, raising the number of solved cases to 16 (out of 116). Finally we note that another 24 cases could be resolved with the ansatz ${}_3F_2\left([A, B, C]\,;[1, D]\,;t\right) = {}_2F_1\left([A, B]\,;[r]\,;t\right) \star {}_2F_1\left([C, r]\,;[D]\,;t\right)$ and finding $r \in \mathbb{Q}$ such that both hypergeometric functions on the right-hand side are algebraic. The remaining 76 cases, however, seem to be much more difficult.

[7] Rivoal and Roques proved in [273, Proposition 1] that one of the 3 cases in [53, §5.2], namely ${}_3F_2\left(\left[\frac{1}{7}, \frac{2}{7}, \frac{4}{7}\right]\,;\left[1, \frac{1}{2}\right]\,;t\right)$, has infinite grade assuming the Rohrlich–Lang conjecture [322, Conj. 22]; the status of the analogous statement for Christol's ${}_3F_2\left(\left[\frac{1}{9}, \frac{4}{9}, \frac{5}{9}\right]\,;\left[1, \frac{1}{3}\right]\,;t\right)$ is still unclear.

# Chapter 3

# Dubrovin-Yang-Zagier numbers and algebraicity of D-finite functions

*"So this is a very mysterious example [...]*
*of numbers defined by recursions with polynomial coefficients."*
Don Zagier, *The arithmetic and topology of differential equations*, 2018

This chapter deals with the task of proving that a given D-finite function is algebraic. We explore some of the known methods on the very explicit example of two generating functions of the so-called Dubrovin-Yang-Zagier numbers $(a_n)_{n\geq 0}$, $(b_n)_{n\geq 0}$. Specifically, we are able to prove with an unified algorithmic method that both generating functions are algebraic. Moreover, we discover 7 other similar sequences and conjecture algebraicity for all of the corresponding generating functions, and prove it for some of them. Finally, using numerical methods we can (heuristically) predict algebraicity degrees for of all these functions.

The contents of this chapter are a combination of §3.1 in [335] as well as current work in progress with A. Bostan and J.-A. Weil. Except for the Section 3.1.1 this chapter did not appear anywhere before.

## 3.1 The family of Dubrovin-Yang-Zagier sequences

The original motivation for this chapter was to resolve the "very mysterious example" stated by Zagier in [338, p. 769]. There he defines the sequence of numbers $(u_n)_{n\geq 0}$ (originally called $c_n$ by Zagier) by the recursion

$$\begin{aligned} 80352000n(5n-1)(5n-2)(5n-4)u_n+ \\ 25\left(2592000n^4-16588800n^3+39118320n^2-39189168n+14092603\right)u_{n-1}+ \\ 20\left(4500n^2-18900n+19739\right)u_{n-2}+u_{n-3}=0, \end{aligned} \tag{3.1}$$

with initial conditions $u_0=1, u_1=-161/(2^{10}3^5)$ and $u_2=26605753/(2^{23}3^{12}5^2)$. We remark that via the equation

$$12n=5(s+1)+m+2, \quad s\geq 0 \text{ and } 0\leq m\leq 3$$

the numbers $u_n$ correspond to the case $r = 5$ of the one-point $r$-spin intersection numbers $\langle \tau_{s,m} \rangle$ introduced by Witten in [331] in the generalization of his (later proven by Kontsevich [211]) conjecture [330] on the $\psi$-class intersection numbers. In the recent work [126] by Dubrovin, Yang and Zagier the authors denote the numbers $u_n$ by $\tau_{A_4}(5n)$. We refer to that article for more information on the origin of the sequence $(u_n)_{n\geq 0}$ and many properties of it and related sequences.

The sequence $(u_n)_{n\geq 0}$ starts as follows:

$$(u_n)_{n\geq 0} = \left(1, \frac{161}{2^{10}3^5}, \frac{26605753}{2^{23}3^{12}5^2}, -\frac{538156369}{2^{32}3^{17}5^2}, \frac{551033855470217}{2^{47}3^{23}5^4 19}, \cdots\right).$$

It is noted by Zagier [338] and proved in [126] that the numbers $u_n$ decay roughly like $n!^{-2}$, hence it is a valid question:

> **Question** (Zagier)**:** Do there exist pairs of positive rational numbers $\alpha, \beta$ and a positive integer $\gamma$ such that
>
> $$\widetilde{u}_n = \gamma^n \cdot u_n \cdot (\alpha)_n (\beta)_n$$
>
> becomes an integer sequence of geometric growth.

Recall that, as usual, we denote by $(x)_n$ the rising factorial: $(x)_n \coloneqq x \cdot (x+1) \cdots (x+n-1)$. Equivalently to the formulation of Zagier's question above one might ask for pairs $(\alpha, \beta) \in \mathbb{Q}^2_{>0}$ such that $u_n(\alpha)_n(\beta)_n \in \mathbb{Z}[1/\widetilde{\gamma}]$ for some $\widetilde{\gamma} \in \mathbb{Z} \setminus \{0\}$, i.e. $u_n(\alpha)_n(\beta)_n$ is a *globally bounded* sequence. Indeed, already in [338] Zagier claims that the two sequences

$$\begin{aligned} a_n &\coloneqq (3/5)_n (4/5)_n \cdot u_n \quad \text{and} \\ b_n &\coloneqq (2/5)_n (9/10)_n \cdot u_n \end{aligned} \tag{3.2}$$

belong to $\mathbb{Z}[1/30]$ and in [126] this claim is proven. Note that, up to a renormalization, the numbers $a_n$ above are called $c_{5n}$ and $b_n$ are denoted $a_{5n}$ in the notation of the latter paper. Moreover, both papers [338, 126] claim that the generating function $F_b(t) = \sum_{n\geq 0} b_n t^n$ has the additional property of being algebraic. This was the main reason why Zagier referred to $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ as a "very mysterious example". As we will see later (Theorem 3.1), in fact, also $F_a(t) = \sum_{n\geq 0} a_n t^n$ is an algebraic function – hence both generating functions are (very special) period functions. Note that in contrast to the first submitted version, in the newest submission of [126] the authors now also claim that that $F_a(t)$ is algebraic, however without a proof.

With a simple search we found seven more pairs $(\alpha, \beta)$ that satisfy the integrality in Zagier's question. More precisely, we first looked at all pairs $(\alpha, \beta) = (\alpha'/60, \beta'/60)$ with $0 < \alpha' \leq \beta' \leq 60$. Then we computed the term $u_{50}(\alpha)_{50}(\beta)_{50}$ and examined its denominator. If we found only the primes $2, 3$ and $5$ in the factorization, we computed more terms and checked their denominators. We call the new sequences $(c_n)_{n\geq 0}, \ldots, (i_n)_{n\geq 0}$ and present the following theorem and the accompanying Table 3.1.

**Theorem 3.1.** *Let $(u_n)_{n\geq 0}$ be defined as in (3.1). If $(\alpha', \beta')$ is a pair in Table 3.1 then*

$$\left(\frac{\alpha'}{60}\right)_n \left(\frac{\beta'}{60}\right)_n \cdot u_n \in \mathbb{Z}[1/30].$$

*The generating function of each such sequence is D-finite with order of the minimal ODE as in Table 3.1. The generating functions of the sequences* $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$, $(c_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$, $(h_n)_{n\geq 0}$ *are algebraic.*

| # | $\alpha'$ | $\beta'$ | ord$(L)$ | # | $\alpha'$ | $\beta'$ | ord$(L)$ |
|---|---|---|---|---|---|---|---|
| $a_n$ | 36 | 48 | 2 | $f_n$ | 19 | 49 | 4 |
| $b_n$ | 24 | 54 | 4 | $g_n$ | 19 | 59 | 4 |
| $c_n$ | 12 | 48 | 2 | $h_n$ | 29 | 49 | 4 |
| $d_n$ | 14 | 54 | 4 | $i_n$ | 29 | 59 | 4 |
| $e_n$ | 34 | 54 | 4 | | | | |

Table 3.1 Pairs $(\alpha', \beta')$ such that $u_n(\alpha'/60)_n(\beta'/60)_n$ is globally bounded.

In [126] the authors also found the pair $(\alpha, \beta) = (1/5, 4/5)$ which corresponds to the sequence $c_n$ in Table 3.1 and it also is proven there that the generating function $F_c(t) = \sum_{n\geq 0} c_n t^n$ is algebraic.

We remark that with our search we found some "fake integrality cases": pairs $(\alpha, \beta)$ that do not give globally bounded sequences, but for many (relatively small) $n$ the denominator of $u_n(\alpha)_n(\beta)_n$ has only $2, 3$ and $5$ as prime factors. One such example is the pair $(\alpha', \beta') = (10, 50)$. The sequence of numbers $n$ for which the denominator of $u_n(1/6)_n(5/6)_n$ is divisible by some prime other than $2, 3$ or $5$ starts as follows: $15$, $25$, $31$, $32$, $39$, $40$, $41$, $49$, $\ldots$. Even though the sequences corresponding to these "fake" pairs are not globally bounded and hence cannot be algebraic, they are still quite interesting sequences from the arithmetic point of view.

After running similar but bigger searches to the described one, we conjecture that the Table 3.1 is complete, that all the corresponding sequences have algebraic generating functions and we also provide conjectural algebraicity degrees for them:

**Conjecture 3.2.** *The Table 3.1 is complete, meaning that if* $u_n(\alpha)_n(\beta)_n$ *is globally bounded for some* $(\alpha, \beta) \in \mathbb{Q}^2_{>0}$ *then the pair* $(\alpha, \beta) \bmod \mathbb{Z}^2$ *is in Table 3.1. Moreover, all generating functions of the sequences in Theorem 3.1 are algebraic with algebraicity degrees as follows (we write* $F_r = \sum_{n\geq 0} r_n t^n$ *for* $r \in \{a, b, \ldots, i\}$*):*

$$\begin{aligned}
&\operatorname{algdeg}(F_a) = \operatorname{algdeg}(F_b) = \operatorname{algdeg}(F_c) = 120 = 2^3 \cdot 3 \cdot 5,\\
&\operatorname{algdeg}(F_g) = \operatorname{algdeg}(F_h) = 46080 = 2^{10} \cdot 3^2 \cdot 5,\\
&\operatorname{algdeg}(F_d) = \operatorname{algdeg}(F_e) = \operatorname{algdeg}(F_f) = \operatorname{algdeg}(F_i) = 155520 = 2^7 \cdot 3^5 \cdot 5.
\end{aligned}$$

The proof of the integrality claim in Theorem 3.1 works along the same lines as the one for $(a_n)_{n\geq 0}$ and $(c_n)_{n\geq 0}$ in [126, §5] and crucially uses the formula obtained by Dubrovin, Yang and Zagier:

$$u_n = 6^{-5n} \cdot \sum_{s=0}^{5n/2} \frac{(-9)^s}{10^{2s}} \cdot \frac{\left(\frac{1}{5}\right)_{3n-s}}{s!(5n-2s)!}. \tag{3.3}$$

To be more precise, instead of proving that $u_n(\alpha)_n(\beta)_n \in \mathbb{Z}[1/30]$, we prove that each summand in (3.3) is almost integral after multiplication with $(\alpha)_n(\beta)_n$. Counting the primes in the numerator and denominator *à la* Legendre and Landau (see [222, 223] and also

[99, §III] by Christol where this computation is attributed to Katz [193, §6]) this can be concluded from non-negativity of the functions

$$f_j(x,y) = \lfloor x + \frac{r_{\alpha',j}}{60} \rfloor + \lfloor x + \frac{r_{\beta',j}}{60} \rfloor + \lfloor 3x - y + \frac{r_{12,j}}{60} \rfloor - \lfloor y \rfloor - \lfloor 5x - 2y \rfloor, \tag{3.4}$$

where $1 \leq j < 60$ is relatively prime to 60 and $0 < r_{i,j} < 60$ such that $r_{i,j} \equiv -ij^{-1} \bmod 60$.

For the sake of completeness we shall give a few more details about why the functions $f_j(x,y)$ appear in (3.4). The computation essentially comes down to estimation of $\nu_p((\gamma)_n)$ where $\gamma \in \mathbb{Q}$ is fixed and $\nu_p$ denotes the $p$-valuation for a prime $p > 5$. Clearly,

$$\nu_p\left(\frac{\left(\frac{1}{5}\right)_{3n-s}(\alpha)_n(\beta)_n}{s!(5n-2s)!}\right) = \nu_p((1/5)_{3n-s}) + \nu_p((\alpha)_n) + \nu_p((\beta)_n) - \nu_p(s!) - \nu_p((5n-2s)!),$$

and we wish to prove that this expression is non-negative for any $n, s \in \mathbb{N}$ and any prime $p > 5$. For $S \subseteq \mathbb{Z}$ and $q \in \mathbb{Z}$ let $u(S,q)$ denote the number of elements in $S$ which are divisible by $q$. Via summing over all $k \geq 0$, the desired non-negativity follows if we can show that

$$\begin{aligned} u(\{1,6,\ldots,5(3n-s-1)+1\},p^k) + u(\{\alpha',60+\alpha',\ldots,60(n-1)+\alpha'\},p^k) \\ +u(\{\beta',60+\beta',\ldots,60(n-1)+\beta'\},p^k) - u(\{1,\ldots,s\},p^k) - u(\{1,\ldots,5n-2s\},p^k) \geq 0 \end{aligned}$$

for all $k \geq 0$. Obviously, $u(\{1,\ldots,\ell\},p^k) = \lfloor \ell/p^k \rfloor$. For the other terms, we need to count for some given $\gamma' \in \mathbb{Z}$ with $0 < \gamma' < 60$ how many terms in

$$S_{\gamma'} = \{\gamma', 60+\gamma', 120+\gamma', \ldots, 60(\ell-1)+\gamma'\}$$

are divisible by $p^k =: q$. Let $p^k \equiv j \bmod 60$ – clearly there are $\varphi(60) = 16$ possible values for $j$. Recall that $r_{i,j}$ was chosen such that $r_{\gamma',j} = -\gamma'/q \bmod 60$. This means that this $r_{\gamma',j}$ is the smallest positive integer solution to $60m - \gamma' = qr_{\gamma',j}$. Our goal is to show that

$$u(S_{\gamma'},q) = \left\lfloor \frac{\ell}{q} + \frac{r_{\gamma',j}}{60} \right\rfloor. \tag{3.5}$$

If we want to compute $u(S_{\gamma'},q)$ we need the smallest positive solution to $60n + \gamma' = qs_{\gamma',j}$, i.e. let $n \equiv -\gamma'/60 \bmod q$ with $1 \leq n < q$. Elementary reasoning shows that $m+n = q$ and $r_{\gamma',j} + s_{\gamma',j} = 60$. Moreover, now clearly, from the set $S_{\gamma'}$ exactly the following elements are divisible by $q$:

$$\{60n+\gamma', 60(n+q)+\gamma', \ldots, 60(n+q(a-1))+\gamma'\},$$

where $a$ is such that $60(n+q(a-1))+\gamma' \leq 60(\ell-1)+\gamma' < 60(n+qa)+\gamma'$, hence $a = \lfloor (\ell-n-1)/q \rfloor + 1$. We therefore proved that

$$u(S_{\gamma'},q) = \left\lfloor \frac{\ell-n-1}{q} \right\rfloor + 1.$$

To see that this is equivalent to (3.5), we replace first $r_{\gamma',j}$ by $60 - s_{\gamma',j}$ and then $s_{\gamma',j}/60$ by $n/q + \gamma'/(60q)$. It then remains to show that

$$\left\lfloor \frac{\ell-n}{q} - \frac{1}{q} \right\rfloor = \left\lfloor \frac{\ell-n}{q} - \frac{\gamma'}{60q} \right\rfloor.$$

This elementary identity follows from the fact $1 \leq \gamma' < 60$.

For each pair $(\alpha, \beta) = (\alpha'/60, \beta'/60)$ we thus have to check non-negativity of (at most) 16 functions $f_j(x, y)$. Moreover, each $f_j$ is periodic mod $\mathbb{Z}^2$ and has jumps only along finitely many lines. Therefore we only need to examine non-negativity of all finitely many regions between all those lines in the unit square. Moreover, it is easy to see that the intersection point of any two such lines can have denominator in absolute value of at most 120. Finally, as a lovely consequence of Pick's theorem [265], it follows that if we evaluate $f_j$ at each point of the grid $(x/360) \times (y/360)$ with $0 \leq x, y \leq 359$, we are guaranteed to check at least one point in every relevant region of $f_j$ in the unit square. This can be done algorithmically very easily and quickly and proves the first part of Theorem 3.1.

### 3.1.1 Computing minimal order ODEs

Since the sequence $(u_n)_{n\geq 0}$ is by definition P-recursive, the same holds for $u_n(\alpha)_n(\beta)_n$ for any $\alpha, \beta \in \mathbb{Q}$ by the closure properties of P-recursive sequences. This proves that the generating functions are D-finite. In order to find the corresponding recursion (and annihilating ODE) we have several different (algorithmic) options:

1. Use effective closure properties of P-recursive sequences.
2. First guess and then prove the recursion.
3. Use the recurrence from 1. and find an equivalent one of minimal order.
4. Guess and prove the differential equation for the generating function. Then convert it into a recurrence.

Following [335, §3.1] will now show on the example of the numbers $a_n$ that all four methods can be easily performed using Maple's `gfun`. We will see that they yield different (but correct) recursions/differential equations – a fact which might look surprising at first glance.

We start by defining in Maple

```
> p0 := 80352000*n*(5*n - 1)*(5*n - 2)*(5*n - 4):
> p1 := 25*(2592000*n^4 - 16588800*n^3 + 39118320*n^2
            - 39189168*n + 14092603):
> p2 := 20*(4500*n^2 - 18900*n + 19739):
```

such that the sequence $(u_n)_{n\geq 0}$ is simply given by

```
> rec_u := {p0*u(n) + p1*u(n - 1) + p2*u(n - 2) + u(n - 3),
u(0) = 1, u(1) = -161/(2^10*3^5), u(2) = 26605753/(2^23*3^12*5^2)}:
```

Using the `gfun` function `rectoproc` we can convert the recurrent definition into a Maple procedure:

```
pro_u := rectoproc(rec_u, u(n)):
```

Now the first few terms of the sequence $(a_n)_{n\geq 0}$ can be computed easily:

```
> seq(pro_u(n)*pochhammer(3/5, n)*pochhammer(4/5, n), n=0..3);
```

$$1, -\frac{161}{518400}, \frac{26605753}{80621568000000}, -\frac{48972229579}{125382662553600000000}$$

**Effective closure properties**

First we will find the recursion for $(a_n)_{n\geq 0}$ using the closure properties implemented in `gfun`. We define the recurrences for the rising factorials:

```
> rec_ph3 := {(n + 3/5)*u(n) = u(n + 1), u(0) = 1}:
> rec_ph4 := {(n + 4/5)*u(n) = u(n + 1), u(0) = 1}:
```

Now we compute the recurrences for $u_n \cdot (3/5)_n$, then for $a_n = u_n \cdot (3/5)_n \cdot (4/5)_n$:

```
> `rec*rec`(rec_u, rec_ph3, u(n)):
> rec_a := `rec*rec`(%, rec_ph4, u(n)):
```

This gives a proof that the sequence $(a_n)_{n\geq 0}$ satisfies the recursion

$$\begin{aligned}
&p_3(n)a_{n+3} + p_2(n)a_{n+2} + p_1(n)a_{n+1} + p_0(n)a_n = 0, \quad \text{where} \\
&p_3(n) = 2^8 3^4 5^9 31(n+3)(5n+11), \\
&p_2(n) = 5^6(2592000n^4 + 14515200n^3 + 29787120n^2 + 27559152n + 10644379), \\
&p_1(n) = 500(5n+8)(5n+9)(4500n^2 + 8100n + 3539), \\
&p_0(n) = (5n+8)(5n+3)(5n+9)(5n+4).
\end{aligned} \tag{3.6}$$

**Guessing the recursion**

A different way to find a recurrence relation for $a_n$ is to guess it first and then prove the guess. We first compute 51 terms of the recursion:

```
> a := [seq(pro_u(n)*pochhammer(3/5,n)*pochhammer(4/5,n), n = 0..50)]:
```

Then we use the function `listtorec` in order to guess a linear relation with polynomial coefficients:

```
> rec_a_guess := listtorec(a,u(n))[1]:
```

We find a smaller recurrence of order 2 (compared to the one proven above of order 3):

$$\begin{aligned}
&\widetilde{p}_2(n)u_{n+2} + \widetilde{p}_1(n)u_{n+1} + \widetilde{p}_0(n)u_n = 0, \quad \text{where} \\
&\widetilde{p}_2(n) = 2^8 3^4 5^6(5n+6)(n+2)(60n+43), \\
&\widetilde{p}_1(n) = 5^4(216000n^3 + 759600n^2 + 836940n + 290603) \\
&\widetilde{p}_0(n) = (5n+4)(5n+3)(60n+103).
\end{aligned} \tag{3.7}$$

This recursion is found in a fraction of a second, however is not yet proven. Because we guessed it using 51 terms, we can only be certain that it gives correct terms $a_n$ for $0 \leq$

$n \leq 50$. One can easily check by computing and comparing terms that this recurrence also holds true for $n \leq 100$ or $n \leq 1000$.

There are several possibilities for proving the guess. Arguably the shortest one is explained below employing Maple's `MinimalRecurrence`. However, for pedagogical reasons, we will first argue on the level of differential operators, since this is exactly the procedure one would follow if trying to prove equality of two D-finite functions. First define $(\widetilde{a}_n)_{n\geq 0}$ as being the unique sequence satisfying equation (3.7) with initial terms $\widetilde{a}_0 = a_0$ and $\widetilde{a}_1 = a_1$. Note that in order to guarantee uniqueness, we use that $\widetilde{p}_2(n)$ is non-zero for $n \geq 0$. Now we rigorously compute the differential equations satisfied by the generating functions of $(a_n)_{n\geq 0}$ and $(\widetilde{a}_n)_{n\geq 0}$ using the `gfun` function `rectodiffeq`:

```
> deq_a := rectodiffeq(rec_a,u(n),y(x)):
> deq_a_guess := rectodiffeq(rec_a_guess,u(n),y(x)):
```

We find different differential equations of order 4 and 3 respectively. Now we translate both equations to differential operators using the Maple package `DEtools`.

```
> L_a := de2diffop(deq_a[1],y(x),[Dx,x]):
> L_a_guess := de2diffop(deq_a_guess,y(x),[Dx,x]):
```

We compute the LCLM $L$ of the two operators, rewrite it as a differential equation and transform it back to a recurrence for the coefficients of the solutions:

```
> L := LCLM(L_a, L_a_guess, [Dx,x]):
> deq := diffop2de(L, y(x), [Dx,x]):
> diffeqtorec(deq, y(x), u(n)):
```

We find exactly the same recurrence relation as in equation (3.6). Notice that the leading coefficient $p_3(n)$ does not vanish for positive $n$. Therefore, if the initial terms are prescribed to be $(a_0, a_1, a_2)$, the differential equation corresponding to the operator $L$ has the unique solution $\sum_{n\geq 0} a_n x^n$. But since $L$ is defined as the LCLM of the operators corresponding to the sequences $(a_n)_{n\geq 0}$ and $(\widetilde{a}_n)_{n\geq 0}$, it also annihilates $\sum_{n\geq 0} \widetilde{a}_n x^n$. This proves that $a_n = \widetilde{a}_n$ for all $n \in \mathbb{N}$ and consequently that our guessed recursion $(3.7)$ is correct.

### Minimal-order recursion

With the Maple's version of 2021 (or later) we have a great shortcut thanks to van Hoeij's improvement in the package `LREtools`. We can namely directly algorithmically find the minimal-order linear recurrence after obtaining one in §3.1.1 by just calling

```
> LREtools[‘MinimalRecurrence‘](rec_a,u(n)):
```

We find exactly $(3.7)$. This not only yields another proof of the correctness of the guessed (and then proven) recurrence, but also proves its minimality. Note that this method does not rely on guessing.

**Guessing the ODE**

We can also guess the differential equation for $\sum_{n\geq 0} a_n x^n$, prove its correctness and transform it to a recursion. This method has the advantage that we might discover a differential equation of smaller order than we would obtain by converting the recurrences above. We simply call the `gfun` function

```
> deq_a_ODEguess := listtodiffeq(a, y(x))[1]:
```

where `a` is the list of the first 51 terms of the sequence $(a_n)_{n\geq 0}$ we computed earlier. We find a small differential equation of order 2 (compared to the differential equations above `deq_a` and `deq_a_guess` of orders 4 and 3).

$$\begin{aligned} &q_2(x)y''(x) + q_1(x)y'(x) + q_0(x)y(x) = 0, \quad \text{where} \\ &q_2(x) = 5x(302400x - 31)(373248000x^2 + 216000x + 1), \\ &q_1(x) = 1354442342400000x^3 + 64571904000x^2 - 61473600x - 31, \\ &q_0(x) = 300(902961561600x^2 - 240974784x - 4991). \end{aligned} \tag{3.8}$$

The proof of the correctness of this guess is similar to the proof in §3.1.1. In this case we actually found a (right) factor of `L_a` as we can see by computing the GCRD (here `L_a_guess2` is the differential operator corresponding to the differential equation (3.8)):

```
> GCRD(L_a, L_a_guess2, [Dx,x]):
```

This gives exactly `L_a_guess2`. Since our solutions to (3.6) and (3.8) agree up to precision 3, they must be equal. Therefore the guess must be correct. Moreover, we can also show that this ODE is minimal by employing `gfun`'s command `minimizediffeq` (which is based on the work [76] by Bostan, Rivoal and Salvy):

```
> minimizediffeq(deq_a_ODEguess, y(x)):
```

finds the same ODE as the input and therefore provides a proof of its minimality.

Transforming this differential equation into a recursion for $(a_n)_{n\geq 0}$ yields yet another recurrence, this time of order 3. This means that we found two different recurrences describing $(a_n)_{n\geq 0}$ and three different differential equations describing the generating function. The orders of these objects are displayed in Table 3.2.

| | Order of recurrence | Order of ODE |
|---|---|---|
| Closure properties | 3 | 4 |
| Guessing the recurrence<br>Computing the minimal rec. | 2 | 3 |
| Guessing the ODE | 3 | 2 |

Table 3.2 Orders of recurrences and ODEs for the sequence $(a_n)_{n\geq 0}$ and its generating function.

The same procedure as above can be used to find the minimal order recurrences and ODEs for all sequences in Theorem 3.1 and their generating functions. We will call the corresponding differential operators $L_a, L_b, \ldots, L_i$. In this case the minimality immediately

follows from irreducibility which can be algorithmically checked using Maple's `DFactor`. Alternatively, as explained above, minimality for recurrences is guaranteed by van Hoeij's `MinimalRecurrence` and for ODEs one can use the minimization algorithm by Bostan, Rivoal and Salvy [76].

In order to prove Theorem 3.1 it remains to justify the algebraicity of $F_a, F_b, F_c, F_g$ and $F_h$. In the next Section 3.2 we will explain why we conjecture the algebraicity of all generating functions in Table 3.1 and we also justify the numbers presented in Conjecture 3.2. In Section 3.3 we sketch a proof that the generating functions of $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$, $(c_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ and $(h_n)_{n\geq 0}$ are algebraic. We demonstrate all ideas very explicitly on the example of the sequence $(a_n)_{n\geq 0}$ since it is the simplest in our list and, from the view of this work, is often equally instructive.

## 3.2 Heuristics based on conjectures and numerical calculations

In this section we present three simple algorithmic methods for conjecturing algebraicity of a given D-finite function $f(x)$. Note that if $f(x)$ is given as the solution of an ODE with enough initial conditions to prescribe $f(x)$ uniquely one can compute ([76, §2.2]) a minimal order differential operator $L_f^{\min}$ annihilating $f(x)$. Then a well-known fact ensures that if $f(x)$ is algebraic then all solutions of $L_f^{\min}$ are algebraic (see, for example, [295, Prop 2.5]; the proof is based on the idea that if $L_f^{\min}$ has not a full basis of algebraic solutions, we can construct a smaller operator annihilating all its algebraic solutions, contradicting minimality). Therefore, up to the (highly non-trivial but efficiently solved) computation of $L_f^{\min}$, deciding whether a D-finite function $f(x)$ is algebraic is equivalent to deciding whether all solutions of a linear ODE are algebraic.

Even though the three approaches discussed in this section do not provide proofs, they are quite strong indicators in practice. The first two procedures are based on the famous *Grothendieck-Katz $p$-curvature conjecture* which relates algebraicity of solutions of linear ODEs to their arithmetic properties. The third strategy is very different in nature: it exploits that, on the one hand, the solutions of $Ly = 0$ are all algebraic if and only if the monodromy group associated to this equation is finite, and, on the other hand, that we can compute numerically the generators of this group efficiently.

### 3.2.1 Methods based on conjectures

The central conjecture in this section is the Grothendieck-Katz conjecture which we shall now briefly recall. Let $L \in \mathbb{Q}(x)\langle\partial\rangle$ be a differential operator. For almost all primes $p$ it makes sense to reduce all coefficients of $L$ mod $p$. We call the resulting operator $L_p$ (or sometimes $L^{(p)}$). As a consequence of the Wronskian lemma it holds that $Ly = 0$ has at most $n$ solutions which are linearly independent over $\mathbb{Q}$. In characteristic $p > 0$ the situation is slightly more delicate, since linear independence of solutions has to be defined over the field of constants $C_p$, i.e. elements in $\mathbb{F}_p[x]$ that are annihilated by $\partial$. It is easy to

see that $C_p = \mathbb{F}_p(x^p)$ and one can prove that $L_p y = 0$ has at most $n$ solutions which are linearly independent over $C_p$.

The following conjecture, attributed to Grothendieck, connects the existence of solutions to $L_p y = 0$ to the solutions of $Ly = 0$.

**Conjecture 3.3** (Grothendieck-Katz). *Let $L \in \mathbb{Q}(x)\langle\partial\rangle$ be a differential operator. The differential equation $Ly = 0$ has a basis of algebraic solutions if and only if $L_p y = 0$ has a full basis of rational solutions for almost all prime numbers $p$.*

Note that the "only if" part of this conjecture is a consequence of Eisenstein's theorem (stated for a special case in [128] by Eisenstein and first proved in [175, 176] by Heine) ensuring that algebraic power series are globally bounded together with the fact that in characteristic $p$ having a basis of rational solutions is equivalent to having a basis of power series solutions, see [181]. Note that Eisenstein's theorem immediately follows from the Newton algorithm if the algebraic function is *étale algebraic*, i.e. if its minimal polynomial $P(x, y)$ satisfies $\partial_y P(0, 0) \neq 0$. The so-called *Simple Root Lemma* [173] allows to reduce the general case to the étale algebraic one.

Except the relatively easy case of first order ODEs, Conjecture 3.3 has been verified by Beukers and Heckman for hypergeometric equations [38] and more generally by Katz [193] for Picard-Fuchs equations.

Importantly for our purposes, it turns out that it is easy to decide in practice whether a given operator $L_p \in \mathbb{F}_p(x)\langle\partial\rangle$ has a full basis of rational solutions. We define the $p$-curvature operator of a linear differential operator $L_p$ as $B \in \mathbb{F}_p(x)\langle\partial\rangle$, where $B$ with $\operatorname{ord}(B) < p$ and monic, is uniquely defined as the remainder in the Euclidean right division

$$\partial^p = AL_p + B.$$

**Lemma 3.4** (Cartier's lemma). *An operator $L_p \in \mathbb{F}_p(x)\langle\partial\rangle$ has a full basis of rational solutions if and only if the $p$-curvature of $L_p$ vanishes.*

This yields a simple testing procedure: compute the $p$-curvatures of mod $p$ reductions a given operator $L \in \mathbb{Q}(x)\langle\partial\rangle$ for many primes $p$, and if "almost all" of them vanish, conclude that the operator likely has a full basis of algebraic solutions.

**Proposition 3.5.** *Let $L_a, \ldots, L_i$ be the monic minimal differential operators annihilating the generating functions of the sequences $(a_n)_{n\geq 0}, \ldots, (i_n)_{n\geq 0}$. The $p$-curvature operators of them vanish for all primes $p$ with $5 < p \leq 101$ except: $\{7\}$ for $L_c$; $\{7, 31\}$ for $L_d$; $\{7, 11\}$ for $L_e$; $\{7, 13, 37\}$ for $L_f$; $\{7, 11, 23, 47\}$ for $L_g$; $\{13, 17, 37\}$ for $L_h$; $\{11, 17, 23, 47\}$ for $L_i$.*

**Example 3.6.** The generating function of the sequence $(1800^n a_n)_{n\geq 0}$ is annihilated by the following minimal differential operator.

$$\begin{aligned}\widetilde{L}_a = {} & 1800x\,(x-2)\left(x^2+50x+20\right)\partial^2 + \\ & 400\left(9x^3+153x^2-846x-108\right)\partial + 288x^2-2971x-8050.\end{aligned}$$

Its reduction mod $p = 7$ reads

$$\widetilde{L}_a^{(7)} = x(x+5)(x^2+x+6)\partial^2 + (2x^3+6x^2+x+4)\partial + x^2+4x.$$

Computing the right division of $\partial^7$ by $\widetilde{L}_a^{(7)}$ in $\mathbb{F}_7(x)\langle\partial\rangle$ using Maple's `rightdivision` from the package `DEtools` we find

$$\partial^7 = A\widetilde{L}_a^{(7)} + 0,$$

for some operator $A \in \mathbb{F}_7(x)\langle\partial\rangle$ of order 5. This means that the 7-curvature of $\widetilde{L}_a$ vanishes.

For the proof of Proposition 3.5 we loop the procedure of the example above over all relevant primes and the 9 operators. The whole computation takes a few minutes on a regular PC.

Proposition 3.5 is the first big indicator that all solutions of all our operators are algebraic functions. We will now explain how to exploit in practice the following weaker conjecture first stated in the literature by Bézivin in [41, p. 299].

**Conjecture 3.7** (Bézivin)**.** *If $Ly = 0$ has a basis of globally bounded power series solutions then all solutions of $Ly = 0$ are algebraic.*

To see that a proof of Conjecture 3.3 implies Conjecture 3.7 one needs the following result by André [18, Prop 5.3.3]: If all solutions of $Ly = 0$ are globally bounded then almost all $p$-curvatures of $L$ vanish. Note that the word "basis" (and "all") is crucial, since it is easy to construct equations with one globally bounded but transcendental solution (for example the generating function of the Apéry numbers).

Using Conjecture 3.7 in practice means computing a basis of solutions at $0$ to $Ly = 0$ up to some (large) order and examining the denominators of all coefficients. If all solutions are algebraic we expect only small primes to appear in the factorizations of these denominators, and Bézivin's conjecture claims that also the converse should hold.

**Example 3.8.** Continuing Example 3.6 but from the viewpoint of Bézivin's conjecture we compute a basis of solutions at $0$ to $\widetilde{L}_a y = 0$ up to order $N = 100$. The command

```
formal_sol(L, x = 0,'order'=100)
```

takes less than a second on a regular PC and finds that such a basis is given by

$$f_1(x) = 1 - \frac{161}{864}x + \frac{26605753}{99532800}x^2 - \frac{41438040413}{85996339200}x^3 + \cdots \quad \text{and}$$
$$f_2(x) = x^{2/5}\left(1 - \frac{311}{288}x + \frac{87572287}{49766400}x^2 - \frac{142563748591}{42998169600}x^3 \cdots\right).$$

Note that in order to use Bézivin's conjecture in this example, we should first perform a substitution $x \mapsto x^5$. However, since such a substitution preserves algebraicity of the solutions, we also see that in Conjecture 3.7 one can equivalently ask for the basis of solutions to be in $x^\rho\mathbb{Z}[1/N][\![x]\!]$ for some $\rho \in \mathbb{Q}$. We remark that it was proven by Fuchs [142, 141] that in general any solution at $0$ of a linear ODE lies in $x^\rho\mathbb{Q}[\log(x)][\![x]\!]$, see also [172, 147] for a modern viewpoint and proof. As we will explain in more detail later, if in the basis of solutions at $0$ of an irreducible differential appears a $\log(x)$ then all solutions must be transcendental. It is a conjecture by André and Christol that the minimal differential operator of a globally bounded but transcendental D-finite function must have a logarithmic singularity, see Conjecture 10.13 and [18, 5.3.2].

The coefficient of $x^{100}$ in $f_1(x)$ has denominator $2^{697} \cdot 3^{298} \cdot 5^{112}$ and the denominator of the coefficient of $x^{100}$ in $f_2(x)x^{-2/5}$ is $2^{694} \cdot 3^{297} \cdot 5^{112}$. A quick check shows that, in fact, up to the order 100 all coefficients of $f_1(x)$ and $f_2(x)x^{-2/5}$ lie in $\mathbb{Z}[1/30]$. In view of Conjecture 3.7 this alone is a strong indication that $f_1(x)$ and $f_2(x)$ are algebraic functions.

We perform a similar test for all our operators $L_a, \ldots, L_i$ and always find bases of Puiseux series solutions in $x^{\ell/5}\mathbb{Z}[1/30][\![x]\!]$ for $\ell = 0, \ldots, 4$. This leads to the conjecture that all the generating functions $F_a(t), \ldots, F_i(t)$ are algebraic.

### 3.2.2 Numerical monodromy group computation

In this section our goal is to briefly explain how we obtained the numbers for the algebraicity degrees presented in Conjecture 3.2. As before, we will present the approach on the example of the differential operator $L_a$.

As already indicated, the first and main step is to compute the monodromy group of $Ly = 0$ numerically. For this purpose we will make use of efficient numeric analytic continuation of D-finite functions. The main ideas of the underlying algorithm were invented by the Chudnovsky brothers in [105, 107] and were then later improved and efficiently implemented by Mezzarobba [242]. We will use the currently fastest implementation provided by the SageMath package `ore_algebra`.

Recall from Table 3.1 that the operators $L_a$ and $L_c$ have order 2 and the other operators have order 4. This means that the monodromy groups of $L_a$ and $L_c$ can be represented by subgroups of $\mathrm{GL}_2(\mathbb{C})$ and the other operators have monodromy groups in $\mathrm{GL}_4(\mathbb{C})$. Like in Example 3.6, after rescaling, the operators $\widetilde{L}_a, \ldots, \widetilde{L}_i$ have 4 singularities: at $0, 2$ and the roots $\phi_1, \phi_2 \in \overline{\mathbb{Q}}$ of $x^2 + 50x + 20$. Therefore the groups are generated by 4 matrices $\widetilde{M}_0, \widetilde{M}_2, \widetilde{M}_{\phi_1}, \widetilde{M}_{\phi_2}$ with the relation that their product is the identity. For simplicity we will work with presentations of the groups given by by the three matrices $M_0, M_1, M_2$ corresponding respectively to the three paths of analytic continuation: (solely) around 0, (solely) around $\phi_1$, and around both $\phi_1$, $\phi_2$.

In SageMath, after defining the corresponding paths, we run the command

```
> L.numerical_transition_matrix(path, eps=2^(-nbits))
```

for each path. We set `nbits` to 800, this means that the three matrices $M_0, M_1, M_2$ will have 800 bits of guaranteed numerical precision.

**Example 3.9.** For the operator $\widetilde{L}_a$ in Example 3.6 we find the monodromy matrices (here capped to just 20 bits of precision):

$$M_0 \approx \begin{pmatrix} 1 & 0 \\ 0 & -0.809017 + 0.587785i \end{pmatrix},$$
$$M_1 \approx \begin{pmatrix} 0.809017 - 0.262866i & 1.00882 - 1.38852i \\ 0.0946567 - 0.130284i & 0.850651i \end{pmatrix},$$
$$M_2 \approx \begin{pmatrix} 0.3090 - 0.9511i & 0 \\ 0 & 0.3090 + 0.9511i \end{pmatrix}.$$

It is easy to check numerically that all eigenvectors of $M_0, M_1, M_2$ have norm 1. With enough digits we can actually (heuristically) identify the numbers appearing in the matrices:

$$(1+\sqrt{5})/4 \approx 0.809017,$$
$$\sqrt{10-2\sqrt{5}}/4 \approx 0.587785,$$
$$\sqrt{50-10\sqrt{5}}/20 \approx 0.262866,$$
$$6\sqrt[5]{600\sqrt{5}-1320}/11 \approx 1.00882, \quad \text{etc.}$$

We are now going to use the following "folklore" fact:

**Proposition 3.10.** *Let $f(x)$ be a solution of $Ly = 0$ for a Fuchsian operator $L \in \mathbb{Q}(x)\langle\partial\rangle$. The algebraicity degree of $f(x)$ is equal to the cardinality of the orbit of $f(x)$ under the action of the monodromy group.*

The generating functions $F_a(t), \ldots, F_i(t)$ are respectively the unique power series solutions of the corresponding operators. In SageMath in the list of solutions at 0 they are therefore represented as the first entry. Thus, we are interested in the orbit of the action of the group generated by $M_0, M_1, M_2$ on the vector $(1,0)^t$ or $(1,0,0,0)^t$ respectively. We compute the orbit numerically by recursively multiplying $M_0, M_1$ and $M_2$ with the vectors in this orbit until no new vector is found. In order not to suffer loss from imprecision, we work with $M_0, M_1, M_2$ computed with 800 bits of precision but identify two vectors as equal if they agree on 10 (decimal) digits (roughly 33 bits of precision). In the end, for each operator, we return the cardinality of the obtained orbit: these are the numbers presented in Conjecture 3.2. The whole procedure takes a few hours computation time on a regular PC.

## 3.3 Proving algebraicity

In this section we will present three practical methods for proving algebraicity of solutions of a differential equation. We recall that the problem of deciding algebraicity of D-finite functions is proven to be decidable [295] by Singer, however this result if only of theoretical importance since its algorithmic complexity is completely impracticable.

Our first strategy exploits the fact that for second-order ODEs the problem of determining existence of algebraic (even Liouvillian) solutions is solved by Kovacic [216] in a very practicable way. This immediately solves our problem for $(a_n)_{n\geq 0}$ and $(c_n)_{n\geq 0}$, and we can moreover express the solutions in terms of pullbacks of Gaussian hypergeometric functions. The second method follows a guess-and-prove paradigm in combination with a famous theorem by the Chudnovsky brothers [103]. Finally, our third strategy exploits invariants (and semi-invariants) in differential Galois theory.

### 3.3.1 Closed form expressions for $F_a(t)$ and $F_c(t)$

We convert the differential operators $L_a$ and $L_c$ to differential equations using Maple's `diffop2de` and subsequently apply `dsolve`. This shows that $F_a(t)$ and $F_c(t)$ are linear combinations of pullbacks of Gaussian hypergeometric functions.

**Example 3.11.** The above procedure implies that $F_a(x) = c_1 A_1^r(x) + c_2 A_2^r(x)$ and $F_c(x) = d_1 A_1^r(x) + d_2 A_2^r(x)$ for some $c_1, c_2, d_1, d_2 \in \overline{\mathbb{Q}}$ and

$$A_1^r(x) \coloneqq u_1^r(x) \cdot {}_2F_1\left[\begin{matrix}-1/60 & 11/60 \\ & 2/3\end{matrix}; \frac{p_1^r(x)}{p_2^r(x)}\right] \quad \text{and}$$
$$A_2^r(x) \coloneqq u_2^r(x) \cdot {}_2F_1\left[\begin{matrix}19/60 & 31/60 \\ & 4/3\end{matrix}; \frac{p_1^r(x)}{p_2^r(x)}\right],$$

where $r \in \{a, c\}$ and $u_1^r(x), u_2^r(x)$ are explicit algebraic functions, $p_1^r(x), p_2^r(x)$ are known polynomials.

In order to prove algebraicity of $F_a(t)$ and $F_c(t)$ we can proceed in two different ways. First, a classical work [288] by Schwarz from 1873, classifies all Gaussian hypergeometric functions that are algebraic. Applying this classification, known as *Schwarz's list*, we can convince ourselves that both ${}_2F_1$'s above are algebraic.

**Lemma 3.12.** *The functions*

$$f_1(x) \coloneqq {}_2F_1\left[\begin{matrix}-1/60 & 11/60 \\ & 2/3\end{matrix}; x\right] \text{ and } f_2(x) \coloneqq {}_2F_1\left[\begin{matrix}19/60 & 31/60 \\ & 4/3\end{matrix}; x\right]$$

*are algebraic.*

A different way to see algebraicity of $f_1(x)$ and $f_2(x)$ is to use the so-called Landau-Errera criterion [223, 224, 133] for Gaussian hypergeometric functions which was later generalized by Beukers and Heckman [38] for ${}_nF_{n-1}$'s and became the "interlacing criterion".

Yet another method, completely different in spirit, but useful also for more general problems of deciding algebraicity, is to use a "guess and prove" approach. We can first try to guess and then to prove minimal polynomials for $f_1(x)$ and $f_2(x)$. This will not only provide a proof for algebraicity of the functions, but also give explicit minimal polynomials. In order to make the computations easier we will actually work with twelfth powers of $f_1(x)$ and $f_2(x)$. Moreover, here we only explain the computations for $f_1(x)$, because the exact same strategy works for $f_2(x)$ as well.

First we compute 100 terms of the series expansion for $f_1(x)^{12}$ in Maple:

```
> f12 := hypergeom([-1/60, 11/60], [2/3], x)^12:
> ser1 := series(f12,x,100):
```

Then we guess an annihilating polynomial for this series using `gfun`:

```
> P := seriestoalgeq(ser1,y(x)):
```

After a few seconds, this finds a polynomial $P(x, y) \in \mathbb{Z}[x, y]$ of degree $20$ in $y$ and degree $4$ in $x$. We note that $\partial_y P(0, 1) \neq 0$, therefore there exists only one power series solution $f(x)$ to $P(x, y) = 0$ such that $f(0) = 1$. Now we will confirm our guess. First we use the effective property that any algebraic function is D-finite:

```
> deq := algeqtodiffeq(P, y(x)):
```

Here we find an inhomogeneous differential equation, which we convert into a homogeneous one using `gfun`'s `diffeqtohomdiffeq`. Let us call the resulting equation `deqh`. It holds that any solution in $y(x)$ to $P(x, y) = 0$ satisfies the differential equation `deqh`. Moreover, we can find the differential equation satisfied by $f_1(x)^{12}$ by simply calling

```
> deqf12 := holexprtodiffeq(f12,y(x)):
```

We find exactly the same differential equation as `deqh`. By uniqueness and after checking enough terms, we can conclude that $P(x, y)$ indeed annihilates $f_1(x)^{12}$. Hence, $f_1(x)$ is algebraic. Moreover, the irreducible polynomial $P(x, y^{12})$ is then clearly the minimal polynomial for $f_1(x)$. This concludes the proof of Lemma 3.12 for the first function, while the second one can be done completely analogously.

Coming back to the generating function $F_a(t) = \sum_{n\geq 0} a_n t^n$, Lemma 3.12 implies that both $A_1^r(t)$ and $A_2^r(t)$ are algebraic. Then any linear combination of them must be an algebraic function as well. This proves Theorem 3.1 for $(a_n)_{n\geq 0}$ and $(c_n)_{n\geq 0}$.

### 3.3.2 Guess-and-prove with Chudnovsky's theorem

As we saw, an often very efficient method for proving algebraicity of solutions of linear differential equations is to guess and then prove the minimal polynomial. The main idea being that computing an ODE satisfied by the solutions of a polynomial equation, i.e. the proving step, can be done very quickly in practice [114, 104].

The practical obstacle to the guess-and-prove approach for proving algebraicity of D-finite functions is that the algebraicity of the solutions of a $n$-th order ODE can be arbitrarily large, even if the degree of the coefficients is bounded by some $k$. A simple example of this phenomenon is the function

$$f(x) = \sqrt[N]{1+x},$$

which obviously satisfies the first order linear differential equation

$$N(x+1)y'(x) - y(x) = 0.$$

Clearly, the algebraicity degree of $f(x)$ is $N$, while the order and degree of the minimal ODE stay bounded by 1. We recall that already in Singer's algorithm [295] (and also in the works of Liouville, Jordan, Painlevé, Boulanger, and others) this first obstacle is overcame by considering the logarithmic derivative of a solution. An application of Jordan's theorem in group theory implies that the algebraicity degree of the logarithmic derivative of some solution can be bounded in terms of the order of $L$:

**Proposition 3.13.** *If* $\operatorname{ord}(L) = n$ *and all solutions of* $Ly = 0$ *are algebraic, then for some solution* $y_0$ *we have* $\operatorname{algdeg}(y_0'/y_0) < (49n)^{n^2}$.

The bound $(49n)^{n^2}$ seems horrible at first glance, however, the strength (and surprising consequence) of this proposition that the algebraicity degree of $y_0'/y_0$ can be bounded solely in terms of the order of the differential equation. We mention that for the small orders $n = 2, 3, 4, 5$ the finite subgroups of $\mathrm{SL}_n(\mathbb{C})$ are classified and therefore much better (tight) bounds are available.

For the sequences in Table 3.1 the minimal polynomials are too large to be guessed. It is therefore natural to try to guess a minimal polynomial for $F_r'(t)/F_r(t)$, especially given that the following theorem holds true

**Theorem 3.14** (Chudnovsky-Chudnovsky [103]). *Let $f(x) \in \mathbb{Q}[\![x]\!]$ be globally bounded and assume that $f'(x)/f(x)$ is algebraic. Then $f(x)$ is algebraic.*

Indeed, for $F_a(t)$, $F_b(t)$ and $F_c(t)$ we are able to guess minimal polynomials for their logarithmic derivatives. If we can prove these guesses then by Theorem 3.14 we will be able conclude that these functions are algebraic, since we proved already that they are globally bounded.

**Example 3.15.** The minimal differential equation for $(1800^n c_n)_{n\geq 0}$ is

$$\widetilde{L}_c = 1800x(x-2)(x^2+50x+20)\partial^2 + 3600(x^3+17x^2-94x-12)\partial + 288x^2 - 2971x - 8050$$

For the power series solution $y_0(x) = 1 - 161/864x + 26605753/99532800x^2 + \cdots$ at 0 we compute the logarithmic derivative $y_0'/y_0$ up to order 500. Then we try to guess a minimal polynomial $P(x,y)$ using `gfun`'s procedure `seriestoalgeq`. We find a polynomial of degree 12 in $y$ and 30 in $x$ which satisfies $P(x, y_0'/y_0) = O(x^{500})$.

The natural question now is: How to prove that the guessed polynomial $P(x,y)$ is the true minimal polynomial of $u = y_0'/y_0$? For this we recall that the logarithmic derivative of any solution of an $n$-th order linear differential equation $Ly = 0$ satisfies the (non-linear) *Riccati equation* $R(u) = 0$ of order $n-1$. It can be found by differentiating the equation $y' = uy$ (for a variable $u = u(x)$), replacing each derivative of $y$ in $Ly$ by the corresponding polynomial expression in $y, u, u', u'', \ldots$, and finally dividing by $y$. If $n = 2$ for $L = \partial^2 + p(x)\partial + q(x)$ we find

$$R(u) = u' + p(x)u + u^2 + q(x). \tag{3.9}$$

For $n = 3$ and $L = \partial^3 + p(x)\partial^2 + q(x)\partial + r(x)$ we find

$$R(u) = u'' + (p(x) + 3u)u' + p(x)u^2 + u^3 + q(x)u(x) + r(x),$$

and finally for $\partial^4 + p\partial^3 + q\partial^2 + r\partial + s$ the Riccati equation is

$$R(u) = u''' + (p+4u)u'' + 3u'^2 + (3pu + 6u^2 + q)u' + pu^3 + u^4 + qu^2 + ru + s. \tag{3.10}$$

In order to prove the correctness of the guessed polynomial $P(x,y)$ it is enough to prove that its (unique) power series solution satisfies (3.9) or (3.10) respectively. This can be done by computing resultants or, much more efficiently, by using the algorithms from [68].

**Example 3.16.** Continuing Example 3.15 let $Q(x,y)$ be the minimal polynomial of $f'(x)$ if $f(x)$ is a solution of $P(x,y) = 0$. We define

$$H(x,y,z) = z + p(x)y + y^2 + q(x),$$

where, in accordance to (3.9) and the equation for $\widetilde{L}_c$,

$$p(x) = \frac{2(x^3+17x^2-94x-12)}{(x-2)(x^2+50x+20)} \quad \text{and} \quad q(x) = \frac{288x^2-2971x-8050}{1800x(x-2)(x^2+50x+20)}.$$

Thus, (3.9) implies that we wish to prove that $H(x, f, f') = 0$ and in the terminology of [68] we wish to compute $P \diamond_H Q$ over the field $\mathbb{Q}(x)$. The naive method of computing

$$(P \diamond_H Q)(x,t) = \mathrm{Res}_y\left(\mathrm{Res}_z(t - H(x,y,z), P(x,z)), Q(x,y)\right)$$

is enough to conclude that $f'(x) + p(x)f(x) + f(x)^2 + q(x) = 0$ in this case.

After a few hours of computation this strategy succeeds in proving that $F_a(t)$ and $F_c(t)$ are algebraic, however for $F_b(t)$ we had to abort the computation after 30 hours.

### 3.3.3 Invariants and semi-invariants

Consider the equation $Ly = 0$ for $L \in \mathbb{C}(x)\langle\partial\rangle$ an $n$-th order differential operator and let $y_1, \ldots, y_n$ be a basis of solutions. Similarly to the first step in classical Galois theory one may define the so-called Picard-Vessiot extension as $K = \mathbb{C}(x, y_1, \ldots, y_n)$. Then the differential Galois group $G$ is the group of field automorphisms of $K$ which commute with the derivation and leave all elements of $\mathbb{C}(x)$ invariant:

$$G \coloneqq \mathrm{Aut}_\partial(K/\mathbb{C}(x)) = \{\sigma \in \mathrm{Aut}(K) \colon \sigma|_{\mathbb{C}(x)} \equiv \mathrm{id} \text{ and } \sigma \circ \partial \equiv \partial \circ \sigma\}.$$

One of the central facts in differential Galois theory is that $G$ is a linear algebraic subgroup of $\mathrm{GL}_n(\overline{\mathbb{Q}})$ [210]. For Fuchsian equations $G$ is isomorphic to the Zariski closure of the monodromy group [316], and importantly for us:

**Theorem 3.17.** *The equation $Ly = 0$ has a basis of algebraic solutions if and only if the group $G$ is finite. In this case $G$ is isomorphic to the monodromy group of $Ly = 0$.*

Hrushovski proved in 2002 that $G$ is computable [183], however, similarly to Singer's algorithm on the decidability of algebraic solutions, the complexity (analyzed and improved in [135, 311]) of this algorithm makes it impracticable.

On the other hand, let $\mathfrak{g} = \mathrm{Lie}(G)$ be Lie algebra of $G$, i.e. the tangent space of the algebraic group $G$ at id. $G$ is finite if and only if $\mathfrak{g} = 0$ and it turns out that proving the latter in practice is sometimes doable [30]. The idea for proving that $\mathfrak{g}$ vanishes for concrete examples is to collect enough information on the group $G$ and then solve a linear system of equations. We recall the following definition [296, 321]:

**Definition 3.18.** A homogeneous polynomial expression over $\mathbb{C}(x)$ in the solutions of $Ly = 0$ that is fixed by $G$ is called an *invariant* of $G$.

Computing an invariant can be done finding rational solutions of the equation $L^{\circledS m}y = 0$, where $L^{\circledS m}$ denotes the *$m$-th symmetric power operator* of $L$ defined as the least order operator monic differential operator which annihilates $P(y_1, \ldots, y_n)$ for any homogeneous polynomial $P \in \mathbb{C}[x_1, \ldots, x_n]$ of degree $m$. Obviously, $L^{\circledS 1} = L$ and Example 3.19 below fully illustrates the situation when $\mathrm{ord}(L) = m = 2$. It holds that $\mathrm{ord}(L^{\circledS m}) \leq \binom{n+m-1}{n-1} =: N$ for an order $n$ operator $L$, since $N$ is precisely the number of monomials of degree $m$ in $n$ variables. Computing $L^{\circledS m}$ in practice can be challenging, however it is observed in [321] that on the level of systems the corresponding $N \times N$ system is easy to construct. Recall that the $n$-th order scalar differential equation $Ly = 0$ can be equivalently rewritten as a first order vectorial differential equation $Y' = AY$, where $A \in \mathbb{C}(x)^{n\times n}$ is the companion matrix of $L$ and $Y = Y(x)$ is a vector of unknown functions. In this way the system $Y' = S^m(A)Y$ corresponds to $L^{\circledS m}$.

**Example 3.19.** Let $L = \partial^2 + p(x)\partial + q(x)$ be a differential operator. Then

$$L^{\circledS 2} = \partial^3 + 3p(x)\partial^2 + (2p(x)^2 + p'(x) + 4q(x))\partial + 4p(x)q(x) + 2q'(x),$$

and if $y_1, y_2$ are the solutions of $Ly = 0$ then $y_1^2, y_1y_2, y_2^2$ are the solutions of $L^{\circledS 2}y = 0$. The equation $Ly = 0$ is equivalent to $Y' = AY$ where

$$A = \begin{pmatrix} 0 & 1 \\ -q(x) & -p(x) \end{pmatrix},$$

and for $S^2(A)$ we find

$$S^2(A) = \begin{pmatrix} 0 & 1 & 0 \\ -2q(x) & -p(x) & 2 \\ 0 & -q(x) & -2p(x) \end{pmatrix}.$$

If $U = (u_{i,j})_{i,j}$ is a fundamental matrix of solutions to some system $Y' = AY$ with $A \in \mathbb{C}(x)^{2\times 2}$ then a fundamental matrix of solutions to $Y' = S^2(A)Y$ is given by given by

$$\mathrm{Sym}^2(U) = \begin{pmatrix} u_{1,1}^2 & u_{1,2}u_{1,1} & u_{1,2}^2 \\ 2u_{1,1}u_{2,1} & u_{1,1}u_{2,2} + u_{1,2}u_{2,1} & 2u_{1,2}u_{2,2} \\ u_{2,1}^2 & u_{2,2}u_{2,1} & u_{2,2}^2 \end{pmatrix},$$

and the formula for $S^2(A)$ is simply

$$\begin{pmatrix} 2A_{1,1} & A_{1,2} & 0 \\ 2A_{1,2} & A_{1,1} + A_{2,2} & 2A_{1,2} \\ 0 & A_{1,2} & A_{2,2} \end{pmatrix}.$$

Assume now that $F$ is an invariant of $Y' = AY$, i.e. $F = (f_1, \dots f_N)^t \in \mathbb{C}(x)$, with $N = \binom{n+m-1}{m-1}$, is a rational solution of $Y' = S^m(A)Y$, where $A \in \mathbb{C}(x)^{n\times n}$. Let $\mathfrak{g}^{\circledS m}$ be the differential Lie algebra of $Y' = S^m(A)Y$. It follows that $M_mF = 0$ for any $M_m \in \mathfrak{g}^{\circledS m}$. Moreover, since $\mathfrak{g}^{\circledS m} = S^m(\mathfrak{g})$, we also have that $M_m = S^m(M)$ for some $M \in \mathfrak{g}$. Let $m_{i,j} \in \mathbb{C}$ be the $n^2$ entries of the matrix $M \in \mathbb{C}^{n\times n}$, viewed as unknowns. The system of equations $M_mF = 0$ yields $N$ linear equations for the $n^2$ unknowns. Hence, if $\binom{n+m-1}{m-1} > n^2$ the system is overdetermined and we can hope to be able to prove that its only solution is $m_{i,j} = 0$ for all $0 \le i, j \le n$. This will prove that the Lie algebra $\mathfrak{g}$ vanishes and consequently that $Y' = AY$ has only algebraic solutions.

**Example 3.20.** Consider $L = 4(x^2 - 1)\partial^2 + 4x\partial - 1$ and consequently $Y' = AY$, where

$$A(x) = \begin{pmatrix} 0 & 1 \\ \frac{1}{4x^2-4} & \frac{-4x}{4x^2-4} \end{pmatrix}.$$

The symmetric square system is given by $Y' = S^2(A)Y$, where

$$S^2(A) = \frac{1}{4(x^2-1)} \begin{pmatrix} 0 & 4(x^2-1) & 0 \\ 2 & -4x & 8(x^2-1) \\ 0 & 1 & -8x \end{pmatrix}.$$

We can find two rational solutions of $Y' = S^2(A)Y$ given by

$$F_1(x) = \begin{pmatrix} 4x \\ 4 \\ \frac{x}{x^2-1} \end{pmatrix} \quad \text{and} \quad F_2(x) = \begin{pmatrix} -4 \\ 0 \\ \frac{1}{x^2-1} \end{pmatrix}.$$

For a matrix $M_2 \in S^2(\mathfrak{g})$ we must have that $M_2F_1(x) = M_2F_2(x) = 0$. It follows that

$$M_2F_\ell = S^2(M)F_\ell = \begin{pmatrix} 2m_{1,1} & m_{1,2} & 0 \\ 2m_{2,1} & m_{1,1}+m_{2,2} & 2m_{1,2} \\ 0 & m_{2,1} & 2m_{2,2} \end{pmatrix} F_\ell = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad m_{i,j} \in \mathbb{C}, \ell = 1, 2.$$

The corresponding linear system has 4 unknowns $m_{i,j}$ and 6 equations. Solving it we find the unique solution given by $M = 0$. Hence, $\mathfrak{g} = 0$ and all solutions of $Ly = 0$ are algebraic.

In the same way as in Example 3.20 we can prove that $L_a, L_b, L_c$ only admit algebraic solutions. For $L_a$ and $L_c$ we find rational solutions of the $20$-th symmetric powers. This means we have to find rational solutions of a $21 \times 21$ first order differential system because $\binom{2+20-1}{2-1} = 21$, and then solve an overdetermined linear system with 21 equations and 4 variables. All together this takes a few seconds on a regular PC. For $L_b$ we find rational solutions of the $5$-th symmetric power (it has size $\binom{4+5-1}{4-1} = 56$) and then solve the corresponding overconstrained (56 equations, 16 variables) linear problem. This computation takes a few minutes.

A similar reasoning can be done with semi-invariants (hyperexponential solutions of $L^{\circledS m}$ or $S^m(A)$) and this allows to conclude algebraicity of all solutions of $L_g$ and $L_h$. This proves the remaining part of Theorem 3.1. To our knowledge the algebraicity of the functions $F_d(t), F_e(t), F_f(t)$ and $F_i(t)$ is still unproven.

# Chapter 4

# On the reduced volume of conformal transformations of tori

*"A computer is used by a pure mathematician in much the same way that a telescope is used by a theoretical astronomer. It shows us 'what's out there'."*
Herbert S. Wilf, *Mathematics: An Experimental Science*, 2005.

This chapter deals with the uniqueness of the solution to the so-called Canham's problem which predicts the shape of biomembranes. More precisely, we compute the reduced volume (i.e. the isoperimetric ratio cubed), denoted by $\mathsf{Iso}$, of any stereographic projection of the Clifford torus to $\mathbb{R}^3$. Most significantly, the chapter contains a short and elementary proof of the main conjectured result of the recent article [334] by Yu and Chen that the function $\mathsf{Iso}(z)$ is bijective on its domain. The key of the new proof is an explicit expression of the central function as a quotient of Gaussian hypergeometric functions. A similar computation (and conclusion) is done for the family of all Möbius transformations of any torus.

This chapter of the thesis is based on joint work with A. Bostan [81], §3.2 in [335] and a work in progress with A. Bostan and T. Yu. The Sections 4.1 and 4.3 appear in this dissertation for the first time.

## 4.1 Introduction

The initial starting point of work that led to this chapter is the question from biology:

*Why do all humans have same shaped red blood cells?*

According to Canham's famous work [89] the answer lies in the bending elasticity induced by curvature, which plays a crucial role in determining any membrane's geometry. More precisely, the model suggests that the shape $S$, being a closed surface in $\mathbb{R}^3$ with fixed genus $g$, area $A_0$ and volume $V_0$, solves the so-called *Canham problem*. This means, it minimizes the *Willmore energy* given by

$$W(S) := \int_S H^2 \mathrm{d}A,$$

where $H = (\kappa_1 + \kappa_2)/2$ is the mean curvature and $g, A_0, V_0$ define constraints. This mathematical reformulation allows to study the initial question from a viewpoint of differential geometry.

Note that $W(S)$ is invariant under Möbius transformations, i.e. invariant under translation, rotation, scaling and sphere inversion in $\mathbb{R}^3$. In particular, the scaling invariance implies that minimizing given $A_0$, $V_0$ is equivalent to minimizing given the *reduced volume*:

$$v_0 = v(S) := \pi^{1/2} \frac{6V_0}{A_0^{3/2}} \in (0, 1].$$

Note that $v(S)$ is the third power of the so-called *isoperimetric ratio*

$$\iota_0 = \iota(S) := \pi^{1/6} \frac{\sqrt[3]{6V_0}}{\sqrt{A_0}} \in (0, 1],$$

a notion usually preferred in differential geometry.

Recall that $v(S) = \iota(S) = 1$ if and only if $S$ is a three-dimensional sphere. Moreover, it is also a well-known fact (observed by Willmore [328]) that the Willmore energy satisfies $W(S) \geq 4\pi$, with equality if and only if $S$ is a sphere. Thus, if $v_0 = 1$ (equivalently, if $\iota_0 = 1$) then the unique solution of Canham's problem is the sphere.

In the case $g = 0$ and $v_0, \iota_0 \in (0, 1)$ Schygulla [289] proved that a solution to Canham's problem exists, and numerical computations [291, 96] suggest that it is unique and is a surface of revolution.

For $g = 1$ finding the (unconstrained) minimizer of $W(S)$ was a long-standing and famous open problem. Willmore computed in 1965 [328] the quantity $W(T)$ for a torus $T = T(R, r)$ with major radius $R$ and minor radius $r$, i.e. for

$$T(R, r) = \left\{ [(R + r\cos u)\cos v, (R + r\cos u)\sin v, r\sin u]^T : u, v \in [0, 2\pi) \right\} \subseteq \mathbb{R}^3.$$

It is not difficult to calculate the coefficients of the first fundamental form in this case: $E = r^2, F = 0$ and $G = (R + r\cos u)^2$. Moreover, the mean curvature is given by

$$H = \frac{R + 2r\cos u}{2r(R + r\cos u)},$$

and then one can further compute

$$\begin{aligned} W(T) &= \int_T H^2 \mathrm{d}A = \int_0^{2\pi}\int_0^{2\pi} H^2 \sqrt{EG - F^2}\,\mathrm{d}u\mathrm{d}v \\ &= \int_0^{2\pi}\int_0^{2\pi} \frac{(R + 2r\cos u)^2}{4r(R + r\cos u)}\mathrm{d}u\mathrm{d}v = \frac{\pi R^2}{2r}\int_0^{2\pi} \frac{1}{R + r\cos u}\mathrm{d}u = \frac{\pi^2 R^2}{r\sqrt{R^2 - r^2}}. \end{aligned}$$

If we assume without loss of generality that $r = 1$ (equivalently consider $\rho = R/r$) then a simple computation now shows that the minimum of $W(T)$ is attained at $R = \sqrt{2}$, i.e. for the "trivial" embedding of the *Clifford torus* (4.1) in $\mathbb{R}^3$:

$$T(\sqrt{2}, 1) = T_{\sqrt{2}} := \left\{ [(\sqrt{2} + \cos u)\cos v, (\sqrt{2} + \cos u)\sin v, r\sin u]^T : u, v \in [0, 2\pi) \right\} \subseteq \mathbb{R}^3.$$

Note that our computation yields $W(T_{\sqrt{2}}) = 2\pi^2$. After doing an analogous computation in 1965, Willmore conjectured:

**Conjecture 4.1** (Willmore's conjecture)**.** *Across all closed surfaces in* $\mathbb{R}^3$ *of genus* $g \geq 1$ *the Willmore energy is minimal for* $T_{\sqrt{2}}$.

This long-standing conjecture has been recently proved by Marques and Neves [236] as an elegant application of the Almgren-Pitts min-max theory. Note that since the Willmore energy is invariant under Möbius transformations, $W(S) = 2\pi^2$ for $S = f(T_{\sqrt{2}})$ for any Möbius transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$. Equivalently, the shape of $S$ is a stereographic image of the Clifford torus in $\mathbb{R}^3$, where the Clifford torus is defined as

$$\left\{ [\cos u, \sin u, \cos v, \sin v]^T / \sqrt{2} : u, v \in [0, 2\pi) \right\} \subseteq \mathbb{S}^3. \tag{4.1}$$

The main theorem of Marques and Neves states that $W(S) \geq 2\pi^2$ with equality only for stereographic projections of the Clifford tours.

A natural question, still in the $g = 1$ case, is whether the shape of the minimizer of the Willmore energy becomes unique if the isoperimetric ratio $\iota_0$ (or equivalently the reduced volume $v_0$) of $S$ is prescribed. In the case $v_0 \in [0.712, 1)$ this question is answered by the following theorem:

**Theorem 4.2.** *The shape of the stereographic projection of the Clifford torus to* $\mathbb{R}^3$ *is uniquely determined by its reduced volume* $v_0 \in [3/(2^{5/4}\pi^{1/2}), 1)$.

In other words, writing $\tau = 3/(2^{5/4}\pi^{1/2})$, if $v_0 \in [\tau, 1) \approx [0.712, 1)$ there exists a unique (up to rotation and translation) stereographic projection of the Clifford torus to $\mathbb{R}^3$ with isoperimetric ratio $v_0$. By the theorem of Marques and Neves it must be the unique solution of Canham's problem for $g = 1$ and the given $v_0$. If $v_0 \in (0, \tau)$ numerical computations [291, 96] suggest that the solution is also unique and a surface of revolution, but this case remains open.

The proof of Theorem 4.2 is a combination of results by Yu and Chen [334] with work by Melczer and Mezzarobba [241] or Bostan and Yurkevich [81]. Denoting by $\mathrm{inv}_{(x,y,z)}(S)$ the inversion of a surface $S$ at the unit sphere centered at $(x, y, z) \in \mathbb{R}^3$, Yu and Chen showed that the shape of $T_{\sqrt{2}}$ is determined by the one parameter family $\mathrm{inv}_{(x,0,0)}(T_{\sqrt{2}})$ for $0 \leq x < \sqrt{2} - 1$. Then they observed that in order to prove Theorem 4.2 it is enough to show that the function $\mathsf{Iso}(z)$ is bijective on its domain, where

$$\begin{aligned} \mathsf{Iso} : [0, \sqrt{2} - 1) &\to \big[\tau, 1\big), \\ x &\mapsto v(\mathrm{inv}_{(x,0,0)}(T_{\sqrt{2}})), \end{aligned} \tag{4.2}$$

maps a number $x$ on the reduced volume of $\mathrm{inv}_{(x,0,0)}(T_{\sqrt{2}})$ – the torus $T_{\sqrt{2}}$ inverted at the unit sphere centered at the vector $(x, 0, 0)$. Moreover, the same authors could write $\mathsf{Iso}(x)$ as the quotient of two D-finite functions and concluded that $\mathsf{Iso}(x)$ is bijective if and only if it is monotonic which holds true if a certain P-recursive sequence (the Taylor coefficients of the numerator of the logarithmic derivative of $\mathsf{Iso}^2$) is positive. This positivity was proved by Melczer and Mezzarobba in [241] while, independently, Bostan and Yurkevich proved directly that $\mathsf{Iso}(x)$ is monotonic by finding and proving its closed form expression in terms of hypergeometric functions:

$$\mathsf{Iso}^2(x) = \frac{9\sqrt{2}}{8\pi} \cdot \frac{{}_2F_1\left[\begin{smallmatrix} -\frac{3}{2} & -\frac{3}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^2}{{}_2F_1\left[\begin{smallmatrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^3} \cdot \left(\frac{1-x^2}{1+x^2}\right)^3.$$

While the proof of Bostan and Yurkevich is completely elementary and self-contained, it should be noted that a similar formula for the reduced volume was found already in 1992 by Fourcade [139, §4.1]. There the author also stated: "It appears that the lowest reduced volume torus corresponds to the axisymmetric case.", however this claim stayed unproven.

## 4.2 Reduced volume of a projected Clifford torus

As explained in the introduction, in their recent paper [334], Yu and Chen needed to prove, as a crucial result, that a certain real-valued function $\mathsf{Iso}$ (related to isoperimetric ratios of Clifford tori) is monotonically increasing. They reduced the proof of this fact to the positivity of a sequence of rational numbers $(d_n)_{n\geq 0}$, defined explicitly in terms of nested binomial sums. This positivity was subsequently proved by Melczer and Mezzarobba [241], who used a computer-assisted approach relying on analytic combinatorics and rigorous numerics, combined with the fact (proved in [334]) that the sequence $(d_n)_{n\geq 0}$ satisfies an explicit linear recurrence of order seven with polynomial coefficients in $n$.

In this section, we provide a shorter and more conceptual proof of the monotonicity of the function $\mathsf{Iso}$. Our approach is different in spirit from the ones in [334] and [241]. Our main result (Theorem 4.4 below) is that the function $\mathsf{Iso}(z)$ can be expressed in terms of Gaussian hypergeometric functions ${}_2F_1$ defined by

$$ {}_2F_1\left[\begin{matrix} a & b \\ & c \end{matrix}; z\right] = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!}, \tag{4.3} $$

where $(a)_n$ denotes, as usual, the rising factorial $(a)_n = a(a+1)\cdots(a+n-1)$ for $n \in \mathbb{N}$.

In the notation of Yu and Chen, the function

$$ \mathsf{Iso} : [0, \sqrt{2}-1) \to [3/2 \cdot (2\pi^2)^{-1/4}, 1) $$

is given as

$$ \mathsf{Iso}(z) = 6\sqrt{\pi} \cdot \frac{V(z)}{A^{3/2}(z)}, \tag{4.4} $$

where $A(z) = \sum_{n\geq 0} a_n z^{2n}$ and $V(z) = \sum_{n\geq 0} v_n z^{2n}$ are complex holomorphic functions in the disk $\{z : |z| < \sqrt{2}-1\}$, given by the power series expansions

$$ A(z) = \sqrt{2}\pi^2 \cdot \left(4 + 52z^2 + 477z^4 + 3809z^6 + \frac{451625}{16} z^8 + \cdots\right), $$
$$ V(z) = \sqrt{2}\pi^2 \cdot \left(2 + 48z^2 + \frac{1269}{2} z^4 + 6600z^6 + \frac{1928025}{32} z^8 + \cdots\right). $$

The precise definitions of $A$ and $V$ are given in Section 4.3 of [334], notably in equations (4.2)–(4.3) and also in Section 4.3 below. Roughly speaking, they denote the surface area and volume of a projection of the Clifford torus to $\mathbb{R}^3$. Since the sequences $(a_n)_{\geq 0}$ and $(v_n)_{\geq 0}$ are expressed in terms of nested binomial sums, $A(z)$ and $V(z)$ satisfy linear differential equations with polynomial coefficients in $z$, that can be found and proved automatically using *creative telescoping* [111] (see Section 4.3). Alternatively, as we will show

in Section 4.3, the functions $A(z)$ and $V(z)$ are defined as integrals of trigonometric functions which implies that they are period functions and are consequently D-finite. Creative telescoping also finds (and proves) the corresponding ODEs directly from the integral expressions. Yu and Chen, resp. Melczer and Mezzarobba, use this methodology to find a linear recurrence satisfied by the coefficients $(d_n)_{n\geq 0}$ of

$$F(z) := \frac{1}{4\pi^4} \cdot \left( \frac{2V'(\sqrt{z})A(\sqrt{z}) - 3V(\sqrt{z})A'(\sqrt{z})}{\sqrt{z}} \right) = 72 + 1932z + 31248z^3 + \cdots,$$

respectively a linear differential equation satisfied by the function $F(z)$.
Similarly, one can compute linear differential equations satisfied individually by

$$\bar{A}(z) := \frac{1}{\sqrt{2}\pi^2} \cdot A(\sqrt{z}) = 4 + 52\,z + 477\,z^2 + 3809\,z^3 + \frac{451625}{16} z^4 + \cdots$$

and by

$$\bar{V}(z) := \frac{1}{\sqrt{2}\pi^2} \cdot V(\sqrt{z}) = 2 + 48\,z + \frac{1269}{2} z^2 + 6600\,z^3 + \frac{1928025}{32} z^4 + \cdots.$$

Concretely, $\bar{A}(z)$ and $\bar{V}(z)$ satisfy second-order linear differential equations:

$$\begin{aligned} z(z-1)(z^2-6z+1)(z+1)^2\bar{A}''(z) + (z+1)(5z^4-8z^3-32z^2+28z-1)\bar{A}'(z) \\ + (4z^4+11z^3-z^2-43z+13)\bar{A}(z) = 0 \end{aligned}$$

and respectively

$$\begin{aligned} z\,(z-1)\,(z+1)\left(z^2-6z+1\right)^2 \bar{V}''(z) \\ + \left(z^2-6z+1\right)\left(7z^4-22z^3-18z^2+26z-1\right)\bar{V}'(z) \\ +3\left(3z^5-24z^4-2z^3+56z^2-25z+8\right)\bar{V}(z) = 0. \end{aligned}$$

From these equations, we deduce the following closed-form expressions:

**Theorem 4.3.** *The following equalities hold for all $z \in \mathbb{R}$ with $0 \leq z \leq \sqrt{2}-1$:*

$$\bar{A}(z) = \frac{4\,(1-z^2)}{\left(z^2-6z+1\right)^2} \cdot {}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; \frac{4z}{(1-z)^2}\right]$$

*and*

$$\bar{V}(z) = \frac{2\,(1-z)^3}{\left(z^2-6z+1\right)^3} \cdot {}_2F_1\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} \\ & 1 \end{matrix}; \frac{4z}{(1-z)^2}\right].$$

*Proof.* It is enough to check that the right-hand side expressions satisfy the same linear differential equations as $\bar{A}$ and $\bar{V}$, with the same initial conditions. □

As a direct consequence of Theorem 4.3 and of definition (4.4) we get:

**Theorem 4.4.** *The function* Iso *admits the following closed-form expression:*

$$\mathsf{Iso}^2(z) = \frac{9\sqrt{2}}{8\pi} \cdot \frac{{}_2F_1\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} \\ & 1 \end{matrix}; \frac{4z^2}{(1-z^2)^2}\right]^2}{{}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; \frac{4z^2}{(1-z^2)^2}\right]^3} \cdot \left(\frac{1-z^2}{1+z^2}\right)^3.$$

Using the expression in Theorem 4.4, we can now prove the main (previously conjectured) result of [334].

**Theorem 4.5.** $\mathsf{Iso}$ *is a monotonic increasing function and* $\lim_{z\to\sqrt{2}-1}\mathsf{Iso}(z)=1$. *In particular,* $\mathsf{Iso}$ *is a bijection.*

*Proof.* The value of $\mathsf{Iso}^2(z)$ at $z=\sqrt{2}-1$ is equal to

$$\mathsf{Iso}^2(\sqrt{2}-1)=\frac{9\sqrt{2}}{8\pi}\cdot\frac{{}_2F_1\left[\begin{smallmatrix}-\frac{3}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{3}{2}\end{smallmatrix};1\right]^2}{{}_2F_1\left[\begin{smallmatrix}-\frac{1}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{1}{2}\end{smallmatrix};1\right]^3}\cdot\frac{\sqrt{2}}{4}.$$

From Gauss's summation theorem [24, Th. 2.2.2] it follows that ${}_2F_1\left[\begin{smallmatrix}-\frac{3}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{3}{2}\end{smallmatrix};1\right]=32/(3\pi)$ and ${}_2F_1\left[\begin{smallmatrix}-\frac{1}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{1}{2}\end{smallmatrix};1\right]=4/\pi$; therefore,

$$\mathsf{Iso}^2(\sqrt{2}-1)=\frac{9\sqrt{2}}{8\pi}\cdot\frac{(32/(3\pi))^2}{(4/\pi)^3}\cdot\frac{\sqrt{2}}{4}=1.$$

It remains to prove that $\mathsf{Iso}$ is monotonic increasing. It is enough to show that

$$z\;\mapsto\;\frac{{}_2F_1\left[\begin{smallmatrix}-\frac{3}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{3}{2}\end{smallmatrix};\frac{4z}{(1-z)^2}\right]^2}{{}_2F_1\left[\begin{smallmatrix}-\frac{1}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{1}{2}\end{smallmatrix};\frac{4z}{(1-z)^2}\right]^3}\cdot\left(\frac{1-z}{1+z}\right)^3$$

is increasing on $[0,3-2\sqrt{2})$. Equivalently, via the change of variables $x=\frac{4z}{(1-z)^2}$, it is enough to prove that the function

$$x\;\mapsto\;\frac{{}_2F_1\left[\begin{smallmatrix}-\frac{3}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{3}{2}\end{smallmatrix};x\right]^2}{{}_2F_1\left[\begin{smallmatrix}-\frac{1}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{1}{2}\end{smallmatrix};x\right]^3}\cdot(x+1)^{-\frac{3}{2}}$$

is increasing on $[0,1)$. Clearly, $h$ can be written as $h=f^3\cdot g^2$, where

$$f(x)=\frac{\sqrt{x+1}}{{}_2F_1\left[\begin{smallmatrix}-\frac{1}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{1}{2}\end{smallmatrix};x\right]}\quad\text{and}\quad g(x)=\frac{{}_2F_1\left[\begin{smallmatrix}-\frac{3}{2}\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-\frac{3}{2}\end{smallmatrix};x\right]}{(x+1)^{\frac{3}{2}}}.$$

Hence, it is enough to prove that both $f$ and $g$ are increasing on $[0,1)$. We will actually prove a more general fact in Proposition 4.6, which may be of independent interest. Using that $w_{1/2}=1/f$ and $w_{3/2}=g$, we deduce from Proposition 4.6 that both $f$ and $g$ are increasing. This concludes the proof of Theorem 4.5. □

**Proposition 4.6.** *Let* $a\geq 0$ *and let* $w_a:[0,1]\to\mathbb{R}$ *be defined by*

$$w_a(x)=\frac{{}_2F_1\left[\begin{smallmatrix}-a\\1\end{smallmatrix}\,\middle|\,\begin{smallmatrix}-a\end{smallmatrix};x\right]}{(x+1)^a}.$$

*Then* $w_a$ *is: decreasing if* $0<a<1$*; increasing if* $a>1$*; constant if* $a\in\{0,1\}$.

*Proof.* Clearly, if $a \in \{0,1\}$, then $w_a(x)$ is constant, equal to $1$ on $[0,1]$.

Consider now the case $a > 0$ with $a \neq 1$. The derivative of $w_a(x)$ satisfies the hypergeometric identity

$$\frac{w_a'(x)\cdot(x+1)^{a+1}}{a\cdot(a-1)\cdot(1-x)^{2a}} = {}_2F_1\left[\begin{matrix}a+1 & a\\ & 2\end{matrix};x\right], \tag{4.5}$$

which is a direct consequence of Euler's transformation formula [24, Eq. (2.2.7), p. 68] and of Lemma 4.7 with $a$ substituted by $-a$.

Since $a > 0$, the right-hand side of (4.5) has only positive Taylor coefficients, therefore it is positive on $[0,1)$. It follows that $w_a'(x) \geq 0$ on $[0,1]$ if $a-1 > 0$, and $w_a'(x) \leq 0$ on $[0,1]$ if $a-1 < 0$. Equivalently, $w_a$ is increasing on $[0,1]$ if $a > 1$, and decreasing on $[0,1]$ if $a < 1$. □

**Lemma 4.7.** *The following identity holds:*

$$(a+1)(1-x)\cdot {}_2F_1\left[\begin{matrix}a+1 & a+2\\ & 2\end{matrix};x\right] = a(x+1)\cdot {}_2F_1\left[\begin{matrix}a+1 & a+1\\ & 2\end{matrix};x\right] + {}_2F_1\left[\begin{matrix}a & a\\ & 1\end{matrix};x\right].$$

*Proof.* We will use two of the classical Gauss' contiguous relations [24, §2.5]:

$${}_2F_1\left[\begin{matrix}a+1 & b+1\\ & c+1\end{matrix};x\right] = \frac{c}{bx}\cdot\left({}_2F_1\left[\begin{matrix}a+1 & b\\ & c\end{matrix};x\right] - {}_2F_1\left[\begin{matrix}a & b\\ & c\end{matrix};x\right]\right) \tag{4.6}$$

and

$$a\cdot\left({}_2F_1\left[\begin{matrix}a+1 & b\\ & c\end{matrix};x\right] - {}_2F_1\left[\begin{matrix}a & b\\ & c\end{matrix};x\right]\right) = \frac{(c-b)\cdot {}_2F_1\left[\begin{smallmatrix}a & b-1\\ & c\end{smallmatrix};x\right] + (b-c+ax)\cdot {}_2F_1\left[\begin{smallmatrix}a & b\\ & c\end{smallmatrix};x\right]}{1-x}. \tag{4.7}$$

Applying (4.6) twice, once with $(b,c) = (a,1)$ and once with $(b,c) = (a+1,1)$, the proof of the lemma is reduced to that of the identity

$$(x-1)\cdot {}_2F_1\left[\begin{matrix}a+1 & a+1\\ & 1\end{matrix};x\right] + 2\cdot {}_2F_1\left[\begin{matrix}a & a+1\\ & 1\end{matrix};x\right] = {}_2F_1\left[\begin{matrix}a & a\\ & 1\end{matrix};x\right],$$

which follows from (4.7) with $(b,c) = (a+1,1)$. □

This concludes the proof the main conjectured statement of [334] and also of the fact that the Canham problem has a unique solution in genus one if the reduced volume is between $3/(2^{5/4}\pi^{1/2})$ and 1. A natural remaining question is what happens with the function lso for a general torus $T(R,r)$ with minor radius $r$ and major radius $R$: can we find a similar closed form expression and prove that the function is increasing on its domain for any $R > r$? This is the content of the next section.

## 4.3 Conformal transformation of any torus

In this section we prove the analogous statement to the Theorems 4.4 and 4.5 for any torus

$$T(R,r) = \left\{[(R+r\cos u)\cos v, (R+r\cos u)\sin v, r\sin u]^T : u,v \in [0,2\pi)\right\} \subseteq \mathbb{R}^3$$

and not only $T(\sqrt{2},1)$. This means that all results of this section contain the previous one by setting $R=\sqrt{2}$ (whenever defined). Recall that $\mathsf{Iso}(z)$ is defined (4.2) as the reduced volume of $\mathrm{inv}_{(x,0,0)}(T_{\sqrt{2}})$, where $\mathrm{inv}_{(x,y,z)}(S)$ denotes the inversion of $S$ at the unit sphere centered at $(x,y,z)\in\mathbb{R}^3$ and $T_{\sqrt{2}} = T(\sqrt{2},1)$. We therefore naturally define

$$\begin{aligned}\mathsf{Iso}_R : [0,1/(R+1)) &\to \big[\tau_R, 1\big),\\ x &\mapsto v(\mathrm{inv}_{(x,0,0)}(T(R,1))),\end{aligned}$$

where for $A_R = 4\pi^2 R$ and $V_R = 2\pi^2 R$ the area and volume of $T(R,1)$ we set $\tau_R = 6\pi^{1/2}\frac{V_R}{A_R^{3/2}} = 3/(2\sqrt{\pi R})$. Here we assumed without loss of generality that $r=1$ because this can be achieved by rescaling. Note that this automatically implies that $R>1$. Similarly to $\mathsf{Iso}_R(x)$ we may define $A_R(x)$ and $V_R(x)$ as the area and volume of $\mathrm{inv}_{(x,0,0)}(T(R,1))$. An analogous computation to [334, §4.1] yields:

$$\begin{aligned}A_R(z) &= \int_0^{2\pi}\int_0^{2\pi} \frac{(R+\sin v)\mathrm{d}u\mathrm{d}v}{(1+2z(R+\sin v)\cos u + z^2(R^2+1+2R\sin v)^2}\\ &= 4\pi^2 R\big(1+(4R^2+5)z^2+(9R^4+36R^2+45/4)z^4+\cdots\big)\end{aligned} \tag{4.8}$$

and

$$\begin{aligned}V_R(z) &= \int_0^1\int_0^{2\pi}\int_0^{2\pi} \frac{r(R+r\sin v)\mathrm{d}u\mathrm{d}v\mathrm{d}r}{(1+2z(R+r\sin v)\cos u + z^2(R^2+r^2+2Rr\sin v))^3}\\ &= 2\pi^2 R\big(1+(9R^2+6)z^2+(36R^4+78R^2+69/4)z^4+\cdots\big).\end{aligned} \tag{4.9}$$

In the next Section 4.3.1 we derive linear differential equations for $A_R(z)$ and $V_R(z)$ using the algorithmic technique creative telescoping. Using these ODEs we then find closed form expressions for these functions in terms of (generalized) hypergeometric functions in Section 4.3.2. Finally, Section 4.3.3 contains a proof that for any $R>1$ the function $\mathsf{Iso}_R(z)$ is increasing.

### 4.3.1 Creative telescoping for $A_R(z)$ and $V_R(z)$

The functions $V_R(z)$ and $A_R(z)$ are initially given as follows:

$$\begin{aligned}A_R(z) &= \int_0^{2\pi}\int_0^{2\pi} \frac{R+\sin v}{(1+2z(R+\sin v)\cos u + z^2(R^2+1+2R\sin v))^2}\mathrm{d}u\mathrm{d}v\\ &= \oint_{|x|=|y|=1} \frac{2x(1+2Ry-y^2)dxdy}{(2x(Ry+1)(R-y)z^2-(x^2+1)(1+2Ry-y^2)z+2xy)^2}\end{aligned}$$

and

$$V_R(z) = \int_0^1 \int_0^{2\pi} \int_0^{2\pi} \frac{r(R + r\sin v)\mathrm{d}u\mathrm{d}v\mathrm{d}r}{(1 + 2z(R + r\sin v)\cos u + z^2(R^2 + r^2 + 2Rr\sin v))^3}$$
$$= \int_0^1 \oint_{|x|=|y|=1} \frac{4r(r + 2Ry - ry^2)yx^2\mathrm{d}x\mathrm{d}y}{(2x(Ry + r)(R - ry)z^2 - (x^2 + 1)(r + 2Ry - ry^2)z + 2xy)^3}\mathrm{d}r.$$

In this section we show how to find (and prove) the differential equations satisfied by $A_R(z)$ and $V_R(z)$ using creative telescoping.

Let $a(z, R, x, y), v(z, R, r, x, y) \in \mathbb{Q}(x, y, z, R, r)$ respectively be the integrands above and let $\gamma$ be the cycle defined in $\mathbb{C}^2$ by $|x| = |y| = 1$. Note that $a$ and $v$ are well-defined at the domain of integration, because:

$$(1 + 2z(R + r\sin v)\cos u + z^2(R^2 + r^2 + 2Rr\sin v) > (1 + z(R + r\sin v)\cos u)^2$$

and $1 + z(R + r\sin v)\cos u > 0$ because $0 < z < 1/(R + 1)$ and $r < 1 < R$.

We wish to prove that $L_a \cdot \oint_\gamma a\mathrm{d}x\mathrm{d}y = 0$ and $L_v \cdot \oint_\gamma v\mathrm{d}x\mathrm{d}y = 0$, where

$$L_a = P_2(z, R)\partial_z^2 + P_1(z, R)\partial_z + P_0(z, R) \in \mathbb{Q}(z, R)\langle\partial_z\rangle, \quad \text{and}$$
$$L_v = Q_3(z, r, R)\partial_z^3 + Q_2(z, r, R)\partial_z^2 + Q_1(z, r, R)\partial_z + Q_0(z, r, R) \in \mathbb{Q}(z, R, r)\langle\partial_z\rangle,$$

for some explicitly computed but rather big polynomials $P_i, Q_i$.

We use Koutschan's improved and implemented version [213, 214] of Chyzak's algorithm [111, 112] for creative telescoping. In a few seconds it finds $L_a, L_v$ together with (huge) rational functions $C_1, C_2 \in \mathbb{Q}(z, R, x, y)$ and $D_1, D_2 \in \mathbb{Q}(z, R, r, x, y)$ such that

$$L_a a = \partial_x C_1 + \partial_y C_2 \text{ and}$$
$$L_v v = \partial_x D_1 + \partial_y D_2.$$

Note that $L_a$ and $L_v$ commute with $\oint_\gamma \cdot\mathrm{d}x\mathrm{d}y$. Moreover, because $\gamma$ is a closed cycle, it holds that $\oint_\gamma \partial_x Q\mathrm{d}x\mathrm{d}y = \oint_\gamma \partial_y Q\mathrm{d}x\mathrm{d}y = 0$ for any rational function $Q = Q(z, R, r, x, y)$ which has no pole in $\gamma$. Each denominator of $C_1, C_2, D_1, D_2$ is a factor of

$$\mathrm{denom}(a) \cdot \mathrm{denom}(v) \cdot x \cdot y \cdot (1 + 2Ry - y^2) \cdot H(z, R, r),$$

for some polynomial $H(z, R, r) \in \mathbb{Q}[z, R, r]$. Since each factor above is non-zero on $\gamma = \{(x, y) \in \mathbb{C}^2 : |x| = |y| = 1\}$ we conclude that indeed

$$L_a \cdot \oint_\gamma a\mathrm{d}x\mathrm{d}y = L_v \cdot \oint_\gamma v\mathrm{d}x\mathrm{d}y = 0.$$

Note that creative telescoping cannot find the differential equation for $V_R(z)$ directly: it namely finds that $v(z, R, r, x, y)$ is a sum of pure derivatives with respect to $r, x, y$ of rational functions and outputs the telescoper $1$. In order words, it holds that

$$1 \cdot v = \partial_x \widetilde{D}_1 + \partial_y \widetilde{D}_2 + \partial_r \widetilde{D}_3,$$

for some rational functions $\widetilde{D}_1, \widetilde{D}_2, \widetilde{D}_3 \in \mathbb{Q}(r, x, y)$. This does not imply that $V_R(z)$ is constant because the defining integral has non-trivial boundary and, in the terminology of [71], the telescoper is not *regular* (i.e. it has new poles compared to $a, v$). We note that this phenomenon was observed already by Picard in [264]. In our particular case, however, we can find a closed form expression for $\oint_\gamma v\mathrm{d}x\mathrm{d}y$ in terms of hypergeometric functions and then integrate $\mathrm{d}r$ "by hand".

### 4.3.2 Closed form expressions for $A_R(z)$ and $V_R(z)$

The function $\mathsf{Iso}_R(z)$ is the reduced volume of $T(R,1)$, i.e. it is given as follows:

$$\mathsf{Iso}_R(z) \coloneqq 6\sqrt{\pi}\cdot\frac{V_R(z)}{A_R^{3/2}(z)},$$

where the functions

$$A_R(z) = 4\pi^2 R\big(1+(4R^2+5)z^2+(9R^4+36R^2+45/4)z^4+\cdots\big)$$

and

$$V_R(z) = 2\pi^2 R\big(1+(9R^2+6)z^2+(36R^4+78R^2+69/4)z^4+\cdots\big)$$

are given in (4.9) and (4.8). For $A_R(z)$ we already proved in Section 4.3.1 that it satisfies the following differential equation:

$$\begin{aligned}0 &= \left(4z^4R_1^5+4R^2+z^2R_1^2\left(4R^2-9\right)+z^3R_1^3\left(12R^2-13\right)-zR_1\left(24R^2-5\right)+5\right)A(z)\\ &+\left(z\left(11R^2+5\right)-4z^2R_1^2+5z^5R_1^5+z^4R_1^3\left(5R^2-13\right)-z^3R_1^2\left(16R^2+8\right)-1\right)A'(z)\\ &+\left(z^2\left(R^2+3\right)-z+z^6R_1^5+2z^3R_1\left(R^2+1\right)-2z^4R_1^2\left(R^2+1\right)-z^5R_1^3\left(R^2+3\right)\right)A''(z),\end{aligned}$$

where $R_1 \coloneqq (R^2-1)$. For $V_R(z)$ we proved that

$$V_R(z) = 2\pi^2 R\int_0^1 w(z,R,r)\mathrm{d}r,$$

where $w = \oint_\gamma v\mathrm{d}x\mathrm{d}y$ satisfies an explicit (but large) second-order differential

$$(S_2(z,R,r)\partial_z^2+S_1(z,R,r)\partial_z+S_0(z,R,r))w(z,R,r)=0, \tag{4.10}$$

equation for $S_0,S_1,S_2\in\mathbb{Q}(z,R,r)$ and with the initial conditions given by:

$$w = 2r\left(1+\left(9R^2+12r^2\right)z^2+\left(36R^4+156R^2r^2+\frac{207}{4}r^4\right)z^4+\cdots\right).$$

In this section we will show how to derive from these expressions the closed form solution in terms of hypergeometric functions:

**Lemma 4.8.** *It holds that*

$$\bar{A}_R(z) \coloneqq \frac{A_R(\sqrt{z})}{\pi^2 R} = \frac{4(1-(R^2-1)^2z^2)}{(1-2(R^2+1)z+(R^2-1)^2z^2)^2}\cdot{}_2F_1\left[\begin{matrix}-\frac{1}{2} & -\frac{1}{2}\\ & 1\end{matrix};\frac{4z}{(1-(R^2-1)z)^2}\right],$$

*and* $\bar{V}_R(z) \coloneqq V_R(\sqrt{z})/(\pi^2R)$ *is given by*

$$\bar{V}_R(z) = \frac{2(1-(R^2-1)z)^3}{(1-2(R^2+1)z+(R^2-1)^2z^2)^3}\cdot{}_3F_2\left[\begin{matrix}-\frac{3}{2} & -\frac{3}{2} & \frac{3}{2R^2-4}+1\\ & 1 & \frac{3}{2R^2-4}\end{matrix};\frac{4z}{(1-(R^2-1)z)^2}\right].$$

From the initial condition $\bar{A}_R(0) = 4$ and the differential equation for $\bar{A}_R(z)$ we immediately obtain using Maple's `dsolve` that

$$\bar{A}_R(z) = \frac{4(1+(R^2-1)z)}{(1-(R^2-1)z)^3} \cdot {}_2F_1\left[\begin{matrix}\frac{3}{2} & \frac{3}{2}\\ 1\end{matrix}; \frac{4z}{(1-(R^2-1)z)^2}\right].$$

We find a small simplification by applying Euler's transformation:

$$\bar{A}_R(z) = \frac{4\left(1-(R^2-1)^2z^2\right)}{(1-2(R^2+1)z+(R^2-1)^2z^2)^2} \cdot {}_2F_1\left[\begin{matrix}-\frac{1}{2} & -\frac{1}{2}\\ 1\end{matrix}; \frac{4z}{(1-(R^2-1)z)^2}\right].$$

For $V_R(z)$ the process of finding a simple expression is more complicated. Maple does not manage directly to find closed form expressions for $w$, however if $R = R_0$ and $r = r_0$ are evaluated to some numbers in $\mathbb{Z}$ then `dsolve` can solve (4.10) in terms of hypergeometric functions. In the first step we therefore perform evaluation-interpolation in order to guess a closed form solution for $w$ and we obtain:

$$w(z,R,r) = q_1(z,R,r){}_2F_1\left[\begin{matrix}1/2 & 1/2\\ 1\end{matrix}; Z\right] + q_2(z,R,r){}_2F_1\left[\begin{matrix}3/2 & 3/2\\ 2\end{matrix}; Z\right], \tag{4.11}$$

where $Z \coloneqq 4r^2z^2/\left(1-z^2\left(R^2-r^2\right)\right)^2$ and $q_1, q_2 \in \mathbb{Q}(z,R,r)$ explicit but big. Proving this guess is trivial, by simply checking that the right-hand side satisfies the ODE (4.10). Maple also does not manage directly to integrate (4.11) with respect to $r$, neither indefinitely nor with the boundaries $0, 1$. We will therefore guess the ODE for $W(z,R,r) = \int w \mathrm{d}r$, a primitive of $w$ with respect to $r$ and then then solve it. So, we expand $w(z,R,r)$ in a series with respect to $r$ up to order 60:

$$w = \frac{-R^4z^4-4R^2z^2-1}{(Rz-1)^5(Rz+1)^5} - 6\frac{\left(R^6z^6+10R^4z^4+12R^2z^2+2\right)z^2}{(Rz-1)^7(Rz+1)^7}r^2 + \cdots,$$

and then call

`seriestodiffeq(int(ser, r), y(r))`,

where `ser` denotes this series expansion. We find a guess for the differential equation (with respect to $r$) for $W(z,R,r)$:

$$\left(T_2(z,R,r)\partial_r^2 + T_1(z,R,r)\partial_r + T_0(z,R,r)\right)W(z,R,r) = 0. \tag{4.12}$$

We solve this ODE with maple (again using evaluation-interpolation) and obtain that

$$W(z,R,r) = Q_1(z,R,r){}_2F_1\left[\begin{matrix}1/2 & 1/2\\ 1\end{matrix}; Z\right] + Q_2(z,R,r){}_2F_1\left[\begin{matrix}3/2 & 3/2\\ 2\end{matrix}; Z\right],$$

for, as before, $Z = 4r^2z^2/\left(1-z^2\left(R^2-r^2\right)\right)^2$ and also $Q_1, Q_2 \in \mathbb{Q}(z,R,r)$. It is easy to check that $\partial_r W = w$, so our guess is confirmed. We find that $W(z,R,0) = 0$ and conclude that $\bar{V}_R(z) = W(\sqrt{z},R,1)$ which is small enough to be displayed here:

$$\bar{V}_R(z) = \frac{2\left(1+\left(R^4-2R^2+1\right)z^2+\left(\frac{10R^2}{3}-2\right)z\right)\cdot {}_2F_1\left[\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ 1\end{matrix}; \frac{4z}{(1-(R^2-1)z)^2}\right]}{\left(1-z\left(R^2-1\right)\right)\left(1-\left(R+1\right)^2z\right)^2\left(1-\left(R-1\right)^2z\right)^2} +$$
$$\frac{4\left(\left(R^6+4R^4-11R^2+6\right)z^2+\left(-2R^4+18R^2-12\right)z+R^2+6\right)z\cdot {}_2F_1\left[\begin{matrix}\frac{3}{2} & \frac{3}{2}\\ 2\end{matrix}; \frac{4z}{(1-(R^2-1)z)^2}\right]}{3\left(1-z\left(R^2-1\right)\right)^3\left(1-\left(R+1\right)^2z\right)^2\left(1-\left(R-1\right)^2z\right)^2}$$

Of course, now proving the expression in Lemma 4.8 is easy. We wish, however, to explain how the expression was found, since this step was not at all trivial.

We first wish to rewrite $V_R(z)$ solely in terms of $x = x(z) := 4z/(1-(R^2-1)z)^2$ in order to find an expression as a single hypergeometric function in $x$. Observe that,

$$\begin{aligned}
1-x &= \frac{\left(-1+(R+1)^2 z\right)^2 \left(-1+(R-1)^2 z\right)^2}{(1-(R^2-1)z)^2},\\
1+(4R^2/3-1)x &= \frac{(1+(R^4-2R^2+1)z^2+\left(\frac{10R^2}{3}-2\right)z}{(1-(R^2-1)z)^2},\\
R^2+6+(7R^2-6)x &= \frac{(R^6+4R^4-11R^2+6)z^2+(-2R^4+18R^2-12)z+R^2+6}{(1-(R^2-1)z)^2}.
\end{aligned}$$

From these cleverly constructed expressions it follows that $(1-(R^2-1)z)^2\bar{V}_R(z)$ is given by

$$\begin{aligned}
\widetilde{V_R}(x) &= \frac{6+(8R^2-6)x}{(1-x)^2}{}_2F_1\left[\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ 1\end{matrix};x\right] + x\frac{R^2+6+(7R^2-6)x}{(1-x)^2}{}_2F_1\left[\begin{matrix}\frac{3}{2} & \frac{3}{2}\\ 2\end{matrix};x\right]\\
&= 6+(27/2+9R^2)x+(675/32+225R^2/8)x^2+(3675/128+3675R^2/64)x^3+\cdots.
\end{aligned}$$

Using already 10 terms in this expansion we can guess with Maple's `gfun` a linear recurrence for the coefficients $u_n$:

$$u_{n+1} = \frac{(2n+3)^2(2R^2n+2R^2+3)}{4(n+1)^2(2R^2n+3)}u_n.$$

If true, this recursion would imply that

$$\widetilde{V_R}(x) = 6\cdot {}_3F_2\left[\begin{matrix}\frac{3}{2} & \frac{3}{2} & \frac{3}{2R^2}+1\\ & 1 & \frac{3}{2R^2}\end{matrix};x\right]. \tag{4.13}$$

Of course, $\widetilde{V_R}(x)$ is D-finite so this guess is easily proven by comparing the differential equations (and initial conditions) for the left-hand and right-hand sides. A generalization of Euler's transformation formula ([243, Thm. 3] or [244, Eq. (1.15)]) now also implies that

$$\widetilde{V_R}(x) = \frac{6}{(1-x)^3}\cdot {}_3F_2\left[\begin{matrix}-\frac{3}{2} & -\frac{3}{2} & \frac{3}{2R^2-4}+1\\ & 1 & \frac{3}{2R^2-4}\end{matrix};x\right].$$

Finally, resubstituting back $x = \frac{4z}{(1-(R^2-1)z)^2}$ in $\bar{V}_R(z) = \widetilde{V_R}(x)/(1-(R^2-1)z)^2$ we find

$$\bar{V}_R(z) = \frac{2\,(1-(R^2-1)z)^3}{(1-2(R^2+1)z+(R^2-1)^2z^2)^3}\cdot {}_3F_2\left[\begin{matrix}-\frac{3}{2} & -\frac{3}{2} & \frac{3}{2R^2-4}+1\\ & 1 & \frac{3}{2R^2-4}\end{matrix};\frac{4z}{(1-(R^2-1)z)^2}\right],$$

as claimed in Lemma 4.8.

Note that it becomes apparent here that $R=\sqrt{2}$ plays a special role since, strictly speaking, $\bar{V}_{\sqrt{2}}(z)$ is not defined because of the term $2R^2-4$ in the denominator of some of the parameters. We also note, however, that this singularity is isolated (for example, because (4.13) is well-defined at $R=\sqrt{2}$).

### 4.3.3 $\mathsf{lso}_R(z)$ is increasing

Combining the definition of $\mathsf{lso}_R$ with Lemma 4.8 we arrive at

$$\mathsf{lso}_R^2(z) = \frac{9}{4\pi R} \frac{{}_3F_2\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2(R^2-2)}+1 \\ & 1 & \frac{3}{2(R^2-2)} \end{matrix}; \frac{4z^2}{(1-(R^2-1)z^2)^2}\right]^2}{{}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; \frac{4z^2}{(1-(R^2-1)z^2)^2}\right]^3} \left(\frac{1-(R^2-1)z^2}{1+(R^2-1)z^2}\right)^3.$$

Note that, analogously to the case $R = \sqrt{2}$ as before in Theorem 4.5,

$$ {}_3F_2\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2(R^2-2)}+1 \\ & 1 & \frac{3}{2(R^2-2)} \end{matrix}; 1\right]^2 = \frac{256R^4}{9\pi^2} \quad \text{and}$$
$$ {}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; 1\right]^3 = \frac{64}{\pi^3},$$

hence $\lim_{z\to 1/(R+1)} \mathsf{lso}(z) = 1$. We will now prove that $\mathsf{lso}_R(z)$ is increasing on $z \in (0, 1/(R+1))$ for any $R > 1$. After the substitution $x = 4z^2/((1-(R^2-1)z^2)^2$ it is enough to show that the function

$$h_R(x) \coloneqq \frac{{}_3F_2\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2(R^2-2)}+1 \\ & 1 & \frac{3}{2(R^2-2)} \end{matrix}; x\right]^2}{{}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; x\right]^3} \cdot (1+(R^2-1)\cdot x)^{-3/2}$$

is increasing on $x \in (0,1)$ for all $R > 1$. Now define two functions

$$f(x) \coloneqq \frac{(x+1)^{1/2}}{{}_2F_1\left[\begin{matrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{matrix}; x\right]} \quad \text{and} \quad g_R(x) \coloneqq \frac{{}_3F_2\left[\begin{matrix} -\frac{3}{2} & -\frac{3}{2} & \frac{3}{2(R^2-2)}+1 \\ & 1 & \frac{3}{2(R^2-2)} \end{matrix}; x\right]}{(1+x)^{3/4}\cdot(1+(R^2-1)\cdot x)^{3/4}}$$

and note that clearly $h_R = f^3 \cdot g_R^2$. We will show that both $f$ and $g_R$ increase on $x \in (0,1)$ for $R > 1$. Since $f$ and $g_R$ are positive on this interval, this will be enough to conclude our main statement.

The fact that $f(x)$ is increasing on $(0,1)$ is already proved in Section 4.2. For $g_R(x)$ a similar argument works, which is, however, not exactly the same method as in Section 4.2, since the parameter $R$ indeed complicates matters. For example, it seems that (at least) the (naive) approach with Sturm-Liouville theory is doomed, because the weight of the resulting differential equation is not strictly positive anymore.

For the derivative $g_R'(x)$ we can still obtain the following: Let

$$\frac{4\cdot g_R'(x)\cdot(1+x)^{7/4}\cdot(1+(R^2-1)\cdot x)^{7/4}}{3\cdot(1-x)^2\cdot(R^2-1)} =: \sum_{n\geq 0} u_n(R)x^n,$$

for some rational functions $u_n(R) \in \mathbb{Q}(R)$, $n \geq 1$ and $u_0(R) = 1$. By observing that the sequence $(u_n(R))_{n\geq 0}$ is hypergeometric, we will prove that $u_n(R) > 0$ for $R > 1$. Clearly, this implies that $g_R'(x) > 0$, i.e. $g_R(x)$ is increasing.

For all $n \geq 0$ define the sequence of polynomials

$$p_n(R) \coloneqq 4(R^4 + 4R^2 - 4)n^3 + 6(R^4 + R^2 - 2)n^2 + (2R^4 - 13R^2 + 10)n - 3R^2 + 3.$$

Then it is not difficult to check with a computer that for $n \geq 0$ we have

$$\frac{u_{n+1}(R)}{u_n(R)} = \frac{(2n-1)(2n+1) \cdot p_{n+1}(R)}{4(n+2)(n+1) \cdot p_n(R)}.$$

For example, $u_1 = -p_1(R)/(4p_0(R)) = p_1(R)/(12(R^2-1))$. Hence, it holds that $u_n(R) > 0$ for $R > 1$ if we can prove that $p_n(R) > 0$ for all $n \geq 1$. Observe that for $n \geq 1$ and $R > 1$:

$$p_n(R) > 4R^4n^3 + (10 - 11R^2)n + 3 - 3R^2.$$

The latter polynomial is strictly increasing in $n$ for $n \geq 1$ since the larger root of its derivative is $\sqrt{36R^2 - 33}/(6R^2) < 1$. Therefore we can set $n = 1$ and obtain

$$p_n(R) > p_1(R) = (2R^2 - 14/4)^2 + 3/4 > 0.$$

# Chapter 5

# Computing terms in $q$-holonomic sequences

*"Someone has to do it"*
D. V. Chudnovsky, G. V. Chudnovsky,
*Talk*[1]*: Calculation of Classical Constants and Special Functions for Fun and Profit*, 2019

In 1977, Strassen invented a famous baby-step/giant-step algorithm that computes the factorial $N!$ in arithmetic complexity quasi-linear in $\sqrt{N}$. In 1988, the Chudnovsky brothers generalized Strassen's algorithm to the computation of the $N$-th term of any holonomic sequence in essentially the same arithmetic complexity. This chapter presents $q$-analogues of these algorithms. We first extend Strassen's algorithm to the computation of the $q$-factorial of $N$, then Chudnovskys' algorithm to the computation of the $N$-th term of any $q$-holonomic sequence. Both algorithms work in arithmetic complexity quasi-linear in $\sqrt{N}$; surprisingly, they are simpler than their analogues in the holonomic case. We provide a detailed cost analysis, in both arithmetic and bit complexity models. Moreover, we describe various algorithmic consequences, including the acceleration of polynomial and rational solving of linear $q$-differential equations, and the fast evaluation of large classes of polynomials, including a family recently considered by Nogneng and Schost.

This chapter consists of the joint work with A. Bostan [83].

## 5.1 Introduction

A classical question in algebraic complexity theory is: how fast can one evaluate a univariate polynomial at one point? The precise formulation of this question depends on the model of computation. We will mainly focus on the *arithmetic complexity* model, in which one counts base field operations at unit cost.

Horner's rule evaluates a polynomial $P$ in $O(\deg(P))$ operations. Ostrowski [251] conjectured in 1954 that this is *optimal for generic polynomials*, *i.e.*, whose coefficients are algebraically independent over the prime subfield. This optimality result was proved a few years later by Pan [256].

[1] At the fantastic event "Transient Transcendence In Transylvania" organized by A. Bostan and K. Raschel.

However, most polynomials that one might wish to evaluate "in practice" have coefficients which are not algebraically independent. Paterson and Stockmeyer [258] showed, using the *baby-step/giant-step* technique, that for any field $\mathbb{K}$, an arbitrary polynomial $P \in \mathbb{K}[x]$ of degree $N$ can be evaluated at any point in an arbitrary $\mathbb{K}$-algebra $A$ using $O(\sqrt{N})$ *nonscalar* multiplications, *i.e.*, multiplications in $A$. However, their algorithm uses a linear amount of scalar multiplications, so it is not well adapted to the evaluation at points from the base field $\mathbb{K}$, since in this case the total arithmetic complexity, counted in terms of operations in $\mathbb{K}$, remains linear in $N$.

For some families of polynomials, one can do much better. Typical examples are $x^N$ and

$$P_N(x) \coloneqq x^{N-1} + \cdots + x + 1,$$

which can be evaluated by the *square-and-multiply* technique in $O(\log N)$ operations[2]. By contrast, a family $(F_n(x))_n$ of univariate polynomials is called *hard to compute* if for large enough $N$, the complexity of the evaluation of $F_N$ grows at least like a power in $\deg(F_N)$, whatever the algorithm used.

Paterson and Stockmeyer [257, 258] proved the existence of polynomials in $\mathbb{Q}[x]$ which are hard to compute (note that this does not follow from Pan's result [256]). However, their proof was based on a non-constructive argument. Specific families of hard-to-compute polynomials were first exhibited by Strassen [306]. For instance, he proved that for large $N$, the polynomial $\sum_{\ell=0}^{N} 2^{2^\ell} x^\ell$ needs at least $\sqrt{N/(3\log N)}$ operations to be evaluated. The techniques were refined and improved by Borodin and Cook [48], Lipton [230] and Schnorr [283], who produced explicit examples of degree-$N$ polynomials whose evaluation requires a number of operations linear in $\sqrt{N}$. Subsequently, various methods have been developed to produce similar results on *lower bounds*, *e.g.*, by Heintz and Sieveking [177] using algebraic geometry, and by Aldaz et al. [13] using a combinatorial approach. The topic is vast and very well summarized in the book by Bürgisser, Clausen and Shokrollahi [87].

In this chapter, we focus on *upper bounds*, that is on the design of fast algorithms for special families of polynomials, which are hard to compute, but easier to evaluate than generic polynomials. For instance, for the degree-$\binom{N}{2}$ polynomial

$$Q_N(x) \coloneqq P_1(x) \cdots P_N(x),$$

a complexity in $O(N)$ is clearly achievable. We will see in §5.2.1 that one can do better, and attain a cost which is almost linear in $\sqrt{N}$ (up to logarithmic factors in $N$). Another striking example is

$$R_N(x) \coloneqq \sum_{\ell=0}^{N} x^{\ell^2},$$

of degree $N^2$, and whose evaluation can also be performed in complexity quasi-linear in $\sqrt{N}$, as shown recently by Nogneng and Schost [249] (see §5.2.2). In both cases, these

[2]For the latter this is obvious, due to the formula $P_N(x) = (x^N - 1)/(x - 1)$ for $x \neq 1$, and it was observed in [51] that this approach requires divisions. We note that, in fact, one can evaluate $P_N(x)$ in $O(\log N)$ ring operations without any divisions, by using the decomposition $P_N(x) = P_{N/2}(x)x^{N/2} + P_{N/2}(x)$ when $N$ is even, and a similar one when $N$ is odd.

complexities are obtained by clever although somehow ad-hoc algorithms. The starting point of our work [83] which led to this chapter was the question whether these algorithms for $Q_N(x)$ and $R_N(x)$ could be treated in a unified way, which would allow to evaluate other families of polynomials in a similar complexity.

The answer to this question turns out to be positive. The key idea, very simple and natural, is to view both examples as particular cases of the following general question:

> Given a $q$-holonomic sequence, that is, a sequence satisfying a linear recurrence with polynomial coefficients in $q$ and $q^n$, how fast can one compute its $N$-th term?

In the more classical case of holonomic sequences (satisfying linear recurrences with polynomial coefficients in the index $n$), fast algorithms exist for the computation of the $N$-th term. They rely on a basic block, which is the computation of the factorial term $N!$ in arithmetic complexity quasi-linear in $\sqrt{N}$, using an algorithm due to Strassen [307]. The Chudnovsky brothers extended in [106] Strassen's algorithm to the computation of the $N$-th term of any holonomic sequence in arithmetic complexity quasi-linear in $\sqrt{N}$.

Our main contribution in this chapter consists in transferring these results to the $q$-holonomic framework. It turns out that the resulting algorithms are actually simpler in the $q$-holonomic case than in the usual holonomic setting, essentially because multipoint evaluation on arithmetic progressions used as a subroutine in Strassen's and Chudnovskys' algorithms is replaced by multipoint evaluation on geometric progressions, which is considerably simpler [77].

A consequence of our results is that the following apparently unrelated polynomials and rational functions can be evaluated fast (note the change in notation, with the variable $x$ denoted now by $q$):

- $A_n(q)$, the generating function of the number of partitions into $n$ positive integers each occurring *at most twice* [332], *i.e.*, the coefficient of $t^n$ in the product
$$\prod_{k\geq 1}(1+q^k t+q^{2k}t^2).$$
- $B_n(q) \coloneqq \prod_{i=1}^{\infty}(1-q^i) \bmod q^n$; by Euler's pentagonal theorem [254, §5],
$$B_n(q) = 1+\sum_{i(3i+1)<2n}(-1)^i\left(q^{\frac{i(3\,i-1)}{2}}+q^{\frac{i(3\,i+1)}{2}}\right).$$
- The number $C_n(q)$ of $2n\times 2n$ upper-triangular matrices over $\mathbb{F}_q$ (the finite field with $q$ elements), whose square is the zero matrix [206]; by [130], $C_n(q)$ is equal to
$$C_n(q)=\sum_j\left[\binom{2n}{n-3j}-\binom{2n}{n-3j-1}\right]\cdot q^{n^2-3j^2-j}.$$

The common feature, exploited by the new algorithm, is that the sequences $(A_n(q))_{n\geq 0}$, $(B_n(q))_{n\geq 0}$, $(C_n(q))_{n\geq 0}$ are all $q$-holonomic. Actually, $q$-holonomic sequences are ubiquitous,

so the range of application of our results is quite broad. This stems from the fact that they are coefficient sequences of power series satisfying $q$-difference equations, or equivalently, $q$-shift (or, $q$-differential) equations. From that perspective, our topic becomes intimately connected with $q$-calculus. The roots of $q$-calculus are in works of famous mathematicians such as Rothe [276], Gauss [160] and Heine [174]. The topic gained renewed interest in the first half of the 20th century, with the work, both on the formal and analytic aspects, of Tanner [312], Jackson [188, 189, 190], Carmichael [92], Mason [237], Adams [8, 9], Trjitzinsky [317], Le Caine [227] and Hahn [165], to name just a few. Modern accounts of the various aspects of the theory (including historical ones) can be found in [125, 207, 132].

One of the reasons for interest in $q$-differential equations is that, formally, as $q$ tends to 1, the $q$-derivative $\frac{f(qx)-f(x)}{(q-1)x}$ tends to $f'(x)$, thus to every differential equation corresponds a $q$-differential equation which goes formally to the differential equation as $q \to 1$. In nice cases, (some of) the solutions of the $q$-difference equation go to solutions of the associated differential equation as $q \to 1$. An early example of such a good deformation behavior is given by the basic hypergeometric equation of Heine [174], see also [207, §1.10].

In computer algebra, $q$-holonomic sequences were considered starting from the early nineties, in the context of computer-generated proofs of identities in the seminal paper by Wilf and Zeilberger [327], notably in Section 5 ("Generalization to $q$-sums and $q$-multisums") and in Section 6.4 ("$q$-sums and integrals"). Creative telescoping algorithms for (proper) $q$-hypergeometric sequences are discussed in various references [263, 47, 93]; several implementations of those algorithms are described for instance in [260, 271, 195, 298]. Algorithms for computing polynomial, rational and $q$-hypergeometric solutions of $q$-difference equations were designed by Abramov and collaborators [3, 2, 5, 204]. These algorithms are important for several reasons. One is that they lie at the heart of the vast generalization by Chyzak [110, 111] of the Wilf and Zeilberger algorithmic theory, for the treatment of general $q$-holonomic (not only $q$-hypergeometric) symbolic summation and integration via creative telescoping. In that context, a multivariate notion of $q$-holonomy is needed; the foundations of the theory were laid by Zeilberger [340] and Sabbah [278] (in the language of D-modules), see also [93, § 2.5] and [156].

The simplest non-trivial holonomic sequence is $(n!)_{n\geq 0}$, whose $n$-th term combinatorially counts the number of permutations of $n$ objects. If instead of direct counting, one assigns to every permutation $\pi$ its number of inversions $\mathrm{inv}(\pi)$, *i.e.*, the number of pairs $1 \leq i < j \leq n$ with $\pi(i) > \pi(j)$, the refined count (by size and number of inversions) is

$$[n]_q! := (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}).$$

This is the $q$-analogue of $n!$; it is the simplest non-trivial $q$-holonomic sequence.

There is also a natural $q$-analog of the binomial coefficients, called the *Gaussian coefficients*, defined by

$$\binom{n}{k}_q := \frac{[n]_q!}{[k]_q![n-k]_q!}.$$

They have many counting interpretations, *e.g.,* they count the $k$-dimensional subspaces of $\mathbb{F}_q^n$ (points on Grassmannians over $\mathbb{F}_q$). There are $q$-analogs to (almost) everything. To select just two more basic examples, the $q$-analog [21, Thm. 3.3] of the binomial theorem

is given by

$$\prod_{k=1}^{n}(1+q^{k-1}x)=\sum_{k=0}^{n}\binom{n}{k}_q q^{\binom{k}{2}}x^k, \tag{5.1}$$

and the $q$-version [21, Thm. 3.4] of the Chu-Vandermonde identity is

$$\sum_{k=0}^{n} q^{k^2}\binom{m}{k}_q\binom{n}{k}_q=\binom{m+n}{n}_q. \tag{5.2}$$

The ubiquity of $q$-holonomic sequences is manifest in plenty of fields: partition theory [293, 21, 22, 332, 254, 231] and other subfields of combinatorics [144, 86, 205, 130, 206, 23, 333]; theta functions and modular forms [33, 336, 219, 218, 158]; special functions [50, 187, 207] and in particular orthogonal polynomials [212]; algebraic geometry [129], representation theory [185]; knot theory [155, 152, 153, 156, 154]; Galois theory [178]; number theory [250, 123, 7].

> The main messages of this chapter are that for any example of a $q$-holonomic sequence occurring in those various fields, *one can compute selected coefficients faster than by a direct algorithm* and that *this fact finds a tremendous number of applications*.

**Complexity basics.** We estimate the arithmetic complexities of algorithms by counting arithmetic operations $(+,-,\times,\div)$ in the base field $\mathbb{K}$ at unit cost. We use standard complexity notation, such as $\mathbf{M}(d)$ for the cost of degree-$d$ multiplication in $\mathbb{K}[x]$, and $\theta$ for feasible exponents of matrix multiplication. The best currently known upper bound is $\theta<2.3729$ [148, 15]. As usual, $O(\cdot)$ stands for the big-Oh notation and $\tilde{O}(\cdot)$ is used to hide polylogarithmic factors in the argument. Most arithmetic operations on univariate polynomials of degree $d$ in $\mathbb{K}[x]$ can be performed in quasi-linear complexity $\tilde{O}(d)$: multiplication, shift, interpolation, gcd, resultant, *etc.* A key feature of these results is the reduction to fast polynomial multiplication, which can be performed in time $\mathbf{M}(d)=O(d\log d\log\log d)$ [285, 90]. Finally, the arithmetic cost of multiplication of polynomial matrices of size $n$ and degree $d$ is denoted by $\mathbf{MM}(n,d)$ and we have $\mathbf{MM}(n,d)=O(n^\theta d+n^2\mathbf{M}(d))=\tilde{O}(n^\theta d)$ [77]. An excellent general reference for these questions is the book by von zur Gathen and Gerhard [159].

A short version of the article [83] has appeared at the ISSAC'20 conference [51]. In [83] as well as in the present chapter, we included the proofs of Theorems 5.7 and 5.9, we added a new Theorem 5.6 containing a detailed complexity analysis of the main algorithm (Algorithm 3) with respect to all parameters, and we displayed pseudo-code for the algorithms as well as figures visualising their performance. We also elaborated on a task which was mentioned as future work in the previous version, namely the application of our methods to the computation of curvatures of $q$-difference equations, see §5.4.4.

The structure of the chapter is as follows: in Section 5.2 we deal with the tasks of evaluating $Q_N(x)$ and $R_N(x)$. We show that these are two instances of the same problem and provide Algorithm 3 which solves both in $O(\mathbf{M}(\sqrt{N}))$ arithmetic complexity. Section 5.3

is devoted to the main results; we prove there that Algorithm 3 can be used for computing terms of any $q$-holonomic sequence with the same cost, and provide extensions and more insight. In the same section we also consider the bit-complexity model. We identify and elaborate on several applications for our result in Section 5.4. In Section 5.5 we report on implementations of our algorithms, which deliver encouraging timings, and we finally describe future tasks and investigation fields in Section 5.6.

## 5.2 Two motivating examples

Before presenting our main results in Section 5.3, we describe in this section the approach and main ideas on two basic examples. Both examples concern the fast evaluation of special families of univariate polynomials. In §5.2.1, we consider polynomials of the form $\prod_\ell (x - q^\ell)$, and in §5.2.2 sparse polynomials of the form $\sum_\ell p^\ell x^{a\ell^2+b\ell}$. In both cases, we first present fast ad-hoc algorithms, then introduce equally fast alternative algorithms, which have the nice feature that they will be generalizable to a broader setting.

### 5.2.1 Evaluation of some structured polynomials

Here is our first example, that emerged from a question asked to the first author by Luca De Feo (private email communication, 10 January 2020); this was the starting point of the article [83] and consequently this chapter.

Let $q$ be an element of the field $\mathbb{K}$, and consider the polynomial

$$F(x) := \prod_{i=0}^{N-1} (x - q^i) \ \in \mathbb{K}[x]. \tag{5.3}$$

Given another element $\alpha \in \mathbb{K}$, how fast can one evaluate $F(\alpha)$?

If $q = 0$, then $F(\alpha) = \alpha^N$ can be computed in $O(\log N)$ operations in $\mathbb{K}$, by binary powering. We assume in what follows that $q$ is nonzero. Obviously, a direct algorithm consists in computing the successive powers $q, q^2, \ldots, q^{N-1}$ using $O(N)$ operations in $\mathbb{K}$, then computing the elements $\alpha - 1, \alpha - q, \ldots, \alpha - q^{N-1}$ in $O(N)$ more operations in $\mathbb{K}$, and finally returning their product. The total arithmetic cost of this algorithm[3] is $O(N)$, linear in the degree of $F$.

Is it possible to do better? The answer is positive, as one can use the following *baby-step/giant-step* strategy, in which, in order to simplify things, we assume that $N$ is a perfect square, $N = s^2$.

Algorithm 1

[3] If $q^n = 1$ for some $n < N$, then it is enough to compute the product of $\alpha - q^i$ for $i = 0, \ldots, n-1$ and its appropriate power. The latter step can be done efficiently (in essentially $\log(N)$ operations) using binary powering. Our main interest lies therefore in $q \in \mathbb{K}$ that are not roots of unity of small order compared to $N$.

1. (Baby-step) Compute the values of $q, q^2, \ldots, q^{s-1}$, and deduce the coefficients of the polynomial
$$G(x) \coloneqq \prod_{j=0}^{s-1}(x - q^j).$$

2. (Giant-step) Compute $Q \coloneqq q^s, Q^2, \ldots, Q^{s-1}$, and deduce the coefficients of the polynomial
$$H(x) \coloneqq \prod_{k=0}^{s-1}(\alpha - Q^k \cdot x).$$

3. Return the resultant $\mathrm{Res}(G, H)$.

By the basic property of resultants, the output of this algorithm is

$$\mathrm{Res}(G,H) = \prod_{j=0}^{s-1} H(q^j) = \prod_{j=0}^{s-1}\prod_{k=0}^{s-1}\left(\alpha - q^{sk+j}\right) = \prod_{i=0}^{N-1}(\alpha - q^i) = F(\alpha).$$

Using the fast subproduct tree algorithm [159, Algorithm 10.3], one can perform the baby-step (1) as well as the giant-step (2) in $O(\mathbf{M}(\sqrt{N})\log N)$ operations in $\mathbb{K}$, and by [159, Corollary 11.19] the same cost can be achieved for the resultant computation in step (3). Using fast polynomial multiplication, we conclude that $F(\alpha)$ can be computed in arithmetic complexity quasi-linear in $\sqrt{N}$.

Note that if $N$ is not a perfect square, then one can compute $F(\alpha)$ as $F(\alpha) = F_1(\alpha)F_2(\alpha)$, where $F_1(\alpha) \coloneqq \prod_{i=0}^{\lfloor\sqrt{N}\rfloor^2-1}(\alpha - q^i)$ is computed as in Algorithm 1, while $F_2(\alpha) \coloneqq \prod_{i=\lfloor\sqrt{N}\rfloor^2}^{N-1}(\alpha - q^i)$ can be computed naively, since $N - \lfloor\sqrt{N}\rfloor^2 = O(\sqrt{N})$.

It is possible to speed up the previous algorithm by a logarithmic factor in $N$ using a slightly different scheme, still based on a *baby-step/giant-step* strategy, but exploiting the fact that the roots of $F$ are in geometric progression. Again, we assume that $N = s^2$ is a perfect square. This alternative algorithm goes as follows. Note that it is very close in spirit to Pollard's algorithm described on page 523 of [266].

Algorithm 2

1. (Baby-step) Compute $q, q^2, \ldots, q^{s-1}$, and deduce the coefficients of the polynomial $P(x) \coloneqq \prod_{j=0}^{s-1}(\alpha - q^j \cdot x)$.

2. (Giant-step) First compute $Q \coloneqq q^s, Q^2, \ldots, Q^{s-1}$, and then evaluate $P$ simultaneously at $1, Q, \ldots, Q^{s-1}$.

3. Return the product $P(Q^{s-1}) \cdots P(Q)P(1)$.

Obviously, the output of this algorithm is

$$\prod_{k=0}^{s-1} P(Q^k) = \prod_{k=0}^{s-1}\prod_{j=0}^{s-1}(\alpha - q^j \cdot q^{sk}) = \prod_{i=0}^{N-1}(\alpha - q^i) = F(\alpha).$$

As pointed out in the remarks after the proof of [77, Lemma 1], one can compute $P(x) = P_s(x) = \prod_{j=0}^{s-1}(\alpha - q^j \cdot x)$ in step (1) without computing the subproduct tree, by using a divide-and-conquer scheme which exploits the fact that $P_{2t}(x) = P_t(q^t x) \cdot P_t(x)$ and $P_{2t+1}(x) = (\alpha - q^{2t}x) \cdot P_t(q^t x) \cdot P_t(x)$. The cost of this algorithm is $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$.

As for step (2), one can use the fast *chirp transform* algorithms of Rabiner, Schafer and Rader [269] and of Bluestein [45]. These algorithms rely on the following observation: writing $Q^{ij} = Q^{\binom{i+j}{2}} \cdot Q^{-\binom{i}{2}} \cdot Q^{-\binom{j}{2}}$ and $P(x) = \sum_{j=0}^{s} c_j x^j$ implies that the needed values $P(Q^i) = \sum_{j=0}^{s} c_j Q^{ij}, 0 \leq i < s$, are

$$P(Q^i) = Q^{-\binom{i}{2}} \cdot \sum_{j=0}^{s} c_j Q^{-\binom{j}{2}} \cdot Q^{\binom{i+j}{2}}, \ 0 \leq i < s,$$

in which the sum is simply the coefficient of $x^{s+i}$ in the product

$$\left(\sum_{j=0}^{s} c_j Q^{-\binom{j}{2}} x^{s-j}\right)\left(\sum_{\ell=0}^{2s} Q^{\binom{\ell}{2}} x^{\ell}\right).$$

This polynomial product can be computed in $2\,\mathbf{M}(s)$ operations (and even in $\mathbf{M}(s) + O(s)$ using the *transposition principle* [166, 73], since only the median coefficients $x^s, \ldots, x^{2s-1}$ are actually needed). In conclusion, step (2) can also be performed in $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$, and thus $O(\mathbf{M}(\sqrt{N}))$ is the total cost of this second algorithm.

We have chosen to detail this second algorithm for several reasons: not only because it is faster by a factor $\log(N)$ compared to the first one, but more importantly because it has a simpler structure, which will be generalizable to the general $q$-holonomic setting. In fact, we do not provide a pseudo-code implementation for this algorithm, since we will do so for the more general case (Algorithm 3).

### 5.2.2 Evaluation of some sparse polynomials

Let us now consider the sequence of sparse polynomial sums

$$v_N^{(p,a,b)}(q) = \sum_{n=0}^{N-1} p^n q^{an^2+bn},$$

where $p \in \mathbb{K}$ and $a, b \in \mathbb{Q}$ such that $2a, a+b$ are both integers. Typical examples are (truncated) modular forms [259], which are ubiquitous in complex analysis [33], number theory [336] and combinatorics [21]. For instance, the *Jacobi theta function* $\vartheta_3$ depends on two complex variables $z \in \mathbb{C}$, and $\tau \in \mathbb{C}$ with $\Im(\tau) > 0$, and it is defined by

$$\vartheta_3(z;\tau) = \sum_{n=-\infty}^{\infty} e^{\pi i(n^2\tau + 2nz)} = 1 + 2\sum_{n=1}^{\infty} \eta^n q^{n^2},$$

where $q = e^{\pi i \tau}$ is the nome ($|q| < 1$) and $\eta = e^{2\pi i z}$. Here, $\mathbb{K} = \mathbb{C}$. Another example is the *Dedekind eta function*, appearing in Euler's famous *pentagonal theorem* [254, §5], which

has a similar form

$$q^{\frac{1}{24}} \cdot \left( 1 + \sum_{n=1}^{\infty} (-1)^n \left( q^{\frac{n(3\,n-1)}{2}} + q^{\frac{n(3\,n+1)}{2}} \right) \right), \quad \text{with } q = e^{2\pi i \tau}.$$

Moreover, sums of the form $v_N^{(1,a,b)}(q) = \sum_{n=0}^{N-1} q^{an^2+bn}$, over $\mathbb{K} = \mathbb{Q}$ or $\mathbb{K} = \mathbb{F}_2$, crucially occur in a recent algorithm by Tao, Crott and Helfgott [313] for the efficient construction of prime numbers in given intervals, *e.g.*, in the context of effective versions of Bertrand's postulate. Actually, (the proof of) Lemma 3.1 in [313] contains the first sublinear complexity result for the evaluation of the sum $v_N^{(p,a,b)}(q)$ at an arbitrary point $q$; namely, the cost is $O(N^{\theta/3})$, where $\theta \in [2,3]$ is any feasible exponent for matrix multiplication. Subsequently, Nogneng and Schost [249] designed a faster algorithm, and lowered the cost down to $\tilde{O}(\sqrt{N})$. Our algorithm is similar in spirit to theirs, as it also relies on a *baby-step/giant-step* strategy.

Let us first recall the principle of the Nogneng-Schost algorithm [249]. Assume as before that $N$ is a perfect square, $N = s^2$. The starting point is the remark that

$$v_N^{(p,a,b)}(q) = \sum_{n=0}^{N-1} p^n q^{an^2+bn} = \sum_{k=0}^{s-1} \sum_{j=0}^{s-1} p^{j+sk} q^{a(j+sk)^2+b(j+sk)}$$

can be written

$$\sum_{k=0}^{s-1} p^{sk} q^{as^2k^2+bsk} \cdot P(q^{2ask}), \text{ where } P(y) \coloneqq \sum_{j=0}^{s-1} p^j q^{aj^2+bj} y^j.$$

Therefore, the computation of $v_N^{(p,a,b)}(q)$ can be reduced essentially to the simultaneous evaluation of the polynomial $P$ at $s = 1 + \deg(P)$ points (in geometric progression), with arithmetic cost $O(\mathbf{M}(\sqrt{N}))$.

We now describe an alternative algorithm, of similar complexity $O(\mathbf{M}(\sqrt{N}))$, with a slightly larger constant in the big-Oh estimate, but whose advantage is its potential of generality.

Let us denote by $u_n(q)$ the summand $p^n q^{an^2+bn}$. Clearly, the sequence $(u_n(q))_{n\geq 0}$ satisfies the recurrence relation

$$u_{n+1}(q) = A(q, q^n) \cdot u_n(q), \quad \text{where} \quad A(x,y) \coloneqq p x^{a+b} y^{2a}.$$

As an immediate consequence, the sequence with general term $v_n(q) \coloneqq \sum_{k=0}^{n-1} u_k(q)$ satisfies a similar recurrence relation

$$v_{n+2}(q) - v_{n+1}(q) = A(q, q^n) \cdot (v_{n+1}(q) - v_n(q)), \tag{5.4}$$

with initial conditions $v_0(q) = 0$ and $v_1(q) = 1$. This scalar recurrence of order two is equivalent to the first-order matrix recurrence

$$\begin{bmatrix} v_{n+2} \\ v_{n+1} \end{bmatrix} = \begin{bmatrix} A(q,q^n)+1 & -A(q,q^n) \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} v_{n+1} \\ v_n \end{bmatrix}.$$

By unrolling this matrix recurrence, we deduce that

$$\begin{bmatrix} v_{n+1} \\ v_n \end{bmatrix} = M(q^{n-1}) \begin{bmatrix} v_n \\ v_{n-1} \end{bmatrix} = M(q^{n-1}) \cdots M(q)M(1) \times \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where

$$M(x) := \begin{bmatrix} pq^{a+b}x^{2a} + 1 & -pq^{a+b}x^{2a} \\ 1 & 0 \end{bmatrix},$$

hence $v_N = \begin{bmatrix} 0 & 1 \end{bmatrix} \times M(q^{N-1}) \cdots M(q)M(1) \times \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Therefore, the computation of $v_N$ reduces to the computation of the "matrix $q$-factorial" $M(q^{N-1}) \cdots M(q)M(1)$, which can be performed fast by using a *baby-step/giant-step* strategy similar to the one of the second algorithm in §5.2.1. Again, we assume for simplicity that $N = s^2$ is a perfect square. The algorithm goes as follows.

Algorithm 3 (matrix $q$-factorial)

(1) (Baby-step) Compute $q, q^2, \ldots, q^{s-1}$; deduce the coefficients of the polynomial matrix $P(x) := M(q^{s-1}x) \cdots M(qx)M(x)$.

(2) (Giant-step) Compute $Q := q^s, Q^2, \ldots, Q^{s-1}$, and evaluate (the entries of) $P(x)$ simultaneously at $1, Q, \ldots, Q^{s-1}$.

(3) Return the product $P(Q^{s-1}) \cdots P(Q)P(1)$.

Clearly, this algorithm generalizes Algorithm 2 in §5.2.1 and, as promised, we also provide a detailed pseudo-code implementation: Step1 and Step2 & Step3:

**Algorithm 3** (Step1)
**Input**: $s, q, M(x)$
**Output**: $M(q^{s-1}x) \cdots M(qx)M(x)$

```
1:  q_s ← [q, q^2, ..., q^{s-1}]
2:  t ← s
3:  function BS(t)
4:      if t = 1 then
5:          return M(x)
6:      end if
7:      if t is even then
8:          p_1(x) ← BS(t/2)
9:          p_2(x) ← p_1(q^{t/2}x)                    ▷ Using q_s
10:         return p_2(x) · p_1(x)                    ▷ Fast polynomial multiplication
11:     else
12:         p_1(x) ← BS((t − 1)/2)
13:         p_2(x) ← p_1(q^{(t−1)/2}x)                ▷ Using q_s
14:         p_3(x) ← M(q^{t−1}x)                      ▷ Using q_s
15:         return p_3(x) · p_2(x) · p_1(x)           ▷ Fast polynomial multiplication
16:     end if
17: end function
```

**Algorithm 3** (Step2 & Step3)
**Input**: $s, Q, P(x)$
**Output**: $P(Q^{s-1}) \cdots P(1)$

**Assumptions:** $Q \neq 0$, $P(x)$ polynomial matrix of size $n \times n$ and degree $d \geq s$.

$Q_d \leftarrow [Q, Q^2, \ldots, Q^{d-1}]$
$Q' \leftarrow 1/Q$
$Q'_d \leftarrow [Q^{-\binom{d}{2}}, \ldots, Q^{-\binom{1}{2}}, Q^{\binom{0}{2}}, \ldots, Q^{\binom{2d}{2}}]$ ▷ Using $Q_d$ and $Q'$
$P_{s-1}, \ldots, P_0$ ▷ Empty $n \times n$ matrices
**for** $i$ from 1 to $n$ **do**
  **for** $j$ from 1 to $n$ **do**
    $p(x) \leftarrow P(x)_{i,j}$ ▷ $p(x) = c_0 + c_1 x + \cdots + c_d x^d$
    $p_1(x) \leftarrow \sum_{\ell=0}^{d} c_\ell Q^{-\binom{\ell}{2}} x^{d-\ell}$ ▷ Using $Q'_d$
    $p_2(x) \leftarrow \sum_{\ell=0}^{2d} Q^{\binom{\ell}{2}} x^\ell$ ▷ Using $Q'_d$
    $p_3(x) \leftarrow P_1(x) \cdot P_2(x)$ ▷ $p_3(x) = \sum_{\ell=0}^{3d} r_\ell x^\ell$; fast multiplication
    **for** $k$ from 0 to $s-1$ **do**
      $(P_k)_{i,j} \leftarrow r_{d+k}$ ▷ $P_\ell = P(Q^\ell)$ for $\ell = 0, \ldots, s-1$
    **end for**
  **end for**
**end for**
$P \leftarrow 1$
**for** $k$ from 0 to $s-1$ **do**
  $P \leftarrow P_k \cdot P$
**end for**
**return** $P$

By the same observations as in Algorithm 2 in §5.2.1, the complexity of Algorithm 3 already is quasi-linear in $\sqrt{N}$. In the next section we will discuss the complexity not only with respect to $N$, but to the matrix size and degree as well.

We remark that when applied to the computation of $v_N^{(p,a,b)}(q)$, the dependence in $a, b$ of Algorithm 3 is quite high (quasi-linear in $a$ and $b$). If $a$ and $b$ are fixed and considered as $O(1)$ this dependence is invisible, but otherwise the following variant has the same complexity with respect to $N$, and a much better cost with respect to $a$ and $b$. It is based on the simple observation that, if $\tilde{M}(x)$ denotes the polynomial matrix

$$\tilde{M}(x) := \begin{bmatrix} prx + 1 & -prx \\ 1 & 0 \end{bmatrix}, \text{ with } r := q^{a+b}, \tag{5.5}$$

and if $\tilde{q} := q^{2a}$, then the following matrix $q$-factorials coincide:

$$M(q^{N-1}) \cdots M(q) M(1) = \tilde{M}(\tilde{q}^{N-1}) \cdots \tilde{M}(\tilde{q}) \tilde{M}(1).$$

Algorithm 4 (matrix $q$-factorial, variant)

(0) (Precomputation) Compute $r := q^{a+b}$, $\tilde{q} := q^{2a}$, and $\tilde{M}$ in (5.5).

(1) (Baby-step) Compute $\tilde{q}, \tilde{q}^2, \ldots, \tilde{q}^{s-1}$; deduce the coefficients of the polynomial matrix

$$\tilde{P}(x) := \tilde{M}(\tilde{q}^{s-1}x) \cdots \tilde{M}(\tilde{q}x)\tilde{M}(x).$$

(2) (Giant-step) Compute $\tilde{Q} := \tilde{q}^s, \tilde{Q}^2, \ldots, \tilde{Q}^{s-1}$, and evaluate (the entries of) $\tilde{P}(x)$ simultaneously at $1, \tilde{Q}, \ldots, \tilde{Q}^{s-1}$.

(3) Return the product $\tilde{P}(\tilde{Q}^{s-1}) \cdots \tilde{P}(\tilde{Q})\tilde{P}(1)$.

Using binary powering, the cost of the additional precomputation in step (0) is only logarithmic in $a$ and $b$. In exchange, the new steps (2) and (3) are performed on matrices whose degrees do not depend on $a$ and $b$ anymore (in the previous, unoptimized, version the degrees of the polynomial matrices were linear in $a$ and $b$). The total arithmetic cost with respect to $N$ is still quasi-linear in $\sqrt{N}$.

In the next section, we will show that Algorithm 3 can be employed for the fast computation of the $N$-th term of *any* $q$-holonomic sequence. Note that the trick in Algorithm 4 relies on the fact that $M(x)$, coming from the recurrence for $v_N^{(p,a,b)}(q)$, contains only pure powers of $x$ and $q$. We cannot hope for this phenomenon in general, however we advise to bear this simplification in mind for some practical purposes. In any case, we can improve on the quasi-linear cost in the degree $d$ of the polynomial matrix $M(x)$ in Algorithm 3, obtaining a complexity of essentially $\sqrt{d}$; in essence, the idea consists in choosing $s = \sqrt{N/d}$ rather than $\sqrt{N}$, see §5.3.4.

## 5.3 Main results

In this section, we generalize the algorithms from §5.2, and show that they apply to the general setting of $q$-holonomic sequences.

### 5.3.1 Preliminaries

A sequence $u_n = u_n(q)$ is $q$-holonomic if it satisfies a nontrivial $q$-recurrence, that is, a linear recurrence with coefficients given by polynomials in $q$ and $q^n$.

**Definition 5.1** ($q$-holonomic sequence)**.** Let $\mathbb{K}$ be a field, and $q \in \mathbb{K}$. A sequence $(u_n(q))_{n\geq 0}$ in $\mathbb{K}^{\mathbb{N}}$ is called *$q$-holonomic* if there exist $r \in \mathbb{N}$ and polynomials $c_0(x,y), \ldots, c_r(x,y)$ in $\mathbb{K}[x,y]$, with $c_r(x,y) \neq 0$, such that

$$c_r(q,q^n)u_{n+r}(q) + \cdots + c_0(q,q^n)u_n(q) = 0, \quad \text{for all } \ n \geq 0. \tag{5.6}$$

The integer $r$ is called the *order* of the $q$-recurrence (5.6). When $r = 1$, we say that $(u_n(q))_{n\geq 0}$ is $q$-hypergeometric.

Note that usually in combinatorics elements of $q$-holonomic sequences are considered as polynomials (or rational functions) in the variable $q$. In this text, however, we let $q$ be an element in the field, so that $u_n(q) \in \mathbb{K}$ as well.

The most basic examples are the *q-bracket* and the *q-factorial*,

$$[n]_q := 1 + q + \cdots + q^{n-1} \quad \text{and} \quad [n]_q! := \prod_{k=1}^{n} [k]_q. \tag{5.7}$$

They are clearly $q$-holonomic, and even $q$-hypergeometric.

The sequences $(u_n)_{n\geq 0} = (q^n)_{n\geq 0}$, $(v_n)_{n\geq 0} = (q^{n^2})_{n\geq 0}$ and $(w_n)_{n\geq 0} = (q^{\binom{n}{2}})_{n\geq 0}$ are also $q$-hypergeometric, since they satisfy the recurrence relations

$$u_{n+1} - qu_n = 0, \quad v_{n+1} - q^{2n+1}v_n = 0, \quad w_{n+1} - q^n w_n = 0.$$

However, the sequence $(q^{n^3})_{n\geq 0}$ is not $q$-holonomic [156, Ex. 2.2(b)]. More generally, this also holds for the sequence $(q^{n^s})_{n\geq 0}$, for any $s > 2$, see [42, Th. 4.1] and also [150, Th. 1.1].

Another basic example is the $q$-Pochhammer symbol

$$(x;q)_n := \prod_{k=0}^{n-1} (1 - xq^k), \tag{5.8}$$

which is also $q$-hypergeometric, since $(x;q)_{n+1} - (1 - xq^n)(x;q)_n = 0$. In particular, the sequence $(q;q)_n := \prod_{k=1}^{n}(1 - q^k)$, also denoted $(q)_n$, is $q$-hypergeometric and satisfies $(q)_{n+1} - (1 - q^{n+1})(q)_n = 0$. In Section §5.2 we encountered $v_n^{(p,a,b)} = \sum_{k=0}^{n} p^k q^{ak^2+bk}$, which is $q$-holonomic (see Eq. (5.4)), but generally not $q$-hypergeometric.

Note that (5.6) reduces to a C-linear recurrence, *i.e.* a linear recurrence with *constant* coefficients, if all polynomials $c_0(x,y), \ldots, c_r(x,y)$ are constant in the variable $y$. For these kinds of sequences there exist quasi-optimal algorithms [245, 136, 74], therefore we assume from now on that the maximal degree $d$ of $c_0(x,y), \ldots, c_r(x,y)$ in $y$ is positive.

As mentioned in the introduction, $q$-holonomic sequences show up in various contexts. As an example, in (quantum) knot theory, the ("colored") Jones function of a (framed oriented) knot (in 3-space) is a powerful knot invariant, related to the Alexander polynomial [29]; it is a $q$-holonomic sequence of Laurent polynomials [155]. Its recurrence equations are themselves of interest, as they are closely related to the A-polynomial of a knot, via the *AJ conjecture* [149, 151, 121], verified in some cases using massive computer algebra calculations [153].

It is well known that the class of $q$-holonomic sequences is closed under several operations, such as addition, multiplication, Hadamard product and monomial substitution [209, 195, 156]. All these closure properties are effective, *i.e.*, they can be executed algorithmically on the level of $q$-recurrences. Several computer algebra packages are available for the manipulation of $q$-holonomic sequences, *e.g.*, the Mathematica packages qGeneratingFunctions [195] and HolonomicFunctions [213], and the Maple packages qsum [47], qFPS [298], qseries and QDifferenceEquations.

A simple but useful fact is that the order-$r$ scalar $q$-recurrence (5.6) can be translated into a first-order recurrence on $r \times 1$ vectors:

$$\begin{bmatrix} u_{n+r} \\ \vdots \\ u_{n+1} \end{bmatrix} = \begin{bmatrix} -\frac{c_{r-1}}{c_r} & \cdots & -\frac{c_1}{c_r} & -\frac{c_0}{c_r} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \times \begin{bmatrix} u_{n+r-1} \\ \vdots \\ u_n \end{bmatrix}. \tag{5.9}$$

In particular, the $N$-th term of the $q$-holonomic sequence $(u_n)$ is simply expressible in terms of the *matrix $q$-factorial*

$$M(q^{N-1}) \cdots M(q)M(1), \tag{5.10}$$

where $M(q^n)$ denotes the companion matrix from equation (5.9). This observation is crucial, since it exposes the connection to the algorithms presented in the previous section.

### 5.3.2 Computation of the $q$-factorial

We now give the promised $q$-analogue of Strassen's result on the computation of $N!$ in $O(\mathbf{M}(\sqrt{N})\log N)$ arithmetic operations. Note that Strassen's case $q = 1$ is also covered by [69, §6], where the cost $O(\mathbf{M}(\sqrt{N}))$ is reached under some invertibility assumptions.

**Theorem 5.2.** *Let $\mathbb{K}$ be a field, let $q \in \mathbb{K} \setminus \{1\}$ and $N \in \mathbb{N}$. The $q$-factorial $[N]_q!$ can be computed using $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$. The same is true for the $q$-Pochhammer symbol $(\alpha; q)_N$ for any $\alpha \in \mathbb{K}$.*

*Proof.* If $\alpha = 0$, then $(\alpha; q)_N = 1$. If $q = 0$, then $[N]_q! = 1$ and $(\alpha; q)_N = 1 - \alpha$. We can assume that $q \in \mathbb{K} \setminus \{0, 1\}$ and $\alpha \in \mathbb{K} \setminus \{0\}$. We have $[N]_q! = r^N \cdot F(q^{-1})$ and $(\alpha; q)_N = \alpha^N \cdot F(\alpha^{-1})$, where $r := q/(1-q)$ and $F(x) := \prod_{i=0}^{N-1}(x - q^i)$. Algorithm 2 can be used to compute $F(q^{-1})$ and $F(\alpha^{-1})$ in $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$. The cost of computing $r^N$ and $\alpha^N$ is $O(\log N)$, and thus it is negligible. □

**Corollary 5.3.** *Under the assumptions of Theorem 5.2 and for any $n \in \mathbb{N}$, one can compute in $O(\mathbf{M}(\sqrt{n}) + \log(N))$ operations in $\mathbb{K}$:*

- *the $q$-binomial coefficient $\binom{N}{n}_q$;*
- *the coefficient of $x^n$ in the polynomial $\prod_{k=1}^{N}(1 + q^{k-1}x)$;*
- *the sum $\binom{N-n}{0}_q \binom{n}{0}_q + q\binom{N-n}{1}_q \binom{n}{1}_q + \cdots + q^{n^2}\binom{N-n}{n}_q \binom{n}{n}_q$.*

*Proof.* The first assertion is a direct consequence of Theorem 5.2 and of the equality

$$\binom{N}{n}_q = \frac{(q^N; q^{-1})_n}{(q; q)_n}.$$

The second assertion is a consequence of the first one, and of (5.1). The third assertion is a consequence of the first one, and of (5.2). □

### 5.3.3 $N$-th term of a $q$-holonomic sequence

We now offer the promised $q$-analogue of Chudnovskys' result on the computation of the $N$-th term of an arbitrary holonomic sequence in $O(\mathbf{M}(\sqrt{N})\log N)$ arithmetic operations. Note that Chudnovskys' case $q = 1$ is also covered by [69, §6], where the improved cost $O(\mathbf{M}(\sqrt{N}))$ is reached under additional invertibility assumptions.

**Theorem 5.4.** *Let $\mathbb{K}$ be a field, $q \in \mathbb{K} \setminus \{1\}$ and $N \in \mathbb{N}$. Let $(u_n(q))_{n\geq 0}$ be a $q$-holonomic sequence satisfying recurrence* (5.6)*, and assume that $c_r(q, q^k)$ is nonzero for $k = 0, \ldots, N-1$. Then, $u_N(q)$ can be computed in $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$.*

*Proof.* Using equation (5.9), it is enough to show that the matrix $q$-factorial

$$M(q^{N-1}) \cdots M(q)M(1)$$

can be computed in $O(\mathbf{M}(\sqrt{N}))$, where $M(q^n)$ denotes the companion matrix from equation (5.9). Algorithm 3 adapts *mutatis mutandis* to this effect. □

Remark that if $q$ is a root of unity of order $n < N$, then the computation of $U_N(q) = M(q^{N-1}) \cdots M(q)M(1)$ can be simplified using

$$U_N(q) = M(q^k) \cdots M(1) \cdot U_n(q)^r,$$

where $r = \lfloor (N-1)/n \rfloor$ and $k = N - 1 - rn$. Algorithm 3 is used to compute $U_n(q)$ and then its $r$-th power is deduced via binary powering. Finally, the product $M(q^k) \cdots M(1)$ is again computed using Algorithm 3. The total cost therefore consists of just $O(\mathbf{M}(\sqrt{n}) + \log(N))$ arithmetic operations. It follows that if, for instance, the base field $\mathbb{K}$ is the prime field $\mathbb{F}_p$, then the prime number $p$ should be larger than $N$ in order to exhibit the full strength of the presented algorithms.

**Corollary 5.5.** *Let $\mathbb{K}$ be a field, $q \in \mathbb{K}$ not a root of unity, and $N \in \mathbb{N}$. Let $e_q(x)$ be the $q$-exponential series*

$$e_q(x) \coloneqq \sum_{n\geq 0} \frac{x^n}{[n]_q!},$$

*and let $E_q^{(N)}(x) \coloneqq e_q(x) \bmod x^N$ be its polynomial truncation of degree $N-1$. If $\alpha \in \mathbb{K}$, then one can compute $E_q^{(N)}(\alpha)$ in $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$.*

*Proof.* Denote the summand $\frac{\alpha^n}{[n]_q!}$ by $u_n(q)$. Then $(u_n(q))_{n\geq 0}$ is $q$-hypergeometric, and satisfies the recurrence $[n+1]_q u_{n+1}(q) - \alpha u_n = 0$, therefore $v_N(q) \coloneqq \sum_{i=0}^{N-1} u_i(q)$ satisfies the second-order recurrence $[n+1]_q(v_{n+2}(q) - v_{n+1}(q)) - \alpha(v_{n+1}(q) - v_n(q)) = 0$. Applying Theorem 5.4 to $v_N(q)$ concludes the proof. □

The same result holds true if $e_q(x)$ is replaced by any power series satisfying a $q$-difference equation. For instance, one can evaluate fast all truncations of Heine's $q$- hypergeometric series

$${}_2\phi_1([a,b],[c];q;x) \coloneqq \sum_{n\geq 0} \frac{(a;q)_n(b;q)_n}{(c;q)_n} \cdot \frac{x^n}{(q)_n}.$$

### 5.3.4 Complexity analysis and computation of several terms

Theorem 5.4 established an $O(\mathbf{M}(\sqrt{N}))$ cost of the presented method for computing the $N$-th term of a $q$-holonomic sequence. We now aim at performing a detailed complexity analysis with respect to all input parameters. So we need to discuss the complexity of

Algorithm 3, where we assume that $M(x) \in \mathcal{M}_n(\mathbb{K}[x]_d)$ is an $n \times n$ polynomial matrix of degree $d \geq 1$. We wish to examine the amount of field operations in $\mathbb{K}$ needed for the computation of $M(q^{N-1}) \cdots M(1)$ in terms of $N, d$ and $n$. Recall that $\mathbf{MM}(n,d)$ controls the arithmetic complexity of the product in $\mathcal{M}_n(\mathbb{K}[x]_d)$ and it holds that $\mathbf{MM}(n,d) = O(n^\theta d + n^2\,\mathbf{M}(d)) = \tilde{O}(n^\theta d)$.

First, we will examine the direct application of Algorithm 3, where $s = \sqrt{N}$, to $M(x)$. As it turns out that the dominating part is step (1), where we compute the polynomial $P_s(x) = M(q^{s-1}x) \cdots M(x)$ using the divide-and-conquer scheme $P_{2t}(x) = P_t(q^t x) \cdot P_t(x)$ and $P_{2t+1}(x) = M(q^{2t}x) \cdot P_t(q^t x) \cdot P_t(x)$. Note that $P_t(x)$ is an $n \times n$ polynomial matrix of degree at most $td$ and therefore the cost of this step is $O(\mathbf{MM}(n, sd)) = \tilde{O}(n^\theta d\sqrt{N})$. Step (2) is done component-wisely at each entry of $P(x)$. By the explained fast chirp transform algorithms, it essentially boils down to $n^2$ multiplications of two polynomials, one of degree $sd$ and the other of degree $2sd$. The cost of the second step is therefore $O(n^2\,\mathbf{M}(sd)) = \tilde{O}(n^2 d\sqrt{N})$. The last step is the multiplication of $N/s = s$ matrices with entries in $\mathbb{K}$ and has therefore an arithmetic complexity of $O(n^\theta s) = \tilde{O}(n^\theta\sqrt{N})$.

If $d < N$ is a parameter of interest, then there is a better choice of $s$ rather than $\sqrt{N}$. We saw that the polynomial $P_s(x)$ has degree $sd$ and we must evaluate it at $N/s$ points. The optimal pick for $s$ is therefore $s = \sqrt{N/d}$, which we again can assume to be integer[4]. Then, by the same arguments as above, the costs of the three steps are $O(\mathbf{MM}(n, \sqrt{Nd})) = \tilde{O}(n^\theta\sqrt{Nd})$, $O(n^2\mathbf{M}(\sqrt{Nd})) = \tilde{O}(n^2\sqrt{Nd})$ and $O(n^\theta\sqrt{Nd})$ respectively.

Now, we address specifically the computation of the $N$-th term in a $q$-holonomic sequence. If $(u_n(q))_{n\geq 0}$ is given by a $q$-recurrence

$$c_r(q, q^n)u_{n+r}(q) + \cdots + c_0(q, q^n)u_n(q) = 0,$$

for $q \in \mathbb{K}$ and for polynomials $c_j(x,y) \in \mathbb{K}[x,y]$, then as observed before, we can compute $u_N(q)$ via

$$\frac{1}{c_r(q,q^{N-1})\cdots c_r(q,q)c_r(q,1)} \cdot \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \times \tilde{M}(q^{N-1}) \cdots \tilde{M}(q)\tilde{M}(1) \times \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix},$$

where now

$$\tilde{M}(x) \coloneqq c_r(q,x) \cdot M(x) = \begin{bmatrix} -c_{r-1}(q,x) & \cdots & -c_1(q,x) & -c_0(q,x) \\ c_r(q,x) & \cdots & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \cdots & c_r(q,x) & 0 \end{bmatrix}.$$

Hence, we are interested in $\tilde{M}(q^{N-1}) \cdots \tilde{M}(1)$ and $c_r(q,q^{N-1}) \cdots c_r(q,1)$. If the degrees of $c_0(q,y), \ldots, c_r(q,y)$ are bounded by $d$, then the considerations above imply that the two $q$-factorials can be computed in $O(\mathbf{MM}(r, \sqrt{Nd}) + r^2\mathbf{M}(\sqrt{Nd}))$ and $O(\mathbf{M}(\sqrt{Nd}))$ operations in $\mathbb{K}$, respectively. We obtain the following theorem (compare with [67, Thm. 2]).

[4]Similarly as before, if $\sqrt{N/d}$ is not an integer, then we can compute $u_{N_1}(q)$ first, where $N_1 = \lfloor\sqrt{N/d}\rfloor^2 d$, and then proceed "naively". Note that $N - N_1 < 2\sqrt{Nd} - d = O(\sqrt{Nd})$.

**Theorem 5.6.** *Under the assumptions of Theorem 5.4, let $d \geq 1$ be the maximum of the degrees of $c_0(q,y),\dots,c_r(q,y)$. Then, for any $N > d$, the term $u_N(q)$ can be computed in $O(r^\theta\sqrt{Nd} + r^2\,\mathbf{M}(\sqrt{Nd}))$ operations in $\mathbb{K}$.*

Theorem 5.4 can be adapted to the computation of *several coefficients* of a $q$-holonomic sequence. The proof is similar to that of Theorem 15 in [69], however simpler, because we deal with geometric progressions instead of arithmetic ones.

**Theorem 5.7.** *Under the assumptions of Theorem 5.6, let $N_1 < N_2 < \cdots < N_n = N$ be positive integers, where $n \leq \sqrt{N}$. Then, the terms $u_{N_1}(q),\dots,u_{N_n}(q)$ can be computed altogether in $O(\mathbf{M}(\sqrt{N})\log N)$ operations in $\mathbb{K}$.*

*Proof.* As before, we assume that $N$ is a perfect square; let $s = \sqrt{N}$. Examining the presented algorithms, we notice that on the way of computing the matrix $q$-factorial $U_N := M(q^{N-1})\cdots M(q)M(1)$ we obtain the evaluated polynomials

$$P(1), P(Q), \dots, P(Q^{s-1}),$$

where $Q := q^s$ and $P(x) := M(q^{s-1}x)\cdots M(qx)M(x)$. The $q$-factorial $U_N$ is then found by step (3) by trivially multiplying $P(Q^{s-1})\cdots P(Q)P(1)$. Observe that while multiplying together from right to left we actually also automatically compute

$$P(Q^{j-1})\cdots P(Q)P(1) = M(q^{sj-1})\cdots M(q)M(1) = U_{sj},$$

for every $j = 1,\dots,s$. It follows that employing Algorithm 3 and by simply taking the top right element of each $U_{sj}$, we find not only $u_N$, but actually $u_s, u_{2s}, \dots, u_{s^2} = u_N$. This already indicates that simultaneous computation of $s$ terms is achievable in similar complexity after some "distillation". In general, we are interested in the sequence of $u_i(q)$ at indices $i = N_1,\dots,N_n$, hence we need to perform the following refinement step.

Let $d_0 \in \mathbb{N}$ be a positive integer with $d_0 \leq n \leq \sqrt{N}$ and assume that for some $k_1^{(0)},\dots,k_n^{(0)}$ with $k_j^{(0)} \leq N_j < k_j^{(0)} + 2d_0$ we already know the values $U_{k_1^{(0)}},\dots,U_{k_n^{(0)}}$. Then we can use a similar strategy as in step (1) of Algorithm 3 and deduce the polynomial matrix $P_{d_0}(x) = M(q^{d_0-1}x)\cdots M(qx)M(x)$. Compute then the values $q^{k_1^{(0)}},\dots,q^{k_n^{(0)}}$ and evaluate $P_{d_0}(x)$ simultaneously at them. For each $j = 1,\dots,n$ it holds that

$$P_{d_0}(q^{k_j^{(0)}})\cdot U_{k_j^{(0)}} = U_{k_j^{(0)}+d_0}.$$

We perform this multiplication for those indices $j$ for which $k_j^{(0)} + d_0 \leq N_j < k_j^{(0)} + 2d_0$. For these $j$ we then set $k_j^{(1)} = k_j^{(0)} + d_0$ and let $k_j^{(1)} = k_j^{(0)}$ for the other indices; moreover $d_1 := \lceil d_0/2\rceil$. We iterate this process at most $\ell := \lceil\log(d_0)\rceil$ many times until $d_\ell = 1$. Then we can easily find $U_{N_1},\dots,U_{N_n}$, from which we finally deduce $u_{N_1},\dots,u_{N_n}$.

Each such step has a cost of at most $O(\mathbf{M}(n)) = O(\mathbf{M}(\sqrt{N}))$ base field operations. Moreover, after first employing Algorithm 3 and by the consideration above, we compute $U_1, U_s, \dots, U_{s^2}$ in $O(\mathbf{M}(\sqrt{N}))$ base operations (Theorem 5.6). Hence, we may choose $d_0 = s$ and for each $j = 1,\dots,n$ let $k_j^{(0)}$ be the largest element in $\{1, s, \dots, (s-1)s\}$ such that $k_j^{(0)} \leq N_j$. Clearly, all conditions of the above refinement step are satisfied and we need at most $\lceil\log(s)\rceil = O(\log N)$ many such steps. The total complexity is henceforth $O(\mathbf{M}(\sqrt{N})) + O(\mathbf{M}(\sqrt{N})\log N) = O(\mathbf{M}(\sqrt{N})\log N)$. □

The same idea applies in the following corollary which states that if $n < \sqrt{N}/N^{\varepsilon}$ for some $\varepsilon > 0$, then we can omit the $\log$-factor in $N$:

**Corollary 5.8.** *Under the assumptions of Theorem 5.6, let $N_1 < N_2 < \cdots < N_n = N$ be positive integers, where $n < N^{\frac{1}{2}-\varepsilon}$ for some $0 < \varepsilon < \frac{1}{2}$. Then, the terms $u_{N_1}(q), \ldots, u_{N_n}(q)$ can be computed altogether in $O(\mathbf{M}(\sqrt{N}))$ operations in $\mathbb{K}$.*

*Proof.* Here, we follow the exact same procedure as in the proof before. Then, regarding complexity, we use $O(\mathbf{M}(n)) = O(\mathbf{M}(N^{\frac{1}{2}-\varepsilon})) = O(\mathbf{M}(\sqrt{N})N^{-\varepsilon})$ and obtain that the same method yields a total arithmetic cost of $O(\mathbf{M}(\sqrt{N})) + O(\mathbf{M}(\sqrt{N})N^{-\varepsilon} \log N) = O(\mathbf{M}(\sqrt{N}))$. □

Remark that regarding a detailed complexity analysis for the computation of several coefficients, we have the following trade-off: either we compute at most $\sqrt{N}$ terms in the arithmetic complexity $O((r^{\theta}d\sqrt{N} + r^2\mathbf{M}(d\sqrt{N})) \log N)$, or at most $\sqrt{N/d}$ terms, but in a better cost of $O((r^{\theta}\sqrt{Nd} + r^2\mathbf{M}(\sqrt{Nd})) \log N)$. The proofs combine the considerations above with setting $s = \sqrt{N}$ and $s = \sqrt{N/d}$ respectively. In both cases we can get rid of the log-factor in $N$ like in Corollary 5.8 by computing a factor of $N^{\varepsilon}$ less terms.

### 5.3.5 The case $q$ is an integer: bit complexity

Until now, we only considered the arithmetic complexity model, which is very well-suited to measure the algorithmic cost when working in algebraic structures whose basic internal operations have constant cost (such as finite fields, or floating point numbers).

Now we discuss here the case where $q$ is an integer (or rational) number. The arithmetic complexity model needs to be replaced by the bit-complexity model.

Recall that the most basic operations on integer numbers can be performed in quasi-optimal time, that is, in a number of bit operations which is almost linear, up to logarithmic factors, in (the maximum of) their bit size. The most basic operation is integer multiplication, for which quasi-linear time algorithms are known since the early seventies, starting with the famous paper by Schönhage and Strassen [286] who showed that two $n$-bit integers can be multiplied in $O(n \log n \log\log n)$ bit operations. After several successive improvements, e.g., [143, 171], we know as of 2020 that two $n$-bit integers can be multiplied in time $O(n \log n)$ [169]. We shall call the cost of multiplying two $n$-bit integers $\mathbf{M}_{\mathbb{Z}}(n)$.

In this context, the matrix $q$-factorials from §5.3.1 are computed by *binary splitting* rather than by baby-step/giant-steps. Recall that this phenomenon already occurs in the usual holonomic setting. For example, the bitsize of $u_N = N!$ is $O(N \log N)$, however both, the "naive" method of computing it using $u_n = nu_{n-1}$, or Strassen's baby-steps/giant-steps method, yield worse bit complexity. In the naive approach the problem is that integers of unbalanced bitsize are multiplied together and hence not the full power of fast integer multiplication techniques can be employed. A more clever and very simple way is to just use the fact that

$$N! = (1 \cdots \lfloor N/2 \rfloor) \times ((\lfloor N/2 \rfloor + 1) \cdots N),$$

and that the bitsizes of both factors have magnitude $O(N/2 \log(N)) = \tilde{O}(N)$. Thus, fast integer multiplication can be used to multiply them in $\tilde{O}(N)$ bit-complexity. This idea results in the binary splitting Algorithm 5. This algorithm is very classical, and we only recall

it for completeness. See [35, §12] for a good survey on this technique, and its applications.

**Algorithm 5** (BinSplit)
**Input**: $A = [a_1, \ldots, a_N]$ list of elements from some arbitrary ring $R$
**Output**: $a_N \cdots a_1$

```
1: function F(A)
2:     if N = 1 then
3:         return A[1]
4:     end if
5:     return F(A[⌊N/2⌋ + 1, ..., N]) · F(A[1, ..., ⌊N/2⌋])
6: end function
```

Assume that each $a_i$ in the input of BinSplit has at most $k$ bits and let $C(n)$ be the complexity of BinSplit if $A = [a_1, \ldots, a_n]$ is $n$-dimensional. It follows that

$$C(N) \leq 2C(\lceil N/2 \rceil) + \mathbf{M}_{\mathbb{Z}}(N/2 \cdot k),$$

where $\mathbf{M}_{\mathbb{Z}}(n) = \tilde{O}(n)$ is the cost of multiplication of integers with $n$ bits. We obtain $C(N) = \tilde{O}(Nk)$. Hence, using this method, the computation of $u_N = N!$ has $\tilde{O}(N)$ bit-complexity, which is quasi-optimal. Moreover, the same idea applies to any holonomic sequence, by deducing the first order matrix recurrence and computing the matrix product using BinSplit.

Now we shall see that the $q$-holonomic case is similar. First, let $q$ be a positive integer of $B$ bits and consider the computation of the $q$-factorial

$$u_N(q) = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{N-1}),$$

as an illustrative example. For each factor, assuming $q > 0$, we have the trivial inequalities

$$q^n < 1 + q + \cdots + q^n < q^{n+1},$$

meaning that $q^{N(N-1)/2} < u_N(q) < q^{N(N+1)/2}$, so the bitsize of $u_N(q)$ is of magnitude $N^2B$. The "naive" algorithm of deducing $u_N(q)$ by first computing the integers $q^i$, then the corresponding sums and products, has $\tilde{O}(N^3B)$ binary complexity. This method is not (quasi-)optimal with respect to the output size. It is also easy to see that the presented baby-steps/giant-steps based algorithms yield bad bit-complexity as well, despite their good arithmetic cost.

Similarly, if

$$u_N(q) = \sum_{n=0}^{N-1} q^{n^2},$$

then the integer $u_N(q)$ is bounded in absolute value from above by $Nq^{(N-1)^2}$ and by $q^{(N-1)^2}$ from below, so its bitsize is again of magnitude $N^2B$. The "naive" algorithm consisting of computing the terms $q^i$ one after the other before summing, has again non-optimal bit-complexity $\tilde{O}(N^3B)$.

Can one do better? The answer is "yes" and one can even achieve a complexity which is quasi-linear in the bitsize of the output. Similarly to the holonomic setting, it is sufficient to use the $q$-holonomic character of $u_N(q)$, and to reduce its computation to that of a $q$-factorial

matrix as in §5.2.2, which can then be handled with BinSplit. To be more precise, given an integer or rational number $q$ of bitsize $B$ and any $q$-holonomic sequence $(u_n(q))_{n\geq 0}$ defined by polynomials $c_0, \ldots, c_r \in \mathbb{Z}[x, y]$ with $c_r(q, q^n) \neq 0$ for any $n \in \mathbb{N}$, we define $M(x) \in \mathbb{Q}(x)$ by

$$\begin{bmatrix} -\frac{c_{r-1}(q,x)}{c_r(q,x)} & \cdots & -\frac{c_1(q,x)}{c_r(q,x)} & -\frac{c_0(q,x)}{c_r(q,x)} \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}.$$

Then, as observed before, $u_N$ can be read off from

$$M(q^{N-1}) \cdots M(q) M(1),$$

which we aim to compute efficiently. Again, instead of using baby-steps/giant-steps, it is a better idea to use binary splitting by applying Algorithm 5 to $A = [M(1), \ldots, M(q^{N-1})]$. Note that obviously, any element in $A$ is a matrix with rational entries of bitsize bounded by $O(NB)$. Therefore, the complexity of BinSplit does not exceed $\tilde{O}(N^2 B)$ by the same argument as before, now using fast multiplication of rational numbers. These considerations prove

**Theorem 5.9.** *Under the assumptions of Theorem 5.4, with $\mathbb{K} = \mathbb{Q}$, the term $u_N(q)$ can be computed in $\tilde{O}(N^2 B)$ bit operations, where $B$ is the bitsize of $q$.*

As a corollary, (truncated) solutions of $q$-difference equations can be evaluated using the same (quasi-linear) bit-complexity. This result should be viewed as the $q$-analogue of the classical fact that holonomic functions can be evaluated fast using binary splitting, a 1988 result by the Chudnovsky brothers [106, §6], anticipated a decade earlier (without proof) by Schroeppel and Salamin in Item 178 of [32].

## 5.4 Applications

### 5.4.1 Combinatorial $q$-holonomic sequences

As already mentioned, many $q$-holonomic sequences arise in combinatorics, for example in connection with the enumeration of lattice polygons, where $q$-analogues of the Catalan numbers $\frac{1}{n+1}\binom{2n}{n}$ occur naturally [161, 144], or in the enumeration of special families of matrices with coefficients in the finite field $\mathbb{F}_q$ [205, 206, 333], where sequences related to the Gaussian coefficients $\binom{n}{k}_q$ also show up.

A huge subfield of combinatorics is the theory of partitions [21], where $q$-holonomic sequences occur as early as in the famous Rogers-Ramanujan identities [275, 274], see also [21, Ch. 7], *e.g.*,

$$1 + \sum_{n\geq 1} \frac{q^{n^2}}{(1-q)\cdots(1-q^n)} = \prod_{n\geq 0} \frac{1}{(1-q^{5n+1})(1-q^{5n+4})}$$

which translates the fact that the number of partitions of $n$ into parts that differ by at least $2$ is equal to the number of partitions of $n$ into parts congruent to $1$ or $4$ modulo $5$.

Andrews [19, 20], see also [21, Chapter 8], laid the foundations of a theory able to capture the $q$-holonomy of any generating function of a so-called *linked partition ideal*.

As a consequence, a virtually infinite number of special families of polynomials coming from partitions can be evaluated fast. For instance, the family of truncated polynomials

$$F_n(x) := \prod_{k=1}^{\infty}(1-x^k)^3 \bmod x^n,$$

can be evaluated fast due to our results and to the identity [254, §6]

$$F_N(q) = \sum_{\binom{n+1}{2}<N} (-1)^n(2n+1)q^{\binom{n+1}{2}}.$$

### 5.4.2 Evaluation of $q$-orthogonal polynomials

In the theory of special functions, *orthogonal polynomials* play a fundamental role. There exists an extension to the $q$-framework of the theory, see, for example, Chapter 9 in Ernst's book [132]. Amongst the most basic examples, the *discrete $q$-Hermite polynomials* [12, 27] are defined by their $q$-exponential generating function

$$\sum_{n\geq 0} F_{n,q}(x)\frac{t^n}{[n]_q!} = \frac{e_q(xt)}{e_q(t)e_q(-t)},$$

and therefore they satisfy the second-order linear $q$-recurrence

$$F_{n+1,q}(x) = xF_{n,q}(x) - (1-q^n)q^{n-1}F_{n-1,q}(x), \quad n \geq 1,$$

with initial conditions $F_{0,q}(x) = 1, F_{1,q}(x) = x$. From there, it follows that for any $\alpha \in \mathbb{K}$, the sequence $(F_{n,q}(\alpha))_{n\geq 0}$ is $q$-holonomic, thus the evaluation of the $N$-th polynomial at $x = \alpha$ can be computed fast. The same is true for the *continuous $q$-Hermite polynomials*, for which $2\alpha H_{n,q}(\alpha) = H_{n+1,q}(\alpha) + (1-q^n)H_{n-1,q}(\alpha)$ for $n \geq 1$, and $H_{0,q}(\alpha) = 1, H_{1,q}(\alpha) = 2\alpha$. More generally, our results in §5.3 imply that any family of $q$-orthogonal polynomials can be evaluated fast.

### 5.4.3 Polynomial and rational solutions of $q$-difference equations

The computation of polynomial and rational solutions of linear differential equations lies at the heart of several important algorithms, for computing hypergeometric, d'Alembertian and Liouvillian solutions, for factoring and for computing differential Galois groups [320, 6, 5]. Creative telescoping algorithms (of second generation) for multiple integration with parameters [111, 213] also rely on computing rational solutions, or deciding their existence. The situation is completely similar for $q$-difference equations, *i.e.* equations of the form

$$Ly = a_\nu(x)y(q^\nu x) + a_{\nu-1}(x)y(q^{\nu-1}x) + \cdots + a_0(x)y(x) = 0, \tag{5.11}$$

with $a_j(x) \in \mathbb{Q}[x]$, for all $j = 0, \ldots, \nu$, such that $a_0(x)a_\nu(x)$ is not identically zero. Improving algorithms for polynomial and rational solutions of such equations is important for finding $q$-hypergeometric solutions [5], for computing $q$-difference Galois groups [178, 26], and for performing $q$-creative telescoping [212, 111, 213].

In both differential and $q$-differential cases, algorithms for computing polynomial solutions proceed in two distinct phases: (i) compute a degree bound $N$, potentially exponentially large in the equation size; (ii) reduce the problem of computing polynomial solutions of degree at most $N$ to linear algebra. Abramov, Bronstein and Petkovšek showed in [2] that, in step (ii), linear algebra in size $N$ can be replaced by solving a much smaller system, of polynomial size. However, setting up this smaller system still requires linear time in $N$, essentially by unrolling a ($q$-)linear recurrence up to terms of indices close to $N$. For differential (and difference) equations, this step has been improved in [67, 62], by using Chudnovskys' algorithms for computing fast the $N$-th term of a holonomic sequence. This allows for instance to decide (non-)existence of polynomial solutions in sublinear time $\tilde{O}(\sqrt{N})$. Moreover, when polynomial solutions exist, one can represent/manipulate them in *compact form* using the recurrence and initial terms as a compact data structure. Similar ideas allow to also compute rational solutions in compact form in the same complexity, see [63, Chap. 17].

The same improvements can be transferred to linear $q$-difference equations, in order to improve the existing algorithms [3, 2, 204]. In this case, setting up the smaller system in phase (ii) amounts to computing the $N$-th term of a $q$-holonomic sequence, and this can be done fast using our results in §5.3. A technical subtlety is that, as pointed out in [2, §4.3], it is not obvious in the $q$-difference case how to guarantee the non-singularity of the $q$-recurrence on the coefficients of the solution. This induces potential technical complications similar to the ones for polynomial solutions of differential equations in small characteristic, which can nevertheless be overcome by adapting the approach described in [78, §3.2]. Similar improvements can be also transferred to systems [4, 31].

Let us finish this discussion by pointing out briefly an application of these improvements. Desingularizing a linear differential operator $L(x, \partial_x)$ consists in computing a left multiple with all apparent singularities removed. It is a central task for determining the Weyl closure of $L$ [318]. The computation of polynomial solutions of Fourier dual operators is a basic step for performing desingularization [113]. By duality, the order $N$ of the desingularization corresponds to the degree of polynomial solutions of the dual $L^\star$ of $L$. This remark in conjunction with the fast algorithms for polynomial solutions [67], themselves based on the fast computation of matrix factorials, allows to speed up the computation of desingularizations. The situation is similar in other Ore algebras, and in particular for $q$-difference equations. Therefore, our algorithmic improvements in the computation of polynomial solutions of $q$-difference equations, themselves based on the fast computation of matrix $q$-factorials, have a direct impact on the acceleration of the desingularizations process for $q$-difference equations. It is less obvious to us whether other desingularization algorithms, such as the one from [215], could also benefit from these remarks.

### 5.4.4 Computing curvatures of $q$-difference equations

A natural application of the fast computation of matrix $q$-factorials is the computation of curvatures of $q$-difference equations, since in this area these objects appear quite inherently. Another strong motivation comes from the fact that the $q$-analogue [41] of Grothendieck's conjecture (relating solutions of equations over $\mathbb{Q}$ with their reductions modulo primes $p$, see Conjecture 3.3 in Chapter 3) is proved [123, 124], while the classical differential case is widely open [193]. Still, in the latter setting algorithms have been developed allowing to compute $p$-curvatures fast [78, 56, 57, 58]. They allow, for example, to perform a quick heuristic, but reliable in practice, test for the existence of a basis of algebraic solutions of a linear differential operator [70].

The $q$-analogue of Grothendieck's conjecture investigates rational solutions of $q$-difference equations. Similarly to the differential setting, one naturally associates to the equation (5.11) the linear $q$-difference system $Y(qx) = A(x)Y(x)$, where

$$A(x) = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & 1 \\ -a_0/a_\nu & -a_1/a_\nu & \cdots & -a_{\nu-1}/a_\nu \end{bmatrix}.$$

Then it easily follows that if $z(x)$ solves (5.11), then $(z(x), z(qx), \ldots, , z(q^{\nu-1}x))^t$ is a solution of $Y(qx) = A(x)Y(x)$. Moreover, these equations are in some sense equivalent through the $q$-analogues of the Wronskian Lemma and the Cyclic Vector Lemma whenever $q$ is not a root of unity of order smaller than $\nu$; see [124] for details. If $q$ is considered as a variable, then Di Vizio and Hardouin proved that (5.11) admits a full set of solutions in $\mathbb{Q}(q, x)$ if and only if for almost all natural numbers $n$,

$$C_n(x) \coloneqq A(q^{n-1}x) \cdots A(qx)A(x) \equiv \mathrm{Id}_\nu \mod \mathrm{GL}_\nu(R_n(x)),$$

where $R_n = \mathbb{Q}[q]/\Phi_n(q)$, with $\Phi_n(x)$ the $n$-th cyclotomic polynomial. The elements in the sequence $(C_n(x))_{n\geq 1}$ are known as curvatures of the $q$-difference system and are clearly just matrix $q$-factorials.

On the other hand, if $q \in \mathbb{Q}$ then it already follows from main result of [123] that (5.11) admits a basis of rational solutions in $\mathbb{Q}(x)$ if and only if for almost all primes $p$,

$$A(q^{\kappa_p-1}x) \cdots A(qx)A(x) \equiv \mathrm{Id}_\nu \mod p^\ell,$$

where $\kappa_p = \mathrm{ord}_p(q)$ and $\ell_p \in \mathbb{Z}$ such that $1 - q^{\kappa_p} = p^{\ell_p}\frac{h}{g}$, with $h, g \in \mathbb{Z}$ coprime to $p$.

On these types of questions there is more progress in the $q$-difference setting than in the classical differential one. Yet, unfortunately, the theorems above are not proven to be effective in the sense that still infinitely many conditions need to be checked in order to conclude the implication we are mostly interested in. Therefore, the computation of any finite number of curvatures only provides a heuristic for the existence of rational solutions of a $q$-difference equation. Moreover, the mentioned algorithms in §5.4.3 compute rational solutions of equations of type (5.11) and therefore allow to decide rigorously about the existence of such a basis. However, all these methods have a cost which is potentially

exponential in the size of the input. Our goal in this section is to design a fast heuristic test for the existence of a basis of rational solutions of a $q$-difference equation using curvatures.

If $q$ is a variable, we want to check whether $C_n(x) \equiv \mathrm{Id}_\nu \bmod \mathrm{GL}_\nu(R_n(x))$ for many $n$. Clearly, after the reduction mod $\Phi_n(q)$, the polynomial $C_n(x)$ has arithmetic size $n^2$ over $\mathbb{Q}$ and $\tilde{O}(n^3)$ bitsize. Hence, computing $C_n(x)$ is unnecessarily costly. We propose to work over $R_{n,p} := \mathbb{F}_p[q]/\Phi_n(q)$ for some (large) prime $p$; moreover, we compute $C_n(x_0)$ for some randomly chosen $x_0 \in \mathbb{F}_p$. Furthermore, in order to avoid computing general cyclotomic polynomials, we compute $C_n(x)$ only for prime numbers $n$. After these considerations, it is easy to see that Algorithm 3 applies and allows to deduce $C_n(x_0)$ modulo $\mathrm{GL}_\nu(R_{n,p})$ in $\tilde{O}(\sqrt{n})$ arithmetic operations in $R_{n,p}$, hence $\tilde{O}(n^{3/2})$ operations in $\mathbb{F}_p$. Finally, if we want to deduce this quantity for all primes $n$ between 2 and some $N \in \mathbb{N}$, it is wiser to apply the accumulating remainder tree method presented in [116, 167], which allows for quasi-optimal complexity of $\tilde{O}(N^2)$ in this case.

If $q$ is some rational number and the goal is to test whether $Y(qx) = A(x)Y(x)$ has a full set of rational solutions, one may check

$$C_{\kappa_p}(x) = A(q^{\kappa_p - 1}x) \cdots A(qx)A(x) \equiv \mathrm{Id}_\nu \mod p^\ell,$$

for many primes $p$. Unfortunately, finding the order $\kappa_p$ in practice may be costly, therefore we shall check the weaker assumption $C_{p-1}(x) \equiv \mathrm{Id}_\nu \bmod p$. Conjecturally, this equality for sufficiently many primes $p$ is also enough to conclude on the existence of a basis of rational solutions. The presented Algorithm 3 allows to compute $C_{p-1}(x_0) \bmod p$ for some $x_0 \in \mathbb{F}_p$ in $\tilde{O}(\sqrt{p})$ arithmetic cost. Finally, again, if we choose $N, x_0 \in \mathbb{N}$, then the accumulating remainder tree allows to deduce $C_{p-1}(x_0) \bmod p$ for all primes $p$ between 2 and $N$ quasi-optimally in $\tilde{O}(N)$ bit operations.

### 5.4.5 $q$-hypergeometric creative telescoping

In the case of differential and difference hypergeometric creative telescoping, it was demonstrated in [62] that the compact representation for polynomial solutions can be used as an efficient data structure, and can be applied to speed up the computation of Gosper forms and Zeilberger's classical summation algorithm [263, §6]. The key to these improvements lies in the fast computation of the $N$-th term of a holonomic sequence, together with the close relation between Gosper's algorithm and the algorithms for rational solutions.

Similarly, in the $q$-difference case, Koornwinder's $q$-Gosper algorithm [212, §5] is closely connected to Abramov's algorithm for computing rational solutions [3, §2], and this makes it possible to transfer the improvements for rational solutions to the $q$-Gosper algorithm. This leads in turn to improvements upon Koornwinder's algorithm for $q$-hypergeometric summation [212], along the same lines as in the differential and difference cases [62].

## 5.5 Experiments

Algorithms 1 and 2 were implemented in Magma and Algorithm 3 in Maple. All implementations deliver some encouraging timings. Of course, since these algorithms are designed

| degree $N$ | Naive algorithm | Algorithm 1 | Algorithm 2 |
|---|---|---|---|
| $2^{16}$ | 0.04 | 0.03 | 0.00 |
| $2^{18}$ | 0.18 | 0.03 | 0.01 |
| $2^{20}$ | 0.72 | 0.06 | 0.01 |
| $2^{22}$ | 2.97 | 0.14 | 0.02 |
| $2^{24}$ | 11.79 | 0.32 | 0.04 |
| $2^{26}$ | 47.16 | 0.73 | 0.08 |
| $2^{28}$ | 188.56 | 1.68 | 0.15 |
| $2^{30}$ | 755.65 | 3.84 | 0.31 |
| $2^{32}$ | 3028.25 | 8.65 | 0.64 |
| $2^{34}$ | | 19.65 | 1.41 |
| $2^{36}$ | | 44.42 | 2.96 |
| $2^{38}$ | | 101.27 | 6.36 |
| $2^{40}$ | | 228.58 | 14.99 |
| $2^{42}$ | | 515.03 | 29.76 |
| $2^{44}$ | | 1168.51 | 61.69 |
| $2^{46}$ | | 2550.28 | 137.30 |
| $2^{48}$ | | | 297.60 |
| $2^{50}$ | | | 731.63 |
| $2^{52}$ | | | 1395.33 |
| $2^{54}$ | | | 3355.39 |

Table 5.1 Comparative timings (in seconds) for the computation of $\prod_{i=0}^{N-1}(\alpha - q^i) \in \mathbb{F}_p$, with $p = 2^{30} + 3$ and $(\alpha, q)$ randomly chosen in $\mathbb{F}_p \times \mathbb{F}_p$. All algorithms were executed on the same machine, running Magma v. 2.24. For each target degree $N$, each execution was limited to 1 hour. Naive algorithm could reach degree $N = 2^{32}$, Algorithm 1 degree $N = 2^{46}$, and Algorithm 2 degree $N = 2^{54} = 8\,014\,398\,509\,481\,984$. By extrapolation, the Naive algorithm would have needed $\approx 4^{11} \times 3028.25$ sec. $\approx 400$ years on the same instance, and Algorithm 2 approximately 18 hours.

to be fast in the *arithmetic model*, it is natural to make experiments over a finite field $\mathbb{K}$, or over truncations of real/complex numbers, as was done in [249] for the problem in §5.2.2.

Recall that both Algorithms 1 and 2 compute $\prod_{i=0}^{N-1}(\alpha - q^i) \in \mathbb{K}$, given $\alpha, q$ in a field $\mathbb{K}$, and $N \in \mathbb{N}$, whereas Algorithm 3 finds $M(q^{N-1}) \cdots M(q)M(1) \in \mathbb{K}^{n\times n}$ for a given polynomial matrix $M(x) \in \mathbb{K}[x]$ of size $n \times n$ and $q \in \mathbb{K}, N \in \mathbb{N}$. In our experiments for Algorithms 1 and 2 (Table 5.1) we choose $\mathbb{K}$ to be the finite field $\mathbb{F}_p$ with $p = 2^{30} + 3$ elements, while in the experiments for Algorithm 3 (Figures 5.1 and 5.2) $\mathbb{K} = \mathbb{F}_p$, where $p = 2^{40} + 15$. Naturally, when performing experiments for Algorithm 5, we work over $\mathbb{Z}$.

Timings for Algorithms 1 and 2 are presented in Table 5.1. We compare the straightforward iterative algorithm (column Naive), to the fast baby-step/giant-step algorithms, one based on subproduct trees and resultants (column Algorithm 1), the other based on multipoint evaluation on geometric sequences (column Algorithm 2).

Some conclusions can be drawn by analyzing these timings:

- The theoretical complexities are perfectly reflected in practice: as $N$ is increased from $2^{2k}$ to $2^{2k+2}$, timings are also multiplied (roughly) by 4 in column Naive, and (roughly) by 2 in columns Algorithm 1 and Algorithm 2.
- The asymptotic regime is reached from the very beginning.

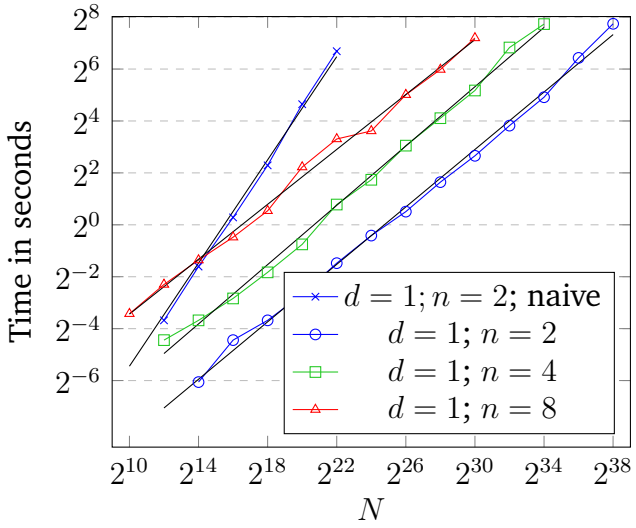



Figure 5.1: Timings of Algorithm 3 implemented in Maple 2020.2. We compare $d = 1$ and $n = 2, 4, 8$ for various values of $N$.

- Algorithm 2 is always faster than Algorithm 1, which is itself much faster than the Naive algorithm, as expected.
- A closer look into the timings shows that for Algorithm 1, $\approx 80\%$ of the time is spent in step (3) (resultant computation), the other steps taking $\approx 10\%$ each; for Algorithm 2, step (1) takes $\approx 25\%$, step (2) takes $\approx 75\%$, and step (3) is negligible.

In order to visualize the performance of Algorithm 3, we work with random polynomial matrices in $\mathcal{M}_n(\mathbb{F}_p[x])$ of degree $d$. Figure 5.1 compares the time needed to compute the Matrix $q$-factorial for $d = 1$ and $n = 2, 4, 8$ as $N$ growths. The black lines represent the best linear fits to the data points and are given by the equations $y = 0.55x - 13.7, y = 0.57x - 11.8$ and $y = 0.53x - 8.7$ respectively. Note that we are plotting on a log-log scale, therefore the established complexity of $\tilde{O}(N^{1/2})$ is indicated by the coefficient of $x$ in these linear fits which is always only slightly greater than $1/2$. In the same figure, we also show the timings for $d = 1$ and $n = 2$ of the naive algorithm given by successively computing and multiplying $M(q^i)$ together. The best linear fit almost perfectly describes this data and has a slope of $0.99 \approx 1$, in line with the linear complexity in $N$.

In Figure 5.2 we show similar timings, however now for $n = 2$ and $d = 2, 4, 8$. The linear fits to the data are now given by $y = 0.53x - 12.4, y = 0.53x - 11.5$ and $y = 0.59x - 11.9$. Again, they describe the observations very well and the coefficients of the regressions are in line with the proven complexity. We observe that, as expected, the lines have slopes of roughly $1/2$ and are closer together than in the previous figure. The naive method has a slope of $1.01 \approx 1$.

Finally, Figure 5.3 shows the timings for computing $q$-holonomic integer sequences. We compute $N$-th terms of $u_n = [n]_3!$ and $v_n = \sum_{i=1}^{n-1} 2^{i^2}$ in $\mathbb{Z}$ for $N$ between $2^{10}$ and $2^{16}$ with the naive approach (green) and using the binary splitting Algorithm 5 (blue). Like before, the black lines are best linear fits to the data. As explained in Section 5.3.5, we expect the slopes of the lines corresponding to the naive method to be around 3 and the slopes of the quasi-optimal computations to be roughly 2. Indeed, we find $y = 3.59x - 39.9$,

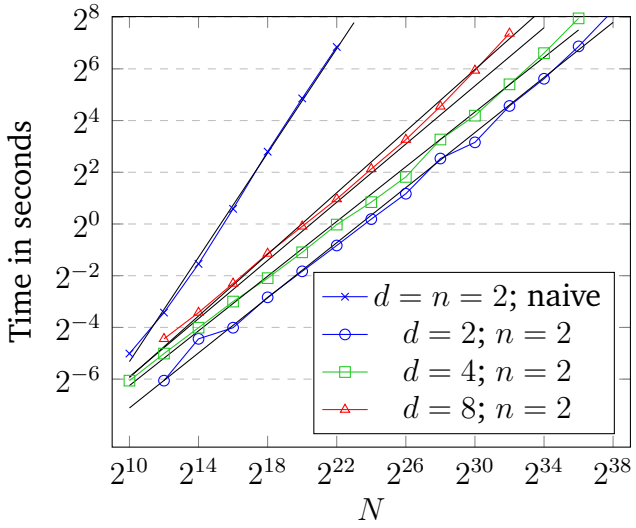



Figure 5.2: Timings of Algorithm 3 implemented in Maple 2020.2. We compare $n = 2$ and $d = 2, 4, 8$ for various values of $N$.

$y = 2.26x - 28.9$ for $u_n$, and $y = 3.9x - 39.6$, $y = 2.36x - 26.9$ for $v_n$.

## 5.6 Conclusion and future work

We have shown that selected terms of $q$-holonomic sequences can be computed fast, both in theory and in practice, the key being the extension of classical algorithms in the holonomic ("$q = 1$") case. We have demonstrated through several examples that this basic algorithmic improvement has many other algorithmic implications, notably on the faster evaluation of many families of polynomials and on the acceleration of algorithms for $q$-difference equations.

Here are some questions that should be investigated in the future.

1. (Counting points on $q$-curves) Counting efficiently points on (hyper-)elliptic curves leads to questions like: for $a, b \in \mathbb{Z}$, compute the coefficient of $x^{\frac{p-1}{2}}$ in $G_p(x) := (x^3+ax+b)^{\frac{p-1}{2}}$ modulo $p$, for one [69] or several [167] primes $p$. A natural extension is to ask the same with $G_p(x)$ replaced by $\prod_{k=1}^{\frac{p-1}{2}}(q^{3k}x^3 + aq^kx + b)$. This might have applications related to §5.4.4, or to counting points on $q$-deformations [284].

2. (Computing $q$-deformed real numbers) Recently, Morier-Genoud and Ovsienko [248] introduced $q$-analogues of real numbers, see also [228, 247]. How fast can one compute (truncations or evaluations of) quantized versions of numbers like $e$ or $\pi$?

3. (Evaluating more polynomials) Is it possible to evaluate fast polynomials of the form $\sum_{\ell=0}^{N} x^{\ell^s}$, for $s \geq 3$, and many others that escape the $q$-holonomic class? *E.g.*, [36] presents a beautiful generalization of Algorithm 1 to the fast evaluation of isogenies between elliptic curves, by using *elliptic resultants*, with applications to isogeny-based cryptography.

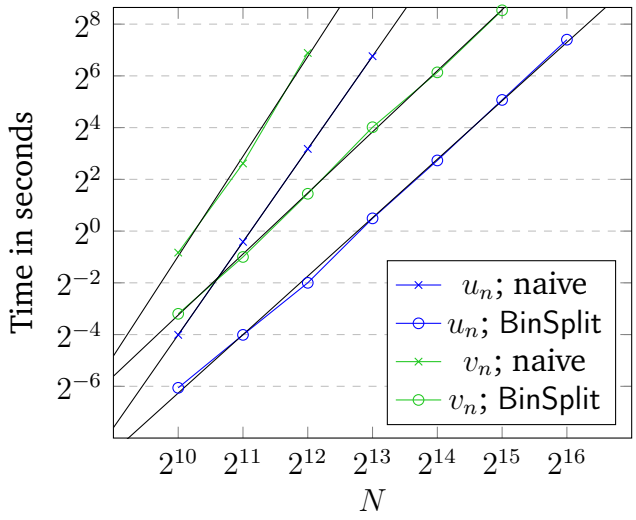



Figure 5.3: Timings for computing the $N$-th term of integer $q$-holonomic sequences $u_n = [n]_3!$ and $v_n = \sum_{i=1}^{n-1} 2^{i^2}$ for various values of $N$ with Algorithm 5 (BinSplit) and naively.

4. ("Precise" complexity of $q$-holonomic sequences) Can one prove non-trivial lower bounds, ideally matching the upper bounds, on examples treated in this chapter?

# Chapter 6

# Computing terms in polynomial C-finite sequences

*"Rule 2: Do not waste a factor of two!"*
*"Rule 8: The development of fast algorithms is slow!"*
Arnold Schönhage, *Fast Algorithms*, 1994

The $N$th power of a polynomial matrix of fixed size and degree can be computed by binary powering as fast as multiplying two polynomials of linear degree in $N$. When Fast Fourier Transform (FFT) is available, the resulting arithmetic complexity is *softly linear* in $N$, i.e. linear in $N$ with extra logarithmic factors. In this chapter we show that it is possible to beat binary powering, by an algorithm whose complexity is *purely linear* in $N$, even in absence of FFT. The key result making this improvement possible is that the entries of the $N$th power of a polynomial matrix satisfy linear differential equations with polynomial coefficients whose orders and degrees are independent of $N$. Similar algorithms are proposed for two related problems: computing the $N$th term of a C-recursive sequence of polynomials, and modular exponentiation to the power $N$ for bivariate polynomials.

This chapter consists of the joint work with A. Bostan and V. Neiger [75].

## 6.1 Introduction

A sequence $(u_n)_{n\geq 0}$ is called *C-recursive* if it satisfies a linear recurrence relation whose coefficients are constant with respect to $n$. The famous sequence $(f_n)_{n\geq 0}$ of Fibonacci numbers, defined by the recurrence $f_{n+2} = f_{n+1} + f_n$ and the initial values $f_0 = 0, f_1 = 1$, is perhaps the most basic example of a C-recursive sequence after the geometric ones $(q^n)_{n\geq 0}$. It is classical that the term $f_N$ can be computed in $O(\log N)$ arithmetic operations, thus as fast as $q^N$. This can be achieved by *binary powering* for $q^N$, and in fact for $f_N$ as well, since it is the top-right entry of $C^N$ where $C = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is a $2 \times 2$ matrix. This idea generalizes to any C-recursive sequence $(u_n)_{n\geq 0}$: a recurrence of order $r \geq 1$ for $(u_n)_{n\geq 0}$ can be encoded, via its *companion matrix*, into an $r \times r$ matrix recurrence of order $1$. Then the term $u_N$ of the sequence appears as an entry of the $N$th power of this $r \times r$ companion matrix multiplied by the vector of initial values $(u_0, \ldots, u_{r-1})$ [245, 136]. Then $u_N$ can be computed in $O(\log N)$ arithmetic operations, and in $O(N \log(N))$ bit operations if $(u_n)_{n\geq 0}$ is an integer sequence,

using fast integer multiplication [169]. Here $r$ is considered as a constant parameter, i.e., $r \in O(1)$.

Fibonacci polynomials $F_n(x)$ are a natural generalization of Fibonacci numbers (see e.g. [88]). They are defined by the recurrence

$$F_{n+2}(x) = xF_{n+1}(x) + F_n(x) \quad \text{for } n \geq 0 \tag{6.1}$$

and the initial values $F_0(x) = 0, F_1(x) = 1$. The first few terms are

$$(F_n)_{n\geq 0} = (0, 1, x, x^2 + 1, x^3 + 2x, x^4 + 3x^2 + 1, \dots).$$

Obviously, for all $n \geq 1$, the polynomial $F_n(x)$ is monic of degree $n - 1$ and the sum of its coefficients is $F_n(1) = f_n$.

Given $N \in \mathbb{N}$, the direct iterative algorithm for computing $F_N(x)$ has arithmetic complexity $O(N^2)$. It computes, for each $n \leq N$, all the $n$ coefficients of the intermediate polynomial $F_n(x)$; in total this amounts to $\Theta(N^2)$ coefficients. Therefore, if one wants to compute all of $(F_0, \dots, F_N)$ then this direct method is optimal with respect to the total arithmetic size of the output. However, it becomes quadratic if one is only interested in determining $F_N(x)$ alone.

To compute the polynomial $F_N(x)$ faster, one can use, as in the scalar case, the reformulation of the second-order recurrence (6.1) as a first-order (polynomial) matrix recurrence:

$$\begin{pmatrix} F_{n+2} & F_{n+1} \\ F_{n+1} & F_n \end{pmatrix} = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}. \tag{6.2}$$

This shows that $F_n(x)$ is the top-right entry of the matrix $C(x)^n$, where $C(x)$ is the $2 \times 2$ polynomial matrix $C(x) = \left(\begin{smallmatrix} x & 1 \\ 1 & 0 \end{smallmatrix}\right)$. One can again compute $C(x)^N$ using binary powering, whose costliest step is the multiplication of two polynomial matrices each with degree about $N/2$. This leads to an algorithm that finds $F_N(x)$ in arithmetic complexity $O(\mathsf{M}(N))$, where $\mathsf{M}(N)$ denotes the arithmetic cost of polynomial multiplication in degree at most $N$.

Using FFT-based polynomial multiplication, this amounts to a number of arithmetic operations in the base field $\mathbb{K}$ which is quasi-linear in $N$ [90]. Not only does this compare favorably to the complexity $O(N^2)$ of the direct iterative algorithm, but this is even quasi-optimal (i.e., optimal up to logarithmic factors) with respect to the arithmetic size $\Theta(N)$ of the output polynomial $F_N(x)$.

In this context, the idea also generalizes to any C-recursive sequence $(u_n(x))_{n\geq 0}$ of polynomials in $\mathbb{K}[x]$, which we will call *polynomial C-recursive sequences*. Indeed, one can encode any recurrence of arbitrary (but independent of $n$) order $r \geq 1$ and coefficients in $\mathbb{K}[x]$ into a polynomial $r \times r$ matrix recurrence of order 1, and the $N$th term of the sequence, $u_N(x)$, can be computed as an element in the $N$th power of an $r \times r$ polynomial matrix multiplied by the polynomial vector of initial values. Conversely, computing the $N$th power of any polynomial matrix can be reduced to computing terms in polynomial C-recursive sequences (see the introduction of Section 6.4). Binary powering allows to solve both problems in $O(\mathsf{M}(N))$ arithmetic operations, and in $O(N^2 \log(N))$ bit operations if $\mathbb{K} = \mathbb{Q}$, considering both the recurrence order (or the matrix size) $r$ and the recurrence degree (or the matrix degree) $d$ as constant parameters, i.e., $r, d \in O(1)$. The main question addressed in this chapter is:

*Can one achieve a better complexity for these tasks?*

As far as scalar C-recursive sequences are concerned, the arithmetic complexity of $O(\log N)$ seems very difficult (if not impossible) to beat, but it is perhaps not impossible to improve the bit complexity $O(N \log(N))$ towards $O(N)$. While we do not achieve this, our results provide polynomial analogues for this type of improvement. As frequently noticed in computer algebra, polynomials are "computationally easier" to deal with than integers. In our case, philosophically, this comes from the fact that we can benefit from an additional operation on polynomials: differentiation.

This possibly cryptic remark will hopefully become clear throughout the next section. There, using the specific case of Fibonacci polynomials as a test bench, we argue why it is indeed legitimate to hope for algorithms of arithmetic complexity $O(N)$ for computing the $N$th term of a polynomial C-recursive sequence.

**Main result** Recall that a *C-finite sequence* is a sequence $(u_n)_{n\geq 0}$ of elements $u_n$ in some ring $R$ which satisfies a recurrence equation

$$u_{n+r} = c_{r-1}u_{n+r-1} + \cdots + c_0 u_n \quad \text{for all } n \geq 0, \tag{6.3}$$

for $c_0, \ldots, c_{r-1} \in R$. In this chapter we consider *polynomial* C-finite sequences, i.e., the case $R = \mathbb{K}[x]$ for some (effective) field $\mathbb{K}$ of characteristic zero; thus $u_n = u_n(x) \in \mathbb{K}[x]$. The customary data structure for representing such a sequence consists of the polynomials $c_0(x), \ldots, c_{r-1}(x)$ defining the recurrence and the $r$ initial conditions $u_0(x), \ldots, u_{r-1}(x) \in \mathbb{K}[x]$. The *order* of the recurrence is $r$ while its *degree* is the maximum of the degrees of the $c_i$'s.

**Theorem 6.1.** *Let $\mathbb{K}$ be an effective field of characteristic 0. Let $d$ and $r$ be fixed positive integers. For each of the following problems, there exists an algorithm solving it in $O(N)$ operations $(\pm, \times, \div)$ in $\mathbb{K}$:*

- *SeqTerm: Given a polynomial C-finite sequence $(u_n(x))_{n\geq 0}$ of order and degree at most $r$ and $d$, compute the $N$th term $u_N(x)$.*
- *BivModPow: Given polynomials $Q(x, y)$ and $P(x, y)$ in $\mathbb{K}[x, y]$ of degrees in $y$ and $x$ at most $r$ and $d$, with $P(x, y)$ monic in $y$, compute $Q(x, y)^N \bmod P(x, y)$.*
- *PolMatPow: Given a square polynomial matrix $M(x)$ over $\mathbb{K}[x]$ of size and degree at most $r$ and $d$, compute $M(x)^N$.*

Our algorithms for SeqTerm, BivModPow, and PolMatPow make essential use of divisions in $\mathbb{K}$. We do not know if the complexity $O(N)$ can be achieved using only operations $(+, -, \times)$ in $\mathbb{K}$.

**Previous work** As already mentioned, the classical way of computing the $N$th term of a given C-finite sequence goes via binary powering of the companion matrix, see e.g. [245]. Another algorithm, due to Fiduccia [136], works via binary powering in a polynomial quotient ring, and has a better complexity with respect the order $r$, but not with respect to $N$.

The fastest existing algorithm [74] for computing the $N$th term of a C-finite sequence (counting arithmetic operations in $\mathbb{K}$) beats Fiduccia's algorithm by a constant factor. In our polynomial C-finite case, and under the assumption $r, d \in O(1)$, all these algorithms achieve an arithmetic complexity in $O(\mathsf{M}(N))$.

Beyond this classical approach, the previous work on the aforementioned problems consists of two distinct directions. The special case of Chebyshev polynomials of the second kind $U_n(x) = (-i)^n F_{n+1}(2ix)$ (with $F_n(x)$ the $n$th Fibonacci polynomial and $i$ the imaginary unit) was considered in [208] (and later in [118]). These references present various methods for the computation of the Chebyshev polynomials (of the first and second kind) with arithmetic complexity ranging from $O(N)$ to $O(N^3)$. The results in [208, 118] exploit the particular structure of these polynomials; except for possibly other families of classical orthogonal polynomials, for which explicit (hypergeometric) formulas exist, the methods in [208, 118] do not admit obvious generalizations.

An idea closely connected to a fundamental building block of our algorithms is explained in [138, Pbm. 4]. There, Flajolet and Salvy exploit the fact that, given a polynomial $P(x)$ in $\mathbb{K}[x]$, the coefficient sequence of the $n$th power $P(x)^n$ satisfies a linear recurrence of order independent of $n$, and with coefficients in $\mathbb{K}[x, n]$ of degree independent of $n$; this recurrence allows them to compute (a selected coefficient of) $P(x)^N$ more efficiently than by binary powering. This idea has been applied in [69, §8] to count points on hyperelliptic curves over finite fields, with applications to cryptography. The technique also yields a general solution to SeqTerm when $r = 1$.

**Outline** The following observation generalizes that in [138]: the coefficient sequence of the $n$th power of any *algebraic function* satisfies a recurrence of order and degree independent of $n$. From this, in Section 6.3, we give algorithms for SeqTerm with cost $O(N)$.

To complete the proof of Theorem 6.1, we design reductions between the three problems. Obviously PolMatPow $\Rightarrow$ SeqTerm, i.e., any algorithm for PolMatPow with cost $O(N)$ induces one for SeqTerm with cost $O(N)$ as well. Indeed, the $N$th term of a polynomial C-recursive sequence is equal to an entry of the product of the $N$th power of a companion matrix and the vector of initial values, and this polynomial matrix-vector multiplication costs $O(N)$. Conversely, it also holds that SeqTerm $\Rightarrow$ PolMatPow. One natural way to see this is to consider $r^2$ sequences corresponding to each entry of $M(x)^n$, with recurrence given by the characteristic polynomial of $M(x)$; see the introduction of Section 6.4. In Section 6.4.2, we give a more efficient algorithm for this reduction, based on an algorithm for SeqTerm $\Rightarrow$ BivModPow described in Section 6.4.1.

**Basics of complexity and holonomic functions** Throughout the text, $\mathbb{K}$ denotes an effective field of characteristic zero. For analyzing the performance of algorithms we will use the arithmetic complexity model, meaning that arithmetic operations $(+, -, \times, \div)$ in the base field $\mathbb{K}$ are counted at unit cost. As before, $\mathsf{M}(N)$ stands for the complexity of multiplication of two polynomials in $\mathbb{K}[x]$ of degree at most $N$. With FFT-based multiplication $\mathsf{M}(N) \in O(N \log(N) \log\log(N))$ [90], improved to $O(N \log(N))$ if $\mathbb{K}$ contains suitable roots of unity [115] or if $\mathbb{K}$ is any finite field [170].

A power series $f(x) \in \mathbb{K}[\![x]\!]$ is called *D-finite* (or *holonomic*) if it satisfies a linear differ-

ential equation (ODE) of the form

$$q_\ell(x)f^{(\ell)}(x) + \cdots + q_0(x)f(x) = 0, \tag{6.4}$$

for some $q_0(x), \ldots, q_\ell(x) \in \mathbb{K}[x]$ with $q_\ell(x) \neq 0$. Equivalently, writing $f(x) = \sum_{k\geq 0} f_k x^k$, the sequence $(f_k)_{k\geq 0}$ is *P-recursive*, i.e., it satisfies a linear recurrence equation (LRE)

$$p_s(k)f_{k+s} + \cdots + p_0(k)f_k = 0 \quad \text{for all } k \geq 0,$$

with polynomial coefficients $p_0(x), \ldots, p_s(x) \in \mathbb{K}[x]$, and $p_s \neq 0$. Note that $s$ and $\ell$ may differ in general, but $s \leq \ell + \max_i(\deg q_i(x))$. It also holds that $\max_i(\deg p_i(x)) \leq \ell$.

It is often useful to write (6.4) as $Lf(x) = 0$, where

$$L = q_\ell(x)\partial_x^\ell + \cdots + q_0(x)$$

is an element in the non-commutative Weyl algebra $\mathbb{K}[x]\langle\partial_x\rangle$ of linear differential operators with multiplication governed by the Leibniz rule $\partial_x x = x\partial_x + 1$. The *order* $\ell$ of a differential operator $L$ is the highest power of $\partial_x$ occurring in $L$, and the *degree* of $L$ is the highest power of $x$. We recall that the least common left multiple (LCLM) of two operators $L_1, L_2 \in \mathbb{K}[x]\langle\partial_x\rangle$ is the unique monic operator $L = \mathrm{LCLM}(L_1, L_2)$ of minimal order such that there exist two nonzero operators $A, B \in \mathbb{K}(x)\langle\partial_x\rangle$ with $L = AL_1 = BL_2$. The LCLM can be computed efficiently [66].

## 6.2 The case of Fibonacci polynomials

Before solving the first part (SeqTerm) of Theorem 6.1 in general, we propose in this section three different approaches that can be used to compute the $N$th Fibonacci polynomial $F_N(x)$ in arithmetic complexity $O(N)$. Two of these methods have the advantage that they generalize to the case of arbitrary C-finite sequences.

The starting point of all that follows is the observation that the generating function $F(x, y) := \sum_{n\geq 0} F_n(x)y^n \in \mathbb{K}[x][\![y]\!]$ of the sequence $(F_n(x))_{n\geq 0}$ is *rational*, and equal to $y/(1 - xy - y^2)$.

### 6.2.1 First method via a closed-form expression

By using the partial fraction decomposition

$$\frac{y}{1 - xy - y^2} = \frac{1}{\varphi_+(x) - \varphi_-(x)} \cdot \left(\frac{1}{1 - \varphi_+(x)y} - \frac{1}{1 - \varphi_-(x)y}\right)$$

where $\varphi_\pm(x) = (x \pm \sqrt{x^2 + 4})/2$ are the roots of $\varphi^2 - x\varphi - 1 = 0$, and by applying the geometric series, we get the closed form expression

$$F_n(x) = \frac{\varphi_+(x)^n - \varphi_-(x)^n}{\varphi_+(x) - \varphi_-(x)} \quad \text{for all } n \geq 0. \tag{6.5}$$

Now, using the binomial formula twice, we obtain the formula

$$F_n(x) = \frac{1}{2^{n-1}} \cdot \sum_{\ell \geq 0} 4^\ell \left( \sum_{k \geq 0} \binom{n}{2k+1} \binom{k}{\ell} \right) x^{n-2\ell-1}. \tag{6.6}$$

The identity [163, 3.121] implies a "magic" simplification:

$$\sum_{k \geq 0} \binom{n}{2k+1} \binom{k}{\ell} = 2^{n-1-2\ell} \binom{n-\ell-1}{\ell}. \tag{6.7}$$

In conclusion, from (6.6) and (6.7) it follows that

$$F_n(x) = \sum_{\ell \geq 0} \binom{n-\ell-1}{\ell} x^{n-2\ell-1}. \tag{6.8}$$

With this expression at hand, it becomes transparent that one can compute $F_N(x)$ efficiently. Indeed, by writing $F_N(x) = \sum_{k=0}^{N-1} f_k x^k$, it follows from (6.8) that $(f_n)_{n \geq 0}$ satisfies the recurrence relation

$$f_{k+2} = \frac{(N+k+1)(N-k-1)}{4(k+1)(k+2)} f_k \quad \text{for all } k \geq 0. \tag{6.9}$$

Moreover, (6.8) also gives $(f_0, f_1) = (1, 0)$ for odd $N$ and otherwise $(f_0, f_1) = (0, N/2)$. With these initial conditions, it is now clear that $F_N(x)$ can be computed in $O(N)$ by unrolling the recurrence (6.9).

As mentioned in the introduction, the analogue of formula (6.8) for the case of Chebyshev polynomials of the first kind $T_n(x)$ was already exploited in [208, §1.9]. The disadvantage of this approach is that for general polynomial C-finite sequences there is no hope for a closed form expression like (6.8).

### 6.2.2 Second method via algebraic substitution

There is another method for computing $F_N(x)$ in $O(N)$, which has the advantage that it generalizes to any C-recursive sequence, as we will show in Section 6.3.1.

The crucial remark (Lemma 6.3) is that since $\varphi_\pm(x)$ is algebraic, $\varphi_\pm(x)^n$ satisfies a "small" linear differential equation: an ODE of order and degree independent of $n$. The same holds true for $1/(\varphi_+(x) - \varphi_-(x))$, therefore for $F_n(x)$ as well. More precisely, $\varphi_\pm(x)^n$ satisfies the linear differential equation

$$(x^2+4)y''(x) + xy'(x) - n^2 y(x) = 0,$$

and $1/(\varphi_+(x) - \varphi_-(x)) = (x^2+4)^{-1/2}$ satisfies the ODE

$$(x^2+4)y'(x) + xy(x) = 0.$$

Using (6.5), it then follows that the polynomial $F_n(x)$ satisfies

$$(x^2+4)y''(x) + 3xy'(x) + (1-n^2)y(x) = 0. \tag{6.10}$$

Writing $F_N(x) = \sum_{k=0}^{N-1} f_k x^k$, plugging into (6.10) for $n = N$ and extracting the $(k+2)$nd coefficient, it now follows that the sequence $(f_k)_{k\geq 0}$ satisfies recurrence (6.9). The initial conditions $f_0, f_1$ are given by $F_N(x) \bmod x^2$ which can be found in complexity $O(\log N)$ by computing the $N$th power of the companion matrix (6.2) in $\mathbb{K}[x]/(x^2)$ by binary powering and reducing mod $x^2$ in each step. As before, unrolling recurrence (6.9), with these initial terms, provides a way to compute $F_N(x)$ in complexity $O(N)$.

### 6.2.3 Third method via Creative Telescoping

Writing $F(x, y) = y/(1 - xy - y^2)$ we are interested in a differential equation for the coefficient of $y^N$ in $F(x, y)$. By Cauchy's integral formula, we have for sufficiently small $\epsilon > 0$:

$$F_N(x) = [y^N]F(x, y) = \frac{1}{2\pi i} \oint_{|y|=\epsilon} \frac{y}{(1 - xy - y^2)y^{N+1}} \mathrm{d}y.$$

Then the method of creative telescoping can be used to find a differential equation for the integral above. For example, the command

```
DEtools[Zeilberger](1/(1-x*y-y^2)/y^n, x, y, Dx);
```

in Maple immediately finds that

$$\left((x^2 + 4)\partial_x^2 + 3x\partial_x + 1 - n^2\right) \frac{F(x, y)}{y^{n+1}} = \partial_y \left( \frac{F(x, y)}{y^n} C(x, y) \right),$$

where $C(x, y) = \frac{n+1-nxy-(n-1)y^2}{1-xy-y^2}$. By Cauchy's integral theorem, the contour integral of the right-hand side vanishes, and thus (6.10) follows. Then one can conclude in the same way as in the previous method and compute $F_N(x)$ in complexity $O(N)$.

### 6.2.4 Comments on the three approaches

It is natural to ask ourselves what in these approaches was just luck, what was truly specific to the particular example of the Fibonacci polynomials, and what can be extended to the general case.

It is clear that the key for computing $F_N(x)$ in complexity $O(N)$ is the existence of the recurrence (6.9) (or equivalently the ODE (6.10)). Even though there is no hope for a closed form solution in general, we shall prove that such a recurrence always exists for polynomial C-finite sequences. We should, however, definitely be careful and avoid proving tautologic statements. Since $u_N(x)$ is a polynomial, it does satisfy *a priori* a first-order linear differential equation with polynomial coefficients, $u_N(x)y'(x) - u'_N(x)y(x) = 0$, but this one is trivial for our purposes. Indeed, converting this differential equation into a recurrence satisfied by the sequence of coefficients of $u_N(x)$ yields a recurrence of order $\deg(u_N)$, which is obviously useless for computing the coefficients of $u_N$. Therefore, we need to ensure existence of a recurrence/ODE whose order and degree are independent of $N$. This is the purpose of the next section. Specifically, in §6.3.1 we will explain how it can be found by algebraic substitution (generalizing §6.2.2) and in §6.3.2 we will show that it can also be found in general via creative telescoping (generalizing §6.2.3).

## 6.3 Polynomial C-finite sequences

Recall that a polynomial C-finite sequence $(u_n(x))_{n\geq 0}$ is a sequence of polynomials $u_n(x) \in \mathbb{K}[x]$ that satisfies a recurrence

$$u_{n+r}(x) = c_{r-1}(x)u_{n+r-1}(x) + \cdots + c_0(x)u_n(x), \tag{6.11}$$

of some order $r \in \mathbb{N}$ and polynomial coefficients $c_0(x), \ldots, c_{r-1}(x) \in \mathbb{K}[x]$. The *degree* of (6.11) is $d = \max_i(\deg c_i(x))$. Eq. (6.11) defines the sequence $(u_n(x))_{n\geq 0}$ uniquely if $r$ initial terms $u_0(x), \ldots, u_{r-1}(x)$ are prescribed. The characteristic polynomial for (6.11) is defined as

$$\chi(y) = y^r - c_{r-1}(x)y^{r-1} - \cdots - c_1(x)y - c_0(x) \in \mathbb{K}[x, y].$$

It is well-known and not difficult to see that the generating function $U(x, y) \coloneqq \sum_{n\geq 0} u_n(x)y^n$ is a rational function and given by

$$U(x, y) = \frac{v_0(x) + \cdots + v_{r-1}(x)y^{r-1}}{y^r\chi(1/y)}, \tag{6.12}$$

with $v_k(x) \coloneqq u_k(x) - c_{r-1}(x)u_{k-1}(x) - \cdots - c_{r-k}(x)u_0(x)$.

Let $a_1(x), \ldots, a_k(x) \in \overline{\mathbb{K}(x)}$ be the roots of $\chi(y) = 0$, and $m_1, \ldots, m_k$ be their multiplicities. It follows from the partial fraction decomposition and geometric series that any sequence $(u_n(x))_{n\geq 0}$ satisfying (6.11) is of the form

$$u_n(x) = q_1(n, x)a_1(x)^n + \cdots + q_k(n, x)a_k(x)^n, \tag{6.13}$$

where $k \leq r$ and each $q_i(n, x) \in \mathbb{K}(a_1(x), \ldots, a_n(x))[n]$ is a polynomial in $n$ of degree at most $m_i - 1$, for $i = 1, \ldots, k$.

### 6.3.1 Computing $u_N(x)$ in $O(N)$

By generalizing the ideas of Section 6.2.2, it is not difficult to prove that the $n$th term of a polynomial C-recursive sequence $(u_n(x))_{n\geq 0}$ satisfies a linear ODE whose degree and order are independent of $n$, and consequently, that there exists a linear recurrence relation for the coefficient sequence of $u_n(x)$ whose order (say $s$) and degree are again independent of $n$. Then, for a given $N \in \mathbb{N}$, first computing initial terms by binary powering of the companion matrix in $\mathbb{K}[x]/(x^s)$ and then unrolling this recurrence for $n = N$, we achieve a complexity $O(N)$ for the computation of $u_N(x)$.

**Theorem 6.2.** *Let $(u_n(x))_{n\geq 0}$ be a polynomial C-recursive sequence. Then there exists $L_n \in \mathbb{K}[n, x]\langle\partial_x\rangle$ with order and degree independent of $n$, and such that $L_n(u_n(x)) = 0$. Consequently, writing $u_n(x) = \sum_{k\geq 0} c_{n,k}x^k$, there exist, for some $s \in \mathbb{N}$ independent of $n$, polynomials $p_0(n, x), \ldots, p_s(n, x) \in \mathbb{K}[n, x]$ of degrees independent of $n$, and such that the sequence $(c_{n,k})_{k\geq 0}$ satisfies the recurrence*

$$p_s(n, k)c_{n,k+s} + \cdots + p_0(n, k)c_{n,k} = 0, \quad k \geq 0. \tag{6.14}$$

**Algorithm 6** SeqTermAS$((u_n)_n, N)$

*Input:* A polynomial C-finite sequence $(u_n(x))_{n\geq 0}$ given by (6.11) with initial conditions, and $N \in \mathbb{N}$.

*Output:* The polynomial $u_N(x)$.

1: $\chi(y) \leftarrow$ the characteristic polynomial of $(u_n(x))_{n\geq 0}$
2: $a_1(x), \ldots, a_k(x) \leftarrow$ the roots of $\chi(y)$
3: Compute minimal polynomials for $q_1(x,n), \ldots, q_k(x,n) \in \mathbb{K}(a_1(x), \ldots, a_k(x))[n]$ such that (6.13) holds.
4: For each $i$ deduce an ODE $L_{i,n} \in \mathbb{K}[n,x]\langle\partial_x\rangle$ with order and degree independent of $n$ such that $L_{i,n}(q_i(x,n)a_i(x)^n) = 0$
5: $L_n \leftarrow \mathrm{LCLM}(L_{1,n}, \ldots, L_{k,n}) \in \mathbb{K}[n,x]\langle\partial_x\rangle$
6: Compute a recurrence $p_s(n,k)c_{n,k+s} + \cdots + p_0(n,k)c_{n,k} = 0$ satisfied by any solution $f_n(x) = \sum_{k\geq 0} c_{n,k}x^k$ of $L_n y = 0$
7: Using binary powering of the companion matrix of the initial recurrence mod $x^s$, compute the values $c_{N,0}, \ldots, c_{N,s-1}$
8: Unroll the recurrence from Line 6 for $n = N$ and with initial terms from Line 7
9: **return** $\sum_{k=0}^{Nd} c_{N,k}x^k$, where $d = \deg_x(\chi(y))$

In the theorem above it is crucial that neither the order nor the degree of $L_n$ depend on $n$. Since each $u_n(x)$ is a polynomial, it is a tautology to say that it satisfies *some* linear differential equation with polynomial coefficients: one may simply take $L = \partial_x^\alpha$, where $\alpha > \deg(u_n(x))$ or $L = u_n(x)\partial_x - u_n'(x)$. However, it is a nontrivial fact that $u_n(x)$ satisfies an ODE of the form

$$p_\ell(n,x)u_n^{(\ell)}(x) + \cdots + p_0(n,x)u_n(x) = 0$$

for some $p_i(n,x) \in \mathbb{K}[n,x]$ (with $\ell$ and $\deg_x p_i$ independent of $n$).

The most direct proof of Theorem 6.2 uses the explicit expression (6.13) for $u_n(x)$ and the following classical fact about algebraic substitution into D-finite functions. Recall that a function $a(x)$ is called *algebraic* over $\mathbb{K}(x)$ if it satisfies a non-trivial polynomial relation $P(x,a(x)) = 0$ for some $P(x,y) \in \mathbb{K}[x,y]$. Size and complexity bounds on differential equations for algebraic functions, and more generally on algebraic substitution, are given in [65, 201].

**Lemma 6.3.** *Let $a(x)$ be an algebraic function over $\mathbb{K}(x)$ and let $g(x)$ be D-finite. Then $f(x) = g(a(x))$ is D-finite. In particular, $a(x)^n$ satisfies a linear ODE of order and degree independent of $n$.*

*Proof.* The first part is a classical result, see for example [299, Thm. 2.7]. In the proof one shows that the vector space spanned over $\mathbb{K}(x)$ by $(f^{(i)}(x))_{i\geq 0}$ is finite-dimensional over $\mathbb{K}(x, a(x))$ which is itself finite-dimensional over $\mathbb{K}(x)$. For the second part, it is enough to set $g(x) = x^n$ which satisfies $xg'(x) = ng(x)$. □

**Example 6.4.** Like in Section 6.2 let $\varphi_\pm(x) = (x \pm \sqrt{x^2+4})/2$ be the roots of $y(x)^2 + xy(x) - 1 = 0$. Then $\varphi_\pm(x)^n$ satisfy the ODE

$$(x^2+4)y''(x) + xy'(x) - n^2y(x) = 0.$$

*Proof of Theorem 6.2.* Write $u_n(x)$ as in (6.13). By Lemma 6.3, each $a_i(x)^n$ satisfies a linear differential equation of order and degree independent of $n$, hence the same holds for $q_i(n,x)a_i(x)^n$, and finally for $u_n(x)$. It follows that the coefficient sequence of $u_n(x)$ is P-recursive with order and degree independent of $n$. □

Since all steps in the proofs above are effective and independent of $N$, this leads to Algorithm 6. Its Lines 1 to 6 can be seen as "precomputations" since they do not depend on $N$. As already mentioned, Line 7 has complexity $O(\log N)$ and Line 8 has complexity $O(N)$. Thus, Algorithm 6 solves SeqTerm in complexity $O(N)$, up to a potential issue during the unrolling at Line 8 of the recurrence from Line 6. Indeed, this unrolling may be impossible for some values $k$, namely those for which $p_s(N,k)$ vanishes. We will explain how to overcome this problem in Section 6.3.3.

For practical applications, however, computing the polynomials $q_i(x,n)$ in Line 3 as well as the LCLM in Line 5 is algorithmically somewhat cumbersome. Thus, generalizing the approach in Section 6.2.3, we now propose a variant of Algorithm 6 which replaces Lines 1 to 5 by an algorithm based on creative telescoping.

### 6.3.2 Computing $L_n$ with Creative Telescoping

Let $U(x,y) = \sum_{n\geq 0} u_n(x)y^n \in \mathbb{K}[x][\![y]\!]$ be the generating function (6.12) of $(u_n(x))_{n\geq 0}$. The sequence is C-finite, so $U(x,y)$ is a rational function. Moreover, the Cauchy integral formula implies that

$$u_n(x) = \frac{1}{2\pi i}\oint_{|y|=\epsilon} \frac{U(x,y)}{y^{n+1}}\mathrm{d}y. \tag{6.15}$$

A *telescoper* of $U(x,y)/y^{n+1}$ is a differential operator

$$L = p_k(x)\partial_x^k + \cdots + p_0(x) \in \mathbb{K}[x]\langle\partial_x\rangle,$$

such that $L$ applied to $U(x,y)/y^{n+1}$ is $\partial_y(C(x,y))$ for some rational function $C(x,y)$ called the *certificate*. By the Cauchy integral theorem, $\oint_{|y|=\epsilon}\partial_y(C(x,y))\mathrm{d}y = 0$, and it follows that $Lu_n(x) = 0$, i.e., $L$ yields a differential equation for $u_n(x)$. In this section we will prove that for $U(x,y)/y^{n+1}$ there exists a telescoper $L_n \in \mathbb{K}[n,x]\langle\partial_x\rangle$ whose order and degree do not depend on $n$. Our proof relies on reduction-based creative telescoping and repeatedly uses the *Hermite reduction* algorithm [59, 60, 64].

We now introduce the necessary definitions and recall the Hermite reduction method. For a more detailed introduction, a full complexity analysis, and applications of reduction-based creative telescoping to integration of bivariate rational functions, we refer to [59]. Let $\mathbb{L} = \mathbb{K}(x)$. For a polynomial $Q(y) \in \mathbb{L}[y]$, let $Q = Q_1Q_2^2\cdots Q_k^k$ be its squarefree factorization and let $Q^* = Q_1\cdots Q_k$ denote the squarefree part of $Q$. We set $Q^- := Q/Q^*$. Recall that, given $P, Q \in \mathbb{L}[y]$, the *Hermite reduction* algorithm computes two polynomials $A, a \in \mathbb{L}[y]$ with $\deg_y a < \deg_y Q^*$ such that

$$\frac{P}{Q} = \partial_y\left(\frac{A}{Q^-}\right) + \frac{a}{Q^*}.$$

Given a bivariate rational function $H(x,y) = P(y)/Q(y) \in \mathbb{K}(x,y)$, one may compute the Hermite reduction $(A_i, a_i)$ of $\partial_x^i H$ for $i = 0, 1, \ldots$. Since $\deg_y a_i$ is uniformly bounded

by $d^* = \deg_y Q^*$ for each $i$, the $d^* + 1$ functions $\{a_i(x, y) \colon 0 \le i \le d^*\}$ will be linearly dependent over $\mathbb{K}(x)$. Hence one can find $q_0(x), \cdots, q_{d^*}(x) \in \mathbb{K}(x)$ not all zero, such that $\sum_{i=0}^{d^*} q_i(x) a_i(x) = 0$. It follows then that $L = \sum_{i=0}^{d^*} q_i(x) \partial_x^i$ is a telescoper for $H$.

This procedure cannot be directly applied to $U(x, y)/y^{n+1}$ if $n$ is an indeterminate. At the same time, if $n = N \in \mathbb{N}$ is fixed, it is *a priori* not obvious that $\deg_x q_i(x)$ will be independent of $N$. Moreover, the complexity of the algorithm will depend on $N$, which we want to avoid. As we will now explain, to achieve this, one should see $U(x, y)/y^{n+1}$ not as a rational function in $x$ and $y$ with potentially large degree in the numerator, but as a hyperexponential function with the parameter $n$ appearing solely as a coefficient in the logarithmic derivative. Recall that $H(x, y)$ is called *hyperexponential* if $\partial_x H/H$ and $\partial_y H/H$ belong to $\mathbb{K}(x, y)$.

For hyperexponential functions, the Almkvist-Zeilberger algorithm [16] was the first practical method to find telescopers and certificates. Indeed, as we mentioned in Section 6.2.3, the command

```
DEtools[Zeilberger](1/(1-x*y-y^2)/y^n, x, y, Dx);
```

in Maple immediately finds the differential equation for the $n$th Fibonacci polynomial for a variable $n$. Note that if $n$ is specialized to an integer $N$ before the execution of the command above, the implemented algorithm becomes slower as $N$ growths.

It is, however, not clear that the Almkvist-Zeilberger algorithm applied to $U(x, y)/y^{n+1}$ will always find a telescoper whose degree and order are independent of $n$, even though we know from Section 6.3.1 that an ODE with this property exists. Therefore, to have a complete algorithm based on creative telescoping, we will invoke the reduction-based method for hyperexponential functions first introduced and analyzed in [60]. Using the implementation of the latter work, the command in Maple

```
HermiteTelescoping(1/(1-x*y-y^2)/y^n, x, y, Dx);
```

also immediately finds the correct ODE for $F_n(x)$. The practical advantage for our purpose of using the reduction based algorithm in comparison to the Almkvist-Zeilberger method is shown in Section 6.5 (Table 6.1). The theoretical advantage comes from the following lemma, which guarantees that the algorithm will find a telescoper for $U(x, y)/y^{n+1}$, and consequently an ODE for $u_n(x)$, whose order and degree do not depend on $n$.

**Lemma 6.5.** *Let $P(y) \in \mathbb{L}[n, y]$ and $Q(y) \in \mathbb{L}[y]$ with $Q(0) \neq 0$. Set $d_n \coloneqq \deg_n P(y)$, $d^* \coloneqq \deg_y Q^*(y)$ and let $k$ be the highest pure power in the square free factorization of $Q(y)$. Then there exist $B(n, y) \in \mathbb{L}(n)[y]$ and $b(n, y) \in \mathbb{L}[n, y]$ with $\deg_y b(n, y) \le d^*$ and $\deg_n b(n, y) \le d_n + k$ such that*

$$\frac{P(y)}{Q(y)y^{n+1}} = \partial_y \left( \frac{B(n, y)}{Q^-(y)y^n} \right) + \frac{b(n, y)}{Q^*(y)y^{n+1}}. \tag{6.16}$$

*Proof.* We are going prove the statement by induction on $d^- \coloneqq \deg_y Q^-$. If $d^- = 0$, then $Q^* = Q$ and the Euclidean division gives $P = P_1 Q + b_1$ with $\deg_y b_1 < d^*$. Moreover,

$$\frac{P_1(y)}{y^{n+1}} = \partial_y \left( \frac{B_1(n, y)}{y^n} \right),$$

where $B_1(n,y)$ is $P_1(y)$ with the $k$th coefficient $p_k$ replaced by $p_k/(k-n)$. Setting $B = B_1$ and $b = b_1$ proves the induction basis.

Now assume that $d^- > 0$ and note that

$$\partial_y\left(\frac{B(n,y)}{Q^-(y)y^n}\right) = y^{-n}\partial_y\left(\frac{B(n,y)}{Q^-(y)}\right) - y^{-n-1}n\frac{B(n,y)}{Q^-(y)},$$

so equation (6.16) is equivalent to

$$\frac{P(y)}{Q(y)y} = \partial_y\left(\frac{B(n,y)}{Q^-(y)}\right) - n\frac{B(n,y)}{Q^-(y)y} + \frac{b(y)}{Q^*(y)y}. \tag{6.17}$$

The Hermite reduction algorithm applied to $P(y)/(Q(y)y)$ yields polynomials $A(y), a(y)$ with $\deg_y a(y) \le d^*$ such that

$$\frac{B(y)}{Q(y)y} = \partial_y\left(\frac{A(y)}{Q^-(y)}\right) + \frac{a(y)}{Q^*(y)y}. \tag{6.18}$$

Comparing (6.17) and (6.18), we now look at

$$H(y) \coloneqq \frac{a(y)}{Q^*(y)y} + n\frac{A(y)}{Q^-(y)y}.$$

The denominator of $H(y)$ is $R(y)y \coloneqq \operatorname{lcm}(Q^*, Q^-)y$. Clearly, $R^* = Q^*$ and $\deg_y R^- < d^-$. The highest pure power in the square free factorization of $R(y)$ is at most $k-1$ and the degree of the numerator in $n$ of $H(y)$ is at most $d_n + 1$. Hence, by induction, we may write

$$\frac{a(y)}{Q^*(y)y} + n\frac{A(y)}{Q^-(y)y} = \partial_y\left(\frac{C(n,y)}{Q^-(y)}\right) - n\frac{C(n,y)}{Q^-(y)y} + \frac{c(n,y)}{Q^*(y)y}$$

with $\deg_y c(n,y) < d^*$ and $\deg_n c(n,y) \le d_n + k$. Setting $B(n,y) = A(y) + C(n,y)$ and $b(y) = a(y) + c(n,y)$ finishes the proof. □

The proof of Lemma 6.5 induces an algorithm for the computation of $B(n,y)$ and $b(n,y)$ given $P(y), Q(y) \in \mathbb{K}[x,y]$ such that (6.16) holds, $\deg_y b \le \deg_y Q^*$ and also $\deg_n b$ bounded in terms of $Q$. It can be seen as a special case of the procedure HermiteReduction in [60]. The ODE for $u_n(x)$ can be now found as in Algorithm 7.

**Algorithm 7** TelescNthTerm$(U(x,y))$

*Input:* A rational function $U(x,y) \in \mathbb{K}(x,y) \cap \mathbb{K}[\![x,y]\!]$.
*Output:* A diff. operator in $\mathbb{K}[n,x]\langle\partial_x\rangle$ for $u_n(x) = [y^n]U(x,y)$.

1: Write $U(x,y) = P(x,y)/Q(x,y)$ and let $d^* = \deg_y Q^*(x,y)$
2: For each $i = 0, \ldots, d^*$ compute the polynomial $b_i(n,x,y) = b(n,y)$ as in Lemma 6.5 applied to $\partial_x^i U(x,y)/y^{n+1}$
3: Find a linear relation of $\{b_i(n,x,y)\colon 0 \le i \le d^*\}$ over $\mathbb{K}(n,x)$, that is polynomials $q_0(n,x), \ldots, q_{d^*}(n,x)$ not all zero with $\sum_{i=0}^{d^*} q_i(n,x)b_i(n,x,y) = 0$
4: Return the differential operator $\sum_{i=0}^{d^*} q_i(n,x)\partial_x^i \in \mathbb{K}[n,x]\langle\partial_x\rangle$

Note that, as in the usual reduction based creative telecoping, the linear relation at Line 3 exists because $\deg_y(b_i(n,x,y))$ is uniformly bounded by $d^*$. Writing $U(x,y) = P(x,y)/Q(x,y)$, the operator $L = \sum_{i=0}^{d^*} q_i(n,x)\partial_x^i$ annihilates $u_n(x) = [y^n]U(x,y)$ since

$$2\pi i \cdot L_n u_n(x) = L_n \oint \frac{U(x,y)}{y^{n+1}}\mathrm{d}y = \oint L_n \frac{P(x,y)}{Q(x,y)y^{n+1}}\mathrm{d}y$$
$$= \oint \partial_y \frac{\sum_{i=1}^{d^*} q_i(x,y)B_i(n,y)}{Q^-(x,y)y^n}\mathrm{d}y + \oint \frac{\sum_{i=0}^{d^*} q_i(n,x)b_i(n,x,y)}{Q^*(x,y)y^{n+1}}\mathrm{d}y;$$

the first integral vanishes by Cauchy's integral theorem, and the second integral vanishes by construction of the $q_i(n,x)$.

This provides a variant for Lines 1 to 5 of Algorithm 6, as described in Algorithm 8.

**Algorithm 8** SeqTermCT$((u_n)_n, N)$

*Input:* A polynomial C-finite sequence $(u_n(x))_{n\geq 0}$ given by (6.11) with initial conditions, and $N \in \mathbb{N}$.

*Output:* The polynomial $u_N(x)$.

1: $U(x,y) \leftarrow$ the rational generating function of $u_n(x)$ in (6.12)
2: $L_n \leftarrow \mathrm{TelescNthTerm}U(x,y)$
3: ▷ *follow Lines 6 to 9 of Algorithm 6*

As above, a potential issue is that the unrolling step is only possible for $p_s(N,k) \neq 0$. The next section deals with this problem.

### 6.3.3 The singular case

In this section we discuss the potential issue of our algorithm that can occur if the sequence for the coefficients of $u_n(x)$ cannot be unrolled due to singularities. We shall first highlight this problem and its solution by means of an example.

Consider the polynomial C-recurrent sequence $u_n(x)$ given by

$$u_{n+3}(x) - (x^2+x+2)u_{n+2}(x) + x(x^2+2x+2)u_{n+1}(x) - 2x^3u_n(x) = 0,$$

for all $n \geq 0$ with initial conditions $u_0 = 3, u_1 = x^2+x+2, u_2 = x^4+x^2+4$. The characteristic polynomial of the defining recurrence is easily computed and turns out to factor completely over $\mathbb{K}[x][y]$:

$$\chi(x,y) = (y-2)(y-x)(y-x^2).$$

With the initial conditions and after a partial fraction decomposition it follows that the generating function of $u_n(x)$ is given by

$$U(x,y) = \frac{1}{1-2y} + \frac{1}{1-x^2y} + \frac{1}{1-xy}.$$

Hence, the solution is $u_n(x) = 2^n + x^n + x^{2n}$, and can be even written down in $O(\log(N))$ arithmetic operations for any $N$. However, as we shall explain now, the direct application of any of the methods described earlier fails.

According to Theorem 6.2, $u_n(x)$ satisfies an ODE whose degree and order are independent of $n$. Indeed, using creating telescoping one quickly finds an annihilator for $u_n(x) = \oint U(x,y)/y^{n+1}\mathrm{d}x$:

$$\big(x^2\partial_x^3 - 3x(n-1)\partial_x^2 + (2n-1)(n-1)\partial_x\big)u_n(x) = 0.$$

Converting this ODE to a recurrence for the coefficient sequence of $u_n(x) = \sum_{k\geq 0} c_{n,k}x^k$ we find

$$(2n-k)(n-k)kc_{n,k} = 0, \quad k \geq 0. \tag{6.19}$$

In other words, $c_{n,k} = 0$ for all $k \in \mathbb{N}$ except $k \in \{0, n, 2n\}$. In order to "unroll" this recurrence we need to know $c_{n,0}, c_{n,n}$ and $c_{n,2n}$. However, it is not immediately clear how to compute those terms for $n = N$ in $O(N)$ arithmetic operations from the initial input (without using the explicit solution).

We propose the following easily generalizable solution: consider $v_n(x) = u_n(x+1)$. Then the ODE for $v_n(x)$ is given by

$$\big((x+1)^2\partial_x^3 - 3(x+1)(n-1)\partial_x^2 + (2n-1)(n-1)\partial_x\big)v_n(x) = 0,$$

and for the coefficient sequence of $v_n(x) = \sum_{k\geq 0} d_{n,k}x^k$ we find

$$(k+1)(k+2)d_{n,k+2} - (k+1)(3N-2k-1)d_{n,k+1} + (2n-k)(n-k)d_{n,k} = 0.$$

Now the leading coefficient of the recurrence is $(k+1)(k+2) \neq 0$, so we can easily unroll it after determining the first two terms, by computing them via binary powering of the corresponding companion matrix mod $x^2$. Having computed $v_N(x)$, it remains to find $u_N(x) = v_N(x-1)$. Note that expanding the polynomial results in an $O(\mathsf{M}(N))$ algorithm. However, recall from (6.19) that we only need to compute $c_{N,N}$ and $c_{N,2N}$, or, in other words, the coefficients of $x^N$ and $x^{2N}$ in $v_N(x-1)$. For any $i$ it holds that

$$c_{N,i} = \sum_{k\geq 0} d_{N,k}\binom{k}{i}(-1)^{k-i}, \tag{6.20}$$

and the sum is finite because $v_N(x)$ is a polynomial. Clearly, it can be computed in complexity $O(N)$ for any $i$.

Generally speaking, an issue with unrolling the recurrence for $(c_{n,k})_{k\geq 0}$ occurs if the roots of the leading polynomial are positive integers that depend on $n$. Let $S$ be the set of these roots. Note that the size of $S$ is independent of $n$ and $S$ can be non-empty only if the ODE for $u_n(x)$ is singular at 0 (that is, if $q_\ell(x)$ in (6.4) vanishes at $0$). In this case, one can always define $v_n(x) = u_n(x+c)$ for $c \in \mathbb{K}$ a non-singular point of the ODE ($q_\ell(c) \neq 0$). Then the coefficients $d_{n,k}$ of $v_n(x)$ can be computed from $O(1)$ initial conditions via unrolling a recurrence. Using the formula (6.20) (with $-c$ instead $-1$) and the fact that $v_n(x)$ is a polynomial, one can compute the coefficients $c_{N,i}$ for $i \in S$. With these, it is then possible to unroll the recurrence for $(c_{N,k})_{k\geq 0}$ and find $u_N(x)$ in complexity $O(N)$.

## 6.4 Impact on polynomial matrix power

Here is an algorithm for PolMatPow using SeqTerm. Let $M(x)$ in $\mathbb{K}[x]^{r\times r}_{\leq d}$ and $p_{i,j,n}(x)$ be the $(i,j)$ entry of $M(x)^n$, for $n \geq 0$ and $i$ and $j$ in $\{1,\ldots,r\}$. The sequence $(p_{i,j,n}(x))_{n\geq 0}$ is polynomial C-recursive, with a recurrence given by the characteristic polynomial

$$\chi_M(x,y) \coloneqq \det(y\,I_r - M(x)) = y^r - c_{r-1}(x)y^{r-1} - \cdots - c_0(x).$$

That is, $p_{i,j,n+r}(x) = c_{r-1}(x)p_{i,j,n+r-1}(x) + \cdots + c_0(x)p_{i,j,n}(x)$ for all $n \geq 0$. Thus, to compute $M(x)^N$, it is enough to find $\chi_M(x,y)$ (in $O(1)$, i.e. independent of $N$), to compute the polynomials $p_{i,j,n}(x)$ for $1 \leq i,j \leq r$ and $0 \leq n < r$ (also in $O(1)$) and to return the entries $p_{i,j,N}(x)$ of $M(x)^N$ via SeqTerm. As such, this approach uses $r^2$ calls to SeqTerm, with total cost $O(N)$.

This section describes an algorithm for PolMatPow which uses only $r$ such calls, through a direct reduction to BivModPow (see Section 6.4.2). Our solution for BivModPow, via $r$ calls to SeqTerm, is presented in Section 6.4.1 and completes the proof of Theorem 6.1.

### 6.4.1 Computing bivariate modular powers

Let $\mathbb{L} = \mathbb{K}[x]$ and $P, Q \in \mathbb{L}[y]$. Assume that $P$, seen as a univariate polynomial in $y$ of degree $r$, is monic. For $N \in \mathbb{N}$, Euclidean division in $\mathbb{L}[y]$ ensures the existence of unique $S, R \in \mathbb{L}[y]$ such that $\deg_y(R) < r$ and $Q^N = SP + R$. The polynomial $R$ is $Q^N \bmod P$. Assume that $P$ and $Q$ are fixed, and let $d \coloneqq \deg_x(P)$ (which is thus in $O(1)$). Then, writing $R = \sum_{i=0}^{d-1} r_i(x)y^i$, it holds that $\deg_x r_i(x) = O(N)$. The efficient computation of $R$ when $Q = y$, given $P(x,y)$ and $N$, is the first step for proving BivModPow in Theorem 6.1.

We shall first illustrate the connection of SeqTerm and BivModPow by means of an example. Let $P(x,y) = y^2 - xy - 1$ and $Q(x,y) = y$, i.e., we are looking for $F_{n-1}(x), F_n(x) \in \mathbb{L}$ such that

$$y^n = S(x,y)(y^2 - xy - 1) + yF_n(x) + F_{n-1}(x), \tag{6.21}$$

for some polynomial $S(x,y) \in \mathbb{L}[y]$. Replace $y$ by $1/y$ in (6.21) and then multiply by $y^{n+1}/(1 - xy - y^2)$ to obtain

$$\frac{y}{1-xy-y^2} = Q(x,1/y)y^{n-1} + y^n\frac{F_n(x) + yF_{n-1}(x)}{1-xy-y^2}.$$

Now observe that $\deg_y(Q(x,1/y)y^{n-1}) \leq n-1$, hence by extracting the $n$th and $(n+1)$st coefficients,

$$F_n(x) = [y^n]\frac{y}{1-xy-y^2} \text{ and}$$
$$F_{n-1}(x) + xF_n(x) = [y^{n+1}]\frac{y}{1-xy-y^2}.$$

We conclude that $F_k(x)$ is the $k$th Fibonacci polynomial, for $k = n$ and $k = n-1$. In particular, each $F_k(x)$ satisfies a linear recurrence with constant polynomials and can be found in $O(k)$ by SeqTerm.

This strategy, outlined on an example, generalizes in the obvious way. Explicitly, we have the following lemma (see [74, Lem. 2]).

**Lemma 6.6.** *Let $P \in \mathbb{K}[x,y]$ and $r := \deg_y(P)$, with $P(x,0) \neq 0$ and reversal $\bar{P}(x,y) := y^r P(x, \frac{1}{y})$. Write $\frac{1}{\bar{P}(x,y)} =: \sum_{k\geq 0} u_k(x)y^k$. Finally, let $v(x,y)$ be the polynomial given by $(u_{N-r+1}(x) + \cdots + u_N(x)y^{r-1})\bar{P}(x,y) \bmod y^r$. Then $y^N \bmod P(x,y) = v(1/y)y^{r-1}$.*

The sequence $(u_k(x))_{k\geq 0}$ in Lemma 6.6 is C-recursive because its generating function is rational. Hence, using SeqTerm, the $r = O(1)$ many terms $u_{N-r+1}(x), \ldots, u_N(x)$ can be computed in complexity $O(N)$. It follows that the case $Q(x,y) = y$ of BivModPow can be solved in $O(N)$ steps as well.

Finally, the computation of $Q(x,y)^N \bmod P(x,y)$ can be reduced to $y^N \bmod P(x,y)$ with a resultant precomputation (see Lemma 6.7). This leads to Algorithm 9, which solves BivModPow in $O(N)$.

**Lemma 6.7.** *Let $P(y), Q(y) \in \mathbb{L}[y]$. Define $A(t), B(t) \in \mathbb{L}[t]$ by $A(t) = \mathrm{Res}_y(P(y), t - Q(y))$ and $B(t) = t^N \bmod A(t)$. Then*

$$Q(y)^N \bmod P(y) = B(Q(y)) \bmod P(y).$$

*Proof.* By the definition of the resultant, $A(t) = \prod_i (t - Q(a_i))$ where $a_i \in \overline{\mathbb{L}}$ are the solutions of $P(y) = 0$. Hence, $P(y)$ divides $A(Q(y))$, which, by construction, divides $B(Q(y)) - Q(y)^N$. □

**Algorithm 9** BivModPow$(P(x,y), Q(x,y), N)$

*Input:* $P(x,y), Q(x,y) \in \mathbb{L}[y]$ with $P(x,y)$ monic in $y$, and $N \in \mathbb{N}$.
*Output:* $Q(x,y)^N \bmod P(x,y)$.

1: $A(x,t) \leftarrow \mathrm{Res}_y(P(x,y), t - Q(x,y))$
2: $\bar{A}(x,t) \leftarrow t^r A(x, 1/t)$, where $r := \deg_t A(x,t)$
3: **for** $i = N - r + 1, \ldots, N$ **do**
4: $\quad u_i(x) \leftarrow [t^i] \frac{1}{\bar{A}(x,t)}$ using SeqTerm
5: **end for**
6: $u(x,t) \leftarrow u_{N-r+1}(x) + \cdots + u_N(x)t^{r-1}$
7: $v(x,t) \leftarrow u(x,t)\bar{A}(x,t) \bmod t^r$; $B(x,t) \leftarrow v(1/t)t^{r-1}$
8: **return** $B(x, Q(x,y)) \bmod P(x,y)$

### 6.4.2 Computing polynomial matrix powers

Let $M \in \mathbb{K}[x]^{r\times r}$ be an $r \times r$ polynomial matrix of degree at most $d$. Its power $M^N$ has degree at most $Nd \in O(N)$. Let $P(x,y)$ be the characteristic polynomial of $M$. Since $P(x,M) = 0$ by the Cayley-Hamilton theorem, we get $M^N = R(x,M)$ where $R(x,y) = y^N \bmod P(x,y)$. The polynomial $R$ can be computed in $O(N)$ via BivModPow. Then evaluating $R(x,y)$ at $y = M(x)$ has cost $O(N)$ because $\deg_x(R) \in O(N)$, $\deg_y(R) = r \in O(1)$ and $\deg_x(M) = d \in O(1)$. Hence Algorithm 10 is correct and has complexity $O(N)$.

## 6.5 Experiments

The main precomputation step for all our algorithms consists in starting with a rational function $U \in \mathbb{K}(x,y) \cap \mathbb{K}[\![x,y]\!]$ and in finding a differential operator $L_n$ that annihilates

**Algorithm 10** PolMatPow$(M, N)$

*Input:* matrix $M(x) \in \mathbb{K}[x]^{r\times r}$, integer $N \in \mathbb{N}$.
*Output:* $M(x)^N \in \mathbb{K}[x]^{r\times r}$.

1: $P(x, y) \leftarrow$ the characteristic polynomial of $M(x)$
2: $R(x, y) \leftarrow y^N \bmod P(x, y)$ ▷ instance of BivModPow
3: **return** $R(x, M(x))$

$u_n(x) = [y^n]U(x, y)$ and whose degree and order are independent of $n$. For this task, in practice, we may either use the method described in §6.3.1, or creative telescoping algorithms for hyperexponential functions. Table 6.1 summarizes timings for a variety of implementations.

The table reveals that, among these implementations, the fastest one for computing a telescoper of $U(x, y)/y^{n+1}$ is the reduction-based creative telescoping in Maple. More specifically, `redct` is the fastest, followed by HT. The implementation in *ore_algebra* [198] in SageMath competes best with reduction-based methods.

Table 6.2 gives timings of an efficient implementation of the remaining stages after precomputations: computing initial terms (IT), and unrolling (UR). We observe that IT takes negligible time compared to UR, except for extreme parameter ranges where, simultaneously, $r$ and $d$ are large and $N$ is small; this is expected since these ranges correspond to cases where the order of the recurrence to be unrolled is close to $N$. We also see that binary powering is always slower, often by a factor more than $5$, than the addition of IT and UR. The speed-up factor is summarized in Figs. 6.1 to 6.3; as expected it grows when $N$ grows, with $r$ and $d$ fixed.

For large $N$, in most of the reported cases, performing both the precomputation and IT+UR is much faster than using binary powering. Still, this is not always true, e.g. for $r = 5$. One has to keep in mind that `redct` is not implemented in low-level Maple, and targets rational coefficients: for a more meaningful assessment of the precomputation part, it would be interesting to have an implementation of creative telescoping which is fully optimized and specialized to coefficients in a word-size prime field.

## 6.6 Conclusion and future work

We have shown that it is possible to beat, both in theory and in practice, the very basic and powerful binary powering method for computing: (i) powers of polynomial matrices, (ii) terms in polynomial C-finite sequences and (iii) modular exponentiation for bivariate polynomials. We describe below several lines of work, including possible optimizations and generalizations, that we leave for future investigations.

**More detailed complexity analysis.** The first and most natural direction for future work is to analyze and improve the complexity of the algorithms in Theorem 6.1 with respect to the parameters $r$ and $d$. For simplicity, these parameters were assumed to be $O(1)$ in this chapter. For the $N$th power of a polynomial matrix $M(x)$ of size $r$ and degree $d$, binary splitting has arithmetic complexity $O(\mathsf{M}(Nd)r^2 + Ndr^\omega)$, where $\omega \in [2, 3]$ is a feasible exponent of matrix multiplication over $\mathbb{K}$. With our approach, it is legitimate to target

a differential equation satisfied by the entries of $M(x)^N$ of order $r$ with coefficients in $x$ of degree $O(dr^3)$, yielding a recurrence of order $O(dr^3)$ and coefficients in $n$ of degree at most $r$. For large $N$, the complexity of the new algorithm would thus be $O(Ndr^2\mathsf{M}(r))$. Using different ODEs, of order $O(r)$ and coefficients of degree $O(dr^2)$ could even lead to a complexity $O(Ndr\mathsf{M}(r))$.

**The $K$th coefficient of the $N$th term** For some (large) integers $N, K \in \mathbb{N}$, one might be interested in computing the single coefficient $[x^K y^N]U(x,y)$ of a rational function $U \in \mathbb{K}(x,y) \cap \mathbb{K}[\![x,y]\!]$. Equivalently it is natural to wonder: How fast can one compute the $K$th coefficient of the $N$th term of a C-recursive sequence $(u_n(x))_{n\geq 0}$? Using our method, a recurrence with initial conditions for the coefficients of $u_N(x)$ can be deduced in $O(\log N)$ arithmetic operations. Then (assuming that the recurrence is non-singular) the $K$th coefficient can be found in $O(\mathsf{M}(\sqrt{K}))$ operations by using baby-steps/giant-steps [106, 69]. We expect that, at least under a genericity assumption, this problem can be solved in arithmetic complexity $O(\log(N)+\mathsf{M}(\sqrt{K}))$ which is a big improvement compared to the previous best $O(N+K)$ by [238].

**Polynomial P-recursive sequences** A somewhat related task is to study the analogous problem to SeqTerm for polynomial P-recursive sequences, that is for $(u_n(x))_n \in \mathbb{K}[x]^{\mathbb{N}}$ satisfying

$$p_r(x,n)u_{n+r}(x) + \cdots + p_0(x,n)u_n(x) = 0,$$

for $p_i(x,n) \in \mathbb{K}[x,n]$. We expect that, at least under a genericity assumption, a generalization of Lemma 6.5 (based on results in [64]) should exist, implying in particular that $u_N(x)$ satisfies a linear ODE of order and degree independent of $N$. Generalizing this even further, one might study the Creative Telescoping problem for rational functions of the form $H(\mathbf{x}) = \frac{P(x_1,\ldots,x_s)}{Q(x_1,\ldots,x_s)R(x_1,\ldots,x_s)^n}$. We expect that (at least generically) the minimal telecoper for $H(\mathbf{x})$ has order and degree independent of $n$ and can be found via a Griffiths-Dwork reduction type approach, based on ideas from [71].

**Connection to the Jordan–Chevalley decomposition** A different approach for computing powers of matrices uses the *Jordan–Chevalley decomposition* (also called *SN decomposition*), see e.g. [184, 131, 281, 117]. It ensures that any polynomial matrix $M \in \mathbb{K}[x]^{r\times r}$ can be written as $M = S + Z$, where $S \in \mathbb{K}(x)^{r\times r}$ is diagonalizable over $\overline{\mathbb{K}(x)}$, $Z \in \mathbb{K}(x)^{r\times r}$ is nilpotent, and $SZ = ZS$. From this decomposition it follows that $M^N = \sum_{i=0}^{r-1} \binom{N}{i} S^{N-i}Z^i$. After a change of basis, this reduces to computing a power of a diagonal matrix with algebraic functions coefficients. Using Lemma 6.3 this can be performed efficiently in $O(N)$ operations. It would be certainly interesting to compare this approach with the other methods.

**A PDE approach for SeqTerm** There is yet another method to deduce recurrence (6.9). The starting point is that the generating function $F(x,y) = y/(1-xy-y^2)$ of $F_n(x)$ satisfies the linear PDE

$$(x^2+4)\frac{\partial^2 F}{\partial x^2} + 3x\frac{\partial F}{\partial x} - y^2\frac{\partial^2 F}{\partial y^2} - y\frac{\partial F}{\partial y} + F = 0, \tag{6.22}$$

and extracting the coefficient of $x^k y^n$ in (6.22) immediately gives (6.9). More generally, such a PDE translates into a recurrence if it is linear with polynomial coefficients in $x$ and $y$ and if additionally all terms of the form $x^i y^\ell \frac{\partial^k F}{\partial x^k} \frac{\partial^j F}{\partial y^j}$ have $\ell = j$. A dimension counting argument in the spirit of [229, Lem. 3] proves that such a PDE exists for *any* rational function $F(x, y)$. The existence proof is effective and amounts to linear algebra. A natural question is whether it is possible to compute such a PDE via Creative Telescoping (either Almkvist-Zeilberger [16] or reduction based [59]), and how the corresponding method compares to the aforementioned ones.

Table 6.1 Timings in seconds for creative telescoping to find a telescoper $L_n$ of $P(x, y)/(y^{n+1}Q(x, y))$. Here $P(x, y)/Q(x, y)$ is the generating function for the sequence of the top-right entry of the powers of a randomly chosen matrix in $\mathbb{F}_p[x]^{r\times r}$ of degree $d$, for a 50-bit prime $p$, with $Q(x, y)$ the $y$-reversal of the characteristic polynomial of this matrix. The order of $L_n$ is $\ell$, its degree in $n$ is $\mathrm{d}_n$, and $\mathrm{d}_x = \deg_x(L_n)$. A blank space means that the computation took more than 1000 seconds. We observe empirically that the degree in $x$ is $dr(r + 1)(2r - 1)/2 - r(r - 1)$ while its degree in $n$ is $(r - 1)(r + 2)/2$; this is expected asymptotically by [59, Thm. 25] and Lemma 6.5, because $\deg_y Q(x, y) = r$ and $\deg_x Q(x, y) = dr$. The tested implementations are

- in Maple: `redct` [64]; `HermiteTelescoping` (HT) [71]; `Zeilberger` (ZB) [16] in DEtools; `creative_telescoping` (c_t) [111];
- in SageMath: creative telescoping (`ct`) from the *ore_algebra* package [198];
- In Mathematica: `FindCreativeTelescoping` (FCT), `CreativeTelescoping` (CT), and `HermiteTelescoping` (HCT), see [213].

| | | Maple | | | | Sage | Mathematica | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $d$ | redct | HT | ZB | c_t | ct | FCT | CT | HCT | $\ell$ | $\mathrm{d}_n$ | $\mathrm{d}_x$ |
| 2 | 2 | 0.0 | 0.1 | 0.0 | 0.1 | 0.5 | 0.2 | 0.2 | 0.2 | 2 | 2 | 16 |
| | 4 | 0.0 | 0.0 | 0.0 | 0.1 | 0.6 | 0.4 | 0.4 | 0.3 | 2 | 2 | 34 |
| | 6 | 0.0 | 0.0 | 0.0 | 0.1 | 0.6 | 0.7 | 0.5 | 0.5 | 2 | 2 | 52 |
| | 8 | 0.0 | 0.0 | 0.0 | 0.1 | 0.8 | 1.0 | 0.7 | 0.7 | 2 | 2 | 70 |
| 3 | 1 | 0.0 | 0.2 | 0.0 | 0.5 | 2.0 | 2.0 | 1.3 | 1.3 | 3 | 5 | 24 |
| | 2 | 0.0 | 0.1 | 0.8 | 3.4 | 3.1 | 4.0 | 2.6 | 2.5 | 3 | 5 | 54 |
| | 3 | 0.1 | 0.2 | 0.8 | 9.3 | 5.6 | 10 | 5.7 | 5.4 | 3 | 5 | 84 |
| | 4 | 0.1 | 0.5 | 18 | 19 | 8.2 | 17 | 9.4 | 8.9 | 3 | 5 | 114 |
| | 5 | 0.2 | 1.1 | 5.1 | 32 | 12 | 25 | 14 | 14 | 3 | 5 | 144 |
| | 6 | 0.5 | 1.7 | 9.8 | 49 | 17 | 35 | 19 | 20 | 3 | 5 | 174 |
| 4 | 1 | 0.4 | 2.9 | 23 | 117 | 20 | 31 | 25 | 25 | 4 | 9 | 58 |
| | 2 | 1.7 | 17 | 410 | 749 | 45 | 101 | 96 | 95 | 4 | 9 | 128 |
| | 3 | 4.4 | 43 | | | 89 | 295 | 376 | 373 | 4 | 9 | 198 |
| | 4 | 12 | 82 | | | 172 | 388 | 752 | 693 | 4 | 9 | 268 |
| | 5 | 18 | 128 | | | 280 | 635 | | | 4 | 9 | 338 |
| 5 | 1 | 11 | 34 | 538 | | 163 | 847 | 780 | | 5 | 14 | 115 |
| | 2 | 64 | 183 | | | 515 | | | | 5 | 14 | 250 |
| | 3 | 159 | 526 | | | | | | | 5 | 14 | 385 |
| | 4 | 345 | | | | | | | | 5 | 14 | 520 |

Table 6.2 Timings in seconds, using the C++ library NTL [294] and PML [186], for computing the top-right entry of the $N$th power of a randomly chosen matrix in $\mathbb{F}_p[x]^{r\times r}$ of degree $d$, for a 50-bit prime $p$. The first tested method is to directly apply binary powering (BP); in the present context, the polynomial matrix multiplication of PML is based on evaluation-interpolation and 3-prime FFT. The second tested method uses Algorithm 8 and we do not count "precomputations" (already showed in Table 6.1), i.e. we only report timings for the two non-negligible steps that depend on $N$, namely Line 8 (UR, unrolling) and Line 7 (IT, initial terms) from Algorithm 6.

| | | $N=2^{10}$ | | | $N=2^{12}$ | | | $N=2^{14}$ | | | $N=2^{16}$ | | | $N=2^{18}$ | | | $N=2^{20}$ | | | $N=2^{22}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | $d$ | BP | UR | IT | BP | UR | IT | BP | UR | IT | BP | UR | IT | BP | UR | IT | BP | UR | IT | BP | UR | IT |
| 2 | 2 | 1.2e-3 | 5.7e-4 | 3.7e-5 | 5.3e-3 | 2.4e-3 | 4.3e-5 | 2.5e-2 | 9.7e-3 | 4.9e-5 | 1.1e-1 | 3.9e-2 | 5.5e-5 | 5.3e-1 | 1.5e-1 | 6.2e-5 | 3.3e+0 | 6.2e-1 | 6.7e-5 | 1.5e+1 | 2.5e+0 | 7.5e-5 |
| | 4 | 2.6e-3 | 1.3e-3 | 7.8e-5 | 1.2e-2 | 5.2e-3 | 9.4e-5 | 5.2e-2 | 2.1e-2 | 1.1e-4 | 2.4e-1 | 8.4e-2 | 1.3e-4 | 1.4e+0 | 3.4e-1 | 1.4e-4 | 7.2e+0 | 1.4e+0 | 1.6e-4 | 3.1e+1 | 5.4e+0 | 1.8e-4 |
| | 6 | 3.8e-3 | 2.1e-3 | 1.2e-4 | 1.7e-2 | 8.7e-3 | 1.5e-4 | 7.9e-2 | 3.5e-2 | 1.8e-4 | 3.6e-1 | 1.4e-1 | 2.1e-4 | 2.3e+0 | 5.5e-1 | 2.4e-4 | 1.0e+1 | 2.2e+0 | 2.7e-4 | 4.6e+1 | 8.9e+0 | 3.0e-4 |
| | 8 | 5.3e-3 | 3.1e-3 | 1.9e-4 | 2.4e-2 | 1.2e-2 | 2.4e-4 | 1.1e-1 | 5.0e-2 | 2.8e-4 | 5.3e-1 | 2.0e-1 | 3.3e-4 | 3.3e+0 | 8.0e-1 | 3.8e-4 | 1.5e+1 | 3.2e+0 | 4.3e-4 | 7.0e+1 | 1.2e+1 | 4.9e-4 |
| 3 | 1 | 1.4e-3 | 3.0e-4 | 1.3e-4 | 6.0e-3 | 1.3e-3 | 1.7e-4 | 2.6e-2 | 5.5e-3 | 2.1e-4 | 1.2e-1 | 2.2e-2 | 2.4e-4 | 5.8e-1 | 8.8e-2 | 2.8e-4 | 3.4e+0 | 3.5e-1 | 3.1e-4 | 1.6e+1 | 1.4e+0 | 3.5e-4 |
| | 2 | 2.9e-3 | 7.8e-4 | 4.0e-4 | 1.2e-2 | 3.2e-3 | 5.3e-4 | 5.6e-2 | 1.3e-2 | 6.5e-4 | 2.6e-1 | 5.2e-2 | 7.8e-4 | 1.5e+0 | 2.1e-1 | 9.1e-4 | 7.6e+0 | 8.4e-1 | 1.0e-3 | 3.4e+1 | 3.3e+0 | 1.2e-3 |
| | 3 | 4.3e-3 | 1.4e-3 | 7.4e-4 | 1.9e-2 | 5.8e-3 | 9.9e-4 | 8.4e-2 | 2.3e-2 | 1.2e-3 | 3.9e-1 | 9.3e-2 | 1.5e-3 | 2.2e+0 | 3.7e-1 | 1.7e-3 | 1.1e+1 | 1.5e+0 | 2.0e-3 | 4.9e+1 | 6.0e+0 | 2.2e-3 |
| | 4 | 6.0e-3 | 2.1e-3 | 8.0e-4 | 2.6e-2 | 8.8e-3 | 1.0e-3 | 1.2e-1 | 3.5e-2 | 1.3e-3 | 5.8e-1 | 1.4e-1 | 1.5e-3 | 3.5e+0 | 5.7e-1 | 1.8e-3 | 1.7e+1 | 2.3e+0 | 2.0e-3 | 7.1e+1 | 9.1e+0 | 2.3e-3 |
| | 5 | 7.4e-3 | 3.0e-3 | 1.0e-3 | 3.3e-2 | 1.2e-2 | 1.3e-3 | 1.5e-1 | 5.0e-2 | 1.7e-3 | 7.2e-1 | 2.0e-1 | 2.0e-3 | 4.3e+0 | 7.9e-1 | 2.3e-3 | 2.0e+1 | 3.2e+0 | 2.6e-3 | 8.8e+1 | 1.3e+1 | 2.9e-3 |
| | 6 | 9.1e-3 | 4.0e-3 | 1.2e-3 | 4.0e-2 | 1.6e-2 | 1.6e-3 | 1.8e-1 | 6.6e-2 | 1.9e-3 | 8.2e-1 | 2.7e-1 | 2.3e-3 | 5.3e+0 | 1.1e+0 | 2.7e-3 | 2.3e+1 | 4.2e+0 | 3.1e-3 | 1.1e+2 | 1.7e+1 | 3.4e-3 |
| 4 | 1 | 2.7e-3 | 4.2e-4 | 7.8e-4 | 1.1e-2 | 1.8e-3 | 1.1e-3 | 4.9e-2 | 7.5e-3 | 1.4e-3 | 2.2e-1 | 3.0e-2 | 1.7e-3 | 1.1e+0 | 1.2e-1 | 2.0e-3 | 6.2e+0 | 4.8e-1 | 2.3e-3 | 2.9e+1 | 1.9e+0 | 2.6e-3 |
| | 2 | 5.5e-3 | 1.2e-3 | 1.3e-3 | 2.4e-2 | 5.2e-3 | 1.8e-3 | 1.1e-1 | 2.1e-2 | 2.3e-3 | 4.9e-1 | 8.6e-2 | 2.8e-3 | 2.8e+0 | 3.4e-1 | 3.2e-3 | 1.4e+1 | 1.4e+0 | 3.7e-3 | 6.2e+1 | 5.5e+0 | 4.2e-3 |
| | 3 | 8.2e-3 | 2.4e-3 | 2.1e-3 | 3.6e-2 | 1.0e-2 | 2.9e-3 | 1.6e-1 | 4.2e-2 | 3.7e-3 | 7.3e-1 | 1.7e-1 | 4.5e-3 | 4.4e+0 | 6.7e-1 | 5.3e-3 | 2.1e+1 | 2.7e+0 | 6.1e-3 | 9.3e+1 | 1.1e+1 | 6.9e-3 |
| | 4 | 1.1e-2 | 4.1e-3 | 3.0e-3 | 5.0e-2 | 1.7e-2 | 4.1e-3 | 2.3e-1 | 6.9e-2 | 5.2e-3 | 1.1e+0 | 2.8e-1 | 6.4e-3 | 6.6e+0 | 1.1e+0 | 7.5e-3 | 3.1e+1 | 4.5e+0 | 8.6e-3 | 1.3e+2 | 1.8e+1 | 9.7e-3 |
| | 5 | 1.4e-2 | 6.0e-3 | 3.7e-3 | 6.3e-2 | 2.5e-2 | 5.2e-3 | 2.8e-1 | 1.0e-1 | 6.6e-3 | 1.3e+0 | 4.1e-1 | 8.0e-3 | 7.7e+0 | 1.6e+0 | 9.4e-3 | 3.6e+1 | 6.5e+0 | 1.1e-2 | 1.6e+2 | 2.6e+1 | 1.2e-2 |
| 5 | 1 | 4.4e-3 | 6.0e-4 | 1.8e-3 | 1.8e-2 | 2.7e-3 | 2.5e-3 | 8.2e-2 | 1.1e-2 | 3.3e-3 | 3.7e-1 | 4.5e-2 | 4.1e-3 | 1.7e+0 | 1.8e-1 | 4.9e-3 | 1.0e+1 | 7.3e-1 | 5.7e-3 | 4.7e+1 | 2.9e+0 | 6.5e-3 |
| | 2 | 9.1e-3 | 2.0e-3 | 3.5e-3 | 3.9e-2 | 9.1e-3 | 5.1e-3 | 1.8e-1 | 3.7e-2 | 6.7e-3 | 8.1e-1 | 1.5e-1 | 8.3e-3 | 4.6e+0 | 6.0e-1 | 9.9e-3 | 2.3e+1 | 2.4e+0 | 1.2e-2 | 1.0e+2 | 9.6e+0 | 1.3e-2 |
| | 3 | 1.3e-2 | 4.3e-3 | 5.7e-3 | 5.8e-2 | 1.9e-2 | 8.3e-3 | 2.6e-1 | 7.8e-2 | 1.1e-2 | 1.2e+0 | 3.2e-1 | 1.4e-2 | 7.1e+0 | 1.3e+0 | 1.6e-2 | 3.4e+1 | 5.1e+0 | 1.9e-2 | 1.5e+2 | 2.0e+1 | 2.2e-2 |
| | 4 | 1.8e-2 | 7.4e-3 | 7.8e-3 | 8.0e-2 | 3.3e-2 | 1.2e-2 | 3.8e-1 | 1.3e-1 | 1.5e-2 | 1.8e+0 | 5.4e-1 | 1.9e-2 | 1.1e+1 | 2.2e+0 | 2.3e-2 | 4.9e+1 | 8.7e+0 | 2.6e-2 | 2.1e+2 | 3.5e+1 | 3.0e-2 |

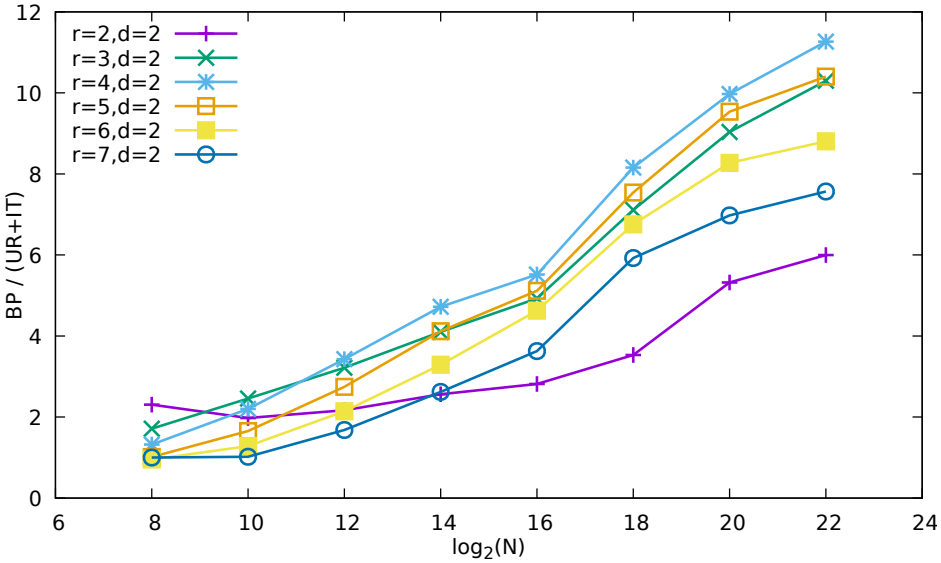



Figure 6.1: Speed-up versus binary powering, not counting precomputations, for $r = 2\ldots7$, $N = 2^8, 2^{10}, \ldots, 2^{22}$, and fixed $d = 2$.

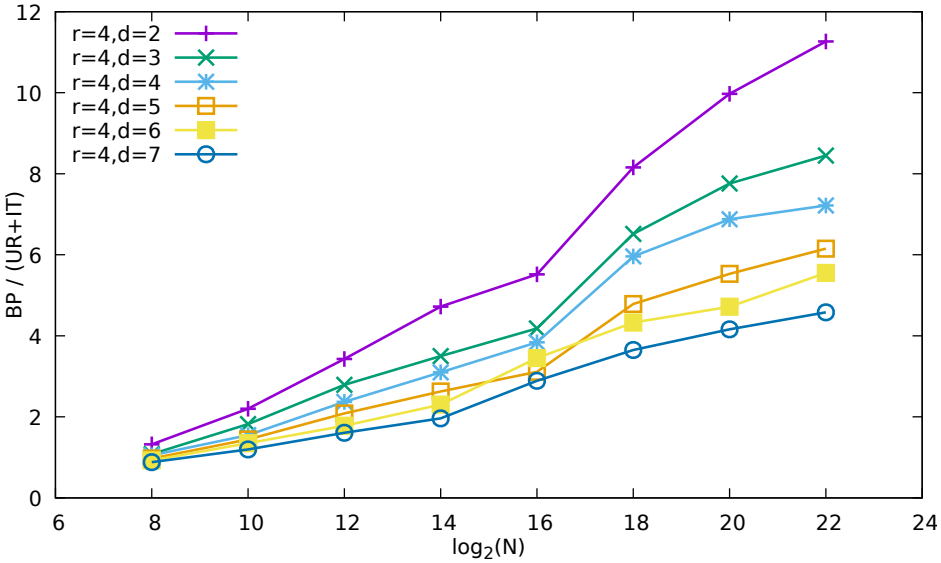



Figure 6.2: Speed-up versus binary powering, not counting precomputations, for $d = 2 \ldots 7$, $N = 2^8, 2^{10}, \ldots, 2^{22}$, and fixed $r = 4$.

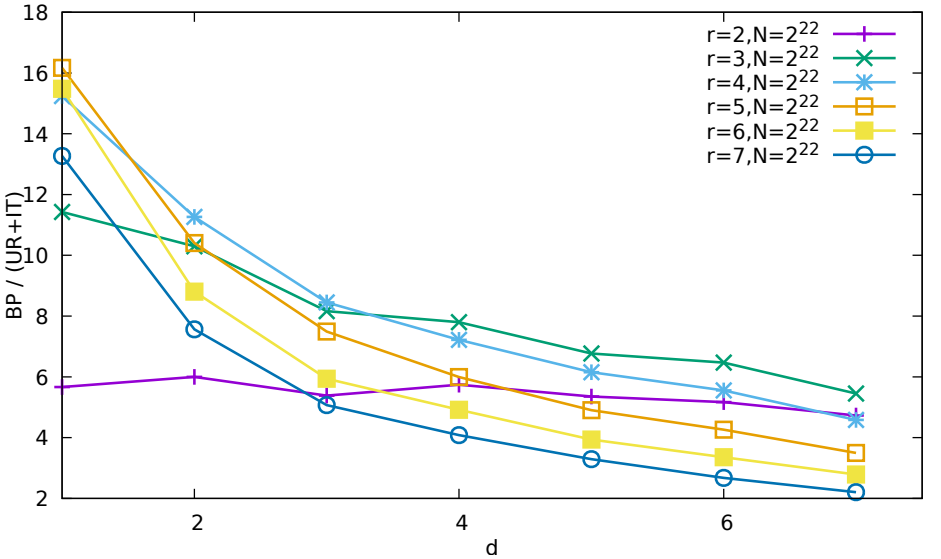



Figure 6.3: Speed-up versus binary powering, not counting precomputations, for $r = 2 \ldots 7$, $d = 1 \ldots 7$, and fixed $N = 2^{22}$.

# Chapter 7

# On the $q$-analogue of Pólya's Theorem

*"Finished mathematics consists of proofs,*
*But mathematics in the making consists of guesses."*

*"In mathematics often the simplest is the best."*
George Pólya, *Video: Teaching us a lesson*, 1966.

In this chapter we answer a question posed by Michael Aissen in 1979 about the $q$-analogue of a classical theorem of George Pólya (1922) on the algebraicity of (generalized) diagonals of bivariate rational power series. In particular, we prove that the answer to Aissen's question, in which he considers $q$ as a variable, is negative in general. Moreover, we show that when $q$ is a complex number, the answer is positive if and only if $q$ is a root of unity.

This part consists of joint work with A. Bostan [84].

## 7.1 Introduction

A beautiful but rather unknown theorem of Pólya [267] states the following:

> Given two algebraic power series[1] $\varphi(x)$ and $\Phi(x)$, let $A_{(i,j)}$ be the coefficient of $x^j$ in $\Phi(x)\varphi(x)^i$. Consider a straight line in the plane and let $(p_n)_{n\geq 0}$ be the sequence of non-negative integer lattice points in $\mathbb{Z}^2$ lying on this line. Then $F(x) = \sum_{n\geq 0} A_{p_n} x^n$ is algebraic.

In particular, this theorem implies that the generalized diagonal $\Delta_{a,b}$ of a bivariate rational power series is algebraic, where for $f(x,y) = \sum_{i,j\geq 0} f_{i,j} x^i y^j$, we define $\Delta_{a,b}(f) := \sum_{n\geq 0} f_{an,bn} x^n$. For example, one finds

$$\Delta_{1,1}\left(\frac{1}{1-x-y}\right) = \Delta_{1,1}\left(\sum_{i,j\geq 0}\binom{i+j}{i} x^i y^j\right) = \sum_{n\geq 0}\binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}},$$

and the latter is a root of $P(x,z) = (1-4x)z^2 - 1$.

[1] Recall that $f(x) \in \mathbb{C}[\![x]\!]$ is called *algebraic* if there exists a bivariate non-zero polynomial $P(x,z)$ in $\mathbb{C}[x,z]$ such that $P(x,f(x)) = 0$. A non-algebraic series is called *transcendental*.

In Pólya's formulation, this example is obtained by choosing $\Phi(x) = 1$, $\varphi(x) = 1 + x$ and the main diagonal $\{x = y\}$ of $\mathbb{Z}^2$. In fact, this special case is the main foundation for the observation and question which led to the article [84] and this chapter.

There is a more combinatorial rephrasing of this example. Arrange Pascal's triangle in the following way

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0 | 1 | | | | | | |
| 1 | 1 | 1 | | | | | |
| 2 | 1 | 2 | 1 | | | | |
| 3 | 1 | 3 | 3 | 1 | | | |
| 4 | 1 | 4 | 6 | 4 | 1 | | |
| 5 | 1 | 5 | 10 | 10 | 5 | 1 | |
| 6 | 1 | 6 | 15 | 20 | 15 | 6 | 1 |
| ⋮ | | | … | … | | | |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 |

and consider a line passing through infinitely many lattice points of the above triangle. If we denote the resulting sequence of values on these lattice points by $(u_j)_{j\geq 0}$, then Pólya's theorem ensures that the generating function $f(x) = \sum_{j\geq 0} u_j x^j$ is algebraic.

It is easy to see that the above condition on the line can be reformulated into the existence of non-negative integers $n, k, a, b$ with $n \geq k, a \geq b$, $\gcd(a, b) = 1$ and that either $n - a < k - b$ or $k - b < 0$, such that $u_j = \binom{n+aj}{k+bj}$. So this special case of Pólya's theorem simply asserts that

$$\sum_{j\geq 0} \binom{n+aj}{k+bj} x^j \in \mathbb{C}[\![x]\!] \quad \text{is algebraic.}$$

It is well-known and easy to see that the binomial coefficient $\binom{x+y}{x}$ counts lattice paths from the origin to $(x, y) \in \mathbb{Z}^2$ with only North and East steps. The observation above therefore implies that the generating function of the number of such paths, as $(x, y)$ increases on a line, is algebraic. Now recall the definition of the $q$-analogue of the binomial coefficient, called the *$q$-binomial* coefficient:

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{[n]_q!}{[k]_q![n-k]_q!}, \qquad \text{where}$$
$$[n]_q! := (1+q)\cdots(1+q+\cdots+q^{n-1}),$$

for a variable $q$ and integers $n, k$ with $0 \leq k \leq n$. It is not difficult to check that $\begin{bmatrix} n \\ k \end{bmatrix}_q \in \mathbb{Z}[q]$ is a polynomial in $q$ of degree $k(n-k)$. Arranged as in the figure above, these $q$-binomial coefficients give rise to the so-called $q$-Pascal triangle.

Pólya showed [268] that the coefficient of $q^j$ in $\begin{bmatrix} x+y \\ x \end{bmatrix}_q$ counts lattice paths in $\mathbb{Z}^2$ from the origin to $(x, y)$ with same steps as before and with area underneath equal to $j$, see also [11]. Aissen asked in [91, p. 585] the natural question whether the following $q$-analogue of Pólya's statement about algebraicity of such path generating functions holds:

> Fix integers $n, k, a, b$ with $n \geq k \geq 0, a \geq b \geq 0$ and $\gcd(a, b) = 1$. Moreover assume that either $n - a < k - b$ or $k - b < 0$. Let
>
> $$F(x, q) := \sum_{j\geq 0} \begin{bmatrix} n+aj \\ k+bj \end{bmatrix}_q x^j \in \mathbb{C}[q][\![x]\!].$$

Is the power series $F(x,q)$ algebraic? That is, does there exist a non-zero polynomial $P(x,q,z) \in \mathbb{C}[x,q,z]$ such that $P(x,q,F(x,q)) = 0$?

The inequality conditions $n \geq k \geq 0$ and $a \geq b \geq 0$ ensure that $F(x,q)$ is well-defined and not a polynomial. The condition $\gcd(a,b) = 1$ means that the line passing through the $q$-Pascal triangle does not "skip" terms; as we will see, it does not affect the algebraicity of $F(x,q)$. Also the condition $n - a < k - b$ or $k - b < 0$ is just a translation of the geometric picture that $F(x,q)$ collects all terms on that line. Aissen noticed the following fact: if the line is parallel to an edge of the $q$-Pascal triangle (i.e., if $a = b$ or $b = 0$), then $F(x,q)$ is trivially algebraic, because it is actually a rational function of $x$ and $q$. Hence, in what follows, we will assume that $a \neq b$ and $b \neq 0$. More precisely, we will only consider *admissible integers* $n, k, a, b$, in the following sense: $n \geq k \geq 0$, $a > b > 0$, $\gcd(a,b) = 1$, and either $n - k < a - b$ or $k < b$.

## 7.2 Results

Using elementary asymptotic estimates on the coefficients, we can show that the answer to Aissen's question is negative, because for any fixed $z \in \mathbb{C}$ of absolute value larger than $1$ the coefficients of $F(x,z)$ grow too fast for this series to be algebraic. However, we notice that the same argument does not apply for the univariate power series $h_z(x) := F(x,z)$, where $z \in \mathbb{C}$ with $|z| \leq 1$. In fact, our main result (Theorem 7.4 below) is that $h_z(x)$ is algebraic if and only if $z$ is a root of unity.

We start with a particular case of our main result and the answer to Aissen's question: the generating function of the central $q$-binomial coefficients is not algebraic.

**Proposition 7.1.** *For $n = k = 0$ and $a = 2, b = 1$ the series $F(x,q)$ is not algebraic.*

*Proof.* Assume that $F(x,q)$ is algebraic with minimal polynomial $P(x,q,z)$. Then the series $h_2(x) := F(x,2)$ must be algebraic as well, since $P_2(x,z) := P(x,2,z) \not\equiv 0$ satisfies $P_2(x, h_2(x)) = 0$. We have

$$\begin{aligned} h_2(x) &= \sum_{j \geq 0} \frac{(2^{j+1}-1)(2^{j+2}-1)\cdots(2^{2j}-1)}{(2-1)(2^2-1)\cdots(2^j-1)} x^j \\ &= 1 + 3x + 35x^2 + 1395x^3 + 200787x^4 + \cdots . \end{aligned}$$

Using the obvious inequality $(2^{j+k}-1)/(2^k-1) > 2^j$, we see that the $j$-th coefficient of $h_2(x)$ is greater than $2^{j^2}$. This growth rate is too fast for $h_2(x)$ to be algebraic, see e.g., Theorem D in [137], or Theorem 3 in [273]. □

A more elementary way to see that the growth rate of the coefficients of $h_2(x)$ is incompatible with algebraicity of the function is to notice that this rate is too fast even for D-finite functions. Recall [299] that a power series $f(x) = \sum_{i \geq 0} u_i x^i$ is called *D-finite* if it satisfies a linear differential equation with polynomial coefficients:

$$p_n(x) f^{(n)}(x) + \cdots + p_0(x) f(x) = 0.$$

A classical theorem ensures that any algebraic function is D-finite [299, Thm. 2.1] (see also [65]), whereas the latter class of functions is clearly much larger. An equivalent characterization of D-finite series [299, Thm. 1.5] states that the coefficients sequence satisfies a linear recurrence with polynomial coefficients:

$$u_{j+r}c_r(j) + \cdots + u_j c_0(j) = 0, \quad j \geq 0.$$

A sequence $(u_j)_{j\geq 0}$ is called *P-recursive* if it satisfies a recurrence as above.

A simple estimation on the growth rate of such sequences shows that any P-recursive sequence $(u_j)_{j\geq 0}$ grows at most like a power of $j!$ which is slower than $2^{j^2}$. We will use the fact that the coefficient sequence of an algebraic function is necessarily P-recursive again later.

We have just proved that the bivariate series $F(x, q)$ cannot be algebraic for all admissible $n, k, a, b$. In the same manner, we can prove that $F(x, q)$ is not algebraic for *any* admissible $n, k, a, b$. This is easily reduced to the following task: when is the univariate power series $h_z(x) := F(x, z)$ (for a fixed $z \in \mathbb{C} \setminus \{0\}$) algebraic? Obviously, the same argument as in the proof of Proposition 7.1 applies for any $z$ with $|z| > 1$: the growth rate of $z^{j^2}$ is incompatible with algebraicity. However, for $z = \omega$ on the unit circle or for $|z| < 1$ the same argument does not work; in fact we have the following result:

**Proposition 7.2.** *Let $n, k, a, b$ be admissible integers and $\omega \in \mathbb{C}$ be a root of unity. Then $h_\omega(x)$ is algebraic.*

Before proving it, we recall the $q$-Lucas theorem (see [120, p. 22] for an algebraic proof and [279, p. 131–132] for a combinatorial proof). It is the $q$-analogue of the well-known Lucas theorem, one form of which states that

$$\binom{n}{m} \equiv \binom{\lfloor \frac{n}{p} \rfloor}{\lfloor \frac{m}{p} \rfloor} \cdot \binom{n - p\lfloor \frac{n}{p} \rfloor}{m - p\lfloor \frac{m}{p} \rfloor} \mod p,$$

where $p$ is a prime number and $m, n$ are non-negative integers; here $\lfloor x \rfloor$ denotes the integer part of $x$, that is the largest integer at most equal to $x$.

**Theorem 7.3** ($q$-Lucas Theorem)**.** *Let $x, y$ be non-negative integers and $\omega \in \mathbb{C}$ be a root of unity of order $s$. Then*

$$\begin{bmatrix} x \\ y \end{bmatrix}_\omega = \binom{\lfloor \frac{x}{s} \rfloor}{\lfloor \frac{y}{s} \rfloor} \cdot \begin{bmatrix} x - s\lfloor \frac{x}{s} \rfloor \\ y - s\lfloor \frac{y}{s} \rfloor \end{bmatrix}_\omega.$$

Now we can show that for roots of unity $\omega \in \mathbb{C}$, the series $h_\omega(x)$ is algebraic.

*Proof of Proposition 7.2.* Let $s$ be the order of $\omega$. We have

$$h_\omega(x) = \sum_{j\geq 0} \begin{bmatrix} n + aj \\ k + bj \end{bmatrix}_\omega x^j = \sum_{r=0}^{s-1} \sum_{\substack{j\geq 0 \\ j\equiv r \bmod s}} \begin{bmatrix} n + aj \\ k + bj \end{bmatrix}_\omega x^j.$$

Let us examine the $s$ summands separately using the $q$-Lucas theorem:

$$\begin{aligned}\sum_{\substack{j\geq 0\\ j\equiv r \bmod s}} \begin{bmatrix} n+aj \\ k+bj \end{bmatrix}_\omega x^j &= \sum_{\ell\geq 0} \begin{bmatrix} n+a(\ell s+r) \\ k+b(\ell s+r) \end{bmatrix}_\omega x^j \\ &= \sum_{\ell\geq 0} \binom{\lfloor \frac{n+ar}{s} \rfloor + a\ell}{\lfloor \frac{k+br}{s} \rfloor + b\ell} \begin{bmatrix} n+ar-s\lfloor \frac{n+ar}{s} \rfloor \\ k+br-s\lfloor \frac{k+br}{s} \rfloor \end{bmatrix}_\omega x^j \\ &= \begin{bmatrix} n+ar-s\lfloor \frac{n+ar}{s} \rfloor \\ k+br-s\lfloor \frac{k+br}{s} \rfloor \end{bmatrix}_\omega \cdot \sum_{\ell\geq 0} \binom{\lfloor \frac{n+ar}{s} \rfloor + a\ell}{\lfloor \frac{k+br}{s} \rfloor + b\ell} x^j.\end{aligned}$$

By Pólya's theorem (see Section 7.1) the last sum is an algebraic series, and therefore obviously the whole last expression is also algebraic. Then, $h_\omega(x)$ is the sum of $s$ algebraic power series, hence it is algebraic as well. □

In the remaining part of the article we will show that if $0 < |z| \leq 1$ but $z$ is not a root of unity, then $h_z(x)$ cannot algebraic. This will prove our main theorem:

**Theorem 7.4.** *Let $n, k, a, b$ be admissible integers and let $q \in \mathbb{C} \setminus \{0\}$. Then $h_q(x)$ is an algebraic power series if and only if $q$ is a root of unity.*

A natural approach to prove this theorem is to use a result originating from Ramis' work [270], see also [280, Corollary 2], or [43, Theorem 7.1]. It says that if a function $f$ is algebraic and at the same time satisfies a linear $q$-difference equation, that is, an equation of the form

$$f(q^m x) + b_{m-1}(x) f(q^{m-1}x) + \cdots + b_0(x) f(x) = 0$$

for some rational functions $b_0, \ldots, b_m \in \mathbb{C}(x)$ not all zero and $q \in \mathbb{C}\setminus\{0\}$ not a root of unity, then $f$ is actually a rational function. Clearly, $h_q(x)$ satisfies a linear $q$-difference equation. Hence, the result described above ensures that if $h_q(x)$ is algebraic, it must already be a rational function. One can prove that for any non-zero $q$, $h_q(x)$ is rational if and only if $a = b$ or $b = 0$. Then, altogether, these facts imply Theorem 7.4.

It turns out that a simple modification of the proof that $h_q(x)$ is never rational for admissible integers already implies a much more general fact which does not require the theory of $q$-difference equations in order to prove our main theorem. More precisely, we will prove directly the following stronger result.

**Theorem 7.5.** *Let $n, k, a, b$ be admissible integers and let $q \in \mathbb{C} \setminus \{0\}$. Then $h_q(x)$ is D-finite if and only if $q$ is a root of unity.*

Recall that an algebraic function is always D-finite, therefore Theorem 7.5 along with Proposition 7.2 will allow us to conclude the validity of Theorem 7.4.

We will make use of the following elementary proposition, that We will prove and use the following elementary proposition: we suspect to be well-known although we could not locate it in the literature.

**Proposition 7.6.** *Let $p(x, y) \in \mathbb{C}[x, y]$ be a bivariate polynomial and assume that for some $q \in \mathbb{C} \setminus \{0\}$ not a root of unity, we have $p(j, q^j) = 0$ for all $j \in \mathbb{N}$. Then $p(x, y) = 0$.*

*Proof.* We distinguish three cases: $|q| < 1$, $|q| = 1$ and $|q| > 1$. For the first case, write $p(x, y) = p_0(x) + r(x, y)y^n$ for some natural number $n$ and $p_0(x) = p(x, 0) \in \mathbb{C}[x]$, $r(x, y) \in \mathbb{C}[x, y]$ such that $r(x, 0) \neq 0$. It follows that

$$0 = p(j, q^j) = p_0(j) + r(j, q^j)q^{nj}.$$

Since $|q| < 1$, we must have $r(j, q^j)q^{nj} \to 0$ as $j \to \infty$. Therefore, $\lim_{j\to\infty} p_0(j) = 0$ and we obtain that $p_0(x) = 0$. Hence, $r(j, q^j) = 0$ for $j \geq 0$. But $r(x, y) = r_0(x) + s(x, y)y$ for some polynomial $s(x, y)$ and non-zero $r_0(x)$. By the same argument, $\lim_{j\to\infty} r_0(j) = 0$, however this contradicts $r_0(x) \neq 0$.

If $|q| = 1$, write $p(x, y) = p_0(y) + p_1(y)x + \cdots + p_d(y)x^d$ for some natural number $d$ and polynomials $p_0(y), \ldots, p_d(y) \in \mathbb{C}[y]$, such that $p_d(y) \neq 0$. We have

$$|p_d(q^j)|j^d = \left|\sum_{k=0}^{d-1} p_k(q^j)j^k\right|. \tag{7.1}$$

Now the idea is that for some sequence $(j_n)_{n\geq 0}$, the terms $|p_0(q^{j_n})|, \ldots, |p_{d-1}(q^{j_n})|$ can be bounded by a constant from above and $|p_d(q^{j_n})|$ is bounded from below by a non-zero constant – this contradicts (7.1) because the left-hand side becomes too large. More precisely, choose $\xi$ on the unit circle which is not a root of $p_d(x)$. Then there exists $\varepsilon > 0$ such that $|p_k(\xi)| < 1/\varepsilon$ for all $k = 0, \ldots, d$ and also $\varepsilon < |p_d(\xi)|$. Moreover, since $q$ is not a root of unity, Jacobi's Theorem implies that the set $\{q^j | j \in \mathbb{N}\}$ is dense on the unit circle [122, Thm 3.13] (see also [119]), consequently there are infinitely many $j$ such that $q^j$ is arbitrarily close to $\xi$. Henceforth, there also exist infinitely many $j$ for which $|p_k(q^j)| < 1/\varepsilon$ for all $k = 0, \ldots, d-1$ and $\varepsilon < |p_d(q^j)|$. However, at the sequence of these $j$, this contradicts (7.1) since then the left-hand side grows at least like $j^d\varepsilon$ and the right-hand side is bounded by $dj^{d-1}/\varepsilon$.

Finally, if $|q| > 1$, write $p(x, y) = r_0(x) + r_1(x)y + \cdots + r_n(x)y^n$ for polynomials $r_0(x), \ldots, r_n(x) \in \mathbb{C}[x]$ and some natural number $n$. Clearly, if $p(x, y)$ is non-zero, $n$ must be positive. But then we have

$$|r_0(j) + r_1(j)q^j + \cdots + r_{n-1}(j)q^{(n-1)j}| \leq cj^m|q|^{(n-1)j},$$

for some constants $c, m > 0$. This contradicts

$$r_n(j)q^{nj} = -r_0(j) - r_1(j)q^j - \cdots - r_{n-1}(j)q^{(n-1)j}$$

for big enough $j$ and finishes the proof. □

**Remark 7.7.** An alternative, purely algebraic, proof of Proposition 7.6 follows from the fact that for any $d \geq 0$, writing $D = \binom{d+2}{2}$, the determinant of the $D \times D$ matrix

$$M(z) = (n^i z^{nj})_{\substack{0\leq i+j\leq d,\\ 1\leq n\leq D}}$$

is given by a constant times a power of $z$ and a product of cyclotomic polynomials in $z$. More precisely, assuming that the total degree of $p(x, y) = \sum_{i,j} c_{i,j}x^i y^j$ is $d$, the equations

$p(j, q^j) = 0$ for $j = 1, \dots, D$ yield the following linear system of equations for the vector of unknowns $c_{i,j}$:

$$M(q) \cdot (c_{i,j})_{0 \le i+j \le d} = 0.$$

To see why $\det M(z)$ only vanishes for $z$ a root of unity, it is useful to rewrite $M(z) = N(1, z, \dots, z^d)$, where $N(z_0, \dots, z_d) = (n^i z_j^n)_{\substack{0 \le i+j \le d, \\ 1 \le n \le D}} \in \mathbb{C}[z_1, \dots, z_d]^{D \times D}$. Then it remains to prove that $\det N(z_0, \dots, z_d)$ is a constant times a product of $z_i$'s times a product of $(z_i - z_j)$ for $i \neq j$. This follows from the observation that the transpose of $N$ is a generalized Vandermonde matrix; more precisely, it is a matrix corresponding to the linear map from the space of polynomials with no constant term and degree at most $D$ to $\mathbb{C}^D$ given by

$$P(x) \mapsto (P(z_0), \dots, P(z_d), \vartheta P(z_0), \dots, \vartheta P(z_{d-1}), \dots, \vartheta^d P(z_0)),$$

where $\vartheta = x\frac{x}{\mathrm{d}x}$ is the Euler derivative. After a change of basis from the monomials to the basis $1, (x - z_0), (x - z_0)(x - z_1), \dots, (x - z_0)^{d-1}(x - z_1)^{d-1}(x - z_2)^{d-2} \cdots (x - z_d)$ the matrix becomes lower-triangular and the determinant evaluation follows trivially. This purely algebraic proof shows that in Proposition 7.6 one can replace "for all $j \in \mathbb{N}$" by the weaker condition "for $j = 1, \dots, (d+1)(d+2)/2$, where $d$ is the total degree of $p(x, y)$". The proof also shows that the conclusion of the proposition holds as well if $q \in \mathbb{C} \setminus \{0\}$ is assumed not to be a root of unity of order at most $d$.

Note that this proposition immediately implies that the function $f(x) = q^x$ is transcendental for any non-zero $q \in \mathbb{C}$ which is also not a root of unity, because an annihilating polynomial $P(x, z)$ would need to satisfy $P(x, q^x) = 0$ and hence this would hold at all integers $x$. In particular, Proposition 7.6 contains the classical and well-known fact that $\exp(x)$ is not algebraic.

Now we are ready to prove Theorem 7.5: we will show that $h_q(x)$ is D-finite (and hence algebraic) if and only if $q$ is a root of unity. This answers Aissen's question completely.

*Proof of Theorem 7.5.* We already observed that if $q$ is a root of unity, the series $h_q(x)$ is algebraic and hence D-finite. Therefore, one direction is clear and we assume now that $q \in \mathbb{C} \setminus \{0\}$ is not a root of unity.

Assume by contradiction that $h_q(x) = \sum_{j \ge 0} u_j x^j$ is D-finite. Then $(u_j)_{j \ge 0}$ is P-recursive and there exist a positive integer $r$ and $c_0(x), \dots, c_r(x) \in \mathbb{C}[x]$ with $c_0(x)c_r(x) \neq 0$ such that

$$u_{j+r}c_r(j) + \cdots + u_j c_0(j) = 0, \quad \text{for all } j \ge 0. \tag{7.2}$$

A simple computation shows that

$$u_{j+1} = u_j \frac{\prod_{\ell=1}^{a}(q^{n+aj+\ell} - 1)}{\prod_{\ell=1}^{b}(q^{k+bj+\ell} - 1)\prod_{\ell=1}^{a-b}(q^{n-k+(a-b)j+\ell} - 1)}.$$

Then it follows by iteration

$$\begin{aligned} u_{j+i} &= u_j \prod_{m=0}^{i-1} \frac{\prod_{\ell=1}^{a}(q^{n+a(j+m)+\ell} - 1)}{\prod_{\ell=1}^{b}(q^{k+b(j+m)+\ell} - 1)\prod_{\ell=1}^{a-b}(q^{n-k+(a-b)(j+m)+\ell} - 1)} \\ &= u_j \frac{\prod_{\ell=1}^{ia}(q^{n+aj+\ell} - 1)}{\prod_{\ell=1}^{ib}(q^{k+bj+\ell} - 1)\prod_{\ell=1}^{i(a-b)}(q^{n-k+(a-b)j+\ell} - 1)}. \end{aligned}$$

Using this, we may rewrite equation (7.2) and obtain

$$u_j \left( \sum_{i=0}^{r} c_i(j) \frac{\prod_{\ell=1}^{ia} (q^{n+aj+\ell} - 1)}{\prod_{\ell=1}^{ib} (q^{k+bj+\ell} - 1) \prod_{\ell=1}^{i(a-b)} (q^{n-k+(a-b)j+\ell} - 1)} \right) = 0, \tag{7.3}$$

for all integers $j \geq 0$. Note that $u_j \neq 0$, since $q$ is not a root of unity, hence already the sum above is identical to $0$ for all $j \in \mathbb{N}$. We define

$$P_i(y) \coloneqq \prod_{\ell=1}^{ia} (y^a q^{n+\ell} - 1) \prod_{\ell=ib+1}^{rb} (y^b q^{k+\ell} - 1) \prod_{\ell=i(a-b)+1}^{r(a-b)} (y^{a-b} q^{n-k+\ell} - 1) \in \mathbb{C}[y],$$

so that after multiplication with the common denominator, equation (7.3) implies that $\sum_{i=1}^{r} c_i(j) P_i(q^j) = 0$. By Proposition 7.6 we now obtain that $p(x, y) \coloneqq \sum_{i=1}^{r} c_i(x) P_i(y)$ must be identically $0$.

We will show however that $p(x, y)$ cannot be the zero polynomial if the integers $n, k, a, b$ are admissible, more precisely if $a > b > 0$. Set first $d \coloneqq \max(\deg(c_i(x)), i = 0, \ldots, r)$ and write $c_i(x) = \sum_{k=0}^{d} c_{i,k} x^k$ for some $c_{i,k} \in \mathbb{C}$. Moreover, let $m$ be an integer such that $c_{r,m} \neq 0$ and denote by $p_m(y)$ the coefficient of $x^m$ in $p(x, y)$. We claim that $p_m(y) \neq 0$. We namely have:

$$p_m(y) = \sum_{i=0}^{r} c_{i,m} \prod_{\ell=1}^{ia} (y^a q^{n+\ell} - 1) \prod_{\ell=ib+1}^{rb} (y^b q^{k+\ell} - 1) \prod_{\ell=i(a-b)+1}^{r(a-b)} (y^{a-b} q^{n-k+\ell} - 1),$$

and the exponent of $y$ of the leading monomial of each summand is $ia^2 + (r-i)b^2 + (r-i)(a-b)^2$. Since we assume that $a > b > 0$, it follows that this expression is maximal only for $i = r$, and hence the leading monomial of $p_m(y)$ is $c_{r,m} y^{a^2 r} q^{ran + ra(ra+1)/2} \neq 0$. □

# Chapter 8

# Representation of sequences as constant terms

*“Но гениальный всплеск похож на бред,*
*В рожденье смерть проглядывает косо.*
*А мы всё ставим каверзный ответ*
*И не находим нужного вопроса.”*[1]
Владимир Высоцкий, *Мой Гамлет*, 1972

A constant term sequence is a sequence of rational numbers whose $n$-th term is the constant term of $P^n(\boldsymbol{x})Q(\boldsymbol{x})$, where $P(\boldsymbol{x})$ and $Q(\boldsymbol{x})$ are multivariate Laurent polynomials. While the generating functions of such sequences are invariably diagonals of multivariate rational functions, and hence special period functions, it is a famous open question, raised by Don Zagier, to classify those diagonals which are constant terms. In this chapter, we provide such a classification in the case of sequences satisfying linear recurrences with constant coefficients. We further consider the case of hypergeometric sequences and, for a simple illustrative family of hypergeometric sequences, classify those that are constant terms.

This chapter of the thesis incorporates joint work with A. Bostan and A. Straub [79].

## 8.1 Introduction

Recognizing and interpreting integrality of sequences defined by recursions is at the same time an extensively studied and a hardly understood topic in number theory. Even for the case of sequences $A(n)$ defined by linear recurrences with polynomial coefficients, the so-called *P-recursive* sequences,

$$p_r(n)A(n+r) = p_{r-1}(n)A(n+r-1) + \cdots + p_0(n)A(n), \quad p_i(n) \in \mathbb{Z}[n],$$

neither a criterion nor even an algorithm is known for classifying/deciding integrality. An attempt for such a classification is the famous and widely open conjecture by Christol [101,

[1] Translation (S. Roy): *“A genius bursts like a delirious cry. At birth, death shows his visage grim and leery. We pose again the tricky old reply, And cannot find the necessary query.”* Vladimir Vysotsky, *My Hamlet*.

Conjecture 4, p. 55]. Roughly speaking, it states that a P-recursive sequence $(A(n))_{n\geq 0}$ with (at most) geometric growth is integral if and only if $(A(n))_{n\geq 0}$ is the coefficient sequence of the diagonal of a rational function $R(\boldsymbol{x}) \in \mathbb{Z}(x_1,\dots,x_d) \cap \mathbb{Z}[\![x_1,\dots,x_d]\!]$ for some $d \geq 1$. Recall that the diagonal of a multivariate power series

$$R(\boldsymbol{x}) = \sum_{n_1,n_2,\dots,n_d\geq 0} c(n_1,n_2,\dots,n_d) x_1^{n_1} x_2^{n_2} \cdots x_d^{n_d} \tag{8.1}$$

is the univariate power series $\mathrm{Diag}(R)$ whose coefficient sequence is given by $A(n) = c(n,n,\dots,n)$. For a precise statement of Christol's conjecture see Conjecture 8.19 below.

Often integrality of sequences can be explained by the underlying combinatorial nature. For example, the Catalan numbers $C(n)$ satisfying

$$(n+2)C(n+1) = 2(2n+1)C(n), \quad C(0)=1,$$

are clearly integers because they count triangulations of convex polygons with $n+2$ vertices. On the other hand, for many other integral and P-recursive sequences, combinatorial interpretations are not *a priori* known; this is the case, for instance, for the Apéry numbers $A(n)$ (associated with the irrationality proof of $\zeta(3)$) defined by

$$(n+1)^3 A(n+1) = (2n+1)(17n^2+17n+5)A(n) - n^3 A(n-1), \\ A(0)=1, A(1)=5.$$

In both examples above, integrality can be seen from the explicit formulas

$$C(n) = \binom{2n}{n} - \binom{2n}{n+1} \quad \text{and} \quad A(n) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2.$$

Putting Christol's conjecture in practice gives a different justification for the integrality of these two examples. It namely holds that

$$\sum_{n\geq 0} C(n)t^n = \mathrm{Diag}\left(\frac{1-y}{1-x(y+1)^2}\right) \quad \text{and}$$
$$\sum_{n\geq 0} A(n)t^n = \mathrm{Diag}\left(\frac{1}{1-(xy+x+y)(zw+z+w)}\right),$$

and the integrality of $C(n)$ and $A(n)$ follows from that of the coefficients in the Taylor expansions of the corresponding multivariate rational functions.

In the context of the current text, however, we would like to emphasize a slightly different viewpoint, which does not only justify integrality of the two examples, but also implies some interesting arithmetic properties. Writing $\mathrm{ct}[P(\boldsymbol{x})]$ for the constant term of a Laurent polynomial $P(\boldsymbol{x}) \in \mathbb{Q}[x_1^{\pm 1},\dots,x_d^{\pm 1}]$, one can prove that [308, Rem. 1.4]

$$C(n) = \mathrm{ct}\left[(x^{-1}+2+x)^n(1-x)\right] \quad \text{and}$$
$$A(n) = \mathrm{ct}\left[\left(\frac{(x+y)(z+1)(x+y+z)(y+x+1)}{xyz}\right)^n\right].$$

Similar identities as in the examples of the Catalan and Apéry numbers can be deduced for many other integral P-recursive sequences. This motivates the following definition and the subsequent natural question.

**Definition 8.1.** A sequence $A(n)$ is a *constant term* if it can be represented as

$$A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})], \tag{8.2}$$

where $P, Q \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$ are Laurent polynomials in $\boldsymbol{x} = (x_1, \ldots, x_d)$.

Using the geometric series it is easy to see that generating functions of constant term sequences can be expressed as diagonals of rational functions. The converse is, however, not true in general. This leads to the following question which was raised by Zagier [338, p. 769, Question 2] and Gorodetsky [162] in the case $Q = 1$ (see Proposition 8.15 below for an indication why this case is of particular arithmetic significance).

**Question 1.** *Which P-recursive sequences are constant terms?*

To our knowledge, Question 1 is widely open. In fact, the initial motivation for the present text was the goal of answering the following very particular sub-question asked by the second author in [309, Question 5.1]:

**Question 2.** *Is the Fibonacci sequence $(F(n))_{n\geq 0}$ a constant term sequence?*

Recall that the Fibonacci sequence $(F(n))_{n\geq 0}$ is the coefficient sequence in the Taylor expansion of the univariate rational function $x/(1-x-x^2)$, or equivalently the P-recursive sequence $(F(n))_{n\geq 0}$ defined by $F(n+2) = F(n+1) + F(n)$ and $F(0) = 0, F(1) = 1$.

Already in [309] the second author noted that a representation of the Fibonacci numbers as constant terms with $Q = 1$ is impossible since $(F(n))_{n\geq 0}$ does not satisfy the so-called *Gauss congruences* (see (8.9)). Exploiting the fact that for any prime $p$, the value $F(p) \pmod p$ depends on $p \pmod 5$, we can show (see Example 8.8) that the answer to Question 2 is negative. The reason for this is that, as we will prove, for any constant term sequence $A(n)$, the sequence $A(p) \pmod p$ must be constant for large enough primes $p$. Note that this is not a sufficient criterion, since already the Lucas numbers $L(n)$ (defined by the same recursion as the Fibonacci numbers, but with different initial terms $L(0) = 2, L(1) = 1$, see (8.8)) do satisfy the Gauss congruences but are not constant terms (see Example 8.12).

In the present text, we are able to answer Question 1 in the case of diagonals of rational functions $F(x) \in \mathbb{Q}(x)$ in a single variable. Such sequences are precisely the (rational) *C-finite sequences* (also known as *C-recursive* sequences), and are characterized by the fact that they satisfy a linear recursion with constant rational coefficients. More explicitly, we define a sequence $A(n)$ of rational numbers to be C-finite if there exists a polynomial $P(x) \in \mathbb{Q}[x]$ such that for every $n \geq 0$ we have

$$P(N)A(n) = 0, \tag{8.3}$$

where $N$ denotes the shift operator $N^\ell(A(n)) := A(n+\ell)$ for all $\ell \geq 0$. Equivalently, there exist integers $r > 0$ and $n_0 \geq 0$, and complex numbers $c_0, \ldots, c_{r-1}$ with $c_0 \neq 0$ such that

$$A(n+r) = c_{r-1}A(n+r-1) + \cdots + c_0 A(n) \quad \text{for all } n \geq n_0. \tag{8.4}$$

We recall that associated to the recursion (8.4), the *characteristic roots* are usually defined as the roots of

$$\chi(\lambda) := \lambda^r - c_{r-1}\lambda^{r-1} - \cdots - c_0.$$

For our purpose, however, it is useful to define the characteristic roots of a C-finite sequence $A(n)$ as the roots of $P(x)$, where $P(x)$ is chosen with minimal degree such that (8.3) holds. Note that the only difference between considering roots of $\chi$ and $P$ is that $0$ can be a root of the latter. Equivalently, 0 is defined to be a characteristic root of $A(n)$ of multiplicity $m_0$ if the minimal $n_0$ in (8.4) (chosen so that $r$ is minimal) equals $m_0$. With these definitions we obtain the following:

**Proposition 8.2.** *Let $A(n)$ be a C-finite sequence. $A(n)$ is a constant term if and only if it has a single characteristic root $\lambda$ and $\lambda \in \mathbb{Q}$.*

This proposition immediately answers Question 2 but also shows that, for example, the sequence $A(n) = 2^n + 1$ is not a constant term sequence either (in both of these cases, there are two different characteristic roots). Evidently, however, it is the sum of two constant terms: we see that the class of constant term sequences is not a ring. Therefore, to fix this issue, it is natural to consider the class of sequences given as $\mathbb{Q}$-linear combinations of constant terms:

**Question 3.** *Which P-recursive sequences are finite $\mathbb{Q}$-linear combinations of constant terms?*

Again in the case of C-finite sequences, we can answer this question completely with the main result of the present work:

**Theorem 8.3.** *Let $A(n)$ be a C-finite sequence. Then $A(n)$ is an $r$-term $\mathbb{Q}$-linear combination of constant terms if and only if it has at most $r$ distinct characteristic roots, all of which are rational.*

Having completed the classification of C-finite sequences that can be written as (sums of) constant terms, there are two most natural directions for further work. On the one hand, it is reasonable to go from diagonals in one variable to diagonals in two variables. By the combination of results due to Pólya [267] and Furstenberg [145] this is known to be exactly the class of algebraic generating functions. One is then lead to the following question which we leave for future work:

**Question 4.** *Which sequences $A(n)$ with algebraic generating function are constant terms?*

Another reasonable direction is to try to classify those hypergeometric sequences which are constant terms. Recall that a P-recursive sequence $A(n)$ is called *hypergeometric* if it satisfies a recursion of order one, i.e. $\alpha(n)A(n+1) = \beta(n)A(n)$ for some polynomials $\alpha(n), \beta(n) \in \mathbb{Q}[n]$. In this sense, this class of sequences is arguably the simplest (and best understood) among P-recursive ones. Still, Christol's conjecture remains open even in this very special case. In fact, it is still an open question whether the generating function of the sequence

$$A(n) = \frac{\left(\frac{1}{9}\right)_n \left(\frac{4}{9}\right)_n \left(\frac{5}{9}\right)_n}{n!^2 \left(\frac{1}{3}\right)_n}$$

can be represented as the diagonal of a rational function. Recall that $(x)_n \coloneqq x(x+1)\cdots(x+n-1)$ denotes the rising factorial. We can use the same methods as in the C-finite case to prove that $A(n)$ is not a constant term sequence (see Lemma 8.22). By classifying when the family (8.21) of hypergeometric sequences is a constant term, we are further able to conclude that not all hypergeometric diagonals are constant terms. The following question, however, remains open in general:

**Question 5.** *Which hypergeometric sequences are constant terms?*

The organization of the paper is as follows: In Section 8.2, we review properties of C-finite sequences that will be important for our purposes. In particular, we state Theorem 8.5 which is due to Minton [246] and which is a crucial ingredient of our approach. In Section 8.3, we derive certain congruences that are satisfied by any constant term sequence; these are already enough to answer Question 2. By combining these congruences with Minton's Theorem, we prove in Section 8.4 our main Theorem 8.3, thus answering Question 1 and Question 3 in the case of C-finite sequences. In the short Section 8.5 we prove a statement which is pleasingly similar to Minton's theorem and which allows to classify the constant terms with $Q = 1$ among all constant terms. Finally, in Section 8.6, we turn our attention to hypergeometric sequences and discuss Question 5.

Throughout the article, $p$ denotes a prime number, $\mathbb{F}_p$ the finite field with $p$ elements and $\mathbb{Z}_p$ the ring of $p$-adic integers.

## 8.2 Trace sequences

Let $A(n)$ be a C-finite sequence. Denote by $\lambda_1, \lambda_2, \ldots, \lambda_d \in \overline{\mathbb{Q}}$ the characteristic roots, and let $m_j$ be the multiplicity of the root $\lambda_j$. Recall that $\lambda_0 = 0$ is defined to be a characteristic root of $A(n)$ of multiplicity $m_0$ if the minimal $n_0$ in (8.4) equals $m_0$. $A(n)$ can be written as a linear combination

$$A(n) = A_0(n) + \sum_{j=1}^{d} \sum_{r=0}^{m_j-1} c_{j,r} n^r \lambda_j^n \tag{8.5}$$

for certain coefficients $c_{j,r} \in \overline{\mathbb{Q}}$ (more precisely, $c_{j,r} \in \mathbb{Q}(\lambda_1, \ldots, \lambda_d)$) and $A_0(n)$ a sequence of finite support $\{0, 1, \ldots, m_0 - 1\}$. We refer to [134] or [199, Chapter 4] for introductions to C-finite sequences. Note that allowing $0$ as a characteristic root is equivalent to not restricting the numerator of the rational generating function of $A(n)$ to have degree less than the degree of its denominator. In the following, we will refer to

$$A^{\mathrm{sep}}(n) = A_0(0) + \sum_{j=1}^{d} c_{j,0} \lambda_j^n \tag{8.6}$$

as the *separable part* of $A(n)$. We note that, if $A(n) \in \mathbb{Q}$, then $A^{\mathrm{sep}}(n) \in \mathbb{Q}$.

A sequence $A(n)$ is said to be a *trace sequence* if it is a $\mathbb{Q}$-linear combination of traces $\mathrm{Tr}(\theta^n) = \theta_1^n + \cdots + \theta_r^n$ of algebraic numbers $\theta$ with Galois conjugates $\theta_1 = \theta, \theta_2, \ldots, \theta_r$ (with the understanding that $\mathrm{Tr}(0^n)$ is $1$ for $n = 0$ and $0$ otherwise). Equivalently, a trace sequence is a C-finite sequence for which the multiplicity of each characteristic root is $m_j = 1$ and for which $c_{i,0} = c_{j,0}$ in (8.5) whenever $\lambda_i$ and $\lambda_j$ have the same minimal polynomial. We further note as in [39] that the condition to be a trace sequence is equivalent to the property that the generating function $F(x)$ is $F(0)$ plus a $\mathbb{Q}$-linear combination of functions of the form $xu'(x)/u(x)$, where $u \in \mathbb{Q}[x]$ is irreducible and $u(0) = 1$.

**Example 8.4.** For the Fibonacci numbers $F(n)$, the representation (8.5) takes the form

$$F(n) = \frac{\varphi_+^n - \varphi_-^n}{\sqrt{5}}, \quad \varphi_\pm = \frac{1 \pm \sqrt{5}}{2}. \tag{8.7}$$

Because the coefficients of $\varphi_+^n$ and $\varphi_-^n$ differ in sign, the Fibonacci numbers $F(n)$ are not a trace sequence. On the other hand, the Lucas numbers

$$L(n) = \varphi_+^n + \varphi_-^n = \operatorname{tr}[M^n], \quad M = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \tag{8.8}$$

which satisfy the same recurrence as the Fibonacci numbers, are a trace sequence. In particular, it follows from Theorem 8.5 that the Lucas numbers $L(n)$ satisfy the Gauss congruences (8.9).

Minton [246] classified those C-finite sequences that satisfy the Gauss congruences (8.9) (see [39] for another proof of Minton's result).

**Theorem 8.5** (Minton, 2014)**.** *Let $A(n)$ be C-finite. Then the following are equivalent:*

1. *For all large enough primes $p$ and for all $r \geq 1$, $A(n)$ satisfies the Gauss congruences*

$$A(p^r n) \equiv A(p^{r-1} n) \pmod{p^r}. \tag{8.9}$$

2. *For all large enough primes $p$, $A(n)$ satisfies the congruences*

$$A(pn) \equiv A(n) \pmod{p}. \tag{8.10}$$

3. *$A(n)$ is a trace sequence.*

We conclude from Minton's Theorem 8.5 the following result, which we employ in the proof of our main result (Theorem 8.3). To see the importance of Lemma 8.6, we note that, as we will show later (in Corollary 8.10), the sequences $A(n)$ which are linear combinations of constant terms satisfy the congruences $A(p^r n) \equiv A(pn) \pmod{p}$ for all $r \geq 1$ and large enough primes $p$.

**Lemma 8.6.** *Let $A(n)$ be C-finite. If $A(n)$ satisfies the congruences*

$$A(p^r n) \equiv A(pn) \pmod{p} \tag{8.11}$$

*for all $r \geq 1$ and for all large enough primes $p$, then the separable part $A^{\mathrm{sep}}(n)$ is a trace sequence.*

*Proof.* It follows from comparing (8.5) with (8.6) that for $n$ large enough

$$A(n) = A^{\mathrm{sep}}(n) + n\tilde{A}(n),$$

where $A^{\mathrm{sep}}(n)$ and $\tilde{A}(n)$ are rational and satisfy the minimal recurrence for $A(n)$. In particular, each of these sequences is in $\mathbb{Z}_p$ for large enough $p$, since denominators can only arise from the coefficients of the recurrence and the initial conditions. It follows that

$$A^{\mathrm{sep}}(pn) \equiv A(pn) \pmod{p}$$

for all large enough $p$. Consequently, the congruences (8.11) are also satisfied by the C-finite sequence $A^{\mathrm{sep}}(n)$. That is, for all $r \geq 1$ and large enough $p$

$$A^{\mathrm{sep}}(p^r n) \equiv A^{\mathrm{sep}}(pn) \pmod{p}. \tag{8.12}$$

On the other hand, let us consider the C-finite sequence $A^{\mathrm{sep}}(n)$ in $\mathbb{F}_p$. To avoid confusion, we denote this reduced sequence by $a_p^{\mathrm{sep}}(n)$. Since the characteristic polynomial of $A^{\mathrm{sep}}(n)$ over $\mathbb{Q}$ is separable, it is also separable for all large enough primes $p$ (this can be seen by looking at the discriminant which, if nonzero over $\mathbb{Q}$, can only vanish modulo finitely many primes). Consequently, we have a version of (8.6) with coefficients and roots in $\overline{\mathbb{F}}_p$. Namely,

$$a_p^{\mathrm{sep}}(n) = \sum_{j=1}^{d} d_j \mu_j^n, \quad d_j, \mu_j \in \overline{\mathbb{F}}_p.$$

Denoting with $\varphi_p : \overline{\mathbb{F}}_p \to \overline{\mathbb{F}}_p$ the Frobenius automorphism defined by $\varphi_p(z) = z^p$, we therefore have

$$a_p^{\mathrm{sep}}(p^s n) = \sum_{j=1}^{d} d_j \mu_j^{p^s n} = \sum_{j=1}^{d} d_j (\varphi_p^s(\mu_j))^n$$

for each $s \in \mathbb{Z}_{>0}$. Note that $\varphi_p$ acts as a permutation on the roots $\mu_j$. Writing $m$ for the order of this permutation, we have $\varphi_p^m(\mu_j) = \mu_j$ and thus

$$a_p^{\mathrm{sep}}(p^m n) = a_p^{\mathrm{sep}}(n).$$

Consequently, the corresponding sequence $A^{\mathrm{sep}}(n)$ satisfies

$$A^{\mathrm{sep}}(p^m n) \equiv A^{\mathrm{sep}}(n) \pmod{p}.$$

Combined with the congruences (8.12), this implies that

$$A^{\mathrm{sep}}(pn) \equiv A^{\mathrm{sep}}(n) \pmod{p}$$

for all large enough $p$. Theorem 8.5 therefore implies that $A^{\mathrm{sep}}(n)$ is a trace sequence. □

## 8.3 Congruences for constant terms

In this section we will show that if $A(n)$ is a constant term sequence then it must satisfy certain congruences for large enough primes $p$. As a consequence, this allows us to conclude that the Fibonacci numbers are not a constant term sequence, thus answering Question 2 from the introduction.

For a Laurent polynomial $P \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$, let $\deg(P)$ denote the maximal degree with which any variable or its inverse appears in $P$.

**Lemma 8.7.** *Let $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ with $P, Q \in \mathbb{Z}_p[\boldsymbol{x}^{\pm 1}]$. Then*

$$A(p^r n + k) \equiv A(k) \,\mathrm{ct}[P(\boldsymbol{x})^{p^{r-1} n}] \pmod{p^r}$$

*for all integers $n, k \geq 0$ and $r \geq 1$, provided that $p > \deg(P^k Q)$.*

*Proof.* Recall that (see, for instance, [277, Proposition 1.9]), for any Laurent polynomial $F \in \mathbb{Z}_p[\boldsymbol{x}^{\pm 1}]$,

$$F(\boldsymbol{x})^{p^r} \equiv F(\boldsymbol{x}^p)^{p^{r-1}} \pmod{p^r}. \tag{8.13}$$

As in [309], it follows from (8.13) that

$$\begin{aligned} A(p^r n + k) &= \operatorname{ct}[P(\boldsymbol{x})^{p^r n} P(\boldsymbol{x})^k Q(\boldsymbol{x})] \\ &\equiv \operatorname{ct}[P(\boldsymbol{x}^p)^{p^{r-1} n} P(\boldsymbol{x})^k Q(\boldsymbol{x})] \pmod{p^r} \\ &= \operatorname{ct}[P(\boldsymbol{x})^{p^{r-1} n} \Lambda_p[P(\boldsymbol{x})^k Q(\boldsymbol{x})]], \end{aligned}$$

where $\Lambda_p$ denotes the Cartier operator

$$\Lambda_p\left[\sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{\boldsymbol{k}} \boldsymbol{x}^{\boldsymbol{k}}\right] = \sum_{\boldsymbol{k} \in \mathbb{Z}^d} a_{p\boldsymbol{k}} \boldsymbol{x}^{\boldsymbol{k}}.$$

If $p > \deg(P^k Q)$, then

$$\Lambda_p[P(\boldsymbol{x})^k Q(\boldsymbol{x})] = \operatorname{ct}[P(\boldsymbol{x})^k Q(\boldsymbol{x})] = A(k)$$

and the claim follows. □

**Example 8.8.** For the Fibonacci numbers $F(n)$, it is a well-known consequence of (8.7) that, modulo any prime $p$, we have the congruences

$$F(p) \equiv \begin{cases} 1, & \text{if } p \equiv 1, 4 \bmod 5, \\ -1, & \text{if } p \equiv 2, 3 \bmod 5, \end{cases} \pmod{p}.$$

Since this is incompatible with Lemma 8.7 (setting $r = n = 1$ and $k = 0$ implies that $A(p) \equiv A(0) \cdot c \pmod{p}$ for some $c \in \mathbb{Q}$ that is independent of $p$), we see that $F(n)$ is not a constant term sequence.

On the other hand, by Theorem 8.5, the Lucas numbers $L(n)$ (from (8.8)) satisfy the congruences $L(p^r n) \equiv L(pn) \pmod{p}$ for $r \geq 1$ and $p$ large enough. As such, Lemma 8.7 is not sufficient to conclude that $L(n)$ is not a constant term sequence. However, we will be able to conclude in Example 8.12 the stronger result that both the Fibonacci numbers and the Lucas numbers cannot be expressed as a $\mathbb{Q}$-linear combination of constant terms.

**Corollary 8.9.** *Let $A(n) = \operatorname{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ with $P, Q \in \mathbb{Z}_p[\boldsymbol{x}^{\pm 1}]$. Then*

$$A(p^s n + k) \equiv A(p^r n + k) \pmod{p^r}$$

*for all integers $n, k \geq 0$ and $s \geq r \geq 1$, provided that $p > \deg(P^k Q)$.*

*Proof.* It follows from Lemma 8.7 and (8.13) that

$$\begin{aligned} A(p^s n + k) &\equiv A(k) \operatorname{ct}[P(\boldsymbol{x})^{p^{s-1} n}] \pmod{p^s} \\ &\equiv A(k) \operatorname{ct}[P(\boldsymbol{x}^{p^{s-r}})^{p^{r-1} n}] \pmod{p^r} \\ &= A(k) \operatorname{ct}[P(\boldsymbol{x})^{p^{r-1} n}], \end{aligned}$$

as claimed. □

The simple but useful special case $r = 1$ and $k = 0$ of the corollary above takes the following form. Here, $p$ is large enough if $p > \deg(Q)$ and $P, Q \in \mathbb{Z}_p[\boldsymbol{x}^{\pm 1}]$.

**Corollary 8.10.** *Let $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ with $P, Q \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$. If $p$ is large enough, then, for all integers $n \geq 0$ and $r \geq 1$,*

$$A(p^r n) \equiv A(pn) \pmod{p}.$$

## 8.4 $C$-finite sequences that are constant terms

In this section, we prove our main result, Theorem 8.3 stated in the introduction, thus classify those C-finite sequences that are constant terms or linear combinations of such. We start by proving the following weaker version, since it illustrates well our approach and the usefulness of the congruences proved in Section 8.3. We then extend the argument to prove Theorem 8.3 in full generality.

**Proposition 8.11.** *Let $A(n)$ be a C-finite sequence. $A(n)$ is a $\mathbb{Q}$-linear combination of constant terms if and only if all characteristic roots are rational.*

*Proof.* For one direction, note that

$$\mathrm{ct}[(x+\lambda)^n (\lambda/x)^r] = \binom{n}{r} \lambda^n = \frac{n(n-1)\cdots(n-r+1)}{r!} \lambda^n. \tag{8.14}$$

Varying $r$, the right-hand side forms a basis for the span of the sequences $(n^r \lambda^n)_{n \geq 0}$. A sequence $A_0(n)$ of finite support can be represented as

$$A_0(n) = \mathrm{ct}[x^n (A(0) + A(1)x^{-1} + \cdots + A(N)x^{-N})],$$

where $N$ is the largest integer for which $A_0(N)$ is non-zero. It therefore follows with (8.5) that, if all characteristic roots $\lambda$ are rational, then $A(n)$ can be represented as a linear combination of constant terms.

On the other hand, suppose that $A(n)$ is a linear combination of constant terms. Note that this implies that any shift $A(n+k)$, where $k \in \mathbb{Z}_{\geq 0}$, is a linear combination of constant terms as well. These shifts generate the space $V_A$ of rational solutions of the minimal constant-coefficient recursion satisfied by $A(n)$. Thus, any sequence in $V_A$ is a linear combination of constant terms. Assume, for contradiction, that there is a characteristic root $\lambda$ that is not rational. Then among the sequences in $V_A$ there is always a sequence $B(n)$ of the form (8.6) (that is, $B(n)$ equals its separable part) which is not a trace sequence.

For instance, if $\lambda_1, \ldots, \lambda_d$ are the roots of the minimal polynomial of $\lambda$, then the space $V_\lambda$ of rational sequences of the form $b(n) = c_1 \lambda_1^n + \cdots + c_d \lambda_d^n$, with $c_1, \ldots, c_d \in \overline{\mathbb{Q}}$, is a $d$-dimensional subspace of $V_A$. Clearly, each sequence in $V_\lambda$ is of the form (8.6). Note that $\lambda_1^n + \cdots + \lambda_d^n$ and its multiples are the only trace sequences in $V_\lambda$. Since $d \geq 2$, we can therefore choose a sequence $B(n)$ in $V_\lambda$ that is not a trace sequence.

It follows from Corollary 8.10 that $B(n)$ satisfies the congruences

$$B(p^r n) \equiv B(pn) \pmod{p}$$

for all $r \geq 1$ and all large enough primes $p$. Lemma 8.6 therefore implies that $B^{\mathrm{sep}}(n) = B(n)$ is a trace sequence. This is a contradiction, and we conclude that all characteristic roots must be rational. $\square$

**Example 8.12.** Recall from (8.7) that the Fibonacci numbers $F(n)$ are C-finite with characteristic roots $(1 \pm \sqrt{5})/2$. Since these are not rational, it follows from Proposition 8.11 that $F(n)$ cannot be expressed as a linear combination of constant terms.

The same argument applied to (8.8) shows that the Lucas numbers $L(n)$ cannot be expressed as a linear combination of constant terms as well. Alternatively, this can also be concluded from the relationship

$$2L(n+1) - L(n) = 5F(n)$$

combined with the fact that Fibonacci numbers are not a sum of constant terms.

We next prove the case $r = 1$ of Theorem 8.3, that is Proposition 8.2, stating that a C-finite sequence $A(n)$ is a single constant term if and only if it has a single characteristic root $\lambda$ and $\lambda \in \mathbb{Q}$.

*Proof of Proposition 8.2.* It follows from (8.5) and (8.14) that if $A(n)$ is a C-finite sequence with the single characteristic root $\lambda \in \mathbb{Q}$ (possibly repeated or possibly 0), then $A(n)$ is a constant term, namely $A(n) = \mathrm{ct}[(x+\lambda)^n Q(x^{-1})]$ for a suitable polynomial $Q(x)$.

On the other hand, suppose that $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ is a single constant term. Since $A(n)$ is a C-finite sequence, it has a representation of the form (8.5) or, equivalently,

$$A(n) = A_0(n) + \sum_{j=1}^{d} \lambda_j^n p_j(n) \tag{8.15}$$

for pairwise distinct $\lambda_j \in \overline{\mathbb{Q}}^{\times}$ and nonzero $p_j(n) \in \overline{\mathbb{Q}}[n]$. As before, $A_0(n)$ is a sequence with finite support, corresponding to the characteristic root $0$. It follows from Proposition 8.11 that all characteristic roots $\lambda_j$ are rational, and this further implies that $p_j(n) \in \mathbb{Q}[n]$.

Let $c_0 = \mathrm{ct}[P(\boldsymbol{x})] \in \mathbb{Q}$. From Lemma 8.7 (with $r = 1$ and $n = 1$) it follows that

$$A(p+n) \equiv A(n) \cdot c_0 \pmod{p}$$

for all $n \geq 0$ and all large enough primes $p$ (namely, $p > \deg(P^n Q)$ and large enough so that $c_0 \in \mathbb{Z}_p$). Combining this congruence with (8.15) and applying Fermat's little theorem to reduce $\lambda_j^{p+n}$ and $p_j(p+n)$ modulo $p$ to $\lambda_j^{n+1}$ and $p_j(n)$ respectively, we find that

$$\sum_{j=1}^{d} \lambda_j^{n+1} p_j(n) \equiv c_0 \left[ A_0(n) + \sum_{j=1}^{d} \lambda_j^n p_j(n) \right] \pmod{p} \tag{8.16}$$

for all large enough $p$ (in particular, so that $p$ is larger than any denominator occuring in the $p_j(n)$ and so that $A_0(p+n) = 0$). Note that both sides of (8.16) are independent of $p$. Since they agree modulo any large enough $p$, it follows that they must be equal (for each fixed value of $n$). Accordingly, we have the identity

$$\sum_{j=1}^{d} \lambda_j^{n+1} p_j(n) = c_0 \left[ A_0(n) + \sum_{j=1}^{d} \lambda_j^n p_j(n) \right] \quad \text{for all } n \geq 0. \tag{8.17}$$

Note that both sides of (8.17) are C-finite sequences so that, because the representation (8.15) is unique, we must have, in particular, $c_0 A_0(n) = 0$. If $c_0 = 0$ then it follows by comparison with the left-hand side of (8.17) that $d = 0$ so that $A(n) = A_0(n)$ with the single characteristic root $\lambda = 0$. In the other case, that is if $c_0 \neq 0$, we have $A_0(n) = 0$, so $0$ is not a characteristic root. Further comparing both sides of (8.17), we find that $\lambda_j = c_0$ for all $j$. Since the $\lambda_j$ are distinct, we conclude that $d = 1$ so that $A(n) = \lambda_1^n p_1(n)$ with the single characteristic root $\lambda_1 \in \mathbb{Q}^\times$. □

We now extend Proposition 8.2 to the case of $r$-term $\mathbb{Q}$-linear combinations of constant terms, thus proving our main result Theorem 8.3. We recall that its statement is that a C-finite $A(n)$ sequence is an $r$-term $\mathbb{Q}$-linear combination of constant terms if and only if it has at most $r$ characteristic roots, all of which are rational.

*Proof of Theorem 8.3.* The case $r = 1$ is proved by Proposition 8.2. With the same argument as in (8.14) it follows that any C-finite sequence with $r$ characteristic roots, all of which are rational, can be represented as a linear combination of $r$ constant terms.

Therefore, suppose that $r > 1$ and that

$$A(n) = \operatorname{ct}[P_1(\boldsymbol{x})^n Q_1(\boldsymbol{x})] + \cdots + \operatorname{ct}[P_r(\boldsymbol{x})^n Q_r(\boldsymbol{x})]$$

is an $r$-term $\mathbb{Q}$-linear combination of constant terms with $P_j, Q_j \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$. We need to show that $A(n)$ has at most $r$ characteristic roots, all of which are rational. As in the proof of Proposition 8.2, we find that all characteristic roots of $A(n)$ are rational and that $A(n)$ can be represented in the form (8.15) with $p_j(n) \in \mathbb{Q}[n]$.

Let $c_j = \operatorname{ct}[P_j(\boldsymbol{x})] \in \mathbb{Q}$. It follows from Lemma 8.7 that

$$A(p+n) \equiv c_1 \operatorname{ct}[P_1(\boldsymbol{x})^n Q_1(\boldsymbol{x})] + \cdots + c_r \operatorname{ct}[P_r(\boldsymbol{x})^n Q_r(\boldsymbol{x})] \pmod{p}$$

for all $n \geq 0$ and all large enough primes $p$. On the other hand, for large $p$, by Fermat's little theorem,

$$A(p+n) \equiv \sum_{j=1}^{d} \lambda_j^{n+1} p_j(n) \pmod{p}.$$

Note that the right-hand sides of the last two congruences are independent of $p$. Since the congruences hold modulo all large enough primes, we conclude that

$$\sum_{j=1}^{d} \lambda_j^{n+1} p_j(n) = c_1 \operatorname{ct}[P_1(\boldsymbol{x})^n Q_1(\boldsymbol{x})] + \cdots + c_r \operatorname{ct}[P_r(\boldsymbol{x})^n Q_r(\boldsymbol{x})].$$

Note that the sequence

$$B(n) \coloneqq \sum_{j=1}^{d} \lambda_j^{n+1} p_j(n) - c_1 A(n) = \sum_{j=1}^{d} (\lambda_j - c_1) \lambda_j^n p_j(n) - c_1 A_0(n)$$

is C-finite and is an $(r-1)$-term $\mathbb{Q}$-linear combination of constant terms. By induction, we may conclude that $B(n)$ has at most $r-1$ characteristic roots, all of which are rational. By comparison with (8.15), we see that $A(n)$ has at most one more characteristic root than $B(n)$. Thus $A(n)$ has at most $r$ characteristic roots, which is what we had to show. □

Theorem 8.3 classifies those rational recursive sequences with constant coefficients which can be represented as a linear combination of $r$ constant terms. In particular, a rational C-finite sequence $A(n)$ is a linear combination of constant terms if and only if all of its characteristic roots are rational. It is natural to wonder whether we can restrict to integer sequences and conclude that all characteristic roots must be integral. This can be achieved by using the following proposition[2], that we could not locate in the vast literature on C-finite sequences.

**Proposition 8.13.** *Let $A(n)$ be a C-finite sequence with characteristic roots $\lambda_1, \ldots, \lambda_d \in \mathbb{Q}$. If $A(n)$ is an integer sequence, then $\lambda_1, \ldots, \lambda_d \in \mathbb{Z}$.*

*Proof.* By assumption, $A(n)$ is equal to $\sum_{i=1}^{d} p_i(n)\lambda_i^n$, where the $\lambda_i$'s are mutually distinct rational numbers and the $p_i(x)$'s are polynomials in $\mathbb{Q}[x]$. We will prove that if for some $N \in \mathbb{Z} \setminus \{0\}$ we have that $A(n) \in \frac{1}{N}\mathbb{Z}$ for all $n$, then all the $\lambda_i$'s are integers.

Let us start with the observation that this is true if $d = 1$; indeed, if $\lambda_1 = a/b$ with coprime integers $a, b$, and $p_1(n) = q(n)/c$ with $q(x) \in \mathbb{Z}[x]$ and $c \in \mathbb{Z}$, then the assumption "$A(n) = p_1(n)\lambda_1^n \in \frac{1}{N}\mathbb{Z}$ for all $n$" implies that $b^n$ divides $Nq(n)$ for all $n$, hence $b = 1$.

Let us now treat the general case. Denote by $V$ the $d \times d$ Vandermonde matrix attached with the $\lambda_i$'s, that is $V = (\lambda_i^{j-1})_{1 \le i,j \le d}$. Since the $\lambda_i$'s are mutually distinct, $V$ is in $\mathrm{GL}_d(\mathbb{Q})$. Therefore, the equality

$$[p_1(n)\lambda_1^n, \ldots, p_d(n)\lambda_d^n] \cdot V = [A(n), \ldots, A(n+d-1)]$$

implies that each term $p_i(n)\lambda_i^n$ is equal to $1/\det(V)$ times a linear combination of the integers $A(n), \ldots, A(n+d-1)$ with fixed rational coefficients. In other terms, each $p_i(n)\lambda_i^n$ is in $\frac{1}{N_i}\mathbb{Z}$ for some $N_i \in \mathbb{Z} \setminus \{0\}$ independent of $n$. By the case $d = 1$, this implies that all the $\lambda_i$'s are integers. □

We conclude this section with the following immediate consequence of Proposition 8.2, Theorem 8.3 and Proposition 8.13:

**Corollary 8.14.** *Let $A(n) \in \mathbb{Z}$ be a C-finite sequence. $A(n)$ is a constant term if and only if it has a single characteristic root $\lambda$ and $\lambda \in \mathbb{Z}$. More generally, $A(n)$ is an $r$-term $\mathbb{Q}$-linear combination of constant terms if and only if it has at most $r$ distinct characteristic roots, all of which are integral.*

## 8.5 An analog of Minton's theorem

In this section, we record the following result which, though having a much simpler proof, is pleasingly similar to Theorem 8.5 due to Minton [246]. Moreover, this result gives a classification of constant term sequences of the form $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n]$ among all constant term sequences $\mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$.

**Proposition 8.15.** *Suppose $A(n) = \mathrm{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ with $P, Q \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$. Then the following are equivalent:*

[2]The proof of Prop. 8.13 was communicated by Carlo Sanna (Politecnico di Torino).

1. *For all large enough primes $p$ and for all $r \geq 1$, $A(n)$ satisfies the Gauss congruences* (8.9).

2. *For all large enough primes $p$, $A(n)$ satisfies the congruences* (8.10).

3. $A(n) = A(0)\operatorname{ct}[P(\boldsymbol{x})^n]$.

*Proof.* We conclude from Lemma 8.7 with $r = 1$ and $k = 0$ that

$$A(pn) \equiv A(0)\operatorname{ct}[P(\boldsymbol{x})^n] \pmod{p}$$

for large enough $p$ (namely, if $p > \deg(Q)$). If $A(n)$ satisfies the congruences (8.10), we find that, for large enough $p$,

$$A(n) \equiv A(0)\operatorname{ct}[P(\boldsymbol{x})^n] \pmod{p}.$$

In that case, since this congruence holds modulo infinitely many $p$, we conclude the equality $A(n) = A(0)\operatorname{ct}[P(\boldsymbol{x})^n]$. Thus the third condition follows from the second.

To complete the proof, we need to show that the third condition implies the first. This follows from Lemma 8.7 with $k = 0$ and $Q = 1$. □

**Remark 8.16.** Note that Proposition 8.15 does not imply that if $A(n) = \operatorname{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ satisfies the Gauss congruences (8.9) for large enough primes, then $Q$ must be constant. For instance, for any $P(x) \in \mathbb{Z}[x^{\pm 1}]$, the constant terms $\operatorname{ct}[P(x^2)^n(1+x)] = \operatorname{ct}[P(x^2)^n] = \operatorname{ct}[P(x)^n]$ satisfy the Gauss congruences for all primes $p$, even though the first constant term has a non-constant $Q$. Proposition 8.15 rather shows that if (a) or (b) are fulfilled, then $Q$ can be replaced by $\operatorname{ct}[Q]$.

## 8.6 Hypergeometric constant terms

Exiting the class of C-finite sequences, we find it natural to ask (Question 5 in the introduction): Which hypergeometric sequences[3] $A(n)$ are constant term sequences?

The reason for the specialization to hypergeometric sequences is threefold. First, it can be argued that it is the easiest P-recursive case. Second, similar to constant terms, hypergeometric sequences are not stable under addition. Finally, as we will see below in Lemma 8.17, the congruences proven in Section 8.3 behave nicely with the hypergeometric assumption.

It follows from Lemma 8.7 (specialized to $n = 1$ and $r = 1$) that if $A(n) = \operatorname{ct}[P(\boldsymbol{x})^n Q(\boldsymbol{x})]$ with $P, Q \in \mathbb{Q}[\boldsymbol{x}^{\pm 1}]$, then

$$A(p+k) \equiv A(k)\operatorname{ct}[P(\boldsymbol{x})] \pmod{p}$$

for all integers $k \geq 0$, provided that $p > \deg(P^k Q)$ and $P, Q \in \mathbb{Z}_p[\boldsymbol{x}^{\pm 1}]$. In other words, if $A(n)$ is a constant term, then there exists a constant $c \in \mathbb{Q}$ such that, for each $k \in \mathbb{Z}_{\geq 0}$, the congruences

$$A(p+k) \equiv A(k) \cdot c \pmod{p} \tag{8.18}$$

[3]Recall that a sequence $A(n)$ is hypergeometric if it satisfies a first-order recurrence $\alpha(n)A(n+1) = \beta(n)A(n)$ for some polynomials $\alpha(n), \beta(n) \in \mathbb{Q}[n]$. For our purposes we will assume $\alpha(n) \neq 0$ for all $n \geq 0$.

hold for all large enough $p$. We shall now show that, for hypergeometric sequences, the congruences (8.18) follow from the base case $k = 0$.

**Lemma 8.17.** *Let $A(n)$ be a hypergeometric sequence. Suppose that there exists a constant $c \in \mathbb{Q}$ such that*

$$A(p) \equiv c \pmod{p} \tag{8.19}$$

*for all large enough $p$. Then, for each $k \in \mathbb{Z}_{\geq 0}$, the congruence* (8.18) *holds for sufficiently large $p$.*

*Proof.* Since $A(n)$ is hypergeometric, we have $A(n+1) = \rho(n)A(n)$ for a rational function $\rho(n) = \beta(n)/\alpha(n)$ with $\alpha(n), \beta(n) \in \mathbb{Z}[n]$. Fix $k \in \mathbb{Z}_{\geq 0}$ and suppose that the congruence (8.18) holds for all large enough $p$. By applying the hypergeometric recurrence twice, we obtain

$$A(p+k+1) = \rho(p+k)A(p+k) \equiv \rho(k)cA(k) = cA(k+1) \pmod{p},$$

which is (8.18) with $k+1$ in place of $k$. Here we used that $\rho(p+k) \equiv \rho(k) \pmod{p}$, which holds true provided that $\alpha(k) \not\equiv 0 \pmod{p}$. The latter is true for all sufficiently large $p$ since, by assumption, $\alpha(k) \neq 0$. The claim therefore follows by induction on $k$. □

**Remark 8.18.** Note that Lemma 8.17 does not hold for non-hypergeometric sequences in general. For instance, it does not hold for the Lucas numbers $L(n)$ as defined in (8.8). These form a trace sequence so that, by Minton's Theorem 8.5, the Gauss congruences (8.9) are satisfied. In the case $n = 1$, these imply the congruences (8.19). However, the Lucas numbers do not satisfy the congruences (8.18) for $k > 0$.

Lemma 8.7 gives a necessary condition for $A(n)$ to be a constant term sequence. It is natural to wonder whether, or to which degree, this condition is sufficient: Is an integer hypergeometric sequence $A(n)$ that satisfies the congruences (8.19) a constant term? Natural sources of potential counterexamples to this question are families of integer sequences that are quotients of binomial coefficients but cannot be written as products of those, for example $A(n) = \binom{8n}{4n}\binom{4n}{n}\binom{2n}{n}^{-1}$ (see [46, Thm. 1.2]).

We recall that the corresponding question for diagonals (8.1) (namely, to classify which hypergeometric sequences $A(n)$ are coefficients of diagonals) also remains open. The following conjecture due to Christol [101, Conjecture 4, p. 55] attempts such a classification. In its statement, we call a sequence $(A(n))_{n\geq 0}$ *almost integral* if there exists a positive integer $K$ such that $K^{n+1}A(n) \in \mathbb{Z}$ for all integers $n \geq 0$. An almost integral sequence with (at most) geometric growth is called *globally bounded*.

**Conjecture 8.19** ([101])**.** *Let $(A(n))_{n\geq 0}$ be sequence of rational numbers. The generating function $\sum_{n\geq 0} A(n)t^n$ is the diagonal of a rational function if and only if $(A(n))_{n\geq 0}$ is P-recursive and globally bounded.*

Any hypergeometric sequence is P-recursive since it satisfies, by definition, a recurrence with polynomial coefficients of order one. Moreover, thanks to a result of Christol [99], [101] it is easy to check when a hypergeometric sequence is integral (in the case when $\alpha(n)$ and $\beta(n)$ in the definition split in $\mathbb{Q}[n]$). This makes hypergeometric sequences a natural

source of potential counterexamples to Conjecture 8.19. We refer to [53], [1] and [82] for recent progress in this area. Here, we only mention that even for

$$A(n) = \frac{\left(\frac{1}{9}\right)_n \left(\frac{4}{9}\right)_n \left(\frac{5}{9}\right)_n}{n!^2 \left(\frac{1}{3}\right)_n} \tag{8.20}$$

the conjecture is open. In other words, it is an open question whether the sequence (8.20) is the diagonal of a rational function. On the other hand, we will show in this section that (8.20) is not a constant term. Before doing so, we first prove the following result answering Question 5 for a special family of hypergeometric sequences.

**Lemma 8.20.** *Let $m \geq 2$ be an integer and consider the sequence*

$$A_m(n) = \frac{\left(\frac{1}{m}\right)_n \left(1 - \frac{1}{m}\right)_n}{n!^2}. \tag{8.21}$$

1. *$A_m(n)$ is a diagonal for all $m \geq 2$.*
2. *$A_m(n)$ is a constant term if and only if $m \in \{2, 3, 4, 6\}$.*

Note that the classification in Lemma 8.20 suggests that constant term sequences are special among diagonals and often have significant additional arithmetic properties. Indeed, the cases $m \in \{2, 3, 4, 6\}$ (see A002894, A006480, A000897 and A113424 in the on-line encyclopedia of integer sequences [297]) correspond precisely to those special hypergeometric functions underlying Ramanujan's theories of elliptic functions ($m = 2$ being the classical case and $m = 3, 4, 6$ corresponding the alternative bases). We refer to [34] for more information.

**Example 8.21.** The hypergeometric sequence

$$B(n) = 5^{3n} \frac{\left(\frac{1}{5}\right)_n \left(\frac{4}{5}\right)_n}{n!^2} = 1, 20, 1350, 115500, 10972500, \ldots \tag{8.22}$$

is an integer sequence and grows at most exponentially. As suggested by Christol's Conjecture 8.19 and stated in Lemma 8.20, the sequence $B(n)$ is a diagonal. However, $B(n)$ is not a constant term. The proof of Lemma 8.20 in this case proceeds by showing that we have the congruences

$$B(p) \equiv \begin{cases} 20, & \text{if } p \equiv \pm 1 \bmod 5, \\ 30, & \text{otherwise}, \end{cases} \pmod p,$$

which contradict Lemma 8.17.

*Proof of Lemma 8.20.* Part (a) follows from the fact that the generating function of $A_m(n)$ is the Hadamard (term-wise) product of $(1-x)^{-1/m}$ and $(1-x)^{1/m-1}$. The latter are algebraic functions and hence diagonals by a result of Furstenberg [145]. Since diagonals are closed under Hadamard products [100], it follows that $A_m(n)$ is a diagonal.

That $A_m(n)$ is a constant term if $m \in \{2,3,4,6\}$ follows from the following alternative representations as products of binomial coefficients:

$$\begin{aligned}
2^{4n}A_2(n) &= \frac{(2n)!^2}{n!^4} = \binom{2n}{n}^2,\\
3^{3n}A_3(n) &= \frac{(3n)!}{n!^3} = \binom{3n}{2n}\binom{2n}{n},\\
4^{3n}A_4(n) &= \frac{(4n)!}{(2n)!n!^2} = \binom{4n}{2n}\binom{2n}{n},\\
2^{4n}3^{3n}A_6(n) &= \frac{(6n)!}{(3n)!(2n)!n!} = \binom{6n}{3n}\binom{3n}{n}.
\end{aligned}$$

In the remainder, we will show that $A_m(n)$ is not a constant term if $m \notin \{2,3,4,6\}$. If $m$ is coprime to $p$ (as it is for large enough $p$) then the right-hand side of

$$m^p \left(\tfrac{1}{m}\right)_p = 1 \cdot (m+1)(2m+1)\cdots((p-1)m+1)$$

is a product of all the residues modulo $p$. In particular, exactly one factor is of the form $ap$ where $a \in \{1,2,\ldots,m-1\}$ is characterized by $ap \equiv 1 \pmod m$. By Wilson's theorem, we therefore have

$$m^p \left(\tfrac{1}{m}\right)_p \equiv -ap \pmod{p^2}$$

or, equivalently,

$$\frac{m^p \left(\frac{1}{m}\right)_p}{p!} \equiv a \pmod p.$$

Similarly,

$$m^p \left(1-\tfrac{1}{m}\right)_p = (m-1)(2m-1)\cdots(pm-1)$$

and, again, the right-hand side features a product of all residues modulo $p$. Exactly one factor is of the form $bp$ where $b \in \{1,2,\ldots,m-1\}$ is characterized by $bp \equiv -1 \pmod m$. It follows that $b = m-a$. Combined, we conclude that

$$m^{2p}A_m(p) = \frac{m^{2p}\left(\frac{1}{m}\right)_p \left(1-\frac{1}{m}\right)_p}{p!^2} \equiv a(m-a) \pmod p. \tag{8.23}$$

Since $a \in \{1,2,\ldots,m-1\}$ is characterized by $a \equiv 1/p \pmod m$ it, in particular, depends only on the residue class of $p$ modulo $m$. As $p$ ranges through all primes, it follows from Dirichlet's theorem on primes in arithmetic progressions, that each value $a \in \{1,2,\ldots,m-1\}$ with $a$ coprime to $m$ appears infinitely many times. There are $\phi(m)$ many such values of $a$, where $\phi$ is Euler's totient function. Consequently, the quantity $a(m-a)$ on the right-hand side of (8.23) takes $\phi(m)/2$ many different values as $p$ ranges through all primes $p > m$.

On the other hand, if $A_m(n)$ is a constant term sequence, then by (8.19) there exists a constant $c \in \mathbb{Q}$ such that $m^{2p}A_m(p) \equiv c \pmod p$ for all large enough primes. If (8.23) holds for infinitely many $p$, we necessarily have $c = a(m-a)$, which is only possible if $\phi(m)/2 = 1$.

Thus, if $\phi(m) > 2$ then $m^{2p}A_m(p)$ cannot satisfy the congruences (8.19) for all large enough primes and, hence, the sequences $m^{2n}A_m(n)$ and $A_m(n)$ cannot be constant terms. Since $\phi(m) > 2$ for all integers $m \geq 2$ except for $m \in \{2,3,4,6\}$, the claim follows. □

For hypergeometric sequences, we therefore have the following inclusions

$$\{\text{constant terms}\} \subsetneq \{\text{diagonals}\} \subseteq \{\text{P-recursive \& globally bounded seq's}\}\,.$$

We note that these inclusions are also true for C-finite as well as for P-recursive sequences. An example for the strictness of the first inclusion in the realm of hypergeometric sequences is given by the sequence (8.22) and in the class of C-finite sequences by the Fibonacci numbers. The second inclusion is a consequence of a result due to Lipshitz [229] and it is strict if and only if Christol's Conjecture 8.19 (restricted to hypergeometric sequences) is false. A potential candidate of a globally bounded hypergeometric sequence that is not a diagonal is sequence (8.20). We now show that this sequence is not a constant term.

**Lemma 8.22.** *The hypergeometric sequence $A(n)$ defined in* (8.20) *is not a constant term sequence.*

*Proof.* Proceeding as in the proof of Lemma 8.20, we find

$$m^p \left(\tfrac{r}{m}\right)_p = r(m+r)\cdots((p-1)m+r),$$

where the right-hand side is a product over all residues modulo $p$. Exactly one factor is of the form $ap$ where $a \in \{1,2,\ldots,m-1\}$ is characterized by $ap \equiv r \pmod m$. In that case,

$$\frac{m^p \left(\tfrac{r}{m}\right)_p}{p!} \equiv a \pmod p.$$

If $p \equiv 1 \pmod 9$, we therefore find

$$\frac{9^p \left(\tfrac{1}{9}\right)_p}{p!} \equiv \frac{3^p \left(\tfrac{1}{3}\right)_p}{p!} \equiv 1, \quad \frac{9^p \left(\tfrac{4}{9}\right)_p}{p!} \equiv 4, \quad \frac{9^p \left(\tfrac{5}{9}\right)_p}{p!} \equiv 5 \pmod p,$$

which combine to

$$3^{5p} A(p) = 3^{5p} \frac{\left(\tfrac{1}{9}\right)_p \left(\tfrac{4}{9}\right)_p \left(\tfrac{5}{9}\right)_p}{p!^2 \left(\tfrac{1}{3}\right)_p} \equiv \frac{1\cdot 4\cdot 5}{1} = 20 \pmod p.$$

On the other hand, if $p \equiv -1 \pmod 9$, then

$$\frac{9^p \left(\tfrac{1}{9}\right)_p}{p!} \equiv 8, \quad \frac{3^p \left(\tfrac{1}{3}\right)_p}{p!} \equiv 2, \quad \frac{9^p \left(\tfrac{4}{9}\right)_p}{p!} \equiv 5, \quad \frac{9^p \left(\tfrac{5}{9}\right)_p}{p!} \equiv 4 \pmod p,$$

which combine to

$$3^{5p} A(p) = 3^{5p} \frac{\left(\tfrac{1}{9}\right)_p \left(\tfrac{4}{9}\right)_p \left(\tfrac{5}{9}\right)_p}{p!^2 \left(\tfrac{1}{3}\right)_p} \equiv \frac{8\cdot 4\cdot 5}{2} = 80 \pmod p.$$

As in the proof of Lemma 8.20 we conclude that $3^{5n} A(n)$ and, hence, $A(n)$ cannot be a constant term. □

# Chapter 9

# On Rupert's problem

*"Though this be madness, yet there is a method in't."*
William Shakespeare, *Hamlet*, Act II, Scene 2.

A polyhedron $\mathbf{P} \subset \mathbb{R}^3$ has Rupert's property if a hole can be cut into it, such that a copy of $\mathbf{P}$ can pass through this hole. There are several works investigating this property for some specific polyhedra: for example, it is known that all 5 Platonic and 9 out of the 13 Archimedean solids admit Rupert's property. A commonly believed conjecture states that every convex polyhedron is Rupert. In this chapter prove that Rupert's problem is algorithmically decidable for polyhedra with algebraic coordinates. We also design a probabilistic algorithm which can efficiently prove that a given polyhedron is Rupert. Using this algorithm we not only confirm this property for the known Platonic and Archimedean solids, but also prove it for one of the remaining Archimedean polyhedra and many others. Moreover, we significantly improve on almost all known Nieuwland numbers and finally conjecture, based on statistical evidence, that the Rhombicosidodecahedron is in fact *not* Rupert.

This chapter of the dissertation combines the joint works with J. Steininger [302, 303].

## 9.1 Introduction

Undoubtedly the following fact is surprising when being first encountered with:

> *It is possible to cut a hole in the unit cube such that another unit cube can pass through it.*

Indeed, Prince Rupert of the Rhine won a wager in the 17th century by betting on the validity of this claim. An elegant and simple way to see why this assertion is true is presented in Figure 9.1; indeed, it is easy to verify that the projection of the unit cube in the direction of a main diagonal yields a regular hexagon of side length $\sqrt{2/3}$ and the unit square (a different projection of the cube) fits inside that hexagon. These two observations are already enough to win Rupert's bet, however at the same time they also open a whole world of interesting questions, conjectures and studies.

For instance, a subsequent natural question was investigated by Pieter Nieuwland a century after Prince Rupert's death:

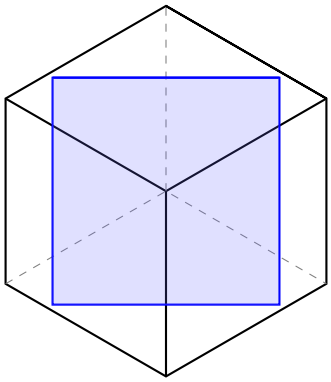


Figure 9.1: The unit square fits inside the regular hexagon of side length $\sqrt{2/3}$.

*How large can the second cube maximal be in order to still fit inside a hole of the unit cube?*

Quite surprisingly, Nieuwland could show that in terms of this question, the solution presented in Figure 9.1 is not optimal. If viewed from a slightly different angle, a "tunnel" inside the unit cube can be constructed such that a cube with side length $1.06$ can be moved through it[1]. Nieuwland could even find the exact maximal side length of the "fitting" cube which turns out to be $3\sqrt{2}/4 \approx 1.06066$ (for a proof of this fact see [40]); later this constant was given the name *Nieuwland's constant*.

Analogously to the cube, Rupert's property can be defined for any polyhedron in $\mathbb{R}^3$. A somewhat imprecise definition of this property is: a polyhedron $\mathbf{P} \subset \mathbb{R}^3$ has Rupert's property if a hole (with the shape of a straight tunnel) can be cut into it such that a copy of $\mathbf{P}$ can be moved through this hole. In the next section we will first recall a rigorous (but rather non-transparent) definition from [191] (see Definition 9.3) and then an easy end explicit reformulation using projections to $\mathbb{R}^2$ (in the spirit of Figure 9.1). In the same way, the *Nieuwland number* can also be generalized for any polyhedron $\mathbf{P}$, see Definition 9.5.

For a historic overview on these questions we refer to [287]; for more recent contributions see [290, 191, 94, 180, 226]. Scriba showed in 1968 that the Tetrahedron and Octahedron have Rupert's property. Half a century later and hence already quite recently, Jerrard, Wetzel and Yuan, the authors of the second paper, built on Scriba's work and investigated Rupert's property of Platonic solids further: they could prove that all five of them are Rupert. Moreover, they also gave lower bounds on Nieuwland numbers for them. One year later Chai, Yuan and Zamfirescu looked at Archimedean solids from "Rupert's perspective", showed that 8 out of 13 have Rupert's property and also provided lower bounds for the corresponding Nieuwland numbers. Finally another year later, Hoffmann [180] and Lavau [226] showed in 2019 Rupert's property for the Truncated tetrahedron, thus enhancing the number to 9 out of 13. Theorem 9.13 in this chapter "resolves" the Truncated icosidodecahedron, pushing the number of settled down Archimedean solids to 10.

After the submission of [303], a preprint [315] by Tonpho and Wichiramala appeared on the internet in which the authors study Rupert's problem in $n$ dimensions and also quote results from a master thesis by Tonpho [314] from 2018. There, a relatively similar to parts of [303] (but solely numerical) approach is used to find solutions to Rupert's problem for all

[1]The side length of the blue square in Figure 9.1 is at most $\sqrt{6} - \sqrt{2} \approx 1.0353$.

Platonic and some Archimedean solids. Even more recently, Fredriksson [140] built on the ideas of our work, applied the algorithm for placements of convex polygons from [10], and used non-linear optimization techniques like SLSQP and Nelder-Mead in order to obtain new results. Most notably, he was able to prove that the Catalan solids Triakis tetrahedron and Pentagonal icositetrahedron have Rupert's property.

**Contribution and structure of the chapter**
In Section 9.2 we introduce the necessary elementary definitions and concepts. We rigorously define Rupert's property of a polyhedron **P** and then show that it is equivalent to the existence of a septuple of real numbers satisfying a simple property depending on **P**. In the same section we recall the notion of the Nieuwland number of a polyhedron.

Contrary to the existing methods for proving that a polyhedron has Rupert's property, we present a new algorithmic approach to this problem in Section 9.3. Roughly speaking, our probabilistic (*Las Vegas* type) algorithm draws pairs of random projections of a given polyhedron and decides whether the chosen directions yield a solution – if they do not, the algorithm draws another pair, and so on. Moreover, by constructing a deterministic algorithm, we also prove that Rupert's question for most interesting polyhedra is algorithmically decidable. However, we also infer that at least for now this algorithm is only of theoretical value, since it is not yet practical because of its bad complexity. In the same section we explain a simple algorithmic idea which allows to significantly improve on known lower bounds for the Nieuwland numbers. Finally, we also define the concept of the *Rupertness*, measuring the likelihood for finding a solution to Rupert's problem of a (centrally symmetric) polyhedron.

It turns out that in practice our probabilistic approach finds solutions to Rupert's problem very efficiently: all 5 Platonic and 10 Archimedean solids can be resolved in less than one minute on a regular computer. We present our new explicit results in Section 9.4: we prove that the Truncated icosidodecahedron is Rupert (Theorem 9.13), show this property for many Catalan and Johnson solids (Theorems 9.14 and 9.15), and significantly improve on all known Nieuwland numbers (Table 9.1), except the Cube, Octahedron and Cuboctahedron. As mentioned, the Nieuwland constant for the Cube is proven to be optimal and the Nieuwland numbers for the Octahedron and Cuboctahedron are conjectured to be optimal as well [191, p. 91]; our findings support this conjecture.

In [191, p. 87] the authors suggest the possible non-existence of "non-Rupert" convex polyhedra in $\mathbb{R}^3$ and infer that in any case any such example would be of considerable interest. The authors of [94] go even further and state the following conjecture.

**Conjecture 9.1** (Chai, Yuan, Zamfirescu [191, 94])**.** *Every convex polyhedron has Rupert's property.*

Also in Section 9.4 we provide statistical evidence for a counter-example to this conjecture (Conjecture 9.16).

Appendix 9.5 contains our solutions for Platonic, Archimedean and Catalan solids and corresponding lower bounds for the Nieuwland numbers. All these solutions are given in a uniform way in one table. Together with the exact coordinates for the Platonic and Archimedean polyhedra we used (also in the appendix) these solutions can be easily verified by the reader; for the coordinates of Catalan solids we refer to the wonderful website

www.dmccooey.com/polyhedra/. For the reader's convenience, we also provide our source code which written in the programming language R and the computer algebra software Maple: www.github.com/Vog0/RupertProblem.

## 9.2 Preliminaries

In order to avoid confusion, let us first collect some elementary definitions.

**Definition 9.2.** The following classical notions we will use throughout the text:

- A *polyhedron*, in this text usually denoted by **P** or **Q**, is a finite non-degenerate set of points in $\mathbb{R}^3$ in convex position. We denote by $\overline{\mathbf{P}}$ the smallest convex set containing all points of **P** (i.e. including the interior) and by $\mathbf{P}^\circ$ its interior.
- A *polygon*, usually denoted by $\mathcal{P}$ or $\mathcal{Q}$, is a finite set of points in $\mathbb{R}^2$ that not all lie on the same line. Similar to polyhedra, we denote by $\overline{\mathcal{P}}$ the convex hull of $\mathcal{P}$ and by $\mathcal{P}^\circ$ the interior of $\overline{\mathcal{P}}$.
- We call $\Sigma$ the set of isometries of $\mathbb{R}^2$ that do not include reflections, i.e. length preserving mappings from $\mathbb{R}^2$ onto itself not including reflections. It is well-known that any element $\sigma \in \Sigma$ can be represented by a rotation about the origin followed by a translation, and also the other way around. We will let $\sigma \in \Sigma$ act on a set of points in the plane elementwise. Furthermore, we parametrize all translations of $\mathbb{R}^2$ by $T_{x,y}$: $\mathbb{R}^2 \to \mathbb{R}^2$,
$$T_{x,y}((a,b)^t) \coloneqq (a+x, b+y)^t.$$
Similarly, the rotation mapping $R_\alpha$: $\mathbb{R}^2 \to \mathbb{R}^2$ is defined by
$$R_\alpha((a,b)^t) \coloneqq \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a\cos(\alpha) - b\sin(\alpha) \\ a\sin(\alpha) + b\cos(\alpha) \end{pmatrix}.$$
Clearly, $T_{x,y}$ translates points in $\mathbb{R}^2$ by the vector $(x,y)^t$ and $R_\alpha$ rotates a point counterclockwise by an angle $\alpha$ about the origin.
- We say that a polygon $\mathcal{P}$ *lies inside* a polygon $\mathcal{Q}$ if $\mathcal{P} \subset \mathcal{Q}^\circ$. Moreover, we say that a polygon $\mathcal{P}$ *fits in* a polygon $\mathcal{Q}$ if there exists an isometry $\sigma \in \Sigma$ such that $\sigma(\mathcal{P})$ lies inside $\mathcal{Q}$.
- A polyhedron **P** is called *centrally symmetric* with respect to $O_\mathbf{P} \in \mathbb{R}^3$ if for each $A \in \mathbf{P}$, the point $2O_\mathbf{P} - A$ belongs to **P**. Analogously, a polygon $\mathcal{P}$ is centrally symmetric about $O_\mathcal{P} \in \mathbb{R}^2$ if $2O_\mathcal{P} - A \in \mathcal{P}$ for each point $A \in \mathcal{P}$. A polyhedron or a polygon is called *point symmetric* if it is centrally symmetric with respect to some point.

Usually Rupert's property is explained as follows: a polyhedron **P** is Rupert if a hole with the shape of a straight tunnel can be cut into it such that a copy of **P** can be moved through this hole. While this definition explains well why this notion is geometrically intriguing, it is admittedly not quite mathematically precise. A rigorous definition is given for example in [191] and we will state here a slightly reformulated version. First, let us set the notion of

a *set with a hole*: we mean a set of points in $\mathbb{R}^3$ whose interior is connected but not simply connected. Given a polyhedron **Q**, we may move it along a straight line in the direction of a vector $v \in \mathbb{R}^3$; taking the convex hull of the union of all these translations we obtain the set $\{\overline{\mathbf{Q}} + tv \in \mathbb{R}^3 \colon t \in \mathbb{R}\}$. Rupert's property of a polyhedron may be defined as follows (see [191]).

**Definition 9.3** (Rupert's property)**.** A polyhedron[2] **P** has *Rupert's property* (or **P** is *Rupert*) if there exists a polyhedron **Q** of the same shape and size as **P** and a vector $v \in \mathbb{R}^3$ such that $\overline{\mathbf{P}} \setminus \{\overline{\mathbf{Q}} + tv \in \mathbb{R}^3 \colon t \in \mathbb{R}\}$ is a set with a hole. *Rupert's problem* is the task to decide whether a given polyhedron is Rupert.

Luckily, the definition of Rupert's property can be reformulated in a much easier criterion on the level of projections to the plane $\mathbb{R}^2$. The idea is that looking from the direction of the vector $v$ in the definition above, we must see the two shadows (normal projections) of the polyhedra **P**, **Q** as two polygons $\mathcal{P}, \mathcal{Q}$, one lying inside the other: $\mathcal{P} \subset \mathcal{Q}^\circ$. This is the core of Theorem 1 in [191] and the reason why Figure 9.1 in the introduction is a proof that the Cube is Rupert. Now we will make this idea even more explicit.

As we are dealing with projections, we first parametrize the set of all those. We define the mapping $X : [0, 2\pi) \times [0, \pi] \to \{x \in \mathbb{R}^3 : \|x\| = 1\}$ by

$$X(\theta, \varphi) \coloneqq (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)^t. \tag{9.1}$$

This gives a way to parametrize the points on the 3-dimensional sphere in terms of two unknowns. It is well-known that drawing $\theta$ uniformly on $(0, 2\pi)$, that is $\theta \sim U(0, 2\pi)$, and $\varphi \sim \arccos(U(-1, 1))$ results in a uniformly distributed $X(\theta, \varphi)$ on the unit sphere.

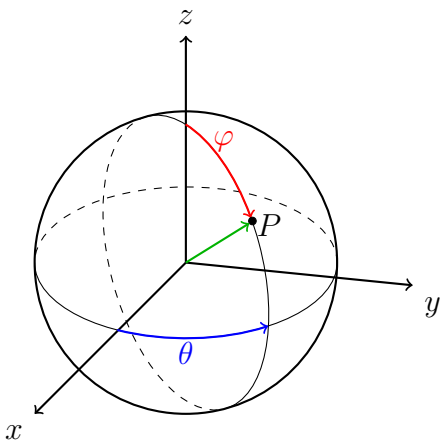



Figure 9.2: Meaning of $\theta$ and $\varphi$ in (9.1) in spherical coordinates.

It follows that a projection onto a plane orthogonal to $X(\theta, \varphi)$ can be given by

$$M_{\theta,\varphi} \coloneqq \begin{pmatrix} -\sin(\theta) & \cos(\theta) & 0 \\ -\cos(\theta)\cos(\varphi) & -\sin(\theta)\cos(\varphi) & \sin(\varphi) \end{pmatrix}. \tag{9.2}$$

[2]Note that, as defined in Definition 9.2, a polyhedron in this text is always convex.

Like the mappings $R_\alpha, T_{x,y}$ we extend the map $M_{\theta,\varphi}$ to act on sets of points in $\mathbb{R}^3$ elementwise. Thus, all parallel projections of the vertices of a polyhedron **P** onto $\mathbb{R}^2$ can now be expressed as

$$(T_{x,y} \circ R_\alpha \circ M_{\theta,\varphi})(\mathbf{P}).$$

It follows that an equivalent characterization of Rupert's property for a polyhedron **P** is the existence of two quintuples of parameters $(x_i, y_i, \alpha_i, \theta_i, \varphi_i)$, $i = 1, 2$, such that

$$(T_{x_1,y_1} \circ R_{\alpha_1} \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset (T_{x_2,y_2} \circ R_{\alpha_2} \circ M_{\theta_2,\varphi_2})(\mathbf{P})^\circ.$$

In other words, the polygon on the left-hand side lies inside the polygon on the right-hand side and both polygons are obtained by some orthogonal projection, rotation and translation of **P**. Moreover, this condition can be rewritten as

$$(R_{-\alpha_2} \circ T_{x_1-x_2,y_1-y_2} \circ R_{\alpha_1} \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ.$$

Note, that $R_{-\alpha_2} \circ T_{x_1-x_2,y_1-y_2} \circ R_{\alpha_1}$ is an isometry on $\mathbb{R}^2$ and thereby may be expressed as the composition of a single rotation and a translation. Hence, we obtain the following equivalent characterization of Rupert's property.

**Proposition 9.4.** *A polyhedron* **P** *satisfies Rupert's property, if and only if there exist 7 parameters* $x, y \in \mathbb{R}$, $\alpha, \theta_1, \theta_2 \in [0, 2\pi)$ *and* $\varphi_1, \varphi_2 \in [0, \pi]$ *such that*

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ. \tag{9.3}$$

Clearly, any solution of Rupert's property can be translated into these 7 parameters and vice versa. Henceforth, we will encode a solution to Rupert's problem by a vector $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$.

Note that from Proposition 9.4 it is evident that Rupert's property is a statement about containment of points inside an open set. Since the projection, rotation and translation mappings are continuous, it follows that if there exists a solution to Rupert's problem $v = (x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$, then there must exist an open ball in $\mathbb{R}^7$ around $v$ of solutions. In other words, if a solution exists, then there is a set of solutions with positive (Lebesgue) measure. We will use this observation several times throughout the text.

Now let us recall the *Nieuwland number* of a polyhedron. If **P** is Rupert, then by the consideration above there exists a hole in it in which even a slightly larger copy of **P** can pass through. Naturally, one may ask for the largest polyhedron similar to **P** which can also be moved through such a hole. In other words, what is the largest (supremum) number $\nu$ for which there exists a copy of **P**, say **Q**, such that $\nu$**Q** can be moved in a straight tunnel through **P**? This number $\nu$ is called the *Nieuwland number* of **P**. It can be defined as in Definition 9.3, but in view of the more concrete and useful equivalent formulation in Proposition 9.4, we will define it directly via projections.

**Definition 9.5** (Nieuwland number)**.** The *Nieuwland number* $\nu = \nu(\mathbf{P})$ of a polyhedron **P** is the supremum over all $\mu \in \mathbb{R}$ for which there exist $x, y \in \mathbb{R}$, $\alpha, \theta_1, \theta_2 \in [0, 2\pi)$ and $\varphi_1, \varphi_2 \in [0, \pi]$ such that

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mu\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ. \tag{9.4}$$

Clearly, **P** is Rupert if and only if $\nu(\mathbf{P}) > 1$. We note that a typo in [191, p. 88] incorrectly states "$\geq$" in this inequality. In fact, $\nu(\mathbf{P}) \geq 1$ holds for every polyhedron, since if $\mu < 1$ in (9.4), one can take all 7 parameters to be equal to 0 (in other words $\mathbf{P} = \mathbf{Q}$) and the inclusion holds. As mentioned in the introduction, for a Cube **P** the Nieuwland number $\nu(\mathbf{P})$ is proven to be $3\sqrt{2}/4$.

Now we will prove that in the case when the polyhedron is point symmetric, the number of parameters in Proposition 9.4 can be reduced to 5. This significantly simplifies the algorithms in the next section in the point symmetric case.

**Proposition 9.6.** *The following two statements hold:*

1) *Let $\mathcal{P}$ and $\mathcal{Q}$ be convex polygons which are centrally symmetric around $O_\mathcal{P}$ and $O_\mathcal{Q}$ respectively. Then $\mathcal{P}$ fits in $\mathcal{Q}$ if and only there exists a $\sigma \in \Sigma$ such that $\sigma(\mathcal{P})$ lies inside $\mathcal{Q}$ and $\sigma(O_\mathcal{P}) = O_\mathcal{Q}$.*

2) *Let **P** by a polyhedron that is centrally symmetric about the origin. Then **P** satisfies Rupert's property if and only if there are 5 parameters $\alpha \in \mathbb{R}$, $\theta_i, \in [0, 2\pi)$ and $\varphi_i \in [0, \pi]$ for $i = 1, 2$ such that*
$$(R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ.$$

*Proof.* For the first statement it suffices to show that if there exists $\sigma \in \Sigma$ such that $\sigma(\mathcal{P})$ lies inside $\mathcal{Q}$, then there also exists $\sigma' \in \Sigma$ such that $\sigma'(\mathcal{P})$ is inside $\mathcal{Q}$ and $\sigma'(O_\mathcal{P}) = O_\mathcal{Q}$. Let $\tau$ be the translation in $\mathbb{R}^2$ which maps $\sigma(O_\mathcal{P})$ to $O_\mathcal{Q}$. We claim that $\sigma' = \tau \circ \sigma$ satisfies the required conditions.

Obviously, $\sigma'(O_\mathcal{P}) = \tau(\sigma(O_\mathcal{P})) = O_\mathcal{Q}$, hence we are left to show that $\sigma'(O_\mathcal{P})$ lies inside $\mathcal{Q}$. Let $\mathcal{P}'$ be the reflection of $\sigma(\mathcal{P})$ around $O_\mathcal{Q}$. Because $\sigma(\mathcal{P})$ is inside $\mathcal{Q}$ and $\mathcal{Q}$ is centrally symmetric, $\mathcal{P}'$ also lies inside $\mathcal{Q}$. Since $\mathcal{P}$ is centrally symmetric, it follows that $\mathcal{P}'$ can be obtained from $\sigma(\mathcal{P})$ by a translation. Moreover, $\sigma'(P)$ is given by the aritmetic mean between $\sigma(P)$ and this translation $\mathcal{P}'$. Now convexity of $\mathcal{Q}$ implies that $\sigma'(P)$ lies inside $\mathcal{Q}$.

We will now prove the second assertion. As both $M_{\theta,\varphi}$ and $R_\alpha$ are linear mappings, one has for any given point $p \in \mathbb{R}^3$ that

$$(R_\alpha \circ M_{\theta,\varphi})(-p) = -(R_\alpha \circ M_{\theta,\varphi})(p).$$

Therefore, any pair of antipodal points of the polyhedron is mapped to antipodal points in $\mathbb{R}^2$, resulting in a centrally symmetric polygon about the origin. So the claim follows from the first assertion. □

In order to keep the notation of Proposition 9.4, we will encode a solution to Rupert's problem of a point symmetric polyhedron by a 7-dimensional vector $(0, 0, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ as well.

## 9.3 The algorithms

In this section we present algorithmic ideas for proving or disproving that a given polyhedron **P** is Rupert. We start by introducing a naive algorithm which searches for a solution to

Rupert's problem for a given polyhedron. Then we gradually expand its sophistication and significantly improve the performance. Furthermore, we introduce a method for finding solutions with a high Nieuwland number. We note that all practical algorithms we present are probabilistic of *Las Vegas* type: If a solution is found, it is easy to check (rigorously) its correctness, however the search running time is probabilistic and cannot be known for sure in advance. We explain another viewpoint in §9.3.3, where we construct a deterministic algorithm, thus prove that Rupert's problem is algorithmically decidable. However, we also explain that in practice this algorithm is not (yet) useful. Finally, in §9.3.4 we introduce the probabilistic concept of the *Rupertness* of a (point symmetric) polyhedron as the likelihood of finding a solution to the corresponding Rupert's problem.

### 9.3.1 Probabilistic algorithm for solving Rupert's problem

Proposition 9.4 states that a polyhedron **P** satisfies Rupert's property if and only if there are $x, y \in \mathbb{R}$, $\alpha, \theta_i, \in [0, 2\pi)$ and $\varphi_i \in [0, \pi]$ for $i = 1, 2$ such that

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ.$$

It seems at first that the two parameters $x$ and $y$ are unbounded. For the first upcoming algorithm it is however necessary to bound all parameters. Hence, we prove the following proposition.

**Proposition 9.7.** *Let* **P** *be a polyhedron containing the origin and let $R \in \mathbb{R}$ be the maximal distance of its vertices to the origin. Assume that a solution to the corresponding Rupert's problem*

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ.$$

*is given. Then $|x|, |y| \leq R$.*

*Proof.* As **P** lies inside the ball with radius $R$ centered at the origin, we have

$$M_{\theta_2,\varphi_2}(\mathbf{P}) \subset \{a \in \mathbb{R}^2 : \|a\| \leq R\}.$$

Since the origin is in the interior of **P**, we have

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})((0,0,0)^t) = (T_{x,y} \circ R_\alpha)((0,0)^t) = T_{x,y}((0,0)^t) = (x,y)^t,$$

hence

$$(x,y)^t \in M_{\theta_2,\varphi_2}(\mathbf{P})^\circ \subset \{a \in \mathbb{R}^2 : \|a\| \leq R\}.$$

Therefore $x^2 + y^2 \leq R^2$ and in particular $|x|, |y| \leq R$. □

Now the interval for each of the 7 parameters in (9.3) is bounded and we can create a first version of our probabilistic deciding algorithm.

Algorithm 1
Input: A polyhedron **P**.
Output: The solution encoded by $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ if **P** is Rupert.

(1) Find $R$ like in Proposition 9.7. Draw $x$ and $y$ uniformly in $[-R, R]$, $\theta_1$, $\theta_2$ and $\alpha$ uniformly in $[0, 2\pi)$, and $\varphi_1$, $\varphi_2$ uniformly in $[0, \pi]$.

(2) Construct the two $3\times 2$ matrices $A$ and $B$ corresponding to the linear maps $R_\alpha \circ M_{\theta_1,\varphi_1}$ and $M_{\theta_2,\varphi_2}$. Compute the two projections of $\mathbf{P}$ given by $\mathcal{P}' \coloneqq T_{x,y}(A\cdot\mathbf{P}) = A\cdot\mathbf{P}+(x,y)$ and $\mathcal{Q}' \coloneqq B\cdot\mathbf{P}$.

(3) Find vertices on the convex hulls of $\mathcal{P}'$ and $\mathcal{Q}'$; denote them by $\mathcal{P}$ and $\mathcal{Q}$.

(4) Decide whether $\mathcal{P}$ lies inside of $\mathcal{Q}$ by checking each vertex of $\mathcal{P}$.

(5) If Step (4) yields a True, return the solution $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2)$. Otherwise, repeat Steps (1)-(5).

Here is a pseudocode for this algorithm:

**Algorithm 11** Probabilistic algorithm for deciding whether **P** is Rupert.
**Input:** Polyhedron **P** given by an $N \times 3$ matrix for some $N \in \mathbb{N}$.
**Output:** The solution encoded by $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ if **P** is Rupert.

$R \leftarrow \sqrt{\max_i(\mathbf{P}[i,1]^2 + \mathbf{P}[i,2]^2 + \mathbf{P}[i,3]^2)}$
isRupert $\leftarrow$ False
**while** isRupert $=$ False **do**
  Draw $x$ and $y$ uniformly in $[-R, R]$
  Draw $\theta_1$, $\theta_2$ and $\alpha$ uniformly in $[0, 2\pi)$
  Draw $\varphi_1$ and $\varphi_2$ uniformly in $[0, \pi]$
  $A \leftarrow R_\alpha \circ M_{\theta_1,\varphi_1}$ and $B \leftarrow M_{\theta_2,\varphi_2}$ ▷ $A, B$ are $3 \times 2$ matrices
  $\mathcal{P}' \leftarrow A \cdot \mathbf{P} + (x, y)$ and $\mathcal{Q}' \leftarrow B \cdot \mathbf{P}$
  $\mathcal{P} \leftarrow$ `ConvexHullPoints`$(\mathcal{P}')$ and $\mathcal{Q} \leftarrow$ `ConvexHullPoints`$(\mathcal{Q}')$
  $n \leftarrow$ # rows of $\mathcal{P}$ and $m \leftarrow$ # rows of $\mathcal{Q}$
  **for** $i$ from 1 to $n$ **do**
    $P \leftarrow \mathcal{P}[i,]$ ▷ $P \in \mathbb{R}^2$, $i$th row of $\mathcal{P}$ is the $i$th vertex of $\mathcal{P}$
    **if** $P$ is not inside $\mathcal{Q}$ **then**
      isRupert $\leftarrow$ False
      Break the **For Loop**
    **end if**
    isRupert $\leftarrow$ True
  **end for**
**end while**
**return** $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2)$

Already this very simple algorithm is able to find solutions for many polyhedra. However, it is quite slow, mostly because the 7-dimensional search space for $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2)$ is large. The first and most significant improvement to Algorithm 1 is to reduce the parameter search space from $\mathbb{R}^7$ to $\mathbb{R}^4$ by algorithmically finding $x, y$ and $\alpha$ for given $\theta_1, \theta_2, \varphi_1, \varphi_2$. Chazelle [95] found an efficient algorithm for deciding polygon containment under translation and rotation, which we may conveniently apply. Let us call Chazelle's algorithm Chazelle; it takes as input two polygons $\mathcal{P}$ and $\mathcal{Q}$ and outputs $(x, y, \alpha)$ such that $(T_{x,y} \circ R_\alpha)(\mathcal{P}) \subset \mathcal{Q}$, and `False` if no such triple exists.

Exploiting Proposition 9.6, namely that if **P** is point symmetric then one can choose $x = y = 0$, one can significantly simplify the algorithm in the point symmetric case. Namely, one needs to solve the polygon containment problem only under rotation (and not additionally translation) which is a much easier task: we will call this algorithm ChazelleR: its input are two polygons $\mathcal{P}$ and $\mathcal{Q}$ and the output is $(0, 0, \alpha)$ such that $R_\alpha(\mathcal{P}) \subset \mathcal{Q}$, and `False` if no such $\alpha$ exists.

We also note that choosing $\theta_1$, $\theta_2$ uniformly in $[0, 2\pi)$, and $\varphi_1, \varphi_2$ uniformly in $[0, \pi]$ is slightly unnatural, since this does not give a uniform distribution on the sphere. As explained in §9.2, we will rather draw $\theta_i \sim U(0, 2\pi)$ and $\varphi \sim \arccos(U(-1, 1))$. We obtain the following improvement to our Algorithm 1:

Algorithm 2 (Using Chazelle)
Input: A polyhedron **P**.
Output: The solution encoded by $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ if **P** is Rupert.

(1) For each $i \in \{1, 2\}$ draw $\theta_i$ uniformly in $[0, 2\pi)$, and $\widetilde{\varphi}_i$ uniformly in $[-1, 1]$. Set $\varphi_i \coloneqq \arccos(\widetilde{\varphi}_i)$.

(2) Construct the two $3 \times 2$ matrices $A$ and $B$ corresponding to the linear maps $M_{\theta_1,\varphi_1}$ and $M_{\theta_2,\varphi_2}$. Compute the two projections of **P** given by $\mathcal{P}' \coloneqq A \cdot \mathbf{P}$ and $\mathcal{Q}' \coloneqq B \cdot \mathbf{P}$.

(3) Find vertices on the convex hulls of $\mathcal{P}'$ and $\mathcal{Q}'$; denote them by $\mathcal{P}$ and $\mathcal{Q}$.

(4) Call Chazelle$(\mathcal{P}, \mathcal{Q})$ (or ChazelleR$(\mathcal{P}, \mathcal{Q})$ if **P** is point symmetric).

(5) If Step (4) yields a solution $(x, y, \alpha)$, return $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2)$. Otherwise, repeat Steps (1)-(5).

The algorithm above can find solutions to Rupert's problem for many solids in fractions of seconds and is able to solve one of the previously unsolved Archimedean polyhedra (see Theorem 9.13). However, we can improve it even further. Analyzing its practical performance, it is clear that the most time consuming part is Step (4). Heuristically, this is expected because the theoretical complexity of Chazelle is $O(pq^2)$, if $p$ is the number of vertices of $\mathcal{P}$ and $q$ the number of vertices of $\mathcal{Q}$ [95], while all other steps in Algorithm 2 are at most linear in $N$, the number of vertices of **P**. Therefore, a natural practical improvement to this algorithm would be to discard pairs $(\mathcal{P}, \mathcal{Q})$ already before Step (4) if it can be algorithmically easily seen that $\mathcal{P}$ cannot fit inside $\mathcal{Q}$. Indeed, we can do so by first computing elementary geometric invariants of the polygons. Moreover, these invariants can be computed for a large batch of polygons coming from randomly drawn projections; then we can discard most pairs and need to test only the remaining ones.

Define *area* and *perimeter* of a polygon in the obvious way and we call the longest line segment inside $\mathcal{P}$ *the diameter of* $\mathcal{P}$. Denote the three by $\mathrm{Area}(\mathcal{P})$, $\mathrm{Peri}(\mathcal{P})$ and $\mathrm{Dia}(\mathcal{P})$ respectively. The following easy lemma allows to speed up our search.

**Lemma 9.8.** *Assume that a convex polygon $\mathcal{P}$ fits in a polygon $\mathcal{Q}$ then:*

1. *The area of $\mathcal{P}$ is smaller than the area of $\mathcal{Q}$:* $\mathrm{Area}(\mathcal{P}) < \mathrm{Area}(\mathcal{Q})$.
2. *The diameter of $\mathcal{P}$ is smaller than the diameter of $\mathcal{Q}$:* $\mathrm{Dia}(\mathcal{P}) < \mathrm{Dia}(\mathcal{Q})$.

3. *The perimeter of $\mathcal{P}$ is smaller than the perimeter of $\mathcal{Q}$:* $\mathrm{Peri}(\mathcal{P}) < \mathrm{Peri}(\mathcal{Q})$.

*Proof.* Since area, perimeter and diameter are invariant under translation and rotation, we may assume that $\mathcal{P}$ not only fits inside $\mathcal{Q}$ but already lies inside $\mathcal{Q}$. Then the statements 1 and 2 become evident. Figure 9.3 proves part 3 of the lemma. Note that convexity is important only for this part. □

Obviously the perimeter of a polygon can be computed in linear time depending on the number of vertices. The Shoelace formula allows for the same complexity for the area. The method of *rotating calipers* allows to compute the diameter of a (convex) polygon in linear time as well [292, Cor. 3.1]. We obtain the following efficient algorithm.

Algorithm 3 (Using Chazelle and Lemma 9.8)
Input: A polyhedron **P**, a batch size $M \in \mathbb{N}$.
Output: The solution encoded by $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ if **P** is Rupert.

(1) For each $j \in \{1, \ldots, M\}$ draw $\theta_j$ uniformly in $[0, 2\pi)$, and $\widetilde{\varphi}_j$ uniformly in $[-1, 1]$. Set $\varphi_j \coloneqq \arccos(\widetilde{\varphi}_j)$.

(2) For each $j \in \{1, \ldots, M\}$ construct the $3 \times 2$ matrix $A_j$ corresponding to the linear map $M_{\theta_j, \varphi_j}$. Compute the projection of **P** given by $\mathcal{P}'_j \coloneqq A_j \cdot \mathbf{P}$. Find the vertices on the convex hull of $\mathcal{P}'_j$ and denote them by $\mathcal{P}_j$. Compute and store: $\mathrm{Area}(\mathcal{P}_j)$, $\mathrm{Peri}(\mathcal{P}_j)$ and $\mathrm{Dia}(\mathcal{P}_j)$.

(3) For each $j \in \{1, \ldots, M\}$ and $k \in \{1, \ldots, M\}$ such that $k \neq j$: if $\mathrm{Area}(\mathcal{P}_j) < \mathrm{Area}(\mathcal{P}_k)$ and $\mathrm{Peri}(\mathcal{P}_j) < \mathrm{Peri}(\mathcal{P}_k)$ and $\mathrm{Dia}(\mathcal{P}_j) < \mathrm{Dia}(\mathcal{P}_k)$ then call Chazelle$(\mathcal{P}_j, \mathcal{P}_k)$ (or ChazelleR$(\mathcal{P}_j, \mathcal{P}_k)$ if **P** is point symmetric).

(4) If for some pair $(j, k)$ Step (3) yields a solution $(x, y, \alpha)$, return $(x, y, \alpha, \theta_j, \theta_j, \varphi_k, \varphi_k)$. Otherwise, repeat steps (1)-(4).

We ran our implementations on all 5 Platonic, 13 Archimedean, 13 Catalan and 92 Johnson polyhedra. The results are presented in Section 9.4.

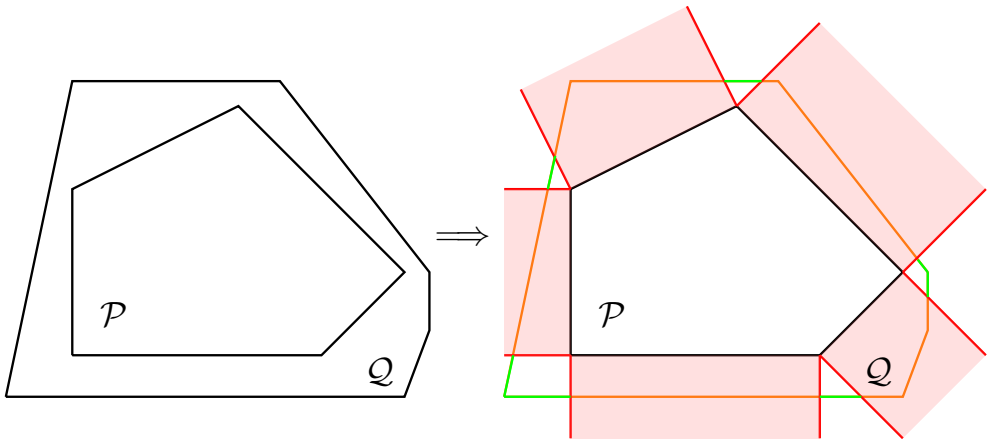



Figure 9.3: Proof that if convex $\mathcal{P}$ lies inside $\mathcal{Q}$ then $\mathrm{Peri}(\mathcal{P}) < \mathrm{Peri}(\mathcal{Q})$.

### 9.3.2 Finding and improving Nieuwland's numbers

In this short section, we briefly explain an algorithmic method which yields lower bounds on Nieuwland numbers of polyhedra and a simple procedure for finding "good" solutions to Rupert's problem. Recall from Definition 9.5 that if there exist $x, y \in \mathbb{R}$, $\alpha, \theta_1, \theta_2 \in [0, 2\pi)$, $\varphi_1, \varphi_2 \in [0, \pi]$ and $\mu \geq 1$ such that

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mu \mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ \tag{9.5}$$

then the Nieuwland number of $\mathbf{P}$ is greater than $\mu$, that is $\nu(\mathbf{P}) > \mu$. We will say that the Nieuwland number of *a solution* $(x, y, \alpha, \theta_1, \varphi_1, \theta_2, \varphi_2)$ to Rupert's problem of some polyhedron is the largest real number $\mu$ such that (9.5) holds. Naturally we will say that a solution is better than another if it has a larger Nieuwland number. Clearly, the Nieuwland number of a solution to Rupert's problem for some polyhedron $\mathbf{P}$ gives a lower bound on $\nu(\mathbf{P})$.

In the previous section we introduced algorithms for finding solutions to Rupert's problem, i.e. finding solutions to (9.5) with $\mu > 1$. Given such a solution $(x, y, \alpha, \theta_1, \varphi_1, \theta_2, \varphi_2)$, it is easy to efficiently find (numerically) an approximation with any given precision for its Nieuwland number using a binary search method: Given $\mathbf{P}$, $v = (x, y, \alpha, \theta_1, \varphi_1, \theta_2, \varphi_2)$ and some $\mu$ it is easy to check whether (9.5) holds, therefore one can search for the correct $\mu$ by constantly halving the interval which it contains. We will denote this procedure $\mu(v, \mathbf{P})$.

Since we are also interested in "optimal" solutions to Rupert's problem, i.e. solutions with maximal Nieuwland number, we will briefly explain a procedure to improve a found solution. The idea is simple: starting with a solution $v = (x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$, we first compute $\mu(v, \mathbf{P})$ and then perturb all parameters by small random numbers $r_1, \ldots, r_7$: $v'_i = v_i + r_i$. If by chance we find a better solution, i.e. if $\mu(v', \mathbf{P}) > \mu(v, \mathbf{P})$, we continue with $v'$, otherwise we choose another random vector $(r_1, \ldots, r_7)$. In practice, the numbers $r_i$ are drawn uniformly from some small intervals around 0, which are steadily narrowed down if no improvement was observed for a long time. Also, in order to avoid convergence to "local minima", one should run this procedure on several different initial solutions. Similarly to the considerations before, if $\mathbf{P}$ is point symmetric, we can choose $x = y = 0$ and then draw only $(r_1, \ldots, r_5) \in \mathbb{R}^5$.

We note that this method is indeed rather naive and probably may be improved easily. For example, in the very recent work [140] Fredriksson uses non-linear optimization methods like SLSQP and Nelder-Mead to find numerically optimal solutions for Rupert's problem. Still, in practice we observed that our approach performs quite well. For example, after less than one minute of computational time on a regular computer, we found a solution to Rupert's problem for the Cube and improved it to have Nieuwland's number of 1.06058. As mentioned in the introduction, it is known that the optimal solution for the Cube has Nieuwland's number $3\sqrt{2}/4 \approx 1.06066$.

We will present our results on improved lower bounds for Nieuwland numbers for various solids in §9.4.2.

### 9.3.3 Deterministic algorithm

In this section we will design a deterministic algorithm for deciding whether a given polyhedron satisfies Rupert's property. The main idea is to transform the problem into systems

of polynomial inequalities and consequently into the decidability problem of emptiness of semi-algebraic sets.

The first step towards this algorithm is to develop an algebraic formulation for expressing the containment of a point $B \in \mathbb{R}^2$ in the convex hull formed by some points $A_1, \dots, A_n \in \mathbb{R}^2$.

**Lemma 9.9.** *Let $A_1, \dots, A_n \in \mathbb{R}^2$ be the vertices of a convex polygon ordered in counterclockwise direction and $B \in \mathbb{R}^2$ a point strictly inside this polygon. Set $A_{n+1} \coloneqq A_1$. Then $\det(A_i - B, A_{i+1} - B) > 0$ for $i = 1, \dots, n$.*

*Proof.* As $B$ lies inside the described polygon, the oriented angles $\sphericalangle A_i B A_{i+1}$ lie in the open interval $(0, \pi)$. This implies

$$\det(A_i - B, A_{i+1} - B) = \underbrace{\|A_i - B\|}_{>0} \cdot \underbrace{\|A_{i+1} - B\|}_{>0} \cdot \underbrace{\sin(\sphericalangle A_i B A_{i+1})}_{>0} > 0. \qquad \square$$

**Lemma 9.10.** *Let $A_1, \dots, A_n, B \in \mathbb{R}^2$, set $A_{n+1} \coloneqq A_1$ and assume that for $i = 1, \dots, n$ it holds that $\det(A_i - B, A_{i+1} - B) > 0$. Then $B$ lies strictly inside the convex hull spanned by $A_1, \dots, A_n$.*

*Proof.* Assume that $B$ is not inside the interior of the convex hull. By the continuity of the determinant, if $B$ lies exactly on the border, there exists a $B' \in \mathbb{R}^2$ outside the convex hull, still satisfying all (strict) inequalities. So we can assume that $B$ lies outside the convex hull. Then there exists a $v \in \mathbb{R}^2$ such that all $A_i$ strictly lie on the same side of the line $\{B + vt \colon t \in \mathbb{R}\} \subset \mathbb{R}^2$. Hence, there is a $w \in \mathbb{R}^2$ perpendicular to $v$ such that every $A_i$ can be written as $A_i = B + t_i v + s_i w$, with $t_i \in \mathbb{R}$ and $s_i \in \mathbb{R}^+$. Let $U = (v, w) \in \mathbb{R}^{2\times 2}$; it follows that

$$(A_i - B, A_{i+1} - B) = U \cdot \begin{pmatrix} t_i & t_{i+1} \\ s_i & s_{i+1} \end{pmatrix}.$$

Taking the determinant, we find $0 < \det(A_i - B, A_{i+1} - B) = \det(U)(t_i s_{i+1} - s_i t_{i+1})$. Dividing by $s_i s_{i+1} > 0$ yields $0 < \det(U)\,(t_i/s_i - t_{i+1}/s_{i+1})$. Finally, summing over all these inequalities gives the desired contradiction. $\square$

Let $\mathbf{P} = \{P_1, \dots, P_n\}$ be a convex polyhedron with $n$ enumerated vertices and further let a parallel projection $\mathcal{P} = (T_{x,y} \circ R_\alpha \circ M_{\theta,\varphi})(\mathbf{P})$ of the polyhedron be given. Only a subset of the projected $P_i$ lie on the boundary of $\overline{\mathcal{P}}$. Let those be $P_{s_1}, \dots, P_{s_k}$ ordered in counterclockwise direction as they appear along the boundary. We call the cycle $s = (s_1, \dots, s_k)$ the *silhouette* of the polyhedron under the projection.

Note that two projections $(T_{x,y} \circ R_\alpha \circ M_{\theta,\varphi})(\mathbf{P})$ and $(M_{\theta,\varphi})(\mathbf{P})$ always have the same silhouette, as translations and rotations do not influence which points of a polygon are on its boundary.

Furthermore, we define $S_n$ to be the set of all non-empty cycles of any (non-empty) subset of the numbers from 1 to $n$. For instance, we have $S_3 = \{(1), (2), (3), (1,2), (2,3), (1,3), (1,2,3), (1,3,2)\}$.

Clearly, the silhouette of a polyhedron with $n$ enumerated vertices under any projection is an element of $S_n$. The following argument bounds $|S_n|$ from above: Denote by $k$ the

length of a cycle and recall that there are $(k-1)!$ cycles of $k$ elements. Hence we have

$$|S_n| = \sum_{k=1}^{n} \binom{n}{k}(k-1)! = \sum_{k=1}^{n} \frac{n!}{k(n-k)!} < n! \sum_{k=1}^{n} \frac{1}{(n-k)!} < e \cdot n!, \tag{9.6}$$

where $e \approx 2.72$ is Euler's number.

**Theorem 9.11.** *Let* $\mathbf{P}$ *be a convex polyhedron with* $n$ *vertices having integer coordinates, whose absolute value is bounded by* $m$*. There exists a deterministic algorithm with running time* $(\log(m) \cdot n)^{O(1)} \cdot n!$ *deciding whether* $\mathbf{P}$ *is Rupert and finding a solution if it exists.*

*Proof.* We start by enumerating the vertices of the polyhedron $\mathbf{P} = \{P_1, \ldots, P_n\}$. The algorithm we will present can decide whether there exists a solution to Rupert's problem

$$(T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ$$

for any possible silhouette $s \in S_n$ of the projection on the right-hand side. Then the full algorithm will run over all elements of $S_n$.

Let $x, y, \alpha, \theta_1, \varphi_1, \theta_2, \varphi_2$ be variables. Given a silhouette $s = (s_1, s_2, \ldots, s_k)$, let $Q_i \coloneqq M_{\theta_2,\varphi_2}(\mathbf{P}_{s_i})$ and $P_j \coloneqq (T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}_j)$ for $i = 1, \ldots, k$ and $j = 1, \ldots, n$. We also set $Q_{k+1} \coloneqq Q_1$. In other words, $Q_1, \ldots, Q_k$ denote the vertices on the boundary of $M_{\theta_2,\varphi_2}(\mathbf{P})^\circ$ given a solution with silhouette $s$. Recall that by definition the vertices $Q_{s_1}, \ldots, Q_{s_k}$ are in ordered in counter-clockwise direction. We define the system of $kn$ inequalities in the seven unknowns $x, y, \alpha, \theta_1, \varphi_1, \theta_2, \varphi_2$:

$$\det(Q_{s_i} - P_j, Q_{s_{i+1}} - P_j) > 0 \qquad j = 1, \ldots, n \text{ and } i = 1, \ldots, k. \tag{9.7}$$

Now there are two important observations:

1. If this system has a solution $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$, then by Lemma 9.10 all $P_j$ lie in the interior of the convex hull of the $Q_{s_i}$. Therefore this septuple gives a solution to Rupert's problem for $\mathbf{P}$ (not necessarily for the silhouette $s$).

2. If the system (9.7) does not have a solution, then there does not exist a solution to Rupert's problem with the silhouette $s$. In other words, if Rupert's problem for $\mathbf{P}$ has a solution $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2)$ for some $s$, then (9.7) must hold at this point. Since in the definition of silhouette, the vertices are required to be ordered in counter-clockwise direction, we can apply Lemma 9.9 and the observation follows.

Therefore solving the system (9.7) is of crucial importance. Denote by $Z_{j,i}$ the matrices $(Q_{s_i} - P_j, Q_{s_{i+1}} - P_j)$, i.e. write (9.7) as $\det(Z_{j,i}) > 0$.

Now we would like to employ algorithms for deciding existence of solutions to systems of polynomial inequalities, but the system (9.7) involves trigonometric functions. However, it is also easy to see that (9.7) is a polynomial system in the "variables" $x, y, \sin(\alpha)$, $\cos(\alpha), \sin(\theta_i), \cos(\theta_i), \sin(\varphi_i), \cos(\varphi_i)$, $i = 1, 2$. Henceforth, we shall apply the following rational parametrization of the circle:

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R}^2 \\ t &\mapsto \left( \frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2} \right). \end{aligned}$$

It is well-known that not only $||f(t)|| = 1$ for all $t$, but also that $f$ is a bijection between $\mathbb{R}$ and $S^2 \setminus \{(-1, 0)\}$. We will substitute the variables $\alpha, \theta_i, \varphi_i$, $i = 1, 2$ with the variables $a, b_1, b_2, c_1, c_2 \in \mathbb{R}$ by

$$\begin{aligned}\big(\cos(\alpha), \sin(\alpha)\big) &=: f(a),\\ \big(\cos(\theta_i), \sin(\theta_i)\big) &=: f(b_i),\\ \big(\cos(\varphi_i), \sin(\varphi_i)\big) &=: f(c_i),\end{aligned}$$

for $i = 1, 2$. Now all entries of $Z_{j,i}$ are rational functions and so are also the inequalities $\det(Z_{j,i}) > 0$. Next, for each $i$ and $j$ we define the matrix $\widetilde{Z_{j,i}}$ as the matrix $Z_{j,i}$ multiplied by $(1 + a^2)(1 + b_1^2)(1 + b_2^2)(1 + c_1^2)(1 + c_2^2) > 0$. Each entry of the matrix $\widetilde{Z_{j,i}}$ is a polynomial in $x, y, a, b_1, b_2, c_1, c_2$.

Note that the determinants of $Z_{j,i}$ and $\widetilde{Z_{j,i}}$ have the same sign by the linearity of the determinant. Therefore, the system (9.7) is equivalent to the system $\det(\widetilde{Z_{j,i}}) > 0$. Expanding $\det(\widetilde{Z_{j,i}})$ shows that its coefficients are bounded by $O(m^2)$ and the polynomials have a total degree of at most 22.

Therefore, we are left with a system $\det(\widetilde{Z_{j,i}}) > 0$ consisting of $kn$ polynomial inequalities in 7 variables, each of them having total degree of at most 22 and integer coefficients bounded in absolute value by $O(m^2)$. According to [164], this system can be solved in a complexity that is polynomial in $\log(m^2)(nk \cdot 22)^{7^2}$, i.e. polynomial in $\log(m)(nk)^{7^2}$. Using $k \leq n$ the complexity simplifies to $(\log(m)n)^{O(1)}$.

Finally, in the worst case, we need to solve such a system for every possible cycle in $S_n$ of possible silhouettes, so using the observation (9.6), we get the total upper bound for the running time complexity: $(\log(m) \cdot n)^{O(1)} \cdot n!$. □

The described algorithm can be summarized as follows:

Algorithm 4
Input: A polyhedron $\mathbf{P} = \{P_1, \ldots, P_n\} \subseteq \mathbb{Z}^3$.
Output: The solution encoded by $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ if $\mathbf{P}$ is Rupert.

For every possible silhouette $s = (s_1, \ldots, s_k) \in S_n$:

(1) Define the system of inequalities $\det(Q_{s_i} - P_j, Q_{s_{i+1}} - P_j) > 0$ for $j = 1, \ldots, n$ and $i = 1, \ldots, k$, where $Q_i := M_{\theta_2,\varphi_2}(\mathbf{P}_{s_i})$ and $P_j := (T_{x,y} \circ R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}_j)$ as well as $Q_{k+1} := Q_1$.

(2) Substitute the variables $\alpha, \theta_i, \varphi_i$ with $a, b_i, c_i$, $i = 1, 2$, using the above defined function $f$. This yields a system of rational inequalities.

(3) Multiply each inequality by $\big((1 + a^2)(1 + b_1^2)(1 + b_2^2)(1 + c_1^2)(1 + c_2^2)\big)^2$, to get a system of polynomial inequalities with integer coefficients.

(4) Search for a solution using the algorithm described in [164].

(5) If (4) yielded a solution: Transform the found solution back to the original variables $(x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2) \in \mathbb{R}^7$ using $f^{-1}$. Break the loop and return this septuple as a solution to Rupert's Problem.

Note that Theorem 9.11 above can easily be extended to incorporate polyhedra having rational coordinates, as these can be stretched by the least common multiple of the denominators of **P** in order to have integer coefficients. Moreover, if the coordinates of the polyhedron are not rational but algebraic numbers (like for most Platonic and Archimedean solids) the algorithm above can be adapted as well. The trick is to add to the system of inequalities (9.7) new variables and equations given by minimal polynomials encoding these coordinates.

We remark that the bound $O(n!)$ on the possible number of silhouettes is very pessimistic. For example, up to (isomorphic) permutations, the Cube has essentially only one silhouette, while $8!$ is quite huge. We are confident that by a closer inspection one can show that the number of possible silhouettes actually growths polynomially in $n$ and for regular polyhedra, like the Platonic or Archimedean solids, is quite small. In practice, however, this does not change much, because the complexity to solve already one single system of inequalities corresponding to a silhouette seems to be infeasible (we will address this issue in §9.4 more explicitly). Therefore, any possible way to reduce the number of silhouettes one needs to check still leads to an algorithm that is unlikely to determine the existence or non-existence of a solution for a non-trivial polyhedron. So we conclude that, at least for now, the described deterministic algorithm is only of theoretical value.

### 9.3.4 Rupertness

In this section we will quantify the likelihood of finding a solution to Rupert's problem by a randomly chosen projection. For a given polyhedron **P** we will define the Rupertness $\mathrm{Rup}(\mathbf{P})$ as the probability that two random projections of it yield a solution to Rupert's problem. We already discussed that point symmetry is advantageous in general for proving Rupert's property, as it decreases the search space from $\mathbb{R}^7$ to $\mathbb{R}^5$. Keeping that in mind, we will only focus on comparing point symmetric polyhedra and define Rupertness only in this setting:

**Definition 9.12.** Let **P** be a centrally symmetric polyhedron. The Rupertness of **P**, denoted $\mathrm{Rup}(\mathbf{P})$, is the probability that two uniformly chosen projections $M_{\theta_1,\varphi_1}(\mathbf{P}), M_{\theta_2,\varphi_2}(\mathbf{P})$ can be extended to a solution of Rupert's problem for **P**, i.e. there exists some $\alpha \in [0,\pi)$ such that $(R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ$.

Note that, naturally, in this definition we draw $\theta_i$ and $\varphi_i$ $(i = 1, 2)$ not uniformly on the intervals $[0, 2\pi)$ and $[0, \pi]$ but in a way such that the projections are uniformly distributed on the sphere. As mentioned in §9.2 this can be modeled by choosing $\theta_i \sim U(0, 2\pi)$ uniformly and $\varphi \sim \arccos(U(-1, 1))$.

As observed in Section 9.2, if **P** is Rupert then there must already exist a set of solutions with positive measure. Therefore, a point symmetric polyhedron **P** is Rupert if and only if $\mathrm{Rup}(\mathbf{P}) > 0$. This also proves that if a solution to Rupert's problem of a polyhedron exists, Algorithm 3 will find it eventually.

As we will elaborate in §9.4, our algorithms can solve all Archimedean polyhedra except three: The Rhombicosidodecahedron (RID in short), Snub cube and Snub dodecahedron. Hence, the RID is the only remaining point symmetric Archimedean polyhedron, for which

Rupert's problem is open. Our main application of the notion of Rupertness is to statistically show that the RID is significantly different from the solved Archimedean polyhedra.

Using the algorithms from §9.3.1 and elementary statistics, we can estimate confidence intervals of Rupertness for various solids. For example, if 1000 random pairs of projections of the Cube gave 65 solutions, the probability estimate would be 6.5% and since this can be viewed as a Bernoulli experiment, one can also calculate the $1-\alpha$ confidence interval $6.5\% \pm \epsilon$ for this probability for any $\alpha \in (0,1)$. More precisely, if $n$ random pairs of projections $M_{\theta_1,\varphi_1}(\mathbf{P}), M_{\theta_2,\varphi_2}(\mathbf{P})$ gave $k > 0$ solutions then the Clopper-Pearson formula implies that the $1-\alpha$ confidence interval for the underlying probability is given by $(S_{\min}, S_{\max})$, where

$$S_{\min} = \left(1 + \frac{n-k+1}{k \cdot F(\alpha/2\,;\, 2k, 2(n-k+1))}\right)^{-1}, \tag{9.8}$$
$$S_{\max} = \left(1 + \frac{n-k}{(k+1) \cdot F(1-\alpha/2\,;\, 2(k+1), 2(n-k))}\right)^{-1},$$

and $F(q\,;\, d_1, d_2)$ is the $q$ quantile of the $F$-distribution with $d_1$ and $d_2$ degrees of freedom. In the case $k = 0$, the probability is between $0$ and $1 - \sqrt[n]{\alpha/2}$ with a certainty of $1-\alpha$.

## 9.4 Explicit results

In this section we collect the explicit results of our work [303]. We prove Rupert's property for a tenth Archimedean solid, show that most Catalan and Johnson solids are Rupert, improve on almost all known Nieuwland numbers and estimate the Rupertness of all point symmetric Platonic and Archimedean polyhedra. The solutions described below in Theorem 9.13 and 9.14 as well as in §9.4.2 are found using the probabilistic and numerical algorithms from the previous section in the programming language R and then verified with rigorous bounds in Maple.

### 9.4.1 Rupert solids

We start by resolving a new Archimedean solid:

**Theorem 9.13.** *The Truncated icosidodecahedron has Rupert's property.*

*Proof.* Since this polyhedron is centrally symmetric, we can set $x = y = 0$ by Proposition 9.6. So we just need to find the five parameters $\alpha, \theta_1, \theta_2, \varphi_1, \varphi_2$ as in Proposition 9.4. They can be found quickly by applying Algorithm 3 to the list of coordinates of the vertices of the Truncated icosidodecahedron (see Table 9.4). Here is an improved solution (after application of the method described in §9.3.2):

$$\begin{aligned} \alpha &= 0.43584, \\ \theta_1 = 2.77685, &\qquad \theta_2 = 0.79061, \\ \varphi_1 = 2.09416, &\qquad \varphi_2 = 2.89674. \end{aligned}$$

A rigorous verification in Maple proves that this quintuple indeed corresponds to a solution of Rupert's problem for the Truncated icosidodecahedron. The visualization of this solution

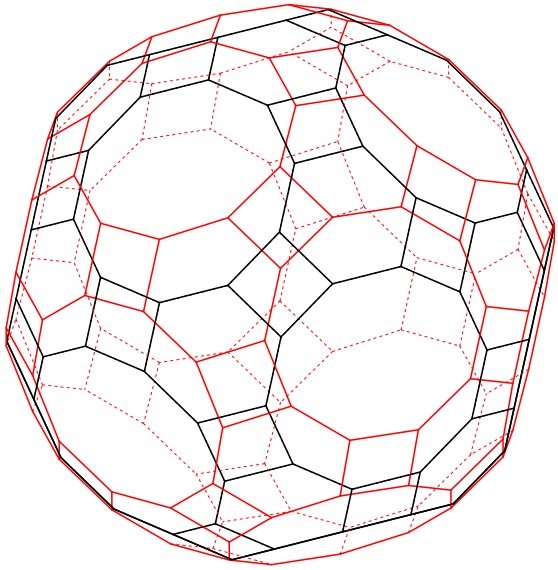


Figure 9.4: The Truncated icosidodecahedron is Rupert.

is presented in Figure 9.4 where the two projections of the polyhedron are plotted such that the black one lies inside the red one. □

We explained in the introduction that of the 13 Archimedean solids 8 were proven to be Rupert in [94] and an additional one in [180, 226]. The theorem above raises this number up to 10, leaving only three Archimedean solids open: Snub cube, Rhombicosidodecahedron and Snub dodecahedron.

The same method and proof as above can be applied to the family of dual solids to the Archimedean ones, called the Catalan solids[3]. We obtain:

**Theorem 9.14.** *The Rhombic dodecahedron, Triakis octahedron, Tetrakis hexahedron, Deltoidal icositetrahedron, Disdyakis dodecahedron, Rhombic triacontahedron, Triakis icosahedron, Pentakis dodecahedron and Disdyakis triacontahedron all have Rupert's property.*

*Proof.* The parameters for the solution of each solid are displayed in Table 9.3. □

Interestingly, this theorem shows that, similarly to Archimedean solids, 9 of the 13 Catalan solids admit Rupert's property. Except for the Triakis tetrahedron (Figure 9.6, left), the remaining unresolved ones are precisely the dual polyhedra of the unsolved Archimedean solids. This raises the question on connectivity of the notions of duality and Rupert's property; we will state it precisely in §9.4.4.

As mentioned in the introduction, Fredriksson [140] could prove that the Triakis tetrahedron and Pentagonal icositetrahedron are Rupert.

In order to test the power of the presented algorithms, we ran our implementation on the family of 92 Johnson solids[4]. We let the algorithm search for a solution for each polyhedron for at most an hour. The result is as follows.

[3]For the coordinates of Catalan solids we refer to www.dmccooey.com/polyhedra/Catalan.html.
[4]Exact coordinates taken from www.dmccooey.com/polyhedra/Johnson.html.

**Theorem 9.15.** *Out of the 92 Johnson solids (at least) 82 admit Rupert's property. The remaining ones are:* $J25$, $J45$, $J47$, $J71$, $J72$, $J73$, $J74$, $J75$, $J76$, $J77$.

Note that $J71$, $J72$, $J73$, $J74$, $J75$, $J76$, $J77$ are all closely connected to the Rhombicosidodecahedron which we conjecture to be not Rupert (Conjecture 9.16).

### 9.4.2 Lower bounds on Nieuwland numbers

Running the algorithm from §9.3.2 for a few hours on the solved Platonic and Archimedean solids, we could significantly improve most of the previously known lower bounds for their Nieuwland numbers. Table 9.1 summarizes these results. Like before, these numbers are found numerically in R and then verified rigorously in Maple.

For the Platonic solids Dodecahedron and Icosahedron we have found solutions with Nieuwland numbers 1.010818 and 1.010805 respectively. These figures are lower bounds for the Nieuwland numbers of these polyhedra. The numerical similarity of these numbers suggests that possibly they agree completely, like it is (conjecturally [191]) the case for the Cube and Octahedron. We address this question again in §9.4.4. Figure 9.5 is a visualization of our solutions to Rupert's problem for the Dodecahedron and Icosahedron. In both cases we plot different projections of the solids in red and black such that the black projection lies inside the red one.

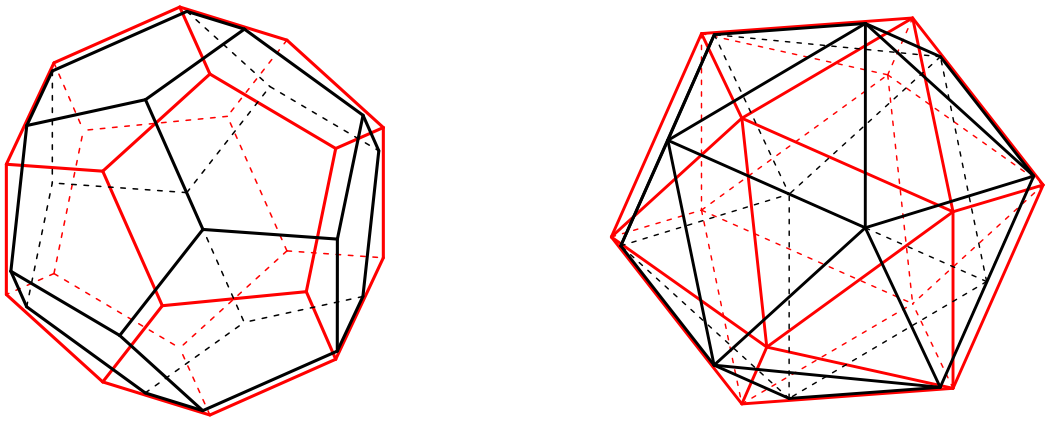


Figure 9.5: Solution of the Rupert's problem for the Dodecahedron (left) with Nieuwland number 1.010818 and for the Icosahedron (right) with Nieuwland number 1.010805.

### 9.4.3 Estimating Rupertness

Recall from Definition 9.12 that the Rupertness of a point symmetric polyhedron is the probability that a pair of uniformly random projections of it can be extended to a solution of Rupert's problem. Like we explained in §9.3.4, we can estimate this probability by randomly drawing projections $M_{\theta_1,\varphi_1}(\mathbf{P}), M_{\theta_2,\varphi_2}(\mathbf{P})$ and then searching for $\alpha \in (0,\pi)$ such that $(R_\alpha \circ M_{\theta_1,\varphi_1})(\mathbf{P}) \subset M_{\theta_2,\varphi_2}(\mathbf{P})^\circ$ holds. For each of the 14 point symmetric Platonic and Archimedean we drew at least 10 million pairs of random projections and for each pair decided on the existence of such an $\alpha \in (0,\pi)$. The quantities of corresponding solutions are summarized in Table 9.3. For example, the first row means that out of our $10^7$ random

| Name of solid | Old best $\mu$ | New best $\mu$ | Improvement |
|---|---|---|---|
| Tetrahedron | 1.004 235 | 1.014 473 | 3.42 |
| Cube | 1.060 660 | 1.060 659 | – |
| Octahedron | 1.060 660 | 1.060 640 | – |
| Dodecahedron | 1.005 882 | 1.010 818 | 1.84 |
| Icosahedron | 1.009 107 | 1.010 805 | 1.19 |
| Truncated tetrahedron | $> 1$ | 1.014 210 | – |
| Cuboctahedron | 1.014 61 | 1.014 571 | – |
| Truncated cube | 1.020 36 | 1.030 659 | 1.51 |
| Truncated octahedron | 1.008 15 | 1.014 602 | 1.79 |
| Rhombicuboctahedron | 1.006 09 | 1.012 819 | 2.10 |
| Truncated cuboctahedron | 1.003 70 | 1.006 563 | 1.77 |
| Snub cube | — | — | |
| Icosidodecahedron | 1.000 15 | 1.000 878 | 5.85 |
| Truncated dodecahedron | 1.000 14 | 1.001 612 | 11.51 |
| Truncated icosahedron | 1.000 04 | 1.001 955 | 48.88 |
| Rhombicosidodecahedron | — | — | – |
| Truncated icosidodecahedron | — | 1.002 048 | – |
| Snub dodecahedron | — | — | – |

Table 9.1 Improved Nieuwland numbers for Platonic and Archimedean solids. The old best lower bounds for the Nieuwland numbers are taken from [191] and [94]. The improvement is calculated using $(\mu_{\text{new}} - 1)/(\mu_{\text{old}} - 1)$.

projections of the Cube precisely 657337 can be extended to a solution of Rupert's problem. This means that the Rupertness of the Cube is approximately 6.57% and the 99.9% confidence interval calculated with the Clopper-Pearson formula (9.8) is $(0.0655, 0.0659)$.

One notices immediately that the Rhombicosidodecahedron (Figure 9.6 right) is not only still unsolved regarding Rupert's property, since out of 100 million tries 0 could have been extended to a solution, but also that its Rupertness is (with confidence of 99.9%) significantly lower than the Rupertness of any other point symmetric Platonic or Archimedean solid. In fact, with probability 99.9%, the Rupertness of the RID is less than 1/10000 of the Rupertness of the Truncated dodecahedron, the one with the smallest figure. Based on Table 9.2 we state the following surprising conjecture which contradicts Conjecture 9.1 taken from [94, Open problem, Conjecture, p. 503].

**Conjecture 9.16.** *The Rhombicosidodecahedron does not have Rupert's property.*

A natural attempt to prove Conjecture 9.16 would be to employ the deterministic algorithm in Theorem 9.11, or rather its extension to polyhedra with coordinates given by algebraic numbers (see §9.3.3). Like we already explained in the remark at the end of §9.3.3, the bound $n!$ is very pessimistic also in this case. To be precise, we are confident that it should not be difficult to prove that (accounting for symmetries) there are not more than 50 possible silhouettes to consider for the Rhombicosidodecahedron. Since the RID has 60 vertices, it follows that we would need to prove emptiness of 50 semi-algebraic sets defined by at most $60^2 = 3600$ polynomial inequalities in $7 - 2 + 1 = 6$ variables (we can

| Name of solid | $n$ | $k$ | $k/n$ (in %) | Confidence interval ($\alpha = 99.9\%$) |
|---|---|---|---|---|
| Cube | $10^7$ | 657337 | 6.57 | $(0.0655, 0.0659)$ |
| Octahedron | $10^7$ | 1195417 | 11.95 | $(0.119, 0.120)$ |
| Dodecahedron | $10^7$ | 230918 | 2.31 | $(0.0230, 0.0232)$ |
| Icosahedron | $10^7$ | 295645 | 2.96 | $(0.0294, 0.0297)$ |
| Cuboctahedron | $10^7$ | 390404 | 3.90 | $(0.0389, 0.0392)$ |
| Truncated cube | $10^7$ | 335602 | 3.36 | $(0.0334, 0.0337)$ |
| Truncated octahedron | $10^7$ | 149188 | 1.49 | $(0.0148, 0.0150)$ |
| Rhombicuboctahedron | $10^7$ | 131176 | 1.31 | $(0.0130, 0.0132)$ |
| Truncated cuboctahedron | $10^7$ | 46044 | 0.460 | $(0.00455, 0.00466)$ |
| Icosidodecahedron | $10^7$ | 40046 | 0.400 | $(0.00395, 0.00406)$ |
| Truncated dodecahedron | $10^7$ | 7583 | 0.0758 | $(0.000736, 0.000781)$ |
| Truncated icosahedron | $10^7$ | 10813 | 0.108 | $(0.00105, 0.00111)$ |
| *Rhombicosidodecahedron* | $10^8$ | 0 | 0 | $[0, 0.000000053)$ |
| Truncated icosidodecahedron | $10^7$ | 16394 | 0.164 | $(0.00161, 0.00167)$ |

Table 9.2 Estimation of the Rupertness of point symmetric Platonic and Archimedean solids. The column $k$ says how many of the $n$ randomly chosen projections can be extended to solutions. $k/n$ is the estimate of the Rupertness and the last column is the $99.9\%$ confidence interval for it.

set $x = y = 0$ but we need a variable for the golden ratio) of total degree of at most 22. Unfortunately, it seems that these numbers are too big for current algorithms and implementations: in order to have a chance for termination in reasonable time, we would need to reduce the number of inequalities to below 20. Therefore, Conjecture 9.16 is still open.

Initially we were quite skeptical that the other unsolved Archimedean solids (Snub cube and Snub Dodecahedron) as well as for the four unsolved Catalan solids (numbers 19, 25, 29, 31 in Table 9.3) and the 10 open Johnson solids (see Theorem 9.15) admit Rupert's property. For these solids we did not estimate the Rupertness and hence have no statistical evidence; so we concluded that it is very much possible that one should just execute the algorithms for a longer time in order to find a solution. Indeed, Fredriksson [140] was able to improve on our methods and show Rupert's property for the Catalan solids 19 and 25 in Table 9.3, as well as for the Johnson solids J25, J45, J47, J71 and J76. We concentrated our search on the RID, since it is the smallest point symmetric solid for which we could not find a solution to Rupert's problem.

### 9.4.4 Concluding remarks and future work

One may notice a surprising fact in Theorem 9.14: A point symmetric Archimedean solid is proven to be Rupert if and only if its dual solid is. While this is only a small indication for the connectivity of duality and Rupert's property, Table 9.1 provides more evidence: the Cube and the Octahedron are conjectured to have the same Nieuwland number and the same seems to hold for the other pair of dual Platonic solids: the Dodecahedron and Icosahedron.

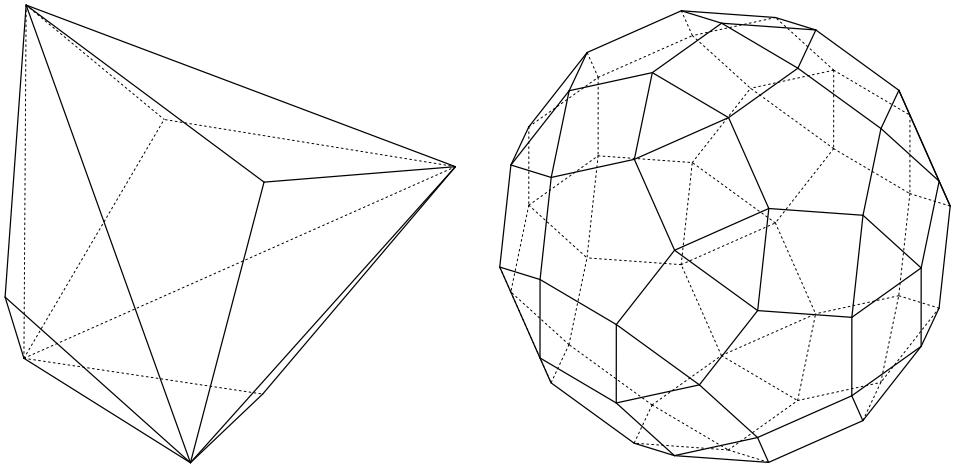


Figure 9.6: Triakis tetrahedron (left), Rhombicosidodecahedron (right)

Based on these observations we formulate natural and interesting but apparently not easy-to-answer questions:

1. Is a point symmetric Archimedean solid Rupert if and only if its dual Catalan solid has Rupert's property?

2. Do dual Platonic solids have the same Nieuwland number? If so, is there a geometric reason for this?

3. What are the exact Nieuwland numbers of the Dodecahedron and Icosahedron? Do they also admit simple algebraic expressions like the (conjectured) $3\sqrt{2}/4$ for the Cube and Octahedron?[5]

If a solution to Rupert's problem of a Platonic or Archimedean solid is given in $\mathbb{R}^3$ by **P** and its copy **Q**, one can look at the duals of both polyhedra. It is quite intriguing that it seems that the dual of an "optimal" solution (i.e. one with highest Nieuwland number) of a Platonic solid yields an "optimal" solution for the dual solid. However, we could not find a (geometric) explanation for this. Moreover, the dual of *some* solution of a Platonic or Archimedean solid is not necessarily a solution at all.

As already mentioned, Conjecture 9.16 contradicts current beliefs on Rupert's property for polyhedra, but at the same time we have statistical reasons to believe in our conjecture. Assuming its validity, further natural questions are:

4. What distinguishes the RID from other point symmetric Archimedean solids and prevents this polyhedron to have Rupert's property? Is there an easy criterion for Rupert polyhedra?

5. How can one prove Conjecture 9.16? Are the remaining Archimedean and Catalan solids (12, 16, 18, 19, 25, 29, 31 in Table 9.3) Rupert?[6]

[5] In [302] we conjecture that the minimal polynomial for these numbers is given by $2025x^8 - 11970x^6 + 17009x^4 - 9000x^2 + 2000$.

[6] The Catalan solids 19 and 25 are now resolved in [140].

## 9.5 Appendix

In the appendix we most importantly present Table 9.3 which summarizes our solutions to Rupert's problem for all Platonic, 10 Archimedean and 9 Catalan solids. According to Proposition 9.4, any solution can be encoded by seven parameters $x, y, \alpha, \theta_1, \theta_2, \varphi_1, \varphi_2$. So for each solved polyhedron we provide these numbers in the corresponding columns. Proposition 9.6 implies that if a polyhedron is point symmetric, one can choose $x = y = 0$, so in these cases $x$ and $y$ are zero. The right column of Table 9.3 shows the Nieuwland number of the solution.

Finally, Table 9.4 incorporates the exact coordinates that we have used for the Platonic and Archimedean solids. The coordinates for Catalan and Johnson solids can be found at www.github.com/Vog0/RupertProblem and are taken, as mentioned, from the website www.dmccooey.com/polyhedra. The first link also contains the source code in R and Maple we used to find and then verify solutions.

Table 9.3 Solutions to Rupert's problem for Platonic, Archimedean and Catalan solids.

| Nr. | Name of solid | $x$ | $y$ | $\alpha$ | $\theta_1$ | $\varphi_1$ | $\theta_2$ | $\varphi_2$ | $\mu(v, \mathbf{P})$ |
|---|---|---|---|---|---|---|---|---|---|
| 1. | Tetrahedron | 0.1788244 | -0.0976062 | 1.0372426 | 5.3278439 | 1.5713832 | 3.9444529 | 0.9501339 | 1.014473 |
| 2. | Cube | 0 | 0 | 2.4840821 | 1.9060829 | 3.1415929 | 5.8188256 | 2.3004443 | 1.060659 |
| 3. | Octahedron | 0 | 0 | 3.1415873 | 5.4977985 | 1.9105975 | 6.2808288 | 1.5701448 | 1.060640 |
| 4. | Dodecahedron | 0 | 0 | 1.0378047 | 0.8553414 | 2.108091 | 4.918788 | 2.0545287 | 1.010818 |
| 5. | Icosahedron | 0 | 0 | 2.7276836 | 2.7732324 | 2.6181502 | 2.3091726 | 2.2712915 | 1.010805 |
| 6. | Truncated tetrahedron | 0.160858 | -0.164724 | 4.7775741 | 6.2831072 | 0.7854425 | 2.0992734 | 1.3849498 | 1.014210 |
| 7. | Cuboctahedron | 0 | 0 | 3.1386793 | 2.5259348 | 1.5710827 | 0.7902177 | 0.9351593 | 1.014571 |
| 8. | Truncated cube | 0 | 0 | 2.298646 | 4.3427928 | 3.1415862 | 2.089632 | 2.2876946 | 1.030659 |
| 9. | Truncated octahedron | 0 | 0 | 1.5690349 | 3.1415601 | 0.785367 | 5.3243536 | 2.0933886 | 1.014602 |
| 10. | Rhombicuboctahedron | 0 | 0 | 0.017061 | 2.9503929 | 3.1415921 | 4.1693802 | 0.636201 | 1.012819 |
| 11. | Truncated cuboctahedron | 0 | 0 | 0.2396229 | 3.1416249 | 0.785486 | 4.4525352 | 0.429099 | 1.006563 |
| 12. | Snub cube | – | – | – | – | – | – | – | – |
| 13. | Icosidodecahedron | 0 | 0 | 1.578603 | 2.7736451 | 0.7120286 | 4.7086522 | 2.1263666 | 1.000878 |
| 14. | Truncated dodecahedron | 0 | 0 | 2.2092757 | 4.3599229 | 1.5508055 | 1.6477247 | 1.0979977 | 1.001612 |
| 15. | Truncated icosahedron | 0 | 0 | 0.9547212 | 4.7124428 | 1.470154 | 0.8649729 | 2.0954566 | 1.001955 |
| 16. | Rhombicosidodecahedron | – | – | – | – | – | – | – | – |
| 17. | Truncated icosidodecahedron | 0 | 0 | 0.4358364 | 2.7768504 | 2.0941596 | 0.79061 | 2.8967442 | 1.002048 |
| 18. | Snub dodecahedron | – | – | – | – | – | – | – | – |
| 19. | Triakis tetrahedron | – | – | – | – | – | – | – | – |
| 20. | Rhombic dodecahedron | 0 | 0 | 0.2389694 | 3.926939 | 0.9553557 | 5.171164 | 1.3442843 | 1.027201 |
| 21. | Triakis octahedron | 0 | 0 | 0.3562255 | 5.7674031 | 2.2867379 | 0.0005374 | 1.5665899 | 1.030648 |
| 22. | Tetrakis hexahedron | 0 | 0 | 0.1945682 | 3.4241341 | 1.1711373 | 0.0040963 | 2.3603178 | 1.009632 |
| 23. | Deltoidal icositetrahedron | 0 | 0 | 0.6277374 | 0.6012867 | 1.4476059 | 6.1255227 | 3.1382821 | 1.007632 |
| 24. | Disdyakis dodecahedron | 0 | 0 | 0.1178211 | 6.1466092 | 2.5957828 | 1.5695218 | 0.7842378 | 1.002500 |
| 25. | Pentagonal icositetrahedron | – | – | – | – | – | – | – | – |
| 26. | Rhombic triacontahedron | 0 | 0 | 0.231712 | 2.84e-05 | 0.5535717 | 1.9227518 | 2.1379305 | 1.007037 |
| 27. | Triakis icosahedron | 0 | 0 | 2.5481489 | 3.3133906 | 0.4995076 | 2.3963212 | 2.1824603 | 1.001304 |
| 28. | Pentakis dodecahedron | 0 | 0 | 3.1547479 | 5.4202246 | 2.1024926 | 4.2553188 | 2.4568193 | 1.001845 |
| 29. | Deltoidal hexecontahedron | – | – | – | – | – | – | – | – |
| 30. | Disdyakis triacontahedron | 0 | 0 | 2.5886126 | 4.2871288 | 0.7860227 | 5.917639 | 2.107937 | 1.000210 |
| 31. | Pentagonal hexecontahedron | – | – | – | – | – | – | – | – |

| Name of solid | Coordinates |
|---|---|
| 1. Tetrahedron | $(\pm1, \pm1, \pm1)$ with an even number of "$-$" signs |
| 2. Cube | $(\pm1, \pm1, \pm1)$ |
| 3. Octahedron | all permutations of $(0, 0, \pm1)$ |
| 4. Dodecahedron | $(\pm1, \pm1, \pm1)$ and all even permutations of $(0, \pm\Phi^{-1}, \pm\Phi)$ |
| 5. Icosahedron | even permutations of $(0, \pm\Phi, \pm1)$ |
| 6. Truncated tetrahedron | all permutations of $(\pm1, \pm1, \pm3)$ with an even number of "$-$" signs |
| 7. Cuboctahedron | all permutations of $(\pm1, \pm1, 0)$ |
| 8. Truncated cube | all permutations of $(\pm1, \pm1, \pm(\sqrt{2}-1))$ |
| 9. Truncated octahedron | all permutations of $(0, \pm1, \pm2)$ |
| 10. Rhombicuboctahedron | all permutations of $(\pm1, \pm1, \pm(1+\sqrt{2}))$ |
| 11. Truncated cuboctahedron | all permutations of $(\pm1, \pm(1+\sqrt{2}), \pm(1+2\sqrt{2}))$ |
| 12. Snub cube | all even permutations of $(\pm1, \pm1/t, \pm t)$ with an even number of plus signs and all odd permutations with an odd number of plus signs. $t$ is the tribonacci constant |
| 13. Icosidodecahedron | all permutations of $(0, 0, \pm\Phi)$ and all even permutations of $(\pm\frac{1}{2}, \pm\frac{\Phi}{2}, \pm\frac{\Phi^2}{2})$ |
| 14. Truncated dodecahedron | all even permutations of $(0, \pm1/\Phi, \pm(2+\Phi)), (\pm\frac{1}{\Phi}, \pm\Phi, \pm2\Phi)$ and $(\pm\Phi, \pm2, \pm(\Phi+1))$ |
| 15. Truncated icosahedron | all odd permutations of $(0, \pm1, \pm3\Phi), (\pm1, \pm(2+\Phi), \pm2\Phi)$ and $(\pm\Phi, \pm2, \pm(2\Phi+1))$ |
| 16. Rhombicosidodecahedron | all even permutations of $(\pm1, \pm1, \pm\Phi^3), (\pm\Phi^2, \pm\Phi, \pm2\Phi)$ and $(\pm(2+\Phi), 0, \pm\Phi^2)$ |
| 17. Truncated icosidodecahedron | all even permutations of $(\pm\frac{1}{\Phi}, \pm\frac{1}{\Phi}, \pm(3+\Phi))$, $(\pm\frac{2}{\Phi}, \pm\Phi, \pm(1+2\Phi)), (\pm\frac{1}{\Phi}, \pm\Phi^2, \pm(-1+3\Phi))$, $(\pm(2\Phi-1), \pm2, \pm(2+\Phi))$ and $(\pm\Phi, \pm3, \pm2\Phi)$ |
| 18. Snub dodecahedron | all even permutations of $(2\alpha, 2, 2\beta)$, $(\alpha+\frac{\beta}{\Phi}+\Phi, -\alpha\Phi+\beta+\frac{1}{\Phi}, \frac{\alpha}{\Phi}+\beta\Phi-1)$, $(\alpha+\frac{\beta}{\Phi}-\Phi, \alpha\Phi-\beta+\frac{1}{\Phi}, \frac{\alpha}{\Phi}+\beta\Phi+1)$, $(-\frac{\alpha}{\Phi}+\beta\Phi+1, -\alpha+\frac{\beta}{\Phi}-\Phi, \alpha\Phi+\beta-\frac{1}{\Phi})$, $(-\frac{\alpha}{\Phi}+\beta\Phi-1, \alpha-\frac{\beta}{\Phi}-\Phi, \alpha\Phi+\beta+\frac{1}{\Phi})$ with an odd number of sign changes of the coordinates, where $\xi = \sqrt[3]{\frac{\Phi}{2}+\frac{1}{2}\sqrt{\Phi-\frac{5}{27}}} + \sqrt[3]{\frac{\Phi}{2}-\frac{1}{2}\sqrt{\Phi-\frac{5}{27}}}$, $\alpha = \xi - 1/\xi$ and $\beta = \xi\Phi + \Phi^2 + \Phi/\xi$ |

Table 9.4 Used coordinates of all Platonic and Archimedean solids, as used in the Maple Package geom3d (for verification), except Snub Cube and Snub Dodecahedron, which are not needed for our results. $\Phi = (\sqrt{5}+1)/2 \approx 1.62$ is the golden ratio.

# Chapter 10

# Open questions related to the thesis

*Problems, problems,*
*problems all day long.*
*Will my problems work out right or wrong?*
The Everly Brothers, *Problems*, 1958

In this final chapter I collect some of the open problems, questions and conjectures related to the topic of the thesis that I found most interesting and intriguing. As mentioned in the introduction, I am convinced that it is precisely open questions like these that drive mathematical research forward. Some of stated problems are well-known, long standing and notoriously difficult conjectures, others have the chance to be resolved in the near future. The reader will notice a big variety in the questions and I hope to catch everyone's taste at least somewhere.

In the beginning problems and conjectures are stated that are related to each chapter of the thesis, then I list questions that I encountered in the previous three years, found very interesting but which are not directly connected to any of the thesis' chapters.

## Questions related to Chapter 2

One motivation of Chapter 2 was to come a step closer to Christol's famous conjecture:

**Conjecture 10.1** (Christol's conjecture)**.** *If $f \in \mathbb{Q}[\![t]\!]$ is D-finite and globally bounded, then $f = \mathrm{Diag}(g)$ for some $n \in \mathbb{N}$ and some rational power series $g \in \mathbb{Q}[\![x_1, \ldots, x_n]\!] \cap \mathbb{Q}(x_1, \ldots, x_n)$.*

We note that the same conjecture also plays an interesting role in Chapter 8. Its first explicit appearance it made in the late 1980s [99] and in the same *exposé* Christol provided the following potential counter-example to his conjecture:

**Question 10.2.** *Does there exist a rational function $R \in \mathbb{Q}[\![x_1, \ldots, x_n]\!]$ such that*

$$\mathrm{Diag}(R) = {}_3F_2\left[\begin{matrix} 1/9 & 4/9 & 5/9 \\ & 1/3 & 1 \end{matrix}; t\right]?$$

Since 35 years this conjecture stands open, even in the particular case of this ${}_3F_2$. Moreover, in 2011 the authors of [52, 53] created a list of 116 similar potential counter-examples

to Christol's conjecture (18 of them are displayed in [52, Appendix F]). Chapter 2 of the present thesis deals with 40 of the 116 (and 4 of the 18) by proving that those are diagonals of explicit algebraic functions. Naturally extending Question 10.2, we formulate:

**Question 10.3.** *In the list of Bostan, Boukraa, Christol, Hassani and Maillard, are the remaining 70 globally bounded ${}_3F_2$'s diagonals?*

Related to Christol's conjecture one might ask for the minimal number of variables to represent a D-finite function as a diagonal of a rational function. Denoting by $\mu(f)$ the minimal $n \in \mathbb{N}$ such that there exists $R \in \mathbb{Q}(x_1, \ldots, x_n)$ with $\mathrm{Diag}(R) = f(t)$ (and setting $n = \infty$ if no such representation exists), we formulate the following conjecture:

**Conjecture 10.4.** *For any $n \geq 1$ there exists a D-finite function $f(t)$ with $\mu(f) = n$.*

Currently it is even open whether $\mu(f) > 3$ for *some* diagonal $f(t)$ [240, Open Problem 3.1]. Note that a combination of results by Pólya [267] and Furstenberg [145] implies that the diagonals of bivariate rational functions are precisely the algebraic functions: $\mu(f) = 2 \Leftrightarrow f(t)$ algebraic. A possible explicit family of D-finite functions satisfying the condition in Conjecture 10.4 is given by the generating functions of the generalized Apéry numbers; more precisely, we conjecture:

**Conjecture 10.5.** *Let $A_{a,b}(t) := \sum_{n\geq 0} \sum_{k\geq 0} \binom{n}{k}^a \binom{n+k}{k}^b t^n \in \mathbb{Z}[\![t]\!]$. Then $\mu(A_{a,b}) = a + b$.*

Most notably, the case $a = b = 2$ represents the "usual" Apéry numbers and it is current work in progress by van Straten and the author to show that $\mu(A_{2,2}) = 4$. The family $A_{a,0}(t)$ for $a \in \mathbb{N}$ was studied by Franel and many others since (see Conjecture 10.16). In general, it is not difficult to show that $\mu(A_{a,b}) \leq a + b$, e.g. if $a \leq b$ then the formula

$$A_{a,b}(t) = \mathrm{Diag}\left(\left(\prod_{i=1}^{a-b}(1-x_i)\cdot\prod_{i=1}^{b}(1-y_i-z_i) - \prod_{i=1}^{a-b}x_i\cdot\prod_{i=1}^{b}y_iz_i\right)^{-1}\right)$$

provides a diagonal representation for $A_{a,b}$ in $a + b$ variables. We also state the identity

$$A_{a,b}(t) = \mathrm{Diag}\frac{1}{((1-y_1)\cdots(1-y_b)-x_1)(1-x_2)\cdots(1-x_a)-x_1\cdots x_a y_1\cdots y_b}$$

which was found by Zudilin (in a private communication) and does not require $a \leq b$.

We recall the notion of Hadamard grade of a power series $f(t)$ as the least number $n$ such that $f(t)$ can be written as the Hadamard product of $n$ algebraic functions. Assuming the Rohrlich-Lang conjecture (Conjecture 10.7 below, see also [322, 225]), Rivoal and Roques proved in [273] that there exist diagonals of any prescribed finite grade. Under the same assumption the authors show that ${}_3F_2([1/7, 2/7, 4/7], [1/2, 1], t)$ has infinite Hadamard grade. It is therefore natural to ask whether one can drop the dependency of this result on the (widely open) Rohrlich-Lang conjecture.

**Conjecture 10.6.** *The function ${}_3F_2\left[\begin{smallmatrix}1/7 & 2/7 & 4/7\\ 1/2 & 1\end{smallmatrix}; t\right]$ has infinite Hadamard grade.*

**Conjecture 10.7** (Rohrlich-Lang). *Let $a_1, \ldots, a_n \in \mathbb{Q}$, $m_1, \ldots, m_n, b \in \mathbb{Z}$ and assume:*

$$\pi^{b/2} \prod_{i=1}^{n} \Gamma(a_i)^{m_i} \in \overline{\mathbb{Q}}.$$

*Then this fact can be proved using only the well-known rules: $\Gamma(a+1) = a\Gamma(a)$, $\Gamma(a)\Gamma(1-a) = \pi/\sin(\pi a)$ and $\prod_{j=0}^{k-1} \Gamma(a + j/k) = (2\pi)^{(k-1)/2} k^{-ka+1/2} \Gamma(ka)$.*

On the same note, we mention the following conjecture about exponential periods and values of the Gamma function (a consequence of Conjecture 10.7):

**Conjecture 10.8.** *The number $\Gamma(1/5)$ is irrational. In fact, $\Gamma(1/5)$ is transcendental and, moreover, the numbers $\Gamma(1/5), \Gamma(2/5), e^{\pi\sqrt{5}}$ are algebraically independent.*

We remark that by a theorem of Chudnovsky the values $\Gamma(1/3), \Gamma(1/4)$ are transcendental and algebraically independent with $\pi$, see [108] and [109, Corollary 2].

Coming back to diagonals and the notion of Hadamard grade, we mention that a proof of Conjecture 10.6 or even Conjecture 10.7 would not imply that the hypergeometric function ${}_3F_2\left[\begin{smallmatrix} 1/7 & 2/7 & 4/7 \\ 1/2 & 1 \end{smallmatrix}; t\right]$ is not a diagonal, since diagonals might (in principle) have infinite Hadamard grade:

**Question 10.9.** *Do there exist diagonals with infinite Hadamard grade?*

# Questions related to Chapter 3

Chapter 3 deals with the question on how to prove that a given D-finite function is algebraic or transcendental. The following three famous conjectures try to classify algebraicity of D-finite functions using arithmetic criteria. The first two relate arithmetic properties of the basis of solutions to a differential equation to their algebraicity. The third conjecture has the interesting feature that it has no explicit condition on the whole basis of solutions.

**Conjecture 10.10** (Grothendieck-Katz). *Let $L \in \mathbb{Q}(x)\langle\partial\rangle$ be a differential operator. If the reduction of $Ly = 0$ mod $p$ has a full basis of rational solutions in $\mathbb{F}_p(x)$ for almost all primes $p$ then $Ly = 0$ has a basis of algebraic solutions.*

This conjecture was formulated first by Grothendieck in the late 1960s. It has been verified for equations of order one (where it is proven [181, 103] to be equivalent to Kronecker's theorem, a consequence of Chebotarev's density theorem in number theory), for hypergeometric equations [38] and more generally for Picard-Fuchs equations [193], and, relatively recently, also the $q$-analog of this conjecture was proven by Di Vizio [123]. We remark that algebraic functions are very special G-functions which are conjecturally period functions:

**Conjecture 10.11** (Bombieri-Dwork). *Every G-function is a solution to a Picard-Fuchs differential equation.*

For the topic of G-functions we refer the reader to the book [17] by André (specifically chapter V in view of the conjecture above).

Another conjecture relating arithmetic properties of solutions to differential equations to their algebraicity is formulated by Bézivin in [41, p. 299]:

**Conjecture 10.12** (Bézivin). *If $Ly = 0$ has a basis of globally bounded solutions then all solutions are algebraic.*

As mentioned in Chapter 3, this conjecture is a consequence of the Grothendieck-Katz conjecture thanks to [18, Prop 5.3.3]. In the preceding "Remarque 5.3.2" of the same reference André mentiones the following conjecture and attributes it to Christol:

**Conjecture 10.13** (André-Christol). *Assume that $f(x)$ is D-finite, globally bounded and that $L_f^{\min}$ is regular at 0. Then $f(x)$ is algebraic.*

We also mention the following "folklore" conjecture on the existence of invariants to a differential Galois group and the implication of its finiteness:

**Conjecture 10.14.** *Let $L$ be an irreducible differential operator of order $n$ and assume that for some $m$ such that $\binom{n+m-1}{n-1} > n^2$ the equation $L^{\circledS m}y = 0$ has a rational non-zero solution. Then all solutions of $Ly = 0$ are algebraic.*

As explained in Section 3.2, an invariant of degree $m$ implies a linear system for the differential Lie algebra with $n^2$ variables and $\binom{n+m-1}{n-1}$ equations, so in the situation of Conjecture 10.14 it is overdetermined and one could expect that this implies the vanishing of the Lie algebra and consequently the finiteness of the differential Galois group.

A related question, more on the computational side, is the following:

**Question 10.15.** *Given a differential operator $L$ of order $r$ and degree $d$ how can one bound the order and degree of $L^{\circledS m}$?*

As remarked in Chapter 3 the order of $L^{\circledS m}$ is at most $N := \binom{n+m-1}{n-1}$. For the degree there also exist some bounds in the literature: Kauers showed in [194, Cor. 9] that $\deg_x L^{\circledS m}$ is bounded by $mdN^2$. However, we observe on generic examples that this bounded is far from being optimal.

Quite related but on the level of sequences and very explicitly, Franel's conjecture predicts the order of minimal recurrences for a family of P-recursive sequences:

**Conjecture 10.16** (Franel). *For $s \in \mathbb{N}_{\geq 1}$ let $A_n^{(s)} = \sum_{k=0}^{n} \binom{n}{k}^s$. The minimal linear recurrence relation with polynomial coefficients for $(A_n^{(s)})_{n\geq 0}$ has order exactly $\lfloor (s+1)/2 \rfloor$.*

Note that the numbers $A_n^{(s)}$ are precisely the coefficients of $A_{s,0}(t)$ in Conjecture 10.5. It is clear that $A_n^{(1)} = 2^n$ and $A_n^{(2)} = \binom{2n}{n}$ are hypergeometric (satisfy linear recurrences of order 1). The Franel numbers $A_n^{(3)}$ as well as the numbers $A_n^{(4)}$ satisfy a recurrence of order 2 and are not hypergeometric (as can be easily verified with Petkovšek's algorithm [262]). Most notably, Stoll proved in [304] that $\lfloor (s+1)/2 \rfloor$ is an upper bound for the minimal recurrence for $A_n^{(s)}$ and Straub and Zudilin proved in [310] that any *telescoping recurrence*

for $A_n^{(s)}$ has order at least $\lfloor (s+1)/2 \rfloor$.

Regarding proving algebraicity for explicit examples, we recall that the following conjecture, which is central in Chapter 3, is still open and currently work in progress.

**Conjecture 10.17.** *The functions $F_d(t), F_e(t), F_f(t)$ and $F_i(t)$ introduced in Chapter 3 are algebraic of algebraicity degree $155520$.*

Of course, as explained in that chapter, the most plausible way to prove this algebraicity would be to find invariants of the corresponding differential operators. The author and his collaborators have guessed rational functions which should be solutions of the 20-th symmetric powers of the operators, however proving these guesses turns out to be very challenging in practice.

Finally, we recall that the following statement is also conjectured in Chapter 3:

**Conjecture 10.18.** *The Table 3.1 presents the complete answer to Zagier's question in Chapter 3, i.e. if $u_n(\alpha)_n(\beta)_n$ is globally bounded for $\alpha, \beta \in \mathbb{Q}_{>0}$ then $(\alpha, \beta) \bmod \mathbb{Z}^2$ is in Table 3.1.*

# Questions related to Chapter 4

One of the main contributions of Chapter 4 is a proof that the function

$$\mathsf{Iso}^2(x) = \frac{9\sqrt{2}}{8\pi} \cdot \frac{{}_2F_1\left[\begin{smallmatrix} -\frac{3}{2} & -\frac{3}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^2}{{}_2F_1\left[\begin{smallmatrix} -\frac{1}{2} & -\frac{1}{2} \\ & 1 \end{smallmatrix}; \frac{4x^2}{(1-x^2)^2}\right]^3} \cdot \left(\frac{1-x^2}{1+x^2}\right)^3$$

is increasing on $[0, \sqrt{2}-1)$. One general approach to prove monotonicity is to show positivity of the derivative. If, like Iso, the function of interest is not D-finite but given as the quotient of two D-finite functions, one can equivalently try to prove positivity of the numerator of the derivative. In both cases, this yields the following question:

**Question 10.19.** *Given a D-finite function $f(x)$ defined on some interval in $I \subseteq \mathbb{R}$, can one decide that $f(x)$ is positive on $I$?*

The approach in [334, 241] is slightly different: Melczer and Mezzarobba prove the stronger statement that all Taylor coefficients of some function $D(x)$ are positive. This also implies positivity of $D(x)$ and gives rise to the following related question.

**Question 10.20.** *For a P-recursive sequence $(u_n)_{n\geq 0}$ can one decide $u_n \geq 0$ for all $n \geq n_0$?*

Solving Question 10.20 in general is a very ambitious task. Some progress was made by Kauers and Pillwein [200] for second-order recurrences or special recurrences of order $3$ in both cases with generic initial conditions, see also [232] and [261]. We also mention a work in progress by Alaa Ibrahim and Bruno Salvy where the authors show that the positivity problem is decidable for P-recursive sequences of arbitrary order with generic initial conditions and under the assumption that the limit as $n \to \infty$ of the companion

matrix of the recursion is an invertible matrix over $\mathbb{Q}$ with one single dominant eigenvalue. This work was presented at JNCF23.

Recall that the analogous problem for the very special case of C-finite sequences is also still open (for recursions of order bigger than 9) [253]. There are even explicit C-finite sequences for which positivity is open as the following conjecture (from [203]) shows:

**Conjecture 10.21.** *The Taylor coefficients $(f_n)_{n\geq 0}$ of*

$$\begin{aligned} f(x) &= \frac{24x^7 - 162x^6 + 330x^5 + 124x^4 - 232x^3 - 114x^2 - 44x - 10}{x^8 - 12x^7 + 34x^6 - 60x^5 - 5x^4 + 15x^3 + 4x^2 + 3x + 1} + \frac{10x+2}{x^2-3x+1} \\ &+ \frac{11x^7 - 135x^6 + 165x^5 - 97x^4 - 30x^3 - 46x^2 - 7x - 3}{x^8 - 11x^7 + 89x^6 - 71x^5 + 65x^4 - 16x^3 + 14x^2 - x + 1} \\ &= -11 - 8x + 240x^3 + 704x^4 - 20x^5 + 192x^6 + 5508x^7 + \cdots \end{aligned}$$

*are positive for $n \geq 6$.*

On the same note we also recall the famous Skolem problem (see, for example, [252] for some references and progress). A sequence $(u_n)_{n\geq 0}$ is said to *have a zero* if there exists a $k \in \mathbb{N}$ such that $u_k = 0$; we also define $\mathcal{Z} := \{k \colon u_k = 0\}$.

**Conjecture 10.22** (Skolem problem)**.** *The problem whether or not a given C-finite sequence has a zero is decidable.*

It was inferred by Tao that it is "*faintly outrageous that this problem is still open*". The well-known Skolem-Mahler-Lech *theorem*, on the other hand, ensures that the set of zeros of a C-finite sequence is the union of a finite set and finitely many arithmetic progressions. It is an open question whether the same conclusion can be made about the zeros of a sequence whose generating function is algebraic (see [339, §4.4]):

**Conjecture 10.23.** *Let $\sum_{n\geq 0} u_n t^n \in \mathbb{Q}[\![t]\!]$ be an algebraic power series. The set $\mathcal{Z}$ is the union of a finite set and finitely many arithmetic progressions.*

We also mention a related conjecture that was formulated by Bostan in 2015:

**Conjecture 10.24** (Bostan)**.** *Let $f(t) \in \mathbb{Q}[\![t]\!]$ be an algebraic power series with minimal polynomial $P(t,T) \in \mathbb{Q}[t,T]$ of degree $d_t$ in $t$ and degree $d_T$ in $T$. If $f$ admits $d_t(d_T-1)$ consecutive zero coefficients then $f$ is a polynomial.*

Also the following related conjecture by Furter from [146, Conj. 1.6 $R(m)$] is still open:

**Conjecture 10.25.** *Let $f(t) \in \mathbb{Q}[t]$ be a polynomial of degree $m$ with $f(0) = 0$ and denote by $f^{[-1]} \in \mathbb{Q}[\![t]\!]$ its compositional inverse. If $m$ consecutive coefficients of $f^{[-1]}$ vanish then $f = t$.*

Coming back to the initial motivation for Chapter 4, we state the following open problem regarding Canham's problem.

**Conjecture 10.26.** *The solution to Canham's problem for genus one and $v_0 \in (0, 3/(2^{5/4}\pi^{1/2})]$ is unique and a surface of revolution.*

Recall that for $v_0 \in [3/(2^{5/4}\pi^{1/2}), 1) \approx [0.712, 1)$ the unique solution is a conformal transformation of a projection of the Clifford torus, however if $v_0$ is less than $0.712$ only existence is guaranteed and only experimental results hint for the shape [334, 202].

# Questions related to Chapter 5 and Chapter 6

In Chapter 6 we showed an algorithm that computes the $N$-th term in a polynomial C-finite sequence (i.e. a C-finite sequence over the ring $R = \mathbb{K}[x]$) in $O(N)$ arithmetic operations in $\mathbb{K}$. Following the analogy between arithmetic complexity in $\mathbb{K}[x]$ and bit complexity in $\mathbb{Z}$ or $\mathbb{Q}$ it is natural to wonder:

**Question 10.27.** *Is it possible to compute the $N$-th term of a C-finite sequence over $\mathbb{Q}$ in bit complexity $O(N)$? Specifically, it is possible to compute $3^N$ in base $2$ or the $N$-th Fibonacci number in base 2 in $O(N)$ bit operations?*

Clearly, binary splitting provides an algorithm to compute the $N$-th term of a C-finite sequence in bit complexity $O(\mathbf{M}(N))$, where $\mathbf{M}(N)$ denotes the complexity of multiplication of two numbers with bit-size $N$. The algorithm in Chapter 6 exploits the fact that if $(u_n)_{n\geq 0}$ is a polynomial C-finite sequence then the coefficient sequence of $u_N(x)$ satisfies a linear recurrence relation with polynomial coefficients, i.e. is P-recursive. Such a behaviour is not observed in $\mathbb{Z}$: the digits of $3^N$ in base $2$, for example, seem to have no recurrence (or any other kind or regularity) one can try to exploit. Still, of course, one can hope to find a method to answer Question 10.27 affirmatively.

Another direction opened by the work in Chapter 6 to the following conjecture:

**Conjecture 10.28.** *Let $(u_n(x))_{n\geq 0}$ be a P-recursive polynomial sequence. Prove that one can compute $u_N(x)$ in $O(N)$ arithmetic operations.*

To prove this statement it is enough to show that the minimal differential equation of

$$f_n(x) := \frac{1}{2\pi i}\oint_\gamma \frac{f(x,y)}{y^{n+1}}\mathrm{d}y,$$

where $f(x,y)$ is D-finite, has order and degree independent of $n$. For the case $f(x,y)$ is a rational function this statement is the content of Lemma 6.5 and a very similar proof should be possible if $f(x,y)$ is hyperexponential.

The main message of Chapter 5 is that the $N$-th term of a $q$-holonomic sequence can be computed faster than naively and that this fact has many applications. One of these applications is that the sparse polynomial $p_n(x) = \sum_{\ell=0}^{N} x^{\ell^2}$ can be evaluated in complexity $\tilde{O}(\sqrt{N})$. Naturally, the following explicit question arises:

**Question 10.29.** *Is it possible to evaluate fast polynomials of the form $\sum_{\ell=0}^{N} x^{\ell^s}$, for $s \geq 3$?*

Another problem posed in the same chapter is the following question on complexity lower bounds:

**Question 10.30.** *Can one prove non-trivial lower bounds of computational complexity for the $N$-th term of a $q$-holonomic sequence or a P-recursive sequence?*

The idea for computing the $N$-th term of a $q$-holonomic sequence is based on ideas by Strassen [307] and the Chudnovsky brothers [106], namely to use a baby-step/giant-step technique together with polynomial multipoint evaluation/interpolation. For P-recursive sequences this yields a method to compute $u_N$ in arithmetic complexity $\tilde{O}(\sqrt{N})$ and the same complexity we achieve for $q$-holonomic sequences (for $q \in \mathbb{K}$). Naturally, one may ask whether it is possible to compute the $N$-th term of a P-recursive sequence even faster:

**Question 10.31.** *Is it possible to compute the $N$-th term of a P-recursive sequence $(u_n)_{n\geq 0}$ faster than in $\tilde{O}(\sqrt{N})$ arithmetic operations? For example, is it possible to compute $N!$ in fewer than $\tilde{O}(\sqrt{N})$ operations?*

As already noted in [307, 106], an affirmative answer to this question would give a deterministic algorithm for factoring $M \in \mathbb{Z}$ faster than in $\tilde{O}(\sqrt[4]{M})$ bit operations. We remark that the exponent $1/4$ for the factoring problem was the best possible until it was improved relatively recently to $9/2$ by Hittmeir [179] and then to $1/5$ by Harvey [168] who found an algorithm that can factor $M \in \mathbb{Z}$ in bit-complexity $\tilde{O}(M^{1/5})$.

Of course, Question 10.31 may also be asked for $q$-holonomic sequences:

**Question 10.32.** *For $q \in \mathbb{K}$ is it possible to compute the $N$-th term of a $q$-holonomic sequence $(u_n(q))_{n\geq 0}$ faster than in $\tilde{O}(\sqrt{N})$ arithmetic operations? For example, is it possible to compute $[N]_q!$ in fewer than $\tilde{O}(\sqrt{N})$ operations?*

While solving these problems (Question 10.31 and Question 10.32) is very ambitious, we mention that in many cases additional structure for the sequence is known. For example, if $\sum_{n\geq 0} u_n t^n$ is a G-function or even algebraic, the best known current method is to use the same baby-step/giant-step technique mentioned earlier, without exploiting the added structure. The following question is therefore also natural to ask:

**Question 10.33.** *Assume that $\sum_{n\geq 0} u_n t^n$ is algebraic. Is it possible to compute $u_N$ faster than with $\tilde{O}(\sqrt{N})$ operations? Specifically, is it possible to compute $\binom{2N}{N}$ with fewer than $\tilde{O}(\sqrt{N})$ arithmetic operations?*

We remark that this question is solved for finite fields $\mathbb{K}$ in [61, 55]. There it is shown that if $\mathbb{K} = \mathbb{F}_{p^k}$ and $\sum_{n\geq 0} u_n t^n \in \overline{\mathbb{K}(t)}$ then $u_N$ can be computed in arithmetic complexity which is linear in $\log N$ and quasi-linear in $\sqrt{p}$.

So far we mostly focused on arithmetic complexity, however many very interesting questions also concern the bit-complexity model. As explained in Chapter 5, given some $N \in \mathbb{Z}$, one can compute $N!$ faster than naively using *binary splitting*; this results in the quasi-optimal complexity $O(\mathbf{M}_{\mathbb{Z}}(N \log N) \log N)$. Borwein found an algorithm [49] that computes $N!$ in $O(\mathbf{M}_{\mathbb{Z}}(N \log N) \log\log N)$, i.e. he could replace the $\log N$ by $\log\log N$. At the end of his short paper Borwein asks:

**Question 10.34.** *Can one compute $N!$ faster than in $O(\mathbf{M}_{\mathbb{Z}}(N \log N) \log\log N)$ binary steps?*

Recall that the same binary splitting algorithm that works for $N!$ adapts to matrix factorials and consequently allows to compute $N$-th terms in P-recursive sequences in $\mathbb{Q}$ quasi-optimally. Borwein's trick does not translate to matrices and even to hypergeometric sequences, since it relies on the fact that $N! = \prod p^{\alpha_p}$, where the product runs over all primes $p$ and $\alpha_p = \lfloor N/p \rfloor + \lfloor N/p^2 \rfloor + \cdots$. Therefore the following question remains open:

**Question 10.35.** *It is possible to compute the $N$-th term of a P-recursive sequence $(u_n)_{n\geq 0}$ over $\mathbb{Q}$ in $O(\mathbf{M}_{\mathbb{Z}}(N \log N) \log\log N)$ bit-complexity?*

Similarly to the arithmetic complexity model and Questions 10.31 and 10.32, often additional arithmetic information on the sequence is available.

**Question 10.36.** *Assume that $\sum_{n\geq 0} u_n t^n \in \mathbb{Q}[\![t]\!]$ is a G-function or even algebraic. Is it possible to compute $u_N$ faster than in $O(\mathbf{M}_{\mathbb{Z}}(N \log N) \log N)$ operations?*

Especially for $u_n = \binom{2n}{n} \in \mathbb{Z}$ it is unsatisfying that no better way is known than to use either binary splitting or Borwein's algorithm; both methods compute the huge numerator $(2n)!$ and denominator $n!^2$ and return their quotient which has bit-size only $O(N)$ instead $O(N \log N)$. We remark that the Chudnovsky brothers claim in [106, p. 459] that Question 10.36 admits an affirmative answer, however, to our knowledge, this claim stayed without proof until today.

Coming back to the arithmetic model and the contributions of Chapters 5 and 6, we recall that the complexity of both algorithms for computing $N$-th terms of $q$-holonomic and polynomial C-finite sequences can be analyzed also with respect to the other appearing parameters. The latter problem is equivalent to computing the $N$-th power of a polynomial matrix and for this task we conjecture:

**Conjecture 10.37.** *Given $M(x) \in \mathbb{K}(x)^{r\times r}$ of degree $d$ and $N \in \mathbb{N}$, it is possible to compute $M(x)^N$ in complexity $O(Ndr\mathsf{M}(r))$.*

For the $N$-th term of a $q$-holonomic sequence, we showed in Theorem 5.4 that $u_N(q)$ can be computed in $O(r^\theta\sqrt{Nd} + r^2\mathsf{M}(\sqrt{Nd}))$ operations in $\mathbb{K}$ if $(u_n(q)_{n\geq 0})$ is given by a recursion of order $r$ and degree $d$. Modulo the Question 10.31 and improvement on the feasible exponent of matrix multiplication $\theta$ this seems optimal. Recall that the best current bound for $\theta$ is $\theta < 2.3729$ [15] and its lower bound is the central open question in algebraic complexity theory:

**Question 10.38.** *Can two $n \times n$ matrices be multiplied in $n^{2+o(1)}$ operations, i.e. is $\omega = 2$?*

Studying the same question not for asymptotically large $n$ but for the small cases $n = 2, 3, 4, \ldots$ opens another world of research and highly interesting questions. Mentioning just a tiny bit of the current state of the art, we recall that Strassen's famous algorithm [305] allows to multiply two $2 \times 2$ matrices in 7 multiplications instead of 8 (and since it works over non-commutative rings this implies that $\omega < \log_2(7) \approx 2.81$). It was shown indecently by Hopcroft and Kerr [182] and Winograd [329] that 7 is optimal for $2 \times 2$ matrices regardless whether non-commutativity is required or not. For multiplying $3 \times 3$ matrices the question on the minimal number of multiplications in $\mathbb{K}$ remains completely open with upper bound 22 [235] (commutative case) or 23 [220] (non-commutative case) and lower bound 19 [44].

**Conjecture 10.39.** *It is possible to multiply two $3\times 3$ matrices over a non-commutative ring $R$ using less than 23 multiplications in $R$.*

Note that even if this conjecture were proven, it would not suffice not beat Strassen's algorithm asymptotically, since $\log_2(7) \approx 2.807 < 2.814 \approx \log_3(22)$. On the other hand, if two $3 \times 3$ matrices can be even multiplied in 21 multiplications in a non-commutative ring $R$, this would yield an asymptotic improvement, since $\log_3(21) \approx 2.771$.

# Questions related to Chapter 7

As an intermediate result, Chapter 7 provides a quick proof that the function

$$F(x,q) = \sum_{n\geq 0} \begin{bmatrix} 2n \\ n \end{bmatrix}_q x^n \in \mathbb{Q}[q][\![x]\!] \tag{10.1}$$

is transcendental over $\mathbb{Q}(q,x)$. Recall that $\begin{bmatrix} x \\ y \end{bmatrix}_q = \frac{[n]_q!}{[k]_q![n-k]_q!}$ denotes the $q$-binomial coefficient and $[n]! = [n]_q! = \prod_{i=1}^{n} \frac{1-q^i}{1-q}$ denotes the $q$-factorial. The following question was raised by Kontsevich during a talk by Bostan at the "Équations différentielles motiviques et au-delà" seminar in April 2023.

**Question 10.40** (Kontsevich)**.** *Does there exist a suitable notion of $q$-algebraicity for which $q$-analogues of algebraic hypergeometric functions become $q$-algebraic. Specifically, for the $q$-Chebyshev numbers*

$$C_q(n) = \frac{[30n]![n]!}{[15n]![10n]![6n]!} \tag{10.2}$$

*what is the appropriate notion of $q$-algebraicity such that $\sum_{n\geq 0} C_q(n)t^n$ is $q$-algebraic?*

Of course, a first step would be to answer the same question for the generating function of the central $q$-binomial coefficients (10.1). Not too far away topic-wise is a conjecture by Warnaar and Zudilin stated in the beautiful and short article [325], see also [326].

**Conjecture 10.41.** *Assume that the numbers $a_1,\ldots,a_r,b_1,\ldots,b_s \in \mathbb{Z}_{>0}$ satisfy the inequality $\sum_{i=1}^{r}\lfloor a_i x\rfloor - \sum_{j=1}^{s}\lfloor b_j x\rfloor \geq 0$ for all $x \geq 0$. Then for any $n \geq 0$ all coefficients of the polynomial*

$$D_n(q) = \frac{[a_1 n]!\cdots[a_r n]!}{[b_1 n]!\cdots[b_s n]!} \in \mathbb{Q}[q]$$

*are non-negative. In particular, all coefficients of $C_q(n)$ in (10.2) are non-negative.*

By Landau's criterion [222], the inequality condition in this conjecture makes the sequence of numbers $D_n$ defined by

$$D_n = \frac{(a_1 n)!\cdots(a_r n)!}{(b_1 n)!\cdots(b_s n)!}$$

an integer sequence. For more general hypergeometric sequences Christol's condition [99] is a criterion for integrality and the interlacing criterion [38] by Beukers and Heckman characterizes hypergeometric sequences with algebraic generating function. In the view of Conjecture 10.41 one might naturally wonder whether these conditions imply some positivity for $q$-analogues expressed in terms of $q$-Pochhammer symbols $(a;q)_n := \prod_{k=0}^{n-1}(1-aq^k)$. Here, however, more care is needed since, for example, the power series in $q$ given by

$$\frac{(1/3;q)_n(2/3;q)_n}{(1/2;q)_n[n]!} \in \mathbb{Q}[\![x]\!]$$

has also negative coefficients, even though ${}_2F_1([1/3,2/3],[1/2],t)$ is algebraic.

A possible way of attacking (at least special cases of) Conjecture 10.41 is by analytic means, in a similar way like the very recent proof of Borwein's conjecture by Wang [323], see also [324]. In the latter the following, still open, conjecture appears and is labeled as "cubic version of Borwein's conjecture":

**Conjecture 10.42.** *For $n > 0$ let $P_n(q) \in \mathbb{Z}[q]$ be defined as*

$$P_n(q) = (1-q)(1-q^2)(1-q^4)(1-q^5)\cdots(1-q^{3n-2})(1-q^{3n-1}).$$

*Then the sign pattern of the coefficient expansion of $P_n^3(q)$ is $+--+--+--\cdots$, where zeros are considered as both $+$ and $-$.*

In [324] Wang and Krattenthaler provide uniform proofs for the first and second Borwein conjectures (which predict sign patterns of $P_n(q)$ and $P_n^2(q)$) based on analytic methods and also prove "two thirds" of Conjecture 10.42, in the sense that they verify that $[q^{3k}]P_n^3(q) \geq 0$ for all $k$ and $[q^{3k+1}]P_n^3(q) \leq 0$ for $k \leq \deg P_n/2$ and consequently $[q^{3k+2}]P_n^3(q) \leq 0$ for $k \geq \deg P_n/2$.

Finally, the following question by Capobianco is only peripherally related to Chapter 7; it appears in the same collection of open problems [91] as Aissen's question on the $q$-analogue of Pólya's theorem and is the item right after it ("30. The Star of David Identity").

**Question 10.43** (Capobianco)**.** *The following is known as* The Star of David Identity*:*

$$\binom{n}{r}\binom{n+1}{r+2}\binom{n+2}{r+1} = \binom{n}{r+1}\binom{n+1}{r}\binom{n+2}{r+2} \quad r,n \in \mathbb{N}.$$

*An algebraic proof is trivial. Does there exist a combinatorial proof?*

# Questions related to Chapter 8

The following question was raised by Zagier [338, p. 769, Question 2] and Gorodetsky [162] and was the main motivation for work that led to Chapter 8.

**Question 10.44.** *Which P-recursive sequences are constant terms?*

The definition of "constant term sequence" is varying from source to source; for this work we defined it as a sequence $(A_n)_{n\geq 0}$ for which there exist Laurent polynomials $P, Q \in \mathbb{Q}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ such that $A_n = \mathrm{ct}[P^n Q]$. Zagier and Gorodetsky assumed that $Q = 1$ which, for example, does not capture the Catalan numbers. Since the sum of two constant term sequences is not necessarily a constant term anymore, we proposed to also look at finite sums of constant terms; this induces the following, more general, question:

**Question 10.45.** *Which P-recursive sequences are linear combinations of constant terms?*

For the case when the sequence satisfies a linear recurrence with constant coefficients (i.e. is C-finite) the answer to this question is given in Chapter 8, however, in general, it remains widely open. After C-finite sequences the next natural class to examine would be hypergeometric sequences or coefficients of algebraic functions:

**Question 10.46.** *Can one classify hypergeometric constant terms or constant term sequences whose generating function is algebraic? Specifically, are the numbers*

$$A(n) = \binom{8n}{4n}\binom{4n}{n}\binom{2n}{n}^{-1}$$

*a constant term sequence?*

We also mention that one may draw natural connections to "multiple binomial sums", i.e. sequences which are coefficients of diagonals of rational functions [72] and consequently to Christol's conjecture (Conjecture 10.1).

# Questions related to Chapter 9

The main (still open) question addressed in Chapter 9 is the following:

**Question 10.47.** *Does any polyhedron have Rupert's property?*

We recall that Rupert's property has been verified for all five Platonic solids, 10 (out of 13) Archimedean solids, 10 (out of 13) Catalan solids and 82 (out of 92) Johnson solids. Our methods find all these solutions in seconds (or sometimes minutes) but fail to prove that the Rhombiicosidodecahedron (RID) is Rupert which should be the easiest of the remaining open solids, since it is point symmetric. After doing some statistical analysis of the *Rupertness* in Section 9.3.4 we conjecture:

**Conjecture 10.48.** *The Rhombiicosidodecahedron is not Rupert.*

Moreover, there is an apparent connection between Rupert's property and the duality of polyhedra, most obvious expressed via Nieuwland numbers. We would like to understand this connection better:

**Question 10.49.** *What is the connection between Nieuwland numbers of dual solids?*

# Other related open questions

In this final section we collect interesting open problems which are related to the topic of the thesis however do not match precisely any of the chapters' contents. We start with an ambitious but very interesting question whether it is possible to algorithmically decide integrality of P-recursive sequences:

**Question 10.50.** *Given a P-recursive sequence $(u_n)_{n\geq 0}$ (by a defining recursion and initial conditions) is it possible to decide whether $u_n \in \mathbb{Z}$ for all $n \geq 0$?*

This problem is essentially solved for hypergeometric sequences by Christol [99] however it remains completely open for higher order recurrences. One such instance are the following two conjectures by Zagier formulated in the delightful paper [337]:

**Conjecture 10.51** (Zagier). *For $A, B, \lambda \in \mathbb{Q}$ consider the recurrence*

$$(n+1)^2 u_{n+1} - An(n+1)u_n + Bn^2 u_{n-1} = \lambda u_n \quad n \geq 0. \tag{10.3}$$

*If $u_0 = 1$ and $u_n \in \mathbb{Z}$ then $A, B, \lambda \in \mathbb{Z}$.*

**Conjecture 10.52** (Zagier). *Up to normalizations, excluding trivial cases, the only parameters $A, B, \lambda$ that yield integral solutions of (10.3) are two families $(A, B, \lambda) = (-1, 0, d^2 + d)$ and $(A, B, \lambda) = (2, 1, d^2 + d + 1)$ (for $d \in \mathbb{Q}, d \geq -1/2$) and seven sporadic cases.*

The following related very short and explicit conjecture was asked by López-Aguayo in [233] (generalizing Chiriţă's proposed problem which supposed $r = 1$).

**Conjecture 10.53.** *For $n, r \in \mathbb{N}_{>0}$ let $S_r(n) := \sum_{k=1}^{n} \frac{k}{k+r}\binom{n}{k}$. Then $S_r(n) \notin \mathbb{Z}$.*

Using elementary techniques, López-Aguayo proved that $S_r(n) \notin \mathbb{Z}$ for $r = 1, 2, 3, 4$ and in [221] this list was extended up to 22. We also mention [234] where Luca and Pomerance prove that the set of $n \in \mathbb{N}$ for which $S_r(n) \in \mathbb{Z}$ for some $r$ has density 0 in the natural numbers.

Closely related to the problem of integrality of P-recursive sequences is the problem of "inverse creative telescoping" (see [97, §8]). For instance, recall that Zeilberger's algorithm is able to deduce the recurrence

$$(n+1)^3 A_{n+1} - (2n+1)(17n^2 + 17n + 5)A_n + n^3 A_{n-1} = 0$$

from the "explicit formula" for the Apéry numbers

$$A_n = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2. \tag{10.4}$$

Roughly speaking, the inverse problem asks for an algorithm which, on this example, would find the representation (10.4) from the recurrence and corresponding initial conditions. More precisely:

**Question 10.54.** *Given a P-recursive sequence $(u_n)_{n\geq 0}$ decide whether there exists a bivariate hypergeometric term $a_{n,k}$ such that*

$$u_n = \sum_{k=0}^{n} a_{n,k} \quad \textit{for all } n \geq 0.$$

In this formulation, a bivariate hypergeometric term is a function $a(n, k) = a_{n,k}$ such that $a(n+1, k)/a(n, k) \in \mathbb{Q}(n, k)$ and $a(n, k+1)/a(n, k) \in \mathbb{Q}(n, k)$.

For algebraic series the relationship between global boundedness and integrality is made via the so-called Eisenstein constant:

**Question 10.55.** *By Eisenstein's theorem, an algebraic power series $f(x) \in \mathbb{Q}[\![x]\!]$ is globally bounded. Given the minimal polynomial for $f$, how can one bound the* Eisenstein constant *which is the least non-zero natural number $\alpha$ such that $f(\alpha x) - f(0) \in \mathbb{Z}[\![x]\!]$.*

There are many approaches to this problem and the best proven bound is by Dwork and van der Poorten [127] (based on ideas and conjectures of Schmidt [282]); it says:

$$\alpha \leq 4.8(8e^{1.22n-3}n^{4+2.74\log n})^n H^{2n-1},$$

where $n$ is the degree in $y$ of the minimal polynomial $P(x,y) \in \mathbb{Z}$ and $H$ is the maximum of the absolute values of its coefficients. The algorithmic computation of $\alpha$ is discussed in [239].

Still on the side of algebraic series but more on the arithmetic side is the following deceptively easy conjecture by Bostan.

**Conjecture 10.56** (Bostan)**.** *Let $(C_n)_{n\geq 0}$ be the sequence of Catalan numbers. Then:*

*(i) The last decimal digit of $C_n$ is never 3;*

*(ii) For $n > 255$, the last digit of an odd $C_n$ is always 5.*

At first glance, this conjecture looks very surprising because on the one hand the generating function of the Catalan numbers is algebraic and on the other hand the behaviour of $p$-automatic sequences (in particular reduction of algebraic functions' coefficients $\bmod p$) is well understood [98, 309]. The catch in Conjecture 10.56 is that the statement is about $C_n \bmod 10$ and 10 a composite number. Surprisingly, this breaks the known theory completely. It is not difficult to see that $C_n$ is odd if and only if $n = 2^k - 1$ for some $k \in \mathbb{N}$. Then by Lucas' theorem it follows that Conjecture 10.56 is equivalent to saying that in the base 5 expansion of $2^k - 1$ with $k > 8$ always appears at least one 3 or 4. Heuristically, this is expected, but proving such a statement might be very difficult.

We note that this conjecture does not predict an isolated fact and the Catalan numbers are not at all special regarding the theme of this statement. For example, the coefficient sequence of $A_{1,2}(t)$ in Conjecture 10.5 given by

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

seems to satisfy that for each $n \geq 0$ the last decimal digit of $u_n$ is 1,5 or 9. To our knowledge this is also an open question.

Another intriguing conjecture concerning the Catalan numbers is mentioned in Stanley's book [301, A28] and goes back to Garrabrant and Pak [157]. For its statement, define a power series $H(x_1,\ldots,x_n) \in \mathbb{N}[\![x_1,\ldots,x_n]\!]$ to be $\mathbb{N}$-rational if it can be obtained from $0, x_1, \ldots, x_n$ by the operations $F+G, F\cdot G$ and $1/(1-F)$. For example, $R = 1/(1-x-y)$ is clearly $\mathbb{N}$-rational, and note that $\mathrm{Diag}(R) = \sum_{n\geq 0}\binom{2n}{n}t^n$.

**Conjecture 10.57.** *The Catalan numbers are not the diagonal of a $\mathbb{N}$-rational power series.*

We mention that the irrationality of Catalan's constant is still open; note that except the name it has nothing in common with the numbers from the previous two conjectures.

**Conjecture 10.58.** *The Catalan constant $G := \sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^2}$ is irrational.*

It was stated in [28] that $G$ is "*arguably the most basic constant whose irrationality and transcendence [...] remain unproven*". The following representation of $G$ immediately follows from a geometric series expansion:

$$G = \int_0^1 \int_0^1 \frac{\mathrm{d}x\mathrm{d}y}{1+x^2y^2}.$$

Given the similarity to $\zeta(2) = \pi^2/6$ for which irrationality is known since centuries, and also relatively recently proven for $\zeta(3)$ [25, 319], it is surprising that for $G$ it is an open problem, especially given Beukers' beautiful and short proof in [37] and Rivoal's equally short but conceptual approach in [272].

The following surprising and very interesting open problem appears in [255] and is attributed to Pak and Yeliussizov.

**Conjecture 10.59** (Pak-Yeliussizov)**.** *Assume that $(u_n)_{n\geq 0}$ is a sequence of rational numbers such that both the ordinary and exponential generating functions (i.e. $\sum_{n\geq 0} a_n t^n$ and $\sum_{n\geq 0} a_n t^n/n!$) are D-algebraic (i.e. satisfy algebraic differential equations, not necessarily linear), then both generating functions are D-finite.*

We refer to the marvellous article [80] for some recent progress on "D-transcendence".

Finally, the following conjecture was stated by Kauers and Koutschan in [197]:

**Conjecture 10.60** (Kauers, Koutschan)**.** *Let $(a_n)_{n\geq 0}$ be the sequence given by*

$$\sum_{n\geq 0} a_n t^n = \mathrm{Diag} \frac{1}{1 - x_1x_2 - \cdots - x_5x_6 - x_1^2 - \cdots - x_6^2}.$$

*The numbers $\tilde{a}_n := a_n n!^3/(3n!)$ satisfy the recurrence given in [197, Conjecture 7].*

The conjecture's authors found the guessed recurrence using a method they call "guessing with little data" [196]. We show below additional insight that allows to compute hundreds of terms and guess the recurrence with "traditional methods", e.g. Maple's `gfun`:

$$x_1x_2 + \cdots + x_5x_6 + x_1^2 + \cdots + x_6^2 = (x_1 + \cdots + x_6)^2/2 + (x_1^2 + \cdots + x_6^2)/2,$$

hence with the geometric expansion and binomial theorem:

$$\begin{aligned} 8^n a_n &= [x_1^n \cdots x_6^n](1 - (x_1 + \cdots + x_6)^2 - x_1^2 - \cdots - x_6^2)^{-1} \\ &= [x_1^n \cdots x_6^n] \sum_{\substack{k\geq 0 \\ j_1,\ldots,j_6\geq 0 \\ i_1,\ldots,i_6\geq 0}} \binom{k+j_1+\cdots+j_6}{k, j_1, \ldots, j_6} \binom{2k}{i_1, \ldots, i_6} x_1^{2j_1+i_1} \cdots x_6^{2j_6+i_6} \\ &= \sum_{j_1,\ldots,j_6\geq 0} \binom{3n+2j_1+\cdots+2j_6}{3n+j_1+\cdots+j_n, j_1, \ldots, j_6} \binom{6n+2j_1+\cdots+2j_6}{n-2j_1, \ldots, n-2j_6}. \end{aligned}$$

Moreover, after a few easy manipulations one sees that the last sum can be expressed as the coefficient of $x^{3n}$ in

$$(3n)!\left(\sum_{\ell\geq 0}\frac{(2\ell)!}{\ell!}x^{\ell}\right)\left(\sum_{k\geq 0}\frac{x^k}{k!(n-2k)!}\right)^6. \tag{10.5}$$

This observation provides an easy and quick way to compute many of sequence elements. For example, putting (10.5) into Maple we find the first 100 terms of $(a_n)_{n\geq 0}$ in 10 seconds on a regular PC and then we guess the recurrence from [197, Conjecture 7] in a few seconds using `gfun`. In principle, creative telescoping allows to deduce the recurrence for $(a_n)_{n\geq 0}$ from (10.5), however using this expression may be still too optimistic since the computation did not terminate after 24 hours after which we aborted it.

*"Der einzige Ausweg wär aus diesem Ungemach:*
*Sie selber dächten auf der Stelle nach*
*Auf welche Weis dem guten Menschen man*
*Zu einem guten Ende helfen kann.*
*Verehrtes Publikum, los, such dir selbst den Schluß!*
*Es muß ein guter da sein, muß, muß, muß!"*[1]
Bertolt Brecht, *Der gute Mensch von Sezuan*, Epilog.

[1] Translation (E. Bentley):
*"How could a better ending be arranged?*
*Could one change people? Can the world be changed?*
*It is for you to find a way, my friends,*
*To help good men arrive at happy ends.*
*You write the happy ending to the play!*
*There must, there must, there's got to be a way!"*

# Bibliography


[1] Y. Abdelaziz, C. Koutschan, and J.-M. Maillard. On Christol's conjecture. *J. Phys. A*, 53(20):205201, 16 pages, 2020.

[2] S. Abramov, M. Bronstein, and M. Petkovšek. On polynomial solutions of linear operator equations. In *ISSAC'95*, pages 290–296. ACM Press, 1995.

[3] S. A. Abramov. Rational solutions of linear difference and $q$-difference equations with polynomial coefficients. *Programmirovanie*, (6):3–11, 1995.

[4] S. A. Abramov. A direct algorithm to compute rational solutions of first order linear $q$-difference systems. *Discrete Math.*, 246(1-3):3–12, 2002.

[5] S. A. Abramov, P. Paule, and M. Petkovšek. $q$-hypergeometric solutions of $q$-difference equations. *Discrete Math.*, 180(1-3):3–22, 1998.

[6] S. A. Abramov and E. V. Zima. D'Alembertian solutions of inhomogeneous linear equations (differential, difference, and some other). In *ISSAC'96*, page 232–240. ACM Press, 1996.

[7] B. Adamczewski, J. P. Bell, É. Delaygue, and F. Jouhet. Congruences modulo cyclotomic polynomials and algebraic independence for $q$-series. *Sém. Lothar. Combin.*, 78B:Art. 54, 12, 2017.

[8] C. R. Adams. On the linear ordinary $q$-difference equation. *Ann. of Math. (2)*, 30(1-4):195–205, 1928/29.

[9] C. R. Adams. Linear $q$-difference equations. *Bull. AMS*, 37(6):361–400, 1931.

[10] P. K. Agarwal, N. Amenta, and M. Sharir. Largest placement of one convex polygon inside another. *Discrete Comput. Geom.*, 19(1):95–104, 1998.

[11] M. Aissen. Variations on a theme of Pólya. In *Second International Conference on Combinatorial Mathematics (New York, 1978)*, volume 319 of *Ann. New York Acad. Sci.*, pages 1–6. 1979.

[12] W. A. Al-Salam and L. Carlitz. Some orthogonal $q$-polynomials. *Math. Nachr.*, 30:47–61, 1965.

[13] M. Aldaz, G. Matera, J. L. Montaña, and L. M. Pardo. A new method to obtain lower bounds for polynomial evaluation. *Theoret. Comput. Sci.*, 259(1-2):577–596, 2001.

[14] J.-P. Allouche and M. Mendès France. Hadamard grade of power series. *J. Number Theory*, 131(11):2013–2022, 2011.

[15] J. Alman and V. Vassilevska Williams. A refined laser method and faster matrix multiplication. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 522–539. [Society for Industrial and Applied Mathematics (SIAM)], Philadelphia, PA, 2021.

[16] G. Almkvist and D. Zeilberger. The method of differentiating under the integral sign. *J. Symbolic Comput.*, 10(6):571–591, 1990.

[17] Y. André. *$G$-functions and geometry*. Aspects of Mathematics, E13. Friedr. Vieweg & Sohn, Braunschweig, 1989.

[18] Y. André. Sur la conjecture des $p$-courbures de Grothendieck-Katz et un problème de Dwork. In *Geometric aspects of Dwork theory. Vol. I, II*, pages 55–112. Walter de Gruyter, Berlin, 2004.

[19] G. E. Andrews. Partition identities. *Advances in Math.*, 9:10–51, 1972.

[20] G. E. Andrews. A general theory of identities of the Rogers-Ramanujan type. *Bull. AMS*, 80:1033–1052, 1974.

[21] G. E. Andrews. *The theory of partitions*. Addison-Wesley Publishing Co., Reading,, 1976. Encyclopedia of Mathematics and its Applications, Vol. 2.

[22] G. E. Andrews. The fifth and seventh order mock theta functions. *Trans. Amer. Math. Soc.*, 293(1):113–134, 1986.

[23] G. E. Andrews. $q$-Catalan identities. In *The legacy of Alladi Ramakrishnan in the mathematical sciences*, pages 183–190. Springer, 2010.

[24] G. E. Andrews, R. Askey, and R. Roy. *Special functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.

[25] R. Apéry. Irrationalité de $\zeta(2)$ et $\zeta(3)$. *Astérisque*, (61):11–13, 1979.

[26] C. E. Arreche and Y. Zhang. Computing differential Galois groups of second-order linear $q$-difference equations. *Adv. in Appl. Math.*, 132:Paper No. 102273, 48, 2022.

[27] R. Askey. Continuous $q$-Hermite polynomials when $q > 1$. In *$q$-series and partitions*, volume 18 of *IMA Vol. Math. Appl.*, pages 151–158. Springer, 1989.

[28] D. H. Bailey, J. M. Borwein, A. Mattingly, and G. Wightwick. The computation of previously inaccessible digits of $\pi^2$ and Catalan's constant. *Notices Amer. Math. Soc.*, 60(7):844–854, 2013.

[29] D. Bar-Natan and S. Garoufalidis. On the Melvin-Morton-Rozansky conjecture. *Invent. Math.*, 125(1):103–133, 1996.

[30] M. Barkatou, T. Cluzeau, L. Di Vizio, and J.-A. Weil. Reduced forms of linear differential systems and the intrinsic Galois-Lie algebra of Katz. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 16:Paper No. 054, 13, 2020.

[31] M. Barkatou, T. Cluzeau, and A. El Hajj. Simple forms and rational solutions of pseudo-linear systems. In *ISSAC'19*, pages 26–33. ACM Press, 2019.

[32] M. Beeler, R. Gosper, and R. Schroeppel. *HAKMEM*. Artificial Intelligence Memo No. 239. MIT, 1972. http://www.inwap.com/pdp10/hbaker/hakmem/algorithms.

[33] R. Bellman. *A brief introduction to theta functions*. Athena Series: Selected Topics in Mathematics. Holt, Rinehart and Winston, 1961.

[34] B. C. Berndt, S. Bhargava, and F. G. Garvan. Ramanujan's theories of elliptic functions to alternative bases. *Transactions of the American Mathematical Society*, 347(11):4163, Nov. 1995.

[35] D. J. Bernstein. Fast multiplication and its applications. In *Algorithmic number theory: lattices, number fields, curves and cryptography*, volume 44 of *Math. Sci. Res. Inst. Publ.*, pages 325–384. Cambridge Univ. Press, 2008.

[36] D. J. Bernstein, L. De Feo, A. Leroux, and B. Smith. Faster computation of isogenies of large prime degree. In *ANTS XIV—Proceedings of the Fourteenth Algorithmic Number Theory Symposium*, volume 4 of *Open Book Ser.*, pages 39–55. Math. Sci. Publ., Berkeley, CA, 2020.

[37] F. Beukers. A note on the irrationality of $\zeta(2)$ and $\zeta(3)$. *Bull. London Math. Soc.*, 11(3):268–272, 1979.

[38] F. Beukers and G. Heckman. Monodromy for the hypergeometric function ${}_nF_{n-1}$. *Invent. Math.*, 95(2):325–354, 1989.

[39] F. Beukers, M. Houben, and A. Straub. Gauss congruences for rational functions in several variables. *Acta Arithmetica*, 184:341–362, 2018.

[40] A. Bezdek, Z. Guan, M. Hujter, and A. Joós. Cubes and boxes have Rupert's passages in every nontrivial direction. *Amer. Math. Monthly*, 128(6):534–542, 2021.

[41] J.-P. Bézivin. Les suites $q$-récurrentes linéaires. *Compositio Math.*, 80(3):285–307, 1991.

[42] J.-P. Bézivin. Sur les équations fonctionnelles aux $q$-différences. *Aequationes Math.*, 43(2-3):159–176, 1992.

[43] J.-P. Bézivin and A. Boutabaa. Sur les équations fonctionelles $p$-adiques aux $q$-différences. *Collect. Math.*, 43(2):125–140, 1992.

[44] M. Bläser. On the complexity of the multiplication of matrices of small formats. *J. Complexity*, 19(1):43–60, 2003.

[45] L. I. Bluestein. A linear filtering approach to the computation of the discrete Fourier transform. *IEEE Trans. Electroacoustics*, AU-18:451–455, 1970.

[46] J. W. Bober. Factorial ratios, hypergeometric series, and a family of step functions. *J. Lond. Math. Soc. (2)*, 79(2):422–444, 2009.

[47] H. Böing and W. Koepf. Algorithms for $q$-hypergeometric summation in computer algebra. *J. Symbolic Comput.*, 28(6):777–799, 1999.

[48] A. Borodin and S. Cook. On the number of additions to compute specific polynomials. *SIAM J. Comput.*, 5(1):146–157, 1976.

[49] P. B. Borwein. On the complexity of calculating factorials. *J. Algorithms*, 6(3):376–380, 1985.

[50] P. B. Borwein. Padé approximants for the $q$-elementary functions. *Constr. Approx.*, 4(4):391–402, 1988.

[51] A. Bostan. Computing the $N$-th Term of a $q$-Holonomic Sequence. In *ISSAC'20*, pages 46–53. ACM, 2020.

[52] A. Bostan, S. Boukraa, G. Christol, S. Hassani, and J.-M. Maillard. Ising $n$-fold integrals as diagonals of rational functions and integrality of series expansions: integrality versus modularity, 2012. arXiv:1211.6031. Extended version of [53].

[53] A. Bostan, S. Boukraa, G. Christol, S. Hassani, and J.-M. Maillard. Ising $n$-fold integrals as diagonals of rational functions and integrality of series expansions. *J. Phys. A*, 46(18):185202, 44, 2013.

[54] A. Bostan, S. Boukraa, J.-M. Maillard, and J.-A. Weil. Diagonals of rational functions and selected differential Galois groups. *J. Phys. A*, 48(50):504001, 29, 2015.

[55] A. Bostan, X. Caruso, G. Christol, and P. Dumas. Fast coefficient computation for algebraic power series in positive characteristic. In *Proceedings of the Thirteenth Algorithmic Number Theory Symposium*, volume 2 of *Open Book Ser.*, pages 119–135. Math. Sci. Publ., Berkeley, CA, 2019.

[56] A. Bostan, X. Caruso, and É. Schost. A fast algorithm for computing the characteristic polynomial of the $p$-curvature. In *ISSAC'14*, pages 59–66. ACM Press, 2014.

[57] A. Bostan, X. Caruso, and É. Schost. A fast algorithm for computing the $p$-curvature. In *ISSAC'15*, pages 69–76. ACM, New York, 2015.

[58] A. Bostan, X. Caruso, and É. Schost. Computation of the similarity class of the $p$-curvature. In *ISSAC'16*, pages 111–118. ACM Press, 2016.

[59] A. Bostan, S. Chen, F. Chyzak, and Z. Li. Complexity of creative telescoping for bivariate rational functions. In *ISSAC'10*, pages 203–210. ACM, 2010.

[60] A. Bostan, S. Chen, F. Chyzak, Z. Li, and G. Xin. Hermite reduction and creative telescoping for hyperexponential functions. In *ISSAC'13*, pages 77–84. ACM, 2013.

[61] A. Bostan, G. Christol, and P. Dumas. Fast computation of the $N$th term of an algebraic series over a finite prime field. In *Proceedings of the 2016 ACM International Symposium on Symbolic and Algebraic Computation*, pages 119–126. ACM, New York, 2016.

[62] A. Bostan, F. Chyzak, T. Cluzeau, and B. Salvy. Low complexity algorithms for linear recurrences. In *ISSAC'06*, pages 31–38. ACM Press, 2006.

[63] A. Bostan, F. Chyzak, M. Giusti, R. Lebreton, G. Lecerf, B. Salvy, and É. Schost. *Algorithmes Efficaces en Calcul Formel*. Palaiseau, Sept. 2017. 686 pages, in French. Printed by CreateSpace. Available in electronic form.

[64] A. Bostan, F. Chyzak, P. Lairez, and B. Salvy. Generalized Hermite reduction, creative telescoping and definite integration of D-finite functions. In *ISSAC'18*, pages 95–102. ACM, 2018.

[65] A. Bostan, F. Chyzak, G. Lecerf, B. Salvy, and É. Schost. Differential equations for algebraic functions. In *ISSAC 2007*, pages 25–32. ACM, New York, 2007.

[66] A. Bostan, F. Chyzak, Z. Li, and B. Salvy. Fast computation of common left multiples of linear ordinary differential operators. In *ISSAC'12*, pages 99–106. ACM, 2012.

[67] A. Bostan, T. Cluzeau, and B. Salvy. Fast algorithms for polynomial solutions of linear differential equations. In *ISSAC'05*, pages 45–52. ACM Press, 2005.

[68] A. Bostan, P. Flajolet, B. Salvy, and É. Schost. Fast computation of special resultants. *J. Symbolic Comput.*, 41(1):1–29, 2006.

[69] A. Bostan, P. Gaudry, and É. Schost. Linear recurrences with polynomial coefficients and application to integer factorization and Cartier-Manin operator. *SIAM J. Comput.*, 36(6):1777–1806, 2007.

[70] A. Bostan and M. Kauers. Automatic classification of restricted lattice walks. In *FPSAC 2009*, Discrete Math. Theor. Comput. Sci. Proc., AK, pages 201–215. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2009.

[71] A. Bostan, P. Lairez, and B. Salvy. Creative telescoping for rational functions using the Griffiths-Dwork method. In *ISSAC'13*, pages 93–100. ACM, 2013.

[72] A. Bostan, P. Lairez, and B. Salvy. Multiple binomial sums. *J. Symbolic Comput.*, 80(part 2):351–386, 2017.

[73] A. Bostan, G. Lecerf, and É. Schost. Tellegen's principle into practice. In *Proceedings of ISSAC'03*, pages 37–44. ACM Press, 2003.

[74] A. Bostan and R. Mori. A simple and fast algorithm for computing the $N$-th term of a linearly recurrent sequence. In *4th SIAM Symposium on Simplicity in Algorithms*, pages 118–132. 2021.

[75] A. Bostan, V. Neiger, and S. Yurkevich. Beating binary powering for polynomial matrices. Accepted for publication in ISSAC'23, Feb. 2023.

[76] A. Bostan, T. Rivoal, and B. Salvy. Minimization of differential equations and algebraic values of $E$-functions. arXiv: 2209.01827, Sept. 2022.

[77] A. Bostan and É. Schost. Polynomial evaluation and interpolation on special sets of points. *J. Complexity*, 21(4):420–446, 2005.

[78] A. Bostan and É. Schost. Fast algorithms for differential equations in positive characteristic. In *ISSAC'09*, pages 47–54. ACM Press, 2009.

[79] A. Bostan, A. Straub, and S. Yurkevich. On the representability of sequences as constant terms. arXiv: 2212.10116, Dec. 2022.

[80] A. Bostan, L. D. Vizio, and K. Raschel. Differential transcendence of Bell numbers and relatives: a Galois theoretic approach, 2021. arXiv: 2012.15292.

[81] A. Bostan and S. Yurkevich. A hypergeometric proof that Iso is bijective. *Proc. Amer. Math. Soc.*, 150(5):2131–2136, 2022.

[82] A. Bostan and S. Yurkevich. On a class of hypergeometric diagonals. *Proc. Amer. Math. Soc.*, 150(3):1071–1087, 2022.

[83] A. Bostan and S. Yurkevich. Fast computation of the $N$-th term of a $q$-holonomic sequence and applications. *J. Symbolic Comput.*, 115:96–123, 2023.

[84] A. Bostan and S. Yurkevich. On the $q$-Analogue of Pólya's Theorem. *Electron. J. Combin.*, 30(2):Paper No. 2.9, 2023.

[85] D. Boucher. About the polynomial solutions of homogeneous linear differential equations depending on parameters. In *Proceedings of the 1999 International Symposium on Symbolic and Algebraic Computation (Vancouver, BC)*, pages 261–268. ACM, New York, 1999.

[86] M. Bousquet-Mélou. Convex polyominoes and algebraic languages. *J. Phys. A*, 25(7):1935–1944, 1992.

[87] P. Bürgisser, M. Clausen, and M. A. Shokrollahi. *Algebraic complexity theory*, volume 315 of *Grundlehren der Mathematischen Wissenschaften*. Springer, 1997.

[88] P. F. Byrd. Expansion of analytic functions in polynomials associated with Fibonacci numbers. *Fibonacci Quart.*, 1(1):16–29, 1963.

[89] P. Canham. The minimum energy of bending as a possible explanation of the biconcave shape of the human red blood cell. *Journal of Theoretical Biology*, 26(1):61–81, 1970.

[90] D. G. Cantor and E. Kaltofen. On fast multiplication of polynomials over arbitrary algebras. *Acta Inform.*, 28(7):693–701, 1991.

[91] M. Capobianco, S. Maurer, D. McCarthy, and J. Molluzzo. A collection of open problems. *Annals of the New York Academy of Sciences*, 319(1):565–592, 1979.

[92] R. D. Carmichael. The General Theory of Linear $q$-Difference Equations. *Amer. J. Math.*, 34(2):147–168, 1912.

[93] P. Cartier. Démonstration "automatique" d'identités et fonctions hypergéométriques (d'après D. Zeilberger). *Astérisque*, (206):Exp. No. 746, 3, 41–91, 1992. Séminaire Bourbaki, Vol. 1991/92.

[94] Y. Chai, L. Yuan, and T. Zamfirescu. Rupert property of Archimedean solids. *Amer. Math. Monthly*, 125(6):497–504, 2018.

[95] B. Chazelle. The polygon containment problem. *Advances in Computing Research I*, pages 1–33, 1983.

[96] J. Chen, T. Yu, P. Brogan, R. Kusner, Y. Yang, and A. Zigerelli. Numerical methods for biomembranes: conforming subdivision methods versus non-conforming PL methods. *Math. Comp.*, 90(328):471–516, 2021.

[97] S. Chen and M. Kauers. Some open problems related to creative telescoping. *J. Syst. Sci. Complex.*, 30(1):154–172, 2017.

[98] G. Christol. Ensembles presque periodiques $k$-reconnaissables. *Theoret. Comput. Sci.*, 9(1):141–145, 1979.

[99] G. Christol. Fonctions hypergéométriques bornées. *Groupe de travail d'analyse ultramétrique*, 14:1–16, 1986-1987. Talk no. 8.

[100] G. Christol. Diagonales de fractions rationnelles. In *Séminaire de Théorie des Nombres, Paris 1986–87*, volume 75 of *Progr. Math.*, pages 65–90. Birkhäuser Boston, Boston, MA, 1988.

[101] G. Christol. Globally bounded solutions of differential equations. In *Analytic number theory (Tokyo, 1988)*, volume 1434 of *Lecture Notes in Math.*, pages 45–64. Springer, Berlin, 1990.

[102] G. Christol. Diagonals of rational fractions. *Eur. Math. Soc. Newsl.*, (97):37–43, 2015.

[103] D. V. Chudnovsky and G. V. Chudnovsky. Applications of Padé approximations to the Grothendieck conjecture on linear differential equations. In *Number theory (New York, 1983–84)*, volume 1135 of *Lecture Notes in Math.*, pages 52–100. Springer, Berlin, 1985.

[104] D. V. Chudnovsky and G. V. Chudnovsky. On expansion of algebraic functions in power and Puiseux series. I. *J. Complexity*, 2(4):271–294, 1986.

[105] D. V. Chudnovsky and G. V. Chudnovsky. Computer assisted number theory with applications. In *Number theory (New York, 1984–1985)*, volume 1240 of *Lecture Notes in Math.*, pages 1–68. Springer, Berlin, 1987.

[106] D. V. Chudnovsky and G. V. Chudnovsky. Approximations and complex multiplication according to Ramanujan. In *Ramanujan revisited (Urbana-Champaign, Ill., 1987)*, pages 375–472. Academic Press, Boston, MA, 1988.

[107] D. V. Chudnovsky and G. V. Chudnovsky. Computer algebra in the service of mathematical physics and number theory. In *Computers in mathematics (Stanford, CA, 1986)*, volume 125 of *Lecture Notes in Pure and Appl. Math.*, pages 109–232. Dekker, New York, 1990.

[108] G. Chudnovsky. Algebraic independence of the values of elliptic function at algebraic points. *Invent. Math.*, 61(3):267–290, 1980. Elliptic analogue of the Lindemann-Weierstrass theorem.

[109] G. V. Chudnovsky. Algebraic independence of values of exponential and elliptic functions. In *Contributions to the theory of transcendental numbers*, volume 19 of *Math. Surveys Monogr.*, pages 1–26. Amer. Math. Soc., Providence, RI, 1984.

[110] F. Chyzak. Gröbner bases, symbolic summation and symbolic integration. In *Gröbner bases and applications (Linz, 1998)*, volume 251 of *London Math. Soc. Lecture Note Ser.*, pages 32–60. Cambridge Univ. Press, 1998.

[111] F. Chyzak. An extension of Zeilberger's fast algorithm to general holonomic functions. *Discrete Math.*, 217(1-3):115–134, 2000.

[112] F. Chyzak. *The ABC of Creative Telescoping: Algorithms, Bounds, Complexity*. HDR (accreditation to supervise research), University Paris-Sud 11, Apr. 2014. 64 pages.

[113] F. Chyzak, P. Dumas, H. Le, J. Martin, M. Mishna, and B. Salvy. Taming apparent singularities via Ore closure. Manuscript, 2016.

[114] L. Comtet. Calcul pratique des coefficients de Taylor d'une fonction algébrique. *Enseign. Math. (2)*, 10:267–270, 1964.

[115] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comp.*, 19:297–301, 1965.

[116] E. Costa, R. Gerbicz, and D. Harvey. A search for Wilson primes. *Math. Comp.*, 83(290):3071–3091, 2014.

[117] D. Couty, J. Esterle, and R. Zarouf. Décomposition effective de Jordan-Chevalley. *Gaz. Math.*, (129):29–49, 2011.

[118] S. Czirbusz. Comparing the computation of Chebyshev polynomials in computer algebra systems. *Ann. Univ. Sci. Budapest. Sect. Comput.*, 36:23–39, 2012.

[119] A. Denjoy. Sur les courbes definies par les équations différentielles à la surface du tore. *J. Math. Pures Appl. (9)*, 11:333–375, 1932.

[120] J. Désarménien. Un analogue des congruences de Kummer pour les $q$-nombres d'Euler. *European J. Combin.*, 3(1):19–28, 1982.

[121] R. Detcherry and S. Garoufalidis. A diagrammatic approach to the AJ conjecture. *Math. Ann.*, 378(1-2):447–484, 2020.

[122] R. L. Devaney. *An introduction to chaotic dynamical systems*. Addison-Wesley Studies in Nonlinearity. Addison-Wesley Publishing Company, Advanced Book Program, Redwood City, CA, second edition, 1989.

[123] L. Di Vizio. Arithmetic theory of $q$-difference equations: the $q$-analogue of Grothendieck-Katz's conjecture on $p$-curvatures. *Invent. Math.*, 150(3):517–578, 2002.

[124] L. Di Vizio and C. Hardouin. Intrinsic approach to Galois theory of $q$-difference equations. *Mem. Amer. Math. Soc.*, 279(1376):xii+70, 2022. With a preface to part 4 by Anne Granier.

[125] L. Di Vizio, J.-P. Ramis, J. Sauloy, and C. Zhang. Équations aux $q$-différences. *Gaz. Math.*, (96):20–49, 2003.

[126] B. Dubrovin, D. Yang, and D. Zagier. Geometry and arithmetic of integrable hierarchies of KdV type. I. Integrality, 2021.

[127] B. M. Dwork and A. J. van der Poorten. The Eisenstein constant. *Duke Math. J.*, 65(1):23–43, 1992.

[128] G. Eisenstein. Über eine allgemeine Eigenschaft der Reihen-Entwicklungen aller algebraischen Funktionen. Berichte Königl. Preuss. Akad. Wiss. Berlin, pages 441–443, 1852.

[129] T. Ekedahl and G. van der Geer. Cycle classes on the moduli of K3 surfaces in positive characteristic. *Selecta Math. (N.S.)*, 21(1):245–291, 2015.

[130] S. B. Ekhad and D. Zeilberger. The number of solutions of $X^2 = 0$ in triangular matrices over $\mathrm{GF}(q)$. *Electron. J. Combin.*, 3(1):Research Paper 2, approx. 2, 1996.

[131] S. N. Elaydi and W. A. Harris, Jr. On the computation of $A^n$. *SIAM Rev.*, 40(4):965–971, 1998.

[132] T. Ernst. *A comprehensive treatment of $q$-calculus*. Birkhäuser/Springer, 2012.

[133] A. Errera. Zahlentheoretische Lösung einer functionentheoretischen Frage. *Rend. Circ. Mat. Palermo*, 35:107–144, 1913.

[134] G. Everest, A. van der Poorten, I. Shparlinski, and T. Ward. *Recurrence Sequences*, volume 104 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2003.

[135] R. Feng. Hrushovski's algorithm for computing the Galois group of a linear differential equation. *Adv. in Appl. Math.*, 65:1–37, 2015.

[136] C. M. Fiduccia. An efficient formula for linear recurrences. *SIAM J. Comput.*, 14(1):106–112, 1985.

[137] P. Flajolet. Analytic models and ambiguity of context-free languages. volume 49, pages 283–309. 1987. Twelfth international colloquium on automata, languages and programming (Nafplion, 1985).

[138] P. Flajolet and B. Salvy. The SIGSAM challenges: symbolic asymptotics in practice. *ACM SIGSAM Bull.*, 31(4):36–47, 1997.

[139] B. Fourcade. Theoretical results on toroidal vesicles. *J. Physique II*, 2(9):1705–1724, 1992.

[140] A. Fredriksson. The triakis tetrahedron and the pentagonal icositetrahedron are Rupert, 2022. arXiv: 2210.00601.

[141] G. Frobenius. Ueber die Integration der linearen Differentialgleichungen durch Reihen. *J. Reine Angew. Math.*, 76:214–235, 1873.

[142] L. Fuchs. Zur Theorie der linearen Differentialgleichungen mit veränderlichen Coefficienten. *J. Reine Angew. Math.*, 66:121–160, 1866.

[143] M. Fürer. Faster integer multiplication. *SIAM J. Comput.*, 39(3):979–1005, 2009.

[144] J. Fürlinger and J. Hofbauer. $q$-Catalan numbers. *J. Combin. Theory Ser. A*, 40(2):248–264, 1985.

[145] H. Furstenberg. Algebraic functions over finite fields. *J. Algebra*, 7:271–277, 1967.

[146] J.-P. Furter. Polynomial composition rigidity and plane polynomial automorphisms. *J. Lond. Math. Soc. (2)*, 91(1):180–202, 2015.

[147] F. Fürnsinn and H. Hauser. Normal forms of ordinary linear differential equations in arbitrary characteristic, 2023. Preprint.

[148] F. L. Gall. Powers of tensors and fast matrix multiplication. In *ISSAC'14*, pages 296–303, 2014.

[149] S. Garoufalidis. On the characteristic and deformation varieties of a knot. In *Proceedings of the Casson Fest*, volume 7 of *Geom. Topol. Monogr.*, pages 291–309. Geom. Topol. Publ., Coventry, 2004.

[150] S. Garoufalidis. The degree of a $q$-holonomic sequence is a quadratic quasi-polynomial. *Electron. J. Combin.*, 18(2):Paper 4, 23, 2011.

[151] S. Garoufalidis. Quantum knot invariants. *Res. Math. Sci.*, 5(1):Paper No. 11, 17, 2018.

[152] S. Garoufalidis and C. Koutschan. Twisting $q$-holonomic sequences by complex roots of unity. In *ISSAC'12*, pages 179–186. ACM Press, 2012.

[153] S. Garoufalidis and C. Koutschan. Irreducibility of $q$-difference operators and the knot $7_4$. *Algebr. Geom. Topol.*, 13(6):3261–3286, 2013.

[154] S. Garoufalidis, A. D. Lauda, and T. T. Q. Lê. The colored HOMFLYPT function is $q$-holonomic. *Duke Math. J.*, 167(3):397–447, 2018.

[155] S. Garoufalidis and T. T. Q. Lê. The colored Jones function is $q$-holonomic. *Geom. Topol.*, 9:1253–1293, 2005.

[156] S. Garoufalidis and T. T. Q. Lê. A survey of $q$-holonomic functions. *Enseign. Math.*, 62(3-4):501–525, 2016.

[157] S. Garrabrant and I. Pak. Counting with irrational tiles, 2014. arXiv: 1407.8222.

[158] F. G. Garvan. New fifth and seventh order mock theta function identities. *Ann. Comb.*, 23(3-4):765–783, 2019.

[159] J. von zur Gathen and J. Gerhard. *Modern computer algebra*. Cambridge Univ. Press, third edition, 2013.

[160] C. F. Gauss. Summatio quarundam serierum singularium. Opera, Vol. 2, Göttingen: Gess. d. Wiss., 1863.

[161] I. Gessel. A noncommutative generalization and $q$-analog of the Lagrange inversion formula. *Trans. Amer. Math. Soc.*, 257(2):455–482, 1980.

[162] O. Gorodetsky. New representations for all sporadic Apéry-like sequences, with applications to congruences. *Experimental Mathematics*, 2021.

[163] H. W. Gould. *Combinatorial identities*. Henry W. Gould, Morgantown, W.Va., 1972. A standardized set of tables listing 500 binomial coefficient summations.

[164] D. Y. Grigor'ev and N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *Journal of Symbolic Computation*, 5(1):37–64, 1988.

[165] W. Hahn. Über die höheren Heineschen Reihen und eine einheitliche Theorie der sogenannten speziellen Funktionen. *Math. Nachr.*, 3:257–294, 1950.

[166] G. Hanrot, M. Quercia, and P. Zimmermann. The middle product algorithm. I. *Appl. Algebra Engrg. Comm. Comput.*, 14(6):415–438, 2004.

[167] D. Harvey. Counting points on hyperelliptic curves in average polynomial time. *Ann. of Math. (2)*, 179(2):783–803, 2014.

[168] D. Harvey. An exponent one-fifth algorithm for deterministic integer factorisation. *Math. Comp.*, 90(332):2937–2950, 2021.

[169] D. Harvey and J. van der Hoeven. Integer multiplication in time $O(n \log n)$. *Ann. of Math. (2)*, 193(2):563–617, 2021.

[170] D. Harvey and J. van der Hoeven. Polynomial multiplication over finite fields in time $O(n \log n)$. *J. ACM*, 69(2):Art. 12, 40, 2022.

[171] D. Harvey, J. van der Hoeven, and G. Lecerf. Even faster integer multiplication. *J. Complexity*, 36:1–30, 2016.

[172] H. Hauser. How regular are regular singularities? In *Transcendence in algebra, combinatorics, geometry and number theory*, volume 373 of *Springer Proc. Math. Stat.*, pages 261–273. Springer, Cham, 2021.

[173] H. Hauser. Fuchsian differential equations, 2023. In preparation.

[174] E. Heine. Untersuchungen über die Reihe $1 + \frac{(1-q^\alpha)(1-q^\beta)}{(1-q)(1-q^\gamma)}x + \frac{(1-q^\alpha)(1-q^{\alpha+1})(1-q^\beta)(1-q^{\beta+1})}{(1-q)(1-q^2)(1-q^\gamma)(1-q^{\gamma+1})}x^2 + \ldots$ *J. reine angew. Math.*, 34:285–328, 1847.

[175] E. Heine. Der Eisensteinsche Satz über Reihen-Entwickelung algebraischer Functionen. *J. Reine Angew. Math.*, 45:285–302, 1853.

[176] E. Heine. Ueber die Entwickelung von Wurzeln algebraischer Gleichungen in Potenzreihen. *J. Reine Angew. Math.*, 48:267–275, 1854.

[177] J. Heintz and M. Sieveking. Lower bounds for polynomials with algebraic coefficients. *Theoret. Comput. Sci.*, 11(3):321–330, 1980.

[178] P. A. Hendriks. An algorithm for computing a standard form for second-order linear $q$-difference equations. *J. Pure Appl. Algebra*, 117/118:331–352, 1997.

[179] M. Hittmeir. A time-space tradeoff for Lehman's deterministic integer factorization method. *Math. Comp.*, 90(330):1999–2010, 2021.

[180] B. Hoffmann. Rupert properties of polyhedra and the generalised Nieuwland constant. *J. Geom. Graph.*, 23(1):29–35, 2019.

[181] T. Honda. Algebraic differential equations. In *Symposia Mathematica, Vol. XXIV (Sympos., INDAM, Rome, 1979)*, pages 169–204. Academic Press, London-New York, 1981.

[182] J. E. Hopcroft and L. R. Kerr. On minimizing the number of multiplications necessary for matrix multiplication. *SIAM J. Appl. Math.*, 20:30–36, 1971.

[183] E. Hrushovski. Computing the Galois group of a linear differential equation. In *Differential Galois theory (Bedlewo, 2001)*, volume 58 of *Banach Center Publ.*, pages 97–138. Polish Acad. Sci. Inst. Math., Warsaw, 2002.

[184] P.-F. Hsieh, M. Kohno, and Y. Sibuya. Construction of a fundamental matrix solution at a singular point of the first kind by means of the $SN$ decomposition of matrices. *Linear Algebra Appl.*, 239:29–76, 1996.

[185] J. Hua. Counting representations of quivers over finite fields. *J. Algebra*, 226(2):1011–1033, 2000.

[186] S. G. Hyun, V. Neiger, and É. Schost. Implementations of efficient univariate polynomial matrix algorithms and application to bivariate resultants. In *ISSAC'19*, pages 235–242. ACM, 2019.

[187] M. E. H. Ismail. Lectures on $q$-orthogonal polynomials. In *Special functions 2000: current perspective and future directions (Tempe, AZ)*, volume 30 of *NATO Sci. Ser. II Math. Phys. Chem.*, pages 179–219. Kluwer Acad. Publ., 2001.

[188] F. H. Jackson. On $q$-functions and a certain difference operator. *Trans. Roy. Soc. Edin.*, 46(11):253–281, 1909.

[189] F. H. Jackson. On $q$-definite integrals. *Quar. J. Pure Appl. Math.*, 41:193–203, 1910.

[190] F. H. Jackson. $q$-Difference Equations. *Amer. J. Math.*, 32(4):305–314, 1910.

[191] R. P. Jerrard, J. E. Wetzel, and L. Yuan. Platonic passages. *Math. Mag.*, 90(2):87–98, 2017.

[192] R. Jungen. Sur les séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence. *Comment. Math. Helv.*, 3(1):266–306, 1931.

[193] N. M. Katz. Algebraic solutions of differential equations ($p$-curvature and the Hodge filtration). *Invent. Math.*, 18:1–118, 1972.

[194] M. Kauers. Bounds for D-finite closure properties. In *ISSAC'14*, pages 288–295. ACM, 2014.

[195] M. Kauers and C. Koutschan. A Mathematica package for $q$-holonomic sequences and power series. *Ramanujan J.*, 19(2):137–150, 2009.

[196] M. Kauers and C. Koutschan. Guessing with little data. In *ISSAC '22—Proceedings of the 2022 International Symposium on Symbolic and Algebraic Computation*, pages 83–90. ACM, New York, 2022.

[197] M. Kauers and C. Koutschan. Some D-finite and some possibly D-finite sequences in the OEIS. *Journal of Integer Sequences*, 26(4):Article 23.4.5, 2023.

[198] M. Kauers and M. Mezzarobba. Multivariate Ore polynomials in SageMath. *ACM Commun. Comput. Algebra*, 53(2):57–60, 2019.

[199] M. Kauers and P. Paule. *The Concrete Tetrahedron*. Springer-Verlag, 2011.

[200] M. Kauers and V. Pillwein. When can we detect that a P-finite sequence is positive? In *ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, pages 195–201. ACM, New York, 2010.

[201] M. Kauers and G. Pogudin. Bounds for substituting algebraic functions into D-finite functions. In *ISSAC'17*, pages 245–252. ACM, 2017.

[202] L. G. A. Keller, A. Mondino, and T. Rivière. Embedded surfaces of arbitrary genus minimizing the Willmore energy under isoperimetric constraint. *Arch. Ration. Mech. Anal.*, 212(2):645–682, 2014.

[203] G. Kenison, J. Nieuwveld, J. Ouaknine, and J. Worrell. Positivity problems for reversible linear recurrence sequences, 2023. Preprint.

[204] D. E. Khmel′ nov. Improved algorithms for solving difference and $q$-difference equations. *Programmirovanie*, (2):70–78, 2000.

[205] A. A. Kirillov. On the number of solutions of the equation $X^2 = 0$ in triangular matrices over a finite field. *Funktsional. Anal. i Prilozhen.*, 29(1):82–87, 1995.

[206] A. A. Kirillov and A. Melnikov. On a remarkable sequence of polynomials. In *Algèbre non commutative, groupes quantiques et invariants (Reims, 1995)*, volume 2 of *Sémin. Congr.*, pages 35–42. Soc. Math. France, Paris, 1997.

[207] R. Koekoek, P. A. Lesky, and R. F. Swarttouw. *Hypergeometric orthogonal polynomials and their $q$-analogues*. Monographs in Mathematics. Springer, 2010.

[208] W. Koepf. Efficient computation of Chebyshev polynomials in computer algebra. In *Computer Algebra Systems: A Practical Guide*, pages 79–99. Wiley, 1999.

[209] W. Koepf, P. M. Rajković, and S. D. Marinković. Properties of $q$-holonomic functions. *J. Difference Equ. Appl.*, 13(7):621–638, 2007.

[210] E. R. Kolchin. Algebraic matric groups and the Picard-Vessiot theory of homogeneous linear ordinary differential equations. *Ann. of Math. (2)*, 49:1–42, 1948.

[211] M. Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Comm. Math. Phys.*, 147(1):1–23, 1992.

[212] T. H. Koornwinder. On Zeilberger's algorithm and its $q$-analogue. *J. Comput. Appl. Math.*, 48(1-2):91–111, 1993.

[213] C. Koutschan. A fast approach to creative telescoping. *Math. Comput. Sci.*, 4(2-3):259–266, 2010.

[214] C. Koutschan. Creative telescoping for holonomic functions. In *Computer algebra in quantum field theory*, Texts Monogr. Symbol. Comput., pages 171–194. Springer, Vienna, 2013.

[215] C. Koutschan and Y. Zhang. Desingularization in the $q$-Weyl algebra. *Adv. in Appl. Math.*, 97:80–101, 2018.

[216] J. J. Kovacic. An algorithm for solving second order linear homogeneous differential equations. *J. Symbolic Comput.*, 2(1):3–43, 1986.

[217] E. E. Kummer. Über die hypergeometrische Reihe $1 + \frac{\alpha\cdot\beta}{1\cdot\gamma}x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{1\cdot2\cdot\gamma(\gamma+1)}x^2 + \frac{\alpha(\alpha+1)(\alpha+2)\beta(\beta+1)(\beta+2)}{1\cdot2\cdot3\cdot\gamma(\gamma+1)(\gamma+2)}x^3 + \ldots$. (Fortsetzung). *J. Reine Angew. Math.*, 15:127–172, 1836.

[218] H. Labrande. Computing Jacobi's theta in quasi-linear time. *Math. Comp.*, 87(311):1479–1508, 2018.

[219] H. Labrande and E. Thomé. Computing theta functions in quasi-linear time in genus two and above. *LMS J. Comput. Math.*, 19(suppl. A):163–177, 2016.

[220] J. D. Laderman. A noncommutative algorithm for multiplying $3 \times 3$ matrices using $23$ multiplications. *Bull. Amer. Math. Soc.*, 82(1):126–128, 1976.

[221] S. Laishram, D. López-Aguayo, C. Pomerance, and T. Thongjunthug. Progress towards a nonintegrality conjecture. *Eur. J. Math.*, 6(4):1496–1504, 2020.

[222] E. Landau. Sur les conditions de divisibilité d'un produit de factorielles par un autre. *Nouvelles annales de mathématique*, 3e série, 19:344–362, 1900.

[223] E. Landau. Eine Anwendung des Eisensteinschen Satzes auf die Theorie der Gaussschen Differentialgleichung. *J. Reine Angew. Math.*, 127:92–102, 1904.

[224] E. Landau. Über einen zahlentheoretischen satz und seine anwendung auf die hypergeometrische reihe. *Sitzungsber. Heidelb. Akad. Wiss. Math.-Natur. Kl.*, (18):3–38, 1911.

[225] S. Lang. Relations de distributions et exemples classiques. In *Séminaire Delange-Pisot-Poitou, 19e année: 1977/78, Théorie des nombres, Fasc. 2*, pages Exp. No. 40, 6. Secrétariat Math., Paris, 1978.

[226] G. Lavau. The truncated tetrahedron is Rupert. *Amer. Math. Monthly*, 126(10):929–932, 2019.

[227] J. Le Caine. The linear $q$-difference equation of the second order. *Amer. J. Math.*, 65:585–600, 1943.

[228] B. Le Stum and A. Quirós. On quantum integers and rationals. In *Trends in number theory*, volume 649 of *Contemp. Math.*, pages 107–130. Amer. Math. Soc., Providence, RI, 2015.

[229] L. Lipshitz. The diagonal of a $D$-finite power series is $D$-finite. *J. Algebra*, 113(2):373–378, 1988.

[230] R. J. Lipton. Polynomials with $0 - 1$ coefficients that are hard to evaluate. *SIAM J. Comput.*, 7(1):61–69, 1978.

[231] J.-C. Liu. Some finite generalizations of Euler's pentagonal number theorem. *Czechoslovak Math. J.*, 67(142)(2):525–531, 2017.

[232] L. L. Liu. Positivity of three-term recurrence sequences. *Electron. J. Combin.*, 17(1):Research Paper 57, 10, 2010.

[233] D. López-Aguayo. Non-integrality of binomial sums and Fermat's little theorem. *Math. Mag.*, 88(3):231–234, 2015.

[234] F. Luca and C. Pomerance. On a nonintegrality conjecture. *Eur. J. Math.*, 8(2):634–639, 2022.

[235] O. M. Makarov. An algorithm for multiplication of $3 \times 3$ matrices. *Zh. Vychisl. Mat. i Mat. Fiz.*, 26(2):293–294, 320, 1986.

[236] F. C. Marques and A. Neves. Min-max theory and the Willmore conjecture. *Ann. of Math. (2)*, 179(2):683–782, 2014.

[237] T. E. Mason. On Properties of the Solutions of Linear $q$-Difference Equations with Entire Function Coefficients. *Amer. J. Math.*, 37(4):439–444, 1915.

[238] P. Massazza and R. Radicioni. On computing the coefficients of bivariate holonomic formal series. *Theoret. Comput. Sci.*, 346(2-3):418–438, 2005.

[239] R. Mechik. Sur la constante d'Eisenstein. *Ann. Math. Blaise Pascal*, 15(1):87–108, 2008.

[240] S. Melczer. *Algorithmic and symbolic combinatorics—an invitation to analytic combinatorics in several variables*. Texts and Monographs in Symbolic Computation. Springer, Cham, 2021.

[241] S. Melczer and M. Mezzarobba. Sequence positivity through numeric analytic continuation: uniqueness of the Canham model for biomembranes. *Comb. Theory*, 2(2):Paper No. 4, 20, 2022.

[242] M. Mezzarobba. NumGfun: a package for numerical and analytic computation and D-finite functions. In *ISSAC 2010—Proceedings of the 2010 International Symposium on Symbolic and Algebraic Computation*, pages 139–146. ACM, New York, 2010.

[243] A. R. Miller and R. B. Paris. Euler-type transformations for the generalized hypergeometric function ${}_{r+2}F_{r+1}(x)$. *Z. Angew. Math. Phys.*, 62(1):31–45, 2011.

[244] A. R. Miller and R. B. Paris. Transformation formulas for the generalized hypergeometric function with integral parameter differences. *Rocky Mountain J. Math.*, 43(1):291–327, 2013.

[245] J. C. P. Miller and D. J. S. Brown. An algorithm for evaluation of remote terms in a linear recurrence sequence. *Comput. J.*, 9:188–190, 1966.

[246] G. T. Minton. Linear recurrence sequences satisfying congruence conditions. *Proceedings of the American Mathematical Society*, 142(7):2337–2352, Apr. 2014.

[247] S. Morier-Genoud and V. Ovsienko. $q$-deformed rationals and $q$-continued fractions. *Forum Math. Sigma*, 8:Paper No. e13, 55, 2020.

[248] S. Morier-Genoud and V. Ovsienko. On $q$-deformed real numbers. *Exp. Math.*, 31(2):652–660, 2022.

[249] D. Nogneng and É. Schost. On the evaluation of some sparse polynomials. *Math. Comp.*, 87(310):893–904, 2018.

[250] C. F. Osgood. On the diophantine approximation of values of functions satisfying certain linear $q$-difference equations. *J. Number Theory*, 3:159–177, 1971.

[251] A. Ostrowski. On two problems in abstract algebra connected with Horner's rule. In *Studies in mathematics and mechanics presented to Richard von Mises*, pages 40–48. Academic Press Inc., 1954.

[252] J. Ouaknine and J. Worrell. Decision problems for linear recurrence sequences. In *Reachability problems*, volume 7550 of *Lecture Notes in Comput. Sci.*, pages 21–28. Springer, Heidelberg, 2012.

[253] J. Ouaknine and J. Worrell. On the positivity problem for simple linear recurrence sequences. In *Automata, languages, and programming. Part II*, volume 8573 of *Lecture Notes in Comput. Sci.*, pages 318–329. Springer, Heidelberg, 2014.

[254] I. Pak. Partition bijections, a survey. *Ramanujan J.*, 12(1):5–75, 2006.

[255] I. Pak. Complexity problems in enumerative combinatorics. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 3153–3180. World Sci. Publ., Hackensack, NJ, 2018.

[256] V. Y. Pan. Methods of computing values of polynomials. *Russian Mathematical Surveys*, 21(1):105–136, 1966.

[257] M. Paterson and L. Stockmeyer. Bounds on the evaluation time for rational polynomial. In *12th Annual Symposium on Switching and Automata Theory (SWAT 1971)*, pages 140–143, 1971.

[258] M. S. Paterson and L. J. Stockmeyer. On the number of nonscalar multiplications necessary to evaluate polynomials. *SIAM J. Comput.*, 2:60–66, 1973.

[259] P. Paule and S. Radu. Rogers-Ramanujan functions, modular functions, and computer algebra. In *Advances in computer algebra*, volume 226 of *Springer Proc. Math. Stat.*, pages 229–280. Springer, Cham, 2018.

[260] P. Paule and A. Riese. A Mathematica $q$-analogue of Zeilberger's algorithm based on an algebraically motivated approach to $q$-hypergeometric telescoping. In *Special functions, $q$-series and related topics*, volume 14 of *Fields Inst. Commun.*, pages 179–210. Amer. Math. Soc., 1997.

[261] Y. Pei, Y. Wang, and Y. Wang. Positivity problem of three-term recurrence sequences. *Linear Algebra Appl.*, 668:93–109, 2023.

[262] M. Petkovšek. Hypergeometric solutions of linear recurrences with polynomial coefficients. *J. Symbolic Comput.*, 14(2-3):243–264, 1992.

[263] M. Petkovšek, H. S. Wilf, and D. Zeilberger. $A = B$. A K Peters, 1996.

[264] E. Picard. Sur les intégrales doubles de seconde espèce dans la théorie des surfaces algébriques (premier mémoire). *Journal de Mathématiques Pures et Appliquées*, 5e série, 5, 1899.

[265] G. Pick. Geometrisches zur Zahlenlehre. Sonderabdr. Naturw.-medizin. Verein f. Böhmen "Lotos" Nr. 8, 9 S. $8^\circ$., 1899.

[266] J. M. Pollard. Theorems on factorization and primality testing. *Proc. Cambridge Philos. Soc.*, 76:521–528, 1974.

[267] G. Pólya. Sur les séries entières, dont la somme est une fonction algébrique. *Enseignement Math.*, 22:38–47, 1921/1922.

[268] G. Pólya. On the number of certain lattice polygons. *J. Combinatorial Theory*, 6:102–105, 1969.

[269] L. R. Rabiner, R. W. Schafer, and C. M. Rader. The chirp $z$-transform algorithm and its application. *Bell System Tech. J.*, 48:1249–1292, 1969.

[270] J.-P. Ramis. About the growth of entire functions solutions of linear algebraic $q$-difference equations. *Ann. Fac. Sci. Toulouse Math. (6)*, 1(1):53–94, 1992.

[271] A. Riese. qMultiSum—a package for proving $q$-hypergeometric multiple summation identities. *J. Symbolic Comput.*, 35(3):349–376, 2003.

[272] T. Rivoal. La fonction zêta de Riemann prend une infinité de valeurs irrationnelles aux entiers impairs. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(4):267–270, 2000.

[273] T. Rivoal and J. Roques. Hadamard products of algebraic functions. *J. Number Theory*, 145:579–603, 2014.

[274] L. Rogers and S. Ramanujan. Proof of certain identities in combinatory analysis. *Proc. Cambridge Philos. Soc.*, 19:211–214, 1919.

[275] L. J. Rogers. Second Memoir on the Expansion of certain Infinite Products. *Proc. Lond. Math. Soc.*, 25:318–343, 1893/94.

[276] H. A. Rothe. Formulae de serierum reversione demonstratio universalis signis localibus combinatorico-analyticorum vicariis exhibita. Leipzig, 1793.

[277] E. Rowland and R. Yassawi. Automatic congruences for diagonals of rational functions. *Journal de Théorie des Nombres de Bordeaux*, 27(1):245–288, 2015.

[278] C. Sabbah. Systèmes holonomes d'équations aux $q$-différences. In *$D$-modules and microlocal geometry (Lisbon, 1990)*, pages 125–147. de Gruyter, 1993.

[279] B. E. Sagan. Congruence properties of $q$-analogs. *Adv. Math.*, 95(1):127–143, 1992.

[280] R. Schäfke and M. Singer. Consistent systems of linear differential and difference equations. *J. Eur. Math. Soc. (JEMS)*, 21(9):2751–2792, 2019.

[281] D. Schmidt. Construction of the Jordan decomposition by means of Newton's method. *Linear Algebra Appl.*, 314(1-3):75–89, 2000.

[282] W. M. Schmidt. Eisenstein's theorem on power series expansions of algebraic functions. *Acta Arith.*, 56(2):161–179, 1990.

[283] C.-P. Schnorr. Improved lower bounds on the number of multiplications / divisions which are necessary to evaluate polynomials. *Theoret. Comput. Sci.*, 7(3):251–261, 1978.

[284] P. Scholze. Canonical $q$-deformations in arithmetic geometry. *Ann. Fac. Sci. Toulouse Math. (6)*, 26(5):1163–1192, 2017.

[285] A. Schönhage. Schnelle Multiplikation von Polynomen über Körpern der Charakteristik 2. *Acta Informatica*, 7:395–398, 1977.

[286] A. Schönhage and V. Strassen. Schnelle Multiplikation grosser Zahlen. *Computing (Arch. Elektron. Rechnen)*, 7:281–292, 1971.

[287] D. Schreck. Prince Rupert's problem and its extension by Pieter Nieuwland. *Scripta Math.*, 16:73–80 and 261–267, 1950.

[288] H. A. Schwarz. Über diejenigen Fälle, in welchen die Gaußische hypergeometrische Reihe einer algebraische Funktion ihres vierten Elementes darstellt. *J. Reine Angew. Math.*, 75:292–335, 1873.

[289] J. Schygulla. Willmore minimizers with prescribed isoperimetric ratio. *Arch. Ration. Mech. Anal.*, 203(3):901–941, 2012.

[290] C. J. Scriba. Das Problem des Prinzen Ruprecht von der Pfalz. *Praxis Math.*, 10(9):241–246, 1968.

[291] U. Seifert. Configurations of fluid membranes and vesicles. *Advances in Physics*, 46(1):13–137, 1997.

[292] M. I. Shamos. *Computational Geometry*. PhD thesis, USA, 1978. AAI7819047.

[293] D. Shanks. A short proof of an identity of Euler. *Proc. AMS*, 2:747–749, 1951.

[294] V. Shoup. NTL: A library for doing number theory, v11.5.1, 2021. https://libntl.org.

[295] M. F. Singer. Algebraic solutions of $n$th order linear differential equations. In *Proceedings of the Queen's Number Theory Conference, 1979 (Kingston, Ont., 1979)*, volume 54 of *Queen's Papers in Pure and Appl. Math.*, pages 379–420. Queen's Univ., Kingston, Ont., 1980.

[296] M. F. Singer and F. Ulmer. Linear differential equations and products of linear forms. volume 117/118, pages 549–563. 1997. Algorithms for algebra (Eindhoven, 1996).

[297] N. J. A. Sloane and The OEIS Foundation Inc. The on-line encyclopedia of integer sequences. http://oeis.org/, 2020.

[298] T. Sprenger and W. Koepf. Algorithmic determination of $q$-power series for $q$-holonomic functions. *J. Symbolic Comput.*, 47(5):519–535, 2012.

[299] R. P. Stanley. Differentiably finite power series. *European J. Combin.*, 1(2):175–188, 1980.

[300] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[301] R. P. Stanley. *Catalan numbers*. Cambridge University Press, New York, 2015.

[302] J. Steininger and S. Yurkevich. Extended Abstract for: Solving Rupert's Problem Algorithmically. *ACM Commun. Comput. Algebra*, 56(2):32–35, nov 2022.

[303] J. Steininger and S. Yurkevich. An algorithmic approach to Rupert's problem. *Math. Comp.*, 92(342):1905–1929, 2023.

[304] M. Stoll. Bounds for the length of recurrence relations for convolutions of $P$-recursive sequences. *European J. Combin.*, 18(6):707–712, 1997.

[305] V. Strassen. Gaussian elimination is not optimal. *Numer. Math.*, 13:354–356, 1969.

[306] V. Strassen. Polynomials with rational coefficients which are hard to compute. *SIAM J. Comput.*, 3:128–149, 1974.

[307] V. Strassen. Einige Resultate über Berechnungskomplexität. *Jber. Deutsch. Math.-Verein.*, 78(1):1–8, 1976/77.

[308] A. Straub. Multivariate Apéry numbers and supercongruences of rational functions. *Algebra Number Theory*, 8(8):1985–2007, 2014.

[309] A. Straub. On congruence schemes for constant terms and their applications. *Research in Number Theory*, 8(3), May 2022.

[310] A. Straub and W. Zudilin. Sums of Powers of Binomials, Their Apéry Limits, and Franel's Suspicions. *International Mathematics Research Notices*, pages 1–19, 05 2022.

[311] M. Sun. A new bound on Hrushovski's algorithm for computing the Galois group of a linear differential equation. *Comm. Algebra*, 47(9):3553–3566, 2019.

[312] H. W. L. Tanner. On the Enumeration of Groups of Totitives. *Proc. Lond. Math. Soc.*, 27:329–352, 1895/96.

[313] T. Tao, E. Croot, III, and H. Helfgott. Deterministic methods to find primes. *Math. Comp.*, 81(278):1233–1246, 2012.

[314] P. Tonpho. Covering of objects related to Rupert property, 2018. Master Thesis, http://cuir.car.chula.ac.th/handle/123456789/73138.

[315] P. Tonpho and W. Wichiramala. Rupert property of some particular $n$-simplex and $n$-octahedrons, June 2022. Preprint.

[316] C. Tretkoff and M. Tretkoff. Solution of the inverse problem of differential Galois theory in the classical case. *Amer. J. Math.*, 101(6):1327–1332, 1979.

[317] W. J. Trjitzinsky. Analytic theory of linear $q$-difference equations. *Acta Math.*, 61(1):1–38, 1933.

[318] H. Tsai. Weyl closure of a linear differential operator. volume 29, pages 747–775. 2000. Symbolic computation in algebra, analysis, and geometry (Berkeley, CA, 1998).

[319] A. van der Poorten. A proof that Euler missed. . .Apéry's proof of the irrationality of $\zeta(3)$. *Math. Intelligencer*, 1(4):195–203, 1978/79. An informal report.

[320] M. van der Put and M. F. Singer. *Galois theory of linear differential equations*, volume 328 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2003.

[321] M. van Hoeij and J.-A. Weil. An algorithm for computing invariants of differential Galois groups. *J. Pure Appl. Algebra*, 117/118:353–379, 1997. Algorithms for algebra (Eindhoven, 1996).

[322] M. Waldschmidt. Transcendence of periods: the state of the art. *Pure Appl. Math. Q.*, 2(2, Special Issue: In honor of John H. Coates. Part 2):435–463, 2006.

[323] C. Wang. An analytic proof of the Borwein conjecture. *Adv. Math.*, 394:Paper No. 108028, 54, 2022.

[324] C. Wang and C. Krattenthaler. An asymptotic approach to Borwein-type sign pattern theorems, 2022. arXiv: 2201.12415.

[325] S. O. Warnaar and W. Zudilin. A $q$-rious positivity. *Aequationes Math.*, 81(1-2):177–183, 2011.

[326] S. O. Warnaar and W. Zudilin. $q$-rious and $q$-riouser, 2019. In Dick Askey's Liber Amicorum, Contribution 77. arXiv:1909.07045.

[327] H. S. Wilf and D. Zeilberger. An algorithmic proof theory for hypergeometric (ordinary $\&$ $q$) multisum/integral identities. *Invent. Math.*, 108(3):575–633, 1992.

[328] T. J. Willmore. Note on embedded surfaces. *An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat. (N.S.)*, 11B:493–496, 1965.

[329] S. Winograd. On multiplication of $2 \times 2$ matrices. *Linear Algebra Appl.*, 4:381–388, 1971.

[330] E. Witten. Two-dimensional gravity and intersection theory on moduli space. In *Surveys in differential geometry (Cambridge, MA, 1990)*, pages 243–310. Lehigh Univ., Bethlehem, PA, 1991.

[331] E. Witten. Algebraic geometry associated with matrix models of two-dimensional gravity. In *Topological methods in modern mathematics (Stony Brook, NY, 1991)*, pages 235–269. Publish or Perish, Houston, TX, 1993.

[332] K.-W. Yang. On the product $\prod_{n\geq 1}(1 + q^n x + q^{2n}x^2)$. *J. Austral. Math. Soc. Ser. A*, 48(1):148–151, 1990.

[333] M. Yip. Rook placements and Jordan forms of upper-triangular nilpotent matrices. *Electron. J. Combin.*, 25(1):Paper 1.68, 25, 2018.

[334] T. Yu and J. Chen. Uniqueness of Clifford torus with prescribed isoperimetric ratio. *Proc. Amer. Math. Soc.*, 150(4):1749–1765, 2022.

[335] S. Yurkevich. The art of algorithmic guessing in gfun. *Maple Transactions*, 2(1):14421:1–14421:19, 2022.

[336] D. Zagier. Elliptic modular forms and their applications. In *The 1-2-3 of modular forms*, Universitext, pages 1–103. Springer, 2008.

[337] D. Zagier. Integral solutions of Apéry-like recurrence equations. In *Groups and symmetries*, volume 47 of *CRM Proc. Lecture Notes*, pages 349–366. Amer. Math. Soc., Providence, RI, 2009.

[338] D. Zagier. The arithmetic and topology of differential equations. In *European Congress of Mathematics*, pages 717–776. Eur. Math. Soc., Zürich, 2018.

[339] U. Zannier. *Lecture notes on Diophantine analysis*, volume 8 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, 2009.

[340] D. Zeilberger. A holonomic systems approach to special functions identities. *J. Comput. Appl. Math.*, 32(3):321–368, 1990.

[341] D. Zeilberger. The method of creative telescoping. *J. Symbolic Comput.*, 11(3):195–204, 1991.