DE GRUYTER

Bruce M. Landman, Florian Luca, Melvyn B. Nathanson, Jaroslav Nešetřil and Aaron Robertson (Eds.)

NUMBER THEORY AND COMBINATORICS

A COLLECTION IN HONOR OF THE MATHEMATICS OF RONALD GRAHAM

PROCEEDINGS IN MATHEMATICS

Bruce M. Landman, Florian Luca, Melvyn B. Nathanson, Jaroslav Nešetřil, Aaron Robertson (Eds.)
Number Theory and Combinatorics

De Gruyter Proceedings in Mathematics

Number Theory and Combinatorics

A Collection in Honor of the Mathematics of Ronald Graham

Edited by
Bruce M. Landman, Florian Luca, Melvyn B. Nathanson, Jaroslav Nešetřil, and Aaron Robertson

DE GRUYTER

Mathematics Subject Classification 2020
11xxx, 05xxx

Editors

Bruce M. Landman
Augusta University
Department of Mathematics
1120 15th Street
Augusta
GA 30912
USA
blandman@augusta.edu

Florian Luca
University of Witwatersrand
School of Mathematics
1 Jan Smuts Avenue
Johannesburg 2000
Republic of South Africa
florian.luca@wits.ac.za

Melvyn B. Nathanson
The City University of New York
Lehman College (CUNY)
Department of Mathematics
250 Bedford Park Boulevard West
Bronx
NY 10468
USA
nathansn@alpha.lehman.cuny.edu

Jaroslav Nešetřil
Charles University
Computer Science Institute (IUUK)
Malostranske nam. 25
118 00 Praha
Czech Republic
nesetril@iuuk.mff.cuni.cz

Aaron Robertson
Colgate University
Department of Mathematics
219 McGregory Hall
Hamilton
NY 13346
USA
arobertson@colgate.edu

ISBN 978-3-11-075343-1
e-ISBN (PDF) 978-3-11-075421-6
e-ISBN (EPUB) 978-3-11-075426-1

Library of Congress Control Number: 2022930168

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Foreword

Ron Graham finished his PhD in combinatorial number theory in 1962 under the direction of Derrick Lehmer and then immediately proceeded to work at Bell Labs. In the summer of 1963, a professor from Purdue reached out to Ron to encourage him to take an assistant professorship and told him, "You'll be dead mathematically in a couple years if you stick to industry." Ron did not take that advice and stayed at Bell Labs for another 36 years before retiring (also holding visiting positions at Stanford, UCLA, Princeton, Rutgers, and others during this time); after this first retirement, he took a position at UC San Diego for another 20 years. Over the course of nearly six decades, Ron had one of the most amazing mathematical lives of the twentieth century. During his career, he produced over 400 publications, held positions on numerous editorial boards and committee assignments, gave countless talks around the world, was awarded multiple honors and degrees, and helped to firmly bring discrete mathematics to the prominence it experiences today.

Ron's mathematical interests were not easy to pigeonhole. He started off working in combinatorial number theory, including Egyptian fraction problems (which is how he first connected with Paul Erdős). At Bell Labs, he quickly expanded into a combination of the practical (which included scheduling, bin packing, vertex labelings, and Steiner tree problems), as well as the more esoteric[1] (including finite semigroups, geometrical packing problems, and Ramsey theory (which became a major focus of his work)). Many of his papers had a strong geometrical flavor including an efficient algorithm for finding a convex hull (the "Graham scan," which became one of his most cited works and helped to open up the field of computational geometry; though when Ron wrote the paper he considered it a "throw-away result"), the largest small hexagon (the hexagon of unit diameter with maximum area[2]—it is not the regular hexagon!), Apollonian circle packings, and many more. Ron also was able to mathematically explore "fun" topics that included card shuffling and magic,[3] guessing games, and juggling.[4]

The contributions in this volume reflect some of the diverse range of mathematical interests of Ron; he would have delighted to leaf through these papers, see their results, and talk about them with friends and colleagues. These contributions also speak of Ron's ability to reach out and touch so many lives. Among mathematicians, he had a nearly singular prowess in being able to connect with people, and then to

1 The fact that AT&T enjoyed a monopoly resulted in giving the Bell Labs employees some flexibility in their research.

2 Joel Spencer liked this result so much he built a sandbox using this shape for his children.

3 This is exemplified most prominently in the book *Magical Mathematics* written by Persi Diaconis and Ron Graham.

4 Ron at one point served as president of the International Jugglers Association and was also instrumental in the development of the juggling pattern Mill's Mess.

https://doi.org/10.1515/9783110754216-201

help push them to be better, to connect them with problems and opportunities, and to delight and amaze them. His zest for playful, yet focused learning, whether it be of mathematics, Chinese, table tennis, trampolining or *Dance, Dance, Revolution*, was infectious.

We were fortunate to have been inspired by the work and life of Ron Graham, and by his example of how transformative a mathematician can be.

Steve Butler and Glenn Hurlbert

Contents

Foreword —— V

Ayomikun Adeniran, Lauren Snider, and Catherine Yan
Multivariate difference Gončarov polynomials —— 1

J.-P. Allouche
On an inequality in a 1970 paper of R. L. Graham —— 21

Noga Alon, Ryan Alweiss, Yang P. Liu, Anders Martinsson, and Shyam Narayanan
Arithmetic progressions in sumsets of sparse sets —— 27

Michael A. Bennett, Greg Martin, and Kevin O'Bryant
Multidimensional Padé approximation of binomial functions: equalities —— 35

Lars Blomberg, S. R. Shannon, and N. J. A. Sloane
Graphical enumeration and stained glass windows, 1: rectangular grids —— 65

Tom C. Brown and Shahram Mohsenipour
Two extensions of Hilbert's cube lemma —— 99

Mark Budden
The Gallai–Ramsey number for a tree versus complete graphs —— 109

Joe Buhler, Chris Freiling, Ron Graham, Jonathan Kariv, James R. Roche, Mark Tiefenbruck, Clint Van Alten, and Dmytro Yeroshkin
On Levine's notorious hat puzzle —— 115

Joshua Cooper and Grant Fickes
Recurrence ranks and moment sequences —— 167

Andrzej Dudek, Jarosław Grytczuk, and Andrzej Ruciński
On weak twins and up-and-down subpermutations —— 187

Sohail Farhangi and Jarosław Grytczuk
Distance graphs and arithmetic progressions —— 203

Michael Filaseta and Jacob Juillerat
Consecutive primes which are widely digitally delicate —— 209

Jerrold R. Griggs
Spanning trees and domination in hypercubes —— 249

Heiko Harborth and Hauke Nienborg
Rook domination on hexagonal hexagon boards —— 259

Neil Hindman and Dona Strauss
Strongly image partition regular matrices —— 267

Brian Hopkins
Introducing shift-constrained Rado numbers —— 285

Jared Duker Lichtman
Mertens' prime product formula, dissected —— 297

Melvyn B. Nathanson
Curious convergent series of integers with missing digits —— 311

Carl Pomerance
A note on Carmichael numbers in residue classes —— 321

I. D. Shkredov and J. Solymosi
Tilted corners in integer grids —— 329

Noga Alon, Tom C. Brown, Steve Butler, Jerrold R. Griggs, Neil Hindman, Veselin Jungic, Bruce M. Landman, and Jaroslav Nešetřil
Remembrances —— 339

Steve Butler
A selected bibliography of Ron Graham —— 355

Ayomikun Adeniran, Lauren Snider, and Catherine Yan

Multivariate difference Gončarov polynomials

In Memory of Ron Graham

Abstract: Univariate delta Gončarov polynomials arise when the classical Gončarov interpolation problem in numerical analysis is modified by replacing derivatives with delta operators. When the delta operator under consideration is the backward difference operator, we acquire the univariate difference Gončarov polynomials, which have a combinatorial relation to lattice paths in the plane with a given right boundary. In this paper, we extend several algebraic and analytic properties of univariate difference Gončarov polynomials to the multivariate case. We then establish a combinatorial interpretation of multivariate difference Gončarov polynomials in terms of certain constraints on d-tuples of nondecreasing integer sequences. This motivates a connection between multivariate difference Gončarov polynomials and a higher-dimensional generalized parking function, the $\boldsymbol{U}$-parking function, from which we derive several enumerative results based on the theory of multivariate delta Gončarov polynomials.

1 Introduction

The primary goal of this paper is to extend results on univariate difference Gončarov polynomials to multiple variables, as well as show that such polynomials have a combinatorial interpretation related to integer sequences and generalized parking functions.

Central to these problems is the theory of Gončarov polynomials, which arose in the fields of numerical analysis and approximation theory from an interpolation problem posed by Gončarov [6].

Gončarov interpolation. Find a degree n polynomial $p(x)$ such that for $i = 0, 1, \ldots$, the ith derivative $p^{(i)}(x)$ evaluated at a given point a_i has a prescribed value b_i.

The solution to this interpolation problem consists of a linear combination of the Gončarov polynomials $\{g_n(x; a_0, a_1, \ldots, a_{n-1})\}_{n\in\mathbb{N}}$, where $g_n(x; a_0, a_1, \ldots, a_{n-1})$ is the

Acknowledgement: The third author is supported in part by Simons Collaboration Grant for Mathematics 704276.

Ayomikun Adeniran, Department of Mathematics, Pomona College, Claremont, CA, USA, e-mail: ayomikun.adeniran@pomona.edu
Lauren Snider, Catherine Yan, Department of Mathematics, Texas A&M University, College Station, TX, USA, e-mails: lsnider@math.tamu.edu, cyan@math.tamu.edu

https://doi.org/10.1515/9783110754216-001

unique polynomial of degree n satisfying the biorthogonality condition

$$g_n^{(i)}(a_i; a_0, a_1, \ldots, a_{n-1}) = n!\delta_{in}.$$

The Gončarov polynomials have been extensively studied for their analytical properties, but their remarkable application to parking functions, which have a vast literature in combinatorics, was surely an unforeseen consequence by Gončarov. A (classical) parking function is a sequence $(x_1, x_2, \ldots, x_n)$ of nonnegative integers whose nondecreasing rearrangement $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ satisfies $x_{(i)} < i$ for all i. The sequence $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ is referred to as the order statistics of $(x_1, x_2, \ldots, x_n)$. More generally, given a vector $\boldsymbol{u} = (u_1, \ldots, u_n)$, a $\boldsymbol{u}$-parking function is a sequence $(x_1, x_2, \ldots, x_n)$ of nonnegative integers whose order statistics satisfy $x_{(i)} < u_i$. Kung and Yan [10] showed that Gončarov polynomials are in direct correspondence with $\boldsymbol{u}$-parking functions, and hence the numerous algebraic and analytic properties of the former extend necessarily to the latter. Classical parking functions correspond to the case $\boldsymbol{u} = (1, 2, \ldots, n)$.

When the Gončarov interpolation problem is extended to multiple variables with partial derivatives $\partial_{x_1}, \ldots, \partial_{x_d}$, a basis of the solutions is the set of multivariate Gončarov polynomials. Khare, Lorentz, and Yan provide a thorough treatment of bivariate Gončarov polynomials in [7], establishing numerous properties analogous to those of the univariate case, and showing that a bivariate Gončarov polynomial counts pairs of integer sequences whose order statistics satisfy certain constraints. This work naturally extends to d-dimensions and leads to a notion of higher-dimensional generalized parking functions, namely, the $\boldsymbol{U}$-parking functions, where $\boldsymbol{U}$ is a set of nodes in $\mathbb{N}^d$.

Another profound generalization of Gončarov polynomials is obtained by applying the rich theory of delta operators and finite operator calculus, which is a unified theory on linear operators analogous to the differentiation operator D and special polynomials developed by Rota, Kahaner, and Odlyzko [16]. Replacing D with an arbitrary delta operator in the Gončarov interpolation problem, Lorentz, Tringali, and Yan [11, 12] introduced the delta Gončarov polynomials and extended many of the algebraic properties of (classical) Gončarov polynomials to this generalized case. They also studied multivariate delta Gončarov polynomials and characterized those that are of binomial type. A complete combinatorial interpretation for univariate delta Gončarov polynomials was given by Adeniran and Yan [1] in terms of weighted enumerators in partition lattices and exponential families.

Of particular interest to us are the difference Gončarov polynomials, which are closely related to lattice paths and integer sequences. Here, the delta operator is the backward difference operator Δ. In [9], the algebraic and combinatorial properties of the univariate difference Gončarov polynomials are presented. In the current paper, we seek to extend these properties to the multivariate case and investigate their combinatorial significance. The remainder of the paper is organized as follows. Section 2

recalls the basic definition and properties of univariate difference Gončarov polynomials. In Section 3, we specifically examine bivariate difference Gončarov polynomials and extend the algebraic and analytic properties of their univariate analogues to two variables. Section 4 characterizes the relationship between bivariate difference Gončarov polynomials and integer sequences. Finally, in Section 5 we state the corresponding results in higher dimensions.

2 Univariate difference Gončarov polynomials

We begin by briefly summarizing the theory of delta Gončarov polynomials with a focus on univariate difference Gončarov polynomials. The detailed theory on delta operators is developed by Mullin and Rota in [14], and the theory of delta Gončarov polynomials is introduced in [12].

Consider the vector space $\mathbb{F}[x]$ of all polynomials in the variable x over a field $\mathbb{F}$ of characteristic zero. For $a \in \mathbb{F}$, let $E_a : \mathbb{F}[x] \to \mathbb{F}[x]$ be the shift operator defined by $E_a(f)(x) = f(x+a)$, and let $\varepsilon(a) : \mathbb{F}[x] \to \mathbb{F}$ be the linear functional that evaluates $p(x) \in \mathbb{F}[x]$ at $a \in \mathbb{F}$. A delta operator is a linear operator $\mathfrak{d} : \mathbb{F}[x] \to \mathbb{F}[x]$ that is shift-invariant, i. e., $\mathfrak{d}E_a = E_a\mathfrak{d}$ for all $a \in \mathbb{F}$, and satisfies $\mathfrak{d}(x) = c$ for some nonzero constant c. The differentiation operator D is one example of a delta operator. Another example is the *backward difference operator* $\Delta = I - E_{-1}$, which is defined by $\Delta p(x) = p(x) - p(x-1)$.

Every delta operator $\mathfrak{d}$ has a unique polynomial sequence $(p_n(x))_{n\in\mathbb{N}}$ such that $p_n(x)$ is of degree n, $p_n(0) = \delta_{0n}$, and $\mathfrak{d}p_n(x) = np_{n-1}(x)$. Such a sequence is called the basic sequence associated to $\mathfrak{d}$. Moreover, any shift-invariant operator T can be expanded as a formal power series of $\mathfrak{d}$ by the formula

$$T = \sum_{k\geq 0} \frac{a_k}{k!}\mathfrak{d}^k,$$

where $a_k = \varepsilon_0(T(p_k(x)))$.

For a delta operator $\mathfrak{d}$, suppose that $(\psi_s(\mathfrak{d}))_{s\in\mathbb{N}}$ is a sequence of linear operators of the form

$$\psi_s(\mathfrak{d}) = \mathfrak{d}^s \sum_{r=0}^{\infty} b_{s,r}\mathfrak{d}^r,$$

where $b_{s,r} \in \mathbb{F}$ and $b_{s,0} \neq 0$. Then there exists a unique sequence of polynomials $(f_n(x))_{n\in\mathbb{N}}$ in $\mathbb{F}[x]$ such that each $f_n(x)$ has degree n and satisfies

$$\varepsilon(0)\psi_s(\mathfrak{d})f_n(x) = n!\delta_{sn} \quad \text{for all } s \in \mathbb{N},$$

where δ_{sn} is the Kronecker delta. In this case, we say that the polynomial sequence $(f_n(x))_{n\in\mathbb{N}}$ is *biorthogonal* to the sequence of linear operators $(\psi_s(\mathfrak{d}))_{s\in\mathbb{N}}$. In fact, the polynomials $(f_n(x))_{n\in\mathbb{N}}$ form a basis of $\mathbb{F}[x]$.

Let the delta operator $\mathfrak{d}$ be the backward difference operator Δ. Then the sequence of upper factorial functions $(x^{(n)})_{n\in\mathbb{N}}$ defined by $x^{(0)} = 1$ and $x^{(n)} = x(x+1)\cdots(x+n-1)$ for $n \geq 1$ is the basic sequence associated to Δ. Given a sequence $a_0, a_1, \ldots$ of *nodes* in $\mathbb{F}$, let $(\psi_s(\Delta))_{s=0}^{\infty}$ be the sequence of linear operators given by the equation

$$\psi_s(\Delta) = \Delta_s \sum_{r=0}^{\infty} \frac{a_s^{(r)}}{r!}\Delta^r = E_{a_s}\Delta^s.$$

The sequence of *difference Gončarov polynomials* is the unique sequence of polynomials biorthogonal to $(\psi_s(\Delta))_{s=0}^{\infty}$. That is, the nth difference Gončarov polynomial $\tilde{g}_n(x; a_0, a_1, \ldots, a_{n-1})$ is the unique polynomial of degree n satisfying

$$\varepsilon(a_s)\Delta^s g_n(x; a_0, \ldots, a_{n-1}) = n!\delta_{s,n}, \quad \text{for all } s \in \mathbb{N}.$$

It is the difference analog of the classical univariate Gončarov polynomial, which has been comprehensively studied in interpolation theory and approximation theory [3].

The notation for the nth difference Gončarov polynomial $\tilde{g}(x; a_0, a_1, \ldots, a_{n-1})$ reflects its dependence on only the nodes $a_0, a_1, \ldots, a_{n-1}$. The preprint [9] contains a set of algebraic and analytic properties for $\tilde{g}_n(x; a_0, a_1, \ldots, a_{n-1})$. Since [9] has never been published, we include those results here for completeness.

1. (*Determinant formula*) For any $n \in \mathbb{N}$, $\tilde{g}_n(x; a_0, a_1, \ldots, a_{n-1}) = n!\det M$ where M is an $(n+1)\times(n+1)$ matrix whose (i,j)-entry, $0 \leq i, j \leq n$, is given by

$$m_{i,j} = \begin{cases} \frac{a_i^{(j-i)}}{(j-i)!}, & \text{if } 0 \leq i \leq j \text{ and } i \leq n-1 \\ \frac{x^{(j)}}{j!} & \text{if } i = n \\ 0 & \text{otherwise.} \end{cases}$$

2. (*Expansion formula*) For $p(x) \in \mathbb{F}[x]$ of degree n,

$$p(x) = \sum_{i=0}^{n} \frac{\varepsilon(0)\psi_i(\Delta)p(x)}{i!}\tilde{g}_i(x; a_0, a_1, \ldots, a_{i-1}).$$

3. (*Linear recursion*)

$$x^{(n)} = \sum_{i=0}^{n} \binom{n}{i} a_i^{(n-i)}\tilde{g}_i(x; a_0, a_1, \ldots, a_{i-1}).$$

4. (*Appell relation*)

$$(1-t)^{-x} = \sum_{n=0}^{\infty} \tilde{g}_n(x; a_0, a_1, \dots, a_{n-1}) \frac{t^n}{n!(1-t)^{a_n}}.$$

5. (*Difference relation*) For any $n \in \mathbb{N}$,

$$\Delta \tilde{g}_n(x; a_0, a_1, \dots, a_{n-1}) = n! \tilde{g}_{n-1}(x; a_1, a_2, \dots, a_{n-1})$$

and

$$\tilde{g}_n(a_0; a_0, a_1, \dots, a_{n-1}) = \delta_{0n},$$

which together uniquely determine the sequence of difference Gončarov polynomials.

6. (*Shift-invariant formula*)

$$\tilde{g}_n(x + t; a_0 + t, a_1 + t, \dots, a_{n-1} + t) = \tilde{g}_n(x; a_0, a_1, \dots, a_{n-1}).$$

7. (*Perturbation formula*) For positive integers m and n with $m < n$,

$$\begin{aligned} &\tilde{g}_n(x; a_0, \dots, a_{m-1}, a_m + \delta_m, a_{m+1}, \dots, a_{n-1}) \\ &\quad = \tilde{g}_n(x; a_0, \dots, a_{m-1}, a_m, a_{m+1}, \dots, a_{n-1}) \\ &\qquad - \binom{n}{m} \tilde{g}_{n-m}(a_m + \delta_m; a_m, a_{m+1}, \dots, a_{n-1}) \tilde{g}_m(x; a_0, a_1, \dots, a_{m-1}). \end{aligned}$$

8. (*Sheffer relation*)

$$\tilde{g}_n(x + y; a_0, a_1, \dots, a_{n-1}) = \sum_{i=0}^{n} \binom{n}{i} \tilde{g}_{n-i}(y; a_i, \dots, a_{n-1}) x^{(i)}.$$

In particular, letting $y = 0$ we obtain the expansion of $\tilde{g}_n(x; a_0, a_1, \dots, a_{n-1})$ under the basis $(x^{(n)})_{n \in \mathbb{N}}$.

Difference Gončarov polynomials are useful in combinatorics due to their connection with lattice paths in the plane with a given right boundary. Let x, n be positive integers. A lattice path in $\mathbb{Z}^2$ from $(0,0)$ to $(x-1, n)$ with steps $(1,0)$ and $(0,1)$ can be recorded by a nondecreasing integer sequence $(x_0, x_1, \dots, x_{n-1})$, where (x_i, i) is the coordinate of the rightmost point on the lattice path and the line $y = i$. Given $a_0 \le a_1 \le \cdots \le a_{n-1} \in [0, x]^n$, let $LP_n(a_0, a_1, \dots, a_{n-1})$ be the number of lattice paths $(x_0, x_1, \dots, x_{n-1})$ from $(0,0)$ to $(x-1, n)$ such that $0 \le x_i < a_i$ for $0 \le i \le n$. Then we have the following theorem.

Theorem 1 ([9]).

$$LP_n(a_0, a_1, \ldots, a_{n-1}) = \frac{1}{n!}\tilde{g}_n(x; x-a_0, x-a_1, \ldots, x-a_{n-1}) = \frac{1}{n!}\tilde{g}_n(0; -a_0, -a_1, \ldots, -a_{n-1}).$$

When $a_i = a$ for all i, $\tilde{g}_n(x; a, \ldots, a) = (x-a)^{(n)}$. Hence $LP_n(a, \ldots, a) = \frac{a^{(n)}}{n!} = \binom{a+n-1}{n}$, which is clearly the number of lattice paths from $(0,0)$ to $(a-1, n)$. When $a_i = a+(i-1)b$, $\tilde{g}_n(x; a, a+b, \ldots, a+(n-1)b) = (x-a)(x-a-nb+1)^{(n-1)}$ for $n > 0$. In particular, for $a = b = 1$, $\tilde{g}_n(0; -a_0, -a_1, \ldots, -a_{n-1})$ is the Catalan number $\frac{1}{n+1}\binom{2n}{n}$; when $a = 1$, $b \in \mathbb{N}$, we get the Fuss–Catalan number $\frac{1}{bn+1}\binom{(b+1)n}{n}$. For general values of a_i's, $\tilde{g}_n(0; a_0, \ldots, a_{n-1})$ can be computed by the determinant formula or the linear recursion.

Lattice paths are a classical subject of study in combinatorics, having a vast literature with applications in many fields of mathematics, computer science, physics, and statistics. For the combinatorial theory of lattice paths, see the monograph [13] by Mohanty and the more recent comprehensive survey [8] by Krattenthaler. In addition to being a basic but useful tool in lattice path counting, difference Gončarov polynomials provide a new perspective to lattice paths and connect them to other combinatorial structures that are associated with general delta operators. The most notable examples are various generalization of parking functions, which are the combinatorial structures associated with the differential operator. In fact, difference Gončarov polynomials have already appeared in enumerating parking distributions over a caterpillar graph [4], and in enumerating increasing parking sequences [2].

3 Bivariate difference Gončarov polynomials

By replacing the difference operator Δ with a set of difference operators $\{\Delta_{x_i}\}_{i=1}^{d}$, where d is a positive integer, we can define a system of multivariate biorthogonal polynomials in $\mathbb{F}[x_1, \ldots, x_d]$ that naturally extend the univariate difference Gončarov polynomials to multiple variables. A general theory of systems of delta operators and delta Gončarov polynomials in multivariables was introduced in [11]. In this paper, we only need a special case: the system of delta operators is $(\Delta_{x_1}, \Delta_{x_2}, \ldots, \Delta_{x_d})$, where Δ_{x_i} is the backward difference operator with respect to the variable x_i. We will first state the definition and the basic properties from the general theory established in [11]. Then we present some special algebraic properties of multivariate difference Gončarov polynomials. In the next section, we discuss the combinatorial significance of such multivariate polynomials.

For simplicity and clarity, in Sections 3 and 4 we restrict our attention to the bivariate case. All the results can be extended easily to the multivariate cases, which we describe briefly in Section 5.

Fix positive integers m and n. We write $(i,j) \preceq (m,n)$ if $i \leq m$ and $j \leq n$. Let $S_{m,n}$ denote the poset $\{(i,j) : (0,0) \preceq (i,j) \preceq (m,n)\}$ and denote the space of all bivariate polynomials having coordinate degree (m,n) by $\Pi^2_{m,n}$. That is,

$$\Pi^2_{m,n} = \left\{ \sum_{(i,j)\in S_{m,n}} b_{i,j} x^i y^j : b_{i,j} \in \mathbb{F} \right\}.$$

The following is a bivariate variation of the Gončarov interpolation problem, with difference operators replacing differential operators.

Bivariate Gončarov interpolation with difference operators
Fix a node-set $\boldsymbol{Z} = \{z_{i,j} = (x_{i,j}, y_{i,j}) : (i,j) \in S_{m,n}\}$. Given a set of numbers $\{b_{i,j} \in \mathbb{F} : (i,j) \in S_{m,n}\}$, find a polynomial $p(x,y) \in \Pi^2_{m,n}$ such that, for all $(i,j) \in S_{m,n}$,

$$\varepsilon(z_{i,j})\Delta_x^i \Delta_y^j p(x,y) = b_{i,j}.$$

From the general theory developed in [11], we have that for any values $\{b_{i,j} : (i,j) \in S_{m,n}\}$, the bivariate Gončarov interpolation problem with difference operators has a unique solution in the space $\Pi^2_{m,n}$. In particular, by taking all but one of $\{b_{i,j} : (i,j) \in S_{m,n}\}$ to be 0, we can define the bivariate difference Gončarov polynomials.

Definition 1. Let $\boldsymbol{Z} = \{z_{i,j} = (x_{i,j}, y_{i,j}) : (i,j) \in S_{m,n}\}$ be a set of nodes. The *bivariate difference Gončarov polynomial* $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ is the unique polynomial in $\Pi^2_{m,n}$ satisfying

$$\varepsilon(z_{i,j})\Delta_x^i \Delta_y^j \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = m!n!\delta_{m,i}\delta_{n,j} \tag{1.1}$$

for all $(i,j) \in S_{m,n}$.

It follows that $\tilde{g}_{0,0}((x,y);\boldsymbol{Z}) = 1$. For the special grid $\boldsymbol{O} = \{z_{i,j} = (0,0) : (i,j) \in S_{m,n}\}$, the set of difference Gončarov polynomials $\{\tilde{g}_{m,n}((x,y);\boldsymbol{O})\}_{m,n\in\mathbb{N}}$ is called the basic sequence of the system (Δ_x, Δ_y). From the interpolation conditions given in equation (1.1), it is easy to check that

$$\tilde{g}_{m,n}\big((x,y);\boldsymbol{O}\big) = x^{(m)} y^{(n)}.$$

In general, the set $\{\tilde{g}_{i,j}((x,y);\boldsymbol{Z}) : (i,j) \in S_{m,n}\}$ forms a basis to the solutions of the bivariate Gončarov interpolation problem with difference operators. Next, we discuss the algebraic properties of bivariate difference Gončarov polynomials, analogous to those of the univariate case. We remark that Theorems 2, 3, and 9 are special cases of Propositions 3.5, 3.6, and Theorem 5.1 in [11], and the other results are new.

Theorem 2 (Expansion formula). *For any* $p(x,y) \in \Pi^2_{m,n}$,

$$p(x,y) = \sum_{i=0}^{m}\sum_{j=0}^{n} \frac{1}{i!j!}[\varepsilon(z_{i,j})\Delta_x^i\Delta_y^j p(x,y)]\tilde{g}_{i,j}((x,y);\boldsymbol{Z}).$$

Proof. This property follows immediately from the definition of bivariate difference Gončarov polynomials and the fact that $\{\tilde{g}_{i,j}((x,y);\boldsymbol{Z})\}_{(i,j)\preceq(m,n)}$ forms a basis of $\Pi^2_{m,n}$. □

Theorem 3 (Linear recursion).

$$x^{(m)}y^{(n)} = \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}x_{i,j}^{(m-i)}y_{i,j}^{(n-j)}\tilde{g}_{i,j}((x,y);\boldsymbol{Z}). \tag{1.2}$$

Proof. It is obtained by letting $p(x,y) = x^{(m)}y^{(n)}$ in the expansion formula. □

Theorem 4 (Appell relation).

$$(1-s)^{-x}(1-t)^{-y} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\tilde{g}_{m,n}((x,y);\boldsymbol{Z})\frac{s^m}{(1-s)^{x_{m,n}}m!}\frac{t^n}{(1-t)^{y_{m,n}}n!}.$$

Proof. Using Taylor expansion and the linear recursion formula, we have

$$\begin{aligned}
\frac{1}{(1-s)^x(1-t)^y} &= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{x^{(m)}s^m}{m!}\frac{y^{(n)}t^n}{n!}\\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{s^m t^n}{m!n!}\sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}x_{i,j}^{(m-i)}y_{i,j}^{(n-j)}\tilde{g}_{i,j}((x,y);\boldsymbol{Z})\\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\tilde{g}_{i,j}((x,y);\boldsymbol{Z})\sum_{m=i}^{\infty}\frac{1}{m!}\binom{m}{i}x_{i,j}^{(m-i)}s^m\sum_{n=j}^{\infty}\frac{1}{n!}\binom{n}{j}y_{i,j}^{(n-j)}t^n\\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\tilde{g}_{i,j}((x,y);\boldsymbol{Z})\frac{s^i}{i!}\frac{t^j}{j!}\sum_{m=i}^{\infty}\frac{x_{i,j}^{(m-i)}s^{m-i}}{(m-i)!}\sum_{n=j}^{\infty}\frac{y_{i,j}^{(n-j)}t^{n-j}}{(n-j)!}\\
&= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\tilde{g}_{i,j}((x,y);\boldsymbol{Z})\frac{s^i}{i!}\frac{t^j}{j!}\frac{1}{(1-s)^{x_{i,j}}}\frac{1}{(1-t)^{y_{i,j}}}.
\end{aligned}$$

The last step uses the identity $\frac{1}{(1-x)^n} = \sum_{k\geq 0}\binom{n+k-1}{k}x^k = \sum_{k\geq 0}\frac{n^{(k)}x^k}{k!}$. □

The following two formulas are analogues of the differential and integral relations of the classical bivariate Gončarov polynomials studied in [7]. For a node-set $\boldsymbol{Z} = \{z_{i,j} : i,j \in \mathbb{N}\}$, let $\boldsymbol{LZ} = \{w_{i,j} : w_{i,j} = z_{i+1,j}, i,j \in \mathbb{N}\}$ and $\boldsymbol{DZ} = \{w_{i,j} : w_{i,j} = z_{i,j+1}, i,j \in \mathbb{N}\}$. From here on, we will assume $\boldsymbol{Z}$ is an infinite grid with indices $i,j \in \mathbb{N}$, and $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ is determined by the subset $\{z_{i,j} \in \boldsymbol{Z} : (i,j) \in S_{m,n}\}$.

Theorem 5 (Difference relations). *For any $m, n \in \mathbb{N}$,*

$$\Delta_x \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = m\tilde{g}_{m-1,n}\big((x,y);\boldsymbol{LZ}\big),$$
$$\Delta_y \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = n\tilde{g}_{m,n-1}\big((x,y);\boldsymbol{DZ}\big).$$

Proof. We will only prove the first relation, as the second follows by symmetry. We wish to show that $\Delta_x \tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ and $m\tilde{g}_{m-1,n}((x,y);\boldsymbol{LZ})$ satisfy the same biorthogonality conditions. Now the definition of $\tilde{g}_{m,n}((x,y);\boldsymbol{LZ})$ implies that

$$\varepsilon(z_{i+1,j})\Delta_x^i\Delta_y^j\big[\Delta_x \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big)\big] = \varepsilon(z_{i+1,j})\Delta_x^{i+1}\Delta_y^j \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = 0$$

when $(i,j) \preceq (m-1,n)$ with $(i,j) \neq (m-1,n)$. When $(i,j) = (m-1,n)$,

$$\varepsilon(z_{i+1,j})\Delta_x^{i+1}\Delta_y^j \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = \varepsilon(z_{m,n})\Delta_x^m\Delta_y^n \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = m!n!.$$

Since $m\tilde{g}_{m-1,n}((x,y);\boldsymbol{LZ})$ satisfies these same conditions, uniqueness of the interpolation yields the first difference relation. □

Corollary 1. *The general difference formula is*

$$\Delta_x^i\Delta_y^j \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = (m)_i(n)_j \tilde{g}_{m-i,n-j}\big((x,y);\boldsymbol{L}^i\boldsymbol{D}^j\boldsymbol{Z}\big),$$

where $(t)_k = t(t-1)\cdots(t-k+1)$ is the kth lower factorial of t.

Theorem 6 (Shift-invariant formula). *Given a node-set $\boldsymbol{Z} = \{z_{i,j} = (x_{i,j}, y_{i,j}) : i,j \in \mathbb{N}\}$, let $\boldsymbol{Z} + (\xi,\eta)$ denote the set $\{(x_{i,j}+\xi, y_{i,j}+\eta) : i,j \in \mathbb{N}\}$. Then we have*

$$\tilde{g}_{m,n}\big((x+\xi, y+\eta);\boldsymbol{Z} + (\xi,\eta)\big) = \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big).$$

Proof. By definition, $\tilde{g}_{m,n}((x,y);\boldsymbol{Z}+(\xi,\eta))$ is the unique polynomial in $\Pi_{m,n}^2$ satisfying interpolation conditions

$$\varepsilon(z_{i,j})E_x^\xi E_y^\eta \Delta_x^i\Delta_y^j \tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}+(\xi,\eta)\big) = m!n!\delta_{im}\delta_{jn}$$

for $(i,j) \preceq (m,n)$, where E_x^a and E_y^b are the shift operators $(E_x^a f)(x,y) = f(x+a,y)$ and $(E_y^b f)(x,y) = f(x,y+b)$, respectively. Since these shift operators commute with the difference operators Δ_x and Δ_y, we may equivalently express the interpolation conditions as

$$\varepsilon(z_{i,j})\Delta_x^i\Delta_y^j \tilde{g}_{m,n}\big((x+\xi, y+\eta);\boldsymbol{Z}+(\xi,\eta)\big) = m!n!\delta_{im}\delta_{jn},$$

which are the precise conditions satisfied by $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$. □

Theorem 7 (Perturbation formula). *Given a set of nodes $\boldsymbol{Z}$, suppose we perturb the (i_0, j_0)-th node of $\boldsymbol{Z}$ to $z^*_{i_0,j_0}$. Let $\boldsymbol{Z}^*$ be the new set of nodes. Then for $(i_0, j_0) \preceq (m, n)$ but $(i_0, j_0) \neq (m, n)$, we have*

$$\begin{aligned}\tilde{g}_{m,n}((x,y);\boldsymbol{Z}^*) = \tilde{g}_{m,n}((x,y);\boldsymbol{Z}) \\ - \binom{m}{i_0}\binom{n}{j_0}\tilde{g}_{m-i_0,n-j_0}(z^*_{i_0,j_0};\boldsymbol{L}^{i_0}\boldsymbol{D}^{j_0}\boldsymbol{Z})\tilde{g}_{i_0,j_0}((x,y);\boldsymbol{Z}).\end{aligned}$$

Proof. Let

$$h_{m,n}(x,y) = \tilde{g}_{m,n}((x,y);\boldsymbol{Z}^*) + \frac{1}{i_0!j_0!}[\varepsilon(z^*_{i_0,j_0})\Delta_x^{i_0}\Delta_y^{j_0}\tilde{g}_{m,n}((x,y);\boldsymbol{Z})]\tilde{g}_{i_0,j_0}((x,y);\boldsymbol{Z}).$$

One can easily check that $h_{m,n}(x,y)$ and $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ satisfy the same interpolation conditions and so are equal by uniqueness. Using the difference relations, we may rewrite $h_{m,n}(x,y)$ as

$$\begin{aligned}&\tilde{g}_{m,n}((x,y);\boldsymbol{Z}^*) + \binom{m}{i_0}\left[\varepsilon(z^*_{i_0,j_0})\frac{1}{j_0!}\Delta_y^{j_0}\tilde{g}_{m-i_0,n}((x,y);\boldsymbol{L}^{i_0}\boldsymbol{Z})\right]\tilde{g}_{i_0,j_0}((x,y);\boldsymbol{Z}) \\ &= \tilde{g}_{m,n}((x,y);\boldsymbol{Z}^*) + \binom{m}{i_0}\binom{n}{j_0}\tilde{g}_{m-i_0,n-j_0}(z^*_{i_0,j_0};\boldsymbol{L}^{i_0}\boldsymbol{D}^{j_0}\boldsymbol{Z})\tilde{g}_{i_0,j_0}((x,y);\boldsymbol{Z}),\end{aligned}$$

and the statement is proved. □

Remark. In Theorem 7 if $(i_0, j_0) = (m, n)$, then $\tilde{g}_{m,n}((x,y);\boldsymbol{Z}^*) = \tilde{g}_{m,n}((x,y);\boldsymbol{Z})$. This is because the difference operators are degree-reducing, in the sense that for a polynomial $p(x,y)$ of coordinate degree (a,b), $\Delta_x p(x,y)$ is of degree $(a-1,b)$ and $\Delta_y p(x,y)$ is of degree $(a,b-1)$. Hence $\Delta_x^m\Delta_y^n\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ is always a constant, which must be equal to $m!n!$ by the interpolation conditions. It also implies that the formula of $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ does not depend on the node $z_{m,n}$.

Theorem 8 (Sheffer relation). *For any nonnegative integers m, n, we have*

$$\tilde{g}_{m,n}((x+b,y+c);\boldsymbol{Z}) = \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}\tilde{g}_{m-i,n-j}((b,c);\boldsymbol{L}^i\boldsymbol{D}^j\boldsymbol{Z})x^{(i)}y^{(j)}.$$

Proof. Expanding the polynomial $\tilde{g}_{m,n}((x+b,y+c);\boldsymbol{Z})$ under the basis $\{\tilde{g}_{i,j}((x,y);\boldsymbol{O})\}_{(i,j)\preceq(m,n)}$ by Theorem 2 and noting that $\tilde{g}_{i,j}((x,y);\boldsymbol{O}) = x^{(i)}y^{(j)}$, we have

$$\begin{aligned}&\tilde{g}_{m,n}((x+b,y+c);\boldsymbol{Z}) \\ &= \sum_{i=0}^{m}\sum_{j=0}^{n}\frac{1}{i!j!}[\varepsilon(0)\Delta_x^i\Delta_y^j\tilde{g}_{m,n}((x+b,y+c);\boldsymbol{Z})]x^{(i)}y^{(j)}\end{aligned}$$

$$= \sum_{i=0}^{m}\sum_{j=0}^{n} \frac{1}{i!j!}[\varepsilon(0)(m)_i(n)_j \tilde{g}_{m-i,n-j}((x+b,y+c);\boldsymbol{L}^i\boldsymbol{D}^j\boldsymbol{Z})]x^{(i)}y^{(j)}$$
$$= \sum_{i=0}^{m}\sum_{j=0}^{n} \binom{m}{i}\binom{n}{j}\tilde{g}_{m-i,n-j}((b,c);\boldsymbol{L}^i\boldsymbol{D}^j\boldsymbol{Z})x^{(i)}y^{(j)}. \qquad \square$$

The following observations give a relation between the univariate and bivariate difference Gončarov polynomials. Both can be checked easily using Definition 1.

1. When $m = 0$ or $n = 0$, we have

$$\tilde{g}_{m,0}((x,y);\boldsymbol{Z}) = \tilde{g}_m(x;x_{0,0},x_{1,0},\dots,x_{m-1,0}),$$
$$\tilde{g}_{0,n}((x,y);\boldsymbol{Z}) = \tilde{g}_n(y;y_{0,0},y_{0,1},\dots,y_{0,n-1}).$$

2. If there exist some sequences $\{\alpha_i\}$ and $\{\beta_j\}$ such that $z_{i,j} = (x_{i,j},y_{i,j}) = (\alpha_i,\beta_j)$, then

$$\tilde{g}_{m,n}((x,y);\boldsymbol{Z}) = \tilde{g}_m(x;\alpha_0,\dots,\alpha_{m-1})\tilde{g}_n(y;\beta_0,\dots,\beta_{n-1})$$

is the product of univariate difference Gončarov polynomials.

In general, the closed formula of $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ is quite involved. However, a special case in which we have an elegant closed formula of $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ occurs when the node $z_{i,j}$ is a linear transformation of (i,j). The resulting polynomials $\{\tilde{g}_{m,n}((x,y);\boldsymbol{Z}) : m,n \in \mathbb{N}\}$ are called *delta Abel polynomials*, since they are analogs of the Abel polynomial $A(x) = x(x-a)^{n-1}$ and satisfy a multivariate identity of binomial type.

Theorem 9. *Assume that $\boldsymbol{Z}$ is a linear transformation of $\mathbb{N}^2$ by a 2×2 matrix A, i. e., there are constants a,b,c,d such that $x_{i,j} = ai + bj$ and $y_{i,j} = ci + dj$ for all $i,j \in \mathbb{N}$. Then*

$$\tilde{g}_{m,n}((x,y);\boldsymbol{Z}) = (xy - x_{0,n}y - y_{m,0}x)(x - x_{m,n} + 1)^{(m-1)}(y - y_{m,n} + 1)^{(n-1)}. \tag{1.3}$$

Proof. It follows from Theorem 5.1 of [11] and the fact that $(x^{(n)})_{n\in\mathbb{N}}$ is the basic sequence of the delta operator Δ_x. $\square$

4 Bivariate difference Gončarov polynomials and integer sequences

In this section, we focus on the combinatorial significance of bivariate difference Gončarov polynomials. Just as the univariate difference Gončarov polynomials describe lattice paths with a right boundary, or equivalently nondecreasing integer sequences with an upper bound, the bivariate difference Gončarov polynomials capture

the structure of a pair of nondecreasing integer sequences whose joint distribution is bounded by a set of constraints. First, we introduce the combinatorial model and the necessary notations.

Let $m, n \in \mathbb{N}$, and suppose $\boldsymbol{U}$ is a set of weight-vectors,

$$\boldsymbol{U} = \{(u_{i,j}, v_{i,j}) \in \mathbb{N}^2 : i, j \in \mathbb{N},\ u_{i,j} \le u_{i',j'},\ v_{i,j} \le v_{i',j'} \text{ whenever } (i,j) \preceq (i',j')\}.$$

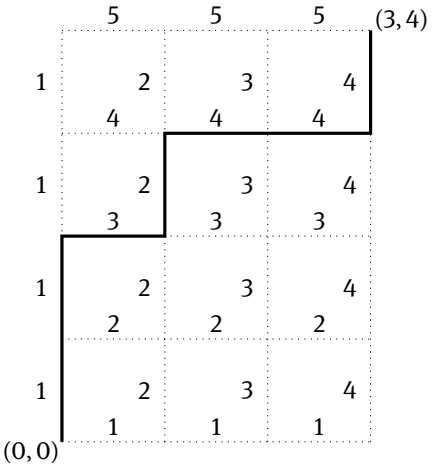
Define $D_{m,n}$ to be the directed graph having as vertices the points $\{(i,j) : 0 \le i \le m, 0 \le j \le n\}$ and having as edges all north steps $N = (0,1)$ and east steps $E = (1,0)$ connecting its vertices. Assign every edge e of $D_{m,n}$ a weight $wt(e)$ by letting

$$wt(e) = \begin{cases} u_{i,j} & \text{if } e \text{ is an east step from } (i,j) \text{ to } (i+1,j), \\ v_{i,j} & \text{if } e \text{ is a north step from } (i,j) \text{ to } (i,j+1). \end{cases}$$

Given a lattice path P from the origin $O = (0,0)$ to the point (m,n), we write $P = e_1 e_2 \dots e_{m+n}$, where $e_i \in \{E, N\}$, to record the sequence of steps of P. Thus, P must have exactly m E-steps and n N-steps. Consider a pair of nondecreasing integer sequences $(\boldsymbol{a}, \boldsymbol{b})$ with $\boldsymbol{a} = (a_0, a_1, \dots, a_{m-1})$ and $\boldsymbol{b} = (b_0, b_1, \dots, b_{n-1})$. We say that *the pair* $(\boldsymbol{a}, \boldsymbol{b})$ *is bounded by* P *with respect to the set* $\boldsymbol{U}$ if and only if, for $r = 1, 2, \dots, m+n$,

$$\begin{cases} a_i < u_{i,j} & \text{if } e_r \text{ is an E-step from } (i,j) \text{ to } (i+1,j), \\ b_j < v_{i,j} & \text{if } e_r \text{ is a N-step from } (i,j) \text{ to } (i,j+1). \end{cases}$$

Example. Let $\boldsymbol{U} = \{(u_{i,j}, v_{i,j}) : 0 \le i \le 3,\ 0 \le j \le 4\}$ be given by $u_{i,j} = j + 1$ and $v_{i,j} = i + 1$. The pair $(\boldsymbol{a}, \boldsymbol{b})$ with $\boldsymbol{a} = (2,2,3)$ and $\boldsymbol{b} = (0,0,1,3)$ is bounded by the lattice path $P = NNENEEN$ in bold in the figure below. Note that the lattice path bounding $(\boldsymbol{a}, \boldsymbol{b})$ may not be unique. For example, $P' = NNEENEN$ is another such path.



Let $\mathcal{I}(m,n)$ be the set of pairs of integer sequences $(\boldsymbol{a},\boldsymbol{b})$ such that

1. $\boldsymbol{a} = (a_0, a_1, \dots, a_{m-1})$ satisfies $0 \le a_0 \le a_1 \le \cdots \le a_{m-1} < x$, and
2. $\boldsymbol{b} = (b_0, b_1, \dots, b_{n-1})$ satisfies $0 \le b_0 \le b_1 \le \cdots \le b_{n-1} < y$.

Denote by $\mathcal{I}_{m,n}(P;\boldsymbol{U})$ the subset of $\mathcal{I}(m,n)$ consisting of the pairs of sequences $(\boldsymbol{a},\boldsymbol{b})$ that are bounded by P with respect to $\boldsymbol{U}$. Our main result is the following theorem.

Theorem 10. *Assume x,y are positive integers. The bivariate difference Gončarov polynomial $\tilde{g}_{m,n}((x,y);\boldsymbol{Z})$ counts the number of pairs of sequences in $\mathcal{I}(m,n)$ that are bounded by some lattice path from O to $A=(m,n)$. Explicitly, we have*

$$\frac{1}{m!n!}\tilde{g}_{m,n}\big((x,y);\boldsymbol{Z}\big) = \Big|\bigcup_{P:O\to A} \mathcal{I}_{m,n}(P;\boldsymbol{U})\Big|,$$

where P ranges over all lattice paths from O to A which use N- and E-steps only, and the set $\boldsymbol{U} = \{(u_{i,j}, v_{i,j}) : 0 \le i \le m, 0 \le j \le n\}$ is determined by $\boldsymbol{Z}$ according to the relations $u_{i,j} = x - x_{i,j}$, $v_{i,j} = y - y_{i,j}$.

Note: For the validity of the combinatorial interpretation, in Theorem 10, we assume that $x_{i,j}, y_{i,j} \in \mathbb{N}$, $0 \le x_{i,j} < x$, $0 \le y_{i,j} < y$, and $x_{i,j} < x_{i'j'}$, $y_{i,j} < y_{i',j}$ for all $(i',j') \preceq (i,j) \preceq (m,n)$.

Proof. Our proof uses a construction similar to that in Section 6 of [7]. For any pair of sequences $\boldsymbol{c} = (\boldsymbol{a},\boldsymbol{b}) \in \mathcal{I}(m,n)$, we construct a subgraph $G(\boldsymbol{c})$ of $D_{m,n}$ as follows:

- $O = (0,0)$ is a vertex of $G(\boldsymbol{c})$.
- For any vertex (i,j) of $G(\boldsymbol{c})$,
  - if $a_i < u_{i,j}$, then add the vertex $(i+1,j)$ and the E-step $\{(i,j),(i+1,j)\}$ to $G(\boldsymbol{c})$.
  - if $b_j < v_{i,j}$, then add the vertex $(i,j+1)$ and the N-step $\{(i,j),(i,j+1)\}$ to $G(\boldsymbol{c})$.

By definition $G(\boldsymbol{c})$ is a connected graph containing at least the vertex O. By Lemmas 6.3 and 6.4 of [7], we have that if edges $\{(i,j),(i+1,j)\}$ and $\{(i,j),(i,j+1)\}$ are both in $G(\boldsymbol{c})$, $\{(i+1,j),(i+1,j+1)\}$ and $\{(i,j+1),(i+1,j+1)\}$ are also in $G(\boldsymbol{c})$. Furthermore, the set of vertices of $G(\boldsymbol{c})$ has a unique maximal vertex $v(\boldsymbol{c})$ under the order $\preceq$.

Define the set $K_{m,n}(i,j) = \{\boldsymbol{c} \in \mathcal{I}(m,n) : v(\boldsymbol{c}) = (i,j)\}$, and let $k_{m,n}(i,j) = |K_{m,n}(i,j)|$. Then $\mathcal{I}(m,n)$ is the disjoint union of all $K_{m,n}(i,j)$ for $0 \le i \le m$ and $0 \le j \le n$, and

$$k_{m,n}(m,n) = \big|K_{m,n}(m,n)\big| = \Big|\bigcup_{P:O\to A} \mathcal{I}_{m,n}(P;\boldsymbol{U})\Big|.$$

Now a pair of sequences $\boldsymbol{c} = (\boldsymbol{a},\boldsymbol{b})$ is in $K_{m,n}(i,j)$ if and only if there exists a lattice path $P : O \to (i,j)$ satisfying the following:

- The initial segments $\boldsymbol{a}' = (a_0,\dots,a_{i-1})$ and $\boldsymbol{b}' = (b_0,\dots,b_{j-1})$ are bounded by P with respect to $\boldsymbol{U}$. That is, $(\boldsymbol{a}',\boldsymbol{b}')$ is in $K_{i,j}(i,j)$. There are $k_{i,j}(i,j)$ such pairs of initial segments.

– The integer sequence $(a_i, \ldots, a_{m-1})$ satisfies $u_{i,j} \le a_i \le \cdots \le a_{m-1} \le x-1$.
– The integer sequence $(b_j, \ldots, b_{n-1})$ satisfies $v_{i,j} \le b_j \le \cdots \le b_{n-1} \le y-1$.

Thus,

$$\begin{aligned} k_{m,n}(i,j) &= \#\{\boldsymbol{c} = (\boldsymbol{a},\boldsymbol{b}) : \boldsymbol{a}', \boldsymbol{b}' \text{ satisfy the three above conditions}\} \\ &= k_{i,j}(i,j)\binom{x-1-u_{i,j}+m-i}{m-i}\binom{y-1-v_{i,j}+n-j}{n-j} \\ &= k_{i,j}(i,j)\frac{(x-u_{i,j})^{(m-i)}}{(m-i)!}\frac{(y-v_{i,j})^{(n-j)}}{(n-j)!}. \end{aligned}$$

Hence

$$\begin{aligned} \frac{x^{(m)}}{m!}\frac{y^{(n)}}{n!} &= \left|\mathcal{I}(m,n)\right| \\ &= \sum_{i=0}^{m}\sum_{j=0}^{n} k_{m,n}(i,j) \\ &= \sum_{i=0}^{m}\sum_{j=0}^{n} \frac{(x-u_{i,j})^{(m-i)}}{(m-i)!}\frac{(y-v_{i,j})^{(n-j)}}{(n-j)!}k_{i,j}(i,j), \end{aligned}$$

or

$$x^{(m)}y^{(n)} = \sum_{i=0}^{m}\sum_{j=0}^{n}\binom{m}{i}\binom{n}{j}(x-u_{i,j})^{(m-i)}(y-v_{i,j})^{(n-j)}i!j!k_{i,j}(i,j).$$

Comparing this to the linear recursion formula in equation (1.2) and using the initial values $\tilde{g}_{0,0}((x,y);\boldsymbol{Z}) = k_{0,0}(0,0) = 1$, we conclude that $\tilde{g}_{m,n}((x,y);\boldsymbol{Z}) = m!n!k_{m,n}(m,n)$, where $z_{i,j} = (x_{i,j}, y_{i,j})$ with $x_{i,j} = x - u_{i,j}$ and $y_{i,j} = y - v_{i,j}$. □

Corollary 2. *Under the same assumptions of Theorem* 10*, we have*

$$\left|\bigcup_{P:O\to A}\mathcal{I}_{m,n}(P;\boldsymbol{U})\right| = \frac{1}{m!n!}\tilde{g}_{m,n}((0,0);-\boldsymbol{U}).$$

Proof. It follows from Theorem 6, the shift-invariant formula, and the relation $\boldsymbol{Z} = (x,y) - \boldsymbol{U}$. □

Recall that for a sequence of real numbers $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, the ith order statistic, $x_{(i)}$, is the ith term in the nondecreasing rearrangement $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ of $\mathbf{x}$. In [7], a generalized notion of *2-dimensional parking functions* was introduced in terms of order statistic constraints on a double sequence.

Definition 2. Given a set of nodes $\boldsymbol{U} = \{(u_{i,j}, v_{i,j}) : i,j \in \mathbb{N}\} \subset \mathbb{N}^2$ satisfying $u_{i,j} \le u_{i',j'}$ and $v_{i,j} \le v_{i',j'}$ when $(i,j) \preceq (i',j')$, a pair of nonnegative integer sequences $(\boldsymbol{a},\boldsymbol{b})$ with $\boldsymbol{a} = (a_0, a_1, \ldots, a_{m-1})$ and $\boldsymbol{b} = (b_0, b_1, \ldots, b_{n-1})$ is a *2-dimensional* $\boldsymbol{U}$*-parking function*

if and only if the order statistics of $(\boldsymbol{a}, \boldsymbol{b})$ are bounded by some lattice path from the origin to (m, n) with respect to $\boldsymbol{U}$.

Clearly, the set $K_{m,n}(m, n)$ consists of those 2-dimensional $\boldsymbol{U}$-parking functions with nondecreasing sequences $\boldsymbol{a}$ and $\boldsymbol{b}$. Following the convention in the univariate case, elements in $K_{m,n}(m, n)$ are called *2-dimensional increasing $\boldsymbol{U}$-parking functions*, and we replace the notation $K_{m,n}(m, n)$ with $\mathcal{IPF}^{(2)}_{m,n}(\boldsymbol{U})$. Hence Corollary 2 gives a formula that enumerates the set $\mathcal{IPF}^{(2)}_{m,n}(\boldsymbol{U})$.

If there exist some sequences $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \dots)$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots)$ such that $(u_{i,j}, v_{i,j}) = (\alpha_i, \beta_j)$, then

$$\frac{1}{m!n!}\tilde{g}_{m,n}\big((0,0); -\boldsymbol{U}\big) = \frac{1}{m!}\tilde{g}_m(0; -\boldsymbol{\alpha}) \cdot \frac{1}{n!}\tilde{g}_n(0; -\boldsymbol{\beta}).$$

In this case, $\mathcal{IPF}^{(2)}_{m,n}(U)$ is the direct product of the set of nondecreasing integer sequences of length m bounded by $\boldsymbol{\alpha}$ and the set of nondecreasing integer sequences of length n bounded by $\boldsymbol{\beta}$.

A more interesting case is when the node-set $\boldsymbol{U}$ is obtained from $\mathbb{N}^2$ by an affine transformation, i. e., there is a 2×2 matrix A such that

$$\begin{bmatrix} u_{i,j} \\ v_{i,j} \end{bmatrix} = A \begin{bmatrix} i \\ j \end{bmatrix} + \begin{bmatrix} s \\ t \end{bmatrix}. \tag{1.4}$$

Theorem 11. *Let $\boldsymbol{U}$ be given by equation* (1.4) *and*

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

with $a, b, c, d, s, t \in \mathbb{N}$. Then

$$\big|\mathcal{IPF}^{(2)}_{m,n}(\boldsymbol{U})\big| = \frac{1}{m!n!}(st + bnt + scm)(s + am + bn + 1)^{(m-1)}(t + cm + dn + 1)^{(n-1)}.$$

Proof. Using the shift invariant formula, we have

$$\tilde{g}_{m,n}\big((0,0); -\boldsymbol{U}\big) = \tilde{g}_{m,n}\big((s,t); \boldsymbol{Z}\big),$$

where the grid $\boldsymbol{Z}$ has nodes $\{z_{i,j} = (x_{i,j}, y_{i,j}) : (0,0) \preceq (i,j) \preceq (m,n)\}$ given by $x_{i,j} = -ai - bj$ and $y_{i,j} = -ci - dj$. Using equation (1.3) in Theorem 9, we obtain

$$\tilde{g}_{m,n}\big((x,y); Z\big) = (xy + xy_{m,0} + yx_{0,n})(x + x_{m,n} + 1)^{(m-1)}(y + y_{m,n} + 1)^{(n-1)},$$

which leads to the desired formula when substituted with $x = s$ and $y = t$. □

Corollary 3. *Let $\boldsymbol{U}$ be given by*

$$\begin{bmatrix} u_{i,j} \\ v_{i,j} \end{bmatrix} = \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Then

$$\left|\mathcal{IPF}^{(2)}_{m,n}(\boldsymbol{U})\right| = \frac{1+bn+cm}{(1+bn)(1+cm)}\binom{bn+m}{m}\binom{cm+n}{n}.$$

In particular, when $b = c = 1$ in Corollary 3, the 2-dimensional increasing $\boldsymbol{U}$-parking functions coincide with the *increasing* (p,q)*-parking functions* defined by Cori and Poulalhon [5], and Corollary 3 gives the Narayana number

$$\frac{1+m+n}{(1+m)(1+n)}\binom{m+n}{m}\binom{m+n}{n} = \frac{1}{1+m+n}\binom{1+m+n}{m}\binom{1+m+n}{n},$$

agreeing with Proposition 14 of [5].

5 Multivariate cases

Let $d \geq 1$ be a fixed integer. For a vector $\mathbf{v} \in \mathbb{F}^d$, we denote by v_j the jth component of $\mathbf{v}$. Given $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$, we set $\mathbf{n}! = n_1! n_2! \cdots n_d!$. For $\mathbf{k}, \mathbf{n} \in \mathbb{N}^d$, $\mathbf{k} \preceq \mathbf{n}$ means $k_i \leq n_i$ for all $1 \leq i \leq d$, and $\binom{\mathbf{n}}{\mathbf{k}} = \binom{n_1}{k_1} \cdots \binom{n_d}{k_d}$. With such notation in place, we can define the d-dimensional difference Gončarov polynomials with respect to the system of difference operators $(\Delta_{x_1}, \ldots, \Delta_{x_d})$. Given a grid $\boldsymbol{Z} = \{z_{\mathbf{k}} \in \mathbb{F}^d : \mathbf{k} \in \mathbb{N}^d\}$, there is a unique polynomial $t_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z})$ of coordinate degree $\mathbf{n} \in \mathbb{N}^d$ satisfying

$$\varepsilon(z_{\mathbf{k}})\Delta_{x_1}^{k_1} \cdots \Delta_{x_d}^{k_d}(t_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z})) = \mathbf{n}!\delta_{\mathbf{k},\mathbf{n}}$$

for all $\mathbf{k} \preceq \mathbf{n}$. This polynomial is the multivariate difference Gončarov polynomial indexed by $\mathbf{n}$, which we will denote by $\tilde{g}_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z})$. Theorems 2–9 can all be extended straightforwardly to d dimensions, where the summations are over the set $\{\mathbf{k} \in \mathbb{N}^d : \mathbf{k} \preceq \mathbf{n}\}$, and the binomial coefficient $\binom{m}{i}\binom{n}{j}$ is replaced with $\binom{\mathbf{n}}{\mathbf{k}}$.

To generalize all definitions and results of Section 4 to d dimensions for $d > 2$, we must first define the d-dimensional analogs of the sets $\mathcal{I}(m,n)$ and $\mathcal{I}_{m,n}(P; \boldsymbol{U})$. Let $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$, and fix a d-dimensional set of weight-vectors,

$$\boldsymbol{U} = \{u_{\mathbf{k}} \in \mathbb{N}^d : \mathbf{k} \in \mathbb{N}^d, u_{\mathbf{k},i} \leq u_{\mathbf{k}',i} \text{ whenever } \mathbf{k} \preceq \mathbf{k}' \text{ and } 1 \leq i \leq d\},$$

where $u_{\mathbf{k},i}$ is the ith entry of the point $u_{\mathbf{k}}$. We can extend the weighted directed graph $D_{m,n}$ to d dimensions by taking as vertices the points $\{(k_1, \ldots, k_d) : 0 \leq k_i \leq n_i \text{ for all } i = 1, \ldots, d\}$ and as edges all steps $\mathbf{e}_i$, $1 \leq i \leq d$, where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ has 1 in the

ith entry. An edge ℓ from $\mathbf{k} = (k_1, \ldots, k_i, \ldots, k_d)$ to $(k_1, \ldots, k_i + 1, \ldots, k_d)$ is assigned the weight $wt(\ell) = u_{\mathbf{k},i}$.

Suppose $P = \ell_1 \ell_2 \ldots \ell_n$ is any lattice path from the origin $O = (0, \ldots, 0)$ to the point $(n_1, \ldots, n_d)$, where each $\ell_i \in \{\mathbf{e}_j : 1 \leq j \leq d\}$ and $n = n_1 + \cdots + n_d$. Consider a d-tuple of nondecreasing integer sequences $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)})$, where $\boldsymbol{a}^{(i)} = (a_0^{(i)}, a_1^{(i)}, \ldots, a_{n_i-1}^{(i)})$ for $1 \leq i \leq d$. Then we say that $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)})$ is bounded by P with respect to the set $\boldsymbol{U}$ if the following condition is satisfied for each $r = 1, 2, \ldots, n$: $a_{k_i}^{(i)} < u_{\mathbf{k},i}$ if ℓ_r is an $\mathbf{e}_i$–step from $\boldsymbol{k} = (k_1, \ldots, k_i, \ldots, k_d)$ to $(k_1, \ldots, k_i + 1, \ldots, k_d)$.

For a fixed $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{N}^d$, let the set $\mathcal{I}(\mathbf{n})$ consist of all d-tuples of nondecreasing integer sequences $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)}) \in \mathbb{N}^{n_1} \times \cdots \mathbb{N}^{n_d}$ such that $0 \leq \boldsymbol{a}_j^{(i)} < x_i$ for all $1 \leq i \leq d$ and $0 \leq j < n_i$, and let $\mathcal{I}_{\mathbf{n}}(P; \boldsymbol{U})$ be the subset of $\mathcal{I}(\mathbf{n})$ containing all d-tuples that are bounded by P with respect to $\boldsymbol{U}$. The following are the d-dimensional analogs of Theorem 10 and Corollary 2.

Theorem 12. *Let $\mathbf{x} = (x_1, \ldots, x_d)$, $\mathbf{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$ and $\boldsymbol{Z} = \{z_{\mathbf{i}} : \mathbf{i} \in \mathbb{N}^d\} \subset \mathbb{N}^d$. The multivariate difference Gončarov polynomial $\tilde{g}_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z})$ gives the number of d-tuples of sequences in $\mathcal{I}(\mathbf{n})$ that are bounded by some lattice path from the origin to $A = (n_1, \ldots, n_d)$. In particular,*

$$\left| \bigcup_{P: O \to A} \mathcal{I}_{\mathbf{n}}(P; U) \right| = \frac{1}{\mathbf{n}!} \tilde{g}_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z}) = \frac{1}{\mathbf{n}!} \tilde{g}_{\mathbf{n}}((0, \ldots, 0); -\boldsymbol{U}), \tag{1.5}$$

where the set $\boldsymbol{U} = \{u_{\mathbf{k}} \in \mathbb{N}^d : \mathbf{k} \preceq \mathbf{n}\}$ is defined by $u_{\mathbf{k},i} = x_i - z_{\mathbf{k},i}$.

As in the 2-dimensional case, we can define a d-dimensional $\boldsymbol{U}$-parking function according to certain constraints imposed on the order statistics of a d-tuple.

Definition 3. Let $d > 2$ be an integer. Given a set of nodes $\boldsymbol{U} = \{u_{\mathbf{k}} \in \mathbb{F}^d : \mathbf{k} \in \mathbb{N}^d\}$ such that $u_{\mathbf{k},i} \leq u_{\mathbf{k}',i}$ whenever $\mathbf{k} \preceq \mathbf{k}'$ and $1 \leq i \leq d$, a d-tuple of integer sequences $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)}) \in \mathbb{N}^{n_1} \times \cdots \times \mathbb{N}^{n_d}$ is a *d-dimensional $\boldsymbol{U}$-parking function* if and only if the order statistics of $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)})$ are bounded by some lattice path from the origin to $(n_1, \ldots, n_d)$ with respect to $\boldsymbol{U}$.

Of particular importance to us are the *d-dimensional increasing $\boldsymbol{U}$-parking functions* $(\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)})$, which have nondecreasing constituent sequences $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(d)}$. Using notation consistent with the bivariate case, we will denote the set of all d-dimensional increasing $\boldsymbol{U}$-parking functions by $\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U})$. Note then that the set $\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U})$ is equivalent to the union of the sets $\mathcal{I}_{\mathbf{n}}(P; U)$ over all lattice paths P from the origin to $A = (n_1, \ldots, n_d)$, so that Theorem 12 also yields a formula for $\#\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U})$ in terms of a multivariate difference Gončarov polynomial.

In the following, we enumerate the set $\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U})$ when the node-set $\boldsymbol{U}$ is affine, meaning there exists a $d \times d$ matrix A and vector $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{F}^d$ such that $u_{\mathbf{k}} = A\mathbf{k} + \mathbf{s}$ for all $\mathbf{k} \in \mathbb{N}^d$, where $\mathbf{k}$, $u_{\mathbf{k}}$, and $\mathbf{s}$ are treated as column vectors. We express such an affine node-set $\boldsymbol{U}$ by $\boldsymbol{U} = A\mathbb{N}^d + \mathbf{s}$. Increasing $\boldsymbol{U}$-parking functions associated

to affine $\boldsymbol{U}$ relate to the notion of $(p_1, p_2, \ldots, p_d)$-parking functions in [5] and the notion of G-parking functions [15], when G is a complete d-partite graph with a distinguished root.

We use a closed formula for d-dimensional delta Abel polynomials proved in Theorem 6.1 of [11]. The following statement is specialized to the system of operators $(\Delta_{x_1}, \ldots, \Delta_{x_d})$.

Theorem 13 ([11]). *Let* $\mathbf{x}$, $\mathbf{n}$, *and* $\boldsymbol{Z}$ *be as in Theorem* 12, *and suppose* $\boldsymbol{Z} = A\mathbb{N}^d$ *for some* $d \times d$ *matrix* $A = (a_{i,j})$. *Let* $B = (b_{i,j})$ *be the* $d \times d$ *diagonal matrix defined by* $b_{i,i} = x_i - z_{\mathbf{n},i}$, *and let* $C = (c_{i,j})$ *be the* $d \times d$ *matrix defined by* $c_{i,j} = z_{n_i \mathbf{e}_i, j}$. *Then*

$$\tilde{g}_{\mathbf{n}}(\mathbf{x}; \boldsymbol{Z}) = \det(B + C) \prod_{i=1}^{d} (x_i - z_{\mathbf{n},i} + 1)^{(n_i - 1)}.$$

Theorem 13 yields the following result on the enumeration of d-dimensional increasing $\boldsymbol{U}$-parking functions.

Corollary 4. *Let* $\mathbf{x}$ *and* $\mathbf{n}$ *be as in Theorem* 12, *and suppose* $\boldsymbol{U} = A\mathbb{N}^d + \mathbf{s}$ *for some* $d \times d$ *matrix* $A = (a_{i,j})$ *and* $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{Z}^d$. *Let* $B = (b_{i,j})$ *be the* $d \times d$ *diagonal matrix defined by* $b_{i,i} = x_i + z_{\mathbf{n},i}$, *and let* $C = (c_{i,j})$ *be the* $d \times d$ *matrix defined by* $c_{i,j} = -z_{n_i \mathbf{e}_i, j}$, *where* $\boldsymbol{Z} = A\mathbb{N}^d$. *Then*

$$\#\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U}) = \frac{1}{\mathbf{n}!} \det(B + C) \prod_{i=1}^{d} (s_i + z_{\mathbf{n},i} + 1)^{(n_i - 1)}.$$

Proof. From Theorem 12 and the shift-invariant property of Gončarov polynomials, we have

$$\#\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U}) = \frac{1}{\mathbf{n}!} \tilde{g}_{\mathbf{n}}(\mathbf{0}; -\boldsymbol{U}) = \frac{1}{\mathbf{n}!} \tilde{g}_{\mathbf{n}}(\mathbf{s}; -\boldsymbol{Z}).$$

Then we use Theorem 13 and notice that the transition matrix for the grid $-\boldsymbol{Z}$ is $-A$. □

Corollary 5. *Suppose* $\boldsymbol{U} = A\mathbb{N}^d + \mathbf{s}$, *where* $A = (a_{i,j})$ *is a* $d \times d$ *matrix with*

$$a_{i,j} = \begin{cases} \alpha_j & \text{if } i \neq j \\ 0 & i = j \end{cases}$$

and $\mathbf{s} = (s_1, \ldots, s_d) \in \mathbb{Z}^d$. *Then we have*

$$\#\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U}) = \frac{1}{\mathbf{n}!} \left(1 - \sum_{j=1}^{d} \frac{\alpha_j n_j}{s_j + N}\right) \prod_{i=1}^{d} (s_i + N)(s_i + N - \alpha_i n_i + 1)^{(n_i - 1)},$$

where $N = \sum_{i=1}^{d} \alpha_i n_i$.

Proof. The result follows from Corollary 4 and the computation of $\det(B+C)$, where B and C are the $d \times d$ matrices defined in Theorem 13. Let $N = \sum_{i=1}^{d} \alpha_i n_i$. We have

$$B + C = \begin{bmatrix} s_1 + N - \alpha_1 n_1 & -\alpha_1 n_1 & \cdots & -\alpha_1 n_1 \\ -\alpha_2 n_2 & s_2 + N - \alpha_2 n_2 & \cdots & -\alpha_2 n_2 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_d n_d & -\alpha_d n_d & \cdots & s_d + N - \alpha_d n_d \end{bmatrix}.$$

Subtracting column $j-1$ from column j for $j = d, d-1, \ldots, 2$, we get that $\det(B+C)$ equals

$$\begin{vmatrix} s_1 + N - \alpha_1 n_1 & -s_1 - N & 0 & \cdots & 0 & 0 \\ -\alpha_2 n_2 & s_2 + N & -s_2 - N & \cdots & 0 & 0 \\ -\alpha_3 n_3 & 0 & s_3 + N & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -\alpha_{d-1} n_{d-1} & 0 & 0 & \cdots & s_{d-1} + N & -s_{d-1} - N \\ -\alpha_d n_d & 0 & 0 & \cdots & 0 & s_d + N \end{vmatrix},$$

which is

$$\begin{aligned} &(s_1 + N - \alpha_1 n_1) \prod_{i=2}^{d} (s_i + N) + \sum_{i=2}^{d} (-1)^{1+i} (-\alpha_i n_i) \left[\prod_{j=1}^{i-1} (-s_j - N) \right] \left[\prod_{j=i+1}^{d} (s_j + N) \right] \\ &= \left(1 - \sum_{j=1}^{d} \frac{\alpha_j n_j}{s_j + N} \right) \prod_{i=1}^{d} (s_i + N). \end{aligned}$$

Then Corollary 5 is obtained from Corollary 4 with $z_{\mathbf{n},i} = u_{\mathbf{n},i} - s_i = N - \alpha_i n_i$. □

When $\alpha_i = s_i = 1$ for all $1 \le i \le d$ in Corollary 18, the second and the third authors showed in [17] that the set of d-dimensional increasing $\boldsymbol{U}$-parking functions is precisely the set of increasing $(p_1, p_2, \ldots, p_d)$-parking functions described by Cori and Poulalhon [5]. According to Proposition 19 of [5], the latter are counted by the formula

$$\frac{1}{N+1} \prod_{i=1}^{d} \binom{N+1}{n_i},$$

which matches the value for $\#\mathcal{IPF}_{\mathbf{n}}^{(d)}(\boldsymbol{U})$ given by Corollary 5 after simple algebraic manipulation.

Bibliography

[1] A. Adeniran and C. Yan, Gončarov polynomials in partition lattices and exponential families, *Adv. Appl. Math.*, **126**, 102045 (2021).

[2] A. Adeniran and C. Yan, On increasing and invariant parking sequences, *Australas. J. Comb.*, **79**(1) (2021), 167–182.

[3] R. P. Boas and R. C. Buck, *Polynomial Expansion of Analytic Functions*, Springer, Heiderberg, 1958.

[4] S. Butler, R. Graham and C. Yan, Parking distributions on trees, *Eur. J. Comb.*, **65** (2017), 168–185.

[5] R. Cori and D. Poulalhon, Enumeration of (p, q)-parking functions, *Discrete Math.*, **256** (2002), 609–623.

[6] V. L. Gončarov, *The Theory of Interpolation and Approximation of Functions*, Gostekhizdat, Moscow, 1954.

[7] N. Khare, R. Lorentz and C. Yan, Bivariate Gončarov polynomials and integer sequences, *Sci. China Math.*, **57** (2014), 1561–1578.

[8] C. F. Krattehthaler, Lattice path enumeration, in *Handbook of Enumerative Combinatorics*, pp. 589–678, Discrete Math. Appl. CRC Press, Boca Raton, FL, 2015.

[9] J. P. S. Kung, X. Sun and C. Yan, Gončarov-type polynomials and applications in combinatorics, preprint (2006). Available at the http://www.math.tamu.edu/~cyan/Files/DGP.pdf.

[10] J. P. S. Kung and C. Yan, Gončarov polynomials and parking functions, *J. Comb. Theory, Ser. A*, **102**(1) (2003), 16–37.

[11] R. Lorentz, S. Tringali and C. Yan, Multivariate delta Gončarov and Abel polynomials, *J. Math. Anal. Appl.*, **446**(1) (2017), 663–680.

[12] R. Lorentz, S. Tringali and C. Yan, Generalized Gončarov polynomials, in S. Butler et al. (eds.) *Connections in Discrete Mathematics: A Celebration of the Work of Ron Graham*, pp. 56–85, Cambridge University Press, Cambridge, 2018.

[13] S. G. Mohanty, *Lattice Path Counting and Applications*, Probability and Mathematical Statistics. Academic Press, New York, 1979.

[14] R. Mullin and G.-C. Rota, On the foundations of combinatorial theory. III. Theory of binomial enumeration, in B. Harris (ed.) *Graph Theory and Its Applications*, pp. 167–213, Academic Press, New York, 1970.

[15] A. Postnikov and B. S. Trees, Parking functions, syzygies, and deformations of monomial ideals, *Trans. Am. Math. Soc.*, **356**(8) (2004), 3109–3142.

[16] G.-C. Rota, D. Kahaner and A. Odlyzko, On the foundations of combinatorial theory. VII. Finite operator calculus, *J. Math. Anal. Appl.*, **42**(3) (1973), 684–760.

[17] L. Snider and C. Yan, On 2-dimensional parking functions, preprint.

J.-P. Allouche

On an inequality in a 1970 paper of R. L. Graham

Dedicated to the memory of Ron Graham

Abstract: Having recently come across a 1970 paper of R. L. Graham about cube-numbering and its generalizations, we found that one proof uses an inequality about the summatory function of the sum of binary digits of integers. Graham gave a very elegant and somehow unexpected proof of this inequality. We propose a more "pedestrian"—and somehow more standard—proof of this inequality, as well as questions about possible generalizations.

1 Introduction

R. L. Graham, in a study of cube-numbering and generalizations [5] used a curious inequality for the summatory function of the sum of binary digits, which we now describe. Let $w(k)$ be the number of 1's in the binary expansion of the integer k. Let $W(n) := \sum_{0\le k\le n} w(k)$. Then, for all n_1, n_2 with $0 < n_1 \le n_2$,

$$W(n_1 - 1) + W(n_2 - 1) + n_1 < W(n_1 + n_2 - 1) + 1. \tag{2.1}$$

The very elegant proof of this inequality given by Graham uses the following lemma.

Lemma 1 (Graham [5]). *Let r and s be nonnegative integers. If φ is a one-to-one map from $[0, r]$ to $[s, s + r]$, define*

$$\delta(\varphi) := \min_{0\le k\le r} \{w(\varphi(k)) - w(k)\}.$$

Then

(i) *There exists φ such that $\delta(\varphi) \ge 0$.*

(ii) *If $s > r$, then there exists φ such that $\delta(\varphi) \ge 1$.*

The proof of this lemma and of the fact that it implies Inequality (2.1) can be found in [5]—but also see [6]. We could not guess how Graham had the idea of this lemma. Thus we wondered whether there could be a more direct, possibly "pedestrian," proof.

Acknowledgement: We would like to thank the referee for their pertinent remarks.

J.-P. Allouche, CNRS, IMJ-PRG, Sorbonne Université, Paris, France, e-mail: jean-paul.allouche@imj-prg.fr

https://doi.org/10.1515/9783110754216-002

The point is that the sum of digit sequence (sequence A000120 in [11]), hence its summatory function (sequence A000788 in [11]), are 2-regular sequences in the sense of [1, 2]. Recall that a sequence $(u(n))_{n\geq 0}$ is called 2-regular if the $\mathbb{Z}$-module generated by its 2-kernel (i. e., the set of subsequences $\{(u(2^k n + a))_{n\geq 0},\ k \geq 0,\ a \in [0, 2^k - 1]\}$) is a finitely generated $\mathbb{Z}$-module. In our case, this is just a consequence of the fact that $(w(2n))_{n\geq 0}$ and $(w(2n+1))_{n\geq 0}$ are linear combinations of the sequence $(w(n))_{n\geq 0}$ and the constant sequence $(1)_{n\geq 0}$. Namely, for all $n \geq 0$, we have $w(2n) = w(n)$ and $w(2n+1) = w(n) + 1$. These equalities imply equalities of a similar type for $W(n)$, and will be the basis of our proof.

2 Proof of Graham's inequality (2.1)

First, we rewrite inequality (2.1) as: for all (n_1, n_2) with $1 \leq n_1 \leq n_2$,

$$W(n_1 - 1) + W(n_2 - 1) + n_1 < W(n_1 + n_2 - 1) + 1.$$

Defining $m := n_1 - 1$ and $n := n_2 - 1$, inequality (2.1) is equivalent to inequality (2.2): for all (m, n) with $0 \leq m \leq n$,

$$W(m) + W(n) + m < W(m + n + 1). \tag{2.2}$$

Lemma 2. *The following equalities hold:*

$$\begin{aligned} &\text{for all } n \geq 1, && W(2n) = W(n) + W(n-1) + n \\ &\text{for all } n \geq 0, && W(2n+1) = 2W(n) + n + 1. \end{aligned}$$

Let $A(m, n) := W(m + n + 1) - W(m) - W(n) - m$. Then the following equalities hold:

$$\begin{aligned} &\text{for } n \geq 0, && A(n, n) = 1 \quad \text{(hence inequality (2.2) is sharp)} \\ &\text{for } m \geq 1 \text{ and } n \geq 1, && A(2m, 2n) = A(m-1, n) + A(m, n-1) \\ &\text{for } m \geq 0 \text{ and } n \geq 1, && A(2m+1, 2n) = A(m, n) + A(m, n-1) - 1 \\ &\text{for } m \geq 1 \text{ and } n \geq 0, && A(2m, 2n+1) = A(m, n) + A(m-1, n) - 1 \\ &\text{for } m \geq 0 \text{ and } n \geq 0, && A(2m+1, 2n+1) = 2A(m, n) - 1. \end{aligned}$$

Proof. We write, for $\ell \geq 1$,

$$\begin{aligned} W(2\ell) = \sum_{k=0}^{2\ell} w(k) &= \sum_{j=0}^{\ell} w(2k) + \sum_{j=0}^{\ell-1} w(2k+1) \\ &= \sum_{j=0}^{\ell} w(k) + \sum_{j=0}^{\ell-1} (w(k) + 1) \\ &= W(\ell) + W(\ell - 1) + \ell \end{aligned}$$

and, for $\ell \geq 0$,

$$\begin{aligned} W(2\ell+1) = \sum_{k=0}^{2\ell+1} w(k) &= \sum_{j=0}^{\ell} w(2k) + \sum_{j=0}^{\ell} w(2k+1) \\ &= \sum_{j=0}^{\ell} w(k) + \sum_{j=0}^{\ell} \big(w(k)+1\big) \\ &= 2W(\ell) + \ell + 1. \end{aligned}$$

Using these relations for $W(2\ell)$ and $W(2\ell+1)$, we obtain successively:

$$\text{For all } n \geq 0,\ A(n,n) = W(2n+1) - 2W(n) - n = 1$$

and the following relations.

– For all (m,n) with $m \geq 1$, $n \geq 1$,

$$\begin{aligned} A(2m,2n) &= W(2m+2n+1) - W(2m) - W(2n) - 2m \\ &= 2W(m+n) - W(m) - W(m-1) \\ &\quad - W(n) - W(n-1) - 2m + 1 \\ &= A(m-1,n) + A(m,n-1). \end{aligned}$$

– For all (m,n) with $m \geq 0$, $n \geq 1$,

$$\begin{aligned} A(2m+1,2n) &= W(2m+2n+2) - W(2m+1) - W(2n) - 2m - 1 \\ &= W(m+n+1) + W(m+n) - 2W(m) \\ &\quad - W(n) - W(n-1) - 2m - 1 \\ &= A(m,n) + A(m,n-1) - 1. \end{aligned}$$

– For all (m,n) with $m \geq 1$, $n \geq 0$,

$$\begin{aligned} A(2m,2n+1) &= W(2m+2n+2) - W(2m) - W(2n+1) - 2m \\ &= W(m+n+1) + W(m+n) - W(m) \\ &\quad - W(m-1) - 2W(n) - 2m \\ &= A(m,n) + A(m-1,n) - 1. \end{aligned}$$

– For all (m,n) with $m \geq 0$, $n \geq 0$,

$$\begin{aligned} A(2m+1,2n+1) &= W(2m+2n+3) - W(2m+1) - W(2n+1) - 2m - 1 \\ &= 2W(m+n+1) - 2W(m) - 2W(n) - 2m - 1 \\ &= 2A(m,n) - 1. \end{aligned}$$

□

Now we are ready to prove the following proposition which is the result of Graham described above.

Proposition. *Inequality* (2.1) *holds.*

Proof. As we have seen, it suffices to prove inequality (2.2), namely for all (m, n) with $0 \le m \le n$, one has

$$W(m) + W(n) + m < W(m + n + 1).$$

We will prove the statement $(\mathcal{H}_n)$:

$$(\mathcal{H}_n): \quad \text{for all } m \in [0, n],\ A(m, n) = W(m + n + 1) - W(m) - W(n) - m > 0.$$

We will actually check that $(\mathcal{H}_0)$ holds, and prove that, if $(\mathcal{H}_r)$ holds for all $r \in [0, n]$, then $(\mathcal{H}_{2n})$ and $(\mathcal{H}_{2n+1})$ hold. Note that $(\mathcal{H}_0)$ holds trivially. Now suppose that, for some $n \ge 0$, $(\mathcal{H}_r)$ holds for all $r \in [0, n]$, i. e., $A(q, r) > 0$ for all $r \in [0, n]$ and for all $q \in [0, r]$. We look at $A(k, 2n)$ and $A(k, 2n + 1)$ for $k \le 2n$, respectively, $k \le 2n + 1$.

- $A(k, 2n)$
  We can (and will) suppose that $n > 0$.
  - If $k \in [0, 2n]$ is even, say $k = 2\ell$, we may suppose that $k \in [1, 2n - 1]$, since the case $k = 0$ is trivial, and the case $k = n$ is deduced from $A(2n, 2n) = 1$ as seen above. Thus $\ell \in [1, n - 1]$. We have

    $$A(k, 2n) = A(2\ell, 2n) = A(\ell - 1, n) + A(\ell, n - 1) > 0$$

    using first Lemma 2, then the induction hypothesis.
  - If $k \in [0, 2n]$ is odd, say $k = 2\ell + 1$, we have that $k \in [1, 2n - 1]$. Thus $\ell \in [0, n - 1]$. We have

    $$A(k, 2n) = A(2\ell + 1, 2n) = A(\ell, n) + A(\ell, n - 1) - 1 > 0$$

    using first Lemma 2, then the induction hypothesis.
- $A(k, 2n + 1)$
  - If $k \in [0, 2n + 1]$ is even, say $k = 2\ell$. We may suppose $k \neq 0$ since the case $k = 0$ is trivial. Then $\ell \in [1, n]$. Thus

    $$A(k, 2n + 1) = A(2\ell, 2n + 1) = A(\ell, n) + A(\ell - 1, n) - 1 > 0$$

    using first Lemma 2, then the induction hypothesis.
  - If $k \in [0, 2n + 1]$ is odd, say $k = 2\ell + 1$. Thus $\ell \in [0, n]$. Thus

    $$A(k, 2n + 1) = A(2\ell + 1, 2n + 1) = 2A(\ell, n) - 1 > 0$$

    using first Lemma 2, then the induction hypothesis. □

Remark 1. It is noted in [6] that $W(n) \le \frac{1}{2} n \log_2 n$. Actually we know more: from [3] (also see [13]), we have $W(n) = \frac{1}{2} n \log_2 n + nF(\log_2 n)$, where F is a periodic, continuous, nowhere differentiable function with an absolutely convergent Fourier series. The function F, called the Trollope–Delange function, is closely related to the Takagi function [12] (see, e. g., [7, 8], and [10]). The evaluation of $W(n)$, F, and generalizations is still an active subject of research (see, e. g., [4]).

Remark 2. It is possible to obtain an upper bound for $A(m,n)$ by using the inequality $w(m+n) \le w(m) + w(n)$ for all m, n. This inequality can be found, e. g., in a comment by Shevelev about sequence A000120 [11]. Namely, one has $w(m) + w(n) - w(n+m) = v_2(\binom{n+m}{n})$, which is a consequence of Legendre's result [9, p. 10–12]: $w(n) = n - v_2(n!)$, where $v_2(k)$ is the 2-adic valuation of k. Hence

$$\begin{aligned} A(m,n) &= \sum_{j=0}^{m+n+1} w(j) - \sum_{j=0}^{n} w(j) - \sum_{j=0}^{m} w(j) - m \\ &= \sum_{j=n+1}^{m+n+1} w(j) - \sum_{j=0}^{m} w(j) - m = \sum_{j=0}^{m} w(j+n+1) - \sum_{j=0}^{m} w(j) - m \\ &\le \sum_{j=0}^{m} w(n+1) - m = (m+1)w(n+1) - m. \end{aligned}$$

Note that this inequality is sharp (e. g., take $m = n = 2^k - 1$ for some $k \ge 1$).

3 Conclusion

A first possible generalization of this inequality is to replace base 2 with base $d \ge 2$, and w with the sum of digits in base d. But, since our proof essentially uses the 2-regularity of the sequence $(w(n))_{n\ge 0}$, and (thus) of its summatory function $(W(n))_{n\ge 0}$, one can ask whether some similar "super-superadditivity" result holds for summatory functions of all 2-regular (resp., d-regular) sequences. A more reasonable question could be whether summatory functions of pattern-counting sequences have such a property: the sequence $(w(n))_{n\ge 0}$ counts the number of 1's in the binary expansion of n, thus one of the first examples to test would be the sequence $(u(n))_{n\ge 0}$ that counts the number of possibly overlapping blocks 11 in the binary expansion of n (this is sequence A014081 in [11]). The difficulty is then that, instead of having the relations $w(2n) = w(n)$ and $w(2n+1) = w(n)+1$, we have relations for a three-dimensional vector $z(n)$:

$$z(n) := \begin{pmatrix} u(n) \\ u(2n+1) \\ 1 \end{pmatrix} \Rightarrow z(2n) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} z(n), \quad z(2n+1) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} z(n).$$

Bibliography

[1] J.-P. Allouche and J. Shallit, The ring of k-regular sequences, *Theor. Comput. Sci.*, **98** (1992), 163–197.
[2] J.-P. Allouche and J. Shallit, The ring of k-regular sequences, II, *Theor. Comput. Sci.*, **307** (2003), 3–29.
[3] H. Delange, Sur la fonction sommatoire de la fonction "somme des chiffres", *Enseign. Math.*, **21** (1975), 31–47.
[4] O. E. Galkin and S. Yu. Galkina, Global extrema of the Delange function, estimations of digital sums and concave functions, *Mat. Sb.*, **211** (2020), 32–70. Translated in Sb. Math. **211** (2020), 336–372.
[5] R. L. Graham, On primitive graphs and optimal vertex assignments, *Ann. N.Y. Acad. Sci.*, **175** (1970), 170–186.
[6] J. C. Jones and B. F. Torrence, The case of the missing case: the completion of a proof by R. L. Graham, *Pi Mu Epsilon J.*, **10** (1999), 772–778.
[7] M. Krüppel, Takagi's continuous nowhere differentiable function and binary digital sums, *Rostock. Math. Kolloqu.*, **63** (2008), 37–54.
[8] J. C. Lagarias, The Takagi function and its properties, in *Functions in Number Theory and their Probabilistic Aspects*, RIMS Kôkyûroku Bessatsu, B34, Res. Inst. Math. Sci., pp. 153–189, RIMS, Kyoto, 2012.
[9] A.-M. Legendre, *Théorie des Nombres*, Firmin Didot Frères, Paris, 1830.
[10] T. Okada, T. Sekiguchi and Y. Shiota, Applications of binomial measures to power sums of digital sums, *J. Number Theory*, **52** (1995), 256–266.
[11] On-Line Encyclopedia of Integer Sequences, founded by N. J. A. Sloane. Electronically available at https://oeis.org.
[12] T. Takagi, A simple example of the continuous function without derivative, *Tokio Math. Ges.*, **1** (1903), 176–177. https://doi.org/10.11429/subutsuhokoku1901.1.F176.
[13] J. R. Trollope, An explicit expression for binary digital sums, *Math. Mag.*, **41** (1968), 21–25.

Noga Alon, Ryan Alweiss, Yang P. Liu, Anders Martinsson, and Shyam Narayanan

Arithmetic progressions in sumsets of sparse sets

Dedicated to the memory of Ron Graham

Abstract: A set of positive integers $A \subset \mathbb{Z}_{>0}$ is *log-sparse* if there is an absolute constant C so that for any positive integer x the sequence contains at most C elements in the interval $[x, 2x)$. In this note, we study arithmetic progressions in sums of log-sparse subsets of $\mathbb{Z}_{>0}$. We prove that for any log-sparse subsets $S_1, \dots, S_n$ of $\mathbb{Z}_{>0}$, the sumset $S = S_1 + \cdots + S_n$ cannot contain an arithmetic progression of size greater than $n^{(1+o(1))n}$. We also show that this is nearly tight by proving that there exist log-sparse sets $S_1, \dots, S_n$ such that $S_1 + \cdots + S_n$ contains an arithmetic progression of size $n^{(1-o(1))n}$.

1 Introduction

Arithmetic progressions have been one of the favorite research topics of Ron Graham. See [3] for a lecture he has given on the subject. The term "arithmetic progression" appears in the title of seven of his papers, and he has written, with András Hajnal, the proof of Szemerédi's theorem on arithmetic progressions in sets of integers of positive upper density [7]. In the present note, dedicated to his memory, we study the maximum possible length of arithmetic progressions in sumsets of very sparse sets.

Waring's problem, first proven by Hilbert [6], states that there exists a function $f(k)$ so that any positive integer can be written as the sum of at most $f(k)$ perfect k-powers. From a crude heuristic perspective, the density of perfect k-powers makes

Acknowledgement: We thank Christian Elsholtz for helpful comments. Research of Noga Alon was supported in part by NSF grant DMS-1855464 and the Simons Foundation. Research of Ryan Alweiss was supported by the NSF Graduate Fellowship (GRFP). Research of Yang P. Liu was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program. Research of Shyam Narayanan was supported by the NSF Graduate Fellowship (GRFP) and a Simons Investigator Award.

Noga Alon, Ryan Alweiss, Department of Mathematics, Princeton University, Princeton, NJ, USA, e-mails: nalon@math.princeton.edu, alweiss@math.princeton.edu
Yang P. Liu, Department of Mathematics, Stanford University, Stanford, CA, USA, e-mail: yangpliu@stanford.edu
Anders Martinsson, Institut für Theoretische Informatik, ETH Zürich, Zürich, Switzerland, e-mail: anders.martinsson@inf.ethz.ch
Shyam Narayanan, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA, e-mail: shyamsn@mit.edu

https://doi.org/10.1515/9783110754216-003

this result plausible. As the number of ways to write integers between 1 and n as sums of r perfect k-powers is asymptotically $\Theta(n^{r/k})$ if r and k are fixed, on average one may expect any (large) n to have some such representation as long as $r > k$. However, for many values of k there are congruence obstructions, so that certain arithmetic progressions cannot be reached by a sum of $k+1$ k-th powers. In the literature on Waring's problem, the worst cases of the congruence obstructons are summarized in a variable $\Gamma(k)$ and the common belief is that any large n has a representation provided $r \geq \max(k+1, \Gamma(k))$.

In this note, instead of perfect k-powers, we consider sums of sets with much lower density. Namely, we consider logarithmically sparse sets, or sets of positive integers where the number of elements less than n grows as $\log n$ or slower rather than as a fractional power of n, and their sumsets. We define a log-sparse set and the sumset of sets formally as follows.

Definition 1. A subset T of $\mathbb{Z}_{>0}$ is (C-)*log-sparse* if for all positive integers x, $|T \cap [x, 2x)| \leq C$.

Definition 2. Given sets $S_1, S_2, \ldots, S_n \subset \mathbb{Z}_{>0}$, the set $S = S_1 + S_2 + \cdots + S_n$ is the *sumset* of $S_1, S_2, \ldots, S_n$ if S is the set of all positive integers x such that $x = x_1 + x_2 + \cdots + x_n$ for some $x_1 \in S_1, x_2 \in S_2, \ldots, x_n \in S_n$.

We note that the specific constant C in Definition 1 is not crucial as long as it is at least 2.

Note that it is impossible for all integers to be in the sumset of r log-sparse sets for any fixed r, as by a counting argument, such a sumset cannot contain more than $(O(\log N))^r$ integers below N. Here, we consider the maximum possible length of *arithmetic progressions* in sumsets of such logarithmically sparse sets.

Sizes of arithmetic progressions have been well studied in various cases. In particular, a well-known result of Szemerédi [7] shows that any subset A of $\mathbb{Z}_{>0}$ with positive upper density contains arbitrarily long arithmetic progressions. Even for sparser sets, such as the set of primes, it is known that they contain arbitrarily long arithmetic progressions [5], and a well-known conjecture due to Erdős states that as long as $A = \{a_1, a_2, \ldots\}$ satisfies $\sum \frac{1}{a_i} = \infty$, A contains arbitrarily long arithmetic progressions.

Arithmetic progressions in sumsets have also been studied. Bourgain [1] proved that when $|A| = \alpha N$ and $|B| = \beta N$ are subsets of $[N] = \{1, 2, \ldots, N\}$, $A+B$ must contain an arithmetic progression of size at least $\exp(\Omega_{\alpha,\beta}(\log n)^{1/3})$. Bourgain's result was subsequently improved by Green [4] and by Croot, Laba, and Sisak [2] to $A + B$ containing an arithmetic progression of size at least $\exp(\Omega_{\alpha,\beta}(\log n)^{1/2})$.

Sumsets of log-sparse sets do not have positive density, but trivially there do exist sparse sets containing arbitrarily long arithmetic progressions, such as the set $S = \{2^a + b : 1 \leq b \leq a\}$, which contains the k-term progression $2^k+1, 2^k+2, \ldots, 2^k+k$ for all k. This set, of course, is not log-sparse. This raises the following question: does there

exist an integer n and n log-sparse sets $S_1, S_2, \ldots, S_n$ such that the sumset $S_1+S_2+\cdots+S_n$ contains arbitrarily long arithmetic progressions?

In this note, we answer this question in the negative, by showing that for all $n \geq 1$ and all log-sparse sets $S_1, S_2, \ldots, S_n$, the maximum possible size of an arithmetic progression in $S_1 + S_2 + \cdots + S_n$ is at most $n^{(1+o(1))n}$, where the $o(1)$ term tends to 0 as $n \to \infty$ and is independent of the choice of the sets $S_1, S_2, \ldots, S_n$. We also establish a nearly matching lower bound, by proving that for all n, there exist log-sparse sets $S_1, S_2, \ldots, S_n$ whose sum contains an arithmetic progression of length at least $n^{(1-o(1))n}$.

This question is also motivated by the following problem from the 2009 China Team Selection Test: Prove that the set $\{2^a+3^b : a, b \geq 0\}$ has no arithmetic progression of length 40. Note that this set can be written as the sum of two log-sparse sets: the set of powers of 2 and the set of powers of 3, so a direct corollary of our upper bound is that the longest arithmetic progression in $\{2^a + 3^b : a, b \geq 0\}$ is bounded.

Throughout this note, $\log(n)$ always denotes $\log_2(n)$, and $[n]$ denotes the set $\{1, 2, \ldots, n\}$ of the first n positive integers.

2 The upper bound

In this section, we prove an upper bound on the size of the longest arithmetic progression in the sumset of n log-sparse sets.

Theorem 1. *Let $S_1, \ldots S_n$ be C-log-sparse sets, for any fixed $C > 0$, and let T be any arithmetic progression in $S = S_1 + \cdots + S_n$. Then $|T| \leq n^{(1+O(\lg\lg n/\lg n))n}$.*

Proof. For any $x \in T$, we fix a representation $x = x_1 + x_2 + \cdots + x_n$. We will bound the number of elements in T by finding an efficient encoding for an arbitrary $x \in T$. To this end, let $\Delta := \max_{y\in T} y - \min_{y\in T} y$ and let δ be the step-length in T (i. e., $\delta := \Delta/(|T|-1)$). For the fixed representation $x = x_1 + \cdots + x_n$ for any $x \in T$, we say that x_i is large if $x_i > \Delta$, small if $x_i < \delta/2n$, and medium otherwise. Observe that as the sum of all small terms is less than $\delta/2$, $x \in T$ is uniquely determined by the values of all its large and medium terms.

We can encode an arbitrary $x \in T$ as follows. First, we choose which terms are large, medium, and small. There are at most 3^n choices for this. Let a, b, and c denote the chosen number of terms of each respective type.

For the a large terms, we first choose their internal order from largest to smallest, and then choose the value of each of these terms in decreasing order. We claim that having fixed the order, there are at most $O(\log n)$ choices for each term. To see this, we may, without loss of generality, assume that the large terms and internal order are given by $x_1 \geq x_2 \geq \cdots \geq x_a$. Having already chosen $x_1, \ldots, x_{i-1}$ where $i \leq a$, we let

$$M := \max_{y\in T} y - x_1 + \cdots + x_{i-1}.$$

Clearly, we must choose $x_i \le M$. On the other hand, we must also have

$$x_1 + \cdots + x_{i-1} + n \cdot x_i \ge \min_{y \in T} y.$$

Rewriting this, using the definition of Δ, we get $n \cdot x_i + \Delta \ge M$. Since $x_i \ge \Delta$ we can conclude that $(n+1)x_i \ge M$ and so any valid choice for x_i is contained in $S_i \cap [M/(n+1), M]$. Thus by log-sparseness there are at most $O(\log n)$ options, as desired. So in total, we have $O(n \log n)^a$ choices for the large terms.

For each medium term x_i, we know that it is contained in $S_i \cap [\delta/2n, \Delta]$, where the lower and upper bounds differ by a factor $2n(|T| - 1)$. Thus again by log-sparseness, there are at most $O(\log n + \log |T|)$ options for each. So in total $O(\log n + \log |T|)^b$ possibilities.

Combining this, we conclude that

$$|T| \le 3^n \cdot O\big(\max(n \log n, \log n + \log |T|)\big)^n = O(n \log n + \log |T|)^n.$$

But this cannot hold if $|T|$ is too large. Assuming $|T| = (nf(n))^n$ where $f(n) \ge 1$ yields $f(n) \le O(\log n + \log f(n))$, which implies that $f(n) = O(\log n)$, or

$$|T| \le n^{n(1+\lg\lg n/\lg n + O(1/\lg n))},$$

as desired. □

3 The lower bound

In this section, we provide a probabilistic construction of n log-sparse sets whose sumset contains an arithmetic progression of length $n^{(1-o(1))n}$.

Theorem 2. *For any $\varepsilon > 0$, there is some positive $n_0 = n_0(\varepsilon)$ so that for all $n \ge n_0(\varepsilon)$, there exists log-sparse S_i for $1 \le i \le n$ so that the sumset $S = S_1 + S_2 + \cdots + S_n$ contains an arithmetic progression of length at least $n^{(1-\varepsilon)^2 n}$.*

Proof. Begin by splitting the integers from 0 to $(1-\varepsilon)^2 n \log n - 1$ into $(1-\varepsilon)n$ blocks of $(1-\varepsilon)\log n$ consecutive integers. Denote the blocks as $b_1, \ldots, b_m$, where $m = (1-\varepsilon)n$, so

$$b_i = \{(i-1)(1-\varepsilon)\log n, (i-1)(1-\varepsilon)\log n + 1, \ldots, i(1-\varepsilon)\log n - 1\}.$$

For each $i \le m$, let B_i be the set of all positive integers which are sums of distinct powers of 2 with exponents in b_i. Then $|B_i| = 2^{(1-\varepsilon)\log n} - 1 = n^{1-\varepsilon} - 1$. Furthermore, every integer from 0 to $2^{(1-\varepsilon)^2 n \log n} - 1 = n^{(1-\varepsilon)^2 n} - 1$ can be uniquely written as the sum of at most

one element from each B_i, by just looking at the integer's binary representation and splitting it into blocks of size $(1-\varepsilon)\log n$.

We first create sets $S_1, \dots, S_n$, each of size $m+1 = (1-\varepsilon)n+1$. For each $1 \le i \le m$ and each $1 \le j \le n$, we uniformly at random choose one element in B_i to be in S_j. Also, allow each S_j to contain 0. This is not important since at the end we can shift all the elements of each S_i up by 1, and clearly there are at most 2 elements in $[x, n^{1-\varepsilon}x) \supset [x, 2x)$ for each integer x, both before and after the shift. Therefore, we have that each S_i is log-sparse.

We show that with positive probability, $[0, 2^{(1-\varepsilon)^2 n \log n}) \subset S$, which clearly concludes the proof. For an integer $0 \le a < 2^{(1-\varepsilon)^2 n \log n}$, write a as $x_1 + \cdots + x_m$, where $x_i \in B_i \cup \{0\}$. Consider a bipartite graph G with nodes $x_1, \dots, x_m$ and $S_1, \dots, S_n$ such that there is an edge from x_i to S_j if and only if $x_i \in S_j$. Then suppose that for any $k \le m$ and $1 \le i_1 < \cdots < i_k \le m$, there exist k integers $1 \le j_1 < \cdots < j_k \le n$ such that S_{j_r} contains some x_{i_t} for all $r \le k$. This implies that for any subset $\{x_{i_1}, \dots, x_{i_k}\}$, the total number of S_j's that some x_{i_t} is connected to in G is at least k. Therefore, by Hall's marriage theorem, there is some matching from $x_1, \dots, x_m$ to $S_1, \dots, S_n$, i. e., there is a permutation $\sigma : [n] \to [n]$ such that $x_i \in S_{\sigma(i)}$ for all $i \le m$, and thus, $a = x_1 + \cdots + x_n \in S_1 + \cdots + S_n$.

Therefore, it suffices to show that the probability of there existing some $1 \le k \le m$, some subset $\{B_{i_1}, \dots, B_{i_k}\} \subset \{B_1, \dots, B_m\}$, some $x_{i_1} \in B_{i_1}, \dots, x_{i_k} \in B_{i_k}$, and some $\{S_{j_1}, \dots, S_{j_{n-k+1}}\} \subset \{S_1, \dots, S_n\}$ such that no x_{i_t} is contained in any S_{j_r}, is less than 1. This follows from the union bound. We can upper bound the probability by at most

$$\sum_{k=1}^{m} \binom{m}{k} \cdot \left(n^{1-\varepsilon}\right)^k \cdot \binom{n}{n-k+1} \cdot \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{k(n-k+1)}.$$

The $\binom{m}{k}$ comes from choosing the subset $\{B_{i_1}, \dots, B_{i_k}\}$, the $(n^{1-\varepsilon})^k$ comes from choosing each x_{i_r}, the $\binom{n}{n-k+1}$ comes from choosing the S_{j_t}'s and the $(1 - \frac{1}{n^{1-\varepsilon}})^{k(n-k+1)}$ is the probability that every S_{j_r} does not contain any x_{i_t}.

Now, using the fact that $\binom{m}{k} \le m^k \le n^k$ and $\binom{n}{n-k+1} \le n^{k-1} \le n^k$, this sum is at most

$$\sum_{k=1}^{m} \left(n^2 \cdot n^{1-\varepsilon} \cdot \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{n-k+1} \right)^k.$$

But since $k \le (1-\varepsilon)n$, we know that $n - k + 1 \ge \varepsilon n$, so this sum is at most

$$\begin{aligned} \sum_{k=1}^{m} \left(n^2 \cdot n^{1-\varepsilon} \cdot \left(1 - \frac{1}{n^{1-\varepsilon}}\right)^{\varepsilon n} \right)^k &\le \sum_{k=1}^{m} \left(n^2 \cdot n^{1-\varepsilon} \cdot e^{-\varepsilon n^{\varepsilon}} \right)^k \\ &\le \sum_{k=1}^{\infty} \left(n^3 \cdot e^{-\varepsilon n^{\varepsilon}} \right)^k < 1, \end{aligned}$$

assuming n is sufficiently large. This concludes the proof. □

4 Explicit construction

The proof of the lower bound above is probabilistic. It is not difficult to derandomize this proof and give an explicit construction containing a progression of length $2^{\Omega(n \log n)}$ using quadratic polynomials over a finite field. The construction is described in what follows. It is possible to use other known explicit bipartite graphs known as condensers to get similar constructions, but the one below is probably the simplest to describe. See, e. g., [8] and its references for some more sophisticated constructions of condensers.

Let $F = F_q$ be the finite field of size q. Define a bipartite graph $G = G_q$ with classes of vertices A and B as follows. $A = F \times F$ is simply the cartesian product of F with itself. B is the disjoint union of q^2 sets $B_{a,b}$ with $a, b \in F$. Each set $B_{a,b}$ consists of the q polynomials $P_{a,b,c}(x) = ax^2 + bx + c$ where c ranges over all elements of F. Each vertex $P = P_{a,b,c} \in B$ is connected to all vertices $(x, P(x)) \in A$. Therefore, the degree of each vertex in B is exactly q. Note that for every fixed a, b, the sets of neighbors of the q vertices $P_{a,b,c}$ as c ranges over all elements of F are pairwise disjoint, and each vertex of A is connected to exactly one of them.

Proposition 1. *Let $G = G_q$, A, and B be as above. Then for every $x \leq q^2/4$, every set of at most x vertices of B has at least x neighbors in A. Therefore, for each such subset of x vertices in B there is a matching in G saturating it, i. e., each vertex in the subset of B is matched.*

Proof. Every two vertices of B have at most 2 common neighbors in A, since any two distinct quadratic polynomials can be equal on at most 2 points. Therefore, if $x \leq (q + 1)/2$ then for every set $X \subset B$ of size $|X| = x$, the number of its neighbors in A is at least

$$q + (q - 2) + (q - 4) + \cdots + (q - 2x + 2) = x(q - x + 1).$$

This is (much) larger than x for all $x \leq (q+1)/2$. For $x = \lfloor (q+1)/2 \rfloor$, this number exceeds $q^2/4$, implying that every set of at least $\lfloor (q + 1)/2 \rfloor$ vertices of B has more than $q^2/4$ neighbors, completing the proof. □

Returning to our sumset problem, put $n = q^2$. Split the integers in $[0, q^2 \log q/4)$ into $q^2/4$ blocks, each of size $\log_2 q$. Each set S_i contains, as in the probabilistic proof, the integer 0 and one sum of the powers of 2 corresponding to each block. The assignment is determined by the induced subgraph of the graph G_q described above on the classes of vertices A and the union of some $q^2/4$ subsets $B_{a,b}$. The proposition ensures that $S_1 + \ldots + S_n$ contains all integers from 0 to $2^{q^2 \log q/4} = 2^{n \log n/8}$.

Bibliography

[1] J. Bourgain, On arithmetic progressions in sums of sets of integers, in *A Tribute to Paul Erdős*, pp. 105–110, Cambridge Univ. Press, Cambridge, 1990.
[2] E. Croot, I. Łaba and O. Sisask, Arithmetic progressions in sumsets and L_p-almost-periodicity, *Comb. Probab. Comput.*, **22**(3) (2013), 351–365.
[3] R. Graham, *Arithmetic Progressions: from Hilbert to Shelah*, American Mathematical Society, Providence, RI, 1989.
[4] B. Green, Arithmetic progressions in subsets, *Geom. Funct. Anal.*, **12** (2002), 584–597.
[5] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, *Ann. Math.*, **167**(2) (2008), 481–547.
[6] D. Hilbert, Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem), *Math. Ann.*, **67** (1909), 281–300.
[7] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.*, **27** (1975), 199–245.
[8] A. Ta-Shma and C. Umans, Better condensers and new extractors from Parvaresh-Vardy codes, in *Conference on Computational Complexity*, pp. 309–315, 2012.

Michael A. Bennett, Greg Martin, and Kevin O'Bryant

Multidimensional Padé approximation of binomial functions: equalities

Abstract: Let $\omega_0, \dots, \omega_M$ be complex numbers. If $H_0, \dots, H_M$ are polynomials of degree at most $\rho_0, \dots, \rho_M$, and $G(z) = \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$ has a zero at $z = 0$ of maximal order (for the given ω_m, ρ_m), we say that $H_0, \dots, H_M$ are a *multidimensional Padé approximation of binomial functions*, and call G the Padé remainder. We collect here with proof all of the known expressions for G and H_m, including a new one: the Taylor series of G. We also give a new criterion for systems of Padé approximations of binomial functions to be perfect (a specific sort of independence used in applications).

1 Introduction

Fix complex functions $f_0, f_1, \dots, f_M$ (all analytic in a neighborhood of 0) and nonnegative integers $\rho_0, \dots, \rho_M$. The set of functions

$$X := \left\{ \sum_{m=0}^{M} H_m(z) f_m(z) : H_m \in \mathbb{C}[z], \deg(H_m) \le \rho_m \right\}$$

forms a finite dimensional vector space, and the subsets of functions

$$X_s := \{G \in X : \operatorname{ord}_{z=0}(G) \ge s\}$$

with a zero at $z = 0$ of order at least s are subspaces. Trivially, $X_0 \supseteq X_1 \supseteq X_2 \supseteq \cdots$. Let σ be the least integer with X_σ having dimension 0, if such σ exists. Then $X_{\sigma-1}$ has positive dimension, and the functions in $X_{\sigma-1}$ are of particular interest, and are called the *Padé remainders of* $f_0, \dots, f_M$.

The $M = 1$ case is the standard tool in numerical analysis known as Padé approximation [2], which generalizes Taylor series. In particular, if $f_0(z) = -1$ identically, and

Acknowledgement: Michael A. Bennett was supported in part by a Natural Sciences and Engineering Research Council of Canada Discovery grant. Greg Martin was supported in part by a Natural Sciences and Engineering Research Council of Canada Discovery grant. Support for this project was provided by a PSC-CUNY Award, jointly funded by The Professional Staff Congress and The City University of New York.

Michael A. Bennett, Greg Martin, Department of Mathematics, University of British Columbia, Vancouver, British Columbia, Canada, e-mails: bennett@math.ubc.ca, gerg@math.ubc.ca
Kevin O'Bryant, Department of Mathematics, City University of New York, The College of Staten Island and The Graduate Center, Staten Island, NY, USA, e-mail: kevin.obryant@csi.cuny.edu

https://doi.org/10.1515/9783110754216-004

$\rho_1 = 0$, then

$$X = \{-H_0(z) + H_1 \cdot f_1(z) : H_0 \in \mathbb{C}[z], \deg(H_0) \le \rho_0, H_1 \in \mathbb{C}\}.$$

Taking $H_0(z)/H_1$ to be the ρ_0-th Taylor polynomial of $f_1(z)$, we find that the Padé remainders are the constant multiples of the Taylor polynomial remainder. Letting $\rho_1 \ge 0$ leads to rational functions $H_0(z)/H_1(z)$ that approximate $f_1(z)$ at least as well as Taylor polynomials. If f_1 has poles near 0, then this rational approximation is typically much sharper than the Taylor's polynomial approximation.

When $M > 1$, we include the adjective "multidimensional." This setting has not been exploited as systematically as the $M = 1$ case. For a few particular choices of $f_0, \dots, f_M$, there is enough structure that we can work out explicit formulae for the Padé remainders and for the system of Padé approximants, i. e., generating polynomials $H_0, \dots, H_M$. In this paper, we take the binomials $f_m(z) := (1 - z)^{\omega_m}$ for complex numbers $\omega_0, \dots, \omega_M$, no pair of which has an integer difference. The resulting system of equations was studied by Riemann [10], Thue [13], Siegel [11], Mahler [7], Baker [1], Chudnovsky [4], Bennett [3], and many others, and the use of these Padé approximations for Diophantine analysis is known as the method of Thue–Siegel.

We present in this article our exposition of these classic results on multidimensional Padé approximation of binomial functions. We combine, and in some cases, simplify the work of Mahler and Jager [6, 7]. While there are some original results here, e. g., Theorem 4(iv) and some cases of Theorem 6, we see the main value of this work as collating the work of many people over many years with common notation, complete proofs, and specialization to the choice $f_m(z)$. The results presented in this work are equalities, and so as a check against off-by-one errors, one can implement the various forms given and directly check the equations for randomly chosen parameters. We have done so in Mathematica; a notebook containing these calculations is on the arXiv.

The current work focuses on various expressions for the Padé remainders and approximants. In subsequent works, we will provide new bounds, both archimedian and non-archimedian, on the size of the approximant polynomials $H_0, \dots, H_M$ and on the Padé remainder, and will exploit those bounds to give new irrationality measures for some numbers of the form $(a/b)^{s/n}$.

2 Statement of results

Let M be a nonnegative integer. Consider $\vec{\omega} := \langle \omega_0, \omega_1, \dots, \omega_M \rangle$, a vector of $M + 1$ distinct complex numbers, no pair of which has a difference that is an integer, and $\vec{\rho} := \langle \rho_0, \dots, \rho_M \rangle$, a vector of $M + 1$ nonnegative integers (typically not distinct). We index the vectors $\vec{\omega} \in \mathbb{C}^{M+1}, \vec{\rho} \in \mathbb{N}^{M+1}$ with $0, 1, \dots, M$; for example, the 0th coordinate

of $\vec{\rho}$ is ρ_0 and the Mth coordinate is ρ_M. We will only consider M, $\vec{\omega}$, $\vec{\rho}$ satisfying these constraints. Two fundamental parameters are

$$\sigma = \sigma(\vec{\rho}) := \sum_{m=0}^{M} (\rho_m + 1), \quad \text{and} \quad \vec{\rho}! := \prod_{m=0}^{M} \rho_m!.$$

Some notation used in Theorem 1 is both standard and uncommon; we give definitions in the next section. When we add a scalar to a vector, we mean that the scalar is added to each coordinate, such as $\vec{\rho} + 1 = \langle \rho_0 + 1, \rho_2 + 1, \ldots, \rho_M + 1 \rangle$. When we delete the mth coordinate, reducing the length of the vector by 1, we use a "$\star m$" exponent, such as

$$\vec{\omega}^{\star m} = \langle \omega_0, \ldots, \omega_{m-1}, \omega_{m+1}, \ldots, \omega_M \rangle.$$

The standard basis vectors are denoted $\vec{e}_0, \vec{e}_1, \ldots, \vec{e}_M$.

Theorem 1. *Let $\vec{\rho}$ and $\vec{\omega}$ be fixed vectors as above.*

(i) *(Existence) There are polynomials H_m in z of degree at most ρ_m, with at least one H_m not identically 0, and with*

$$G(z) := \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$$

having a zero of order at least $\sigma - 1$ at $z = 0$.

(ii) *(Uniqueness) For such $G(z)$, the function $G(z)$ necessarily has a zero of order exactly $\sigma - 1$ at $z = 0$, and furthermore the polynomials $H_m(z)$ are uniquely determined given the additional constraint that*

$$\lim_{z \to 0} \frac{G(z)}{z^{\sigma-1}} = \frac{1}{(\sigma-1)!}.$$

Each $H_m(z)$ has degree exactly ρ_m. There is no $\alpha \in \mathbb{C}$ with

$$H_0(\alpha) = \cdots = H_M(\alpha) = 0.$$

(iii) *(Domain) $G(z)$ is analytic on $\mathbb{C} \setminus [1, \infty)$.*

Theorem 1 allows us to make the following definition of Padé approximants and remainders.

Definition 2. Let $\vec{\rho}$ and $\vec{\omega}$ be fixed vectors as above. The $M + 1$ *Padé approximants* $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ (with $0 \le m \le M$) are the polynomials with degrees ρ_m, and with *Padé remainder*

$$\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) := \sum_{m=0}^{M} \textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_m}$$

both having a zero of order $\sigma - 1$ at $z = 0$, and satisfying

$$\lim_{z\to 0} \frac{\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho})}{z^{\sigma-1}} = \frac{1}{(\sigma-1)!}.$$

In Proposition 3, we draw attention to some obvious symmetries, immediate from Theorem 1, whose proofs we do not spell out.

Proposition 3 (Permutation and shift symmetry). *If π is any permutation of $0, 1, \ldots, M$, then*

$$\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho}) = \mathrm{Rem}\left(z \;\middle|\; \begin{matrix}\langle \omega_0,\omega_1,\ldots,\omega_M\rangle \\ \langle \rho_0,\rho_1,\ldots,\rho_M\rangle\end{matrix}\right) = \mathrm{Rem}\left(z \;\middle|\; \begin{matrix}\langle \omega_{\pi(0)},\omega_{\pi(1)},\ldots,\omega_{\pi(M)}\rangle \\ \langle \rho_{\pi(0)},\rho_{\pi(1)},\ldots,\rho_{\pi(M)}\rangle\end{matrix}\right)$$

and

$$\mathrm{Poly}_m(z \mid \vec{\omega}\,/\,\vec{\rho}) = \mathrm{Poly}_m\left(z \;\middle|\; \begin{matrix}\langle \omega_0,\omega_1,\ldots,\omega_M\rangle \\ \langle \rho_0,\rho_1,\ldots,\rho_M\rangle\end{matrix}\right) = \mathrm{Poly}_{\pi^{-1}(m)}\left(z \;\middle|\; \begin{matrix}\langle \omega_{\pi(0)},\omega_{\pi(1)},\ldots,\omega_{\pi(M)}\rangle \\ \langle \rho_{\pi(0)},\rho_{\pi(1)},\ldots,\rho_{\pi(M)}\rangle\end{matrix}\right).$$

For any α, we have

$$(1-z)^{\alpha}\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho}) = \mathrm{Rem}(z \mid {}^{\alpha+\vec{\omega}}/_{\vec{\rho}}) \quad \textit{and} \quad \mathrm{Poly}_m(z \mid \vec{\omega}\,/\,\vec{\rho}) = \mathrm{Poly}_m(z \mid {}^{\alpha+\vec{\omega}}/_{\vec{\rho}}).$$

The purpose of the current work is to collect together various explicit formulae for the Padé remainder $\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho})$ and the Padé approximants $\mathrm{Poly}_m(z \mid \vec{\omega}\,/\,\vec{\rho})$, in a common notation, and with complete proofs. Formulae for the Padé remainder are given in Theorem 4, and formulae for the Padé approximants are given in Theorem 5.

Theorem 4 (Forms for the Padé remainder). *The following five expressions give* $\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho})$*:*

(i) *The Padé remainder* $\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho})$ *is given by the iterated integral*

$$\frac{(1-z)^{\omega_0}}{\vec{\rho}!} \int_0^z \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_{M-1}} \mathcal{G}(z, t_1, t_2, \ldots, t_M)\, dt_M \cdots dt_3\, dt_2\, dt_1,$$

where

$$\mathcal{G}(t_0, t_1, \ldots, t_M) = t_M^{\rho_M}\left(\prod_{h=1}^{M}\left(\frac{t_{h-1}-t_h}{1-t_h}\right)^{\rho_{h-1}}\right)\left(\prod_{h=1}^{M}(1-t_h)^{\omega_h-\omega_{h-1}-1}\right).$$

(ii) *The Padé remainder* $\mathrm{Rem}(z \mid \vec{\omega}\,/\,\vec{\rho})$ *is given by the M-dimensional integral*

$$z^{\sigma-1}\frac{(1-z)^{\omega_0}}{\vec{\rho}!} \int_{[0,1]^M} U_M^{-1} \prod_{h=1}^{M} \frac{U_h^{1+\rho_h}}{(1-zU_h)^{1-\omega_h+\omega_{h-1}}}\left(\frac{1-u_h}{1-zU_h}\right)^{\rho_{h-1}} d\vec{u},$$

where $U_m = \prod_{h=1}^{m} u_h$.

(iii) *The Padé remainder* $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is the contour integral*

$$\frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma}(1-z)^{\xi}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\overline{\rho_k+1}}}\,d\xi,$$

where γ *is any simple positively oriented contour enclosing all* σ *of the complex numbers* $\omega_m + r$ $(0 \le m \le M, 0 \le r \le \rho_m)$.

(iv) *The Maclaurin series for* $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is*

$$\sum_{n=0}^{\infty}(-1)^n\sum_{m=0}^{M}\frac{1}{\rho_m!}\sum_{r=0}^{\rho_m}\binom{\rho_m}{r}\frac{(-1)^r(\omega_m+r)^{\underline{n}}}{\prod_{\substack{k=0\\k\neq m}}^{M}(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}\frac{z^n}{n!},$$

which converges for $|z| < 1$.

(v) *Finally,* $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is the special value of Meijer's G-function given by*

$$G^{M+1,0}_{M+1,M+1}\left(1-z\,\middle|\,\begin{matrix}\vec{\omega}+\vec{\rho}+1\\ \vec{\omega}\end{matrix}\right).$$

In addition to the formulae for $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ given in Theorem 4, we note that

$$\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \sum_{m=0}^{M}\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_m},$$

and so any formula for $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ generates a formula for $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$. Theorem 5 gives a number of useful representations of $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$.

Theorem 5 (Forms for the Padé approximants). *The following five expressions give* $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$*:*

(i) *Let* γ_m *be a simple positively oriented contour enclosing all* $\rho_m + 1$ *of the complex numbers* $\omega_m + r$ $(0 \le r \le \rho_m)$ *and none of* $\omega_k + r$ $(0 \le k \le M, k \ne m, 0 \le r \le \rho_k)$. *Then* $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is given by*

$$\frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma_m}(1-z)^{\xi-\omega_m}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\overline{\rho_k+1}}}\,d\xi.$$

(ii) *The Padé approximant* $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is equal to*

$$\frac{1}{\rho_m!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\frac{1}{(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}.$$

(iii) *For $M \geq 1$, the Padé approximant $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ is the M-fold iterated integral*

$$\frac{Q_m}{\vec{\rho}!} \int_{(G)} T_m^{-\omega_m - 1} \left(\prod_{\substack{k=0 \\ k \neq m}}^{M} t_k^{\omega_k} (1 + t_k)^{\rho_k} \right) \left(1 - (-1)^M \frac{1 - z}{T_m} \right)^{\rho_m} d\vec{t},$$

where $\int_{(G)} \cdots d\vec{t}$ integrates each of $t_0, \ldots, t_M$ (except t_m) counterclockwise on the unit circle from $-\pi$ radians to π radians (i. e., the principal value),

$$Q_m := \prod_{\substack{k=0 \\ k \neq m}}^{M} \frac{1}{2i \sin(\pi(\omega_k - \omega_m))}, \quad \text{and} \quad T_m := \prod_{\substack{k=0 \\ k \neq m}}^{M} t_k.$$

(iv) *The Padé approximant is a scaled generalized hypergeometric function:*

$$\frac{1}{\rho_m!} \left(\prod_{\substack{k=0 \\ k \neq m}}^{M} \frac{1}{(\omega_k - \omega_m)^{\overline{\rho_k + 1}}} \right) {}_{M+1}F_M \left[\begin{matrix} \omega_m - \vec{\omega} - \vec{\rho} \\ (1 + \omega_m - \vec{\omega})^{\star m} \end{matrix}; 1 - z \right].$$

(v) *Set $W := W(m, k) = \omega_k - \omega_m$, and define $C_{m,k,r}$ by*

$$C_{m,k,r} := \binom{\rho_k}{r},$$

if $m = k$, by

$$C_{m,k,r} := (-1)^{\rho_k + 1} \binom{r}{\rho_k}^{-1} \frac{\Gamma(r+1)}{\Gamma(r+1-W)} \frac{\Gamma(r - \rho_k - W)}{\Gamma(r - \rho_k + 1)}$$

if $m \neq k$ and $\rho_k < r$, and by

$$C_{m,k,r} := (-1)^r \binom{\rho_k}{r} \frac{\Gamma(r+1)}{\Gamma(r+1-W)} \frac{\Gamma(\rho_k - r + 1)}{\Gamma(\rho_k - r + 1 + W)} \frac{\pi}{\sin(\pi W)}$$

if $m \neq k$ and $\rho_k \geq r$. Then we have

$$\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \frac{1}{\vec{\rho}!} \sum_{r=0}^{\rho_m} (z-1)^r \prod_{k=0}^{M} C_{m,k,r}.$$

Theorem 6 precisely states that notion that the approximants for nearby $\vec{\rho}$ are independent. This property is referred to as "perfect approximation," and relies mostly on $\deg(\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})) = \rho_m$ and $\operatorname{ord}_{z=0}(\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})) = \sigma - 1$. Recall that our $M + 1$ dimensional vectors have coordinates indexed from 0 through M.

Theorem 6 (Approximants are perfect). *Fix $\vec{\rho} \in \mathbb{N}^{M+1}$ and $\vec{\epsilon}_0, \vec{\epsilon}_1, \ldots, \vec{\epsilon}_M \in \mathbb{Z}^{M+1}$ with each $\vec{\rho} + \vec{\epsilon}_k$ having nonnegative coordinates, and denote the jth coordinate of $\vec{\epsilon}_i$ as $\vec{\epsilon}_{i,j}$.*

Let S be maximum of $\sum_{i=0}^{M} \vec{\epsilon}_{i,\beta(i)}$ taken over all permutations β of $0,1,\dots,M$, and let T be the minimum of $\sum_{j=0}^{M} \vec{\epsilon}_{i,j}$ taken over $0 \le i \le M$. Suppose the following two conditions are satisfied:

(i) *There is a unique permutation α of $0,1,\dots,M$ with $S = \sum_{i=0}^{M} \vec{\epsilon}_{i,\alpha(i)}$;*

(ii) $T + M = S$.

Then the $(M+1)\times(M+1)$ matrix whose (k,m) coordinate is the polynomial $\mathrm{POLY}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}+\vec{\epsilon}_k})$ has determinant

$$Cz^{\sigma(\vec{\rho})+T-1},$$

where C does not depend on z.

The most startling aspect of Theorem 6 is that $\vec{\omega}$ plays no role in the hypotheses nor in the conclusion.

We note that (in Theorem 6) with $\vec{\epsilon}_k = \vec{e}_k$ one has $T = 1, S = M+1$, and the conditions in Theorem 6 are satisfied. This recovers a result stated and used by Mahler, Chudnovsky, and Bennett [3, 4, 7]. If one takes $I_k \subseteq \{0,1,\dots,k-1\}$ and $\vec{\epsilon}_k = \vec{e}_k + \sum_{i\in I_k} \vec{e}_i$, one recovers a result of Jager [6]. Our result covers many more examples than we found in the literature, but it is not exhaustive.

3 More notation

We denote the rising and falling factorials as

$$x^{\overline{r}} := x\cdot(x+1)^{\overline{r-1}} = x\cdot(x+1)\cdot(x+2)\cdots(x+r-1),$$
$$x^{\underline{r}} := x\cdot(x-1)^{\underline{r-1}} = x\cdot(x-1)\cdot(x-2)\cdots(x-r+1),$$

for positive integers r, and define $x^{\overline{0}} = x^{\underline{0}} = 1$. We use the following trivial identities without comment (provided $x-r+1 \notin \{0,-1,-2,\dots\}$):

$$x^{\underline{r}} = \frac{\Gamma(x+1)}{\Gamma(x-r+1)}, \quad x^{\underline{r}} = (x-r+1)^{\overline{r}} = (-1)^r(-x)^{\overline{r}},$$

and typically choose to eliminate ratios of Γ functions in preference for the more computationally friendly rising and falling factorials. All of our functions will be analytic in a complex neighborhood of $z = 0$. We use $\deg(f(z))$ to be the degree of f, which is ∞ if f is not a polynomial. We use $\mathrm{ord}_{z=0}(f(z))$ to denote the order of the zero of f at $z = 0$, and we use $O(z^k)$ to denote a function that has a zero at $z = 0$ of order at least k.

We shall briefly encounter the generalized hypergeometric function (defined for $|z| < 1$, $q < p$, and appropriate integers a_i, b_i),

$$ {}_pF_q\left[\begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix}; z\right] = \sum_{n=0}^{\infty} \frac{a_1^{\overline{n}} a_2^{\overline{n}} \cdots a_p^{\overline{n}}}{b_1^{\overline{n}} b_2^{\overline{n}} \cdots b_q^{\overline{n}}} \frac{z^n}{n!}, $$

and also the Meijer G-function [8],

$$ G_{p,q}^{m,n}\left(z \,\middle|\, \begin{matrix} a_1, a_2, \dots, a_p \\ b_1, b_2, \dots, b_q \end{matrix} \right) $$

(defined for natural numbers m, n, p, q, provided $m \le q$ and $n \le p$, although we only encounter it in this work with $n = 0, m = p = q = M + 1$), defined by

$$ \frac{1}{2\pi i} \int_C \frac{\prod_{k=1}^{m} \Gamma(s + b_k) \prod_{k=1}^{n} \Gamma(1 - a_k - s)}{\prod_{k=n+1}^{p} \Gamma(s + a_k) \prod_{k=m+1}^{q} \Gamma(1 - b_k - s)} z^{-s} \, ds, $$

where C is an infinite contour that separates the poles of $\Gamma(1 - a_k - s)$ from those of $\Gamma(b_k + s)$; the particular contour required for convergence varies depending on m, n, p, q, z.

4 Claims and proofs

It is at least plausible that there are polynomials $H_0, \dots, H_M$ with degrees $\rho_0, \dots, \rho_M$ and

$$ G(z) := \sum_{m=0}^{M} H_m(z)(1 - z)^{\omega_m} = \frac{z^{\sigma-1}}{(\sigma - 1)!} + O(z^{\sigma}), \tag{4.1} $$

where $O(z^{\sigma})$ refers to $z \to 0$. After all, the polynomials have a total of σ coefficients, and we may choose them so that $G(z)$ has a zero at $z = 0$ of order $\sigma - 1$, and the first nonzero coefficient in the power series expansion of $G(z)$ is according to our choosing. Establishing this rigorously is the point to our first claims.

In all of the claims in this section, we assume that M is a nonnegative integer, and that $0 \le m \le M$. We assume that $\vec{\rho} = \langle \rho_0, \dots, \rho_M \rangle$ is vector of $M + 1$ nonnegative integers, and that $\vec{\omega} = \langle \omega_0, \dots, \omega_M \rangle$ is a vector of $M + 1$ distinct complex numbers, no two of which have a difference that is an integer. Both $\vec{\rho}$ and $\vec{\omega}$ (and vectors derived from them) are indexed 0 through M.

4.1 Existence and uniqueness

The following claim is used implicitly frequently throughout this work.

Claim 7. *For any polynomials $H_m(z)$ (not all zero), the sum*

$$G(z) := \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$$

is not identically 0.

Proof. Since no two ω_i have difference that is an integer, there is a unique k with

$$\omega_k + \deg(H_k) = \max\{\omega_i + \deg(H_i) : H_i \neq 0\}.$$

Then

$$\lim_{z\to-\infty} \frac{G(z)}{H_k(z)(1-z)^{\omega_k}} = 1 + \sum_{\substack{m=0\\ m\neq k}}^{M} \lim_{z\to-\infty} \frac{H_m(z)(1-z)^{\omega_m}}{H_k(z)(1-z)^{\omega_k}} = 1.$$

Consequently, G cannot be identically 0. □

Claim 8. *There are polynomials $H_0(z), \dots, H_M(z)$ of degrees at most $\rho_0, \dots, \rho_M$, respectively, not all identically 0, such that*

$$\operatorname{ord}_{z=0}\left(\sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}\right) \geq \sigma - 1.$$

Proof. Consider polynomials $H_0(z), \dots, H_M(z)$ of degrees $\rho_0, \dots, \rho_M$ with unknown coefficients, a total of σ unknowns. Recall Newton's binomial theorem: for $|z| < 1$ and any complex ω, we have

$$(1-z)^{\omega} = \sum_{i=0}^{\infty} (-1)^i \frac{\omega^{\underline{i}}}{i!} z^i.$$

Considering the coefficient of z^j, for $0 \leq j \leq \sigma - 2$, on both sides of the desired equality

$$\sum_{m=0}^{M} H_m(z) \sum_{i\geq 0} (-1)^i \frac{\omega_m^{\underline{i}}}{i!} z^i = O(z^{\sigma-1})$$

yields a homogeneous linear equation in the unknowns, a total of $\sigma - 1$ equations. By linear algebra, there is a choice of the σ unknowns, not all zero, which satisfies all of the equations. In other words, there are polynomials $H_0(z), \dots, H_M(z)$ (not all zero)

with degrees at most $\rho_0, \dots, \rho_M$, such that

$$\sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$$

has a zero of order at least $\sigma - 1$ at $z = 0$. □

Claim 8 establishes Theorem 1(i).

The next claim is slightly stronger than the $M = 0$ case of Theorem 1, in that explicit formulae are given, and is used as a base case for subsequent induction arguments.

Claim 9. *The $M = 0$ Padé approximant and remainder are given by the formulae* $\textsc{Poly}_0(z \mid \frac{\langle \omega_0 \rangle}{\langle \rho_0 \rangle}) = \frac{z^{\rho_0}}{\rho_0!}$, *and* $\textsc{Rem}(z \mid \frac{\langle \omega_0 \rangle}{\langle \rho_0 \rangle}) = \frac{z^{\rho_0}}{\rho_0!}(1-z)^{\omega_0}$.

Proof. We need to show that the only nonzero polynomials H_0 with $\mathrm{ord}_{z=0}(H_0(z)(1-z)^{\omega_0}) \geq \sigma - 1$ and degree at most ρ_0 are $H_0(z) = Cz^{\rho_0}$. First, observe that $\sigma = \rho_0 + 1$. As $\mathrm{ord}_{z=0}((1-z)^{\omega_0}) = 0$, we know that $\mathrm{ord}_{z=0}(H_0(z)(1-z)^{\omega_0}) = \mathrm{ord}_{z=0}(H_0)$. That is, H_0 must be a nonzero polynomial with $\mathrm{ord}_{z=0}(H_0) \geq \rho_0$ and $\deg(H_0) \leq \rho_0$. The only candidates are $\textsc{Poly}_0(z \mid \frac{\langle \omega_0 \rangle}{\langle \rho_0 \rangle}) = Cz^{\rho_0}$ and $\textsc{Rem}(z \mid \frac{\langle \omega_0 \rangle}{\langle \rho_0 \rangle}) = Cz^{\rho_0}(1-z)^{\omega_0}$.

Now, observe that

$$C = \lim_{z \to 0} \frac{Cz^{\rho_0}(1-z)^{\omega_0}}{z^{\rho_0}} = \frac{1}{(\sigma-1)!} = \frac{1}{\rho_0!}.$$

Thus Theorem 1(ii) is proved in the $M = 0$ case, and the values of $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ and $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ are as claimed here. □

Claim 10. *If $\deg(H_m(z)) \leq \rho_m$, and some $H_m \neq 0$, then*

$$\mathrm{ord}_{z=0}\left(\sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m} \right) \leq \sigma - 1.$$

Proof. Suppose $M = 0$. With H_0 a nonzero polynomial with degree at most ρ_0, we have

$$\mathrm{ord}_{z=0}(H_0(z)(1-z)^{\omega_0}) = \mathrm{ord}_{z=0}(H_0(z)) \leq \deg(H_0) \leq \rho_0 = \sigma - 1.$$

So, the claim holds for $M = 0$.

Assume the claim is false, and let M be the smallest positive integer for which this claim does not hold, and let ρ_0 correspond to the first counterexample: that is, for any $\vec{\omega}, \vec{\rho}$ that has a smaller M, or the same M but smaller ρ_0, the claim holds. Let

$$G(z) := \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$$

be a counterexample, i. e., $\mathrm{ord}_{z=0}(G) \geq \sigma$. As multiplying by $(1-z)^{-\omega_0}$ does not change $\mathrm{ord}_{z=0}(G(z))$, we may assume that $\omega_0 = 0$.

If $\rho_0 = 0$, so that $H_0(z)$ is a constant, we have

$$\begin{aligned}\frac{d}{dz}G(z) &= \frac{d}{dz}H_0(z) + \sum_{m=1}^{M} \frac{d}{dz}H_m(z)(1-z)^{\omega_m} \\ &= \sum_{m=1}^{M} (H_m'(z)(1-z) - H_m(z)\omega_m)(1-z)^{\omega_m - 1}.\end{aligned}$$

Note that $\deg(H_m'(z)(1-z) - H_m(z)\omega_m) \leq \deg(H_m) \leq \rho_m$, for $1 \leq m \leq M$. Thus $\frac{d}{dz}G(z)$ has a smaller M and the same ρ_m. By assumption on $G(z)$,

$$\mathrm{ord}_{z=0}\left(\frac{d}{dz}G(z)\right) \geq \sigma - 1,$$

but by our assumption of the minimality of $G(z)$, we know that

$$\mathrm{ord}_{z=0}\left(\frac{d}{dz}G(z)\right) \leq \sigma - 2.$$

This contradiction shows that $\rho_0 \neq 0$. But even in the case that $\rho_0 > 0$,

$$\begin{aligned}\frac{d}{dz}G(z) &= H_0'(z) + \sum_{m=1}^{M} \frac{d}{dz}H_m(z)(1-z)^{\omega_m} \\ &= H_0'(z) + \sum_{m=0}^{M} (H_m'(z)(1-z) - H_m(z)\omega_m)(1-z)^{\omega_m - 1}.\end{aligned}$$

As above, our assumption on the minimality of ρ_0, as $\deg(H_0') = \deg(H_0) - 1$, implies a contradiction. □

The proof of the next claim establishes the rest of Theorem 1(ii), and justifies Definition 2.

Claim 11. *Suppose that H_m (with $0 \leq m \leq M$) are polynomials with degree at most ρ_m, and that $G(z) := \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m}$ has a zero of order at least $\sigma - 1$ at $z = 0$. Then $G(z)$ has an order of exactly $\sigma - 1$ at $z = 0$. Suppose further that*

$$\lim_{z \to 0} \frac{G(z)}{z^{\sigma-1}} = \frac{1}{(\sigma-1)!}.$$

Then G and H_m are uniquely determined by these constraints. The polynomial $H_m(z)$ has degree exactly ρ_m, and there is no $\alpha \in \mathbb{C}$ with $H_0(\alpha) = \cdots = H_m(\alpha) = 0$.

Proof. By Claims 8 and 10, we can take $\operatorname{ord}_{z=0}(G(z))$ to be at least $\sigma - 1$, and can never have it be larger than $\sigma - 1$, so there are polynomials H_m with

$$G(z) = \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m} = Cz^{\sigma-1} + O(z^{\sigma}).$$

By multiplying through by a constant, we can take

$$C = \frac{1}{(\sigma-1)!}.$$

If both $G_1(z)$ and $G_2(z)$ have this form, then their difference would have a zero of order greater than $\sigma - 1$, and by Claim 10 this is not possible unless $G_1(z) - G_2(z)$ is identically 0. By Claim 7, however, this is only possible if all of the polynomials are identically 0. That is, only if $G_1(z) = G_2(z)$. Thus, G is uniquely defined and the definition of $\textsc{Rem}(z \mid \vec{\omega}_{\vec{\rho}})$ is justified.

If

$$\textsc{Rem}(z \mid \vec{\omega}_{\vec{\rho}}) = \sum_{m=0}^{M} H_m(z)(1-z)^{\omega_m} = \sum_{m=0}^{M} B_m(z)(1-z)^{\omega_m}$$

for polynomials H_m, B_m of degree at most ρ_m, then

$$0 = \sum_{m=0}^{M} (H_m(z) - B_m(z))(1-z)^{\omega_m}.$$

But by Claim 7, this implies that $H_m(z) = B_m(z)$. Thus H_m is uniquely defined and the definition of $\textsc{Poly}_m(z \mid \vec{\omega}_{\vec{\rho}})$ is justified.

Suppose that $\textsc{Poly}_m(z \mid \vec{\omega}_{\vec{\rho}})$ has degree strictly less than ρ_m, which in particular means that $\rho_m \geq 1$. Let $\vec{e}_m$ be the $M+1$-dimensional unit vector in the mth coordinate direction. Then $\textsc{Rem}(z \mid \vec{\omega}_{\vec{\rho}-\vec{e}_m})$ is a constant multiple of $\textsc{Rem}(z \mid \vec{\omega}_{\vec{\rho}})$, which has a zero of order $\sigma(\vec{\rho}) - 1 > \sigma(\vec{\rho} - \vec{e}_m) - 1$, contradicting Claim 10.

Finally, if $H_m(\alpha) = 0$ for $0 \leq m \leq M$, then $H_m(z)/(z-\alpha)$ are polynomials with degree $\rho_m - 1$, and $G(z)/(z-\alpha)$ has a zero of order $\sigma(\vec{\rho}) - 1$ at $z = 0$, contradicting the uniqueness of G. □

4.2 Respectful differential operators

Claim 12 (Differentiation to reduce ρ). *Define the operators*

$$d_\omega := (1-z)^{\omega+1}\left(\frac{d}{dz}\right)(1-z)^{-\omega}.$$

If $\rho_i > 0$, then d_{ω_i} reduces ρ_i by 1 and increases ω_i by 1, i. e., if $\rho_i > 0$, then

$$d_{\omega_i}\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \textsc{Rem}(z \mid \tfrac{\vec{\omega}+\vec{e}_i}{\vec{\rho}-\vec{e}_i}).$$

If $\rho_i = 0$, then d_{ω_i} eliminates the ith coordinates of $\vec{\omega}$ and $\vec{\rho}$, i. e., if $\rho_i = 0$, then

$$d_{\omega_0}\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \textsc{Rem}(z \mid \tfrac{\vec{\omega}^{*i}}{\vec{\rho}^{*i}}).$$

Consequently, for any ρ_0,

$$(1-z)^{\omega_0+\rho_0+1}\left(\frac{d}{dz}\right)^{\rho_0+1}(1-z)^{-\omega_0}\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \textsc{Rem}\left(z \,\middle|\, \begin{smallmatrix}\langle\omega_1,\dots,\omega_M\rangle\\ \langle\rho_1,\dots,\rho_M\rangle\end{smallmatrix}\right).$$

Proof. As d_ω is linear and

$$\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \textsc{Poly}_i(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_i} + \sum_{\substack{m=0\\ m\neq i}}^{M}\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_m},$$

we can assess the impact of d_{ω_i} on the two pieces separately. First,

$$\begin{aligned} d_{\omega_i}\,\textsc{Poly}_i(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_i} &= (1-z)^{\omega_i+1}\left(\frac{d}{dz}\right)(1-z)^{-\omega_i}\cdot\textsc{Poly}_i(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_i}\\ &= \left(\frac{d}{dz}\,\textsc{Poly}_i(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})\right)(1-z)^{\omega_i+1}. \end{aligned}$$

This is 0 if $\rho_i = 0$, and if $\rho_i > 0$ it has the form $P_i(z)(1-z)^{\omega_i+1}$ with P_i a polynomial of degree $\rho_i - 1$. The other piece is more involved (for the sake of the margins, we let $H(z) := \textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})$ in the following displayed equations):

$$\begin{aligned} &d_{\omega_i}\sum_{\substack{m=0\\ m\neq i}}^{M}\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}})(1-z)^{\omega_m}\\ &= (1-z)^{\omega_i+1}\left(\frac{d}{dz}\right)(1-z)^{-\omega_i}\sum_{\substack{m=0\\ m\neq i}}^{M}H(z)(1-z)^{\omega_m}\\ &= (1-z)^{\omega_i+1}\sum_{\substack{m=0\\ m\neq i}}^{M}\frac{d}{dz}H(z)(1-z)^{\omega_m-\omega_i}\\ &= (1-z)^{\omega_i+1}\sum_{\substack{m=0\\ m\neq i}}^{M}(1-z)^{\omega_m-\omega_i}\frac{d}{dz}H(z) - H(z)(\omega_m-\omega_i)(1-z)^{\omega_m-\omega_i-1} \end{aligned}$$

$$= (1-z)^{\omega_i+1} \sum_{\substack{m=0 \\ m\neq i}}^{M} \left((1-z)\frac{d}{dz}H(z) - (\omega_m - \omega_i)H(z) \right)(1-z)^{\omega_m-\omega_i-1}$$
$$= \sum_{\substack{m=0 \\ m\neq i}}^{M} \left((1-z)\frac{d}{dz}\,\textsc{Poly}_m\big(z \mid \vec{\tilde{\rho}}\big) - (\omega_m - \omega_i)\,\textsc{Poly}_m\big(z \mid \vec{\tilde{\rho}}\big) \right)(1-z)^{\omega_m}.$$

This has the form

$$\sum_{\substack{m=0 \\ m\neq i}}^{M} P_m(z)(1-z)^{\omega_m}$$

with P_m a polynomial of degree at most ρ_m. To wit, $d_{\omega_i}\,\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ has the correct form to be $\textsc{Rem}(z \mid \vec{\tilde{\rho}-\tilde{e}_i})$ if $\rho_i > 0$, and the correct form to be $\textsc{Rem}(z \mid \vec{\tilde{\rho}^{*i}})$ if $\rho_i = 0$. By our earlier uniqueness result, it remains only to check that $d_{\omega_i}\,\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ has a zero (at $z = 0$) of order one less than $\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ and the correct scaling. These are both clear, as $(1-z)^{\omega_i+1}$ and $(1-z)^{-\omega_i}$ have no zero at $z = 0$, and the $\frac{d}{dz}$ reduces the order of the zero by one and the scaling coefficient is multiplied by $\sigma - 1$.

The last sentence of Claim 12 is now immediate, as the product of operators telescopes

$$d_{\omega_0+\rho_0} \cdots d_{\omega_0+1} d_{\omega_0} = (1-z)^{\omega_0+\rho_0+1} \left(\frac{d}{dz} \right)^{\rho_0+1} (1-z)^{-\omega_0}. \qquad \square$$

The previous claim establishes $\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ as the solution of a differential equation (henceforth DE), which we make explicit next. Then we solve the DE to express $\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ as an M-fold iterated integral.

Claim 13. *Let $D_0, \ldots, D_M$ be the operators*

$$D_i := (1-z)^{\omega_i+\rho_i+1} \left(\frac{d}{dz} \right)^{\rho_i+1} (1-z)^{-\omega_i}.$$

With this notation, $G(z) = \textsc{Rem}(z \mid \vec{\tilde{\rho}})$ is the unique analytic solution to the differential equation

$$D_{M-1} \cdots D_1 D_0 G(z) = \frac{z^{\rho_M}}{\rho_M!}(1-z)^{\omega_M}$$

with initial conditions

$$G^{(\sigma-1)}(0) = 1, \quad G^{(m)}(0) = 0, \quad (0 \le m \le \sigma - 2).$$

Proof. That $\textsc{Rem}(z \mid \vec{\tilde{\rho}})$ satisfies the DE is a consequence of Claim 12, and the initial conditions are part of the definition of $\textsc{Rem}(z \mid \vec{\tilde{\rho}})$.

As for uniqueness, suppose that $G(z) = \sum_{i=0}^{\infty} g_i z^i$. Observe that the initial conditions force

$$g_{\sigma-1} = \frac{1}{(\sigma-1)!}, \quad g_i = 0, \quad (0 \le i \le \sigma-2).$$

The DE then forces the value of g_i for $i \ge \sigma$. □

It would be interesting to use the proof of the above claim to work out the full power series of $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$.

4.3 Iterated integrals

Claim 14 is Theorem 4(i), and Claim 15 is Theorem 4(ii).

Claim 14 (Mahler). *We can represent* $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *as an M-fold integral as*

$$\begin{aligned}&\vec{\rho}! \cdot \textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) \\ &\quad = (1-z)^{\omega_0} \int_0^z \int_0^{t_1} \int_0^{t_2} \cdots \int_0^{t_{M-1}} \mathcal{G}(z, t_1, t_2, \ldots, t_M)\, dt_M \cdots dt_3\, dt_2\, dt_1,\end{aligned}$$

where

$$\mathcal{G}(t_0, t_1, \ldots, t_M) = t_M^{\rho_M} \left(\prod_{h=1}^{M} \left(\frac{t_{h-1} - t_h}{1 - t_h} \right)^{\rho_{h-1}} \right) \left(\prod_{h=1}^{M} (1-t_h)^{\omega_h - \omega_{h-1} - 1} \right).$$

Mahler's proof [7]. "This [Claim 13] can easily be brought to the following form." □

This result allows one to produce to an efficient bound for $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$, and is thereby a linchpin in applications. Other authors cite Mahler, or cite authors who cite Mahler. We did not find it easy, and hence indicate in some detail how to arrive at Mahler's conclusion.

Proof. We begin with the differential equation given in Claim 12:

$$(1-z)^{\omega_0+\rho_0+1} \left(\frac{d}{dz} \right)^{\rho_0+1} (1-z)^{-\omega_0}\, \textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \textsc{Rem}\left(z \,\middle|\, \begin{smallmatrix} \langle \omega_1, \ldots, \omega_M \rangle \\ \langle \rho_1, \ldots, \rho_M \rangle \end{smallmatrix} \right).$$

Hence

$$\begin{aligned}\left(\frac{d}{dz} \right)^{\rho_0+1} (1-z)^{-\omega_0}\, \textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) &= (1-z)^{-(\omega_0+\rho_0+1)}\, \textsc{Rem}\left(z \,\middle|\, \begin{smallmatrix} \langle \omega_1, \ldots, \omega_M \rangle \\ \langle \rho_1, \ldots, \rho_M \rangle \end{smallmatrix} \right) \\ &= \textsc{Rem}\left(z \,\middle|\, \begin{smallmatrix} \langle \omega_1, \ldots, \omega_M \rangle - \omega_0 - \rho_0 - 1 \\ \langle \rho_1, \ldots, \rho_M \rangle \end{smallmatrix} \right),\end{aligned}$$

where the second equality follows from Proposition 3.

We observe that

$$\frac{d}{dz}\int_0^z \frac{(z-t)^k}{k!}f(t)\,dt = \begin{cases} f(z), & k=0; \\ \int_0^z \frac{(z-t)^{k-1}}{(k-1)!}f(t)\,dt, & k>0. \end{cases}$$

It then follows by repetition that

$$\left(\frac{d}{dz}\right)^{\rho_0+1}\int_0^z \frac{(z-t)^{\rho_0}}{\rho_0!}f(t)\,dt = f(z).$$

Thus $(1-z)^{-\omega_0}\,\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho})$ and $\int_0^z \frac{(z-t)^{\rho_0}}{\rho_0!}\,\textsc{Rem}\left(t \mid {\langle \omega_1,\ldots,\omega_M\rangle - \omega_0 - \rho_0 - 1 \atop \langle \rho_1,\ldots,\rho_M\rangle}\right)dt$ have the same (ρ_0+1)-th derivative. Therefore, they differ by a polynomial with degree at most ρ_0.

From the definition of $\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho})$, we have

$$\operatorname{ord}_{z=0}\left(\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho})\right) = \rho_0 + 1 + \operatorname{ord}_{z=0}\left(\textsc{Rem}\left(t \,\middle|\, {\langle \omega_1,\ldots,\omega_M\rangle - \omega_0 - \rho_0 - 1 \atop \langle \rho_1,\ldots,\rho_M\rangle}\right)\right),$$

which dictates that the degree-at-most-ρ_0 polynomial is identically 0. We can thus undo the differential operators as

$$(1-z)^{-\omega_0}\,\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho}) = \int_0^z \frac{(z-t)^{\rho_0}}{\rho_0!}\,\textsc{Rem}\left(t \,\middle|\, {\langle \omega_1,\ldots,\omega_M\rangle - \omega_0 - \rho_0 - 1 \atop \langle \rho_1,\ldots,\rho_M\rangle}\right)dt,$$

whence

$$\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho}) = \frac{(1-z)^{\omega_0}}{\rho_0!}\int_0^z (z-t)^{\rho_0}\,\textsc{Rem}\left(t \,\middle|\, {\langle \omega_1,\ldots,\omega_M\rangle - \omega_0 - \rho_0 - 1 \atop \langle \rho_1,\ldots,\rho_M\rangle}\right)dt. \tag{4.2}$$

We wish to apply equation (4.2) inductively to express $\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho})$ as an iterated integral. So that the notation will fit on the page, we define for $1 \le i \le M$,

$$\begin{aligned} S_0 &:= 0, \\ S_i &:= \omega_{i-1} + \rho_{i-1} + 1, \\ G_0(z) &:= \textsc{Rem}(z \mid \vec{\omega}/\vec{\rho}), \\ G_i(z) &:= \textsc{Rem}\left(z \,\middle|\, {\langle \omega_i,\omega_{i+1},\ldots,\omega_M\rangle - S_i \atop \langle \rho_i,\ldots,\rho_M\rangle}\right). \end{aligned}$$

Note that $G_M(z) = \frac{z^{\rho_M}}{\rho_M!}(1-z)^{\omega_M - S_M}$ by Claim 9, while equation (4.2) gives

$$G_i(z) = \frac{(1-z)^{\omega_i - S_i}}{\rho_i!}\int_0^z (z-t)^{\rho_i} G_{i+1}(t)\,dt$$

for $0 \le i < M$. Now, iterating equation (4.2) gives

$$\begin{aligned}
\textsc{Rem}(t_0 \mid \tfrac{\vec{\omega}}{\vec{\rho}}) &= G_0(t_0) \\
&= \frac{(1-t_0)^{\omega_0}}{\rho_0!} \int_0^{t_0} (t_0 - t_1)^{\rho_0} G_1(t_1)\, dt_1 \\
&= \frac{(1-t_0)^{\omega_0}}{\rho_0!\rho_1!} \int_0^{t_0} (t_0 - t_1)^{\rho_0} \cdot (1-t_1)^{\omega_1 - S_1} \int_0^{t_1} (t_1 - t_2)^{\rho_1} G_2(t_2)\, dt_2\, dt_1 \\
&\;\;\vdots \\
&= \frac{(1-t_0)^{\omega_0}}{\rho_0! \cdots \rho_M!} \int_0^{t_0} \int_0^{t_1} \cdots \int_0^{t_{M-1}} \mathcal{G}(t_0, t_1, \ldots, t_M)\, dt_M \cdots dt_2\, dt_1,
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{G}(t_0, t_1, \ldots, t_M) &= \left(\prod_{h=0}^{M-1} (t_h - t_{h+1})^{\rho_h} \right) \left(\prod_{h=1}^{M-1} (1-t_h)^{\omega_h - S_h} \right) t_M^{\rho_M} (1-t_M)^{\omega_M - S_M} \\
&= t_M^{\rho_M} \left(\prod_{h=1}^{M} \left(\frac{t_{h-1} - t_h}{1-t_h} \right)^{\rho_{h-1}} \right) \left(\prod_{h=1}^{M} (1-t_h)^{\omega_h - \omega_{h-1} - 1} \right),
\end{aligned}$$

as claimed. □

Claim 15. *The Padé remainder* $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ *is given by the M-dimensional integral*

$$z^{\sigma-1} \frac{(1-z)^{\omega_0}}{\vec{\rho}!} \int_{[0,1]^M} \left(U_M^{-1} \prod_{h=1}^{M} U_h^{1+\rho_h} \left(\frac{1-u_h}{1-zU_h} \right)^{\rho_{h-1}} (1-zU_h)^{\omega_h - \omega_{h-1} - 1} \right) d\vec{u},$$

where $U_m = \prod_{h=1}^{m} u_h$.

Proof. This follows from the previous claim upon the substitutions

$$t_h = z \prod_{i=1}^{h} u_i = zU_h, \quad dt_M dt_{M-1} \cdots dt_2 dt_1 = z^M \prod_{h=1}^{M-1} U_h d\vec{u},$$

and the obvious algebraic manipulations. □

4.4 Contour integrals and derived expressions

Claim 16 is Theorem 4(iii). Claim 17 is Theorem 5(i). Claim 18 is Theorem 5(ii). Claim 19 is Theorem 5(v). Claim 20 is Theorem 5(iii).

Claim 16. *Let γ be a simple positively oriented contour enclosing all σ of the complex numbers $\omega_m + r$ $(0 \le m \le M, 0 \le r \le \rho_m)$. Then*

$$\mathrm{REM}(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) = \frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma}(1-z)^{\xi}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi.$$

Proof. Set

$$I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) := \frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma}(1-z)^{\xi}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi,$$

and, as in Claim 13,

$$D_i := (1-z)^{\omega_i+\rho_i+1}\left(\frac{d}{dz}\right)^{\rho_i+1}(1-z)^{-\omega_i}.$$

We will show that

$$D_0 I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) = I(z \mid \begin{smallmatrix}\vec{\omega}^{*0}\\ \vec{\omega}^{*0}\end{smallmatrix}).$$

Substituting yields

$$\begin{aligned} D_0 I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) &= (1-z)^{\omega_0+\rho_0+1}\left(\frac{d}{dz}\right)^{\rho_0+1}(1-z)^{-\omega_0} I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) \\ &= (1-z)^{\omega_0+\rho_0+1}\left(\frac{d}{dz}\right)^{\rho_0+1}\frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma}(1-z)^{\xi-\omega_0}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi. \end{aligned}$$

As

$$\left(\frac{d}{dz}\right)^{\rho_0+1}(1-z)^{\xi-\omega_0} = (-1)^{\rho_0+1}(\xi-\omega_0)^{\underline{\rho_0+1}}(1-z)^{\xi-\omega_0-\rho_0-1},$$

differentiating under the integral eliminates the $k=0$ factor in the product, giving

$$D_0 I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) = \frac{(-1)^{\sigma-\rho_0-2}}{2\pi i}\int_{\gamma}(1-z)^{\xi}\prod_{k=1}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi = I(z \mid \begin{smallmatrix}\vec{\omega}^{*0}\\ \vec{\rho}^{*0}\end{smallmatrix}).$$

We iterate, using $D_1, \ldots, D_{M-1}$ successively to remove all but the final coordinates of $\vec{\omega}, \vec{\rho}$, arriving at

$$D_{M-1}\cdots D_1 D_0 I(z \mid \begin{smallmatrix}\vec{\omega}\\ \vec{\rho}\end{smallmatrix}) = I\left(z \,\middle|\, \begin{smallmatrix}\langle\omega_M\rangle\\ \langle\rho_M\rangle\end{smallmatrix}\right) = \frac{(-1)^{\rho_M}}{2\pi i}\int_{\gamma}\frac{(1-z)^{\xi}}{(\xi-\omega_M)^{\underline{\rho_M+1}}}\,d\xi.$$

By partial fractions [5, equation (5.41) in Section 5.3],

$$\frac{(-1)^{\rho_M}}{(\xi-\omega_M)^{\underline{\rho_M+1}}} = \frac{1}{\rho_M!}\sum_{r=0}^{\rho_M}\frac{(-1)^r}{\xi-\omega_M-r}\binom{\rho_M}{r},$$

and with Cauchy's integral formula, we conclude

$$\begin{aligned}
\frac{(-1)^{\rho_M}}{2\pi i}\int_\gamma \frac{(1-z)^\xi}{(\xi-\omega_M)^{\underline{\rho_M+1}}}\,d\xi &= \frac{1}{2\pi i}\int_\gamma \frac{(1-z)^\xi}{\rho_M!}\sum_{r=0}^{\rho_M}\frac{(-1)^r}{\xi-\omega_M-r}\binom{\rho_M}{r}\,d\xi \\
&= \frac{1}{\rho_M!}\sum_{r=0}^{\rho_M}\binom{\rho_M}{r}(-1)^r\frac{1}{2\pi i}\int_\gamma \frac{(1-z)^\xi}{\xi-\omega_M-r}\,d\xi \\
&= \frac{1}{\rho_M!}\sum_{r=0}^{\rho_M}\binom{\rho_M}{r}(-1)^r(1-z)^{\omega_M+r} \\
&= \frac{(1-z)^{\omega_M}}{\rho_M!}\sum_{r=0}^{\rho_M}\binom{\rho_M}{r}(z-1)^r \\
&= \frac{(1-z)^{\omega_M}}{\rho_M!}z^{\rho_M}.
\end{aligned}$$

Thus $I(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ satisfies the DE in Claim 13. We now show that it also satisfies the initial conditions given there, and so by Claim 13 we will have $I(z \mid \frac{\vec{\omega}}{\vec{\rho}}) = \textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$.

As for the initial conditions, it remains to show that $\frac{d^r}{dz^r}I(z \mid \frac{\vec{\omega}}{\vec{\rho}})|_{z=0} = 0$ for $0 \le r \le \sigma-2$, and for $r = \sigma-1$ we get 1. We start with

$$\frac{d^r}{dz^r}I(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \frac{(-1)^{\sigma-1}}{2\pi i}\int_\gamma (-1)^r\xi^{\underline{r}}(1-z)^{\xi-r}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi$$

and evaluating this at $z = 0$ gives

$$\frac{(-1)^{\sigma-r-1}}{2\pi i}\int_\gamma \frac{\xi^{\underline{r}}}{\prod_{k=0}^{M}(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi. \tag{4.3}$$

We may take γ to be a circle with large radius N, where $N > \omega_k + r$ for $0 \le k \le M$ and $0 \le r \le \rho_k + 1$. We now appeal to an argument that has little to do with our particular integrand, and so we generalize. Let $P(\xi) = \prod(\xi - p_j)$ be a monic polynomial of degree r, and let $Q(\xi) = \prod(\xi - q_j)$ be a monic polynomial of degree σ with all of its roots inside $|\xi| = N$. Then, using the substitution $\xi \mapsto N^2/u$, which reverses the

orientation of the contour,

$$\begin{aligned}\frac{1}{2\pi i}\int_{|\xi|=N}\frac{P(\xi)}{Q(\xi)}\,d\xi &= \frac{1}{2\pi i}\int_{|u|=N}\frac{P(N^2/u)}{Q(N^2/u)}\,\frac{N^2}{u^2}\,du\\ &= \frac{N^2}{2\pi i}\int_{|u|=N}\frac{\prod(N^2/u-p_j)}{\prod(N^2/u-q_j)}\,\frac{du}{u^2}\\ &= \frac{N^2}{2\pi i}\int_{|u|=N}\frac{\prod(N^2-up_j)}{\prod(N^2-uq_j)}u^{\sigma-r-2}\,du.\end{aligned}$$

As all the roots of the denominator $\prod(N^2-uq_j)$ are outside the contour, this integral is 0 provided that $\sigma-r-2\geq 0$, i. e., provided $r\leq\sigma-2$. If $r=\sigma-1$, then

$$\begin{aligned}\frac{1}{2\pi i}\int_{|\xi|=N}\frac{P(\xi)}{Q(\xi)}\,d\xi &= \frac{N^2}{2\pi i}\int_{|u|=N}\frac{\prod(N^2-up_j)}{\prod(N^2-uq_j)}u^{\sigma-r-2}\,du\\ &= N^2\cdot\frac{\prod(N^2-0\cdot p_j)}{\prod(N^2-0\cdot q_j)} = N^2\cdot\frac{(N^2)^r}{(N^2)^\sigma} = 1.\end{aligned}$$

□

Claim 17. *Let γ_m be a simple positively oriented contour enclosing all ρ_m+1 of the complex numbers ω_m+r ($0\leq r\leq\rho_m$) and none of ω_k+r ($0\leq k\leq M, k\neq m, 0\leq r\leq\rho_m$). Then*

$$\textsc{Poly}_m\big(z\mid\tfrac{\vec{\omega}}{\vec{\rho}}\big) = \frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma_m}(1-z)^{\xi-\omega_m}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi.$$

Proof. As no pair of the ω_i has a difference that is an integer, the σ numbers ω_m+r, where $0\leq m\leq M, 0\leq r\leq\rho_m$, are distinct. Set

$$\Phi_{r,m}(\xi) := (\xi-\omega_m-r)\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}},$$

where we understand the removable singularity to be removed. Observe that each $\Phi_{r,m}$ has $\sigma-1$ simple poles. We will evaluate $\textsc{Rem}(z\mid\frac{\vec{\omega}}{\vec{\rho}})$ using Cauchy's integral formula. Let $\gamma_{r,m}$ be a simple closed contour enclosing ω_m+r, but none of the roots of $\Phi_{r,m}$. From Claim 16, we find that

$$\begin{aligned}\textsc{Rem}\big(z\mid\tfrac{\vec{\omega}}{\vec{\rho}}\big) &= \frac{(-1)^{\sigma-1}}{2\pi i}\int_{\gamma}(1-z)^{\xi}\prod_{k=0}^{M}\frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\,d\xi\\ &= (-1)^{\sigma-1}\sum_{m=0}^{M}\sum_{r=0}^{\rho_m}\frac{1}{2\pi i}\int_{\gamma_{r,m}}\frac{(1-z)^{\xi}\Phi_{r,m}(\xi)}{\xi-\omega_m-r}\,d\xi\end{aligned}$$

$$= (-1)^{\sigma-1} \sum_{m=0}^{M} \sum_{r=0}^{\rho_m} (1-z)^{\omega_m+r} \Phi_{r,m}(\omega_m + r)$$

$$= \sum_{m=0}^{M} \left((-1)^{\sigma-1} \sum_{r=0}^{\rho_m} (1-z)^{r} \Phi_{r,m}(\omega_m + r) \right) (1-z)^{\omega_m}.$$

We now notice that

$$\operatorname{Poly}_m\left(z \mid \frac{\vec{\omega}}{\vec{\rho}}\right) = (-1)^{\sigma-1} \sum_{r=0}^{\rho_m} (1-z)^{r} \Phi_{r,m}(\omega_m + r), \tag{4.4}$$

as this is a polynomial of the required degree.

Also,

$$\begin{aligned}
&\frac{(-1)^{\sigma-1}}{2\pi i} \int_{\gamma_m} (1-z)^{\xi-\omega_m} \prod_{k=0}^{M} \frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}}\, d\xi \\
&\qquad = (-1)^{\sigma-1} \sum_{r=0}^{\rho_m} \frac{1}{2\pi i} \int_{\gamma_{m,r}} \frac{(1-z)^{\xi-\omega_m} \Phi_{r,m}(\xi)}{\xi-\omega_m-r}\, d\xi \\
&\qquad = (-1)^{\sigma-1} \sum_{r=0}^{\rho_m} (1-z)^{(\omega_m+r)-\omega_m} \Phi_{r,m}(\omega_m + r) \\
&\qquad = (-1)^{\sigma-1} \sum_{r=0}^{\rho_m} (1-z)^{r} \Phi_{r,m}(\omega_m + r) \\
&\qquad = \operatorname{Poly}_m\left(z \mid \frac{\vec{\omega}}{\vec{\rho}}\right).
\end{aligned}$$

□

Claim 18. $\operatorname{Poly}_m\left(z \mid \frac{\vec{\omega}}{\vec{\rho}}\right) = \frac{1}{\rho_m!} \sum_{r=0}^{\rho_m} (z-1)^{r} \binom{\rho_m}{r} \prod_{\substack{k=0 \\ k\neq m}}^{M} \frac{1}{(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}$.

Proof. We continue with the notation of the proof of Claim 17. In particular, we simplify the expression (4.4). Observe that

$$\Phi_{r,m}(\xi) = \left[\prod_{\substack{k=0 \\ k\neq m}}^{M} \frac{1}{(\xi-\omega_k)^{\underline{\rho_k+1}}} \right] \cdot \left[\prod_{\substack{r'=0 \\ r'\neq r}}^{\rho_m} \frac{1}{\xi-\omega_m-r'} \right]$$

so that

$$\Phi_{r,m}(\omega_m + r) = \left[\prod_{\substack{k=0 \\ k\neq m}}^{M} \frac{1}{(\omega_m-\omega_k+r)^{\underline{\rho_k+1}}} \right] \cdot \left[\prod_{\substack{r'=0 \\ r'\neq r}}^{\rho_m} \frac{1}{r-r'} \right].$$

Now,

$$\prod_{\substack{r'=0\\ r'\neq r}}^{\rho_m}(r-r') = (r)(r-1)\cdots(2)(1)(-1)(-2)\cdots(r-\rho_m) = (-1)^{r-\rho_m}r!(\rho_m-r)!,$$

so that

$$\frac{1}{\prod_{\substack{r'=0\\ r'\neq r}}^{\rho_m}(r-r')} = \frac{(-1)^{r-\rho_m}}{\rho_m!}\binom{\rho_m}{r}.$$

Also,

$$\prod_{\substack{k=0\\ k\neq m}}^{M}(\omega_m-\omega_k+r)^{\overline{\rho_k+1}} = \prod_{\substack{k=0\\ k\neq m}}^{M}(-1)^{\rho_k+1}(\omega_k-\omega_m-r)^{\overline{\rho_k+1}} = (-1)^{\sigma-\rho_m-1}\prod_{\substack{k=0\\ k\neq m}}^{M}(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}.$$

We now have $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ as claimed. □

Claim 19. *Set $W := W(m,k) = \omega_k - \omega_m$, and define $C_{m,k,r}$ by*

$$C_{m,k,r} := \binom{\rho_k}{r},$$

if $m = k$, by

$$C_{m,k,r} := (-1)^{\rho_k+1}\binom{r}{\rho_k}^{-1}\frac{\Gamma(r+1)}{\Gamma(r+1-W)}\frac{\Gamma(r-\rho_k-W)}{\Gamma(r-\rho_k+1)}$$

if $m \neq k$ and $\rho_k < r$, and by

$$C_{m,k,r} := (-1)^r\binom{\rho_k}{r}\frac{\Gamma(r+1)}{\Gamma(r+1-W)}\frac{\Gamma(\rho_k-r+1)}{\Gamma(\rho_k-r+1+W)}\frac{\pi}{\sin(\pi W)}$$

if $m \neq k$ and $\rho_k \geq r$. Then we have

$$\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \frac{1}{\vec{\rho}!}\sum_{r=0}^{\rho_m}(z-1)^r\prod_{k=0}^{M}C_{m,k,r}.$$

Proof. We begin from Claim 18, writing W in place of $\omega_k - \omega_m$:

$$\begin{aligned}\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) &= \frac{1}{\rho_m!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\ k\neq m}}^{M}\frac{1}{(W-r)^{\overline{\rho_k+1}}}\\ &= \frac{1}{\vec{\rho}!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\ k\neq m}}^{M}\frac{\rho_k!}{(W-r)^{\overline{\rho_k+1}}}.\end{aligned}$$

If $k \neq m$ and $\rho_k < r$, then

$$\rho_k! = \binom{r}{\rho_k}^{-1} \frac{r!}{(r-\rho_k)!} = \binom{r}{\rho_k}^{-1} \frac{\Gamma(r+1)}{\Gamma(r-\rho_k+1)}$$

and

$$(W-r)^{\overline{\rho_k+1}} = (-1)^{\rho_k+1}(r-W)^{\underline{\rho_k+1}} = (-1)^{\rho_k+1}\frac{\Gamma(r-W+1)}{\Gamma(r-W-\rho_k)}.$$

Combining these,

$$\begin{aligned}\frac{\rho_k!}{(W-r)^{\overline{\rho_k+1}}} &= (-1)^{\rho_k+1}\binom{r}{\rho_k}^{-1}\frac{\Gamma(r+1)}{\Gamma(r-\rho_k+1)}\frac{\Gamma(r-W-\rho_k)}{\Gamma(r-W+1)}\\ &= (-1)^{\rho_k+1}\binom{r}{\rho_k}^{-1}\frac{\Gamma(r+1)}{\Gamma(r+1-W)}\frac{\Gamma(r-\rho_k-W)}{\Gamma(r-\rho_k+1)}\\ &= C_{m,k,r}.\end{aligned}$$

If $k \neq m$ and $\rho_k \geq r$, then

$$\rho_k! = \binom{\rho_k}{r} r!(\rho_k - r)! = \binom{\rho_k}{r}\Gamma(r+1)\Gamma(\rho_k - r + 1)$$

and

$$\begin{aligned}(W-r)^{\overline{\rho_k+1}} &= (W-r)^{\overline{r}} \cdot W^{\overline{\rho_k+1-r}}\\ &= (-1)^r(r-W)^{\underline{r}} \cdot (W+\rho_k-r)^{\underline{\rho_k-r+1}}\\ &= (-1)^r\frac{\Gamma(r-W+1)}{\Gamma(r-W-r+1)} \cdot \frac{\Gamma(\rho_k-r+W+1)}{\Gamma(\rho_k-r+W-(\rho_k-r+1)+1)}\\ &= (-1)^r\frac{\Gamma(r+1-W)\Gamma(\rho_k-r+1+W)}{\Gamma(1-W)\Gamma(W)}.\end{aligned}$$

By Euler's reflection formula for the Gamma function, $\Gamma(1-W)\Gamma(W) = \pi/\sin(\pi W)$. Combining these,

$$\begin{aligned}\frac{\rho_k!}{(W-r)^{\overline{\rho_k+1}}} &= (-1)^r\frac{\pi}{\sin(\pi W)}\frac{\Gamma(r+1)}{\Gamma(r+1-W)}\frac{\Gamma(\rho_k-r+1)}{\Gamma(\rho_k-r+1+W)}\binom{\rho_k}{r}\\ &= C_{m,k,r}.\end{aligned}$$

This concludes the proof. □

Claim 20. *For $M \geq 1$, we can represent $\operatorname{Poly}_m(z \mid \vec{\tilde{\rho}}^{\,\vec{\omega}})$ as an M-fold iterated (principal value) contour integral as*

$$\operatorname{Poly}_m(z \mid \vec{\tilde{\rho}}^{\,\vec{\omega}}) = \frac{Q_m}{\vec{\rho}!}\int_{(G)} T_m^{-\omega_m-1}\left(\prod_{\substack{k=0\\k\neq m}}^{M} t_k^{\omega_k}(1+t_k)^{\rho_k}\right)\left(1-(-1)^M\frac{1-z}{T_m}\right)^{\rho_m} d\vec{t},$$

where

$$Q_m := \prod_{\substack{k=0\\k\neq m}}^{M} \frac{1}{2i\sin(\pi(\omega_k-\omega_m))}$$

$$T_m := \prod_{\substack{k=0\\k\neq m}}^{M} t_k$$

$$\int_{(G)} := \int_{|t_0|=1} \cdots \int_{|t_{m-1}|=1} \int_{|t_{m+1}|=1} \cdot \int_{|t_M|=1}$$

$$d\vec{t} := dt_M \cdots dt_{m+1}\, dt_{m-1} \cdots t_1.$$

Proof. By induction and integration-by-parts, we notice that

$$\text{P.V.} \int_{|t|=1} t^{x-1}(1+t)^{\rho}\, dt = \int_{-\pi}^{\pi} e^{i(x-1)t}\big(1+e^{it}\big)^{\rho} ie^{it} dt = \frac{2i\sin(\pi x)\rho!}{x^{\overline{\rho+1}}}, \tag{4.5}$$

provided that ρ is a nonnegative integer and $x \in \mathbb{C}\setminus\{0,-1,-2,\dots\}$. In this claim and its proof, all integrals are understood to be principal values.

Beginning with Claim 18, we may write $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ as

$$\begin{aligned}
&\textsc{Poly}_m\big(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}\big)\\
&\quad= \frac{1}{\rho_m!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\frac{1}{(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}\\
&\quad= \frac{1}{\vec{\rho}!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\frac{1}{2i\sin(\pi(\omega_k-\omega_m-r))}\frac{2i\sin(\pi(\omega_k-\omega_m-r))\rho_k!}{(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}.
\end{aligned}$$

We now use equation (4.5) to continue

$$\begin{aligned}
&\textsc{Poly}_m\big(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}\big)\\
&\quad= \frac{1}{\vec{\rho}!}\sum_{r=0}^{\rho_m}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}(-1)^r Q_m\int_{|t_k|=1} t_k^{\omega_k-\omega_m-r-1}(1+t_k)^{\rho_k}\, dt_k\\
&\quad= \frac{Q_m}{\vec{\rho}!}\sum_{r=0}^{\rho_m}(-1)^{rM}(z-1)^r\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\int_{|t_k|=1} t_k^{\omega_k-\omega_m-r-1}(1+t_k)^{\rho_k}\, dt_k
\end{aligned}$$

Compressing the product of integrals using the $\int_{(G)}$ notation, we continue with

$$\begin{aligned}
&\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) \\
&= \frac{Q_m}{\vec{\rho}!} \int_{(G)} \sum_{r=0}^{\rho_m} (-1)^{rM} (z-1)^r \binom{\rho_m}{r} \prod_{\substack{k=0 \\ k\neq m}}^{M} t_k^{\omega_k - \omega_m - r - 1} (1+t_k)^{\rho_k} \, d\vec{t} \\
&= \frac{Q_m}{\vec{\rho}!} \int_{(G)} \sum_{r=0}^{\rho_m} (-1)^{rM} (z-1)^r \binom{\rho_m}{r} T_m^{-r} T_m^{-\omega_m - 1} \prod_{\substack{k=0 \\ k\neq m}}^{M} t_k^{\omega_k} (1+t_k)^{\rho_k} \, d\vec{t} \\
&= \frac{Q_m}{\vec{\rho}!} \int_{(G)} T_m^{-\omega_m - 1} \left(\prod_{\substack{k=0 \\ k\neq m}}^{M} t_k^{\omega_k} (1+t_k)^{\rho_k} \right) \sum_{r=0}^{\rho_m} \binom{\rho_m}{r} \left((-1)^M \frac{z-1}{T_m} \right)^r d\vec{t} \\
&= \frac{Q_m}{\vec{\rho}!} \int_{(G)} T_m^{-\omega_m - 1} \left(\prod_{\substack{k=0 \\ k\neq m}}^{M} t_k^{\omega_k} (1+t_k)^{\rho_k} \right) \left(1 - (-1)^M \frac{1-z}{T_m} \right)^{\rho_m} d\vec{t},
\end{aligned}$$

as asserted in the claim. □

4.5 Hypergeometric functions

Claim 21 below is Theorem 4(v), and Claim 23 is Theorem 5(iv).

The Meijer G-function [8] is defined for natural numbers m, n, p, q, provided $m \le q$ and $n \le p$, although we only encounter it here with $n = 0, m = p = q = M + 1$. It is denoted

$$G_{p,q}^{m,n}\left(z \,\middle|\, \begin{matrix} a_1, a_2, \ldots, a_p \\ b_1, b_2, \ldots, b_q \end{matrix} \right)$$

and defined as

$$\frac{1}{2\pi i} \int_C \frac{\prod_{k=1}^{m} \Gamma(s + b_k) \prod_{k=1}^{n} \Gamma(1 - a_k - s)}{\prod_{k=n+1}^{p} \Gamma(s + a_k) \prod_{k=m+1}^{q} \Gamma(1 - b_k - s)} z^{-s} \, ds,$$

where C is a particular infinite contour that separates the poles of $\Gamma(1 - a_k - s)$ from those of $\Gamma(b_k + s)$; the particular contour required for convergence varies depending on m, n, p, q, z.

Claim 21. $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$, *when* $|z| < 1$ *and* $|1-z| < 1$, *is a special value of Meijer's G-function,*

$$\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = G_{M+1,M+1}^{M+1,0}\left(1 - z \,\middle|\, \begin{matrix} \vec{\omega} + \vec{\rho} + 1 \\ \vec{\omega} \end{matrix} \right).$$

Sketch of proof. With $m = p = q = M + 1, n = 0$, we see that

$$\prod_{k=1}^{n} \Gamma(1 - a_k - s) = \prod_{k=m+1}^{q} \Gamma(1 - b_k - s) = 1.$$

Further, with $a_{k+1} = \omega_k + \rho_k + 1, b_{k+1} = \omega_k$,

$$\frac{\prod_{k=1}^{m} \Gamma(s + b_k)}{\prod_{k=n+1}^{p} \Gamma(s + a_k)} = \prod_{k=0}^{M} \frac{\Gamma(s + \omega_k)}{\Gamma(s + \omega_k + \rho_k + 1)} = \prod_{k=0}^{M} \frac{1}{(s + \omega_k)^{\overline{\rho_k + 1}}}.$$

We now have

$$\begin{aligned} G_{M+1,M+1}^{M+1,0}\left(1 - z \,\middle|\, \begin{matrix} \vec{\omega} + \vec{\rho} + 1 \\ \vec{\omega} \end{matrix}\right) &= \frac{1}{2\pi i} \int_C (1 - z)^{-s} \prod_{k=0}^{M} \frac{1}{(s + \omega_k)^{\overline{\rho_k + 1}}}\, ds \\ &= \frac{1}{2\pi i} \int_C (1 - z)^{\xi} \prod_{k=0}^{M} \frac{(-1)^{\rho_k + 1}}{(\xi - \omega_k)^{\overline{\rho_k + 1}}}\, (-d\xi) \\ &= \frac{(-1)^{\sigma - 1}}{2\pi i} \int_C (1 - z)^{\xi} \prod_{k=0}^{M} \frac{1}{(\xi - \omega_k)^{\overline{\rho_k + 1}}}\, d\xi \\ &= \textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}). \end{aligned}$$

Admittedly, we have played fast-and-loose with the contour and, therefore, the conditions $|z| < 1$ and $|1 - z| < 1$ are not explained. □

Theorem 22 (Slater's theorem [12]). *Provided that $a_j - b_h$ is not a positive integer (with $j \le n, h \le m$), and $b_j - b_k$ is not an integer (with $1 \le j < k \le q$), and $0 < |z| < 1$,*

$$\begin{aligned} &G_{p,q}^{m,n}\left(z \,\middle|\, \begin{matrix} a_1, a_2, \ldots, a_p \\ b_1, b_2, \ldots, b_q \end{matrix}\right) \\ &\quad = \sum_{h=1}^{m} \frac{\prod_{\substack{k=1 \\ k \ne h}}^{m} \Gamma(b_k - b_h) \prod_{k=1}^{n} \Gamma(1 + b_h - a_k)}{\prod_{k=m+1}^{q} \Gamma(1 + b_h - b_k) \prod_{k=n+1}^{p} \Gamma(a_k - b_h)} {}_pF_{q-1}\left[\begin{matrix} \vec{a} \\ \vec{b} \end{matrix}; (-1)^{m+n-p} z\right] z^{b_h}, \end{aligned}$$

where $\vec{a_h} = \langle 1 + b_h - a_1, \ldots, 1 + b_h - a_p \rangle$ and $\vec{b_h} = \langle 1 + b_h - b_1, \ldots, 1 + b_h - b_q \rangle$ (with the $1 + b_h - b_h$ term omitted).

Claim 23. *The Padé approximant $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ is associated with a generalized hypergeometric function by*

$$\textsc{Poly}_m(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) = \frac{1}{\rho_m!} \left(\prod_{\substack{k=0 \\ k \ne m}}^{M} \frac{1}{(\omega_k - \omega_m)^{\overline{\rho_k + 1}}} \right) {}_{M+1}F_M\left[\begin{matrix} \omega_m - \vec{\omega} - \vec{\rho} \\ (1 + \omega_m - \vec{\omega})^{\star m} \end{matrix}; 1 - z\right].$$

Proof. Using Claim 21 and Theorem 22, we may write $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ as

$$\sum_{m=0}^{M} \frac{\prod_{\substack{k=0\\k\neq m}}^{M} \Gamma(\omega_k - \omega_m)}{\prod_{k=0}^{M} \Gamma(\omega_k + \rho_k + 1 - \omega_m)} {}_{M+1}F_M\left[\begin{matrix} 1+\omega_m-\vec{\omega}-\vec{\rho}-1 \\ (1+\omega_m-\vec{\omega})^{\star m}\end{matrix}; 1-z\right](1-z)^{\omega_m},$$

which we manipulate into the form

$$\sum_{m=0}^{M} \frac{1}{\rho_m!}\left(\prod_{\substack{k=0\\k\neq m}}^{M} \frac{1}{(\omega_k-\omega_m)^{\overline{\rho_k+1}}}\right) {}_{M+1}F_M\left[\begin{matrix} \omega_m-\vec{\omega}-\vec{\rho} \\ (1+\omega_m-\vec{\omega})^{\star m}\end{matrix}; 1-z\right](1-z)^{\omega_m}.$$

One coordinate of $\omega_m - \vec{\omega} - \vec{\rho}$ is $-\rho_m$, a nonpositive integer. Consequently,

$$\frac{1}{\rho_m!}\left(\prod_{\substack{k=0\\k\neq m}}^{M} \frac{1}{(\omega_k-\omega_m)^{\overline{\rho_k+1}}}\right) {}_{M+1}F_M\left[\begin{matrix} \omega_m-\vec{\omega}-\vec{\rho} \\ (1+\omega_m-\vec{\omega})^{\star m}\end{matrix}; 1-z\right]$$

is a polynomial with degree at most ρ_m. Therefore, it must be $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$. □

4.6 Power series

Claim 24 is Theorem 4(iv).

Claim 24. *Let g_n be the coefficients in the power series expansion of $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ at $z = 0$, i. e., $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}}) = \sum_{n=0}^{\infty} g_n \frac{z^n}{n!}$. Then for $n \geq 0$, we have*

$$g_n = (-1)^n \sum_{m=0}^{M} \frac{1}{\rho_m!} \sum_{r=0}^{\rho_m} \binom{\rho_m}{r} \frac{(-1)^r (\omega_m + r)^{\underline{n}}}{\prod_{\substack{k=0\\k\neq m}}^{M} (\omega_k - \omega_m - r)^{\overline{\rho_k+1}}}.$$

In particular, $g_n = 0$ for $0 \leq n \leq \sigma - 2$ and $g_{\sigma-1} = 1$.

Proof. We begin with the contour integral representation of $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ given in Claim 16, replace $(1-z)^{\xi}$ with its power series, and then integrate term by term, obtaining

$$\begin{aligned}
\textsc{Rem}(z \mid \tfrac{\vec{\omega}}{\vec{\rho}}) &= \frac{(-1)^{\sigma-1}}{2\pi i} \int_{\gamma} (1-z)^{\xi} \prod_{k=0}^{M} \frac{1}{(\xi-\omega_k)^{\overline{\rho_k+1}}}\, d\xi \\
&= \frac{(-1)^{\sigma-1}}{2\pi i} \int_{\gamma} \left(\sum_{n=0}^{\infty} (-1)^n \xi^{\underline{n}} \frac{z^n}{n!}\right) \prod_{k=0}^{M} \frac{1}{(\xi-\omega_k)^{\overline{\rho_k+1}}}\, d\xi \\
&= \sum_{n=0}^{\infty} (-1)^n \left(\frac{1}{2\pi i} \int_{\gamma} \frac{(-1)^{\sigma-1} \xi^{\underline{n}}}{\prod_{k=0}^{M} (\xi-\omega_k)^{\overline{\rho_k+1}}}\, d\xi\right) \frac{z^n}{n!}.
\end{aligned}$$

Continuing as in the proof of Claims 17 and 18, we see that

$$
\begin{aligned}
g_n &= (-1)^n\left(\frac{1}{2\pi i}\int_\gamma \frac{(-1)^{\sigma-1}\xi^{\underline{n}}}{\prod_{k=0}^{M}(\xi-\omega_k)^{\overline{\rho_k+1}}}\,d\xi\right) \qquad (4.6)\\
&= (-1)^n\sum_{m=0}^{M}\sum_{r=0}^{\rho_m}(-1)^{\sigma-1}(\omega_m+r)^{\underline{n}}\Phi_{r,m}(\omega_m+r)\\
&= (-1)^n\sum_{m=0}^{M}\sum_{r=0}^{\rho_m}(-1)^{\sigma-1}(\omega_m+r)^{\underline{n}}\frac{(-1)^{r-\rho_m}}{\rho_m!}\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\frac{1}{(\omega_m+r-\omega_k)^{\overline{\rho_k+1}}}\\
&= (-1)^n\sum_{m=0}^{M}\sum_{r=0}^{\rho_m}(-1)^{\sigma-1}(\omega_m+r)^{\underline{n}}\frac{(-1)^{r-\rho_m}}{\rho_m!}\binom{\rho_m}{r}\prod_{\substack{k=0\\k\neq m}}^{M}\frac{(-1)^{\rho_k+1}}{(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}\\
&= (-1)^n\sum_{m=0}^{M}\frac{1}{\rho_m!}\sum_{r=0}^{\rho_m}\binom{\rho_m}{r}\frac{(-1)^r(\omega_m+r)^{\underline{n}}}{\prod_{\substack{k=0\\k\neq m}}^{M}(\omega_k-\omega_m-r)^{\overline{\rho_k+1}}}.
\end{aligned}
$$

That $g_n = 0$ for $0 \le n \le \sigma-2$ and $g_{\sigma-1} = 1$ follow from the definition of $\textsc{Rem}(z \mid \vec{\omega}/\vec{\rho})$. Alternatively, the expression on line (4.6) is shown directly to have these values in the proof of Claim 16, beginning with equation (4.3). □

4.7 Perfection

We remind our reader that our vectors are indexed from 0, so that the jth coordinate of $\langle\rho_0,\dots,\rho_M\rangle$ is ρ_j. The coordinates of the $(M+1)\times(M+1)$ matrix $\mathbf{H}$ in the next claim is indexed in the same manner.

Claim 25 is Theorem 6.

Claim 25. *Fix $\vec{\rho}\in\mathbb{N}^{M+1}$ and $\vec{\epsilon}_0,\vec{\epsilon}_1,\dots,\vec{\epsilon}_M\in\mathbb{Z}^{M+1}$ with each $\vec{\rho}+\vec{\epsilon}_k$ having nonnegative coordinates, and denote the jth coordinate of $\vec{\epsilon}_i$ as $\vec{\epsilon}_{i,j}$. Let S be maximum of $\sum_{i=0}^{M}\vec{\epsilon}_{i,\beta(i)}$ taken over all permutations β of $0,1,\dots,M$, and let T be the minimum of $\sum_{j=0}^{M}\vec{\epsilon}_{i,j}$ taken over $0\le i\le M$. Suppose the following two conditions are satisfied:*

1. *There is a unique permutation α of $0,1,\dots,M$ with $S=\sum_{i=0}^{M}\vec{\epsilon}_{i,\alpha(i)}$;*
2. *$T+M=S$.*

Then the $(M+1)\times(M+1)$ matrix $\mathbf{H}$, whose (k,m) coordinate is the polynomial $\textsc{Poly}_m(z\mid\vec{\omega}/\vec{\rho}+\vec{\epsilon}_k)$, has determinant

$$Cz^{\sigma(\vec{\rho})+T-1},$$

where C is nonzero and does not depend on z.

Proof. The determinant of $\mathbf{H}$, by the familiar permutation expansion, is

$$\det(\mathbf{H}) = \sum_{\beta\in\Sigma_{[0,M]}} (-1)^{\operatorname{sgn}(\beta)} \prod_{k=0}^{M} \textsc{Poly}_{\beta(k)}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_k}),$$

which is clearly a polynomial. Notice that

$$\begin{aligned}\deg\left(\prod_{k=0}^{M} \textsc{Poly}_{\beta(k)}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_k})\right) &= \sum_{k=0}^{M} \deg\big(\textsc{Poly}_{\beta(k)}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_k})\big)\\ &= \sum_{k=0}^{M} (\rho_{\beta(k)} + \vec{\epsilon}_{k,\beta(k)}) \le \sigma(\vec{\rho}) - (M+1) + S,\end{aligned}$$

with equality achieved for (and only for) $\beta = \alpha$. Consequently,

$$\deg\big(\det(\mathbf{H})\big) = \sigma(\vec{\rho}) - M - 1 + S = \sigma(\vec{\rho}) + T - 1,$$

and in particular $\det(\mathbf{H})$ is not identically 0.

Let $\vec{v}$ be the column vector $\langle (1-z)^{\omega_0}, (1-z)^{\omega_1}, \dots, (1-z)^{\omega_M}\rangle^T$. By definition $\mathbf{H}\vec{v}$ is a column of $M+1$ functions of z: in row k it is $\textsc{Rem}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_k})$, which has a zero of order $\sigma(\vec{\rho}+\vec{\epsilon}_k) - 1 = \sigma(\vec{\rho}) + (\sum_{j=0}^{M} \vec{\epsilon}_{k,j}) - 1 \ge \sigma(\vec{\rho}) + T - 1$. Now multiply $\mathbf{H}\vec{v}$ by the adjoint of $\mathbf{H}$, which is also a matrix of polynomials. We have

$$\begin{aligned}\det(\mathbf{H})\vec{v} &= \operatorname{adj}(\mathbf{H})\mathbf{H}\vec{v}\\ &= \operatorname{adj}(\mathbf{H}) \begin{pmatrix} \textsc{Rem}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_0}) \\ \textsc{Rem}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_1}) \\ \vdots \\ \textsc{Rem}(z \mid {}^{\vec{\omega}}_{\vec{\rho}+\vec{\epsilon}_M}) \end{pmatrix}\\ &= \operatorname{adj}(\mathbf{H}) \left(z^{\sigma(\vec{\rho})+T-1} \sum_{n=0}^{\infty} \vec{v}_n z^n \right)\\ &= z^{\sigma(\vec{\rho})+T-1} \sum_{n=0}^{\infty} (\operatorname{adj}(\mathbf{H})\vec{v}_n) z^n\end{aligned}$$

for some column vectors $\vec{v}_0 \neq \vec{0}, \vec{v}_1, \dots$. That $\vec{v}_0 \neq \vec{0}$ follows from the definition of T. Each coordinate of $\det(\mathbf{H})\vec{v}$ has the form $\det(\mathbf{H})(1-z)^{\omega}$, and so has a zero at $z = 0$ of order at most $\deg(\det(\mathbf{H})) = \sigma(\vec{\rho}) - M - 1 + S$. By the above displayed equations, each coordinate of $\det(\mathbf{H})\vec{v}$ has a zero at $z = 0$ of order at least $\sigma(\vec{\rho}) + T - 1$ with equality for some coordinate. But $T + M = S$, by hypothesis, so that $\det(\mathbf{H})$ is a polynomial whose degree coincides with the order of its zero at $z = 0$. Therefore,

$$\det(\mathbf{H}) = Cz^{\sigma(\vec{\rho})+T-1} = Cz^{\sigma(\vec{\rho})+S-M-1},$$

as claimed. The constant C is nonzero as $\det(\mathbf{H})$ is not identically 0. □

5 Opportunities for further work

1. Is there a nice iterated integral representation of $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$ without contours, similar to the representation in Theorem 4(i) for $\textsc{Rem}(z \mid \frac{\vec{\omega}}{\vec{\rho}})$?
2. For fixed $\vec{\omega}$, which degree vectors $\vec{\rho}^{(0)}, \vec{\rho}^{(1)}, \ldots, \vec{\rho}^{(M)}$ lead to a perfect system? There seems to be some geometry involved. That is, a modest amount of computation suggests that for each M there is B such that if any coordinate of any $\vec{\epsilon}_k - \vec{\epsilon}_j$ is not between $-B$ and B, then the resulting system is not perfect for any ρ (the determinant of $\mathbf{H}$ does not have the form Cz^n).
3. What is the value of C in Theorem 6?
4. What is the nice power series expression for $\textsc{Poly}_m(z \mid \frac{\vec{\omega}}{\vec{\rho}})$? For $M = 1$, this is an important part of the best explicit irrationality measure for $2^{1/3}$.

Bibliography

[1] A. Baker, Rational approximations to $\sqrt[3]{2}$ and other algebraic numbers, *Q. J. Math. Oxf. Ser. (2)*, **15** (1964), 375–383. https://doi.org/10.1093/qmath/15.1.375.

[2] G. A. Baker Jr. and P. Graves-Morris, *Padé Approximants*, 2nd ed., Encyclopedia of Mathematics and its Applications, vol. **59**, Cambridge University Press, Cambridge, 1996, xiv+746 pp. https://doi.org/10.1017/CBO9780511530074.

[3] M. A. Bennett, Rational approximation to algebraic numbers of small height: the Diophantine equation $|ax^n - by^n| = 1$, *J. Reine Angew. Math.*, **535** (2001), 1–49. https://doi.org/10.1515/crll.2001.044.

[4] G. V. Chudnovsky, On the method of Thue-Siegel, *Ann. Math. (2)*, **117**(2) (1983), 325–382. https://doi.org/10.2307/2007080.

[5] R. L. Graham, D. E. Knuth and O. Patashnik, Concrete Mathematics, in *A foundation for computer science*, 2nd ed., Addison-Wesley Publishing Company, Reading, MA, 1994, xiv+657 pp.

[6] H. Jager, A multidimensional generalization of the Padé table. I, II, III, IV, V, VI, *Ned. Akad. Wet. Proc. Ser. A*, **26** (1964), 193–249.

[7] K. Mahler, Ein Beweis des Thue-Siegelschen Satzes über die Approximation algebraischer Zahlen für binomische Gleichungen, language=German, *Math. Ann.*, **105**(1) (1931), 267–276. https://doi.org/10.1007/BF01455819. English translation by Karl Levy at https://arxiv.org/abs/1507.01447.

[8] NIST Digital Library of Mathematical Functions, http://dlmf.nist.gov/. Online companion to [9], Release 1.0.8 of 2014-04-25.

[9] F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark, *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, NY, 2010. Print companion to [8].

[10] B. Riemann, *Oeuvres Mathimatiques*, Albert Blanchard, Paris, 1968.

[11] C. L. Siegel, Approximation algebraische zahlen, *Math. Z.*, **10** (1921), 127–213.

[12] L. J. Slater, *Generalized Hypergeometric Functions*, Cambridge University Press, Cambridge, 1966, xiii+273 pp.

[13] A. Thue, Bemerkung über gewisse Nährungsbrüche algebraischer Zahlen. *Skrifter udigvne af Videnskabsselskabet i Christiania* (1908).

Lars Blomberg, S. R. Shannon, and N. J. A. Sloane

Graphical enumeration and stained glass windows, 1: rectangular grids

Dedicated to the memory of Ronald Lewis Graham (1935–2020)

Abstract: This is a survey of enumeration problems arising from the study of planar graphs formed when the edges of a polygon are marked with evenly spaced points and every pair of points is joined by a line. A few of these problems have been solved, a classical example being the graph K_n formed when all pairs of vertices of a regular n-gon are joined by chords, which was analyzed by Poonen and Rubinstein in 1998. Most of these problems are unsolved, however, and this two-part article provides data from a number of such problems as well as colored illustrations, which are often reminiscent of stained glass windows. The polygons considered include rectangles, hollow rectangles (or frames), triangles, pentagons, pentagrams, crosses, etc., as well as figures formed by drawing semicircles joining equally-spaced points on a line. Part 1 discusses planar graphs that are based on rectangular grids.

1 Introduction

In 1998, Poonen and Rubinstein [17] (see also [22]) solved the problem of finding the numbers of intersection points and cells in a regular drawing of the complete graph K_n, and in 2009–2010 Legendre [10] and Griffiths [7] solved a similar problem for the complete bipartite graph $K_{n,n}$. Stated another way, [17] analyzes the planar graph formed by joining all pairs of vertices of a regular n-gon, while [7, 10] analyze the graph formed by taking a row of $n-1$ identical squares and drawing lines between every pair of boundary nodes.

One motivation for the present work was to see if these investigations could be extended to graphs formed from other structures, such as an $m-1 \times n-1$ array of identical squares. Take a rectangle of size $m \times n$, and place $m-1$ equally spaced points on the two vertical sides, and $n-1$ equally spaced points on the two horizontal sides. Then draw lines between every pair of the $2(m+n)$ boundary points, and place a node

Acknowledgement: We thank Max Alekseyev, Gareth McCaughan, Ed Pegg, Jr., and Jinyuan Wang for their assistance during the course of this work. Tom Duff and Keith F. Lynch carried out extensive computations to check how well Sylvestor's theorem applied in practice (answer: very well, see Section 8). We made frequent use of the *gfun* Maple program [18] and the *TikZ* Latex package [5, 23]. We also thank a referee for many helpful comments, including a remark, which strengthened a result in Section 8.

Lars Blomberg, Linghem, Sweden, e-mail: lars.blomberg2@hotmail.com
S. R. Shannon, Rowville, Victoria, Australia, e-mail: scott_r_shannon@hotmail.com
N. J. A. Sloane, The OEIS Foundation Inc., Highland Park, NJ, USA, e-mail: njasloane@gmail.com

https://doi.org/10.1515/9783110754216-005

at each point where these lines intersect. The resulting planar graph, which we denote by $BC(m,n)$, is the main subject of Part 1 of this paper.

Although we have not been very successful in analyzing these graphs, we have collected a great deal of data, which has been entered into various sequences in the *On-Line Encyclopedia of Integer Sequences* [14].

In Part 2 of this paper, we plan to consider other structures such as hollow squares (or "frames"), triangles, pentagons, hexagons, pentagrams, etc., as well as figures formed by drawing semicircles joining equally-spaced points on an interval. The last-mentioned figures are reminiscent of juggling patterns,[1] as studied by Ron Graham in [4] and other papers, and we regret that now it is too late to ask him for help for finding a formula for those numbers.

We were also motivated by memories of stained glass windows seen in the great Gothic cathedrals of northern Europe. In 2019, we made a colored drawing of K_{23} (Figure 5.1) which was reminiscent of a rose window, and we were curious to see what colored versions of other graphs would look like. Informally, our philosophy has been, if we cannot solve it, make art. We make no great claims for artistic merit, but the images are certainly colorful.

Space limitations have restricted the number and quality of the images that we could include here. The corresponding entries in [14] (*A007678*[2] in the case of Figure 5.1) contain a large number of other images, with better resolution. We are especially fond of the three images of K_{41} in *A007678*, and in *A331452* the reader should not miss the images labeled $T(10,2)$, $T(6,6)$, $T(7,7)$, which are drawings of the graphs $BC(10,2)$, $BC(6,6)$, and $BC(7,7)$ discussed below.

This paper is arranged as follows. The last section of this Introduction establishes the notation we will use, especially the terms *nodes*, *chords*, and *cells*, and provides some examples. Section 2 deals with the graphs $BC(1,n)$ (or equivalently $BC(n,1)$), where the bounding polygon is a rectangle of size $1\times n$ (or $n\times 1$). Theorem 2.1 gives Legendre and Griffiths's enumeration of the nodes and cells in $BC(1,n)$. In 2019, Max Alekseyev (personal communication; see also *A306302*) pointed out that the Legendre–Griffiths results are essentially the same as results that he and his coauthors obtained in connection with the enumeration of threshold functions [1, 2]. The family of isosceles triangle graphs $IT(n)$ (Section 3) provides a bridge between the graphs $BC(1,n)$ and two-dimensional threshold functions. Alekseyev also mentioned that their work implies a result that was apparently overlooked in the Legendre and Griffiths papers: the cells in $BC(1,n)$ are always triangles or quadrilaterals. See Theorem 3.1. The proof of this fact in [2] depends on a theorem about teaching sets for threshold functions [19, 26]. We feel that such a elementary property should have a purely geometrical proof,

1 See entry *A290447* in [14].

2 Six-digit numbers prefixed by A refer to entries in the On-Line Encyclopedia of Integer Sequences [14].

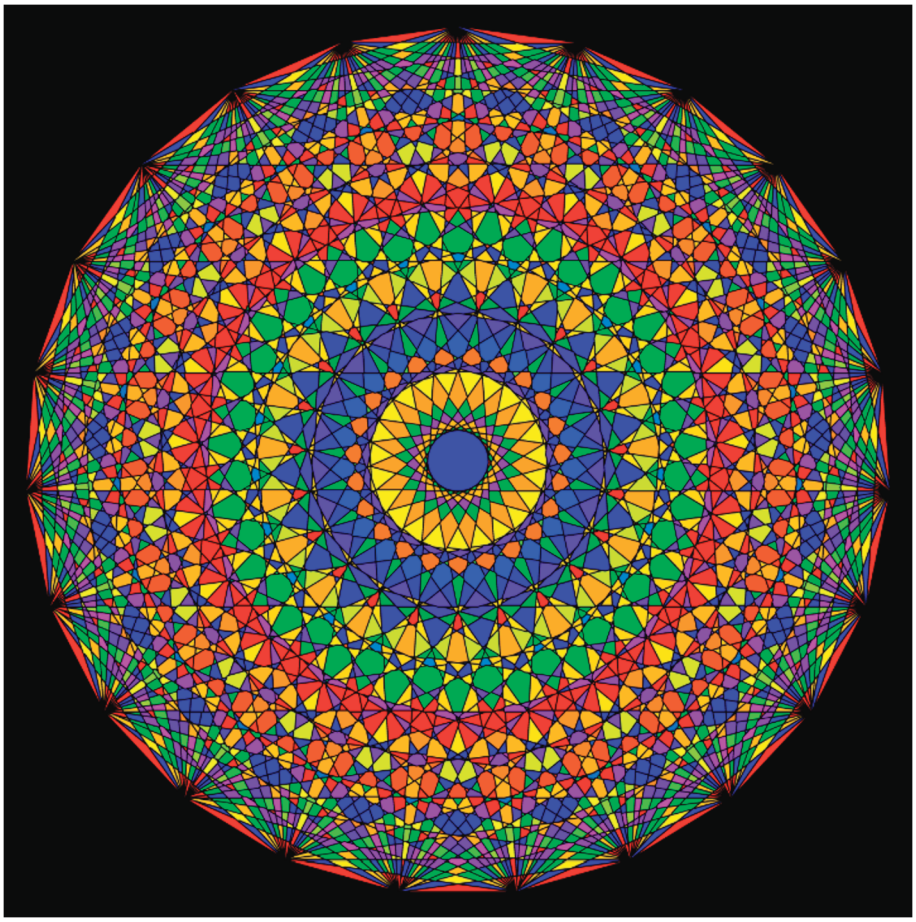


Figure 5.1: Colored drawing of complete graph K_{23} (see Section 10.3 for the coloring algorithm). Entry *A007678* in [14] has many similar images (which are also of higher quality).

although no such proof is presently known. We state this question as Open Problem 3.2.

One possible attack on this problem is to study the distribution of cells in each of the n squares of $BC(1, n)$; see Section 4, and especially Tables 5.2, 5.3, 5.4. The *gfun* Maple program [18] suggests a form for the generating functions of the columns of these tables, but so far this is only a conjecture.

In Section 5, we consider the number of interior nodes in $BC(1, n)$ where c chords meet (Table 5.5). The number of simple nodes, where just two chords cross, is of the greatest interest, since these seem to dominate. But even though we have calculated 500 terms of this sequence (Table 5.6 and *A334701*) we have been unable to find a formula or recurrence (Open Problem 5.2). There have been several similar occasions

during this project when we have regretted not having an oracle that would take a few hundred terms of a simple, well-defined sequence and suggest some kind of formula.[3]

The graph $BC(1,n)$ has bounding polygon, which is a $1 \times n$ rectangle. If we start instead from an $m \times n$ rectangle, with m and $n > 1$, there are three natural ways to define a planar graph, which we will denote by $BC(m,n)$, $AC(m,n)$, and $LC(m,n)$. These are the subjects of Sections 6, 8, and 9, respectively. For these families, we have plenty of data and pictures, but not many results. In Section 6, we conjecture that the cells in $BC(2,n)$ have at most eight sides, and for $n \geq 19$, at most six sides (Conjecture 6.2). Our main result concerning $BC(m,n)$ is an upper bound on the numbers of nodes and cells in $BC(m,n)$, presented in Section 7, which appears to be reasonably close to the true values.

The final section (Section 10) describes how we colored the graphs.

Terminology

Our subject is planar graphs, as shown in most of the figures. We start with a grid or lattice in the plane, draw a polygon on this grid, mark points along the boundary of the polygon, and form the graph by joining these boundary points by lines and creating nodes where these lines cross.

Since all three concepts, grids, polygons, and graphs, involve points and lines, we will establish our terminology with some care, hoping to avoid confusion without being too pedantic.

In graph theory, many different terms are used for the basic notions of node (or point, or vertex) and edge (or line). We will use *node* and *edge* for these specifically graph-theoretic terms. A planar graph divides the plane into cells (or regions, or chambers). We will use the term *cell*, with the understanding that the unbounded region exterior to the graph is not considered to be a cell.[4] Our graphs are also *maps* in the sense of Tutte [24, 25], but we will refer to them simply as planar graphs.

To construct our graph, we usually start from the boundary curve of some connected polygon P in the plane, and mark various points on this boundary. We call these the *boundary points* and we call P the *defining polygon*. We assume P is connected, but not necessarily simply connected. In Part 2, for example, P may be a square annulus. Our graph is then constructed by joining pairs of boundary points by line segments, according to some specific rule. For the graphs $BC(m,n)$, every pair of distinct boundary points is joined by a line segment, which starts at one boundary point and ends

3 The oracle might compare the sequence with shifted versions of each of the hundreds of thousands of entries in [14], and ask Bruno Salvy and Paul Zimmermann's program *gfun*, or Harm Derksen's program *guesss*, or Christian Kratthentaler's program *Rate*, or one of the other programs used by *Superseeker* [20] if there is a formula for the difference.

4 In Part 2, when we consider graphs that have a "hole" in them (such as a square annulus or frame), the hole is also not considered to be a cell.

at another. A line segment is called a *chord* if (apart from its end-points) it lies in the interior of P. Our graph then has *nodes,* which are all the boundary points together with all the points where the chords cross. A node, which is not a boundary point, is called an *interior node*. When referring to the geometry of the polygon P, we will speak of its *sides* and *vertices*. Some of the line segments may coincide with the sides of P; these are not called chords.

We usually subdivide the sides of the polygon by dividing them into equal parts. To divide a side into k equal parts, we insert $k-1$ equally spaced boundary points along that side, so that the side contains a total of $k+1$ boundary points, the two vertices plus the $k-1$ additional points. We say that the side has been *k-reticulated*.

Finally, we give coordinates for our polygons P by defining them in terms of some underlying grid or lattice, which in Part 1 will be the simple square lattice $\mathbb{Z} \times \mathbb{Z}$. The *grid points* have integer coordinates (i, j).

As an example, Figure 5.2 shows the graph $BC(1,1)$ (defined in Section 2), which has 5 nodes, 8 edges, and 4 cells. The polygon is a square and there are two chords which meet at the central node. Figure 5.3 shows the graph $BC(2,2)$, constructed from a square in which each side has been 2-reticulated. There are 16 chords. This graph has 37 nodes, 92 edges, and 56 cells.[5] Note that for a connected planar graph, Euler's formula states that the numbers of nodes, edges, and cells are related by

$$|\text{nodes}| - |\text{edges}| + |\text{cells}| = 1. \tag{5.1}$$

Figure 5.4 shows a colored version of $BC(2,2)$. The principles used to color these graphs are discussed in Section 10. For any undefined terms from graph theory, see [3, 9].

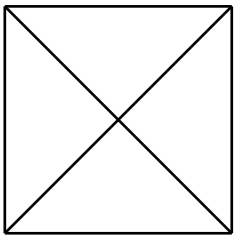


Figure 5.2: The planar graph $BC(1,1)$, a 1-reticulated square. There are four boundary points and two chords, and the graph has five nodes, eight edges, and four cells.

5 $BC(3,3)$ is shown in Figure 5.14 in Section 6 and has 340 cells. There is no known formula for the sequence $4, 56, 340, 1120, 3264, \ldots$ (*A255011*), the number of cells in $BC(n,n)$, even though we have 52 terms.

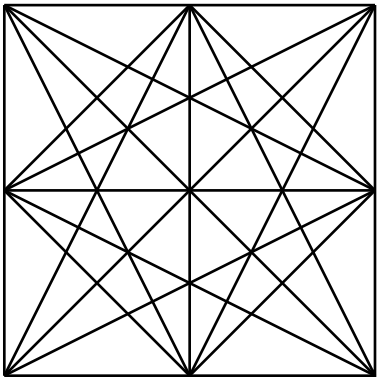


Figure 5.3: $BC(2,2)$: a 2-reticulated square with 56 cells.

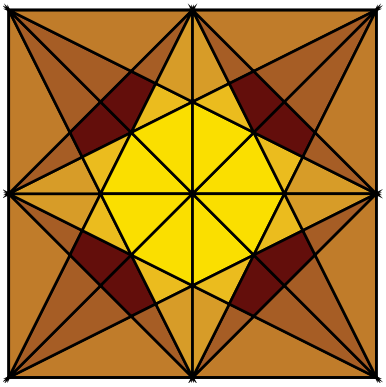


Figure 5.4: The same $BC(2,2)$ drawn with colored cells. See Section 10.2 for coloring scheme.

2 $BC(1,n)$: $1 \times n$ rectangular windows

The defining polygon for the graph $BC(1,n)$ ($n \geq 1$) is a $1 \times n$ rectangle, which we take to have vertices $(0,0)$, $(n,0)$, $(0,1)$, and $(n,1)$. The boundary points for $BC(1,n)$ are the points $\{(i,0),(i,1) : 1 \leq i \leq n\}$, and we join every pair of distinct boundary points by a line segment. Some of these line segments lie on the sides of the rectangle, and there are in addition n^2+2n-1 chords. Figures 5.2, 5.5, and 5.6 show $BC(1,n)$ for $n = 1, 2$, and 3.

Of course, we could equally well have started with a vertical rectangle of size $n \times 1$, in which case the graph would be denoted by $BC(n,1)$. Since this work was partly inspired by the windows of Gothic cathedrals, we admit to a slight preference for $BC(n,1)$ over $BC(1,n)$, although as graphs they are isomorphic. Figures 5.7 and 5.8 show our stained glass window $BC(4,1)$ using two different coloring schemes.

We will continue to discuss $BC(1,n)$, but the reader should remember that the results apply equally well to $BC(n,1)$.

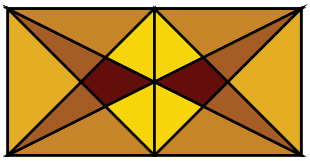


Figure 5.5: $BC(1,2)$.

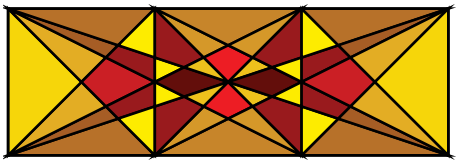


Figure 5.6: $BC(1,3)$.

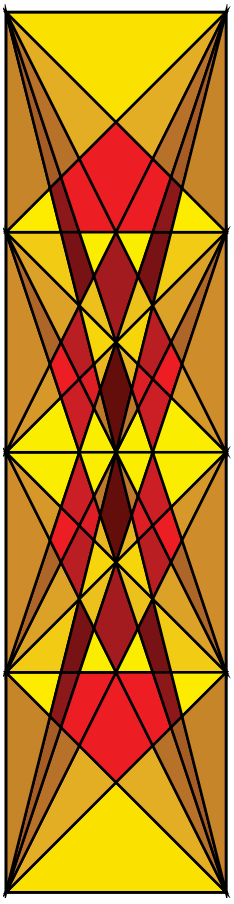


Figure 5.7: $BC(4,1)$, colored using the red and yellow palettes (see Section 10.2).

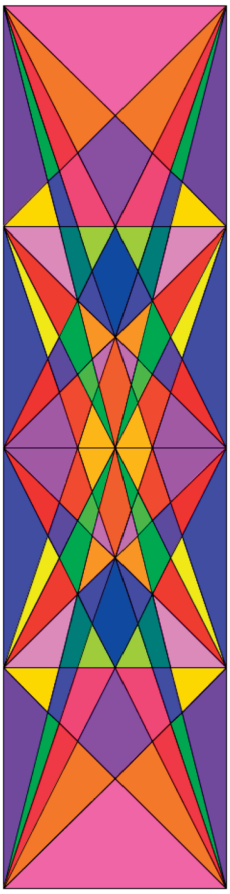


Figure 5.8: A version of $BC(4,1)$ colored by our "random coloring" algorithm (see Section 10.3).

Another way to construct $BC(1,n)$ is to start with the complete bipartite graph $K_{n+1,n+1}$ formed by taking $n+1$ equally spaced points in each of two horizontal rows, joining every upper point to every lower point by a line segment, placing a node at each point where these lines intersect, and then adding the line segments through the two rows of points. Thus $BC(1,2)$ in Figure 5.5 is the well-known nonplanar "utilities" graph $K_{3,3}$ if the two horizontal lines, the seven interior nodes, and the colors are ignored.

The graphs $BC(1,n)$ are one of the few families where explicit formulas are known for the numbers of nodes ($\mathcal{N}(1,n)$), edges ($\mathcal{E}(1,n)$), and cells ($\mathcal{C}(1,n)$). The initial values of these quantities are shown in Table 5.1, along with the A-numbers of the corresponding sequences.

Since by Euler's formula (5.1), $\mathcal{E}(1,n) = \mathcal{N}(1,n) + \mathcal{C}(1,n) - 1$, there is no need to tabulate $\mathcal{E}(1,n)$, and in the future we shall omit those numbers.

Table 5.1: Numbers of nodes, edges, cells in $BC(1,n)$.

n :	1	2	3	4	5	6	7	8	9	10	...	[14]
$\mathcal{N}(1,n)$:	5	13	35	75	159	275	477	755	1163	1659	...	*A331755*
$\mathcal{E}(1,n)$:	8	28	80	178	372	654	1124	1782	2724	3914	...	*A331757*
$\mathcal{C}(1,n)$:	4	16	46	104	214	380	648	1028	1562	2256	...	*A306302*

The following theorem is due to Legendre (2009) [10] and Griffiths (2010) [7], who discuss the problem from the point of view of $K_{n+1,n+1}$. First, we introduce an expression that will frequently appear in these formulas. For $m,n,q \geq 1$, let

$$V(m,n,q) = \sum_{a=1..m} \sum_{\substack{b=1..n \\ \gcd\{a,b\}=q}} (m+1-a)(n+1-b). \tag{5.2}$$

Theorem 2.1 (Legendre [10, Proposition 6], Griffiths [7, Theorem 3]). *For $n \geq 1$, the number of nodes in $BC(1,n)$i is*

$$\mathcal{N}(1,n) = 2(n+1) + V(n,n,1) - V(n,n,2), \tag{5.3}$$

and the number of cells is

$$\mathcal{C}(1,n) = n^2 + 2n + V(n,n,1). \tag{5.4}$$

Remarks. (i) A key step in the proof of (5.3) (see [10]) is finding a condition for three chords to meet at a point. (ii) The starting point for the proof of (5.4) (see [7]) is the observation that in the graph $BC(1,n)$ the chords contain no edges that are parallel to the two long sides of the rectangle. This means that every cell has a unique node that is closest to the upper side of the rectangle. (iii) The term $2(n+1)$ on the right-hand side of (5.3) is the number of boundary points. The difference between the other two terms is therefore the number of interior nodes in $BC(1,n)$ (*A159065*):

$$1, 7, 27, 65, 147, 261, 461, 737, 1143, \ldots. \tag{5.5}$$

(iv) The values of $\mathcal{N}(1,n)$ and $\mathcal{C}(1,n)$ are given in *A331755* and *A306302*.

3 The isosceles triangle graph *IT*(*n*)

In 2019, Max Alekseyev added a comment to *A306302* pointing out that the results in Theorem 2.1 are essentially the same as the results he and his coauthors had obtained in [2] (2015) for the isosceles triangle graphs $IT(n)$.

The definition of the *isosceles triangle graph* $IT(n)$, $n \geq 1$, starts with an isosceles right triangle with vertices $(0,0)$, $(0,1)$, and $(1,0)$. On the vertical side of the triangle, we place n additional boundary points at the points

$$\left(0, \frac{1}{2}\right), \left(0, \frac{1}{3}\right), \left(0, \frac{1}{4}\right), \ldots, \left(0, \frac{1}{n+1}\right),$$

and similarly on the horizontal side we place n additional boundary points at the points

$$\left(\frac{1}{2}, 0\right), \left(\frac{1}{3}, 0\right), \left(\frac{1}{4}, 0\right), \ldots, \left(\frac{1}{n+1}, 0\right).$$

Including the three vertices, there are a total of $2n + 3$ boundary points. We then join every pair of distinct boundary points by a line segment. Besides the three sides of the triangle there are $n^2 + 2n$ chords, and there are interior nodes at the points where the chords intersect. Figures 5.9, 5.10, 5.11 show $IT(2)$, $IT(3)$, and $IT(4)$. The latter two graphs have been colored using the red and yellow palettes (Section 10.2). (There are no boundary points on the hypotenuse other than its two endpoints. In Part 2 of this paper, we will discuss graphs formed by placing n equally-spaced boundary points on all three sides of an equilateral triangle.)

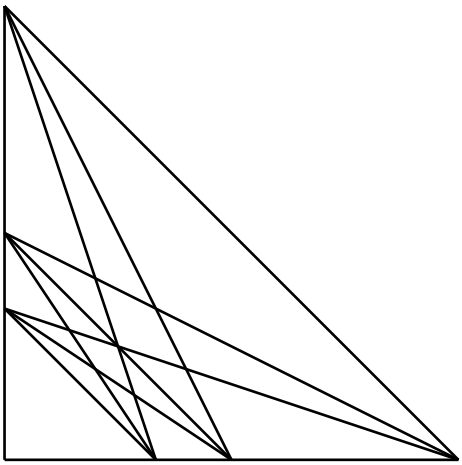


Figure 5.9: The isosceles triangle graph $IT(2)$. There are 14 nodes (7 on boundary, 7 in interior), 30 edges, and 17 cells (15 triangles and 2 quadrilaterals).

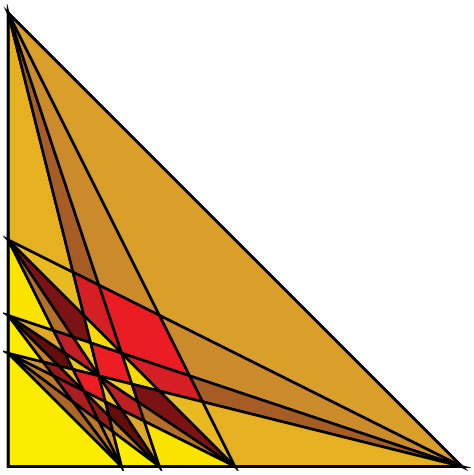


Figure 5.10: $IT(3)$ (33 triangles, 14 quadrilaterals).

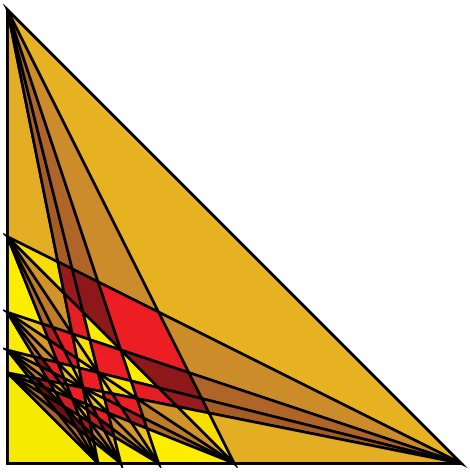


Figure 5.11: $IT(4)$ (71 triangles, 34 quadrilaterals).

Alekseyev pointed out that if we take the boundary points of $BC(1,n)$ to be the points $(i,0)$ and $(i,1)$ for $i = 0,\ldots,n$, then the map

$$(x,y) \longmapsto \left(\frac{1-y}{x+1}, \frac{y}{x+1}\right), \tag{5.6}$$

maps $BC(1,n)$ onto $IT(n)$ minus the node and cell at the origin. Figure 5.12 illustrates this in the case $n = 2$. The six boundary points A,B,C,E,F,G of $BC(1,2)$ are mapped to six of the seven boundary points of $IT(2)$. The point D, the point at infinity on the pos-

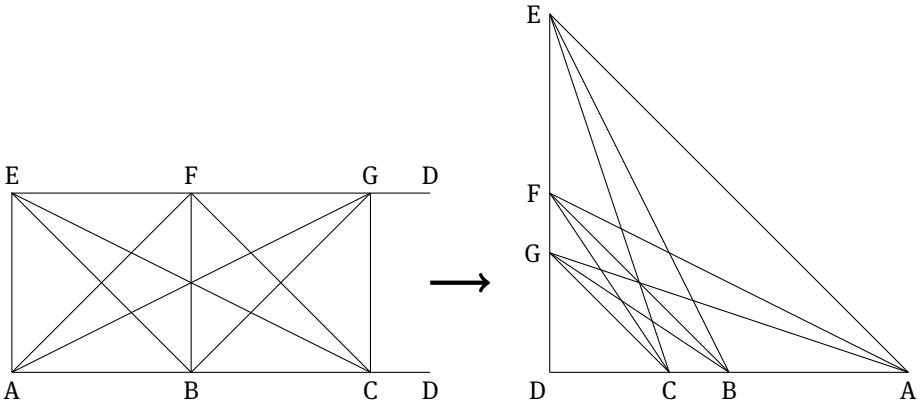


Figure 5.12: Illustrating the map (5.6) from $BC(1,2)$ to $IT(2)$.

itive x axis (not part of $BC(1,2)$), is mapped to the origin in $IT(2)$. The region D, C, G, D to the right of $BC(1,2)$ is mapped to the triangular cell D, G, C, D at the origin in $IT(2)$.

A similar thing happens in the general case: $IT(n)$ always has one more node than $BC(1,n)$, two more edges, and one more cell. When these adjustments are made to the formulas in Theorem 2.1, we obtain the formulas in Theorem 13 of [2]. The counts for nodes, edges, and cells in $IT(n)$ are given in *A332362*, *A332360*, and *A332358*.

However, Alekseyev (personal communication) also pointed out that Theorem 13 of [2] mentions an additional property of $IT(n)$—and hence of $BC(1,n)$—that seems to have been overlooked in [10] and [7].

Theorem 3.1 (Alekseyev et al. [2]). *The cells in $IT(n)$, and hence $BC(1,n)$ are either triangles or quadrilaterals.*

That is, no cell in $BC(1,n)$ has five or more edges. The proof in [2] depends on a theorem about teaching sets for threshold function [19, 26]. No other proof seems to be known.

Open Problem 3.2. Find a purely geometrical proof of Theorem 3.1.

4 The cells in $BC(1,n)$

From Theorems 2.1 and 3.1, we can determine the numbers of triangular and quadrilateral cells in $BC(1,n)$ (sequences *A324042* and *A324043*).

Theorem 4.1. *The $C(1,n)$ cells in $BC(1,n)$ are made up of*

$$T(n) = 2V(n,n,2) + 2n(n+1) \tag{5.7}$$

triangles and

$$Q(n) = V(n,n,1) - 2V(n,n,2) - n^2 \tag{5.8}$$

quadrilaterals.

Proof. The sum $3T(n)+4Q(n)$ double-counts the edges in $BC(1,n)$ except that the $2n+2$ boundary edges are counted only once. Therefore,

$$3T(n) + 4Q(n) + (2n+2) = 2\mathcal{E}(1,n) = 2\big(\mathcal{N}(1,n) + \mathcal{C}(1,n) - 1\big), \tag{5.9}$$

and of course by Theorem 3.1, $T(n) + Q(n) = \mathcal{C}(1,n)$.

The proof is completed by solving these two equations for $T(n)$ and $Q(n)$ and using (5.3), (5.4). □

Figures 5.2, 5.5, 5.6, and 5.7 show the triangles and quadrilaterals for $n = 1, \ldots, 4$.

One way to attack Open Problem 3.2 is to try to understand the distribution of cells in each of the n squares of $BC(1,n)$. Let $t_{n,k}$, $q_{n,k}$, and $c_{n,k}$ denote the numbers of triangles, quadrilaterals, and cells in the kth square of $BC(1,n)$ for $1 \le k \le n$ (so $t_{n,k} + q_{n,k} = c_{n,k}$ and $\sum_k c_{n,k} = \mathcal{C}(1,n)$). From Figure 5.5, for example, we see that $t_{1,1} = t_{1,2} = 7$, $q_{1,1} = q_{1,2} = 1$, and $c_{1,1} = c_{1,2} = 8$.

The two end squares of $BC(1,n)$ are easily understood, and for future reference we state the result as the following.

Theorem 4.2. *For $n \ge 2$, the two end squares of $BC(1,n)$ both contain $2n+3$ triangles and $2n-3$ quadrilaterals.*

Tables 5.2, 5.3, and 5.4 show the values of $t_{n,k}$, $q_{n,k}$, and $c_{n,k}$ for $n \le 10$. More extensive tables, for $n \le 80$, are given in entries *A333286*, *A333287*, *A333288*. However, even with 80 rows of data, we have been unable to find formulas for these numbers.

Table 5.2: Number $t_{n,k}$ of triangles in kth square in $BC(1,n)$ (*A333286*).

$n\backslash k$	1	2	3	4	5	6	7	8	9	10
1	4									
2	7	7								
3	9	14	9							
4	11	24	24	11						
5	13	30	38	30	13					
6	15	38	60	60	38	15				
7	17	44	76	86	76	44	17			
8	19	52	92	120	120	92	52	19		
9	21	58	106	146	158	146	106	58	21	
10	23	66	126	178	216	216	178	126	66	23

Table 5.3: Number $q_{n,k}$ of quadrilaterals in kth square in $BC(1,n)$ (*A333287*).

$n\backslash k$	1	2	3	4	5	6	7	8	9	10
1	0									
2	1	1								
3	3	8	3							
4	5	12	12	5						
5	7	22	32	22	7					
6	9	28	40	40	28	9				
7	11	38	58	74	58	38	11			
8	13	46	74	98	98	74	46	13		
9	15	58	92	130	152	130	92	58	15	
10	17	68	104	150	180	180	150	104	68	17

Table 5.4: Total number $c_{n,k}$ of cells in kth square in $BC(1,n)$ (*A333288*).

$n\backslash k$	1	2	3	4	5	6	7	8	9	10
1	4									
2	8	8								
3	12	22	12							
4	16	36	36	16						
5	20	52	70	52	20					
6	24	66	100	100	66	24				
7	28	82	134	160	134	82	28			
8	32	98	166	218	218	166	98	32		
9	36	116	198	276	310	276	198	116	36	
10	40	134	230	328	396	396	328	230	134	40

There is certainly a lot of structure in these tables. Using the Salvy–Zimmermann *gfun* Maple program [18], we attempted to find generating functions for the columns of these tables. On the basis of admittedly little evidence, we make the following conjecture.

Conjecture 4.3. In all three of Tables 5.2, 5.3, and 5.4, the kth column for $k \geq 3$ has a rational generating function, which can be written with denominator $(1-x^{k-2})(1-x^{k-1})(1-x^k)$.

For example, column 3 of Table 5.2, the sequence $\{t_{n,3}\}$, appears to have generating function

$$x^3 \frac{9 + 15x + 5x^2 - 2x^3 - 13x^4 - 11x^5 - 9x^6 + 2x^7 + 8x^8 - 4x^{10} + 4x^{12}}{(1-x)(1-x^2)(1-x^3)}. \tag{5.10}$$

It would be nice to know more about these quantities.

5 The nodes in $BC(1, n)$

Besides looking at the cells of $BC(1, n)$, it is also interesting to study the nodes. For $n \geq 2$, $BC(1, n)$ has four boundary points of degree $n + 1$ and $2n - 2$ boundary points of degree $n + 2$. An interior node formed when c chords (say) cross has degree $2c$. Let $v_{n,c}$ denote the number of interior nodes of degree $2c$, for $2 \leq c \leq n + 1$. Table 5.5 shows the values of $v_{n,c}$ for $n \leq 10$. A more extensive table, for $n \leq 100$, is given in *A333275*.

Table 5.5: Number $v_{n,c}$ of interior nodes in $BC(1, n)$ where c chords cross (*A333275*).

$n\backslash c$	2	3	4	5	6	7	8	9	10	11
1	1									
2	6	1								
3	24	2	1							
4	54	8	2	1						
5	124	18	2	2	1					
6	214	32	10	2	2	1				
7	382	50	22	2	2	2	1			
8	598	102	18	12	2	2	2	1		
9	950	126	32	26	2	2	2	2	1	
10	1334	198	62	20	14	2	2	2	2	1

Theorem 5.1. *For $n \geq 2$, the numbers $v_{n,c}$ satisfy:*

$$\sum_{c=2}^{n+1} v_{n,c} + 2n + 2 = \mathcal{N}(1, n), \tag{5.11}$$

$$\sum_{c=2}^{n+1} c v_{n,c} + n^2 + 4n + 1 = \mathcal{E}(1, n), \tag{5.12}$$

$$\sum_{c=2}^{n+1} \binom{c}{2} v_{n,c} = \binom{n+1}{2}^2. \tag{5.13}$$

Proof. The first equation simply gives the total number of nodes in $BC(1, n)$. For (5.12) we count pairs (α, β), where α is a cell and β is a node, in two ways, obtaining

$$3T(n) + 4Q(n) = 4(n + 1) + (2n - 2)(n + 2) + \sum_{c} 2c v_{n,c},$$

and use (5.9). To establish (5.13), we start with the observation that if all the $2n + 2$ boundary points of $BC(1, n)$ are perturbed by small random amounts, there will be no triple or higher-order intersection points, all the interior nodes will be simple, and there will be $\binom{n+1}{2}^2$ of them (since any pair of nodes on the upper side of the rectangle

and any pair of nodes on the lower side will determine a unique intersection point). As the boundary points are returned to their true positions, the interior nodes coalesce. If there is an interior point where c chords intersect, the $\binom{c}{2}$ interior nodes there coalesce into one, and we lose $\binom{c}{2}-1$ intersections. We are left with the $\mathcal{N}(1,n)-(2n+2)$ interior intersection points. Thus

$$\sum_{c=2}^{n+1}\left(\binom{c}{2}-1\right)v_{n,c}+\mathcal{N}(1,c)-(2n+2)=\binom{n+1}{2}^2,$$

which simplifies to give (5.13). □

However, we do not even have a formula for the number of simple interior intersection points in $BC(1,n)$ (the first column of Table 5.5, the sequence $\{v_{n,2}\}$, *A334701*), although we have computed 500 terms. The first 100 terms are shown in Table 5.6. We feel that a formula should exist.

Table 5.6: The first 100 terms of the number of simple interior intersection points in $BC(1,n)$.

Terms 1–25	26–50	51–75	76–100
1	49246	679040	3264422
6	57006	732266	3438642
24	65334	790360	3616430
54	75098	849998	3805016
124	85414	914084	3998394
214	97384	980498	4202540
382	110138	1052426	4408406
598	124726	1125218	4626162
950	139642	1203980	4850198
1334	156286	1285902	5085098
1912	174018	1374300	5321854
2622	194106	1463714	5571470
3624	214570	1559064	5826806
4690	237534	1657422	6095870
6096	261666	1762004	6369534
7686	288686	1869106	6655902
9764	316770	1983922	6948566
12010	348048	2102162	7256076
14866	380798	2228512	7565826
18026	416524	2356822	7889032
21904	452794	2493834	8220566
25918	492830	2635310	8568428
30818	534962	2786090	8919298
36246	580964	2938326	9285288
42654	627822	3099230	9658638

Open Problem 5.2. Find a formula for the number of simple interior intersection points in $BC(1,n)$ (see Table 5.6 for 100 terms, or *A334701* for 500 terms).

6 $BC(m,n)$: $m \times n$ rectangular windows

The graph $BC(1,n)$ ($n \geq 1$) is based on a $1 \times n$ rectangle. In this section, we consider what happens if we start more generally from an (m,n)-reticulated rectangle (where $m \geq 1$, $n \geq 1$): this is a rectangle of size $m \times n$ in which both vertical sides are divided into m equal parts, and both horizontal sides into n equal parts. There are $m-1$ nodes on each vertical side and $n-1$ nodes on each horizontal side, for a total of $4+2(m-1)+2(n-1) = 2(m+n)$ boundary points.

We will discuss three families of graphs based on these rectangles, which will be denoted by $BC(m,n)$, $AC(m,n)$, and $LC(m,n)$. The graph $BC(m,n)$ is formed by joining every pair of boundary points by a line segment and placing a node at each point where two or more line segments intersect. Figures 5.3 and 5.4 show $BC(2,2)$, and Figure 5.14 shows $BC(3,3)$ ("*BC*" stands for "boundary chords").

Alternatively, we could have constructed $BC(m,n)$ by starting with an $m \times n$ grid of equal squares, and then joining each pair of boundary grid points by a line segment. However, if we include the interior grid points, there are $(m+1)(n+1)$ grid points in all, and if we join *each* pair of grid points by a line segment, we obtain the graph $AC(m,n)$ ("*AC*" stands for "all chords"). These graphs are discussed by Huntington T. Hall [8], Marc E. Pfetsch and Günter M. Ziegler [15], and Hugo Pfoertner (entry *A288187* in [14]). We shall say more about $AC(m,n)$ in Section 8.

A third family of graphs, $LC(m,n)$, arises if we extend each line segment in $AC(m,n)$ in both directions until it reaches the boundary of the grid ("*LC*" stands for "long chords"). These graphs are discussed by Seppo Mustonen [11–13]. We say more about $LC(m,n)$ in Section 9.

Figure 5.13 shows the differences between the three definitions in the case of a $(3,2)$ reticulated rectangle, the first time the definitions differ. The black lines (both thick and thin) form the graph $BC(3,2)$. The four red lines are the additional line segments that appear when we construct $AC(3,2)$. They start at an interior grid point and so are not present in $BC(3,2)$. The four blue lines extend the red chords until they reach the boundary of the grid, and form $AC(3,2)$.

The numbers of nodes $\mathcal{N}(m,n)$ and cells $\mathcal{C}(m,n)$ in $BC(m,n)$ are shown for $m,n \leq 37$ in *A331453* and *A331452*, respectively, and the initial terms are shown in Table 5.7.

Regrettably, except when m or n is 1, we have been unable to find formulas for any of these quantities. The diagonal case, when $m = n$, is the most interesting (because the most symmetrical), but is also probably the hardest to solve. In accordance with our philosophy of "if you cannot solve it, make art," Figure 5.14 shows our stained

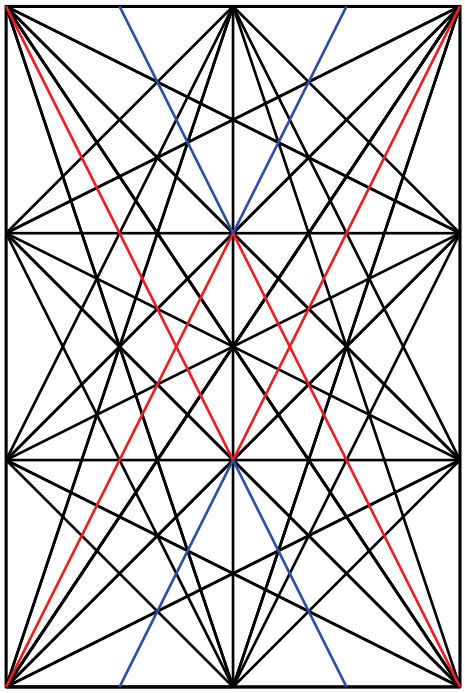


Figure 5.13: Comparison of the graphs $BC(3,2)$ (black lines), $AC(3,2)$ (add the red lines), and $LC(3,2)$ (also add the blue lines).

Table 5.7: Numbers of nodes $\mathcal{N}(m,n)$ and cells $\mathcal{C}(m,n)$ in $BC(m,n)$ for $1 \le m,n \le 7$.

$m\backslash n$	1	2	3	4	5	6	7
1	5, 4	13, 16	35, 46	75, 104	159, 214	275, 380	477, 648
2	13, 16	37, 56	99, 142	213, 296	401, 544	657, 892	1085, 1436
3	35, 46	99, 142	257, 340	421, 608	881, 1124	1305, 1714	2131, 2678
4	75, 104	213, 296	421, 608	817, 1120	1489, 1916	2143, 2820	3431, 4304
5	159, 214	401, 544	881, 1124	1489, 1916	2757, 3264	3555, 4510	5821, 6888
6	275, 380	657, 892	1305, 1714	2143, 2820	3555, 4510	4825, 6264	7663, 9360
7	477, 648	1085, 1436	2131, 2678	3431, 4304	5821, 6888	7663, 9360	12293, 13968

glass window $BC(3,3)$, and entry *A331452* has a large number of larger and even more striking examples which space restrictions do not permit us to show here.

Out of all of these unsolved problems, the case when m (or n) is fixed at 2 would seem to be the most amenable to analysis,[6] perhaps by extending the work of Legendre

6 Since $BC(1,n)$ is not a subgraph of $BC(2,n)$, it may be that $AC(2,n)$ (see Section 8), which *does* have $BC(1,n)$ as a subgraph, may be easier to analyze than $BC(2,n)$.

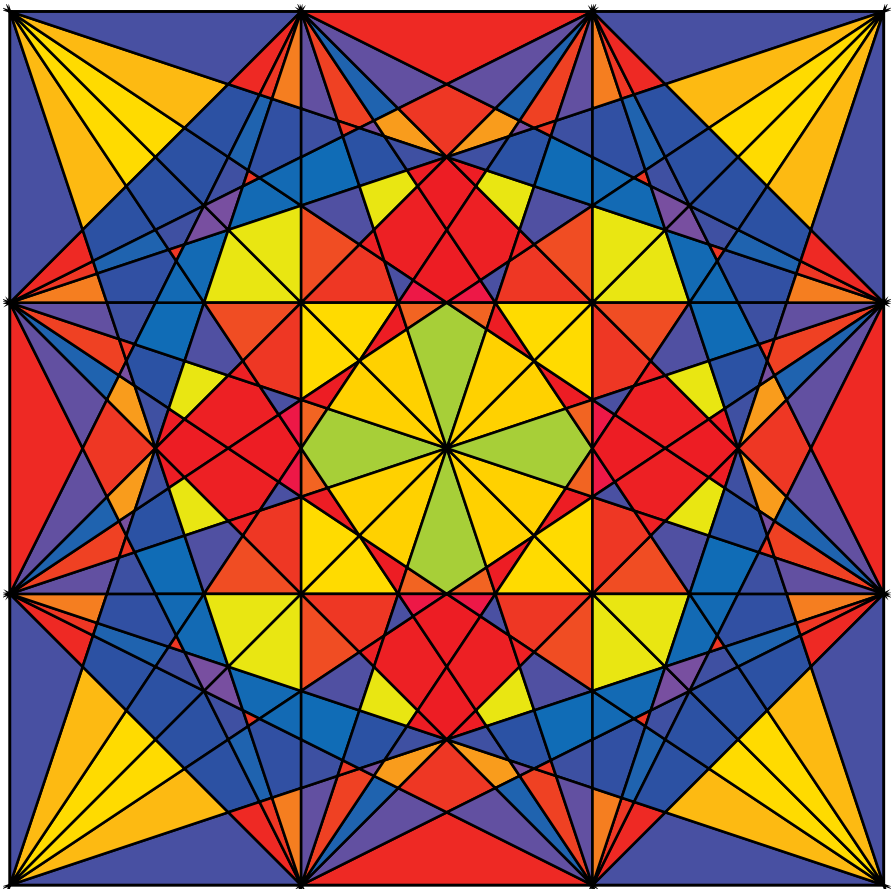


Figure 5.14: The graph $BC(3,3)$. There are 257 nodes and 340 cells.

[10] and Griffiths [7]. For instance, what are the conditions for three chords in $BC(2,n)$ to intersect at a common point? We emphasize this by stating the following.

Open Problem 6.1. Find formulas for the numbers of nodes ($\mathcal{N}(2,n)$, *A331763*) and cells ($\mathcal{C}(2,n)$, *A331766*) in $BC(2,n)$.

The first 10 terms are given in Table 5.8, and 100 terms are given in the entries for these two sequences in [14].

Table 5.8: Numbers of nodes and cells in $BC(2,n)$.

n:	1	2	3	4	5	6	7	8	9	10	...	[14]
$\mathcal{N}(2,n)$:	13	37	99	213	401	657	1085	1619	2327	3257	...	*A331763*
$\mathcal{C}(2,n)$:	16	56	142	296	544	892	1436	2136	3066	4272	...	*A331766*

In $BC(1, n)$, the cells are always triangles or quadrilaterals (Theorem 3.1). It appears that a similar phenomenon holds for $BC(2, n)$. The data strongly suggests the following conjecture.

Conjecture 6.2. The cells in $BC(2, n)$ have at most eight sides, and for $n \geq 19$, at most six sides.

We have verified the conjecture for $n \leq 106$.

Row n of Table 5.9 gives the number of cells in $BC(2, n)$ with k sides, for $k \geq 3$ and $n \leq 20$. For rows $n = 1, 2$, and 3 of this table, see Figures 5.5, 5.2, and 5.13 (black lines only). For row 4 see Figure 5.15, where one can see that $BC(4, 2)$ has 192 triangular cells (red), 92 quadrilaterals (yellow), and 12 pentagons (blue). Entry *A335701* gives the first 106 rows of this table, and has many further illustrations. The row sums in Table 5.9 are the numbers $C(2, n)$ given in column 2 of Table 5.7 and *A331766*.

Table 5.9: Row n gives the number of cells in $BC(2, n)$ with k sides, for $k \geq 3$. It appears that for $n \geq 19$, no cell has more than six sides (see *A335701*).

$n\backslash k$	3	4	5	6	7	8
1	14	2				
2	48	8				
3	102	36	4			
4	192	92	12			
5	326	194	24			
6	524	336	28	4		
7	802	554	80			
8	1192	812	128	4		
9	1634	1314	112	0	4	2
10	2296	1756	200	20		
11	3074	2508	236	22		
12	4052	3252	356	28		
13	5246	4348	472	28		
14	6740	5464	652	28		
15	8398	7054	656	74		
16	10440	8760	940	52		
17	12770	11050	1040	58		
18	15512	13324	1300	60	4	
19	18782	16162	1600	70		
20	22384	19256	1948	104		

Open Problem 6.3. For $BC(m, n)$, m fixed, is there an upper bound on the number of sides of a cell as n varies?

We are at least able to analyze the corner squares of $BC(2, n)$.

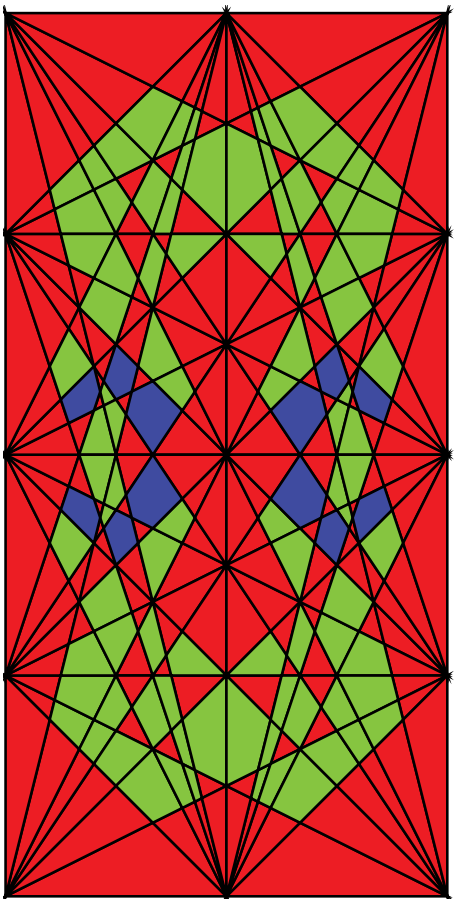


Figure 5.15: $BC(4,2)$ with cells color-coded to distinguish triangles (red), quadrilaterals (yellow), and pentagons (blue).

Theorem 6.4. *For $n = 2$ the four corner squares of $BC(2,n)$ (and $BC(n,2)$) each contain 12 triangles and 4 quadrilaterals, while for $n = 3$ they contain 15 triangles, 6 quadrilaterals, and (exceptionally) one pentagon. For $n \geq 4$, the corner squares each contain $7n+1$ cells, consisting of $2n+9$ triangles and $5n-8$ quadrilaterals.*

Proof. For the proof, we choose a local coordinate system for $BC(2,n)$ with $(0,0)$ at the top left, with the x-axis directed to the right, and the y-axis directed downwards. The four vertices of the rectangle defining $BC(n,2)$ are $(0,0)$, $(2,0)$, $(2,n)$, and $(0,n)$. The top left corner square has vertices A, B, C, D with coordinates $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$, respectively. We assume $n \geq 4$. (The case $n = 4$, where the corner squares contain 17 triangles and 12 quadrilaterals, is shown in Figure 5.15.)

We dissect this square into regions, in each of which the cell structure is apparent (and is such that the boundaries of the regions do not cross any cell boundaries). This

is done as indicated in Figure 5.16. There are six regions, labeled a through f, which are defined as follows. The chord from A to the grid point $(2,1)$ meets BC at its midpoint E, and the chord from A to $(2,4)$ meets CD at its midpoint F. The lines AE, AC, AF, BD, and BF define the six regions.

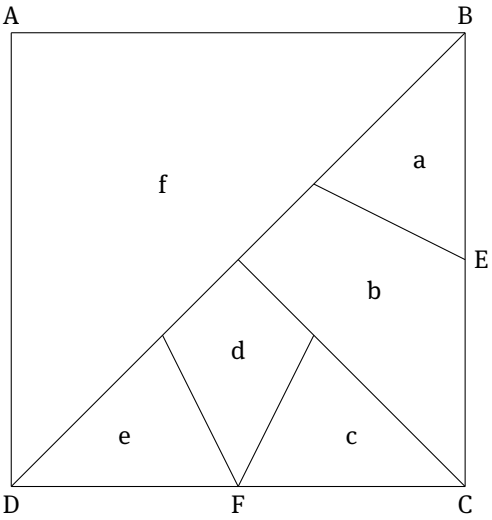


Figure 5.16: Dissection of corner square of $BC(n,2)$, $n \geq 4$, used in proof of Theorem 6.4.

The pencil of $n-1$ chords from A to the grid points $(1,n)$, $(2,n)$, $(2,n-1)$, $(2,n-2)$, $\ldots$, $(2,3)$ cuts CD at the points $(\frac{1}{n},1)$, $(\frac{2}{n},1)$, $(\frac{2}{n-1},1)$, $\ldots$, $(\frac{2}{4},1) = F$, $(\frac{2}{3},1)$. The pencil of $n-1$ chords from B to the grid points $(0,k)$, $2 \leq k \leq n$ cuts CD at the points $F = (\frac{1}{2},1)$, $(\frac{2}{3},1)$, $(\frac{3}{4},1)$, $\ldots$, $(\frac{n-1}{n},1)$. The chord from D to E intersects both pencils of chords. The reader will now have no difficulty in verifying that the numbers of the cells in regions a, b, c, d, e, f are as shown in Table 5.10. □

Table 5.10: Numbers of triangles and quadrilaterals in the regions shown in Figure 5.16.

Region	Triangles	Quadrilaterals
a	n	0
b	2	$2n-3$
c	2	$n-2$
d	1	3
e	2	$2n-6$
f	$n+2$	0
Total	$2n+9$	$5n-8$

7 $BC(m, n)$ in general position

We can obtain reasonably good upper bounds on $\mathcal{N}(m, n)$ and $\mathcal{C}(m, n)$ by analyzing what would happen if all the intersection points in $BC(m, n)$ were simple intersection, that is, if there was no interior point where three or more chords met.

We use $BC_{GP}(m, n)$ to denote a graph obtained by perturbing the boundary points of $BC(m, n)$ (excluding the four vertices) by small random sideways displacements along the boundaries. That is, if a boundary point was a fraction $\frac{i}{j}$ say of the way along a side, we move it to a point $\frac{i}{j} + \epsilon$ of the way along the side, where ϵ is a small random real number. If the ϵ's are chosen independently, the new graph will be in "general position," and there will be no multiple intersection points in the interior.

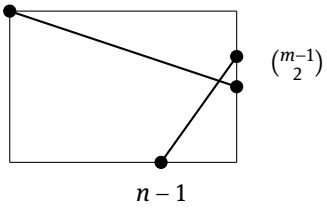
To illustrate the perturbing process, in Figure 5.17 below one can see (ignoring for now the supporting strut on the left) a perturbed version of $BC(1, 2)$ obtained by slightly displacing just one boundary point (labeled 4) so as to avoid the triple intersection point at the center (see Figure 5.5).

Let $\mathcal{N}_{GP}(m, n)$ and $\mathcal{C}_{GP}(m, n)$ denote the numbers of nodes and cells, respectively, in the perturbed graph. The perturbations increase the numbers of nodes and cells, so $\mathcal{N}_{GP}(m, n) \geq \mathcal{N}(m, n)$ and $\mathcal{C}_{GP}(m, n) \geq \mathcal{C}(m, n)$

Theorem 7.1. *For $m, n \geq 1$, the number of interior nodes in $BC_{GP}(m, n)$ is*

$$\frac{1}{4}\{(m + n)(m + n - 1)^2(m + n - 4) + 2mn(2m + n - 1)(m + 2n - 1)\}. \tag{5.14}$$

Proof. We start with the observation that any four boundary points of the rectangle, no three of which are on an side, determine a unique intersection point in the interior of the rectangle. There are several ways to choose these four points. They might be the four vertices of the rectangle, which can be done in just one way. They might consist of three vertices and a single node on one of the other two sides, which can be done in $4(m_1 + n_1)$ ways, where $m_1 = m - 1$ and $n_1 = n - 1$ are the numbers of ways of choosing a single nonvertex point on a side. A more typical example consists of one vertex, and one, respectively, two, points on the two opposite sides, as shown in the following drawing. This can be done in $4(m_1n_2+m_2n_1)$ ways, where $m_2 = (m-1)(m-2)/2$, $n_2 = (n - 1)(n - 2)/2$ are the numbers of ways of choosing two nonvertex nodes from the sides.



There are in all seventeen different configurations for choosing four points, and when the seventeen counts are added up the result is the expression given in (5.14). We omit the details. □

Remarks.

(i) Since there are $2(m+n)$ boundary points, the total number of nodes in $BC_{GP}(m,n)$ is

$$\mathcal{N}_{GP}(m,n) = \frac{1}{4}\{(m+n)(m+n-1)^2(m+n-4) + 2mn(2m+n-1)(m+2n-1)\} + 2(m+n). \tag{5.15}$$

This is our upper bound for $\mathcal{N}(m,n)$.

(ii) Another way to interpret $\mathcal{N}_{GP}(m,n)$ is that this is the number of nodes in $BC(m,n)$ counted with multiplicity (meaning that if there is an interior node where c chords meet, it contributes $\binom{c}{2}$ to the total).

(iii) When $m = n$, (5.15) simplifies to

$$\frac{n}{2}(17n^3 - 30n^2 + 19n + 4), \tag{5.16}$$

which is our upper bound for $\mathcal{N}(n,n)$. For $n = 52$, $\mathcal{N}(n,n) = 52484633$ (from *A331449*), while (5.16) gives 60065408, too large by a factor of 1.14, which is not too bad. The moral seems to be that most interior nodes are simple.

(iv) When $m = 1$, (5.14) becomes $n^2(n+1)^2/4$, which agrees with the number mentioned in the proof of Theorem 5.1.

(v) For large m and n, the expression (5.15) is dominated by the degree 4 terms, which are

$$\frac{1}{4}(m^4 + n^4 + 8mn(m^2 + n^2) + 16m^2n^2). \tag{5.17}$$

Setting $m = n$, we get $\mathcal{N}_{GP}(n,n) \sim 17n^4/2$ as $n \to \infty$. We can confirm this by looking at the number of ways to choose four nodes out of the $4n$ boundary points so that no three are on a side. This is (essentially)

$$\binom{4n}{4} - 4\binom{n}{4} - 12n\binom{n}{3} \sim \frac{17}{2}n^4. \tag{5.18}$$

(vi) From (v), we have $\mathcal{N}(n,n) = O(n^4)$. In fact, we conjecture that $\mathcal{N}(n,n) \sim \mathcal{N}_{GP}(n,n) \sim 17n^4/2$. But to establish this we would need better information about the number of interior nodes in $BC(n,n)$ with a given multiplicity.

Now that we know the number of nodes, we can also find the number $\mathcal{C}_{GP}(m,n)$ of cells in $BC_{GP}(m,n)$. For this, we use a method described by Freeman [6]. The following is a slight modification of his procedure. $BC_{GP}(m,n)$ has $2(m+n)$ boundary points. We label the top left corner vertex 0, and the bottom right corner vertex $2(m+n)-1$. The

nodes along the top side we label 0, 1, 3, 5 . . .,$2n - 1$, continuing along the right-hand side with $2n + 1$, $2n + 3$ . . ., $2(m + n) - 1$. Along the left-hand side, we place the labels 0, 2, 4, . . ., $2m$, continuing along the bottom side with $2m + 2$, $2m + 4$, . . ., $2m + 2n - 2$, $2(m + n) - 1$.

Next, we raise the bottom left corner of the rectangle until the boundary points are at different heights, and so that the order of the heights matches the order of the labels (node 0 becomes the highest point, followed by nodes 1, 2, . . . in order). Figure 5.17 illustrates the case $BC_{GP}(1, 2)$. The black strut raises the bottom left corner so that the heights of the nodes are in the correct order.

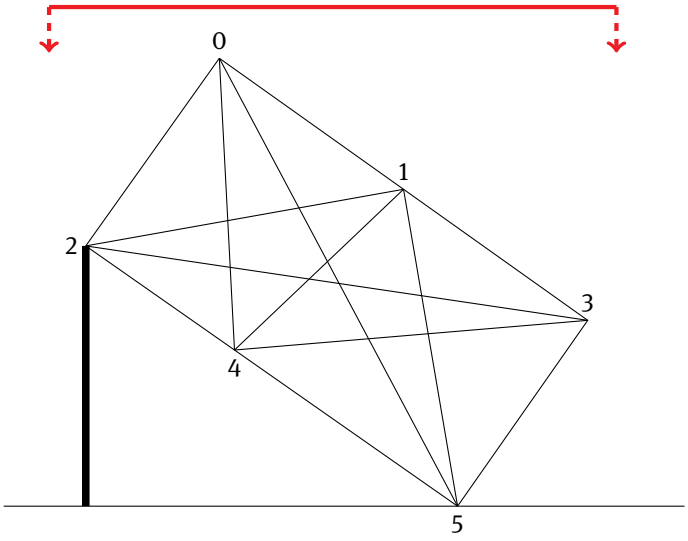


Figure 5.17: $BC(1, 2)$ in general position: node 4 has been displaced slightly so as to avoid the triple intersection point at the center. The strut on the left tilts the figure so that the ordinates of the boundary points are in the same order as the labels. The red line is the "counting line," which descends across the picture in order to count the cells.

We now take a horizontal line (Freeman calls it a "counting line"), and slide it downwards from the top of the figure to the bottom, recording each time it cuts a new cell. The counting line is shown in red in the figure.

When the counting line reaches a boundary point, with label k (say), the count is increased by the number of cells originating at k that have not yet been counted. This number is equal to the number of boundary points with label greater than k, which are not on the same side as k. On the other hand, when the counting line reaches an interior node the count increases by exactly 1 (this is because there is no point where three chords meet). So the contribution to the count from the interior nodes is simply the number of interior nodes, which is known from Theorem 7.1.

In Figure 5.17, the count goes up by 3 at node 0, by 3 at node 1, 1 at node 2, and 1 at node 3, for a subtotal of 8. There are 9 interior nodes, so the total number of cells is 17.

From a careful study of a tilted version of the general case $BC_{GP}(m,n)$, combined with (5.15), we obtain the following.

Theorem 7.2. *For $m,n \geq 1$, the number of cells in $BC_{GP}(m,n)$ is*

$$\mathcal{C}_{GP}(m,n) = \frac{1}{4}\{(m-1)^2(m-2)^2 + (n-1)^2(n-2)^2\} + 2mn\left(m+n-\frac{3}{2}\right)^2 + \frac{9mn}{2} - 1. \quad (5.19)$$

Remark. Asymptotically, $\mathcal{C}_{GP}(m,n)$ and $\mathcal{N}_{GP}(m,n)$ behave in the same way. In fact the difference $\mathcal{C}_{GP}(m,n) - \mathcal{N}_{GP}(m,n)$ is only $m^2+4mn+n^2-4m-4n+1$, a quadratic function of m and n.

8 The graphs $AC(m,n)$

The graph $AC(m,n)$ was defined in Section 6. We take an $(m+1)\times(n+1)$ square grid of nodes, and draw a line segment between *every* pair of distinct grid nodes. (If we only joined pairs of boundary points we would get $BC(m,n)$.)

Figure 5.13 shows $AC(3,2)$ (take the black and red lines only, not the blue lines). Hugo Pfoertner has made black and white drawings of $AC(m,n)$ for $1 \leq m,n \leq 5$ in *A288187*. Figure 5.18 shows a black and white drawing of $AC(3,3)$ made using *TikZ* [5, 23].

The numbers of nodes $\mathcal{N}_{AC}(m,n)$ and cells $\mathcal{C}_{AC}(m,n)$ in $AC(m,n)$ are given for $m,n \leq 9$ in *A288180* and *A288187*, respectively, and the initial terms are shown in Table 5.11. The first row and column of Table 5.11 are the same as the first row and column of Table 5.7 but are included for completeness.

It is clear (compare Figures 5.14 and 5.18) that $AC(m,n)$ contains far more nodes and cells than $BC(m,n)$. We may obtain an upper bound on $\mathcal{N}_{AC}(n,n)$ as follows. The graph $AC(n,n)$ has $(n+1)^2$ grid points. The number of ways of choosing four grid points is $\binom{(n+1)^2}{4}$, and except for a vanishingly small fraction of cases, no three points will be collinear. There are then two possibilities: the four points may form a convex quadrilateral, or a triangle with the fourth point in its interior. In the first case, the intersection of the two diagonals of the quadrilaterals is a node of $AC(n,n)$ (which may or may not be a new node), but in the second case no new node is formed.

If four points in the plane are chosen at random from a square, by what is known as "Sylvester's theorem," the probability that they form a convex quadrilateral is 25/36 and the probability that they form a triangle with an interior point is 11/36 (see [16, Table 4], [21, Table 3, p. 114] for the complicated history of this result). If we rescale our $(n+1)\times(n+1)$ grid into the unit square and let n go to infinity, the uniform distribution on the grid converges to Lebesgue measure. We can then conclude from Sylvester's

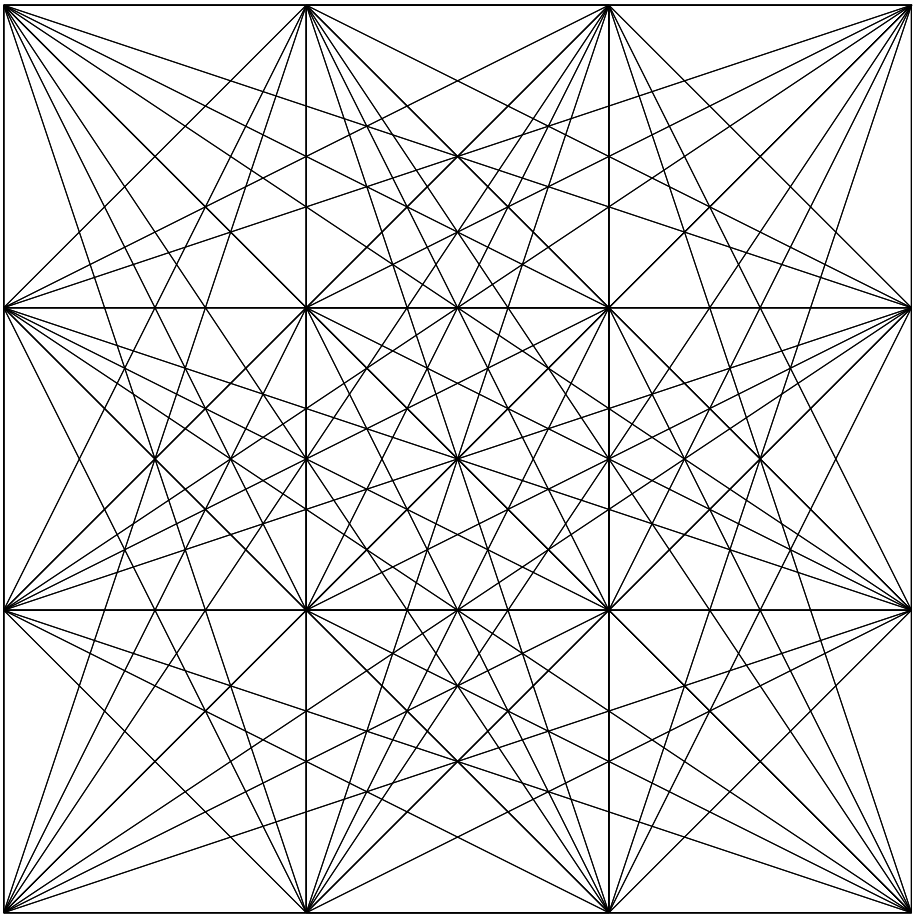


Figure 5.18: The graph $AC(3,3)$. There are 353 nodes and 520 cells.

theorem that the number of nodes in $AC(n,n)$ counted with multiplicity is asymptotically

$$\frac{25}{36}\binom{(n+1)^2}{4} \sim \frac{25}{864}n^8 = 0.0289\ldots n^8. \tag{5.20}$$

So $\mathcal{N}_{AC}(n,n) = O(n^8)$, compared with $\mathcal{N}(n,n) = O(n^4)$ for $BC(n,n)$.

However, in contrast to our upper bound (5.18) for $BC(n,n)$, where the constant 17/2 seems correct, the constant 25/864 in (5.20) seems far from the truth (at $n = 25$ it is too big by a factor of about 4). But to improve it we would need further information about the multiplicity of the chord-intersections than we have now.

Both Tom Duff (personal communication) and Keith F. Lynch (personal communication) have carried out extensive experiments, studying what happens when four

Table 5.11: Numbers of nodes $\mathcal{N}_{AC}(m,n)$ and cells $\mathcal{C}_{AC}(m,n)$ in $AC(m,n)$ for $1 \le m,n \le 7$.

$m\backslash n$	1	2	3	4	5	6	7
1	5, 4	13, 16	35, 46	75, 104	159, 214	275, 380	477, 648
2	13, 16	37, 56	121, 176	265, 388	587, 822	1019, 1452	1797, 2516
3	35, 46	121, 176	353, 520	771, 1152	1755, 2502	3075, 4392	5469, 7644
4	75, 104	265, 388	771, 1152	1761, 2584	4039, 5700	7035, 9944	12495, 17380
5	159, 214	587, 822	1755, 2502	4039, 5700	8917, 12368	15419, 21504	27229, 37572
6	275, 380	1019, 1452	3075, 4392	7035, 9944	15419, 21504	26773, 37400	47685, 65810
6	477, 648	1797, 2516	5469, 7644	12495, 17380	27229, 37572	47685, 65810	84497, 115532

points are chosen from an $m \times n$ grid, and have confirmed that there is excellent agreement with the predictions of Sylvester's theorem.

In a remarkable calculation, Tom Duff enumerated and classified all sets of four points chosen from an $m \times n$ grid for $m, n \leq 349$. In a 349×349 grid, there are 6366733094048270910 strictly convex quadrilaterals out of 9170030499095875150 total. The fraction is 0.6942979, just a little short of Sylvester's $25/36 = 0.694444\ldots$. The deficit is explained by the not quite negligible counts of quadrilaterals with at least three collinear points. If those are included with the strictly convex quadrilaterals, the ratio is 0.6945982, slightly more than 25/36.

9 The graphs $LC(m, n)$

The graph $LC(m, n)$ was defined in Section 6. We take an $(m+1) \times (n+1)$ square grid of nodes, draw a line segment between *every* pair of grid nodes, and extend these lines until they meet the boundaries of the grid. These graphs were discussed by Mustonen [11–13]. Figure 5.13 shows $LC(3, 2)$ (take the black, red, and blue lines), and Figure 5.19 shows our stained glass coloring of $LC(3, 3)$.

The numbers of nodes $\mathcal{N}_{LC}(m, n)$ and cells $\mathcal{C}_{LC}(m, n)$ in $LC(m, n)$ are given for $m, n \leq 8$ in *A333284* and *A333282*, respectively, and the initial terms are shown in Table 5.12. Again the first row and column are the same as in Table 5.7. Mustonen [12, Table 3] gives the first 29 terms of the diagonal sequence $\mathcal{N}_{LC}(n, n)$ (*A333285*). For this problem, we can also give an upper bound on the number of nodes counted with multiplicity, only now we have no need of Sylvester's theorem. Consider four points chosen from the $(n+1) \times (n+1)$ grid points, with no three points collinear. If the points form a triangle with a point in the interior, joining the three vertices of the triangle to the interior point and then extending these chords until they meet the sides of the triangle (something we were not allowed to do in the previous case) will produce three potentially new nodes. If the four points form a convex quadrilateral, there are also potentially three nodes that could be created: the intersection of the two diagonals, and the two points where pairs of opposite sides meet when extended. Figure 5.20 shows the two cases. The black nodes are the four grid points and the red nodes are the potential new nodes. Of course, in the second case, the two external red points may be outside the grid (or at infinity), and so would not be counted.

In any case, the maximum number of new nodes that are created is $3\binom{(n+1)^2}{4} + 2\binom{(n+1)^2}{2}$ (the latter term coming from the intersections of the chords with the boundaries). Asymptotically, this is $n^8/8$. This is an upper bound on $\mathcal{N}_{LC}(n, n)$, because we do not always get three new nodes for each 4-tuple of grid points, and because multiple intersection points are counted multiple times. Based on his data for $n \leq 29$, Mustonen [12] makes an empirical estimate that $\mathcal{N}_{LC}(n, n) \sim Cn^8$, where C is about 0.0075. So our constant, 1/8 is, unsurprisingly, an overestimate.

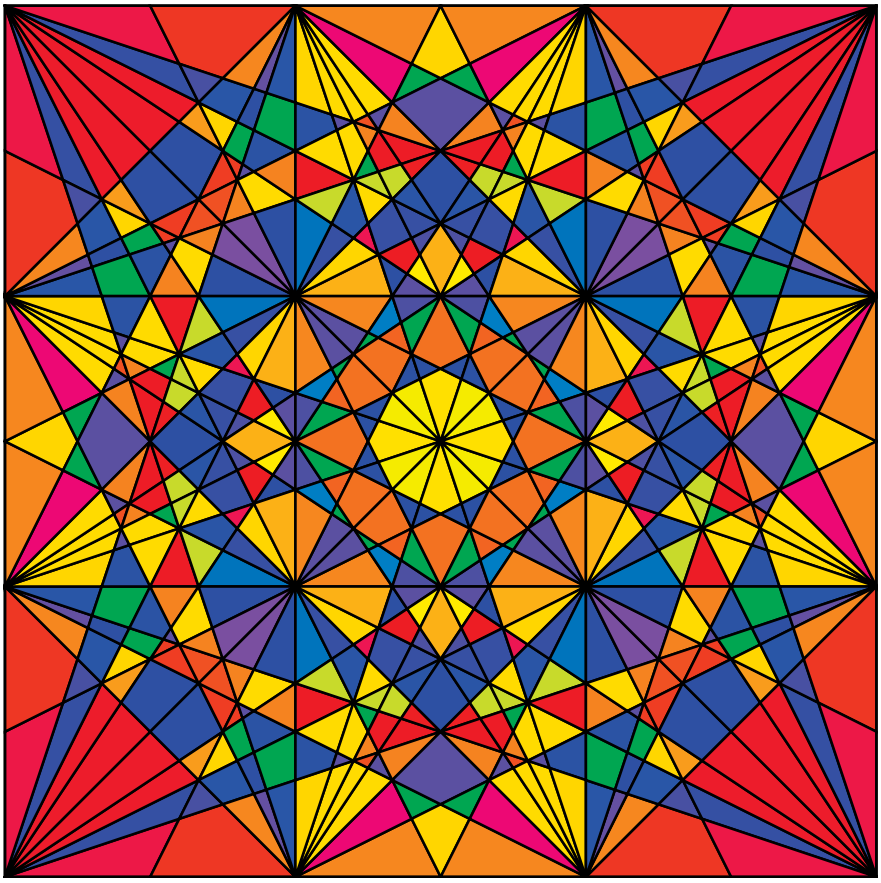


Figure 5.19: The graph $LC(3, 3)$. There are 405 nodes and 624 cells.

We conclude that as we progress from $BC(n, n)$ to $AC(n, n)$ to $LC(n, n)$, the graphs become progressively more dense, and so counting the nodes with multiplicity gives a steadily weaker upper bound on their number.

10 Choosing the colors

We used three different coloring schemes.

10.1 Number-of-sides coloring

The simplest scheme colors the cells according to the number of sides, with randomly chosen colors. This is used in Figure 5.15 and in figures in [14] (entries *A333282*, *A335701*, for example) when studying the distribution of cells by number of sides.

Table 5.12: Numbers of nodes $\mathcal{N}_{LC}(m,n)$ and cells $\mathcal{C}_{LC}(m,n)$ in $LC(m,n)$ for $1 \le m,n \le 7$.

$m\backslash n$	1	2	3	4	5	6	7
1	5, 4	13, 16	35, 46	75, 104	159, 214	275, 380	477, 648
2	13, 16	37, 56	129, 192	289, 428	663, 942	1163, 1672	2069, 2940
3	35, 46	129, 192	405, 624	933, 1416	2155, 3178	3793, 5612	6771, 9926
4	75, 104	289, 428	933, 1416	2225, 3288	5157, 7520	9051, 13188	16129, 23368
5	159, 214	663, 942	2155, 3178	5157, 7520	11641, 16912	20341, 29588	36173, 52368
6	275, 380	1163, 1672	3793, 5612	9051, 13188	20341, 29588	35677, 51864	63987, 92518
7	477, 648	2069, 2940	6771, 9926	16129, 23368	36173, 52368	63987, 92518	114409, 164692

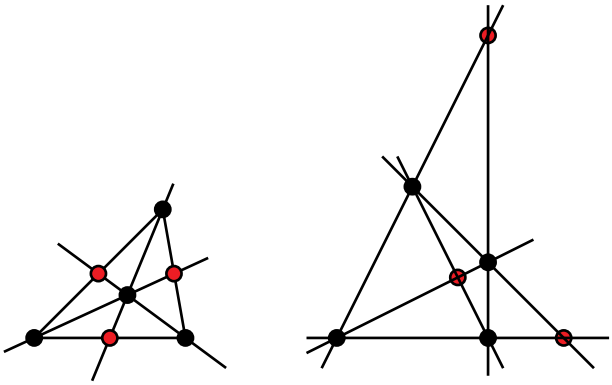
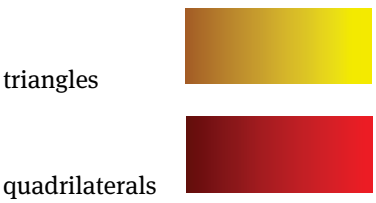


Figure 5.20: The two possibilities for choosing four noncollinear points from an $m \times n$ grid.

10.2 The yellow and red palettes

This is a refinement of the previous scheme, which modifies the color according to the shape of the cell. For Figures 5.4, 5.5, 5.6, 5.7, 5.10, 5.11, the cells are either triangles or quadrilaterals, and we use colors which darken as the cell becomes more irregular. More precisely, the cells are colored according to the following rule. If the cell has n sides (where n is 3 or 4), let λ be the area of the cell divided by the area of an n-sided regular polygon with the same circumradius. Then the cell is assigned color number $\sqrt{\lambda}$ from the following palettes:



10.3 Random colorings

For Figures 5.1, 5.7, 5.14, etc., the color of a cell is assigned by first computing the average distance of the nodes of the cell from the center of the picture. These average distances are then grouped into a certain number of bins (we used 1000 bins), and the nonempty bins are assigned a random color from the standard spectrum from red to violet. This ensures a symmetrical coloring with contrasting colors for neighboring cells. In practice, we do this several times and then choose the most appealing picture. We also have the option of restricting the color palette to achieve certain effects (reds, blues, and greens for a cathedral-like window, or various shades of browns for the frames that we will see in Part 2).

Bibliography

[1] M. A. Alekseyev, On the number of two-dimensional threshold functions, *SIAM J. Discrete Math.*, **24**(4) (2010), 1617–1631.

[2] M. A. Alekseyev, M. Basova and N. Yu. Zolotykh, On the minimal teaching sets of two-dimensional threshold functions, *SIAM J. Discrete Math.*, **29**(1) (2015), 157–165.

[3] B. Bollobás, *Graph Theory: An Introductory Course*, Springer, 1979.

[4] F. Chung and R. Graham, Primitive juggling sequences, *Am. Math. Mon.*, **115**(3) (2008), 185–194.

[5] J. Crémer, A very minimal introduction to TikZ. March 11, 2011, https://cremeronline.com/LaTeX/minimaltikz.pdf.

[6] J. W. Freeman, The number of regions determined by a convex polygon, *Math. Mag.*, **49**(1) (1975), 23–26.

[7] M. Griffiths, Counting the regions in a regular drawing of $K_{n,n}$, *J. Integer Seq.*, **13**, Article #10.8.5 (2010).

[8] H. T. Hall, *Counterexamples in Discrete Geometry*. Ph.D. Dissertation, Mathematics Department, University of California Berkeley, 2004.

[9] F. Harary, *Graph Theory*, Addison-Wesley, Reading MA, 1969.

[10] S. Legendre, The number of crossings in a regular drawing of the complete bipartite graph, *J. Integer Seq.*, **12**, Article #09.5.5 (2009).

[11] S. Mustonen, Statistical accuracy of geometric constructions, September, 2, 2008. http://www.survo.fi/papers/GeomAccuracy.pdf.

[12] S. Mustonen, On lines and their intersection points in a rectangular grid of points, April 16, 2009. http://www.survo.fi/papers/PointsInGrid.pdf.

[13] S. Mustonen, On lines going through a given number of points in a rectangular grid of points, May 12, 2010. http://www.survo.fi/papers/LinesInGrid2.pdf.

[14] The OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, 2021. https://oeis.org.

[15] M. E. Pfetsch and G. M. Ziegler, Large chambers in a lattice polygon, December 13, 2004. http://www.mathematik.tu-darmstadt.de/~pfetsch/chambers.

[16] R. E. Pfiefer, The historical development of J. J. Sylvester's Four Point Theorem, *Math. Mag.*, **62**(5) (1989), 309–317.

[17] B. Poonen and M. Rubinstein, The number of intersection points made by the diagonals of a regular polygon, *SIAM J. Discrete Math.*, **11**(1) (1998), 135–156.

[18] B. Salvy and P. Zimmermann, GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable, *ACM Trans. Math. Softw.*, **20** (1994), 163–177.

[19] V. N. Shevchenko and N. Yu. Zolotykh, On the complexity of deciphering the threshold functions of k-valued logic, *Dokl. Akad. Nauk*, **362**(5) (1998), 606–608 (Russian), (English translation) Dokl. Math. **58** (1998), 268–270.

[20] N. J. A. Sloane, The email servers and Superseeker, 2010. https://oeis.org/ol.html.

[21] H. Solomon, *Geometric Probability*, SIAM, Philadelphia, 1978.

[22] S. E. Sommars and T. Sommars, Number of triangles formed by intersecting diagonals of a regular polygon, *J. Integer Seq.*, **1**, Article #98.1.5 (1998).

[23] T. Tantau, *The PGF/TikZ Programming Language*, Version 2.10, CTAN Org., October 25, 2010.

[24] W. T. Tutte, A census of planar maps, *Can. J. Math.*, **15** (1963), 249–271.

[25] W. T. Tutte, On the enumeration of planar maps, *Bull. Am. Math. Soc.*, **74** (1968), 64–74.

[26] N. Yu. Zolotykh, On the complexity of deciphering threshold functions in two variables, in Proc. 11th Internat. School Seminar "Synthesis and complexity of control systems," Part I, pp. 74–79, Center of Applied Research, Moscow State Univ. Faculty of Mechanics and Mathematics, Moscow, Russia, 2001.

Tom C. Brown and Shahram Mohsenipour

Two extensions of Hilbert's cube lemma

Dedicated to the memory of Ron Graham

Abstract: "Hilbert's cube lemma" states that for every finite coloring of $\mathbb{N}$ (the set of positive integers) and every $n \in \mathbb{N}$, there exist $d_1, d_2, \ldots, d_n \in \mathbb{N}$ such that infinitely many translates of

$$\left\{\sum_{i=1}^{n} \epsilon_i d_i : 0 \le \epsilon_1, \ldots, \epsilon_n \le 1\right\}$$

are monochromatic. (Given the coloring, $d_1, d_2, \ldots, d_n$ depend on n.) We show that for every finite coloring of $\mathbb{N}$ and all $k \ge 2$ there exist $d_1 < d_2 < \cdots \in \mathbb{N}$ such that for each $n \ge 1$, infinitely many translates of

$$P_n = \left\{\sum_{i=1}^{n} \epsilon_i d_i : 0 \le \epsilon_1, \ldots, \epsilon_n \le k-1\right\}$$

are monochromatic, and $|P_n| = k^n$. (Given the coloring, the sequence $d_1, d_2, \ldots$ depends only on k. That is, $P_1 \subset P_2 \subset \cdots \subset P_n \subset \cdots$.) We also show that for every finite coloring of $\mathbb{N}$ and all $n, k \in \mathbb{N}$, there exist $a, d_1, d_2, \ldots, d_n$ such that $d_1 < d_2 < \cdots < d_n$ and

$$\left\{a + \sum_{i=1}^{n} \epsilon_i d_i : 0 \le \epsilon_1, \ldots, \epsilon_n \le k-1\right\} \cup \{d_1, \ldots, d_n\}$$

is monochromatic. (Given the coloring, $a, d_1, d_2, \ldots, d_n$ depend on n, k.)

1 Introduction

Hilbert's cube lemma appeared in 1892 [8] and is sometimes viewed as the first theorem in Ramsey theory. See [1], and especially [11], for some background. Both [6] and

Acknowledgement: The research of the second author was in part supported by a grant from IPM (No. 99030403). We thank the referee for a careful reading of the manuscript and for several helpful suggestions.

Tom C. Brown, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada, e-mail: tbrown@sfu.ca
Shahram Mohsenipour, School of Mathematics, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran, e-mail: sh.mohsenipour@gmail.com

https://doi.org/10.1515/9783110754216-006

[11] contain statements of the "density version" of this result, and proofs of the density version can be found in [5, 6].

The results described in this note are presented as extensions of Hilbert's cube lemma. They may also be viewed as generalizations of van der Waerden's theorem on arithmetic progressions [12, 13], which says that for every finite coloring of $\mathbb{N}$ and all $k \in \mathbb{N}$ there exist a, d such that $\{a + \epsilon d : 0 \leq \epsilon < k\}$ is monochromatic. Brauer strengthened this [6, p. 70] to the van der Waerden–Brauer theorem, which says that $\{a + \epsilon d : 0 \leq \epsilon < k\} \cup \{d\}$ is monochromatic.

For a given fixed value of k and a given finite coloring of $\mathbb{N}$, one may note that (for a single value of n) the existence of $d_1, \ldots, d_n$ and a monochromatic translate of

$$P_n = \left\{ \sum_{i=1}^{n} \epsilon_i d_i : \epsilon_i \in \{0, 1, \ldots, k-1\} \right\}$$

does follow directly from the extended Hales–Jewett theorem [6, 7]. Indeed, given n, r (k is fixed at the beginning of this paragraph), the extended Hales–Jewett theorem says that if m is sufficiently large and the set A_k^m of all words of length m on the alphabet $\{0, 1, \ldots, k-1\}$ is r-colored, then there is a monochromatic combinatorial n-space. If now the elements of A_k^m are viewed as the base k representations of the elements of $[0, k^m - 1]$, a combinatorial n-space is precisely a translate of P_n (where each d_i is a sum of distinct powers of k).

However, in order to obtain $P_1 \subset P_2 \subset \cdots \subset P_n \subset \cdots$, as in the first extension described in the Abstract (Theorem 2 below), we need to use a different approach.

We use Theorem 1 below, which involves "uniform reccurrence" of factors in certain infinite words on a finite alphabet [9, 10], together with van der Waerden's theorem on arithmetic progressions. We also indicate (Theorem 3) that the fixed "k" in Theorem 2 can be replaced by any sequence $k_i, i \geq 1$, of positive integers.

The second extension described in the Abstract (Theorem 4 below) is proved using only the van der Waerden theorem.

2 Uniformly recurrent infinite words all of whose factors are factors of a given infinite word

The crucial result we need concerning infinite words is Theorem 1 below. For completeness, we include a proof, based on one due to J. Justin and G. Pirillo [9]. (A more labor-intensive proof can be obtained using the elaborate methods of symbolic dynamics. See, for example, pp. 213–215 of [4].)

We begin with some terminology.

Let A be a finite set. We denote by A^ω the set of all infinite sequences of elements of A, or *infinite words* on the "alphabet" A. If $c \in A^\omega$, we write $c = c(1)c(2)c(3)\cdots$, and

we regard $c : \mathbb{N} \to A$ as a *coloring* of $\mathbb{N}$, where A is the set of colors and for $n \in \mathbb{N}$, $c(n)$ is the color assigned to n.

We denote by A^* the set of all *finite* sequences of elements of A, or *words* on the alphabet A, including the empty word. If $u, v \in A^*$, say $u = a_1 a_2 \cdots a_n$ and $v = b_1 b_2 \cdots b_m$, with $a_i, b_j \in A$, $1 \le i \le n$, $1 \le j \le m$, then their *product* $uv \in A^*$ is the word $uv = a_1 a_2 \cdots a_n b_1 b_2 \cdots b_m$. A word $v \in A^*$ is a *factor* of the word $w \in A^*$ if there exist (possibly empty) words $p, q \in A^*$ such that $w = pvq$. A word $v \in A^*$ is a factor of an infinite word $s \in A^\omega$ if there are $p \in A^*$ and $s' \in A^\omega$ such that $s = pvs'$.

If $w \in A^*$, then $F(w)$ denotes the set of all factors of w. If $c \in A^\omega$, then $F(c)$ denotes the set of all factors of c.

If $u = a_1 a_2 \cdots a_n \in A^*$, where $a_i \in A$, $1 \le i \le n$, then we say w has *length* n and write $|w| = n$. (The empty word has length 0.)

Definition 1. Let $c \in A^\omega$ and let u be a factor of c. We define

$$k(c, u) = \sup\{|v| : v \in F(c) \text{ and } u \notin F(v)\}.$$

Thus if (and only if) u is "missing" from arbitrarily long factors of c, we have $k(c, u) = \infty$.

Definition 2. If $c \in A^\omega$, $u \in F(c)$, and $k(c, u) < \infty$, that is, if every sufficiently long factor w of c contains u as a factor, we say that the factor u of $c \in A^\omega$ is *uniformly recurrent* (in c). If *every* factor u of c is uniformly recurrent in c, then we say that c itself is *uniformly recurrent*.

First we need what is essentially König's lemma.

Lemma 1. *Let L be any infinite subset of A^*, where A is a finite set. Then there is an infinite word $t \in A^\omega$ such that each factor of t is a factor of infinitely many words of L.*

Proof. Since A is finite, there is a letter in A, call it $t(1)$, which is the first letter in each word of an infinite subset L_1 of L. Similarly, there is a letter in A, call it $t(2)$, such that $t(1)t(2)$ are the first two letters of each word of an infinite subset L_2 of L_1. Continuing in this way, we produce an infinite word $t = t(1)t(2)t(3)\cdots \in A^\omega$ such that each "prefix" $t(1)t(2)t(3)\cdots t(n)$ of t is a prefix of an infinite subset L_n of $L_{n-1} \subseteq \cdots \subseteq L_2 \subseteq L_1 \subseteq L$. Since each factor of t is a factor of a prefix of t, every factor of t is a factor of infinitely many words of L. □

Definition 3. Let A be a finite set, and let $c \in A^\omega$. We define a sequence of infinite words $t_0, t_1, t_2, \ldots$ inductively as follows. We set $t_0 = c$. Let the factors of c be $F(c) = \{w_1, w_2, w_3, \ldots\}$. (This is an arbitrary enumeration.) For $r > 0$, assume $t_0, t_1, t_2, \ldots, t_{r-1}$ have been defined. Let E_r be the set of all those factors of t_{r-1} which do not contain w_r as a factor. Thus

$$E_r = \{v \in F(t_{r-1}) : w_r \notin F(v)\}.$$

If E_r is finite, we set $t_r = t_{r-1}$.

If E_r is infinite, we obtain t_r by using Lemma 1 in the following way. We set $L = E_r$ in the hypothesis of Lemma 1 and conclude (by Lemma 1) there is an infinite word $t_r \in A^\omega$ such that each factor of t_r is a factor of (infinitely many) words of E_r. (Thus $F(t_r) \subseteq E_r$.)

This concludes our definition of $\{t_r\}_{r=0}^{\infty}$.

Lemma 2. *For all $r > 0$, if E_r is finite, then $k(t_{r-1}, w_r) < \infty$.*

Proof. By Definition 1, $k(t_{r-1}, w_r) = \sup\{|v| : v \in F(t_{r-1}), w_r \notin F(v)\} = \sup\{|v| : v \in E_r\}$. □

Lemma 3. *We have*

$$\cdots \subseteq F(t_r) \subseteq F(t_{r-1}) \subseteq \cdots \subseteq F(t_2) \subseteq F(t_1) \subseteq F(t_0) = F(c).$$

Proof. If E_r is finite, then $t_r = t_{r-1}$. If E_r is infinite, then $F(T_r) \subseteq E_r \subseteq F(t_{r-1})$. □

Lemma 4. *If $w_r \in F(t_r)$, then $k(t_r, w_r) < \infty$.*

Proof. By the definition of E_r, $w_r \notin E_r$. If E_r is infinite, then by Definition 3, $F(t_r) \subseteq E_r$. Thus, if E_r is infinite, then $w_r \notin F(t_r)$. Therefore, $w_r \in F(t_r)$ implies E_r is finite. Since E_r is finite, Lemma 2 gives $k(t_{r-1}, w_r) < \infty$ and Definition 3 gives $t_r = t_{r-1}$, hence $k(t_r, w_r) < \infty$. □

We are now ready to prove the crucial Theorem 1. Recall that the term "uniformly recurrent" is defined in Definition 2.

Theorem 1. *Given an arbitrary infinite word $c \in A^\omega$, there exists an infinite word $s \in A^\omega$ such that s is uniformly recurrent and $F(s) \subseteq F(c)$.*

Proof. We make use of the sequence of infinite words $t_0, t_1, t_2, \ldots$ defined in Definition 3. For each $i \geq 1$, let u_i be any factor of t_i with length i:

$$u_i \in F(t_i), \quad |u_i| = i.$$

Using Lemma 1, we let $s \in A^\omega$ be any infinite word such that every factor of s is a factor of infinitely many u_i. From Lemma 3, it is clear that $F(s) \subseteq F(c)$. In fact, something stronger is true, namely

$$F(s) \subseteq \bigcap \{F(t_j) : j \geq 0\}.$$

To see this, let $w \in F(s)$, and let j be arbitrary. Since w is a factor of infinitely many u_i, choose $j_0 \geq j$ such that w is a factor of u_{j_0}. Since $u_{j_0} \in F(t_{j_0})$, we have $w \in F(t_{j_0}) \subseteq F(t_j)$.

Now we can show that every $w \in F(s)$ is uniformly recurrent (in s). Since $F(s) \subseteq F(c)$, let $w = w_r$. We have just seen that $w_r \in F(t_r)$, and by Lemma 4, this gives

$k(t_r, w_r) < \infty$. Finally, since $F(s) \subseteq F(t_r)$, we have $k(s, w_r) \leq k(t_r, w_r) < \infty$, which means (Definition 2) that w_r is uniformly recurrent. □

3 The first extension

Theorem 2. *Let c be an arbitrary finite coloring of $\mathbb{N}$. Then for every $k \geq 2$ there exists a sequence of positive integers $d_1 < d_2 < d_3 < \cdots$ such that for all $n \geq 1$, infinitely many translates of the set*

$$P_n = \left\{ \sum_{i=1}^{n} \epsilon_i d_i : 0 \leq \epsilon_1, \ldots, \epsilon_n \leq k-1 \right\}$$

are monochromatic. The $\{d_i\}$ depend only on k (and the coloring), and can be chosen so that $|P_n| = k^n$.

Proof. As mentioned in the Introduction, we regard the coloring c as an infinite word $c = c(1)c(2)c(3)\cdots \in A^\omega$. Here, A is the set of "colors," and $c(n)$ is the color assigned to n, for all $n \in \mathbb{N}$. Throughout the proof, $k \geq 2$ is fixed.

Using Theorem 1, we let s be any infinite word $s \in A^\omega$ such that s is uniformly recurrent and $F(s) \subseteq F(c)$. To prove Theorem 2, we only need to show (for each n) that a single translate of P_n is monochromatic under the coloring s. Suppose that $[p, q]$ is an interval, which contains a translate $\overline{P_n}$ of P_n, and that the coloring s, restricted to $[p, q]$, is constant on $\overline{P_n}$; then the word

$$w = s(p)s(p+1)s(p+2)\cdots s(q)$$

occurs infinitely often in the word s, and hence occurs infinitely often in the word c (since $F(s) \subseteq F(c)$). Thus c is constant on infinitely many translates of P_n.

Let $k \in \mathbb{N}$ be fixed throughout the remainder of this argument.

We use induction on n. For $n = 1$, let a be any element of A which occurs in s.

Since s is uniformly recurrent, there are $D, x_1, x_2, x_3, \cdots \in \mathbb{N}$ with $x_1 < x_2 < x_3 < \cdots$ and $x_{j+1} - x_j \leq D$ for all $j \geq 1$, such that the word a (of length 1) occurs at each of the positions $x_1, x_2, x_3, \cdots$ in s (i. e., s is constant on $\{x_1, x_2, x_3, \ldots\}$).

Let $V = \{x_1, x_2, x_3, \cdots\}$. Then $V \cup (V+1) \cup (V+2) \cup \cdots \cup (V+D-1) = [x_1, \infty)$. Now set $V_0 = V$, $V_1 = (V+1) - V_0$, $V_2 = (V+2) - (V_0 \cup V_1)$, $\ldots$, $V_{D-1} = V + (D-1) - (V_0 \cup V_1 \cup \cdots \cup V_{D-2})$, to obtain a partition $V_0, V_1, V_2, \ldots, V_{D-1}$ (or "coloring with D colors") of $[x_1, \infty)$. By van der Waerden's theorem on arithmetic progressions, some V_i contains a k-term arithmetic progression. Since V_i is a translate of V, V itself contains a k-term arithmetic progression.

Thus there exists d_1 such that s is constant on a translate of

$$P_1 = \{\epsilon_1 d_1 : 0 \leq \epsilon_1 \leq k-1\}.$$

For the induction step, let $n \geq 1$ and assume that $d_1 < d_2 < \cdots < d_n$ exist so that s is constant on a translate of

$$P_n = \left\{ \sum_{i=1}^{n} \epsilon_i d_i : 0 \leq \epsilon_1, \ldots, \epsilon_n \leq k-1 \right\},$$

and $|P_n| = k^n$. To be specific, assume that s is constant on the set

$$\overline{P_n} = m + P_n.$$

The set $\overline{P_n}$ is contained in the interval $[m,\ m + (k-1)(d_1 + d_2 + \cdots + d_n)]$. Let $q = (k-1)(d_1 + d_2 + \cdots + d_n)$, so that $\min \overline{P_n} = m$, $\max \overline{P_n} = m + q$, and $\overline{P_n}$ is contained in the interval $[m,\ m+q]$.

Let $w = s(m)s(m+1)s(m+2)\cdots s(m+q)$. Thus w is a certain factor of s with length $|w| = q+1$. Moreover, $\overline{P_n}$ is contained in $[m, m+q]$, and s is constant on $\overline{P_n}$. We will use the abbreviated notation

$$w = s[m, m+q].$$

Since s is uniformly recurrent, there are $D, x_1, x_2, x_3, \cdots \in \mathbb{N}$ with $x_1 < x_2 < x_3 < \cdots$ and $x_{j+1} - x_j \leq D$ for all $j \geq 1$, such that the factor w *begins* at each of the positions $x_1, x_2, x_3, \cdots$ in s.

Again using van der Waerden's theorem on arithmetic progressions, there exists d_{n+1} such that $x_1, x_2, x_3, \cdots$ contains a translate of the arithmetic progression $\{\epsilon_{n+1} d_{n+1} : \epsilon_{n+1} \in \{0, 1, \ldots, k-1\}\}$. We can assume that $d_{n+1} > q = (k-1)(d_1 + \cdots + d_n)$.

Thus for some $m' \in \mathbb{N}$, the factor w begins at each of the positions $m', m'+d_{n+1}, m'+2d_{n+1}, \ldots, m' + (k-1)d_{n+1}$. We can assume that $m' > m$.

For $0 \leq i \leq k-1$, let I_i denote the interval $I_i = [m' + id_{n+1}, m' + id_{n+1} + q]$. Since the factor $w = s[m, m+q]$ begins at each number $m' + id_{n+1}$, we have that for $0 \leq i \leq k-1$,

$$w = s[m, m+q] = s(I_i).$$

Since $I_i = (m' - m) + id_{n+1} + [m, m+q]$ and s is constant on $\overline{P_n} \subseteq [m, m+q]$, it follows that s is constant on $(m' - m) + id_{n+1} + \overline{P_n} = m' + id_{n+1} + P_n \subseteq I_i$.

Thus s is constant on

$$\bigcup \{m' + id_{n+1} + P_n : 0 \leq i \leq k-1\} = m' + P_{n+1}.$$

Since $d_{n+1} > q = (k-1)(d_1 + \cdots + d_n)$, the intervals $I_i, 0 \leq i \leq k-1$, are pairwise disjoint, hence $|P_{n+1}| = k|P_n| = k^{n+1}$. □

Theorem 3. *Let c be an arbitrary finite coloring of $\mathbb{N}$. Then for every sequence of positive integers $\{k_i\}_{i=1}^{\infty}$ there exists a sequence of positive integers $d_1 < d_2 < d_3 < \cdots$ such that for all $n \geq 1$, infinitely many translates of the set*

$$P_n = \left\{ \sum_{i=1}^{n} \epsilon_i d_i : \epsilon_i \in \{0, 1, \ldots, k_i - 1\} \right\}$$

are monochromatic. Furthermore, the d_i can be chosen so that $|P_n| = k_1 k_2 \cdots k_n$.

Proof. Trivial modifications of the proof of Theorem 2 give a proof of Theorem 3. □

4 The second extension

Let $W(l, r)$, i. e., the van der Waerden number, be the least positive integer m such that if $[1, m]$ is r-colored, then there exist a, d such that $a, a+d, \ldots, a+(l-1)d$ have the same color.

Theorem 4. *Let n, k, r be positive integers. There exists a least positive integer $m = WB_n(k, r)$ such that if $[1, m]$ is r-colored, there exist $a, d_1, d_2, \ldots, d_n$ such that $d_1 < d_2 < \cdots < d_n$, $\mathcal{P}_n = \{a + x_1 d_1 + \cdots + x_n d_n : 0 \leq x_1, \ldots, x_n \leq k-1\} \subseteq [1, m]$, $|\mathcal{P}_n| = k^n$, and $\mathcal{P}_n \cup \{d_1, \ldots, d_n\}$ is monochromatic.*

Proof. The proof is by double induction on n, r. The case $r = 1$ is trivial, so let $r \geq 2$ and the nonnegative integer n be also fixed through the reminder of the proof. Now fix a positive integer k and assume that the following numbers exist: $WB_n(l, r)$ for all l and $WB_{n+1}(k, r-1)$ (note that for $n = 0$ we put $WB_0(l, r) = W(l, r)$). We will show that $WB_{n+1}(k, r)$ exists.

Let

$$m = WB_n\big(WB_{n+1}(k, r-1) \cdot k(k-1) + k, r\big).$$

We show that $WB_{n+1}(k, r) \leq m$, which will complete the argument.

Let $[1, m]$ be r-colored, using the colors $\{1, 2, \ldots, r\}$. By the definition of m, there are $a, d_1, \ldots, d_n$ such that $d_1 < d_2 < \cdots < d_n$,

$$\mathcal{Q}_n = \{a + y_1 d_1 + \cdots + y_n d_n : 0 \leq y_1, \ldots, y_n \leq WB_{n+1}(k, r-1) \cdot k(k-1) + k - 1\} \subseteq [1, m],$$

and $\mathcal{Q}_n \cup \{d_1, \ldots, d_n\}$ is monochromatic, say with color r. Now consider the following (monochromatic) subsets of $\mathcal{Q}_n$:

$$\mathcal{Q}_{n+1}^1 = \{a + x_1 d_1 + \cdots + x_n d_n + x_{n+1}[k(d_1 + \cdots + d_n)] : 0 \leq x_1, \ldots, x_{n+1} \leq k-1\}$$
$$\mathcal{Q}_{n+1}^2 = \{a + x_1 d_1 + \cdots + x_n d_n + x_{n+1}[2k(d_1 + \cdots + d_n)] : 0 \leq x_1, \ldots, x_{n+1} \leq k-1\}$$

$$\mathcal{Q}_{n+1}^{3} = \{a + x_1 d_1 + \cdots + x_n d_n + x_{n+1}[3k(d_1 + \cdots + d_n)] : 0 \le x_1, \ldots, x_{n+1} \le k-1\}$$
$$\cdots$$
$$\mathcal{Q}_{n+1}^{w} = \{a + x_1 d_1 + \cdots + x_n d_n + x_{n+1}[wk(d_1 + \cdots + d_n)] : 0 \le x_1, \ldots, x_{n+1} \le k-1\}$$

where $w = WB_{n+1}(k, r-1)$.

(One may note that $\max \mathcal{Q}_{n+1}^{w} = \max \mathcal{Q}_n$.)

There are now two cases.

If at least one of $\{k(d_1 + \cdots + d_n), 2k(d_1 + \cdots + d_n), \ldots, wk(d_1 + \cdots + d_n)\}$, say $j_0 k(d_1 + \cdots + d_n)$, has the color r, we are done by setting $d_{n+1} = j_0 k(d_1 + \cdots + d_n)$, for then

$$\mathcal{Q}_{n+1}^{j_0} \cup \{d_1, \ldots, d_{n+1}\}$$

is our desired monochromatic set.

If none of $\{k(d_1 + \cdots + d_n), 2k(d_1 + \cdots + d_n), \ldots, wk(d_1 + \cdots + d_n)\}$ have color r, then recalling that $w = WB_{n+1}(k, r-1)$, all elements of

$$k(d_1 + \cdots + d_n)[1, WB_{n+1}(k, r-1)]$$

have been colored with $r-1$ colors, and we are done by the induction hypothesis on r. □

5 Remarks and addendum

1. In proving Theorem 2, we do not actually need the full strength of Theorem 1, which says that each factor w of the infinite word s occurs *syndetically* in s. In order to apply van der Waerden's theorem, it is enough to know that each factor w of the infinite word s occurs *piecewise syndetically* in s. This means that for each factor w of the infinite word s, there exists a fixed $d = d(w)$ such that for arbitrarily large m, there is a factor $q_1 q_2 \cdots q_m$ of s with $|q_i| = d$, $1 \le i \le m$, such that w is a factor of each q_i, $1 \le i \le m$.
2. Theorem 2 has the same relation with van der Waerden's theorem as Carlson–Simpson's theorem [2] has with Hales–Jewett's theorem. It is possible to deduce Theorem 2 from Carlson–Simpson's theorem by generalizing the usual proof of deducing van der Waerden's theorem from the Hales–Jewett theorem. Similarly, Theorem 3 can be deduced from a generalized version of Carlson–Simpson's theorem [3, Section 10].
3. As the proof given for Theorem 4 uses a double induction argument, it gives no primitive recursive upper bound for $WB_n(k, r)$. We noticed that Theorem 4 is a special case of the theorem on the existence of monochromatic (m, p, c)-sets [6, Chapter 3, Theorem 10] for $c = 1$. Checking the standard proof given there shows that in

fact $WB_n(k, r)$ has a primitive recursive bound belonging to the class of WOW functions [6, 2.7]. On the other hand by a straightforward modification of our proof for Theorem 4, we get a new and simpler proof of the above mentioned theorem on (m, p, c)-sets.

Bibliography

[1] T. C. Brown, F. R. K. Chung, P. Erdős and R. L. Graham, Quantitative Forms of a Theorem of Hilbert, *J. Comb. Theory, Ser. A*, **38** (1985), 210–216.

[2] T. J. Carlson and S. G. Simpson, A dual form of Ramsey's theorem, *Adv. Math.*, **53** (1984), 265–290.

[3] T. J. Carlson, Some unifying principles in Ramsey theory, *Discrete Math.*, **68** (1988), 117–169.

[4] H. Furstenberg, Poincaré Recurrence and Number Theory, *Bull., New Ser., Am. Math. Soc.*, **5** (1981), 211–234.

[5] R. R. Graham, *Rudiments of Ramsey Theory*, American Mathematical Society, Providence, 1983.

[6] R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey Theory*, 2nd ed., John Wiley & Sons, New York, 1990.

[7] A. W. Hales and R. I. Jewett, Regularity and Positional Games, *Trans. Am. Math. Soc.*, **106** (1963), 222–229.

[8] D. Hilbert, Über die Irreducibilität Ganzer Rationaler Functionen mit Ganzzahligen Coefficienten, *J. Reine Angew. Math.*, **110** (1892), 104–129.

[9] J. Justin and G. Pirillo, Shirshov's Theorem and ω-Permutability of Semigroups, *Adv. Math.*, **87** (1991), 151–159.

[10] A. de Luca and S. Varricchio, *Finiteness and Regularity in Semigroups and Formal Languages*, Springer, Berlin, 1999.

[11] M. B. Villarino, W. Gasarch and K. W. Regan, Hilbert's Proof of his Irreducibility Theorem, *Am. Math. Mon.*, **125** (2018), 513–530.

[12] B. L. van der Waerden, Beweis einer Baudet'schen Vermutung, *Nieuw Arch. Wiskd.*, **15** (1927), 212–216.

[13] B. L. van der Waerden, How the proof of Baudet's conjecture was found, in *Studies in Pure Mathematics*, pp. 251–260, Academic Press, New York, 1971.

Mark Budden

The Gallai–Ramsey number for a tree versus complete graphs

Dedicated to the memory of Ron Graham

Abstract: For a collection of graphs $G_1, G_2, \ldots, G_t$, the Gallai–Ramsey number

$$gr(G_1, G_2, \ldots, G_t)$$

is the least positive integer p such that every Gallai t-coloring of the edges of K_p contains a subgraph isomorphic to G_i spanned by edges in color i, for some $1 \leq i \leq t$. This note focuses on the evaluation of the Gallai–Ramsey number

$$gr(T, K_{s_1}, K_{s_2}, \ldots, K_{s_t}),$$

where T is a tree. We offer several exact evaluations that build off of known results and conclude with an overview of critical colorings for such Gallai–Ramsey numbers.

1 Introduction

Gallai–Ramsey numbers are a common variation of graph Ramsey numbers. Their name is derived from the close connection that rainbow triangle-free colorings share with Gallai's foundational paper [8] on transitively orientable graphs (comparability graphs). An English translation of [8] by F. Maffray and M. Preissmann can be found in [13]. This note focuses on the evaluation of the Gallai–Ramsey number for a tree versus a collection of complete graphs, and a description of the critical colorings associated with this number. We begin with an overview of the terminology and background required for our investigation.

If G is a simple graph (avoiding loops and multiedges), we denote by $V(G)$ and $E(G)$ its vertex and edge sets, respectively. A *t-coloring* of G is a function

$$c : E(G) \longrightarrow \{1, 2, \ldots, t\}.$$

In general, we do not assume that a t-coloring is surjective. A *Gallai t-coloring* is a t-coloring that avoids rainbow triangles. That is, there are no instances of distinct vertices x, y, and z such that $|\{c(xy), c(yz), c(xz)\}| = 3$. When $t = 1$ or $t = 2$, observe that every t-coloring is a Gallai t-coloring.

Mark Budden, Department of Mathematics and Computer Science, Western Carolina University, Cullowhee, NC, USA, e-mail: mrbudden@email.wcu.edu

https://doi.org/10.1515/9783110754216-007

If $G_1, G_2, \ldots, G_t$ are graphs, then the *Ramsey number* $r(G_1, G_2, \ldots, G_t)$ is defined to be the least positive integer p such that every t-coloring of the complete graph K_p of order p contains a subgraph isomorphic to G_i spanned by edges in color i, for some $1 \leq i \leq p$. The existence of Ramsey numbers follows from the ubiquitous theorem of Frank Ramsey [14]. Analogously, the *Gallai–Ramsey number* $gr(G_1, G_2, \ldots, G_t)$ is the least positive integer p such that every Gallai t-coloring of K_p contains a subgraph isomorphic to G_i spanned by edges in color i, for some $1 \leq i \leq t$. Since every Gallai t-coloring is a t-coloring, it follows that

$$gr(G_1, G_2, \ldots, G_t) \leq r(G_1, G_2, \ldots, G_t).$$

If $G = G_1 = G_2 = \cdots = G_t$, then we write $gr^t(G)$ for the corresponding t-color Gallai–Ramsey number. Most research on Gallai–Ramsey numbers has focused on the "diagonal" case $gr^t(G)$ (e. g., see [2, 5, 7, 9, 11]). One of the earliest known results in this area is due to Chung and Graham [2], where in 1983, they proved a result equivalent to the statement

$$gr^t(K_3) = \begin{cases} 5^{t/2} + 1 & \text{if } t \text{ is even} \\ 2 \cdot 5^{(t-1)/2} + 1 & \text{if } t \text{ is odd,} \end{cases}$$

whenever $t \geq 2$. This result will prove to be useful to us in Section 2.

Recall that a *tree* T is a connected acyclic graph. Throughout the remainder of this note, assume that T_m is any tree of order m. In 1972, Chvátal and Harary [4] proved a general lower bound for 2-color Ramsey numbers that implied

$$r(T_m, K_n) \geq (m-1)(n-1) + 1.$$

Five years later, Chvátal [3] was able to complete the proof that

$$r(T_m, K_n) = (m-1)(n-1) + 1. \tag{7.1}$$

Our main result concerns the evaluation of the $(t+1)$-colored Gallai–Ramsey number $gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t})$. Specifically, in Theorem 1, we prove that

$$gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) = (m-1)\big(gr(K_{s_1}, K_{s_2}, \ldots, K_{s_t}) - 1\big) + 1.$$

Known evaluations of $gr(K_{s_1}, \ldots, K_{s_t})$ then allow us to obtain explicit evaluations. Finally, we consider the critical colorings for $gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t})$ and discuss the "goodness" of graphs in this setting.

2 The evaluation of $gr(T, K_{s_1}, K_{s_2}, \ldots, K_{s_t})$

We begin this section with the main result of this note.

Theorem 1. *Let $t \geq 2$ and $m \geq 1$. Then*

$$gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) = (m-1)\big(gr(K_{s_1}, \ldots, K_{s_t}) - 1\big) + 1.$$

Proof. Let $n = gr(K_{s_1}, K_{s_2}, \ldots, K_{s_t})$ and fix a Gallai t-coloring of K_{n-1} that avoids a monochromatic copy of K_{s_i} in color i, for all $1 \leq i \leq t$. Replace each of the vertices in this K_{n-1} with complete red copies of K_{m-1} to form a $(t+1)$-colored $K_{(m-1)(n-1)}$. Clearly, no red T_m exists since the largest red component only contains $m-1$ vertices. The largest complete subgraph in colors other than red contain at most one vertex from each K_{m-1}, so this construction lacks monochromatic copies of K_{s_i} in colors $1 \leq i \leq t$. It is also easy to verify that the resulting coloring is a Gallai coloring. It follows that

$$gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) \geq (m-1)(n-1) + 1.$$

To prove the other direction, consider a Gallai $(t+1)$-coloring of $K_{(m-1)(n-1)+1}$. If we identify the last t colors together, we obtain a 2-coloring of $K_{(m-1)(n-1)+1}$. By equation (7.1), it follows that there is a red T_m or a copy of K_n spanned by edges using only colors $1 \leq i \leq t$. In the former case, we are done. In the latter case, the K_n is Gallai t-colored, and since $gr(K_{s_1}, K_{s_2}, \ldots, K_{s_t}) = n$, it follows that there is a monochromatic copy of K_{s_i} in color i, for some $1 \leq i \leq t$. Hence

$$gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) \leq (m-1)(n-1) + 1,$$

completing the proof of the theorem. □

When $t = 2$, observe that $r(K_{s_1}, K_{s_2}) = gr(K_{s_1}, K_{s_2})$. This allows us to apply known nontrivial 2-color classical Ramsey numbers to obtain 3-color Gallai–Ramsey numbers (see Section 2.1 of [12]). A list of these results are contained in Table 7.1.

Next, we apply Chung and Graham's result [2]:

$$gr^t(K_3) = \begin{cases} 5^{t/2} + 1 & \text{if } t \text{ is even} \\ 2 \cdot 5^{(t-1)/2} + 1 & \text{if } t \text{ is odd.} \end{cases}$$

Theorem 1 implies that the $(t+1)$-color Gallai–Ramsey number satisfies

$$gr(T_m, \underbrace{K_3, \ldots, K_3}_{t \text{ terms}}) = \begin{cases} (m-1)5^{t/2} + 1 & \text{if } t \text{ is even} \\ 2(m-1)5^{(t-1)/2} + 1 & \text{if } t \text{ is odd.} \end{cases}$$

Table 7.1: Gallai–Ramsey numbers that follow from the known nontrivial 2-color classical Ramsey numbers compiled in Radziszowski's dynamic survey [12].

$r(K_{s_1}, K_{s_2})$	$gr(T_m, K_{s_1}, K_{s_2})$
$r(K_3, K_3) = 6$	$gr(T_m, K_3, K_3) = 5m - 4$
$r(K_3, K_4) = 9$	$gr(T_m, K_3, K_4) = 8m - 7$
$r(K_3, K_5) = 14$	$gr(T_m, K_3, K_5) = 13m - 12$
$r(K_3, K_6) = 18$	$gr(T_m, K_3, K_6) = 17m - 16$
$r(K_3, K_7) = 23$	$gr(T_m, K_3, K_7) = 22m - 21$
$r(K_3, K_8) = 28$	$gr(T_m, K_3, K_8) = 27m - 26$
$r(K_3, K_9) = 36$	$gr(T_m, K_3, K_9) = 35m - 34$
$r(K_4, K_4) = 18$	$gr(T_m, K_4, K_4) = 17m - 16$
$r(K_4, K_5) = 25$	$gr(T_m, K_4, K_5) = 24m - 23$

Similarly, the recent evaluation

$$gr^t(K_4) = \begin{cases} 17^{t/2} + 1 & \text{if } t \text{ is even} \\ 3 \cdot 17^{(t-1)/2} + 1 & \text{if } t \text{ is odd,} \end{cases}$$

by Liu, Magnant, Saito, Schiermeyer, and Shi [11] implies that

$$gr(T_m, \underbrace{K_4, \ldots, K_4}_{t \text{ terms}}) = \begin{cases} (m-1)17^{t/2} + 1 & \text{if } t \text{ is even} \\ 3(m-1)17^{(t-1)/2} + 1 & \text{if } t \text{ is odd.} \end{cases}$$

A well-known conjecture of Fox, Grinshpun, and Pach (Conjecture 1.7 of [6]) states that

$$gr^t(K_n) = \begin{cases} (r(K_n, K_n) - 1)^{t/2} + 1 & \text{if } t \text{ is even} \\ (n-1)(r(K_n, K_n) - 1)^{(t-1)/2} + 1 & \text{if } t \text{ is odd,} \end{cases}$$

which, if proved, would imply a similar result as in the cases $n = 3, 4$.

3 Critical colorings and good graphs

The construction given in the proof of Theorem 1 to obtain the lower bound for $gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t})$ turns out to be the only such construction. To be precise, if $p = gr(G_1, G_2, \ldots, G_t)$, then a *critical coloring* of K_{p-1} is a t-coloring that lacks a subgraph isomorphic to G_i spanned by edges in color i, for all $1 \le i \le t$. To determine a critical coloring for $gr(T_m, K_{s_1}, K_{s_2} \ldots, K_{s_t})$, let $n = gr(K_{s_1}, K_{s_2}, \ldots, K_{s_t})$ and identify the last t-colors together. We know from Theorem 1 and equation (7.1) that

$$gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) = (m-1)(n-1) + 1 = r(T_m, K_n).$$

It was proved by Hook and Isaak (Proposition 2.4 of [10]) that the only critical colorings for $r(T_m, K_n)$ are formed by taking a blue K_{n-1} and replacing each of its vertices with a red K_{m-1}. Thus the only critical colorings for $gr(T_m, K_{s_1}, K_{s_2}, \ldots, K_{s_t})$ are formed by taking a Gallai t-coloring of K_{n-1} that lacks a subgraph isomorphic to K_{s_i} in color i, for all $1 \le i \le t$, and replacing each vertex with a red copy of K_{m-1}.

Since every connected graph G contains a spanning tree, it follows that if G has order m, then

$$gr(G, K_{s_1}, K_{s_2}, \ldots, K_{s_t}) \ge (m-1)\big(gr(K_{s_1}, \ldots, K_{s_t}) - 1\big) + 1. \tag{7.2}$$

Building on the concept of "goodness" introduced by Burr and Erdős [1], we say that G is *Gallai-*$\{K_{s_1}, K_{s_2}, \ldots, K_{s_t}\}$*-good* if equality holds in inequality (7.2). At the present time, the determination of which G are Gallai-$\{K_{s_1}, K_{s_2} \ldots, K_{s_t}\}$-good is an open problem. A good starting point for investigating this problem is motivated by the work of Chung and Graham [2]: identify the connected graphs G of order m that satisfy

$$gr(G, \underbrace{K_3, \ldots, K_3}_{t \text{ terms}}) = \begin{cases} (m-1)5^{t/2} + 1 & \text{if } t \text{ is even} \\ 2(m-1)5^{(t-1)/2} + 1 & \text{if } t \text{ is odd.} \end{cases}$$

Bibliography

[1] S. Burr and P. Erdős, Generalizations of a Ramsey-theoretic result of Chvátal, *J. Graph Theory*, **7** (1983), 39–51.

[2] F. Chung and R. Graham, Edge-colored complete graphs with precisely colored subgraphs, *Combinatorica*, **3** (1983), 315–324.

[3] V. Chvátal, Tree-complete graph Ramsey numbers, *J. Graph Theory*, **1** (1977), 93.

[4] V. Chvátal and F. Harary, Generalized Ramsey theory for graphs III. Small off-diagonal numbers, *Pac. J. Math.*, **41** (1972), 335–345.

[5] R. Faudree, R. Gould, M. Jacobson and C. Magnant, Ramsey numbers in rainbow triangle free colorings, *Australas. J. Comb.*, **46** (2010), 269–284.

[6] J. Fox, A. Grinshpun and J. Pach, The Erdős–Hajnal conjecture for rainbow triangles, *J. Comb. Theory, Ser. B*, **111** (2015), 75–125.

[7] S. Fujita, C. Magnant and K. Ozeki, Rainbow generalizations of Ramsey theory – a dynamic survey, Theory Appl. Graphs (2014). https://doi.org/10.20429/tag.2014.000101.

[8] T. Gallai, Transitiv orientierbare graphen, *Acta Math. Acad. Sci. Hung.*, **18** (1967), 25–66.

[9] A. Gyárfás and G. Simonyi, Edge colorings of complete graphs without tricolored triangles, *J. Graph Theory*, **46**(3) (2004), 211–216.

[10] J. Hook and G. Isaak, Star-critical Ramsey numbers, *Discrete Appl. Math.*, **159** (2011), 328–334.

[11] H. Liu, C. Magnant, A. Saito, I. Schiermeyer and Y. Shi, Gallai–Ramsey number for K_4, *J. Graph Theory*, **94** (2020), 192–205.

[12] S. Radziszowski, Small Ramsey numbers – Revision 16, *Electron. J. Comb.*, **DS1.16** (2021), 116 pp.

[13] J. L. Ramírez Alfonsín and B. Reed, *Perfect Graphs*, John Wiley & Sons, Inc., New York, 2001.

[14] F. Ramsey, On a Problem of Formal Logic, *Proc. Lond. Math. Soc. (2)*, **30** (1929), 264–286.

Joe Buhler, Chris Freiling, Ron Graham, Jonathan Kariv,
James R. Roche, Mark Tiefenbruck, Clint Van Alten, and
Dmytro Yeroshkin

On Levine's notorious hat puzzle

Abstract: The Levine hat game requires n players, each wearing an infinite random stack of black and white hats, to guess the location of a black hat on their own head seeing only the hats worn by all the other players. They are allowed a strategy session before the game, but no further communication. The players collectively win if and only if all their guesses are correct. In this paper, we give an overview of what is known about strategies for this game, including an extended discussion of the case with $n = 2$ players (and a conjecture for an optimal strategy in this case). We also prove that V_n, the optimal value of the joint success probability in the n-player game, is a strictly decreasing function of n.

1 Introduction

In her blog in 2011, Tanya Khovanova [5] described a "hat puzzle" from Lionel Levine involving n people, each wearing a stack of infinitely many black and white hats; she also gave a problem of her own inspired by that puzzle. Although superficially recreational [1, 7], the Levine puzzle became notorious because of the difficulty of giving definitive answers to any of the questions it raised.

Acknowledgement: We have corresponded or spoken with Aaron Atlee, Larry Carter, Joseph DeVincentis, Eric Egge, Ehud Friedgut, Jerry Grossman, Gil Kalai, Tanya Khovanova, Sandy Kutin, Lionel Levine, Stephen Morris, Rob Pratt, Jay-C Reyes, Joel Rosenberg, Walter Stromquist, Alan Taylor, Dan Velleman, Stan Wagon, Peter Winkler, Chen Yan, Piotr Zielinski, and no doubt others. The eighth author received funding from Excellence of Science grant number 30950721, "Symplectic Techniques."

Joe Buhler, Reed College, Portland, OR, USA, e-mail: jpb@reed.edu
Chris Freiling, Project Inertia, San Diego, CA, USA, e-mail: chris.freiling@projectinertia.com
Ron Graham, UCSD, San Diego, CA, USA
Jonathan Kariv, Isazi Consulting, Johannesburg, South Africa, e-mail: jkariv@isaziconsulting.co.za
James R. Roche, Department of Defense, Fort Meade, MD, USA, e-mail: juggling.jim.roche@ieee.org
Mark Tiefenbruck, IDA, Center for Communications Research, La Jolla, CA, USA, e-mail: mgtiefe@ccrwest.org
Clint Van Alten, School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa, e-mail: clint.vanalten@wits.ac.za
Dmytro Yeroshkin, Géométrie Différentielle, Université Libre de Bruxelles, Brussels, Belgium, e-mail: Dmytro.Yeroshkin@ulb.ac.be

https://doi.org/10.1515/9783110754216-008

Levine hat puzzle

A team with n players has an initial strategy session, after which a referee places a stack of h hats on each player's head. Each hat is either black (a. k. a. 1) or white (a. k. a. 0). The players must name a position on their own stack, and they collectively win if and only if they all name the position of a black hat on their own head. Players cannot see the hats on their own heads, but they can see all of the other players' hats. No communication between the players is allowed after the strategy session. Each of the $n \cdot h$ hats placed by the referee is chosen by an independent flip of a fair coin (probability 1/2 of each color). Each player i must communicate a positive integer x_i to the referee (without learning any of the values x_j communicated to the referee by the other players). The players win if and only if for all i, the hat in position x_i in the stack on the i^{th} player's head is black. What (joint) strategy should the players use to maximize their chance of winning?

This puzzle seems to have arisen out of Levine's work with Tobias Friedrich [4] on fast simulations of certain growth models. It is sometimes stated with wardens/prisoners or sultans/wise men instead of referees/players.

What can the players possibly do at the strategy session? They have to agree on a collection $\{f_i\}$ of "strategy" functions f_i, one for each player, that map the possible stacks that player i might see to positive integers. Each such (joint) strategy has a probability of success (based on the coin flips that will determine hat colors). Let $V_n^{(h)}$ denote the maximum value, over all possible strategies, of the probability of success for any given n and h. It is easy to see that $V_n^{(h)}$ is nondecreasing as a function of h. Let

$$V_n := \lim_{h\to\infty} V_n^{(h)} = \sup_h V_n^{(h)}.$$

Readers who like to work on puzzles themselves before seeing hints or solutions are recommended to set this paper down immediately and prove the following three statements (which are easy, moderately challenging, and difficult, respectively):

(1) $V_n \geq 1/2^n$,
(2) $V_n \geq 1/(n+1)$,
(3) $V_n \geq c/\log(n)$ for some $c > 0$.

The notoriety of the puzzle arises from the difficulty of answering the most basic questions. In particular, no value of V_n is known exactly for any $n > 1$, and the limiting behavior of V_n for large n—which was perhaps of primary interest to Levine—is unknown. He made the following conjecture.

Conjecture 1 (Levine). The optimal success probability in the n-player game is $o(1)$ as n goes to infinity; i. e.,

$$\lim_{n\to\infty} V_n = 0.$$

If you thought we were going to answer this question or find the value of V_2, say, then you would, alas, be mistaken. The goal of this paper is to describe several results that are aimed at these two big questions, in the hope that this will spur people to answer them.

Our main results are as follows: (1) a proof that V_n is strictly decreasing, i. e.,

$$V_{n+1} < V_n \quad \text{for all } n;$$

(2) the inequalities

$$\frac{7}{20} = 0.35 \le V_2 \le 0.3616\ldots$$

(we conjecture that equality holds on the left); (3) a technique of "matrix hints" that, at least in principle, can be used to give arbitrarily good upper bounds on the V_n; (4) an analysis of what happens when a fair coin is replaced by a Bernoulli coin that yields heads with probability p; and (5) various ancillary results and data that might help someone who wants to answer any of the various open questions!

We learned about a fascinating recent preprint by Friedgut, Kalai, and Kindler [3] on the same day that they learned of ours; their paper also proves that V_n is strictly decreasing, and conjectures generalizations that situate the problem in an interesting combinatorial and graph-theoretic context.

By way of introducing some of the basic techniques that will arise later, we now focus on the case $n = 2$. Each of the players A and B (whom the reader may think of as Alice and Betty if that makes the problem seem more compelling) has a large stack of h hats on her head. Asking A to say something about the random stack of hats on her head, using only the information in the completely independent stack on B's head, seems a bit unfair; indeed, it feels like a mysterious game show, perhaps run by mathematicians with a strange sense of humor. If A and B choose random strategy functions f_A and f_B, then each has an independent probability of success $1/2$ of naming the position of a black hat on her head, so their joint probability of success is $1/4$. Of course, the trick is that they should jointly choose their strategy functions before the game so as to make their choices correlate in a useful way.

The following warm-up theorem proves weaker versions of the inequalities stated in (2) above, giving elementary precursors of ideas that will be developed later. The lower bound is straightforward and has no doubt been found by many people who have looked at the puzzle. The upper bound is trickier and has been discovered by (at least) several different people. It seems likely that Noga Alon was the first; several of us first learned of it from a letter that Walter Stromquist wrote to Ron Graham.

Before proving these bounds, we make two remarks that will be used in the proof and will be assumed at many points later in the paper.

Remark 1. The reader might wonder whether the players could do better with a probabilistic strategy. The success probability for any probabilistic strategy is a convex linear combination of those for deterministic strategies, so (assuming that the source of randomness is uncorrelated with the hat placements) there is always a deterministic strategy that is at least as good as any probabilistic strategy, and it suffices to consider deterministic strategies throughout. (This is also true if the players receive "hints" as in the proof below).

Remark 2. It is important to remember that V_n is defined as a limit of success probabilities for finite stacks of h hats. For instance, in the proof of the next theorem there is a stack of h hats on each player's head, and the case in which some player has hats of only one color can be ignored. This case has probability at most $2n/2^h$ for any given n, which vanishes in the limit as h goes to infinity. So we can, variously, speak of finite or infinite stacks of hats, but must always remember that the infinite case is defined as a limit of finite cases. If strategies are actually allowed to use functions defined on infinite sets, the situation is entirely different. Any reader who can really see an infinite stack of hats all at once (as well as perform computations on infinite sets in finite time), and who believes in the Axiom of Choice and is not squeamish about nonmeasurable sets, is advised to read Appendix A. In that Appendix, we consider the 1-person (!) version of the game and describe a strategy for which the player is "virtually guaranteed" to win.

Theorem 1. *The optimal probability of success, V_2, in the 2-player game is bounded as follows:*

$$\frac{1}{3} \leq V_2 \leq \frac{3}{8}.$$

Proof. The lower bound is proved by using the following "first-black" strategy: Each player finds the position of the first (lowest) black hat on her partner's head and names that same position on her own head! For example, suppose that the hat stacks begin as in the diagram below (with black hats represented by 1s and white hats by 0s).

$$\begin{array}{ccc} \vdots & & \vdots \\ 0 & & 1 \\ 1 & \rightarrow & 1 \\ 0 & \leftarrow & 1 \\ 0 & & 0 \\ A & & B \end{array}$$

Positions are numbered from 1 starting at the bottom, as is fitting for stacks of hats. Then A will say 2, B will say 3, and they will lose. (Although B happens to have a black hat on level 3, A does *not* have a black hat on level 2.)

What is the probability that they will win using the first-1 (i. e., first-black) strategy? Consider the lowest level where at least one of the players has a 1. (By Remark 2, we may assume that there is such a level.) There are 3 possible hat pairs for A and B at the critical level: 01, 10, and 11. These are equally likely, so the first-black strategy has a success probability (or, as we sometimes say, "value") of 1/3. (Wow! Even though neither player has any information about the color of any hat on her own head, the team can do significantly better with correlated strategies than by making random guesses.)

Now we turn to the upper bound. To prove it, we use the curious device of allowing the players to get a hint from the referee. (This is a simple example of the "matrix hints" technique to be described later for finding upper bounds.) This extra information certainly cannot lower the optimal success probability of A and B, since they are free to ignore the information if they wish.

Suppose that the referee takes pity on A and B and, before their strategy session, shows the players a bit string s (each bit having been determined by a flip of a fair coin) and says that just before the game he will flip a fair coin one more time and then, based on that flip, put the hat sequence corresponding either to s or to its bitwise complement s' on A's head. For instance, if s is the bit string 101110 . . ., then A knows that her hat sequence will be either $s = 101110\ldots$ or $s' = 010001\ldots$, each with probability 1/2. (Of course, after the game starts, B will actually see the sequence s or s' on A's head.) The players are given no information about the hat sequence on B's head.

Note that if A and B refrain from looking at s, this game is exactly equivalent to the originally described game: The referee is now determining A's hat sequence in a two-stage process, but in the end, for any height h, all 2^h of her possible sequences are equally likely. To establish the desired upper bound on V_2, we will show that for each possible value of s and every possible joint strategy, the players' success probability is at most 3/8.

Because of the referee's hint, both players know during the strategy session that B will see one of only two possible sequences on A's head. Thus B's strategy is completely determined by the integers $x = f_B(s)$ and $y = f_B(s')$ that she will announce according to whether she sees s or s' as A's stack. (It will turn out to be better for B to choose $x \neq y$, but the possibility that $x = y$ must also be considered.) On the other hand, A's strategy could be any function f_A of the stack that she sees on B's head.

For the rest of the proof, we suppose that a particular (though arbitrary) hint s has been given. Then there are 2 possible values of A's stack and 2^h possible values of B's stack, and all 2^{h+1} joint possibilities are equally likely. We also suppose that the team has chosen particular (though arbitrary) strategy functions f_A and f_B.

For each of the 2^h possible hat sequences $b \in \{0,1\}^h$ that B could be given, player A will choose some level $f_A(b) \in \{1, 2, \ldots, h\}$ on her own head. Whatever level she

chooses, her hat color at that level is determined by a fair coin flip independent of the coin flips used to determine s and b. Thus we have the following "atomic" probability result:

$$\mathbb{P}(B \text{ has stack } b; A \text{ chooses a black hat}) = 1/2^{h+1} \quad \text{for every } b. \tag{8.1}$$

If $x = y$, then B points to the same level whether she sees s or s' as A's stack. Thus half of all values of b (those with component $b_x = 0$) cause B to choose a white hat, in which case the team fails. The other 2^{h-1} values of b (those with $b_x = 1$) cause B to choose a black hat, in which case the team wins if and only if A chooses a black hat. It follows from equation (8.1) that

$$\begin{aligned}
&\mathbb{P}(A \text{ and } B \text{ both choose black hats}) \\
&\quad = \sum_{b:b_x=1} \mathbb{P}(B \text{ has stack } b; A \text{ chooses a black hat}) \\
&\quad = \left(2^{h-1}\right) \cdot \left(1/2^{h+1}\right) = 1/4 < 3/8.
\end{aligned}$$

If $x \neq y$, then 2^{h-2} values of b have $b_x = b_y = 0$, in which case B will point to a white hat and the team will fail. For the remaining $3 \cdot 2^{h-2}$ values of b, the team has a chance. Again using equation (8.1), we have

$$\begin{aligned}
&\mathbb{P}(A \text{ and } B \text{ both choose black hats}) \\
&\quad = \sum_{b:(b_x,b_y)\neq(0,0)} \mathbb{P}(B \text{ has stack } b; A \text{ and } B \text{ both choose black hats}) \\
&\quad \leq \sum_{b:(b_x,b_y)\neq(0,0)} \mathbb{P}(B \text{ has stack } b; A \text{ chooses a black hat}) \\
&\quad = \left(3 \cdot 2^{h-2}\right) \cdot \left(1/2^{h+1}\right) = 3/8,
\end{aligned}$$

finishing the proof of the theorem. □

We invite the reader to verify the following two claims: If the levels x and y are chosen in the above proof so that the hats in the hint s at those positions are of opposite colors, then A may be assumed without loss of generality to choose from those same two levels on her own head. Furthermore, if the two players follow a "first-black" strategy within those two levels when given the hint s, they can achieve the upper bound 3/8 on their success probability in the limit as h goes to infinity.

The main sections of this paper (a) consider strategies for 2 players in detail (giving the lower bound 7/20, and considering what happens when the hat colors are determined by biased coins), (b) use "matrix hints" to give upper bounds, including the 0.36... stated above, (c) give results for n players, and (d) give lower bounds found by computer for smallish n.

Some of the results in this paper are recent, but many of the results and techniques here arose in an extensive series of emails in 2013–2015 between a somewhat amorphous group of people that included the authors as well as others; naturally, they referred to themselves as "the Mad Hatters." The origins of this paper lie in the Hatters' desire to collect the useful information in those much earlier emails. Many results, though not essential to the main flow of this paper, may be of interest or use to someone wanting to look more deeply into details, so they are included here as Appendices. Because preliminary drafts of the current paper were developed independently by at least two different subsets of authors, there are varying conventions (e. g., in the players' appellations and genders). We have, however, tried to maintain local consistency.

At the beginning of the process of revising, polishing, and extending the paper for the sake of publication, Ron Graham (1935–2020) left us. The Levine puzzle is the kind of question that delighted him, and he was fond of challenging people with this particular problem. In addition to pushing for greater clarity and more cleverness, he also repeatedly asked for more data. Partly at his urging, a rather large number of computational experiments were done on strategies for this puzzle in 2014, by many people. The other authors dedicate this article to the memory of Ron's exuberance, mathematical and otherwise.

2 New results for two players

2.1 Proof that $V_2 \geq 7/20$

We begin our discussion of two-player strategies with the promised 7/20 lower bound.

Theorem 2. *The optimal success probability V_2 satisfies the inequality*

$$V_2 \geq \frac{7}{20}.$$

Proof. To prove the lower bound, we exhibit a strategy with value $v = 0.35 = 7/20$.

Consider the following characterization of the first-black strategy. Each player looks at the first hat on the other player's head. If he does not like what he sees, then he skips it and looks at the next hat. The two players might not skip the same number of hats, but if they do, then they have an unusually high chance of winning. We can generalize this idea by letting each player look at the first k hats before deciding whether to skip them. The simplest 7/20 strategies use $k = 3$.

Players A and B each look at the lowest three hats on their partner's head. They advance over a triple if it is all-0 or all-1 (or, as we will say, is "monochromatic") until they arrive at a nonmonochromatic triple. Since there are 2 possible monochromatic triples and 6 nonmonochromatic triples, the chance that a player will skip over a triple

is 1/4. They each will stop on a nonmonochromatic triple, and then use the following algorithm: If there is a single 1 in the triple that they see, then they announce that the corresponding bit in their own string is also a 1. (So if A skips twice and then the first bit in the next triple is a 1, A would say 7.) If there are two 1-bits and one 0-bit in the lowest nonmonochromatic triple that they see, then A names the position Above the 0-bit, and B names the position Below the 0-bit. Here, above/below in the triple are interpreted cyclically; e. g., if B skips one triple and then sees 011 (where 0, in the fourth position, is the lowest hat in that triple), B will say 6.

What is the value v of this strategy? If r denotes the probability that they win when both of their bottom triples are nonmonochromatic, then

$$v = \frac{1}{4} \cdot \frac{1}{4} \cdot v + \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{1}{4} + \frac{3}{4} \cdot \frac{3}{4} \cdot r.$$

The first term represents the case where they both see initial monochromatic triples (no harm, no foul, they just skip those and are then playing the same game). The second and third terms are the cases where one of A or B skips an initial triple but the other does not; this is the perhaps unfortunate case in which there can be no correlation because they are looking at different triples, so they are both making random uncorrelated guesses and have probability 1/4 of winning. The fourth term represents the case where neither skips the initial triple. In order to solve this equation for v, we have to calculate r.

Call a nonmonochromatic triple a *single* if it has a single 1 bit, and a *double* if it has exactly two 1 bits. There are 36 possible pairs of nonmonochromatic triples: 9 cases where both are singles, 18 in which one is a single and one is a double (in one order or the other), and 9 in which both are doubles. Immediately below, we show example pairs of all three types. (As it happens, all three examples are losing pairs for the above strategy.)

0		0	0		0	1		0
0	↔	1	1	↔	1	1	↔	1
1		0	0		1	0		1
A		B	A		B	A		B

In the 9 cases where they both have singles, they win exactly in the 3 cases where the 1s are in the same location. In the 18 cases where they have a single and a double (for one or the other order), they are both correct only in the 6 cases where the 1 bit in the single is located in exactly the right location with respect to the two 1 bits in the double. In the case when A and B both have doubles, they fail *only* in the three cases where, as in the case pictured above, the bits are exactly aligned so that they are both wrong. This means that they are both correct in all 6 of the other cases, as the reader can check by trying the other two possibilities for B when A has 110 as pictured above.

Thus

$$r = \frac{3+6+6}{36} = \frac{5}{12}.$$

Solving the earlier equation for v gives $v = 7/20$ as claimed. □

The "reset" on black (not just on white) in the 7/20 strategy for the 2-person game was counterintuitive and surprisingly difficult to find; apparently it eluded discovery for 3 years after the puzzle was popularized in 2011 on Tanya Khovanova's blog [5]. It was finally found in 2014 by a California group (Larry Carter, Jay-C Reyes, Joel Rosenberg, and M. Tiefenbruck) and by a Pennsylvania group (J. Kariv and D. Yeroshkin). The fact that the description above might be judged by some as "easy to remember" is probably a red herring, since a strategy amounts to nothing more than an arbitrary function from what a player sees to what they are supposed to do, and there is no reason that this needs to be structured in any memorable way whatsoever. In particular, there are multiple 7/20 strategies based on 3 hats (skipping monochromatic triples), and some are symmetric in the sense that both players have the same function.

It might seem paradoxical to skip over the case of 3 black hats, where one of the players is guaranteed to be right, but it seems to be necessary. One possible rationale is that it is harder to correlate with monochromatic triples, so it might be better just to skip them. A more detailed explanation is given immediately below; it emphasizes why A should reset when seeing a monochromatic triple, but the same argument applies when the roles of A and B are reversed.

If B has a white triple, the team can win only if B resets, so A should hope for the best and reset as well. (This argument is easy to find.) If B has a black triple and A has a nonmonochromatic triple, then B is not going to reset in any case, so the team will on average win half the time whether or not A resets. Resetting upon seeing a black triple thus makes a net difference only when *both* players have monochromatic triples, at least one of them black. (As we will see when considering the n-player game, the actual requirement is that at *most* one of these monochromatic "tiers" be *white*.) Within this subset of 3 cases, a shared strategy to reset on seeing a black triple (as well as on seeing a white triple, which is assumed) gives up 1 sure win and 2 sure losses in order to get 3 fresh starts at a new 3-level tier. As long as the team has a basic (nonresetting) 3-level strategy that wins strictly more than 1/3 of the time (it is 22/64 for the 2-player game), incorporating the reset on seeing a black triple is a net win.

One might expect that strategy functions based on larger numbers of hats would yield increased probability of success. We do not know for sure, but we think otherwise and have the vague intuition that with larger clumps it is harder to usefully correlate assignments of probability mass to various choices. Many attempts were made in 2014 to find better strategies, but none succeeded. On these grounds, we have come to believe that the 7/20 strategies are quite possibly optimal.

Conjecture 2. We conjecture that

$$V_2 = \frac{7}{20}.$$

2.2 The two-player game with a biased coin

The 7/20 strategy for the 2-player Levine puzzle has been a sticking point for about seven years. No one has found a better strategy or proved that it is optimal. Sometimes it is useful to change a problem when stuck, and one natural idea here is to replace the flip of a fair coin, used by the referee to determine hat color, by a biased coin which has probability p of giving a black hat. This turns out to have several virtues, one of which is to show that there are actually several distinct optimal strategies for $p = 1/2$ that are *not* equivalent to each other when $p \neq 1/2$. As is usual, we set $q = 1 - p$.

The special case with $p = \frac{a}{b}$ rational (for some $b > 2$) corresponds to the natural extension of having b hat colors, of which some (a) are considered good and the remaining ones are considered bad. The goal is then for each player to choose a hat of a good color (equivalently, for none of them to choose a bad color). The special cases of $a = 1$ and $a = b - 1$ are of particular interest as they respectively correspond to the case of a single good color and a single bad (nuclear) color.

In this section, we will consider general values of $p \in (0, 1)$ and compare the team's success probability, or value, using different strategies S. Thus we extend the notation from the previous section to consider quantities

$$V_n^{(h)}(p; S).$$

We will omit the superscript h when considering the limiting case as $h \to \infty$ and will omit the argument S when referring to the supremum over all strategies. If the argument p is also omitted, then the default value $p = 1/2$ is understood.

Using the naive strategy of each player choosing the first hat corresponding to a black hat on the other player's head, we can obtain a probability of winning of $\frac{p}{2-p}$. If each player instead chooses the first hat corresponding to a *white* hat on the partner's head, the probability of winning is $\frac{2p^2}{1+p}$.

As for the special case of $p = \frac{1}{2}$, these strategies, while easy to state and better than random, are not optimal. We construct four distinct strategies based upon the 3-hat strategy that all achieve a performance of 0.35 for the special case of $p = 0.5$. However, they all perform differently for general p.

Theorem 3. *There exist at least four distinct strategies S that achieve $V_2(1/2; S) = 7/20$ but are inequivalent for $V_2(p; S)$ when $p_2 \neq 1/2$.*

We shall construct the four strategies mentioned above and will denote them by S_1, S_2, S_3, and S_4. We provide their respective win rates here for the reader's conve-

nience:

$$V_2(p;S_1) = \frac{p(1+p+p^2+3p^3-3p^4+p^5)}{2+p+p^2+p^3-p^4};$$

$$V_2(p;S_2) = \frac{p(1-p+p^2+p^3)}{2-3p+3p^2};$$

$$V_2(p;S_3) = \frac{p(1+5p-10p^2+10p^3-5p^4+p^5)}{(2-2p+p^2)(1+p)(2-p)};$$

$$V_2(p;S_4) = \frac{p(1+7p-21p^2+35p^3-20p^4-14p^5+40p^6-48p^7+40p^8-22p^9+7p^{10}-p^{11})}{(1-p+p^2)(1+p-p^2)(2-2p+p^2)(1+p^2)(1+p)(2-p)}.$$

These combine to give a lower bound for $V_2(p)$ of

1. $\frac{p(1+p+p^2+3p^3-3p^4+p^5)}{(1+p)(2-p)(1+p^2)} \le V_2(p)$ for $p \le \frac{1}{2}$;
2. $\frac{p(1+5p-10p^2+10p^3-5p^4+p^5)}{(2-2p+p^2)(1+p)(2-p)} \le V_2(p)$ for $\frac{1}{2} \le p$.

The curve given by this theorem is provided in Figure 8.1, in the next section.

2.3 Constructing strategies

The strategies described in this subsection are based on ones found by a computer search for the game with only finitely many hats on each player's head. In particular, an exhaustive search was run to find the optimal strategy with three hats and the optimal "symmetric" strategy (i. e., both players use the same strategy) with four hats. Beyond these two cases, the authors ran hill-climbing and genetic algorithms for up to 12 hats; no strategies were found with better performance than the ones described here. In Section 2.3.1, we describe the outcome of the search of strategies with only three hats, and then in Section 2.3.2 we adapt the results to obtain the best known strategies for infinitely many hats. For related work on applying genetic algorithms to hat problems, see [2], which considers a different hat game.

2.3.1 Basic strategy for three hat levels

We denote by S_0 the symmetric strategy for 3 hat levels defined by Table 8.1. As shown in Table 8.2, strategy S_0 wins in 22 of 64 cases. The likelihood of each case depends on p as is also shown in Table 8.2. For convenience, we assume that a player who sees all hats of the same color points to the first hat on his or her own head. The columns correspond to the distribution of black hats on the first player's head and the rows to the distribution on the second player's. The cells are blank when the players lose. It is useful to note that this strategy is the unique optimal strategy for every value of p, up to reordering the hats on one or both of the players' heads.

Table 8.1: S_0, an optimal strategy on 3 hats.

Black hats	**∅**	**{1}**	**{2}**	**{1, 2}**	**{3}**	**{1, 3}**	**{2, 3}**	**{1, 2, 3}**
Picture	□□□	■□□	□■□	■■□	□□■	■□■	□■■	■■■
Choice	any	1	3	1	2	2	3	any

Table 8.2: Winning combinations for S_0 with probability of each event given in the block.

	□□□	■□□	□■□	■■□	□□■	■□■	□■■	■■■
□□□								
■□□		p^2q^4		p^3q^3				p^4q^2
□■□					p^2q^4	p^3q^3		
■■□		p^3q^3		p^4q^2		p^4q^2		p^5q
□□■			p^2q^4				p^3q^3	
■□■			p^3q^3	p^4q^2			p^4q^2	p^5q
□■■					p^3q^3	p^4q^2	p^4q^2	
■■■		p^4q^2		p^5q		p^5q		p^6

The sum of all the winning probabilities is as follows:

$$V_2^{(3)}(p; S_0) = 3p^2q^4 + 6p^3q^3 + 8p^4q^2 + 4p^5q + p^6 = 3p^2 - 6p^3 + 8p^4 - 6p^5 + 2p^6.$$

2.3.2 Adaptation to infinitely many hat levels

We give four adaptations of the above 3-hat strategy to the general game that performed well in cases of up to 12 hats for various values for p. As explained below, we assume without loss of generality that the first 3 hats considered are those in positions 1 to 3, that the second group of 3 hats occupies either positions 3 to 5 or 4 to 6, and so on.

Strategy S_1:

1. If the first three hats of the other player are not monochromatic, play the 3-hat strategy S_0.
2. If the first three hats of the other player are BBB or WWW, repeat S_1 on hats 3 through ∞.

Strategy S_2:

1. If the first three hats of the other player are not monochromatic, play the 3-hat strategy S_0.
2. If the first three hats of the other player are BBB or WWW, repeat S_2 on hats 4 through ∞.

The strategies S_3 and S_4 are constructed the same way as S_1 with different, but equivalent, symmetric 3-hat strategies. We provide those 3-hat strategies in Tables 8.3 and 8.4.

Table 8.3: 3-hat strategy that produces S_3.

Black hats	{1}	{2}	{1, 2}	{3}	{1, 3}	{2, 3}
Picture	■□□	□■□	■■□	□□■	■□■	□■■
Choice	3	2	2	1	3	1

Table 8.4: 3-hat strategy that produces S_4.

Black hats	{1}	{2}	{1, 2}	{3}	{1, 3}	{2, 3}
Picture	■□□	□■□	■■□	□□■	■□■	□■■
Choice	2	1	1	3	2	3

Remark 3. Iterating over all the basic 3-hat strategies equivalent to S_0, if one uses the shift-by-3-levels construction for S_2, the strategies will once again be equivalent (this includes the asymmetric strategy described in Theorem 2). On the other hand, if one undertakes the shift-by-2-levels construction for S_1, then the win rate of the strategy will be one of $V_2(p; S_1)$, $V_2(p; S_3)$, or $V_2(p; S_4)$.

We close this section by computing the success probability (or value) for the S_2 strategy, which has the most concise presentation. The computation for S_1, which yields our best lower bound for $p \leq 1/2$, is more complicated and is deferred to Appendix B. The computations for S_3 and S_4 are omitted, since they follow the same procedure as that for S_1.

2.4 Computing the performance of S_2

Consider the following cases:

(a) Neither player has WWW or BBB as their first three hats. The probability of this occurring and the players' winning is $3p^2q^4 + 6p^3q^3 + 6p^4q^2$ (see Table 8.2).

(b) One player has BBB as their first three hats, and the other does not have either WWW or BBB. Either of the two players can have the BBB stack, which occurs with probability of p^3, and the probability of not having WWW or BBB is $1 - p^3 - q^3$. The probability of this case occurring is therefore $2p^3(1 - p^3 - q^3)$. The probability of winning given this case is p since the player with BBB guesses correctly, and the other player chooses a hat in position 4 or greater, which has probability p of being B. Thus $2p^3(1 - p^3 - q^3)p$ is the probability of this case occurring and the players' winning.

(c) One player has WWW as their first three hats, and the other does not have either WWW or BBB. In this case, the players are certain to lose.

(d) Both players have either BBB or WWW as their first three hats, which occurs with probability $p^6 + 2p^3q^3 + q^6 = (p^3 + q^3)^2$. Then the strategy looks to the next three hats, and this repeats, giving an infinite sum with ratio $(p^3 + q^3)^2$.

The overall probability of a win is therefore

$$\begin{aligned} V_2(p; S_2) &= \frac{(3p^2q^4 + 6p^3q^3 + 6p^4q^2) + 2p^3(1 - p^3 - q^3)p}{1 - (p^3 + q^3)^2} \\ &= \frac{p(1 - p + p^2 + p^3)}{2 - 3p + 3p^2}. \end{aligned}$$

3 The matrix game and upper bounds for two players

3.1 An informal introduction to the matrix game

In the introductory section of this paper, we considered giving a simple hint to the two players. This allowed us to compute an upper bound of 3/8 for the value of the two-person game where the probability of each hat color is 1/2. In the current section, we present a scheme to construct more elaborate hints that prove better upper bounds. For variety, we refer to the players as Alice and Bob in the present subsection. In the later subsections of Section 3, we will use the more rigorous-sounding names "Player 1" and "Player 2" but will continue to think of the two players as female and male, respectively.

Recall that in Theorem 1 in the Introduction, the referee helped the players by revealing a bit string that was guaranteed to be either the first player's exact hat sequence or its bitwise complement. One way for the referee to produce a pair of complementary sequences is to randomly choose rows from the 2×2 identity matrix,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Note that in order to conform to the placing of hats on heads, all rows in this section will be indexed from the bottom up, starting with index 1. For example, choosing row 2, row 1, row 1, row 2, row 1, ...would produce the following complementary pair of hat sequences:

$$\begin{pmatrix} \vdots & \vdots \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

To construct a more elaborate hint, we could repeat this procedure using a different matrix, perhaps

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

We hope that a larger matrix will amount to a weaker hint given by the referee, which might give us a tighter upper bound. Choosing rows randomly gives us a collection of six infinite hat sequences. For example,

$$\begin{pmatrix} \vdots \\ x_4 \\ x_3 \\ x_2 \\ x_1 \end{pmatrix} \in \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

would be produced if the referee chose row 2, row 3, row 2, row 4, ... The hint comes when the referee reveals these six sequences and promises that Alice's sequence is among them.

Notice that in the example above, hats x_1 and x_3 are guaranteed to be identical, no matter which of the six sequences is chosen. More generally, when Alice chooses an index, it only matters which row was used to produce the hats at that level, and this row is known to all. Thus, once the matrix has been fixed, the game on Alice's head is finite. The referee's task is merely to choose a random column of the matrix, and Alice's decision is reduced to identifying a row of the matrix. For the players to have a chance of winning, there must be a 1 in the resulting row and column.

In the case of this 4×6 hint, it may be convenient to imagine that Alice has a single hat in the shape of a tetrahedron. The referee chooses an edge on the hat, and Alice hopes to choose a vertex adjacent to that edge.

So, the matrix hint greatly simplifies the game for Alice. But what about Bob, who receives no such hint? Bob will see which of the six columns is randomly chosen by the referee and will base his decision on this observation. Assuming a deterministic strategy, there are at most six hats that Bob will ever use. It may be fewer than six, because Bob may decide to use the same hat index for several different columns. In fact, Bob's strategy boils down to choosing some partition of the six columns of the matrix. So, for example, if the six columns are $c_1, c_2, c_3, c_4, c_5, c_6$, then Bob may decide on the partition $\{\{c_1, c_4, c_5\}, \{c_2, c_3\}, \{c_6\}\}$, which means that he will choose a certain hat—which may as well be x_1—when he sees c_1, c_4, or c_5; choose x_2 when he sees c_2 or c_3; and choose x_3 when he sees c_6. So the game for Bob also reduces to choosing from among a fixed, finite set of strategies.

Once Bob's strategy is fixed, there is an obvious best strategy for Alice, which we now describe. Alice observes the hats on Bob's head. Knowing Bob's strategy, Alice can see which of the six columns would cause Bob to succeed. Call these "winning columns." In order to survive, they need the referee to choose one of these winning columns, resulting in a 1 on Bob's head. But they also need a 1 on Alice's head, so they also need the referee to choose a column with a 1 in Alice's chosen row, whatever that may be. The probability of winning, then, is $k/6$, where k is the number of winning columns with a 1 in Alice's row. So Alice simply chooses any row that maximizes this probability. Finding the expected value of this probability over all of Bob's hat assignments gives the value of Bob's strategy.

For a concrete example, suppose once more that Bob chooses the partition $\{\{c_1, c_4, c_5\}, \{c_2, c_3\}, \{c_6\}\}$. Then Bob is using a 3-hat strategy. We consider the eight possible assignments of these three hat colors, shown in Table 8.5. So the value of this strategy is 17/48.

Table 8.5: Computing Alice's strategy.

Bob's colors	Winning columns	Best row for Alice	$\mathbb{P}$(win)
000	$\{\}$	does not matter	0/6
001	$\{c_6\}$	row 3	1/6
010	$\{c_2, c_3\}$	row 1	2/6
011	$\{c_2, c_3, c_6\}$	row 1	2/6
100	$\{c_1, c_4, c_5\}$	row 2	3/6
101	$\{c_1, c_4, c_5, c_6\}$	row 2	3/6
110	$\{c_1, c_2, c_3, c_4, c_5\}$	row 1	3/6
111	$\{c_1, c_2, c_3, c_4, c_5, c_6\}$	does not matter	3/6

To find the best overall strategy, loop over all partitions of the columns of the matrix. For each of these "Bob strategies," compute its value. Then choose the partition with the best value. This maximum value gives an upper bound for the two-player game.

One might think that after all this work, the 4×6 hint would give an improvement on our 3/8 upper bound. In that case, one would be wrong! All we get is another way to prove the 3/8 bound. However, applying an even more elaborate 8×14 hint,

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix},$$

does finally give an improvement of $81/224 = 0.361607\ldots$. It is worth mentioning that there are 3920 different partitions that provide this bound, which are of 8 different types once we account for equivalences under permutations of rows and columns.

There are a couple of things to keep in mind when constructing hints. First, the players are free to ignore the hint if they wish. Second, for a matrix hint each column in the matrix must have equal numbers of zeros and ones. Otherwise, the referee will not fulfill her obligation that the final sequence be random with $p = 1/2$. Third, the stronger the hint, the weaker the upper bound will be. So, to produce good bounds, we would like a matrix that is short (few rows) and wide (many columns). Also, we would like to give such hints to as few players as possible. The catch is that we loop over all partitions of the matrix columns, and this quickly becomes infeasible as the number of columns is increased.

The 8×14 hint was constructed by considering the fourteen nonconstant affine functions on three bits. The $81/224$ upper bound it produces remains the best provable upper bound to date for V_2. A reader interested in beating this record may be tempted to try the shorter and wider 6×20 matrix formed by the 20 three-element subsets of six elements. Surprisingly, this matrix does not do as well. Although an exhaustive search over column partitions was not performed, strategies achieving $117/320 = 0.365625$ have already been found, demonstrating that the upper bound will be worse. It is possible that a 16×30 matrix derived from nonconstant affine functions on four bits might yield a better upper bound, since hill climbs have found no success rate higher than 0.35625 for this larger matrix. However, an exhaustive search over all partitions of the 30 columns of this matrix currently seems infeasible.

The rest of this section is dedicated to formalizing these hints, using them to compute some bounds, applying them to other values of p, and providing a proof that these upper bounds converge to the optimal success probability.

3.2 The matrix game

We now turn our attention to computing upper bounds on $V_2(p)$. We do this by introducing a matrix game, which will be isomorphic to the hats game with the players given a little extra information. This extra information allows us to compute upper bounds. Furthermore, we shall show that by choosing a sufficiently large matrix, we can make these upper bounds arbitrarily tight.

Note

Throughout most of the paper, the variable n is used for the number of players. In this section, however, the number of players is always 2, and we use n to refer to the number of columns in a matrix.

Given a Bernoulli parameter $p \in (0,1)$ and a matrix of $\{0,1\}$-valued entries,

$$M = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{pmatrix},$$

we can play the following two-player cooperative game.

Player 2

The second player chooses an equivalence relation, $\sim$, on the set $\{1, \ldots, n\}$ (or, equivalently, a partition of the set of columns of M). Each possible choice for $\sim$ yields C equivalence classes (or disjoint subsets of the n columns) for some C with $1 \leq C \leq n$. (This equivalence relation corresponds to how Player 2 will choose a hat level on his own head based on which of n "hint columns" the referee ultimately selects for the stack of hats on Player 1's head.)

Referee

The referee independently flips a Bernoulli(p) coin C times, once for each of the C column subsets chosen by Player 2, assigning the value 1 with probability p and the value 0 with probability $1 - p$ each time. (For the hat game, these correspond to the hat colors on the C distinguished levels from which Player 2 will choose a hat on his own head.)

The referee then forms a vector $v = (v_1, \ldots, v_n)$, where each entry v_j is the binary value assigned above to the component of the partition that contains column j. We say that the vector v *respects* the equivalence relation $\sim$, because equivalent columns are assigned the same binary value.

Player 1

The first player observes the now-specified vector v assigned by the referee and chooses a row, r, of M and reports the dot product $r \cdot v$.

Players 1 and 2 wish to maximize the dot product. As Player 1 can easily compute the dot product for each possible row r, it is trivial for her to choose the maximal dot product for any given v.

Conditional value

The conditional value $V(M; p, \sim)$ of the matrix game on M—abbreviated to $V(M; \sim)$ when p is understood—for a particular equivalence relation is the expected value (over all possible realizations of the referee's Bernoulli(p) coin flips) of the reported (maxi-

mal) dot product $r \cdot v$ divided by n, the number of columns. That is,

$$V(M; p, \sim) = \frac{1}{n} \sum_{v} \max_{i} (r_i \cdot v) \mathbb{P}(v),$$

where the probability of each vector, $\mathbb{P}(v)$, depends on the value of p and on the particular choice of $\sim$. (The normalization by $1/n$ corresponds in the hats-with-hints game to the fact that the referee will create n possible stacks for Player 1's head and reveal this set of n possibilities to both players the night before the game; on the day of the game, the referee will uniformly at random choose one of these n possible stacks, which Player 2 will see but Player 1 will not.)

Value

The (unconditional) value of the matrix game $V(M; p)$—abbreviated to $V(M)$ when p is understood—is the maximum of the conditional value of the matrix game over all choices of equivalence relation:

$$V(M; p) = \max_{\sim} V(M; p, \sim).$$

The matrix game was devised primarily as a way of formalizing the hints that the referee can give the players to yield an upper bound on the value of the original hat game. However, by using matrices with slightly different characteristics, we can also encode the original hat game *without* hints and derive *lower* bounds on the value of the original hat game.

Theorem 4 asserts that these two bounds converge asymptotically, so that we could, in principle, approximate $V_2(p)$ arbitrarily closely (at least for rational p) by choosing suitable large matrices.

3.3 Lower bounds from the matrix game

For a fixed rational value of p, we can get a lower bound on the probability of winning the original hat game by playing the matrix game on a matrix with appropriately repeated columns, such as

$$L_{3,\frac{1}{2}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

or

$$L_{2,\frac{2}{3}} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

This is equivalent to playing the hat game where Player 2 looks only at the first few hats of Player 1, but Player 1 can look at all of Player 2's hats. The columns of $L_{3,\frac{1}{2}}$ represent the possible colorings of Player 1's first three hats with $p = \frac{1}{2}$, while $L_{2,\frac{2}{3}}$ represents the possible colorings of her first 2 hats with $p = \frac{2}{3}$ with the repeated columns representing the proportionate likelihood of the colorings. Player 2's hat strategy assigns one of his own hat positions to each distinct coloring. This gives an equivalence relation on the columns, where two columns are equivalent if they are assigned to the same hat position on Player 2's head. (Repeated instances of the same column may without loss of generality be assigned to the same equivalence class, corresponding to the fact that the players' strategies for the hat game may be taken to be deterministic.)

A vector v corresponds to a coloring on Player 2's head; v_i is the value of the hat in the position on Player 2's head chosen by the given strategy. When Player 1 chooses the best row, she is really choosing a hat on her own head that is most likely to give a pair of matching ones when Player 2 uses his strategy.

The conditional value $V(M;\sim)$ is the probability of both players guessing black hats for a given choice of equivalence relation, and the value $V(M)$ is the probability of this occurring for an optimal strategy.

This gives a lower bound for the complete hat game because it restricts the hats that Player 2 can look at, but does not give any advantage over the original game.

3.4 Upper bounds from the matrix game

To get an upper bound with $p = \frac{a}{b}$, we take any matrix of 0s and 1s in which the proportion of 1s is p in every column. The matrix can have duplicate columns. Then, when we play the game on this matrix, we get an upper bound. As an illustration, consider the following matrix:

$$U = \begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}.$$

We randomly generate an infinite sequence of rows from U. This produces, in this case, five infinite (vertical) sequences of bits. Since exactly two-thirds of the bits in each column are ones, each of the five infinite sequences has each bit independently equal to 1 with probability $p = \frac{2}{3}$. We then reveal to the two players that Player 1's hat sequence will be randomly chosen from this set of five. Although the five infinite sequences are not independent, this final selection will still be a randomly chosen sequence of hats. Player 2's hat strategy will assign one of his own hat positions to each of these 5 possible sequences for Player 1. As before, this creates an equivalence relation on the five columns of U, where two columns are equivalent if the sequences

they generated are assigned by Player 2 to the same hat position on his own head. Since we generated infinitely many independent selections from the rows of U, by the second Borel–Cantelli lemma each row is almost surely chosen infinitely often. Therefore, when Player 1 chooses one of her own hats, it is equivalent to choosing one of the rows of U. The reason that this is an upper bound for the hat game is because the players are given extra information. They are not restricted in any way and they do not have to use this extra information if they do not want to, so this cannot hurt them. But it may help.

3.5 The meeting of upper and lower bounds

For each rational probability $p = a/b$ in lowest terms and each positive integer m that is a multiple of b, we define two matrices, $L_{m,p}$ and $U_{m,p}$. The columns of these matrices will be elements of $\{0,1\}^m$. In $L_{m,p}$ all 2^m such columns appear, and each column with t 1s occurs $a^t(b-a)^{m-t}$ times, for a total of b^m columns. In $U_{m,p}$ only the $\binom{m}{mp}$ columns with mp 1s occur, and there is no repetition of columns. For both matrices, the columns may be ordered arbitrarily.

As argued in the previous sections, $V(L_{m,p})$ is a lower bound for the value of the two-person hat game with black-hat probability p, and $V(U_{m,p})$ is an upper bound. We now state the main theorem of this section, that the matrix-based upper and lower bounds converge to $V_2(p)$, the value of the 2-player hat game with (rational) black-hat probability $p \in (0,1)$.

Theorem 4 (Convergence theorem). *Let $L_{m,p}$ and $U_{m,p}$ be defined as above. Then*

$$\lim_{m\to\infty} V(U_{m,p}) - V(L_{m,p}) = 0.$$

We prove this result in Appendix C.

For certain values of p (in particular, those of the form $\frac{1}{k}$ or $\frac{k-1}{k}$), the general techniques above can be applied to particular matrices that yield rather good upper bounds without requiring too much work. In Appendix D, we prove the following general theorem and display hint matrices that yield even better bounds for the special cases $p = 1/3$ and $p = 2/3$.

Theorem 5. *For $p = \frac{a}{b} \le \frac{1}{2}$, we have $V_2(p) \le \frac{a}{b} - (1-\frac{a}{b})^b(\frac{a}{b})$.*
For $p = \frac{a}{b} \ge \frac{1}{2}$, we have $V_2(p) \le \frac{a}{b} - (1-\frac{a}{b})(\frac{a}{b})^b$.

The upper bounds from Theorem 5 for selected rational values of p are plotted as dots in Figure 8.1, with the lower bounds shown as a continuous curve obtained from Theorem 3.

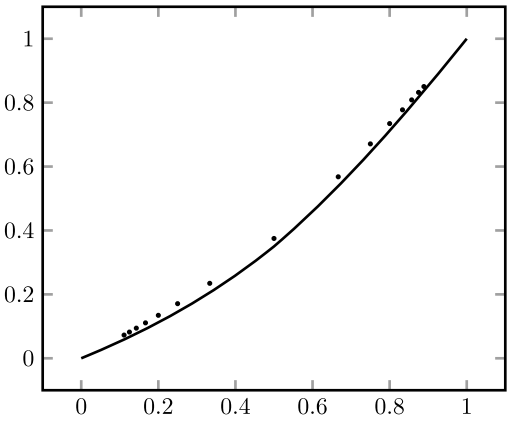


Figure 8.1: Bounds on $V_2(p)$.

4 Results for *n* players

4.1 Terminology and preliminary results

The n-player Levine hat game generalizes the 2-player game analyzed in the preceding sections. In the general version of the game, each of the n players can see all the hats on all of the other players' heads but not on his own. Once again, the players succeed if and only if each of them chooses a level on his own head that has a black hat (which we represent by "B" or "1," with a white hat being represented by "W" or "0"). In the main version of this game, and throughout the current section unless stated otherwise, we suppose that the referee (or sultan, or warden) chooses all hat colors on all players' (or wise men's, or prisoners') heads independently to be black with probability $p = 1/2$. Within this section, n will always refer to the number of players. Since it is sometimes important to distinguish clearly between one-person and multiperson subsets, we will think of the players as the wise men of Tanya Khovanova's problem statement [5] and use singular masculine pronouns when referring to individual players.

As usual, for any fixed value of n, we suppose that there are h hats per head (h stands both for "hats" and for "height") and let S be any joint strategy for the players. Within this section, we are usually interested in $V_n(p)$, the optimal limiting n-player success probability (or "value") as $h \to \infty$ for some given black-hat probability p, always with $0 < p < 1$. (As usual, we define $q := 1 - p$ throughout this section.) Often we are specifically interested in V_n, the value of the game for the default case $p = 1/2$. As noted near the beginning of Section 2, other cases of particular interest have p of the form $1/m$ (or $(m-1)/m$) for integer m, corresponding to a game in which each player must pick (or avoid) one of m colors.

In the course of describing strategies or proving results, we sometimes refer to restricted quantities such as

$$V_n^{(h)}(p; S),$$

the probability that the n players win with a particular strategy S when there are h hats on each player's head. It is not hard to see that $V_n^{(h)}(p)$ is nondecreasing in h (we formally state this in Lemma 1), so we have

$$V_n(p) := \lim_{h\to\infty} V_n^{(h)}(p) = \sup_h V_n^{(h)}(p).$$

Remark 4. Recall from Section 1 that the optimal success probability, or value, of any randomized strategy is always matched or exceeded by that of some deterministic strategy. We therefore assume without loss of generality that all players choose deterministic (pure) strategies, with each player's choice of level depending only on the (ordered) collection of $(n-1)h$ hat colors that he sees. Since, for each fixed n, p, and h, there are only finitely many possible strategies, the maximal value $V_n^{(h)}(p)$ is actually achieved for some such strategy.

Lemma 1. *We have $V_n^{(1)}(p) = p^n$ for all $n \geq 1$, and the values $V_n^{(h)}(p)$ are nondecreasing in h for each fixed n and p.*

Proof. The argument is straightforward. The players are free to ignore hats above any level, so increasing h cannot hurt. □

Lemma 2. *For each $h \geq 1$ and $p \in (0,1)$, the values $V_n^{(h)}(p)$ and $V_n(p)$ are nonincreasing in n.*

Proof. If an $(n+1)^{\text{st}}$ player is added, his hat colors are independent of those for the first n players, and thus cannot help the first n players to guess correctly with probability greater than $V_n^{(h)}(p)$ for finite h. Thus the full $(n+1)$-player set certainly cannot win with probability greater than $V_n^{(h)}(p)$. The result for $V_n(p)$ follows after taking suprema over $h \geq 1$. □

4.2 Some general bounds on $V_n(p)$

One of our main results for the multiplayer game will be that $V_n := V_n(1/2)$ is actually *strictly* decreasing in n. Before proving this, however, we derive some lower and upper bounds on $V_n(p)$, sometimes focusing on the case $p = 1/2$.

We already know from Lemma 1 that $V_n(p) \geq p^n$, but we will see now that we can do much better than random guessing of levels.

Theorem 6. *We have $V_n \geq 1/(n+1)$ and $V_n(p) \geq \frac{p/q}{n+p/q}$ for all $n \geq 1$.*

Proof. Each player finds the first level at which the other $n-1$ players all have black hats, and he chooses that same level for his own head. (With probability 1 as $h \to \infty$,

each player can find such a level.) The players succeed if and only if the first level that contains at least $n-1$ black hats actually contains n black hats. There are $n+1$ possible arrangements of hats on this level, exactly one of them leads to success, and when $p = 1/2$ they are all equally likely, yielding the result for $p = 1/2$. For general p, the bound in the theorem follows from the posterior probability that the first candidate level actually has n black hats. □

Thus we see that we can do dramatically better than with random guessing. In order to approach $1/(n+1)$ with the strategy above, however, we need h to grow exponentially in n. The next strategy, discovered by Peter Winkler [5] for the usual case $p = 1/2$ and generalized below for arbitrary p, requires h to grow only logarithmically in n and, quite surprisingly, yields another dramatic improvement in the success probability for large n.

Definition 1. For the next two theorems, given any $p \in (0,1)$, we let $q := 1-p$ as usual and define r to be the reciprocal $r := 1/q$.

Theorem 7. *For every p such that $0 < p < 1$, we have*

$$V_n(p) \geq \frac{1 - 1/\ln(n)}{\lceil \log_r n + \log_r \ln n \rceil} \quad \textit{for all } n \geq 3.$$

Thus for each $\epsilon > 0$,

$$V_n(p) \geq (1-\epsilon)/\log_r(n) \quad \textit{for all sufficiently large } n.$$

Proof. Each player will attempt to choose the *first* level on which he has a black hat. (There might also be some serendipitous success probability coming from cases in which some players choose black hats but not the first black hats on their respective heads. We get a valid lower bound on V_n by considering only the joint probability that every player chooses the *first* black hat on his own head.) We will further have each player choose only from the first t levels, where $t = \lceil \log_r n + \log_r \ln n \rceil$. The players will hope that they each have at least one black hat within the first t levels and that the sum of their n first-black-hat levels is congruent (mod t) to some specified residue s. Given the values of n and t, they will choose the residue s during their strategy session to maximize their probability of hitting that residue. Conditioned on the assumption that all n players have at least one black hat apiece within the first t levels, the best residue will certainly occur with probability at least $1/t$.

The probability that any given player is bereft of black hats on the first t levels is

$$q^t \leq q^{\log_r n + \log_r \ln n} = 1/(n \ln n),$$

and now, by a union bound,

$$\mathbb{P}(\text{At least one player has no usable black hats}) \leq n/(n \ln n) = 1/\ln(n).$$

Thus $\mathbb{P}(\text{Every player has a usable black hat}) \geq 1 - 1/\ln(n)$, which is positive if $n \geq 3$.

Conditioned on every player having a usable black hat, the best residue $s \pmod t$ for the sum of the n lowest-black-hat levels occurs with probability at least $1/t$, and each player guesses the appropriate level on his own head to make the sum of all n lowest-black-hat levels congruent to the target residue. The theorem follows. □

In the next section, we discuss several refinements to the basic Winkler strategy that improve the lower bounds by 5 to 10 % for moderate values of n and outperform the $1/(n+1)$ lower bound for all $n \geq 3$. However, the following result (attributed in essence to Ori Gurel–Gurevich for the case $p = 1/2$ in a comment by hatmeister Lionel Levine on Tanya Khovanona's blog [5], with no details given for the proof) shows that as long as each player is required to choose the *lowest* level at which he has a black hat, the lower bound from Theorem 7 is asymptotically tight. Thus it seems likely that any substantial improvements upon the $1/\log_r(n) = \ln(1/q)/\ln(n)$ approximate lower bound will require new ideas.

Theorem 8. *Suppose that each player must choose the* lowest *level on his own head that has a black hat (with the team failing if any player has only white hats). Then, using $\tilde{V}_n(p)$ to refer to the optimal probability as $h \to \infty$ that the players succeed (for any given black-hat probability $p \in (0,1)$) under this more stringent requirement, for each ϵ with $0 < \epsilon \leq 1/4$ we have*

$$\tilde{V}_n(p) \leq (1+\epsilon)/\log_r(n) \text{ for all sufficiently large } n.$$

Proof (overview). We restrict attention to the first $t \approx \log_r(n)$ levels. An accomplice will uniformly at random choose a level k from $\{1, \ldots, t\}$ and then uniformly at random choose a player j from among those who happen to have lowest-black-hat level Y_j equal to k. The accomplice will then inform the chosen player of his special status but will not tell him his lowest-black-hat level k. This special player is the only one required to guess his level Y_j; the accomplice will tell the other $(n-1)$ players (including those with $Y_j > t$) their own values Y_j.

As shown in the complete proof in Appendix E.1, even with this substantial help (which the players can ignore if they wish), the one player who is not told his own level still cannot pick the correct level of his first black hat with probability much greater than $1/t$. □

For $n \geq 5$, the best lower bounds on V_n that we know come from small refinements to Winkler's order-$(1/\log(n))$ strategy. As n grows, our best lower bounds on V_n are asymptotically equal to $1/\log_2(n)$. For upper bounds, we can use the fact that V_n is nonincreasing in n (from Lemma 2) to see that

$$V_n \leq V_2 \leq 81/224 \approx 0.361607 \quad \text{for all } n \geq 2,$$

but we would like our upper bounds actually to (strictly) decrease as n increases. A first step in this direction is the following theorem, which establishes a gap between V_2 and V_3.

Theorem 9. *We have*

$$V_3 \le 89/256 = 0.34765625 < 0.35 \le V_2.$$

Proof. We generalize the "hint" technique that was used earlier for 2 players to show that $V_2 \le 81/224$. Given Players A, B, and C, we suppose that the referee gives A a 2×2 hint with associated matrix that contains the two balanced binary columns of Hamming weight 1. In addition, the referee gives B a 4×6 hint whose associated binary matrix contains all $\binom{4}{2} = 6$ distinct weight-2 columns of 4 bits each. The referee gives C no hint.

It follows, much as for the 2-player upper bounds, that C sees one of $6 \cdot 2 = 12$ column pairs for (A, B) jointly and can then restrict attention to at most the first 12 levels on his own head, with each of the 12 possible (A, B) column pairs associated with one of these levels on his own head. The (at most) 12 levels on C's head can be permuted arbitrarily without loss of generality, leaving us with several million inequivalent strategies for C, which we put into the outer loop of a computer search. For each choice of C's strategy, together with the realization of C's hats within his (at most) 12 distinguished levels, A and B are left with a 2-player game with hints, and they choose their best (conditional) strategy so as to win with conditional probability $k/12$ for some integer k.

Given C's strategy and one of (at most) 2^{12} hat vectors on his distinguished levels, B conditionally has only $4^2 = 16$ distinct strategies to consider, depending on which of the 4 rows he chooses on his own head for each of the 2 possible columns the referee could have placed on A's head. For each joint choice of a B, C strategy, together with what A sees on B's and C's heads, A guesses whichever of the 2 rows on his own head yields a higher (conditional) probability that A, B, and C will each point to a black hat (a '1'), breaking ties arbitrarily.

It turns out that, up to isomorphism, C's unique best strategy uses only 6 levels on his own head. This best strategy can be represented by the following 2×6 matrix, indexed by the "hint column"of A and the hint column of B:

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 3 & 3 \\ 2 & 4 & 3 & 5 & 1 & 6 \end{pmatrix}.$$

When this optimal C-strategy is combined with optimal strategies for B and A, the team wins with probability $267/(64\cdot 12) = 89/256 = 0.34765625$. Since we already know that $V_2 \ge 7/20 = 0.35$, we see that $V_2 - V_3 > 0.0023$, a nonzero gap. □

Shortly, we will prove that V_n is strictly decreasing in n. The proof will require a nontrivial upper bound on the probability that n players can *avoid* all choosing black

hats on their own heads, which one might call the "misère" version of the game. (Stan Rabinowitz and Stan Wagon originally suggested this version of the problem.) More generally, one can try to bound

$$\mathbb{P}(\text{At least } k \text{ of the } n \text{ players choose black hats})$$

or

$$\mathbb{P}(\text{At most } k \text{ of the } n \text{ players choose black hats})$$

for any value k from 0 through n.

Still more generally, for any subset $\mathcal{S}$ of $\{0, 1, 2, \ldots, n\}$, one could ask for

$$\mathbb{P}(k \in \mathcal{S}), \text{ where } k \text{ is the number of players who choose black hats.}$$

(For example, $\mathcal{S}$ might be the set of all even integers in $\{0, 1, 2, \ldots, n\}$.)

When $n = 2$, all of the problems above are equivalent to each other, as we now argue.

For any value of p (not necessarily 1/2) and any joint strategy S for the 2 players, we write $\mathbb{P}(WW)$ for the probability that both players point to white hats, $\mathbb{P}(WB)$ for the probability that Players 1 and 2 point to white and black hats, respectively, on their own heads, etc. We use '$*$' as a wildcard; e. g., $\mathbb{P}(*B)$ is the probability that Player 1 points to a hat of either color and Player 2 points to a black hat.

Once we know $V_2(p; S)$ for any given strategy S, we know the entry $\mathbb{P}(BB)$ in a 2×2 matrix of probabilities with $\mathbb{P}(WW), \mathbb{P}(WB), \mathbb{P}(BW)$, and $\mathbb{P}(BB)$, and both row sums and both column sums are fixed (equal to p for $\mathbb{P}(B*)$ and $\mathbb{P}(*B)$, equal to $1 - p$ for $\mathbb{P}(W*)$ and $\mathbb{P}(*W)$). Thus the probability of every possible outcome or set of outcomes for any given strategy S with 2 players can be calculated given $\mathbb{P}(BB)$. For example, our upper and lower bounds on $V_2(p)$ translate directly to upper and lower bounds on the probability that both players choose the same color hat, on the probability that the players choose hats of different colors, and on the probability that at least one player chooses a white hat (the latter being the misère version of the original game).

When $n > 2$, however, these different generalizations of the original game appear to be essentially different from one another. We will consider only the misère game (in which at least one player is supposed to choose a white hat), and only for the case $p = 1/2$. Letting W_n be the maximal probability of winning the n-player misère game (where W stands for "white" and is the letter following the V used for the original version of the game), we will see that $W_n \to 1$ as $n \to \infty$ but that W_n is bounded away from 1 for each fixed n.

We begin by defining an infinite sequence of pairs of integers (r_j, s_j) for $j \geq 1$ as below; these will be the dimensions of "hint matrices" given to the various players for the n-player misère game.

Definition 2. Let $(r_1, s_1) = (2, 2)$, $(r_2, s_2) = (4, 6)$, and for each $k \geq 3$, define

$$r_k = 2\prod_{j=1}^{k-1} s_j \quad \text{and} \quad s_k = \binom{r_k}{r_k/2}.$$

Theorem 10. *Letting* $W_n = \sup_S \mathbb{P}(\textit{At least 1 of n players chooses a white hat})$ *over all possible n-player strategies as h, the number of hats per head, grows to infinity, and we have the following:*

$$W_2 = 1/2 + V_2 \in [17/20, 193/224] \quad (\textit{hence in [0.85, 0.8616...]})$$

and, for $n \geq 3$,

$$1 - (1/2)^n \leq W_n \leq 1 - \frac{1}{(r_n/2)2^{(r_n/2)}}.$$

Proof. It is straightforward to show that

$$W_2 = \sup_S\big(\mathbb{P}(WW) + \mathbb{P}(WB) + \mathbb{P}(BW)\big) = \sup_S\big(\mathbb{P}(W*) + \mathbb{P}(BW)\big).$$

For $p = 1/2$, though, $\mathbb{P}(W*) = 1/2$ for every strategy S, and since $p = 1 - p$,

$$\sup_S \mathbb{P}(BW) = \sup_S \mathbb{P}(BB) = V_2,$$

and the bounds for W_2 follow immediately.

The lower bounds on W_n for all $n \geq 3$ follow by letting each player just choose the first level on his own head. The upper bounds follow from a cascading "hint" technique generalizing the upper-bounding techniques used earlier for V_2 and V_3. The remaining details of the proof are given in Appendix E.2. □

Now that we have bounded W_n away from 1 for every n (for the misère game), we are finally ready to show that V_n (for the original game) is strictly decreasing in n.

Theorem 11. *We have* $V_{n+1} < V_n$ *for all* $n \geq 1$.

Proof. Given any h and strategy S for a game with $n+1$ players, let us write $\mathbb{P}(B\ldots B, B)$ to refer to the probability that Players 1 through $n+1$ all choose black hats. Let us write $\mathbb{P}(B\ldots B, W)$ for the probability that the first n players choose black hats while Player $n+1$ chooses a white hat, and write

$$\mathbb{P}(B\ldots B, *) := \mathbb{P}(B\ldots B, B) + \mathbb{P}(B\ldots B, W)$$

for the probability that the first n players choose black hats while Player $n+1$ chooses a hat of arbitrary color.

Now, for any given $(n+1)$-player strategy, we have

$$\mathbb{P}(B\dots B,B) = \mathbb{P}(B\dots B,*) - \mathbb{P}(B\dots B,W).$$

Thus, taking suprema over strategies and over h, we have

$$\begin{aligned} V_{n+1} &= \sup \mathbb{P}(B\dots B,B) \\ &= \sup\{\mathbb{P}(B\dots B,*) - \mathbb{P}(B\dots B,W)\} \\ &\le \sup \mathbb{P}(B\dots B,*) + \sup\{1-\mathbb{P}(B\dots B,W)\} - 1 \\ &= V_n + \sup\{1-\mathbb{P}(B\dots B,B)\} - 1, \end{aligned}$$

where the last equality is because $p = 1/2$, so the maximal probability of avoiding any given ordered sequence of chosen hat colors for the $n+1$ players is the same as the maximal probability for any other ordered sequence of chosen hat colors.

Now we have

$$V_{n+1} \le V_n + W_{n+1} - 1 = V_n - (1 - W_{n+1}),$$

so $(V_n - V_{n+1}) \ge (1 - W_{n+1})$, which by Theorem 10 is positive and bounded away from 0 for each fixed n. □

Remark 5. A similar argument shows that W_n for the misère game is strictly increasing in n. See Theorem 13 in Appendix E for the proof.

One of our main unresolved questions concerns Levine's original conjecture, which we restate here.

Conjecture 1 (Levine). The optimal success probability in the n-player game is $o(1)$ as n goes to infinity; i. e.,

$$\lim_{n\to\infty} V_n = 0.$$

Remark 6. Although the conjecture above seems likely to be true, the rate of decrease provided in Theorem 11 is insufficient to prove it, even with tighter bounds on $(1-W_n)$. The upper bound that Theorem 11 yields on $\lim_{n\to\infty} V_n$ is

$$V_2 - \sum_{k=3}^{\infty}(1-W_k),$$

which is at least

$$\frac{7}{20} - \sum_{k=3}^{\infty}\frac{1}{2^k} = \frac{7}{20} - \frac{1}{4} = \frac{1}{10}.$$

5 Best current lower bounds

For puzzle aficionados (or prisoners' advocates) who would like to improve upon existing strategies for various numbers of players (or prisoners), we collect here the best lower bounds we know on V_2 through V_{12} (all for $p = 1/2$). The result for V_2 comes from Section 2 and is conjectured to be optimal. The results for V_3 and V_4 essentially come from hill climbs over symmetric strategies with 5 or 4 hats, respectively, per player.

The bounds on V_5 through V_{12} come from generalized Winkler-style strategies (as described in Theorem 7) in which the players focus on the first $t \approx \log_2(n)$ levels and all try to select the *lowest* levels on which they have black hats. However, the n players use general $t \times t$ Latin-square operations, not necessarily mod-t addition, in order to construct a "sum" in $\{0, 1, \ldots, t-1\}$ of their n respective lowest-black-hat levels. Furthermore, in assessing each candidate strategy, we take into account all "bonus" success probability that arises when the players miss their target "sum" (mod t) but nonetheless all serendipitously point to black hats.

Finally, all t-level strategies—whether found by hill climbing or derived from Latin-square operations—are then augmented by working with t-level "tiers" and having each player recursively "reset" (shifting up t levels at a time) whenever he sees at least one other player with an all-white stack of hats in the current tier. A second, smaller, improvement comes from also recursively resetting whenever a player sees *only* black hats on *all* other players' heads within the current tier. Using both "white" and "black" recursive resets leads to rational lower bounds with denominators of the form

$$(2^t - 1)^n + n(2^t - 1)^{n-1} - (n + 1),$$

and we retain these unreduced fractions in the table below.

Additional details about the best strategies known are given in Appendix E.4. The resulting values of V_n for $2 \le n \le 12$ are shown below, with all decimal values rounded down to 6 decimal places to provide true lower bounds:

$$\begin{aligned}
V_2 &\ge 21/60 = 0.350000,\\
V_3 &\ge 9119/32670 = 0.279124\ldots,\\
V_4 &\ge 14844/64120 = 0.231503\ldots,\\
V_5 &\ge 205447/1012494 = 0.202911\ldots,\\
V_6 &\ge 2984604/15946868 = 0.187159\ldots,\\
V_7 &\ge 43930663/250593742 = 0.175306\ldots,\\
V_8 &\ge 651583632/3929765616 = 0.165807\ldots,\\
V_9 &\ge 0.158764\ldots,\\
V_{10} &\ge 0.153517\ldots,\\
V_{11} &\ge 0.149025\ldots,\\
V_{12} &\ge 0.145047\ldots.
\end{aligned}$$

6 Future directions

This paper leaves open certain questions that might be of interest to other hatters, mad or otherwise. In particular, there are two conjectures that we stated earlier, phrased below as questions:

1. Is V_2 exactly equal to 0.35?
2. Does V_n approach 0 as $n \to \infty$?

We also note that for the upper-bound results in Section 3, one promising family of matrices has size $2^k \times (2^{k+1} - 2)$; the cases for $k = 1, 2$, and 3 are presented in that section. For any such matrix, one must consider $\text{Bell}(2^{k+1} - 2)$ partitions of the columns, a task that seems infeasible beyond $k = 3$. One might look for ways to reduce the computational difficulty of this approach. For example, we can safely assume that none of the column subsets within an optimal partition contains both a column and its bitwise complement. Unfortunately, the resulting reduction in work is not very significant.

There are also several generalizations one could consider, some of which are presented below:

1. In the n-player case, one might require at least k players to pick a black hat (the $k = n$ and $k = 1$ cases are discussed above).
2. One could allow more than 2 hat colors, perhaps with a different payoff system. For example, one might consider a grayscale version of the game, where each hat has a value in $[0, 1]$, with 0 being white and 1 being black, and with the value of a joint guess taken to be the product of the values of the selected hats.

Appendix A. The 1-player game

Now we consider the curious case of the 1-player (or "solo") version of the Levine hat game, in which Sol must try to point to a black hat on his own head. Imagine, if you will, that Sol cannot see any of the infinitely many black and white hats on his own head, but that he is endowed with infinite computational abilities and armed with a secure faith in the Axiom of Choice (AC). In order to describe Sol's strategy, we first look at an auxiliary game with countably many players.

We begin with some definitions. Throughout this section, I denotes a fixed countable set, which we call the *set of players*. A *hat-stack* is an infinite sequence of zeros and ones, or if preferred, an infinite sequence from $\{white, black\}$. Each member of a hat-stack will be called a *hat*. A *hat-assignment* is a mapping that assigns a hat-stack to each element of I. Note that the image of a hat-assignment is a set of hat-stacks. Given a set of hat-stacks, M, a *black level* of M is a natural number, i, such that $m(i) = 1$ for every $m \in M$. In addition, M will be called *generic* if every finite subset of M has a black level, and M will be called *almost-generic* if some cofinite subset of M is generic.

In a hat-assignment there are countably many hats (each element of I is assigned a hat-stack, and each hat-stack has countably many hats). A hat-assignment gives each of these hats a color. A cylinder is an assignment of colors to some finite subset of these hats. The standard Bernoulli measure ($p = 1/2$) is a probability measure defined on the σ-algebra of sets generated by the cylinders. With this probability space, the following proposition is standard and is stated without proof.

Proposition 1. *In the standard Bernoulli measure ($p = 1/2$), almost every hat-assignment is one-to-one (i. e., with no two elements of I receiving the same hat-stack) and has a generic image.*

Theorem 12. *Assume the Axiom of Choice. Then there exists a set C whose elements are countable sets of hat-stacks, and C has the following property: For every countable set of hat-stacks, T, there is a unique element $R \in C$ such that*

1. *$T \backslash R$ is finite, and*
2. *$R \backslash T$ is finite.*

Furthermore, if T is generic, then $R \backslash T$ has a black level.

Proof. Let two countable sets of hat-stacks be equivalent if they differ by a finite set of hat-stacks. Using *AC*, let E be a choice set for the set of equivalence classes. Using *AC* a second time, replace each $D \in E$ that is almost-generic, with a cofinite generic subset of D. Since this does not change the equivalence class of D, the resulting set C is also a choice set. In addition, any almost-generic element of C is generic. Therefore, given a countable set of hat-stacks, T, there is a unique element $R \in C$ that is equivalent to T. This gives the first two properties. Furthermore, if T is generic, then any subset of T is also generic. So, removing the finite set $R \backslash T$ from R results in the generic set $R \cap T$. Therefore, R is almost-generic. But $R \in C$, so R is generic. Since $R \backslash T$ is a finite subset of R, it has a black level. □

We now interpret the previous results as a hat game.

The *Auxiliary Game* has a countably infinite set of players, each with an infinite sequence of hats. The usual rules apply, including "No looking at your own hats." The players win if all but finitely many of them are able to point to a black hat. This game is similar to a puzzle described by Greg Muller [6], who attributed the earlier puzzle to Mike O'Connor. In that game, there is a fixed ordering of the players. Here, we want the players to be indistinguishable.

Our previous results provide a strategy for winning the Auxiliary Game with probability 1. Let I be the set of players. Consider a random hat-assignment, chosen according to the probability space defined above. Using the proposition and ignoring a measure-zero event, we find that the assignment is one-to-one and its image, T, is generic. Let C be as in the previous theorem, and let $R \in C$ be the unique set satisfying the conclusions of the theorem.

To describe the strategy, fix a player $i \in I$. Player i cannot determine T, since no players see their own hats. Nevertheless, Player i can determine T_i, the set of hat-stacks assigned to the other players. The theorem applies to T as well as to T_i, and by uniqueness, both yield the same element R. So, Player i reports a black level of the finite set $R \backslash T_i$.

Suppose that each player follows the strategy above. If Player i has a hat stack in R, then since the assignment is one-to-one, this hat stack will also be in $R \backslash T_i$. By choosing a black level of $R \backslash T_i$, any such player is guaranteed to choose a black hat. Since $T \backslash R$ is finite, the game is won.

The Auxiliary Game also shows how Sol can win the solo game. On the fateful day, he brings with him infinitely many friends, each with his own referee and coin. This is not explicitly against the rules. All referees simultaneously select random hat sequences for their respective players. Sol and his friends play the Auxiliary Game. With probability 1, all but finitely many players choose a black hat. Sol is confident that he will not be one of the unlucky ones. Now, he might have a persnickety logician friend who warns against depending on an event whose probability cannot be precisely measured. Sol should ignore this advice! Since only finitely many players fail, and all players are essentially identical, he is virtually guaranteed to win the game!

Appendix B. Computing the performance of 2-player strategy S_1

Below we compute the success probability, or value, $V_2(p; S_1)$ for the strategy S_1 from Section 2. We break the calculation of $V_2(p; S_1)$ into 7 cases.

Case 1: Both players are monochromatic to the same odd position. By "monochromatic up to an odd position $2\ell + 1$," we mean that a player has either all W or all B up to position $2\ell + 1$ but not up to position $2\ell + 3$.

Case 1(a): Both players start with B and have B hats up to position $2\ell + 1$ but not to position $2\ell + 3$, for some integer ℓ. The probability of winning conditioned on this is given in Table 8.6, which shows hats in positions $2\ell + 1$, $2\ell + 2$, and $2\ell + 3$ for both players.

Table 8.6: Case 1(a).

	■■□	■□■	■□□
■■□	p^2q^2	p^2q^2	pq^3
■□■	p^2q^2		
■□□	pq^3		q^4

The probability of this case occurring and the players' winning is

$$p^2(3p^2q^2 + 2pq^3 + q^4) + p^6(3p^2q^2 + 2pq^3 + q^4) + p^{10}(3p^2q^2 + 2pq^3 + q^4) + \cdots,$$

which we can simplify by summing the geometric series to obtain

$$\frac{(pq)^2(1 + 2p^2)}{1 - p^4}.$$

Case 1(b): Both players start with W and have W hats up to position $2\ell + 1$ but not to position $2\ell + 3$, for some integer ℓ. The probability of winning conditioned on this is given in Table 8.7, which shows hats in positions $2\ell + 1$, $2\ell + 2$, and $2\ell + 3$.

Table 8.7: Case 1(b).

	□■■	□■□	□□■
□■■	p^4		p^3q
□■□			p^2q^2
□□■	p^3q	p^2q^2	

The probability of this case occurring and the players' winning is

$$q^2(p^4 + 2p^3q + 2p^2q^2) + q^6(p^4 + 2p^3q + 2p^2q^2) + q^{10}(p^4 + 2p^3q + 2p^2q^2) + \cdots,$$

which simplifies to

$$\frac{(pq)^2}{1 - q^2}.$$

Case 1(c): One player starts W and the other starts B, or vice versa, and both are monochromatic to the same odd position (see Table 8.8).

Table 8.8: Case 1(c).

	■■□	■□■	■□□
□■■		p^3q	
□■□		p^2q^2	
□□■			

The probability of this case occurring and the players' winning is

$$pq(p^3q + p^2q^2) + p^3q^3(p^3q + p^2q^2) + p^5q^5(p^3q + p^2q^2) + \cdots,$$

which simplifies to

$$\frac{(pq)^2(p^2+pq)}{1-(pq)^2} = \frac{p^3q^2}{1-(pq)^2}.$$

Finally, notice that the roles of the two players could be interchanged here, so we double the above probability to get

$$\frac{2p^3q^2}{1-(pq)^2}.$$

Case 2: Players are monochromatic to different odd positions.

Case 2(a): The taller monochromatic stack is W. In this case the player with the taller W stack will always choose a W hat, so the probability of winning is 0.

Case 2(b): The taller monochromatic stack is B.

Case 2(b)(i): The taller monochromatic B stack is taller than the shorter one by at least 2 odd positions. In this case, the player with the taller B stack always guesses correctly. The player with the shorter stack guesses correctly with probability p since his guess is uncorrelated with the other player's guess. To calculate the probability of winning in this case, note that the probability of a player being monochromatic up to odd position $2\ell+1$ and not to $2\ell+3$ is $p^{2\ell+1}(1-p^2)+q^{2\ell+1}(1-q^2)$. The probability of the other player being monochromatic B up to at least position $2\ell+5$ is $p^{2\ell+5}$. Thus, the probability of this case occurring and the players' winning is

$$\sum_{\ell=0}^{\infty}(p^{2\ell+1}(1-p^2)+q^{2\ell+1}(1-q^2))p^{2\ell+5},$$

which simplifies to

$$\frac{p^6(1-p^2)}{1-p^4}+\frac{p^5q(1-q^2)}{1-(pq)^2}.$$

Thus, the probability of winning in this case is obtained by multiplying by p and by 2, giving

$$\frac{2p^7(1-p^2)}{1-p^4}+\frac{2p^6q(1-q^2)}{1-(pq)^2}.$$

Case 2(b)(ii): The taller stack is monochromatic B to odd position $2\ell+3$ and the shorter stack is monochromatic B to position $2\ell+1$, for some integer ℓ. The player with the taller B stack always guesses correctly. The player with the shorter B stack guesses correctly according to Table 8.9, which shows hats in position $2\ell+1$ to $2\ell+5$ for the player with the taller B stack, and $2\ell+1$ to $2\ell+3$ for the other player. Note that in the middle row, the player with the shorter stack guesses the hat in position $2\ell+4$

Table 8.9: Case 2(b)(ii).

	■■□	■□■	■□□
■■■■□		p^2q^2	
■■■□■	p^3q^2	p^3q^2	p^2q^3
■■■□□		pq^3	

and so has probability p of being correct, which is multiplied by the probability of the situation occurring.

If this case occurs, then the probability of winning is given by

$$p^2q^2 + 2p^3q^2 + p^2q^3 + pq^3 = pq^2(1 + p + p^2).$$

The probability of this case occurring and the players' winning is, therefore, given by

$$p^4pq^2(1 + p + p^2) + p^8pq^2(1 + p + p^2) + p^{12}pq^2(1 + p + p^2) + \cdots,$$

which simplifies to

$$\frac{p^5q(1 - p^3)}{1 - p^4}.$$

Taking into account the fact that either player could have the taller stack, we get a probability of winning as

$$\frac{2p^5q(1 - p^3)}{1 - p^4}.$$

Case 2(b)(iii): The taller B stack is monochromatic B to odd position $2\ell + 3$ and the shorter stack is monochromatic W to position $2\ell + 1$, for some integer ℓ. The player with the taller B stack always guesses correctly. The player with the shorter W stack guesses correctly according to Table 8.10.

Table 8.10: Case 2(b)(iii).

	□■■	□■□	□□■
■■■■□	p^3q		p^2q^2
■■■□■	p^4q	p^3q^2	p^3q^2
■■■□□	p^2q^2		pq^3

If this case occurs, then the probability of winning is given by

$$p^4q + 2p^3q^2 + p^3q + 2p^2q^2 + pq^3 = pq(1 + p - pq^2).$$

The probability of this case occurring and the players' winning is, therefore, given by

$$p^3qpq(1+p-pq^2)+p^5q^3pq(1+p-pq^2)+p^7q^5pq(1+p-pq^2)+\cdots,$$

which simplifies to

$$\frac{p^4q^2(1+p-pq^2)}{1-(pq)^2}.$$

Taking into account the fact that either player could have the taller stack, we get a probability of winning as

$$\frac{2p^4q^2(1+p-pq^2)}{1-(pq)^2}.$$

Summing the success probabilities from all the cases above and replacing q with $1-p$ yields

$$V_2(p;S_1)=\frac{p(1+p+p^2+3p^3-3p^4+p^5)}{2+p+p^2+p^3-p^4}.$$

Appendix C. Proof that the matrix-based upper and lower bounds converge to $V_2(p)$

For each rational probability $p=a/b$ in lowest terms and each positive integer m that is a multiple of b, we define two matrices, $L_{m,p}$ and $U_{m,p}$. The columns of these matrices will be elements of $\{0,1\}^m$. In $L_{m,p}$ all 2^m such columns appear, and each column with t 1s occurs $a^t(b-a)^{m-t}$ times, for a total of b^m columns. In $U_{m,p}$ only the $\binom{m}{mp}$ columns with mp 1s occur, and there is no repetition of columns. For both matrices, the columns may be ordered arbitrarily.

As argued in Section 3, $V(L_{m,p})$ is a lower bound for the value of the two-person hat game with black-hat probability p, and $V(U_{m,p})$ is an upper bound. Below we prove Theorem 4, the main theorem of that section, that as $m\to\infty$, the matrix-based upper and lower bounds converge to $V_2(p)$, the value of the 2-player hat game with (rational) black-hat probability $p\in(0,1)$.

The following three lemmas establish background results needed to prove Theorem 4.

Lemma 3 (Determinism lemma). *If there are duplicate columns in M, then Player 2 may as well put them into the same equivalence class when playing the matrix game on M.*

Proof. This is just a simple convexity argument. More precisely, let $\sim$ be an equivalence relation chosen by Player 2. Let c and d be two duplicate columns in two different

equivalence classes, C and D, respectively. Consider the two equivalence relations, $\sim_C$ and $\sim_D$, which are identical to $\sim$ except that in $\sim_C$ we move column d to C and in $\sim_D$ we move column c to D. Let v be any vector that respects $\sim$. We form two new vectors v^C and v^D that are identical to v except in positions c and d, where we have $v^C_c = v^C_d = v_c$ and $v^D_c = v^D_d = v_d$. Note that v^C respects $\sim_C$ and v^D respects $\sim_D$. Note also that these mappings might not be one-to-one; if d is the only member of D, then it is possible that $v^C = w^C$ even though $v \neq w$. In this case, we will count v^C and w^C as two different vectors, so we can write $\mathbb{P}(v) = \mathbb{P}(v^C) = \mathbb{P}(v^D)$. Let r be any row of M. Then $r \cdot c = r \cdot d$ because c and d are duplicate columns.

So $r \cdot v = (r \cdot v^C + r \cdot v^D)/2$. If we let $r(v)$ denote the row that Player 1 assigns to v, then we have

$$2\sum_v (r(v) \cdot v)\mathbb{P}(v) = \sum_v (r(v) \cdot v^C)\mathbb{P}(v^C) + \sum_v (r(v) \cdot v^D)\mathbb{P}(v^D).$$

Therefore, at least one of sums on the right is at least as big as the (undoubled) sum on the left. Assume, without loss of generality, that

$$\sum_v (r(v) \cdot v)\mathbb{P}(v) \leq \sum_v (r(v) \cdot v^C)\mathbb{P}(v^C).$$

Then $V(M;\sim) \leq V(M;\sim^C)$. □

Lemma 4 (Replication lemma). *If all columns are replicated the same number of times, this does not change the value of the matrix.*

Proof. Suppose M_2 is an $m \times kn$ matrix formed from the $m \times n$ matrix M_1 by including each column of M_1 a total of k times. Let $\sim_2$ be an equivalence relation chosen by Player 2 on M_2. By the determinism lemma, we can assume that $\sim_2$ assigns each set of identical columns to the same equivalence class. Let $\sim_1$ be the restriction of $\sim_1$ to the original matrix M_1. Let v_1 be any vector of numbers that respects $\sim_1$, and let v_2 be the k-fold expansion of v_2. Let r be a row of M_2. Then $r \cdot v_2 = kr \cdot v_1$ Therefore, $V(M_2;\sim_2) = V(M_1;\sim_1)$. □

Lemma 5 (Erasure lemma). *If we remove a small proportion ϵ of the columns of M, the value $V(M)$ changes by at most $\frac{2\epsilon}{1-\epsilon}$.*

Proof. Let r be any row of M with n columns, and v any vector of size n whose entries are zeros and ones. Then when the columns are removed, $r \cdot v$, which is at most n, becomes $r' \cdot v'$ and decreases by some amount k, where $0 \leq k \leq \epsilon n$. So

$$\left|\frac{r \cdot v}{n} - \frac{r' \cdot v'}{n(1-\epsilon)}\right| = \left|\frac{r \cdot v}{n} - \frac{r \cdot v - k}{n(1-\epsilon)}\right|$$
$$= \left|\frac{-\epsilon r \cdot v + k}{n(1-\epsilon)}\right|$$

$$\leq \left|\frac{-\epsilon}{1-\epsilon}\right| + \left|\frac{\epsilon}{1-\epsilon}\right|$$
$$= \frac{2\epsilon}{1-\epsilon}.$$

□

Lemma 6 (Perturbation lemma). *Let $\epsilon > 0$ and let M and N be two matrices of the same size with elements in $\{0,1\}$. Let*

$$E := N - M = \begin{pmatrix} \epsilon_{1,1} & \epsilon_{1,2} & \cdots & \epsilon_{1,n} \\ \epsilon_{2,1} & \epsilon_{2,2} & \cdots & \epsilon_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{m,1} & \epsilon_{m,2} & \cdots & \epsilon_{m,n} \end{pmatrix},$$

with each entry $\epsilon_{i,j} \in \{-1,0,1\}$. If the average of the absolute values of the entries in each row of E satisfies

$$\frac{1}{n}\sum_{j=1}^{n} |\epsilon_{i,j}| < \epsilon,$$

then $|V(M+E) - V(M)| < \epsilon$.

Proof. When E is added to M, each dot product $v \cdot r_i$ changes by less than $n\epsilon$. Thus the value of the best dot product for v_i changes by less than $n\epsilon$. So the value of the matrix changes by less than $\frac{n\epsilon}{n} = \epsilon$. □

We now prove that the upper and lower matrix-based bounds converge for the 2-player game.

Theorem 4 (Convergence theorem). *Let $L_{m,p}$ and $U_{m,p}$ be defined as above. Then*

$$\lim_{m\to\infty} V(U_{m,p}) - V(L_{m,p}) = 0.$$

Proof. We first give a brief outline. We shall begin with $L_{m,p}$ and, using the replication, erasure, and perturbation lemmas, will move to $U_{m,p}$ and notice that the value of V will not have changed much. Each entry of $L_{m,p}$ is in $\{0,1\}$. We will first remove a small proportion of the columns from $L_{m,p}$, ones that are far out of balance. Then we will replicate the remaining columns (each column replicated the same number of times). Then we will flip some values of the remaining columns, in order to bring them into balance. Each entry is changed by a perturbation $\epsilon_{i,j} \in \{1,0,-1\}$. In order to appeal to the perturbation lemma, we will need to keep the average absolute values of these perturbations small along each row, which is the same as keeping the proportion of changes small along each row.

But we will keep a symmetry of the rows, so that each row will receive the same number of changes. So it suffices to keep small the proportion of the matrix that is changed. The resulting matrix will be a replication of $U_{m,p}$, finishing the proof.

More precisely, let us call a column *balanced* if the proportion of ones in the column is p. Fix ϵ and using the law of large numbers, let m be large enough that a proportion of at most ϵ of the columns of $L_{m,p}$ are not within $m\epsilon$ bits being balanced (i. e., if 1 appears in a column t times then $|t - mp| \geq m\epsilon$). We delete any column that satisfies this condition. This deletes at most a proportion ϵ of the columns and so by the erasure lemma, this changes the value V by at most $2\epsilon/(1-\epsilon)$. For each of the remaining columns, replicate it a total of $m!$ times. By the replication lemma, this does not change the value of the matrix. Now replace each of these duplicates by one of its nearest balanced neighbors. For a column with t "1" bits the number of balanced neighbors to choose from is $\binom{t}{t-mp}$ in the case that $t \geq mp$, and $\binom{m-t}{mp-t}$ otherwise. In any case, the number of choices divides $m!$. Therefore, we can make sure that each neighbor is used the same whole number of times.

Recall that in $L_{m,p}$, each column with the same number of ones occurred the same number of times. By construction, the same will be true of our resulting matrix, N. In other words, N is just a replication of $U_{m,p}$. Furthermore, the maximum number of changes made to any column is bounded by $m\epsilon$. By symmetry, each row will receive the same number of changes, so the proportion of changes in each row is also at most ϵ. Therefore, the perturbation lemma applies, and $|V(U_{m,p}) - V(L_{m,p})| \leq 2\epsilon/(1-\epsilon) + \epsilon$. Since ϵ is arbitrary, the theorem follows. □

Appendix D. Upper bounds on $V_2(p)$ for rational p

In this Appendix, we discuss some specific upper-bound results for the 2-player game, along with proving the upper bounds claimed in Theorem 5.

D.1 Dual strategies

One tool that is useful for both upper and lower bounds is the notion of a dual strategy. For a given strategy S, we denote by S^d the dual strategy to S, where "dual" refers to switching the roles of W and B. Equivalently, we view a strategy S as a pair of functions (one for each player) $f_S^1, f_S^2 : P(X) \to X$ where X is the set of hats being considered, and we view the two functions as taking the set of black hats on the partner's head as input and outputting the player's guess. Then the dual strategy S^d has as its functions $f_{S^d}^i(A) = f_S^i(X \setminus A)$.

Remark 7. In reference to the 4 optimal 2-player strategies, it is worth noting that $S_3 = S_1^d$, $S_2 = S_2^d$ (up to reordering of hats), while $V_2(p; S_4) = V_2(p; S_4^d)$, but we do not currently know whether $S_4^d = S_4$.

The following lemma gives a formula for calculating the value of a dual strategy in terms of the value of the original strategy. For a given strategy S for the hat game with probability p of a B hat, let $\mathbb{P}_S(x_1, x_2)(p)$ be the probability that Player 1 chooses an x_1 hat on her head ($x_1 \in \{W, B\}$) and Player 2 chooses an x_2 hat on his head.

Lemma 7. *Given any strategy S, let S^d be its dual strategy. Then, for any $p \in (0,1)$,*

$$V_2(p; S^d) = 2p - 1 + V_2(1-p; S).$$

Proof. First, observe that by the definition of a dual strategy, $\mathbb{P}_{S^d}(B,B)(p) = \mathbb{P}_S(W, W)(q)$. This can be seen by pairing scenarios where all hat colors are reversed.

Observe that if the players utilize strategy S when the probability of a black hat is q, then $\mathbb{P}_S(W,W)(q) + \mathbb{P}_S(W,B)(q) = 1 - q = p$ is the probability of the first player selecting a white hat. Similarly, $\mathbb{P}_S(W,B)(q) + \mathbb{P}_S(B,B)(q) = q$ is the probability of the second player selecting a black hat.

Combining these observations, we have

$$\begin{aligned} V_2(p; S^d) &= \mathbb{P}_{S^d}(B,B)(p) \\ &= \mathbb{P}_S(W,W)(q) \\ &= p - \mathbb{P}_S(W,B)(q) \\ &= p - \big(q - \mathbb{P}_S(B,B)(q)\big) \\ &= p - q + \mathbb{P}_S(B,B)(q) \\ &= 2p - 1 + V_2(q; S). \end{aligned}$$

□

D.2 Proofs of upper bounds on $V_2(p)$

In order to derive a general upper bound on $V_2(p)$ for any rational $p \in (0,1)$, we can apply the methods of Section 3 to a $b \times b$ matrix where the first column is a 1s followed by $b - a$ 0s, and the other columns are cyclic permutations of it. We compute upper bounds on the value of this matrix game, which are in turn upper bounds on the value $V_2(p)$ of the hat game for $p = \frac{a}{b}$.

Below we restate Theorem 5 from the end of Section 3 and then prove the theorem with the help of a lemma.

Theorem 5. *For $p = \frac{a}{b} \le \frac{1}{2}$, we have $V_2(p) \le \frac{a}{b} - (1 - \frac{a}{b})^b(\frac{a}{b})$.*
For $p = \frac{a}{b} \ge \frac{1}{2}$, we have $V_2(p) \le \frac{a}{b} - (1 - \frac{a}{b})(\frac{a}{b})^b$.

Using the notation from Section 4 (e. g., with $\mathbb{P}(BB)$ for the probability that each player picks a black hat), we have

$$\mathbb{P}(BB) = \mathbb{P}(B*) - \mathbb{P}(BW).$$

We know that $\mathbb{P}(B*) = p$. Thus all that remains is to bound the value of $\mathbb{P}(BW)$.

Lemma 8. *With the additional matrix-based information at the start of this section provided to the players, there is no strategy that wins with probability greater than*

$$\frac{a}{b} - \left(\frac{a}{b}\right)^b \left(1 - \frac{a}{b}\right).$$

Proof. We use the observation above that

$$\mathbb{P}(BB) = p - \mathbb{P}(BW).$$

In particular, suppose that the players' agreed-upon strategy is that Player 1 will choose a hat from a fixed list $x_1, \ldots, x_k$ ($k \le b$). It may happen that all of these hats are black, which happens with probability $p^k \ge p^b$, in which case Player 2 picks a white hat with probability $1 - p$. This means that $\mathbb{P}(BW) \ge p^b(1 - p)$. Therefore,

$$V_2\left(\frac{a}{b}\right) \le \frac{a}{b} - \left(\frac{a}{b}\right)^b \left(1 - \frac{a}{b}\right). \qquad \square$$

To complete the proof of Theorem 5, we observe that by the duality discussed above, we also have

$$V_2\left(\frac{a}{b}\right) \le \frac{a}{b} - \left(1 - \frac{a}{b}\right)^b \left(\frac{a}{b}\right).$$

We compare the two bounds and discover that they are stronger on the intervals claimed in Theorem 5.

Remark 8. The upper bound in Lemma 8 is sharp for the case when $p = \frac{b-1}{b}$ in the sense that one can describe a strategy for the $b \times b$ matrix game that succeeds with this probability. (However, this remains only an upper bound on $V_2(p)$ for the original hat game, in which the players do not receive hints.)

Remark 9. Supposing that $V(p)$ is differentiable, using Theorem 5, one can calculate an upper bound on the derivative of the function $V(p)$ at 0, and a lower bound at 1. Also, one can use the strategies described in Section 2 to calculate a lower bound at 0 and upper at 1. One obtains $\frac{1}{2} \le V'(0) \le 1 - \frac{1}{e}$ and $1 + \frac{1}{e} \le V'(1) \le \frac{3}{2}$.

The upper bounds on $V_2(p)$ in Theorem 5 are quite good for p of the form $\frac{1}{b}$ or $\frac{b-1}{b}$, as indicated in Figure 8.1 (at the end of Section 3) by their proximity to the continuous lower-bound curve. However, they can be improved by using larger and less structured hint matrices. We give examples of this improvement below for the cases $p = 1/3$ and $p = 2/3$, whose upper bounds from Theorem 5 are $19/81 = 0.234567\ldots$ and $46/81 =$

0.567901..., respectively. (The respective lower bounds from Theorem 3 in Section 2 are 0.205555... and 0.538888....)

The hint matrix

$$U = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}$$

yields an upper bound of 0.221307... for $p = \frac{1}{3}$.

The hint matrix

$$U = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

yields an upper bound of 0.554641... for $p = \frac{2}{3}$.

Appendix E. *n*-player results

Note

As in Section 4, throughout Appendix E we will think of the players as the wise men of Tanya Khovanova's problem statement [5] and use singular masculine pronouns when referring to individual players.

E.1 Proof of Theorem 8

Below we prove the partial converse Theorem 8 from Section 4 to Peter Winkler's order-$(1/\log(n))$ strategy for n players with arbitrary black-hat probability p. As in Section 4, we define $q := 1 - p$ and $r := 1/q$.

Theorem 8. *Suppose that each player must choose the* lowest *level on his own head that has a black hat (with the team failing if any player has only white hats). Then, using* $\tilde{V}_n(p)$ *to refer to the optimal probability as* $h \to \infty$ *that the players succeed (for any given black-hat probability* $p \in (0,1)$*) under this more stringent requirement, for each* ϵ *with* $0 < \epsilon \le 1/4$ *we have*

$$\tilde{V}_n(p) \le (1+\epsilon)/\log_r(n) \text{ for all sufficiently large } n.$$

Proof. For $j = 1, \dots, n$, let the random variable Y_j be the lowest level on which Player j has a black hat. We will restrict attention to the first $t = \lceil (1-\epsilon/2)\log_r(n) \rceil$ levels, and we will establish the desired upper bound on $\tilde{V}_n^{(h)}$ for all $h \ge t$, from which the bound will follow immediately for $\tilde{V}_n = \sup_h \tilde{V}_n^{(h)}$. (If any players have all-white h-hat stacks, the team would automatically lose the actual game, but such players will be considered to have $Y_j > t$ and will generously be exempted from having to guess at all for the purpose of this upper bound.)

An accomplice will uniformly at random choose a level k from $\{1, \dots, t\}$ and then uniformly at random choose a player (j^*, say) from among those who happen to have $Y_j = k$ for the given realization. The accomplice will then inform the chosen player j^* of his special status but will not tell him his lowest-black-hat level k. Player j^* is the only player required to guess his level Y_j; the accomplice will tell the other $(n-1)$ players (including those with $Y_j > t$) their own values Y_j, and if any players have all-white h-hat stacks, they will be exempted from having to guess their values Y_j.

We argue below that, with very high probability, all values $k \in \{1, \dots, t\}$ occur and, in fact, occur about as many times as expected at random. It then follows that, even with this substantial help from the accomplice (which any or all of the players can ignore if they wish), the one player who is not told his own level still cannot pick the correct level with probability greater than $(1+\epsilon)/\log_r(n)$ as $n \to \infty$.

Now we formalize the claim above. For each $k \in \{1, \dots, t\}$, let X_k be the number of players j with $Y_j = k$. Then μ_k, the expected value of X_k, satisfies

$$\mu_k = npq^{k-1} \ge npq^{t-1} \ge pn^{\epsilon/2}.$$

By a 2-sided multiplicative Chernoff bound for each value k, followed by a union bound over the t possible values for k, we have, for each δ with $0 < \delta \le 1$,

$$\begin{aligned} \mathbb{P}\big((1-\delta)\mu_k \le X_k \le (1+\delta)\mu_k \quad \text{for all } k \in \{1,\dots,t\}\big) &\ge 1 - 2t\exp\big(-(\delta^2/3)pn^{\epsilon/2}\big) \\ &= 1 - o\big(1/n^C\big) \quad \text{for every } C > 0 \\ &= 1 - o\big(1/\log_r(n)\big). \end{aligned}$$

Thus, for the purpose of proving Theorem 8, we can neglect the probability that some value $k \in \{1, \dots, t\}$ fails to occur or occurs with relative frequency significantly different from its expected value.

Now the chosen player j^* sees the lowest-black-hat level Y_j for each of the other players j and knows how he was chosen; this allows him to compute a well-defined posterior probability distribution for his own level $Y_{j^*} \in \{1, \dots, t\}$. Since $X_k/\mu_k \in [1-\delta, 1+\delta]$ for all $k \in \{1, \dots, t\}$ with all but asymptotically negligible probability, it follows readily from Bayes' theorem that the ratios of posterior probabilities

$$\mathbb{P}_{\text{post}}(Y_{j^*} = k_1)/\mathbb{P}_{\text{post}}(Y_{j^*} = k_2)$$

are in $[(1-\delta)/(1+\delta), (1+\delta)/(1-\delta)]$ for all $k_1, k_2 \in \{1, \dots, t\}$. If we let $\delta = \epsilon/16$, say, it then follows readily that, for each $k \in \{1, \dots, t\}$,

$$\mathbb{P}_{\text{post}}(Y_{j^*} = k) \le \frac{1+\epsilon/4}{t} \le \frac{(1+\frac{\epsilon}{4})/(1-\frac{\epsilon}{2})}{\log_r(n)}.$$

Recalling that $0 < \epsilon \le 1/4$ and taking into account the asymptotically negligible probability of atypical events, we find that, even with the help from the accomplice (which cannot hurt the players, since they are free to ignore extra information),

$$\tilde{V}_n(p) \le \frac{1+\epsilon}{\log_r n}$$

for all sufficiently large n. □

E.2 Proof that misère success probability is bounded away from 1

Now we complete the proof of Theorem 10 from Section 4 to show that W_n, the probability of success for n players in the misère game, is bounded away from 1 for each n. We begin by recalling the definition of the pairs of integers (r_j, s_j) for $j \ge 1$ as below; these will be the dimensions of "hint matrices" given to the various players for the n-player misère game.

Definition 3. Let $(r_1, s_1) = (2, 2)$, $(r_2, s_2) = (4, 6)$, and for each $k \ge 3$, define

$$r_k = 2\prod_{j=1}^{k-1} s_j \quad \text{and} \quad s_k = \binom{r_k}{r_k/2}.$$

Theorem 10. *Letting $W_n = \sup_S \mathbb{P}(\text{At least 1 of n players chooses a white hat})$ over all possible n-player strategies as h, the number of hats per head, grows to infinity, we have the following:*

$$W_2 = 1/2 + V_2 \in [17/20, 193/224] \quad (\textit{hence in [0.85, 0.8616...]})$$

and, for $n \geq 3$,

$$1 - (1/2)^n \leq W_n \leq 1 - \frac{1}{(r_n/2)2^{(r_n/2)}}.$$

Proof. The bounds on W_2 and the lower bounds on W_n for all $n \geq 3$ were already established in Section 4. The upper bounds follow from a "hint" technique generalizing the upper-bounding techniques used earlier for V_2 and V_3.

For $1 \leq j \leq n-1$, Player j is given an $r_j \times s_j$ hint matrix with r_j and s_j as defined just before the statement of the current theorem. Player n is given no hint, but by the usual argument, he can without loss of generality restrict attention to (at most) the first $r_n/2$ levels on his own head, since this is the product of the number of columns in the other $n-1$ players' hint matrices, which is the total number of distinguishable situations in which Player n can find himself. With probability at least $1/(2^{r_n/2})$ (strictly greater than this if Player n does not actually use all possible $r_n/2$ levels on his own head), he will have black hats on all of the levels from which he chooses, and Players 1 through $n-1$ will all know when they are in this situation. In this case, it is up to the first $n-1$ players to choose at least 1 white hat.

Now, proceeding inductively downstream from Player $n-1$ through Player 2, conditioning on what is seen on the heads of all the upstream players $k+1, \ldots, n$, each Player k sees one of $s_1 \cdot s_2 \ldots \cdot s_{k-1} = r_k/2$ possible joint column choices for Players 1 through $k-1$. Even if Player k assigns a different row of his own hint matrix to each of these $r_k/2$ distinguishable downstream possibilities, one of his $\binom{r_k}{r_k/2}$ columns will contain 1s in all $r_k/2$ of these rows. (If Player k sometimes assigns the same row to different distinguishable downstream observations, there will be multiple columns of his hint matrix that contain 1s in all rows that he actually uses.)

Thus, conditioned on whatever Player k observes upstream and downstream (and whatever strategy he has committed himself to), with probability at least $1/s_k$ he will have been assigned a column by the referee that forces him to choose a black hat, inductively leaving the downstream Players 1 through $k-1$ with the responsibility of choosing at least one white hat. Finally, if Players 2 through n have all been assigned these most unfavorable columns by the referee, Player 1 will know this fact and will have the burden of choosing a white hat on his own head. However, his 2 possible hint columns are equally probable and differ from each other on every level, so Player 1 will fail with probability 1/2. Multiplying all n of the players' respective conditional failure probabilities together, we see that they must lose the misère game with probability at least

$$(1/2^{r_n/2}) \cdot (1/s_{n-1})(1/s_{n-2}) \ldots (1/s_1).$$

Since r_n is defined as $2s_1s_2 \ldots s_{n-1}$, the claimed upper bound on W_n follows immediately. □

E.3 Proof that misère success probability W_n decreases in n

The next result, mentioned in Section 4, shows that the optimal success probability W_n for the n-player misère game with $p = 1/2$ (in which *at least* one player must point to a *white* hat) is strictly increasing in n.

Theorem 13. *The optimal misère success probabilities satisfy* $W_{n+1} > W_n$ *for all* $n \geq 1$, *with* $\lim_{n\to\infty} W_n = 1$.

Proof. The fact that $\lim_{n\to\infty} W_n = 1$ follows immediately from Theorem 10. Now, much as in the proof of the previous theorem, we have

$$\mathbb{P}(B\ldots B,*) = \mathbb{P}(B\ldots B,B) + \mathbb{P}(B\ldots B,W)$$

for any $(n+1)$-player strategy, so

$$1 - \mathbb{P}(B\ldots B,*) = \big(1 - \mathbb{P}(B\ldots B,B)\big) + \big(1 - \mathbb{P}(B\ldots B,W)\big) - 1$$

for any $(n+1)$-player strategy. Thus

$$\sup\{1 - \mathbb{P}(B\ldots B,*)\} = \sup\{\big(1 - \mathbb{P}(B\ldots B,B)\big) + \big(1 - \mathbb{P}(B\ldots B,W)\big)\} - 1,$$

where the supremum is over all $(n+1)$-player strategies. Then

$$\sup\{1 - \mathbb{P}(B\ldots B,*)\} \leq \sup\{1 - \mathbb{P}(B\ldots B,B)\} + \sup\{1 - \mathbb{P}(B\ldots B,W)\} - 1.$$

Since the maximal probability of avoiding $(B\ldots B,B)$ is the same as the maximal probability of avoiding $(B\ldots B,W)$ when $p = 1/2$, we have

$$W_n \leq 2W_{n+1} - 1,$$

so $W_{n+1} \geq (1 + W_n)/2 > (W_n + W_n)/2$, where the last inequality is because $W_n < 1$ for each n.

Thus we have the strict inequality $W_{n+1} > W_n$ for all $n \geq 1$. □

E.4 Details of best strategies known for n players

Now we give details about the best lower bounds known for V_3 through V_{12} (all for black-hat probability $p = 1/2$) that were omitted in Section 5. We discuss strategies found by hill climbing for V_3 and V_4 and describe generalizations of the basic Winkler strategy from Theorem 7 in Section 4. Finally, we analyze (recursive) white-reset and black-reset enhancements for t-hat tiers that improve all of the strategies above.

In an email sent in 2014 to various hats enthusiasts, Jay-C Reyes and Larry Carter reported what were then the best lower bounds known on V_3 and V_4 for $p = 1/2$. Their

bounds were constructive and came from hill-climbing on symmetric strategies for 4 hats per player. (Since we will later extend these strategies by considering 4-hat *tiers*, we use t rather than h to refer to the number of hats per player in each basic strategy.) The Carter–Reyes 4-player result leads to what is still the best known lower bound for V_4. Their 3-player search has since been adapted to consider $t = 3, 4, 5$, or 6 hats per player. Before incorporating the "reset" enhancements, we obtain the following lower bounds from these searches:

$$V_3 \geq 9120/(2^5)^3 = 9120/32768 = 0.278320\ldots,$$
$$V_4 \geq 14845/(2^4)^4 = 14845/65536 = 0.226516\ldots.$$

Remark 10. The 3-player strategy, which uses 5 levels, can be described by a symmetric 32×32 matrix with values in $\{1, 2, 3, 4, 5\}$ that is used by all 3 players. The 4-player strategy, which uses 4 levels, can be described by a $16 \times 16 \times 16$ symmetric tensor with values in $\{1, 2, 3, 4\}$ that is used by all 4 players.

For $n \geq 5$, our best lower bounds on V_n come from Winkler-style strategies in which players focus on the first $t \approx \log_2(n)$ levels. With Y_j defined as the lowest level on which Player j has a black hat, players using the original Winkler strategy hoped that $Y_1 + \cdots + Y_n$ would have some particular residue (mod t) with probability as much above the guaranteed $1/t$ as possible. This best-residue probability can be improved a little if each player is allowed to apply some permutation π_j to his value Y_j before the values are summed. (The players' permutations on $\{1, \ldots, t\}$ can be different for different players, as long as they are fixed during the strategy session.) For example, for the usual case $p = 1/2$, one appears to do better by computing the alternating sum $(Y_1 - Y_2 + Y_3 - \cdots + (-1)^{n+1} Y_n) \mod t$ than the straight sum (mod t), since the alternating sum concentrates the probability mass more effectively.

Still more generally, one can use $t \times t$ Latin squares to "add in" one player's level at a time to the "running sum," possibly permuting the output symbols after each new player's level is folded in. When $t = 4$, for example, one does better by using Latin squares corresponding to the Klein 4-group than by using squares corresponding to addition or subtraction (mod 4).

The bounds for $5 \leq n \leq 8$ all come from applying the Winkler idea to 4-hat tiers, representing each level within a tier by a dibit in $\{00, 01, 10, 11\}$, and XORing the n dibits corresponding to the lowest level in the tier (if any) on which each player has a black hat, resetting to the next tier if necessary. The players hope that they will each have at least one black hat within the 4-hat tier and that the XOR of the n resulting dibits will be 00, and they each choose the corresponding one of 4 levels on their own heads. It turns out that for this strategy, the players win if and only the mod-2 sum of all their dibits really is 00.

However, for many strategies (most notably, for the original strategy of straight addition of lowest-black-hat levels (mod t)), there is "secondary success probability"

(or "bonus" probability) coming from cases in which the actual "sum" of the n lowest levels differs from the targeted value but the players nonetheless serendipitously each point to a black hat. (In fact, because of this bonus probability, it turns out that straight mod-t addition almost always yields more overall success than alternating addition and subtraction (mod t), even though the latter almost always yields higher "primary" success probability than the former.) We have accounted for this bonus probability in all of our best known lower bounds in Section 5.

The XOR strategies with $t = 4$ equivalently use the Latin square

$$\begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix}$$

associated with $Z_2 \times Z_2$ to combine values corresponding to the lowest-black-hat level within the tier for each successive player. These strategies outperform strategies based on mod-4 addition and subtraction (and many permutation-based generalizations thereof), apparently because the Latin square associated with $Z_2 \times Z_2$ concentrates probability within the first two categories more effectively than the Latin squares associated with those other arithmetic operations.

For $9 \leq n \leq 12$, our best strategies use 5-hat tiers (i. e., $t = 5$). There are two distinct isotopy classes of 5×5 Latin squares (as there are for 4×4 Latin squares), and once again, it appears that the isotopy class *not* associated with mod-t addition or subtraction does a better job of concentrating probability within 2 of the t categories. The 5×5 isotopy class that we found to work best corresponds to a nonassociative quasigroup (in fact, to a loop). Our best results are probably not optimal even within the class of strategies we considered, since we used a greedy search algorithm, but they all begin by combining the first two mod-5 values using the Latin square

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 4 & 2 \\ 2 & 4 & 0 & 1 & 3 \\ 3 & 2 & 4 & 0 & 1 \\ 4 & 3 & 1 & 2 & 0 \end{pmatrix}.$$

Later mod-5 values are folded in using Latin squares isotopic to this first square, but not the same square (or quasigroup) for each new player.

All of the strategies above are improved slightly by working with t-hat "tiers" and recursively "resetting" (shifting up t levels at a time) whenever a player sees another player with an all-white stack of hats within the current tier. In any such situation, the players would certainly lose without resetting, and if there is only one player with an all-white stack within the current tier, he will fail to get the memo, so the team will still lose. However, if 2 or more players have all-white stacks in the current t-level tier, all n

players will reset together and give themselves an independent chance of winning at the next tier. By recursively resetting, the players succeed for all placements of hats on the lowest t levels that would have won without resetting, but now the denominator of their success probability is reduced from 2^{tn} to

$$(2^t - 1)^n + \binom{n}{1}(2^t - 1)^{n-1}.$$

One can refine the reset strategy a bit. If any player (call him Player j) sees only *black* hats within the first t-level tier on *all* of the other $n-1$ heads, he can reset (shifting up t levels on all of the stacks that he sees).

If he has neither all-black nor all-white on his own first t levels, then the other $n-1$ players will stay on the first t levels and will all be guaranteed to guess correctly. In this case, the team wins with average conditional probability 1/2 whether Player j (the nonmonochromatic player) guesses from his first t-level tier or from any other t-level tier.

If Player j has only white hats in his first tier, then the other $n-1$ players will also reset (since they see his all-white stack), and in this case, the team will get a fresh start at the next tier of t levels, whereas they would have lost for sure if they had used no resets or only the reset-on-white strategy. If Player j has all black hats in his first t-level tier, then all n players have all-black first tiers, so they will all reset to the next tier and give themselves a fresh start of winning with conditional probability V_n rather than the conditional probability of 1 that they would have enjoyed if they had stayed put.

We see that this reset-on-black strategy will change the probability of winning (with respect to the earlier strategy of resetting only when at least one player has an all-white first tier) only when everyone has a monochromatic first tier, with at most one player having an all-white first tier. This situation occurs with probability $(n+1)(1/2)^{nt}$, and the conditional net gain in success probability in this case is at least

$$\frac{n}{n+1}(\underline{V}_n - 0) + \frac{1}{n+1}(\underline{V}_n - 1) = \underline{V}_n - 1/(n+1),$$

where $\underline{V}_n$ is any valid lower bound on V_n. Thus as long as we have a starting strategy that achieves success probability strictly greater than $1/(n+1)$, as we do for all $n \geq 2$, this augmented resetting strategy helps.

When this reset-on-black policy is implemented recursively, the denominator of the players' success probability is reduced by $n+1$, to

$$(2^t - 1)^n + n(2^t - 1)^{n-1} - (n+1),$$

and the numerator is reduced by 1.

With resetting incorporated into the basic t-hat strategies, we obtain the following lower bounds on V_3 and V_4.

Lemma 9. *For the 3-player and 4-player games, we have*

$$V_3 \geq 9119/32670 = 0.279124\ldots \quad \textit{and} \quad V_4 \geq 14844/64120 = 0.231503\ldots.$$

Proof. For $n = 3$, we used 5 levels and found a symmetric strategy (described by a symmetric 32×32 matrix with values in $\{1, 2, 3, 4, 5\}$ that is used by all 3 players) that wins 9120 times out of 32768. With white and black resets, this yields success probability

$$\frac{9120 - 1}{31^3 + \binom{3}{1}(31)^2 - (3 + 1)} = \frac{9119}{(31)^2(31 + 3) - 4} = \frac{9119}{32670}.$$

For $n = 4$, Reyes and Carter used 4 levels and found a symmetric strategy that wins with probability 14845/65536, which is improved to 14844/64120 with recursive resetting. □

For large values of n, the best lower bounds we know how to achieve are only very slightly above the obvious lower bounds from the basic Winkler strategy with t-hat tiers and a reset to the next tier if a player sees at least one other player who has only white hats in his first tier. (Resetting when one sees only *black* hats on *all* other players' heads provides much less help for large n.) With just the reset on white, we obtain

$$V_n \geq \max_t \frac{1}{(2^t - 1)^n + n(2^t - 1)^{n-1}} \frac{(2^t - 1)^n}{t} = \max_t \frac{2^t - 1}{2^t - 1 + n} \frac{1}{t},$$

which decreases asymptotically as $1/\log_2(n)$, essentially as argued in Theorem 7.

Bibliography

[1] E. Berlekamp and J. Buhler, Puzzles Column. *Emissary* (2014). www.msri.org.

[2] E. Burke, S. Gustafson and G. Kendall, A puzzle to challenge genetic programming, in *Genetic Programming*, pp. 238–247, Springer, 2002.

[3] E. Friedgut, G. Kalai and G. Kindler, The success probability in Lionel Levine's hat problem is strictly decreasing with the number of players, and this is related to interesting questions regarding Hamming powers of Kneser graphs and independent sets in random subgraphs, 2021. Preprint, arXiv:2103.01541 [math.CO].

[4] T. Friedrich and L. Levine, Fast simulation of large-scale growth models, *Random Struct. Algorithms*, **42** (2013), 185–213.

[5] T. Khovanova, How many hats can fit on your head? 2011. blog.tanyakhovanova.com/2011/04.

[6] G. Muller, The Axiom of Choice is Wrong. The Everything Seminar, 2007. https://cornellmath.wordpress.com/2007/09/13/the-axiom-of-choice-is-wrong/.

[7] D. Velleman and S. Wagon, *Bicycle or Unicycle*, MAA Press, 2020.

Joshua Cooper and Grant Fickes

Recurrence ranks and moment sequences

This work is dedicated to the memory of Ron Graham, a gentle giant of the highest scholarly caliber, a peerless and playful teacher, and an extraordinarily generous person. His scientific contributions will live on in innumerable ways, particularly in his endless demonstrations that theory and application are not just complementary, but profoundly interwoven. Here, we invoke three persistent themes of his work: recurrence relations, algorithmic thinking, and expansive TFAE statements.

Abstract: We introduce the "moment rank" and "unitary rank" of numerical sequences, close relatives of linear-recursive order. We show that both parameters can be characterized by a broad set of criteria involving moments of measures, types of recurrence relations, Hankel matrix factorizations, Waring rank, analytic properties of generating functions, and algebraic properties of polynomial ideals. In the process, we solve the "complex finite-atomic" and "integral finite-atomic" moment problems: which sequences arise as the moments of a finite-atomic complex-/integer-valued measures on $\mathbb{C}$?

1 Introduction

We begin with a motivating problem, the original impetus for this work: Suppose that G is a finite, simple graph; then G is associated with a *characteristic polynomial* whose roots are its adjacency eigenvalues. This polynomial, despite substantial attention in the literature dating back to at least 1957 [6], is still the subject of many open problems. One example is the question of describing the multiplicity of zero as a root, i.e, the nullity of the adjacency matrix $A(G)$. If $\phi(x) = \det(xI - A(G))$ is the polynomial, then this multiplicity is the largest m so that $\phi(x)/x^m$ is also a polynomial; thus the degree of $\phi(x)/x^m$ (or, for this application, its normalized reciprocal polynomial $\overline{\phi}(x) = \det(I - A(G)x)/C$, where C is chosen so that $\overline{\phi}$ is monic) then encodes this quantity as $m = \deg\phi - \deg\overline{\phi}$. Note that

$$\log\overline{\phi}(x) = \sum_{i=1}^{r}\log(1-b_ix) = \sum_{i=1}^{r}\sum_{j=1}^{\infty}\frac{b_i^jx^j}{j}$$

where $\{b_i\}_{i=1}^{\deg\overline{\phi}}$ are the nonzero roots of ϕ, whereupon the question becomes of bounding the smallest r so that c_j can be written as a sum of r jth powers, where jc_j is the jth coefficient of $\log\overline{\phi}$. As will be defined below, this is exactly the "unitary rank" of the

Joshua Cooper, Grant Fickes, Department of Mathematics, University of South Carolina, Columbia, SC, USA, e-mails: cooper@math.sc.edu, gfickes@email.sc.edu

https://doi.org/10.1515/9783110754216-009

sequence $(jc_j)_{j\geq 1}$. Lest this seem like a roundabout way to study the quantity m, note that, using the log det = tr log identity, it is straightforward to see that $jc_j = \mathrm{tr}(A(G)^j)$, the number of closed walks in G of length j, for each $j \geq 1$. See condition (6) in Theorem 2 for this connection with log-polynomial degree.

More generally, one might ask for the simplest recurrence that a combinatorial sequence $\mathcal{C}$ satisfies, as a kind of measure of complexity. If $\mathcal{C}$ is "C-finite," then it satisfies a linear recurrence with constant coefficients, and the order of that recurrence captures this complexity. It is natural then to ask how to compute this order, or even if it is finite. Famously, for example, this is an open question for the sequence A_n equal to the number of permutations of n with no 1324 pattern (i. e., $\sigma \in S_n$ so that there exist no $a < b < c < d$ so that $\sigma(a) < \sigma(c) < \sigma(b) < \sigma(d)$). The theory of such sequences is extremely well-trodden territory, and it is simplest in the case that the characteristic polynomial—the polynomial whose coefficients are the same as those of the recurrence—has no repeated roots. We focus on this case presently.

Another way in which the smallest order of a linear recurrence satisfied by a sequence appears in the literature is in the context of the venerable "moment problem." Here, one asks whether a sequence can arise as the sequence of moments of various kinds of distributions: important instances include (positive) measures on $\mathbb{R}$, $[0,\infty)$, $[0,1]$, or $\mathbb{T} = \exp(i\mathbb{R})$ (the "Hamburger," "Stieltjes," "Hausdorff," and "Toeplitz"/"trigonometric" moment problems, respectively); signed measures on $\mathbb{R}$ (already considered by Hausdorff [14]); or atomic measures [7, 10]. Important related lines of research in this area include truncated and multidimensional moment problems [8, 9] and generalized moment problems and their numerical solution [17]. See [18, 24] for extensive explorations of this old and very broad range of topics. The question of when a sequence does *not* arise as moments of a finite-atomic measure was recently addressed [3], a topic with a long history connected with totally positive matrices [12, 21], strong log-concavity/unimodality [4, 5], continued fractions and Padé approximants [25]. Here, we add to the literature on moment problems by addressing the case of the underlying space being $\mathbb{C}$ with the two conditions that either (1) the measures are complex-valued and finite-atomic, or (2) the (positive) measures are finite-atomic with integer masses.

Yet another large constellation of topics closely connected with recurrence rank is the theory of Hankel matrices [19, 28], matrices which are constant on anti-diagonals. These matrices—and, more generally, Hankel operators—play an important role in combinatorial sequence transforms [13, 19], numerical methods in signal processing [5], and Riordan arrays [20]. The determinants of Hankel matrices, known as "catalecticants," are objects of study going back as far as Sylvester's work in the 1850s [27], and lives on in invariant theory [26], polynomial positivity [1], Waring rank and binary forms [23], and the theory of orthogonal polynomials [16].

Clearly, the subject matters connected with linear recurrence order are vast, and there is not space here to discuss them all (and many important references are therefore omitted, although they can be found by following threads in the aforementioned

references). Indeed, so much work has been done on related topics over such a long period of time that it is difficult to trace their history. This work, in addition to presenting several new results, is an attempt to relate and distill these perspectives into one focused on the matter of linear recurrence rank, the order of the shortest recurrence a sequence satisfies, in the particularly interesting cases we term "moment rank" and "unitary rank." We attempt to keep the below exposition mostly self-contained, which entails borrowing a variety of arguments from the literature, indicated whenever possible.

In the next section, we introduce notation, definitions, and state some basic results. In Section 3, we present our first main theorem, a wide-ranging TFAE statement about moment rank, and discuss a few consequences. In Section 4, we present our second main theorem, another TFAE statement about unitary rank, and some consequences thereof.

2 Preliminaries

Suppose the sequence $\mathcal{C} = (c_n)_{n=0}^{\infty}$ satisfies an rth order linear recurrence relation

$$\sum_{n=0}^{r} a_n c_n = 0 \tag{9.1}$$

Then, by classical results [11], the elements of $\mathcal{C}$ can be expressed as

$$c_n = \sum_{i=1}^{r} \alpha_i \beta_i^r$$

for some $\{\alpha_i\}_{i=1}^r$, where $\{\beta_i\}_{i=1}^r$ are the roots of the degree-r polynomial $p(x) = a_0 \prod_{i=1}^r (x - \beta_i) = \sum_{i=0}^{r-1} a_i x^r$, as long as the β_i are distinct. The $\{\alpha_i\}_{i=1}^r$ can be obtained by solving the linear system

$$\forall j \in \{0, \ldots, r-1\}, \quad \sum_{i=1}^{r} \alpha_i \beta_i^{j+1} = c_j \tag{9.2}$$

These observations motivate the following definition.

Definition 1. The sequence $\mathcal{C} = (c_n)_{n=0}^N$ (with $N = \infty$ allowed) is said to have *recurrence rank* r if r is the smallest positive integer so that $\mathcal{C}$ satisfies a linear recurrence of order r. If $N = \infty$, we write $\operatorname{rrank}(\mathcal{C})$ for the recurrence rank.

Definition 2. The sequence $\mathcal{C} = (c_n)_{n=0}^N$ (with $N = \infty$ allowed) is said to have *moment rank* r if r is the smallest positive integer so that there exists a set of nonzero complex numbers $\{\alpha_i\}_{i=1}^r$ and distinct nonzero $\{\beta_i\}_{i=1}^r$ so that $c_n = \sum_{i=1}^r \alpha_i \beta_i^{n+1}$ for all $0 \le n \le N$. If $N = \infty$, we write $\operatorname{mrank}(\mathcal{C})$ for the moment rank.

Use of this definition depends on the uniqueness of the quantity for a given sequence. This motivates the following lemma.

Lemma 1. $\operatorname{mrank}((c_n)_{n\geq 0})$ *is well-defined.*

Proof. Suppose, by way of contradiction, that there are two sets $\{\beta_i\}_{i=1}^r$ and $\{\beta'_j\}_{j=1}^s$ of nonzero distinct complex numbers along with sets of nonzero complex numbers $\{\alpha_i\}_{i=1}^r$ and $\{\alpha'_j\}_{j=1}^s$ so that

$$c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1} = \sum_{j=1}^{s} \alpha'_j \beta'^{(n+1)}_j$$

Write $f(z) = \sum_{n=0}^{\infty} c_n z^n$. Because $\max_i |\beta_i|$ and $\max_j |\beta'_j|$ are finite, the following is true around a sufficiently small ball about $z = 0$:

$$\sum_{n=0}^{\infty}\sum_{i=1}^{r} z^n \alpha_i \beta_i^{n+1} = \sum_{n=0}^{\infty}\sum_{j=1}^{s} z^n \alpha'_j \beta'^{(n+1)}_j$$

$$\sum_{i=1}^{r}\sum_{n=0}^{\infty} z^n \alpha_i \beta_i^{n+1} = \sum_{j=1}^{s}\sum_{n=0}^{\infty} z^n \alpha'_j \beta'^{(n+1)}_j$$

$$\sum_{i=1}^{r} \frac{\alpha_i \beta_i}{1 - z\beta_i} = \sum_{j=1}^{s} \frac{\alpha'_j \beta'_j}{1 - z\beta'_j}$$

These two functions are equal, so they have the same set of (simple) poles; thus $\{\beta_i\}_i = \{\beta'_j\}_j$ and $r = s$. Furthermore, since the residues of these poles are proportional to the multiplicity of the $\alpha_i\beta_i$ and $\alpha'_j\beta'_j$ values, we also have that $\{\alpha_i\}_{i=1}^r = \{\alpha'_j\}_{j=1}^s$. □

The beginning of this section motivates the moment rank definition by the satisfaction of a linear recurrence. The following definition introduces a specific kind of linear recurrence we consider throughout the work.

Definition 3. An rth order linear recurrence of the form $\sum_{j=0}^{r} a_j c_{j+t} = 0$ satisfied by the sequence $(c_j)_{j\geq 0}$ for all $t \geq 0$ is said to be *simple* if the characteristic polynomial $\sum_{j=0}^{r} a_j x^j$ has distinct roots.

The characteristic polynomial of a recurrence provides a way to translate between polynomials and recurrences. Given a sequence which satisfies a linear recurrence, there are methods to define other recurrences of higher orders, which the given sequence satisfies. We consider this idea in the context of characteristic polynomials, motivating the following definition and the lemma that follows.

Definition 4. Given a complex sequence $\mathcal{C} = (c_n)_{n=0}^{\infty}$, let

$$R_{\mathcal{C}} = \left\{ \sum_{j=0}^{r} a_j x^j : a_j \in \mathbb{C} \text{ and } \sum_{j=0}^{r} a_j c_{j+t} = 0 \text{ for all } t \geq 0 \right\}$$

be the set of characteristic polynomials of arbitrary finite order linear recurrences the sequence $\mathcal{C}$ satisfies for all $t \geq 0$. We note that the zero polynomial is a trivial element of $R_{\mathcal{C}}$.

The set above is a subset of one variable polynomials with complex coefficients. We explore useful algebraic properties of this subset of $\mathbb{C}[x]$.

Lemma 2. *Given a complex sequence $\mathcal{C} = (c_n)_{n=0}^{\infty}$, the set $R_{\mathcal{C}}$ is an ideal in $\mathbb{C}[x]$.*

Proof. We know the zero polynomial is in $R_{\mathcal{C}}$ by definition. Let $a(x) = \sum_{i=0}^{r} a_i x^i$ and $b(x) = \sum_{j=0}^{s} b_j x^j$ be arbitrary elements of $R_{\mathcal{C}}$. Then $\sum_{i=0}^{r} a_i c_{i+t} = 0$ and $\sum_{j=0}^{s} b_j c_{j+t} = 0$, giving that

$$\sum_{i=0}^{r} a_i c_{i+t} + \sum_{j=0}^{s} b_j c_{j+t} = 0.$$

Thus $R_{\mathcal{C}}$ is closed under addition.

Now let $\{d_k\}_{k=0}^{q}$ be complex constants so that $d(x) \in R_{\mathcal{C}}$ where $d(x) = \sum_{k=0}^{q} d_k x^k$. Let $p(x) = \sum_{l=0}^{m} p_l x^l$ be an arbitrary polynomial with complex coefficients. If $d(x)$ or $p(x)$ is the zero polynomial, then $p(x) \cdot d(x) \in R_{\mathcal{C}}$. Suppose now that $d(x)$ and $p(x)$ are not identically zero. Since the sequence $\mathcal{C}$ satisfies a recurrence with characteristic polynomial $d(x)$, the generating function $\Phi(x)$ of the sequence has denominator $d(x)$. Therefore, $p(x) \cdot d(x) \cdot \Phi(x)$ is a polynomial, so $p(x) \cdot d(x) \cdot \Phi(x)$ is the characteristic polynomial for a recurrence satisfied by $\mathcal{C}$, giving that $R_{\mathcal{C}}$ is closed under multiplication by elements of $\mathbb{C}[x]$. □

Note that since $\mathbb{C}[x]$ is a principal ideal domain, $R_{\mathcal{C}}$ is generated by one complex polynomial. Moreover, we call $R_{\mathcal{C}}$ the *recurrence ideal* of the sequence $\mathcal{C}$.

Corollary 1. *Given a complex sequence $\mathcal{C} = (c_n)_{n=0}^{\infty}$ let $R_{\mathcal{C}}$ be generated by $p(x)$. If $p(x)$ has repeated roots, then $\mathcal{C}$ does not satisfy a simple linear recurrence of any order. Thus $\mathcal{C}$ satisfying a simple linear recurrence implies that $p(x)$ has distinct roots, i. e., if* $\operatorname{mrank}(\mathcal{C}) < \infty$, *then* $\operatorname{rank}(\mathcal{C}) = \operatorname{mrank}(\mathcal{C})$.

The algebraic structure of the recurrence ideal gives rise to useful properties of simple linear recurrences, some of which are investigated by Lemma 3. The properties addressed in the following two lemmas are useful in the proof of Theorem 1.

Lemma 3. *Let r be the smallest positive integer so that the sequence $(c_j)_{j\geq 0}$ satisfies the simple rth order linear recurrence $\sum_{j=0}^{r} a_j c_{j+t} = 0$. Then the following observations hold:*

1. $a_r \neq 0$.
2. *The roots of $p(x) = \sum_{j=0}^{r} a_j x^j$ are nonzero.*
3. *The polynomial $q(x) = \sum_{j=0}^{r} a_j x^{r-j}$ (the "reciprocal" of the characteristic polynomial) has r distinct, nonzero roots.*

Proof. (1) This follows trivially from the minimality of r.

(2) Due to Corollary 1, $p(x)$ is the generator of the recurrence ideal, $R_{\mathcal{C}}$. If x is a factor of $p(x)$, that corresponds to an index shift in the corresponding recurrence. Then $a(x) = p(x)/x$ is also the characteristic polynomial of a linear recurrence satisfied by $\mathcal{C}$, so $a(x) \in R_{\mathcal{C}}$, contradicting the minimality of $\deg(p)$.

(3) Since $a_r \neq 0$ by (1), $q(x)$ has nonzero constant term so $q(0) \neq 0$. Therefore, the reciprocal of $q(x)$ is well-defined, with $p(x)$ the reciprocal of $q(x)$. The polynomial $p(x)$ has r distinct nonzero roots by (2) and simplicity, so the same can be said of q, whose roots are the reciprocals of the roots of p. □

Lemma 4. *Let the sequence $\mathcal{C} = (c_n)_{n\geq 0}$ satisfy two rth order linear recurrences, namely the minimal order recurrence $\sum_{n=0}^{r} a_n c_{n+t} = 0$ and another rth order recurrence $\sum_{n=0}^{r} b_n c_{n+t} = 0$, for all $t \geq 0$. Then $(a_1, a_2, \ldots, a_r) = \lambda(b_1, b_2, \ldots, b_r)$, where $\lambda \neq 0$ is a scalar.*

Proof. Let $a(x)$ be the characteristic polynomial of the recurrence $\sum_{n=0}^{r} a_n c_{n+t} = 0$. Since r is minimal, we have that $a(x)$ generates $R_{\mathcal{C}}$. If $b(x)$ is the characteristic polynomial of the recurrence $\sum_{n=0}^{r} b_n c_{n+t} = 0$, it must be that $a(x) = \lambda b(x)$ for some $\lambda \in \mathbb{C} \setminus \{0\}$ since $R_{\mathcal{C}}$ is principal and $\deg(a) = \deg(b)$, from which the result follows. □

We have already seen that the roots of characteristic polynomials associated with simple linear recurrences are distinct. The discriminant is a polynomial in the coefficients of univariate complex polynomials, whose kernel is exactly the set of polynomials with a repeated root. This kernel is the "discriminant variety."

Definition 5. Fix the natural number $r \geq 1$. Then the (affine) *r-discriminant variety*, denoted ∇_r, is the closure of

$$\left\{(b_0, \ldots, b_r) \in \mathbb{C}^{r+1} : f(x) = \sum_{i=0}^{r} b_i x^i \text{ has a repeated root}\right\}.$$

It is also common to consider a Hankel matrix whose entries are given by the elements of a sequence. Both finite and infinite dimensional square Hankel matrices are considered throughout the paper. In [2], the authors show that all infinite Hankel matrices have generalized Vandermonde decompositions of a specified form, dependent on the recurrences the original sequence satisfies. Our investigation into simple linear recurrences invites the question of which additional matrix properties are satisfied by Hankel matrices generated by sequences satisfying simple linear recurrences. The following lemma and subsequent definition provide tools necessary to analyze the structure of these matrices.

Lemma 5. *Let V be an $r \times n$ Vandermonde matrix where the (i,j) entry is a_i^{j-1}, and let D be an $r \times r$ diagonal matrix with (i,i) entry b_i. Take a_i and b_i for $1 \leq i \leq r$ to be complex scalars. Then the matrix $V^T DV$ is a Hankel matrix.*

Proof. We simply perform the matrix multiplication, showing the form of each product along the way. Let the matrices $D_{r\times r}$ and $V_{n\times r}$ be defined as follows, where $\{a_i\}_{i=1}^r$ and $\{b_i\}_{i=1}^r$ are complex scalars. Let $D_{r\times r} = \operatorname{diag}(b_i)$ and

$$V = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_r & a_r^2 & \cdots & a_r^{n-1} \end{pmatrix}.$$

We see that

$$V^T D = \begin{pmatrix} b_1 & b_2 & b_3 & \cdots & b_r \\ b_1a_1 & b_2a_2 & b_3a_3 & \cdots & b_ra_r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b_1a_1^{n-1} & b_2a_2^{n-1} & b_3a_3^{n-1} & \cdots & b_ra_r^{n-1} \end{pmatrix},$$

and furthermore that

$$V^T DV = \begin{pmatrix} \sum_{i=1}^r b_i & \sum_{i=1}^r b_ia_i & \sum_{i=1}^r b_ia_i^2 & \cdots & \sum_{i=1}^r b_ia_i^{n-1} \\ \sum_{i=1}^r b_ia_i & \sum_{i=1}^r b_ia_i^2 & \sum_{i=1}^r b_ia_i^3 & \cdots & \sum_{i=1}^r b_ia_i^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^r b_ia_i^{n-1} & \sum_{i=1}^r b_ia_i^n & \sum_{i=1}^r b_ia_i^{n+1} & \cdots & \sum_{i=1}^r b_ia_i^{2n-2} \end{pmatrix}.$$

Then $c_j = \sum_{i=1}^r b_ia_i^j$ for $0 \le j \le 2n-2$ is the sequence which populates the Hankel matrix V^TDV. □

We are specifically interested in Vandermonde matrices with no zero entries, motivating the following definition.

Definition 6. Let $H_{n\times n}$, where n is allowed to be ∞, be a complex matrix. We say H has a *nondegenerate Vandermonde decomposition* if there exists a Vandermonde matrix $V_{r\times n}$ with all entries nonzero and a diagonal matrix $D_{r\times r}$ so that $H_\infty = V^TDV$.

3 Moment rank

Finally, before presenting the main theorem of this section, we describe an algorithm which returns the moment rank of a sequence and the coefficients of a linear combination of powers witnessing to this rank. Denote the $r\times r$ (modified) Vandermonde matrix with (i,j) entry β_i^j for $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_r)$ by $\mathrm{VDM}'(\boldsymbol{\beta})$, and the $r\times r$ (ordinary) Vandermonde matrix with (i,j) entry β_i^{j-1} by $\mathrm{VDM}(\boldsymbol{\beta})$.

Algorithm MRANK($\mathcal{C}$). Given the sequence $\mathcal{C} = (c_n)_{n\geq 0}$, set $r = 0$. Then:

1. $r \leftarrow r + 1$
2. Let $\mathbf{c}_t = (c_t, \ldots, c_{t+r-1})$, and $\mathbf{C} = (\mathbf{c}_0^T, \ldots, \mathbf{c}_{r-1}^T)$.
3. If $\det(\mathbf{C}) = 0$, `goto` step 1. Else `continue`.
4. Let $(a_0, \ldots, a_{r-1})^T = -\mathbf{C}^{-1}\mathbf{c}_r^T$, and define $p(x) = \sum_{n=0}^{r} a_n x^n$, where $a_r = 1$.
5. If $(c_n)_{n\geq 0}$ does not satisfy the recurrence $\sum_{i=0}^{r} a_i c_{i+t} = 0$ for all $t \geq 0$, `goto` step 1. Else `continue`.
6. If p has repeated roots, `throw ErrorNotSimple` and `terminate`. Else `continue`.
7. `return` r and $\boldsymbol{\alpha} = (\mathbf{c}_0 \operatorname{VDM}'(\boldsymbol{\beta})^{-1})^T$, where $\boldsymbol{\beta}$ is the vector of roots of $p(x)$.

Note that the above is in truth only a *template* for an actual executable algorithm, since some steps involve unspecified subroutines, such as computation of the determinant in Step 3, or checking for repeated roots in Step 6. Indeed, Step 5 involves checking whether a sequence is satisfied by a given recurrence, a task which could range from very straightforward (e. g., if the sequence was given as the solution to a linear recurrence) to undecidable (e. g., if the sequence is not a computable function of its index). We therefore make no attempt to analyze the complexity of MRANK($\cdot$) and instead treat constitutive subproblems as black boxes. However, we do assume that each step is indeed computable in the sense that there exists an algorithm which will, in finite time, return `True` or `False` correctly.

The following is our main theorem concerning sequences of finite moment rank. We show that all of the above contexts provide interpretations of the moment rank. Recall that an S-measure on a space $(\mathcal{X}, \Sigma)$, where Σ is a σ-algebra on $\mathcal{X}$, is a countably additive function from Σ to S, where S is an additive monoid with limits such as $[0, \infty)$ (positive measure), $\mathbb{R}$ (signed measure), or $\mathbb{C}$ (complex measure). The tth moment m_t of an S-measure $d\mu$ on $\mathcal{X}$ is the quantity $\int_{\mathcal{X}} x^t \, d\mu$, and the sequence $\{m_t\}_{t=0}^{\infty}$ is its "moment sequence." We call a measure "r-atomic" if there exists an $A \subset \mathcal{X}$ with $|A| = r$ so that $\mu(x) \neq 0$ for $x \in A$, and $\mu(B) = 0$ for any $B \subset \mathcal{X} \setminus A$ in Σ.

Theorem 1. *Suppose $\mathcal{C} = (c_n)_{n=0}^{\infty}$ is a sequence in $\mathbb{C}$, and $r \in \mathbb{N}$. Let $H_{m,t}$ denote the $(m+1) \times (m+1)$ Hankel matrix whose entries come from the sequence $(c_n)_{n=t}^{2m+t}$, and let $f = \sum_{n\geq 0} c_n z^n$ denote the ordinary generating function of $\mathcal{C}$. Then the following are equivalent:*

1. *The sequence $\mathcal{C}$ has moment rank r.*
2. *$\mathcal{C}$ satisfies a simple rth order linear recurrence, and r is the smallest positive integer so that $\mathcal{C}$ has this property.*
3. *The matrices $H_{m,t}$ satisfy $\det(H_{r-1,t}) \neq 0$, $\operatorname{null}(H_{r,0}) = 1$, $\ker(H_{r,t}) = \ker(H_{r,0})$ for every $t \geq 0$, and $\ker(H_{r,0}) \not\subseteq \nabla_r$.*
4. *There exist $\{\alpha_1, \ldots, \alpha_r\}, \{\beta_1, \ldots, \beta_r\}, \{\lambda_1, \ldots, \lambda_r\} \subset \mathbb{C} \setminus \{0\}$ so that, for each $t \geq 0$, the polynomial $\sum_{j=0}^{2r} \binom{2r}{j} c_{j+t} x^{2r-j} y^j = \sum_{j=1}^{r} \lambda_j (\beta_j/\alpha_j)^t (\alpha_j x + \beta_j y)^{2r}$ and $\{\alpha_j/\beta_j\}_{j=1}^{r}$ is a set of r distinct values.*

5. *The ordinary generating function* $\Phi(z) = \sum_{n\geq 0} c_n z^n$ *of* $\mathcal{C}$ *is a rational function with exactly r simple poles.*
6. *The infinite Hankel matrix* H_∞ *has rank r and admits a nondegenerate Vandermonde decomposition.*
7. *The ideal* $R_\mathcal{C}$ *is radical, and the least degree of any nonzero element is r.*
8. *The sequence* $\mathcal{C}$ *is the moment sequence for a complex r-atomic measure on* $\mathbb{C}$.
9. *The algorithm* MRANK($\mathcal{C}$) *returns the parameter r.*

Proof.

$2 \Leftrightarrow 7$: Suppose 7. Let $p(x)$ be a monic generator of $R_\mathcal{C}$, which exists because $\mathbb{C}[x]$ is a PID. We know that p is the characteristic polynomial for a linear recurrence of minimal order satisfied by $\mathcal{C}$. Note that $\deg(p) = r$. Let $\{\beta_i\}_{i=1}^s$ be the distinct roots of p, and let m be the maximum multiplicity of a root of p. Consider the polynomial $q(x) = \prod_{i=1}^s (x-\beta_i)^m$. Clearly, p divides q, giving that $q(x) \in R_\mathcal{C}$. Since the ideal is radical, we have that $\sqrt[m]{q(x)} = \prod_{i=1}^s (x-\beta_i)$ is an element of $R_\mathcal{C}$. Thus $\deg(\sqrt[m]{q(x)}) \leq \deg(p(x))$, and $p(x)$ generating the recurrence ideal implies $\deg(\sqrt[m]{q(x)}) = \deg(p(x))$, so $p(x)$ has distinct roots.

We prove the reverse direction by contraposition. Let $p(x)$ be the generator of $R_\mathcal{C}$ and suppose $R_\mathcal{C} = \langle p(x)\rangle$ is not radical. Let $f(x) \in R_\mathcal{C}$ and $m \geq 2$ be an integer so that $\sqrt[m]{f(x)}$ is an element of $\mathbb{C}[x] \setminus R_\mathcal{C}$. Since $R_\mathcal{C}$ is principal, we have that p divides f, but p does not divide $\sqrt[m]{f}$. Since the distinct roots of f and $\sqrt[m]{f}$ are the same, we have that p does not have distinct roots. Since p does not have distinct roots, the same can be said of every polynomial in $R_\mathcal{C}$ and so $\mathcal{C}$ does not satisfy a simple linear recurrence.

$2 \Rightarrow 5 \Rightarrow 4$: We adapt an argument from [23], ultimately drawing upon Sylvester's legendary manuscript [27]. Suppose (2), so for $\mathcal{C}$ we have $\sum_{n=0}^r a_n c_{n+t} = 0$ for $t \geq 0$, where r is the smallest order simple recurrence the sequence satisfies. By Lemma (3), the polynomial $g(x) = \sum_{n=0}^r a_n x^{r-n}$ has no repeated roots, and $a_r \neq 0$. Define $h(x,y) = \sum_{n=0}^r a_n x^{r-n} y^n$. Without loss of generality, let $a_r = 1$. Let α_n and β_n be complex numbers for $1 \leq n \leq r$ so that $h(x,y) = \prod_{n=1}^r (-\beta_n x + \alpha_n y)$. Note that the $\frac{\alpha_n}{\beta_n}$ are distinct since $h(x,1) = g(x)$ has distinct roots.

Let $\Phi(T) = \sum_{m=0}^\infty c_m T^m$. Then we have the following, which converges within a positive-radius disk about zero:

$$\left(\sum_{n=0}^r a_{r-n} T^n\right)\Phi(T) = \sum_{j=0}^{r-1}\sum_{k=0}^{j} a_{r-(j-k)} c_k T^j + \sum_{j=r}^{\infty}\sum_{k=0}^{r} a_{r-k} c_{j-k} T^j.$$

In the second term above, we have $j - r \geq 0$. Therefore, $\sum_{k=0}^r a_{r-k} c_{j-k} = \sum_{n=0}^r a_n c_{n+(j-r)}$ and the second term vanishes, leaving

$$\left(\sum_{n=0}^r a_{r-n} T^n\right)\Phi(T) = \sum_{j=0}^{r-1}\sum_{k=0}^{j} a_{r-(j-k)} c_k T^j.$$

Thus $\Phi(T)$ is a rational function with denominator $\sum_{n=0}^{r} a_{r-n}T^n = \sum_{n=0}^{r} a_n T^{r-n} = h(T,1) = \prod_{n=1}^{r}(\alpha_n - \beta_n T)$. Since the $\frac{\alpha_n}{\beta_n}$ are distinct, this completes the proof of (5). Continuing from here, by partial fractions, since the $\frac{\alpha_n}{\beta_n}$ are distinct there exist λ_n for $1 \le n \le r$ so that (choosing numerators with foresight)

$$\Phi(T) = \sum_{n=1}^{r} \frac{\lambda_n \alpha_n^{2r+1}}{\alpha_n - \beta_n T} \Rightarrow c_m = \sum_{n=1}^{r} \lambda_n \alpha_n^{2r} \left(\frac{\beta_n}{\alpha_n}\right)^m.$$

The following computation completes the proof of (4), noting that the minimality of r in this setting is due to the construction in $4 \Rightarrow 2$ (see below):

$$\begin{aligned}\sum_{j=0}^{2r} \binom{2r}{j} c_{j+t} x^{2r-j} y^j &= \sum_{n=1}^{r} \lambda_n \alpha_n^{2r} \left(\frac{\beta_n}{\alpha_n}\right)^t \sum_{j=0}^{2r} \binom{2r}{j} \left(\frac{\beta_n}{\alpha_n}\right)^j x^{2r-j} y^j \\ &= \sum_{n=1}^{r} \lambda_n (\beta_n/\alpha_n)^t (\alpha_n x + \beta_n y)^{2r}.\end{aligned}$$

$4 \Rightarrow 2$: Suppose (4), giving that there exist nonzero $\{\alpha_1, \ldots, \alpha_r\}$, $\{\beta_1, \ldots, \beta_r\}$, and $\{\lambda_1, \ldots, \lambda_r\}$ so that, for each $t \ge 0$, the polynomial $\sum_{j=0}^{2r} \binom{2r}{j} c_{j+t} x^{2r-j} y^j = \sum_{j=1}^{r} \lambda_j (\beta_j/\alpha_j)^t (\alpha_j x + \beta_j y)^{2r}$, the set $\{\alpha_j/\beta_j\}_{j=1}^{r}$ consists of r distinct values, and r is the smallest positive integer for which this property holds. Then for $0 \le j \le 2r$ we have

$$c_{j+t} = \sum_{j=1}^{r} \lambda_j (\beta_j/\alpha_j)^t (\alpha_j^{2r-j} \beta_j^j) = \sum_{j=1}^{r} \lambda_j \alpha_j^{2r-j-t} \beta_j^{j+t}.$$

Let $h(x,y) = \prod_{n=1}^{r}(-\beta_n x + \alpha_n y)$. Moreover, let a_n for $0 \le n \le r$ so that $h(x,y) = \sum_{n=0}^{r} a_n x^{r-n} y^n$. Note that $a_r = \prod_{n=1}^{r} \alpha_n$. Since $\alpha_n \ne 0$ for all $1 \le n \le r$, we have that $a_r \ne 0$. Continuing, we have

$$\begin{aligned}\sum_{n=0}^{r} a_n c_{n+t} &= \sum_{j=1}^{r} \sum_{n=0}^{r} a_n \lambda_j \alpha_j^{2r-n-t} \beta_j^{n+t} \\ &= \sum_{j=1}^{r} \lambda_j \alpha_j^{r-t} \beta_j^t \sum_{n=0}^{r} a_n \alpha_j^{r-n} \beta_j^n \\ &= \sum_{j=1}^{r} \lambda_j \alpha_j^{r-t} \beta_j^t h(\alpha_j, \beta_j) = 0.\end{aligned}$$

Therefore, the sequence satisfies an rth order linear recurrence. Notice that the recurrence is simple since $h(x,1) = \prod_{n=1}^{r}(-\beta_n x + \alpha_n)$ has distinct, nonzero roots. Moreover, the minimality of r in this setting is due to the construction in the preceding argument.

$5 \Rightarrow 7$: Suppose (5). Let $\Phi(z) = \sum_{n=0}^{\infty} c_n z^n$. Let the r simple poles of $\Phi(z)$ be γ_n for $1 \le n \le r$ so that $g(z) = \prod_{n=1}^{r}(z - \gamma_n)$ is the denominator of $\Phi(z)$. It is a well-known

result [11] that the polynomial $f(z) = \sum_{i=0}^{m} a_i z^i$ is the characteristic polynomial of a recurrence satisfied by $\mathcal{C}$ if and only if $f(z)\Phi(z)$ is a polynomial. Thus $f(z) \in R_{\mathcal{C}}$ if and only if $g(z)$ divides $f(z)$, giving that $g(z)$ (a polynomial of degree r) is a generator for $R_{\mathcal{C}}$. The distinctness of the roots of $g(z)$ implies $R_{\mathcal{C}}$ is radical.

$1 \Leftrightarrow 2$: Suppose (1), meaning there exists a set of nonzero complex numbers $\{\alpha_i\}_{i=1}^{r}$ and distinct nonzero $\{\beta_i\}_{i=1}^{r}$ so that $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ for all $n \geq 0$. We multiply both sides of the equation by z^n and take the sum over n to examine the ordinary generating functions of both sides. We have the following:

$$\Phi(z) = \sum_{n=0}^{\infty} \sum_{i=1}^{r} \alpha_i \beta_i^{n+1} z^n = \sum_{i=1}^{r} \sum_{n=0}^{\infty} \alpha_i \beta_i^{n+1} z^n = \sum_{i=1}^{r} \frac{\alpha_i \beta_i}{1 - \beta_i z}.$$

Thus, if we let $g(z) = \prod_{i=1}^{r}(1 - \beta_i z) = \sum_{i=0}^{r} a_i z^i$ for some coefficients a_i, then $\Phi(z)g(z) = \sum_{i=0}^{r-1} \lambda_i z^i$ is a polynomial of degree $r-1$ where $\lambda_{r-1} \neq 0$, i. e.,

$$\begin{aligned} \sum_{i=0}^{r-1} \lambda_i z^i &= \sum_{n=0}^{\infty} c_n z^n g(z) = \sum_{n=0}^{\infty} c_n z^n \sum_{i=0}^{r} a_i z^i \\ &= \sum_{n=r}^{\infty} z^n \sum_{i=0}^{r} a_i c_{n-i} + \sum_{n=0}^{r-1} z^n \sum_{i=0}^{n} a_i c_{n-i}. \end{aligned}$$

By matching coefficients of z^n on each side, we see, for $n \geq r$,

$$0 = \sum_{i=0}^{r} a_i c_{n-i}.$$

But, $a_0 = 1$, so

$$c_n = -\sum_{i=1}^{r} a_i c_{n-i}, \tag{9.3}$$

i. e., if $\boldsymbol{\alpha} = (a_r, a_{r-1}, \ldots, a_1)$, we have $\mathbf{C}\boldsymbol{\alpha}^T = -\mathbf{c}_r^T$, where we define $\mathbf{c}_t := (c_t, \ldots, c_{t+r-1})$ and $\mathbf{C} = (\mathbf{c}_0, \ldots, \mathbf{c}_{r-1})^T$. So $\boldsymbol{\alpha}^T = -\mathbf{C}^{-1}\mathbf{c}_r^T$, and $q(x) = \sum_{i=0}^{r} a_i x^{r-i}$ is the characteristic polynomial of the recurrence in equation (9.3). Since the recurrence is solved by $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$, the roots of $p(x)$ are the distinct nonzero values $\{\beta_i\}_{i=1}^{r}$. Moreover, since each $\beta_i \neq 0$, we have that $a_r \neq 0$. Thus $\mathcal{C}$ satisfies a simple rth order linear recurrence, and the minimality of r in this setting is given by the construction in the other direction of the proof (below) together with Lemma 1.

Suppose (2), giving that $\sum_{n=0}^{r} a_n c_{n+t} = 0$ for all $t \geq 0$, and r is the smallest order recurrence the sequence satisfies. By Lemma 3, let $\{\beta_i\}_{i=1}^{r}$ be the distinct nonzero roots of $p(x)$, the characteristic polynomial of the recurrence. It follows from standard facts about rational functions that $p(x)$ having distinct roots implies $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ for all $n \geq 0$ where each $\alpha_i \in \mathbb{C}$. If any of the α_i are zero,

the construction in the other direction of the proof would contradict the minimality of r in this direction. Therefore, all α_i are nonzero and $\mathcal{C}$ has moment rank r.

$3 \Leftrightarrow 2$: Suppose (2), giving that $\sum_{n=0}^{r} a_n c_{n+t} = 0$ for all $t \geq 0$ with $a_r = 1$, and r is the smallest order recurrence the sequence satisfies. The rth order linear recurrence implies $H_{r,t}A = 0$, where $A = [a_0, a_1, \ldots, a_r]^T$. Let M be the $(r+1) \times (r+1)$ matrix given by $M_{i,j} = 1$ if $j = i+1$, $M_{r+1,j} = -a_{j-2} - a_{j-1}$ for $1 \leq j \leq r+1$ where we define $a_{-1} = 0$, and $M_{i,j} = 0$ otherwise. Then M is invertible, because $M_{r+1,1} \neq 0$, the $(r+1,1)$-minor of M is 1, and $M(c_t, \ldots, c_{r+t})^T = (c_{t+1}, \ldots, c_{r+t+1})^T$ because M acts as a left-shift on the first r coordinates, and the $r+1$-st coordinate is given by

$$\begin{aligned}\sum_{j=1}^{r+1} -(a_{j-1} + a_{j-2})c_{j+t-1} &= \sum_{j=1}^{r+1} -a_{j-2}c_{j+t-1} - \sum_{j=1}^{r+1} a_{j-1}c_{j+t-1} \\ &= a_r c_{r+t+1} - \sum_{n=0}^{r} a_n c_{n+t+1} - \sum_{n=0}^{r} a_n c_{n+t} \\ &= c_{r+t+1}.\end{aligned}$$

Therefore, $H_{r,t} = M^k H_{r,t-k}$ for any integers $t \geq \max\{0, k\}$, so $\ker(H_{r,t}) = \ker(H_{r,t'})$ for every $t, t' \geq 0$. Suppose that there exists a nontrivial vector $B = [b_0, b_1, \ldots, b_r]^T$ contained within $\ker(H_{r,t})$. Then $\sum_{n=0}^{r} b_n c_{n+t} = 0$ for all $t \geq 0$, and Lemma 4, gives that B is a scalar multiple of A. Thus $\text{null}(H_{r,t}) = 1$. The polynomial $p(x) = \sum_{i=0}^{r} a_i x^r$ has no repeated roots, implying that $\ker(H_{r,t}) \notin \nabla_r$.

Finally, the minimality of r gives that $\text{null}(H_{m,t}) = 0$ for all $1 \leq m \leq r-1$, implying that $\det(H_{r-1,t}) \neq 0$ for all $t \geq 0$.

Suppose (3). There exists $A = [a_0, a_1, \ldots, a_r]^T$ so that $A \in \ker(H_{r,t})$ for all $t \geq 0$. Moreover, $(a_0, \ldots, a_r) \notin \nabla_r$, so the polynomial $a(x) = \sum_{i=0}^{r} a_i x^i$ has distinct complex roots. Let the first row of $H_{r,t} = \mathbf{c}_t = [c_t, \ldots, c_{t+r}]$. Then $\mathbf{c}_t \cdot A = 0$, giving that $0 = \sum_{i=0}^{r} a_i c_{i+t}$ for all $t \geq 0$, and the original sequence satisfies an rth order linear recurrence. The minimality of r is given by $\det(H_{r-1,t}) \neq 0$ for all $t \geq 0$, completing the proof.

$1 \Leftrightarrow 6$: Suppose (6). Let H_∞ (with entries from the sequence $\mathcal{C}$) have rank r and admit a nondegenerate Vandermonde decomposition. Let $\{a_i\}_{i=1}^{r}$ and $\{b_i\}_{i=1}^{r}$ be sets of complex scalars so that $D_{r\times r} = \text{diag}(b_i)$,

$$V_{r\times\infty} = \begin{pmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} & \cdots \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \cdots \\ 1 & a_r & a_r^2 & \cdots & a_r^{n-1} & \cdots \end{pmatrix},$$

and $H_\infty = V^T D V$. Since $\text{rank}(H_\infty) = r$, $\text{rank}(V) \leq r$, and $\text{rank}(D) \leq r$, we have that $\text{rank}(V) = \text{rank}(D) = r$, and furthermore that $b_i \neq 0$ for $1 \leq i \leq r$. By the

computation given in Lemma 5, we see that

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} & \cdots \\ c_1 & c_2 & \cdots & c_n & \cdots \\ \vdots & \vdots & \ddots & \vdots & \cdots \\ c_{n-1} & c_{n+1} & \cdots & c_{2n-2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$= H_\infty = V^T D V = \begin{pmatrix} \sum_{i=1}^r b_i & \sum_{i=1}^r b_i a_i & \cdots & \sum_{i=1}^r b_i a_i^{n-1} & \cdots \\ \sum_{i=1}^r b_i a_i & \sum_{i=1}^r b_i a_i^2 & \cdots & \sum_{i=1}^r b_i a_i^n & \cdots \\ \vdots & \vdots & \ddots & \vdots & \cdots \\ \sum_{i=1}^r b_i a_i^{n-1} & \sum_{i=1}^r b_i a_i^n & \cdots & \sum_{i=1}^r b_i a_i^{2n-2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Thus equating the entries of the two representations of H_∞ yields $c_n = \sum_{i=1}^r b_i a_i^n$ for all $n \geq 0$. Using the transformation $\beta_i = a_i$ and $\frac{b_i}{\beta_i} = \alpha_i$ for $1 \leq i \leq r$, we obtain the desired form. Note that this transformation is well-defined since $a_i \neq 0$ is a consequence of the definition of a nondegenerate Vandermonde decomposition. Moreover, all $a_i = \beta_i$ are distinct, since $a_i = a_j$ for $i \neq j$ implies the ith and jth rows of V are equal, contradicting $\operatorname{rank}(V) = r$.

Suppose (1) by letting $\mathcal{C}$ have moment r. Then there exists a set of nonzero complex numbers $\{\alpha_i\}_{i=1}^r$ and distinct nonzero $\{\beta_i\}_{i=1}^r$ so $c_n = \sum_{i=1}^r \alpha_i \beta_i^{n+1}$ for all n. We see from the computation in Lemma 5 that $D = \operatorname{diag}(\alpha_i \beta_i)$ and the (i, j) entry of $V_{r\times\infty}$ given by β_i^{j-1} is a nondegenerate Vandermonde decomposition of H_∞.

It only remains to show that H_∞ has rank r. It is immediately clear that $\operatorname{rank}(H_\infty) \leq r$. Moreover, Frobenius's inequality [15, 0.4.5(e)] implies that for the product $V^T D V$,

$$\operatorname{rank}(V^T D) + \operatorname{rank}(DV) \leq \operatorname{rank}(D) + \operatorname{rank}(V^T D V).$$

We show now that $\operatorname{rank}(V^T D) = r$. Suppose that $\operatorname{rank}(V^T D) < r$. Then, not all columns of $V^T D$ are linearly independent, and there exists $\{\gamma_i\}_{i=1}^{r-1} \subset \mathbb{C}$ so that $\sum_{i=1}^{r-1} \gamma_i C_i = C_r$, where C_i denotes the ith column of $V^T D$. By examining the entries within the columns of $V^T D$, we see that $\sum_{i=1}^{r-1} \gamma_i C_i = C_r$ implies $\sum_{i=1}^{r-1} \gamma_i \alpha_i \beta_i^{n+1} = \alpha_r \beta_r^{n+1}$ for all $n \geq 0$, leading to the following representation of c_n:

$$c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1} = \sum_{i=1}^{r-1} (1 + \gamma_i) \alpha_i \beta_i^{n+1}$$

This implies $\mathcal{C}$ has description complexity at most $r - 1$, contradicting Lemma 1. Therefore, $\operatorname{rank}(V^T D) = r = \operatorname{rank}(DV)$, and Frobenius's inequality produces the desired result. (It is also possible to argue this via Sylvester's law of inertia applied to the leading principal $r \times r$ submatrices of H_∞, V, and D.)

$1 \Leftrightarrow 8$: It is clear that $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ is the $(n+1)$-st moment of the r-atomic complex measure $\mu = \sum_{i=1}^{r} \alpha_i \delta_{\beta_i}$, where δ_{β_i} denotes the standard Dirac delta function.

$1 \Leftrightarrow 9$: We argue that the algorithm returns the parameter r if and only if r is the smallest positive integer so that every term of the sequence (c_n) can be expressed as $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ for a set of nonzero complex α_i and distinct β_i (since $\alpha_i = 0$ or $\beta_i = \beta_j$ for $i \neq j$ contradict the minimality of r).

Suppose the algorithm succeeds. Then $(c_n)_{n\geq 0}$ satisfies $\sum_{n=0}^{r} a_n c_{n+t} = 0$, where the $\{a_n\}$ are defined by step 4; this is well-defined because step 3 says that $\det(\mathbf{C}) \neq 0$. The characteristic polynomial of the recurrence $p(x)$ is well-defined and factors into r distinct linear factors $x - \beta_i$ by step 6. Thus, by standard results in the theory of linear recurrence relations,

$$c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$$

for some $\{\alpha_i\}_{i=1}^{r} \subset \mathbb{C}$, which are nonzero by the minimality of r (since $\textsc{mrank}(\mathcal{C})$ did not terminate on any $r' < r$). So, $\operatorname{mrank}(\mathcal{C}) = r$.

Now, suppose that $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ with all distinct nonzero β_i and nonzero values α_i. By $1 \Rightarrow 3$ above, the matrix $\mathbf{C} = H_{r-1,0}$ introduced in step 2 is invertible, so we pass step 3. Write $\Phi(z) = \sum_{n=0}^{\infty} c_n z^n$. Because $\max_i |\beta_i| < \infty$, the following is true in a sufficiently small ball about $z = 0$:

$$\begin{aligned}\Phi(z) &= \sum_{n=0}^{\infty} \sum_{i=1}^{r} z^n \alpha_i \beta_i^{n+1} \\ &= \sum_{i=1}^{r} \frac{\alpha_i \beta_i}{1 - \beta_i z}.\end{aligned}$$

Thus, if we let $g(z) = \prod_{i=1}^{r}(1 - \beta_i z) =: \sum_{i=0}^{r} a_{r-i} z^i$, then $\Phi(z)g(z)$ is a polynomial of degree $r-1$, i. e.,

$$\begin{aligned}\Phi(z)g(z) &= \sum_{n=0}^{\infty} c_n z^n g(z) = \sum_{n=0}^{\infty} c_n z^n \sum_{i=0}^{r} a_{r-i} z^i \\ &= \sum_{n=r}^{\infty} z^n \sum_{i=0}^{r} a_{r-i} c_{n-i} + \sum_{n=0}^{r-1} z^n \sum_{i=0}^{n} a_{r-i} c_{n-i} \\ &= \sum_{n=r}^{\infty} z^n \sum_{i=0}^{r} a_i c_{i+n-r} + \sum_{n=0}^{r-1} z^n \sum_{i=0}^{n} a_{i+r-n} c_i.\end{aligned}$$

By matching coefficients of z^n on each side, we see, for $n \geq r$, $\sum_{i=0}^{r} a_i c_{i+n-r} = 0$, so we pass step 5. But, $a_r = 1$, so setting $n = r$,

$$c_r = -\sum_{i=0}^{r-1} a_i c_i, \tag{9.4}$$

i. e.,

$$\mathbf{C}\begin{bmatrix} a_0 \\ \vdots \\ a_{r-1} \end{bmatrix} = -\mathbf{c}_r^T$$

so

$$\begin{bmatrix} a_0 \\ \vdots \\ a_{r-1} \end{bmatrix} = -\mathbf{C}^{-1}\mathbf{c}_r^T.$$

Thus $p(x) = \sum_{i=0}^{r} a_i x^i$ is the characteristic polynomial of the recurrence in equation (9.4). Since the recurrence is solved by $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$, the roots of $p(x)$ are the distinct values $\{\beta_i\}_{i=1}^{r}$, and we pass step 6. Since Lemma 1 implies that the algorithm does not succeed for any $r' < r$, the result follows. □

In condition (4), the quantity r is known as the "Waring rank" of this polynomial. Pratt [22] presents a history and many interesting results about this classical invariant.

Note that, when mrank C is finite, MRANK($\mathcal{C}$) returns r (the order of the recurrence satisfied by $\mathcal{C}$) and a vector $\boldsymbol{\alpha} = (\mathbf{c}_0\,\mathrm{VDM}(\boldsymbol{\beta})^{-1})^T$. Since $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ for some nonzero values $\{\alpha_i\}_{i=1}^{r}$, we may write

$$\mathrm{VDM}(\boldsymbol{\beta})^T \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_r \\ \beta_1^2 & \beta_2^2 & \cdots & \beta_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_1^r & \beta_2^r & \cdots & \beta_r^r \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_r \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{r-1} \end{bmatrix} = \mathbf{c}_0^T.$$

Since the $\{\beta_i\}$ are distinct when r is minimal, $\mathrm{VDM}(\boldsymbol{\beta})$ is invertible, so the vector of α_i values equals $(\mathrm{VDM}(\boldsymbol{\beta})^{-1})^T \mathbf{c_0}$, i. e., $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r)$.

Corollary 2. *For a sequence $\mathcal{C} = (c_n)_{n=0}^{\infty}$ satisfying a simple rth order linear recurrence for r minimal,*

$$\mathrm{null}\, H_{m,t} = \max\{0, m - r + 1\}.$$

Proof. The rank of $H_{m,t}$ is equal to the order r of the smallest linear recurrence satisfied by $\mathcal{C}$ for $m \geq r$ and is $m + 1$ for $m < r$, since $H_{r,t}$ is invertible. Therefore, $\mathrm{null}(H_{m,t}) = \max\{0, m - r + 1\}$. □

4 Unitary rank

Given a sequence $\mathcal{C} = (c_n)_{n\geq 0}$ with moment rank r, we know that $c_n = \sum_{i=1}^{r} \alpha_i \beta_i^{n+1}$ for nonzero $\{\alpha_i\}$ and nonzero, distinct $\{\beta_i\}$. It will be useful for some applications to isolate and consider the case when $\alpha_i = 1$ for all i, meaning $c_n = \sum_{i=1}^{r} \beta_i^{n+1}$ for $n \geq 0$. We proceed with the following definition (changing the sequence index for ease of use later).

Definition 7. The sequence $\mathcal{C} = (c_n)_{n=1}^{N}$ (with $N = \infty$ allowed) is said to have *unitary rank* r if r is the smallest positive integer so that there exists a multiset of nonzero complex values $\{\beta_i\}_{i=1}^{r}$ so that $c_n = \sum_{i=1}^{r} \beta_i^n$ for all $1 \leq n \leq N$. If $N = \infty$, we write $\operatorname{urank}((c_n)_{n\geq 1})$ for the unitary rank.

Lemma 6. $\operatorname{urank}((c_n)_{n\geq 1})$ *is well-defined.*

Proof. Suppose, by way of contradiction, that there are two distinct multisets $\{\beta_i\}_{i=1}^{r}$ and $\{\gamma_j\}_{j=1}^{s}$ of nonzero complex numbers so that

$$c_n = \sum_{i=1}^{r} \beta_i^n = \sum_{j=1}^{s} \gamma_j^n.$$

Write $f(z) = \sum_{n=1}^{\infty} c_n z^n$. Because $\max_i |\beta_i|$ and $\max_j |\gamma_j|$ are finite, the following is true around a sufficiently small ball about $z = 0$:

$$\sum_{n=1}^{\infty} \sum_{i=1}^{r} z^n \beta_i^n = \sum_{n=1}^{\infty} \sum_{j=1}^{s} z^n \gamma_j^n$$
$$\sum_{i=1}^{r} \frac{z\beta_i}{1 - z\beta_i} = \sum_{j=1}^{s} \frac{z\gamma_j}{1 - z\gamma_j}.$$

These two functions are equal, so they have the same set of poles; thus, as sets, $\{\gamma_i\}_i = \{\beta_i\}_i$. Furthermore, since the residues of these poles are proportional to the multiplicity of the γ_i and β_i values, they also occur the same number of times. Thus, as multisets as well, $\{\gamma_i\}_i = \{\beta_i\}_i$. □

Theorem 2. *Let $\mathcal{C} = (c_n)_{n=1}^{\infty}$ be a sequence and r a positive integer. Let $H_{r-1,t}$ denote the $r \times r$ Hankel matrix whose entries come from the sequence $(c_n)_{n=t}^{2r-2+t}$. Lastly, let $\Phi(z) = \sum_{n\geq 1} c_n z^n$ denote the ordinary generating function of the sequence $\mathcal{C}$. Then the following are equivalent:*

1. *The sequence $\mathcal{C}$ has unitary rank r.*
2. *The algorithm MRANK($T[\mathcal{C}]$) succeeds on the sequence $T[\mathcal{C}] = (c_{n+1})_{n\geq 0}$ and returns r' and $\boldsymbol{\alpha} \in \mathbb{N}^{r'}$ so that $\boldsymbol{\alpha} \cdot \mathbf{1} = r$, where $\mathbf{1}$ is the all-ones vector of dimension r'.*
3. *$\mathcal{C}$ satisfies an r'th order simple linear recurrence for some $r' \leq r$ with coefficients $\{\alpha_i\}_{i=1}^{r'}$, $\alpha_i \in \mathbb{N}$ for all $1 \leq i \leq r'$, and $\sum_{i=1}^{r'} \alpha_i = r$.*

4. *Let $\mathcal{C}' = (c'_n)_{n\geq 0}$ be a sequence so that $c'_0 = r$ and $c'_n = c_n$ for $n \geq 1$. If $H'_{r-1,t}$ is the $r \times r$ Hankel matrix whose first row consists of $c'_t, \ldots, c'_{t+r-1}$, then there exists an $r \times r$ Vandermonde matrix $V = \mathrm{VDM}(\boldsymbol{\beta})$ so that $H'_{r-1,t} = V^T D^t V$ where $D = \mathrm{diag}(\boldsymbol{\beta})$, for each $t \geq 0$.*
5. *Let $\mathcal{C}' = (c'_n)_{n\geq 0}$ be a sequence so that $c'_0 = r$ and $c'_n = c_n$ for $n \geq 1$. If H'_∞ is the infinite Hankel matrix with entries from $\mathcal{C}'$, then there exists a Vandermonde matrix $V'_{r,\infty}$ so that $H'_\infty = (V')^T V'$.*
6. *$\exp(\int -x^{-1}\Phi(x)\,dx)$ is a polynomial of degree r with nonzero roots.*
7. *There exist nonzero $r' \in \mathbb{N}$ and $\{\alpha_1, \ldots, \alpha_{r'}\}$, $\{\beta_1, \ldots, \beta_{r'}\}$, and $\{\lambda_1, \ldots, \lambda_{r'}\}$ so that, for each $t \geq 0$, the polynomial $\sum_{j=0}^{2r'} \binom{2r'}{j} c_{j+t} x^{2r'-j} y^j = \sum_{j=1}^{r'} \lambda_j (\beta_j/\alpha_j)^t (\alpha_j x + \beta_j y)^{2r'}$, and the $\{\alpha_j\}_{j=1}^{r'}$ are positive naturals that sum to r.*
8. *The sequence $\mathcal{C}$ is the moment sequence for a complex finite-atomic measure on $\mathbb{C}$ where the masses of the atoms are positive naturals with sum r.*

Proof.

$1 \Leftrightarrow 3$: This is clear from the definitions of urank and mrank, taking into account the shift in index.

$1 \Leftrightarrow 2$: The sequence $\mathcal{C}$ having unitary rank r implies there exists $r' \leq r$ so that $\mathrm{mrank}(T[\mathcal{C}]) = r'$, since we may take the same β_i with α_i equal to the multiplicity of β_i in the representation $c_n = \sum_{i=1}^{r} \beta_i^n$. Thus the algorithm applied to $T[\mathcal{C}]$ with Hankel matrix of size r' succeeds and returns r' and $\boldsymbol{\alpha} = (\mathbf{c}_0\,\mathrm{VDM}(\boldsymbol{\beta})^{-1})^T$, which, by the note following the proof of Theorem 1, is the vector of coefficients α_i in the representation $c_n = \sum_{i=1}^{r'} \alpha_i \beta_i^n$. (Note that the exponent of β_i is n instead of $n+1$ because $T[\mathcal{C}]$ is the input to MRANK here.) But then, the sum $\sum_{i=1}^{r'} \alpha_i$ over the values generated by the algorithm is r, since this is the number of terms in the representation $c_n = \sum_{i=1}^{r} \beta_i^n$.

To establish the converse direction, suppose the algorithm returns r' and $\boldsymbol{\alpha} \in \mathbb{N}^{r'}$ with $\boldsymbol{\alpha} \cdot \mathbf{1} = r$. Note that none of the α_i are zero, since then MRANK would have terminated on a smaller value of r'. Thus the α_i are positive integers summing to r, so we may write

$$c_{n+1} = \sum_{j=1}^{r'} \sum_{i=1}^{\alpha_i} \beta_j^{n+1},$$

i. e., $c_n = \sum_{j=1}^{r} {\beta'_j}^n$ if $\{\beta_j\}_{j=1}^{r'} = \{\beta'_j\}_{j=1}^{r}$ with β_j appearing with multiplicity α_j on the right-hand side.

$3 \Leftrightarrow 8$: This is a direct consequence of Theorem 1.

$1 \Leftrightarrow 5$: Suppose condition (1) holds and let $\{\beta_i\}_{i=1}^{r}$ be given to satisfy the definition of unitary rank for the sequence $\mathcal{C}$. To adopt the notation of Lemma 5, letting $b_i = 1$ and $a_i = \beta_i$ for $1 \leq i \leq r$ proves (5).

Now, suppose (5) holds. Letting $b_i = 1$ for each $1 \le i \le r$ in the form of $V^T DV$ given in Lemma 5, we have that the (i,j) entry of H'_∞ is $\sum_{i=1}^{r} a_i^{i+j-2}$. Clearly, then, we have that $c'_n = \sum_{i=1}^{r} a_i^n$, further implying that $c_n = \sum_{i=1}^{r} a_i^{n+1}$. The uniqueness of unitary rank given by Lemma 6 completes the proof.

$1 \Leftrightarrow 6$: Suppose (1) holds. Let $c_n = \sum_{i=1}^{r} \beta_i^n$. Since $f(x) = \sum_{n\ge1} c_n x^n$, we have that $\int -x^{-1} f(x)\,dx = -\sum_{n\ge1} \frac{c_n x^n}{n} + C$. We have the following computation:

$$\begin{aligned}
\exp\Big(-\sum_{n\ge1} \frac{c_n x^n}{n}\Big) &= \exp\Big(-\sum_{n\ge1}\sum_{i=1}^{r} \frac{\beta_i^n x^n}{n}\Big) \\
&= \exp\Big(-\sum_{i=1}^{r}\sum_{n\ge1} \frac{(\beta_i x)^n}{n}\Big) \\
&= \prod_{i=1}^{r} \exp\Big(-\sum_{n\ge1} \frac{(\beta_i x)^n}{n}\Big) \\
&= \prod_{i=1}^{r} \exp(\log[1-\beta_i x]) \\
&= \prod_{i=1}^{r} (1-\beta_i x).
\end{aligned}$$

Therefore, we see that $\int -x^{-1} f(x)\,dx$ is the log of a polynomial of degree r, as desired.

Suppose (6) holds. Since it is only possible to take the log of a polynomial if the polynomial has nonzero constant term, we have that there exists nonzero $\{\beta_i\}_{i=1}^{r}$ so that $\exp(-\int x^{-1} f(x)\,dx) = \prod_{i=1}^{r}(1-\beta_i x)$. By letting $c_n = \sum_{i=1}^{r} \beta_i$ and working backwards through the computation given in the first direction of the proof shows

$$\exp\Big(-\int x^{-1} f(x)\,dx\Big) = e^C \exp\Big(-\sum_{n\ge1} \frac{c_n x^n}{n}\Big).$$

Taking log of both sides and differentiating gives that $f(x) = \sum_{n\ge1} c_n x^n$, showing that $f(x)$ is the generating function for a sequence with unitary rank r, completing the proof.

$1 \Leftrightarrow 7$: Suppose (1) holds. Having already proved $1 \Leftrightarrow 3$, we appeal to (3) and Theorem 1, to give that (despite the shift in index, which is taken care of by the variable t) there exist nonzero $\{\alpha_1, \ldots, \alpha_{r'}\}$, $\{\beta_1, \ldots, \beta_{r'}\}$, and $\{\lambda_1, \ldots, \lambda_{r'}\}$ so that, for each $t \ge 0$, the polynomial $\sum_{j=0}^{2r'} \binom{2r'}{j} c_{j+t} x^{2r'-j} y^j = \sum_{j=1}^{r'} \lambda_j (\beta_j/\alpha_j)^t (\alpha_j x + \beta_j y)^{2r'}$. The fact that $\sum_{i=1}^{r'} \alpha_i = r$ is given by the assumption of condition (1).

Now suppose (7) is true. By Theorem 1, we have that $\operatorname{mrank}(\mathcal{C}) = r'$. The assumption in (7) that $\{\alpha_i\}_{i=1}^{r'}$ is a set of positive naturals summing to r gives that $\operatorname{urank}(\mathcal{C}) = r$.

$1 \Leftrightarrow 4$: Suppose (4). By the computation given in Lemma 5, it is clear that the (i,j) entry of $V^T D^t V$ is given by $\sum_{i=1}^{r} \beta_i^{t+i+j-2}$. By letting $c_n = \sum_{i=1}^{r} \beta_i^n$, we have that $\mathcal{C} = (c_n)_{n=1}^{\infty}$ has unitary rank at most r. We note that Lemma 6 completes the proof that $\operatorname{urank}(\mathcal{C}) = r$.

Now, suppose (1). By the computation given in Lemma 5, we see that letting the ith row of V be generated by β_i and defining $D = \operatorname{diag}(\beta_i)$, the result follows immediately. □

Suppose A is an infinite positive-semidefinite matrix (aka positive-type kernel for ℓ^2); then A can be written as $A = M^*M$, which we will refer to as a "Gramian representation" (since M^*M is a Gramian matrix in the finite-dimensional case). In general, this representation is unique up to unitary conjugation.

Corollary 3. *The property that a real sequence $\mathcal{C}$ is the moment sequence of a finite-atomic measure on $\mathbb{C}$ with integer masses is equivalent to H'_∞ being positive semidefinite of finite rank with a Vandermonde Gramian representation.*

Proof. By Theorem 2, H'_∞ has a factorization as $V^T V$ for some Vandermonde kernel V. Note that, if $H'_\infty = M^*M$ for some M, then H'_∞ is Hermitian and symmetric, which implies that $\mathcal{C}$ is real. Thus $V^T = V^*$. □

Bibliography

[1] G. Blekherman, Nonnegative polynomials and sums of squares, Semidefinite optimization and convex algebraic geometry, *MOS-SIAM Ser. Optim.*, **13** (2013), 159–202.

[2] D. L. Boley, F. T. Luk and D. Vandervoorde, Vandermonde factorization of a Hankel matrix, Sci. Comput. (1997), 27–39.

[3] A. Bostan, A. Elvey-Price, A. Guttmann and J. Maillard, Stieltjes moment sequences for pattern-avoiding permutations, *Electron. J. Comb.*, **27**(4) (2020), Article #4.20, 59 pp.

[4] F. Brenti, Unimodal, log-concave and Pólya frequency sequences in combinatorics, *Mem. Am. Math. Soc.*, **81** (1989).

[5] M. T. Chu and M. M. Lin, On the finite rank and finite-dimensional representation of bounded semi-infinite Hankel operators, *IMA J. Numer. Anal.*, **35** (2015), 1256–1276.

[6] L. Collatz and U. Sinogowitz, Spektren endlicher Grafen, *Abh. Math. Semin. Univ. Hamb.*, **21** (1957), 63–77.

[7] R. E. Curto, L. A. Fialkow and H. M. Möller, The extremal truncated moment problem, *Integral Equ. Oper. Theory*, **60** (2008), 177–200.

[8] R. E. Curto and L. A. Fialkow, Solution of the truncated complex moment problem for flat data, *Mem. Am. Math. Soc.*, **119**(568) (1996).

[9] R. E. Curto and L. A. Fialkow, Flat extensions of positive moment matrices: recursively generated relations, *Mem. Am. Math. Soc.*, **136**(648) (1998).

[10] P. J. di Dio and K. Schmüdgen, The multidimensional truncated moment problem: atoms, determinacy, and core variety, *J. Funct. Anal.*, **274** (2018), 3124–3148.

[11] G. Everest, A. van der Poorten, I. Shparlinski and T. Ward, *Recurrence Sequences*, American Mathematical Society, Providence, RI, 2003.

[12] S. Fomin and A. Zelevinsky, Total positivity: tests and parametrizations, *Math. Intell.*, **22** (2000), 23–33.
[13] C. French, Transformations preserving the Hankel transform, *J. Integer Seq.*, **10** (2007), Article #07.7.3.
[14] F. Hausdorff, Momentprobleme für ein endliches Intervall, *Math. Z.*, **16** (1923), 220–248.
[15] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, Cambridge, 2013.
[16] A. Junod, Hankel determinants and orthogonal polynomials, *Expo. Math.*, **21** (2003), 63–74.
[17] J. B. Lasserre, *Moments, positive polynomials and their applications*, Imperial College Press Optimization Series, Imperial College Press, London, 2010.
[18] M. Laurent, Sums of squares, moment matrices and optimization over polynomials. Emerging applications of algebraic geometry, *IMA Vol. Math. Appl.*, **149** (2009), 157–270.
[19] J. Layman, The Hankel transform and some of its properties, *J. Integer Seq.*, **4** (2001), Article #01.1.5.
[20] P. Peart and W. Woan, Generating functions via Hankel and Stieltjes matrices, *J. Integer Seq.*, **3** (2000), Article #00.2.1.
[21] A. Pinkus, *Totally Positive Matrices*, Cambridge University Press, Cambridge, 2010.
[22] K. Pratt, Waring rank, parameterized and exact algorithms, in *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 806–823, 2019.
[23] B. Reznick, On the length of binary forms, in *Quadratic and higher degree forms*, vol. **31**, pp. 207–232, 2013.
[24] K. Schmüdgen, *The Moment Problem*, Springer, Cham, Switzerland, 2017.
[25] J. Shohat and J. Tamarkin, *The Problem of Moments*, American Mathematical Society, Providence, RI, 1943.
[26] B. Sturmfels, *Algorithms in invariant theory*, Springer, Wien, 2008.
[27] J. J. Sylvester, LX. On a remarkable discovery in the theory of canonical forms and of hyperdeterminants, *Philos. Mag.*, **2** (1851), 391–410.
[28] H. Widom, Hankel matrices, *Trans. Am. Math. Soc.*, **121** (1966), 1–35.

Andrzej Dudek, Jarosław Grytczuk, and Andrzej Ruciński

On weak twins and up-and-down subpermutations

Dedicated to the memory of Ron Graham

Abstract: Two permutations $(x_1, \dots, x_w)$ and $(y_1, \dots, y_w)$ are *weakly similar* if $x_i < x_{i+1}$ if and only if $y_i < y_{i+1}$ for all $1 \leqslant i \leqslant w$. Let π be a permutation of the set $[n] = \{1, 2, \dots, n\}$ and let $wt(\pi)$ denote the largest integer w such that π contains a pair of *disjoint* weakly similar subpermutations (called *weak twins*) of length w. Finally, let $wt(n)$ denote the minimum of $wt(\pi)$ over all permutations π of $[n]$. Clearly, $wt(n) \leq n/2$. In this paper, we show that $\frac{n}{12} \leq wt(n) \leq \frac{n}{2} - \Omega(n^{1/3})$. We also study a variant of this problem. Let us say that $(\pi(i_1), \dots, \pi(i_j))$, $i_1 < \cdots < i_j$, is an *alternating* (or *up-and-down*) subpermutation of π if $\pi(i_1) > \pi(i_2) < \pi(i_3) > \cdots$ or $\pi(i_1) < \pi(i_2) > \pi(i_3) < \cdots$. Let Π_n be a random permutation selected uniformly from all $n!$ permutations of $[n]$. Stanley has shown that the length of a longest alternating permutation in Π_n is asymptotically almost surely (a. a. s.) close to $2n/3$. We study the maximum length $\alpha(n)$ of a pair of disjoint alternating sub-permutations in Π_n and show that there are two constants $1/3 < c_1 < c_2 < 1/2$ such that a. a. s. $c_1 n \leq \alpha(n) \leq c_2 n$. In addition, we show that the alternating shape is the most popular among all permutations of a given length.

1 Introduction

Looking for twin objects in mathematical structures has a long and rich tradition going back to ancient geometric dissection problems and culminating in the famous Banach–Tarski paradox (see [18]). From that research we know, for instance, that two very different looking objects, like the Sun and an apple, or the square and the circle, can be split into finitely many pairwise identical pieces. A general problem is to partition a given structure (or structures) into possibly few pairwise similar substructures.

Acknowledgement: We would like to thank the anonymous referee for careful reading and helpful suggestions, in particular, for raising the question about the ratio A_n/B_n. The first author was supported in part by Simons Foundation Grant #522400. The second author was supported in part by Narodowe Centrum Nauki, grant 2015/17/B/ST1/02660. The third author was supported in part by Narodowe Centrum Nauki, grant 2018/29/B/ST1/00426.

Andrzej Dudek, Department of Mathematics, Western Michigan University, Kalamazoo, MI, USA, e-mail: andrzej.dudek@wmich.edu
Jarosław Grytczuk, Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland, e-mail: j.grytczuk@mini.pw.edu.pl
Andrzej Ruciński, Dept. of Discrete Mathematics, Adam Mickiewicz University, Poznań, Poland, e-mail: rucinski@amu.edu.pl

https://doi.org/10.1515/9783110754216-010

A related issue is to find, in a given structure, a pair of twin substructures, as large as possible.

Despite such "continuous" origins, questions of that sort can be studied in diverse discrete contexts, with various types of similarity specified between the objects. For instance, Chung, Graham, Erdős, Ulam, and Yao [5] studied edge decompositions of pairs of graphs into pairwise isomorphic subgraphs (see also [6, 12]), while Erdős, Pach, and Pyber [8] looked for twins in a single graph (defined as a pair of edge disjoint isomorphic subgraphs). Axenovich, Person, and Puzynina [2, 3] investigated twins in words, and Gawron [11], inspired by their work, initiated exploration of twins in permutations (defined as a pair of disjoint *order-isomorphic* subpermutations).

Let us dwell on this last problem for a while. By a *permutation*, we mean any finite sequence of distinct positive integers. Let $t(n)$ be the maximum number k such that every permutation of length n has a pair of twins, each of length k. By a probabilistic argument, Gawron [11] proved that $t(n) = O(n^{2/3})$ and made a conjecture that this is best possible, that is, $t(n) = \Theta(n^{2/3})$. We confirmed this conjecture in [7] (up to a logarithmic factor) for a *random* permutation. A refinement of our result (getting rid of the logarithmic factor) was then obtained by Bukh and Rudenko [4]. In the deterministic case, the $t(n) \geqslant \Omega(\sqrt{n})$ follows immediately from the famous result of Erdős and Szekeres [9] on monotone subsequences in permutations. Currently, the best lower bound $t(n) = \Omega(n^{3/5})$ is due to Bukh and Rudenko [4].

In this paper, we consider a weaker type of similarity of permutations than order-isomorphism in which we only look at the relations between neighboring elements. We say that two permutations $(x_1, \ldots, x_w)$ and $(y_1, \ldots, y_w)$ are *weakly similar* if $x_i < x_{i+1}$ if and only if $y_i < y_{i+1}$ for all $1 \leqslant i \leqslant w$.

This notion can be equivalently defined in terms of shapes. For our purposes, *the shape* of a permutation $\pi = (x_1, \ldots, x_w)$ is defined as a binary sequence $s(\pi) = (s_1, \ldots, s_{w-1})$ with elements from the set $\{+, -\}$, where $s_i = +$ if and only if $x_i < x_{i+1}$, $i = 1, \ldots, w-1$. For instance, $s(6, 1, 4, 3, 7, 9, 8, 2, 5) = (-, +, -, +, +, -, -, +)$. Then permutations $\pi_x = (x_1, \ldots, x_w)$ and $\pi_y = (y_1, \ldots, y_w)$ are weakly similar if $s(\pi_x) = s(\pi_y)$.

Let $[n] = \{1, 2, \ldots, n\}$ and let π be a permutation of $[n]$, called also an *n-permutation*. Two weakly similar disjoint subpermutations of π are called *weak twins* and the *length of the twins* is defined as the number of elements in just *one* of the subpermutations. For example, in permutation

$$(6, \boxed{1}, \boxed{4}, 3, \boxed{7}, 9, \boxed{8}, \boxed{2}, \boxed{5}),$$

the blue $(1, 4, 2)$ and red $(7, 8, 5)$ subsequences form weak twins of length 3 (with a common shape $(+, -)$).

Let $wt(\pi)$ denote the largest integer w such that π contains weak twins of length w. Further, let $wt(n)$ denote the minimum of $wt(\pi)$ over all n-permutations π. In other words, $wt(n)$ is the largest integer w such that every n-permutation contains weak

twins of length w. Our aim is to estimate this function which, unlike its stronger version $t(n)$, turns out to be linear in n.

Theorem 1. *For n large enough,*

$$\frac{n}{12} \le wt(n) \le \frac{n}{2} - \Omega(n^{1/3}). \tag{10.1}$$

Turning to our second result, note that given a sequence $s^{(n)}$ of length $n-1$, it is quite nontrivial to determine the number $N(s^{(n)})$ of n-element permutations with the shape $s^{(n)}$. Of course, there is just one permutation with a given monotone shape, $(+,\dots,+)$ and $(-,\dots,-)$. But already for the alternating shapes, $a_+^{(n)} = (+,-,+,\dots)$ and $a_-^{(n)} = (-,+,-,\dots)$, this is so called André's problem [1], which was solved asymptotically in the 19th century and exactly, in terms of a finite sum of Stirling numbers, only in the 21th century [15] (see also [17]).

The asymptotic formula of André says that, setting $A_n := N(a_+^{(n)}) = N(a_-^{(n)})$,

$$A_n \sim 2(2/\pi)^{n+1} n!.$$

In other words, the probability that a random n-permutation Π_n is alternating (either way) is only $\sim 4(2/\pi)^{n+1}$. On the other hand, by the result of Stanley [16], we know that a. a. s. a random n-permutation contains an alternating subsequence of length at least $\sim 2n/3$, yielding alternating twins of length at least $\sim n/3$ (just split in half a longest alternating subpermutation in Π_n). In Theorem 2, we show, however, that a. a. s. one can get substantially longer alternating twins; on the other hand, they are much shorter than $n/2$, the absolute upper bound.

To state this result, let $\alpha(\pi)$ be the largest integer w such that π contains weak twins of length w with an alternating shape, $a_+^{(w)}$ or $a_-^{(w)}$. We will call them *alternating twins*. Further, set $\alpha_n := \alpha(\Pi_n)$, where Π_n is a random n-permutation.

Theorem 2. *A. a. s.*

$$\left(\frac{1}{3} + \frac{1}{60} + o(1)\right)n \le \alpha_n \le \left(\frac{1}{2} - \frac{1}{120} + o(1)\right)n. \tag{10.2}$$

We end this paper by proving that, in fact, permutations with alternating shapes are the most popular ones. This result, not directly related to our main theorems, may be of independent interest.

Proposition 1. *For every n and every shape $s^{(n)}$ of length $n-1$, we have $N(s^{(n)}) \le A_n$.*

The proof of Proposition 1 can be found in Section 3.

Note

We believe that Ron Graham would like the topic of this paper. Not only was he among those who planted the idea of twins into the combinatorial soil, but he also wrote several papers devoted to permutations, both with and without connections to juggling

(see, e. g., http://www.math.ucsd.edu/~ronspubs/ for the entire collection of Ron's publications).

2 Proofs of Theorems 1 and 2

2.1 Extremal points

In our proofs, a decisive role is played by local extremes. We call the element i *maximal* in π if $i = 1$ and $\pi(1) > \pi(2)$, or $i = n$ and $\pi(n-1) < \pi(n)$, or $1 < i < n$ and $\pi(i-1) < \pi(i) > \pi(i+1)$. By swapping all signs $<$ and $>$ around, we obtain the notion of *a minimal* point i in π. Maximal and minimal points alternate and are jointly referred to as *extremal*. The points 1 and n are always extremal. Clearly, all extremal points of π form an alternating sequence in π. In fact, as shown by Bóna (see [17], and [13] for a proof), it is the longest one.

Let $E = \{j_1, \dots, j_k\}$ be the set of all extremal points in π. These points divide the whole range $[n]$ into monotone segments

$$\pi_i = \big(\pi(j_i), \pi(j_i+1), \dots, \pi(j_{i+1})\big), \quad i = 1, \dots, k-1, \tag{10.3}$$

which, however, share their endpoints. For a true partition, we define

$$\bar{\pi}_i = \big(\pi(j_i), \pi(j_i+1), \dots, \pi(j_{i+1}-1)\big), \quad i = 1, \dots, k-2,$$

and leave the last one unchanged, that is, $\bar{\pi}_{k-1} = \pi_{k-1}$.

2.2 Weak twins

Proof of Theorem 1, *lower bound.* As the extremal points themselves form an alternating subsequence E of π, by splitting it evenly, we obtain a pair of weak twins of length $\lfloor k/2 \rfloor$. Thus, we may assume that $k - 1 \le n/6$, since otherwise $\lfloor k/2 \rfloor \ge k/2 - 1/2 \ge n/12$ and we are done.

Let $Q_1, \dots, Q_\ell$ be those segments among $\bar{\pi}_1, \dots, \bar{\pi}_{k-1}$, which contain at least 4 elements each. It is easy to check that

$$|Q_1| + \cdots + |Q_\ell| \ge \frac{1}{2}n.$$

Indeed, otherwise we would have

$$n = \sum_{i=1}^{k-1} |\bar{\pi}_i| < 3(k-1) + \frac{1}{2}n \le n,$$

a contradiction. All we need now is the following proposition.

Proposition 2. *One can find weak twins in $Q_1 \cup \cdots \cup Q_\ell$ of length at least*

$$\frac{1}{2}\sum_{i=1}^{\ell} |Q_i| - \ell.$$

Before proving the proposition, let us finish the proof of the lower bound in (10.1). By Proposition 2, there is in π a pair of weak twins of length

$$\frac{1}{2}\sum_{i=1}^{\ell} |Q_i| - \ell \geq \frac{1}{4}n - (k-1) \geq \frac{1}{4}n - \frac{1}{6}n = \frac{n}{12}. \qquad \square$$

Proof of Proposition 2. We begin with the following observation. We say that weak twins (A, B), where $A = (\pi(i_1), \ldots, \pi(i_k))$, $i_1 < \cdots < i_k$, and $B = (\pi(j_1), \ldots, \pi(j_k))$, $j_1 < \cdots < j_k$, are *aligned upward*, respectively, *downward* if the two right-most elements of A and the two right-most elements of B interwind and form a monotone subsequence, that is, $j_{k-1} < i_{k-1} < j_k < i_k$ and $\pi(j_{k-1}) < \pi(i_{k-1}) < \pi(j_k) < \pi(i_k)$, or, respectively, $\pi(j_{k-1}) > \pi(i_{k-1}) > \pi(j_k) > \pi(i_k)$.

Claim 1. *Let (A, B) be aligned weak twins in π and let $Q = (\pi(m_1), \ldots, \pi(m_s))$, $s \geq 4$, be a monotone subsequence of π completely to the right of (A, B), that is, $m_1 > i_k$. Then one can extend (A, B) to a new pair of aligned weak twins (A', B') which contains all elements of A, B and Q except for at most 2 elements. The lost elements are either all from Q (the first or the last or both) or one from Q (the last one) and one from A (the last one).*

Proposition 2 follows quickly from the above claim. Indeed, by its repeated applications, beginning with selecting a pair of aligned weak twins (A_1, B_1) within Q_1 (here we lose one element in the case when $|Q_1|$ is odd), we recursively construct the desired object losing along the way at most $1 + 2(\ell - 1) < 2\ell$ elements. $\square$

Proof of Claim 1. Without loss of generality, assume that the weak twins (A, B) are aligned upward. However, with respect to Q, we have to consider both cases of its monotonicity. We first assume that Q is increasing.

We are going to examine 4 cases of how the two bottom values in Q position themselves with respect to the two top ones in (A, B) (see Figure 10.1). Set $a = \pi(i_k)$, $\bar{a} = \pi(i_{k-1})$, $b = \pi(j_k)$, $\bar{b} = \pi(j_{k-1})$, and $q_i = \pi(m_j)$, $j = 1, 2, \ldots, s$. Recall that $\bar{b} < \bar{a} < b < a$.

`Case 1:` $q_1 < b, q_2 < a$. We extend A and B as follows:

$$A' = A, q_2, q_4, \ldots, \quad B' = B, q_1, q_3, \ldots.$$

If s is odd, the point q_s is not used (we say it is lost). Note that due to the order of q_1, q_2, q_3, q_4, the new pair (A', B') is indeed aligned.

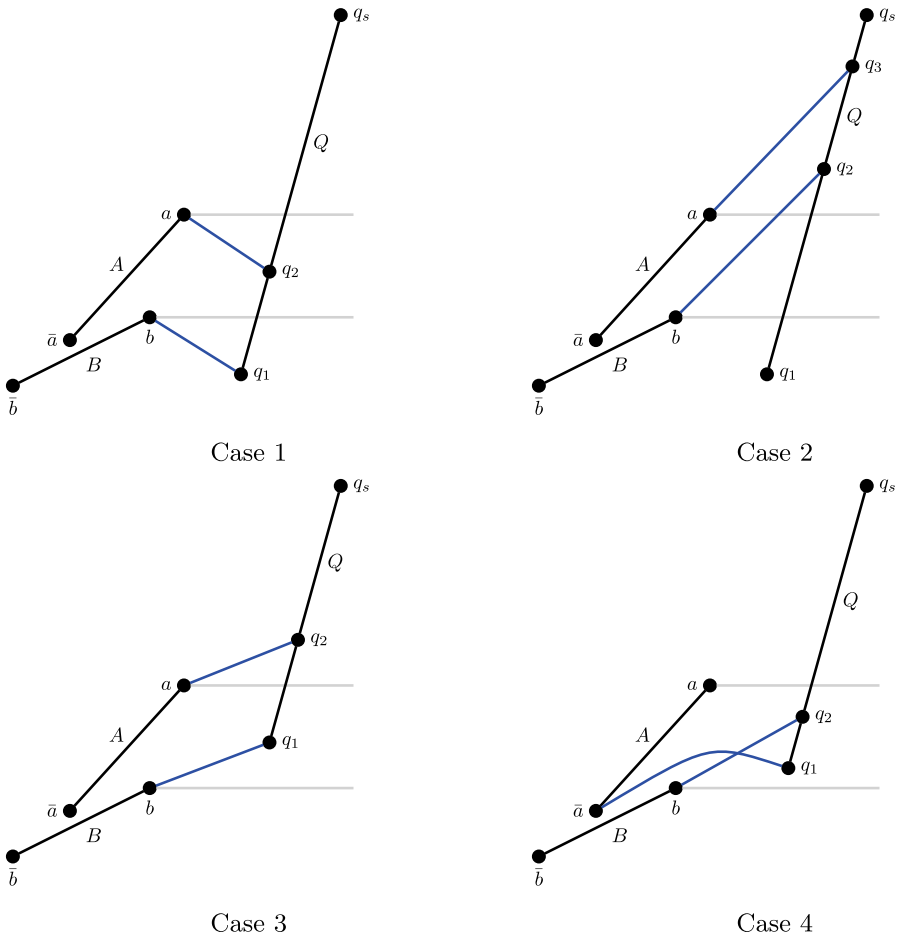


Figure 10.1: Extending twins in the proof of Claim 1 with increasing Q.

Case 2: $q_1 < b < a < q_2$. Here, we set

$$A' = A, q_3, q_5, \ldots, \quad B' = B, q_2, q_4, \ldots.$$

We definitely lose q_1 and, if s is even, we also lose q_s. For $s = 4$ or $s = 5$, the last 4 points of (A', B') are thus b, a, q_2, q_3, which are aligned upward. If $s \geq 6$, then (A', B') is aligned as well.

Case 3: $q_1 > b, q_2 > a$. This case is very similar to Case 1, so we omit the details.

Case 4: $b < q_1 < q_2 < a$. This is the only case when we lose a point of (A, B). Let A^- denote the subsequence A without the last element, a. We set

$$A' = A^-, q_1, q_3, \ldots, \quad B' = B, q_2, q_4, \ldots.$$

Besides a, we may also lose q_s, provided s is even. Observe that for $s = 4$, b, q_1, q_2, q_3 are aligned upward. This exhaust the case when Q is increasing.

For decreasing Q, there are also 4 cases to examine. However, three of them, namely, (i) $a > q_1, b > q_2$, (ii) $q_1 > a > b > q_2$, and (iii) $q_1 > a, q_2 > b$ are very similar to those for increasing Q, so we leave them for the reader. The only somewhat different case is when (iv) $a > q_1 > q_2 > b$ (see Figure 10.2). Then, denoting by B^- the subsequence B without its last element, b, we set

$$A' = A, q_2, q_4, \dots, \quad B' = B^-, q_1, q_3, \dots.$$

Besides b, we may also lose q_s, provided s is even. Finally, observe that for $s = 4$, a, q_1, q_2, q_3 are aligned downward, though with the roles of A' and B' switched (which does not really matter to us; formally we should swap A' and B' around). □

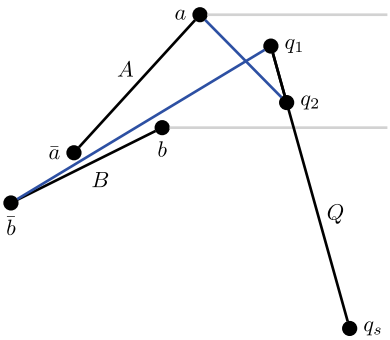


Figure 10.2: Extending twins in the proof of Claim 1 with decreasing Q.

Proof of Theorem 1, *upper bound.* We are going to construct a permutation π on $[n]$, n large enough, with no weak twins longer than $n/2 - cn^{1/3}$ for some $c > 0$. This permutation will consist of $k' \le k := \lceil n^{1/3} \rceil$ consecutive increasing segments $P_1, \dots, P_{k'}$ with $\max P_{i+1} < \min P_i$, of diminishing lengths, which have to be chosen carefully. For $i = 1, \dots, k$, set

$$x_i = 2k^2 - 2(i-1)k - 1.$$

Note that for all $i = 1, \dots, k$, $x_i > 0$ and x_i is an odd integer. Moreover,

$$\sum_{i=1}^{k} x_i = 2k^3 - 2k\binom{k}{2} - k = k^3 + k^2 - k > n.$$

Let $k' = \min\{j : \sum_{i=1}^{j} x_i \ge n\}$. Then we set $|P_i| = x_i$, $i = 1, \dots, k'-1$, and $|P_{k'}| = n - \sum_{i=1}^{k'-1} x_i$. Since $\sum_{i=1}^{k} x_i = n + O(k^2)$, with a big margin we have, say, $k' \ge 0.99k$. Also, what is

crucial here, for all $i = 1, \ldots, k' - 1$, we have $x_i - x_{i+1} \geq 2k$, in fact, with equality except for $i = k' - 1$.

So, we define $\pi = (P_1, \ldots, P_{k'})$ in the following manner. We set

$$\pi(1) = n - x_1 + 1, \pi(2) = n - x_1 + 2, \ldots, \pi(x_1) = n \quad \text{and} \quad A_1 = \big(\pi(1), \ldots, \pi(x_1)\big).$$

Then we dip down and set

$$\pi(x_1 + 1) = n - x_1 - x_2 + 1, \pi(x_1 + 2) = n - x_1 - x_2 + 2, \ldots, \pi(x_1 + x_2) = n - x_1$$

and

$$A_2 = \big(\pi(x_1 + 1), \ldots, \pi(x_1 + x_2)\big),$$

and so on, and so forth.

We now state a proposition from which the desired bound follows quickly.

Proposition 3. *Let (A, B) be weak twins in the permutation π defined above of length $|A| = |B| \geq n/2 - k/3$. Then, for all $1 \leq i < k'$, we have $|A \cap P_i| = |B \cap P_i|$.*

Before proving the proposition, let us finish the proof of the upper bound in Theorem 1. Suppose there is in π a pair of weak twins of length at least $n/2 - k/2$. Since for all $1 \leq i < k'$, $|P_i|$ is odd, in view of Proposition 3, at least one point of each such P_i is missing from (A, B). Hence,

$$|A| = |B| \leq n/2 - \big(k' - 1\big)/2 \leq n/2 - (0.99k - 1)/2 < n/2 - k/3,$$

a contradiction. □

Proof of Proposition 3. We proceed by (strong) induction on $i = 1, \ldots, k' - 1$. Let us start with the base case $i = 1$. Since at most $2k/3$ points of π are not in $A \cup B$, while $|P_1| > 2k/3$, without loss of generality, $A \cap P_1 \neq \emptyset$. It suffices to prove that also $B \cap P_1 \neq \emptyset$, since then, due to the fact that the rest of π lies totally below P_1, A and B must have the same number of elements in P_1. Suppose to the contrary that $B \cap P_1 = \emptyset$. But then

$$|A \cap P_1| \geq |P_1| - \frac{2}{3}k > |P_2| > |P_3| > \cdots,$$

so A begins with a longer increasing segment than B does, a contradiction with the notion of weak twins.

For the induction step, which is similar to the base step, assume that $|A \cap P_j| = |B \cap B_j|$, for $j = 1, \ldots, i \leq k' - 2$. If $|A \cap P_{i+1}| = |B \cap B_{i+1}| = 0$, then we are done. Without loss of generality, assume that $|A \cap P_{i+1}| > 0$. As before, it suffices to show that also $|B \cap P_{i+1}| > 0$; suppose otherwise. Then, since at most $2k/3$ points of π are not in $A \cup B$,

we have

$$|A \cap P_{i+1}| \geq |P_{i+1}| - \frac{2}{3}k > |P_{i+2}| > \cdots.$$

This means, however, that A and B will differ in the length of the first increasing segment commencing to the right of the point $\sum_{j=1}^{i} x_j$. This yields a contradiction with (A, B) being weak twins and completes the proof. □

2.3 Alternating weak twins

Recall that the extremal points of π form an alternating subsequence. In the proof of the lower bound in Theorem 2, we are going to use this fact and then reiterate it for the subpermutation π' obtained from π by removing all the extremal points of π. As a crucial tool, we invoke the Azuma–Hoeffding inequality for random permutations (see, e. g., Lemma 11 in [10] or Section 3.2 in [14]).

Theorem 3. *Let $h(\pi)$ be a function of n-permutations such that if permutation π_2 is obtained from permutation π_1 by swapping two elements, then $|h(\pi_1) - h(\pi_2)| \leq 1$. Then, for every $\eta > 0$,*

$$\mathbb{P}\big(|h(\Pi_n) - \mathbb{E}[h(\Pi_n)]| \geq \eta\big) \leq 2\exp(-\eta^2/(2n)).$$

Proof of Theorem 2, *lower bound.* We are going to show that extremal points are evenly distributed in both "halves" of Π_n. For mere convenience, we assume that n is even.

Let X_1 and X_2 be the numbers of extremal points in Π_n among, respectively, $\{1, \ldots, n/2\}$ and $\{n/2{+}1, \ldots, n\}$. Note that the probability that a given point i, $2 \leq i \leq n{-}1$, is extremal is $2 \times \frac{1}{3} = \frac{2}{3}$. Thus,

$$\mathbb{E}(X_1) = \mathbb{E}(X_2) = 1 + \left(\frac{n}{2} - 1\right) \times \frac{2}{3} = \frac{n+1}{3}.$$

Now we apply Theorem 3 to show that this expectation is highly concentrated about its mean. To verify the Lipschitz assumption, note that if π_2 is obtained from a permutation π_1 by swapping any two of its elements, then trivially $|X_j(\pi_1) - X_j(\pi_2)| \leq 6$, $j = 1, 2$. (A detailed analysis shows that 6 can be replaced by 4, which is optimal.) Consequently, Theorem 3 applied with $h(\pi) = X_j(\pi)/6$ and $\eta = n^{3/5}$ implies

$$\mathbb{P}\big(|X_j(\Pi_n) - \mathbb{E}[X_j(\Pi_n)]| \geq n^{3/5}\big) = o(1)$$

implying that a. a. s. $X_j = (1 + o(1))\frac{n}{3}$, $j = 1, 2$.

It is quite hard to characterize the extremal points of π'. Unable to do so, we instead identify a 6-point configuration in π which contains an extremal point of π'. A

6-tuple $\{i, i+1, i+2, i+3, i+4, i+5\}$, $1 \le i \le n-5$, is called *a lucky six* if $\pi(i) < \pi(i+1) < \pi(i+2) < \pi(i+3) > \pi(i+4) > \pi(i+5)$ and $\pi(i+2) > \pi(i+4)$, or when all signs $<$ and $>$ are swapped. It should be clear that in a lucky six $i+3$ is an extremal point of π and, most importantly, $i+2$ is an extremal point in π'. Of course, the same property is enjoyed by the symmetrical structures where $\pi(i) < \pi(i+1) < \pi(i+2) > \pi(i+3) > \pi(i+4) > \pi(i+5)$ and $\pi(i+1) < \pi(i+3)$ (and, again, with signs $<$ and $>$ swapped). So, we also call them *lucky sixes*. See Figure 10.3 for all 4 types of lucky sixes.

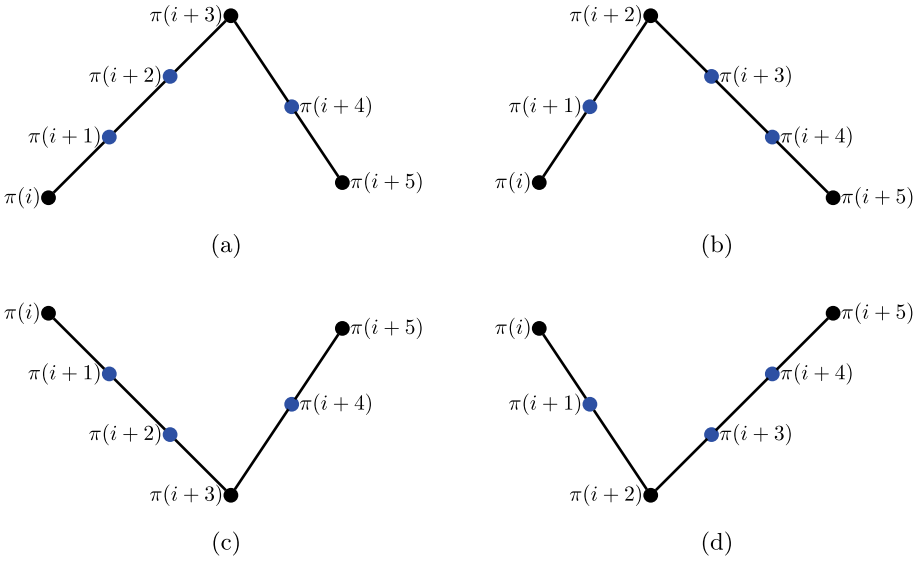


Figure 10.3: Lucky sixes. The blue points appear in π' as consecutive ones.

Let Y_1 and Y_2 be the numbers of lucky sixes $\{i, i+1, i+2, i+3, i+4, i+5\}$ in Π_n for, respectively, $1 \le i \le n/2-3$ and $n/2-1 \le i \le n-5$. Note that the probability that a given 6-tuple is a lucky six is

$$4 \times \frac{\binom{4}{2}}{6!} = \frac{1}{30}.$$

Indeed, considering, for instance, the number of ways to label by $1, \ldots, 6$, the lucky six in Figure 10.3(a), there is no question that 6 must be at the top, while 5 to its left. The remaining 4 values can be, however, distributed freely between the two pairs, $i, i+1$ and $i+4, i+5$. This explains $\binom{4}{2}$. Thus

$$\mathbb{E}(Y_1) = \mathbb{E}(Y_2) \sim \frac{n}{60}.$$

Again, a standard application of the Azuma inequality (Theorem 3) yields that a. a. s. $Y_j = (1 + o(1))\frac{n}{60}$, $j = 1, 2$.

Let A_j, $j = 1, 2$, be the alternating subsequences in, respectively, $\{1, \dots, n/2\}$ and $\{n/2 + 1, \dots, n\}$, consisting of the extremal points of Π_n. Further, let B_j, $j = 1, 2$, be alternating subsequences in, respectively, $\{1, \dots, n/2\}$ and $\{n/2 + 1, \dots, n\}$, consisting of the extremal points of Π_n'. By losing at most one point each, one can concatenate A_j with B_{3-j}, $j = 1, 2$, obtaining the desired pair of alternating twins. Noting that $|A_j \cup B_{3-j}| \sim \frac{n}{3} + \frac{n}{60}$ completes the proof of the lower bound in (10.2). □

Proof of Theorem 2, upper bound. For the proof of the upper bound, we need to consider two kinds of special 5-tuples. A 5-tuple $\{i, i + 1, i + 2, i + 3, i + 4\}$ is called *cornered* if either the first or the last four consecutive points form a monotone sub-sequence but all five do not (see Figure 10.4). A 5-tuple $\{i, i + 1, i + 2, i + 3, i + 4\}$ is called *crooked* if the three middle points form a monotone subsequence but no four points do (see Figure 10.5). Given a permutation π, let $e(\pi)$ be the number of extremal points in π, and let $co(\pi)$ and $cr(\pi)$ be, respectively, the number of cornered 5-tuples and the number of crooked 5-tuples in π. The following crucial lemma sets an upper bound on the number of elements in two disjoint alternating subsequences of π in terms of the three defined above parameters.

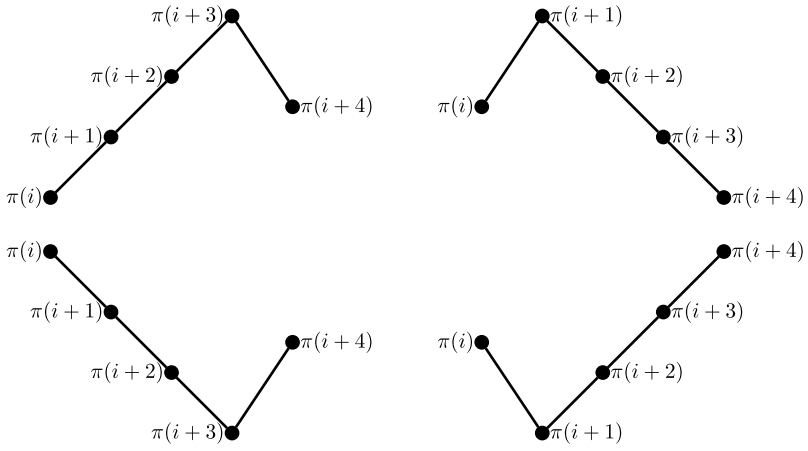


Figure 10.4: Cornered 5-tuples.

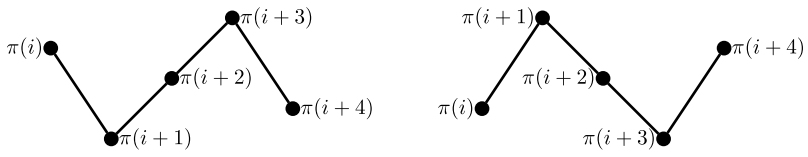


Figure 10.5: Crooked 5-tuples.

Lemma 1. *Let A and B be two disjoint alternating subsequences in a permutation π of $[n]$. Then*

$$|A| + |B| \leq e(\pi) + co(\pi) + cr(\pi). \tag{10.4}$$

Deferring the proof of Lemma 1 for later, we now deduce from it the upper bound in (10.2). Let L count the cornered 5-tuples in the random permutation Π_n and let Z count the crooked 5-tuples in Π_n. Note that the probability that a given 5-tuple is cornered is $4 \times \binom{4}{3}/5! = \frac{8}{60}$ and so, $\mathbb{E}(L) = \frac{8}{60} \times (n-4)$. Note also the that the probability that a given 5-tuple is crooked is $2 \times \frac{11}{5!} = \frac{11}{60}$ (see Figure 10.6) and so, $\mathbb{E}(W) = \frac{11}{60} \times (n-4)$. Another application of the Azuma inequality (Theorem 3) yields that a. a. s. $L = (1+o(1))\frac{8n}{60}$, while $Z = (1+o(1))\frac{11n}{60}$. Plugging into (10.4), we finally obtain that

$$\alpha_n \leq \frac{1}{2}(1+o(1))\left(\frac{2}{3} + \frac{8}{60} + \frac{11}{60}\right)n = \left(\frac{1}{2} - \frac{1}{120} + o(1)\right)n. \qquad \square$$

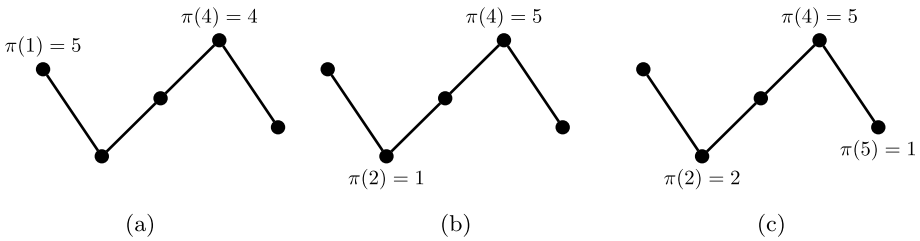


Figure 10.6: (a) If $\pi(1) = 5$, then $\pi(4) = 4$ and the remaining number of choices is $\binom{3}{2}$. (b) If $\pi(4) = 5$ and $\pi(2) = 1$, then we have 3! choices. (c) Finally, if $\pi(4) = 5$ and $\pi(5) = 1$, then there are 2! remaining choices.

It remains to prove Lemma 1.

Proof of Lemma 1. Let A and B be given as in the lemma. Let E be the set of extremal points in π and F the set of points neighboring the extremal points but not extremal themselves. We are going to construct an injective mapping $\phi : A \cup B \to E \cup F$. Then, noting that $|F| = co(\pi) + cr(\pi)$, completes the proof.

Let $j_1, \ldots, j_k$ be all extremal points in π. These points divide the whole range $[n]$ into monotone segments defined in (10.3), which we now express in terms of the numbers of their inner points ℓ_i:

$$\pi_i = \big(\pi(j_i), \pi(j_i+1), \ldots, \pi(j_i+\ell_i), \pi(j_{i+1})\big)$$

$i = 1, \ldots, k-1$. Note that ℓ_i can equal 0. Before constructing the desired mapping ϕ, let us examine the distribution of the set $A \cup B$ among the segments π_i. Our first observation is that each segment contains at most two elements of A and at most two elements

of B. Moreover, if π_i contains exactly two elements of A, then one of them is a minimal element of A and the other—a maximal element of A, and the same is true for B. But most crucial is the following property concerning a pair of consecutive segments π_i and π_{i+1}. If j_{i+1} is maximal, respectively, minimal in π, then there is in total at most one maximal, respectively, minimal element of A on these segments.

Knowing all this, it is easy to see that the following construction is, indeed, an injection. If π_i is increasing, then to the maximal elements of $A \cup B$ lying on π_i, assign the top-most two elements of π_i, that is, to $\pi(j_i + \ell_i), \pi(j_{i+1})$, in any feasible fashion. While to the minimal elements of $A \cup B$ lying on π_i assign the two down-most elements of π_i, that is, to $\pi(j_i), \pi(j_i+1)$. If π_i is decreasing, we proceed similarly, but with the pairs $\pi(j_i + \ell_i), \pi(j_{i+1})$ and $\pi(j_i), \pi(j_i + 1)$ swapped. □

3 Proof of Proposition 1

Given a sequence $s = (s_1, \dots, s_r)$ with $s_i \in \{+, -\}$ and a linearly ordered set S of size $|S| = r + 1$, denote by $\mathcal{N}_S(s)$ the set of all permutations π of S with the shape $s(\pi) = s$. If $S = [r+1]$, then we abbreviate $\mathcal{N}(s) := \mathcal{N}_{[r+1]}(s)$. Further, let $N_S(s) = |\mathcal{N}_S(s)|$. Observe that $N_S(s)$ does not depend on S, so we skip the subscript $_S$ altogether here.

The complement of a sequence $s = (s_1, \dots, s_r)$ is naturally defined as the sequence $\bar{s} = (\bar{s}_1, \dots, \bar{s}_r)$, where $\{s_i, \bar{s}_i\} = \{+, -\}$ for each i. In other words, one replaces each $+$ in s with $-$, and vice versa. It is easy to see that $N(s) = N(\bar{s})$.

Recall that $A_n = N(a_+^{(n)}) = N(a_-^{(n)})$. Our proof of Proposition 1 is by induction on n and, in its final accord, utilizes the following known identity involving the sequence A_n (see, e. g., [17]):

$$\sum_{k=0}^{n} \binom{n}{k} A_k A_{n-k} = 2A_{n+1}. \tag{10.5}$$

What is more, our proof is also inspired by the idea behind the proof of (10.5), which is to build a permutation of $[n+1]$ beginning with positioning the element $n+1$, and then separately counting the completions to the left and to the right of it. Also, as the RHS of (10.5) is a double of what we want, we are doomed to count in permutations with the complementary shape as well.

Proof of Proposition 1. For $n \le 3$, the proposition follows by inspection. Fix $n \ge 3$ and assume it is true for all $n' \le n$. Given a shape $s^{(n+1)} := s = (s_1, \dots, s_n)$ our goal is to show that $N(s) \le A_{n+1}$.

For each $k = 0, 1, \dots, n$, let $\mathcal{N}_k(s) = \{\pi \in \mathcal{N}(s) : \pi(k+1) = n+1\}$. As $n+1$ is always a maximum element of π, we have $\mathcal{N}_k(s) \ne \emptyset$ if and only if $s_k = +$ and $s_{k+1} = -$. Thus, setting $K^\wedge = \{k : s_k = + \text{ and } s_{k+1} = -\}$, we have $\mathcal{N}(s) = \bigcup_{k \in K^\wedge} \mathcal{N}_k(s)$, and, as the sets under the union are obviously disjoint, $N(s) = \sum_{k \in K^\wedge} N_k(s)$, where $N_k(s) = |\mathcal{N}_k(s)|$. For

a fixed k, let us focus on the number $N_k(s)$. Every permutation in $\mathcal{N}_k(s)$ consists of a "prefix" u, followed by $n+1$, followed by a suffix v. Introducing "truncated" shapes $s'_k = (s_1, \ldots, s_{k-1})$ and $s''_k = (s_{k+2}, \ldots, s_n)$, u and v must satisfy $s(u) = s'_k$ and $s(v) = s''_k$. Hence, using also the induction assumption,

$$N_k(s) = \binom{n}{k} N(s'_k) N(s''_k) \leq \binom{n}{k} A_k A_{n-k}.$$

The same is true for the complementary shape $\bar{s}$ as well. Recalling that $N(\bar{s}) = N(s)$ and noticing that the set

$$\{k : \bar{s}_k = + \text{ and } \bar{s}_{k+1} = -\} = \{k : s_k = - \text{ and } s_{k+1} = +\} =: K^{\vee}$$

is disjoint from $K^{\wedge}$, we thus conclude that

$$2N(s) \leq \sum_{k \in K^{\wedge} \cup K^{\vee}} \binom{n}{k} A_k A_{n-k} \leq \sum_{k=0}^{n} \binom{n}{k} A_k A_{n-k} = 2A_{n+1},$$

where the last equality is (10.5). □

4 Concluding remarks

We believe that the lower bound in Theorem 1 can be improved and it is plausible to conjecture that $wt(n) \sim \frac{n}{2}$. As a matter of fact, if π happens to be an n-permutation with $e(\pi) = o(n)$, then the construction used in the proof of (10.1) yields $wt(\pi) \sim \frac{n}{2}$.

It is also not difficult to see that the lower bound on α_n in Theorem 2 can be improved. Let π' be the subpermutation obtained from π by removing all the extremal points of π. Recall that in the proof of the lower bound (10.2) we estimated $e(\pi')$ by using the lucky six tuples. But one can also consider more "lucky" structures. This can be done by incorporating zigzags into the lucky six tuples as in Figure 10.7, for example. This already gives an improvement on the lower bound on α_n:

$$\alpha_n \geq \left(\frac{1}{3} + \frac{1}{60} + \frac{1}{2} \cdot 4 \cdot \frac{117 + 105}{8!} + o(1) \right) n.$$

Now we can consider longer lucky tuples (of length $10, 12, 14, \ldots$) and use computer to calculate the corresponding expectations. Computer simulations suggest that

$$\alpha_n \geq (1/3 + 0.1006\ldots)n.$$

We do not know what the exact value of the second term is here, since it is not even clear how to compute the expected value $\mathbb{E}(e(\Pi'_n))$.

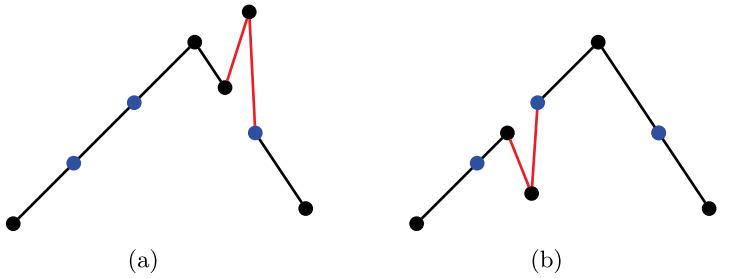


Figure 10.7: (a) 117 choices; (b) 105 choices.

Another direction of related studies would be to consider a more general notion of *weak r-twins*, defined as r pairwise disjoint subsequences of a permutation with the same shape. One naturally expects that the analogous function $wt^{(r)}(n)$ should satisfy $wt^{(r)}(n) \sim \frac{n}{r}$.

Finally, let us point at some natural counting problems involving the notion of weak similarity of permutations. For instance, define, for even n, the sequence T_n which counts all n-permutations that are weak twins of length $n/2$, that is, all n-permutations that can be split into two subpermutations with the same shape. What is the asymptotic growth of the sequence T_n?

A related problem stems from Proposition 1. We proved there that no shape is more represented among all permutations of length n than the alternating ones. It follows from the proof of Proposition 1 that actually the number A_n is strictly bigger than B_n—the largest among the numbers $N(s)$ with s being a nonalternating sequence of length $n-1$. It can be shown that $A_n/B_n \leq 2$. Indeed, by swapping the first two elements, we see that $N(b^{(n)}) = N'(a_+^{(n)})$, where $b^{(n)} = (-,-,+,-,+,\dots)$ is the shape of permutations which begin with a decreasing triple and then alternate, while $N'(a_+^{(n)})$ counts those alternating permutations π of length n for which $\pi(2) > \pi(1) > \pi(3)$. In turn, swapping the first and the third element of an alternating permutation counted by $N(a_+^{(n)})$–$N'(a_+^{(n)})$, that is, one for which $\pi(2) > \pi(3) > \pi(1)$, yields that $N'(a_+^{(n)}) \geq \frac{1}{2}A_n$. It would be interesting to compute $\lim_{n\to\infty} A_n/B_n$, if it exists.

Bibliography

[1] D. André, Développement de sec x and tg x, *C. R. Math. Acad. Sci. Paris*, **88** (1879), 965–979.
[2] M. Axenovich, Repetitions in graphs and sequences, in *Recent Trends in Combinatorics*, Springer, 2016.
[3] M. Axenovich, Y. Person and S. Puzynina, Regularity lemma and twins in sequences, *J. Comb. Theory, Ser. A*, **120** (2013), 733–743.
[4] B. Bukh and O. Rudenko, Order-isomorphic twins in permutations, *SIAM J. Discrete Math.*, **34** (2020), 1620–1622.

[5] F. R. K. Chung, P. Erdős, R. L. Graham, S. M. Ulam and F. F. Yao, Minimal decompositions of two graphs into pairwise isomorphic subgraphs, *Congr. Numer.*, **23** (1979), 3–18.
[6] F. R. K. Chung, P. Erdős and R. L. Graham, Minimal decompositions of graphs into mutually isomorphic subgraphs, *Combinatorica*, **1** (1981), 13–24.
[7] A. Dudek, J. Grytczuk and A. Ruciński, Variations on twins in permutations, *Electron. J. Comb.*, **28**(3), Paper No. 3.19 (2021), 18 pp.
[8] P. Erdős, J. Pach and L. Pyber, Isomorphic subgraphs in a graph, in *Combinatorics (Eger, 1987)*, Colloquia Mathematica Societatis János Bolyai, vol. **52**, pp. 553–556, North-Holland, Amsterdam, 1988.
[9] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compos. Math.*, **2** (1935), 463–470.
[10] A. Frieze and B. Pittel, Perfect matchings in random graphs with prescribed minimal degree, in *Mathematics and Computer Science III*, Trends Math., pp. 95–132, Birkhäuser, Basel, 2004.
[11] M. Gawron, *Izomorficzne podstruktury w słowach i permutacjach*, Master Thesis, Uniwersytet Jagielloński, 2014.
[12] R. L. Graham, Reflections on a theme of Ulam, in *Graph Theory: Favorite Conjectures and Open Problems*, vol. **II**, Springer, 2018.
[13] C. Houdré and R. Restrepo, A probabilistic approach to the asymptotics of the length of the longest alternating subsequence, *Electron. J. Comb.*, **17**, Article #R168 (2010).
[14] C. McDiarmid, Concentration, in Probabilistic Methods for Algorithmic Discrete Mathematics. Algorithms Combin., vol. **16**, Springer, Berlin, 1998.
[15] A. Mendes, A note on alternating permutations, *Am. Math. Mon.*, **114** (2007), 437–440.
[16] R. Stanley, Longest alternating subsequences of permutations, *Mich. Math. J.*, **57** (2008), 675–687.
[17] R. Stanley, A survey of alternating permutations, *Contemp. Math.*, **531** (2010), 165–196.
[18] G. Tomkowicz and S. Wagon, *The Banach-Tarski Paradox*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2016.

Sohail Farhangi and Jarosław Grytczuk

Distance graphs and arithmetic progressions

Dedicated to the memory of Ron Graham

Abstract: A set of positive integers D is called *lonely* if there exist real numbers $\alpha, \delta \in (0,1)$ such that each point of the dilation αD is at distance at least δ from the nearest integer. We prove that for every lonely set there is a 2-coloring of the integers without arbitrarily long monochromatic arithmetic progressions with steps $d \in D$. This result is a step toward a more general conjecture by Brown, Graham, and Landman, stating that a similar 2-coloring exists whenever the set of allowable steps D violates the restricted version of van der Waerden's theorem.

1 Introduction

Let $D \subseteq \mathbb{N}$ be a fixed subset of the set of positive integers. Consider a graph G_D on the set of vertices $\mathbb{N}$ in which two vertices $a, b \in \mathbb{N}$, with $a < b$, are joined by an edge if and only if $b - a \in D$. Investigations of such graphs were initiated by Eggleton, Erdős, and Skilton [4] in connection with the famous Hadwiger–Nelson problem concerning the chromatic number of the plane (see [16]).

Let $\chi(D)$ denote the chromatic number of G_D. A challenging problem is to characterize sets D with finite chromatic number. For example, if D consists of all even integers, then $\chi(D) = \infty$, since there is an infinite clique in G_D. On the other hand, if D consists of all odd numbers, then $\chi(D) = 2$, since the sets of even and odd integers are independent in G_D. This shows that the chromatic number $\chi(D)$ may radically differ for sets that are just translates of one another.

The main conjecture in this matter was posed independently by Katznelson [11] and Ruzsa (personal communication), using different terminology of topological dynamics and additive number theory, respectively. To make a precise statement, let us denote by $\|x\|$ the distance from x to the nearest integer. A set D is called *lonely*, if there exist real numbers $\alpha, \delta > 0$ such that the inequality $\|\alpha d\| \geqslant \delta$ is satisfied for all $d \in D$. For example, the set of all odd integers is lonely, as can be seen by taking $\alpha = \delta = 1/2$.

Conjecture 1 (Katznelson–Ruzsa). The chromatic number $\chi(D)$ is finite if and only if the set D is a finite union of lonely sets.

Acknowledgement: S. Farhangi was supported by the National Science Center of Poland, grant 2015/17/B/ST1/02660. We would like to thank Vitaly Bergelson for making us aware of Pollington [14].

Sohail Farhangi, The Ohio State University, Columbus, OH, USA, e-mail: farhangi.3@osu.edu
Jarosław Grytczuk, Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw, Poland, e-mail: j.grytczuk@mini.pw.edu.pl

https://doi.org/10.1515/9783110754216-011

That the loneliness condition on D is sufficient for the finiteness of $\chi(D)$ was proved by Katznelson [11] and independently (implicitly) by Ruzsa, Tuza, and Voigt [15]. Both papers solve a problem posed by Erdős whether $\chi(D)$ is finite for sets with exponential growth (so-called *lacunary* sets).

In this note, we prove a result concerning arithmetic progressions whose steps are restricted to lonely sets. The celebrated theorem of van der Waerden [17] asserts that any finite coloring of $\mathbb{N}$ admits arbitrarily long monochromatic arithmetic progressions. Clearly, this is not necessarily true if we restrict the set of allowable steps of arithmetic progressions to some fixed set D. In particular, in a proper coloring of G_D with $\chi(D)$ colors, there are no nontrivial monochromatic arithmetic progressions (with steps from D) at all.

The problem of characterizing sets D for which the restricted van der Waerden's theorem holds was undertaken by Brown, Graham, and Landman [2]. They posed the following intriguing conjecture.

Conjecture 2 (Brown, Graham, Landman, [2]). Let D be a set of positive integers. Suppose that there is a finite coloring of $\mathbb{N}$ such that all monochromatic arithmetic progressions with steps in D have bounded length. Then there is a 2-coloring of $\mathbb{N}$ with the same property.

In this note, we prove that lonely sets satisfy this conjecture. In particular, this solves an open problem, posed in [1] (see also [12]), of determining the least number of colors needed to avoid long monochromatic arithmetic progressions with steps belonging to the set of Fibonacci numbers.

2 The result

Let us consider the *torus* $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, which is geometrically just a circle of unit circumference with a distinguished point 0. For $x \in \mathbb{R}$, there is a unique point on $\mathbb{T}$ corresponding to x whose circular coordinate is the factional part $\{x\}$ of x. We will denote real numbers and their corresponding points on the torus by the same symbols. By $\|x\|$, we denote the circular distance from $x \in \mathbb{T}$ to the point 0. More precisely, $\|x\|$ equals $\{x\}$ or $1 - \{x\}$, where $\{x\}$ is the fractional part of x. Notice also that for any two numbers $x, y \in \mathbb{R}$, the number $\|x - y\|$ is equal to the circular distance between the corresponding points on the torus $\mathbb{T}$.

Suppose now that D is a lonely set, i. e., the inequality $\|\alpha d\| \geqslant \delta$ holds for some $\alpha, \delta \in (0,1)$ and all $d \in D$. For a fixed $\alpha \in (0,1)$, let $\delta_\alpha(D) = \inf\{\|\alpha d\| : d \in D\}$, and let $\lambda(D) = \sup\{\delta_\alpha(D) : \alpha \in (0,1)\}$. We shall call it the *loneliness constant* of the set D.

Theorem 1 (Katznelson [11]). *Let D be a lonely set with the loneliness constant $\lambda(D)$. Then $\chi(D) \leqslant \lceil \frac{1}{\lambda(D)} \rceil$.*

Proof. Put $k = \lceil \frac{1}{\lambda(D)} \rceil$ and partition the torus $\mathbb{T}$ into half-open arcs $A_j = [\frac{j}{k}, \frac{j+1}{k})$, for $j = 0, 1, \ldots, k-1$. Define a k-coloring $c : \mathbb{N} \to \{0, 1, \ldots, k-1\}$ by $c(n) = j$ if and only if $\alpha n \in A_j$. We claim that this is a proper coloring of the distance graph G_D. Indeed, suppose that $c(a) = c(b) = j$ for some $a, b \in \mathbb{N}$. Then $\alpha a, \alpha b \in A_j$, which implies that $\|\alpha a - \alpha b\| < \frac{1}{k}$. Hence, $\|\alpha(a-b)\| < \frac{1}{k}$, and in consequence $a - b$ cannot be an element of D (since every $d \in D$ satisfies $\|\alpha d\| \geq \lambda(D) \geq \frac{1}{k}$). □

Now we use similar methods to prove Theorem 2, which is an effective version of both Lemma 7.4 in [9] and Corollary 8.11 in [6].

Theorem 2. *Let D be a lonely set with the loneliness constant λ. Then there exists a 2-coloring of $\mathbb{N}$ such that no arithmetic progression of length $\ell = \lceil \frac{1}{2\lambda} \rceil + 1$ and step $d \in D$ is monochromatic.*

Proof. Let D be a lonely set with loneliness $\lambda > 0$. Consider a red-blue coloring of the torus $\mathbb{T} = R \cup B$, where $R = [0, \frac{1}{2})$ and $B = [\frac{1}{2}, 1)$. Define a red-blue coloring of $\mathbb{N}$ so that the color of a number $n \in \mathbb{N}$ coincides with the color of the corresponding point $\alpha n \in \mathbb{T}$ on the torus. Here, α is a constant satisfying the loneliness condition $\|\alpha d\| \geqslant \lambda$ for all $d \in D$.

Now, consider any arithmetic progression $a, a+d, a+2d, \ldots, a+kd$ with $d \in D$. We claim that it is not monochromatic, provided that $k \geqslant \frac{1}{2\lambda}$. Indeed, consider the corresponding points $\alpha a, \alpha(a+d), \alpha(a+2d), \ldots, \alpha(a+kd)$. These points also form an arithmetic progression with step $s = \|\alpha d\| \geqslant \lambda$ on the torus $\mathbb{T}$ (going clockwise or counterclockwise). So, the length of the whole arc spanned between the first and the last point of this progression equals at least $ks \geqslant \frac{1}{2\lambda}\lambda = \frac{1}{2}$. On the other hand, $s \leqslant \frac{1}{2}$, so, there must exist two points in the progression dropping into different parts of the partition of the torus. This proves the theorem. □

Recall that a set $D = \{d_1, d_2, \ldots\}$ is *lacunary* if there exists a real number $\theta > 0$ such that $\frac{d_{i+1}}{d_i} \geqslant 1 + \theta$, for all $i = 1, 2, \ldots$. Pollington [14] was the first to show that every lacunary set is lonely, thus answering a question of Erdős [5] (see also de Mathan [3]). This fact was independently rediscovered by Katznelson [11], and it was also independently shown by Ruzsa, Tuza, and Voigt [15] that sufficiently quickly growing lacunary sequences are lonely. For instance, the set of Fibonacci numbers $F = \{1, 2, 3, 5, 8, 13, \ldots\}$ is well known to be lacunary. Hence F is lonely and satisfies the assertion of Theorem 2. As mentioned at the end of the Introduction, this answers a question posed in [1] (see also [12]).

Moreover, as proved by Peres and Schlag in [13], every finite union of lacunary sets is a lonely set. By their results, one may derive the dependence between loneliness and lacunary constants and obtain thereby an upper bound on the length of monochromatic arithmetic progressions in this case. More specifically, let $t \geqslant 1$ be a fixed integer, and suppose that D_i is a lacunary set with the lacunary constant $1 + \theta_i$, for $i = 1, 2 \ldots, t$. Then, as proved in [13], the set $D = \bigcup_{i=1}^{t} D_i$ is lonely with the loneliness

constant satisfying

$$\lambda(D) \geqslant \frac{1}{240M \log_2 M},$$

where $M = \max(\sum_{i=1}^{t} \lceil \theta_i^{-1} \rceil, 4)$.

Finally, notice that Theorem 2 has the following consequence for finite unions of lonely sets.

Corollary 1. *Let $t \geqslant 1$ be an integer and let $D = D_1 \cup D_2 \cup \cdots \cup D_t$, where each D_i is a lonely set. Then there exists a 2^t-coloring of $\mathbb{N}$ and a constant $\ell = \ell(D)$ such that no arithmetic progression of length ℓ and step $d \in D$ is monochromatic.*

Proof. It is enough to take the product coloring whose components are the 2-colorings guaranteed by Theorem 2 for each of the sets D_i. Clearly, there can be no monochromatic progression with step in D of length greater than $\lceil \frac{1}{2\lambda} \rceil$, where $\lambda = \min \lambda(D_i)$. □

3 Further remarks

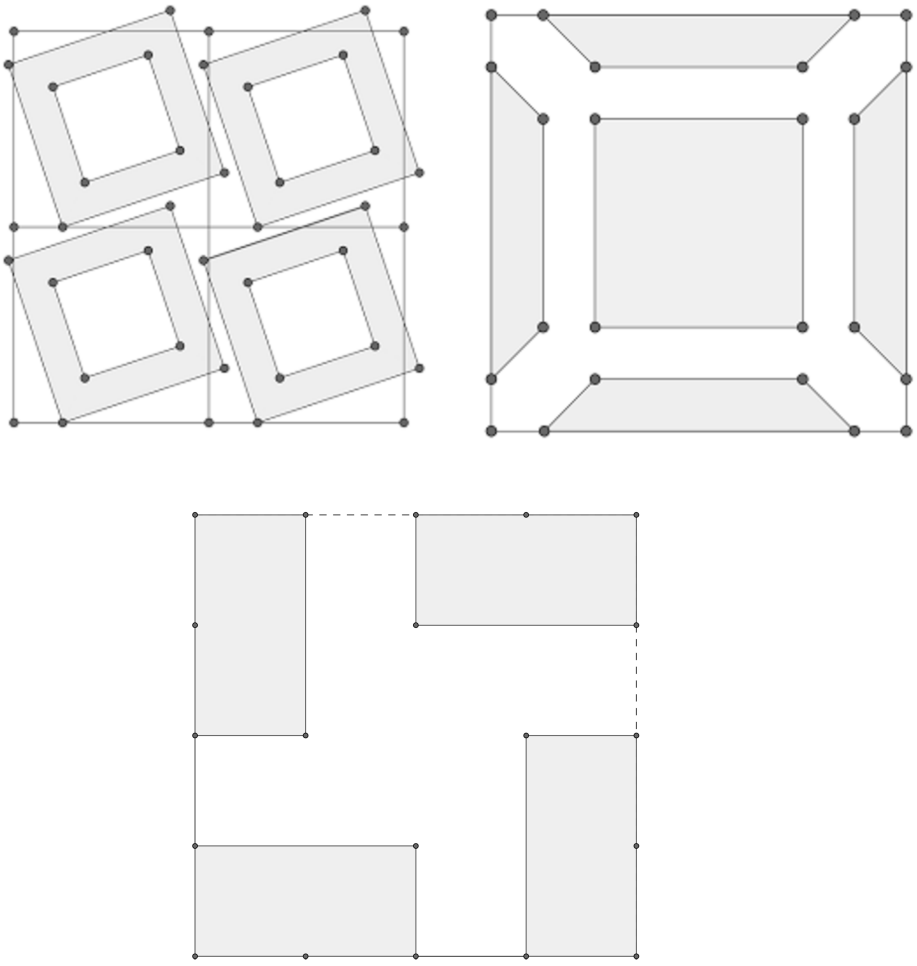
We conclude this short note with some reflections on further possible applications of the torus coloring method to other problems for distance graphs.

The first and natural attempt would be to extend Theorem 2 to all sets D with finite chromatic number $\chi(D)$. A natural attempt here is to consider finite unions of lonely sets (cf. Conjecture 8.13 in [6]). By Corollary 1, we know that a desired coloring (avoiding long arithmetic progressions with steps in D) can be obtained by using at most 2^t colors, where t is the number of lonely sets in the union.

We will now propose a series of 2-colorings of the integers which we believe to show (but are unable to prove) that a union of two lonely sets is not 2-large. Given a partition of the 2-Torus $\mathbb{T}^2 = A \cup B$, and some $(\alpha, \beta) \in \mathbb{T}^2$, we may define the coloring $f_{A,B,\alpha,\beta} : \mathbb{N} \to \{1, 2\}$ by

$$f_{A,B,\alpha,\beta}(n) = \begin{cases} 1 & \text{if } (n\alpha, n\beta) \in A \\ 2 & \text{if } (n\alpha, n\beta) \in B. \end{cases} \tag{11.1}$$

Informally, Theorem 2 was proven by analyzing the coloring $f_{R,B,\alpha}$. Now let $S = L_1 \cup L_2$ be a union of two lonely sets. Suppose that $\alpha, \beta \in \mathbb{R}$ and $\delta > 0$ is such that $||n\alpha|| > \delta$ for every $n \in L_1$ and $||n\beta|| > \delta$ for every $n \in L_2$. We conjecture that for at least one of the partitions of $\mathbb{T}^2$ that are shown below, the coloring $f_{A,B,\alpha,\beta}$ will not contain long monochromatic arithmetic progressions with steps in S.



We hope that at least one of the above partitions will also give insight into partitions of higher dimensional Tori that will allow us to generalize Theorem 2 to finite unions of lonely sets.

Another direction could be to look at arbitrary forward paths in distance graphs, not only following arithmetic progressions. It is known, for instance, that for the set F of Fibonacci numbers, there is a 6-coloring of G_F avoiding arbitrarily long monochromatic forward paths. On the other hand, two colors are not sufficient for this property (see [1]). Other related problems and results can be found in [8].

It is also known that there is a set H with infinite chromatic number $\chi(H)$ and a finite coloring of $\mathbb{N}$ avoiding 3-term monochromatic arithmetic progressions with steps in H (see [10] or Section 9.1 of [7]). On the other hand, there is no finite coloring of $\mathbb{N}$ avoiding arbitrarily long monochromatic forward paths in the graph G_H.

Bibliography

[1] H. Ardal, D. Gunderson, V. Jungić, B. Landman and K. Williamson, Ramsey results involving the Fibonacci numbers, *Fibonacci Q.*, **46/47** (2008/2009), 10–17.

[2] T. Brown, R. Graham and B. Landman, On the set of common differences in van der Waerden's theorem on arithmetic progressions, *Can. Math. Bull.*, **42** (1999), 25–36.

[3] B. de Mathan, Numbers contravening a condition in density modulo 1, *Acta Math. Acad. Sci. Hungar.*, **36**(3–4) (1980), 237–241.

[4] R. Eggleton, P. Erdős and P. Skilton, Colouring the real line, *J. Comb. Theory, Ser. B*, **39** (1985), 86–100.

[5] P. Erdős, Problems and results on Diophantine approximations. II, in *Répartition modulo 1 (Actes Colloq., Marseille-Luminy, 1974)*, pp. 89–99, Lecture Notes in Math., vol. **475**, 1975.

[6] S. Farhangi, *On Refinements of van der Waerden's Theorem*, Masters thesis, Virginia Polytechnic Institute and State University, 2016.

[7] H. Furstenberg, Recurrence in ergodic theory and combinatorial number theory, in *M. B. Porter Lectures*, Princeton University Press, 1981.

[8] J. Guerreiro, I. Z. Ruzsa and M. Silva, Monochromatic paths for the integers, *Eur. J. Comb.*, **58** (2016), 283–288.

[9] B. Host, B. Kra and A. Maass, Variations on topological recurrence, *Monatshefte Math.*, **179** (2016), 57–89.

[10] V. Jungić, On a conjecture of Brown concerning accessible sets, *J. Comb. Theory, Ser. A*, **110** (2005), 175–178.

[11] Y. Katznelson, Chromatic numbers of Cayley graphs on $\mathbb{Z}$ and recurrence, in *Paul Erdős and his mathematics (Budapest, 1999)*, pp. 211–219, vol. **21**, 2001.

[12] B. L. Landman and A. Robertson, *Ramsey Theory on the Integers*, 2nd ed., Student Mathematical Library, vol. **73**, The American Mathematical Society, 2014.

[13] Y. Peres and W. Schlag, Two Erdős problems on lacunary sequences: chromatic number and Diophantine approximation, *Bull. Lond. Math. Soc.*, **42** (2010), 295–300.

[14] A. D. Pollington, On the density of sequence $\{n_k\xi\}$, *Illinois J. Math.*, **23**(4) (1979), 511–515.

[15] I. Z. Ruzsa, Zs. Tuza and M. Voigt, Distance graphs with finite chromatic number, *J. Comb. Theory, Ser. B*, **85** (2002), 181–187.

[16] A. Soifer, The Mathematical Coloring Book: Mathematics of Coloring and the Colorful Life of Its Creators, Springer, New York, 2008.

[17] B. L. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wiskd.*, **15** (1927), 212–216.

Michael Filaseta and Jacob Juillerat

Consecutive primes which are widely digitally delicate

Dedicated to the fond memory of Ronald Graham

Abstract: We show that for every positive integer k, there exist k consecutive primes having the property that if any digit of any one of the primes, including any of the infinitely many leading zero digits, is changed, then that prime becomes composite.

1 Introduction

In 1978, M. S. Klamkin [20] posed the following problem:

> Does there exist any prime number such that if any digit (in base 10) is changed to any other digit, the resulting number is always composite?

In addition to computations establishing the existence of such a prime, the published solutions in 1979 to this problem included a proof by P. Erdős [7] that there exist infinitely many such primes. Borrowing the terminology from J. Hopper and P. Pollack [15], we call such primes *digitally delicate*. The first digitally delicate prime is 294001. Thus 294001 is a prime and, for every $d \in \{0, 1, \dots, 9\}$, each of the numbers

$$d94001,\ 2d4001,\ 29d001,\ 294d01,\ 2940d1,\ 29400d$$

is either equal to 294001 or composite. The proof provided by Erdős consisted of creating a partial covering system of the integers (defined in the next section) followed by a sieve argument. In 2011, T. Tao [32] showed by refining the sieve argument of Erdős that a positive proportion (in terms of asymptotic density) of the primes are digitally delicate. In 2013, S. Konyagin [22] pointed out that a similar approach implies that a positive proportion of composite numbers n, coprime to 10, satisfy the property that if any digit in the base 10 representation of n is changed, then the resulting number remains composite. For example, the number $n = 212159$ satisfies this property. Thus,

Acknowledgement: The authors express their gratitude for the anonymous referee for providing the references [3] and [13].

Michael Filaseta, Department of Mathematics, University of South Carolina, Columbia, SC, USA, e-mail: filaseta@math.sc.edu
Jacob Juillerat, Department of Mathematics, University of North Carolina at Pembroke, Pembroke, NC, USA, e-mail: jacob.juillerat@uncp.edu

https://doi.org/10.1515/9783110754216-012

every number in the set

$$\{d12159, 2d2159, 21d159, 212d59, 2121d9, 21215d : d \in \{0, 1, 2, \ldots, 9\}\}$$

is composite. Later, in 2016, J. Hopper and P. Pollack [15] resolved a question of Tao's on digitally delicate primes allowing for an arbitrary but fixed number of digit changes to the beginning and end of the prime. All of these results and their proofs hold for numbers written in an arbitrary base b rather than base 10, though the proof provided by Erdős [7] only addresses the argument in base 10.

In 2020, the first author and J. Southwick [12] showed that a positive proportion of primes p, are *widely digitally delicate*, which they define as having the property that if any digit of p, *including any one of the infinitely many leading zeros of* p, is replaced by any other digit, then the resulting number is composite. The proof was specific to base 10, though they elaborate on other bases for which the analogous argument produces a similar result, including for example base 31; however, it is not even clear whether widely digitally delicate primes exist in every base. Observe that the first digitally delicate prime, 294001, is not widely digitally delicate since 10294001 is prime. A specific example of a widely digitally delicate prime, which has 4030 digits, was provided by Jon Grantham in the comment section of the article [25]. Recently, the authors with J. Southwick [10] obtained a related result showing that there are infinitely many (not necessarily a positive proportion) of composite numbers n in base 10 such that when any digit is inserted in the decimal expansion of n, including between two of the infinitely many leading zeros of n and to the right of the units digit of n, the number n remains composite (see also [11]).

In this paper, we show the following.

Theorem 1. *For every positive integer k, there exist k consecutive primes all of which are widely digitally delicate.*

Let $\mathcal{P}$ be a set of primes. It is not difficult to see that if $\mathcal{P}$ has an asymptotic density of 1 in the set of primes, then there exist k consecutive primes in $\mathcal{P}$ for each $k \in \mathbb{Z}^+$. On the other hand, for every $\varepsilon \in (0, 1)$, there exists $\mathcal{P}$ having asymptotic density $1-\varepsilon$ in the set of primes such that there do not exist k consecutive primes in $\mathcal{P}$ for k sufficiently large (more precisely, for $k \geq 1/\varepsilon$). Thus, the prior results stated above are not sufficient to establish Theorem 1. The main difficulty in using the prior methods to obtain Theorem 1 is in the application of sieve techniques in the prior work. We want to bypass the use of sieve techniques and instead give complete covering systems to show that there is an arithmetic progression containing infinitely many primes such that every prime in the arithmetic progression is a widely digitally delicate prime. This then gives an alternative proof of the result in [12]. After that, the main driving force behind the proof of Theorem 1, work of D. Shiu [29], can be applied. D. Shiu [29] showed that in any arithmetic progression containing infinitely many primes (i. e., $an+b$ with $\gcd(a, b) = 1$ and $a > 0$) there are arbitrarily long strings of consecutive primes—appropriately coined

"Shiu strings" by T. Freiberg [13]. Thus, once we establish through covering systems that such an arithmetic progression exists where every prime in the arithmetic progression is widely digitally delicate, D. Shiu's result immediately applies to finish the proof of Theorem 1.

Our main focus in this paper is on the proof of Theorem 1. However, in part, this paper is to emphasize that the remarkable work of Shiu [29] provides for a nice application to a number of results established via covering systems. One can also take these applications further by looking at the strengthening of Shiu's work by W. D. Banks, T. Freiberg, and C. L. Turnage-Butterbaugh [3] and J. Maynard [24] (also see T. Freiberg [13]). To illustrate the application of Shiu's work in other context, we give some further examples before closing this Introduction.

A Riesel number is a positive odd integer k with the property that $k \cdot 2^n - 1$ is composite for all positive integers n. A Sierpiński number is a positive odd integer k with the property that $k \cdot 2^n + 1$ is composite for all nonnegative integers n. The existence of such k were established in [28] and [30], respectively, though the former is a rather direct consequence of P. Erdős's work in [6] and the latter is a somewhat less direct application of this same work, an observation made by A. Schinzel (cf. [9]). A Brier number is a number k which is simultaneously Riesel and Sierpiński, named after Eric Brier who first considered them (cf. [9]). The smallest known Brier number, discovered by Christophe Clavier in 2014 (see [31]), is

$$3316923598096294713661.$$

As is common with all these numbers, examples typically come from covering systems giving an arithmetic progression of examples. In particular, Clavier established that every number in the arithmetic progression

$$377021473959660125796259470411 0\, n + 3316923598096294713661, \quad n \in \mathbb{Z}^+ \cup \{0\}$$

is a Brier number. Since the numbers 377021473959660125796259470411 0 and 3316923598096294713661 are coprime, Shiu's theorem gives the following.

Theorem 2. *For every positive integer k, there exist k consecutive primes all of which are Brier numbers.*

Observe that as an immediate consequence the same result holds if Brier numbers are replaced by Riesel or Sierpiński numbers.

As another less obvious result to apply Shiu's theorem to, we recall a result of R. Graham [14] from 1964. He showed that there exist relatively prime positive integers a and b such that the recursive Fibonacci-like sequence

$$u_0 = a, \quad u_1 = b, \quad \text{and} \quad u_{n+1} = u_n + u_{n-1} \quad \text{for integers } n \geq 1, \tag{12.1}$$

consists entirely of composite numbers. The size of known values for admissible a and b have decreased over the years through the work of others including D. Knuth [21], J. W. Nicol [26], and M. Vsemirnov [33], the latter giving the smallest known such a and b (but notably the same number of digits as the a and b in [26]). The result has also been generalized to other recursions; see A Dubickas, A. Novikas, and J. Šiurys [5], D. Ismailescu, A. Ko, C. Lee, and J. Y. Park [18] and I. Lunev [23]. As the Graham result concludes with all u_i being composite, the initial elements of the sequence, a and b, are composite. However, there is still a sense in which one can apply Shiu's result. To be precise, the smallest known example given by Vsemirnov is done by taking

$$a = 106276436867 \quad \text{and} \quad b = 35256392432.$$

With u_j defined as above, one can check that each u_j is divisible by a prime from the set

$$\mathcal{P} = \{2, 3, 5, 7, 11, 17, 19, 23, 31, 41, 47, 61, 107, 181, 541, 1103, 2521\}.$$

Setting

$$N = \prod_{p \in \mathcal{P}} p = 182189589586035679089873123 0,$$

the value of a and b can be replaced by any integers a and b satisfying

$$a \equiv 106276436867 \pmod{N} \quad \text{and} \quad b \equiv 35256392432 \pmod{N}.$$

As $\gcd(106276436867, N) = 31$ and $\gcd(35256392432, N) = 2$, these congruences are equivalent to taking $a = 31a'$ and $b = 2b'$ where a' and b' are integers satisfying

$$a' \equiv 3428272157 \pmod{587708353503340900289913 30}$$

and

$$b' \equiv 17628196216 \pmod{910947947930178395449365615}.$$

As a direct application of D. Shiu's result, we have the following.

Theorem 3. *For every $k \in \mathbb{Z}^+$, there are k consecutive primes $p_1, p_2, \dots, p_k$ and k consecutive primes $q_1, q_2, \dots, q_k$ such that for any $i \in \{1, 2, \dots, k\}$, the numbers $a = 31p_i$ and $b = 2q_i$ satisfy* $\gcd(a, b) = 1$ *and have the property that the u_n defined by* (12.1) *are all composite.*

This latter result is not meant to be particularly significant but rather an indication that Shiu's work does provide information in cases where covering systems are used to form composite numbers.

Regarding open problems, given the recent excellent works surrounding the nonexistence of covering systems of particular forms (cf. [1, 2, 16, 17]), the authors are not convinced that widely digitally delicate primes exist in every base. Thus, a tantalizing question is whether they exist or whether a positive proportion of the primes in every base are widely digitally delicate. In the opposite direction, as noted in [12], Carl Pomerance has asked for an unconditional proof that there exist infinitely many primes which are not digitally delicate or which are not widely digitally delicate. For other open problems in this direction, see the end of the introductions in [10] and [12].

Before closing this Introduction, we note that Matt Parker has recently done an excellent presentation [27] on his Stand-up Maths YouTube channel of the material in this paper.

2 The first steps of the argument

As noted in the Introduction, to prove Theorem 1, the work of D. Shiu [29] implies that it suffices to obtain an arithmetic progression $An + B$, with A and B relatively prime positive integers, such that every prime in the arithmetic progression is widely digitally delicate. We will determine such an A and B by finding relatively prime positive integers A and B satisfying property $(*)$ given by

$(*)$ If $d \in \{-9, -8, \ldots, -1\} \cup \{1, 2, \ldots, 9\}$, then each number in the set

$$\mathcal{A}_d = \{An + B + d \cdot 10^k : n \in \mathbb{Z}^+, k \in \mathbb{Z}^+ \cup \{0\}\}$$

is composite.

As changing a digit of $An + B$, including any one of its infinitely many leading zero digits, corresponds to adding or subtracting one of the numbers $1, 2, \ldots, 9$ from a digit of $An + B$, we see that relatively prime positive integers A and B satisfying property $(*)$ also satisfy the property we want, that every prime in $An + B$ is widely digitally delicate.

To find relatively prime positive integers A and B satisfying property $(*)$, we make use of covering systems which we define as follows.

Definition 1. A *covering system* (or *covering*) is a finite set of congruences

$$x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2}, \quad \ldots, \quad x \equiv a_r \pmod{m_r},$$

where $r \in \mathbb{Z}^+$, each $a_j \in \mathbb{Z}$, and each $m_j \in \mathbb{Z}^+$, such that every integer satisfies at least one congruence in the set of congruences.

In other contexts in the literature, further restrictions can be made on the m_j, so we emphasize here that we want to allow for $m_j = 1$ and for repeated moduli (so that

the m_j are not necessarily distinct). There will be restrictions on the m_j that will arise in the covering systems we build due to the approach we are using. We will see these as we proceed.

For each $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$, we will create a separate covering system to show that the elements of $\mathcal{A}_d$ in $(*)$ are composite. Table 12.1 indicates, for each d, the number of different congruences in the covering system corresponding to d.

Table 12.1: Number of congruences for each covering.

d	# cong.	d	# cong.	d	# cong.
−9	232	−3	739	4	26
−8	441	−2	289	5	1
−7	1	−1	1	6	19
−6	257	1	37	7	137
−5	268	2	1	8	1
−4	1	3	203	9	4

The integers we are covering for each d are the exponents k on 10 in the definition of $\mathcal{A}_d$. In other words, we will want to view each exponent k as satisfying one of the congruences in our covering system for a given $\mathcal{A}_d$. In the end, the values of A and B will be determined by the congruences we choose for the covering systems as well as certain primes that arise in our method.

We clarify that the work on digitally delicate primes in prior work mentioned in the Introduction used a partial covering of the integers k, that is a set of congruences where most but not all integers k satisfy at least one of the congruences, together with a sieve argument. The work in [12] on widely digitally delicate primes used covering systems for $d \in \{1, 2, \dots, 9\}$ and the same approach of partial coverings and sieves for $d \in \{-9, -8, \dots, -1\}$. The work in [10], like we will use in this paper, made use of covering systems for all $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$. For [10], some of the covering systems could be handled rather easily by taking advantage of the fact that we were looking for composite numbers satisfying a certain property rather than primes.

Next, we explain more precisely how we create and take advantage of a covering system for a given fixed $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$. We begin with a couple illustrative examples. Table 12.1 indicates that a number of the d are handled with just one congruence. This is accomplished by taking

$$A \equiv 0 \pmod 3 \quad \text{and} \quad B \equiv 1 \pmod 3.$$

Observe that each element of $\mathcal{A}_d$ in $(*)$ is divisible by 3 whenever $d \equiv 2 \pmod 3$. Thus, since A and B are positive, as long as we also have $B > 3$, the elements of $\mathcal{A}_d$ for such d are all composite, which is our goal. Note the crucial role of the order of 10 modulo the prime 3. The order is 1, and the covering system for each of these d is

simply $k \equiv 0 \pmod 1$. Every integer satisfies this congruence, so it is a covering system. The modulus corresponds to the order of 10 modulo 3. Note also that we cannot use the prime 3 in an analogous way to cover another digit d because the choices for A and B, and hence the congruences on A and B above, are to be independent of d. For example, if $d = 4$, then $An + B + d \cdot 10^k \equiv 1 + 4 \equiv 2 \pmod 3$, and hence, $An + B + d \cdot 10^k$ will not be divisible by 3.

As a second illustration, we see from Table 12.1 that we handle the digit $d = 9$ with 4 congruences. The congruences for $d = 9$ are

$$k \equiv 0 \pmod 2, \quad k \equiv 3 \pmod 4, \quad k \equiv 1 \pmod 8, \quad k \equiv 5 \pmod 8.$$

One easily checks that this is a covering system, i. e., that every integer k satisfies one of these congruences. To take advantage of this covering system, we choose a different prime p for each congruence with 10 having order modulo p equal to the modulus. We used the prime 11 with 10 of order 2, the prime 101 with 10 of order 4, the prime 73 with 10 of order 8, and the prime 137 with 10 of order 8. We take A divisible by each of these primes. For $(*)$, with $d = 9$, we want $An + B + 9 \cdot 10^k$ composite. For $k \equiv 0 \pmod 2$, we accomplish this by taking $B \equiv 2 \pmod{11}$ and $B > 11$ since then $An + B + 9 \cdot 10^k \equiv B + 9 \equiv 0 \pmod{11}$. For $k \equiv 3 \pmod 4$, we accomplish this by taking $B \equiv 90 \pmod{101}$ and $B > 101$ since then $An + B + 9 \cdot 10^k \equiv 90 + 9 \cdot 10^3 \equiv 9090 \equiv 0 \pmod{101}$. Similarly, for $k \equiv 1 \pmod 8$ and $B \equiv 56 \pmod{73}$, we obtain $An + B + 9 \cdot 10^k \equiv 0 \pmod{73}$; and for $k \equiv 5 \pmod 8$ and $B \equiv 90 \pmod{137}$, we obtain $An + B + 9 \cdot 10^k \equiv 0 \pmod{137}$. Thus, taking $B > 137$, we see that $(*)$ holds with $d = 9$.

Of some significance to our explanations later, we note that we could have interchanged the roles of the primes 73 and 137 since 10 has the same order for each of these primes. In other words, we could associate 137 with the congruence $k \equiv 1 \pmod 8$ above and associate 73 with the congruence $k \equiv 5 \pmod 8$. Then for $k \equiv 1 \pmod 8$ and $B \equiv 47 \pmod{137}$, we would have $An + B + 9 \cdot 10^k \equiv 0 \pmod{137}$; and for $k \equiv 5 \pmod 8$ and $B \equiv 17 \pmod{73}$, we would have $An + B + 9 \cdot 10^k \equiv 0 \pmod{73}$. In general, in our construction of widely digitally delicate primes, we want each congruence $k \equiv a \pmod m$ in a covering system associated with a prime p for which the order of 10 modulo p is m, but how we choose the ordering of those primes (which prime goes to which congruence) for a fixed modulus m is irrelevant.

For each $d \in \{-9, -8, \ldots, -1\} \cup \{1, 2, \ldots, 9\}$, we determine a covering system of congruences for k, where each modulus m corresponds to the order of 10 modulo some prime p. This imposes a condition on A, namely that A is divisible by each of these primes p. Fixing d, a congruence from our covering system $k \equiv a \pmod m$, and a corresponding prime p with 10 having order m modulo p, we determine B such that $An + B + d \cdot 10^k \equiv B + d \cdot 10^a \equiv 0 \pmod p$. Note that the values of d, a, and p dictate the congruence condition for B modulo p. Each prime p will correspond to a unique congruence condition $B \equiv -d \cdot 10^a \pmod p$, so the Chinese remainder theorem implies the

existence of a $B \in \mathbb{Z}^+$ simultaneously satisfying all the congruence conditions modulo primes on B. As long as B is large enough, then the condition $(*)$ will hold.

To make sure that there is a prime of the form $An+B$, we will want $\gcd(A, B) = 1$. For $k \equiv a \pmod{m}$ and a corresponding prime p as above, we will have A divisible by p and $B \equiv -d \cdot 10^a \pmod{p}$. Since $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$, if $p \geq 11$, then we see that $p \nmid B$. We will not be using the primes $p \in \{2, 5\}$ as 10 does not have an order modulo these primes. We have already seen that we are using the prime $p = 3$ for $d \equiv 2 \pmod{3}$, so this ensures that $3 \nmid B$. We will use $p = 7$ for $d \in \{-9, -8, -6, -5, -3, 3, 4\}$, which then implies $7 \nmid B$. Therefore, the condition $\gcd(A, B) = 1$ will hold.

Recall that we used the same congruence and corresponding prime in our covering system for each $d \equiv 2 \pmod{3}$. There is no obstacle to repeating a congruence for different d if the corresponding prime, having 10 of order the modulus, is different. But in the case of $d \equiv 2 \pmod{3}$, the same prime 3 was used for different d. To illustrate how we can repeat the use of a prime, we return to how we used the prime $p = 11$ above for $d = 9$. We ended up with $A \equiv 0 \pmod{11}$ and $B \equiv 2 \pmod{11}$. In order for us to take advantage of the prime $p = 11$ for d, we therefore want $An + B + d \cdot 10^k \equiv 2 + d \cdot 10^k \equiv 0 \pmod{11}$. It is easy to check that this holds for $(d, k) \in \{(-9, 1), (-2, 0), (2, 1), (9, 0)\}$. The case $(d, k) = (9, 0)$ is from our example with $d = 9$ above. The case $(d, k) = (2, 1)$ does not serve a purpose for us as $d = 2$ was covered by our earlier example using the prime 3 for all $d \equiv 2 \pmod{3}$. The cases where $(d, k) \in \{(-9, 1), (-2, 0)\}$ are significant, and we make use of congruences modulo 11 in the covering systems for $d = -9$ and $d = -2$. Thus, we are able to repeat the use of some primes for different values of d. However, this is not the case for most primes we used. A complete list of the primes which we were able to use for more than one value of d is given in Table 12.2, together with the list of corresponding d's. The function $\rho(m, p)$ in this table will be explained in the next section.

Recalling that the modulus in a covering system is equal to the order of 10 modulo a prime p, the role of primes and the order of 10 modulo those primes is significant in coming up with covering systems to deduce $(*)$. A modulus m can be used in a given covering system as many times as there are primes with 10 of order m. Thus, for the covering system for $d = 9$, we saw the modulus 8 being used twice as there are two primes with 10 of order 8, namely the primes 73 and 137. One can look at a list of primitive prime factors of $10^k - 1$ such as in [4], but we needed much more extensive data than what is contained there. Our approach uses that the complete list of primes for which 10 has a given order m is the same as the list of primes dividing $\Phi_m(10)$ and not dividing m where $\Phi_m(x)$ is the mth cyclotomic polynomial (cf. [4, 10, 12]). We used Magma V2.23-1 on a 2017 MacBook Pro to determine different primes dividing $\Phi_m(10)$. We did not always get a complete factorization but used that if the remaining unfactored part of $\Phi_m(10)$ is composite, relatively prime to the factored part of $\Phi_m(10)$ and m, and not a prime power, then there must be at least two further distinct prime factors of $\Phi_m(10)$. This allowed us then to determine a lower bound on the number

Table 12.2: Primes used for more than one digit d.

Prime	d's	$\rho(m,p)$	Prime	d's	$\rho(m,p)$
3	$-7,-4,-1,2,5,8$	1	199	$-6,-3,7$	1
7	$-9,-8,-6,-5,-3,3,4$	1	211	$-6,6$	1
11	$-9,-2,9$	1	241	$-6,6$	2
13	$-9,-3,3,4$	2	331	$-8,7$	1
17	$-8,-6,-3,-2,7$	1	353	$-6,7$	1
19	$-6,4$	1	409	$-8,-3$	1
23	$-9,-8,-6,-3,3,7$	1	449	$-9,7$	2
29	$-9,-8,-6,1,3$	1	2161	$-6,6$	3
31	$-8,-2,6$	1	3541	$-6,6$	1
37	$3,4$	1	9091	$-6,6$	1
43	$-8,-3,1$	1	27961	$-6,6$	2
53	$-8,-5,3$	1	1676321	$-6,6$	1
61	$-6,3,6$	1	3762091	$-6,6$	2
67	$-9,7$	1	4188901	$-6,6$	2
79	$-9,-5$	2	39526741	$-6,6$	3
89	$-6,-3,7$	1	5964848081	$-6,6$	2
103	$-9,-8,-3$	1			

of distinct primes of a given order m. Though we used most of these in our coverings, sometimes we found extra primes that we did not need to use.

In total, we made use of 673 different moduli m and 2596 different primes dividing $\Phi_m(10)$ for such m. Of the 2596 different primes, there are 590 which came from 295 composite numbers arising from an unfactored part of some $\Phi_m(10)$, and there are 63 other composite numbers for which only one prime factor of each of the composite numbers was used. The largest explicit prime (not coming from the $295 + 63 = 358$ composite numbers) has 1700 digits, arising from testing what was initially a large unfactored part of $\Phi_m(10)$ for primality and determining it is a prime. The largest of the 358 composite numbers has 17234 digits. For obvious reasons, we will avoid listing these primes and composites in this paper, though to help with verification of the results, we are providing the data from our computations in [8]; more explicit tables can also be found in [19].

Table 12.4 in the Appendix gives, for each of the 673 different moduli m, the detailed information on the number of distinct primes we used with 10 of order m, which we denote by $L(m)$. Thus $L(m)$ is a lower bound on the total number of distinct primes with 10 of order m. Note that $L(m)$ is less than or equal to the number of distinct primes dividing $\Phi_m(10)$ but not dividing m.

For each $d \in \{-9,-8,\dots,-1\} \cup \{1,2,\dots,9\}$, the goal is to find a covering system so that $(*)$ holds. We have already given the covering systems we obtained for $d \equiv 2 \pmod 3$ and for $d = 9$. In the next section and the Appendix, we elaborate on the covering systems for the remaining d. We also explain how the reader can verify the data showing these covering systems satisfy the conditions needed for $(*)$.

3 Finishing the argument

To finish the argument, we need to present a covering system for each value of d in $\{-9,-8,\dots,-1\}\cup\{1,2,\dots,9\}$ as described in the previous section. For the purposes of keeping the presentation of these covering systems manageable, for each m listed in Table 12.4, we take the $L(m)$ primes we found with 10 of order m and order them in some way. Corresponding to the discussion concerning $d = 9$ and the primes 73 and 137, the particular ordering is not important to us (e. g., increasing order would be fine). Suppose the primes corresponding to m are ordered in some way as $p_1, p_2, \dots, p_{L(m)}$. We define $\rho(p_j, m) = j$. Thus, if p_j is the jth prime in our ordering of the primes with 10 of order m, we have $\rho(p_j, m) = j$. The particular values we used for the $\rho(p, m)$ are not important to the arguments. So as to make the entries in Table 12.2 correct, the entries for $\rho(p, m)$ indicate the values we used for those primes. For example, Table 12.4 indicates there are 2 primes with 10 of order 6. One of them is 7. Table 12.2 indicates then that $\rho(7, 6) = 1$. Thus we put 7 as the first of the 2 primes with 10 of order 6. The other prime with 10 of order 6 is 13, and as Table 12.2 indicates we set 13 as the second of the 2 primes with 10 of order 6.

Tables 12.5–12.16 give the covering systems used for each $d \in \{-9,-8,\dots,-1\}\cup\{1,2,\dots,9\}$ with $d \not\equiv 2 \pmod 3$. Rather than indicating the prime, which in some cases has thousands of digits, corresponding to each congruence $k \equiv a \pmod m$ listed, we simply wrote the value of $\rho(m, p)$. As m corresponds to the modulus used in the given congruence $k \equiv a \pmod m$ and the ordering of the primes is not significant to our arguments (any ordering will do), this is enough information to confirm the covering arguments.

That said, the time consuming task of coming up with the $L(m)$ primes to order for each m is nontrivial (at least at this point in time). So that this work does not need to be repeated, a complete list of the $L(m)$ primes for each m is given in [8]. Further, the tables in the form of lists can be found there as well, with the third column in each case replaced by the prime we used with 10 of order the modulus of the congruence in the second column. In the way of clarity, recall that the primes were not explicitly computed in the case that the unfactored part of $\Phi_m(10)$ was determined to be composite; instead the composite number is listed in place of both primes in [8].

For the remainder of this section, we clarify how to verify the information in Tables 12.5–12.16. We address both verification of the covering systems and the information on the primes as listed in [8].

3.1 Covering verification

The most direct way to check that a system $\mathcal{C}$ of congruences

$$x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2}, \quad \dots, \quad x \equiv a_s \pmod{m_s}$$

is a covering system is to set $\ell = \operatorname{lcm}(m_1, m_2, \ldots, m_s)$ and then to check if every integer in the interval $[0, \ell - 1]$ satisfies at least one congruence in $\mathcal{C}$. If not, then $\mathcal{C}$ is not a covering system. If on the other hand, every integer in $[0, \ell - 1]$ satisfies a congruence in $\mathcal{C}$, then $\mathcal{C}$ is a covering system. To see the latter, let n be an arbitrary integer, and write $n = \ell q + r$ where q and r are integers with $0 \le r \le \ell - 1$. Since $r \in [0, \ell - 1]$ satisfies some $x \equiv a_j \pmod{m_j}$ and since $\ell \equiv 0 \pmod{m_j}$, we deduce for this same j that $n = \ell q + r \equiv a_j \pmod{m_j}$.

The above is a satisfactory approach if ℓ is not too large. For the values of d in $\{-9, -8, \ldots, -1\} \cup \{1, 2, \ldots, 9\}$ with $d \not\equiv 2 \pmod 3$, the least common multiple ℓ given by the congruences in Tables 12.5–12.16 are listed in Table 12.3. The maximum prime divisor of ℓ is also listed in the fourth column of Table 12.3. The value of ℓ can exceed 10^{12}, so we found a more efficient way to test whether one of our systems $\mathcal{C}$ of congruences, where ℓ is large, is a covering system.

Table 12.3: Least common multiple of the moduli for the coverings in each table.

d	Table	ℓ	max p	d	Table	ℓ	max p
−9	5	14433138720	31	1	11	5040	7
−8	6	699847948800	17	3	12	133333200	37
−6	7	1045044000	29	4	13	1296	3
−5	8	56216160	13	6	14	360	5
−3	9	1486147703040	19	7	15	18295200	11
−2	10	321253732800	23	9	16	8	2

Suppose $\ell > 10^6$ in Table 12.3 and the corresponding collection of congruences coming from the table indicated in the second column is $\mathcal{C}$. Let q be the largest prime divisor of ℓ as indicated in the fourth column. Let $w = 4 \cdot 3 \cdot 5 \cdot q$. This choice of w was selected on the basis of some trial and error; other choices are certainly reasonable. We do however want and have that w divides ℓ. Based on the comments above, we would like to know if every integer in the interval $[0, \ell - 1]$ satisfies at least one congruence in $\mathcal{C}$. The basic idea is to take each $u \in [0, w - 1]$ and to consider the integers that are congruent to u modulo w in $[0, \ell - 1]$. One advantage of doing this is that not every congruence in $\mathcal{C}$ needs to be considered. For example, take $d = -3$. Then Table 12.3 indicates $\ell = 1486147703040$ and Table 12.1 indicates the number of congruences in $\mathcal{C}$ is 739. From Table 12.9, the first few of the congruences in $\mathcal{C}$ are

$$k \equiv 4 \pmod 6, \quad k \equiv 5 \pmod 6, \quad k \equiv 0 \pmod{16}, \quad k \equiv 11 \pmod{21}.$$

Here, $w = 4 \cdot 3 \cdot 5 \cdot 19 = 1140$. If we take $u = 0$, then only the third of these congruences can be satisfied by an integer k congruent to u modulo w, as each of the other ones requires $k \not\equiv 0 \pmod 3$ whereas $k \equiv u \pmod w$ requires $k \equiv 0 \pmod 3$. Let $\mathcal{C}'$ be the congruences in $\mathcal{C}$ which are consistent with $k \equiv u \pmod w$. One can determine

these congruences by using that there exist integers satisfying both $k \equiv a \pmod{m}$ and $k \equiv u \pmod{w}$ if and only if $a \equiv u \pmod{\gcd(m,w)}$.

Observe that, with $u \in [0, w-1]$ fixed, we would like to know if each integer v of the form

$$v = wt + u, \quad \text{with } 0 \le t \le (\ell/w) - 1 \tag{12.2}$$

satisfies at least one congruence in $\mathcal{C}'$. The main advantage of this approach is that, as we shall now see, not all ℓ/w values of t need to be considered. First, we note that if $\mathcal{C}'$ is the empty set, then the integers in (12.2) are not covered and, therefore, $\mathcal{C}$ is not a covering system. Suppose then that $|\mathcal{C}'| \ge 1$. Let ℓ' denote the least common multiple of the moduli in $\mathcal{C}'$. Let $\delta = \gcd(w, \ell')$. We claim that we need only consider $v = wt + u$ where $0 \le t \le (\ell'/\delta) - 1$. To see this, suppose we know that every $v = wt + u$ with $0 \le t \le (\ell'/\delta) - 1$ satisfies one of the congruences in $\mathcal{C}'$. There are integers q, q', r and r' satisfying $t = \ell' q' + r'$ where $0 \le r' \le \ell' - 1$ and $r' = (\ell'/\delta)q + r$, where $0 \le r \le (\ell'/\delta) - 1$. Then

$$v = wt + u = w\ell' q' + wr' + u = w\ell' q' + (w/\delta)\ell' q + wr + u.$$

The definition of δ implies that $w/\delta \in \mathbb{Z}$. As each modulus in $\mathcal{C}'$ divides ℓ', we see that v satisfies a congruence in $\mathcal{C}'$ if and only if $wr + u$ does. Here, w and u are fixed and $0 \le r \le (\ell'/\delta) - 1$. Thus, we see that for each $u \in [0, w-1]$, we can restrict to determining whether v in (12.2) satisfies a congruence in $\mathcal{C}'$ for $0 \le t \le (\ell'/\delta) - 1$. Returning to the example of $d = -3$, $\ell = 1486147703040$ and $|\mathcal{C}| = 739$, where $w = 1140$ and we considered $u = 0$, one can check that $|\mathcal{C}'| = 19$, $\ell' = 12640320$, $\delta = w$ and $\ell'/\delta = 11088$. Thus what started out as ominously checking whether over 10^{12} integers each satisfy at least one of 739 different congruences is reduced in the case of $u = 0$ to looking at whether 11088 integers each satisfy at least one of 19 different congruences. As $u \in [0, w-1]$ varies, the number of computations does as well. An extreme case for $d = -3$ occurs for $u = 75$, where we get $\ell'/\delta = 14325696$ and $|\mathcal{C}'| = 47$. As d and u vary, though, this computation becomes manageable for determining that we have covering systems for each d in $\{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$ with $d \not\equiv 2 \pmod 3$ and $\ell > 10^6$. On a 2017 MacBook Plus running Maple 2019 with a 2.3 GHz Dual-Core Intel Core i5 processor, the total cpu time for determining the systems of congruences in Tables 12.5–12.16 are all covering systems took approximately 2.9 cpu hours, with almost all of this time spent on the case $d = -3$, which took 2.7 hours. The largest value of ℓ'/δ encountered was $\ell'/\delta = 14325696$, which occurred precisely for $d = -3$ and $u \in \{75, 303, 531, 759, 987\}$.

3.2 Data check

The most cumbersome task for us was the determination of the data in Table 12.4. As noted earlier, although the reader can check the data there directly, we have made

the list of primes corresponding to each m available through [8]. With the list of such primes for each m, it is still worth indicating how the data can be checked. Recall, in particular, the list of primes is not explicit in the case that there was an unfactored part of $\Phi_m(10)$. In this subsection, we elaborate on what checks should be and were done. All computations below were done with the MacBook Pro mentioned at the end of the last subsection and using Magma V2.23-1.

For each modulus m used in our constructions (listed in Table 12.4), we made a list of primes $p_1, p_2, \ldots, p_s$, written in increasing order, together with up to two additional primes q_1 and q_2, included after p_s on the list but not written explicitly (as we will discuss). Each prime came from a factorization or partial factorization of $\Phi_m(10)$. The primes $p_1, p_2, \ldots, p_s$ are the distinct primes appearing in the factored part of $\Phi_m(10)$, and as noted earlier do not include primes dividing m. In some cases, a complete factorization was found for $\Phi_m(10)$. For such m, there are no additional primes q_1 and q_2. If $\Phi_m(10)$ had an unfactored part $Q > 1$ (already tested to be composite), then we checked that Q is relatively prime to $mp_1p_2\cdots p_s$ and that Q is not of the form N^k with $N \in \mathbb{Z}^+$ and k an integer greater than or equal to 2. As this was always the case for the Q tested, we knew each such Q had two distinct prime factors q_1 and q_2. We deduce that there are at least two more primes q_j, $j \in \{1, 2\}$, different from $p_1, p_2, \ldots, p_s$ for which 10 has order m modulo q_j. As the data only contains the primes used in the covering systems, we only included the primes q_1 and q_2 that were used. Thus, despite Q having at least two distinct prime divisors, we may have listed anywhere from 0 to 2 of them. The question arises, however, as to how one can list primes that we do not know; there are primes q_1 and q_2 dividing Q, but we were unable to (or chose not to) factor Q to determine them explicitly. Instead of listing q_1 and q_2 then, we opted to list Q. Thus, for each m, we associated a list of one of the forms

$$[p_1, p_2, \ldots, p_s],\ [p_1, p_2, \ldots, p_s, Q],\ [p_1, p_2, \ldots, p_s, Q, Q],$$

depending on whether Q either did not exist or we used no prime factor of Q, we used one prime factor of Q, or we used two prime factors of Q, respectively. It is possible that $s = 0$; for example, the lists associated with the moduli 2888 and 2976 each take the middle form with no p_j and one composite number.

For a fixed m, given such a list, say from [8], one merely needs to check the following:

- Each element of the list divides $\Phi_m(10)$.
- Each element of the list is relatively prime to m.
- There is at most one composite number, say $Q > 1$, in the list, which may appear at most twice. The other numbers in the list are distinct primes.
- If the composite number Q exists, then $\gcd(Q, p_1p_2\cdots p_s) = 1$.
- If the composite number Q exists twice, then $Q^{1/k} \notin \mathbb{Z}^+$ for every integer $k \in [2, \log(Q)/\log(2)]$.

The upper bond in the last item above is simply because $k > \log(Q)/\log(2)$ implies $1 < Q^{1/k} < 2$, and hence $Q^{1/k}$ is not an integer. For each m, the value of $L(m)$ in Table 12.4 is simply the number of elements in the list associated with m.

With the data from the tables in the Appendix, also available in [8] with the indicated primes $p_1, p_2, \dots, p_s, q_1, q_2$ depending on m as above, some further details need to be checked to fully justify the computations. We verified that whenever m is used as a modulus in a table, it was associated with one of the primes dividing $\Phi_m(10)$. Furthermore, for any given $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$, the complete list of primes used as the congruences vary are distinct, noting that q_1 and q_2, for a given m, will be denoted by the same number Q but represent two distinct prime divisors of Q. As d varies, a given modulus m and a prime p dividing $\Phi_m(10)$ can be used more than once as indicated in Table 12.2. To elaborate, suppose such an m and p is used for each $d \in \mathcal{D} \subseteq \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$. For each $d \in \mathcal{D}$, then there corresponds a congruence $k \equiv a \pmod m$, where $a = a(d)$ will depend on d, as well as m and p. As noted earlier, this is permissible if and only if the values of $d \cdot 10^{a(d)}$ are congruent modulo p for all $d \in \mathcal{D}$. Thus, for each p that occurs in more than one table, as in Table 12.2, a check is done to verify the corresponding values of $d \cdot 10^{a(d)}$ are congruent modulo p.

The verification of the covering systems needed for Theorem 1 is complete, and the work of D. Shiu [29] now implies the theorem.

Appendix A

This Appendix begins with Table 12.4[1], giving a lower bound $L(m)$ on the number of distinct prime divisors of $\Phi_m(10)$ coprime to m. The m correspond to moduli in our coverings, and $L(m)$ is a lower bound on the number of primes p with 10 of order m modulo p. After Table 12.4, the remaining tables give the congruences $k \equiv a \pmod m$ that form the covering systems we obtained for $d \in \{-9, -8, \dots, -1\} \cup \{1, 2, \dots, 9\}$ with $d \not\equiv 2 \pmod 3$. For each congruence, there is an associated prime coming from the primes listed in Table 12.4 and that prime is tabulated in the second columns of Tables 12.5–12.16 (using the notation $\rho(m, p)$ discussed earlier in this paper).

Table 12.4: Number of primes used, $L = L(m)$, with 10 of order m.

m	L	m	L	m	L	m	L	m	L	m	L
1	1	56	2	132	3	234	3	361	7	520	2
2	1	57	3	133	3	238	3	363	3	522	4
3	1	58	2	135	5	240	3	364	3	527	4
4	1	60	3	136	2	242	5	368	2	528	3

1 The ordering of data in the tables below has been altered by the publisher. See the article in INTEGERS or [8] for the original lists of data.

Table 12.4 (continued)

m	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*
5	2	62	1	138	3	247	5	370	2	532	5
6	2	63	3	140	5	248	5	372	3	540	7
7	2	64	4	143	4	250	2	374	5	544	3
8	2	65	2	144	2	252	2	377	4	546	2
9	1	66	2	145	4	253	5	380	8	552	4
10	1	68	3	148	6	255	4	384	5	555	6
11	2	69	3	152	4	259	5	390	3	561	4
12	1	70	2	153	6	260	5	391	4	570	4
13	3	72	3	154	6	261	8	396	3	572	5
14	1	74	3	155	3	264	7	399	5	575	6
15	2	75	3	156	3	266	3	403	5	578	5
16	2	76	2	161	5	270	4	406	5	580	4
17	2	77	4	162	4	272	2	407	6	589	2
18	2	78	4	165	3	273	5	408	5	592	2
19	1	80	2	168	3	275	5	414	5	594	5
20	2	81	5	169	3	276	5	416	5	595	4
21	3	84	2	170	3	280	2	418	4	598	3
22	3	85	3	171	3	285	3	420	4	605	2
23	1	87	3	174	3	286	5	425	3	608	3
24	1	88	2	175	3	289	3	429	3	609	3
25	3	90	2	176	2	290	6	432	2	612	6
26	2	91	7	180	3	296	3	434	5	620	2
27	2	92	3	182	2	297	6	435	4	621	4
28	3	93	1	184	2	299	3	437	5	624	2
29	5	95	5	185	4	304	3	440	2	627	4
30	3	96	4	186	4	306	6	442	5	630	5
31	3	99	4	187	3	310	4	444	4	638	5
32	2	100	4	190	3	312	5	455	5	644	3
33	2	102	2	192	4	315	3	456	5	646	7
34	3	104	2	195	3	319	4	459	5	651	3
35	3	105	3	198	2	322	6	460	8	660	8
36	1	108	3	203	4	323	3	462	3	663	3
37	3	110	4	204	6	330	3	464	7	665	3
38	1	111	3	207	3	333	4	465	5	666	4
39	1	112	2	208	6	336	3	476	5	672	6
40	2	114	2	209	5	338	4	480	3	676	3
42	3	115	5	210	3	340	5	483	7	680	4
44	2	116	4	216	2	341	5	484	6	682	6
45	2	117	4	217	3	342	5	494	4	684	4
46	4	119	4	220	6	345	6	495	6	690	5
48	1	120	1	221	3	348	5	496	4	693	6
50	3	121	4	222	4	350	2	506	3	696	3
51	4	124	2	228	7	351	4	507	7	702	4
52	2	125	4	230	2	352	4	510	5	704	4
54	2	126	2	231	6	357	6	513	2	714	2

Table 12.4 (continued)

m	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*
55	4	130	2	232	5	360	3	518	2	715	3
720	2	969	4	1288	7	1767	4	2420	6	3468	4
722	3	988	3	1292	3	1768	4	2432	3	3480	3
726	4	990	4	1302	4	1776	4	2442	3	3496	6
728	2	992	4	1311	3	1785	2	2448	2	3534	4
740	10	1001	4	1320	8	1794	6	2484	2	3549	3
741	5	1012	3	1326	2	1805	5	2508	2	3570	6
744	7	1014	4	1330	3	1824	5	2535	4	3627	5
748	4	1015	7	1332	6	1848	4	2550	2	3648	2
754	2	1020	7	1352	8	1850	5	2565	4	3696	4
759	3	1023	4	1368	4	1860	8	2576	3	3700	2
760	2	1026	4	1380	9	1862	1	2584	8	3720	2
765	2	1035	3	1386	3	1870	2	2601	4	3724	6
768	6	1036	2	1392	3	1885	3	2604	5	3740	4
777	5	1040	4	1395	5	1904	5	2622	6	3770	6
782	6	1044	6	1428	2	1932	4	2652	4	3808	2
792	2	1045	3	1430	4	1938	5	2660	2	3876	4
798	3	1054	2	1440	4	1953	6	2664	4	3960	6
805	4	1056	3	1444	4	1976	4	2704	4	3990	2
806	3	1064	3	1445	6	1980	8	2736	3	4004	2
812	6	1083	7	1452	1	1995	4	2775	7	4046	3
814	5	1085	3	1456	6	2002	5	2790	1	4060	6
816	4	1088	6	1480	5	2023	3	2793	4	4080	4
828	2	1104	7	1482	2	2024	2	2888	1	4104	2
833	5	1105	4	1488	3	2030	4	2890	3	4180	6
836	3	1110	2	1496	4	2040	5	2904	2	4224	2
840	5	1122	2	1508	4	2046	3	2907	3	4256	4
845	2	1131	4	1521	1	2052	1	2912	3	4332	4
850	1	1140	11	1530	6	2070	6	2960	5	4352	3
858	5	1150	2	1547	3	2090	3	2964	3	4356	3
867	8	1156	5	1554	6	2108	1	2976	1	4370	5
868	6	1160	2	1564	5	2112	5	3003	3	4416	6
870	7	1173	4	1581	6	2128	4	3042	4	4420	6
874	5	1178	5	1584	2	2166	5	3060	7	4440	5
880	5	1183	4	1596	4	2170	5	3094	4	4560	4
884	4	1188	2	1610	6	2176	2	3108	1	4641	2
888	4	1190	6	1612	5	2185	5	3128	3	4692	2
897	3	1196	7	1615	2	2208	7	3135	9	4752	3
910	5	1197	3	1632	4	2210	3	3162	2	4788	4
912	2	1209	3	1665	3	2220	5	3179	5	4836	5
918	4	1210	3	1666	4	2262	2	3192	3	4864	4
920	2	1216	9	1672	4	2280	4	3230	3	4896	4
924	5	1221	6	1680	4	2300	4	3249	3	5005	3
925	9	1224	5	1690	4	2312	5	3255	4	5016	6
928	4	1235	5	1700	6	2346	4	3264	4	5070	2

Table 12.4 (continued)

m	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*	*m*	*L*
930	4	1240	6	1716	4	2356	3	3315	5	5130	2
931	7	1242	4	1734	6	2366	4	3330	6	5168	2
935	5	1254	5	1736	2	2380	4	3344	4	5202	4
952	6	1260	7	1740	9	2392	2	3380	4	5320	2
960	2	1275	4	1748	5	2394	5	3420	5	5328	4
966	3	1276	3	1760	2	2418	5	3432	2	5472	7
5544	3	6270	3	7068	5	9405	6	12540	4	25080	1
5550	6	6324	3	7254	2	9537	4	12716	5	25432	6
5586	4	6358	3	7392	4	9576	5	12996	8	25992	6
5776	2	6384	2	7448	7	9792	4	13056	2	30030	6
5780	10	6460	6	7752	5	10010	4	14508	4	37620	2
5808	4	6498	3	7980	10	10032	4	15015	3	51984	2
5814	3	6510	3	8008	4	10140	10	15960	4	60060	6
5928	2	6528	3	8092	3	10260	5	16184	4	75240	6
5985	3	6630	3	8208	4	10336	4	16720	2		
6006	6	6660	3	8360	3	10944	4	18810	3		
6069	2	6840	5	8664	3	11100	4	19074	2		
6188	3	6936	3	8704	2	11172	4	20064	4		
6256	3	6960	2	8880	3	12512	2	22344	4		

Table 12.5: Covering information for $d = -9$.

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 1 \pmod 2$	1	$k \equiv 44 \pmod{744}$	2	$k \equiv 80 \pmod{465}$	4
$k \equiv 0 \pmod 6$	2	$k \equiv 200 \pmod{744}$	4	$k \equiv 266 \pmod{465}$	5
$k \equiv 4 \pmod 6$	1	$k \equiv 386 \pmod{744}$	3	$k \equiv 452 \pmod{930}$	1
$k \equiv 6 \pmod{32}$	2	$k \equiv 572 \pmod{744}$	5	$k \equiv 638 \pmod{930}$	2
$k \equiv 26 \pmod{28}$	1	$k \equiv 728 \pmod{744}$	6	$k \equiv 824 \pmod{930}$	3
$k \equiv 16 \pmod{22}$	1	$k \equiv 170 \pmod{744}$	7	$k \equiv 50 \pmod{930}$	4
$k \equiv 26 \pmod{33}$	1	$k \equiv 414 \pmod{496}$	1	$k \equiv 236 \pmod{620}$	1
$k \equiv 9 \pmod{13}$	2	$k \equiv 446 \pmod{496}$	2	$k \equiv 546 \pmod{620}$	2
$k \equiv 0 \pmod{31}$	1	$k \equiv 478 \pmod{496}$	3	$k \equiv 112 \pmod{1240}$	1
$k \equiv 1 \pmod{31}$	2	$k \equiv 14 \pmod{496}$	4	$k \equiv 1042 \pmod{1240}$	2
$k \equiv 2 \pmod{31}$	3	$k \equiv 662 \pmod{992}$	1	$k \equiv 732 \pmod{1240}$	4
$k \equiv 34 \pmod{62}$	1	$k \equiv 694 \pmod{992}$	2	$k \equiv 422 \pmod{1240}$	3
$k \equiv 35 \pmod{93}$	1	$k \equiv 726 \pmod{992}$	3	$k \equiv 608 \pmod{1240}$	5
$k \equiv 98 \pmod{186}$	1	$k \equiv 758 \pmod{992}$	4	$k \equiv 1228 \pmod{1240}$	6
$k \equiv 68 \pmod{186}$	2	$k \equiv 356 \pmod{1488}$	1	$k \equiv 1538 \pmod{1860}$	1
$k \equiv 38 \pmod{186}$	3	$k \equiv 1100 \pmod{1488}$	2	$k \equiv 1724 \pmod{1860}$	2
$k \equiv 8 \pmod{186}$	4	$k \equiv 542 \pmod{1488}$	3	$k \equiv 794 \pmod{1860}$	4
$k \equiv 40 \pmod{124}$	1	$k \equiv 2774 \pmod{2976}$	1	$k \equiv 20 \pmod{1860}$	3
$k \equiv 102 \pmod{124}$	2	$k \equiv 140 \pmod{155}$	1	$k \equiv 950 \pmod{1860}$	5
$k \equiv 320 \pmod{372}$	1	$k \equiv 16 \pmod{155}$	2	$k \equiv 1136 \pmod{1860}$	6

Table 12.5 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 134 \pmod{372}$	2	$k \equiv 47 \pmod{155}$	3	$k \equiv 206 \pmod{1860}$	7
$k \equiv 104 \pmod{372}$	3	$k \equiv 78 \pmod{310}$	1	$k \equiv 392 \pmod{1860}$	8
$k \equiv 42 \pmod{248}$	1	$k \equiv 264 \pmod{310}$	2	$k \equiv 1322 \pmod{3720}$	1
$k \equiv 136 \pmod{248}$	2	$k \equiv 110 \pmod{310}$	3	$k \equiv 3182 \pmod{3720}$	2
$k \equiv 74 \pmod{248}$	3	$k \equiv 296 \pmod{310}$	4	$k \equiv 578 \pmod{1395}$	1
$k \equiv 12 \pmod{248}$	4	$k \equiv 17 \pmod{465}$	1	$k \equiv 113 \pmod{1395}$	2
$k \equiv 168 \pmod{248}$	5	$k \equiv 203 \pmod{465}$	2	$k \equiv 1043 \pmod{1395}$	3
$k \equiv 602 \pmod{744}$	1	$k \equiv 389 \pmod{465}$	3	$k \equiv 299 \pmod{1395}$	4
$k \equiv 1229 \pmod{1395}$	5	$k \equiv 1574 \pmod{1953}$	6	$k \equiv 6290 \pmod{7254}$	1
$k \equiv 764 \pmod{2790}$	1	$k \equiv 242 \pmod{341}$	1	$k \equiv 3314 \pmod{7254}$	2
$k \equiv 21 \pmod{217}$	1	$k \equiv 56 \pmod{341}$	2	$k \equiv 896 \pmod{14508}$	1
$k \equiv 176 \pmod{217}$	2	$k \equiv 211 \pmod{341}$	3	$k \equiv 8150 \pmod{14508}$	2
$k \equiv 114 \pmod{217}$	3	$k \equiv 25 \pmod{341}$	4	$k \equiv 5732 \pmod{14508}$	3
$k \equiv 52 \pmod{434}$	1	$k \equiv 149 \pmod{341}$	5	$k \equiv 12986 \pmod{14508}$	4
$k \equiv 424 \pmod{434}$	2	$k \equiv 304 \pmod{682}$	1	$k \equiv 10 \pmod{34}$	1
$k \equiv 300 \pmod{434}$	3	$k \equiv 118 \pmod{682}$	2	$k \equiv 153 \pmod{527}$	1
$k \equiv 84 \pmod{434}$	4	$k \equiv 614 \pmod{682}$	3	$k \equiv 494 \pmod{527}$	2
$k \equiv 22 \pmod{434}$	5	$k \equiv 428 \pmod{682}$	4	$k \equiv 308 \pmod{527}$	3
$k \equiv 611 \pmod{651}$	1	$k \equiv 88 \pmod{682}$	5	$k \equiv 122 \pmod{527}$	4
$k \equiv 332 \pmod{651}$	2	$k \equiv 584 \pmod{682}$	6	$k \equiv 990 \pmod{1054}$	1
$k \equiv 53 \pmod{651}$	3	$k \equiv 398 \pmod{1023}$	1	$k \equiv 804 \pmod{1054}$	2
$k \equiv 146 \pmod{1302}$	1	$k \equiv 212 \pmod{1023}$	2	$k \equiv 1145 \pmod{1581}$	1
$k \equiv 1232 \pmod{1302}$	2	$k \equiv 677 \pmod{1023}$	3	$k \equiv 959 \pmod{1581}$	2
$k \equiv 302 \pmod{1302}$	3	$k \equiv 491 \pmod{1023}$	4	$k \equiv 773 \pmod{1581}$	3
$k \equiv 674 \pmod{1302}$	4	$k \equiv 1328 \pmod{2046}$	1	$k \equiv 587 \pmod{1581}$	4
$k \equiv 796 \pmod{868}$	1	$k \equiv 1142 \pmod{2046}$	2	$k \equiv 215 \pmod{1581}$	5
$k \equiv 208 \pmod{868}$	2	$k \equiv 956 \pmod{2046}$	3	$k \equiv 29 \pmod{1581}$	6
$k \equiv 488 \pmod{868}$	3	$k \equiv 182 \pmod{403}$	1	$k \equiv 1424 \pmod{3162}$	1
$k \equiv 768 \pmod{868}$	4	$k \equiv 27 \pmod{403}$	2	$k \equiv 1238 \pmod{3162}$	2
$k \equiv 612 \pmod{868}$	5	$k \equiv 275 \pmod{403}$	3	$k \equiv 1052 \pmod{2108}$	1
$k \equiv 178 \pmod{868}$	6	$k \equiv 120 \pmod{403}$	4	$k \equiv 4214 \pmod{6324}$	1
$k \equiv 116 \pmod{2604}$	1	$k \equiv 368 \pmod{403}$	5	$k \equiv 4028 \pmod{6324}$	2
$k \equiv 1418 \pmod{2604}$	2	$k \equiv 616 \pmod{806}$	1	$k \equiv 866 \pmod{6324}$	3
$k \equiv 860 \pmod{2604}$	3	$k \equiv 58 \pmod{806}$	2	$k \equiv 247 \pmod{589}$	1
$k \equiv 2162 \pmod{2604}$	4	$k \equiv 306 \pmod{806}$	3	$k \equiv 495 \pmod{589}$	2
$k \equiv 644 \pmod{2604}$	5	$k \equiv 554 \pmod{1209}$	1	$k \equiv 154 \pmod{1178}$	1
$k \equiv 210 \pmod{1736}$	1	$k \equiv 647 \pmod{1209}$	2	$k \equiv 402 \pmod{1178}$	2
$k \equiv 1078 \pmod{1736}$	2	$k \equiv 89 \pmod{1209}$	3	$k \equiv 650 \pmod{1178}$	3
$k \equiv 365 \pmod{1085}$	1	$k \equiv 740 \pmod{2418}$	1	$k \equiv 898 \pmod{1178}$	4
$k \equiv 1016 \pmod{1085}$	2	$k \equiv 338 \pmod{2418}$	2	$k \equiv 1146 \pmod{1178}$	5
$k \equiv 582 \pmod{1085}$	3	$k \equiv 2198 \pmod{2418}$	3	$k \equiv 1394 \pmod{1767}$	1
$k \equiv 148 \pmod{2170}$	1	$k \equiv 1640 \pmod{2418}$	4	$k \equiv 464 \pmod{1767}$	2
$k \equiv 1884 \pmod{2170}$	2	$k \equiv 1082 \pmod{2418}$	5	$k \equiv 1301 \pmod{1767}$	3
$k \equiv 520 \pmod{2170}$	3	$k \equiv 524 \pmod{1612}$	1	$k \equiv 371 \pmod{1767}$	4
$k \equiv 86 \pmod{2170}$	4	$k \equiv 772 \pmod{1612}$	2	$k \equiv 1208 \pmod{3534}$	1
$k \equiv 1822 \pmod{2170}$	5	$k \equiv 1020 \pmod{1612}$	3	$k \equiv 278 \pmod{3534}$	2

Table 12.5 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 1388 \pmod{3255}$	1	$k \equiv 1268 \pmod{1612}$	4	$k \equiv 2882 \pmod{3534}$	3
$k \equiv 2039 \pmod{3255}$	2	$k \equiv 1516 \pmod{1612}$	5	$k \equiv 1952 \pmod{3534}$	4
$k \equiv 1760 \pmod{3255}$	3	$k \equiv 2942 \pmod{4836}$	1	$k \equiv 2200 \pmod{2356}$	1
$k \equiv 2411 \pmod{3255}$	4	$k \equiv 4802 \pmod{4836}$	2	$k \equiv 1022 \pmod{2356}$	2
$k \equiv 3062 \pmod{6510}$	1	$k \equiv 1826 \pmod{4836}$	3	$k \equiv 92 \pmod{2356}$	3
$k \equiv 458 \pmod{6510}$	2	$k \equiv 3686 \pmod{4836}$	4	$k \equiv 3626 \pmod{7068}$	1
$k \equiv 4364 \pmod{6510}$	3	$k \equiv 710 \pmod{4836}$	5	$k \equiv 2696 \pmod{7068}$	2
$k \equiv 830 \pmod{1953}$	1	$k \equiv 803 \pmod{3627}$	1	$k \equiv 6230 \pmod{7068}$	3
$k \equiv 1481 \pmod{1953}$	2	$k \equiv 2012 \pmod{3627}$	2	$k \equiv 5300 \pmod{7068}$	4
$k \equiv 179 \pmod{1953}$	3	$k \equiv 3221 \pmod{3627}$	3	$k \equiv 1766 \pmod{7068}$	5
$k \equiv 272 \pmod{1953}$	4	$k \equiv 245 \pmod{3627}$	4		
$k \equiv 923 \pmod{1953}$	5	$k \equiv 1454 \pmod{3627}$	5		

Table 12.6: Covering information for $d = -8$.

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 2 \pmod{13}$	1	$k \equiv 1569 \pmod{2176}$	1	$k \equiv 1454 \pmod{2040}$	5
$k \equiv 3 \pmod{21}$	1	$k \equiv 481 \pmod{2176}$	2	$k \equiv 111 \pmod{612}$	1
$k \equiv 19 \pmod{28}$	1	$k \equiv 73 \pmod{1088}$	1	$k \equiv 9 \pmod{612}$	2
$k \equiv 10 \pmod{22}$	1	$k \equiv 209 \pmod{1088}$	2	$k \equiv 519 \pmod{612}$	3
$k \equiv 9 \pmod{110}$	1	$k \equiv 345 \pmod{1088}$	3	$k \equiv 417 \pmod{612}$	4
$k \equiv 0 \pmod{6}$	1	$k \equiv 617 \pmod{1088}$	4	$k \equiv 315 \pmod{612}$	5
$k \equiv 11 \pmod{15}$	1	$k \equiv 753 \pmod{1088}$	5	$k \equiv 213 \pmod{612}$	6
$k \equiv 13 \pmod{16}$	1	$k \equiv 889 \pmod{1088}$	6	$k \equiv 94 \pmod{306}$	4
$k \equiv 0 \pmod{17}$	1	$k \equiv 56 \pmod{459}$	1	$k \equiv 298 \pmod{306}$	5
$k \equiv 1 \pmod{17}$	2	$k \equiv 413 \pmod{459}$	2	$k \equiv 196 \pmod{306}$	6
$k \equiv 2 \pmod{34}$	1	$k \equiv 311 \pmod{459}$	3	$k \equiv 77 \pmod{1224}$	1
$k \equiv 19 \pmod{34}$	2	$k \equiv 209 \pmod{459}$	4	$k \equiv 1097 \pmod{1224}$	2
$k \equiv 3 \pmod{34}$	3	$k \equiv 107 \pmod{459}$	5	$k \equiv 893 \pmod{1224}$	3
$k \equiv 88 \pmod{204}$	1	$k \equiv 464 \pmod{918}$	1	$k \equiv 689 \pmod{1224}$	4
$k \equiv 20 \pmod{204}$	2	$k \equiv 362 \pmod{918}$	2	$k \equiv 485 \pmod{1224}$	5
$k \equiv 190 \pmod{204}$	3	$k \equiv 260 \pmod{918}$	3	$k \equiv 1505 \pmod{2448}$	1
$k \equiv 122 \pmod{204}$	4	$k \equiv 158 \pmod{918}$	4	$k \equiv 281 \pmod{2448}$	2
$k \equiv 4 \pmod{68}$	1	$k \equiv 124 \pmod{204}$	5	$k \equiv 995 \pmod{1020}$	1
$k \equiv 21 \pmod{68}$	2	$k \equiv 22 \pmod{204}$	6	$k \equiv 791 \pmod{1020}$	2
$k \equiv 38 \pmod{68}$	3	$k \equiv 447 \pmod{1632}$	1	$k \equiv 587 \pmod{1020}$	3
$k \equiv 123 \pmod{136}$	1	$k \equiv 991 \pmod{1632}$	2	$k \equiv 383 \pmod{1020}$	4
$k \equiv 55 \pmod{136}$	2	$k \equiv 1535 \pmod{1632}$	3	$k \equiv 179 \pmod{1020}$	5
$k \equiv 5 \pmod{272}$	1	$k \equiv 40 \pmod{51}$	1	$k \equiv 10 \pmod{85}$	1
$k \equiv 243 \pmod{272}$	2	$k \equiv 23 \pmod{51}$	2	$k \equiv 61 \pmod{85}$	2
$k \equiv 651 \pmod{1632}$	4	$k \equiv 7 \pmod{51}$	3	$k \equiv 27 \pmod{85}$	3
$k \equiv 2827 \pmod{3264}$	1	$k \equiv 41 \pmod{51}$	4	$k \equiv 78 \pmod{170}$	1
$k \equiv 1195 \pmod{3264}$	2	$k \equiv 57 \pmod{102}$	1	$k \equiv 163 \pmod{170}$	2

Table 12.6 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 1739 \pmod{3264}$	3	$k \equiv 75 \pmod{102}$	2	$k \equiv 44 \pmod{170}$	3
$k \equiv 107 \pmod{3264}$	4	$k \equiv 127 \pmod{153}$	1	$k \equiv 129 \pmod{680}$	1
$k \equiv 3099 \pmod{6528}$	1	$k \equiv 110 \pmod{153}$	2	$k \equiv 299 \pmod{680}$	2
$k \equiv 1467 \pmod{6528}$	2	$k \equiv 76 \pmod{153}$	3	$k \equiv 469 \pmod{680}$	3
$k \equiv 6363 \pmod{6528}$	3	$k \equiv 59 \pmod{153}$	4	$k \equiv 639 \pmod{680}$	4
$k \equiv 4731 \pmod{13056}$	1	$k \equiv 25 \pmod{153}$	5	$k \equiv 215 \pmod{255}$	1
$k \equiv 11259 \pmod{13056}$	2	$k \equiv 8 \pmod{153}$	6	$k \equiv 62 \pmod{255}$	2
$k \equiv 379 \pmod{4896}$	1	$k \equiv 297 \pmod{306}$	1	$k \equiv 113 \pmod{255}$	3
$k \equiv 2011 \pmod{4896}$	2	$k \equiv 93 \pmod{306}$	2	$k \equiv 164 \pmod{255}$	4
$k \equiv 3643 \pmod{4896}$	3	$k \equiv 195 \pmod{306}$	3	$k \equiv 45 \pmod{510}$	1
$k \equiv 4187 \pmod{4896}$	4	$k \equiv 145 \pmod{408}$	1	$k \equiv 351 \pmod{510}$	2
$k \equiv 923 \pmod{9792}$	1	$k \equiv 43 \pmod{408}$	2	$k \equiv 147 \pmod{510}$	3
$k \equiv 7451 \pmod{9792}$	2	$k \equiv 349 \pmod{408}$	3	$k \equiv 453 \pmod{510}$	4
$k \equiv 5819 \pmod{9792}$	3	$k \equiv 247 \pmod{408}$	4	$k \equiv 249 \pmod{510}$	5
$k \equiv 2555 \pmod{9792}$	4	$k \equiv 128 \pmod{408}$	5	$k \equiv 640 \pmod{1020}$	6
$k \equiv 39 \pmod{544}$	1	$k \equiv 434 \pmod{816}$	1	$k \equiv 385 \pmod{1020}$	7
$k \equiv 311 \pmod{544}$	2	$k \equiv 26 \pmod{816}$	2	$k \equiv 130 \pmod{4080}$	2
$k \equiv 175 \pmod{544}$	3	$k \equiv 740 \pmod{816}$	3	$k \equiv 3190 \pmod{4080}$	1
$k \equiv 1025 \pmod{8704}$	1	$k \equiv 332 \pmod{816}$	4	$k \equiv 2170 \pmod{4080}$	3
$k \equiv 5377 \pmod{8704}$	2	$k \equiv 230 \pmod{2040}$	1	$k \equiv 1150 \pmod{4080}$	4
$k \equiv 2113 \pmod{4352}$	1	$k \equiv 1046 \pmod{2040}$	2	$k \equiv 2935 \pmod{3060}$	2
$k \equiv 3201 \pmod{4352}$	2	$k \equiv 1862 \pmod{2040}$	3	$k \equiv 895 \pmod{3060}$	1
$k \equiv 4289 \pmod{4352}$	3	$k \equiv 638 \pmod{2040}$	4	$k \equiv 1915 \pmod{3060}$	3
$k \equiv 181 \pmod{765}$	1	$k \equiv 693 \pmod{1904}$	3	$k \equiv 235 \pmod{374}$	4
$k \equiv 436 \pmod{765}$	2	$k \equiv 455 \pmod{1904}$	4	$k \equiv 65 \pmod{374}$	5
$k \equiv 1456 \pmod{1530}$	1	$k \equiv 1407 \pmod{1904}$	5	$k \equiv 456 \pmod{748}$	1
$k \equiv 691 \pmod{1530}$	2	$k \equiv 64 \pmod{357}$	1	$k \equiv 269 \pmod{748}$	2
$k \equiv 1252 \pmod{1530}$	4	$k \equiv 302 \pmod{357}$	2	$k \equiv 82 \pmod{748}$	3
$k \equiv 487 \pmod{1530}$	3	$k \equiv 268 \pmod{357}$	3	$k \equiv 643 \pmod{748}$	4
$k \equiv 742 \pmod{1530}$	5	$k \equiv 149 \pmod{357}$	4	$k \equiv 592 \pmod{1496}$	1
$k \equiv 1507 \pmod{1530}$	6	$k \equiv 115 \pmod{357}$	5	$k \equiv 218 \pmod{1496}$	2
$k \equiv 232 \pmod{3060}$	4	$k \equiv 353 \pmod{357}$	6	$k \equiv 1340 \pmod{1496}$	3
$k \equiv 997 \pmod{3060}$	6	$k \equiv 183 \pmod{714}$	1	$k \equiv 966 \pmod{1496}$	4
$k \equiv 1762 \pmod{3060}$	5	$k \equiv 387 \pmod{714}$	2	$k \equiv 405 \pmod{1870}$	1
$k \equiv 2527 \pmod{3060}$	7	$k \equiv 1305 \pmod{1428}$	1	$k \equiv 31 \pmod{1870}$	2
$k \equiv 28 \pmod{425}$	1	$k \equiv 591 \pmod{1428}$	2	$k \equiv 3397 \pmod{3740}$	1
$k \equiv 283 \pmod{425}$	2	$k \equiv 200 \pmod{595}$	1	$k \equiv 1527 \pmod{3740}$	2
$k \equiv 113 \pmod{425}$	3	$k \equiv 557 \pmod{595}$	2	$k \equiv 1153 \pmod{3740}$	3
$k \equiv 368 \pmod{1700}$	1	$k \equiv 438 \pmod{595}$	3	$k \equiv 3023 \pmod{3740}$	4
$k \equiv 793 \pmod{1700}$	2	$k \equiv 319 \pmod{595}$	4	$k \equiv 490 \pmod{935}$	1
$k \equiv 1218 \pmod{1700}$	3	$k \equiv 676 \pmod{1785}$	1	$k \equiv 116 \pmod{935}$	2
$k \equiv 1643 \pmod{1700}$	4	$k \equiv 81 \pmod{1785}$	2	$k \equiv 677 \pmod{935}$	3
$k \equiv 1048 \pmod{1700}$	5	$k \equiv 880 \pmod{1190}$	2	$k \equiv 303 \pmod{935}$	4
$k \equiv 198 \pmod{1700}$	6	$k \equiv 761 \pmod{1190}$	1	$k \equiv 864 \pmod{935}$	5
$k \equiv 623 \pmod{850}$	1	$k \equiv 642 \pmod{1190}$	3	$k \equiv 337 \pmod{561}$	1
$k \equiv 79 \pmod{1275}$	1	$k \equiv 523 \pmod{1190}$	4	$k \equiv 524 \pmod{561}$	2

Table 12.6 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 334 \pmod{1275}$	2	$k \equiv 404 \pmod{1190}$	5	$k \equiv 184 \pmod{561}$	3
$k \equiv 589 \pmod{1275}$	3	$k \equiv 285 \pmod{1190}$	6	$k \equiv 371 \pmod{561}$	4
$k \equiv 844 \pmod{1275}$	4	$k \equiv 166 \pmod{2380}$	1	$k \equiv 711 \pmod{1122}$	1
$k \equiv 2374 \pmod{2550}$	1	$k \equiv 1356 \pmod{2380}$	2	$k \equiv 1119 \pmod{1122}$	2
$k \equiv 1099 \pmod{2550}$	2	$k \equiv 47 \pmod{2380}$	3	$k \equiv 117 \pmod{221}$	1
$k \equiv 63 \pmod{119}$	1	$k \equiv 1237 \pmod{2380}$	4	$k \equiv 66 \pmod{221}$	2
$k \equiv 29 \pmod{119}$	2	$k \equiv 1118 \pmod{3570}$	1	$k \equiv 185 \pmod{221}$	3
$k \equiv 114 \pmod{119}$	3	$k \equiv 3498 \pmod{3570}$	2	$k \equiv 134 \pmod{442}$	1
$k \equiv 80 \pmod{119}$	4	$k \equiv 2308 \pmod{3570}$	3	$k \equiv 355 \pmod{442}$	2
$k \equiv 46 \pmod{238}$	1	$k \equiv 999 \pmod{3570}$	4	$k \equiv 304 \pmod{442}$	3
$k \equiv 165 \pmod{238}$	2	$k \equiv 3379 \pmod{3570}$	5	$k \equiv 83 \pmod{442}$	4
$k \equiv 12 \pmod{238}$	3	$k \equiv 2189 \pmod{3570}$	6	$k \equiv 32 \pmod{442}$	5
$k \equiv 131 \pmod{476}$	1	$k \equiv 251 \pmod{833}$	1	$k \equiv 1137 \pmod{1768}$	2
$k \equiv 216 \pmod{476}$	2	$k \equiv 13 \pmod{833}$	2	$k \equiv 1579 \pmod{1768}$	1
$k \equiv 97 \pmod{476}$	3	$k \equiv 608 \pmod{833}$	3	$k \equiv 253 \pmod{1768}$	3
$k \equiv 454 \pmod{476}$	4	$k \equiv 370 \pmod{833}$	4	$k \equiv 695 \pmod{1768}$	4
$k \equiv 335 \pmod{476}$	5	$k \equiv 132 \pmod{833}$	5	$k \equiv 644 \pmod{884}$	1
$k \equiv 1797 \pmod{1904}$	1	$k \equiv 1560 \pmod{1666}$	1	$k \equiv 865 \pmod{884}$	2
$k \equiv 1321 \pmod{1904}$	2	$k \equiv 727 \pmod{1666}$	2	$k \equiv 202 \pmod{884}$	3
$k \equiv 2273 \pmod{3808}$	1	$k \equiv 1322 \pmod{1666}$	3	$k \equiv 423 \pmod{884}$	4
$k \equiv 369 \pmod{3808}$	2	$k \equiv 489 \pmod{1666}$	4	$k \equiv 151 \pmod{663}$	1
$k \equiv 336 \pmod{952}$	1	$k \equiv 99 \pmod{187}$	1	$k \equiv 593 \pmod{663}$	2
$k \equiv 217 \pmod{952}$	2	$k \equiv 133 \pmod{187}$	2	$k \equiv 100 \pmod{663}$	3
$k \equiv 98 \pmod{952}$	3	$k \equiv 167 \pmod{187}$	3	$k \equiv 1035 \pmod{1326}$	1
$k \equiv 931 \pmod{952}$	4	$k \equiv 14 \pmod{374}$	1	$k \equiv 321 \pmod{1326}$	2
$k \equiv 812 \pmod{952}$	5	$k \equiv 201 \pmod{374}$	2	$k \equiv 1868 \pmod{2652}$	1
$k \equiv 574 \pmod{952}$	6	$k \equiv 48 \pmod{374}$	3	$k \equiv 1205 \pmod{2652}$	2
$k \equiv 542 \pmod{2652}$	3	$k \equiv 1699 \pmod{2312}$	5	$k \equiv 543 \pmod{6936}$	1
$k \equiv 2531 \pmod{2652}$	4	$k \equiv 696 \pmod{1156}$	1	$k \equiv 5167 \pmod{6936}$	2
$k \equiv 270 \pmod{1105}$	1	$k \equiv 985 \pmod{1156}$	2	$k \equiv 2855 \pmod{6936}$	3
$k \equiv 712 \pmod{1105}$	2	$k \equiv 118 \pmod{1156}$	3	$k \equiv 832 \pmod{3468}$	1
$k \equiv 933 \pmod{1105}$	3	$k \equiv 407 \pmod{1156}$	4	$k \equiv 2566 \pmod{3468}$	2
$k \equiv 49 \pmod{1105}$	4	$k \equiv 1121 \pmod{1156}$	5	$k \equiv 1988 \pmod{3468}$	3
$k \equiv 2701 \pmod{3315}$	1	$k \equiv 424 \pmod{867}$	1	$k \equiv 254 \pmod{3468}$	4
$k \equiv 1596 \pmod{3315}$	2	$k \equiv 713 \pmod{867}$	2	$k \equiv 560 \pmod{2023}$	1
$k \equiv 2871 \pmod{3315}$	3	$k \equiv 730 \pmod{867}$	3	$k \equiv 1716 \pmod{2023}$	2
$k \equiv 661 \pmod{3315}$	4	$k \equiv 152 \pmod{867}$	4	$k \equiv 849 \pmod{2023}$	3
$k \equiv 1324 \pmod{3315}$	5	$k \equiv 169 \pmod{867}$	6	$k \equiv 1138 \pmod{4046}$	1
$k \equiv 440 \pmod{2210}$	1	$k \equiv 458 \pmod{867}$	5	$k \equiv 3161 \pmod{4046}$	2
$k \equiv 1545 \pmod{2210}$	2	$k \equiv 475 \pmod{867}$	7	$k \equiv 3450 \pmod{4046}$	3
$k \equiv 882 \pmod{2210}$	3	$k \equiv 764 \pmod{867}$	8	$k \equiv 5473 \pmod{16184}$	1
$k \equiv 4197 \pmod{4420}$	1	$k \equiv 135 \pmod{1734}$	1	$k \equiv 1427 \pmod{16184}$	2
$k \equiv 1987 \pmod{4420}$	2	$k \equiv 441 \pmod{1734}$	2	$k \equiv 13565 \pmod{16184}$	3
$k \equiv 2208 \pmod{4420}$	5	$k \equiv 747 \pmod{1734}$	4	$k \equiv 9519 \pmod{16184}$	4
$k \equiv 3313 \pmod{4420}$	3	$k \equiv 1053 \pmod{1734}$	3	$k \equiv 6340 \pmod{8092}$	1
$k \equiv 4418 \pmod{4420}$	4	$k \equiv 1359 \pmod{1734}$	5	$k \equiv 4317 \pmod{8092}$	2

Table 12.6 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 1103 \pmod{4420}$	6	$k \equiv 1648 \pmod{1734}$	6	$k \equiv 2294 \pmod{8092}$	3
$k \equiv 219 \pmod{6630}$	1	$k \equiv 1648 \pmod{2601}$	1	$k \equiv 2005 \pmod{6069}$	1
$k \equiv 5744 \pmod{6630}$	2	$k \equiv 2515 \pmod{2601}$	2	$k \equiv 4028 \pmod{6069}$	2
$k \equiv 2429 \pmod{6630}$	3	$k \equiv 781 \pmod{2601}$	3	$k \equiv 1155 \pmod{3179}$	1
$k \equiv 168 \pmod{1547}$	1	$k \equiv 1937 \pmod{2601}$	4	$k \equiv 2311 \pmod{3179}$	2
$k \equiv 610 \pmod{1547}$	2	$k \equiv 2804 \pmod{5202}$	1	$k \equiv 288 \pmod{3179}$	3
$k \equiv 1052 \pmod{1547}$	3	$k \equiv 203 \pmod{5202}$	2	$k \equiv 1444 \pmod{3179}$	4
$k \equiv 1936 \pmod{3094}$	1	$k \equiv 1070 \pmod{5202}$	3	$k \equiv 2600 \pmod{3179}$	5
$k \equiv 389 \pmod{3094}$	2	$k \equiv 3671 \pmod{5202}$	4	$k \equiv 6357 \pmod{6358}$	1
$k \equiv 2820 \pmod{3094}$	3	$k \equiv 220 \pmod{1445}$	3	$k \equiv 3756 \pmod{6358}$	2
$k \equiv 1273 \pmod{3094}$	4	$k \equiv 1376 \pmod{1445}$	1	$k \equiv 577 \pmod{6358}$	3
$k \equiv 5472 \pmod{6188}$	1	$k \equiv 1087 \pmod{1445}$	2	$k \equiv 4912 \pmod{12716}$	1
$k \equiv 3925 \pmod{6188}$	2	$k \equiv 798 \pmod{1445}$	4	$k \equiv 1733 \pmod{12716}$	2
$k \equiv 2378 \pmod{6188}$	3	$k \equiv 509 \pmod{1445}$	5	$k \equiv 11270 \pmod{12716}$	3
$k \equiv 4588 \pmod{4641}$	1	$k \equiv 815 \pmod{1445}$	6	$k \equiv 8091 \pmod{12716}$	4
$k \equiv 3041 \pmod{4641}$	2	$k \equiv 1971 \pmod{2890}$	1	$k \equiv 6068 \pmod{12716}$	5
$k \equiv 16 \pmod{289}$	1	$k \equiv 526 \pmod{2890}$	2	$k \equiv 2889 \pmod{25432}$	2
$k \equiv 33 \pmod{289}$	2	$k \equiv 1682 \pmod{2890}$	3	$k \equiv 12426 \pmod{25432}$	1
$k \equiv 50 \pmod{289}$	3	$k \equiv 3127 \pmod{5780}$	9	$k \equiv 21963 \pmod{25432}$	3
$k \equiv 356 \pmod{578}$	1	$k \equiv 237 \pmod{5780}$	10	$k \equiv 15605 \pmod{25432}$	4
$k \equiv 67 \pmod{578}$	2	$k \equiv 5728 \pmod{5780}$	1	$k \equiv 25142 \pmod{25432}$	5
$k \equiv 84 \pmod{578}$	3	$k \equiv 1393 \pmod{5780}$	2	$k \equiv 9247 \pmod{25432}$	6
$k \equiv 373 \pmod{578}$	4	$k \equiv 2838 \pmod{5780}$	3	$k \equiv 4045 \pmod{9537}$	1
$k \equiv 390 \pmod{578}$	5	$k \equiv 4283 \pmod{5780}$	4	$k \equiv 866 \pmod{9537}$	3
$k \equiv 1257 \pmod{2312}$	2	$k \equiv 1104 \pmod{5780}$	6	$k \equiv 8380 \pmod{9537}$	2
$k \equiv 1835 \pmod{2312}$	1	$k \equiv 2549 \pmod{5780}$	5	$k \equiv 5201 \pmod{9537}$	4
$k \equiv 101 \pmod{2312}$	3	$k \equiv 3994 \pmod{5780}$	7	$k \equiv 16761 \pmod{19074}$	1
$k \equiv 679 \pmod{2312}$	4	$k \equiv 5439 \pmod{5780}$	8	$k \equiv 11559 \pmod{19074}$	2

Table 12.7: Covering information for $d = -6$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 3 \pmod{6}$	1	$k \equiv 319 \pmod{840}$	5	$k \equiv 330 \pmod{464}$	6
$k \equiv 89 \pmod{90}$	2	$k \equiv 439 \pmod{1680}$	1	$k \equiv 446 \pmod{464}$	7
$k \equiv 6 \pmod{16}$	1	$k \equiv 1279 \pmod{1680}$	2	$k \equiv 388 \pmod{928}$	1
$k \equiv 14 \pmod{28}$	1	$k \equiv 1399 \pmod{1680}$	3	$k \equiv 620 \pmod{928}$	2
$k \equiv 9 \pmod{32}$	1	$k \equiv 559 \pmod{1680}$	4	$k \equiv 852 \pmod{928}$	3
$k \equiv 17 \pmod{18}$	1	$k \equiv 49 \pmod{280}$	1	$k \equiv 156 \pmod{928}$	4
$k \equiv 6 \pmod{22}$	1	$k \equiv 189 \pmod{280}$	2	$k \equiv 12 \pmod{87}$	1
$k \equiv 4 \pmod{44}$	1	$k \equiv 64 \pmod{315}$	1	$k \equiv 70 \pmod{87}$	2
$k \equiv 98 \pmod{99}$	1	$k \equiv 274 \pmod{315}$	2	$k \equiv 41 \pmod{87}$	3
$k \equiv 22 \pmod{30}$	1	$k \equiv 169 \pmod{315}$	3	$k \equiv 42 \pmod{174}$	1
$k \equiv 2 \pmod{30}$	2	$k \equiv 289 \pmod{630}$	1	$k \equiv 100 \pmod{174}$	2

Table 12.7 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 12 \pmod{30}$	3	$k \equiv 499 \pmod{630}$	2	$k \equiv 158 \pmod{174}$	3
$k \equiv 7 \pmod{10}$	1	$k \equiv 79 \pmod{630}$	3	$k \equiv 72 \pmod{348}$	1
$k \equiv 13 \pmod{20}$	1	$k \equiv 199 \pmod{630}$	4	$k \equiv 304 \pmod{348}$	2
$k \equiv 3 \pmod{20}$	2	$k \equiv 409 \pmod{630}$	5	$k \equiv 188 \pmod{348}$	3
$k \equiv 1 \pmod{25}$	1	$k \equiv 1249 \pmod{1260}$	1	$k \equiv 16 \pmod{348}$	4
$k \equiv 6 \pmod{25}$	2	$k \equiv 109 \pmod{1260}$	2	$k \equiv 190 \pmod{348}$	5
$k \equiv 11 \pmod{25}$	3	$k \equiv 949 \pmod{1260}$	3	$k \equiv 594 \pmod{696}$	1
$k \equiv 66 \pmod{75}$	1	$k \equiv 529 \pmod{1260}$	4	$k \equiv 130 \pmod{696}$	2
$k \equiv 16 \pmod{75}$	2	$k \equiv 649 \pmod{1260}$	5	$k \equiv 362 \pmod{696}$	3
$k \equiv 41 \pmod{75}$	3	$k \equiv 229 \pmod{1260}$	6	$k \equiv 942 \pmod{1392}$	1
$k \equiv 21 \pmod{125}$	1	$k \equiv 1069 \pmod{1260}$	7	$k \equiv 478 \pmod{1392}$	2
$k \equiv 46 \pmod{125}$	2	$k \equiv 104 \pmod{175}$	1	$k \equiv 14 \pmod{1392}$	3
$k \equiv 71 \pmod{125}$	3	$k \equiv 34 \pmod{175}$	2	$k \equiv 189 \pmod{261}$	1
$k \equiv 96 \pmod{125}$	4	$k \equiv 139 \pmod{175}$	3	$k \equiv 73 \pmod{261}$	2
$k \equiv 246 \pmod{250}$	1	$k \equiv 69 \pmod{350}$	1	$k \equiv 218 \pmod{261}$	3
$k \equiv 121 \pmod{250}$	2	$k \equiv 349 \pmod{350}$	2	$k \equiv 102 \pmod{261}$	4
$k \equiv 25 \pmod{50}$	1	$k \equiv 8 \pmod{40}$	1	$k \equiv 247 \pmod{261}$	5
$k \equiv 5 \pmod{50}$	2	$k \equiv 28 \pmod{40}$	2	$k \equiv 131 \pmod{261}$	6
$k \equiv 35 \pmod{50}$	3	$k \equiv 0 \pmod{29}$	1	$k \equiv 15 \pmod{261}$	7
$k \equiv 65 \pmod{100}$	1	$k \equiv 1 \pmod{29}$	2	$k \equiv 160 \pmod{261}$	8
$k \equiv 15 \pmod{100}$	2	$k \equiv 2 \pmod{29}$	3	$k \equiv 44 \pmod{522}$	1
$k \equiv 45 \pmod{100}$	3	$k \equiv 3 \pmod{29}$	4	$k \equiv 306 \pmod{522}$	2
$k \equiv 95 \pmod{100}$	4	$k \equiv 4 \pmod{29}$	5	$k \equiv 480 \pmod{522}$	3
$k \equiv 149 \pmod{180}$	1	$k \equiv 34 \pmod{58}$	1	$k \equiv 132 \pmod{522}$	4
$k \equiv 59 \pmod{180}$	2	$k \equiv 6 \pmod{58}$	2	$k \equiv 596 \pmod{1044}$	1
$k \equiv 29 \pmod{180}$	3	$k \equiv 36 \pmod{116}$	1	$k \equiv 74 \pmod{1044}$	2
$k \equiv 659 \pmod{720}$	1	$k \equiv 8 \pmod{116}$	2	$k \equiv 248 \pmod{1044}$	4
$k \equiv 299 \pmod{720}$	2	$k \equiv 96 \pmod{116}$	3	$k \equiv 770 \pmod{1044}$	5
$k \equiv 839 \pmod{1440}$	1	$k \equiv 68 \pmod{116}$	4	$k \equiv 944 \pmod{1044}$	3
$k \equiv 1199 \pmod{1440}$	2	$k \equiv 210 \pmod{232}$	1	$k \equiv 422 \pmod{1044}$	6
$k \equiv 119 \pmod{1440}$	3	$k \equiv 66 \pmod{232}$	2	$k \equiv 48 \pmod{60}$	1
$k \equiv 479 \pmod{1440}$	4	$k \equiv 154 \pmod{232}$	3	$k \equiv 28 \pmod{60}$	2
$k \equiv 19 \pmod{360}$	1	$k \equiv 10 \pmod{232}$	4	$k \equiv 8 \pmod{60}$	3
$k \equiv 139 \pmod{360}$	2	$k \equiv 40 \pmod{232}$	5	$k \equiv 75 \pmod{145}$	1
$k \equiv 259 \pmod{360}$	3	$k \equiv 94 \pmod{464}$	1	$k \equiv 104 \pmod{145}$	2
$k \equiv 679 \pmod{840}$	1	$k \equiv 414 \pmod{464}$	2	$k \equiv 105 \pmod{145}$	3
$k \equiv 799 \pmod{840}$	2	$k \equiv 270 \pmod{464}$	3	$k \equiv 134 \pmod{145}$	4
$k \equiv 79 \pmod{840}$	3	$k \equiv 126 \pmod{464}$	4	$k \equiv 280 \pmod{290}$	1
$k \equiv 199 \pmod{840}$	4	$k \equiv 98 \pmod{464}$	5	$k \equiv 164 \pmod{290}$	2
$k \equiv 20 \pmod{290}$	3	$k \equiv 575 \pmod{1015}$	2	$k \equiv 288 \pmod{319}$	3
$k \equiv 194 \pmod{290}$	4	$k \equiv 1010 \pmod{1015}$	3	$k \equiv 201 \pmod{319}$	4
$k \equiv 50 \pmod{290}$	5	$k \equiv 430 \pmod{1015}$	4	$k \equiv 346 \pmod{638}$	1
$k \equiv 224 \pmod{290}$	6	$k \equiv 865 \pmod{1015}$	5	$k \equiv 578 \pmod{638}$	2
$k \equiv 278 \pmod{580}$	1	$k \equiv 285 \pmod{1015}$	6	$k \equiv 172 \pmod{638}$	3
$k \equiv 18 \pmod{580}$	2	$k \equiv 720 \pmod{1015}$	7	$k \equiv 404 \pmod{638}$	4
$k \equiv 338 \pmod{580}$	3	$k \equiv 1764 \pmod{2030}$	1	$k \equiv 636 \pmod{638}$	5

Table 12.7 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 78 \pmod{580}$	4	$k \equiv 1184 \pmod{2030}$	2	$k \equiv 114 \pmod{1276}$	1
$k \equiv 978 \pmod{1160}$	1	$k \equiv 604 \pmod{2030}$	3	$k \equiv 868 \pmod{1276}$	2
$k \equiv 398 \pmod{1160}$	2	$k \equiv 24 \pmod{2030}$	4	$k \equiv 230 \pmod{1276}$	3
$k \equiv 225 \pmod{435}$	1	$k \equiv 3504 \pmod{4060}$	1	$k \equiv 260 \pmod{377}$	1
$k \equiv 370 \pmod{435}$	2	$k \equiv 1474 \pmod{4060}$	2	$k \equiv 144 \pmod{377}$	2
$k \equiv 80 \pmod{435}$	3	$k \equiv 2924 \pmod{4060}$	3	$k \equiv 28 \pmod{377}$	3
$k \equiv 399 \pmod{435}$	4	$k \equiv 894 \pmod{4060}$	4	$k \equiv 289 \pmod{377}$	4
$k \equiv 544 \pmod{870}$	1	$k \equiv 2344 \pmod{4060}$	5	$k \equiv 550 \pmod{754}$	1
$k \equiv 254 \pmod{870}$	2	$k \equiv 314 \pmod{4060}$	6	$k \equiv 434 \pmod{754}$	2
$k \equiv 690 \pmod{870}$	3	$k \equiv 141 \pmod{203}$	1	$k \equiv 1072 \pmod{1508}$	3
$k \equiv 400 \pmod{870}$	4	$k \equiv 170 \pmod{203}$	2	$k \equiv 318 \pmod{1508}$	1
$k \equiv 110 \pmod{870}$	5	$k \equiv 199 \pmod{203}$	3	$k \equiv 956 \pmod{1508}$	2
$k \equiv 864 \pmod{870}$	6	$k \equiv 25 \pmod{203}$	4	$k \equiv 202 \pmod{1508}$	4
$k \equiv 574 \pmod{870}$	7	$k \equiv 54 \pmod{406}$	1	$k \equiv 840 \pmod{1131}$	1
$k \equiv 2024 \pmod{3480}$	1	$k \equiv 286 \pmod{406}$	2	$k \equiv 463 \pmod{1131}$	2
$k \equiv 1154 \pmod{3480}$	2	$k \equiv 316 \pmod{406}$	3	$k \equiv 86 \pmod{1131}$	3
$k \equiv 284 \pmod{3480}$	3	$k \equiv 142 \pmod{406}$	4	$k \equiv 1101 \pmod{1131}$	4
$k \equiv 6374 \pmod{6960}$	1	$k \equiv 374 \pmod{406}$	5	$k \equiv 724 \pmod{2262}$	1
$k \equiv 2894 \pmod{6960}$	2	$k \equiv 112 \pmod{812}$	1	$k \equiv 1478 \pmod{2262}$	2
$k \equiv 138 \pmod{1740}$	1	$k \equiv 84 \pmod{812}$	2	$k \equiv 985 \pmod{1885}$	1
$k \equiv 718 \pmod{1740}$	2	$k \equiv 200 \pmod{812}$	3	$k \equiv 608 \pmod{1885}$	2
$k \equiv 1298 \pmod{1740}$	3	$k \equiv 606 \pmod{812}$	4	$k \equiv 1739 \pmod{1885}$	3
$k \equiv 1038 \pmod{1740}$	4	$k \equiv 432 \pmod{812}$	5	$k \equiv 2000 \pmod{3770}$	1
$k \equiv 1618 \pmod{1740}$	5	$k \equiv 26 \pmod{812}$	6	$k \equiv 3508 \pmod{3770}$	3
$k \equiv 458 \pmod{1740}$	6	$k \equiv 258 \pmod{609}$	1	$k \equiv 2754 \pmod{3770}$	2
$k \equiv 198 \pmod{1740}$	7	$k \equiv 55 \pmod{609}$	2	$k \equiv 1130 \pmod{3770}$	4
$k \equiv 778 \pmod{1740}$	8	$k \equiv 461 \pmod{609}$	3	$k \equiv 2638 \pmod{3770}$	5
$k \equiv 1358 \pmod{1740}$	9	$k \equiv 143 \pmod{319}$	1	$k \equiv 1884 \pmod{3770}$	6
$k \equiv 140 \pmod{1015}$	1	$k \equiv 56 \pmod{319}$	2		

Table 12.8: Covering information for $d = -5$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod{13}$	1	$k \equiv 186 \pmod{208}$	3	$k \equiv 70 \pmod{78}$	1
$k \equiv 1 \pmod{13}$	2	$k \equiv 30 \pmod{208}$	4	$k \equiv 31 \pmod{78}$	2
$k \equiv 2 \pmod{13}$	3	$k \equiv 147 \pmod{208}$	5	$k \equiv 44 \pmod{78}$	3
$k \equiv 16 \pmod{26}$	1	$k \equiv 199 \pmod{208}$	6	$k \equiv 5 \pmod{78}$	4
$k \equiv 3 \pmod{26}$	2	$k \equiv 43 \pmod{416}$	1	$k \equiv 1 \pmod{6}$	1
$k \equiv 4 \pmod{52}$	1	$k \equiv 251 \pmod{416}$	2	$k \equiv 45 \pmod{117}$	1
$k \equiv 17 \pmod{52}$	2	$k \equiv 303 \pmod{416}$	3	$k \equiv 84 \pmod{117}$	2
$k \equiv 82 \pmod{208}$	1	$k \equiv 95 \pmod{416}$	4	$k \equiv 6 \pmod{117}$	3
$k \equiv 134 \pmod{208}$	2	$k \equiv 18 \pmod{39}$	1	$k \equiv 58 \pmod{234}$	1
$k \equiv 136 \pmod{234}$	2	$k \equiv 87 \pmod{91}$	7	$k \equiv 1427 \pmod{6006}$	2

Table 12.8 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 214 \pmod{234}$	3	$k \equiv 49 \pmod{182}$	1	$k \equiv 2428 \pmod{6006}$	3
$k \equiv 110 \pmod{117}$	4	$k \equiv 140 \pmod{182}$	2	$k \equiv 3429 \pmod{6006}$	4
$k \equiv 32 \pmod{351}$	1	$k \equiv 36 \pmod{364}$	1	$k \equiv 4430 \pmod{6006}$	5
$k \equiv 149 \pmod{351}$	2	$k \equiv 127 \pmod{364}$	2	$k \equiv 5431 \pmod{6006}$	6
$k \equiv 266 \pmod{351}$	3	$k \equiv 309 \pmod{364}$	3	$k \equiv 790 \pmod{5005}$	1
$k \equiv 305 \pmod{351}$	4	$k \equiv 218 \pmod{728}$	1	$k \equiv 1791 \pmod{5005}$	2
$k \equiv 71 \pmod{702}$	1	$k \equiv 582 \pmod{728}$	2	$k \equiv 2792 \pmod{5005}$	3
$k \equiv 188 \pmod{702}$	2	$k \equiv 23 \pmod{273}$	1	$k \equiv 3793 \pmod{15015}$	1
$k \equiv 422 \pmod{702}$	3	$k \equiv 114 \pmod{273}$	2	$k \equiv 8798 \pmod{15015}$	2
$k \equiv 539 \pmod{702}$	4	$k \equiv 205 \pmod{273}$	3	$k \equiv 13803 \pmod{15015}$	3
$k \equiv 20 \pmod{65}$	1	$k \equiv 10 \pmod{273}$	4	$k \equiv 4794 \pmod{30030}$	1
$k \equiv 46 \pmod{65}$	2	$k \equiv 192 \pmod{273}$	5	$k \equiv 9799 \pmod{30030}$	2
$k \equiv 7 \pmod{130}$	1	$k \equiv 101 \pmod{546}$	1	$k \equiv 14804 \pmod{30030}$	3
$k \equiv 72 \pmod{130}$	2	$k \equiv 374 \pmod{546}$	2	$k \equiv 19809 \pmod{30030}$	5
$k \equiv 33 \pmod{390}$	1	$k \equiv 88 \pmod{455}$	1	$k \equiv 24814 \pmod{30030}$	4
$k \equiv 163 \pmod{390}$	2	$k \equiv 179 \pmod{455}$	2	$k \equiv 29819 \pmod{30030}$	6
$k \equiv 293 \pmod{390}$	3	$k \equiv 270 \pmod{455}$	3	$k \equiv 1154 \pmod{8008}$	2
$k \equiv 228 \pmod{520}$	1	$k \equiv 361 \pmod{455}$	4	$k \equiv 3156 \pmod{8008}$	1
$k \equiv 488 \pmod{520}$	2	$k \equiv 452 \pmod{455}$	5	$k \equiv 5158 \pmod{8008}$	3
$k \equiv 98 \pmod{1040}$	1	$k \equiv 75 \pmod{910}$	1	$k \equiv 7160 \pmod{8008}$	4
$k \equiv 358 \pmod{1040}$	2	$k \equiv 257 \pmod{910}$	2	$k \equiv 153 \pmod{10010}$	1
$k \equiv 618 \pmod{1040}$	3	$k \equiv 439 \pmod{910}$	3	$k \equiv 2155 \pmod{10010}$	2
$k \equiv 878 \pmod{1040}$	4	$k \equiv 621 \pmod{910}$	4	$k \equiv 4157 \pmod{10010}$	3
$k \equiv 59 \pmod{195}$	1	$k \equiv 803 \pmod{910}$	5	$k \equiv 8161 \pmod{10010}$	4
$k \equiv 124 \pmod{195}$	2	$k \equiv 348 \pmod{1456}$	1	$k \equiv 6159 \pmod{60060}$	1
$k \equiv 189 \pmod{195}$	3	$k \equiv 530 \pmod{1456}$	2	$k \equiv 16169 \pmod{60060}$	2
$k \equiv 8 \pmod{104}$	1	$k \equiv 712 \pmod{1456}$	3	$k \equiv 26179 \pmod{60060}$	3
$k \equiv 60 \pmod{104}$	2	$k \equiv 1076 \pmod{1456}$	4	$k \equiv 36189 \pmod{60060}$	4
$k \equiv 21 \pmod{156}$	1	$k \equiv 1258 \pmod{1456}$	5	$k \equiv 46199 \pmod{60060}$	5
$k \equiv 73 \pmod{156}$	2	$k \equiv 1440 \pmod{1456}$	6	$k \equiv 56209 \pmod{60060}$	6
$k \equiv 125 \pmod{156}$	3	$k \equiv 166 \pmod{2912}$	1	$k \equiv 11 \pmod{143}$	1
$k \equiv 34 \pmod{312}$	1	$k \equiv 1622 \pmod{2912}$	2	$k \equiv 24 \pmod{143}$	2
$k \equiv 86 \pmod{312}$	2	$k \equiv 2350 \pmod{2912}$	3	$k \equiv 89 \pmod{143}$	3
$k \equiv 138 \pmod{312}$	3	$k \equiv 62 \pmod{416}$	5	$k \equiv 102 \pmod{143}$	4
$k \equiv 242 \pmod{312}$	4	$k \equiv 244 \pmod{1001}$	1	$k \equiv 37 \pmod{286}$	1
$k \equiv 294 \pmod{312}$	5	$k \equiv 517 \pmod{1001}$	2	$k \equiv 115 \pmod{286}$	2
$k \equiv 190 \pmod{624}$	1	$k \equiv 608 \pmod{1001}$	3	$k \equiv 180 \pmod{286}$	3
$k \equiv 502 \pmod{624}$	2	$k \equiv 881 \pmod{1001}$	4	$k \equiv 258 \pmod{286}$	4
$k \equiv 47 \pmod{260}$	1	$k \equiv 335 \pmod{2002}$	1	$k \equiv 50 \pmod{429}$	1
$k \equiv 99 \pmod{260}$	2	$k \equiv 972 \pmod{2002}$	2	$k \equiv 193 \pmod{429}$	2
$k \equiv 151 \pmod{260}$	3	$k \equiv 1336 \pmod{2002}$	3	$k \equiv 336 \pmod{429}$	3
$k \equiv 203 \pmod{260}$	4	$k \equiv 1973 \pmod{2002}$	4	$k \equiv 128 \pmod{572}$	1
$k \equiv 255 \pmod{260}$	5	$k \equiv 1700 \pmod{2002}$	5	$k \equiv 271 \pmod{572}$	2
$k \equiv 9 \pmod{91}$	1	$k \equiv 699 \pmod{4004}$	1	$k \equiv 414 \pmod{572}$	3
$k \equiv 22 \pmod{91}$	2	$k \equiv 2701 \pmod{4004}$	2	$k \equiv 557 \pmod{572}$	4
$k \equiv 35 \pmod{91}$	3	$k \equiv 62 \pmod{3003}$	1	$k \equiv 63 \pmod{1716}$	1

Table 12.8 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 48 \pmod{91}$	4	$k \equiv 1063 \pmod{3003}$	2	$k \equiv 635 \pmod{1716}$	2
$k \equiv 61 \pmod{91}$	5	$k \equiv 2064 \pmod{3003}$	3	$k \equiv 1207 \pmod{1716}$	3
$k \equiv 74 \pmod{91}$	6	$k \equiv 426 \pmod{6006}$	1	$k \equiv 349 \pmod{572}$	5
$k \equiv 206 \pmod{286}$	5	$k \equiv 272 \pmod{1014}$	2	$k \equiv 1169 \pmod{1183}$	4
$k \equiv 141 \pmod{858}$	1	$k \equiv 610 \pmod{1014}$	3	$k \equiv 662 \pmod{2366}$	1
$k \equiv 284 \pmod{858}$	2	$k \equiv 779 \pmod{1014}$	4	$k \equiv 831 \pmod{2366}$	2
$k \equiv 427 \pmod{858}$	3	$k \equiv 116 \pmod{676}$	1	$k \equiv 1845 \pmod{2366}$	3
$k \equiv 570 \pmod{858}$	4	$k \equiv 285 \pmod{676}$	2	$k \equiv 2014 \pmod{2366}$	4
$k \equiv 856 \pmod{858}$	5	$k \equiv 454 \pmod{676}$	3	$k \equiv 1000 \pmod{3549}$	1
$k \equiv 713 \pmod{1716}$	4	$k \equiv 623 \pmod{2704}$	1	$k \equiv 2183 \pmod{3549}$	2
$k \equiv 1571 \pmod{3432}$	1	$k \equiv 1299 \pmod{2704}$	2	$k \equiv 3366 \pmod{3549}$	3
$k \equiv 3287 \pmod{3432}$	2	$k \equiv 1975 \pmod{2704}$	3	$k \equiv 168 \pmod{2535}$	1
$k \equiv 76 \pmod{715}$	1	$k \equiv 2651 \pmod{2704}$	4	$k \equiv 675 \pmod{2535}$	2
$k \equiv 362 \pmod{715}$	2	$k \equiv 129 \pmod{1352}$	1	$k \equiv 1182 \pmod{2535}$	3
$k \equiv 505 \pmod{715}$	3	$k \equiv 298 \pmod{1352}$	2	$k \equiv 2196 \pmod{2535}$	4
$k \equiv 219 \pmod{1430}$	1	$k \equiv 467 \pmod{1352}$	3	$k \equiv 1689 \pmod{5070}$	1
$k \equiv 648 \pmod{1430}$	2	$k \equiv 636 \pmod{1352}$	4	$k \equiv 4224 \pmod{5070}$	2
$k \equiv 934 \pmod{1430}$	3	$k \equiv 805 \pmod{1352}$	5	$k \equiv 506 \pmod{1521}$	1
$k \equiv 1363 \pmod{1430}$	4	$k \equiv 974 \pmod{1352}$	6	$k \equiv 1013 \pmod{3042}$	1
$k \equiv 12 \pmod{169}$	1	$k \equiv 1143 \pmod{1352}$	7	$k \equiv 1520 \pmod{3042}$	2
$k \equiv 25 \pmod{169}$	2	$k \equiv 1312 \pmod{1352}$	8	$k \equiv 2534 \pmod{3042}$	3
$k \equiv 38 \pmod{169}$	3	$k \equiv 311 \pmod{845}$	1	$k \equiv 3041 \pmod{3042}$	4
$k \equiv 51 \pmod{338}$	1	$k \equiv 480 \pmod{845}$	2	$k \equiv 844 \pmod{10140}$	1
$k \equiv 64 \pmod{338}$	2	$k \equiv 142 \pmod{1690}$	1	$k \equiv 1858 \pmod{10140}$	2
$k \equiv 220 \pmod{338}$	3	$k \equiv 818 \pmod{1690}$	2	$k \equiv 2872 \pmod{10140}$	3
$k \equiv 233 \pmod{338}$	4	$k \equiv 987 \pmod{1690}$	3	$k \equiv 3886 \pmod{10140}$	4
$k \equiv 77 \pmod{507}$	2	$k \equiv 1663 \pmod{1690}$	4	$k \equiv 4900 \pmod{10140}$	8
$k \equiv 90 \pmod{507}$	1	$k \equiv 649 \pmod{3380}$	1	$k \equiv 5914 \pmod{10140}$	5
$k \equiv 246 \pmod{507}$	3	$k \equiv 1494 \pmod{3380}$	2	$k \equiv 6928 \pmod{10140}$	6
$k \equiv 259 \pmod{507}$	4	$k \equiv 2339 \pmod{3380}$	3	$k \equiv 7942 \pmod{10140}$	7
$k \equiv 415 \pmod{507}$	5	$k \equiv 3184 \pmod{3380}$	4	$k \equiv 8956 \pmod{10140}$	9
$k \equiv 428 \pmod{507}$	6	$k \equiv 155 \pmod{1183}$	1	$k \equiv 9970 \pmod{10140}$	10
$k \equiv 441 \pmod{507}$	7	$k \equiv 324 \pmod{1183}$	2		
$k \equiv 103 \pmod{1014}$	1	$k \equiv 493 \pmod{1183}$	3		

Table 12.9: Covering information for $d = -3$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 4 \pmod{6}$	2	$k \equiv 193 \pmod{420}$	1	$k \equiv 39 \pmod{126}$	2
$k \equiv 5 \pmod{6}$	1	$k \equiv 403 \pmod{420}$	2	$k \equiv 249 \pmod{252}$	1
$k \equiv 0 \pmod{16}$	1	$k \equiv 109 \pmod{420}$	3	$k \equiv 123 \pmod{252}$	2
$k \equiv 11 \pmod{21}$	1	$k \equiv 319 \pmod{420}$	4	$k \equiv 113 \pmod{176}$	1
$k \equiv 14 \pmod{22}$	1	$k \equiv 18 \pmod{336}$	1	$k \equiv 69 \pmod{176}$	2

Table 12.9 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 40 \pmod{44}$	1	$k \equiv 228 \pmod{336}$	2	$k \equiv 201 \pmod{352}$	1
$k \equiv 96 \pmod{99}$	1	$k \equiv 102 \pmod{336}$	3	$k \equiv 333 \pmod{352}$	2
$k \equiv 8 \pmod{64}$	1	$k \equiv 522 \pmod{672}$	1	$k \equiv 25 \pmod{352}$	3
$k \equiv 24 \pmod{64}$	2	$k \equiv 186 \pmod{672}$	2	$k \equiv 157 \pmod{352}$	4
$k \equiv 40 \pmod{64}$	3	$k \equiv 396 \pmod{672}$	3	$k \equiv 3 \pmod{704}$	1
$k \equiv 56 \pmod{64}$	4	$k \equiv 60 \pmod{672}$	4	$k \equiv 531 \pmod{704}$	2
$k \equiv 25 \pmod{210}$	1	$k \equiv 270 \pmod{672}$	5	$k \equiv 355 \pmod{704}$	3
$k \equiv 151 \pmod{210}$	2	$k \equiv 606 \pmod{672}$	6	$k \equiv 179 \pmod{704}$	4
$k \equiv 67 \pmod{210}$	3	$k \equiv 81 \pmod{126}$	1	$k \equiv 267 \pmod{1056}$	1
$k \equiv 619 \pmod{1056}$	2	$k \equiv 4268 \pmod{7392}$	2	$k \equiv 24 \pmod{57}$	3
$k \equiv 971 \pmod{1056}$	3	$k \equiv 6116 \pmod{7392}$	3	$k \equiv 79 \pmod{114}$	1
$k \equiv 795 \pmod{2112}$	1	$k \equiv 572 \pmod{7392}$	4	$k \equiv 98 \pmod{114}$	2
$k \equiv 91 \pmod{2112}$	2	$k \equiv 5060 \pmod{5544}$	1	$k \equiv 61 \pmod{228}$	1
$k \equiv 1499 \pmod{2112}$	3	$k \equiv 3212 \pmod{5544}$	2	$k \equiv 175 \pmod{228}$	2
$k \equiv 1851 \pmod{2112}$	4	$k \equiv 1364 \pmod{5544}$	3	$k \equiv 194 \pmod{228}$	3
$k \equiv 1147 \pmod{2112}$	5	$k \equiv 0 \pmod{19}$	1	$k \equiv 157 \pmod{228}$	4
$k \equiv 443 \pmod{4224}$	1	$k \equiv 1 \pmod{38}$	1	$k \equiv 43 \pmod{228}$	5
$k \equiv 2555 \pmod{4224}$	2	$k \equiv 58 \pmod{76}$	1	$k \equiv 62 \pmod{228}$	6
$k \equiv 135 \pmod{440}$	1	$k \equiv 58 \pmod{76}$	2	$k \equiv 308 \pmod{456}$	1
$k \equiv 311 \pmod{440}$	2	$k \equiv 20 \pmod{152}$	1	$k \equiv 404 \pmod{456}$	2
$k \equiv 487 \pmod{880}$	1	$k \equiv 116 \pmod{152}$	2	$k \equiv 25 \pmod{456}$	3
$k \equiv 47 \pmod{880}$	2	$k \equiv 2 \pmod{152}$	3	$k \equiv 367 \pmod{456}$	4
$k \equiv 663 \pmod{880}$	3	$k \equiv 78 \pmod{152}$	4	$k \equiv 253 \pmod{456}$	5
$k \equiv 223 \pmod{880}$	4	$k \equiv 97 \pmod{304}$	1	$k \equiv 63 \pmod{171}$	1
$k \equiv 839 \pmod{880}$	5	$k \equiv 211 \pmod{304}$	2	$k \equiv 120 \pmod{171}$	2
$k \equiv 399 \pmod{1760}$	1	$k \equiv 21 \pmod{304}$	3	$k \equiv 6 \pmod{171}$	3
$k \equiv 1279 \pmod{1760}$	2	$k \equiv 135 \pmod{608}$	1	$k \equiv 595 \pmod{912}$	1
$k \equiv 0 \pmod{231}$	1	$k \equiv 439 \pmod{608}$	2	$k \equiv 139 \pmod{912}$	2
$k \equiv 99 \pmod{231}$	2	$k \equiv 553 \pmod{608}$	3	$k \equiv 956 \pmod{1368}$	1
$k \equiv 198 \pmod{231}$	3	$k \equiv 857 \pmod{1216}$	1	$k \equiv 44 \pmod{1368}$	2
$k \equiv 66 \pmod{231}$	4	$k \equiv 249 \pmod{1216}$	2	$k \equiv 500 \pmod{1368}$	3
$k \equiv 33 \pmod{231}$	5	$k \equiv 971 \pmod{1216}$	3	$k \equiv 1526 \pmod{5472}$	1
$k \equiv 132 \pmod{231}$	6	$k \equiv 667 \pmod{1216}$	4	$k \equiv 3350 \pmod{5472}$	2
$k \equiv 385 \pmod{462}$	1	$k \equiv 363 \pmod{1216}$	5	$k \equiv 5174 \pmod{5472}$	3
$k \equiv 253 \pmod{462}$	2	$k \equiv 59 \pmod{1216}$	7	$k \equiv 45 \pmod{342}$	1
$k \equiv 121 \pmod{462}$	3	$k \equiv 781 \pmod{1216}$	6	$k \equiv 216 \pmod{342}$	2
$k \equiv 770 \pmod{924}$	1	$k \equiv 477 \pmod{1216}$	8	$k \equiv 273 \pmod{342}$	3
$k \equiv 638 \pmod{924}$	2	$k \equiv 173 \pmod{1216}$	9	$k \equiv 102 \pmod{342}$	4
$k \equiv 506 \pmod{924}$	3	$k \equiv 1085 \pmod{2432}$	1	$k \equiv 159 \pmod{342}$	5
$k \equiv 374 \pmod{924}$	4	$k \equiv 2301 \pmod{2432}$	2	$k \equiv 330 \pmod{3420}$	1
$k \equiv 110 \pmod{924}$	5	$k \equiv 1807 \pmod{2432}$	3	$k \equiv 3066 \pmod{3420}$	2
$k \equiv 308 \pmod{1848}$	1	$k \equiv 63 \pmod{192}$	4	$k \equiv 2382 \pmod{3420}$	3
$k \equiv 1100 \pmod{1848}$	2	$k \equiv 159 \pmod{384}$	1	$k \equiv 1698 \pmod{3420}$	4
$k \equiv 44 \pmod{1848}$	3	$k \equiv 31 \pmod{384}$	2	$k \equiv 1014 \pmod{3420}$	5
$k \equiv 836 \pmod{1848}$	4	$k \equiv 303 \pmod{384}$	3	$k \equiv 5460 \pmod{6840}$	1

Table 12.9 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 451 \pmod{693}$	1	$k \equiv 127 \pmod{384}$	4	$k \equiv 1356 \pmod{6840}$	2
$k \equiv 220 \pmod{693}$	2	$k \equiv 319 \pmod{384}$	5	$k \equiv 4092 \pmod{6840}$	3
$k \equiv 682 \pmod{693}$	3	$k \equiv 559 \pmod{768}$	1	$k \equiv 6828 \pmod{6840}$	4
$k \equiv 649 \pmod{693}$	4	$k \equiv 175 \pmod{768}$	2	$k \equiv 2724 \pmod{6840}$	5
$k \equiv 418 \pmod{693}$	5	$k \equiv 591 \pmod{768}$	3	$k \equiv 577 \pmod{684}$	1
$k \equiv 187 \pmod{693}$	6	$k \equiv 79 \pmod{768}$	4	$k \equiv 7 \pmod{684}$	2
$k \equiv 55 \pmod{1386}$	1	$k \equiv 207 \pmod{768}$	5	$k \equiv 121 \pmod{684}$	3
$k \equiv 979 \pmod{1386}$	2	$k \equiv 463 \pmod{768}$	6	$k \equiv 235 \pmod{684}$	4
$k \equiv 517 \pmod{1386}$	3	$k \equiv 3935 \pmod{4864}$	1	$k \equiv 1033 \pmod{1368}$	4
$k \equiv 1826 \pmod{3696}$	1	$k \equiv 1503 \pmod{4864}$	2	$k \equiv 349 \pmod{2736}$	1
$k \equiv 2750 \pmod{3696}$	2	$k \equiv 2415 \pmod{4864}$	3	$k \equiv 1717 \pmod{2736}$	2
$k \equiv 3674 \pmod{3696}$	3	$k \equiv 4847 \pmod{4864}$	4	$k \equiv 463 \pmod{2736}$	3
$k \equiv 902 \pmod{3696}$	4	$k \equiv 3 \pmod{57}$	1	$k \equiv 5251 \pmod{5472}$	4
$k \equiv 2420 \pmod{7392}$	1	$k \equiv 42 \pmod{57}$	2	$k \equiv 1831 \pmod{5472}$	5
$k \equiv 3883 \pmod{5472}$	6	$k \equiv 10 \pmod{532}$	1	$k \equiv 885 \pmod{1197}$	2
$k \equiv 2515 \pmod{5472}$	7	$k \equiv 390 \pmod{532}$	2	$k \equiv 87 \pmod{1197}$	3
$k \equiv 10039 \pmod{10944}$	1	$k \equiv 314 \pmod{532}$	3	$k \equiv 2215 \pmod{2394}$	1
$k \equiv 6619 \pmod{10944}$	2	$k \equiv 49 \pmod{532}$	4	$k \equiv 1417 \pmod{2394}$	2
$k \equiv 4567 \pmod{10944}$	3	$k \equiv 182 \pmod{532}$	5	$k \equiv 619 \pmod{2394}$	3
$k \equiv 1147 \pmod{10944}$	4	$k \equiv 276 \pmod{1064}$	1	$k \equiv 1550 \pmod{2394}$	4
$k \equiv 140 \pmod{570}$	1	$k \equiv 124 \pmod{1064}$	2	$k \equiv 752 \pmod{2394}$	5
$k \equiv 26 \pmod{570}$	2	$k \equiv 580 \pmod{1064}$	3	$k \equiv 201 \pmod{931}$	1
$k \equiv 482 \pmod{570}$	3	$k \equiv 1379 \pmod{2128}$	1	$k \equiv 600 \pmod{931}$	2
$k \equiv 368 \pmod{570}$	4	$k \equiv 1911 \pmod{2128}$	2	$k \equiv 68 \pmod{931}$	3
$k \equiv 254 \pmod{1140}$	1	$k \equiv 315 \pmod{2128}$	3	$k \equiv 467 \pmod{931}$	4
$k \equiv 824 \pmod{1140}$	2	$k \equiv 847 \pmod{2128}$	4	$k \equiv 866 \pmod{931}$	5
$k \equiv 65 \pmod{95}$	1	$k \equiv 3108 \pmod{4256}$	1	$k \equiv 334 \pmod{931}$	6
$k \equiv 46 \pmod{95}$	2	$k \equiv 4172 \pmod{4256}$	2	$k \equiv 733 \pmod{931}$	7
$k \equiv 27 \pmod{95}$	3	$k \equiv 980 \pmod{4256}$	3	$k \equiv 790 \pmod{1862}$	1
$k \equiv 8 \pmod{95}$	4	$k \equiv 2044 \pmod{4256}$	4	$k \equiv 1721 \pmod{3724}$	1
$k \equiv 84 \pmod{95}$	5	$k \equiv 505 \pmod{665}$	1	$k \equiv 3583 \pmod{3724}$	2
$k \equiv 66 \pmod{285}$	1	$k \equiv 106 \pmod{665}$	2	$k \equiv 2120 \pmod{3724}$	3
$k \equiv 256 \pmod{285}$	2	$k \equiv 372 \pmod{665}$	3	$k \equiv 1189 \pmod{3724}$	4
$k \equiv 161 \pmod{285}$	3	$k \equiv 1303 \pmod{1330}$	1	$k \equiv 258 \pmod{3724}$	5
$k \equiv 142 \pmod{190}$	1	$k \equiv 239 \pmod{1330}$	2	$k \equiv 3051 \pmod{3724}$	6
$k \equiv 123 \pmod{190}$	2	$k \equiv 1094 \pmod{1330}$	3	$k \equiv 657 \pmod{7448}$	1
$k \equiv 104 \pmod{190}$	3	$k \equiv 638 \pmod{2660}$	1	$k \equiv 3450 \pmod{7448}$	2
$k \equiv 85 \pmod{380}$	1	$k \equiv 2234 \pmod{2660}$	2	$k \equiv 6243 \pmod{7448}$	3
$k \equiv 275 \pmod{380}$	2	$k \equiv 4628 \pmod{5320}$	1	$k \equiv 1588 \pmod{7448}$	5
$k \equiv 47 \pmod{380}$	3	$k \equiv 3564 \pmod{5320}$	2	$k \equiv 4381 \pmod{7448}$	4
$k \equiv 237 \pmod{380}$	4	$k \equiv 30 \pmod{1995}$	1	$k \equiv 7174 \pmod{7448}$	6
$k \equiv 218 \pmod{380}$	5	$k \equiv 1626 \pmod{1995}$	2	$k \equiv 2519 \pmod{7448}$	7
$k \equiv 9 \pmod{380}$	6	$k \equiv 1227 \pmod{1995}$	3	$k \equiv 1056 \pmod{2793}$	1
$k \equiv 199 \pmod{380}$	7	$k \equiv 828 \pmod{1995}$	4	$k \equiv 1455 \pmod{2793}$	3
$k \equiv 370 \pmod{380}$	8	$k \equiv 3355 \pmod{3990}$	1	$k \equiv 1854 \pmod{2793}$	2

Table 12.9 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 180 \pmod{760}$	1	$k \equiv 961 \pmod{3990}$	2	$k \equiv 2253 \pmod{2793}$	4
$k \equiv 28 \pmod{760}$	2	$k \equiv 2557 \pmod{7980}$	1	$k \equiv 1987 \pmod{5586}$	1
$k \equiv 145 \pmod{1140}$	3	$k \equiv 6547 \pmod{7980}$	2	$k \equiv 5179 \pmod{5586}$	2
$k \equiv 715 \pmod{1140}$	5	$k \equiv 4153 \pmod{7980}$	3	$k \equiv 2785 \pmod{5586}$	3
$k \equiv 601 \pmod{1140}$	6	$k \equiv 163 \pmod{7980}$	4	$k \equiv 391 \pmod{5586}$	4
$k \equiv 31 \pmod{1140}$	4	$k \equiv 2690 \pmod{7980}$	5	$k \equiv 2918 \pmod{11172}$	2
$k \equiv 1057 \pmod{1140}$	8	$k \equiv 4286 \pmod{7980}$	6	$k \equiv 6110 \pmod{11172}$	1
$k \equiv 487 \pmod{1140}$	7	$k \equiv 5882 \pmod{7980}$	8	$k \equiv 9302 \pmod{11172}$	3
$k \equiv 373 \pmod{1140}$	9	$k \equiv 7478 \pmod{7980}$	7	$k \equiv 1322 \pmod{11172}$	4
$k \equiv 943 \pmod{1140}$	10	$k \equiv 5749 \pmod{7980}$	9	$k \equiv 19676 \pmod{22344}$	2
$k \equiv 829 \pmod{1140}$	11	$k \equiv 1759 \pmod{7980}$	10	$k \equiv 524 \pmod{22344}$	1
$k \equiv 259 \pmod{2280}$	1	$k \equiv 14660 \pmod{15960}$	1	$k \equiv 3716 \pmod{22344}$	3
$k \equiv 1399 \pmod{2280}$	2	$k \equiv 8276 \pmod{15960}$	2	$k \equiv 6908 \pmod{22344}$	4
$k \equiv 105 \pmod{133}$	1	$k \equiv 1892 \pmod{15960}$	3	$k \equiv 126 \pmod{399}$	1
$k \equiv 29 \pmod{133}$	2	$k \equiv 11468 \pmod{15960}$	4	$k \equiv 183 \pmod{399}$	2
$k \equiv 86 \pmod{133}$	3	$k \equiv 4419 \pmod{5985}$	1	$k \equiv 240 \pmod{399}$	3
$k \equiv 143 \pmod{266}$	1	$k \equiv 2424 \pmod{5985}$	2	$k \equiv 297 \pmod{399}$	4
$k \equiv 257 \pmod{266}$	2	$k \equiv 429 \pmod{5985}$	3	$k \equiv 12 \pmod{399}$	5
$k \equiv 181 \pmod{266}$	3	$k \equiv 486 \pmod{1197}$	1	$k \equiv 468 \pmod{798}$	1
$k \equiv 69 \pmod{798}$	2	$k \equiv 3129 \pmod{5130}$	2	$k \equiv 204 \pmod{209}$	5
$k \equiv 734 \pmod{798}$	3	$k \equiv 474 \pmod{540}$	1	$k \equiv 337 \pmod{418}$	1
$k \equiv 1190 \pmod{1596}$	1	$k \equiv 510 \pmod{540}$	2	$k \equiv 261 \pmod{418}$	2
$k \equiv 50 \pmod{1596}$	2	$k \equiv 186 \pmod{540}$	3	$k \equiv 185 \pmod{418}$	3
$k \equiv 506 \pmod{1596}$	3	$k \equiv 402 \pmod{540}$	4	$k \equiv 109 \pmod{418}$	4
$k \equiv 962 \pmod{1596}$	4	$k \equiv 78 \pmod{540}$	5	$k \equiv 546 \pmod{836}$	1
$k \equiv 1988 \pmod{3192}$	1	$k \equiv 294 \pmod{540}$	6	$k \equiv 470 \pmod{836}$	2
$k \equiv 2444 \pmod{3192}$	2	$k \equiv 204 \pmod{540}$	7	$k \equiv 394 \pmod{836}$	3
$k \equiv 2900 \pmod{3192}$	3	$k \equiv 6720 \pmod{10260}$	2	$k \equiv 1154 \pmod{3344}$	1
$k \equiv 164 \pmod{6384}$	1	$k \equiv 2616 \pmod{10260}$	1	$k \equiv 1990 \pmod{3344}$	2
$k \equiv 3356 \pmod{6384}$	2	$k \equiv 8772 \pmod{10260}$	3	$k \equiv 2826 \pmod{3344}$	3
$k \equiv 1874 \pmod{4788}$	1	$k \equiv 4668 \pmod{10260}$	4	$k \equiv 318 \pmod{3344}$	4
$k \equiv 3470 \pmod{4788}$	2	$k \equiv 564 \pmod{10260}$	5	$k \equiv 964 \pmod{1672}$	1
$k \equiv 278 \pmod{4788}$	3	$k \equiv 2 \pmod{48}$	1	$k \equiv 52 \pmod{1672}$	2
$k \equiv 2348 \pmod{4788}$	4	$k \equiv 38 \pmod{96}$	1	$k \equiv 812 \pmod{1672}$	3
$k \equiv 4268 \pmod{9576}$	1	$k \equiv 86 \pmod{96}$	2	$k \equiv 1572 \pmod{1672}$	4
$k \equiv 7460 \pmod{9576}$	2	$k \equiv 74 \pmod{96}$	3	$k \equiv 243 \pmod{627}$	1
$k \equiv 1076 \pmod{9576}$	3	$k \equiv 26 \pmod{96}$	4	$k \equiv 585 \pmod{627}$	2
$k \equiv 9530 \pmod{9576}$	4	$k \equiv 14 \pmod{192}$	1	$k \equiv 15 \pmod{627}$	3
$k \equiv 4742 \pmod{9576}$	5	$k \equiv 158 \pmod{192}$	2	$k \equiv 357 \pmod{627}$	4
$k \equiv 108 \pmod{513}$	1	$k \equiv 110 \pmod{192}$	3	$k \equiv 72 \pmod{2508}$	1
$k \equiv 279 \pmod{513}$	2	$k \equiv 350 \pmod{480}$	1	$k \equiv 1953 \pmod{2508}$	2
$k \equiv 963 \pmod{1026}$	1	$k \equiv 446 \pmod{480}$	2	$k \equiv 8850 \pmod{10032}$	1
$k \equiv 165 \pmod{1026}$	2	$k \equiv 62 \pmod{480}$	3	$k \equiv 6342 \pmod{10032}$	2
$k \equiv 849 \pmod{1026}$	3	$k \equiv 638 \pmod{960}$	1	$k \equiv 3834 \pmod{10032}$	3
$k \equiv 507 \pmod{1026}$	4	$k \equiv 254 \pmod{960}$	2	$k \equiv 1326 \pmod{10032}$	4

Table 12.9 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 18 \pmod{108}$	1	$k \equiv 1153 \pmod{1824}$	1	$k \equiv 15747 \pmod{20064}$	1
$k \equiv 30 \pmod{108}$	2	$k \equiv 241 \pmod{1824}$	2	$k \equiv 10731 \pmod{20064}$	2
$k \equiv 66 \pmod{108}$	3	$k \equiv 1381 \pmod{1824}$	3	$k \equiv 5715 \pmod{20064}$	3
$k \equiv 2046 \pmod{2052}$	1	$k \equiv 469 \pmod{1824}$	4	$k \equiv 699 \pmod{20064}$	4
$k \equiv 1476 \pmod{4104}$	1	$k \equiv 1609 \pmod{1824}$	5	$k \equiv 3207 \pmod{5016}$	1
$k \equiv 3756 \pmod{4104}$	2	$k \equiv 2521 \pmod{3648}$	1	$k \equiv 452 \pmod{5016}$	2
$k \equiv 2388 \pmod{8208}$	1	$k \equiv 697 \pmod{3648}$	2	$k \equiv 4556 \pmod{5016}$	3
$k \equiv 6492 \pmod{8208}$	2	$k \equiv 45 \pmod{80}$	1	$k \equiv 2732 \pmod{5016}$	4
$k \equiv 5124 \pmod{8208}$	3	$k \equiv 61 \pmod{80}$	2	$k \equiv 1820 \pmod{5016}$	5
$k \equiv 1020 \pmod{8208}$	4	$k \equiv 157 \pmod{240}$	1	$k \equiv 908 \pmod{5016}$	6
$k \equiv 60 \pmod{135}$	1	$k \equiv 13 \pmod{240}$	2	$k \equiv 661 \pmod{1254}$	1
$k \equiv 6 \pmod{135}$	2	$k \equiv 109 \pmod{240}$	3	$k \equiv 1003 \pmod{1254}$	2
$k \equiv 87 \pmod{135}$	3	$k \equiv 127 \pmod{228}$	7	$k \equiv 433 \pmod{1254}$	3
$k \equiv 33 \pmod{135}$	4	$k \equiv 20 \pmod{120}$	1	$k \equiv 775 \pmod{1254}$	4
$k \equiv 114 \pmod{135}$	5	$k \equiv 716 \pmod{2280}$	3	$k \equiv 1117 \pmod{1254}$	5
$k \equiv 2445 \pmod{2565}$	1	$k \equiv 1172 \pmod{2280}$	4	$k \equiv 205 \pmod{1045}$	1
$k \equiv 906 \pmod{2565}$	2	$k \equiv 3908 \pmod{4560}$	1	$k \equiv 1041 \pmod{1045}$	2
$k \equiv 1932 \pmod{2565}$	3	$k \equiv 1628 \pmod{4560}$	2	$k \equiv 832 \pmod{1045}$	3
$k \equiv 393 \pmod{2565}$	4	$k \equiv 2084 \pmod{4560}$	3	$k \equiv 623 \pmod{2090}$	1
$k \equiv 69 \pmod{270}$	1	$k \equiv 4364 \pmod{4560}$	4	$k \equiv 1459 \pmod{2090}$	2
$k \equiv 105 \pmod{270}$	2	$k \equiv 166 \pmod{209}$	1	$k \equiv 965 \pmod{2090}$	3
$k \equiv 51 \pmod{270}$	3	$k \equiv 90 \pmod{209}$	2	$k \equiv 1801 \pmod{4180}$	1
$k \equiv 267 \pmod{270}$	4	$k \equiv 147 \pmod{209}$	3	$k \equiv 3891 \pmod{4180}$	2
$k \equiv 2103 \pmod{5130}$	1	$k \equiv 71 \pmod{209}$	4	$k \equiv 3758 \pmod{4180}$	3
$k \equiv 414 \pmod{4180}$	4	$k \equiv 8964 \pmod{9405}$	3	$k \equiv 1308 \pmod{1976}$	4
$k \equiv 2010 \pmod{4180}$	5	$k \equiv 2694 \pmod{9405}$	5	$k \equiv 282 \pmod{741}$	1
$k \equiv 2846 \pmod{4180}$	6	$k \equiv 5829 \pmod{9405}$	6	$k \equiv 529 \pmod{741}$	2
$k \equiv 1668 \pmod{8360}$	1	$k \equiv 901 \pmod{990}$	1	$k \equiv 35 \pmod{741}$	3
$k \equiv 6684 \pmod{8360}$	2	$k \equiv 571 \pmod{990}$	2	$k \equiv 738 \pmod{741}$	4
$k \equiv 4100 \pmod{8360}$	3	$k \equiv 241 \pmod{990}$	3	$k \equiv 453 \pmod{741}$	5
$k \equiv 756 \pmod{16720}$	1	$k \equiv 307 \pmod{990}$	4	$k \equiv 985 \pmod{1482}$	1
$k \equiv 9116 \pmod{16720}$	2	$k \equiv 18787 \pmod{18810}$	1	$k \equiv 1441 \pmod{1482}$	2
$k \equiv 2637 \pmod{3135}$	1	$k \equiv 12517 \pmod{18810}$	2	$k \equiv 4196 \pmod{5928}$	1
$k \equiv 1383 \pmod{3135}$	4	$k \equiv 17533 \pmod{18810}$	3	$k \equiv 4652 \pmod{5928}$	2
$k \equiv 129 \pmod{3135}$	5	$k \equiv 373 \pmod{1980}$	1	$k \equiv 415 \pmod{1235}$	1
$k \equiv 1725 \pmod{3135}$	2	$k \equiv 1363 \pmod{1980}$	3	$k \equiv 1156 \pmod{1235}$	2
$k \equiv 471 \pmod{3135}$	3	$k \equiv 1033 \pmod{1980}$	6	$k \equiv 662 \pmod{1235}$	3
$k \equiv 2352 \pmod{3135}$	6	$k \equiv 43 \pmod{1980}$	5	$k \equiv 168 \pmod{1235}$	4
$k \equiv 1098 \pmod{3135}$	7	$k \equiv 109 \pmod{1980}$	4	$k \equiv 909 \pmod{1235}$	5
$k \equiv 2979 \pmod{3135}$	8	$k \equiv 1099 \pmod{1980}$	2	$k \equiv 29 \pmod{34}$	1
$k \equiv 1440 \pmod{3135}$	9	$k \equiv 769 \pmod{1980}$	7	$k \equiv 128 \pmod{204}$	1
$k \equiv 547 \pmod{6270}$	1	$k \equiv 1759 \pmod{1980}$	8	$k \equiv 17 \pmod{323}$	1
$k \equiv 5563 \pmod{6270}$	2	$k \equiv 35089 \pmod{37620}$	2	$k \equiv 188 \pmod{323}$	2
$k \equiv 4309 \pmod{6270}$	3	$k \equiv 16279 \pmod{37620}$	1	$k \equiv 36 \pmod{323}$	3
$k \equiv 625 \pmod{660}$	1	$k \equiv 956 \pmod{3960}$	1	$k \equiv 530 \pmod{646}$	1
$k \equiv 295 \pmod{660}$	2	$k \equiv 2276 \pmod{3960}$	2	$k \equiv 207 \pmod{646}$	2

Table 12.9 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 361 \pmod{660}$	5	$k \equiv 3596 \pmod{3960}$	3	$k \equiv 55 \pmod{646}$	3
$k \equiv 31 \pmod{660}$	3	$k \equiv 3332 \pmod{3960}$	4	$k \equiv 549 \pmod{646}$	4
$k \equiv 97 \pmod{660}$	4	$k \equiv 692 \pmod{3960}$	5	$k \equiv 397 \pmod{646}$	5
$k \equiv 427 \pmod{660}$	6	$k \equiv 2012 \pmod{3960}$	6	$k \equiv 245 \pmod{646}$	6
$k \equiv 493 \pmod{660}$	7	$k \equiv 37388 \pmod{75240}$	1	$k \equiv 93 \pmod{646}$	7
$k \equiv 163 \pmod{660}$	8	$k \equiv 62468 \pmod{75240}$	2	$k \equiv 378 \pmod{1292}$	1
$k \equiv 889 \pmod{12540}$	1	$k \equiv 12308 \pmod{75240}$	3	$k \equiv 226 \pmod{1292}$	2
$k \equiv 7159 \pmod{12540}$	2	$k \equiv 67484 \pmod{75240}$	4	$k \equiv 74 \pmod{1292}$	3
$k \equiv 2485 \pmod{12540}$	4	$k \equiv 17324 \pmod{75240}$	5	$k \equiv 330 \pmod{340}$	1
$k \equiv 8755 \pmod{12540}$	3	$k \equiv 42404 \pmod{75240}$	6	$k \equiv 126 \pmod{340}$	2
$k \equiv 20 \pmod{1320}$	1	$k \equiv 130 \pmod{247}$	1	$k \equiv 262 \pmod{340}$	3
$k \equiv 1076 \pmod{1320}$	2	$k \equiv 92 \pmod{247}$	3	$k \equiv 58 \pmod{340}$	4
$k \equiv 812 \pmod{1320}$	3	$k \equiv 54 \pmod{247}$	2	$k \equiv 194 \pmod{340}$	5
$k \equiv 548 \pmod{1320}$	4	$k \equiv 16 \pmod{247}$	4	$k \equiv 2354 \pmod{2584}$	1
$k \equiv 284 \pmod{1320}$	5	$k \equiv 225 \pmod{247}$	5	$k \equiv 1062 \pmod{2584}$	2
$k \equiv 932 \pmod{1320}$	6	$k \equiv 187 \pmod{494}$	1	$k \equiv 1746 \pmod{2584}$	5
$k \equiv 668 \pmod{1320}$	7	$k \equiv 149 \pmod{494}$	2	$k \equiv 454 \pmod{2584}$	3
$k \equiv 404 \pmod{1320}$	8	$k \equiv 111 \pmod{494}$	3	$k \equiv 2316 \pmod{2584}$	4
$k \equiv 22340 \pmod{25080}$	1	$k \equiv 73 \pmod{494}$	4	$k \equiv 2164 \pmod{2584}$	6
$k \equiv 351 \pmod{495}$	1	$k \equiv 434 \pmod{988}$	1	$k \equiv 2012 \pmod{2584}$	7
$k \equiv 21 \pmod{495}$	2	$k \equiv 890 \pmod{988}$	2	$k \equiv 1860 \pmod{2584}$	8
$k \equiv 186 \pmod{495}$	3	$k \equiv 358 \pmod{988}$	3	$k \equiv 4292 \pmod{5168}$	1
$k \equiv 252 \pmod{495}$	4	$k \equiv 2790 \pmod{2964}$	1	$k \equiv 1708 \pmod{5168}$	2
$k \equiv 417 \pmod{495}$	5	$k \equiv 814 \pmod{2964}$	2	$k \equiv 3684 \pmod{10336}$	1
$k \equiv 87 \pmod{495}$	6	$k \equiv 1802 \pmod{2964}$	3	$k \equiv 1100 \pmod{10336}$	2
$k \equiv 7083 \pmod{9405}$	1	$k \equiv 1916 \pmod{1976}$	1	$k \equiv 8852 \pmod{10336}$	3
$k \equiv 813 \pmod{9405}$	4	$k \equiv 396 \pmod{1976}$	2	$k \equiv 6268 \pmod{10336}$	4
$k \equiv 3948 \pmod{9405}$	2	$k \equiv 852 \pmod{1976}$	3	$k \equiv 264 \pmod{969}$	1
$k \equiv 435 \pmod{969}$	2	$k \equiv 3399 \pmod{6460}$	6	$k \equiv 1747 \pmod{6498}$	1
$k \equiv 606 \pmod{969}$	3	$k \equiv 18 \pmod{361}$	1	$k \equiv 6079 \pmod{6498}$	2
$k \equiv 948 \pmod{969}$	4	$k \equiv 37 \pmod{361}$	3	$k \equiv 3913 \pmod{6498}$	3
$k \equiv 2088 \pmod{2907}$	1	$k \equiv 56 \pmod{361}$	2	$k \equiv 1405 \pmod{4332}$	1
$k \equiv 1119 \pmod{2907}$	2	$k \equiv 75 \pmod{361}$	4	$k \equiv 3571 \pmod{4332}$	2
$k \equiv 150 \pmod{2907}$	3	$k \equiv 94 \pmod{361}$	5	$k \equiv 3229 \pmod{4332}$	3
$k \equiv 3228 \pmod{3876}$	1	$k \equiv 113 \pmod{361}$	6	$k \equiv 1063 \pmod{4332}$	4
$k \equiv 321 \pmod{3876}$	2	$k \equiv 132 \pmod{361}$	7	$k \equiv 1652 \pmod{8664}$	1
$k \equiv 1290 \pmod{3876}$	3	$k \equiv 151 \pmod{722}$	1	$k \equiv 3476 \pmod{8664}$	2
$k \equiv 2259 \pmod{3876}$	4	$k \equiv 531 \pmod{722}$	2	$k \equiv 5300 \pmod{8664}$	3
$k \equiv 1879 \pmod{1938}$	1	$k \equiv 189 \pmod{722}$	3	$k \equiv 8948 \pmod{12996}$	1
$k \equiv 1081 \pmod{1938}$	2	$k \equiv 1234 \pmod{1444}$	1	$k \equiv 284 \pmod{12996}$	2
$k \equiv 283 \pmod{1938}$	3	$k \equiv 170 \pmod{1444}$	2	$k \equiv 4616 \pmod{12996}$	3
$k \equiv 625 \pmod{1938}$	4	$k \equiv 550 \pmod{1444}$	3	$k \equiv 2108 \pmod{12996}$	5
$k \equiv 1765 \pmod{1938}$	5	$k \equiv 512 \pmod{1444}$	4	$k \equiv 6440 \pmod{12996}$	4
$k \equiv 4843 \pmod{5814}$	1	$k \equiv 892 \pmod{2888}$	1	$k \equiv 10772 \pmod{12996}$	6
$k \equiv 967 \pmod{5814}$	2	$k \equiv 5604 \pmod{5776}$	1	$k \equiv 8264 \pmod{12996}$	7
$k \equiv 2905 \pmod{5814}$	3	$k \equiv 2716 \pmod{5776}$	2	$k \equiv 12596 \pmod{12996}$	8

Table 12.9 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 6572 \pmod{7752}$	1	$k \equiv 930 \pmod{1083}$	1	$k \equiv 3932 \pmod{25992}$	1
$k \equiv 3836 \pmod{7752}$	2	$k \equiv 588 \pmod{1083}$	2	$k \equiv 14420 \pmod{25992}$	2
$k \equiv 6116 \pmod{7752}$	4	$k \equiv 246 \pmod{1083}$	4	$k \equiv 5756 \pmod{25992}$	3
$k \equiv 3380 \pmod{7752}$	3	$k \equiv 987 \pmod{1083}$	3	$k \equiv 23084 \pmod{25992}$	4
$k \equiv 644 \pmod{7752}$	5	$k \equiv 645 \pmod{1083}$	5	$k \equiv 15788 \pmod{25992}$	5
$k \equiv 815 \pmod{1615}$	1	$k \equiv 303 \pmod{1083}$	6	$k \equiv 24452 \pmod{25992}$	6
$k \equiv 1461 \pmod{1615}$	2	$k \equiv 1044 \pmod{1083}$	7	$k \equiv 7124 \pmod{51984}$	1
$k \equiv 492 \pmod{3230}$	1	$k \equiv 702 \pmod{3249}$	1	$k \equiv 33116 \pmod{51984}$	2
$k \equiv 2107 \pmod{3230}$	2	$k \equiv 1785 \pmod{3249}$	2	$k \equiv 360 \pmod{1805}$	1
$k \equiv 1138 \pmod{3230}$	3	$k \equiv 2868 \pmod{3249}$	3	$k \equiv 721 \pmod{1805}$	2
$k \equiv 2753 \pmod{6460}$	2	$k \equiv 1291 \pmod{2166}$	1	$k \equiv 1082 \pmod{1805}$	3
$k \equiv 5983 \pmod{6460}$	1	$k \equiv 949 \pmod{2166}$	2	$k \equiv 1443 \pmod{1805}$	4
$k \equiv 1784 \pmod{6460}$	3	$k \equiv 607 \pmod{2166}$	3	$k \equiv 1804 \pmod{1805}$	5
$k \equiv 169 \pmod{6460}$	4	$k \equiv 265 \pmod{2166}$	4		
$k \equiv 5014 \pmod{6460}$	5	$k \equiv 2089 \pmod{2166}$	5		

Table 12.10: Covering information for $d = -2$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod{2}$	1	$k \equiv 507 \pmod{690}$	4	$k \equiv 53 \pmod{92}$	2
$k \equiv 1 \pmod{16}$	1	$k \equiv 553 \pmod{690}$	5	$k \equiv 77 \pmod{92}$	3
$k \equiv 8 \pmod{15}$	1	$k \equiv 139 \pmod{1380}$	1	$k \equiv 31 \pmod{184}$	1
$k \equiv 0 \pmod{23}$	1	$k \equiv 369 \pmod{1380}$	2	$k \equiv 123 \pmod{184}$	2
$k \equiv 1 \pmod{345}$	1	$k \equiv 599 \pmod{1380}$	3	$k \equiv 9 \pmod{69}$	1
$k \equiv 70 \pmod{345}$	2	$k \equiv 829 \pmod{1380}$	4	$k \equiv 32 \pmod{69}$	2
$k \equiv 116 \pmod{345}$	3	$k \equiv 1059 \pmod{1380}$	5	$k \equiv 55 \pmod{69}$	3
$k \equiv 185 \pmod{345}$	4	$k \equiv 1289 \pmod{1380}$	6	$k \equiv 33 \pmod{138}$	1
$k \equiv 231 \pmod{345}$	5	$k \equiv 3 \pmod{46}$	1	$k \equiv 79 \pmod{138}$	2
$k \equiv 300 \pmod{345}$	6	$k \equiv 5 \pmod{46}$	2	$k \equiv 125 \pmod{138}$	3
$k \equiv 47 \pmod{690}$	1	$k \equiv 27 \pmod{46}$	3	$k \equiv 57 \pmod{276}$	1
$k \equiv 93 \pmod{690}$	2	$k \equiv 29 \pmod{46}$	4	$k \equiv 103 \pmod{276}$	2
$k \equiv 277 \pmod{690}$	3	$k \equiv 7 \pmod{92}$	1	$k \equiv 149 \pmod{276}$	3
$k \equiv 195 \pmod{276}$	4	$k \equiv 85 \pmod{368}$	2	$k \equiv 111 \pmod{1196}$	1
$k \equiv 241 \pmod{276}$	5	$k \equiv 269 \pmod{2576}$	1	$k \equiv 203 \pmod{1196}$	3
$k \equiv 11 \pmod{552}$	1	$k \equiv 591 \pmod{2576}$	2	$k \equiv 617 \pmod{1196}$	4
$k \equiv 287 \pmod{552}$	2	$k \equiv 2201 \pmod{2576}$	3	$k \equiv 709 \pmod{1196}$	5
$k \equiv 12 \pmod{115}$	1	$k \equiv 39 \pmod{966}$	1	$k \equiv 801 \pmod{1196}$	6
$k \equiv 35 \pmod{115}$	2	$k \equiv 361 \pmod{966}$	2	$k \equiv 893 \pmod{1196}$	7
$k \equiv 58 \pmod{115}$	3	$k \equiv 683 \pmod{966}$	3	$k \equiv 295 \pmod{2392}$	1
$k \equiv 81 \pmod{115}$	4	$k \equiv 131 \pmod{805}$	1	$k \equiv 1491 \pmod{2392}$	2
$k \equiv 104 \pmod{115}$	5	$k \equiv 292 \pmod{805}$	2	$k \equiv 88 \pmod{897}$	2
$k \equiv 105 \pmod{230}$	1	$k \equiv 453 \pmod{805}$	3	$k \equiv 387 \pmod{897}$	3
$k \equiv 151 \pmod{230}$	2	$k \equiv 775 \pmod{805}$	4	$k \equiv 686 \pmod{897}$	1

Table 12.10 (continued)

Congruence	*p*	Congruence	*p*	Congruence	*p*
$k \equiv 13 \pmod{460}$	1	$k \equiv 1419 \pmod{1610}$	1	$k \equiv 479 \pmod{1794}$	1
$k \equiv 59 \pmod{460}$	2	$k \equiv 223 \pmod{1610}$	2	$k \equiv 571 \pmod{1794}$	2
$k \equiv 197 \pmod{460}$	3	$k \equiv 545 \pmod{1610}$	3	$k \equiv 1077 \pmod{1794}$	3
$k \equiv 243 \pmod{460}$	4	$k \equiv 867 \pmod{1610}$	4	$k \equiv 1169 \pmod{1794}$	4
$k \equiv 289 \pmod{460}$	5	$k \equiv 1189 \pmod{1610}$	5	$k \equiv 1675 \pmod{1794}$	5
$k \equiv 427 \pmod{460}$	6	$k \equiv 1511 \pmod{1610}$	6	$k \equiv 1767 \pmod{1794}$	6
$k \equiv 14 \pmod{161}$	1	$k \equiv 63 \pmod{207}$	1	$k \equiv 117 \pmod{552}$	3
$k \equiv 37 \pmod{161}$	2	$k \equiv 109 \pmod{207}$	2	$k \equiv 393 \pmod{552}$	4
$k \equiv 60 \pmod{161}$	3	$k \equiv 155 \pmod{207}$	3	$k \equiv 531 \pmod{2208}$	1
$k \equiv 106 \pmod{161}$	4	$k \equiv 201 \pmod{414}$	1	$k \equiv 807 \pmod{2208}$	2
$k \equiv 129 \pmod{161}$	5	$k \equiv 247 \pmod{414}$	2	$k \equiv 1359 \pmod{2208}$	3
$k \equiv 313 \pmod{1288}$	1	$k \equiv 293 \pmod{414}$	3	$k \equiv 1635 \pmod{2208}$	4
$k \equiv 405 \pmod{1288}$	2	$k \equiv 339 \pmod{414}$	4	$k \equiv 1911 \pmod{2208}$	6
$k \equiv 635 \pmod{1288}$	3	$k \equiv 385 \pmod{414}$	5	$k \equiv 2187 \pmod{2208}$	7
$k \equiv 957 \pmod{1288}$	5	$k \equiv 17 \pmod{828}$	1	$k \equiv 255 \pmod{4416}$	1
$k \equiv 1049 \pmod{1288}$	4	$k \equiv 431 \pmod{828}$	2	$k \equiv 1083 \pmod{4416}$	2
$k \equiv 1279 \pmod{1288}$	6	$k \equiv 110 \pmod{253}$	1	$k \equiv 2463 \pmod{4416}$	3
$k \equiv 83 \pmod{644}$	1	$k \equiv 133 \pmod{253}$	2	$k \equiv 3291 \pmod{4416}$	4
$k \equiv 15 \pmod{322}$	1	$k \equiv 156 \pmod{253}$	3	$k \equiv 25 \pmod{1104}$	1
$k \equiv 61 \pmod{322}$	2	$k \equiv 179 \pmod{253}$	4	$k \equiv 163 \pmod{1104}$	2
$k \equiv 107 \pmod{322}$	3	$k \equiv 202 \pmod{253}$	5	$k \equiv 301 \pmod{1104}$	3
$k \equiv 199 \pmod{322}$	4	$k \equiv 225 \pmod{506}$	1	$k \equiv 439 \pmod{1104}$	4
$k \equiv 245 \pmod{322}$	5	$k \equiv 271 \pmod{506}$	2	$k \equiv 577 \pmod{1104}$	5
$k \equiv 291 \pmod{322}$	6	$k \equiv 501 \pmod{506}$	3	$k \equiv 715 \pmod{1104}$	6
$k \equiv 153 \pmod{644}$	2	$k \equiv 41 \pmod{1012}$	1	$k \equiv 853 \pmod{1104}$	7
$k \equiv 475 \pmod{644}$	3	$k \equiv 317 \pmod{1012}$	2	$k \equiv 2095 \pmod{4416}$	6
$k \equiv 85 \pmod{483}$	1	$k \equiv 547 \pmod{1012}$	3	$k \equiv 4303 \pmod{4416}$	5
$k \equiv 154 \pmod{483}$	2	$k \equiv 823 \pmod{2024}$	1	$k \equiv 991 \pmod{2208}$	5
$k \equiv 246 \pmod{483}$	3	$k \equiv 1835 \pmod{2024}$	2	$k \equiv 2 \pmod{621}$	1
$k \equiv 315 \pmod{483}$	4	$k \equiv 87 \pmod{759}$	1	$k \equiv 140 \pmod{621}$	2
$k \equiv 407 \pmod{483}$	5	$k \equiv 340 \pmod{759}$	2	$k \equiv 278 \pmod{621}$	3
$k \equiv 476 \pmod{483}$	6	$k \equiv 593 \pmod{759}$	3	$k \equiv 416 \pmod{621}$	4
$k \equiv 177 \pmod{483}$	7	$k \equiv 65 \pmod{299}$	1	$k \equiv 71 \pmod{1242}$	1
$k \equiv 499 \pmod{1932}$	1	$k \equiv 157 \pmod{299}$	2	$k \equiv 209 \pmod{1242}$	2
$k \equiv 821 \pmod{1932}$	2	$k \equiv 249 \pmod{299}$	3	$k \equiv 347 \pmod{1242}$	3
$k \equiv 1465 \pmod{1932}$	3	$k \equiv 341 \pmod{598}$	1	$k \equiv 1175 \pmod{1242}$	4
$k \equiv 1787 \pmod{1932}$	4	$k \equiv 433 \pmod{598}$	2	$k \equiv 485 \pmod{2484}$	1
$k \equiv 1235 \pmod{1288}$	7	$k \equiv 525 \pmod{598}$	3	$k \equiv 1727 \pmod{2484}$	2
$k \equiv 39 \pmod{368}$	1	$k \equiv 19 \pmod{1196}$	2	$k \equiv 20 \pmod{391}$	1
$k \equiv 204 \pmod{391}$	2	$k \equiv 228 \pmod{437}$	3	$k \equiv 987 \pmod{4370}$	2
$k \equiv 273 \pmod{391}$	3	$k \equiv 251 \pmod{437}$	4	$k \equiv 1861 \pmod{4370}$	3
$k \equiv 342 \pmod{391}$	4	$k \equiv 343 \pmod{437}$	5	$k \equiv 2735 \pmod{4370}$	4
$k \equiv 43 \pmod{782}$	1	$k \equiv 159 \pmod{874}$	1	$k \equiv 3609 \pmod{4370}$	5
$k \equiv 89 \pmod{782}$	2	$k \equiv 389 \pmod{874}$	2	$k \equiv 45 \pmod{460}$	7
$k \equiv 227 \pmod{782}$	3	$k \equiv 481 \pmod{874}$	3	$k \equiv 275 \pmod{460}$	8
$k \equiv 365 \pmod{782}$	4	$k \equiv 711 \pmod{874}$	4	$k \equiv 321 \pmod{1380}$	7

Table 12.10 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 549 \pmod{782}$	5	$k \equiv 803 \pmod{874}$	5	$k \equiv 781 \pmod{1380}$	9
$k \equiv 687 \pmod{782}$	6	$k \equiv 67 \pmod{1748}$	1	$k \equiv 1241 \pmod{1380}$	8
$k \equiv 181 \pmod{1564}$	1	$k \equiv 297 \pmod{1748}$	2	$k \equiv 91 \pmod{920}$	1
$k \equiv 503 \pmod{1564}$	2	$k \equiv 619 \pmod{1748}$	3	$k \equiv 551 \pmod{920}$	2
$k \equiv 963 \pmod{1564}$	3	$k \equiv 941 \pmod{1748}$	4	$k \equiv 137 \pmod{1035}$	1
$k \equiv 1285 \pmod{1564}$	4	$k \equiv 1493 \pmod{1748}$	5	$k \equiv 252 \pmod{1035}$	2
$k \equiv 2205 \pmod{6256}$	1	$k \equiv 849 \pmod{3496}$	1	$k \equiv 712 \pmod{1035}$	3
$k \equiv 3769 \pmod{6256}$	2	$k \equiv 1171 \pmod{3496}$	2	$k \equiv 367 \pmod{2070}$	1
$k \equiv 5333 \pmod{6256}$	3	$k \equiv 1723 \pmod{3496}$	3	$k \equiv 597 \pmod{2070}$	2
$k \equiv 1423 \pmod{1564}$	5	$k \equiv 2597 \pmod{3496}$	4	$k \equiv 827 \pmod{2070}$	3
$k \equiv 319 \pmod{3128}$	1	$k \equiv 2919 \pmod{3496}$	5	$k \equiv 1057 \pmod{2070}$	4
$k \equiv 1101 \pmod{3128}$	2	$k \equiv 3471 \pmod{3496}$	6	$k \equiv 1517 \pmod{2070}$	5
$k \equiv 1883 \pmod{3128}$	3	$k \equiv 90 \pmod{1311}$	1	$k \equiv 1977 \pmod{2070}$	6
$k \equiv 2665 \pmod{12512}$	1	$k \equiv 527 \pmod{1311}$	2	$k \equiv 68 \pmod{575}$	1
$k \equiv 8921 \pmod{12512}$	2	$k \equiv 964 \pmod{1311}$	3	$k \equiv 183 \pmod{575}$	2
$k \equiv 388 \pmod{1173}$	1	$k \equiv 205 \pmod{2622}$	1	$k \equiv 229 \pmod{575}$	3
$k \equiv 779 \pmod{1173}$	2	$k \equiv 757 \pmod{2622}$	2	$k \equiv 298 \pmod{575}$	4
$k \equiv 1170 \pmod{1173}$	3	$k \equiv 1079 \pmod{2622}$	3	$k \equiv 413 \pmod{575}$	5
$k \equiv 457 \pmod{2346}$	1	$k \equiv 1631 \pmod{2622}$	4	$k \equiv 528 \pmod{575}$	6
$k \equiv 1239 \pmod{2346}$	2	$k \equiv 1953 \pmod{2622}$	5	$k \equiv 459 \pmod{1150}$	1
$k \equiv 848 \pmod{1173}$	4	$k \equiv 2505 \pmod{2622}$	6	$k \equiv 689 \pmod{1150}$	2
$k \equiv 135 \pmod{2346}$	3	$k \equiv 435 \pmod{2185}$	1	$k \equiv 919 \pmod{2300}$	1
$k \equiv 1699 \pmod{2346}$	4	$k \equiv 872 \pmod{2185}$	2	$k \equiv 1149 \pmod{2300}$	2
$k \equiv 917 \pmod{4692}$	1	$k \equiv 1309 \pmod{2185}$	3	$k \equiv 2069 \pmod{2300}$	3
$k \equiv 3263 \pmod{4692}$	2	$k \equiv 1746 \pmod{2185}$	4	$k \equiv 2299 \pmod{2300}$	4
$k \equiv 21 \pmod{437}$	1	$k \equiv 2183 \pmod{2185}$	5		
$k \equiv 136 \pmod{437}$	2	$k \equiv 113 \pmod{4370}$	1		

Table 12.11: Covering information for $d = 1$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod{7}$	1	$k \equiv 11 \pmod{35}$	2	$k \equiv 12 \pmod{42}$	3
$k \equiv 1 \pmod{7}$	2	$k \equiv 18 \pmod{35}$	3	$k \equiv 33 \pmod{84}$	1
$k \equiv 9 \pmod{21}$	1	$k \equiv 25 \pmod{70}$	1	$k \equiv 75 \pmod{84}$	2
$k \equiv 2 \pmod{21}$	2	$k \equiv 60 \pmod{70}$	2	$k \equiv 19 \pmod{63}$	1
$k \equiv 16 \pmod{21}$	3	$k \equiv 32 \pmod{105}$	1	$k \equiv 40 \pmod{63}$	2
$k \equiv 3 \pmod{14}$	1	$k \equiv 67 \pmod{105}$	2	$k \equiv 61 \pmod{63}$	3
$k \equiv 10 \pmod{28}$	1	$k \equiv 102 \pmod{105}$	3	$k \equiv 6 \pmod{28}$	3
$k \equiv 24 \pmod{28}$	2	$k \equiv 5 \pmod{42}$	1	$k \equiv 13 \pmod{56}$	1
$k \equiv 4 \pmod{35}$	1	$k \equiv 26 \pmod{42}$	2	$k \equiv 41 \pmod{56}$	2
$k \equiv 20 \pmod{140}$	1	$k \equiv 132 \pmod{140}$	5	$k \equiv 111 \pmod{168}$	2
$k \equiv 48 \pmod{140}$	2	$k \equiv 27 \pmod{112}$	1	$k \equiv 167 \pmod{168}$	3
$k \equiv 76 \pmod{140}$	3	$k \equiv 83 \pmod{112}$	2		
$k \equiv 104 \pmod{140}$	4	$k \equiv 55 \pmod{168}$	1		

Table 12.12: Covering information for $d = 3$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod 3$	1	$k \equiv 611 \pmod{888}$	3	$k \equiv 361 \pmod{740}$	5
$k \equiv 1 \pmod 6$	2	$k \equiv 389 \pmod{888}$	4	$k \equiv 731 \pmod{740}$	7
$k \equiv 2 \pmod 6$	1	$k \equiv 167 \pmod{592}$	1	$k \equiv 657 \pmod{740}$	6
$k \equiv 4 \pmod{12}$	1	$k \equiv 463 \pmod{592}$	2	$k \equiv 287 \pmod{740}$	8
$k \equiv 10 \pmod{24}$	1	$k \equiv 353 \pmod{1776}$	1	$k \equiv 213 \pmod{740}$	9
$k \equiv 22 \pmod{72}$	1	$k \equiv 1241 \pmod{1776}$	2	$k \equiv 583 \pmod{740}$	10
$k \equiv 46 \pmod{72}$	2	$k \equiv 1685 \pmod{1776}$	3	$k \equiv 875 \pmod{2220}$	1
$k \equiv 70 \pmod{72}$	3	$k \equiv 797 \pmod{1776}$	4	$k \equiv 395 \pmod{2220}$	2
$k \equiv 3 \pmod{22}$	1	$k \equiv 1019 \pmod{1332}$	1	$k \equiv 2135 \pmod{2220}$	3
$k \equiv 12 \pmod{13}$	1	$k \equiv 131 \pmod{1332}$	2	$k \equiv 1655 \pmod{2220}$	4
$k \equiv 11 \pmod{28}$	1	$k \equiv 575 \pmod{1332}$	3	$k \equiv 1175 \pmod{2220}$	5
$k \equiv 5 \pmod{60}$	1	$k \equiv 1133 \pmod{1332}$	4	$k \equiv 1435 \pmod{1480}$	1
$k \equiv 0 \pmod{37}$	1	$k \equiv 467 \pmod{1332}$	5	$k \equiv 695 \pmod{1480}$	2
$k \equiv 1 \pmod{37}$	2	$k \equiv 317 \pmod{333}$	1	$k \equiv 955 \pmod{1480}$	3
$k \equiv 2 \pmod{37}$	3	$k \equiv 95 \pmod{333}$	2	$k \equiv 215 \pmod{1480}$	4
$k \equiv 3 \pmod{74}$	1	$k \equiv 206 \pmod{333}$	3	$k \equiv 475 \pmod{1480}$	5
$k \equiv 41 \pmod{74}$	2	$k \equiv 281 \pmod{333}$	4	$k \equiv 4175 \pmod{4440}$	1
$k \equiv 5 \pmod{74}$	3	$k \equiv 59 \pmod{666}$	1	$k \equiv 2729 \pmod{4440}$	2
$k \equiv 80 \pmod{111}$	1	$k \equiv 503 \pmod{666}$	2	$k \equiv 1619 \pmod{4440}$	3
$k \equiv 44 \pmod{111}$	2	$k \equiv 245 \pmod{666}$	3	$k \equiv 509 \pmod{4440}$	4
$k \equiv 8 \pmod{111}$	3	$k \equiv 23 \pmod{666}$	4	$k \equiv 3839 \pmod{4440}$	5
$k \equiv 83 \pmod{222}$	1	$k \equiv 659 \pmod{1332}$	6	$k \equiv 1731 \pmod{2960}$	1
$k \equiv 47 \pmod{222}$	2	$k \equiv 61 \pmod{185}$	1	$k \equiv 2471 \pmod{2960}$	2
$k \equiv 11 \pmod{222}$	3	$k \equiv 172 \pmod{185}$	2	$k \equiv 251 \pmod{2960}$	3
$k \equiv 197 \pmod{222}$	4	$k \equiv 98 \pmod{185}$	3	$k \equiv 991 \pmod{2960}$	4
$k \equiv 13 \pmod{148}$	1	$k \equiv 24 \pmod{185}$	4	$k \equiv 547 \pmod{2960}$	5
$k \equiv 87 \pmod{148}$	2	$k \equiv 321 \pmod{370}$	1	$k \equiv 4247 \pmod{8880}$	1
$k \equiv 125 \pmod{148}$	3	$k \equiv 247 \pmod{370}$	2	$k \equiv 2027 \pmod{8880}$	2
$k \equiv 51 \pmod{148}$	4	$k \equiv 173 \pmod{555}$	1	$k \equiv 8687 \pmod{8880}$	3
$k \equiv 89 \pmod{148}$	5	$k \equiv 284 \pmod{555}$	2	$k \equiv 1361 \pmod{2664}$	1
$k \equiv 15 \pmod{148}$	6	$k \equiv 26 \pmod{555}$	3	$k \equiv 473 \pmod{2664}$	2
$k \equiv 53 \pmod{444}$	1	$k \equiv 137 \pmod{555}$	4	$k \equiv 2249 \pmod{2664}$	3
$k \equiv 275 \pmod{444}$	2	$k \equiv 248 \pmod{555}$	5	$k \equiv 29 \pmod{2664}$	4
$k \equiv 17 \pmod{444}$	3	$k \equiv 359 \pmod{555}$	6	$k \equiv 4469 \pmod{5328}$	1
$k \equiv 239 \pmod{444}$	4	$k \equiv 101 \pmod{1110}$	1	$k \equiv 1805 \pmod{5328}$	2
$k \equiv 129 \pmod{296}$	1	$k \equiv 767 \pmod{1110}$	2	$k \equiv 917 \pmod{5328}$	3
$k \equiv 203 \pmod{296}$	2	$k \equiv 693 \pmod{740}$	1	$k \equiv 3581 \pmod{5328}$	4
$k \equiv 277 \pmod{296}$	3	$k \equiv 323 \pmod{740}$	2	$k \equiv 6023 \pmod{6660}$	1
$k \equiv 647 \pmod{888}$	2	$k \equiv 249 \pmod{740}$	3	$k \equiv 3803 \pmod{6660}$	2
$k \equiv 833 \pmod{888}$	1	$k \equiv 619 \pmod{740}$	4	$k \equiv 1583 \pmod{6660}$	3
$k \equiv 29 \pmod{1665}$	1	$k \equiv 1622 \pmod{2775}$	6	$k \equiv 737 \pmod{1554}$	2
$k \equiv 1139 \pmod{1665}$	2	$k \equiv 1178 \pmod{2775}$	7	$k \equiv 1403 \pmod{1554}$	4
$k \equiv 584 \pmod{1665}$	3	$k \equiv 1733 \pmod{5550}$	1	$k \equiv 1181 \pmod{1554}$	5
$k \equiv 1991 \pmod{3330}$	1	$k \equiv 5063 \pmod{5550}$	2	$k \equiv 293 \pmod{1554}$	6
$k \equiv 3101 \pmod{3330}$	2	$k \equiv 2843 \pmod{5550}$	3	$k \equiv 809 \pmod{1036}$	1
$k \equiv 881 \pmod{3330}$	3	$k \equiv 623 \pmod{5550}$	4	$k \equiv 921 \pmod{1036}$	2
$k \equiv 2657 \pmod{3330}$	4	$k \equiv 179 \pmod{5550}$	5	$k \equiv 2069 \pmod{3108}$	1
$k \equiv 437 \pmod{3330}$	5	$k \equiv 3509 \pmod{5550}$	6	$k \equiv 220 \pmod{407}$	1

Table 12.12 (continued)

Congruence	p	Congruence	p	Congruence	p
$k \equiv 1547 \pmod{3330}$	6	$k \equiv 1289 \pmod{3700}$	1	$k \equiv 331 \pmod{407}$	2
$k \equiv 178 \pmod{925}$	3	$k \equiv 3139 \pmod{3700}$	2	$k \equiv 35 \pmod{407}$	3
$k \equiv 733 \pmod{925}$	2	$k \equiv 10169 \pmod{11100}$	4	$k \equiv 257 \pmod{407}$	4
$k \equiv 363 \pmod{925}$	1	$k \equiv 4619 \pmod{11100}$	1	$k \equiv 368 \pmod{407}$	5
$k \equiv 918 \pmod{925}$	5	$k \equiv 7949 \pmod{11100}$	3	$k \equiv 72 \pmod{407}$	6
$k \equiv 548 \pmod{925}$	4	$k \equiv 2399 \pmod{11100}$	2	$k \equiv 183 \pmod{814}$	1
$k \equiv 104 \pmod{925}$	6	$k \equiv 217 \pmod{259}$	1	$k \equiv 701 \pmod{814}$	3
$k \equiv 659 \pmod{925}$	7	$k \equiv 106 \pmod{259}$	2	$k \equiv 405 \pmod{814}$	2
$k \equiv 289 \pmod{925}$	8	$k \equiv 254 \pmod{259}$	3	$k \equiv 109 \pmod{814}$	4
$k \equiv 844 \pmod{925}$	9	$k \equiv 143 \pmod{259}$	4	$k \equiv 517 \pmod{814}$	5
$k \equiv 1399 \pmod{1850}$	1	$k \equiv 180 \pmod{259}$	5	$k \equiv 221 \pmod{1221}$	1
$k \equiv 401 \pmod{1850}$	2	$k \equiv 69 \pmod{518}$	1	$k \equiv 332 \pmod{1221}$	2
$k \equiv 31 \pmod{1850}$	3	$k \equiv 329 \pmod{518}$	2	$k \equiv 554 \pmod{1221}$	3
$k \equiv 1511 \pmod{1850}$	4	$k \equiv 218 \pmod{777}$	1	$k \equiv 665 \pmod{1221}$	4
$k \equiv 1141 \pmod{1850}$	5	$k \equiv 107 \pmod{777}$	2	$k \equiv 776 \pmod{1221}$	5
$k \equiv 2621 \pmod{2775}$	1	$k \equiv 773 \pmod{777}$	3	$k \equiv 887 \pmod{1221}$	6
$k \equiv 2177 \pmod{2775}$	2	$k \equiv 551 \pmod{777}$	4	$k \equiv 2219 \pmod{2442}$	1
$k \equiv 2732 \pmod{2775}$	3	$k \equiv 440 \pmod{777}$	5	$k \equiv 1109 \pmod{2442}$	2
$k \equiv 512 \pmod{2775}$	5	$k \equiv 959 \pmod{1554}$	1	$k \equiv 2441 \pmod{2442}$	3
$k \equiv 1067 \pmod{2775}$	4	$k \equiv 71 \pmod{1554}$	3		

Table 12.13: Covering information for $d = 4$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 1 \pmod{3}$	1	$k \equiv 53 \pmod{54}$	2	$k \equiv 159 \pmod{162}$	4
$k \equiv 0 \pmod{6}$	2	$k \equiv 15 \pmod{81}$	1	$k \equiv 9 \pmod{36}$	1
$k \equiv 5 \pmod{6}$	1	$k \equiv 33 \pmod{81}$	2	$k \equiv 27 \pmod{144}$	1
$k \equiv 2 \pmod{9}$	1	$k \equiv 51 \pmod{81}$	3	$k \equiv 99 \pmod{144}$	2
$k \equiv 14 \pmod{18}$	1	$k \equiv 69 \pmod{81}$	4	$k \equiv 63 \pmod{216}$	1
$k \equiv 3 \pmod{18}$	2	$k \equiv 6 \pmod{81}$	5	$k \equiv 135 \pmod{216}$	2
$k \equiv 17 \pmod{27}$	1	$k \equiv 105 \pmod{162}$	1	$k \equiv 207 \pmod{432}$	1
$k \equiv 8 \pmod{27}$	2	$k \equiv 123 \pmod{162}$	2	$k \equiv 423 \pmod{432}$	2
$k \equiv 26 \pmod{54}$	1	$k \equiv 141 \pmod{162}$	3		

Table 12.14: Covering information for $d = 6$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod{5}$	1	$k \equiv 27 \pmod{30}$	3	$k \equiv 4 \pmod{15}$	2
$k \equiv 1 \pmod{5}$	2	$k \equiv 8 \pmod{40}$	2	$k \equiv 29 \pmod{45}$	1
$k \equiv 2 \pmod{10}$	1	$k \equiv 28 \pmod{40}$	1	$k \equiv 14 \pmod{45}$	2
$k \equiv 3 \pmod{20}$	1	$k \equiv 18 \pmod{60}$	1	$k \equiv 89 \pmod{90}$	1
$k \equiv 13 \pmod{20}$	2	$k \equiv 58 \pmod{60}$	2	$k \equiv 44 \pmod{90}$	2
$k \equiv 7 \pmod{30}$	1	$k \equiv 38 \pmod{60}$	3		
$k \equiv 17 \pmod{30}$	2	$k \equiv 9 \pmod{15}$	1		

Table 12.15: Covering information for $d = 7$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 16 \pmod{66}$	2	$k \equiv 95 \pmod{275}$	2	$k \equiv 32 \pmod{121}$	3
$k \equiv 136 \pmod{264}$	6	$k \equiv 150 \pmod{275}$	3	$k \equiv 43 \pmod{121}$	4
$k \equiv 202 \pmod{264}$	7	$k \equiv 29 \pmod{165}$	1	$k \equiv 54 \pmod{242}$	1
$k \equiv 70 \pmod{528}$	1	$k \equiv 84 \pmod{165}$	2	$k \equiv 175 \pmod{242}$	2
$k \equiv 268 \pmod{528}$	2	$k \equiv 139 \pmod{165}$	3	$k \equiv 65 \pmod{242}$	3
$k \equiv 334 \pmod{528}$	3	$k \equiv 51 \pmod{220}$	5	$k \equiv 186 \pmod{242}$	4
$k \equiv 532 \pmod{1584}$	1	$k \equiv 161 \pmod{220}$	6	$k \equiv 76 \pmod{242}$	5
$k \equiv 4 \pmod{4752}$	1	$k \equiv 106 \pmod{330}$	1	$k \equiv 197 \pmod{484}$	1
$k \equiv 1588 \pmod{4752}$	2	$k \equiv 216 \pmod{330}$	2	$k \equiv 439 \pmod{484}$	2
$k \equiv 3 \pmod{16}$	2	$k \equiv 326 \pmod{330}$	3	$k \equiv 87 \pmod{484}$	3
$k \equiv 11 \pmod{32}$	1	$k \equiv 8 \pmod{77}$	1	$k \equiv 208 \pmod{484}$	4
$k \equiv 27 \pmod{32}$	2	$k \equiv 19 \pmod{77}$	2	$k \equiv 98 \pmod{363}$	1
$k \equiv 0 \pmod{11}$	1	$k \equiv 30 \pmod{77}$	3	$k \equiv 219 \pmod{363}$	2
$k \equiv 1 \pmod{11}$	2	$k \equiv 41 \pmod{77}$	4	$k \equiv 109 \pmod{726}$	1
$k \equiv 10 \pmod{16}$	1	$k \equiv 52 \pmod{154}$	1	$k \equiv 230 \pmod{726}$	2
$k \equiv 2 \pmod{22}$	1	$k \equiv 129 \pmod{154}$	2	$k \equiv 351 \pmod{726}$	3
$k \equiv 13 \pmod{22}$	2	$k \equiv 63 \pmod{154}$	3	$k \equiv 593 \pmod{1452}$	1
$k \equiv 3 \pmod{22}$	3	$k \equiv 140 \pmod{154}$	4	$k \equiv 1319 \pmod{4356}$	1
$k \equiv 14 \pmod{44}$	1	$k \equiv 74 \pmod{154}$	5	$k \equiv 2771 \pmod{4356}$	2
$k \equiv 36 \pmod{44}$	2	$k \equiv 151 \pmod{154}$	6	$k \equiv 2166 \pmod{5808}$	1
$k \equiv 15 \pmod{33}$	1	$k \equiv 9 \pmod{99}$	1	$k \equiv 5070 \pmod{5808}$	2
$k \equiv 26 \pmod{33}$	2	$k \equiv 20 \pmod{99}$	2	$k \equiv 2892 \pmod{5808}$	3
$k \equiv 37 \pmod{66}$	1	$k \equiv 31 \pmod{99}$	3	$k \equiv 5796 \pmod{5808}$	4
$k \equiv 38 \pmod{132}$	1	$k \equiv 42 \pmod{99}$	4	$k \equiv 120 \pmod{605}$	1
$k \equiv 126 \pmod{132}$	2	$k \equiv 53 \pmod{198}$	1	$k \equiv 362 \pmod{1210}$	1
$k \equiv 60 \pmod{132}$	3	$k \equiv 152 \pmod{198}$	2	$k \equiv 967 \pmod{1210}$	2
$k \equiv 5 \pmod{88}$	1	$k \equiv 64 \pmod{396}$	1	$k \equiv 1088 \pmod{2420}$	5
$k \equiv 49 \pmod{88}$	2	$k \equiv 163 \pmod{396}$	2	$k \equiv 2298 \pmod{2420}$	6
$k \equiv 104 \pmod{264}$	1	$k \equiv 262 \pmod{396}$	3	$k \equiv 604 \pmod{2420}$	1
$k \equiv 236 \pmod{264}$	2	$k \equiv 361 \pmod{792}$	1	$k \equiv 1209 \pmod{2420}$	2
$k \equiv 71 \pmod{264}$	3	$k \equiv 757 \pmod{792}$	2	$k \equiv 1814 \pmod{2420}$	3
$k \equiv 159 \pmod{264}$	4	$k \equiv 75 \pmod{297}$	1	$k \equiv 2419 \pmod{2420}$	4
$k \equiv 247 \pmod{264}$	5	$k \equiv 174 \pmod{297}$	2	$k \equiv 1060 \pmod{1584}$	2
$k \equiv 6 \pmod{55}$	1	$k \equiv 273 \pmod{297}$	3	$k \equiv 3172 \pmod{4752}$	3
$k \equiv 17 \pmod{55}$	2	$k \equiv 86 \pmod{297}$	4	$k \equiv 205 \pmod{275}$	4
$k \equiv 28 \pmod{55}$	3	$k \equiv 185 \pmod{297}$	5	$k \equiv 260 \pmod{275}$	5
$k \equiv 39 \pmod{55}$	4	$k \equiv 284 \pmod{297}$	6	$k \equiv 329 \pmod{484}$	5
$k \equiv 50 \pmod{110}$	1	$k \equiv 97 \pmod{594}$	1	$k \equiv 450 \pmod{484}$	6
$k \equiv 105 \pmod{110}$	2	$k \equiv 196 \pmod{594}$	2	$k \equiv 340 \pmod{363}$	3
$k \equiv 7 \pmod{110}$	3	$k \equiv 295 \pmod{594}$	3	$k \equiv 472 \pmod{726}$	4
$k \equiv 62 \pmod{110}$	4	$k \equiv 394 \pmod{594}$	4	$k \equiv 4223 \pmod{4356}$	3
$k \equiv 18 \pmod{220}$	1	$k \equiv 493 \pmod{594}$	5	$k \equiv 714 \pmod{2904}$	1
$k \equiv 73 \pmod{220}$	2	$k \equiv 592 \pmod{1188}$	1	$k \equiv 1440 \pmod{2904}$	2
$k \equiv 128 \pmod{220}$	3	$k \equiv 1186 \pmod{1188}$	2	$k \equiv 241 \pmod{605}$	2
$k \equiv 183 \pmod{220}$	4	$k \equiv 10 \pmod{121}$	1	$k \equiv 483 \pmod{1210}$	3
$k \equiv 40 \pmod{275}$	1	$k \equiv 21 \pmod{121}$	2		

Table 12.16: Covering information for $d = 9$.

Congruence	p	Congruence	p	Congruence	p
$k \equiv 0 \pmod 2$	1	$k \equiv 1 \pmod 8$	1	$k \equiv 5 \pmod 8$	2
$k \equiv 3 \pmod 4$	1				

Bibliography

[1] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, The Erdős–Selfridge problem with square-free moduli, *Algebra Number Theory*, **15** (2021), 609–626.

[2] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe and M. Tiba, On the Erdős Covering Problem: the density of the uncovered set, arXiv:1811.03547.

[3] W. D. Banks, T. Freiberg and C. L. Turnage-Butterbaugh, Consecutive primes in tuples, *Acta Arith.*, **167** (2015), 261–266.

[4] J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckerman and S. S. Wagstaff Jr., *Factorizations of* $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ *Up to High Powers*, 3rd ed., Contemp. Math., vol. **22**, American Math. Soc., Providence, 2002 (available online).

[5] A. Dubickas, A. Novikas and J. Šiurys, A binary linear recurrence sequence of composite numbers, *J. Number Theory*, **130** (2010), 1737–1749.

[6] P. Erdős, On integers of the form $2^k + p$ and some related problems, *Summa Bras. Math.*, **2** (1950), 113–123.

[7] P. Erdős, Solution to problem 1029: Erdos and the computer, *Math. Mag.*, **52** (1979), 180–181.

[8] M. Filaseta and J. Juillerat, Data for "Consecutive primes which are widely digitally delicate" (containing data for Table 12.4 and Tables 12.5–12.16), https://people.math.sc.edu/filaseta/ConsecutiveWDDPrimes.html.

[9] M. Filaseta, C. Finch and M. Kozek, On powers associated with Sierpiński numbers, Riesel numbers and Polignac's conjecture, *J. Number Theory*, **128** (2008), 1916–1940.

[10] M. Filaseta, J. Juillerat and J. Southwick, Widely Digitally Stable Numbers, in M. Nathanson (ed.) *Combinatorial and Additive Number Theory IV*, Springer Proc. Math. Stat., vol. **347**, pp. 161–193, Springer, Cham, 2021.

[11] M. Filaseta, M. Kozek, C. Nicol and J. Selfridge, Composites that remain composite after changing a digit, *J. Comb. Number Theory*, **2** (2010), 25–36.

[12] M. Filaseta and J. Southwick, Primes that become composite after changing an arbitrary digit, *Math. Comput.*, **90** (2021), 979–993.

[13] T. Freiberg, A note on the Theorem of Maynard and Tao, arXiv:1311.5319.

[14] R. Graham, A Fibonacci-like sequence of composite numbers, *Math. Mag.*, **37** (1964), 322–324.

[15] J. Hopper and P. Pollack, Digitally delicate primes, *J. Number Theory*, **168** (2016), 247–256.

[16] R. D. Hough, Solution of the minimum modulus problem for covering systems, *Ann. Math.*, **181** (2015), 361–382.

[17] R. D. Hough and P. P. Nielsen, Covering systems with restricted divisibility, *Duke Math. J.*, **168** (2019), 3261–3295.

[18] D. Ismailescu, A. Ko, C. Lee and J. Y. Park, On second-order linear sequences of composite numbers, *J. Integer Seq.*, **22** (2019), Article #19.7.2, 16 pp.

[19] J. Juillerat, *Widely digitally stable numbers and irreducibility criteria for polynomials with prime values*, University of South Carolina, dissertation, 2021.

[20] M. S. Klamkin, Problem 1029, *Math. Mag.*, **51** (1978), 69.

[21] D. Knuth, A Fibonacci-like sequence of composite numbers, *Math. Mag.*, **63** (1990), 21–25.

[22] S. V. Konyagin, Numbers that become composite after changing one or two digits, *Pioneer J. Algebra Number Theory Appl.*, **6** (2013), 1–7.
[23] I. Lunev, A tribonacci-like sequence of composite numbers, *J. Integer Seq.*, **20** (2017), Article #17.3.2, 6 pp.
[24] J. Maynard, Dense clusters of primes in subsets, *Compos. Math.*, **152** (2016), 1517–1554.
[25] S. Nadis, Mathematicians find a new class of digitally delicate primes. *Quanta Mag.*, March 30, 2021. https://www.quantamagazine.org/mathematicians-find-a-new-class-of-digitally-delicate-primes-20210330/.
[26] J. W. Nicol, A Fibonacci-like sequence of composite numbers, *Electron. J. Comb.*, **6** (1999), #R44.
[27] M. Parker, Stand-Up Maths: How do you prove a prime is infinitely fragile?, July 28, 2021. https://www.youtube.com/watch?v=p3Khnx0lUDE.
[28] H. Riesel, Några stora primtal, *Elementa*, **39** (1956), 258–260.
[29] D. K. L. Shiu, Strings of congruent primes, *J. Lond. Math. Soc.*, **61** (2000), 359–373.
[30] W. Sierpiński, Sur un problème concernant les nombres $k \cdot 2^n + 1$, *Elem. Math.*, **15** (1960), 73–74.
[31] N. J. A. Sloane (ed.), *The On-Line Encyclopedia of Integer Sequences*, 2020. Published electronically at https://oeis.org/A076335.
[32] T. Tao, A remark on primality testing and decimal expansions, *J. Aust. Math. Soc.*, **91** (2011), 405–413.
[33] M. Vsemirnov, A new Fibonacci-like sequence of composite numbers, *J. Integer Seq.*, **7** (2004), Article #04.3.7, 3 pp.

Jerrold R. Griggs

Spanning trees and domination in hypercubes

Abstract: Let $L(G)$ denote the maximum number of leaves in any spanning tree of a connected graph G. We show the (known) result that for the n-cube Q_n, $L(Q_n) \sim 2^n = |V(Q_n)|$ as $n \to \infty$. Examining this more carefully, consider the minimum size of a connected dominating set of vertices $\gamma_c(Q_n)$, which is $2^n - L(Q_n)$ for $n \geq 2$. We show that $\gamma_c(Q_n) \sim 2^n/n$, which rather surprisingly is no larger than the asymptotic behavior of the domination number $\gamma(Q_n)$. We use Hamming codes and an "expansion" method to construct leafy spanning trees in Q_n.

1 Introduction

The n-cube graph Q_n has 2^n vertices, the strings $a_1 \dots a_n$ on n bits, where two vertices are adjacent if and only if their strings differ in exactly one coordinate (where one vertex has 0 and the other has 1). The n-cube is frequently used as a structure for computer networks, where there are 2^n processors corresponding to the vertices of Q_n. An efficient way to connect all of the processors, so that they all communicate with each other, is to take a spanning tree in Q_n.

With this in mind, S. Bezrukov imagined it would be interesting to construct such spanning trees with many leaves (degree one vertices). At the IWOCA conference (Duluth, 2014), Bezrukov proposed the following problem: Letting $L(G)$ denote the maximum number of leaves in any spanning tree of a connected simple graph G, what can one say about $L(Q_n)$? He shared this problem in notes [2].

Acknowledgement: This paper is dedicated to the memory of Ron Graham. After reading about him often in Martin Gardner's columns, and studying his papers in college, it was a thrill when I realized the person asking insightful questions at my first conference talk was Ron. Over the years, I heard him give many charming colloquia and conference talks, learned news from him about math results and math people, and played with the latest puzzles and gadgets he shared. When I presented my early ideas on this spanning tree problem, at the 2015 conference in his honor, I was proud to see him there again. He had a tremendous impact on so many of us. We express gratitude to Sergei Bezrukov for sharing his questions and notes that led to this study. Particular thanks go to colleague Joshua Cooper (himself a student of Ron and Fan Chung) for bringing attention to the key domination result $\gamma(Q_n) \sim 2^n/n$ as presented in [5]. This research was supported in part by a grant from the Simons Foundation (#282896 to Jerrold Griggs). We appreciate support for travel to Taiwan to work on this project from Mathematics Research Promotion Center Grant 108-17 and Ministry of Science and Technology Grant 107-2115-M-005-002-MY2.

Jerrold R. Griggs, Department of Mathematics, University of South Carolina, Columbia, SC, USA, e-mail: griggs@math.sc.edu

https://doi.org/10.1515/9783110754216-013

For a spanning tree, the nonleaf vertices are connected, so form a tree themselves, which we may think of as the backbone of the tree: All vertices are either in this backbone, or are leaves adjacent to it. Bezrukov's question then is equivalent to constructing a spanning tree of the hypercube with the smallest backbone.

Notice that the opposite question, finding the *minimum* number of leaves in a spanning tree, is easy: By a simple induction, Q_n has a Hamilton path for all $n \geq 1$. This path is a spanning tree with just two leaves. We are interested in the other extreme, *maximizing* the number of leaves.

Our problem is closely related to the subject of domination in graphs. A subset W of the vertex set V of a graph $G = (V, E)$ is a *dominating set* if every vertex is either in W or adjacent to some vertex in W. The *domination number* $\gamma(G)$ is the minimum size of any dominating set.

Note that if one pulls off the leaves from a spanning tree T for a connected graph $G = (V, E)$ with at least three vertices, then the remaining vertices W form a dominating set, and, moreover, what remains of T still connects them. That is, W forms a connected dominating set. Conversely, from any connected dominating set we can span them with a tree and attach any other vertices as leaves to obtain a spanning tree. The minimum size of a connected dominating set of G is called the *connected domination number* $\gamma_c(G)$.

We see that maximizing the number of leaves of any spanning tree of such G corresponds to minimizing the size of a connected dominating set. From this discussion, we obtain for such G

$$L(G) + \gamma_c(G) = |V(G)|.$$

The simple ordering relationship between these parameters is

$$1 \leq \gamma(G) \leq \gamma_c(G) \leq |V|.$$

For example, one can readily check that for the four-cycle Q_2, $\gamma = \gamma_c = L = 2$, while for the ordinary cube Q_3, $\gamma = 2, \gamma_c = L = 4$. For larger n, more than half the vertices can be leaves.

The earliest paper we can find that investigates the connected domination number of a graph is by Sampathkumar and Walikar (1979) [15]. Several studies investigate bounding $L(G)$ for classes of graphs G, such as those with given minimum degree [8, 9, 12, 16]. Caro et al. [3] study both parameters, and provide more references. Many papers concern algorithms for finding leafy trees (or small connected dominating sets).

Searching online, we discovered several papers concerning domination in hypercubes. These were often done independently of other studies. The 1990 dissertation of Jha [10] gives a good general upper bound on $\gamma(Q_n)$, which is just twice the easy lower bound. Arumugam and Kala [1] (1998) focus on domination in hypercubes. Duckworth et al. [6](2001) give good general bounds on $L(Q_n)$. It follows that $L(Q_n) \sim 2^n$. It means

asymptotically there is a spanning tree for the hypercube in which almost all vertices are leaves. It is nicer to restate their results in terms of connected domination.

Theorem 1 ([6]).
- *Lower bound: For* $n \geq 2$, $\frac{\gamma_c(Q_n)}{2^n} \geq \frac{1}{n}$
- *Upper bound: As* $n \to \infty$, $\frac{\gamma_c(Q_n)}{2^n} \leq (1 + o(1))\frac{2}{n}$

Another 2012 study of hypercubes [4] gives values of $\gamma_c(Q_n)$ for small n, but unfortunately its formula for general n, stated without proof, is far from correct. Mane and Waphare [14] investigate several generalizations of domination numbers of hypercubes. The 2017 Master's thesis of Kuboň [13] considers domination in hypergraphs, and uses some of the same methods as in this paper.

In the next section, we present simple general lower bounds on $\gamma(Q_n)$ and $\gamma_c(Q_n)$. In Section 3, we describe the Hamming code construction that gives a "perfect dominating set" for Q_n when n is of the form $2^k - 1$. We give a method to produce a small connected dominating set, given a dominating set, that leads to an upper bound on $\gamma_c(Q_n)$ for $n = 2^k - 1$. A simple inductive method we call doubling is used to give upper bounds on $\gamma(Q_n)$ and $\gamma_c(Q_n)$ for general n in Section 4.

Where we make new progress is by introducing in Section 5, a new method we call expansion, in which we take a minimum dominating set in each of 2^j copies of Q_N and connect them appropriately to obtain a small connected dominating set in Q_n, where $n = N + j$. Choosing N and j wisely improves the best previous upper bound on $\gamma_c(Q_n)$ by a factor of 2. Indeed, in Section 6 we settle the leading asymptotic behavior of $\gamma_c(Q_n)$.

Theorem 2. *As* $n \to \infty$, $\frac{\gamma_c(Q_n)}{2^n} = (1 + o(1))\frac{1}{n}$.

Restating this in terms of the maximum number of leaves, it means

$$L(Q_n) = \left(1 - \frac{1}{n} + o\left(\frac{1}{n}\right)\right)2^n.$$

We conclude with suggestions for further study and acknowledgments of valuable ideas and support of this project.

2 Domination lower bounds

Let us note some easy lower bounds on our domination parameters for the hypercube Q_n.

Proposition 1.
- *For* $n \geq 1$, $\gamma(Q_n) \geq 2^n/(n+1)$.
- *For* $n \geq 2$, $\gamma_c(Q_n) \geq (2^n - 2)/(n-1) \geq 2^n/n$.

Proof. A single vertex can dominate at most itself and its n neighbors, leading to the lower bound on $\gamma(Q_n)$.

Next, consider a connected dominating set of Q_n of size c. There is a tree T on these c vertices using $c-1$ edges from Q_n. The sum of degrees of these c vertices has $2c-2$ accounted for by T. It means that the number of additional vertices (dominated by those in T) is at most $nc-2(c-1)$. But there are 2^n-c vertices besides T. Rearranging terms gives the stated inequality on c, hence the lower bounds on $\gamma_c(Q_n)$. □

3 Hamming code

The famous Hamming code gives an elegant construction of a "perfect dominating set" in Q_n when $n = 2^k - 1$ for some integer $k \geq 1$. This means it achieves the lower bound on $\gamma(Q_n)$ in Proposition 1. Viewing the vertices of Q_n for such an n as n-dimensional vectors over $GF(2)$, the code consists of the 2^{n-k} vectors in the row space of a $(n-k)\times n$ matrix built as follows: The first $n-k$ columns form the identity matrix, while the rows of the other k columns consist of all $n-k = 2^k - k - 1$ vectors of length k with weight (number of ones) at least 2. The difference between any two vectors in this row space is then a nonzero vector in the row space, and hence a nonempty sum of rows of the matrix. By design, such a sum will always have weight at least three.

Consequently, the 2^{n-k} stars in Q_n that are centered at the vectors in the row space are disjoint. Each star is a $K_{1,n}$. By counting, we see that these stars partition the vertices of Q_n. They form a minimum dominating set for Q_n.

As Bezrukov pointed out when he proposed his problem about $L(Q_n)$, for such n we only have to add some edges between leaves of different stars to obtain a spanning tree with many leaves. After all, Q_n is connected, and all edges not used in the stars are between leaves of stars (different stars, in fact). If we have c components, we need to add $c-1$ edges to obtain a spanning tree; here, $c = 2^{n-k}$. At worst, each additional edge costs us two new leaves–it would be less, if we are able to use several edges from the same leaf. When we finish, we have a spanning tree where the non-leaves are the c star centers from the Hamming code, as well as at most $2c-2$ vertices that were star leaves.

In fact, we can use this method for any connected simple graph G to build a spanning tree. Starting from a minimum dominating set of c vertices, the stars centered at those vertices cover the entire vertex set (though in general they are not disjoint, and dominating vertices could even be adjacent). One can add at most $c-1$ edges between stars to create a spanning tree. We obtain this general bound.

Proposition 2. *Let G be a connected simple graph. Then*

$$\gamma_c(G) \leq 3\gamma(G) - 2.$$

Applying this to our Hamming code construction, we obtain the following.

Proposition 3. *Let $n = 2^k - 1$, where the integer $k \geq 1$. Then $\gamma(Q_n) = 2^{n-k} = 2^n/(n+1)$, and $\gamma_c(Q_n) < (3/(n+1))2^n$.*

For this Hamming code case $n = 2^k - 1$, our tree construction can be viewed this way: Starting from a perfect dominating set in Q_n, we take the corresponding $C = 2^n/(n+1)$ stars $K_{1,n}$ and add $C - 1$ edges to form a tree with many leaves. Since all edges for the star centers are used already, each edge we add will join leaves from two different stars. At worst, we give up $2(C-1)$ star leaves (they become part of the tree backbone), plus the backbone contains the C star centers. This gives us a connected dominating set of size at most $3C - 2 \sim 3(2^n/n)$.

If we are fortunate, we do not have to pick two new leaves for each successive additional edge: It could be that one or both leaves are already in the backbone. However, for each of the C stars we must give up at least one leaf, in order that the stars connect in the spanning tree. It means that the connected dominating set we construct will have at least $2C \sim 2(2^n/n)$ vertices.

4 Doubling

So far, we have constructed leafy trees in the n-cube only when n has the special form $2^k - 1$. One can view the $(n+1)$-cube Q_{n+1} as consisting of two copies of Q_n along with a matching in which each vertex of one Q_n is on an edge with its corresponding vertex in the other Q_n. This is true for any value of n, not just the special values where the Hamming code exists.

If we take a dominating set for each copy of Q_n, we get a dominating set for Q_{n+1}. Moreover, if we take the same connected dominating set for each copy, it gives a dominating set for Q_{n+1} that is connected. We see this simply by adding the edge joining the two copies of a vertex in the connected dominating set for Q_n. We record these observations about doubling next.

Proposition 4. *For all $n \geq 1$, $\gamma(Q_{n+1}) \leq 2\gamma(Q_n)$, and $\gamma_c(Q_{n+1}) \leq 2\gamma_c(Q_n)$.*

Now suppose n is between two consecutive values where the Hamming code construction is the last section applies, say $n = N+j$, where $k \geq 1$, $N = 2^k - 1$, and $0 \leq j \leq 2^k$. We apply the doubling proposition j times, starting from Q_N, and obtain

$$\gamma(Q_n) \leq 2^j \frac{2^N}{N+1} = \frac{N+j}{N+1}\frac{2^n}{n} < 2\frac{2^n}{n}.$$

It follows that

$$\gamma(Q_n)/2^n < 2/n \to 0,$$

as $n \to \infty$. This matches the bound given by Jha [10].

For connected domination, we apply Proposition 2 and obtain

$$\gamma_c(Q_n) < 3\gamma(Q_n) < 6\frac{2^n}{n}.$$

It follows that

$$\frac{\gamma_c(Q_n)}{2^n} < \frac{6}{n} \to 0,$$

as $n \to \infty$, confirming our earlier assertion that there are spanning trees for hypercubes with almost all vertices being leaves. Of course, Theorem 1 got a better bound than this on $\gamma_c(Q_n)/2^n$; our main result will do even better.

Let us summarize our findings so far. The domination problem for Q_n is solved by the Hamming code for $n = N = 2^k - 1$. Then as $n = N + j$ grows with $j, 0 \le j \le 2^k$, our upper bound on $\gamma(Q_n)/(2^n/n)$ grows from around 1 to around 2. However, at $j = 2^k$, we have the next Hamming code case, $n = 2^{k+1} - 1$, and it is better to switch again to the Hamming code construction. It means we have a sawtooth function upper bound, rising from 1 to 2 as n increases, then abruptly dropping back down to 1 and rising again. Of course, each tooth covers an interval of length about 2^k, so the teeth get wider with k.

Owing to our upper bound Proposition 2, for connected domination $\gamma_c(Q_n)$ has a similar sawtooth upper bound, but each tooth rises from value 3 to 6.

5 Expansion

We introduce a new method of tree construction that takes advantage of small dominating sets to produce smaller connected dominating sets in Q_n. This will bring down our upper bound for connected domination, and eventually allow us to solve our problem asymptotically.

For constructing a spanning tree, the Hamming code bound punished us by potentially using up so many leaves to connect the stars. If we repeatedly double the construction, then it repeats this penalty over and over. A better idea could then be to select one copy (or "layer") of the base hypercube, add edges to connect the stellar clusters in just that layer, and then connect all the copies of each star center to the one in the special layer.

Describing this explicitly, let $N = 2^k - 1$, and $n = N + j$, where $0 \le j \le 2^k$. Partition the vertices of Q_n into 2^j "layers" according to the last j coordinates of the vertices $(a_1, \ldots, a_n)$. Each layer induces a Q_N, and its vertices are partitioned into $|C| = 2^{N-k}$ stars, according to the Hamming code partition of Q_N. For each star S in the partition of Q_N, there are 2^j vertices, one in each layer, that are centers of the stars corresponding to star S. The centers all agree in their first N coordinates, so together induce a

subgraph Q_j. By adding $2^j - 1$ edges these stars (copies of S) can be connected into a tree. We now have a forest of $|C| = 2^{N-k}$ such trees.

We connect these trees by adding $|C|-1$ edges, which may as well all be in the layer ending with 0's. Each such edge adds at most two vertices to the connected dominating set we construct. It is similar to how we connected the stars in the Hamming code construction. We record the result of our expansion construction.

Proposition 5. *Let $n = N + j$, where $N = 2^k - 1$ and $1 \le j \le 2^k$. Then $\gamma_c(Q_n) \le 2^j|C| + 2(|C| - 1)$, where C is the set of 2^{N-k} codewords for the Hamming code in Q_N.*

We have seen that $\gamma_c(Q_n)/2^n \ge 1/n$ for all $n \ge 2$. It would be nice if we could find a tree construction for Q_n that has so many leaves that its backbone (connected dominating set) comes close to achieving the lower bound, acting asymptotically like a perfect dominating set: What we want is that $\gamma_c(Q_n)/(2^n/n) \to 1$ as $n \to \infty$. Expansion allows us to come much closer to this goal. The next result is what we can show now.

Theorem 3. *For all $n \ge 1$, $\gamma_c(Q_n)/2^n < 2/n$. For all $n \ge 3$, $\gamma_c(Q_n)/2^n > 1/n$. We have* $\liminf_{n\to\infty} \gamma_c(Q_n)/(2^n/n) = 1$.

Proof. We have n, N, K, j as above. Proposition 5 gives us

$$\begin{aligned}\gamma_c(Q_n) &\le 2^j|C| + 2(|C| - 1)\\ &< (2^j + 2)|C|\\ &= (2^j + 2)(2^{N-k})\\ &= (2^n + 2^{N+1})/2^k.\end{aligned}$$

We rewrite this as

$$\frac{\gamma_c(Q_n)}{2^n/n} < \left(1 + \frac{1}{2^{j-1}}\right)\left(1 + \frac{j-1}{2^k}\right).$$

In our range $1 \le j \le 2^k$, the first term in the product on the right starts at 2 and decreases exponentially quickly toward 1. The second term starts at 1 and grows linearly to just below 2 at the end of this range. Throughout this whole range in j, the product is at most 2, giving us the first statement of the theorem.

The second statement, the lower bound on $\gamma_c(Q_n)/2^n$, follows easily from Proposition 1. For the third statement, we select values of n for which we can show $\gamma_c(Q_n)/(2^n/n)$. Specifically, given k take $j = k+1$, so that $n = 2^k + k$. Then in the upper bound inequality above on $\gamma_c(Q_n)/(2^n/n)$, both terms in the product are small (slightly above 1), and their product $\to 1$ as $k \to \infty$. The lim inf statement follows. □

An interesting observation is that for n of the form $2^k - 1$, the Hamming code exists, but the corresponding spanning tree construction for Q_n we described earlier only guarantees that $\gamma_c(n)/(2^n/n)$ is at most 3 for such n. We can do better, constructing a

tree that reduces the bound for such n to 2, by taking the Hamming construction for $2^{k-1} - 1$ and applying the expansion method with $j = 2^{k-1}$. Nevertheless, we are still seeking to do better, aiming to construct trees that bring the bound down to 1 asymptotically.

6 Main result

We have shown how to construct spanning trees for hypercubes Q_n with many leaves– the proportion of the 2^n vertices that are not leaves is at most roughly $2/n$. The idea is to take a Hamming code and then expand.

Now observe that the expansion idea can be used starting from *any* values of N, not just a Hamming code value $2^k - 1$, and from *any* dominating set C in Q_N, to produce a connected dominating set for Q_n, $n = N + j$: Set C gives a partition of Q_N into stars. For each star center (vertex in the dominating set), add edges to connect the 2^j copies of the vertex. In the original Q_N add edges to connect the stars. We now have a spanning tree for Q_n. Denote by D its backbone, a connected dominating set in Q_n. We get an upper bound on $|D|$ as in Proposition 5. Assuming $|C|$ is minimum-sized, we get that

$$\gamma_c(Q_n) < (2^j + 2)\gamma(Q_N).$$

Given n large, let j be an integer close to $\log n$ (logarithm base 2), and take $N = n-j$. Then the display above implies that $\gamma_c(Q_n)/(2^n/n)$ is bounded above approximately by $\gamma(Q_n)/(2^N/N)$. So an upper bound function for the domination number, shifted to the right by $\log n$, yields an approximate upper bound function on connected domination.

In particular, if it holds that for domination $\gamma(Q_n)/(2^n/n)$ tends toward 1, its lower limit, then the same will be true for the similar expression for connected domination! Fortunately, what we need is proven in the 1997 book *Covering Codes* by Cohen, Honkala, Litsyn, and Lobstein [5], page 332. They attribute the result to Kabatyanskii and Panchenko [11] (1988). The proof relies on various coding constructions, including q-ary Hamming codes for prime powers q. It also depends on results on the density of primes.

We include their result on the domination number as the first part of our main theorem below. It is restated for convenient comparison to our result for connected domination number, the second part, which can be viewed as a strengthening of the first part.

Theorem 4. *The domination ratio for hypercubes satisfies* [11]

$$\lim_{n\to\infty} \frac{\gamma(Q_n)}{2^n/n} = 1.$$

The connected domination ratio satisfies

$$\lim_{n\to\infty} \frac{\gamma_c(Q_n)}{2^n/n} = 1.$$

Proof. As noted above, the first statement is proven in the literature. What is new is the second part, which is a stronger statement. Building on Theorem 3, it suffices to give an upper bound on $\gamma_c(Q_n)/(2^n/n)$ that goes to 1 as $n \to \infty$. As in the discussion above, given n we take j is close to $\log n$ and $N = n - j$. Given $\varepsilon > 0$, we have that for all sufficiently large n (and N) that

$$\frac{\gamma(Q_N)}{2^N/N} < 1 + \varepsilon.$$

Applying this in the discussion above, gives us for all sufficiently large n that

$$\frac{\gamma_c(Q_n)}{2^n/n} < (1+\varepsilon)^2,$$

and the second part follows. □

Formulating this equivalently in terms of leaves in spanning trees, we obtain the following.

Corollary 1. *As $n \to \infty$, $L(Q_n) = 2^n(1 - \frac{1}{n} + o(\frac{1}{n}))$.*

7 Further study

Here are some ideas for continuing research. We were not able to give a simple enough proof that the domination number that $\gamma(Q_n)/(2^n/n) \to 1$. We were hoping to give a self-contained proof of our main result. The proof in the literature of this domination result relies on rather technical explicit coding constructions. It would be nice if one could devise an algorithm, or use probabilistic arguments, to prove the existence of dominating sets in the hypercube that are as small as the theorem.

Another question asked by Bezrukov [2] is this: For $n = 2^k - 1$, starting from the stars given by the Hamming code, can one describe nicely how to add edges to form a tree with the most leaves (the smallest connected dominating set)? We have seen that for large k the connected dominating set will have size between 2 and 3 times $2^n/n$. Noga Alon pointed out (pers. comm.) that one only has to take a minimum connected dominating set and add to it the Hamming code to obtain a connected dominating set that, in view of the main theorem, is of size only $\sim 2(2^n/n)$. Still, we ask whether one can construct a connected dominating set of size $\sim 2(2^n/n)$, including the Hamming code, without relying on the other known covering codes (used in the proof of the main theorem).

What can one say about a more general class of graphs? For instance, one could consider domination and connected domination in a generalized grid (box) graph, such as a Cartesian product of n paths on p vertices. This graph on p^n vertices is the hypercube when $p = 2$. Perhaps the more natural graph to study is a product of n cycles on p vertices. Note that for $p = 4$ it is the same graph as Q_{2n}. Edenfield [7] studied products of cycles and products of complete graphs, both generalizations of the hypergraph questions in this paper.

Joshua Cooper suggests considering powers of graphs. That is, for a graph $G = (V, E)$, such as the hypercube, fix integer $r > 0$ and consider the same questions as before, but for the graph G^r: This graph also has vertex set V, but now edges join vertices at distance at most r in G. This is motivated by covering codes of radius r.

Bibliography

[1] S. Arumugam and R. Kala, Domination parameters of hypercubes, *J. Indian Math. Soc.*, **65** (1998), 31–38.

[2] S. Bezrukov, On the number of leaves in a spanning tree of the unit cube, *Notes (unpub.)* (2014), 2 pp.

[3] Y. Caro, D. B. West and R. Yuster, Connected domination and spanning trees with many leaves, *SIAM J. Discrete Math.*, **13** (2000), 202–211.

[4] Y-C. Chen and Y.-L. Syu, Connected dominating set of hypercubes and star graphs, in *IPCSIT*, vol. **41**, pp. 15–19, 2012.

[5] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein, *Covering Codes*, Elsevier Science, Amsterdam, 1997.

[6] W. Duckworth, P. E. Dunne, A. M. Gibbons and M. Zito, Leafy spanning trees in hypercubes, *Appl. Math. Lett.*, **14** (2001), 801–804.

[7] W. C. Edenfield, *Spanning Trees and Domination Numbers in the Cartesian Product of Cycle and Complete Graphs*, Honors Thesis, South Carolina Honors College, University of South Carolina, 2019.

[8] J. R. Griggs, D. J. Kleitman and A. Shastri, Spanning trees with many leaves in cubic graphs, *J. Graph Theory*, **13** (1989), 669–695.

[9] J. R. Griggs and M. Wu, Spanning trees in graphs of minimum degree 4 or 5, *Discrete Math.*, **104** (1992), 167–183.

[10] P. K. Jha, *Hypercubes, median graphs and product of graphs: some algorithmic and combinatorial results*, Ph.D. dissertation, Iowa State University, 1990, p. 26.

[11] G. A. Kabatyanskii and V. I. Panchenko, Unit sphere packings and coverings of the Hamming sphere, *Probl. Inf. Transm.*, **24** (1988), 261–272.

[12] D. J. Kleitman and D. B. West, Spanning trees with many leaves, *SIAM J. Discrete Math.*, **4** (1991), 99–106.

[13] D. Kuboň, *Genetic Approach to Hypercube Problems*, Master thesis, Computer Science, Charles University, Prague, 2017.

[14] S. A. Mane and B. N. Waphare, On independent and (d, n)-domination numbers of hypercubes, *AKCE Int. J. Graphs Comb.*, **9** (2012), 161–168.

[15] E. Sampathkumar and H. B. Walikar, The connected domination number of a graph, *J. Math. Phys. Sci.*, **13** (1979), 607–613.

[16] J. A. Storer, Constructing full spanning trees for cubic graphs, *Inf. Process. Lett.*, **13** (1981), 8–11.

Heiko Harborth and Hauke Nienborg

Rook domination on hexagonal hexagon boards

Abstract: Chess-like game boards B_n are considered, which are hexagonal parts of the Euclidean tessellation of the plane by regular hexagons. For chess-like rooks on B_n the domination number $\gamma(n)$ is determined.

1 Introduction

Corresponding to a classical chessboard a hexagonal hexagon board B_n is defined as the following hexagonal part of the Euclidean tessellation of the plane by regular hexagons: If B_1 is one hexagon and if B_2 consists of three hexagons with a common vertex, then B_n for $n \geq 3$ consists of B_{n-2} together with all neighboring hexagons of B_{n-2} (see Figure 14.1). One may wonder whether Ronald Graham, who liked to look at combinatorial problems for chessboards, would have liked the corresponding problems for these hexagonal boards as well.

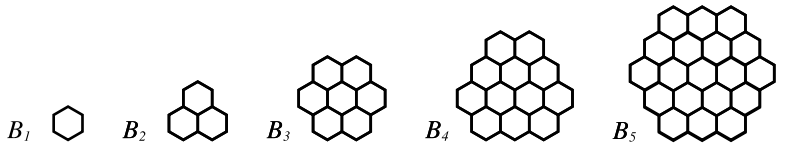


Figure 14.1: Hexagonal hexagon boards.

A rook can move on straight-line sequences of edge-adjacent hexagons (see [3]) as on straight-line sequences of edge-adjacent squares on classical chessboards. Then the domination number $\gamma(n)$ denotes the smallest number of rooks, so that every hexagon of B_n is either occupied or threatened. Here, we want to determine $\gamma(n)$.

Theorem 1. *For all $n \geq 1$, it holds that*

$$\gamma(n) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \pmod 4, \\ \frac{n+1}{2} & \text{if } n \equiv 1 \pmod 4, \\ \frac{n-1}{2} & \text{if } n \equiv 3 \pmod 4. \end{cases}$$

Heiko Harborth, Hauke Nienborg, Diskrete Mathematik, Technische Universität Braunschweig, Braunschweig, Germany, e-mails: h.harborth@tu-bs.de, hauke.nienborg@ewetel.net

https://doi.org/10.1515/9783110754216-014

For the domination number of grids that threaten all neighboring hexagons on B_n, estimates can be found in [5].

The domination number for rooks on triangular hexagon boards seems to be more difficult to determine (see [1, 4]). The independence number $\beta(n)$ corresponding to $\gamma(n)$; that is, the maximum number of pairwise not threatening rooks on B_n is proven in [1] to be $\beta(n) = 2\lceil \frac{n}{2} \rceil - 1$.

In general, results for domination, independence, and other parameters for chessboards can be found in [2, 3].

2 Upper bound

For $n \leq 3$, $\gamma(n)$ is easy to prove, as claimed in Theorem 1. Consider the four consecutive boards $B_n = B_{4t+i}$ for $i = 1, 2, 3, 4$, and $t \geq 1$. If the first row has $\lfloor \frac{n+1}{2} \rfloor$ hexagons at the top, then we put a first rook in row $\lfloor \frac{n+5}{4} \rfloor$ at position $\lfloor \frac{n+5}{4} \rfloor$ from the left and a second rook in the next row at position $\lfloor \frac{n+5}{4} \rfloor$ from the right. Vertically below the first rook, $\lfloor \frac{n-1}{4} \rfloor$ further rooks are placed in the hexagons of the rows $\lfloor \frac{n+5}{4} \rfloor + 2j$ for $1 \leq j \leq \lfloor \frac{n-1}{4} \rfloor$ and vertically below the second rook, $\lfloor \frac{n}{4} \rfloor - 1$ further rooks are placed in the hexagons of the rows $\lfloor \frac{n+5}{4} \rfloor + 2j + 1$ for $1 \leq j \leq \lfloor \frac{n}{4} \rfloor - 1$ (see Figure 14.2 for $t = 3$). Note that this construction is also possible for $n = 4$.

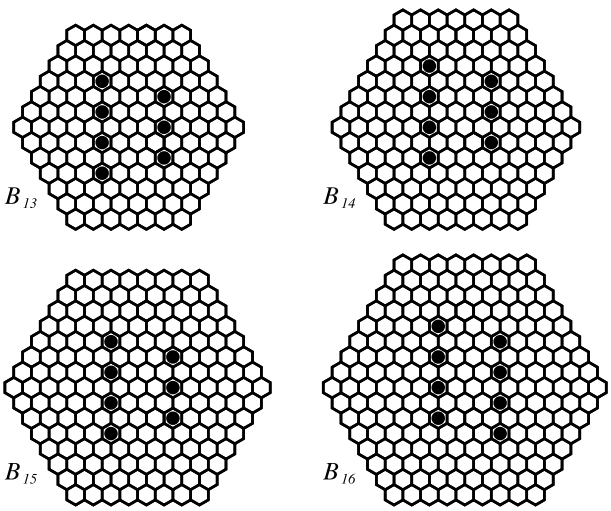


Figure 14.2: Rook domination for $n = 13$, 14, 15, and 16.

Together, there are $1 + \lfloor \frac{n}{4} \rfloor + \lfloor \frac{n-1}{4} \rfloor$ rooks, that is, $2t + 1 = \frac{n+1}{2}$, $2t + 1 = \frac{n}{2}$, $2t + 1 = \frac{n-1}{2}$, and $2t + 2 = \frac{n}{2}$ rooks for $i = 1, 2, 3$, and 4, respectively, as claimed in Theorem 1. Now it

is easy to check that the rooks on the three families of adjacent parallel straight lines of hexagons dominate each hexagon of B_n.

3 Lower bound

The straight-line sequences of edge-adjacent hexagons of B_n are called x-lines if they have x hexagons. There are three n-lines (diagonals) which, for odd n, have the central hexagon in common and which, for even n, pairwise have one of the three central hexagons in common. The six x-lines at the border of B_n have $x = \frac{n+1}{2}$ for odd n and alternatingly $x = \frac{n+2}{2}$ and $x = \frac{n}{2}$ for even n, say $x = \frac{n}{2}$ at the top.

The strategy for the proofs of the lower bounds is to check all x-lines with $x = n, n-1, n-2, \ldots$

(i) whether it must contain a rook,
(ii) whether it must not contain a rook, or
(iii) whether both cases containing a rook or not having a rook are to be distinguished.

There are two possibilities for (i):

$(i)_1$ Since each of the assumed y rooks from outside an x-line can threaten at most two hexagons of the x-line and since each of p pairs of the y rooks that threaten one and the same hexagon threatens at most three hexagons of the x-line, it follows that at most $2y - p$ hexagons are threatened and $2y - p < x$ forces a rook on the x-line if $p < x$.

$(i)_2$ If for a hexagon of an x-line the other two lines through this hexagon both contain no rook, then a rook on this x-line is forced.

If for (ii) an x-line has a hexagon, so that one of the two other lines through the hexagon contains a rook and so that the other line, say an y-line, has been chosen as without a rook because of $2y - p = y$, $p < y$, and this hexagon is threatened only once, then no rook on this x-line is forced.

For (iii), it must hold for the x-line $2y - p \geq x$, $p < x$.

3.1 $n \equiv 3 \pmod 4$

Assume that $\gamma(n) \leq \frac{n-3}{2}$. Because $2\frac{n-3}{2} < n-2$, all $(n-i)$-lines for $i \leq 2$ each have a rook. If a rook is assumed on every $(n-j)$-line for $j \leq i-1$, then on every $(n-i)$-line there are $p = i-1$ hexagons which are threatened twice (see Figure 14.3). Hence, with $(i)_1$ and because of $2y - p \leq 2\frac{n-3}{2} - (i-1) < n-i$ for $i < \frac{n+1}{2}$ it follows that all $(n-i)$-lines must each contain a rook. Since all $(n-i)$-lines for $i \geq \frac{n-3}{4}$ are required to cover all hexagons of B_n, there are at least $2i + 1 \geq 2\frac{n-3}{4} + 1 = \frac{n-1}{2}$ parallel x-lines, each containing a rook, contradicting $y \leq \frac{n-3}{2}$.

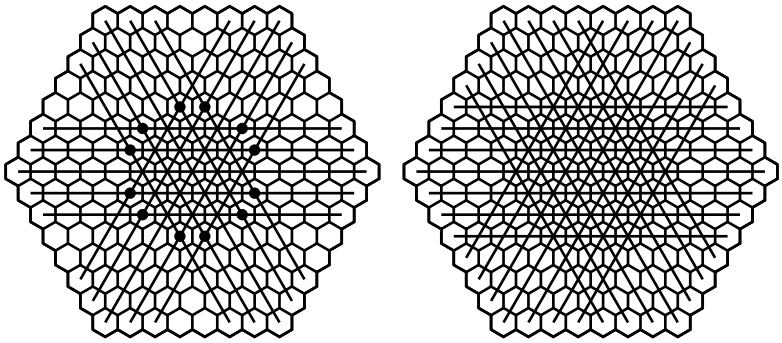


Figure 14.3: Case $i = 3$ for B_{15}.

3.2 $n \equiv 1 \pmod 4$

For $n = 4s + 1$, it is assumed $\gamma(n) \leq \frac{n-1}{2}$. Because $2\frac{n-1}{2} < n$, every n-line must have a rook. If all x-lines for $n \geq x \geq n - t + 1$ each contain a rook, then on each $(n - t)$-line there are $p = t - 1$ hexagons that have already been threatened twice. Hence, with (iii) and because of $2\frac{n-1}{2} - (t - 1) = n - t$ for $t < \frac{n+1}{2}$, every $(n - t)$-line could contain a rook or not contain a rook. Note that in the later case, only exactly one threat is allowed for all hexagons of an $(n - t)$-line, except for those $p = t - 1$ hexagons that are twice threatened.

If there is a rook on one of the six $(n - t)$-lines, then each of the two neighboring $(n - t)$-lines now has t doubly threatened hexagons, so that with $(\mathrm{i})_1$ and because of $2\frac{n-1}{2} - t < n - t$ for $t < \frac{n}{2}$ these $(n - t)$-lines also contain a rook. Now the next two neighboring $(n - t)$-lines also have t doubly threatened hexagons, so they also must have a rook as before. Finally, the sixth $(n - t)$-line then has $t + 1$ doubly threatened hexagons, and it must have a rook as before (see Figure 14.4). So if one of the $(n–t)$-lines contains a rook, then all six must contain a rook. Thus there are only two possibilities: that all $(n - t)$-lines have a rook and that all are without a rook.

Now the six $(n - t)$-lines are supposed to be the largest that are without a rook.

$t = 1$: In this case, all $(n - i)$-lines for $i = 0, 1, 2, \ldots$ are alternatingly with or without a rook if i is even or odd, respectively. This is true for $i = 0$ and $i = 1$ because $t = 1$. If all $(n - 2j - 1)$-lines are without a rook $(j \geq 0)$, then every $(n - 2j - 2)$-line is forced to have a rook because of a common hexagon together with an $(n - 1)$-line (without rook) and an $(n - 2j - 1)$-line (without rook) and with $(\mathrm{i})_2$. If then all $(n - 2j - 2)$-lines have a rook, then every $(n - 2j - 3)$-line is forced to be without a rook because of a common hexagon together with an $(n - 1)$-line (without rook) and an $(n - 2j - 2)$-line (with rook) and with (ii) (see Figure 14.5).

Since then on every x-line with a rook hexagons occur that are threatened only once, all x-lines with a rook are required for a domination of all hexagons of B_n. So there are $\lceil \frac{4s+1}{2} \rceil = 2s + 1 = \frac{n+1}{2}$ parallel x-lines that each have a rook in contradiction

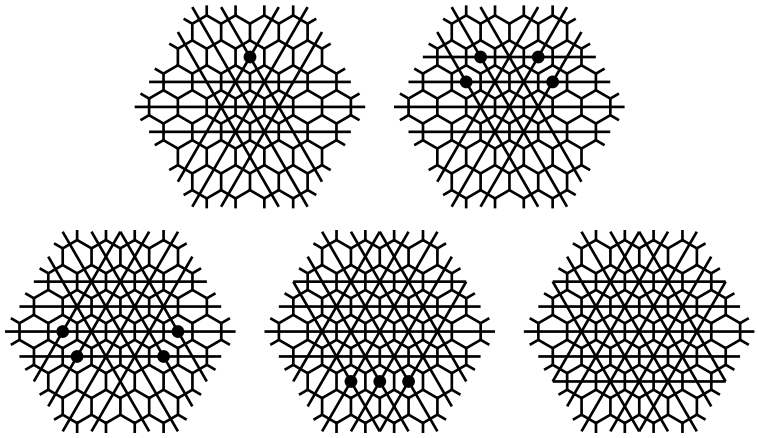


Figure 14.4: All $(n-t)$-lines with a rook for $t = 2$.

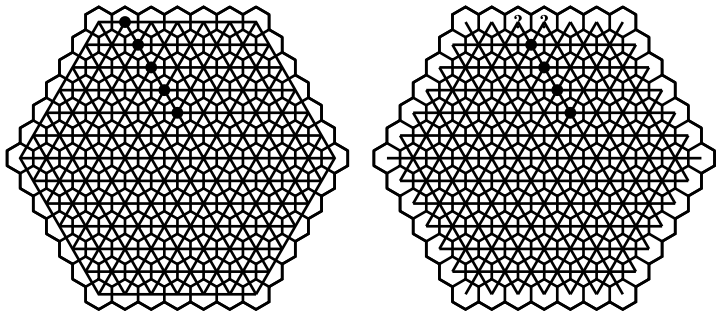


Figure 14.5: Cases $t = 1$ and $t = 2$ for B_{13}.

to $\gamma(n) \leq \frac{n-1}{2}$. It can be noted that for $n = 4s - 1$, two x-lines, each with a rook, are omitted, so that $\gamma(n) = \gamma(4s-1) = \lceil \frac{4s+1}{2} \rceil - 2 = 2s - 1 = \frac{n-1}{2}$ could apply.

$t \geq 2$: Any $(n-x-t)$-line for $1 \leq x \leq t-1$ is forced to be without a rook because of a common hexagon together with an $(n-t)$-line (without rook) and an $(n-x)$-line (with rook) and with (ii). Next, any $(n-x-t)$-line for $t \leq x \leq 2t-1$ is forced to have a rook because of a common hexagon together with an $(n-t)$-line (without rook) and an $(n-x)$-line (without rook) and with $(\mathrm{i})_2$ (see Figure 14.5). Then on the one hand any $(n-3t)$-line is forced to be without a rook because of a common hexagon together with an $(n-t)$-line (without rook) and an $(n-2t)$-line (with rook) and with (ii). On the other hand, any $(n-3t)$-line is forced to have a rook because of a common hexagon together with an $(n-t-1)$-line (without rook) and an $(n-2t+1)$-line (without rook) and with $(\mathrm{i})_2$ (see Figure 14.5). This is a contradiction, and thus a domination cannot exist for $n - 3t \geq \frac{n+1}{2}$, that is, for $n \geq 6t+1$. For $n < 6t+1$, there are rooks on each $(n-x)$-line for $0 \leq x \leq t-1$ and for $2t \leq x \leq 3t-1$. Then B_n has $2(t-1)+1 \geq \frac{n+1}{2}$ rooks on parallel

lines if $n \le 4t-3$ and $2(t-1)+1+2j \ge \frac{n+1}{2}$ rooks on parallel lines if $n \ge 4t+2j-1$, $j \ge 1$. Both cases are contradicting $\gamma(n) \le \frac{n-1}{2}$, which finishes the proof for $n \equiv 1 \pmod 4$.

It can be noted that for $n = 4t-1$, that is, for $j = 0$, there are only $2(t-1)+1 = \frac{n-1}{2}$ rooks on parallel lines so that $\gamma(n) = \frac{n-1}{2}$ could apply. Together with the note in case $t = 1$ there are only two possibilities for $\gamma(n) = \frac{n-1}{2}$ if $n \equiv 3 \pmod 4$ both of which are similar by a factor of 2. Then all maximum independences of rooks for the triple threatened hexagons determine all minimum dominations of rooks for B_n (see Figure 14.6). This gives the following.

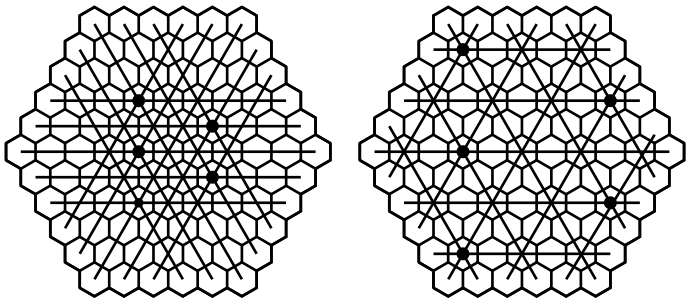


Figure 14.6: Two dominations for $n = 11$.

Corollary 1. *For $n \equiv 3 \pmod 4$, all minimum dominations with rooks on B_n are determined by all maximum independences with rooks on $B_{(n-1)/2}$.*

3.3 $n \equiv 0 \pmod 2$

Assume that $\gamma(n) < \frac{n}{2}$. Then $2(\frac{n}{2}-1) < n-1$ implies that all $(n-i)$-lines for $i \le 1$ must have a rook. If a rook is assumed on every $(n-j)$-line for $j \le i-1$, then on the six $(n-i)$-lines there are alternatingly $p = i$ and $p = i-2$ hexagons, which are threatened twice (see Figure 14.7). With $(\mathrm{i})_1$ and because of $2\frac{n-2}{2} - i < n-i$ for $i < \frac{n}{2}$ those three $(n-i)$-lines with i doubly threatened hexagons also must contain a rook. Then the three remaining $(n-i)$-lines also each have $p = i$ hexagons, which are threatened twice (see Figure 14.7) and as before these $(n-i)$-lines also must have a rook. Now all $(n-i)$-lines for $i \ge \lfloor \frac{n}{4} \rfloor$ are needed to cover all hexagons of B_n so that there are at least $2i+1 \ge 2\lfloor \frac{n}{4} \rfloor + 1 \ge \frac{n}{2}$ parallel lines, each containing a rook, contradicting $\gamma(n) < \frac{n}{2}$. This completes the proof of Theorem 1.

It has to be remarked that after the submission it was noticed that in [6] the domination number has already been determined, but only for odd n and with a completely different proof. The rooks are referred to as queens in [6], although queens should be able to move in six directions.

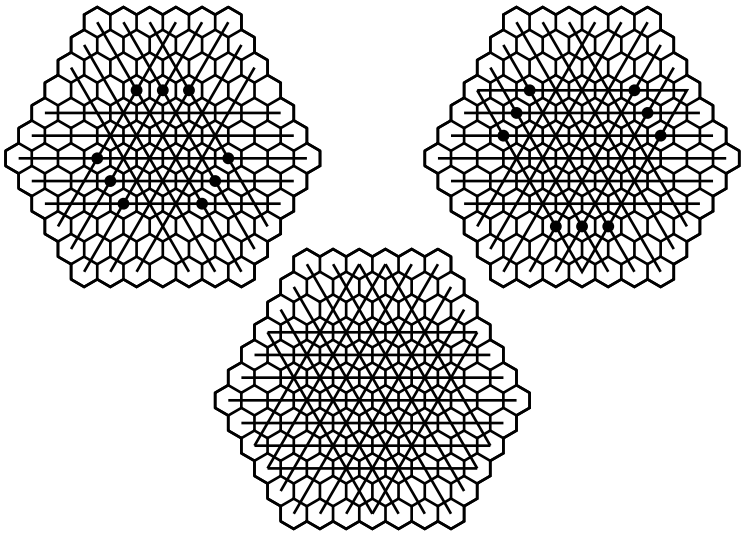


Figure 14.7: Case $i = 3$ for B_{12}.

Bibliography

[1] J. D. E. Konhauser, D. Velleman and S. Wagon, in Which Way did the Bicycle Go? The Mathematical Association of America, Washington DC, MAA, 1996.

[2] S. M. Hedetniemi, S. T. Hedetniemi and R. Reynolds, Combinatorial problems on chessboards, II, in *Domination in Graphs*, pp. 133–162, M. Dekker, New York, 1998.

[3] J.-P. Bode and H. Harborth, Independent chess pieces on euclidean boards, *J. Comb. Math. Comb. Comput.*, **33** (2000), 209–223.

[4] S. Wagon, Graph theory problems from hexagonal and traditional chess, *Coll. Math. J.*, **45** (2014), 278–287.

[5] H. Harborth and H. Nienborg, Grid domination on hexagonal boards, *Congr. Numer.*, **233** (2019), 11–20.

[6] A. P. Burger and C. M. Mynhardt, Queens on hexagonal boards, *J. Comb. Math. Comb. Comput.*, **31** (1999), 97–111.

Neil Hindman and Dona Strauss

Strongly image partition regular matrices

Abstract: A $u \times v$ matrix A with rational entries is *image partition regular over* $\mathbb{N}$ provided that whenever $\mathbb{N}$ is finitely colored, there exists $\vec{x} \in \mathbb{N}^v$ such that the entries of $A\vec{x}$ are monochromatic. We say that A is *strongly image partition regular over* $\mathbb{N}$ provided that for every IP-set C in $\mathbb{N}$ there exists $\vec{x} \in \mathbb{N}^v$ such that the entries of $A\vec{x}$ are in C. Many characterizations of image partition regular matrices are known. We provide here two sufficient conditions and one necessary condition for a matrix with rank u to be strongly image partition regular and show that such matrices can be expanded horizontally at will. We provide several examples showing that our results are sharp.

1 Introduction

We let $\mathbb{N}$ be the set of positive integers and $\omega = \mathbb{N} \cup \{0\}$.

Definition 1.1. Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with rational entries.

(1) The matrix A is *kernel partition regular over* $\mathbb{N}$ provided that whenever $\mathbb{N}$ is finitely colored, there exists monochromatic $\vec{x} \in \mathbb{N}^v$ such that $A\vec{x} = \vec{0}$.

(2) The matrix A is *image partition regular over* $\mathbb{N}$ (IPR/$\mathbb{N}$) provided that whenever $\mathbb{N}$ is finitely colored, there exists $\vec{x} \in \mathbb{N}^v$ such that the entries of $A\vec{x}$ are monochromatic.

In 1933, Richard Rado [11] characterized kernel partition regular matrices in terms of the "columns condition."

Definition 1.2. Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from $\mathbb{Q}$. For $i \in \{1, 2, \ldots, v\}$, let $\vec{c}_i$ be column i of A. Then A satisfies the *columns condition* if and only if there exist $m \in \mathbb{N}$ and a partition $\{I_1, I_2, \ldots, I_m\}$ of $\{1, 2, \ldots, v\}$ such that

(1) $\sum_{i \in I_1} \vec{c}_i = \vec{0}$ and

(2) for each $j \in \{2, 3, \ldots, m\}$, if any, $\sum_{i \in I_j} \vec{c}_i$ is a linear combination over $\mathbb{Q}$ of $\{\vec{c}_i : i \in \bigcup_{t=1}^{j-1} I_t\}$.

Theorem 1.3 (Rado [11]). *Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from $\mathbb{Q}$. Then A is kernel partition regular over $\mathbb{N}$ if and only if A satisfies the columns condition.*

Neil Hindman, Department of Mathematics, Howard University, Washington, DC, USA, e-mail: nhindman@aol.com

Dona Strauss, Department of Pure Mathematics, University of Leeds, Leeds, United Kingdom, e-mail: d.strauss@emeritus.hull.ac.uk

https://doi.org/10.1515/9783110754216-015

Say that a subset C of $\mathbb{N}$ is *large* if for every kernel partition regular matrix A, there exists $\vec{x}$ in the kernel of A with all entries of $\vec{x}$ in C. Rado conjectured that if a large subset of $\mathbb{N}$ is finitely colored, then there will be a monochromatic large subset. This conjecture was proved by Walter Deuber in 1973 [2] using what he called (m,p,c)-*sets*. These (m,p,c)-sets are images of certain "first entries" matrices. Part of Deuber's results included the fact that first entries matrices are image partition regular over $\mathbb{N}$.

We follow the custom of denoting the entries of a matrix by the lower case letter corresponding to the name of the matrix.

Definition 1.4. Let $u,v \in \mathbb{N}$ and let A be a $u \times v$ matrix with rational entries. Then A is a *first entries matrix* if and only if no row of A is $\vec{0}$ and whenever $i,j \in \{1,2,\ldots,u\}$ and $k = \min\{t \in \{1,2,\ldots,v\} : a_{i,t} \neq 0\} = \min\{t \in \{1,2,\ldots,v\} : a_{j,t} \neq 0\}$, then $a_{i,k} = a_{j,k} > 0$. An element b of $\mathbb{Q}$ is a *first entry* of A if and only if there is some row i of A such that $b = a_{i,k}$ where $k = \min\{t \in \{1,2,\ldots,v\} : a_{i,t} \neq 0\}$.

Image partition regular matrices were first characterized in 1993 [5]. One of these characterizations involves first entries matrices.

Theorem 1.5 ([5]). *Let $u,v \in \mathbb{N}$ and let A be a $u \times v$ matrix with rational entries. Then A is image partition regular over $\mathbb{N}$ if and only if there exist $m \in \mathbb{N}$ and a $u \times m$ first entries matrix B such that for each $\vec{y} \in \mathbb{N}^m$ there exists $\vec{x} \in \mathbb{N}^v$ such that $A\vec{x} = B\vec{y}$.*

Since the publication of [5] several other characterizations of IPR/$\mathbb{N}$ matrices have been obtained. Theorem 15.24 in [8] lists twelve statements that are equivalent to IPR/$\mathbb{N}$. Some of these, first obtained in [6], are included in the following theorem. Two that are of interest to us involve "central" sets. Central sets were introduced by Hillel Furstenberg in [4] and defined in terms of topological dynamics.

Theorem 1.6 ([6]). *Let $u,v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from $\mathbb{Q}$. The following statements are equivalent:*

(a) *A is image partition regular over $\mathbb{N}$.*
(b) *For each central set C in $\mathbb{N}$, $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^u\} \neq \emptyset$.*
(c) *For each central set C in $\mathbb{N}$, $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^u\}$ is central in $\mathbb{N}^v$.*
(d) *For each column $\vec{c} \in \mathbb{Q}^u$, $(A\ \vec{c})$ is image partition regular over $\mathbb{N}$.*
(e) *For each row $\vec{r} \in \mathbb{Q}^v \setminus \{\vec{0}\}$, there exists $b \in \mathbb{Q} \setminus \{0\}$ such that $\begin{pmatrix} b\vec{r} \\ A \end{pmatrix}$ is image partition regular over $\mathbb{N}$.*

It was an idea of Vitaly Bergelson [1] to characterize central sets in terms of the algebra of the Stone–Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$. See [8, Definition 4.42] for the algebraic definition of central set and [8, Chapter 19] for a proof of the equivalence of the algebraic and dynamical definitions of central. We will not go into the precise definitions in this paper since we will not be using the algebra of the Stone–Čech compactification of a discrete semigroup here. What is important for us here is that central sets are IP-sets.

Given a set X, we write $\mathcal{P}_f(X)$ for the set of finite nonempty subsets of X.

Definition 1.7. Let $(S,+)$ be a commutative semigroup and let $\langle x_n\rangle_{n=1}^{\infty}$ be a sequence in S. Then $FS(\langle x_n\rangle_{n=1}^{\infty}) = \{\sum_{n\in F} x_n : F \in \mathcal{P}_f(\mathbb{N})\}$. If $k,m \in \mathbb{N}$ and $k \leq m$, then $FS(\langle x_n\rangle_{n=k}^{m}) = \{\sum_{n\in F} x_n : \emptyset \neq F \subseteq \{k,k+1,\ldots,m\}\}$.

Definition 1.8. Let $(S,+)$ be a commutative semigroup and let $C \subseteq S$. Then C is an *IP-set* if and only if there exists a sequence $\langle x_n\rangle_{n=1}^{\infty}$ in S such that $FS(\langle x_n\rangle_{n=1}^{\infty}) \subseteq C$.

For readers familar with the algebra of the Stone–Čech compactification βS of a discete semigroup S, we remark that a subset C of S is an *IP*-set if and only if C is a member of an idempotent in βS; see [8, Theorem 5.12].

Lemma 1.9. *Let C be an IP-set in $\mathbb{N}$ and let $m \in \mathbb{N}$. There is an increasing sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS(\langle x_n\rangle_{n=1}^{\infty}) \subseteq C \cap m\mathbb{N}$.*

Proof. By [8, Lemma 6.6] $C \cap m\mathbb{N}$ is an IP-set so one can pick $\langle x_n\rangle_{n=1}^{\infty}$ with $FS(\langle x_n\rangle_{n=1}^{\infty}) \subseteq C \cap m\mathbb{N}$. By combining successive terms, we may presume that $\langle x_n\rangle_{n=1}^{\infty}$ is increasing. □

Definition 1.10. Let $u,v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from $\mathbb{Q}$. Then A is *strongly image partition regular over* $\mathbb{N}$ (SIPR/$\mathbb{N}$) provided whenever C is an IP-set in $\mathbb{N}$, there exists $\vec{x}$ in $\mathbb{N}^v$ such that $A\vec{x} \in C^u$.

We shall see in Section 2 that strongly image partition matrices are indeed image partition regular. It is easy to see that the converse fails. The simplest nontrivial instance of van der Waerden's theorem [12] tells us that the matrix

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$$

is image partition regular. On the other hand, if $a \in \mathbb{N} \setminus \{1,2\}$ a simple consideration of the base a expansions shows that $FS(\langle a^t\rangle_{t=1}^{\infty})$ does not contain any length 3 arithmetic progressions, so that matrix is not strongly image partition regular over $\mathbb{N}$.

We shall see in Section 3 that, if one adds the assumption that the rank of A is u, where u is the number of rows, one gets a substantial collection of SIPR/$\mathbb{N}$ matrices. Further, in this section we develop sufficient conditions for such a matrix to be SIPR/$\mathbb{N}$ as well as one necessary condition. These conditions are in terms of the inverse of a matrix consisting of u linearly independent columns of A.

Section 4 is primarily devoted to examples.

In Section 5, we will extend the notion of strong image partition regularity to infinite matrices.

2 Strongly image partition regular over S

In this section, we present some results that apply to arbitrary commutative semigroups. Unfortunately, there have been different definitions in the literature for the notion of image partition regularity over a commutative semigroup. We use here the definition that we used in [9]. (See the discussion in [9] for reasons for the choice.)

If a commutative semigroup has an identity, we denote that identity by 0. If not, then of course $S \setminus \{0\} = S$. If S is cancellative and $x \in S$, then by $-x$ we mean the inverse of x in the group of differences of S.

Definition 2.1. Let S be a commutative semigroup, let $u, v \in \mathbb{N}$, and let A be a $u \times v$ matrix. If S is cancellative, and therefore embeddable in a group, then A is *appropriate for S* provided no row of A is zero and the entries of A come from $\mathbb{Z}$. If S is not cancellative, then A is *appropriate for S* provided no row of A is zero and the entries of A come from ω.

Definition 2.2. Let S be a commutative semigroup, let $u, v \in \mathbb{N}$, and let A be a $u \times v$ matrix which is appropriate for S. Then A is *image partition regular over S* (IPR/S) if and only if whenever $S \setminus \{0\}$ is finitely colored, there exists $\vec{x} \in (S \setminus \{0\})^v$ such that the entries of $A\vec{x}$ are monochromatic.

In [8, Definition 5.9] what we are calling *image partition regular* here was called *strongly image partition regular*.

Definition 2.3. Let S be an infinite commutative semigroup, let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix which is appropriate for S. Then A is *strongly image partition regular over S* (SIPR/S) if and only if whenever C is an IP-set contained in $S \setminus \{0\}$, there exists $\vec{x} \in (S \setminus \{0\})^v$ such that $A\vec{x} \in C^u$.

Since we have defined *strongly* image partition regular, we pause to show that very weak hypotheses guarantee that a SIPR/S matrix is in fact IPR/S.

Theorem 2.4. *Let S be an infinite commutative semigroup, let $u, v \in \mathbb{N}$, and let A be a $u \times v$ matrix, which is appropriate for S. Assume that $S \setminus \{0\}$ is an IP-set in S and that A is SIPR/S. Then A is IPR/S.*

Proof. Let $r \in \mathbb{N}$ and assume that $S \setminus \{0\} = \bigcup_{i=1}^{r} D_i$. Pick a sequence $\langle x_n \rangle_{n=1}^{\infty}$ in S such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq S \setminus \{0\}$. Then $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq \bigcup_{i=1}^{r} D_i$ so by [8, Corollary 5.15] pick a sequence $\langle y_n \rangle_{n=1}^{\infty}$ and $i \in \{1, 2, \ldots, r\}$ such that $C = FS(\langle y_n \rangle_{n=1}^{\infty}) \subseteq D_i$. Pick $\vec{x} \in (S \setminus \{0\})^v$ such that $A\vec{x} \in C^u$. Then the entries of $A\vec{x}$ are all in D_i. □

It is easy to see that if S is weakly cancellative, that is if for each $x, y \in S$, $\{z \in S : x + z = y\}$ is finite, then $S \setminus \{0\}$ is an IP-set in S. In fact, if $\{x \in S : \{y \in S : x + y = 0\}$ is infinite$\}$ is finite, then it is routine to construct a sequence $\langle x_n \rangle_{n=1}^{\infty}$ with $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq S \setminus \{0\}$.

We see now that if S satisfies this weak hypothesis, then SIPR/S matrices satisfy a conclusion similar to Theorem 1.6(c).

Theorem 2.5. *Let S be an infinite commutative semigroup, let $u, v \in \mathbb{N}$, and let A be a $u \times v$ matrix, which is appropriate for S. Assume that $S \setminus \{0\}$ is an IP-set in S and that A is SIPR/S. Then for each IP-set C in $S \setminus \{0\}$, $\{x \in S^v : A\vec{x} \in C^u\}$ is an IP-set in S^v.*

Proof. Let C be an IP-set in $S \setminus \{0\}$ and pick a sequence $\langle x_n \rangle_{n=1}^{\infty}$ in S such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq C$. Pick $\vec{y}(1) \in (S \setminus \{0\})^v$ and $F_{1,1}, F_{1,2}, \ldots, F_{1,u}$ in $\mathcal{P}_f(\mathbb{N})$ such that

$$A\vec{y}(1) = \begin{pmatrix} \sum_{t \in F_{1,1}} x_t \\ \vdots \\ \sum_{t \in F_{1,u}} x_t \end{pmatrix}.$$

Let $n \in \mathbb{N}$ and assume we have chosen $\vec{y}(n)$ and $F_{n,1}, F_{n,2}, \ldots, F_{n,u}$. Let $m = \max \bigcup_{i=1}^{u} F_{n,i}$. Then $FS(\langle x_t \rangle_{t=m+1}^{\infty})$ is an IP-set in $S \setminus \{0\}$ so pick $\vec{y}(n+1) \in (S \setminus \{0\})^v$ and $F_{n+1,1}, F_{n+1,2}, \ldots, F_{n+1,u}$ in $\mathcal{P}_f(\mathbb{N})$ such that $\min \bigcup_{i=1}^{u} F_{n+1,i} > m$ and

$$A\vec{y}(n+1) = \begin{pmatrix} \sum_{t \in F_{n+1,1}} x_t \\ \vdots \\ \sum_{t \in F_{n+1,u}} x_t \end{pmatrix}.$$

Given $H \in \mathcal{P}_f(\mathbb{N})$ and $i \in \{1, 2, \ldots, u\}$, let $K_i = \bigcup_{n \in H} F_{n,i}$. Then

$$A\left(\sum_{n \in H} \vec{y}(n) \right) = \begin{pmatrix} \sum_{t \in K_1} x_t \\ \vdots \\ \sum_{t \in K_u} x_t \end{pmatrix}. \qquad \square$$

In the generality of Theorem 2.5, we do not see that we can guarantee that $FS(\langle \vec{y}(n) \rangle_{n=1}^{\infty}) \subseteq (S \setminus \{0\})^v$; that is, that $0 \notin FS(\langle \vec{y}(n) \rangle_{n=1}^{\infty})$.

Definition 2.3 applies to the semigroup $(\mathbb{N}, +)$ and differs from Definition 1.10 because in the latter the entries of A were allowed to be fractions. We see now that this makes no essential difference.

Theorem 2.6. *Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from $\mathbb{Q}$, let $d \in \mathbb{N}$ such that all entries of dA are in $\mathbb{Z}$. If for every IP-set C in $\mathbb{N}$, $\{\vec{a} \in \mathbb{N}^v : (dA)\vec{a} \in C^u\} \neq \emptyset$, then for every IP-set C in $\mathbb{N}$, $\{\vec{b} \in \mathbb{N}^v : A\vec{b} \in C^u\} \neq \emptyset$.*

Proof. Let C be an IP-set in $\mathbb{N}$. Pick $\vec{a} \in \mathbb{N}^v$ such that $(dA)\vec{a} \in C^u$. Let $\vec{b} = d\vec{a}$. Then $A\vec{b} \in C^u$. $\square$

3 Strongly image partition regular over $\mathbb{N}$

We begin by showing that if the rank of the $u \times v$ matrix is u, then the property of being SIPR/$\mathbb{N}$ shares one of the strong conclusions applying to the property of being IPR/$\mathbb{N}$, namely the condition of Theorem 1.6(d).

Definition 3.1. Let S be a semigroup. A subset D of S is an *IP*-set* provided it has nonempty intersection with every IP-set in S.

Lemma 3.2. *Let $k, v \in \mathbb{N}$. Then $\{\vec{x} \in \mathbb{N}^v : \text{for all } i \in \{1, 2, \ldots, v\}, x_i > k\}$ is an IP*-set in $\mathbb{N}^v$.*

Proof. Let $D = \{\vec{x} \in \mathbb{N}^v : \text{for all } i \in \{1, 2, \ldots, v\}, x_i > k\}$ and let C be an IP-set in $\mathbb{N}^v$. Pick a sequence $\langle \vec{x}_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}^v$ such that $FS(\langle \vec{x}_n \rangle_{n=1}^{\infty}) \subseteq C$. Then $\sum_{n=1}^{k+1} \vec{x}_n \in C \cap D$. □

Theorem 3.3. *Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with rational entries such that $\operatorname{rank}(A) = u$ and A is SIPR/$\mathbb{N}$. Let $\vec{y} \in \mathbb{Q}^u$. Then $(A \ \vec{y})$ is SIPR/$\mathbb{N}$.*

Proof. Since the columns of A span $\mathbb{Q}^u$, pick $\vec{z} \in \mathbb{Q}^v$ such that $A\vec{z} = \vec{y}$. Pick $m \in \mathbb{N}$ such that $m\vec{z} \in \mathbb{Z}^v$ and let $k = \max(\{1\} \cup \{mz_i : i \in \{1, 2, \ldots, v\}\})$. Let C be an IP-set in $\mathbb{N}$. Now $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^u\}$ is an IP-set in $\mathbb{N}^v$ by Theorem 2.5 and $\{\vec{x} \in \mathbb{N}^v : \text{for all } i \in \{1, 2, \ldots, v\}, x_i > k\}$ is an IP*-set in $\mathbb{N}^v$ so pick $\vec{x} \in \mathbb{N}^v$ such that $A\vec{x} \in C^u$ and $x_i > k$ for each $i \in \{1, 2, \ldots, u\}$.

Define $\vec{w} \in \mathbb{Q}^{v+1}$ by $w_j = x_j - mz_j$ if $j \leq v$ and $w_{v+1} = m$. Note that $\vec{w} \in \mathbb{N}^{v+1}$. Also $(A \ \vec{y})\vec{w} = A\vec{x} - A(m\vec{z}) + m\vec{y} = A\vec{x} \in C^u$. □

The $\operatorname{rank}(A) = u$ hypothesis cannot be simply omitted as seen by considering the matrix

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

We saw in the Introduction that the matrix

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$$

is not SIPR/$\mathbb{N}$. On the other hand, the $\operatorname{rank}(A) = u$ assumption is not necessary since any column can be added to $\binom{1}{1}$ and the result will be SIPR/$\mathbb{N}$. (We will show in the next section that any $2 \times v$ matrix which is IPR/$\mathbb{N}$ is SIPR/$\mathbb{N}$.)

We have two sufficient conditions for a $u \times v$ matrix with rank u to be SIPR/$\mathbb{N}$ and one necessary condition.

Theorem 3.4. *Let $u, v \in \mathbb{N}$, let A be a $u \times v$ matrix with rational entries and rank u, and assume that B consists of u linearly independent columns of A. Let $D = B^{-1}$ and for $i \in \{1, 2, \ldots, u\}$, let $\vec{c}_i$ be column i of D. Assume there is nonempty $I \subseteq \{1, 2, \ldots, u\}$ such that all entries of $\sum_{i \in I} \vec{c}_i$ are positive. Then A is SIPR/$\mathbb{N}$.*

Proof. By Theorem 3.3, we may presume that $A = B$. Pick $m \in \mathbb{N}$ such that for each $(i, j) \in \{1, 2, \ldots, u\} \times \{1, 2, \ldots, u\}$, $md_{i,j} \in \mathbb{Z}$. Let C be an IP-set in $\mathbb{N}$. By Lemma 1.9, we may pick an increasing sequence $\langle x_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq C \cap m\mathbb{N}$.

Pick $n \in \mathbb{N}$ such that for each $i \in \{1, 2, \ldots, u\}$, $x_n \sum_{j \in I} d_{i,j} + \sum_{j \notin I} d_{i,j} x_1 > 0$.

For $j \in \{1, 2, \ldots, u\}$, let $\alpha_j = x_n$ if $j \in I$ and let $\alpha_j = x_1$ if $j \notin I$. Let

$$\vec{y} = B^{-1} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_u \end{pmatrix}.$$

Then $A\vec{y} \in C^u$ so it suffices to show that $\vec{y} \in \mathbb{N}^u$. Let $i \in \{1, 2, \ldots, u\}$. Then

$$\begin{aligned} y_i &= \sum_{j=1}^{u} d_{i,j} \alpha_j \\ &= \sum_{j \in I} d_{i,j} x_n + \sum_{j \notin I} d_{i,j} x_1. \end{aligned}$$

Since x_n and x_1 are in $m\mathbb{N}$, $y_i \in \mathbb{Z}$. By the choice of x_n, $y_i \in \mathbb{N}$. □

Theorem 3.5. *Let $u, v \in \mathbb{N}$, let A be a $u \times v$ matrix with rational entries and rank u, and assume that B consist of u linearly independent columns of A. Let $D = B^{-1}$ and assume that the first nonzero entry of each row of D is positive. Then A is SIPR/$\mathbb{N}$.*

Proof. By Theorem 3.3, we may presume that $A = B$. Pick $m \in \mathbb{N}$ such that for each $(i, j) \in \{1, 2, \ldots, u\} \times \{1, 2, \ldots, u\}$, $md_{i,j} \in \mathbb{Z}$. Let C be an IP-set in $\mathbb{N}$. By Lemma 1.9, we may pick an increasing sequence $\langle x_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq C \cap m\mathbb{N}$.

For $i \in \{1, 2, \ldots, u\}$, let $\mu(i) = \min\{j \in \{1, 2, \ldots, u\} : d_{i,j} \neq 0\}$. Let $I = \{\mu(i) : i \in \{1, 2, \ldots, u\}\}$, let $k = |I|$, and let $m_1, m_2, \ldots, m_k$ enumerate I in order. Note that $m_1 = 1$.

If $k = 1$, so for all $i \in \{1, 2, \ldots, u\}$, $\mu(i) = 1$, let $\alpha_j = x_1$ if $j > 1$ and pick $n_1 > 1$ such that for all $i \in \{1, 2, \ldots, u\}$, $d_{i,1} x_{n_1} + \sum_{j=2}^{u} d_{i,j} \alpha_j > 0$. Let $\alpha_1 = x_{n_1}$.

Now assume that $k > 1$. For $j \in \{1, 2, \ldots, u\} \setminus I$, if any, let $\alpha_j = x_1$. Pick $n_k > 1$ such that for each i with $\mu(i) = m_k$, $d_{i,m_k} x_{n_k} + \sum_{j=m_k+1}^{u} d_{i,j} \alpha_j > 0$ and let $\alpha_{m_k} = x_{n_k}$.

Given $l \in \{1, 2, \ldots, k-1\}$, having chosen n_{l+1} and $\alpha_{m_{l+1}}$, pick $n_l > 1$ such that for each i with $\mu(i) = m_l$, $d_{i,m_l} x_{n_l} + \sum_{j=m_l+1}^{u} d_{i,j} \alpha_j > 0$ and let $\alpha_{m_l} = x_{n_l}$.

Let

$$\vec{y} = B^{-1}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_u \end{pmatrix}.$$

It suffices to show that $\vec{y} \in \mathbb{N}^u$ so let $i \in \{1, 2, \ldots, u\}$ and pick l such that $\mu(i) = m_l$. Then

$$\begin{aligned} y_i &= \sum_{j=m_l}^{u} d_{i,j}\alpha_j \\ &= d_{i,m_l}x_{n_l} + \sum_{j=m_l+1}^{u} d_{i,j}\alpha_j. \end{aligned}$$

Since each α_j is in $m\mathbb{Z}$, $y_i \in \mathbb{Z}$. By the choice of x_{n_l}, $y_i \in \mathbb{N}$. □

Theorem 3.6. *Let $u \in \mathbb{N}$. Let A be a $u \times u$ matrix with rational entries and rank u, let $D = A^{-1}$, and for $i \in \{1, 2, \ldots, u\}$, let $\vec{c}_i$ be column i of D. If A is SIPR/$\mathbb{N}$, then there exists a nonempty subset I of $\{1, 2, \ldots, u\}$ such that all entries of $\sum_{i\in I}\vec{c}_i$ are nonnegative.*

Proof. Suppose not. For each nonempty $I \subseteq \{1, 2, \ldots, u\}$ let $\vec{f}(I) = \sum_{i\in I}\vec{c}_i$ and pick $s(I) \in \{1, 2, \ldots, u\}$ such that $\vec{f}(I)_{s(I)} < 0$. For $\vec{x}, \vec{y} \in \mathbb{Q}^u$, let $\|\vec{x}-\vec{y}\| = \max\{|x_i - y_i| : i \in \{1, 2, \ldots, u\}\}$.

For this paragraph, fix nonempty $I \subseteq \{1, 2, \ldots, u\}$ and let χ_I be the characteristic function of I. Note that $D\chi_I = \sum_{i\in I}\vec{c}_i = \vec{f}(I)$. Pick $\epsilon(I) > 0$ such that if $\vec{x} \in \mathbb{Q}^u$ and $\|\vec{x} - \chi_I\| < \epsilon(I)$, then $\|D\vec{x} - \vec{f}(I)\| < |\vec{f}(I)_{s(I)}|$.

Let $\epsilon = \min\{\epsilon(I) : \emptyset \neq I \subseteq \{1, 2, \ldots, u\}\}$. Inductively, choose a sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that for each n, $\epsilon x_{n+1} > \sum_{i=1}^{n} x_i$. Pick $F_1, F_2, \ldots, F_u \in \mathcal{P}_f(\mathbb{N})$ and $\vec{y} \in \mathbb{N}^u$ such that

$$A\vec{y} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_u \end{pmatrix}$$

where for each $i \in \{1, 2, \ldots, u\}$, $\alpha_i = \sum_{t\in F_i} x_t$. Pick k such that $\alpha_k = \max\{\alpha_i : i \in \{1, 2, \ldots, u\}\}$ and let $m_k = \max F_k$. We can presume that $m_k > 1$. Now

$$\frac{1}{\alpha_k}D\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_u \end{pmatrix} = D\begin{pmatrix} \alpha_1/\alpha_k \\ \vdots \\ \alpha_u/\alpha_k \end{pmatrix}.$$

Note that, if $i \in \{1, 2, \ldots, u\}$ and $\max F_i < m_k$, then by the choice of the sequence, $0 < \alpha_i/\alpha_k < \epsilon$ while if $\max F_i = m_k$, then $|\alpha_i/\alpha_k - 1| < \epsilon$. To verify the latter statement, note that $\alpha_i/\alpha_k \leq 1$ and $\alpha_i/\alpha_k \geq x_{m_k}/(\sum_{t=1}^{m_k} x_t) = x_{m_k}/(x_{m_k} + \sum_{t=1}^{m_k-1} x_t) > x_{m_k}/(x_{m_k} + \epsilon x_{m_k}) = 1/(1+\epsilon) > 1 - \epsilon$.

Let $I = \{i \in \{1, 2, \ldots, u\} : \max F_i = m_k\}$ and let

$$\vec{x} = \begin{pmatrix} \alpha_1/\alpha_k \\ \vdots \\ \alpha_u/\alpha_k \end{pmatrix}.$$

Then $\|\vec{x} - \chi_I\| < \epsilon \leq \epsilon(I)$ so $\|D\vec{x} - \vec{f}(I)\| < |\vec{f}(I)_{s(I)}|$. Now

$$D\vec{x} = (1/\alpha_k)D\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_u \end{pmatrix} = (1/\alpha_k)\vec{y}$$

so $\|(1/\alpha_k)\vec{y} - \vec{f}(I)\| < |\vec{f}(I)_{s(I)}|$. Then $|(1/\alpha_k)y_{s(I)} - \vec{f}(I)_{s(I)}| < |\vec{f}(I)_{s(I)}|$ so $(1/\alpha_k)y_{s(I)} < 0$, and thus $y_{s(i)} < 0$, a contradiction. □

Finally, we have a special situation where one column of A^{-1} has one zero entry and the rest of its entries are positive.

Theorem 3.7. *Let $u \in \mathbb{N} \setminus \{1\}$, let A be a $u \times u$ matrix with rational entries and rank u, and let $D = A^{-1}$. Assume we have $i, j \in \{1, 2, \ldots, u\}$ such that:*

(1) $d_{i,j} = 0$ *and*

(2) *if $k \in \{1, 2, \ldots, u\} \setminus \{i\}$, then $d_{k,j} > 0$.*

The following statements are equivalent:

(a) *A is SIPR/$\mathbb{N}$.*

(b) *A is IPR/$\mathbb{N}$.*

(c) *There exists $\vec{y} \in \mathbb{N}^u$ such that $A\vec{y} \in \mathbb{N}^u$.*

(d) *There exists $l \in \{1, 2, \ldots, u\}$ such that $d_{i,l} > 0$.*

Proof. That $(a) \Rightarrow (b)$ and $(b) \Rightarrow (c)$ are trivial.

To see that $(c) \Rightarrow (d)$, pick $\vec{y} \in \mathbb{N}^u$ such that $\vec{z} = A\vec{y} \in \mathbb{N}^u$. Suppose that for each $l \in \{1, 2, \ldots, u\}$, $d_{i,l} \leq 0$.

Then $\vec{y} = D\vec{z}$ so $y_i = \sum_{l=1}^{u} d_{i,l}z_i \leq 0$, a contradiction.

To see that $(d) \Rightarrow (a)$, let C be an IP-set in $\mathbb{N}$. Pick $m \in \mathbb{N}$ such that all entries of mD are in $\mathbb{Z}$. By Lemma 1.9, pick an increasing sequence $\langle x_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq C \cap m\mathbb{N}$.

Pick $l \in \{1, 2, \ldots, u\} \setminus \{j\}$ such that $d_{i,l} > 0$. For $t \in \{1, 2, \ldots, u\} \setminus \{j, l\}$ let $\alpha_t = x_1$. Pick n_1 such that $d_{i,l}x_{n_1} + \sum_{t \in \{1,2,\ldots,u\} \setminus \{j,l\}} d_{i,t}\alpha_t > 0$ and let $\alpha_l = x_{n_1}$. Pick n_2 such that for each $k \in \{1, 2, \ldots, u\} \setminus \{i\}$, $d_{k,j}x_{n_2} + \sum_{t \in \{1,2,\ldots,u\} \setminus \{j\}} d_{k,t}\alpha_t > 0$ and let $\alpha_j = x_{n_2}$. If $\vec{y} = D\vec{\alpha}$, then $A\vec{y} = \vec{\alpha} \in C^u$. □

4 Examples

The following theorem will be used in some of the examples of this section.

Theorem 4.1. *Let $v \in \mathbb{N} \setminus \{1\}$ and let A be a $1 \times v$ or $2 \times v$ matrix with rational entries such that A is IPR/$\mathbb{N}$. Then A is SIPR/$\mathbb{N}$.*

Proof. If A has only one row, our claim is immediate from Theorem 3.3 and the trivial fact that the matrix (c) is IPR/$\mathbb{N}$ if and only if $c > 0$, in which case it is also SIPR/$\mathbb{N}$. So we may suppose that A has two rows. By [8, Theorem 15.24(g)], we may pick $m \in \{1, 2\}$, a $v \times m$ matrix G with entries from ω and no row equal to $\vec{0}$, $c \in \mathbb{N}$, and a $2 \times m$ first entries matrix B with entries from ω whose only first entry is c such that $AG = B$. (The fact that $m \leq 2$ is not part of the statement of Theorem 15.24(g), but in the proof that (a) implies (g), $\{I_1, I_2, \ldots, I_m\}$ is a partition of $\{1, 2, \ldots, u\}$.)

Let C be an IP-set in $\mathbb{N}$. We will show that there is some $\vec{y} \in \mathbb{N}^m$ such that $B\vec{y} \in C^2$. Then letting $\vec{x} = G\vec{y}$, we have that $\vec{x} \in \mathbb{N}^v$ and $A\vec{x} \in C^2$.

Assume first that $\text{rank}(B) = 1$ so that $B = \binom{c}{c}$ or for some $b \in \omega$, $B = \left(\begin{smallmatrix} c & b \\ c & b \end{smallmatrix}\right)$. In this case, our claim follows because it holds for matrices with only one row.

So assume that $\text{rank}(B) = 2$. By switching rows if need be, we either have that $B = \left(\begin{smallmatrix} c & a \\ 0 & c \end{smallmatrix}\right)$ for some $a \in \omega$ or $B = \left(\begin{smallmatrix} c & a \\ c & b \end{smallmatrix}\right)$ for some $a, b \in \omega$ with $a < b$. In the first case, our claim follows from Theorem 3.7. So assume that $B = \left(\begin{smallmatrix} c & a \\ c & b \end{smallmatrix}\right)$ for some $a, b \in \omega$ with $a < b$. Pick by Lemma 1.9 a sequence $\langle x_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FS(\langle x_n \rangle_{n=1}^{\infty}) \subseteq c(b-a)\mathbb{N}$. Pick $n \in \mathbb{N} \setminus \{1\}$ such that $x_n > \frac{ax_1}{b-a}$. Let $y_1 = \frac{x_n}{c} - \frac{ax_1}{c(b-a)}$ and let $y_2 = \frac{x_1}{b-a}$. Then $\vec{y} \in \mathbb{N}^2$ and $B\vec{y} = \binom{x_n}{x_n + x_1} \in C^2$. □

Let $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$. A is a first entries matrix so is IPR/$\mathbb{N}$ and so by Theorem 4.1 A is SIPR/$\mathbb{N}$. Now $A^{-1} = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right)$ so A does not satisfy the hypotheses of either Theorem 3.4 or Theorem 3.5 so neither of these sufficient conditions is necessary.

Now let $B = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ and $C = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$. Then $B^{-1} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$ so B satisfies the hypotheses of Theorem 3.4 but not of Theorem 3.5. And $C^{-1} = \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$ so C satisfies the hypotheses of Theorem 3.5 but not of Theorem 3.4. Therefore, the two sufficient conditions are independent.

Let

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Then A is a first entries matrix and

$$A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 2 \\ \frac{1}{2} & -\frac{1}{2} & -1 \end{pmatrix}$$

so A satisfies the hypothesis of Theorem 3.6. It is a consequence of the next theorem that A is not SIPR/$\mathbb{N}$, so the necessary condition of Theorem 3.6 is not sufficient.

Theorem 4.2. *Let*

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

and let $C = FS(\langle 2^{4t} \rangle_{t=1}^{\infty})$. *Then* $\{\vec{y} \in \mathbb{N}^3 : A\vec{y} \in C^3\} = \emptyset$.

Proof. Suppose we have $\vec{y} \in \mathbb{N}^3$ such that

$$A\vec{y} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \in C^3.$$

Pick $F, G, H \in \mathcal{P}_f(\mathbb{N})$ such that $\alpha_1 = \sum_{t\in F} 2^{4t}$, $\alpha_2 = \sum_{t\in G} 2^{4t}$, and $\alpha_3 = \sum_{t\in H} 2^{4t}$ where $F, G, H \in \mathcal{P}_f(\mathbb{N})$.

Then multiplying by A^{-1} we see that $\alpha_2 > 0$, $4\alpha_3 > \alpha_1 - \alpha_2$, and $\alpha_1 - \alpha_2 > 2\alpha_3$.

Let $m = \max H$. Then $2^{4m} \leq \alpha_3 < 2^{4m+1}$ so $2^{4m+1} \leq 2\alpha_3 < 2^{4m+2}$ and $2^{4m+2} \leq 4\alpha_3 < 2^{4m+3}$. Therefore, $2^{4m+1} < \alpha_1 - \alpha_2 < 2^{4m+3}$.

Now $\alpha_1 - \alpha_2 = \sum_{t\in F\setminus G} 2^{4t} - \sum_{t\in G\setminus F} 2^{4t}$. Since $\alpha_1 > \alpha_2$, $F \setminus G \neq \emptyset$. Let $k = \max(F \setminus G)$.

Case 1. $G \setminus F = \emptyset$. Then $2^{4k} \leq \alpha_1 - \alpha_2 < 2^{4k+1}$.

Case 2. $G\setminus F \neq \emptyset$. Let $r = \max(G\setminus F)$ and note that $r < k$. Then $2^{4k} \leq \sum_{t\in F\setminus G} 2^{4t} < 2^{4k+1}$ and $-2^{4r+1} < -\sum_{t\in G\setminus F} 2^{4t} \leq -2^{4r}$ so $2^{4k-1} < 2^{4k} - 2^{4r+1} < \alpha_1 - \alpha_2 < 2^{4k+1} - 2^{4r} < 2^{4k+1}$.

Thus, in either case, $2^{4k-1} < \alpha_1 - \alpha_2 < 2^{4k+1}$. Thus $2^{4k-1} < \alpha_1 - \alpha_2 < 2^{4m+3}$ and $2^{4m+1} < \alpha_1 - \alpha_2 < 2^{4k+1}$. Since $2^{4m+1} < 2^{4k+1}$, $m \leq k-1$. So $2^{4k-1} < 2^{4m+3} \leq 2^{4(k-1)+3} = 2^{4k-1}$, a contradiction. □

We saw in Theorem 3.3 that a strong analogue of Theorem 1.6(d) is valid for SIPR/$\mathbb{N}$ matrices. We shall show now that the natural analogues of Theorem 1.6(e) are not valid for SIPR/$\mathbb{N}$ matrices using two examples. One of these starts with a square matrix and the other ends up with a square matrix. The matrices $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 2 & 4 \end{smallmatrix}\right)$ are first entries matrices so are SIPR/$\mathbb{N}$ by Theorem 4.1. We will see that they cannot be extended by adding a multiple of the row $(1\ 0)$ to $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ nor by adding a multiple of the row $(1\ 0\ 0)$ to $\left(\begin{smallmatrix} 0 & 1 & 1 \\ 1 & 2 & 4 \end{smallmatrix}\right)$.

Let $b \in \mathbb{Q} \setminus \{0\}$, let

$$A = \begin{pmatrix} b & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}$$

and let

$$B = \begin{pmatrix} b & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}.$$

If $b = 1$, we saw in the introduction that A is not SIPR/$\mathbb{N}$ and it is a consequence of Theorem 4.2 that B is not SIPR/$\mathbb{N}$. Further, since

$$B^{-1} = \begin{pmatrix} \frac{1}{b} & 0 & 0 \\ \frac{1}{2b} & 2 & -\frac{1}{2} \\ -\frac{1}{2b} & -1 & \frac{1}{2} \end{pmatrix}$$

it is an immediate consequence of Theorem 3.6 that B is not SIPR/$\mathbb{N}$ if $b \neq 1$. We establish now a stronger result.

Theorem 4.3. *Let $b \in \mathbb{Q} \setminus \{0, 1\}$. Then neither A nor B is IPR/$\mathbb{N}$.*

Proof. First, suppose that A is IPR/$\mathbb{N}$. Then by [8, Theorem 15.24(b)] there exist positive rationals s and t such that

$$D = \begin{pmatrix} bs & 0 & -1 & 0 & 0 \\ s & t & 0 & -1 & 0 \\ s & 2t & 0 & 0 & -1 \end{pmatrix}$$

satisfies the columns condition. For $i \in \{1, 2, 3, 4, 5\}$, let $\vec{c}_i$ be column i of D. In particular, there exists nonempty $I_1 \subseteq \{1, 2, 3, 4, 5\}$ such that $\sum_{i \in I_1} \vec{c}_i = \vec{0}$. One cannot have $I_1 \subseteq \{3, 4, 5\}$. If $2 \in I_1$, then $t = 2t$ contradicting the fact that $t > 0$. So $2 \notin I_1$ and $1 \in I_1$. But then from row 2 one sees that $s = 1$ while from row 1 one sees that $s = \frac{1}{b}$.

Similarly, if one assumes that B is IPR/$\mathbb{N}$ one easily derives a contradiction from the assumption that there exist positive rationals r, s, and t such that

$$\begin{pmatrix} br & 0 & 0 & -1 & 0 & 0 \\ 0 & s & t & 0 & -1 & 0 \\ r & 2s & 4t & 0 & 0 & -1 \end{pmatrix}$$

satisfies the columns condition. □

Our original motive for the current study was [10, Question 4.9].

Definition 4.4. A Q-set in $\mathbb{N}$ is a set which contains a set of the form $\{x_n - x_m : m < n \text{ in } \mathbb{N}\}$ for some increasing sequence $\langle x_n \rangle_{n=1}^{\infty}$ in $\mathbb{N}$.

We remark that every IP-set in $\mathbb{N}$ contains a Q-set in $\mathbb{N}$. Let $\langle x_n \rangle_{n=1}^{\infty}$ be a sequence in $\mathbb{N}$. If $y_n = \sum_{i=1}^{n} x_i$, then $\{y_n - y_m : m, n \in \mathbb{N}, n > m\} \subseteq FS(\langle x_n \rangle_{n=1}^{\infty})$.

Question 4.5. [10] *Let $u, v \in \mathbb{N}$ and let A be a $u \times v$ matrix with entries from ω which is IPR/$\mathbb{N}$ such that* $\operatorname{rank}(A) = u$.
(1) *If C is an IP-set in $\mathbb{N}$, must $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^u\}$ be an IP-set in $\mathbb{N}^v$?*
(2) *If C is a Q-set in $\mathbb{N}$, must $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^u\}$ be a Q-set in $\mathbb{N}^v$?*

By Theorems 2.5 and 4.1, the answer to (1) is "yes" if $u = 2$, even without the rank assumption. By Theorem 4.2, the answer to (1) is "no" if $u = 3$.

The proof of the following theorem is very similar to the proof of Theorem 4.1, but the conclusion is weaker; we cannot assert that $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^2\}$ is a Q-set. That is, we cannot assert the existence of a sequence $\langle \vec{z}(n)\rangle_{n=1}^{\infty}$ in $\mathbb{N}^v$ such that $\vec{z}(n) - \vec{z}(m) \in \{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^2\}$ whenever $m < n$ in $\mathbb{N}$.

Theorem 4.6. *Let $v \in \mathbb{N} \setminus \{1\}$ and let A be a $2 \times v$ matrix with rational entries such that A is IPR/$\mathbb{N}$. If C is a Q-set in $\mathbb{N}$, then $\{\vec{x} \in \mathbb{N}^v : A\vec{x} \in C^2\} \neq \emptyset$.*

Proof. By [8, Theorem 15.24(g)], we may pick $m \in \{1, 2\}$, a $v \times m$ matrix G with entries from ω and no row equal to $\vec{0}$, $c \in \mathbb{N}$, and a $2 \times m$ first entries matrix B with entries from ω whose only first entry is c such that $AG = B$.

Let C be a Q-set in $\mathbb{N}$. We will show that there is some $\vec{y} \in \mathbb{N}^m$ such that $B\vec{y} \in C^2$. Then letting $\vec{x} = G\vec{y}$, we have that $\vec{x} \in \mathbb{N}^v$ and $A\vec{x} \in C^2$.

Assume first that $\operatorname{rank}(B) = 1$ so that $B = \binom{c}{c}$ or for some $b \in \omega$, $B = \binom{c\ b}{c\ b}$. In the former case, let $b = 0$. Pick an increasing sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $\{x_n - x_m : m < n \text{ in } \mathbb{N}\} \subseteq C$. By thinning the sequence, we may assume that for all m and n, $x_n \equiv x_m \pmod{c}$. Pick $n \in \mathbb{N}$ such that $x_n - x_1 > b$. If $B = \binom{c}{c}$, let $y = \frac{x_n - x_1}{c}$ so that $By = \binom{x_n - x_1}{x_n - x_1} \in C^2$. If $B = \binom{c\ b}{c\ b}$, let $\vec{y} = \binom{\frac{x_n - x_1}{c} - b}{c}$. Then $B\vec{y} = \binom{x_n - x_1}{x_n - x_1} \in C^2$.

Now assume that $\operatorname{rank}(B) = 2$. By switching rows if need be, we either have that $B = \binom{c\ a}{0\ c}$ for some $a \in \omega$ or $B = \binom{c\ a}{c\ b}$ for some $a, b \in \omega$ with $a < b$.

Assume first that $B = \binom{c\ a}{0\ c}$ for some $a \in \omega$. Pick an increasing sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $\{x_n - x_m : m < n \text{ in } \mathbb{N}\} \subseteq C$. By thinning the sequence, we may assume that for all m and n, $x_n \equiv x_m \pmod{c^2}$. Pick $n \in \mathbb{N}$ such that $c(x_n - x_1) > a(x_2 - x_1)$. Let $y_1 = \frac{x_n - x_1}{c} - \frac{a(x_2 - x_1)}{c^2}$ and let $y_2 = \frac{x_2 - x_1}{c}$. Then $\vec{y} \in \mathbb{N}^2$ and $B\vec{y} = \binom{x_n - x_1}{x_2 - x_1} \in C^2$.

Now assume that $B = \binom{c\ a}{c\ b}$ for some $a, b \in \omega$ with $a < b$. Pick an increasing sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $\{x_n - x_m : m < n \text{ in } \mathbb{N}\} \subseteq C$. By thinning the sequence, we may assume that for all m and n, $x_n \equiv x_m \pmod{c(b-a)}$. Pick $n \in \mathbb{N} \setminus \{1, 2\}$ such that $x_n - x_2 > \frac{a(x_2 - x_1)}{b-a}$. Let $y_1 = \frac{x_n - x_2}{c} - \frac{a(x_2 - x_1)}{c(b-a)}$ and let $y_2 = \frac{x_2 - x_1}{b-a}$. Then $\vec{y} \in \mathbb{N}^2$ and $B\vec{y} = \binom{x_n - x_2}{x_n - x_1} \in C^2$. □

Theorem 4.7. *Let $A = \binom{1\ 1}{1\ 2}$, let $C = \{2^{2t} - 2^{2s} : s < t \text{ in } \mathbb{N}\}$, and let $B = \{\vec{y} \in \mathbb{N}^2 : A\vec{y} \in C^2\}$. There do not exist $\vec{y}$ and $\vec{z}$ in B such that $\vec{y} + \vec{z} \in B$. In particular, B is not a Q-set.*

Proof. For the "in particular" assertion note that if $\vec{y}(1)$, $\vec{y}(2)$, and $\vec{y}(3)$ are in $\mathbb{N}^2$, then $(\vec{y}(3) - \vec{y}(2)) + (\vec{y}(2) - \vec{y}(1)) = (\vec{y}(3) - \vec{y}(1))$.

For $x \in \mathbb{N}$, let $\phi(x) = \max(\{i \in \omega : 2^i \leq x\})$. Observe that, for every $x, y \in \mathbb{N}$ for which $\phi(x) = \phi(y)$, $\phi(x+y) = \phi(x) + 1$. Observe also that $\phi(x)$ is odd if $x \in C$.

We claim that, if $a, b, a+b \in C$ with $a > b$, there exist $s, t, p \in \mathbb{N}$ with $t > s > p$, such that $a = 2^{2t} - 2^{2s}$ and $b = 2^{2s} - 2^{2p}$. To see this, suppose that $a = 2^{2t} - 2^{2s}$ and $b = 2^{2r} - 2^{2p}$, where $p, r, s, t \in \mathbb{N}$, $t > s$ and $r > p$. Observe that $a + b \in C$ implies that $t > r$, because $\phi(a)$, $\phi(b)$ and $\phi(a+b)$ are odd. Since $2^{2t} - 2^{2s} + 2^{2r} - 2^{2p} = 2^{2n} - 2^{2m}$ for some $m, n \in \mathbb{N}$ with $n > m$, $2^{2t} + 2^{2r} + 2^{2m} = 2^{2s} + 2^{2p} + 2^{2n}$. Now $t > s$ and $t > r > p$. So $t = n$, and hence $2^{2r} + 2^{2m} = 2^{2s} + 2^{2p}$. Since $r > p$, it follows that $r = s$.

Suppose we have $\vec{y}$ and $\vec{z}$ in B such that $\vec{y} + \vec{z} \in B$. Observe that $A^{-1} = \left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right)$. Let $A\vec{y} = \vec{w}$ and $A\vec{z} = \vec{x}$. Then $\vec{w}$ and $\vec{x}$ are in C^2, and $A^{-1}\vec{w}$ and $A^{-1}\vec{x}$ are in $\mathbb{N}^2$. So x_1 and x_2 are in C, and $(1/2)x_2 < x_1 < x_2$. Similarly, w_1 and w_2 are in C, and $(1/2)w_2 < w_1 < w_2$. Now $2^{\phi(x_2)} \leq x_2 < 2^{\phi(x_2)+1}$ so $2^{\phi(x_2)-1} \leq (1/2)x_2 < x_1 < x_2 < 2^{\phi(x_2)+1}$, and thus $\phi(x_1) = \phi(x_2) - 1$ or $\phi(x_1) = \phi(x_2)$. Since $\phi(x_1)$ and $\phi(x_2)$ are odd, $\phi(x_1) = \phi(x_2)$. Also $\phi(w_1) = \phi(w_2)$. This implies that $x_1 = 2^{2t} - 2^{2s}$ and $x_2 = 2^{2t} - 2^{2r}$ for some $s, t, r \in \mathbb{N}$, and $w_1 = 2^{2p} - 2^{2n}$ and $w_2 = 2^{2p} - 2^{2m}$ for some m, n, p in $\mathbb{N}$. Since $A(\vec{y} + \vec{z}) \in C^2$, $\vec{w} + \vec{x} \in C^2$. So $w_1 + x_1$ and $w_2 + x_2$ are in C, and thus $p = s = r$ and $x_1 = x_2$, a contradiction. □

To conclude our discussion of [10, Question 4.9], we show that the matrix of Theorem 4.2 is a strong counterexample to part (2) of that question.

Theorem 4.8. *Let*

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

and let $C = \{2^{4t} - 2^{4s} : s < t \text{ in } \mathbb{N}\}$. *Then* $\{\vec{y} \in \mathbb{N}^3 : A\vec{y} \in C^3\} = \emptyset$.

Proof. Recall that

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

Let $\alpha_1 = 2^{4t} - 2^{4s}$, $\alpha_2 = 2^{4l} - 2^{4k}$, and $\alpha_3 = 2^{4m} - 2^{4r}$.

Suppose we have $\vec{y} \in \mathbb{N}^3$ such that

$$A\vec{y} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix}.$$

Then multiplying by A^{-1} we see that $\alpha_2 > 0$, $4\alpha_3 > \alpha_1 - \alpha_2$ and $\alpha_1 - \alpha_2 > 2\alpha_3$. That is, $2^{4m+2} - 2^{4r+2} > 2^{4t} - 2^{4s} + 2^{4k} - 2^{4l} > 2^{4m+1} - 2^{4r+1}$.

Since $\alpha_1 > \alpha_2$, either both $t = l$ and $k > s$ or $t > l$.

Case 1. $t = l$ and $k > s$. Then $2^{4m+2} > 2^{4m+2} - 2^{4r+2} > 2^{4k} - 2^{4s} > 2^{4m+1} - 2^{4r+1}$ Then $2^{4m+2} + 2^{4s} > 2^{4k}$. The highest power on the left is at least as big as $4k$ and $s < k$ so $4m + 2 \geq 4k$, and thus $m \geq k$.

Also $2^{4k} > 2^{4k} - 2^{4s} > 2^{4m+1} - 2^{4r+1}$ so $2^{4k} + 2^{4r+1} > 2^{4m+1}$. Since $r < m$, we must have $4k \geq 4m + 1$ so $k > m$, a contradiction.

Case 2. $t > l$. Then $2^{4m+2} > 2^{4m+2} - 2^{4r+2} > 2^{4t} - 2^{4s} + 2^{4k} - 2^{4l}$ so $2^{4m+2} + 2^{4s} + 2^{4l} > 2^{4t} + 2^{4k}$. Now $k < l < t$ so the highest power on the right is $4t$. Also $s < t$ and $l < t$ so $4s < 4t$ and $4l < 4t$ and (if $s = l$) $4s + 1 < 4t$ so we must have that $4m + 2 \geq 4t$ and, therefore, $m \geq t$.

Also $2^{4t} + 2^{4k} > 2^{4t} - 2^{4s} + 2^{4k} - 2^{4l} > 2^{4m+1} - 2^{4r+1}$ so $2^{4t} + 2^{4k} + 2^{4r+1} > 2^{4m+1}$. Now $r < m$ and $k < l < t$ so we must have $4t \geq 4m + 1$ so $t > m$, a contradiction. □

It is easy to take a matrix which is not SIPR/$\mathbb{N}$ and make it SIPR/$\mathbb{N}$ by adding a column. For example, $\left(\begin{smallmatrix} 1 & -2 \\ 2 & 1 \end{smallmatrix}\right)$ is not even IPR/$\mathbb{N}$ but $\left(\begin{smallmatrix} 1 & -2 & 0 \\ 2 & 1 & 1 \end{smallmatrix}\right)$ is SIPR/$\mathbb{N}$.

Question 4.9. *Let $u, v \in \mathbb{N}$ with $u < v$ and let A be a $u \times v$ matrix with rational entries and* $\operatorname{rank}(A) = u$ *such that A is SIPR/$\mathbb{N}$. Must there exist u columns of A that form an SIPR/$\mathbb{N}$ matrix with rank u?*

5 Infinite strongly image partition regular matrices

In this section, we allow infinite matrices, so some earlier definitions must be modified. (If u and v are in $\mathbb{N}$, nothing changes.)

Definition 5.1. Let S be a commutative semigroup, let $u, v \in \mathbb{N} \cup \{\omega\}$, and let A be a $u \times v$ matrix. If $S = \mathbb{N}$, then A is *appropriate for S* provided no row of A is zero, the number of nonzero entries in each row is finite, and the entries of A come from $\mathbb{Q}$. If $S \neq \mathbb{N}$ and S is cancellative and, therefore, embeddable in a group, then A is *appropriate for S* provided no row of A is zero, the number of nonzero entries in each row is finite, and the entries of A come from $\mathbb{Z}$. If S is not cancellative, then A is *appropriate for S* provided no row of A is zero, the number of nonzero entries in each row is finite, and the entries of A come from ω.

Except for the fact that the matrix in question is allowed to be infinite, the definitions of IPR/S and SIPR/S remain verbatim the same.

If $S \setminus \{0\}$ is not an IP-set, then any finite matrix which is appropriate for S is vacuously SIPR/S. If $S \setminus \{0\}$ is an IP-set, then any finite identity matrix is SIPR/S. So the number of finite matrices that are SIPR/S is infinite. Since the number of finite matrices with entries from $\mathbb{Q}$ is countable, one can enumerate the finite matrices that are SIPR/S.

We set out to produce an infinite matrix which is SIPR/S. It is based on the results of [3].

Definition 5.2. Let $(S,+)$ be a commutative semigroup. For each $n \in \mathbb{N}$, let $Y_n \in \mathcal{P}_f(S)$. Then

$$FS(\langle Y_n\rangle_{n=1}^{\infty}) = \left\{\sum_{n\in F} x_n : F \in \mathcal{P}_f(\omega) \text{ and for each } n \in F, x_n \in Y_n\right\}.$$

Thus $FS(\langle Y_n\rangle_{n=1}^{\infty})$ is all finite sums choosing at most one term from each Y_n.

The following theorem can be proved using the algebra of βS copying the proof of [8, Theorem 6.16] almost verbatim. We present an elementary proof because it is so simple.

Theorem 5.3. *Let $(S,+)$ be a commutative semigroup such that $S \setminus \{0\}$ is an IP-set. Let $\langle A(n)\rangle_{n=1}^{\infty}$ enumerate the finite matrices that are (appropriate for S and) SIPR/S where each $A(n)$ is a $u(n) \times v(n)$ matrix. Let C be an IP-set contained in $S \setminus \{0\}$. There exists for each $n \in \mathbb{N}$, a choice of $\vec{x}(n) \in (S \setminus \{0\})^{v(n)}$ such that if Y_n is the set of entries of $A(n)\vec{x}(n)$, then $FS(\langle Y_n\rangle_{n=1}^{\infty}) \subseteq C$.*

Proof. Pick a sequence $\langle y_n\rangle_{n=1}^{\infty}$ in S such that $FS(\langle y_n\rangle_{n=1}^{\infty}) \subseteq C$. Pick $\vec{x}(1) \in (S \setminus \{0\})^{v(1)}$ such that $A(1)\vec{x}(1) \in (FS(\langle y_n\rangle_{n=1}^{\infty}))^{u(1)}$. Pick $m(1)$ such that all entries of $A(1)\vec{x}(1)$ are in $FS(\langle y_n\rangle_{n=1}^{m(1)})$. Inductively, let $k \in \mathbb{N}$ and assume we have chosen $\vec{x}(k)$ and $m(k)$. Pick $\vec{x}(k+1) \in (S \setminus \{0\})^{v(k+1)}$ such that $A(k+1)\vec{x}(k+1) \in (FS(\langle y_n\rangle_{n=m(k)+1}^{\infty}))^{u(k+1)}$. Pick $m(k+1)$ such that all entries of $A(k+1)\vec{x}(k+1)$ are in $FS(\langle y_n\rangle_{n=m(k)+1}^{m(k+1)})$. For each $k \in \mathbb{N}$, let Y_k be the set of entries of $A(k)\vec{x}(k)$. Then $FS(\langle Y_n\rangle_{n=1}^{\infty}) \subseteq FS(\langle y_n\rangle_{n=1}^{\infty}) \subseteq C$. □

Corollary 5.4. *Let $(S,+)$ be a commutative semigroup such that $S \setminus \{0\}$ is an IP-set and let $u,v,w \in \mathbb{N}$. Assume that A ia a $u \times v$ matrix which is SIPR/S and B is a $u \times w$ matrix which is SIPR/S. Then the matrix $(A\ B)$ is SIPR/S.*

Proof. Let $\langle A(n)\rangle_{n=1}^{\infty}$ be the enumeration in Theorem 5.3 and pick $n,m \in \mathbb{N}$ such that $A = A(n)$ and $B = A(m)$. Let C be an IP-set contained in $S \setminus \{0\}$. Then all entries of $(A\ B)\begin{pmatrix}\vec{x}(n)\\ \vec{x}(m)\end{pmatrix}$ are in $Y_n + Y_m \subseteq C$. □

The matrix in the following definition is based on the construction of a DH-matrix in [7], which started with an enumeration of all finite matrices with rational entries that are IPR/$\mathbb{N}$.

Definition 5.5. Let $(S,+)$ be a commutative semigroup such that $S \setminus \{0\}$ is an IP-set. A *Strong DH-matrix for S* is an $\omega\times\omega$ matrix **SD** defined as follows. Let $K = \mathbb{Q}$ if $S = \mathbb{N}$, let $K = \mathbb{Z}$ if $S \neq \mathbb{N}$ and S is cancellative, and otherwise let $K = \omega$. First fix an enumeration $\langle A(n)\rangle_{n=0}^{\infty}$ of the finite matrices with entries from K that are SIPR/$\mathbb{N}$. For each n, assume that $A(n)$ is a $u(n)\times v(n)$ matrix. For each $i \in \mathbb{N}$, let $\vec{0}_i$ be the 0 row vector with i entries. Let **SD** be an $\omega\times\omega$ matrix with all rows of the form $\vec{r}_1{}^\frown\vec{r}_2{}^\frown\vec{r}_3{}^\frown\ldots$ where each $\vec{r}_i$ is either $\vec{0}_{v(i)}$ or is a row of $A(i)$, at least one $\vec{r}_i$ is a row of $A(i)$ and for all but finitely many $i \in \mathbb{N}$, $\vec{r}_i = \vec{0}_{v(i)}$.

Corollary 5.6. *Let $(S,+)$ be a commutative semigroup such that $S\setminus\{0\}$ is an IP-set and let* **SD** *be a Strong DH-matrix for S. Then* **SD** *is SIPR/S.*

Proof. Let C be an IP-set contained in $S\setminus\{0\}$. For each $n\in\mathbb{N}$, pick $\vec{x}(n)$ as guaranteed by Theorem 5.3. Then

$$\mathbf{SD}\begin{pmatrix}\vec{x}(1)\\ \vec{x}(2)\\ \vdots\end{pmatrix}\in C^{\omega}.$$

□

We remark that the property of being SIPR/S can be very different for different semigroups S. It follows from the definition of an *IP*-set, that every matrix (finite or infinite) with entries in $\{0,1\}$, which has no row whose entries are all zero and finitely many nonzero entries in each row is SIPR/S for every commutative semigroup S. We will show that these are the only matrices with entries in ω, which have this universal property by considering the semigroup $(\mathbb{N},\cdot)$.

Since the operation is written multiplicatively, some adjustment in notation is required. A set C is an IP-set in $(\mathbb{N},\cdot)$ provided there is a sequence $\langle x_n\rangle_{n=1}^{\infty}$ in $\mathbb{N}$ such that $FP(\langle x_n\rangle_{n=1}^{\infty})\subseteq C$ where $FP(\langle x_n\rangle_{n=1}^{\infty})=\{\prod_{n\in F}x_n : F\in\mathcal{P}_f(\mathbb{N})\}$. The assertion that the $u\times v$ matrix A is SIPR/$(\mathbb{N}\cdot)$ says that whenever C is an IP-set in $(\mathbb{N}\setminus\{1\},\cdot)$ there exists $\vec{x}\in(\mathbb{N}\setminus\{1\})^v$ such that $\vec{x}^A\in C^u$ where the entry in row i of $\vec{x}^A$ is $\prod_{j=1}^{v}x^{a_{i,j}}$.

Assume that A is a $u\times v$ matrix with entries from ω, has no row equal to $\vec{0}$, and has finitely many nonzero entries in each row. Assume that A has some entry $a_{i,j}\in\omega\setminus\{0,1\}$. Let $\langle p_n\rangle_{n=1}^{\infty}$ be the sequence of primes. If $\vec{x}\in(\mathbb{N}\setminus\{1\})^v$, then entry i of $\vec{x}^A$ has a repeated prime factor, so is not in $FP(\langle p_n\rangle_{n=1}^{\infty})$.

The situation is more complicated for matrices with entries in $\mathbb{Z}$. For example, if $A=\left(\begin{smallmatrix}1&-1\\0&1\end{smallmatrix}\right)$, then A is SIPR/S for every commutative cancellative semigroup S because $A\left(\begin{smallmatrix}x_1+x_2\\x_2\end{smallmatrix}\right)=\left(\begin{smallmatrix}x_1\\x_2\end{smallmatrix}\right)$. It would be interesting to characterize the matrices with entries in $\mathbb{Z}$, which have this property.

If S is a Boolean group, then every finite or infinite matrix with entries in $\mathbb{Z}$, is SIPR/S if and only if it has an odd entry in every row. To see this, let B denote the matrix obtained from A by replacing every even entry by 0 and every odd entry by 1. Then, for every column vector $\vec{x}$ with entries in S which has the same number of entries as A has columns, $A\vec{x}=B\vec{x}$.

Bibliography

[1] V. Bergelson and N. Hindman, Nonmetrizable topological dynamics and Ramsey Theory, *Trans. Am. Math. Soc.*, **320** (1990), 293–320.

[2] W. Deuber, Partitionen und lineare Gleichungssysteme, *Math. Z.*, **133** (1973), 109–123.

[3] W. Deuber and N. Hindman, Partitions and sums of (m,p,c)-sets, *J. Comb. Theory, Ser. A*, **45** (1987), 300–302.
[4] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorical Number Theory*, Princeton University Press, Princeton, 1981.
[5] N. Hindman and I. Leader, Image partition regularity of matrices, *Comb. Probab. Comput.*, **2** (1993), 437–463.
[6] N. Hindman, I. Leader and D. Strauss, Image partition regular matrices – bounded solutions and preservation of largeness, *Discrete Math.*, **242** (2002), 115–144.
[7] N. Hindman, I. Leader and D. Strauss, Extensions of infinite partition regular systems, *Electron. J. Comb.*, **22**(2), Paper 2.29 (2015).
[8] N. Hindman and D. Strauss, *Algebra in the Stone-Čech compactification: theory and applications*, 2nd ed., de Gruyter, Berlin, 2012.
[9] N. Hindman and D. Strauss, Image partition regularity of matrices over commutative semigroups, *Topol. Appl.*, **259** (2019), 179–202.
[10] N. Hindman and D. Strauss, Image partition regular matrices and concepts of largeness, *N.Y. J. Math.*, **26** (2020), 230–260.
[11] R. Rado, Studien zur Kombinatorik, *Math. Z.*, **36** (1933), 242–280.
[12] B. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw Arch. Wiskd.*, **19** (1927), 212–216.

Brian Hopkins

Introducing shift-constrained Rado numbers

Abstract: The 2-color Rado number for $ax + by = z$ with positive integers $1 \le a \le b$ is known; this is the least integer such that any 2-coloring of the integers from 1 to the Rado number must include a solution to the equation consisting of numbers that have been assigned the same color. We modify the requirements by introducing constraints on the colorings. These constraints are motivated by symbolic dynamics, specifically the golden mean shift and a version of the even shift, both over a 2-letter alphabet. We establish several initial results, offer some conjectures, and outline possible directions for further research in this new study of shift-constrained Rado numbers.

1 Background

We consider a problem in Ramsey theory on the positive integers informed by symbolic dynamics. We begin with background material for each of these two areas.

For a positive integer r, call a map $\Delta : \{1, \ldots, n\} \to \{0, 1, \ldots, r-1\}$ an r-coloring. A set of integers is monochromatic if Δ assigns the same color to each element of the set. In this article, we limit our attention to $r = 2$. Within the context of a specific coloring, we will write Δ_0 for the set of integers assigned the color zero and similarly Δ_1.

Issai Schur showed that, for any $r \ge 1$, there is a least positive integer $S(r)$ such that any r-coloring of $\{1, \ldots, S(r)\}$ includes a monochromatic solution to the equation $x + y = z$. For example, the second Schur number $S(2) = 5$. There are two steps to verifying such a value. First, one demonstrates that there is a 2-coloring on $\{1, 2, 3, 4\}$ that contains no monochromatic solutions—the coloring given by $\Delta_0 = \{1, 4\}$ and $\Delta_1 = \{2, 3\}$ satisfies this condition, along with the coloring that reverses the assignments to zero and one. Clearly, the same coloring restricted to $\{1, 2, 3\}$ is equally valid, likewise restricted to $\{1, 2\}$ and to $\{1\}$. Second, one argues that any 2-coloring of $\{1, 2, 3, 4, 5\}$ would produce a monochromatic solution: The coloring Δ given for $\{1, 2, 3, 4\}$ (and its reverse) would result in a monochromatic solution for either color assignment of 5 as both $(x, y, z) = (1, 4, 5)$ and $(2, 3, 5)$ are solutions to $x + y = z$, and any other coloring will already have a monochromatic solution on $\{1, 2, 3, 4\}$.

Acknowledgement: The author thanks Katrina Luckenbach and Matthew Vieira who, as undergraduate research students, explored some of these topics. Also, the referee read the manuscript very closely and offered several helpful suggestions. Many computations supporting this research were done with Wolfram Mathematica.

Brian Hopkins, Department of Mathematics, Saint Peter's University, Jersey City, NJ, USA, e-mail: bhopkins@saintpeters.edu

https://doi.org/10.1515/9783110754216-016

Richard Rado extended these ideas to systems of general linear equations and developed a criterion for which systems have a bound analogous to $S(r)$; these are known as Rado numbers. See Landman and Robertson [4, Chapter 9] for additional background. In this article, we restrict our attention to the following parameters, where the Rado numbers are known to be finite.

Definition 1. Given integers a, b with $1 \le a \le b$, the *(2-color) Rado number* is the least positive integer $R(a, b)$ such that any 2-coloring of $\{1, \dots, R(a, b)\}$ includes a monochromatic solution of $ax + by = z$.

Guo and Sun [2] proved

$$R(a, b) = a(a + b)^2 + b, \tag{16.1}$$

confirming a conjecture of the current author and Schaal [3] (the full results allow arbitrarily many variables). Work continues on determining the 2-color Rado number for the fully general linear equation $a_1x_1 + \cdots + a_{m-1}x_{m-1} = a_mx_m + c$; see Thanatipanonda [6] for some recent results towards this goal. There are also related results involving more colors, selected systems and nonlinear equations, and numeric structures beyond the integers. Ron Graham's wide interests in Ramsey theory included an article in this journal with Alexeev and Fox on minimal colorings without rational monochromatic solutions to $x_1 + x_2 + x_3 = 4x_4$ [1].

Moving to our other ingredient, symbolic dynamics begins with a choice of what symbols and patterns are allowed in the underlying system. The full r-shift allows all sequences over the alphabet $\{0, 1, \dots, r - 1\}$. For various theoretical and applied reasons, often only parts of the full shift are used. Subsets of the full shift are called shifts, subshifts, or shift spaces. See Lind and Marcus [5] for additional background.

We focus on two simple binary shifts (so $r = 2$), the golden mean shift and a version of the even shift, described next.

The golden mean shift is defined by forbidding consecutive ones. Table 16.1 shows the short binary sequences that can arise in the golden mean shift; notice that their counts are Fibonacci numbers. Another reason for the name comes from the fact that the entropy of the golden mean shift is $\log(\varphi)$ where $\varphi = (1 + \sqrt{5})/2$; entropy is a useful statistic in symbolic dynamics that is invariant under various operations and measures the information capacity of a shift [5, Chapter 4].

Table 16.1: Allowed length n words in the golden mean shift for small n.

Length	Allowed words	Count
1	0, 1	2
2	00, 01, 10	3
3	000, 001, 010, 100, 101	5
4	0000, 0001, 0010, 0100, 0101, 1000, 1001, 1010	8

The version of the even shift that we use requires that runs of zeros have even length with the possible exception of a terminal run of zeros, which can have any length. Table 16.2 shows the short binary sequences allowed by this shift. For instance, 110 is allowed since the length 1 run of zeros is at the end of the word, but 101 is not allowed since that odd length run of zeros is nonterminal. (In the terminology of symbolic dynamics, this is the follower set of 1 in the even shift [5, Example 3.2.7] or a particular one-sided even shift; we will simply call it the even shift.) Notice that the word counts are the same as for the golden mean shift, which implies that the even shift also has entropy $\log(\varphi)$, but we will see that the two shifts differ notably in other ways.

Table 16.2: Allowed length n words in our even shift for small n.

Length	Allowed words	Count
1	0, 1	2
2	00, 10, 11	3
3	000, 001, 100, 110, 111	5
4	0000, 0010, 0011, 1000, 1001, 1100, 1110, 1111	8

Both of these shifts are closed under taking certain subwords. Specifically, if $x_1 \dots x_n$ is an allowed word in the golden mean shift then, for all $1 \le i \le j \le n$, the word $x_i \dots x_j$ is also allowed. The situation for our even shift is not as strong, but the following holds. If $x_1 \dots x_n$ is an allowed word in the even shift then, for all $1 \le j \le n$, the word $x_1 \dots x_j$ is also allowed.

In Section 2, we combine these ideas to create a new class of Rado number problems. Section 3 presents results related to the shift spaces mentioned here, with proofs in Section 4. We conclude in Section 5 with some conjectures and ideas for further investigations.

2 Shift-constrained Rado numbers

We combine Ramsey theory on the integers and symbolic dynamics by requiring that colorings satisfy the conditions of a shift space. We have seen that shifts can treat different symbols in different ways, thus the Rado numbers for a given equation can vary depending on the color assigned to 1. (In other words, the standard initial step "Without loss of generality, assume that $\Delta(1) = 0$" is no longer valid.)

Definition 2. Given integers a, b with $1 \le a \le b$ and a binary shift S, the *S-constrained Rado number* $R_0^S(a, b)$ is the least positive integer such that any 2-coloring Δ of $\{1, \dots, R_0^S(a, b)\}$ includes a monochromatic solution to $ax + by = z$ where $\Delta(1) = 0$ and Δ satisfies the constraints of the shift S. Define $R_1^S(a, b)$ analogously with $\Delta(1) = 1$.

As a first example, consider $x + y = z$ with the golden mean shift constraint, i. e., there is no i such that $\Delta(i) = \Delta(i+1) = 1$. Note that the coloring detailed above, $\Delta_0 = \{1, 4\}$ and $\Delta_1 = \{2, 3\}$, is no longer valid. Write $R_0^{\varphi}(1, 1)$ and $R_1^{\varphi}(1, 1)$ for these particular shift-constrained Rado numbers; we will call them *golden Rado numbers*.

Suppose $\Delta(1) = 0$. The solution $(1, 1, 2)$ forces $\Delta(2) = 1$. By the golden mean shift requirement, we must have $\Delta(3) = 0$. Then either possible coloring of 4 would give a monochromatic solution, $(1, 3, 4)$ in color zero or $(2, 2, 4)$ in color one. It is easy to check that the 2-coloring on $\{1, 2, 3\}$ given by $\Delta_0 = \{1, 3\}$ and $\Delta_1 = \{2\}$ has no monochromatic solutions to $x + y = z$. We conclude that $R_0^{\varphi}(1, 1) = 4$.

For the other possibility, $\Delta(1) = 1$, the reverse of the first coloring does work as before, since $\Delta_0 = \{2, 3\}$ and $\Delta_1 = \{1, 4\}$ has no consecutive numbers assigned color one. The issues with assigning a color to 5 are the same as before, thus $R_1^{\varphi}(1, 1) = 5$.

For a second example, consider $x + y = z$ with the even shift described above and write $R_0^{e}(1, 1)$, $R_1^{e}(1, 1)$ for these shift-constrained Rado numbers. With $\Delta(1) = 0$, the solution $(1, 1, 2)$ precludes $\Delta(2) = 0$, while $\Delta(2) = 1$ would make an odd length nonterminal run of zeros, so $R^{e}(1, 1) = 1$. For $\Delta(1) = 1$, the now-familiar coloring $\Delta_0 = \{2, 3\}$ and $\Delta_1 = \{1, 4\}$ has only an even length run of zeros and no monochromatic solutions, while no coloring of $\{1, 2, 3, 4, 5\}$ works as before, thus $R_1^{e}(1, 1) = 5$.

In both examples, the closure of the shifts under taking subwords, as discussed in Section 1, means that a coloring of $\{1, \ldots, n\}$ avoiding monochromatic solutions while satisfying the shift constraints restricts to a similarly valid coloring of $\{1, \ldots, k\}$ for all $1 \leq k \leq n - 1$.

The examples are consistent with the following result.

Proposition 1. *Given positive integers a, b with $1 \leq a \leq b$ and a binary shift S, the S-constrained Rado numbers satisfy*

$$R_0^S(a, b) \leq R(a, b) \quad \text{and} \quad R_1^S(a, b) \leq R(a, b).$$

In the case of the full shift F, we have $R_0^F(a, b) = R_1^F(a, b) = R(a, b)$.

Proof. The full shift F introduces no constraints, so the standard Rado number results apply, where color assignments to zero and one are interchangeable. Adding the constraints of a binary shift S may affect colorings such that monochromatic solutions must occur at smaller values, decreasing the Rado number, but requiring additional structure cannot increase the Rado number. □

Note that adding a nontrivial shift constraint does not necessarily decrease the Rado number, as $R_1^{\varphi}(1, 1) = R_1^{e}(1, 1) = R(1, 1) = 5$.

In the next sections, we establish some initial results on $R_0^{\varphi}(a, b)$, $R_1^{\varphi}(a, b)$, and $R_0^{e}(a, b)$, $R_1^{e}(a, b)$.

3 Some golden mean shift and even shift constrained Rado numbers

We begin with equations of the form $x + by = z$ and 2-colorings that satisfy the golden mean shift, i. e., there is no i for which $\Delta(i) = \Delta(i+1) = 1$.

Theorem 1. *Given a positive integer b, the golden Rado numbers are*

$$R_0^{\varphi}(1,b) = \begin{cases} 2b+2 & \text{if } b \text{ is odd,} \\ 2b & \text{if } b \text{ is even;} \end{cases}$$

$$R_1^{\varphi}(1,b) = \begin{cases} 5 & \text{if } b = 1, \\ 3b+1 & \text{if } b \geq 2. \end{cases}$$

The golden Rado numbers for $ax + by = z$ with $a \geq 2$ are more complicated. We prove one result for this shift, the case of equal coefficients. In Section 5, we mention other patterns suggested by computational data.

Theorem 2. *Given a positive integer a, the golden Rado numbers are*

$$R_0^{\varphi}(a,a) = 4a^2,\ R_1^{\varphi}(a,a) = 4a^2 + a.$$

Before considering Rado numbers with colorings constrained by the even shift, it is helpful to recall a 2-coloring detailed by Hopkins and Schaal. The following proposition comes from the proof of [3, Theorem 2] applied to the current case of a three variable equation.

Proposition 2. *Given integers $1 \leq a \leq b$, there are no monochromatic solutions to the equation $ax + by = z$ in the coloring of $\{1, \ldots, a(a+b)^2 + b - 1\}$ specified by*

$$\Delta_0 = \{1, 2, \ldots, a+b-1, (a+b)^2, (a+b)^2+1, \ldots, a(a+b)^2 + b - 1\},$$
$$\Delta_1 = \{a+b, a+b+1, \ldots, (a+b)^2 - 1\}.$$

This coloring consists of length $a+b-1$ and length $(a^2+ab-b)(a+b-1)$ runs of zeros and a length $(a+b)(a+b-1)$ run of ones.

As specified in the next theorem, several even shift constrained Rado numbers, where nonterminal runs of zeroes must have even length, match standard Rado numbers.

Theorem 3. *Given integers $1 \leq a \leq b$, the even shift constrained Rado numbers with $\Delta(1) = 1$ match the standard Rado numbers, i. e.,*

$$R_1^{e}(a,b) = R(a,b) = a(a+b)^2 + b.$$

Also, if $a + b$ is odd, then the same equality holds for colorings with $\Delta(1) = 0$, i. e.,

$$R_0^e(a, b) = R(a, b) = a(a + b)^2 + b.$$

Our final result mirrors Theorem 1, determining Rado numbers for equations of the form $x + by = z$ and 2-colorings that satisfy the even shift constraint.

Theorem 4. *Given a positive integer b, the even shift constrained Rado numbers are*

$$R_0^e(1, b) = \begin{cases} 1 & \text{if } b = 1, \\ b^2 + 3b + 1 & \text{if } b \text{ is even}, \\ b^2 + 2b & \text{if } b \text{ is odd and } b \geq 3; \end{cases}$$
$$R_1^e(1, b) = b^2 + 3b + 1.$$

Note that several cases match the standard Rado number $R(1, b) = b^2 + 3b + 1$.

4 Proofs of results

The proofs of the four theorems are elementary but sometimes a bit lengthy. When arguments are analogous to previous ones, we skip some details.

Proof of Theorem 1. We organize the proof into four cases; it is helpful to split the $R_1^{\varphi}(1, b)$ proof into cases for even b and odd b even though the conclusion is the same for both. Each case requires two things. First, we demonstrate a coloring from 1 to one less than the Rado number that satisfies the conditions of the golden mean shift and contains no monochromatic solutions to $x + by = z$. Second, we show that any coloring from 1 to the Rado number satisfying the shift constraint includes a monochromatic solution.

(a) Consider the case b odd and colorings with $\Delta(1) = 0$. Write $b = 2k - 1$. We show that the coloring of $\{1, \dots, 4k - 1\}$ specified by

$$\Delta_0 = \{1, 3, \dots, 4k - 1\}, \ \Delta_1 = \{2, 4, \dots, 4k - 2\}$$

includes no monochromatic solutions to $x + (2k - 1)y = z$; clearly, it satisfies the golden mean shift condition. By parity arguments, any solution has exactly one or three of x, y, z even. Therefore, there can be no monochromatic solution with $x, y, z \in \Delta_0$. Now suppose $x, y \in \Delta_1$. Since $x + (2k - 1)y > 2 + (4k - 2) = 4k$, beyond the range of the coloring, there is no monochromatic solution in color one. This valid coloring shows that $R_0^{\varphi}(1, 2k - 1) \geq 4k = 2b + 2$.

To complete this case, we show that any coloring of $\{1, \dots, 4k\}$ with $\Delta(1) = 0$ and satisfying the golden mean shift condition includes a monochromatic solution to $x + (2k - 1)y = z$.

If $\Delta(2k) = 0$, then $(1, 1, 2k)$ would be a monochromatic solution in color zero, so we may assume that $\Delta(2k) = 1$. By the golden mean shift constraint, $\Delta(2k-1) = 0$ and $\Delta(2k+1) = 0$.

If $\Delta(2) = 0$, then $(2, 1, 2k+1)$ would be a monochromatic solution, so we may assume that $\Delta(2) = 1$.

Now either color assignment for $4k$ gives a monochromatic solution. If $\Delta(4k) = 0$, then $(2k+1, 1, 4k)$ would be a monochromatic solution in color zero. If $\Delta(4k) = 1$, then $(2, 2, 4k)$ would be a monochromatic solution in color one. This shows that $R_0^{\varphi}(1, 2k-1) \leq 4k$.

Together with the bound from the valid coloring, we conclude $R_0^{\varphi}(1, 2k-1) = 4k$, i. e., $R_0^{\varphi}(1, b) = 2b + 2$ for $b = 2k - 1$.

(b) Consider the case b even and colorings with $\Delta(1) = 0$. Write $b = 2k$. We will show that the coloring of $\{1, \ldots, 4k-1\}$ specified by

$$\Delta_0 = \{1, 3, \ldots, 2k-1, 2k, 2k+2, \ldots, 4k-2\},$$
$$\Delta_1 = \{2, 4, \ldots, 2k-2, 2k+1, 2k+3, \ldots, 4k-1\}$$

includes no monochromatic solutions to $x + 2ky = z$; clearly, it satisfies the golden mean shift condition. It is straightforward to verify that this coloring is valid: For a solution (x, y, z) to have $z < 4k$ requires $x < 2k$ and $y = 1$, and in that range parity arguments similar to those in (a) can be made, etc. The validity of this coloring also follows from the second part of this case.

We show that any coloring of $\{1, \ldots, 4k\}$ with $\Delta(1) = 0$ and satisfying the golden mean shift condition includes a monochromatic solution to $x + 2ky = z$.

If $\Delta(2k+1) = 0$, then $(1, 1, 2k+1)$ would be a monochromatic solution, so we may assume that $\Delta(2k+1) = 1$. By the golden mean shift constraint, $\Delta(2k) = 0$ and $\Delta(2k+2) = 0$.

If $\Delta(2) = 0$, then $(2, 1, 2k+2)$ would be a monochromatic solution, so we may assume that $\Delta(2) = 1$. It follows that $\Delta(3) = 0$.

If $\Delta(2k+3) = 0$, then $(3, 1, 2k+3)$ would be a monochromatic solution, so we may assume that $\Delta(2k+3) = 1$. It follows that $\Delta(2k+4) = 0$.

If $\Delta(4) = 0$, then $(4, 1, 2k+4)$ would be a monochromatic solution, so we may assume that $\Delta(4) = 1$. It follows that $\Delta(5) = 0$.

This bootstrapping continues through $\Delta(4k-2) = 0$ by the golden mean shift constraint.

If $\Delta(2k-2) = 0$, then $(2k-2, 1, 4k-2)$ would be a monochromatic solution, so we may assume that $\Delta(2k-2) = 1$. It follows that $\Delta(2k-1) = 0$.

If $\Delta(4k-1) = 0$, then $(2k-1, 1, 4k-1)$ would a monochromatic solution, so we may assume that $\Delta(4k-1) = 1$.

At this point, we have shown that the coloring described above for $\{1, \ldots, 4b-1\}$ does not contain any monochromatic solutions to $x + 2ky = z$. (It is the unique such coloring, as the color assignments have all been forced.)

Now either color assignment for $4k$ gives a monochromatic solution. If $\Delta(4k) = 0$, then $(2k, 1, 4k)$ would be a monochromatic solution in color zero. If $\Delta(4k) = 1$, then Δ would violate the golden mean shift constraint, as the valid coloring requires $\Delta(4k - 1) = 1$. We conclude that $R_0^{\varphi}(1, 2k) = 4k$, i. e., $R_0^{\varphi}(1, b) = 2b$ for $b = 2k$.

(c) Consider the case b even and colorings with $\Delta(1) = 1$. Write $b = 2k$. The coloring of $\{1, \dots, 6k\}$ specified by

$$\Delta_0 = \{2, 4, \dots, 2k, 2k + 1, 2k + 2, \dots, 4k + 1, 4k + 3, \dots, 6k - 1\},$$
$$\Delta_1 = \{1, 3, \dots, 2k - 1, 4k + 2, 4k + 4, \dots, 6k\}$$

includes no monochromatic solutions to $x + 2ky = z$; clearly, it satisfies the golden mean shift condition. Starting with this case, we omit verifications that the given colorings are valid. As in (b), one can show that this is the unique valid coloring in this case.

We show that any coloring of $\{1, \dots, 6k + 1\}$ with $\Delta(1) = 1$ and satisfying the golden mean shift constraint includes a monochromatic solution to $x + 2ky = z$.

Since $\Delta(1) = 1$, the golden mean shift constraint requires $\Delta(2) = 0$.

If $\Delta(2k + 1) = 1$, then $(1, 1, 2k + 1)$ would be a monochromatic solution, so we may assume that $\Delta(2k + 1) = 0$.

If $\Delta(4k + 2) = 0$, then $(2, 2, 4k + 2)$ would be a monochromatic solution, so we may assume that $\Delta(4k + 2) = 1$. It follows that $\Delta(4k + 3) = 0$.

If $\Delta(3) = 0$, then $(3, 2, 4k + 3)$ would be a monochromatic solution, so we may assume that $\Delta(3) = 1$.

Now either color assignment for $6k + 1$ gives a monochromatic solution. If $\Delta(6k + 1) = 0$, then $(2k + 1, 2, 6k + 1)$ would be a monochromatic solution in color zero. If $\Delta(6k + 1) = 1$, then $(1, 3, 6k + 1)$ would be a monochromatic solution in color one. With the valid coloring above, we conclude that $R_1^{\varphi}(1, 2k) = 6k + 1$, i. e., $R_1^{\varphi}(1, b) = 3b + 1$ for b even.

(d) Consider the case b odd and colorings with $\Delta(1) = 1$. We established in Section 2 that $R_1^{\varphi}(1, 1) = 5$, so we assume $b \geq 3$. Write $b = 2k - 1$. The coloring of $\{1, \dots, 6k - 3\}$ specified by

$$\Delta_0 = \{2, 4, \dots, 2k - 2, 2k, 2k + 1, \dots, 4k - 1, 4k + 1, 4k + 3, \dots, 6k - 3\},$$
$$\Delta_1 = \{1, 3, \dots, 2k - 1, 4k, 4k + 2, \dots, 6k - 4\}$$

includes no monochromatic solutions to $x + (2k - 1)y = z$; clearly, it satisfies the golden mean shift condition. In fact, it is the unique valid coloring in this case.

The argument to show that any coloring of $\{1, \dots, 6k - 2\}$, with $\Delta(1) = 1$ and satisfying the golden mean shift condition, includes a monochromatic solution to $x + (2k - 1)y = z$ is very similar to the steps of (c). One can show

$$2, 2k, 4k + 1 \in \Delta_0,\ 1, 3, 4k \in \Delta_1$$

from which no color assignment for $6k-2$ is valid due to the solutions $(2k, 2, 6k-2)$ and $(1, 3, 6k-2)$. With the coloring above, we conclude that $R_1^{\varphi}(1, 2k-1) = 6k-2$, i. e., $R_1^{\varphi}(1, b) = 3b+1$ for odd $b \geq 3$. □

Proof of Theorem 2. For the equation $ax+ay=z$, first consider colorings with $\Delta(1)=0$. The coloring of $\{1, \ldots, 4a^2-1\}$ specified by

$$\begin{aligned}\Delta_0 = \{&1, 2, \ldots, 2a-1, 2a+1, 2a+2, \ldots, 3a-1, 3a+1, 3a+2, \ldots,\\ &4a^2-a-1, 4a^2-a+1, 4a^2-a+2, \ldots, 4a^2-1\},\\ \Delta_1 = \{&2a, 3a, \ldots, 4a^2-a\}\end{aligned}$$

includes no monochromatic solutions to $ax+ay=z$; clearly, it satisfies the golden mean shift condition. Following the convention adapted in the previous proof, we do not verify that the coloring in valid. We mention in passing that the color assignments are forced except for the integers greater than $4a^2-a$ (although the golden mean shift constraint still applies).

We show that any coloring of $\{1, \ldots, 4a^2\}$ with $\Delta(1)=0$ and satisfying the golden mean shift condition includes a monochromatic solution to $ax+ay=z$.

If $\Delta(2a)=0$, then $(1, 1, 2a)$ would be a monochromatic solution, so we may assume that $\Delta(2a)=1$. It follows that $\Delta(2a-1)=0$ and $\Delta(2a+1)=0$.

Already, either color assignment for $4a^2$ gives a monochromatic solution. If $\Delta(4a^2)=0$, then $(2a-1, 2a+1, 4a^2)$ would be a monochromatic solution in color zero. If $\Delta(4a^2)=1$, then $(2a, 2a, 4a^2)$ would be a monochromatic solution in color one. With the valid coloring above, we conclude that $R_0^{\varphi}(a, a) = 4a^2$.

Second, consider colorings with $\Delta(1)=1$. The coloring of $\{1, \ldots, 4a^2+a-1\}$ specified by

$$\begin{aligned}\Delta_0 = \{&2, 3, \ldots, 4a-1, 4a+1, 4a+2, \ldots, 5a-1, 5a+1, 5a+2, \ldots,\\ &4a^2-1, 4a^2+1, 4a^2+2, \ldots, 4a^2+a-1\},\\ \Delta_1 = \{&1, 4a, 5a, \ldots, 4a^2\}\end{aligned}$$

includes no monochromatic solutions to $ax+ay=z$; clearly, it satisfies the golden mean shift condition. The color assignments are forced except for the integers greater than $4a^2$ (although the golden mean shift constraint still applies).

The argument that any coloring of $\{1, \ldots, 4a^2+a\}$, with $\Delta(1)=1$ and satisfying the golden mean shift condition, includes a monochromatic solution to $ax+ay=z$ is very similar to the $\Delta(1)=0$ case. One can show

$$2, 2a, 4a-1 \in \Delta_0,\ 1, 4a \in \Delta_1$$

from which no color assignment for $4a^2+a$ would be valid due to the solutions $(2, 4a-1, 4a^2+a)$ and $(1, 4a, 4a^2+a)$. With the valid coloring above, we conclude that $R_1^{\varphi}(a, a) = 4a^2+a$. □

Proof of Theorem 3. For the equation $ax+by=z$, first consider colorings with $\Delta(1)=0$ constrained by the even shift. In the cases that the coloring described in Proposition 2 has only even length nonterminal runs of zeros, that coloring shows $R_0^e(a,b) \geq R(a,b)$. That will complete the proof of this case since $R_0^e(a,b) \leq R(a,b)$ by Proposition 1.

As described after Proposition 2, the length of the initial run of zeros is $a+b-1$ which is even exactly when $a+b$ is odd. Thus $R_0^e(a,b)=R(a,b)$ when $a+b$ is odd and equation (16.1) provides the formula.

Second, consider colorings with $\Delta(1)=1$. The coloring described in Proposition 2 is for standard Rado numbers, so the assignments to zero and one are interchangeable. Reversing the zero and one colors to satisfy $\Delta(1)=1$ results in one run of zeros, length $(a+b)(a+b-1)$, which is always even. Therefore, the reversed coloring satisfies the even shift constraint and, as before, we conclude $R_1^e(a,b)=R(a,b)=a(a+b)^2+b$. □

Proof of Theorem 4. The $b=1$ case was treated in Section 2. Theorem 3 applies to the remaining cases except $R_0^e(1,b)$ when b is odd, so we take $b \geq 3$.

Let $b=2k-1$ with $k \geq 2$ and consider colorings with $\Delta(1)=0$. We want to show that the even shift constrained Rado number is $b^2+2b=4k^2-1$. First, we show that the coloring of $\{1,\ldots,4k^2-2\}$ specified by

$$\Delta_0 = \{1,2,\ldots,2k-2,4k^2-2k,4k^2-2k+1,\ldots,4k^2-2\},$$
$$\Delta_1 = \{2k-1,2k,\ldots,4k^2-2k-1\}$$

includes no monochromatic solutions to $x+(2k-1)y=z$. Since the only nonterminal run of zeros has length $2k-2$, the coloring satisfies the even shift condition. The standard argument concerning monochromatic solutions is straightforward (i. e., $x,y \leq 2k-2$ in Δ_0 give $z \in \Delta_1$, then $x,y \in \Delta_1$ give $z \geq 4k^2-2k$, etc.), but we will show that this is the unique valid coloring in the second part of the proof.

We show that any coloring of $\{1,\ldots,4k^2-1\}$ with $\Delta(1)=0$ and satisfying the even shift condition would include a monochromatic solution to $x+by=z$.

Let c be the least integer such that $\Delta(c)=1$. We show that $c=2k-1$. Since $\Delta(1)=0$, we know $c \geq 2$. By the definition of c, we have $\Delta(c-1)=0$.

If $\Delta(2ck)=1$, then $(c,c,2ck)$ would be a monochromatic solution, so we may assume that $\Delta(2ck)=0$.

If $\Delta(2k+c-1)=0$, then $(2k+c-1,c-1,2ck)$ would be a monochromatic solution, so we may assume that $\Delta(2k+c-1)=1$.

Either color assignment for $(2c+2)k-1$ would give a monochromatic solution. If $\Delta((2c+2)k-1)=0$, then $(2ck,1,(2c+2)k-1)$ would be a monochromatic solution in color zero. If $\Delta((2c+2)k-1)=1$, then $(2k+c-1,c,(2c+2)k-1)$ would be a monochromatic solution in color one.

To avoid this problematic $(2c+2)k-1$ in the range of integers of the valid coloring, we need $(2c+2)k-1>4k^2-2$, equivalently $c>2k-1-1/(2k)$. By the solution $(1,1,2k)$,

we must have $\Delta(2k) = 1$. Thus the range for c not leading to a monochromatic solution is

$$2k - 1 - \frac{1}{2k} < c \leq 2k.$$

Of the two possible integer values, it must be that $c = 2k - 1$ since the initial run of zeros needs to have even length to satisfy the shift constraint. Keeping track of the color assignments, so far we have

$$1, \ldots, 2k - 2 \in \Delta_0,\ 2k - 1, 2k \in \Delta_1.$$

In addition to $2k - 1, 2k \in \Delta_1$, the initial run of zeros forces many integers to be assigned color one. Specifically, by the solution $(2, 1, 2k + 1)$ we must have $\Delta(2k + 1) = 1, \ldots,$ by $(2k - 2, 1, 4k - 3)$ we must have $\Delta(4k - 3) = 1$. By the solutions $(1, 2, 4k - 1)$ through $(2k - 2, 2, 6k - 4)$, we must have $4k - 1, \ldots, 6k - 4 \in \Delta_1$. Note that $4k - 2$ has not been assigned a color. If $\Delta(4k - 2) = 0$, then there would be a length one nonterminal run of zeros, so $\Delta(4k - 2) = 1$ by the even shift constraint. All of this continues through the solution $(2k - 2, 2k - 2, 4k^2 - 4k)$ implying $\Delta(4k^2 - 4k) = 1$. That is,

$$2k + 1, \ldots, 4k^2 - 4k \in \Delta_1.$$

Next, we show the color zero assignments that follow from this run of ones. By the solution $(2k - 1, 2k - 1, 4k^2 - 2k)$, we must have $\Delta(4k^2 - 2k) = 0$. By the solutions $(2k, 2k - 1, 4k^2 - 2k + 1)$ through $(4k - 3, 2k - 1, 4k^2 - 2)$, we have the length $2k - 1$ run of zeros

$$4k^2 - 2k, \ldots, 4k^2 - 2 \in \Delta_0. \tag{16.2}$$

These additional color zero assignments allow us to extend the run of ones. If $\Delta(4k^2 - 4k + 1) = 0$, then $(4k^2 - 4k + 1, 1, 4k^2 - 4k)$ would be a monochromatic solution in color zero, so we may assume that $\Delta(4k^2 - 4k + 1) = 1$. This continues through the solution $(4k^2 - 2k - 1, 1, 4k^2 - 2)$, so that

$$4k^2 - 4k + 1, \ldots, 4k^2 - 2k - 1 \in \Delta_1.$$

This establishes the validity and uniqueness of the coloring given at the beginning of the proof.

To complete the proof, we show that neither color assignment for $4k^2 - 1$ is valid. If $\Delta(4k^2 - 1) = 0$, then $(4k^2 - 2k, 1, 4k^2 - 1)$ would be a monochromatic solution in color zero. If $\Delta(4k^2 - 1) = 1$, then equation (16.2) would describe an odd length nonterminal run of zeros. (Also, $(4k - 2, 2k - 1, 4k^2 - 1)$ would be a monochromatic solution in color one.)

We conclude that $R_0^e(1, 2k - 1) = 4k^2 - 1$ for $k \geq 2$, i. e., $R_0^e(1, b) = b^2 + 2b$ for odd $b \geq 3$. □

5 Conjectures and other ideas for further study

Here are some possible next investigations in this new study of shift-constrained Rado numbers.

For the golden mean shift, computation data suggest additional identities, two of which we include here as conjectures.

Conjecture 1. Given positive integers a and ℓ, the golden Rado number for the case $\Delta(1) = 0$ is $R_0^{\varphi}(a, \ell a) = (\ell + 1)^2 a^2$.

Theorem 2 confirms the $\ell = 1$ case of Conjecture 1.

By Theorem 1, we know $R_0^{\varphi}(1, b) < R_1^{\varphi}(1, b)$ for all positive integers b. Also, by Theorem 2, $R_0^{\varphi}(a, a) < R_1^{\varphi}(a, a)$ for all positive integers a. In general, though, the relation between $R_0^{\varphi}(a, b)$ and $R_1^{\varphi}(a, b)$ is not clear. That is, which color assignment for 1 has the greater effect on lowering the standard Rado number when applying the golden mean shift constraint? Computational data support the following claim.

Conjecture 2. Given integers $2 \le a \le b$, the golden mean shift constrained Rado numbers satisfy $R_0^{\varphi}(a, b) < R_1^{\varphi}(a, b)$ except when $b = \ell a$ for some integer $\ell \ge 2$, in which case $R_0^{\varphi}(a, \ell a) > R_1^{\varphi}(a, \ell a)$.

Of course, a full understanding of $R_0^{\varphi}(a, b)$, $R_1^{\varphi}(a, b)$ is desired. Similarly, for the even shift, where we have not determined $R_0^{e}(a, b)$ for $a + b$ even with $a > 1$. (The analogue of Conjecture 2 for the even shift constraint is clear, as we know from Proposition 1 and Theorem 3 that $R_0^{e}(a, b) \le R_1^{e}(a, b) = R(a, b)$.)

Following developments in the study of Rado numbers, one can generalize from $ax + by = z$ to equations with arbitrarily many variables, including a constant term, requiring $x \ne y$, etc. In symbolic dynamics, there are many interesting shifts to explore, including run-length limited shifts, charge constrained shifts, and generalized Morse shifts, along with larger alphabets corresponding to more colors.

Bibliography

[1] B. Alexeev, J. Fox and R. Graham, On minimal coloring without monochromatic solutions to a linear equation, *Integers*, **7**(2) (2007), Article #A1.

[2] S. Guo and Z.-W. Sun, Determination of the two-color Rado number for $a_1x_1 + \cdots + a_mx_m = x_0$, *J. Comb. Theory, Ser. A*, **115** (2008), 345–353.

[3] B. Hopkins and D. Schaal, On Rado numbers for $\sum_{i=1}^{m-1} a_i x_i = x_m$, *Adv. Appl. Math.*, **35** (2005), 433–441.

[4] B. Landman and A. Robertson, *Ramsey Theory on the Integers*, 2nd ed., American Mathematical Society, Providence, RI, 2014.

[5] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, 2nd ed., Cambridge University Press, Cambridge, 2021.

[6] T. Thanatipanonda, Rado numbers of regular nonhomogeneous equations, *Discrete Math. Lett.*, **4** (2020), 5–10.

Jared Duker Lichtman

Mertens' prime product formula, dissected

Abstract: In 1874, Mertens famously proved an asymptotic formula for the product of $p/(p-1)$ over all primes p up to x. Observe that this product equals the reciprocal sum of all integers composed of prime factors up to x. It is natural to restrict such series to integers with a fixed number k of prime factors. In this article, we obtain formulae for these series for each k, which together dissect Mertens' original estimate. The proof is by elementary methods of a combinatorial flavor.

1 Introduction

We begin with the Euler–Mascheroni constant $\gamma = 0.57721\ldots$, defined as the limit of the difference between the harmonic series up to x and $\log x$. The ubiquitous constant γ crops up in many contexts, notably, in the third of three results from a celebrated paper of Mertens [6] on the distribution of prime numbers.

As notation, throughout we write $f(x) = O(g(x))$ and $f(x) \ll g(x)$ to mean $|f(x)/g(x)|$ is bounded, while $f(x) \sim g(x)$ means $\lim_{x\to\infty} f(x)/g(x) = 1$. Also, let $\log_2 x = \log\log x$, and let p denote a prime number.

Theorem 1.1 (Mertens (1874)). *There exists a constant $\beta > 0$ for which*

$$\sum_{p\le x} \frac{\log p}{p} = \log x + O(1), \quad \sum_{p\le x} \frac{1}{p} = \log_2 x + \beta + O\Big(\frac{1}{\log x}\Big) \tag{17.1}$$

$$\prod_{p\le x}\Big(1-\frac{1}{p}\Big)^{-1} \sim e^{\gamma}\log x. \tag{17.2}$$

Here, $\beta = 0.26149\ldots$ is Mertens' constant, which is known to satisfy

$$\beta-\gamma = \sum_{p}\Big(\log\Big(1-\frac{1}{p}\Big)+\frac{1}{p}\Big) = -\sum_{p}\sum_{j\ge 2}\frac{p^{-j}}{j} = -\sum_{j\ge 2}\frac{Z(j)}{j}, \tag{17.3}$$

where $Z(s) = \sum_p p^{-s}$ denotes the prime zeta function, for $s > 1$; see, for instance, Theorem 2.7 in [7, p. 50].

Acknowledgement: The author is grateful to Paul Kinlaw and James Maynard for stimulating conversations, as well as to Paul Pollack and Carl Pomerance for valuable feedback. In addition, the author thanks the anonymous referee for comments to clarify the paper. The author is supported by a Clarendon Scholarship at the University of Oxford.

Jared Duker Lichtman, Mathematical Institute, University of Oxford, Oxford, United Kingdom, e-mail: jared.d.lichtman@gmail.com

https://doi.org/10.1515/9783110754216-017

Now by expanding Mertens' prime product in equation (17.2), we have

$$\prod_{p\le x}\left(1-\frac{1}{p}\right)^{-1}=\prod_{p\le x}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right)=\sum_{P^+(n)\le x}\frac{1}{n} \tag{17.4}$$

where $P^+(n)$ denotes the largest prime factor of n.

Consider "dissecting" the sum in equation (17.4) according to the number of prime factors of n with multiplicity, denoted $\Omega(n)$. Our main result is an asymptotic formula for this dissected sum.

Theorem 1.2. *For each fixed $k \ge 1$, we have*

$$\sum_{\substack{\Omega(n)=k\\ P^+(n)\le x}}\frac{1}{n}=\sum_{j=0}^{k}\frac{c_{k-j}}{j!}(\log_2 x+\beta)^j+O_k\left(\frac{(\log_2 x)^{k-1}}{\log x}\right) \tag{17.5}$$

where the sequence $(c_k)_{k=0}^{\infty}$ is recursively defined by $c_0 = 1$ and

$$c_k=\frac{1}{k}\sum_{j=2}^{k}c_{k-j}Z(j). \tag{17.6}$$

Theorem 1.2 may be viewed as a "dissection" of Mertens' prime product formula. Indeed, as shown later in equation (17.18), the main term $e^y \log x$ in equation (17.2) may be expressed as the series over all $k \ge 1$ of the main terms in equation (17.5) (i. e., the sum over $j \le k$). Here, "dissection" is meant to highlight the formal compatibly of main terms. Whereas the estimate equation (17.5) itself does not necessarily hold uniformly over all $k \ge 1$.

We note that this terminology was introduced by Pollack [8], who dissected a classical mean value theorem of Hall and Tenenbaum.

1.1 Uniform estimates via complex analysis

Classically, the analogous series to equation (17.5) has been studied, replacing the condition $P^+(n) \le x$ with the more common $n \le x$.

Mertens' first theorem implies, by induction on each fixed $k \ge 1$,

$$\sum_{\substack{\Omega(n)=k\\ n\le x}}\frac{1}{n}\sim\frac{(\log_2 x)^k}{k!} \tag{17.7}$$

as $x \to \infty$; see [7, p. 228]. Note that equation (17.7) is historically attributed to Landau [4]. This is another example of dissection, as the sum over all k of each side gives

$\sum_{n\le x}\frac{1}{n}$ and $\log x$, respectively. We also note that the asymptotic (17.7) also holds with $\Omega(n)$ replaced by $\omega(n)$, the number of distinct prime factors of n.

However, equation (17.7) only holds for fixed k. The celebrated theorem of Sathe and Selberg implies the following uniform estimate for k less than $2\log_2 x$.

Theorem 1.3 (Sathe–Selberg). *Define* $\nu(z) = \frac{1}{\Gamma(z+1)}\prod_p(1-\frac{z}{p})^{-1}(1-\frac{1}{p})^z$, *and let* $r = k/\log_2 x$. *For any* $\varepsilon > 0$, *as* $x\to\infty$ *we have uniformly for* $r \le 2-\varepsilon$,

$$\sum_{\substack{\Omega(n)=k\\ n\le x}}\frac{1}{n} \sim \nu(r)\frac{(\log_2 x)^k}{k!}. \tag{17.8}$$

To see this, [7, Theorem 7.19] or [12, Theorem 6.5] gives an asymptotic in the stated range

$$\sum_{\substack{\Omega(n)=k\\ n\le x}} 1 = \nu(r)\frac{x}{\log x}\frac{(\log_2 x)^{k-1}}{(k-1)!}\left(1+O_\varepsilon\left(\frac{k}{\log_2 x}\right)\right).$$

Then equation (17.8) follows by partial summation, combined with, e. g., the Erdős–Sarközy upper bound $O(k^4 2^{-k} x\log x)$ uniformly for all $x, k \ge 1$; see [2].

Remark 1. As $\nu(r) = 1$ only when $r = 0, 1$, Landau's estimate (17.7) holds if and only if $k = o(\log_2 x)$ or $k = (1+o(1))\log_2 x$.

Remark 2. [12, Theorem 6.4] gives an analogous result with $\Omega(n)$ replaced by $\omega(n)$, by substituting the function $\nu(z)$ above with $\lambda(z) = \frac{1}{\Gamma(z+1)}\prod_p(1+\frac{z}{p-1})(1-\frac{1}{p})^z$.

Remark 3. The Sathe–Selberg theorem is proved through contour integration in the complex plane. Recently, Popa [9, 10] and Tenenbaum [13] have obtained results by similar analytic methods, for a generalized series that replaces the conditions $\Omega(n) = k$ and $n \le x$ by the condition $p_1\cdots p_k \le x$ over k independent prime variables. Or equivalently, they weight n by the number of its ordered prime factorizations.

The uniformity coming from sophisticated analytic tools exemplifies the larger tension within mathematics, between proving the strongest results and using the simplest arguments. Of particular interest historically is the case $k = 1$, i. e., the prime number theorem. Hadamard and de la Vallée Poussin initially gave proofs in 1896 using complex analysis, and for decades many believed it impossible to prove by elementary means. It came as a great shock when Selberg and Erdős did so in 1948. For an intriguing historical account, see Spencer and Graham [11].

As such, we emphasize that in Theorem 1.2, our particular conditions $\Omega(n) = k$, $P^+(n) \le x$ in equation (17.5) are directly amenable to elementary methods when k is fixed. Nevertheless, applying analytic tools to equation (17.5) do lend the advantage of uniformity in $k < (2-\varepsilon)\log_2 x$.

Theorem 1.4. *Let $r = k/\log_2 x$. For any $\varepsilon > 0$, as $x \to \infty$ we have uniformly for $r \le 2 - \varepsilon$,*

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} \frac{1}{n} \sim \nu(r) e^{r\gamma} \Gamma(r+1) \frac{(\log_2 x)^k}{k!}. \tag{17.9}$$

Hence by comparison with the Sathe–Selberg theorem, we obtain the following elegant relation between sums over $P^+(n) \le x$ with those over $n \le x$.

Corollary 1.1. *Let $r = k/\log_2 x$. For any $\varepsilon > 0$, as $x \to \infty$ we have uniformly for $r \le 2 - \varepsilon$,*

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} \frac{1}{n} \sim e^{r\gamma} \Gamma(r+1) \sum_{\substack{\Omega(n)=k \\ n\le x}} \frac{1}{n}. \tag{17.10}$$

Remark 4. One may prove an analogous result for $\omega(n)$, with the same factor $e^{r\gamma}\Gamma(r+1)$.

Note the factor $e^{r\gamma}\Gamma(r+1) = 1$ if and only if $r = 0$. Hence Corollary 1.1 implies

$$\sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} \frac{1}{n} \sim \sum_{\substack{\Omega(n)=k \\ n\le x}} \frac{1}{n} \tag{17.11}$$

if and only if k is in the uniform range $k = o(\log_2 x)$. This is an example of friable regularity, in the following sense. Recall an integer n with $P^+(n) \le x$ is called x-smooth or x-friable.

Definition 1. A sequence $(a_n)_{n\in\mathbb{N}}$ is *friably regular*[1] if $\sum_{n\le x} a_n \sim \sum_{P^+(n)\le x} a_n$ as $x \to \infty$.

For example, the friable regularity of $(\mu(n)/n)_{n\in\mathbb{N}}$ is equivalent to the prime number theorem. We also extend the definition to families of sequences.

Definition 2. A one-parameter family $(a_{n,x})_{n\in\mathbb{N}}$, indexed by $x \in \mathbb{R}$, is *friably regular* if $\sum_{n\le x} a_{n,x} \sim \sum_{P^+(n)\le x} a_{n,x}$ as $x \to \infty$.

In particular, Corollary 1.1 implies the family $(\mathbf{1}_{\Omega(n)=k}/n)_{n\in\mathbb{N}}$, indexed by $k = k(x)$, is friably regular if and only if $k = o(\log_2 x)$.

1.2 The coefficients c_k

Finally, we emphasize an important feature of the combinatorial approach in Theorem 1.2. The recursion in equation (17.6) enables rapid computation of the coefficients

1 This extends the notion of friable regularity as in [1],[3], from equality of limits of convergent series to asymptotic equality of (possibly nonconvergent) partial sums.

c_k to high precision, the first few displayed below.

k	c_k	k	c_k
0, 1	1, 0	6	0.0108213…
2	0.226123…	7	0.0054110…
3	0.058254…	8	0.0027375…
4	0.044814…	9	0.0013752…
5	0.020323…	10	0.0006903…

At first glance, one might not expect the coefficients c_k arising from equation (17.6) to exhibit any particular structure. However, the combinatorial approach shows c_k to satisfy exponentially precise asymptotics.

Theorem 1.5. *The coefficients satisfy* $c_k = \eta 2^{-k} + O(3^{-k})$. *Here, the constant* η *is given by* $\eta = e^{-1}\prod_{p>2}(1-\frac{2}{p})^{-1}e^{-2/p} = 0.71206\ldots$.

2 Elementary combinatorial proof for *k* fixed

In this section, we prove Theorem 1.2. For $x, s > 0$, define the (truncated) zeta functions

$$Z_k(s,x) = \sum_{\substack{\Omega(n)=k \\ P^+(n)\le x}} n^{-s}, \quad Z(s,x) = Z_1(s,x) = \sum_{p\le x} p^{-s}.$$

We first express $Z_k(s,x)$ in terms of $Z(s,x)$.

Proposition 2.1. *For each* $k \ge 1$ *and any* $x, s > 0$, *we have the identity*

$$Z_k(s,x) = \sum_{n_1+2n_2+\cdots=k} \prod_{j\ge1} \frac{1}{n_j!}(Z(js,x)/j)^{n_j} \tag{17.12}$$

where the sum ranges over all partitions of k.

Proof. For any $x, s > 0$, we have a formal power series identity in z,

$$\sum_{k\ge0} Z_k(s,x)z^k = \sum_{P^+(n)\le x} \frac{z^{\Omega(n)}}{n^s} = \prod_{p\le x}\left(1+\frac{z}{p^s}+\frac{z^2}{p^{2s}}+\cdots\right) = \prod_{p\le x}\left(1-\frac{z}{p^s}\right)^{-1}$$

since the function $n \mapsto z^{\Omega(n)}/n^s$ is completely multiplicative. Thus expanding Taylor series,

$$\begin{aligned}\sum_{k\ge0} Z_k(s,x)z^k &= \exp\Big(-\sum_{p\le x}\log(1-zp^{-s})\Big) = \exp\Big(\sum_{p\le x}\sum_{j\ge1}\frac{(zp^{-s})^j}{j}\Big)\\ &= \exp\Big(\sum_{j\ge1}\frac{Z(js,x)}{j}z^j\Big) = \prod_{j\ge1}\exp\Big(\frac{Z(js,x)}{j}z^j\Big)\end{aligned}$$

$$= \prod_{j\geq 1} \sum_{n_j\geq 0} \frac{1}{n_j!}\left(\frac{Z(js,x)}{j} z^j\right)^{n_j}$$
$$= \sum_{k\geq 0} z^k \sum_{n_1+2n_2+\cdots=k} \prod_{j\geq 1} \frac{1}{n_j!}(Z(js,x)/j)^{n_j}. \tag{17.13}$$

Now equation (17.12) follows by comparing the coefficients of z^k. □

Remark 5. This proposition generalizes [5, Proposition 3.1].

Next, the recursion for c_k in equation (17.6) leads to the explicit formula,

$$c_k = \sum_{2n_2+3n_3\ldots=k} \prod_{j\geq 2} \frac{1}{n_j!}(Z(j)/j)^{n_j}, \tag{17.14}$$

by the following lemma, for the choices $A_1 = 0$ and $A_j = Z(j)$ when $j \geq 2$.

Lemma 2.1. *Given any sequence $(A_k)_{k=1}^{\infty}$, the sequence $(b_k)_{k=0}^{\infty}$ is given recursively by $b_0 = 1$ and $b_k = \frac{1}{k}\sum_{j=1}^{k} b_{k-j}A_j$, if and only if $(b_k)_{k=0}^{\infty}$ is given explicitly as*

$$b_k = \sum_{n_1+2n_2+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!}.$$

Note that the (unique) partition of $k = 0$ has $n_j = 0$ for all $j \geq 1$, so indeed $b_0 = \prod_{j\geq 1}(A_j/j)^0/0! = 1$.

Proof. We prove the forward direction by induction on k (the reverse direction is similar). For $k = 1$, we have $b_1 = b_0A_1 = A_1$.

Then assuming the claim for each $r < k$,

$$kb_k = \sum_{r=1}^{k} b_{k-r}A_r = \sum_{r=1}^{k} A_r \sum_{n_1+\cdots=k-r} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!}$$
$$= \sum_{r=1}^{k} \sum_{n_1+\cdots=k-r} \frac{A_r^{n_r+1}}{r^{n_r}n_r!} \prod_{j\neq r} \frac{(A_j/j)^{n_j}}{n_j!} = \sum_{r=1}^{k} \sum_{\substack{n_1+\cdots=k \\ n_r\geq 1}} rn_r \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!}$$
$$= \sum_{n_1+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!} \sum_{\substack{1\leq r\leq k \\ n_r\geq 1}} rn_r = k \sum_{n_1+\cdots=k} \prod_{j\geq 1} \frac{(A_j/j)^{n_j}}{n_j!}$$

In the last step, we dropped the condition $n_r \geq 1$ (since $rn_r = 0$ for $n_r = 0$), which gives $\sum_{r=1}^{k} rn_r = k$. Dividing by k completes the induction. □

Now equipped with Proposition 2.1 and equation (17.14) for c_k, we now prove Theorem 1.2.

Proof of Theorem 1.2. For $Z(j,x)$ with $j \geq 2$, we trivially bound $\sum_{p>x} p^{-j}$ by $x^{2-j} \sum_{n>x} n^{-2} = O(x^{1-j})$, which gives

$$Z(j,x) := \sum_{p \leq x} p^{-j} = Z(j) - \sum_{p>x} p^{-j} = Z(j) + O(x^{1-j}) \quad \text{for } j \geq 2.$$

Thus plugging into the identity for $Z_k(1,x)$, Proposition 2.1 with $s = 1$ gives

$$Z_k(1,x) = \sum_{n_1+2n_2+\cdots=k} \frac{Z(1,x)^{n_1}}{n_1!} \prod_{j\geq 2} \frac{1}{n_j!} \left(\frac{Z(j) + O(x^{1-j})}{j} \right)^{n_j}.$$

For any partition of k, the binomial theorem implies $\prod_{j\geq 2} \frac{1}{n_j!}([Z(j) + O(x^{1-j})]/j)^{n_j}$ equals $\prod_{j\geq 2} \frac{1}{n_j!}(Z(j)/j)^{n_j}$ at negligible cost $O_k(1/x)$. Thus

$$Z_k(1,x) = \sum_{n_1+2n_2+\cdots=k} \frac{Z(1,x)^{n_1}}{n_1!} \prod_{j\geq 2} \frac{1}{n_j!} (Z(j)/j)^{n_j} + O_k\left(\frac{Z(1,x)^k}{x} \right). \tag{17.15}$$

Then for $Z(1,x)$, we recall Mertens' second theorem,

$$Z(1,x) := \sum_{p \leq x} \frac{1}{p} = \log_2 x + \beta + E(x), \quad \text{with } E(x) = O\left(\frac{1}{\log x} \right), \tag{17.16}$$

so plugging in above gives

$$\begin{aligned} Z_k(1,x) &= \sum_{n_1+2n_2+\cdots=k} \frac{1}{n_1!} (\log_2 x + \beta + E(x))^{n_1} \prod_{j\geq 2} \frac{1}{n_j!} (Z(j)/j)^{n_j} + O_k\left(\frac{(\log_2 x)^k}{x} \right) \\ &= \sum_{n_1=0}^{k} \frac{1}{n_1!} (\log_2 x + \beta)^{n_1} \sum_{2n_2+\cdots=k-n_1} \prod_{j\geq 2} \frac{1}{n_j!} (Z(j)/j)^{n_j} + O_k(E(x)(\log_2 x)^{k-1}) \end{aligned} \tag{17.17}$$

again by the binomial theorem. Here, we used $\prod_{j\geq 2} \frac{1}{n_j!}(Z(j)/j)^{n_j} = O_k(1)$.

Now recalling equation (17.14) and $E(x) = O(1/\log x)$ completes the proof of Theorem 1.2. □

From here, we may "dissect" Mertens' third theorem. Indeed by equation (17.4),

$$\prod_{p \leq x} \left(1 - \frac{1}{p}\right)^{-1} = \sum_{P^+(n) \leq x} \frac{1}{n} = \sum_{k \geq 0} Z_k(1,x)$$

and using the asymptotic formula for $Z_k(1,x)$ from Theorem 1.2,

$$\begin{aligned} \sum_{k\geq 0} \sum_{j=0}^{k} \frac{c_{k-j}}{j!} (\log_2 x + \beta)^j &= \sum_{j\geq 0} \frac{1}{j!} (\log_2 x + \beta)^j \sum_{k\geq j} c_{k-j} = e^{\beta} \log x \sum_{m\geq 0} c_m \\ &= e^{\gamma} \log x, \end{aligned} \tag{17.18}$$

as desired, provided $\sum_{m\geq 0} c_m = e^{\gamma-\beta}$. This follows in turn by equation (17.14),

$$\sum_{m\geq 0} c_m = \sum_{m\geq 0} \sum_{2n_2+3n_3\ldots=m} \prod_{j\geq 2} \frac{1}{n_j!}(Z(j)/j)^{n_j} = \prod_{j\geq 2} \sum_{n_j\geq 0} \frac{1}{n_j!}(Z(j)/j)^{n_j}$$
$$= \prod_{j\geq 2} \exp(Z(j)/j) = \exp\left(\sum_{j\geq 2} \frac{Z(j)}{j}\right) = e^{\gamma-\beta}$$

recalling equation (17.3). This shows the claim.

3 Combinatorial proof of asymptotics for coefficients c_k

In this section, we prove a strengthening of Theorem 1.5. To this, we first rephrase the recursion for c_k in equation (17.6).

Let $A_1 = 0$ and $A_k = \sum_p p^{-k}$ for $k \geq 2$. Then c_k is recursively defined by $c_0 = 1$ and

$$kc_k = \sum_{j=1}^{k} c_{k-j}A_j. \tag{17.19}$$

Consider the following induced sequences $A_{k,q}, c_{k,q}$ for each prime q: let $A_{k,2} = A_k$, $c_{k,2} = c_k$; and if p is the prime preceding $q > 2$, let

$$A_{k,q} = A_{k,p} - p^{-k} \quad \text{for } k \geq 1, \tag{17.20}$$
$$c_{k,q} = c_{k,p} - p^{-1}c_{k-1,p} \quad \text{for } k \geq 1, \quad \text{and} \quad c_{0,q} = c_0. \tag{17.21}$$

Explicitly, we have

$$A_{k,q} = \sum_{r\geq q} r^{-k} \quad \text{for } k \geq 2, \quad \text{and} \quad A_{1,q} = -\sum_{p<q} p^{-1}. \tag{17.22}$$

Lemma 3.1. *For each prime q and $k \geq 0$, we have the recursion*

$$kc_{k,q} = \sum_{j=1}^{k} c_{k-j,q}A_{j,q}. \tag{17.23}$$

Proof. We proceed by induction on the prime q. The base case $q = 2$ holds by equation (17.19).

Now assume equation (17.23) for $p < q$. The difference of recursions in equation (17.23) for $c_{k,p}$ and $p^{-1} \cdot c_{p,k-1}$ is

$$\begin{aligned}
&kc_{k,p} - (k-1)p^{-1}c_{k-1,p} \\
&= \sum_{j=1}^{k-1} (c_{k-j,p} - p^{-1}c_{k-j-1,p})A_{j,p} + c_{0,p}A_{k,p} \\
&= \sum_{j=1}^{k-1} (c_{k-j,p} - p^{-1}c_{k-j-1,p})(p^{-j} + A_{j,q}) + c_{0,p}(p^{-k} + A_{k,q}) \\
&= \sum_{j=1}^{k-1} (p^{-j}c_{k-j,p} - p^{-j-1}c_{k-j-1,p}) + c_{0,p}p^{-k} + \sum_{j=1}^{k-1} c_{k-j,q}A_{j,q} + c_{0,q}A_{k,q} \\
&= p^{-1}c_{k-1,p} + \sum_{j=1}^{k} c_{k-j,q}A_{j,q}
\end{aligned}$$

using equations (17.20), (17.21) and telescoping series. Subtracting $p^{-1}c_{k-1,p}$ gives

$$kc_{k,q} = k(c_{k,p} - p^{-1}c_{k-1,p}) = \sum_{j=1}^{k} c_{k-j,q}A_{j,q}. \qquad \square$$

Note that Lemmas 2.1 and 3.1 together imply

$$c_{k,q} = \sum_{n_1+2n_2\ldots=k} \prod_{j\geq 1} \frac{(A_{j,q}/j)^{n_j}}{n_j!} \tag{17.24}$$

for each prime q, $k \geq 1$.

Now with the recursion in hand, we bound the induced sequence $c_{k,q}$.

Lemma 3.2. *For each prime q, we have $c_{k,q} \ll_q q^{-k}$ as $k \to \infty$.*

Proof. Fix q and let $m_k = \max_{j\leq k} q^j|c_{j,q}|$. We shall prove $m_k \ll_q 1$, and it suffices to show this along a subsequence, since m_k is itself a nondecreasing sequence. Namely, we consider the indices k for which $m_k = q^k|c_{k,q}|$.

Recalling (17.22), we have for all $n \geq 1$,

$$\sum_{1\leq j\leq n} q^j A_{j,q} = qA_{1,q} + (n-1) + \sum_{2\leq j\leq n} \sum_{r>q} (q/r)^j = n + O_q(1),$$

by summing the geometric series, and so the recursion in equation (17.23) gives

$$\begin{aligned}
kq^k|c_{k,q}| = \left|\sum_{j=1}^{k} q^{k-j}c_{k-j,q} \cdot q^j A_{j,q}\right| &\leq m_{k/2}\left|\sum_{k/2<j\leq k} q^j A_{j,q}\right| + m_k\left|\sum_{1\leq j\leq k/2} q^j A_{j,q}\right| \\
&= m_{k/2}(k/2 + O(1)) + m_k(k/2 + O(1)).
\end{aligned}$$

And by our choice of k, we have $m_k = q^k|c_{k,q}|$ and so

$$m_k \le m_{k/2}(1 + O(1/k)). \tag{17.25}$$

Hence by induction on k, we conclude

$$m_k \ll m_1 \prod_{2^i \le k}(1 + O(2^{-i})) \ll \exp \sum_{2^i \le k} O(2^{-i}) \ll 1. \qquad \square$$

Since Lemma 3.2 holds for every prime q and the sequences $c_{k,q}$ are defined inductively on primes, Lemma 3.2 is self-improving. Indeed, for each pair of consecutive primes $p < q$,

$$c_{k,p} - p^{-1}c_{k-1,p} = c_{k,q} = O(q^{-k}).$$

In other words, multiplying above by p^k the modified sequence $c'_{k,p} := c_{k,p}p^k$ satisfies $c'_{k,p} - c'_{k-1,p} = O((p/q)^k)$, so $(c'_{k,p})_{k\ge1}$ is a Cauchy sequence for each prime p. Hence the limit

$$\eta_p := \lim_{k\to\infty} c'_{k,p} = \lim_{k\to\infty} c_{k,p}p^k$$

exists with $c'_{k,p} = \eta_p + O((p/q)^k)$. That is,

$$c_{k,p} = \eta_p p^{-k} + O_p(q^{-k}) \quad \text{for each prime } p. \tag{17.26}$$

To summarize, we expanded the definition of $c_{k,q}$ and used a zeroth-order expansion for each prime (Lemma 3.2) to prove a first-order expansion for every prime simultaneously.

Continuing in this way, we obtain a hth order expansion for $c_{k,q}$ by induction on the order $h \ge 1$, at each step proving the respective expansion for every prime simultaneously.

Proposition 3.1. *For any $h \ge 1$, we have*

$$c_{k,p_n} = \sum_{l=0}^{h-1} \eta^{(n)}_{p_{n+l}} p_{n+l}^{-k} + O_{n,h}(p_{n+h}^{-k}), \quad \text{where } \eta^{(n)}_{p_{n+l}} = \eta_{p_{n+l}} / \prod_{i=0}^{l-1}\left(1 - \frac{p_{n+l}}{p_{n+i}}\right) \tag{17.27}$$

for all n as $k \to \infty$. Here, p_n denotes the nth prime, and $\eta_p = \lim_{k\to\infty} c_{k,p}p^k$ as in Equation (17.26).

Proof. We proceed by induction on h. The base $h = 1$ holds for all n by equation (17.26), since $\eta^{(n)}_{p_n} = \eta_{p_n}$. Now assume equation (17.27) holds with h for every n, and write $c_{k,p_n} = \sum_{l=0}^{h} \eta^{(n)}_{p_{n+l}} p_{n+l}^{-k} + E_{k,n,h}$. By assumption $E_{k,n,h} \ll p_{n+h}^{-k}$, and we aim to show $E_{k,n,h} \ll p_{n+1+h}^{-k}$.

By equation (17.21) and the induction hypothesis (17.27) for $c_{k,p_{n+1}}$,

$$c_{k,p_{n+1}} = c_{k,p_n} - p_n^{-1} c_{k-1,p_n}$$
$$\sum_{l=0}^{h-1} \eta_{p_{n+1+l}}^{(n+1)} p_{n+1+l}^{-k} + O(p_{n+1+h}^{-k}) = \sum_{l=0}^{h} \eta_{p_{n+l}}^{(n)} \Big(1 - \frac{p_{n+l}}{p_n}\Big) p_{n+l}^{-k} + (E_{k,n,h} - p_n^{-1} E_{k-1,n,h})$$

Note that by definition $\eta_{p_{n+l}}^{(n+1)} = \eta_{p_{n+l}}^{(n)}(1 - \frac{p_{n+l}}{p_n})$, and so the above simplifies as

$$O(p_{n+1+h}^{-k}) = E_{k,n,h} - p_n^{-1} E_{k-1,n,h}.$$

Thus, similarly as with equation (17.26), the modified sequence $E'_{k,n,h} := E_{k,n,h} p_n^k$ converges as $k \to \infty$ to some limit $\ell_{n,h}$, with $E'_{k,n,h} = \ell_{n,h} + O((p_n/p_{n+1+h})^k)$. That is,

$$E_{k,n,h} = \ell_{n,h} p_n^{-k} + O(p_{n+1+h}^{-k})$$

On the other hand, $E_{k,n,h} \ll p_{n+h}^{-k}$ forces $\ell_{n,h} = 0$. Hence $E_{k,n,h} = O(p_{n+1+h}^{-k})$ as desired. □

Next, we determine the expansion coefficients η_p from equation (17.26).

Proposition 3.2. *For any prime p, the coefficient $\eta_p = \lim_{k\to\infty} c_{k,p} p^k$ equals*

$$\eta_p = e^{-\sum_{q\le p} p/q} \prod_{q>p} \Big(1 - \frac{p}{q}\Big)^{-1} e^{-p/q}. \tag{17.28}$$

Proof. Consider the generating function $C_p(z) = \sum_{k\ge 0} c_{k,p} z^k$. On one hand, the explicit formula for $c_{k,p}$ in equation (17.24) implies

$$C_p(z) = \sum_{k\ge 0} c_{k,p} z^k = \sum_{k\ge 0} z^k \sum_{n_1+2n_2\ldots=k} \prod_{j\ge 1} \frac{(A_{j,p}/j)^{n_j}}{n_j!} = \prod_{j\ge 1} \sum_{n_j\ge 0} \frac{1}{n_j!} (A_{j,p} z^j/j)^{n_j}$$
$$= \prod_{j\ge 1} \exp(A_{j,p} z^j/j) = \exp\Big(\sum_{j\ge 1} A_{j,p} z^j/j\Big).$$

Then recalling $A_{j,p} = \sum_{q\ge p} q^{-j}$ for $j \ge 2$,

$$C_p(z) = \exp\Big(zA_{1,p} + \sum_{q\ge p} \sum_{j\ge 2} (z/q)^j/j\Big) = e^{zA_{1,p}} \exp\Big(-\sum_{q\ge p} [\log(1 - z/q) + z/q]\Big)$$
$$= e^{zA_{1,p}} \prod_{q\ge p} (1 - z/q)^{-1} e^{-z/q}. \tag{17.29}$$

On the other hand, by the expansion for c_k in equation (17.26) we have

$$C_p(z) = \sum_{k\geq 0} c_k z^k = \eta_p \sum_{k\geq 0} (z/p)^k + O_p\bigg(\sum_{k\geq 0}(z/q)^k\bigg) = \frac{\eta_p}{1-z/p} + \frac{O_p(1)}{1-z/q}, \tag{17.30}$$

since $A_{1,p} = -\sum_{q<p} q^{-1}$. So comparing $C_p(z)$ from equations (17.29) and (17.30) at the pole $z = p$,

$$\eta_p = \lim_{z\to p} C_p(z)(1-z/p) = e^{pA_{1,p}-1}\prod_{q>p}(1-p/q)^{-1}e^{-p/q}.$$

Hence the result follows since $A_{1,p} = -\sum_{q<p} q^{-1}$. □

Finally, we obtain an expansion for the original sequence $c_{k,p_1} = c_k$ to arbitrary order, which gives a considerable refinement of Theorem 1.5.

Theorem 3.1. *For each prime q,*

$$c_k = \sum_{p<q} \alpha_p p^{-k} + O_q(q^{-k})$$

where $\alpha_p := e^{-1}\prod_{q\neq p}(1-\frac{p}{q})^{-1}e^{-p/q}$. In particular $c_k = \alpha_2 2^{-k} + O(3^{-k})$.

Proof. Setting $n = 1$ in Proposition 3.1, the sequence $c_{k,p_1} = c_k$ satisfies

$$c_k = \sum_{p<q} \eta_p^{(1)} p^{-k} + O_q(q^{-k})$$

where Proposition 3.2 gives, by definition of $\eta_p^{(1)}$ in equation (17.27),

$$\eta_p^{(1)} := \eta_p / \prod_{q<p}\bigg(1-\frac{p}{q}\bigg) = e^{-1}\prod_{q\neq p}\bigg(1-\frac{p}{q}\bigg)^{-1} e^{-p/q} = \alpha_p.$$ □

4 Analytic proof for *k* in uniform range

We prove Theorem 1.4, which quantitative error, which we state below.

Theorem 4.1. *Let $r = k/\log_2 x$ and define $\eta(z) = e^{\gamma z}\prod_p (1-\frac{1}{p})^z(1-\frac{z}{p})^{-1}$. For any $\varepsilon > 0$, as $x\to\infty$ we have uniformly for $r \leq 2-\varepsilon$,*

$$\sum_{\substack{\Omega(n)=k\\ P^+(n)\leq x}} \frac{1}{n} = \eta(r)\frac{(\log_2 x)^k}{k!}\bigg(1 + O_\varepsilon\bigg(\frac{k}{(\log_2 x)^2}\bigg)\bigg).$$

Proof. By Cauchy's residue formula, we have for any $r < 2$,

$$Z_k(1,x) = \frac{1}{2\pi i}\int\limits_{|z|=r} f_x(z)\frac{dz}{z^{k+1}}, \tag{17.31}$$

where f_x is given by the power series

$$\begin{aligned} f_x(z) &= \sum_{k\geq 0} Z_k(1,x)z^k = \sum_{P^+(n)\leq x}\frac{z^{\Omega(n)}}{n} \\ &= \prod_{p\leq x}\left(1+\frac{z}{p}+\frac{z^2}{p^2}+\cdots\right) = \prod_{p\leq x}\left(1-\frac{z}{p}\right)^{-1} \\ &= \big(1+O(E(x))\big)e^{z\gamma}(\log x)^z\prod_{p\leq x}\left(1-\frac{z}{p}\right)^{-1}\left(1-\frac{1}{p}\right)^{z} \\ &= \big(1+O(E(x))\big)\eta(z)(\log x)^z, \end{aligned}$$

as $\prod_{p\leq x}(1-\frac{1}{p})^{-1} = (1+E(x))e^{\gamma}\log x$ by Merten's second theorem in quantitative form (this also follows from the prime number theorem.)

Hence equation (17.31) becomes

$$Z_k(1,x) = \frac{1+O(E(x))}{2\pi i}\int\limits_{|z|=r}\eta(z)(\log x)^z\frac{dz}{z^{k+1}}. \tag{17.32}$$

The desired main term in Theorem 1.4 is given by evaluating $\eta(z)$ at $z = r$, namely

$$\frac{\eta(r)}{2\pi i}\int\limits_{|z|=r}(\log x)^z\frac{dz}{z^{k+1}} = \eta(r)\frac{(\log_2 x)^k}{k!}. \tag{17.33}$$

For the error, we follow the argument in [7, p. 233], which we provide for completeness. Recall $E(x) \ll 1/\log x$. For $|z| = r = k/\log_2 x$, integration by parts gives

$$\frac{1}{2\pi i}\int\limits_{|z|=r}(z-r)(\log x)^z\frac{dz}{z^{k+1}} = \frac{(\log_2 x)^{k-1}}{(k-1)!} - \frac{r(\log_2 x)^k}{k!} = 0, \quad \text{and} \tag{17.34}$$

$$\eta(z) - \eta(r) - \eta'(r)(z-r) = \int_r^z (z-w)\eta''(w)dw \ll |z-r|^2. \tag{17.35}$$

Thus subtracting equations (17.33) from (17.32), the error is

$$\begin{aligned} \eta(r)\frac{(\log_2 x)^k}{k!} - Z_k(1,x) &\ll \int\limits_{|z|=r}[\eta(r)-\eta(z)](\log x)^z\frac{dz}{z^{k+1}} \\ &\overset{(17.34)}{=} \int\limits_{|z|=r}[\eta(r)-\eta(z)-\eta'(r)(z-r)](\log x)^z\frac{dz}{z^{k+1}} \end{aligned}$$

$$\begin{aligned}
&\overset{(17.35)}{\ll} \int_{|z|=r} |z-r|^2 (\log x)^z \frac{dz}{z^{k+1}} \\
&\ll r^{2-k} \int_{-1/2}^{1/2} (\sin \pi\theta)^2 e^{k\cos(2\pi\theta)}\, d\theta \\
&\ll r^{2-k} e^k \int_0^{\infty} \theta^2 e^{-8k\theta^2} d\theta \ll r^{2-k} e^k k^{-3/2} \\
&= (\log_2 x)^{k-2} (e/k)^k k^{1/2} \ll k(\log_2 x)^{k-2}/k!
\end{aligned}$$

Here, we used $|\sin x| \le x$, $\cos(2\pi\theta) \le 1 - 8\theta^2$ for $|\theta| \le 1/2$, and Stirling's formula. □

Bibliography

[1] R. J. Duffin, Representation of Fourier integrals as sums, III, *Proc. Am. Math. Soc.*, **8** (1957), 272–277.

[2] P. Erdős and A. Sárközy, On the number of prime factors of integers, *Acta Sci. Math.*, **42** (1980), 237–246.

[3] E. Fouvry and G. Tenenbaum, Entiers sans grand facteur premierè en progressions arithmétiques, *Proc. Lond. Math. Soc.*, **63** (1991), 449–494.

[4] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, Erster Band. B. G. Teubner, Leipzig, Berlin, 1909.

[5] J. D. Lichtman, Almost primes and the Banks–Martin conjecture, *J. Number Theory*, **211** (2020), 513–529.

[6] F. Mertens, Ein Beitrag zur analytischen Zahlentheorie, *J. Reine Angew. Math.*, **78** (1874), 46–62.

[7] H. L. Montgomery and R. C. Vaughan, *Multiplicative Number Theory I: Classical Theory*, Cambridge University Press, Cambridge, 2006.

[8] P. Pollack, A generalization of the Hardy-Ramanujan inequality and applications, *J. Number Theory*, **210** (2020), 171–182.

[9] D. Popa, A double Mertens type evaluation, *J. Math. Anal. Appl.*, **409** (2014), 1159–1163.

[10] D. Popa, A triple Mertens evaluation, *J. Math. Anal. Appl.*, **444** (2016), 464–474.

[11] J. Spencer and R. Graham, The elementary proof of the prime number theorem, *Math. Intell.*, **31** (2009), 18–23.

[12] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Graduate Studies in Mathematics, vol. **163**, Amer. Math. Soc., 2015.

[13] G. Tenenbaum, Generalized Mertens sums, in G. Andrews and F. Garvan (eds.) Analytic Number Theory, Modular Forms and q-hypergeometric Series – in honor of Krishna Alladi's 60th birthday, Springer Proc. Math. Stat., vol. **221**, pp. 477–495, Springer, Cham, 2017.

Melvyn B. Nathanson

Curious convergent series of integers with missing digits

To Ron Graham

Abstract: A classical theorem of Kempner states that the sum of the reciprocals of positive integers with missing decimal digits converges. This result is extended to much larger families of "missing digits" sets of positive integers with both convergent and divergent harmonic series.

1 Kempner's theorem

> "It is well known that the series
>
> $$\sum_{n=1}^{\infty} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots$$
>
> diverges. The object of this Note is to prove that if the denominators do not include all natural numbers $1, 2, 3, \ldots$, but only those which do not contain any figure 9, the series converges. The method of proof holds unchanged if, instead of 9, any other figure $1, 2, \ldots, 8$ is excluded, but not for the figure 0."
>
> A. J. Kempner, *Amer. Math. Monthly* **21** (1914), 48–50.

A *harmonic series* is a series of the form $\sum_{a \in A} 1/a$, where A is a set of positive integers. Mathematicians have long been interested in the convergence or divergence of harmonic series. Let $c \in \{1, 2, \ldots, 9\}$, and let $A_{10}(c)$ be the set of positive integers in which the digit c does not occur in the usual decimal representation. Kempner [6] proved in 1914 that $\sum_{a \in A_{10}(c)} 1/a$ converges. He called this "a curious convergent series." More generally, for every integer $g \geq 2$, every positive integer n has a unique *g-adic representation* of the form $n = \sum_{i=0}^{k} c_i g^i$, with digits $c_i \in \{0, 1, 2, \ldots, g-1\}$ for $i = 0, 1, \ldots, k$ and $c_k \neq 0$. If $A_g(c)$ is the set of integers whose g-adic representation contains no digit c, then the infinite series $\sum_{a \in A_g(c)} 1/a$ converges. This includes the case $c = 0$, which was not discussed by Kempner.

Kempner's theorem has been studied and extended by Baillie [1], Farhi [2], Gordon [3], Irwin [5], Lubeck-Ponomarenko [7], Schmelzer and Baillie [10], and Wadhwa [11, 12]. It is Theorem 144 in Hardy and Wright [4].

The g-adic representation is a special case of a more general method to represent the positive integers. A *$\mathcal{G}$-adic sequence* is a strictly increasing sequence of positive integers $\mathcal{G} = (g_i)_{i=0}^{\infty}$ such that $g_0 = 1$ and g_i divides g_{i+1} for all $i \geq 0$. The integer

Melvyn B. Nathanson, Lehman College (CUNY), Bronx, NY, USA, e-mail: melvyn.nathanson@lehman.cuny.edu

https://doi.org/10.1515/9783110754216-018

quotients

$$d_i = \frac{g_{i+1}}{g_i}$$

satisfy $d_i \geq 2$ and

$$g_{k+1} = g_k d_k = d_0 d_1 d_2 \cdots d_k \tag{18.1}$$

for all $k \geq 0$. Every positive integer n has a unique representation in the form

$$n = \sum_{i=0}^{k} c_i g_i \tag{18.2}$$

where $c_i \in \{0, 1, \ldots, d_i - 1\}$ for all $i \in \{0, 1, \ldots, k\}$ and $c_k \neq 0$. We call (18.2) the *$\mathcal{G}$-adic representation* of n. This is equivalent to de Bruijn's additive system (Nathanson [8, 9]).

Harmonic series constructed from sets of positive integers with missing $\mathcal{G}$-adic digits do not necessarily converge. In Theorem 1, we construct sets of integers with missing $\mathcal{G}$-adic digits whose harmonic series converge, and also sets of integers with missing $\mathcal{G}$-adic digits whose harmonic series diverge.

2 $\mathcal{G}$-adic representations with bounded quotients

Define the *interval of integers*

$$[a, b] = \{n \in \mathbf{Z} : a \leq n \leq b\}.$$

Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with quotients $d_i = g_{i+1}/g_i$. Let I be a set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. For every nonnegative integer k, let A_k be the set of integers $n \in [g_k, g_{k+1} - 1]$ whose $\mathcal{G}$-adic representation $n = \sum_{i=0}^{k} c_i g_i$ satisfies the following *missing digits condition*:

$$c_i \in [0, d_i - 1] \setminus U_i \quad \text{for all } i \in I \cap [0, k]. \tag{18.3}$$

Lemma 1. *The set A_k satisfies:*
(a) *$A_k = \emptyset$ if and only if $k \in I$ and $U_k = [1, d_k - 1]$.*
(b) *If $A_k \neq \emptyset$, then*

$$|A_k| \leq \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \leq 2|A_k|. \tag{18.4}$$

Proof. We use the inequality

$$x \leq 2(x - 1) \quad \text{for } x \geq 2. \tag{18.5}$$

If $n \in [g_k, g_{k+1} - 1]$, then n has the $\mathcal{G}$-adic representation

$$n = \sum_{i=0}^{k-1} c_i g_i + c_k g_k$$

with $c_k \neq 0$ and so $c_k \in [1, d_k - 1]$. It follows that $A_k = \emptyset$ if and only if $k \in I$ and $U_k = [1, d_k - 1]$.

For $A_k \neq \emptyset$, there are three cases.

(i) If $k \in I$ and $0 \in U_k$, then

$$|A_k| = \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i < 2|A_k|.$$

(ii) If $k \in I$ and $0 \notin U_k$, then U_k is a proper subset of $[1, d_k - 1]$ and so $|U_k| \leq d_k - 2$. Inequality (18.5) gives

$$d_k - |U_k| \leq 2(d_k - |U_k| - 1).$$

We obtain

$$\begin{aligned} |A_k| &= (d_k - |U_k| - 1) \prod_{\substack{i=0 \\ i \in I}}^{k-1} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &< \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &\leq 2(d_k - |U_k| - 1) \prod_{\substack{i=0 \\ i \in I}}^{k-1} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &= 2|A_k|. \end{aligned}$$

(iii) We have $d_k \geq 2$ and so $d_k \leq 2(d_k - 1)$ from inequality (18.5). If $k \notin I$, then

$$\begin{aligned} |A_k| &= (d_k - 1) \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k-1} d_i < \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i \\ &= d_k \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k-1} d_i \leq 2(d_k - 1) \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k-1} d_i \\ &= 2|A_k|. \end{aligned}$$

This completes the proof. □

Let A be a set of nonnegative integers, and let $A(n)$ be the number of elements $a \in A$ with $a \leq n$. The *upper asymptotic density* of the set A is $d_U(A) = \limsup_{n\to\infty} A(n)/n$. If $\lim_{n\to\infty} A(n)/n$ exists, then $d(A) = \lim_{n\to\infty} A(n)/n$ is called the *asymptotic density* of the set A. The set A has *asymptotic density zero* if $d(A) = d_U(A) = 0$.

Lemma 2. *Let A be a set of positive integers. If $\sum_{a\in A} 1/a < \infty$, then $d(A) = 0$.*

Proof. We show that $d_U(A) > 0$ implies $\sum_{a\in A} 1/a = \infty$.

If $d_U(A) = \limsup_{n\to\infty} A(n)/n = \alpha > 0$, then, for every $\varepsilon > 0$, we have

$$\frac{A(n)}{n} < \alpha + \varepsilon \quad \text{for all integers } n \geq N(\varepsilon)$$

and

$$\frac{A(n_i)}{n_i} > \alpha - \varepsilon \quad \text{for infinitely many integers } n_i. \tag{18.6}$$

Let $\varepsilon < \alpha/3$. There is a sequence of positive integers $(n_i)_{i=0}^{\infty}$ satisfying inequality (18.6) such that $n_0 \geq N(\varepsilon)$ and $n_i > 2n_{i-1}$ for all $i \geq 1$. We have

$$\begin{aligned} A(n_i) - A(n_{i-1}) &> (\alpha - \varepsilon)n_i - (\alpha + \varepsilon)n_{i-1} \\ &> (\alpha - \varepsilon)n_i - \frac{(\alpha + \varepsilon)n_i}{2} \\ &= n_i\left(\frac{\alpha - 3\varepsilon}{2}\right) \end{aligned}$$

and so

$$\sum_{\substack{a\in A \\ n_{i-1}<a\leq n_i}} \frac{1}{a} \geq \frac{A(n_i) - A(n_{i-1})}{n_i} > \frac{\alpha - 3\varepsilon}{2} > 0.$$

It follows that

$$\sum_{\substack{a\in A \\ 1\leq a\leq n_k}} \frac{1}{a} \geq \sum_{i=1}^{k} \sum_{\substack{a\in A \\ n_{i-1}<a\leq n_i}} \frac{1}{a} > k\left(\frac{\alpha - 3\varepsilon}{2}\right)$$

and the infinite series $\sum_{a\in A} 1/a$ diverges. Equivalently, convergence of the infinite series $\sum_{a\in A} 1/a$ implies $d(A) = 0$. This completes the proof. □

The converse of Lemma 2 is false. The set of prime numbers has asymptotic density zero, but the sum of the reciprocals of the primes diverges.

Theorem 1. *Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with bounded quotients, that is,*

$$d_i = \frac{g_{i+1}}{g_i} \leq d \tag{18.7}$$

for some integer $d \geq 2$ and all $i = 0, 1, 2, \ldots$. Let I be a set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$.

Let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let A be the set of positive integers n that satisfy the missing digits condition (18.3)*. If*

$$I(k) \geq \frac{(1+\delta)\log k}{\log(d/(d-1))} \tag{18.8}$$

for some $\delta > 0$ and all $k \geq k_0 = k_0(\delta)$, then the set A has asymptotic density zero and the harmonic series $\sum_{a \in A} 1/a$ converges.

If

$$I(k) \leq \frac{(1-\delta)\log k}{\log d} \tag{18.9}$$

for some $\delta > 0$ and all $k \geq k_1 = k_1(\delta)$, then the harmonic series $\sum_{a \in A} 1/a$ diverges.

Kempner's theorem is the special case $g_i = 10^i$, $d_i = 10$, and $U_i = \{9\}$ for all $i \in I = \mathbf{N}_0$.

Proof. For all $k \in \mathbf{N}_0$, the finite sets

$$A_k = A \cap [g_k, g_{k+1} - 1]$$

are pairwise disjoint and $A = \bigcup_{k=0}^{\infty} A_k$.

For all $i \in I$, we have

$$1 \leq |U_i| \leq d_i - 1$$

and

$$\frac{1}{d} \leq \frac{1}{d_i} \leq \frac{d_i - |U_i|}{d_i} = 1 - \frac{|U_i|}{d_i} \leq 1 - \frac{1}{d} < 1.$$

Let $I(k)$ satisfy inequality (18.8). We obtain

$$\left(1 - \frac{1}{d}\right)^{I(k)} \leq \left(\frac{d-1}{d}\right)^{\frac{(1+\delta)\log k}{\log(d/(d-1))}} = \frac{1}{k^{1+\delta}}. \tag{18.10}$$

If $a \in A_k$, then $a \geq g_k = d_0 d_1 \cdots d_{k-1}$. By Lemma 1,

$$\sum_{\substack{a \in A \\ a \geq g_{k_0}}} \frac{1}{a} = \sum_{k=k_0}^{\infty} \sum_{a \in A_k} \frac{1}{a} \leq \sum_{k=k_0}^{\infty} \frac{|A_k|}{g_k} \leq \sum_{k=k_0}^{\infty} \frac{d_k}{\prod_{i=0}^{k} d_i} \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i$$

$$\leq d \sum_{k=k_0}^{\infty} \prod_{\substack{i=0 \\ i \in I}}^{k} \frac{d_i - |U_i|}{d_i}$$

$$\leq d \sum_{k=k_0}^{\infty} \left(1 - \frac{1}{d}\right)^{I(k)}$$
$$\leq d \sum_{k=k_0}^{\infty} \frac{1}{k^{1+\delta}} < \infty.$$

Thus, the harmonic series converges. By Lemma 2, the set A has asymptotic density zero.

Let $I(k)$ satisfy inequality (18.9). We obtain

$$\left(\frac{1}{d}\right)^{I(k)} \geq \left(\frac{1}{d}\right)^{\frac{(1-\delta)\log k}{\log d}} = \frac{1}{k^{1-\delta}}. \tag{18.11}$$

If $a \in A_k$, then $a < g_{k+1} = d_0 d_1 \cdots d_{k-1} d_k$. By Lemma 1,

$$\sum_{\substack{a \in A \\ a \geq g_{k_1}}} \frac{1}{a} = \sum_{k=k_1}^{\infty} \sum_{a \in A_k} \frac{1}{a} \geq \sum_{k=k_1}^{\infty} \frac{|A_k|}{g_{k+1}}$$
$$\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \frac{1}{\prod_{i=0}^{k} d_i} \prod_{\substack{i=0 \\ i \in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0 \\ i \notin I}}^{k} d_i$$
$$= \frac{1}{2} \sum_{k=k_1}^{\infty} \prod_{\substack{i=0 \\ i \in I}}^{k} \frac{d_i - |U_i|}{d_i} \geq \frac{1}{2} \sum_{k=k_1}^{\infty} \prod_{\substack{i=0 \\ i \in I}}^{k} \frac{1}{d_i}$$
$$\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \left(\frac{1}{d}\right)^{I(k)}$$
$$\geq \frac{1}{2} \sum_{k=k_1}^{\infty} \frac{1}{k^{1-\delta}}$$

and the harmonic series diverges. This completes the proof. □

Corollary 1. *Let I be a set of nonnegative integers, and let $(v_i)_{i \in I}$ be a sequence of 0s and 1s. Let A be the set of integers n such that, if $n \in [2^k, 2^{k+1}-1]$ has the 2-adic representation $n = \sum_{i=0}^{k} c_i 2^i$, then $c_i = v_i$ for all $i \in I \cap [0, k]$. If*

$$I(k) \geq (1 + \delta) \log_2 k$$

for some $\delta > 0$ and all $k \geq k_0(\delta)$, then the harmonic series $\sum_{a \in A} 1/a$ converges. If

$$I(k) \leq (1 - \delta) \log_2 k$$

for some $\delta > 0$ and all $k \geq k_1(\delta)$, then the harmonic series $\sum_{a \in A} 1/a$ diverges.

Proof. For all $i \in I$, let $u_i = 1 - v_i$ and $U_i = \{u_i\}$. Apply Theorem 1. □

It is an open problem to determine the convergence or divergence of $\sum_{a\in A} 1/a$ if $I(k) \sim \log_2 k$.

3 $\mathcal{G}$-adic representations with unbounded quotients

Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence with quotients $d_i = g_{i+1}/g_i$. Let I be an infinite set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. If the sequence $\mathcal{G} = (g_i)_{i=0}^{\infty}$ has bounded quotients $d_i \le d$, then

$$\frac{|U_i|}{d_i} \ge \frac{1}{d}$$

for all $i \in I$ and the infinite series $\sum_{i\in I} \frac{|U_i|}{d_i}$ diverges. Equivalently, the convergence of this series implies that $\mathcal{G}$ has unbounded quotients.

Let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let A be the set of positive integers whose $\mathcal{G}$-adic representations satisfy the missing digits condition (18.3). The missing digits set A is finite if and only if I is a cofinite set of nonnegative integers and $U_i = [1, d_i - 1]$ for all sufficiently large i. The harmonic series of a finite set of positive integers converges.

Theorem 1 shows that harmonic series constructed from infinite sets of integers with missing $\mathcal{G}$-adic digits do not always converge. It follows from Theorem 1 that if

$$I(k) \ge (\log k)^{1+\delta}$$

for some $\delta > 0$ and all $k \ge k_0(\delta)$, and if $\sum_{a\in A} 1/a$ diverges, then the sequence $\mathcal{G}$ must have *unbounded quotients*, that is,

$$\limsup d_i = \infty.$$

Theorem 2 gives a sufficient condition for the divergence of harmonic series of sets of positive integers constructed from $\mathcal{G}$-adic sequences with unbounded quotients. We use the following inequality, which is easily proved by induction: If $0 \le x_i < 1$ for $i = 1, \dots, n$, then

$$\prod_{i=1}^{n}(1 - x_i) \ge 1 - \sum_{i=1}^{n} x_i. \tag{18.12}$$

Theorem 2. *Let $\mathcal{G} = (g_i)_{i=0}^{\infty}$ be a $\mathcal{G}$-adic sequence, and let $n = \sum_{i=0}^{k} c_i g_i$ be the $\mathcal{G}$-adic representation of the positive integer n. Let I be an infinite set of nonnegative integers, and, for all $i \in I$, let U_i be a nonempty proper subset of $[0, d_i - 1]$. Let A be the set of*

positive integers whose $\mathcal{G}$-adic representations satisfy the missing digits condition (18.3). *If the set A is infinite and if*

$$\sum_{i\in I} \frac{|U_i|}{d_i} < \infty \tag{18.13}$$

then the sequence $\mathcal{G} = (g_i)_{i=0}^{\infty}$ has unbounded quotients and the harmonic series $\sum_{a\in A} 1/a$ diverges.

For example, the "missing digits" set constructed from $\mathcal{G} = (g_i)_{i=0}^{\infty}$ with $g_i = 2^{i(i+1)/2}$ and $d_i = 2^{i+1}$ and with $I = \mathbf{N}_0$ and $U_i = \{0\}$ for all $i \in I$ has a divergent harmonic series.

Proof. Because the infinite series (18.13) converges, there is an integer $i_0 \in I$ such that

$$\sum_{\substack{i=i_0\\ i\in I}}^{\infty} \frac{|U_i|}{d_i} < \frac{1}{2}.$$

Inequality (18.12) implies that, for all $k \in \mathbf{N}_0$,

$$\prod_{\substack{i=i_0\\ i\in I}}^{k} \left(1 - \frac{|U_i|}{d_i}\right) \geq 1 - \sum_{\substack{i=i_0\\ i\in I}}^{k} \frac{|U_i|}{d_i} > \frac{1}{2}$$

and so

$$\begin{aligned}\prod_{\substack{i=0\\ i\in I}}^{k}\left(1 - \frac{|U_i|}{d_i}\right) &= \prod_{\substack{i=0\\ i\in I}}^{i_0-1}\left(1 - \frac{|U_i|}{d_i}\right)\prod_{\substack{i=i_0\\ i\in I}}^{k}\left(1 - \frac{|U_i|}{d_i}\right)\\ &> \frac{1}{2}\prod_{\substack{i=0\\ i\in I}}^{i_0-1}\left(1 - \frac{|U_i|}{d_i}\right) = \delta > 0.\end{aligned}$$

Let $A_k = A \cap [g_k, g_{k+1} - 1]$. The set A is infinite if and only if $A_k \neq \emptyset$ for infinitely many k. Applying inequality (18.4) of Lemma 1, we obtain

$$\begin{aligned}\sum_{a\in A} \frac{1}{a} &= \sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty} \sum_{a\in A_k} \frac{1}{a} \geq \sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty} \frac{|A_k|}{g_{k+1}}\\ &\geq \frac{1}{2}\sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty} \frac{1}{\prod_{i=0}^{k} d_i} \prod_{\substack{i=0\\ i\in I}}^{k} (d_i - |U_i|) \prod_{\substack{i=0\\ i\notin I}}^{k} d_i\\ &= \frac{1}{2}\sum_{\substack{k=0\\ A_k\neq\emptyset}}^{\infty} \prod_{\substack{i=0\\ i\in I}}^{k}\left(1 - \frac{|U_i|}{d_i}\right)\end{aligned}$$

and so the harmonic series $\sum_{a\in A} \frac{1}{a}$ diverges. This completes the proof. □

Bibliography

[1] R. Baillie, Sums of reciprocals of integers missing a given digit, *Am. Math. Mon.*, **86**(5) (1979), 372–374.
[2] B. Farhi, A curious result related to Kempner's series, *Am. Math. Mon.*, **115**(10) (2008), 933–938.
[3] R. A. Gordon, Comments on "Subsums of the harmonic series", *Am. Math. Mon.*, **126**(3) (2019), 275–279.
[4] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 6th ed., Oxford University Press, Oxford, 2008.
[5] F. Irwin, A curious convergent series, *Am. Math. Mon.*, **23**(5) (1916), 149–152.
[6] A. J. Kempner, A curious convergent series, *Am. Math. Mon.*, **21**(2) (1914), 48–50.
[7] B. Lubeck and V. Ponomarenko, Subsums of the harmonic series, *Am. Math. Mon.*, **125**(4) (2018), 351–355.
[8] M. B. Nathanson, Additive systems and a theorem of de Bruijn, *Am. Math. Mon.*, **121**(1) (2014), 5–17.
[9] M. B. Nathanson, Limits and decomposition of de Bruijn's additive systems, in *Combinatorial and Additive Number Theory II (New York)*, Springer Proc. Math. Stat., vol. **220**, pp. 255–267, Springer, 2017.
[10] T. Schmelzer and R. Baillie, Summing a curious, slowly convergent series, *Am. Math. Mon.*, **115**(6) (2008), 525–540.
[11] A. D. Wadhwa, An interesting subseries of the harmonic series, *Am. Math. Mon.*, **82**(9) (1975), 931–933.
[12] A. D. Wadhwa, Some convergent subseries of the harmonic series, *Am. Math. Mon.*, **85**(8) (1978), 661–663.

Carl Pomerance

A note on Carmichael numbers in residue classes

In memory of Ron Graham (1935–2020) and Richard Guy (1916–2020)

Abstract: Improving on some recent results of Matomäki and of Wright, we show that the number of Carmichael numbers to X in a coprime residue class exceeds $X^{1/(6\log\log\log X)}$ for all sufficiently large X depending on the modulus of the residue class.

1 Introduction

The "little theorem" of Fermat asserts that when p is a prime number, we have $b^p \equiv b \pmod p$ for all integers b. Given two integers b, p with $p > b > 0$, it is computationally easy to check this congruence, taking $O(\log p)$ arithmetic operations in $\mathbb{Z}/p\mathbb{Z}$. So, if the congruence is checked and we find that $b^p \not\equiv b \pmod p$ we immediately deduce that p is composite. Unfortunately, there are easily found examples where n is composite and the Fermat congruence holds for a particular b. For example, it always holds when $b = 1$. It holds when $b = 2$ and $n = 341$, and another example is $b = 3$, $n = 91$. In fact, there are composite numbers n where $b^n \equiv b \pmod n$ holds for all b, the least example being $n = 561$. These are the *Carmichael numbers*, named after R. D. Carmichael who published the first few examples in 1910; see [4]. (Interestingly, Šimerka published the first few examples 25 years earlier; see [8].)

We now know that there are infinitely many Carmichael numbers (see [1]) the number of them at most X exceeding X^c for a fixed $c > 0$ and X sufficiently large.

A natural question is if a given residue class contains infinitely many Carmichael numbers. After the work of Matomäki [7] and Wright [9], we now know there are infinitely many in a coprime residue class. More precisely, we have the following two theorems. Let

$$C_{a,M}(X) = \#\{n \le X : n \text{ is a Carmichael number, } n \equiv a \pmod M\}.$$

Theorem M (Matomäki). *Suppose that a, M are positive coprime integers and that a is a quadratic residue* mod M. *Then $C_{a,M}(X) \ge X^{1/5}$ for X sufficiently large depending on the choice of M.*

Carl Pomerance, Mathematics Department, Dartmouth College, Hanover, NH, USA, e-mail: carlp@math.dartmouth.edu

https://doi.org/10.1515/9783110754216-019

Theorem W (Wright). *Suppose that a, M are positive coprime integers. There are positive numbers K_M, X_M depending on the choice of M such that $C_{a,M}(X) \ge X^{K_M/(\log\log\log X)^2}$ for all $X \ge X_M$.*

Thus Wright was able to remove the quadratic residue condition in Matomäki's theorem but at the cost of lowering the count to an expression that is of the form $X^{o(1)}$. The main contribution of this note is to somewhat strengthen Wright's bound.

Theorem 1. *Suppose that a, M are positive coprime integers. Then $C_{a,M}(X) \ge X^{1/(6\log\log\log X)}$ for all sufficiently large X depending on the choice of M.*

That is, we reduce the power of $\log\log\log X$ to the first power and we remove the dependence on M in the bound, though there still remains the condition that X must be sufficiently large depending on M. (It is clear though that such a condition is necessary since if $M > X$ and $a = 1$, then there are no Carmichael numbers $n \le X$ in the residue class $a \bmod M$.)

Our proof largely follows Wright's proof of Theorem W, but with a few differences.

Unlike with primes, it is conceivable that a non-coprime residue class contains infinitely many Carmichael numbers, e. g., there may be infinitely many that are divisible by 3. This is unknown, but seems likely. Let $\lambda(n)$ denote the universal exponent of the group $(\mathbb{Z}/n\mathbb{Z})^*$ (so that a composite number n is a Carmichael number if and only if $\lambda(n) \mid n-1$). For a residue class $a \pmod M$, let $g = \gcd(a, M)$ and let $h = \gcd(\lambda(2g), M)$. A necessary condition that there is a Carmichael number $n \equiv a \pmod M$ is that $h \mid a-1$. I conjecture that if this condition holds then there are infnitely many Carmichael numbers $n \equiv a \pmod M$. (This modifies a similar conjecture in [3].) Though we do not know this for any example with $g > 1$, the old heuristic of Erdős [5] suggests that $C_{a,M}(X) \ge X^{1-o(1)}$ as $X \to \infty$.

2 Proof of Theorem 1

There is an elementary and easily-proved criterion for Carmichael numbers: a composite number n is one if and only if it is squarefree and $p-1 \mid n-1$ for each prime p dividing n. This is due to Korselt, and perhaps others, and is over a century old. In our construction, we will have a number L composed of many primes, a number k coprime to L that is not much larger than L, and primes p of the form $dk+1$ where $d \mid L$. We will show there are many $n \equiv a \pmod M$ that are square-free products of the p's and are $1 \pmod{kL}$. Such n, if they involve more than a single p, will satisfy Korselt's criterion and so are therefore Carmichael numbers.

We may assume that $M \ge 2$. Let $\mu = \varphi(4M)$, so that $4 \mid \mu$. Let y be an independent variable; our other quantities will depend on it. For a positive integer n, let $P(n)$ denote the largest prime factor of n (with $P(1) = 1$), and let $\omega(n)$ denote the number of distinct prime factors of n.

Let

$$\mathcal{Q}_0 = \{q \text{ prime} : y < q \le y\log^2 y,\ q \equiv -1 \pmod{\mu},\ P(q-1) \le y\}.$$

If $q \le y\log^2 y$ and $P(q-1) > y$, then q is of the form $mr+1$, where $m < \log^2 y$ and r is prime. By Brun's sieve (see [6, (6.1)]), the number of such primes q is at most

$$\sum_{m<\log^2 y} \sum_{\substack{r \text{ prime} \\ mr \le y\log^2 y \\ rm+1 \text{ prime}}} 1 \ll \sum_{m<\log^2 y} \frac{y\log^2 y}{\varphi(m)\log^2 y} \ll y\log\log y.$$

Also, the number of primes $q \le y\log^2 y$ with $q \equiv -1 \pmod{\mu}$ is $\sim \frac{1}{\varphi(\mu)} y\log y$ as $y \to \infty$ by the prime number theorem for residue classes. We conclude that

$$\#\mathcal{Q}_0 \sim \frac{1}{\varphi(\mu)} y\log y \quad \text{and} \quad \prod_{q\in\mathcal{Q}_0} q = \exp\Big(\frac{1+o(1)}{\varphi(\mu)} y\log^2 y\Big), \quad y \to \infty. \tag{19.1}$$

We also record that

$$\sum_{q\in\mathcal{Q}_0} \frac{1}{q} = o(1), \quad y \to \infty, \tag{19.2}$$

since this holds for all of the primes in the interval $(y, y\log^2 y]$.

Fix $0 < B < 5/12$; we shall choose a numerical value for B near to $5/12$ at the end of the argument. Let

$$x = M^{1/B} \prod_{q\in\mathcal{Q}_0} q^{1/B}. \tag{19.3}$$

It follows from [1, (0.3)] that there is an absolute constant D and a set $\mathcal{D}(x)$ of at most D integers greater than $\log x$, such that if $n \le x^B$, n is not divisible by any member of $\mathcal{D}(x)$, b is coprime to n, and $z \ge nx^{1-B}$, then the number of primes $p \le z$ with $p \equiv b \pmod{n}$ is $> \frac{1}{2}\pi(z)/\varphi(n)$.

For each number in $\mathcal{D}(x)$, we choose a prime factor and remove this prime from $\mathcal{Q}_0$ if it happens to be there. Let L be the product of the primes in the remaining set $\mathcal{Q}$, so that L is not divisible by any member of $\mathcal{D}(x)$, and $\mathcal{Q}$ satisfies (19.1) and (19.2). In particular,

$$\begin{aligned} &L = \exp\Big(\frac{1+o(1)}{\varphi(\mu)} y\log^2 y\Big), \quad \omega(L) \sim \frac{1}{\varphi(\mu)} y\log y, \quad \text{and} \\ &\sum_{q\mid L} \frac{1}{q} = o(1) \quad \text{as } y \to \infty. \end{aligned} \tag{19.4}$$

In addition, we have $ML \le x^B$.

For each $d \mid L$ and each quadratic residue $b \pmod{L/d}$, we consider the primes

- $p \le dx^{1-B}$,
- $p \equiv a \pmod{M}$,
- $p \equiv 1 \pmod{d}$,
- $p \equiv b \pmod{L/d}$.

Since M is coprime to L, the congruences may be glued to a single congruence modulo ML, and the number of such primes p is

$$> \frac{\pi(dx^{1-B})}{2\varphi(ML)} > \frac{dx^{1-B}}{3\varphi(ML)\log x}$$

for y sufficiently large.

We add these inequalities over the various choices of b, the number of which is $\varphi(L/d)/2^{\omega(L/d)}$, so the number of primes p corresponding to $d \mid L$ is

$$> \frac{dx^{1-B}2^{\omega(d)}}{3\cdot 2^{\omega(L)}\varphi(Md)\log x}.$$

We wish to impose an additional restriction on these primes p, namely that $\gcd((p-1)/d, L) = 1$. For a given prime $q \mid L$, the number of primes p just counted and for which $q \mid (p-1)/d$ is, via the Brun–Titchmarsh inequality,

$$\ll \frac{dx^{1-B}2^{\omega(d)}}{2^{\omega(L)}q\varphi(Md)\log(x/(qML))} \ll \frac{dx^{1-B}2^{\omega(d)}}{2^{\omega(L)}q\varphi(Md)\log x}.$$

Summing this over all $q \mid L$ and using that $\sum_{q\mid L} 1/q = o(1)$, these primes p are seen to be negligible. It follows that for y sufficiently large, there are

$$> \frac{dx^{1-B}2^{\omega(d)}}{2^{\omega(L)+2}\varphi(Md)\log x} > \frac{x^{1-B}2^{\omega(d)}}{2^{\omega(L)+2}\varphi(M)\log x}$$

primes $p \le dx^{1-B}$ with $p \equiv 1 \pmod{d}$, $\gcd((p-1)/d, L) = 1$, $p \equiv a \pmod{M}$, and p is a quadratic residue (mod L) (noting that 1 (mod d) is a quadratic residue (mod d)).

For each pair p, d as above, we map it to $(p-1)/d$ which is an integer $\le x^{1-B}$ coprime to L. The number of pairs p, d is

$$> \frac{x^{1-B}}{2^{\omega(L)+2}\varphi(M)\log x}\sum_{d\mid L} 2^{\omega(d)} = \frac{x^{1-B}3^{\omega(L)}}{2^{\omega(L)+2}\varphi(M)\log x}.$$

We conclude that there is a number $k \le x^{1-B}$ coprime to L, which has more than $(3/2)^{\omega(L)}/(4\varphi(M)\log x)$ representations as $(p-1)/d$. Let $\mathcal{P}$ be the set of primes $p = dk+1$

that arise in this way. Then

$$\#\mathcal{P} > \frac{(3/2)^{\omega(L)}}{4\varphi(M)\log x}. \tag{19.5}$$

For a finite abelian group G, let $n(G)$ denote Davenport's constant, the least number such that in any sequence of group elements of length $n(G)$ there is a non-empty subsequence with product the group identity. It is easy to see that $n(G) \geq \lambda(G)$ (the universal exponent for G), and in general it is not much larger: $n(G) \leq \lambda(G)(1 + \log(\#G))$. This result is essentially due to van Emde Boas–Kruyswijk and Meshulam; see [1].

Let G be the subgroup of $(\mathbb{Z}/kML\mathbb{Z})^*$ of residues $\equiv 1 \pmod{k}$. We have $\#G \leq ML$. Also, $\lambda(G) \leq M\lambda(L)$. (Note that, as usual, we denote $\lambda((\mathbb{Z}/L\mathbb{Z})^*)$ by $\lambda(L)$. It is the lcm of $q-1$ for primes $q \mid L$, using that L is square-free.) Each prime dividing $\lambda(L)$ is at most y and each prime power dividing $\lambda(L)$ is at most $y\log^2 y$, so that

$$\lambda(L) \leq (y\log^2 y)^{\pi(y)}.$$

Thus, for large y, using (19.4),

$$n(G) \leq M(y\log^2 y)^{\pi(y)}\log(ML) \leq e^{2y}. \tag{19.6}$$

For a sequence A of elements in a finite abelian group G, let A^* denote the set of nonempty subsequence products of A. In Baker–Schmidt [2, Proposition 1], it is shown that there is a number $s(G)$ such that if $\#A \geq s(G)$, then G has a nontrivial subgroup H such that $(A \cap H)^* = H$. Further,

$$s(G) \leq 5\lambda(G)^2\Omega(\#G)\log\bigl(3\lambda(G)\Omega(\#G)\bigr),$$

where $\Omega(m)$ is the number of prime factors of m counted with multiplicity. Thus, with G the group considered above, we have

$$s(G) \leq e^{2.5y}$$

for y sufficiently large.

It is this theorem that Matomäki and Wright use in their papers on Carmichael numbers. The role of the sequence A is played by $\mathcal{P}$, the set of primes constructed above of the form $dk+1$ where $d \mid L$. So, if $\#\mathcal{P} > s(G)$ we are guaranteed that every member of a nontrivial subgroup H of G is represented by a subset product of $\mathcal{P} \cap H$.

We do not know precisely what this subgroup H is, but we do know that it is nontrivial and that it is generated by members of $\mathcal{P}$. Well, suppose p_0 is in $\mathcal{P} \cap H$. Then $p_0^m \in H$ for every integer m. Note that by construction, $\gcd(\lambda(L)/2, \varphi(M)) = 1$, so there is an integer $m \equiv 1 \pmod{\varphi(M)}$ and $m \equiv 0 \pmod{\lambda(L)/2}$. Further, since p_0 is a quadratic residue (mod L), it follows that $p_0^{\lambda(L)/2} \equiv 1 \pmod{L}$. Thus $p_0^m \equiv 1 \pmod{L}$ and $p_0^m \equiv a \pmod{M}$ (since $m \equiv 1 \pmod{\varphi(M)}$).

Thus there is a subsequence product n of $\mathcal{P}$ that is 1 (mod kL) and a (mod M). (Note that every member of G is 1 (mod k).) Further, n is square-free and for each prime factor p of n and we have $p-1 \mid kL$. Since $n \equiv 1 \pmod{kL}$, we have $p-1 \mid n-1$. Thus $n \equiv a \pmod{M}$ is either a prime or a Carmichael number.

We actually have many subsequence products n of $\mathcal{P}$ that satisfy these conditions, and $\mathcal{P}$ has at most one element that is 1 (mod L), so we do not need to worry about the case that n is prime. We let $t = \lceil e^{3y} \rceil$, so that $t > s(G)$. As shown in [7, 9], the Baker–Schmidt result implies that $\mathcal{P}$ has at least

$$N := \binom{\#\mathcal{P} - n(G)}{t - n(G)} \Big/ \binom{\#\mathcal{P} - n(G)}{n(G)}$$

subsequence products n of length at most t which are Carmichael numbers in the residue class a (mod M). Thus

$$\begin{aligned} N &> \left(\frac{\#\mathcal{P} - n(G)}{t - n(G)}\right)^{t-n(G)} (\#\mathcal{P})^{-n(G)} \\ &> \left(\frac{\#\mathcal{P}}{t}\right)^{t-n(G)} (\#\mathcal{P})^{-n(G)} > (\#\mathcal{P})^{t-2n(G)} t^{-t}. \end{aligned}$$

Let $X = x^t$. Since each $p \in \mathcal{P}$ has $p \le x$, it follows that all of the Carmichael numbers constructed above are at most X. Using (19.1), (19.3), and (19.6), we have

$$X = \exp\left(\frac{1/B + o(1)}{\varphi(\mu)} ty \log^2 y\right),$$

and using (19.5) and (19.4) gives

$$\begin{aligned} N &\ge \exp\left(\frac{\log(3/2) + o(1)}{\varphi(\mu)} ty \log y - t \log t\right) \\ &= \exp\left(\frac{\log(3/2) + o(1)}{\varphi(\mu)} ty \log y\right). \end{aligned}$$

Thus $N \ge X^{(B \log(3/2) + o(1))/\log y}$. Now,

$$\log X \sim \frac{1}{B\varphi(\mu)} ty \log^2 y,$$

so that using $t = \lceil e^{3y} \rceil$,

$$\log \log X = 3y + O(\log y), \quad \log \log \log X = \log y + O(1).$$

We thus have $N \ge X^{(B \log(3/2) + o(1))/\log \log \log X}$. The number $B < 5/12$ can be chosen arbitrarily close to $5/12$ and since $(5/12)\log(3/2) > 1/6$, the theorem is proved.

Bibliography

[1] W. R. Alford, A. Granville and C. Pomerance, There are infinitely many Carmichael numbers, *Ann. Math. (2)*, **139** (1994), 703–722.
[2] R. C. Baker and W. M. Schmidt, Diophantine problems in variables restricted to the values 0 and 1, *J. Number Theory*, **12** (1980), 460–486.
[3] W. D. Banks and C. Pomerance, On Carmichael numbers in arithmetic progressions, *J. Aust. Math. Soc.*, **28** (2010), 313–321.
[4] R. D. Carmichael, A new number-theoretic function, *Bull., New Ser., Am. Math. Soc.*, **16** (1910), 232–238.
[5] P. Erdős, On pseudoprimes and Carmichael numbers, *Publ. Math. (Debr.)*, **4** (1956), 201–206.
[6] H. Halberstam and H.-E. Richert, *Sieve Methods*, London Mathematical Society Monographs, vol. **4**, Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London–New York, 1974.
[7] K. Matomäki, Carmichael numbers in arithmetic progressions, *J. Aust. Math. Soc.*, **94** (2013), 268–275.
[8] V. Šimerka, Zbytky z arithmetické posloupnosti. (Czech) [On the remainders of an arithmetic progression]. *Čas. Pěst. Math. Fys.*, **14** (1885), 221–225.
[9] T. Wright, Infinitely many Carmichael numbers in arithmetic progressions, *Bull. Lond. Math. Soc.*, **45** (2013), 943–952.

I. D. Shkredov and J. Solymosi

Tilted corners in integer grids

Dedicated to the memory of Ron Graham

Abstract: It was proved by Ron Graham and the second author that for any coloring of the $N \times N$ grid using fewer than $\log\log N$ colors, one can always find a monochromatic isosceles right triangle, a triangle with vertex coordinates $(x, y), (x + d, y)$, and $(x, y + d)$. In this paper, we are asking questions where not only axis-parallel, but tilted isosceles right triangles are considered as well. Both coloring and density variants of the problem will be discussed.

1 Introduction

In this paper, we are going to consider several problems inspired by questions raised by Ron Graham. After learning Szemerédi's proof of the Erdős–Turán conjecture on 4-term arithmetic progressions in dense subsets of integers [24], Graham asked the following question: Is it true that for any real number $\delta > 0$ there is a natural number $N_0 = N_0(\delta)$ such that for $N > N_0$ every subset of $[N] \times [N]$ of size at least δN^2 contains a square, i. e., a quadruple of the form $\{(a, b), (a + d, b), (a, b + d), (a + d, b + d)\}$ for some integer $d \neq 0$? ($[N] = \{0, 1, 2, \dots, N - 1\}$.) Using the full power of Szemerédi's theorem on k-term arithmetic progressions, Ajtai and Szemerédi in [1] proved a simpler statement: for sufficiently large N, every subset of $[N]\times[N]$ of size at least δN^2 contains corners, three points with coordinates $\{(a, b), (a+d, b), (a, b+d)\}$[1] (see also in [25]). Later Fürstenberg and Katznelson proved a much stronger, general theorem in [11], but their proof did not give an explicit bound as it uses ergodic theory. After Tim Gowers gave an analytical proof for Szemerédi's theorem (receiving a $1,000 check from Ron Graham who paid rewards offered by Paul Erdős), he again raised the question of finding a quantitative proof for Graham's question. Such proof was given by the second author

1 Throughout the paper, we are assuming that the corners and squares are not degenerate, $d \neq 0$.

Acknowledgement: Research of the first author received financial support from the Ministry of Educational and Science of the Russian Federation in the framework of MegaGrant no. 075-15-2019-1926. Research of the second author was supported in part by an NSERC Discovery grant, OTKA K 119528 and NKFI KKP 133819 grants. The authors are thankful to the referee for the useful comments and for pointing to important references.

I. D. Shkredov, Steklov Mathematical Institute, Moscow & IITP, Moscow & MIPT, Dolgoprudnii, Russia, e-mail: ilya.shkredov@gmail.com

J. Solymosi, University of British Columbia, Vancouver, Canada; and Obuda University, Budapest, Hungary, e-mail: solymosi@math.ubc.ca

https://doi.org/10.1515/9783110754216-020

in [23] using a hypergraph regularity lemma of Frankl and Rödl [12]. Although it is quantitative, it is still very far from a conjecture of Graham.

Conjecture 1 (Ron Graham [9]). Given a set of lattice points in the plane

$$S = \{p_1, p_2, \dots, p_i, p_{i+1}, \dots\},$$

let us denote the distance of p_i from the origin by d_i. If

$$\sum_{i=1}^{\infty} \frac{1}{d_i^2} = \infty,$$

then S contains the four vertices of an axes-parallel square.

The second author of this paper heard the conjecture from Ron Graham multiple times, with increasing reward offer. Once Ron said *"I think it is a safe bet to offer $1,000 for the solution. I don't think I ever have to pay that."*

Even after the recent breakthrough of Bloom and Sisask, breaking the logarithmic barrier in Roth's theorem on three term arithmetic progressions [3], we are very far from such bounds. We offer a weaker conjecture, changing squares to corners even allowing rotated (tilted) corners. In light of Theorem 7 below, it might be accessible using techniques available now.

Conjecture 2. Given a set of lattice points in the plane

$$S = \{p_1, p_2, \dots, p_i, p_{i+1}, \dots\},$$

let us denote the distance of p_i from the origin by d_i. If

$$\sum_{i=1}^{\infty} \frac{1}{d_i^2} = \infty,$$

then S contains the three vertices of an isosceles right triangle.

If we restrict our attention to axis parallel corners, then the best known density bound for the Ajtai–Szemerédi theorem belongs to the first author.

Theorem 1 (Shkredov [21]). *For sufficiently large N, every subset of $[N] \times [N]$ of size at least $N^2/(\log\log N)^C$ contains corners, three points with coordinates*

$$\{(a, b), (a + d, b), (a, b + d)\}.$$

This problem is one of the few examples where the coloring variant has a better (known) bound than its density version.

Ron Graham, Fan Chung, and Jozsef Solymosi.

Theorem 2 (Graham–Solymosi [10]). *For N large enough, any coloring of the $N \times N$ grid using fewer than* $\log \log N$ *colors, one can always find a monochromatic isosceles right triangle, a triangle with vertex coordinates $(x, y), (x + d, y)$, and $(x, y + d)$.*

In what follows, we will see variants of the above mentioned problems. The next section is about saturated point sets of the integer grid, sets without corners (or squares) which are maximal, adding any further gridpoint will result a corner (or square).

In Section 3, we summarize what are the best density results one can expect using the available techniques. Unfortunately, we cannot provide full proofs here; they are quite technical, but the arguments are hopefully complete enough that experts could reconstruct the proofs.

The last section is about related coloring problems, briefly addressing Euclidean Ramsey type problems, one of the many fields where Ron Graham has made significant impact. We close this Introduction with a nice result of Ron, similar to problems we are going to consider in this paper, finding monochromatic right triangles in integer grids.

Theorem 3 (Graham [8]). *For any r, there exists a positive integer $T(r)$ so that in any r-coloring of the lattice points $\mathbb{Z}^2$ of the plane, there is always a monochromatic right triangle with area exactly $T(r)$.*

2 Square saturated point sets

For technical reasons here and in future sections, we often switch between integer grids, $[n] \times [n]$ and planes over finite fields, $\mathbb{F}_p \times \mathbb{F}_p$.

The next definition we are going to use originates in graph theory. It goes back to a paper from 1964 by Erdős, Hajnal, and Moon [5].

Definition 1. Given a graph H, a graph G is *H-saturated* if G does not contain H but the addition of an edge joining any pair of nonadjacent vertices of G completes a copy of H. The *saturation number* of H, written $\mathrm{sat}(n, H)$ is the minimum number of edges in an H-saturated graph with n vertices (assuming $n \geq |V(H)|$).

Similar definitions can be given for various combinatorial structures. Here, we are going to use the definition for point sets in a plane. The point sets in the definition are subsets of a larger set, a *universe* U, like an integer grid $[n] \times [n]$, or a plane over the finite field $\mathbb{F}_p$. Problems of asking the saturation number for certain subsets of the integer grid $[n] \times [n]$, can be found as early as a paper of Erdős and Guy [6] from 1970, but similar problems were probably considered earlier.

Definition 2. Given a point set Q, another point set P is *Q-saturated* if P does not contain Q but the addition of any point outside of P completes a similar copy of Q. The *saturation number of Q,* written $sat(U, Q)$, is the minimum number of points in a Q-saturated point set in U.

Similarity here means that Q is similar to Q' if there is a transformation T, given by translation rotation and scaling, such that $T(Q) = Q'$.

Let us denote the corner, three points with coordinates $(0,0), (1,0), (0,1)$, by C, and the square, four points with coordinates $(0,0), (1,0), (0,1), (1,1)$, by Q.

Claim 1. *We have the following bounds on the saturation number for sets in $\mathbb{F}_p \times \mathbb{F}_p$ without (tilted) corners:*

$$\frac{p}{\sqrt{3}} \leq \mathrm{sat}(\mathbb{F}_p \times \mathbb{F}_p, C) \leq p.$$

Proof. Let S be a corner saturated set. Two elements of S are vertices of three distinct squares, so there are six points which could form a corner with the two elements. There are p^2 elements of $\mathbb{F}_p \times \mathbb{F}_p$, so

$$p^2 - |S| \leq 6\binom{|S|}{2},$$

providing the lower bound. The upper bound is a simple construction. Set

$$S = \{(0, i) : i \in \mathbb{F}_p\}.$$

Any point outside S with coordinates (a, b) would form a corner with $(0, b), (0, a+b) \in S$ (also with $(0, b), (0, b - a) \in S$). □

Both bounds hold in $[n] \times [n]$ as well. It would be interesting to find the sharp bound, or even just a construction in $\mathbb{F}_p \times \mathbb{F}_p$ where $|S| \leq p - 1$.

Before stating our next result, we recall a nice result of Katz and Tao [18], which will be our main tool bounding $\mathrm{sat}(U, Q)$. It gives a nontrivial bound on a basic quantity in additive combinatorics.

Theorem 4 (Katz–Tao [18]). *Let A, B, be finite subsets of a torsion-free abelian group, and let*

$$G \subset A \times B \quad \text{be such that } |A|, |B|, |\{a+b : (a,b) \in G\}| \le N.$$

Then $|\{a-b : (a,b) \in G\}| \le N^{11/6}$.

The $11/6 = 1.833\ldots$ exponent is not known to be sharp. A lower bound follows from a variant of a construction of Ruzsa [20] showing that the difference set can be as large as $N^{\log(6)/\log(3)} = N^{1.63093\ldots}$.

Theorem 5. *Let p be a prime $p \equiv 3 \pmod 4$. Then $\mathrm{sat}(\mathbb{F}_p \times \mathbb{F}_p, Q) \ge p^{12/11} - p^{3/5}$, i. e., every set which is square-saturated in $\mathbb{F}_p \times \mathbb{F}_p$ has size much larger than the obvious bound, p.*

Proof. In this case, we can write the elements of $\mathbb{F}_p \times \mathbb{F}_p$ similar to Gaussian integers. We can work on the field $F = \{a + ib : a, b \in \mathbb{F}_p\}$. Multiplying by i is a rotation by 90 degrees, so tilted corners are given by α, β, γ triples where

$$\alpha = (a + ib), \beta = (c + id), \gamma = \alpha + i(\alpha - \beta).$$

The key observation is that

$$\alpha = \frac{1+i}{2}\beta + \frac{1-i}{2}\gamma \quad \text{and} \quad -i\left(\frac{1+i}{2}\beta - \frac{1-i}{2}\gamma\right) = \beta + i(\alpha - \beta)$$

which is the fourth point of the square determined by α, β, γ. If S is Q-saturated then every point outside S is the fourth point of a square with the other three points in S. We know that $|S| = o(p^2)$ from Theorem 7, but here we can simply assume that $|S| \le p^{12/11}$ (for otherwise we are done). We have at least $p^2 - p^{12/11}$ points outside of S, all of which are fourth corners of a square with 3 vertices in S. Let us define a graph G with vertex set S and two elements (β, γ) form an edge if and only if they to be diagonals of a corner. Let us consider the sets $A = (1+i)S$, $B = (1-i)S$, and a graph G', defined on $A \times B$ as $(a, b) \in G'$ if and only if $(a/(1+i), b/(1-i)) \in G$. With these definitions, we have $\{a+b : (a,b) \in G'\} \subset 2S$, and $|\{a-b : (a,b) \in G'\}| \ge p^2 - p^{12/11}$. We can apply Theorem 4 with $N = |S|$, so $p^2 - p^{12/11} \le |S|^{11/6}$ giving the desired bound. □

Note that we did not use that S was square-free; all we used is that any point outside of S would form a square with a corner in S. The very same proof works for $[n] \times [n]$ using Gaussian integers.

Theorem 6. *If $S \subset [n]\times[n]$ has the property that for any $a \in ([n]\times[n])\setminus S$ there are three elements in S, which form a square with a, then $|S| \geq n^{12/11} + o(n)$.*

3 Maximum corner-free sets

In the previous section, we gave a bound on the smallest maximal corner-free set; here, we are going investigate what the size of the maximum corner-free set is. This part is not self-contained. We collected references to techniques and analogous results, which can be used to tackle our problem. To follow the arguments here, one should be familiar with Fourier methods used to deal with three- and four-term arithmetic progressions up to the level of use of Gowers norms. It was pointed out by the anonymous referee that Theorem 7 below was obtained in a nice paper of Prendiville [19, Corollary 1.3] and improved in [2, Theorem 2.21] by Bloom. Our proof below is similar to their work. The main goal is to give a better simple upper bound (on the density of sets without tilted corners) than what is known for axis parallel corners [21].

Theorem 7. *Let $A \subseteq [n]^2$ be a set having no isosceles right triangles. Then $|A| = O(n^2/\log^{c_1} n)$. Now if A does not contain squares, then $|A| = O(n^2/(\log\log n)^{c_2})$, where $c_1, c_2 > 0$ are some absolute constants.*

In order to prove the theorem, we will see a more general statement, which shows that the estimates for Szemerédi's theorem on k-term arithmetic progressions can be extended to k-element point sets in dimension two. As we mentioned earlier, this part of the paper is not self-explanatory, the statements are heavily dependent on the contents of the cited papers.

Lemma 1. *Let $k \geqslant 2$ be a positive integer and $M_1, \dots, M_k$ be 2×2 invertible matrices, $M_i \neq M_j$, $i \neq j$. Also, let $A \subseteq [n]^2$ be a set having no configurations $x, x+M_1y, \dots, x+M_ky$. Then there is $c_k > 0$ such that*

$$|A| = O\Big(\frac{n^2}{(\log\log n)^{c_k}}\Big), \tag{20.1}$$

and for $k = 2$ there exists $c > 0$ with

$$|A| = O\Big(\frac{n^2}{(\log n)^{c}}\Big). \tag{20.2}$$

Proof. Consider the quantity

$$\sigma = \sum_{\vec{x},\vec{y}} A(\vec{x})A(\vec{x}+M_1\vec{y})\dots A(\vec{x}+M_k\vec{y}).$$

Let us follow [15, Proposition 5.3] in the process changing the variables: $\vec{x} = \vec{z}_1 + \cdots + \vec{z}_k$, $\vec{x} + M_i\vec{y} = \sum_{j=1}^{k}(I - M_iM_j^{-1})\vec{z}_j$, so $\vec{y} = -\sum_{j=1}^{k} M_j^{-1}\vec{z}_j$. Since $M_i \neq M_j$, $i \neq j$ it follows that this is a uniform cover.[2] Then σ is expressed as

$$\sigma = n^{-k+2} \sum_{\vec{z}_1,\ldots,\vec{z}_k} A(\vec{z}_1 + \cdots + \vec{z}_k) \prod_{j=1}^{k} f_j(\vec{z}_1, \ldots, \vec{z}_k),$$

where the function f_j, $j \in [k]$ does not depend on $\vec{z}_j$. Hence by the characteristic property of Gowers norms, we see that σ is controlled by U^k–uniformity norm of A; see [14]. Notice also, that the quantity σ is affine–invariant. Applying the method from Bourgain's classical paper [4] (or for a sharper bound follow [3]) for $k = 2$, and for $k > 2$ following the steps of [13, 14, 17] we obtain a similar bound as in the case of arithmetic progressions of length k. □

Now we are ready to prove Theorem 7 as an easy corollary of Lemma 1.

Proof of Theorem 7. To calculate the number of isosceles right triangles, we need to consider

$$\sum_{\vec{x},\vec{y}} A(\vec{x})A(\vec{x} + \vec{y})A(\vec{x} + \vec{y}^{\perp}),$$

where $\vec{y} = (y_1, y_2)$ and $\vec{y}^{\perp} = (-y_2, y_1)$. So, in terms of Lemma 1, we have

$$M_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix};$$

hence both matrices are invertible. In the case of squares, the correspondent quantity is

$$\sum_{\vec{x},\vec{y}} A(\vec{x})A(\vec{x} + (y_1, y_2))A(\vec{x} + (-y_2, y_1))A(\vec{x} + (y_1 - y_2, y_1 + y_2)),$$

and hence

$$M_3 = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

is invertible as well so we can apply Lemma 1. □

Remark 1. As we have seen the case of squares corresponds to arithmetic progressions of length four and in this particular case the result can be improved further following

2 If A and B are finite nonempty sets and $\Phi : A \longrightarrow B$ is a map, then we say that Φ is a uniform cover of B by A if Φ is surjective and all the fibers $\{\Phi^{-1}(b) : b \in B\}$ have the same cardinality.

the work of Green and Tao in [16]. Also, it will be interesting to improve Bloom's bound (see [2]) $|A| = O(n^2/\log^{1-o(1)} n)$ for the maximal size of A having no isosceles right triangles to $|A| = O(n^2/\log^{1+c} n)$, $c > 0$, using methods from [3].

4 Coloring problems

In this section, we show two results from Euclidean Ramsey theory related to corners. These results follow almost directly from a more general result of the first author's paper *"On some problems of Euclidean Ramsey theory"* [22]. As in the previous section, we are not going to include the details; however, we give enough references so that with the cited paper the full proof can be recovered. As we stated in Theorem 2, coloring the integer grids with few colors results a monochromatic axis parallel corner. Using two colors and relaxing the axis parallel condition will give many monochromatic corners. The systematic investigation of monochromatic triangles in two coloring of $\mathbb{E}^2$ started in the third paper of the fundamental sequence of papers titled *"Euclidean Ramsey Theorems I, II, III"* [7]. The next result [22, Corollary 6] shows that two coloring of $\mathbb{F}_p \times \mathbb{F}_p$ always gives as many monochromatic corners as one would expect.

Theorem 8 (Shkredov [22]). *Let p be a sufficiently large prime number. Then for any two coloring of the plane $\mathbb{F} \times \mathbb{F}_p$ and any $a, b \neq 0$ such that a/b is a quadratic residue there is a monochromatic collinear triple $\{x, y, z\}$ with $\|y - x\| = a$, $\|z - y\| = b$.*

Actually, by the arguments of the proof of [22, Theorem 4] we consider $\sigma(R,R,R)$, $\sigma(B,B,B)$ the number of ERT at each color $R \bigsqcup B = \mathbb{F}_p \times \mathbb{F}_p$ and obtain

$$\sigma(R,R,R) + \sigma(B,B,B) = p^{-3}\big(|R|^3 + |B|^3\big) + \sigma(R,R,R) + \sigma(B,B,B) + 3\sigma(\delta_R, f_R, f_R) + 3\sigma(\delta_B, f_B, f_B), \tag{20.3}$$

where $f_R(x) = R(x) - |R|/p^2$, $f_B(x) = B(x) - |B|/p^2$ are the balanced functions of the colors B and R, correspondingly. As was shown in [22, Theorem 4], the terms $\sigma(\delta_R, f_R, f_R)$, $\sigma(\delta_B, f_B, f_B)$ in (20.3) are negligible thanks to the bound for the Kloosterman sums, and hence

$$\sigma(R,R,R) + \sigma(B,B,B) = p^{-3}\big(|R|^3 + |B|^3\big) + O(p^{5/2}) \geqslant p^3/4 - Cp^{5/2},$$

where $C > 0$ is an absolute constant. As a consequence, we obtain the following.

Theorem 9. *Let p be a prime number. Then for any two coloring of $\mathbb{F}_p \times \mathbb{F}_p$ the number of monochromatic isosceles right triangles is at least*

$$\frac{p^3}{4} + O(p^{5/2}).$$

A similar argument (now the Kloosterman sums are replaced to the bounds for the zeroth Bessel function) gives the following.

Theorem 10. *Suppose that we have a measurable coloring of the Euclidean plane with two colors. Then the measure of monochromatic isosceles right triangles in any of such coloring is at least* 0.0079.

Proof. To obtain the statement, we use [22, Theorem 10] and derive that the desired measure is at least

$$\frac{1}{4} + \frac{1}{4} \cdot \min_{t \geqslant 0} (2J_0(t) + J_0(\sqrt{2}t)), \tag{20.4}$$

where J_0 is the zeroth Bessel function. Using Maple, we see that the minimum in (20.4) is at least −0.9683275949. This completes the proof. □

Bibliography

[1] M. Ajtai and E. Szemerédi, Sets of lattice points that form no squares, *Studia Sci. Math. Hung.*, **9** (1974), 9–11.

[2] T. F. Bloom, *Quantitative topics in arithmetic combinatorics*, PhD thesis, University of Bristol, 2014, 154 pp. http://thomasbloom.org/thesis.pdf.

[3] T. F. Bloom and O. Sisask, Breaking the logarithmic barrier in Roth's theorem on arithmetic progressions, 2020, arXiv:2007.03528.

[4] J. Bourgain, On triples in arithmetic progression, *Geom. Funct. Anal.*, **9**(5) (1999), 968–984.

[5] P. Erdős, A. Hajnal and J. W. Moon, A problem in graph theory, *Am. Math. Mon.*, **71** (1964), 1107–1110.

[6] P. Erdős and R. K. Guy, Distinct distances between lattice points, *Elem. Math.*, **25** (1970), 121–123.

[7] P. Erdős, R. L. Graham, P. Montgomery, B. L. Rothschild, J. H. Spencer and E. G. Straus, Euclidean Ramsey Theorems. III, in *Infinite and finite sets*, Colloq. Math. Soc. Janos Bolyai, vol. **10**, Colloq., Keszthely, Vol I, 1973, pp. 559–583, North-Holland, Amsterdam, 1975.

[8] R. L. Graham, On partitions of $\mathbb{E}^n$, *J. Comb. Theory, Ser. A*, **28**(1) (1980), 89–97.

[9] R. L. Graham, Conjecture 8.4.6, in J. E. Goodman and J. O'Rourke (eds.) *Discrete and Computational Geometry*, p. 11, CRC Press, Boca Raton, NY, 1997.

[10] R. L. Graham and J. Solymosi, Monochromatic Isosceles Right Triangles on the Integer Grid, in M. Klazar, J. Kratochvil, M. Loebl, J. Matousek, P. Valtr, R. Thomas (eds.), *Topics in Discrete Mathematics*, pp. 129–132. Algorithms and Combinatorics, vol. **26**, Springer, Berlin, 2006.

[11] H. Fürstenberg and Y. Katznelson, A density version of the Hales–Jewett theorem, *J. Anal. Math.*, **57** (1991), 64–119.

[12] P. Frankl and V. Rödl, Extremal problems on set systems, *Random Struct. Algorithms*, **20** (2002), 131–164.

[13] W. T. Gowers, A new proof of Szemerédi's theorem for arithmetic progressions of length four, *Geom. Funct. Anal.*, **8**(3) (1998), 529–551.

[14] W. T. Gowers, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.*, **11** (2001), 465–588.

[15] B. Green and T. Tao, The primes contain arbitrarily long arithmetic progressions, Ann. Math. **167** (2008), 481–547.

[16] B. Green and T. Tao, New bounds for Szemerédi's theorem, III: A polylogarithmic bound for $r_4(N)$, *Mathematika*, **63**(3) (2017), 944–1040.
[17] B. Green, T. Tao and T. Ziegler, An inverse theorem for the Gowers $U^{s+1}[N]$–norm, *Ann. Math.*, **176**(2) (2012), 1231–1372.
[18] N. Katz and T. Tao, Bounds on arithmetic projections, and applications to the Kakeya conjecture, *Math. Res. Lett.*, **6** (1999), 625–630.
[19] S. Prendiville, Matrix progressions in multidimensional sets of integers, *Mathematika*, **61**(1) (2015), 14–48.
[20] I. Z. Ruzsa, in *Sums of finite sets*, Number Theory (New York, 1991–1995), pp. 281–293, Springer, New York, 1996.
[21] I. D. Shkredov, On a two-dimensional analog of Szemerédi's Theorem in Abelian groups, *Izv. Russ. Acad. Sci.*, **73**(5) (2009), 455–505.
[22] I. D. Shkredov, On some problems of Euclidean Ramsey theory, *Anal. Math.*, **41**(4) (2015), 299–310.
[23] J. Solymosi, A note on a question of Erdős and Graham, *Comb. Probab. Comput.*, **13** (2004), 263–267.
[24] E. Szemerédi, On sets of integers containing no four elements in arithmetic progression, *Acta Math. Acad. Sci. Hung.*, **20** (1969), 89–104.
[25] V. Vu, On a Question of Gowers, *Ann. Comb.*, **6**(2) (2002), 229–233.

Remembrances

Noga Alon

Remembrance of Ron Graham[1]

On October 30, 2019, Ron Graham and Fan Chung invited my wife, Nurit, and I to dinner at their place in La Jolla. Their house featured a magnificent ocean view, a vast collection of gadgets for performing magic tricks, and a truly impressive library of mathematical journals. Just before dinner, Ron gave me a reprint of the original foundational paper of Erdős and Rényi on the evolution of random graphs. This was one day before Ron's 84th birthday, which sadly turned out to be his last one.

Ron has been a superb researcher, who obtained fundamental results in discrete mathematics focusing on Ramsey theory and combinatorial number theory, and influential results in the study of scheduling and online algorithms. He served for many years in the advisory board of Random Structures and Algorithms. Not much of his work has been on the probabilistic method, but he has made seminal contributions to many of the topics investigated by probabilistic techniques, and his work on quasirandomness had a profound impact on the subject.

I wrote two joint papers with Ron. Anyone who worked with him or heard him talk about his work quickly realized that for him doing mathematics had always been fun. He clearly enjoyed it, and in fact clearly enjoyed a lot of other activities. Indeed, he served as the President of the American Mathematical Society, the President of the Mathematical Association of America, and the President of the International Jugglers' Association.

I first met him at a conference in 1984 in Kalamazoo. Ron gave the banquet lecture there. He started by saying that he realized everybody must be tired, and he had good news and bad news. The good news was that he only had one slide (this was still the era of physical slides with overhead projectors). The bad news, he added, was the size of this slide. Then he unfolded a slide of one by two meters, and the topic of the lecture turned out to be the Erdős graph. Ron had on his single slide the names of all the participants in the meeting, as well as the names of some others, and much of the induced subgraph of the Erdős graph on this set of roughly 400 vertices. The lecture was absolutely hilarious and yet conveyed the immense impact of Erdős on the development of discrete mathematics.

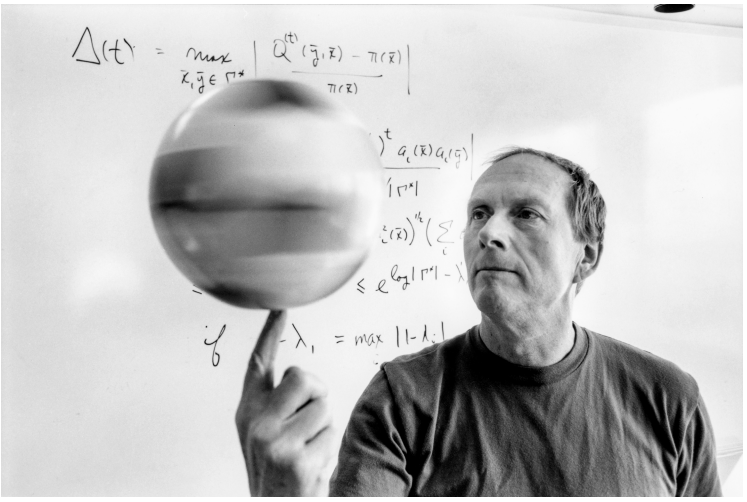
Starting in the mid 1980s, I visited the research group headed by Ron in Bell Labs multiple times. I served on the panel of the combinatorics section in ICM 94 that he

1 This chapter is reprinted courtesy of Wiley Publishing. This originally appeared in Random Structures Algorithms 57 (2020), no. 3.

Noga Alon, Department of Mathematics, Princeton University, Princeton, NJ, USA

https://doi.org/10.1515/9783110754216-021

chaired, visited him and Fan in their place in New Jersey, and later in California, and met him in conferences in numerous countries. One advantage of being a mathematician, that we sometimes forget, is that we get to meet and befriend interesting people. Even among these interesting people, Ron has been truly unique. I feel lucky I had the opportunity to know him for many years as an admirer of his work and personality, and as a friend.



Ron, 1995, New York City (courtesy of Ché Graham).

Tom C. Brown

Remembrance of Ron Graham

It was my good fortune to meet Ron when I was still in my 30s. This was in 1971 at a number theory meeting in Pullman, Washington. At the beginning of his talk, he said something like "First, something a little different," pulled 5 balls from a paper bag, and juggled them for 15 or 20 seconds. (A skill that I myself have still not mastered, after 50 years.) Then he started his talk. (As all his talks were, throughout his whole life, it was extremely clear, exciting, informative, and entertaining.) I had never seen anyone juggle before, and I found this amazing. Later that day or the next, I asked him for a lesson. In 1973, we practiced juggling together at a meeting on the occasion of Paul Erdős's 60th birthday at Keszthely, Hungary, and after that, I would get a juggling lesson from Ron at least once or twice a year. On one occasion, we stood side-by-side,

Tom C. Brown, Department of Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada

facing the same direction, and with my right hand and Ron's left hand, acted like a single person juggling 5 balls.

At that meeting in Keszthely, at lunch one day I happened to sit opposite Bruce Rothschild. He remarked that he could make forty throws with four balls, and someone asked him why so many mathematicians juggled. His reply: "Because Ron does."

In 1996, in an article by Donald J. Albers in Math Horizons, Ron is quoted as saying "I taught Tom Brown to juggle who in turn taught Joe Buhler how to juggle and he now is a better juggler than I am. That's a true mark of your teaching ability if you produce better students than you are." (In fact Ron himself was the main, if not the first, juggling teacher of Joe Buhler.)

If someone expressed awe or wonder at the depth and breadth of Ron's accomplishments (fluency in Mandarin, one-handed handstands, ping-pong at tournament level, and on and on—to say nothing of his vast mathematical output) he might say "Well, there *are* 168 hours in every week."

More than anything else, I remember Ron for his extraordinary generosity and helpfulness. He did so much to further the careers and/or help the lives of innumerable people, including myself.

Ron giving a $1000 reward to Timothy Gowers for the solution of a longstanding Ramsey theory problem at the "Paul Erdős and His Mathematics" conference, Budapest, July, 1999 (courtesy of Tom Brown).

Ron Graham, Fan Chung, and Tom Brown, June 6, 2019 (Courtesy of Tom Brown).

Steve Butler

Working with Ron Graham

The way I started working with Ron was a bit nontraditional; I erased chalkboards. I was a graduate student at UC San Diego and Ron was teaching the introduction to the discrete structures course. I started attending and noticed that the boards from the previous class were not erased and so I started erasing them and Ron and I would talk for a few moments. I eventually found out that his papers had never been posted online, and so then I had a new project; one which would make me have to visit Ron regularly to go through his papers and scan and process them (this list of papers is maintained online at rongraham.org).

Eventually, Ron started feeding me problems and I was able to make progress in some of them. Ron discovered I was fairly good at generating data and okay at understanding and explaining some of it. From this, our research partnership was born and we worked on dozens of papers covering a range of topics (shuffling, juggling, circle packings, guessing games, parking functions, origami, de Bruijn sequences, and more). He was the best of research collaborators, in part because of his steady flow of interesting problems that always seemed to suit our combined talents well.

Over time, we became close as Ron and Fan opened up their house and let me stay in their basements during the weekdays as needed (at the time my wife and I lived a significant distance away). This gave me the chance to see a different side of Ron on a fairly regular basis. Many people might envision Ron as a caring person with good

Steve Butler, Iowa State University, Ames, IA 50011, USA

humor who was always generous with his time with seemingly unending energy and a Starbucks coffee in his hand; and this is highly accurate! It always amazed me how much Ron was able to get done, how much time he spent connecting and reconnecting with people, and how much he loved watching movies and keeping up with sporting events.

Given how much he did every day, I always wondered how he was able to get so much math done. I soon discovered that he could spend the night thinking through math, it was not unusual for him to say "I was lying in bed at 3 a. m. and was thinking about the problem and had this idea...," and often it was the right idea for us to make progress! The other thing I discovered was how meticulous and detailed his notes on problems were. He would fill up folders of notes thinking about problems all carefully organized. These were not just random computations, rather these were a flow of consciousness as he would work through the problem. Reading them you would see him processing his thoughts and how he approached problem solving.[2] Ron worked by breaking problems down, and working through examples, often with a fair amount of computation involved, to make sure he understood them. He had an ability to see patterns in data and then work slowly, and meticulously to explain them. At times, it might seem superhuman, but mostly it was persistence and a lifetime of practice that helped him succeed.

In June 2020, about a month before he passed, Ron called me to talk about a problem involving sums of distinct powers. Namely, let $\lambda(x^n)$ be the largest number which is not the sum of distinct positive nth powers (this is sequence A001661 in the OEIS). So $\lambda(x^2) = 128$ since 128 cannot be written as a sum of distinct square numbers, but every number from $129 (= 10^2 + 5^2 + 2^2)$ and on can be. During the phone call, Ron explained how for very small values of n this could be computed, and mentioned that he believed that when the exponents were powers of 2 that something unusual seemed to happen and thought that it might be possible that $\lambda(x^8) > \lambda(x^9)$.

2 With what seems to be one exception, these notes might be the closest thing that we have to a journal for Ron. The exception was the following he wrote on November 29, 1973:

> Helicopter flight was uneventful and nice. While changing money into Krones (Danish), I found I had been given a counterfeit bill(!) from Summit and Elizabeth. Because of a strike and aggressive picketing the bags were delayed being loaded onto the plane. Consequently, we were 1 1/2 hours late in departing. I hope my bag makes it. In it are the scheduling monographs, all clothes, ping pong paddle, etc. Two seats away is a whining, yelping dog who is losing his mind. Rats, rats, rats.
> An hour later. The dog stopped. A brilliant fellow and his girlfriend seated across the aisle from me decided to celebrate. So, they opened a new bottle of champagne. Unfortunately, the decreased pressure of the air cabin caused the champagne to fly all over the place as it exploded open. What a mess!!

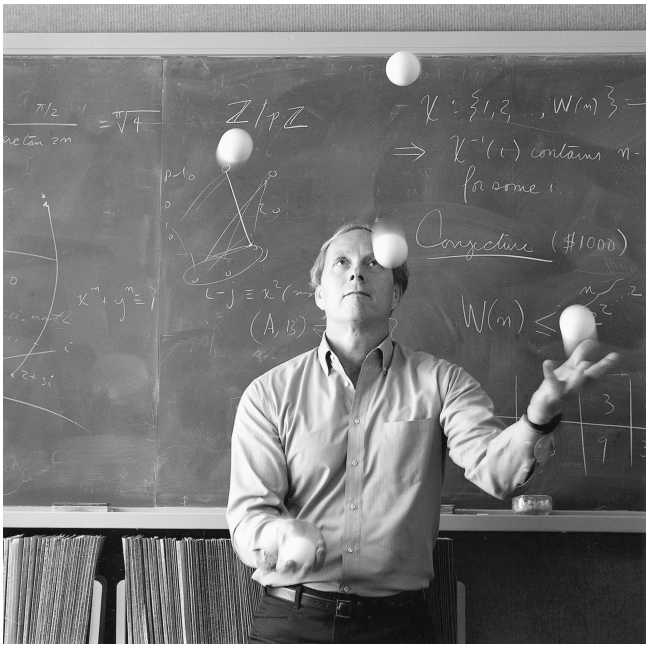
After his passing, I was looking through some of his old correspondence and discovered that this was a problem he had been thinking about for 60 years (making it one of the problems Ron worked the longest on without solving). In particular, in a letter dated May 29, 1969, to the editor for the Journal of Recreational Mathematics (D. L. Silverman), he wrote the following:

> In reference to your problem 71 in Vol. 2, No. 2 of JRM, I have included a reprint which lists some relevant references to this problem. The value for cubes (12758) was done by myself about ten years ago by hand (similar to the method used for squares) and checked by machine by several people in the past few years. The value for fourth powers (5134240) was obtained by S. Lin last year. The value for fifth powers seems feasible by present day techniques and I would guess it to be $\leq 2\times10^7$ or so. I have evidence that $\lambda(x^n)$ is abnormally large when n is a power of 2 and I would even conjecture that $\lambda(x^{2^n}) > \lambda(x^{2^n+1})$ for n sufficiently large (perhaps $n \geq 4$).

So in the spirit of Ron Graham and Paul Erdős, I offer $100 to anyone who can either prove or disprove the following.

Conjecture 1 ($100). *Let $\lambda(x^n)$ be the largest number, which is not the sum of distinct positive nth powers, show that the sequence $\lambda(x^n)$ for $n \geq 2$ is not monotonic.*

Ron Graham, Persi Diaconis, and Ricky Jay (courtesy of Ché Graham).



Ron, Bell Labs, Murray Hill, New Jersey, 1988 (Copyright, Peter Vidor).

Jerrold R. Griggs

Remembrance of Ron Graham

I studied Ron Graham's work on Ramsey theory, posets, and graphs in grad school. Martin Gardner often mentioned him in his Scientific American puzzle column. At my first professional talk, contributed at the annual 1977 AMS meetings, it was thrilling to read the nametag of the man asking me expert questions–Ron Graham. I met him again often at meetings, at colloquia (as he spoke everywhere), and on visits to Bell Labs. My visits with Ron and Fan in La Jolla were special and memorable.

He influenced my own work in extremal set theory, especially on chain partitions of posets and in helping develop what came to be called Ramsey–Sperner theory (his term, I think). I managed to confirm, for any given set X and for infinitely many values of n, a conjecture of Ron's on the maximum size of a family of subsets of an n-set such that any two both contain the same cyclic translate of X. However, Ron's offered cash prize requires a solution for every set for general n.

Besides sharing the latest progress in combinatorics, number theory, and computer science, Ron always had some new toy or gadget to amaze everyone, including an addictive early handheld electronic game (saving babies falling from a burning building), Rubik's cube, and a cellphone that could miraculously display web pages.

Jerrold R. Griggs, Department of Mathematics, University of South Carolina, Columbia, SC, USA

Ron loved to spread math gossip, especially who was moving where or what job opportunities might open up. I asked him what job was available in 1981, as I felt isolated at the University of Hawaii. When he learned of the discrete math search at the University of South Carolina, he encouraged them to call me. I made the move and spent my career here.

He seemed capable of mastering anything that interested him, whether it was speaking Mandarin or playing the piano. He was famous for his juggling, trampolining, and other circus acrobatics. Back when I could still play tennis, he challenged me to a match when he visited. I was proud to actually win the first game. But he found his rhythm and swept me away in the next twelve games. No mercy!

I remember when Ron raced Paul Erdős up the stairs at the Hyatt Regency in Atlanta, during the joint meetings there. I believe Ron agreed to go up ten floors, while "Uncle Paul" went up five. I think Paul was very annoyed that Ron beat him. I know Ron looked after Paul's interests in the USA, and they were very close, but both were very competitive.

Ron was one of the most famous and influential mathematicians in the world, and an unforgettable person. His colloquium and conference talks were fantastic, for the exciting math, the beautiful slides, and the great humor.

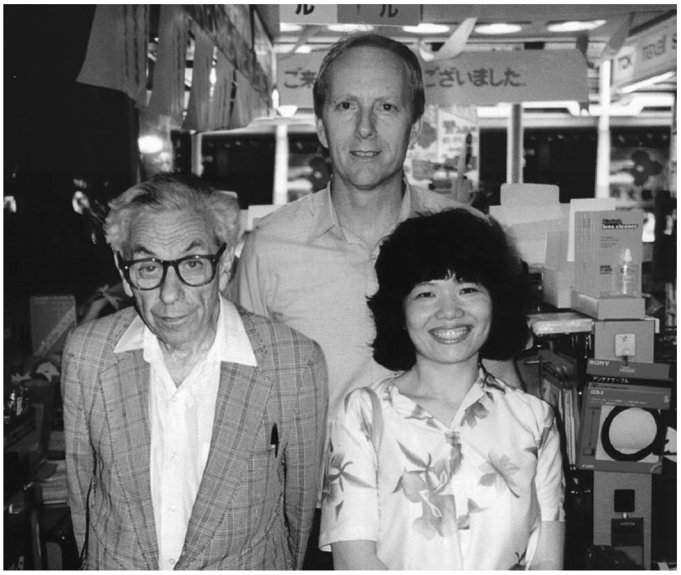
When I last saw him, at the 2015 conference in his honor at SFU, I was delighted that he attended my talk, on the spanning tree project that evolved into my contribution to this memorial volume.

Paul Erdos, Ron Graham, Peter Frankl, Jin Akiyama: First Japan International Conference on Graph Theory and Applications, 1986, Hakone, Japan (courtesy of Jerrold Griggs).

Neil Hindman

Remembrance of Ron Graham[3]

I would like to call attention to some of the greatest prose in the mathematical literature. As you may recall, if α, β, γ, and δ are cardinals and $[A]^\gamma = \{C \subseteq A : |C| = \gamma\}$, then the notation $\alpha \longrightarrow (\beta)^\gamma_\delta$ abbreviates the statement "whenever A is a set with $|A| = \alpha$ and $[A]^\gamma$ is divided into δ classes, there is some $B \in [A]^\beta$ such that $[B]^\gamma$ is contained in one of these classes." In his lovely little book *Rudiments of Ramsey Theory* (American Mathematical Society, Providence (1981)), Ron wrote "We will occasionally use this arrow notation unless there is danger of no confusion."



Ron Graham, Fan Chung, Paul Erdös (courtesy of Ché Graham).

Veselin Jungic

An unexpected encounter with Graham's Number

Fan and Ron frequently visited Vancouver, BC. They owned an apartment at the very edge of Stanley Park facing English Bay, both well-known Vancouver landmarks.

On February 25, 2014, Bojan Mohar from Simon Fraser University hosted a dinner for a group of local discrete mathematicians and a few visitors including Jarik Nešetřil

3 This chapter originally appeared in *Integers* Volume 7(2) (2007), Article A18.

Neil Hindman, 10900 Fiesta Road, 20901 Silver Spring, MD, USA
Veselin Jungic, Simon Fraser University, Burnaby, Canada

and Steve Butler. Since Ron was in Vancouver, he attended the dinner, too. And, as was always the case, in his charming way, Ron was at the center of every conversation, whether the subject was a burning mathematical question of the day, an anecdote about a member of the Ramsey theory community, or the latest blockbuster movie that he, of course, had already seen.

It also happened that, as part of the Math Catcher Outreach Program, the very next day I was scheduled to visit a school on Vancouver Island. During my class visits, I use storytelling, pictures, models, and hands-on activities to give to students an early positive experience with mathematics.

The next day after our dinner, there I was: visiting a school in a remote coastal community. While walking with me to the next class, my host warned me: "This is my 'slow' Grade 10 class. Many of my students struggle to focus for a longer period."

At one point during my workshop, I asked the class to tell me about something math-related but outside of their classroom experience. A student responded immediately: "Graham's number!"

When I asked what Graham's number was, the student started describing it as "three, arrow, arrow, ..." And the student beside him added: "The biggest number that we know."

I asked how they knew all that. "The Internet," they answered.

I don't think that in my professional life I ever gained as much respect from my audience than when I said: "I had dinner with Ron Graham yesterday."

June 16, 2015, in Burnaby, BC. From the right: Fan Chung-Graham, Ron Graham, Peter Borwein, Persi Diaconis, Richard Guy, and Veselin Jungic (Copyright Veselin Jungic).

Bruce M. Landman

Remembrance of Ron Graham

As I reflect back on my career in mathematics, it is clear to me that Ron Graham probably had as much impact as anyone, even though I saw Ron perhaps ten times in total.

When I was a graduate student in the early 1980s, I was looking for possible research problems, and was fortunate to buy a copy of the wonderful little book called "Old and New Problems and Results in Combinatorial Number Theory" by Paul Erdős and Ronald Graham. This book introduced me to van der Waerden's theorem and related questions, which became the basis for my doctoral dissertation and for almost all of my later research work in Ramsey theory.

I met Ron a couple of times over the next 15 years at meetings, and occasionally would write to him with a question regarding a research problem. Although he barely knew me, and I was a virtual "unknown" in research circles, he was always very generous with his take on my question and with suggestions. It was a great honor for me in 1997 when Ron expressed interest in collaborating on a paper (which later appeared in the Canadian Math Bulletin) that I and Tom Brown were working on. At that time, I was busy with a young family and I recall a lengthy phone conversation with Ron, with him very patiently explaining his idea for a proof, even while there were children's voices in the background during much of our conversation.

In 1998, I started thinking about the need for a journal whose focus was in the area of combinatorial number theory. I asked Ron what he thought about this idea. He liked the idea but suggested I talk to Carl Pomerance, Jarik Nesetril, and Mel Nathanson about it. His suggestion turned out to be prophetic as, not only did these three esteemed mathematicians like the idea but, in fact, they ended up serving as the Editors-in-Chief for the first 13 years of the journal, *Integers*. (Jarik and Mel are still in that role.) *Integers*, now in its 22nd year of publication, has given rise to a number of conferences and published proceedings. Indeed, if it were it not for Ron Graham's initial suggestion, it is unlikely that this memorial volume would exist.

I fondly remember the "Integers Conference 2005," held in honor of Ron's 70th birthday. Friends and colleagues of Ron from around the world gathered for the event. Entertainment at the conference banquet was provided by the Dazzling Mills Family, a professional juggling act. One of the most enjoyable parts of their performance was when Ron joined them onstage. The audience was thrilled.

In addition to his many accomplishments as a mathematician, Ron was also past president of the International Jugglers Association. In fact, Steve Mills, the leader of the Dazzling Mills Family and a world-renowned juggler, was taught to juggle by Ron himself.

I am just one of the countless people whose lives were positively affected by Ron Graham's kind and generous nature, and his knack for helping others be successful.

Bruce M. Landman, Augusta University, Augusta, GA, USA

Ron and Fan, Jinan, China, 1986 (courtesy of Ché Graham).

Jaroslav Nešetřil

Ronald Lewis Graham – Just a Few Memories[4]

It was high summer of 1973, Keszthely, Hungary. An unusually large meeting "Finite and Infinite Sets" was held there in Hotel Helikon from June 25 until July 1, on the occasion of Paul Erdős 60th birthday. It was an excellent meeting by any standards then and today, too. It is instructive to page through its 3rd volume proceedings [11]: totaling 1550 pages, containing papers by Rado, Tutte, de Bruijn, Straus, Berge, Galvin, Rudin, Guy, Selfridge, Hilton, McKenzie, Kleitman, Kunen, Milner, Neumann-Lara to name just a few; 12 papers coauthored by Erdős (including a joint paper with Lovász which inaugurated the Lovász Local Lemma), 3 papers by Shelah, 4 papers by Hajnal, 3 papers by Laver to list just a few contributions. And also 3 papers by Ron Graham all related to Ramsey with a total of more than 20 papers dealing with Ramsey type problems.

This was the meeting which for many years set high standards for universal combinatorial conferences which were held in 1970s and 1980s in France, Czechoslovakia, Hungary, Canada, and elsewhere. It was the event of the year.

One of my strong memories of the meeting is a tall athletic man who excelled at everything. His name was known to me as well as some of his work (even in that pre-email and preinternet age). But there he was: running, juggling, frisbeeing, and showing tricks in everything from photography to handling magically an overhead projector

4 This is an expanded version of an article that appeared in the Notices of the American Math. Soc., December, 2021 issue. Reprinted with permission of the American Math. Soc.

Jaroslav Nešetřil, Charles University, Praha, Czech Republic

(as far as I remember there was not a trampoline there). This was Ron Graham at his best, legendary already at that time. There we met for the first time.

My memory is vivid even now after years when in many meetings and collaboration I have seen that this youthful engaged style was Ron Graham's modus operandi. And later we all learned that many of these activities were not mere hobbies but professional level acts. What seemed to be easy and what Ron liked to display easily in his easygoing style was in fact hard learned and hard core. I believe this was symptomatic of his mathematics, too. Ron aimed for substantial and hard, yet concrete problems. He was a problem killer with an easy style. I still hear his "take it easy Jarik"—how helpful this was! He surrounded himself with very good people and aimed for depth and quality. In fact, he was a very *concrete mathematician* in the style of the famous textbook [6].

I have been fortunate to work with Ron on papers and books mostly related to Ramsey's theorem and its variations. Ramsey's theorem is a universal mathematical principle often summarized by Ron as "complete disorder is impossible." This was perhaps Ron's favorite if not key area. In fact, during his time *Ramsey theory* emerged as a "theory" from a mere particular collection of statements of "Ramsey-type" (due to Van den Waerden, Schur, Hilbert, Rado, and others). In this development had the above Keszthely meeting an important crystallizing role and Ron Graham's influence was pivotal. This was particularly true for structural Ramsey theory where the starting group of researchers was of course small. See the preface and the selection of topics covered by [4], the book which became a standard reference for this emerging field.

In this development, a particular place was assumed by the Hales–Jewett theorem [10] and the Graham–Rothschild theorem [2]. These are strong statements, which found many applications and serve as a tool for proving many Ramsey-type statements. Particularly, they led to a solution of Rota's conjecture (which is the analog of Ramsey's theorem for finite vector spaces) by Graham, Leeb, and Rothschild [3]. All five people involved in these early results received the inaugural Pólya Prize in 1971.

These results led to many papers since and blossomed into the whole theory. Today we seem to be witnessing a renaissance of the field in the context of topological dynamics, functional analysis, model theory and, of course, combinatorics. In 2016, there was even a meeting celebrating 50 years of Hales–Jewett theorem in Bellingham.

I cannot resist the temptation to try to outline the mathematical meaning of these results. Ramsey's theorem guarantees certain regularity in large structures. For graphs, this regularity is a complete graph or an empty graph. Ramsey's theorem is in fact a general combinatorial principle useful across mathematics and the theory of computing. Some 50 years later Hales–Jewett and Graham–Rothschild found another such principle, this time both combinatorial and geometrical. It is possible to sketch it as follows:

Think of a finite set A as an alphabet, for example, $A = \{1, 2, \dots, k\}$. The product set A^d is then just a set of vectors $(a_1, \dots, a_d)$ with each $a_i \in A$. Alternatively, we may view A^d as a geometric object: A^d is d-dimensional cube (or rather A-cube) with

sides indexed by A. Thus $\{1,2,3\}^3$ is the popular Rubik's cube, $\{1,2,3\}^2$ is a square lattice like in the game tic-tac-toe. In this way, A^d may be viewed as a board for a d-dimensional version of this game. (In fact, this was one of the motivations of the original paper [10]. As in tic-tac-toe, we are looking for lines, horizontal, vertical, diagonal, and this may be defined for any d-dimensional cube and more generally we can speak about d-dimensional subcubes of a D-dimensional cube. One can express lines and d-dimensional subcubes concisely as *parameter words* (a term coined by Graham–Rothschild) where parameters indicate which coordinates are "moving." (In a square grid, the lines have the form $(a\lambda), (\lambda b)$ and $(\lambda\lambda)$ for the diagonal.) The exact definition is a bit technical but it confirms the above intuition. And this is all that is needed in order to state the result of Graham and Rothschild [2].

Theorem. *For every choice of alphabet A and positive integers d, n, there exists $N = N(A, d, n)$ such that whenever the set of all d-dimensional subcubes of A^N is partitioned in two parts then one of the parts has to contain an n-dimensional A-subcube with all its d-dimensional A-subcubes belonging to one of the classes of the partition.*

(Recall that Ramsey's theorem speaks about subsets instead of subcubes. Hales–Jewett theorem corresponds to $d = 0, n = 1$.)

It is perhaps surprising that such a seemingly technical result plays such an important role. But this is like Ramsey's theorem itself: it is a combinatorial principle, which fits in diverse situations and assumptions. The Graham–Rothschild theorem is a far-reaching generalization of Ramsey's theorem, provides a proper setting for Van der Waerden's theorem, and as it was realized later, it yields a "dual" form of Ramsey's theorem. This inspiration lives on.

The mathematics of Ron Graham is important and it spans many diverse areas. But still I think that Ramsey theory was closest to his heart. It was also the topic of Ron's invited lecture at ICM 82 (held in Warsaw 1983), [5]. Ramsey theory was also dear to Paul Erdős as witnessed by the 2-volume set Mathematics of Paul Erdős where it occupies a whole chapter ([7, 8]; see also [9]). In fact these volumes, which were assembled still under the guidance of Paul Erdős himself, contain many pages written by editors reflecting a long experience of collaboration with Erdős.

The other parts of Ron Graham's activity are reflected by a forthcoming volume of the journal INTEGERS and also by volumes, which were published for his 80th birthday [12, 1]. Ron was public figure and a well-known mathematician, often representing mathematics as a whole. This was nicely documented recently by an article in The New Yorker [13]. But I want to add yet another aspect of Ron's personality. I believe Ron Graham was a patriot. Patriot in a very good and decent sense. He liked very much Bell Labs; he liked his country. Perhaps this was one of the factors why he had such a keen interest in the development of friendship on the other side of the Iron Curtain. This interest was of course motivated by mathematics and it was forged by P. Erdős and the excellence of Hungarian combinatorics. But there was much more on a personal and, yes, human level—he really tried to be helpful. He encouraged us and served as

Elegant easy style (photo by John Gimbel).

a bridge to the world. And this was in those times when there were not many bridges at all, and it needed courage. It would take too long to illustrate this. Let us just mention that he helped to establish DIMATIA (as a "European DIMACS"), steadily invited people to Bell Labs and communicated about chances, possibilities, and simply was spreading informations and books.

There were no obstacles or curtains for Ron. In this, he is a great role model and this is the lasting legacy of his personality. He is and will be remembered by many.

Bibliography

[1] Connections in Discrete Mathematics (a celebration of work of Ron Graham) eds. S. Butler, J. Cooper, G. Hurlbert, Cambridge University Press 2018.
[2] Graham R.L., Rothschild B.L., Ramsey's theorem for n-parameter sets, Trans. Amer. math.Soc. 159 (1971), 257-292.
[3] Graham R.L., Leeb, K., Rothschild B.L., Ramsey's Theorem for a Class of Categories, Adv. in Math. 8 (1972) 417-433. Reprinted with corrections in Classic Papers in Combinatorics, I. Gessel and G.-C. Rota, eds. Birkhauser, Boston, (1987), 431-445.
[4] Graham R.L., Rothschild B.L., Spencer J.H., Ramsey Theory, John Wiley & Sons, Inc., New York 1980, 2nd edition 1990.
[5] Graham R.L., Recent developments in Ramsey Theory, Proc. ICM 82, Polish Sci. Publishers and Elsevier (1983), 1555-1567.
[6] Graham R.L., Knuth D, Potaschnik O, Concrete Mathematics – A foundation for computer science, Addison-Wesley 1989, second printing 1994.
[7] Graham R.L. and Nešetřil J. (eds.), The Mathematics of Paul Erdős. I–II, Springer-Verlag, New York, 1996.

[8] Graham R.L., Nešetřil J., and Butler S. (eds.), The Mathematics of Paul Erdős. I–II (second edition), Springer-Verlag, New York, 2013.
[9] Graham R.L., Nešetřil J., Ramsey theory in the work of Paul Erdős, The Mathematics of Paul Erdős, R. Graham, J. Nešetřil and S. Butler eds., Springer-Verlag, New York, 2013, 171-193.
[10] Hales A.W., Jewett R.I., Regularity and positional games, Trans. Amer. Math. Soc. 106 (1963), 222-229.
[11] Infinite and Finite Sets to Paul Erdős on his 60th birthday, Vol. I, II, III (eds. A. Hajnal, R. Rado, Vera T. Sós), Coll. Math. Societatis Janos Bolyai, North-Holland , 1975.
[12] Journal of Combinatorics, Vol 8, No 3, 2017.
[13] D. Rockmore: Three Mathematicians we lost in 2020, The New Yorker Dec 31, 2020.

Steve Butler

A selected bibliography of Ron Graham

1 Introduction

Ron Graham was a prolific mathematician. During his six decade career, he had over 400 publications;[1] these included pure research papers, surveys (on scheduling and Ramsey theory), and popular expository papers (including four papers in *Scientific American*). In total, his papers combine to over 5000 pages (this figure does not include his books). In Figure 22.1 is a chart of Ron's mathematical output in papers each year starting from his first publication in 1963.

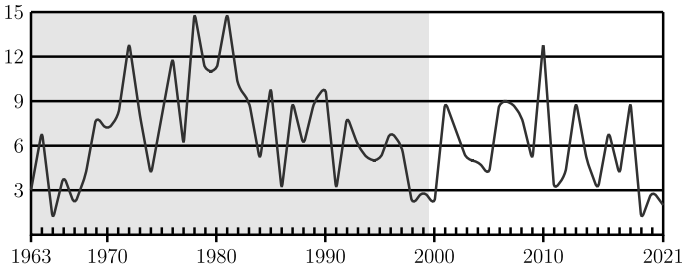


Figure 22.1: The number of published papers per year for Ron Graham from 1963–2021. The shaded region before 2000 marks his time at Bell Labs (and its various incarnations); the unshaded region after 2000 marks his time at UC San Diego.

Ron embodied the collaborative spirit of combinatorics with over 200 coauthors (his top four collaborators were Fan Chung, Paul Erdős, Persi Diaconis, and Steve Butler). He also published under pseudonyms on two occasions.[2]

2 Books written by Ron Graham

In this section, we give a full list of the books written by Ron, together with some information about the contents or stories related to their publication.

1 And still has several papers in the process of being published; a forthcoming book on the mathematics of juggling; and several unpublished papers, which will eventually appear in a volume of collected works.

2 *G. W. Peck*, a name formed from the last initials of the authors Ron Graham, Douglas West, George Purdy, Paul Erdős, Fan Chung, and Daniel Kleitman; and *Tom Odda*, a name which phonetically derives from the Chinese phrase 他媽的.

Steve Butler, Iowa State University, Ames, IA 50011 , USA, e-mail: butler@iastate.edu

https://doi.org/10.1515/9783110754216-022

P. Erdős and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Monographie No 28 de L'Enseignement Mathématique, Université Genève, 1980, 128 pp.

Ron's dissertation in 1962, under the direction of Derrick Lehmer, was *On finite sums of rational numbers*. The dissertation dealt with problems in combinatorial number theory, in particular Egyptian fractions. Around the time he was finishing up his degree, he started a correspondence with Paul Erdős who had written often on the subject of Egyptian fractions. The two met in person in August 1963 at a number theory conference in Boulder, Colorado, USA. While they would go on and collaborate on many topics, both of them always had a love for combinatorial number theory. They compiled this collection of results and open problems that went on and inspired the research of many mathematicians in the field.

Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer, *Ramsey Theory*, Wiley, 1980, ix+174 pp.

Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer, *Ramsey Theory, Second edition*, Wiley, 1990, xiii+196 pp. (The paperback edition in 2013 includes several small additions and corrections.)

Ron became widely known and recognized for his contributions to the field of Ramsey Theory; including being a corecipient of the first Polya prize. In many ways, his work, including these books, helped to bring a collection of various topics together and present them in a cohesive theory, which helped give the field a firm foundation on which to build.
Ron shared the story of how in the first edition all the coauthors had assumed that the other coauthors had checked the proof of Ramsey's theorem; but it turned out that the initial version actually had an incorrect proof! (A good reason for a second edition.)

Ronald L. Graham, *Rudiments of Ramsey Theory*, AMS, 1981, v+65 pp.

Ron Graham and Steve Butler, *Rudiments of Ramsey Theory, Second edition*, AMS, 2015, ix+82 pp.

This volume grew out of a set of CBMS lectures that Ron delivered at St. Olaf College in 1979. In the time between the first and second editions, there had been significant progress and growth in Ramsey theory leading to a new edition, which was meant to be up to date; but even after just a few years, several problems raised in the revised text were answered, showing the growth in the field.

Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics*, Addison Wesley, 1989, xiii+625 pp.

Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics, Second edition*, Addison Wesley, 1994, xiii+657 pp.

This grew out of a course that had been introduced by Donald Knuth at Stanford, which served as an advanced introduction to the tools of discrete mathematics. Ron taught the course on two

occasions when visiting Stanford, in 1979 and 1981 (with Ron teaching the course it gave Donald Knuth some much needed time to work on TEX). Ron was a wildly popular teacher and the students of the class had annual reunions for several years after the course was taught. Ron would later teach much of the same material at Princeton and UC San Diego. This book has proven to be a commercial success and has so far been translated into a dozen languages.
Ron was also particularly proud of the impact that his teaching of the concrete math course had in convincing his TA, Mark Haiman, to switch his study to mathematics.

Fan Chung and Ron Graham, *Erdős on Graphs, His Legacy of Unsolved Problems*, AK Peters, 1999, xiii+142 pp.

Throughout his life, Paul Erdős posed many problems to mathematicians (some even consider Erdős as the king of problem posers). After Erdős passed in 1996, Fan and Ron collected the open problems that Erdős had posed in the field of graph theory into a book for reference and inspiration for the community. Since the publication of this book several of the problems have been solved, but many of them still remain open.
Fan and Ron took good care of Erdős, often hosting him in their home when he visited.[3] From 1973, Ron also took care of Erdős' financial matters and continued to pay out the bounties that Erdős had attached to any of his problems.

Persi Diaconis and Ron Graham, *Magical Mathematics, The Mathematical Ideas that Animate Great Magic Tricks*, Princeton, 2012, xii+244 pp.

This book showed how one could take some nontrivial mathematics and incorporate them into card tricks and other magic-related phenomena. The book was written for a general audience and was well received, earning the 2013 Euler Book Prize. Ron himself loved demonstrating card tricks and had decks of cards all over his office and home, and almost always had one on hand while at conferences.
The book took almost 20 years to write. As the authors were drawing near to having the book ready they reached out to the publisher, only to discover that the publisher had lost the original contract! The publisher was of course happy to make a new contract and publish the book.

In addition to these books, it should be noted that Ron worked as editor for countless numbers of books and book series. Two in particular that stand out for their contribution to the field are the following:

- *Handbook of Combinatorics; Volumes I and II* edited by R. L. Graham, M. Grötschel, and L. Lovász, published in 1995. This is a tour de force of combinatorial surveys and has yet to be equaled in the 25 years since publication.
- *The Mathematics of Paul Erdős; Volumes I and II* edited by R. L. Graham and J. Nešetřil, published in 1996 (a second edition with additional materials was published in 2013).

3 When Erdős would leave their home, Ron would pack up his bags for him and would playfully include a random item, such as a heavy clock, in his luggage to be discovered at his next destination.

3 Selected papers written by Ron Graham

In this section, we highlight several papers written by Ron. These have been chosen to showcase several aspects of Ron's mathematical interests and contributions. The full extent of Ron's contribution to mathematics is much more significant and there are many more papers that could have been included in this list. A complete publication list can be accessed at rongraham.org.

R. L. Graham and H. O. Pollak, On the addressing problem for loop switching, *Bell System Technical Journal* **50** (1971), 2495–2519.

Ron considered this one of his most "fruitful" papers. The problem considered is the addressing of vertices in a graph with $\{0, 1, d\}$ strings so that the distance between any two vertices is the number of times the corresponding entries in the string are 0 and 1 (d is interpreted as "don't care"). Among other things, they showed that there is a connection to the eigenvalues of the distance matrix of the graph (this, together with follow-up papers, helped bring the distance matrix to the attention of the spectral graph theory community). This is also the origination of the Graham–Pollak theorem, which states that a complete graph on n vertices cannot be edge decomposed into fewer than $n - 1$ edge-disjoint complete bipartite graphs. Subsequent work of Peter Winkler would show that $n - 1$ is worst possible for any graph on n vertices.

R. L. Graham and B. L. Rothschild, Ramsey's Theorem for n-parameter sets, *Transactions of the American Mathematical Society* **159** (1971), 257–292.

This paper was a seminal paper in Ramsey theory; it was able to give a generalization of Ramsey's theorem and also a much stronger version of the Hales–Jewett theorem; this was also used to partially prove a conjecture of Rota on finite vector spaces. Ron wrote dozens of papers in the field of Ramsey theory helping to grow the field and also broaden it into new and interesting directions.

R. L. Graham, An efficient algorithm for determining the convex hull of a finite planar set, *Information Processing Letters* **1** (1972), 132–133.

This paper became one of Ron's most cited papers and was the origination of the "Graham scan." Originally though, when he wrote it he thought of it as a "throw-away" result and did not think that much would come out of it. The paper presents an $O(n \log n)$ algorithm for finding the convex hull of a set of n points in the plane. This is the first paper in the field of computational geometry and also one of the first examples of using amortized analysis of an algorithm.

R. L. Graham, The largest small hexagon, *Journal of Combinatorial Theory (A)* **18** (1975), 165–170.

The paper finds the hexagon with maximal area for which no two points of the hexagon are more than distance 1 apart. The actual shape is roughly a pentagon with one side pushed out (though there are slightly more technical details involved, including finding roots of a high-degree polynomial). This is one of many papers which highlight Ron's geometric approach to problems.

After this paper came out, a company contacted Ron about the possibility of using the shape to make memorial markers, making this perhaps one of the more unusual applications of combinatorics.

Persi Diaconis, R. L. Graham, and William M. Kantor, The mathematics of perfect shuffles, *Advances in Applied Mathematics* **4** (1983), 175–196.

This paper looks at the possible ways that one can rearrange a deck using perfect in- and outshuffles, in particular the group structure formed with these two generators. A complete characterization is given, including the quite unexpected result that for decks of 24 cards the result is the Mathieu group M_{12} (a result which fascinated John Conway to no end).

F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362.

This is the first in a series of papers that introduced the notion of quasirandomness. These are collections of properties which (1) a "random" object would be expected to have and (2) if any one property is satisfied, then all properties are satisfied. This has become an important tool for combinatorialists allowing them to split problems into two parts, one which is random-like and can be solved using typical analysis for random graphs, and the other which is nonrandom and so has some rich structure, which can be used.

Fan Chung, Martin Gardner, and Ron Graham, Steiner trees on a checkerboard, *Mathematics Magazine* **62** (1989), 83–96.

Bell Labs from time to time was called on to solve problems with real-world implications, and among these were Steiner trees. This was motivated by litigation, which said that AT&T could only charge for the least costly installation to create a network connecting several sites. This translates into finding the shortest set of connections, which connect the sites (possibly adding "imaginary sites" in the meantime), which is known as the Steiner tree problem. In general it was shown (by Ron and collaborators) that this problem was NP-hard, but several special cases could be considered, e. g., as discussed here on a checkerboard.
This paper won the Carl Allendoerfer Award in 1990, and was the only mathematical paper of Martin Gardner. (Martin Gardner wrote the "Mathematical Games" column in *Scientific American* which helped to popularize mathematics; Martin and Ron were good friends and Ron helped connect Martin to many mathematicians and their results.)

Joe Buhler, David Eisenbud, Ron Graham, and Colin Wright, Juggling drops and descents, *American Mathematical Monthly* **101** (1994), 507–519.

Ron was juggling since the late 1950s, served as the president of the International Jugglers' Association in 1972, and was a cocreator of Mills' Mess (a popular juggling trick). Ron found a way to combine his passion of juggling and mathematics and helped to popularize juggling inside the mathematical community.
This particular paper was one of the first to deal with siteswap notation (which describes patterns by instructing what to do with the ball in the hand). Ron would later write papers exploring juggling through state graphs, juggling cards, multiplex juggling, and more.

Ronald L. Graham, Jeffrey C. Lagarias, Colin L. Mallows, Allan R. Wilks, and Catherine H. Yan, Apollonian circle packings: number theory, *Journal of Number Theory* **100** (2003), 1–45.

Apollonian circle packings start with three mutually tangent circles and then recursively filling in "triangular" holes with a new circle tangent to all three sides. It was known that if the initial curvatures of the three initial circles together with the first circle to fill a hole were all integers, then *every* curvature that would be created would also be an integer. Ron had a huge poster of an Apollonian packing hanging in his office with the curvatures printed. He would look at the poster and over time noticed, among other things, that there were certain modular constraints happening. Eventually, a more full theory of the circle packings was developed and this was the first in a series of several papers on the subject.

De Gruyter Proceedings in Mathematics

Kağan Kurşungöz, Ayberk Zeytin (Eds.)
Number Theory. Proceedings of the Journées Arithmétiques, 2019, XXXI, held at Istanbul University, 2021
ISBN 978-3-11-076029-3, e-ISBN 978-3-11-076111-5

Aref Jeribi (Ed.)
Operator Theory. Proceedings of the International Conference on Operator Theory, Hammamet, Tunisia, April 30–May 3, 2018
ISBN 978-3-11-059686-1, e-ISBN 978-3-11-059819-3

James A. Davis (Ed.)
Finite Fields and their Applications. Proceedings of the 14th International Conference on Finite Fields and their Applications, Vancouver, June 3–7, 2019
ISBN 978-3-11-062123-5, e-ISBN 978-3-11-062173-0

Mahmoud Filali (Ed.)
Banach Algebras and Applications. Proceedings of the International Conference held at the University of Oulu, July 3–11, 2017
ISBN 978-3-11-060132-9, e-ISBN 978-3-11-060241-8

Ioannis Emmanouil, Anargyros Fellouris, Apostolos Giannopoulos, Sofia Lambropoulou (Eds.)
First Congress of Greek Mathematicians. Proceedings of the Congress held in Athens, Greece, June 25-–30, 2018
ISBN 978-3-11-066016-6, e-ISBN 978-3-11-066307-5

Paul Baginski, Benjamin Fine, Anja Moldenhauer, Gerhard Rosenberger, Vladimir Shpilrain (Eds.)
Elementary Theory of Groups and Group Rings, and Related Topics.
Proceedings of the Conference held at Fairfield University and at the Graduate Center, CUNY, November 1–2, 2018
ISBN 978-3-11-063673-4, e-ISBN 978-3-11-063838-7

Galina Filipuk, Alberto Lastra, Sławomir Michalik, Yoshitsugu Takei, Henryk Żołądek (Eds.)
Complex Differential and Difference Equations. Proceedings of the School and Conference held at Będlewo, Poland, September 2–15, 2018

www.degruyter.com