MATHEMATICAL CONSTANTS II

Famous mathematical constants include the ratio of circular circumference to diameter, $\pi = 3.14\ldots$, and the natural logarithm base, $e = 2.718\ldots$. Students and professionals can often name a few others, but there are many more buried in the literature awaiting discovery.

How do such constants arise, and why are they important? Here the author renews the search he began in his book *Mathematical Constants*, adding another 133 essays that broaden the landscape. Topics include the minimality of soap film surfaces, prime numbers, elliptic curves and modular forms, Poisson–Voronoi tessellations, random triangles, Brownian motion, uncertainty inequalities, Prandtl–Blasius flow (from fluid dynamics), Lyapunov exponents, knots and tangles, continued fractions, Galton–Watson trees, electrical capacitance (from potential theory), Zermelo's navigation problem, and the optimal control of a pendulum. Unsolved problems appear virtually everywhere as well. This volume continues an outstanding scholarly attempt to bring together all significant mathematical constants in one place.

Encyclopedia of Mathematics and Its Applications

This series is devoted to significant topics or themes that have wide application in mathematics or mathematical science and for which a detailed development of the abstract theory is less important than a thorough and concrete exploration of the implications and applications.

Books in the **Encyclopedia of Mathematics and Its Applications** cover their subjects comprehensively. Less important results may be summarized as exercises at the ends of chapters. For technicalities, readers are referred to the bibliography, which is expected to be comprehensive. As a result, volumes are encyclopedic references or manageable guides to major subjects.

All the titles listed below can be obtained from good booksellers or from Cambridge University Press. For a complete series listing visit

www.cambridge.org/mathematics.

122 S. Khrushchev *Orthogonal Polynomials and Continued Fractions*
123 H. Nagamochi and T. Ibaraki *Algorithmic Aspects of Graph Connectivity*
124 F. W. King *Hilbert Transforms I*
125 F. W. King *Hilbert Transforms II*
126 O. Calin and D.-C. Chang *Sub-Riemannian Geometry*
127 M. Grabisch *et al. Aggregation Functions*
128 L. W. Beineke and R. J. Wilson (eds.) with J. L. Gross and T. W. Tucker *Topics in Topological Graph Theory*
129 J. Berstel, D. Perrin and C. Reutenauer *Codes and Automata*
130 T. G. Faticoni *Modules over Endomorphism Rings*
131 H. Morimoto *Stochastic Control and Mathematical Modeling*
132 G. Schmidt *Relational Mathematics*
133 P. Kornerup and D. W. Matula *Finite Precision Number Systems and Arithmetic*
134 Y. Crama and P. L. Hammer (eds.) *Boolean Models and Methods in Mathematics, Computer Science, and Engineering*
135 V. Berthé and M. Rigo (eds.) *Combinatorics, Automata and Number Theory*
136 A. Kristály, V. D. Rădulescu and C. Varga *Variational Principles in Mathematical Physics, Geometry, and Economics*
137 J. Berstel and C. Reutenauer *Noncommutative Rational Series with Applications*
138 B. Courcelle and J. Engelfriet *Graph Structure and Monadic Second-Order Logic*
139 M. Fiedler *Matrices and Graphs in Geometry*
140 N. Vakil *Real Analysis through Modern Infinitesimals*
141 R. B. Paris *Hadamard Expansions and Hyperasymptotic Evaluation*
142 Y. Crama and P. L. Hammer *Boolean Functions*
143 A. Arapostathis, V. S. Borkar and M. K. Ghosh *Ergodic Control of Diffusion Processes*
144 N. Caspard, B. Leclerc and B. Monjardet *Finite Ordered Sets*
145 D. Z. Arov and H. Dym *Bitangential Direct and Inverse Problems for Systems of Integral and Differential Equations*
146 G. Dassios *Ellipsoidal Harmonics*
147 L. W. Beineke and R. J. Wilson (eds.) with O. R. Oellermann *Topics in Structural Graph Theory*
148 L. Berlyand, A. G. Kolpakov and A. Novikov *Introduction to the Network Approximation Method for Materials Modeling*
149 M. Baake and U. Grimm *Aperiodic Order I: A Mathematical Invitation*
150 J. Borwein *et al. Lattice Sums Then and Now*
151 R. Schneider *Convex Bodies: The Brunn–Minkowski Theory (Second Edition)*
152 G. Da Prato and J. Zabczyk *Stochastic Equations in Infinite Dimensions (Second Edition)*
153 D. Hofmann, G. J. Seal and W. Tholen (eds.) *Monoidal Topology*
154 M. Cabrera García and Á. Rodríguez Palacios *Non-Associative Normed Algebras I: The Vidav–Palmer and Gelfand–Naimark Theorems*
155 C. F. Dunkl and Y. Xu *Orthogonal Polynomials of Several Variables (Second Edition)*
156 L. W. Beineke and R. J. Wilson (eds.) with B. Toft *Topics in Chromatic Graph Theory*
157 T. Mora *Solving Polynomial Equation Systems III: Algebraic Solving*
158 T. Mora *Solving Polynomial Equation Systems IV: Buchberger Theory and Beyond*
159 V. Berthé and M. Rigo (eds.) *Combinatorics, Words and Symbolic Dynamics*
160 B. Rubin *Introduction to Radon Transforms: With Elements of Fractional Calculus and Harmonic Analysis*
161 M. Ghergu and S. D. Taliaferro *Isolated Singularities in Partial Differential Inequalities*
162 G. Molica Bisci, V. D. Radulescu and R. Servadei *Variational Methods for Nonlocal Fractional Problems*
163 S. Wagon *The Banach–Tarski Paradox (Second Edition)*
164 K. Broughan *Equivalents of the Riemann Hypothesis I: Arithmetic Equivalents*
165 K. Broughan *Equivalents of the Riemann Hypothesis II: Analytic Equivalents*
166 M. Baake and U. Grimm (eds.) *Aperiodic Order II: Crystallography and Almost Periodicity*
167 M. Cabrera García and Á. Rodríguez Palacios *Non-Associative Normed Algebras II: Representation Theory and the Zel'manov Approach*
168 A. Yu. Khrennikov, S. V. Kozyrev and W. A. Zúñiga-Galindo *Ultrametric Pseudodifferential Equations and Applications*

Encyclopedia of Mathematics and its Applications

Mathematical Constants II

STEVEN R. FINCH
Massachusetts Institute of Technology

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108470599
DOI: 10.1017/9781316997741

First published 2019

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library.

ISBN 978-1-108-47059-9 Hardback

Contents

Preface

One reviewer for the first volume of *Mathematical Constants* described the book as "excellent bedtime reading" [1]. My aim here is similar to before: to gather far-flung ideas in one place, focusing on highly concrete, eminently computable results. These essays recount stories that are both successful (with depth of understanding) and tangible (in terms of numerical precision). Much mathematical research these days is necessarily abstract and qualitative, due to the enormous difficulty of the issues under consideration. Here I direct the spotlight to those rare cases when quantitative exactness still appears to be pertinent. My words from fifteen years ago (concerning purpose and scope) apply as well now as then.

A sample problem serves to illustrate my endeavor. While discussing the mass M_n of all nonisomorphic Type I inner product modules of rank n, Milnor & Husemoller [2] gave a plot of M_n on a logarithmic scale for $1 \leq n \leq 30$. They remarked that M_n is asymptotic to $C \cdot F(n)$ as $n \to \infty$, describing the function $F(n)$ exactly, but reporting only that "the constant C is approximately 0.705". Unraveling this enigma – what is the precise nature of C? – is captivating to me. Understandably this question was incidental to the purposes of [2]; it is, however, central here to me [3]. The answer involves a quantity [4] discovered in 1860, as well as something else.

This volume is dedicated to the memory of Philippe Flajolet, a fearless leader and inspiring mentor. It is also a tribute to my parents, Charles Richard Finch and Shirley Peery Finch, and to my siblings, Valerie Jean Bridge, Gregory Charles Finch and William Robert Finch, with love and gratitude. I acknowledge a Book Fellowship from the Clay Mathematics Institute in 2004–2005, long before the magnitude of my present task became clear.

"Open the book at random", the aforementioned reviewer wrote, evoking a few constants from many across the canvas, seedlings drawn from a vast forest. Read, learn, wander, reflect, ... "and so on into the night".

[1] Philip J. Davis, Constants in the universe: their validation, their compilation, and their mystique, *SIAM News*, v. 37 (April 2004) n. 3.
[2] J. Milnor and D. Husemoller, *Symmetric Bilinear Forms*, Springer-Verlag, 1973, p. 50; MR0506372 (58 #22129).
[3] S. R. Finch, Minkowski-Siegel mass constants, *this volume*, §5.6.
[4] S. R. Finch, Glaisher-Kinkelin constant, *first volume*, pp. 135–145.

Notation

$\lfloor x \rfloor$	*floor function:* largest integer $\leq x$
$\lceil x \rceil$	*ceiling function:* smallest integer $\geq x$
$\{x\}$	*fractional part:* $x - \lfloor x \rfloor$
$\ln x$	*natural logarithm:* $\log_e x$
$\binom{n}{k}$	*binomial coefficient:* $\dfrac{n!}{k!(n-k)!}$
$b_0 + \dfrac{a_1\vert}{\vert b_1} + \dfrac{a_2\vert}{\vert b_2} + \dfrac{a_3\vert}{\vert b_3} + \cdots$	*continued fraction:* $b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \cdots}}}$
$f(x) = O(g(x))$	*big O:* $\vert f(x)/g(x)\vert$ is bounded from above as $x \to x_0$
$f(x) = o(g(x))$	*little o:* $f(x)/g(x) \to 0$ as $x \to x_0$
$f(x) \sim g(x)$	*asymptotic equivalence:* $f(x)/g(x) \to 1$ as $x \to x_0$
$\sum_p$	summation over all prime numbers $p = 2, 3, 5, 7, 11, \ldots$ (only when the letter p is used)
$\prod_p$	same as $\sum_p$, with addition replaced by multiplication
$f(x)^n$	*power*: $(f(x))^n$, where n is an integer
$f^n(x)$	*iterate*: $\underbrace{f(f(\cdots f}_{n \text{ times}}(x)\cdots))$, where $n \geq 0$ is an integer
$\mathrm{erf}(x) = \dfrac{2}{\sqrt{\pi}} \displaystyle\int_0^x \exp(-t^2)\,dt$	*error function*: same as $1 - \mathrm{erfc}(x)$

$\Phi(x) = \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2}$ *standard normal distribution function*

${}_pF_q(\cdot;\ \cdot;\ z)$ *generalized hypergeometric function*:

$$
{}_pF_q(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z)
= \frac{\Gamma(b_1)\Gamma(b_2)\cdots\Gamma(b_q)}{\Gamma(a_1)\Gamma(a_2)\cdots\Gamma(a_p)} \sum_{k=0}^{\infty} \frac{\Gamma(a_1+k)\Gamma(a_2+k)\cdots\Gamma(a_p+k)}{\Gamma(b_1+k)\Gamma(b_2+k)\cdots\Gamma(b_q+k)} \frac{z^k}{k!}
$$

1

Number Theory and Combinatorics

1.1 Bipartite, k-Colorable and k-Colored Graphs

A labeled graph G is **bipartite** if its vertex set V can be partitioned into two disjoint subsets A and B, $V = A \cup B$, such that every edge of G is of the form (a, b), where $a \in A$ and $b \in B$.

Let k be a positive integer and $K = \{1, 2, \ldots, k\}$. A labeled graph G is **k-colorable** if there exists a function $V \to K$ with the property that adjacent vertices must be colored differently. Clearly G is bipartite if and only if G is 2-colorable.

Define $c_{n,k}$ to be the number of k-colorable graphs with n vertices. We have $c_{n,1} = 1$ for $n \geq 1$ since a 1-colorable graph G cannot possess any edges. We also have $c_{1,k} = 1$ for $k \geq 1$, $c_{2,k} = 2$ for $k \geq 2$, $c_{3,2} = 7$ by Figure 1.1, $c_{3,3} = 8$, $c_{4,2} = 41$ by Figure 1.2, and $c_{4,3} = 63$. More generally, $c_{n,n-1} = 2^{n(n-1)/2} - 1$ since the total number of labeled graphs with n vertices is $2^{n(n-1)/2}$ and, of these, only the complete graph cannot be $(n-1)$-colored.

Does there exist a formula for $c_{n,k}$? The answer is yes if $k = 2$, but evidently no for $k \geq 3$. We will examine this issue momentarily, but first define a related notion.

A **k-colored graph** is a labeled k-colorable graph together with its coloring function. Let $\gamma_{n,k}$ be the number of k-colored graphs with n vertices. The point is that a k-colorable graph counts several times as a k-colored graph. Clearly $\gamma_{n,1} = 1$, $\gamma_{1,k} = k$, $\gamma_{2,2} = 6$ by Figure 1.3, $\gamma_{2,3} = 15$ by Figure 1.4, and $\gamma_{3,2} = 26$ by Figure 1.5.

When $k = 2$, the following formulas can be proved [1–3]:

$$\gamma_{n,2} = \sum_{j=0}^{n} \binom{n}{j} 2^{j(n-j)},$$

$$c_{n,2} = n! \cdot \left(\text{the } n^{\text{th}} \text{ degree Maclaurin series coefficient of } \sqrt{\Gamma(x)}\right),$$

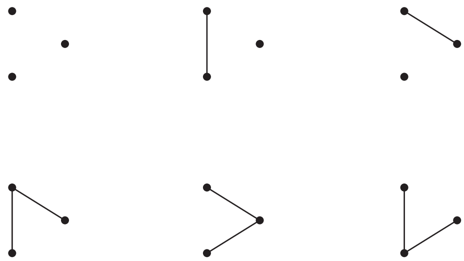

Figure 1.1 There are 7 labeled bipartite graphs with 3 vertices.

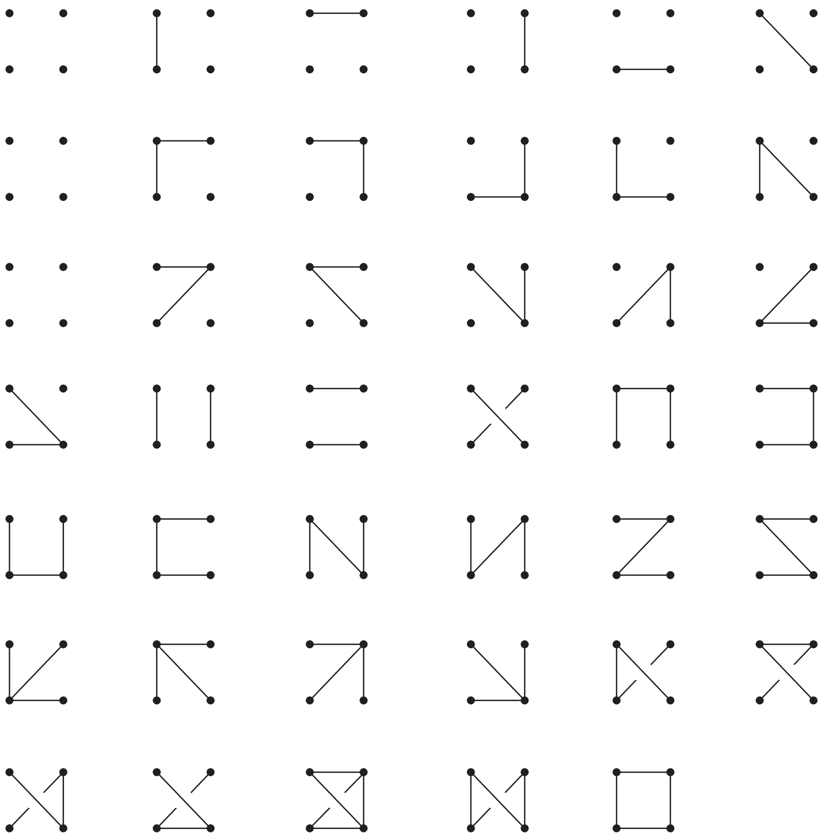

Figure 1.2 There are 41 labeled bipartite graphs with 4 vertices.

1 1   2 2   1 2   2 1   1----2   2----1

Figure 1.3 There are 6 labeled 2-colored graphs with 2 vertices.

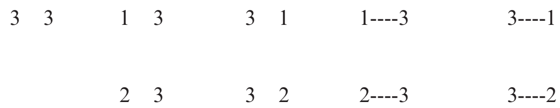

Figure 1.4 There are 15 labeled 3-colored graphs with 2 vertices (these 9 plus the preceding 6).

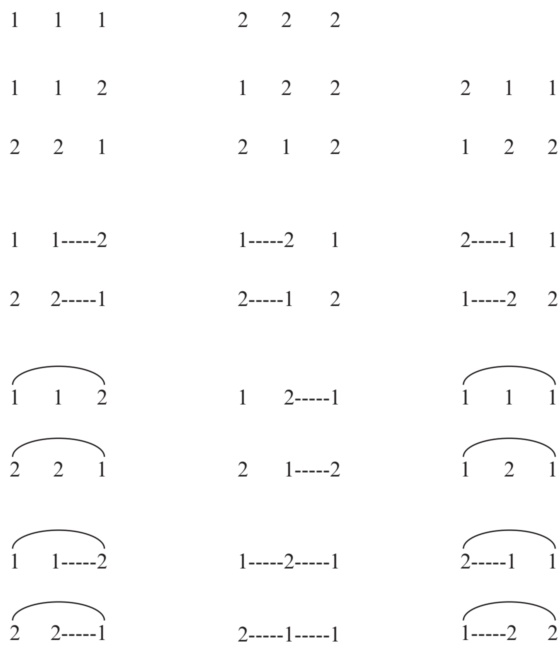

Figure 1.5 There are 26 labeled 2-colored graphs with 3 vertices.

where

$$\Gamma(x) = \sum_{i=0}^{\infty} \gamma_{i,2} \frac{x^i}{i!}.$$

For arbitrary k, we have the following recursion [4, 5]:

$$\gamma_{n,k} = \sum_{j=0}^{n} \binom{n}{j} 2^{j(n-j)} \gamma_{j,k-1}$$

with initial conditions $\gamma_{0,k} = 1$ and $\gamma_{n,0} = 0$ for $n \geq 1$. Alternatively, we have a closed-form expression involving multinomial coefficients:

$$\gamma_{n,k} = \sum_{N} \binom{n}{n_1, n_2, \ldots, n_k} 2^{\frac{1}{2}\left(n^2 - n_1^2 - n_2^2 - \cdots - n_k^2\right)}$$

where the summation is over all nonnegative integer k-vectors $N=(n_1,n_2,\ldots,n_k)$ satisfying $n_1+n_2+\cdots+n_k=n$. There is, however, no known analogous formula for $c_{n,k}$ when $k\geq 3$.

Computations show that [4, 6]

$$\{\gamma_{n,2}\}_{n=1}^{\infty}=\{2,6,26,162,1442,18306,330626,8488962\ldots\},$$

$$\{c_{n,2}\}_{n=1}^{\infty}=\{1,2,7,41,376,5177,103237,2922446\ldots\}$$

and suggest that $\gamma_{n,2}/c_{n,2}\to 2$ as $n\to\infty$. We also have

$$\{\gamma_{n,3}\}_{n=1}^{\infty}=\{3,15,123,1635,35043,1206915,66622083,5884188675,\ldots\},$$

$$\{c_{n,3}\}_{n=1}^{\infty}=\{1,2,8,63,958,27554,\ldots\}$$

but there is insufficient data on $c_{n,3}$ to clearly suggest the asymptotic behavior of $\gamma_{n,3}/c_{n,3}$. Prömel & Steger [7], however, proved that

$$\lim_{n\to\infty}\frac{\gamma_{n,k}}{c_{n,k}}=k!$$

for each $k\geq 2$. In words, a random k-colorable graph is almost surely uniquely k-colorable (up to a permutation of colors). This is an important result since it allows us to utilize at least one term of the $\gamma_{n,k}$ asymptotics to estimate the growth of $c_{n,k}$.

We turn now to a result due to Wright [8–12]: if $n\equiv a \mod k$, where $0\leq a<k$, then

$$\gamma_{n,k}\sim C(k,a)\cdot 2^{\frac{1}{2}(1-\frac{1}{k})n^2}\cdot k^n\cdot\left(\frac{k}{\ln(2)\cdot n}\right)^{\frac{k-1}{2}}$$

as $n\to\infty$, where $C(k,a)$ is a constant that depends on n only via its residue modulo k. In fact,

$$C(k,a)=k^{\frac{1}{2}}\cdot(\ln(2))^{\frac{k-1}{2}}\cdot(2\pi)^{-\frac{k-1}{2}}\cdot L_k(a)$$

and the infinite series $L_k(a)$ will be defined for $k=2$, 3 and 4 shortly.

1.1.1 *2-Colored Graph Asymptotics*

To characterize the growth of $\gamma_{n,k}$, by the above, it is sufficient to determine $C(k,a)$ for each $0\leq a<k$. We have here

$$\begin{aligned}L_2(a)&=\sum_{r=-\infty}^{\infty}2^{-\frac{1}{2}r^2-\frac{1}{2}(a-r)^2+\frac{1}{4}a^2}\\&=\sum_{r=-\infty}^{\infty}2^{-\frac{1}{4}(a-2r)^2}=\begin{cases}2.1289368272\ldots & \text{if } a=0,\\ 2.1289312505\ldots & \text{if } a=1.\end{cases}\end{aligned}$$

These two constants also appear with regard to the asymptotic enumeration of partially ordered sets [13] and of linear subspaces of $\mathbb{F}_2^n$ [14], where $\mathbb{F}_2$ is the binary field with arithmetic modulo 2. Therefore

$$C(2,a)=\begin{cases} 1.0000013097\ldots=1+\varepsilon & \text{if } a=0, \\ 0.9999986902\ldots=1-\varepsilon & \text{if } a=1 \end{cases}$$

where $\varepsilon=1.3097396978\ldots\times 10^{-6}$. In fact, all of the constants $C(k,a)$ we examine are close to 1; thus we shall focus on difference with 1 henceforth.

1.1.2 *3-Colored Graph Asymptotics*

We have here

$$\begin{aligned} L_3(a) &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} 2^{-\frac{1}{2}r^2-\frac{1}{2}s^2-\frac{1}{2}(a-r-s)^2+\frac{1}{6}a^2} \\ &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty} 2^{-\frac{1}{3}\left(a^2-3ar+3r^2-3as+3rs+3s^2\right)} \end{aligned}$$

and therefore

$$C(3,a)=\begin{cases} 1+2\varepsilon & \text{if } a=0, \\ 1-\varepsilon & \text{if } a=1 \text{ or } 2 \end{cases}$$

where $\varepsilon=1.7060611047...\times 10^{-8}$.

1.1.3 *4-Colored Graph Asymptotics*

All planar graphs are 4-colorable by the famous Four Color Theorem. We have here [4, 6]

$$\{\gamma_{n,4}\}_{n=1}^{\infty}=\{4,28,340,7108,254404,15531268,1613235460,284556079108,\ldots\},$$

$$\{c_{n,4}\}_{n=1}^{\infty}=\{1,2,8,64,1023,32596,\ldots\},$$

$$\begin{aligned} L_4(a) &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} 2^{-\frac{1}{2}r^2-\frac{1}{2}s^2-\frac{1}{2}t^2-\frac{1}{2}(a-r-s-t)^2+\frac{1}{8}a^2} \\ &= \sum_{r=-\infty}^{\infty}\sum_{s=-\infty}^{\infty}\sum_{t=-\infty}^{\infty} 2^{-\frac{1}{8}\left(3a^2-8ar+8r^2-8as+8rs+8s^2-8at+8rt+8st+8t^2\right)} \end{aligned}$$

and therefore

$$C(4,a)=\begin{cases} 1+\delta & \text{if } a=0, \\ 1-\varepsilon & \text{if } a=1 \text{ or } 3, \\ 1-\delta+2\varepsilon & \text{if } a=2, \end{cases}$$

where $\delta=4.2421496651\ldots\times 10^{-9}$ and $\varepsilon=2.5731271141...\times 10^{-12}$. A simple relationship between δ and ε is not apparent.

Higher-order asymptotics for $\gamma_{n,k}$ are possible, due to Wright [8]; the corresponding constants await study. Observe that terms beyond the first need not necessarily apply for $c_{n,k}$.

A random k-colorable graph is almost surely connected [10, 12, 15] and is almost surely k-chromatic (meaning that $k-1$ colors will not suffice to color all n vertices). The asymptotics discussed above therefore apply to these important subclasses as well.

Enumerating unlabeled k-colorable graphs (that is, non-isomorphic types of labeled k-colorable graphs) is also a difficult computational problem [16]. A general result due to Prömel [17] provides that $c_{n,k}/n!$ is the associated asymptotic formula.

[1] H. S. Wilf, *generatingfunctionology*, 2nd edn., Academic Press, 1994, pp. 86–89; MR1277813 (95a:05002).
[2] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009, p. 138; MR2483235 (2010h:05005).
[3] D. A. Klarner, The number of graded partially ordered sets, *J. Combin. Theory* 6 (1969) 12–19; MR0236035 (38 #4333).
[4] R. C. Read, The number of k-colored graphs on labelled nodes, *Canad. J. Math.* 12 (1960) 410–414; MR0114775 (22 #5594).
[5] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, 1973, pp. 16–18; MR0357214 (50 #9682).
[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000684, A000685, A000686, A047863, A047864, A084279, and A084280.
[7] H. J. Prömel and A. Steger, Random l-colorable graphs, *Random Structures Algorithms* 6 (1995) 21–37; MR1368832 (96j:05095).
[8] E. M. Wright, Counting coloured graphs, *Canad. J. Math.* 13 (1961) 683–693; MR0133252 (24 #A3086).
[9] E. M. Wright, Counting coloured graphs. II, *Canad. J. Math.* 16 (1964) 128–135; MR0177913 (31 #2171).
[10] R. C. Read and E. M. Wright, Coloured graphs: A correction and extension, *Canad. J. Math.* 22 (1970) 594–596; MR0263695 (41 #8296).
[11] E. M. Wright, Counting coloured graphs. III, *Canad. J. Math.* 24 (1972) 82–89; MR0289368 (44 #6559).
[12] F. Harary and R. W. Robinson, Labeled bipartite blocks, *Canad. J. Math.* 31 (1979) 60–68; MR0518706 (80b:05036).
[13] S. R. Finch, Lengyel's constant: Chains in the subset lattice of S, *first volume*, pp. 317–318.
[14] S. R. Finch, Transitive relations, topologies and partial orders, *this volume*, §1.2.
[15] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001832, A084283, and A084284.
[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A033995, A076315, and A076316.
[17] H. J. Prömel, Counting unlabeled structures, *J. Combin. Theory Ser. A* 44 (1987) 83–93; MR0871390 (87m:05014).

1.2 Transitive Relations, Topologies and Partial Orders

Let S be a set with n elements. A subset R of $S \times S$ is a **binary relation** (or **relation**) on S. The number of relations on S is 2^{n^2}. Equivalently, there are 2^{n^2} labeled bipartite graphs on $2n$ vertices, assuming the bipartition is fixed and equitable.

A relation R on S is **reflexive** if for all $x \in S$, we have $(x, x) \in R$. The number of reflexive relations on S is $2^{n(n-1)}$.

A relation R on S is **antisymmetric** if for all $x, y \in S$, the conditions $(x, y) \in R$ and $(y, x) \in R$ imply that $x = y$. The number of antisymmetric relations on S is $2^n \cdot 3^{n(n-1)/2}$.

A relation R on S is **transitive** if for all $x, y, z \in S$, the conditions $(x, y) \in R$ and $(y, z) \in R$ imply that $(x, z) \in R$. There is no known general formula for the number T_n of transitive relations on S. It is surprising that such a simply-stated counting problem remains unsolved [1–6].

A **topology** on S is a collection Σ of subsets of S that satisfy the following axioms:

- $\emptyset \in \Sigma$ and $S \in \Sigma$
- the union of any two sets in Σ is in Σ
- the intersection of any two sets in Σ is in Σ.

Note that since S is finite, our phrasing of the second axiom is correct. No one knows a general formula for the number U_n of topologies on S. Also, a topology on S is a **T0 topology** if it additionally satisfies a (weak) separation axiom:

- for any pair of distinct points in S, there is a set in Σ containing one point but not the other.

Again, no one knows a general formula for the number V_n of T0 topologies [7].

A **quasi-order** on S is a relation that is both reflexive and transitive. Let Q_n denote the number of such relations. Other uses of the phrase “quasi-order” exist and so care must be taken when reviewing the literature. There is a one-to-one correspondence between the topologies on S and the quasi-orders on S; hence $Q_n = U_n$.

A **partial order** on S is a quasi-order that is antisymmetric as well. Let P_n denote the number of such relations. We usually write $x \le y$ if $(x, y) \in R$ and, moreover, $x < y$ if $x \ne y$. There is a one-to-one correspondence between the T0 topologies on S and the partial orders on S; hence $P_n = V_n$.

Further connections between P_n and Q_n, and between P_n and T_n, can be expressed in terms of Stirling numbers of the second kind [1, 8]:

$$Q_n = \sum_{k=1}^{n} S_{n,k} P_k, \quad T_n = \sum_{k=1}^{n} \left(\sum_{j=0}^{k} \binom{n}{j} S_{n-j,k-j} \right) P_k$$

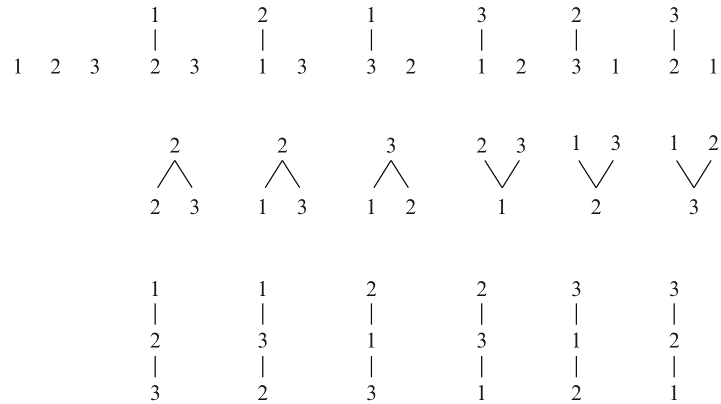

Figure 1.6 There are 19 labeled posets with 3 elements, that is, $P_3 = 19$.

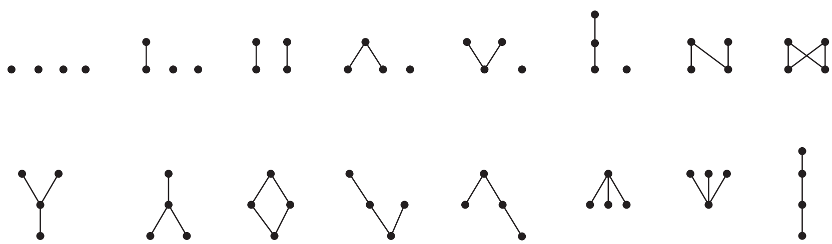

Figure 1.7 There are 16 unlabeled posets with 4 elements, that is, $p_4 = 16$.

and hence [9, 10]

$$Q_n \sim P_n, \quad T_n \sim 2^n P_n$$

as $n \to \infty$. It is therefore sufficient to focus on just one of these sequences; we choose $\{P_n\}$, which enumerates labeled posets (see Figure 1.6) as opposed to $\{p_n\}$, which enumerates unlabeled posets (see Figure 1.7). The existence of an edge (x, y) in any of the graphs pictured here indicates that $x < y$ and y is drawn above x.

Even though a closed-form expression for P_n is unknown, progress has been made in understanding the asymptotics of

$$\{P_n\}_{n=1}^{\infty} = \{1, 3, 19, 219, 4231, 130023, 6129859, 431723379, \ldots\}.$$

Kleitman & Rothschild [11] deduced that

$$\frac{\ln(P_n)}{\ln(2)} = \frac{n^2}{4} + O\left(n^{\frac{3}{2}} \ln(n)\right)$$

and later sharpened this to [12]

$$\frac{\ln(P_n)}{\ln(2)} = \frac{n^2}{4} + \frac{3n}{2} + O\left(\ln(n)\right).$$

Building on their work, several authors [10, 13–16] gave the following improvement:

$$P_n \sim C_a \cdot \sqrt{\frac{2}{\pi}} \cdot 2^{\frac{n^2}{4}+\frac{3n}{2}+\frac{1}{4}} \cdot n^{-\frac{1}{2}}$$

where $n \equiv a \bmod 2$ and $a \in \{0,1\}$, and where

$$C_1 = \sum_{k=-\infty}^{\infty} 2^{-k^2} = 2.1289368272... = \pi \cdot (0.8058800428...) \cdot 2^{-\frac{1}{4}},$$

$$C_0 = \sum_{k=-\infty}^{\infty} 2^{-(k-\frac{1}{2})^2} = 2.1289312505... = \pi \cdot (0.8058779318...) \cdot 2^{-\frac{1}{4}}.$$

It is interesting that the constant depends on the parity of n.

The asymptotics of the unlabeled case [17, 18]:

$$\{p_n\}_{n=1}^{\infty} = \{1, 2, 5, 16, 63, 318, 2045, 16999, \ldots\}$$

turn out to satisfy

$$p_n \sim \frac{P_n}{n!} \sim C_a \cdot \frac{1}{\pi} \cdot 2^{\frac{n^2}{4}+\frac{3n}{2}+\frac{1}{4}} \cdot e^n \cdot n^{-n-1}$$

thanks to a general result due to Prömel [19].

See [20, 21] for more appearances of the constants C_0 and C_1. It's believed that, for any asymptotic enumeration problem where a typical member is based on a bipartite graph, these constants are likely to occur. Alternative representations include [16, 22]:

$$C_1 = \sqrt{\frac{\pi}{\ln(2)}} \sum_{k=-\infty}^{\infty} \exp\left(\frac{-\pi^2}{\ln(2)} k^2\right), \quad C_0 = \sqrt{\frac{\pi}{\ln(2)}} \sum_{k=-\infty}^{\infty} (-1)^k \exp\left(\frac{-\pi^2}{\ln(2)} k^2\right)$$

from which the strict inequality $C_0 < C_1$ becomes obvious.

1.2.1 *Natural Partial Orders*

Consider the set $S = \{1, 2, \ldots n\}$ equipped with the usual total ordering $\leq$. A **natural partial order** $\preceq$ on S is a partial ordering that is compatible with $\leq$ (meaning that if $x \preceq y$, then $x \leq y$). This is equivalent to saying that $(S, \leq)$ is a **linear extension** of $(S, \preceq)$. Define σ_n to be the number of natural partial orders on S, then [23–25]

$$\{\sigma_n\}_{n=1}^{\infty} = \{1, 2, 7, 40, 357, 4824, 96428, 2800472, \ldots\}$$

(see Figure 1.8).

Brightwell, Prömel & Steger [16] proved that

$$\sigma_n \sim \begin{cases} \frac{1}{2}\eta^2 \cdot C_1 \cdot 2^{\frac{n^2}{4}} \cdot n = (12.7636300229...) \cdot 2^{\frac{n^2}{4}} \cdot n & \text{if } n \text{ is even,} \\ \frac{1}{2}\eta^2 \cdot C_0 \cdot 2^{\frac{n^2}{4}} \cdot n = (12.7635965889...) \cdot 2^{\frac{n^2}{4}} \cdot n & \text{if } n \text{ is odd,} \end{cases}$$

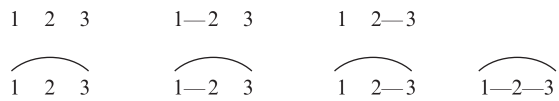

Figure 1.8 There are 7 natural partial orders on {1, 2, 3}, that is, $\sigma_3 = 7$.

where

$$\eta = \prod_{j=1}^{\infty} \left(1 - 2^{-j}\right)^{-1} = 3.4627466194...$$

is a digital search tree constant [26]. These constants also arise when determining the average number λ_n of linear extensions of S, where S is a *random* poset on n points [16, 27]:

$$\lambda_n \sim \begin{cases} \dfrac{\eta^2 C_1}{2^{5/4} C_0} \cdot \left(\frac{n}{2}\right)!^2 \cdot n \cdot 2^{-n/2} = (5.0414454338...) \cdot \left(\frac{n}{2}\right)!^2 \cdot n \cdot 2^{-n/2}, \\[2ex] \dfrac{\eta^2 C_0}{2^{5/4} C_1} \cdot \left(\frac{n-1}{2}\right)! \cdot \left(\frac{n+1}{2}\right)! \cdot n \cdot 2^{-n/2} = (5.0414190220...) \cdot \left(\frac{n-1}{2}\right)! \\ \qquad \cdot \left(\frac{n+1}{2}\right)! \cdot n \cdot 2^{-n/2} \end{cases}$$

when n is even, respectively, n is odd.

Consider instead the set S of all 2^n subsets of $\{1, 2, \ldots, n\}$ equipped with the usual partial ordering $\subseteq$. Define τ_n in a manner analogous to σ_n. We observe that $\lambda_n \cdot P_n \sim n! \cdot \sigma_n$ and wonder what the corresponding asymptotics for τ_n might be.

1.2.2 *Evolving Posets*

An interesting variation is as follows. What is the number N_ρ of partial orders on S with the property that a specified fraction ρ of the $n(n-1)/2$ pairs of distinct points are comparable? (If necessary, $\rho n(n-1)/2$ is rounded to the nearest integer.) Dhar [28, 29] investigated this question in the limit as $n \to \infty$ and proposed a lattice gas model (with infinitely many phase transitions) based on the evolution of N_ρ as ρ increases. A highly intricate analysis of Dhar's model was completed in [30–32].

[1] J. Klaska, Transitivity and partial order, *Math. Bohemica* 122 (1997) 75–82; MR1446401 (98c:05006).
[2] J. Klaska, History of the number of finite posets, *Acta Univ. Mathaei Belii Nat. Sci. Ser. Ser. Math.* 5 (1997) 73–84; MR1618881 (99k:06001).
[3] M. Erné and K. Stege, Counting finite posets and topologies, *Order* 8 (1991) 247–265; MR1154928 (93b:06004).
[4] P. Renteln, Geometrical approaches to the enumeration of finite posets: An introductory survey, *Nieuw Arch. Wisk.* 14 (1996) 349–371; MR1430049 (98g:06010).

[5] M. H. El-Zahar, Enumeration of ordered sets, *Algorithms and Order*, Proc. NATO Advanced Study Institute, Ottawa, 1987, ed. I. Rival, Kluwer, 1989, pp. 327–352; MR1037788 (90m:06005).

[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000112, A000798, A001035, A001930, and A006905.

[7] P. J. Cameron, *Combinatorics: Topics, Techniques, Algorithms*, Cambridge Univ. Press, 1994, pp. 37–38; MR1311922 (95j:05002).

[8] S. R. Finch, Lengyel's constant: Stirling partition numbers, *first volume*, pp. 316–317.

[9] M. Erné, Struktur- und Anzahlformeln für Topologien auf endlichen Mengen, *Manuscripta Math.* 11 (1974) 221–259; MR0360300 (50 #12750).

[10] K. H. Kim and F. W. Roush, Posets and finite topologies, *Pure Appl. Math. Sci.* 14 (1981) 9–22; MR0613626 (82f:54004).

[11] D. J. Kleitman and B. L. Rothschild, The number of finite topologies, *Proc. Amer. Math. Soc.* 25 (1970) 276–282; MR0253944 (40 #7157).

[12] D. J. Kleitman and B. L. Rothschild, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.* 205 (1975) 205–220; MR0369090 (51 #5326).

[13] J. L. Davison, Asymptotic enumeration of partial orders, *Proc. 17th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1986, ed. F. Hoffman, R. C. Mullin, R. G. Stanton and K. Brooks Reid, Congr. Numer. 53, Utilitas Math., 1986, pp. 277–286; MR0885256 (88c:06001).

[14] K. H. Kim, *Boolean Matrix Theory and Applications*, Dekker, 1982, pp. 141–149; MR0655414 (84a:15001).

[15] G. Brightwell, Models of random partial orders, *Surveys in Combinatorics, 1993*, Proc. 14th British Combinatorial conf., Keele, 1993, ed. K. Walker, Cambridge Univ. Press, 1993, pp. 53–83; MR1239232 (94i:06003).

[16] G. Brightwell, H. J. Prömel, and A. Steger, The average number of linear extensions of a partial order, *J. Combin. Theory Ser. A* 73 (1996) 193–206; MR1370128 (97h:06005).

[17] R. P. Stanley, *Enumerative Combinatorics*, v 1., Wadsworth & Brooks/Cole, 1986, pp. 96–154; MR1442260 (98a:05001).

[18] B. S. W. Schröder, *Ordered Sets: An Introduction*, Birkhäuser, 2003, pp. 277–281.

[19] H. J. Prömel, Counting unlabeled structures, *J. Combin. Theory Ser. A* 44 (1987) 83–93; MR0871390 (87m:05014).

[20] S. R. Finch, Lengyel's constant: Chains in the subset lattice of *S*, *first volume*, pp. 317–318.

[21] S. R. Finch, Bipartite, k-colorable and k-colored graphs, *this volume*, §1.1.

[22] W. Feller, *An Introduction to Probability Theory and Its Applications*, v. II, 2nd ed., Wiley, 1971, pp. 342, 628–633; MR0270403 (42 #5292).

[23] S. P. Avann, The lattice of natural partial orders, *Aequationes Math.* 8 (1972) 95–102; MR0311519 (47 #81).

[24] M. Erné, The number of partially ordered sets with more points than incomparable pairs, *Discrete Math.* 105 (1992) 49–60; MR1180192 (93g:05011).

[25] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006455.

[26] S. R. Finch, Digital search tree constants, *first volume*, pp. 351–361.

[27] G. Brightwell, Linear extensions of random orders, *Discrete Math.* 125 (1994) 87–96; MR1263735 (95i:06005).

[28] D. Dhar, Entropy and phase transitions in partially ordered sets, *J. Math. Phys.* 19 (1978) 1711–1713; MR0503271 (58 #20062).

[29] D. Dhar, Asymptotic enumeration of partially ordered sets, *Pacific J. Math.* 90 (1980) 299–305; MR0600632 (82c:06004).

[30] H. J. Prömel, A. Steger, and A. Taraz, Phase transitions in the evolution of partial orders, *J. Combin. Theory Ser. A* 94 (2001) 230–275; MR1825789 (2002e:06004).
[31] H. J. Prömel, A. Steger, and A. Taraz, Counting partial orders with a fixed number of comparable pairs, *Combin. Probab. Comput.* 10 (2001) 159–177; MR1833068 (2002e:06003).
[32] H. J. Prömel, A. Steger, and A. Taraz, Asymptotic enumeration, global structure, and constrained evolution, *Discrete Math.* 229 (2001) 213–233; MR1815608 (2001m:05237).

1.3 Series-Parallel Networks

Series and parallel connections are usually first encountered in the study of electrical circuits. Our approach is to first examine a relevant class of partially ordered sets (posets) and then to define series-parallel networks by analogy [1]. Interesting asymptotic constants appear everywhere, similar to those associated with counting various species of trees [2]. We also talk briefly about the enumeration of Boolean (or switching) functions under different notions of equivalence.

1.3.1 *Series-Parallel Posets*

We introduce two procedures for combining two posets $(S, \leq)$ and $(S', \leq)$ to obtain a new poset, assuming that $S \cap S' = \emptyset$:

- the **disjoint sum** $S \oplus S'$ is the poset on $S \cup S'$ such that $x \leq y$ in $S \oplus S'$ if either $x, y \in S$ and $x \leq y$ in S, or $x, y \in S'$ and $x \leq y$ in S'
- the **linear product** $S \odot S'$ is the poset on $S \cup S'$ such that $x \leq y$ in $S \odot S'$ if $x, y \in S$ and $x \leq y$ in S, or $x, y \in S'$ and $x \leq y$ in S', or $x \in S$ and $y \in S'$.

Clearly $\oplus$ is commutative but $\odot$ is not. A **series-parallel poset** is one that can be recursively constructed by applying the operations of disjoint sum and linear product, starting with a single point [3].

Define a poset to be **N-free** if there is no subset $\{a, b, c, d\}$ whose only nontrivial relations are given by

$$a < c, \quad a < d, \quad b < d.$$

It can be proved that a finite poset is series-parallel if and only if it is N-free [4–6]. Hence there are 15 series-parallel posets with 4 points (see the 16 posets in Figure 1.7 of [7] and eliminate the poset that looks like an "N").

There are two cases we shall consider. The number f_n of unlabeled series-parallel posets with n points has (ordinary) generating function [3, 8–10]

$$\begin{aligned} F(x) &= \sum_{n=0}^{\infty} f_n x^n = 1 + x + 2x^2 + 5x^3 + 15x^4 + 48x^5 + 167x^6 \\ &\quad + 602x^7 + 2256x^8 + \cdots \end{aligned}$$

which satisfies the functional equation

$$F(x)=\exp\left[\sum_{k=1}^{\infty}\frac{1}{k}\left(F(x^k)+\frac{1}{F(x^k)}+x^k-2\right)\right].$$

Alternatively, if the sequence $\{\hat{f}_n\}$ is defined by $1/F(x)=\sum_{n=0}^{\infty}\hat{f}_n x^n$, then

$$F(x)=\prod_{j=1}^{\infty}\left(1-x^j\right)^{-(f_j+\hat{f}_j+\delta_{j,1})}$$

where $\delta_{j,k}=1$ when $j=k$ and $\delta_{j,k}=0$ otherwise. Using such properties, it follows that

$$f_n\sim\beta\cdot n^{-3/2}\cdot\alpha^{-n},$$

where $\alpha=0.2163804273...$ is the unique positive root of $F(x)=\varphi$ and φ is the Golden mean, and

$$\beta=\sqrt{\frac{1}{(3\sqrt{5}-5)\pi}\left[\frac{\alpha}{1-\alpha}+\sum_{i=2}^{\infty}\alpha^i F'(\alpha^i)\left(1-\frac{1}{F(\alpha^i)^2}\right)\right]}=0.2291846208....$$

The number g_n of labeled series-parallel posets with n points has (exponential) generating function [1, 3, 8, 10]

$$\begin{aligned}G(x)&=\sum_{n=1}^{\infty}\frac{g_n}{n!}x^n=x+\frac{3}{2!}x^2+\frac{19}{3!}x^3+\frac{195}{4!}x^4+\frac{2791}{5!}x^5\\&\quad+\frac{51303}{6!}x^6+\frac{1152019}{7!}x^7+\cdots\\&=\left(\ln(1+x)-\frac{x^2}{1+x}\right)^{\langle-1\rangle}=\left(\sum_{k=1}^{\infty}(-1)^{k+1}\frac{k+1}{k}x^k\right)^{\langle-1\rangle}\end{aligned}$$

where the notation $P(x)^{\langle-1\rangle}$ denotes the reversion of the power series $P(x)$. Well-established theory [11, 12] gives

$$g_n\sim\eta\cdot n!\cdot n^{-3/2}\cdot\xi^{-n}$$

where $\xi=\ln(\varphi)-2\varphi+3=0.2451438475...$ and

$$\eta=\sqrt{\frac{\xi}{2\sqrt{5}(2-\varphi)\pi}}=0.2137301074....$$

Now let us define an equivalence relation on the set of series-parallel posets with n points, induced simply by declaring $S\odot S'$ and $S'\odot S$ to be equivalent. (See Figure 1.9.) The equivalence classes correspond to what are called

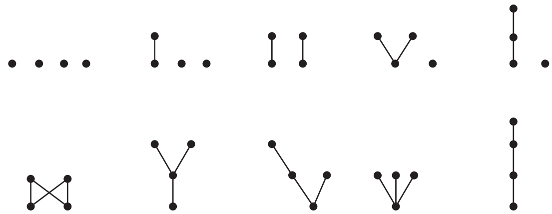

Figure 1.9 There are 10 non-equivalent (unlabeled) series-parallel posets with 4 points. Note the analogy with Figure 1.10.

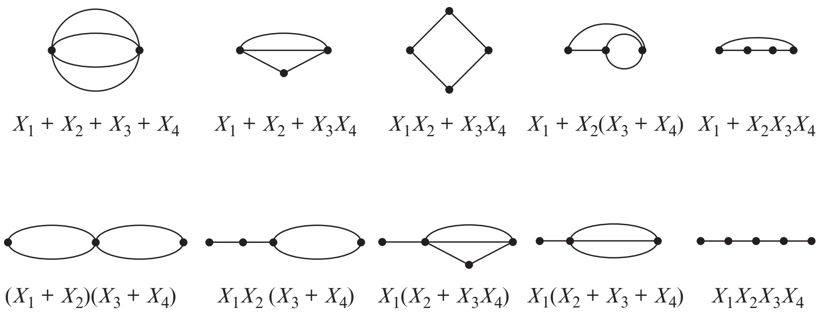

Figure 1.10 There are 10 unlabeled series-parallel networks with 4 edges, that is, $u_4 = 10$. The "essentially parallel" networks constitute the first row and the "essentially series" networks constitute the second row.

two-terminal series-parallel networks with n edges [13–19], with the understanding that

- points of a poset are mapped in a one-to-one manner to edges of the corresponding network
- two points of the poset are comparable if and only if the analogous edges of the network are connected in series
- two points of the poset are incomparable if and only if the analogous edges of the network are connected in parallel.

(See Figures 1.10 and 1.11.) The leftmost and rightmost points are the terminals (two distinguished points playing a role similar to that of the root of a rooted tree). A network, however, is not necessarily a graph since it may possess multiple parallel edges. Observe that an interchange of parts of the network, either in series or in parallel, is immaterial. In other words, when we count series-parallel networks, our tally is unaffected by a permutation of variables in the indicated Boolean representations.

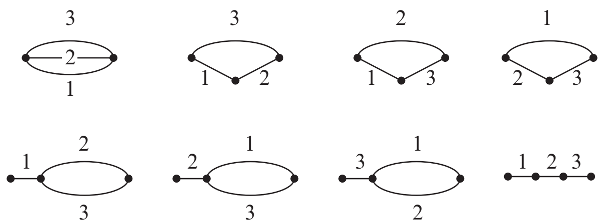

Figure 1.11 There are 8 labeled series-parallel networks with 3 edges, that is, $v_3 = 8$. The "essentially parallel" networks constitute the first row and the "essentially series" networks constitute the second row.

1.3.2 *Series-Parallel Networks*

The number u_n of unlabeled series-parallel networks with n edges has generating function [20]

$$U(x) = \sum_{n=0}^{\infty} u_n x^n = 1 + x + 2x^2 + 4x^3 + 10x^4 + 24x^5 + 66x^6 + 180x^7 + 522x^8 + \cdots$$

which satisfies the functional equation

$$U(x) = \exp\left[\sum_{k=1}^{\infty} \frac{1}{2k}\left(U(x^k) + x^k - 1\right)\right].$$

Alternatively, we have

$$U(x) = \prod_{j=1}^{\infty} \left(1 - x^j\right)^{-(u_j + \delta_{j,1})/2}.$$

Using these properties, it follows that [15, 21–23]

$$u_n \sim \lambda \cdot n^{-3/2} \cdot \kappa^{-n}$$

where $\kappa = 0.2808326669... = (3.5608393095...)^{-1}$ is the unique positive root of $U(x) = 2$ and

$$\lambda = \sqrt{\frac{1}{\pi}\left[\frac{\kappa}{1-\kappa} + \sum_{i=2}^{\infty} \kappa^i U'(\kappa^i)\right]} = 0.4127628892... = 2 \cdot (0.2063814446...).$$

This also gives the number of non-equivalent Boolean functions of n variables, built only with $+$ (disjunction) and $\cdot$ (conjunction).

The number v_n of labeled series-parallel networks with n edges has generating function [1, 24]

$$\begin{aligned} V(x) &= \sum_{n=1}^{\infty} \frac{v_n}{n!} x^n = x + \frac{2}{2!}x^2 + \frac{8}{3!}x^3 + \frac{52}{4!}x^4 + \frac{472}{5!}x^5 \\ &\quad + \frac{5504}{6!}x^6 + \frac{78416}{7!}x^7 + \cdots \\ &= (2\ln(1+x) - x)^{\langle -1 \rangle} = \left(\sum_{k=1}^{\infty} (-1)^{k+1} \frac{2}{k} x^k \right)^{\langle -1 \rangle}. \end{aligned}$$

By techniques similar to those used to analyze $\{g_n\}$, we have [21, 25]

$$v_n \sim \tau \cdot n! \cdot n^{-3/2} \cdot \sigma^{-n}$$

where $\sigma = 2\ln(2) - 1 = 0.3862943611... = (2.5886994495...)^{-1}$ and

$$\tau = \sqrt{\frac{\sigma}{\pi}} = 0.3506584008... = 2 \cdot (0.1753292004...).$$

Related work involves bracketing of n-symbol products [26] and phylogenetic trees [27].

1.3.3 *Series-Parallel Networks Without Multiple Parallel Edges*

If we prohibit multiple parallel edges, so that the networks under consideration are all graphs, different constants arise. (See Figure 1.12.)

The number q_n of such unlabeled series-parallel networks with n edges has generating function [28]

$$\begin{aligned} Q(x) &= \sum_{n=0}^{\infty} q_n x^n = 1 + x + x^2 + 2x^3 + 4x^4 + 8x^5 + 18x^6 + 40x^7 \\ &\quad + 94x^8 + 224x^9 + \cdots \end{aligned}$$

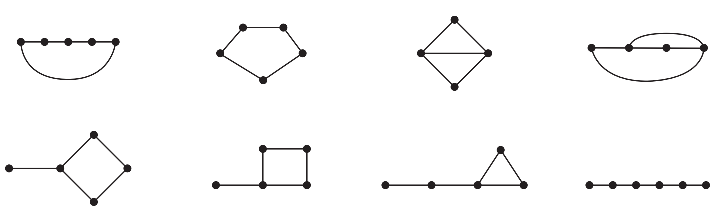

Figure 1.12 There are 8 unlabeled series-parallel networks with 5 edges that obey the prohibition against multiple parallel edges, that is, $q_5 = 8$. The "essentially parallel" networks constitute the first row and the "essentially series" networks constitute the second row.

which satisfies the functional equation

$$Q(x) = \exp\left[\sum_{k=1}^{\infty} \frac{1}{2k}\left(Q(x^k) - x^{2k} + x^k - 1\right)\right].$$

Alternatively, we have

$$Q(x) = \prod_{j=1}^{\infty} \left(1 - x^j\right)^{-(q_j + \delta_{j,1} - \delta_{j,2})/2}.$$

Using these properties, it follows that [21]

$$q_n \sim \nu \cdot n^{-3/2} \cdot \mu^{-n},$$

where $\mu = 0.3462834070...$ is the unique positive root of $Q(x) = 2$ and

$$\nu = \sqrt{\frac{1}{\pi}\left[\frac{\mu}{1+\mu} + \sum_{i=2}^{\infty} \mu^i Q'(\mu^i)\right]} = 0.3945042461... = 2 \cdot (0.1972521230...).$$

The number r_n of such labeled series-parallel networks with n edges has generating function [29]

$$\begin{aligned} R(x) &= \sum_{n=1}^{\infty} \frac{r_n}{n!} x^n = x + \frac{1}{2!}x^2 + \frac{4}{3!}x^3 + \frac{20}{4!}x^4 + \frac{156}{5!}x^5 + \frac{1472}{6!}x^6 + \frac{17396}{7!}x^7 + \cdots \\ &= \left((x+1)^2 \exp(-x) - 1\right)^{\langle -1 \rangle} = \left(\sum_{k=1}^{\infty} (-1)^k \frac{k^2 - 3k + 1}{k!} x^k\right)^{\langle -1 \rangle}. \end{aligned}$$

Proceeding as before, we have [21]

$$r_n \sim \omega \cdot n! \cdot n^{-3/2} \cdot \theta^{-n}$$

where $\theta = 4/e - 1 = 0.4715177646...$ and

$$\omega = \frac{1}{2}\sqrt{\frac{e\theta}{\pi}} = 0.3193679560... = 2 \cdot (0.1596839780...).$$

It follows that the probability that a random n-edge series-parallel network has no multiple parallel edges is asymptotically

$$\left(\frac{\nu}{\lambda}\right)\left(\frac{\kappa}{\mu}\right)^n = (0.9557648142...)(0.8109908278...)^n$$

if the network is unlabeled and

$$\left(\frac{\omega}{\tau}\right)\left(\frac{\sigma}{\theta}\right)^n = (0.9107665899...)(0.8192572794...)^n$$

if the network is labeled. More relevant material is covered in [21].

1.3.4 *Boolean Functions*

We have already enumerated the number u_n of distinct Boolean functions of n variables, built only with $+$ and $\cdot$, under the action of the symmetric group S_n.

Of course, the set of *all* Boolean functions also includes those involving complementation of variables ($\neg X$). Let us examine briefly this larger set [30, 31]. Define two Boolean functions to be **equivalent** if they are identical up to a bijective renaming of the variables. The number of equivalence classes in this case is asymptotically [32–34]

$$2^{2^n}/n!,$$

hence no new constants arise. Define two Boolean functions to be **congruent** if they are identical up to a bijective renaming of the variables and an additional complementation of some of the variables. The number of congruence classes is asymptotically

$$2^{2^n-n}/n!.$$

Other results of this kind are also known, but none contain new constants.

Let us return to our original set of Boolean functions of n variables and let $\mathbb{F}_2$ denote the binary field. S_n is a subgroup of the group T_n of invertible linear transformations $\mathbb{F}_2^n \to \mathbb{F}_2^n$, namely, the $n \times n$ matrices that have exactly one 1 in each row and each column. What can be said about the number $\tilde{u}_n$ of distinct Boolean functions, built only with $+$ and $\cdot$, under the action of the (larger) group T_n? Our experience with u_n leads us to conjecture that the asymptotics of $\tilde{u}_n$ will be quite interesting.

1.3.5 *Irreducible Posets*

Another unsolved problem involves the number a_n of unlabeled $(\oplus,\odot)$-irreducible posets with n points. Such a poset cannot be written as a disjoint union or a linear product of two non-empty posets. It is known that

$$\begin{aligned} A(x) &= \sum_{n=0}^{\infty} a_n x^n = x + x^4 + 12x^5 + 104x^6 + 956x^7 + 10037x^8 \\ &\quad + 126578x^9 + 1971005x^{10} + \cdots \end{aligned}$$

and, further,

$$P(x) = \exp\left[\sum_{k=1}^{\infty} \frac{1}{k}\left(P(x^k) + \frac{1}{P(x^k)} + A(x^k) - 2\right)\right]$$

where

$$\begin{aligned} P(x) &= \sum_{n=0}^{\infty} p_n x^n = 1 + x + 2x^2 + 5x^3 + 16x^4 + 63x^5 + 318x^6 \\ &\quad + 2045x^7 + 16999x^8 + \cdots \end{aligned}$$

is the generating function of (arbitrary) unlabeled posets [3, 7, 10]. What can be said about the asymptotics of a_n? Even a nice functional equation for $A(x)$ in and by itself is probably impossible.

Addendum Bodirsky, Giménez, Kang & Noy [35, 36] determined that the number of labeled series-parallel graphs on n vertices is asymptotically

$$(0.0076388...)n^{-5/2}(0.1102133...)^{-n}n!$$

as $n \to \infty$, but formulas underlying the constants are too elaborate to reproduce here. Special cases of such planar graphs [37] – connected and 2-connected – give rise to

$$(0.0067912...)n^{-5/2}(0.1102133...)^{-n}n!,$$

$$(0.0010131...)n^{-5/2}(0.1280038...)^{-n}n!$$

respectively. The distribution of the number of edges in a random graph with n vertices is asymptotically normal and the distribution of the number of connected components (minus one) is asymptotically Poisson, both with explicit computable parameters.

[1] R. P. Stanley, *Enumerative Combinatorics*, v 2., Cambridge Univ. Press, 1999, pp. 88–89, 133; MR1676282 (2000k:05026).
[2] S. R. Finch, Otter's tree enumeration constants, *first volume*, pp. 295–316.
[3] R. P. Stanley, Enumeration of posets generated by disjoint unions and ordinal sums, *Proc. Amer. Math. Soc.* 45 (1974) 295–299; MR0351928 (50 #4416).
[4] J. Valdes, R. E. Tarjan, and E. L. Lawler, The recognition of series parallel digraphs, *Proc. 11th ACM Symp. on Theory of Computing (STOC)*, Atlanta, ACM, 1979, pp. 1–12; *SIAM J. Comput.* 11 (1982) 298–313; MR0564616 (81d:68088) and MR0652904 (84d:68073)
[5] J. Grabowski, On partial languages, *Fundamenta Informaticae* 4 (1981) 427–498; MR0645249 (83i:68107).
[6] J. L. Gischer, The equational theory of pomsets, *Theoret. Comput. Sci.* 61 (1988) 199–224; MR0980242 (90b:68058).
[7] S. R. Finch, Transitive relations, topologies and partial orders, *this volume*, §1.2.
[8] M. H. El-Zahar, Enumeration of ordered sets, *Algorithms and Order*, Proc. NATO Advanced Study Institute, Ottawa, 1987, ed. I. Rival, Kluwer, 1989, pp. 327–352; MR1037788 (90m:06005).
[9] M. H. El-Zahar and S. M. Khamis, Enumeration of series-parallel posets according to heights, *J. Egyptian Math. Soc.* 8 (2000) 1–7; MR1798694 (2001m:05015).
[10] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003430, A003431, A046906, A053554.
[11] E. A. Bender, Asymptotic methods in enumeration, *SIAM Rev.* 16 (1974) 485–515; errata 18 (1976) 292; MR0376369 (51 #12545) and MR0437344 (55 #10276).
[12] A. Meir and J. W. Moon, On an asymptotic method in enumeration, *J. Combin. Theory Ser. A* 51 (1989) 77–89; erratum 52 (1989) 163; MR0993650 (90e:05007a) and MR1008169 (90e:05007b).

[13] P. A. MacMahon, Yoke-chains and multipartite compositions in connexion with the analytical forms called "trees", *Proc. London Math. Soc.* 22 (1891) 330–346; *Collected Papers*, v. 1, ed. G. E. Andrews, MIT Press, 1978, pp. 600–616.

[14] P. A. MacMahon, The combinations of resistances, *The Electrician* 28 (1892) 601–602; *Collected Papers*, v. 1, ed. G. E. Andrews, MIT Press, 1978, pp. 617–619.

[15] J. Riordan and C. E. Shannon, The number of two-terminal series-parallel networks, *J. Math. Phys. Mass. Inst. Tech.* 21 (1942) 83–93; *Claude Elwood Shannon: Collected Papers*, ed. N. J. A. Sloane and A. D. Wyner, IEEE Press, 1993, pp. 560–570; MR0007511 (4,151c).

[16] W. Knödel, Über Zerfällungen, *Monatsh. Math.* 55 (1951) 20–27; MR0040996 (12,778d).

[17] L. Carlitz and J. Riordan, The number of labeled two-terminal series-parallel networks, *Duke Math. J.* 23 (1956) 435–445; MR0078940 (18,3e).

[18] J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, 1958, pp. 139–143; MR0096594 (20 #3077).

[19] Z. A. Lomnicki, Two-terminal series-parallel networks, *Adv. Appl. Probab.* 4 (1972) 109–150; MR0319662 (47 #8204).

[20] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000084, A000669, A058964, A058965.

[21] J. W. Moon, Some enumerative results on series-parallel networks, *Random Graphs '85*, Proc. 1985 Poznan conf, ed. M. Karonski and Z. Palka, Annals Discrete Math. 33, North-Holland, 1987, pp. 199–226; MR0930492 (89h:05004).

[22] O. Golinelli, Asymptotic behavior of two-terminal series-parallel networks, cond-mat/9707023.

[23] E. M. Rains and N. J. A. Sloane, The series-parallel constant, unpublished note (2001).

[24] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000311, A006351.

[25] A. Meir and J. W. Moon, The asymptotic behaviour of coefficients of powers of certain generating functions, *European J. Combin.* 11 (1990) 581–587; MR1078714 (91m:05014).

[26] J. Riordan, The blossoming of Schröder's fourth problem, *Acta Math.* 137 (1976) 1–16; MR0429582 (55 #2593).

[27] L. R. Foulds and R. W. Robinson, Enumeration of phylogenetic trees without points of degree two, *Ars Combin.* 17A (1984) 169–183; MR0746182 (85f:05045).

[28] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A058385, A058386, A058387.

[29] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A058379, A058380, A058381.

[30] B. Borchert, D. Ranjan, and F. Stephan, On the computational complexity of some classical equivalence relations on Boolean functions, *Theory Comput. Syst.* 31 (1998) 679–693; MR1642457 (2000i:68055).

[31] P. Savický and A. R. Woods, The number of Boolean functions computed by formulas of a given size, *Random Structures Algorithms* 13 (1998) 349–382; MR1662790 (2000e:68077).

[32] M. A. Harrison, On asymptotic estimates in switching and automata theory, *J. ACM* 13 (1966) 151–157; MR0186504 (32 #3964).

[33] M. A. Harrison, Counting theorems and their applications to classification of switching functions, *Recent Developments in Switching Theory*, ed. A. Mukhopadhyay, Academic Press, 1971, pp. 85–120; MR0280279 (43 #6000).

[34] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000231, A003180, A000616.

[35] M. Bodirsky, O. Giménez, M. Kang and M. Noy, On the number of series parallel and outerplanar graphs, *Proc. European Conf. on Combinatorics, Graph Theory, and Applications (EuroComb05)*, DMTCS Proc. AE (2005) 383–388.
[36] M. Bodirsky, O. Giménez, M. Kang and M. Noy, Enumeration and limit laws for series-parallel graphs, *European J. Combin.* 28 (2007) 2091–2105; math.CO/0512435; MR2351512 (2008i:05087).
[37] S. R. Finch, Planar graph growth constants, *this volume*, §1.9.

1.4 Two Asymptotic Series

When enumerating trees [1, 2] or prime divisors [3, 4], the leading term of the corresponding asymptotic series is usually sufficient for practical purposes. Greater accuracy is possible by using several more terms, but the coefficients are not as widely known as one might expect. We briefly provide the formulas required to compute the required constants, as well as some theoretical background.

1.4.1 *Trees*

If T_n is the number of non-isomorphic rooted trees with n vertices, then [5]

$$T_n \sim r^{-n} n^{-3/2} \left(0.4399240125... + \frac{0.0441699018...}{n} + \frac{0.2216928059...}{n^2} + \frac{0.8676554908...}{n^3} + \cdots\right)$$

where $r = 0.3383218568...$ is the unique positive root of the equation $F(x,1) = 0$, where

$$F(x,y) = x \exp\left(y + \sum_{k=2}^{\infty} \frac{T(x^k)}{k}\right) - y$$

and $T(x) = \sum_{n=1}^{\infty} T_n x^n$ is the generating function for $\{T_n\}$. Let us denote the four numerical coefficients by $C_0/(2\sqrt{\pi})$, $C_1/(2\sqrt{\pi})$, $C_2/(2\sqrt{\pi})$ and $C_3/(2\sqrt{\pi})$. Exact formulas for these constants can be written in terms of the partial derivatives

$$F_{i,j} = \frac{\partial^{i+j}}{\partial x^i \partial y^j} F(x,y)\bigg|_{\substack{x=r\\y=1}}$$

via computer algebra. Note that $F_{0,0} = F_{0,1} = 0$,

$$1 = F_{0,2} = F_{0,3} = F_{0,4} = F_{0,5} = \cdots,$$

$$0 < F_{1,0} = F_{1,1} = F_{1,2} = F_{1,3} = F_{1,4} = \cdots,$$

and likewise $F_{i,j}=F_{i,0}$ for all $i\geq 2, j\geq 1$. We have

$$C_0=\sqrt{2\,r\,F_{1,0}},$$

$$C_1=\{9\,r\,F_{1,0}+r^2\,[-11\,F_{1,0}^2+9\,F_{2,0}\,]\}/\{12\,C_0\},$$

$$\begin{aligned} C_2 \;=\; &\{225\,r\,F_{1,0}^2+r^2\,[-990\,F_{1,0}^3+810\,F_{1,0}\,F_{2,0}] \\ &+r^3\,[769\,F_{1,0}^4-990\,F_{1,0}^2\,F_{2,0}-135\,F_{2,0}^2+360\,F_{1,0}\,F_{3,0}]\}/\{576\,F_{1,0}\,C_0\}, \end{aligned}$$

$$\begin{aligned} C_3 \;=\; &\{42525\,r\,F_{1,0}^3+r^2\,[-571725\,F_{1,0}^4+467775\,F_{1,0}^2\,F_{2,0}] \\ &+r^3\,[1211175\,F_{1,0}^5-1559250\,F_{1,0}^3\,F_{2,0}-212625\,F_{1,0}\,F_{2,0}^2+567000\,F_{1,0}^2\,F_{3,0}] \\ &+r^4\,[-680863\,F_{1,0}^6+1211175\,F_{1,0}^4\,F_{2,0}-155925\,F_{1,0}^2\,F_{2,0}^2+42525\,F_{2,0}^3 \\ &-415800\,F_{1,0}^3\,F_{3,0}-113400\,F_{1,0}\,F_{2,0}\,F_{3,0} \\ &+113400\,F_{1,0}^2\,F_{4,0}]\}/\{207360\,F_{1,0}^2\,C_0\}. \end{aligned}$$

The associated formula for t_n, the number of non-isomorphic free trees of order n, is [5]

$$t_n\sim r^{-n}n^{-5/2}\left(0.5349496061...+\frac{0.4853877311...}{n}+\frac{2.379745574...}{n^2}+\cdots\right)$$

where r is as before and the first numerical coefficient is simply $C_0^3/(4\sqrt{\pi})$. Exact formulas for the second and third coefficients are new:

$$\frac{C_0^2(C_0^3+30C_1)}{24\sqrt{\pi}},\quad \frac{C_0(C_0^6+35C_0^3C_1+210C_1^2+126C_0C_2)}{72\sqrt{\pi}}$$

and we wonder what the next few coefficients might look like.

Other varieties of trees examined in [5] include binary trees, identity trees and homeomorphically irreducible trees. Different functional equations apply in each case; for example, we have

$$F(x,y)=x+\frac{1}{2}\left(y^2+B(x^2)\right)-y$$

for the first variety, where $B(x)=\sum_{n=1}^{\infty}B_nx^n$ is the generating function for the number B_n of non-isomorphic rooted strongly binary trees with n leaves ($B_1=B_2=B_3=1$, $B_4=2$, $B_5=3$, ...). One obtains

$$B_n\sim\rho^{-n}n^{-3/2}\left(0.3187766259...+\frac{0.2038317427...}{n}+\frac{0.3682702316...}{n^2}+\frac{1.4768193666...}{n^3}+\cdots\right)$$

with $\rho=0.4026975036...$ as the radius of convergence. The details are omitted.

An intermediate step to studying $\{T_n\}$ involves the analysis of the series [6, 7]

$$\begin{aligned}T(x) &= \sum_{k=0}^{\infty} c_k (r-x)^{k/2} \\ &= 1 - (2.6811281472...)(r-x)^{1/2} + (2.3961493806...)(r-x) \\ &\quad - (1.4507456802...)(r-x)^{3/2} + (1.4447836810...)(r-x)^2 \\ &\quad - (5.1438071207...)(r-x)^{5/2} + \cdots\end{aligned}$$

which is valid as $x \to r^-$, where

$$c_0 = 1, \quad c_1 = -\sqrt{2\,F_{1,0}}, \quad c_2 = 2\,F_{1,0}/3,$$

$$c_3 = \left\{11\,F_{1,0}^2 - 9\,F_{2,0}\right\} / \{18\,c_1\}, \quad c_4 = \left\{43\,F_{1,0}^2 - 45\,F_{2,0}\right\} /135,$$

$$c_5 = \left\{769\,F_{1,0}^4 - 990\,F_{1,0}^2\,F_{2,0} - 135\,F_{2,0}^2 + 360\,F_{1,0}\,F_{3,0}\right\} / \{2160\,F_{1,0}\,c_1\}\,.$$

Note that $c_2 = c_1^2/3$ and $c_4 = (30c_1c_3 - c_1^4)/45$, while c_3 and c_5 cannot be algebraically represented in terms of preceding c_k values. Most of these results are new.

Likewise, in connection with $\{t_n\}$, we have [6, 7]

$$\begin{aligned}t(x) &= \sum_{k=0}^{\infty} d_k (r-x)^{k/2} \\ &= 0.5657439434... - (4.0484928944...)(r-x) - (6.4243835496...)(r-x)^{3/2} \\ &\quad -(5.5810996983...)(r-x)^2 + (7.3498535571...)(r-x)^{5/2} + \cdots\end{aligned}$$

where

$$d_0 = \tfrac{1}{2}\left(1 + T(r^2)\right), \quad d_1 = 0,$$

$$d_2 = -\tfrac{1}{2}\left(c_1^2 + 2rT'(r^2)\right), \quad d_3 = c_1c_2,$$

$$d_4 = \tfrac{1}{2}\left(-c_2^2 - 2c_1c_3 + 2r^2T''(r^2) + T'(r^2)\right), \quad d_5 = -c_2c_3 - c_1c_4$$

and $T'(x)$, $T''(x)$ denote the first and second derivatives of $T(x)$, respectively. The singular part of $t(x)$ (that is, the part corresponding to d_k for odd k) depends just on the coefficients c_j. No analogous simplification of the analytic part of $t(x)$ (d_k for even k) is known.

1.4.2 *Darboux–Pólya Method*

Although the asymptotic series for T_n and t_n are evidently new, the underlying method appears (at least implicitly) in the works of Darboux [8, 9] and Pólya [10]. We give the steps of a straightforward algorithm for computing the m^{th} coefficient C_m of the asymptotic series for T_n.

Define first $z_{i,j}$ to be 0 if ($i \geq 1$ and $j = 2$) or ($j > 2$), and 1 otherwise. Define $P_{i,j}$ and $A_{i,j}$ via the recursions

$$P_{i,j} = z_{i,j} \frac{F_{i,j} - \sum_{p=1}^{i-1} \sum_{q=0}^{j} \binom{i}{p}\binom{j}{q} A_{p,q} P_{i-p,j-q} - \sum_{q=1}^{j} \binom{j}{q} A_{0,q} P_{i,j-q}}{A_{0,0}},$$

$$A_{i,j} = \frac{F_{i,j+2} - \sum_{p=0}^{i-1} \sum_{q=0}^{j+2} \binom{i}{p}\binom{j+2}{q} A_{p,q} P_{i-p,j-q+2}}{(j+1)(j+2)}$$

with initial conditions $P_{0,2} = 2$ and $P_{0,j} = 0$ for all $j \neq 2$. Let

$$p_k = \frac{P_{k,1}(-r)^k}{k!}, \quad q_k = \frac{P_{k,0}(-r)^k}{k!}$$

and define b_l via the recursion

$$b_l = \frac{-\sum_{k=1}^{l-1} b_k b_{l-k} + \frac{1}{4}\sum_{k=1}^{l} p_k p_{l-k+1} - q_{l+1}}{2b_0}$$

with initial condition $b_0 = -\sqrt{-q_1}$.

Define next

$$s_i = 2^{i-2}\binom{2i}{i} - \frac{1}{2}\sum_{j=1}^{i-1} \binom{i-1}{j-1} 2^{3(i-j)} s_j - \sum_{k=1}^{i-1} \sum_{j=1}^{i-k} \binom{i-k-1}{j-1} 2^{3(i-j-k)} s_j s_k$$

with initial condition $s_0 = 1$, and the recursion

$$S_{u,v} = \begin{cases} 1 & \text{if } u = v = 0 \\ (-1)^u 2^{1-4u} s_u & \text{if } u \geq 1 \text{ and } v = 0 \\ -\sum_{w=0}^{u} \left(v - \frac{1}{2}\right)^{w+1} S_{u-w,v-1} & \text{if } u \geq 0 \text{ and } v \geq 1. \end{cases}$$

Finally, we have

$$C_m = 2\sum_{k=0}^{m} b_k S_{m-k,k+1}$$

which completes the algorithm.

Some explanation is clearly needed. We know that $F(x, T(x)) = 0$. The Weierstrass Preparation Theorem implies that, for (x, y) sufficiently close to $(r, 1)$,

$$F(x,y) = A(x,y) \cdot P(x,y)$$

where $A(x,y)$ is analytic, $A(r,1)\neq 0$, and

$$P(x,y)=(y-1)^2+p(x)(y-1)+q(x)$$

where $p(x)$, $q(x)$ are analytic and $p(r)=q(r)=0$. The sequence $\{b_l\}$ arises from setting the various coefficients of the polynomial-like approximation $P(x,T(x))$ equal to zero. By Darboux's theorem,

$$T_n\sim(-1)^n r^{-n}\sum_{k=0}^{\infty}b_k\binom{k+1/2}{n};$$

hence it remains to compute asymptotic series for half-integer binomial coefficients. We know that [11]

$$\begin{aligned}\binom{-1/2}{n} &= \frac{(-1)^n}{\sqrt{\pi n}}\left(1-\frac{1}{8n}+\frac{1}{128n^2}+\frac{5}{1024n^3}-\frac{21}{32768n^4}-\frac{399}{262144n^5}+\cdots\right)\\ &= \frac{(-1)^n}{\sqrt{\pi n}}\sum_{j=0}^{\infty}\frac{S_{j,0}}{n^j}\end{aligned}$$

from which we immediately deduce that

$$\begin{aligned}\binom{1/2}{n} &= \frac{(-1)^{n+1}}{2\sqrt{\pi}n^{3/2}}\left(1+\frac{3}{8n}+\frac{25}{128n^2}+\frac{105}{1024n^3}+\frac{1659}{32768n^4}+\frac{6237}{262144n^5}+\cdots\right)\\ &= \frac{1}{2\sqrt{\pi}}\frac{(-1)^n}{n^{3/2}}\sum_{j=0}^{\infty}\frac{2S_{j,1}}{n^j},\end{aligned}$$

$$\begin{aligned}\binom{3/2}{n} &= \frac{3(-1)^n}{4\sqrt{\pi}n^{5/2}}\left(1+\frac{15}{8n}+\frac{385}{128n^2}+\frac{4725}{1024n^3}+\frac{228459}{32768n^4}+\frac{2747745}{262144n^5}+\cdots\right)\\ &= \frac{1}{2\sqrt{\pi}}\frac{(-1)^n}{n^{3/2}}\sum_{j=0}^{\infty}\frac{2S_{j,2}}{n^{j+1}},\end{aligned}$$

$$\begin{aligned}\binom{5/2}{n} &= \frac{15(-1)^{n+1}}{8\sqrt{\pi}n^{7/2}}\left(1+\frac{35}{8n}+\frac{1785}{128n^2}+\frac{40425}{1024n^3}+\frac{3462459}{32768n^4}+\frac{71996925}{262144n^5}+\cdots\right)\\ &= \frac{1}{2\sqrt{\pi}}\frac{(-1)^n}{n^{3/2}}\sum_{j=0}^{\infty}\frac{2S_{j,3}}{n^{j+2}},\end{aligned}$$

and so forth. The conclusion follows.

Addendum I Philippe Flajolet maintained that the preceding discussion tends to "hide the facts" and provided thoughtful comments. Briefly, the equation $F(x,T(x))=0$ can be rearranged as $T(x)=\xi\exp(T(x))$ with

$$\xi(x)=x\exp\left(\sum_{k=2}^{\infty}\frac{T(x^k)}{k}\right).$$

The inverse function of $y\exp(-y)$ is the well-known Cayley tree function τ, an elementary variant of the Lambert W function:

$$\tau(z) = \sum_{n=1}^{\infty} n^{n-1}\frac{z^n}{n!}$$

on the complex plane. In a small disk around the origin, therefore, $T(z) = \tau(\xi(z))$. From here, singularities are easily accessed, making a full asymptotic expansion possible. Writing such conceptual remarks were, in Flajolet's words, an "enjoyable intermezzo" for him despite limited time. These eventually found their way into his treatise [12] with Sedgewick. For completeness, we mention that $C_0 = 1.5594900203...$ for rooted trees (as presented in [12]) and that the corresponding coefficient is 1.1300337163... for binary trees.

1.4.3 *Prime Divisors*

If $\omega(n)$ is the number of distinct prime divisors of n, and $\Omega(n)$ is the total number (including multiplicity) of prime divisors of n, then

$$\mathrm{E}_n(\omega) \sim \ln(\ln(n)) + 0.2614972128... + \sum_{k=1}^{\infty}\left(-1+\sum_{j=0}^{k-1}\frac{\gamma_j}{j!}\right)\frac{(k-1)!}{\ln(n)^k},$$

$$\begin{aligned}\mathrm{Var}_n(\omega) \;\sim\; & \ln(\ln(n)) - 1.8356842740... + \frac{1.0879488865...}{\ln(n)} \\ & + \frac{3.3231293098...}{\ln(n)^2} + \cdots,\end{aligned}$$

$$\mathrm{E}_n(\Omega) \sim \ln(\ln(n)) + 1.0346538818... + \sum_{k=1}^{\infty}\left(-1+\sum_{j=0}^{k-1}\frac{\gamma_j}{j!}\right)\frac{(k-1)!}{\ln(n)^k},$$

$$\begin{aligned}\mathrm{Var}_n(\Omega) \;\sim\; & \ln(\ln(n)) + 0.7647848097... - \frac{2.8767219464...}{\ln(n)} \\ & - \frac{4.9035933594...}{\ln(n)^2} + \cdots,\end{aligned}$$

where

$$\mathrm{E}_n(X) = \frac{1}{n}\sum_{i=1}^{n} X(i), \quad \mathrm{Var}_n(X) = \mathrm{E}_n(X^2) - \mathrm{E}_n(X)^2$$

and γ_j is the j^{th} Stieltjes constant [13]. The leading numerical terms in each of the four expansions are [4, 14]

$$\lambda = \gamma_0 + \sum_p \left[\ln\left(1-\frac{1}{p}\right)+\frac{1}{p}\right] = \gamma_0 + \sum_{k=2}^{\infty}\frac{\mu(k)}{k}\ln(\zeta(k)),$$

$$\lambda - \sum_p \frac{1}{p^2} - \frac{\pi^2}{6} = \lambda - \sum_{k=1}^{\infty} \frac{\mu(k)}{k} \ln(\zeta(2k)) - \frac{\pi^2}{6},$$

$$\Lambda = \gamma_0 + \sum_p \left[\ln\left(1 - \frac{1}{p}\right) + \frac{1}{p-1} \right] = \gamma_0 + \sum_{k=2}^{\infty} \frac{\varphi(k)}{k} \ln(\zeta(k)),$$

$$\Lambda + \sum_p \frac{1}{(p-1)^2} - \frac{\pi^2}{6} = \Lambda + \sum_{k=2}^{\infty} \frac{\varphi_2(k) - \varphi(k)}{k} \ln(\zeta(k)) - \frac{\pi^2}{6},$$

respectively, where $\zeta(x)$ is the Riemann zeta function, $\mu(k)$ is the Möbius mu function, $\varphi(k)$ is the Euler totient function, and the function $\varphi_l(k)$ is defined by

$$\frac{\varphi_l(k)}{k^l} = \prod_{p|k} \left(1 - \frac{1}{p^l}\right), \quad \frac{\zeta(s-l)}{\zeta(s)} = \sum_{k=1}^{\infty} \frac{\varphi_l(k)}{k^s}$$

(in particular, $\varphi = \varphi_1$).

The second numerical coefficient in $\mathrm{Var}_n(\omega)$ is

$$\gamma_0 - 1 + 2\sum_p \frac{\ln(p)}{p(p-1)} = \gamma_0 - 1 + 2\sum_{k=2}^{\infty} \mu(k) \frac{\zeta'(k)}{\zeta(k)}$$

and the second numerical coefficient in $\mathrm{Var}_n(\Omega)$ is

$$\gamma_0 - 1 - 2\sum_p \frac{\ln(p)}{(p-1)^2} = \gamma_0 - 1 + 2\sum_{k=2}^{\infty} \varphi(k) \frac{\zeta'(k)}{\zeta(k)},$$

where $\zeta'(x)$ is the derivative of the zeta function. This result, as well as the result for means, appears in [14–16] but apparently with errors. Knuth [17] revisited Diaconis' original computations; this essay closely follows [17]. Finally, the third numerical coefficient in $\mathrm{Var}_n(\omega)$ is

$$-\gamma_1 - (\gamma_0 - 1)\left(\gamma_0 + 2\sum_p \frac{\ln(p)}{p(p-1)}\right) + 2\sum_p \frac{(2p-1)\ln(p)^2}{2p(p-1)^2}$$

and the third numerical coefficient in $\mathrm{Var}_n(\Omega)$ is

$$-\gamma_1 - (\gamma_0 - 1)\left(\gamma_0 - 2\sum_p \frac{\ln(p)}{(p-1)^2}\right) - 2\sum_p \frac{p\ln(p)^2}{(p-1)^3}.$$

For completeness' sake, we record the values of six relevant prime series [4, 14, 18]:

$$t = \sum_p \frac{1}{p^2} = 0.4522474200..., \quad T = \sum_p \frac{1}{(p-1)^2} = 1.3750649947...,$$

$$u=\sum_p \frac{\ln(p)}{p(p-1)}=0.7553666108..., \quad U=\sum_p \frac{\ln(p)}{(p-1)^2}=1.2269688056...,$$

$$v=\sum_p \frac{(2p-1)\ln(p)^2}{2p(p-1)^2}=1.1837806913..., \quad V=\sum_p \frac{p\ln(p)^2}{(p-1)^3}=2.0914802823....$$

1.4.4 *Selberg–Delange Method*

The theory here is deeper than what was discussed earlier. It starts with asymptotic formulas for the generating functions [19–21]

$$\frac{1}{N}\sum_{n=1}^{N} z^{\omega(n)}=\ln(N)^{z-1}\left(a_0(z)+\frac{a_1(z)}{\ln(N)}+\frac{a_2(z)}{\ln(N)^2}+\cdots\frac{a_r(z)}{\ln(N)^r}+O\left(\frac{1}{\ln(N)^{r+1}}\right)\right),$$

$$\frac{1}{N}\sum_{n=1}^{N} z^{\Omega(n)}=\ln(N)^{z-1}\left(A_0(z)+\frac{A_1(z)}{\ln(N)}+\frac{A_2(z)}{\ln(N)^2}+\cdots\frac{A_r(z)}{\ln(N)^r}+O\left(\frac{1}{\ln(N)^{r+1}}\right)\right),$$

where if

$$\frac{s-1}{s}\prod_p\left(1-\frac{1}{p^s}\right)^{z-1}\left(1+\frac{z}{p^s-1}\right)=\sum_{k=0}^{\infty} b_k(z)(s-1)^k=b(z),$$

$$\frac{s-1}{s}\prod_p\left(1-\frac{1}{p^s}\right)^{z-1}\left(1-\frac{z}{p^s}\right)^{-1}=\sum_{k=0}^{\infty} B_k(z)(s-1)^k=B(z),$$

then

$$a_j(z)=\frac{b_j(z)}{\Gamma(z-j)}, \quad A_j(z)=\frac{B_j(z)}{\Gamma(z-j)}.$$

Let us focus on $\omega(n)$ for the sake of definiteness. Delange's formula expresses that, asymptotically, if n is uniformly distributed on $\{1,2,...,N\}$, then the distribution of $\omega(n)$ is the convolution of a Poisson random variable with mean $\ln(\ln(N))$ and another random variable X whose generating function is

$$\mathrm{E}(z^X)\sim a_0(z)+\frac{a_1(z)}{\ln(N)}+\frac{a_2(z)}{\ln(N)^2}+\cdots.$$

Thus, the mean of $\omega(n)$ will be $\ln(\ln(N))$ plus the mean of X, and the variance will be $\ln(\ln(N))$ plus the variance of X. We have

$$\mathrm{E}(X)\sim a_0'(1)+\frac{a_1'(1)}{\ln(N)}+\frac{a_2'(1)}{\ln(N)^2}+\cdots,$$

$$\mathrm{E}(X(X-1))\sim a_0''(1)+\frac{a_1''(1)}{\ln(N)}+\frac{a_2''(1)}{\ln(N)^2}+\cdots,$$

hence

$$\mathrm{Var}(X) \sim c_0 + \frac{c_1}{\ln(N)} + \frac{c_2}{\ln(N)^2} + \cdots$$

where

$$c_j = a_j''(1) + a_j'(1) - \sum_{i=0}^{j} a_i'(1)a_{j-i}'(1).$$

The corresponding coefficients for $\Omega(n)$ will be denoted by C_0, C_1, C_2, ... and satisfy similar relations.

To obtain the mean, note that setting $z = 1$ in the formula for $b(z)$ gives

$$\frac{s-1}{s}\zeta(s) = \sum_{k=0}^{\infty} b_k(1)(s-1)^k.$$

Replacing s by $s+1$, we have

$$\left(\sum_{i=0}^{\infty}(-1)^i s^i\right)\left(1 + \sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\gamma_j s^{j+1}\right) = \frac{s}{s+1}\zeta(s+1) = \sum_{k=0}^{\infty} b_k(1)s^k,$$

thus

$$b_0(1) = 1, \quad b_1(1) = \gamma_0 - 1, \quad b_2(1) = -(\gamma_1 + \gamma_0 - 1).$$

Since

$$a_0'(1) = b_0'(1) + \gamma_0 b_0(1) = \lambda \quad \text{(to be proved shortly)},$$

$$a_k'(1) = (-1)^{k-1}(k-1)!b_k(1), \quad k \geq 1,$$

the result follows. This argument also applies verbatim to $B(z)$, but with λ replaced by Λ.

To obtain the variance, differentiate $b(z)$ and set $z = 1$:

$$\begin{aligned} b'(1) &= b(1)\sum_p \left[\ln\left(1 - \frac{1}{p^s}\right) + \frac{1}{p^s}\right] \\ &= \left\{1 + (\gamma_0 - 1)(s-1) - (\gamma_1 + \gamma_0 - 1)(s-1)^2 + \cdots\right\} \\ &\quad \cdot \left\{(\lambda - \gamma_0) + u(s-1) - v(s-1)^2 + \cdots\right\} \end{aligned}$$

thus

$$b_0'(1) = \lambda - \gamma_0, \quad b_1'(1) = (\gamma_0 - 1)(\lambda - \gamma_0) + u,$$

$$b_2'(1) = -v + (\gamma_0 - 1)u - (\gamma_1 + \gamma_0 - 1)(\lambda - \gamma_0).$$

Also

$$\begin{aligned} b''(1) &= b'(1)\sum_p \left[\ln\left(1 - \frac{1}{p^s}\right) + \frac{1}{p^s}\right] - b(1)\sum_p \frac{1}{p^{2s}} \\ &= \{(\lambda - \gamma_0) + \cdots\}\{(\lambda - \gamma_0) + \cdots\} - \{1 + \cdots\}\{t + \cdots\} \end{aligned}$$

therefore $b_0''(1) = (\lambda - \gamma_0)^2 - t$. Since

$$a_0''(1) = b_0''(1) + 2\gamma_0 b_0'(1) + \left(\gamma_0^2 - \tfrac{\pi^2}{6}\right) b_0(1) = \lambda^2 - t - \tfrac{\pi^2}{6},$$

$$a_k''(1) = 2(-1)^{k-1}(k-1)! \left(b_k'(1) + \left(\gamma_0 - \sum_{j=1}^{k-1} \tfrac{1}{j} \right) b_k(1) \right), \quad k \geq 1,$$

the formulas for c_0, c_1, c_2 follow.

In the same way, to obtain the variance for $\Omega(n)$, differentiate $B(z)$ and set $z = 1$:

$$\begin{aligned} B'(1) &= B(1) \sum_p \left[\ln\left(1 - \frac{1}{p^s}\right) + \frac{1}{p^s - 1} \right] \\ &= \left\{ 1 + (\gamma_0 - 1)(s-1) - (\gamma_1 + \gamma_0 - 1)(s-1)^2 + \cdots \right\} \\ &\quad \cdot \left\{ (\Lambda - \gamma_0) - U(s-1) + V(s-1)^2 + \cdots \right\} \end{aligned}$$

thus

$$\begin{aligned} &B_0'(1) = \Lambda - \gamma_0, \quad B_1'(1) = (\gamma_0 - 1)(\Lambda - \gamma_0) - U, \\ &B_2'(1) = V - (\gamma_0 - 1)U - (\gamma_1 + \gamma_0 - 1)(\Lambda - \gamma_0). \end{aligned}$$

Also

$$\begin{aligned} B''(1) &= B'(1) \sum_p \left[\ln\left(1 - \frac{1}{p^s}\right) + \frac{1}{p^s - 1} \right] + B(1) \sum_p \frac{1}{(p^s - 1)^2} \\ &= \{(\Lambda - \gamma_0) + \cdots\} \{(\Lambda - \gamma_0) + \cdots\} + \{1 + \cdots\} \{T + \cdots\} \end{aligned}$$

therefore $B_0''(1) = (\Lambda - \gamma_0)^2 + T$. We have $A_0''(1) = \Lambda^2 + T - \frac{\pi^2}{6}$ and a formula for $A_k''(1)$, $k \geq 1$, identical to that for $a_k''(1)$ earlier; hence the formulas for C_0, C_1, C_2 follow. It is interesting that higher-order terms for $\mathrm{E}_n(\omega)$ and $\mathrm{E}_n(\Omega)$ coincide, but differ for $\mathrm{Var}_n(\omega)$ and $\mathrm{Var}_n(\Omega)$.

We conclude with an unsolved problem. The expressions

$$\sum_{n=1}^{N} 2^{\omega(n)}, \quad \sum_{n=1}^{N} 3^{\omega(n)}, \quad \sum_{n=1}^{N} 2^{\Omega(n)}$$

were mentioned in [22]. Tenenbaum [23] computed that

$$\sum_{n=1}^{N} 3^{\Omega(n)} = N^\theta g\left(\tfrac{\ln(N)}{\ln(2)}\right) + O(N \ln(N)^3)$$

where $\theta = \ln(3)/\ln(2) = 1.5849625007...$ [24] and $g(x)$ is a fractal-like function of period 1 that oscillates between two positive constants. In fact,

$$g(x) = \frac{3}{2} \sum_{\substack{m \geq 1 \\ \gcd(m,6)=1}} \left(\frac{3^{\Omega(m)}}{m^\theta} \cdot \sum_{k \geq 0} 3^{-(\theta-1)k - \left\{x - \frac{\ln(m)}{\ln(2)} - \theta k\right\}} \right)$$

where $\{y\} = y - \lfloor y \rfloor$ for all real numbers y, and

$$3.74... = \lim_{x\to 1^-} g(x) = \inf_x g(x) < \sup_x g(x) = \lim_{x\to 0^+} g(x) = 4.74...$$

It would be good to someday know these bounds to higher precision.

Addendum II Let $P(n)$ be the largest prime factor of n. The average of $P(n)$ satisfies [25]

$$\frac{1}{N}\sum_{n\le N} P(n) \sim \sum_{k=0}^{\infty} k! \frac{N}{\ln(N)^{k+1}}\xi_k$$

as $N\to\infty$, where

$$\xi_k = \frac{1}{2^{k+1}}\sum_{j=0}^{k} \frac{2^j(-1)^j}{j!}\zeta^{(j)}(2)$$

and the median $M(N)$ of $\{P(n) : n \le N\}$ satisfies [26]

$$M(N) \sim \exp\left(\frac{\gamma - 1}{\sqrt{e}}\right) N^{1/\sqrt{e}} = (0.7738078734...)N^{0.6065306597...}$$

(actually, more terms in the asymptotic expansion of $M(N)$ are possible). Clearly the median value grows substantially slower than the mean value. The mode (most frequent value) grows even more slowly [27]. Another interesting asymptotic expansion [26] refines de la Vallée Poussin's average for fractional parts of a large integer N divided by each prime $p \le N$:

$$\sum_{p\le N}\left\{\frac{N}{p}\right\} \sim \sum_{k=0}^{\infty} k! \frac{N}{\ln(N)^{k+1}}\eta_k$$

as $N\to\infty$, where

$$\eta_k = 1 - \sum_{j=0}^{k}\frac{\gamma_j}{j!}.$$

Alternative proofs regarding $\mathrm{E}_n(\omega)$ and $\mathrm{E}_n(\Omega)$ also appear in [26], but not $\mathrm{Var}_n(\omega)$ nor $\mathrm{Var}_n(\Omega)$.

Acknowledgments I am grateful to Donald Knuth for his detailed letter [17], on which my discussion of the Selberg–Delange method is based, and Philippe Flajolet for his supportive guidance. Persi Diaconis kindly provided [15] and Pascal Sebah computed the constants v and V.

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000055 and A000081.
[2] S. R. Finch, Otter's tree enumeration constants, *first volume*, pp. 295–316.
[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001221, A001222, A013939, A022559, and A069811.

[4] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.
[5] J. M. Plotkin and J. W. Rosenthal, How to obtain an asymptotic expansion of a sequence from an analytic identity satisfied by its generating function, *J. Austral. Math. Soc. Ser. A* 56 (1994) 131–143; MR1250997 (94k:05015).
[6] M. Drmota and B. Gittenberger, The distribution of nodes of given degree in random trees, *J. Graph Theory* 31 (1999) 227–253; MR1688949 (2000f:05069).
[7] M. Drmota, Combinatorics and asymptotics on trees, *Cubo* 6 (2004) 105–136.
[8] G. Darboux, Mémoire sur l'approximation des fonctions de très grands nombres, *Journal de Mathématiques* 4 (1878) 5–56, 377–416.
[9] G. Szegö, *Orthogonal Polynomials*, 2nd ed., Amer. Math. Soc., 1975, pp. 205–206; MR0106295 (21 #5029).
[10] G. Pólya and R. C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, 1987; MR0884155 (89f:05013).
[11] D. E. Knuth, I. Vardi and R. Richberg, The asymptotic expansion of the middle binomial coefficient, *Amer. Math. Monthly* 97 (1990) 626–630.
[12] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009, pp. 475–480; MR2483235 (2010h:05005).
[13] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.
[14] D. E. Knuth, *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 338–339; MR1762319 (2001c:68066).
[15] P. Diaconis, Asymptotic expansions for the mean and variance of the number of prime factors of a number n, Dept. of Statistics Tech. Report 96, Stanford Univ., 1976.
[16] P. Diaconis, G. H. Hardy and probability??? *Bull. London Math. Soc.* 34 (2002) 385–402; MR1897417 (2003d:01023).
[17] D. E. Knuth, Asymptotics for $\mathrm{E}_n(\Omega)$ and $\mathrm{Var}_n(\Omega)$, unpublished note (2003).
[18] P. Sebah, Computing $\sum \frac{p}{(p-1)^3}\log^2 p$ and $\sum \frac{2p-1}{2p(p-1)^2}\log^2 p$, unpublished note (2003).
[19] A. Selberg, Note on a paper by L. G. Sathe, *J. Indian Math. Soc.* 18 (1954) 83–87; MR0067143 (16,676a).
[20] B. Saffari, Sur quelques applications de la "méthode de l'hyperbole" de Dirichlet à la théorie des nombres premiers, *Enseign. Math.* 14 (1970) 205–224; MR0268138 (42 #3037).
[21] H. Delange, Sur des formules de Atle Selberg, *Acta Arith.* 19 (1971) 105–146 (errata insert); MR0289432 (44 #6623).
[22] S. R. Finch, Hafner-Sarnak-McCurley constant, *first volume*, pp. 110–112.
[23] G. Tenenbaum, Asymptotics for the sum of $3^{\Omega(n)}$, unpublished note (2001).
[24] S. R. Finch, Stolarsky-Harborth constant, *first volume*, pp. 145–151.
[25] E. Naslund, The average largest prime factor, *Integers* 13 (2013) A81; MR3167928.
[26] E. Naslund, The median largest prime factor, *J. Number Theory* 141 (2014) 109–118; arXiv:1207.0232; MR3195391.
[27] N. McNew, Popular values of the largest prime divisor function, arXiv:1504.05985.

1.5 Multiples and Divisors

Before discussing multiplication, let us speak about addition. The number $A(k)$ of distinct sums $i+j\leq k$ such that $1\leq i\leq k/2$, $1\leq j\leq k/2$ is clearly $2\lfloor k/2\rfloor-1$. Hence the number $A(2n)$ of distinct elements in the $n\times n$ addition table involving $\{1,2,\ldots,n\}$ satisfies $\lim_{n\to\infty}A(2n)/n=2$, as expected.

We turn to multiplication. Let $M(k)$ be the number of distinct products $ij \le k$ such that $1 \le i \le \sqrt{k}$, $1 \le j \le \sqrt{k}$. One might expect that the number $M(n^2)$ of distinct elements in the $n \times n$ multiplication table to be approximately $n^2/2$; for example, $M(10^2) = 42$. In a surprising result, Erdős [1–3] proved that $\lim_{n\to\infty} M(n^2)/n^2 = 0$. More precisely, we have [4]

$$\lim_{k\to\infty} \frac{\ln(M(k)/k)}{\ln(\ln(k))} = -\delta$$

where

$$\delta = 1 - \frac{1 + \ln(\ln(2))}{\ln(2)} = 0.0860713320...$$

In spite of good estimates for $M(k)$, an asymptotic formula for $M(k)$ as $k \to \infty$ remains unknown [5].

Given a positive integer n, define

$$\rho_1(n) = \max_{\substack{d|n,\\ d \le \sqrt{n}}} d, \quad \rho_2(n) = \min_{\substack{d|n,\\ d \ge \sqrt{n}}} d;$$

thus $\rho_1(n)$ and $\rho_2(n)$ are the two divisors of n closest to $\sqrt{n}$. Let

$$R_1(N) = \sum_{n=1}^{N} \rho_1(n), \quad R_2(N) = \sum_{n=1}^{N} \rho_2(n).$$

It is not difficult to prove that

$$\lim_{N\to\infty} \frac{\ln(N)}{N^2} R_2(N) = \frac{\pi^2}{12}.$$

An analogous asymptotic expression for $R_1(N)$ is still open, but Tenenbaum [6–8] proved that

$$\lim_{N\to\infty} \frac{\ln(R_1(N)/N^{3/2})}{\ln(\ln(N))} = -\delta$$

where δ is exactly as before. It is curious that one limit is so much harder than the other, and that the same constant δ appears as with the multiplication table problem.

Erdős conjectured long ago that almost all integers n have two divisors d, d' such that $d < d' \le 2d$. By "almost all", we mean all integers n in a sequence of asymptotic density 1, abbreviated as "p.p." Given n, select divisors $a_n < b_n$ for which b_n/a_n is minimal. To prove the conjecture, it is sufficient to show that $b_n/a_n \to 1^+$ as $n \to \infty$ p.p.; that is, $\ln(\ln(b_n/a_n)) \to -\infty$ p.p. Maier & Tenenbaum [9–11] succeeded in doing this and, further, demonstrated that

$$\lim_{n\to\infty} \frac{\ln(\ln(b_n/a_n))}{\ln(\ln(n))} = -(\ln(3) - 1) = -0.0986122886... \text{ p.p.}$$

Another way of viewing this problem is by counting those integers n up to N without such divisors d and d'. If $\varepsilon(N)$ is the number of these exceptional integers, then [4]

$$\lim_{N\to\infty} \frac{\ln(\varepsilon(N)/N)}{\ln(\ln(\ln(N)))} \le -\beta$$

where

$$\beta = 1 - \frac{1+\ln(\ln(3))}{\ln(3)} = 0.0041547514....$$

As the inequality suggests, we do not know if this constant is necessarily optimal.

Yet another way of viewing this problem is via the Hooley function

$$\Delta(n) = \max_{x\ge 0} \sum_{\substack{d|n, \\ e^x < d \le e^{x+1}}} 1,$$

that is, the greatest number of divisors of n contained in any interval of logarithmic length 1. More interesting constants emerge here, but their optimality is questionable. In fact, it is conjectured [4] that $\Delta(n)/\ln(\ln(n))$ accumulates not at a single point, but over an entire subinterval $(u, v) \subseteq (0, \infty)$. Estimates of u and v would be good to see someday.

Ramanujan [12] studied the asymptotics of $\sum_{n=1}^{N} 1/d(n)$ as $N \to \infty$, where [13] $d(n)$ is the number of distinct divisors of n. See [14] for more details. This is a special case of a result in [4, 15], which is used to prove the following arcsine distributional law for random divisors d of n:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathrm{P}\left(\frac{\ln(d)}{\ln(n)} < x\right) = \frac{2}{\pi} \arcsin\left(\sqrt{x}\right).$$

Consequently, an integer has (on average) many small divisors and many large divisors.

Sita Ramaiah & Suryanarayana [16] found a corresponding formula for $\sum_{n=1}^{N} 1/\sigma(n)$, where [17] $\sigma(n)$ is the sum of all divisors of n. DeKoninck & Ivić [18] had asserted that constants appearing in such a formula would be complicated; they were right! [14] It turns out that the Riemann Hypothesis [19] is true if and only if [20, 21]

$$\sigma(n) < e^{\gamma} n \ln(\ln(n)) \quad \text{for all sufficiently large } n,$$

where γ is the Euler–Mascheroni constant [22].

An integer n is **highly composite** if $d(m) < d(n)$ for all $m < n$. Let $Q(N)$ denote the number of highly composite integers $n \le N$. It is known that [11, 23–26]

$$1.136 \le \liminf_{N\to\infty} \frac{\ln(Q(N))}{\ln(N)} \le 1.44, \quad \limsup_{N\to\infty} \frac{\ln(Q(N))}{\ln(N)} \le 1.71,$$

based on Diophantine approximations of the quantity $\ln(3/2)/\ln(2) = 0.5849625007...$. It is conjectured that the limit exists and

$$\lim_{N\to\infty} \frac{\ln(Q(N))}{\ln(N)} = \frac{\ln(2)+\ln(3)+\ln(5)}{4\ln(2)} = 1.2267226489...$$

but this appears to be difficult.

Let us return to the constant δ, which appears in several other places in the literature [27–34]. We mention only three. With regard to Erdős' conjecture, Roesler [35] added a further constraint that $a_n b_n = n$ when minimizing b_n/a_n; he proved that

$$\lim_{N\to\infty} \frac{\ln\left(\frac{1}{N}\sum\limits_{n=1}^{N}\frac{a_n}{b_n}\right)}{\ln(\ln(N))} = -\delta.$$

Hence the integers are fairly quadratic, in the sense that $b_n - a_n$ is quite small on average. We wonder what happens to the limiting ratio if a_n/b_n is replaced in the summation by b_n/a_n.

An odd prime p is said to be **symmetric** [36, 37] if there exists an odd prime q such that $|p-q| = \gcd(p-1, q-1)$. For example, any twin prime is symmetric. It is known that the reciprocal sum of symmetric primes is finite (like Brun's constant [38]). If the twin prime conjecture is true, then there are infinitely many symmetric primes. Let $S(n)$ denote the number of symmetric primes $\le n$. It is conjectured that

$$\lim_{n\to\infty} \frac{\ln(S(n)/n)}{\ln(\ln(n))} = -1-\delta$$

and a heuristic argument supporting this formula appears in [36].

Finally, let $T(N)$ denote the number of integers $n \le N$ satisfying the inequality $d(n) \ge \ln(N)$. Norton [39], responding to a question raised by Steinig, proved that there are positive constants $\xi < \eta$ with

$$\xi \le \rho(N) = \frac{T(N)}{N\ln(N)^{-\delta}\ln(\ln(N))^{-1/2}} \le \eta$$

for all large N. Balazard, Nicolas, Pomerance & Tenenbaum [40] proved that the ratio $\rho(N)$ does not tend to a limit as $N \to \infty$, and that

$$\rho(N) \sim f\left(\frac{\ln(\ln(N))}{\ln(2)}\right) \quad \text{as } N \to \infty$$

where $f(x)$ is an explicit left-continuous function of period 1 with only countably many jump discontinuities. Deléglise & Nicolas [41] further computed that

$$\xi = \lim_{x\to 0^+} f(x) = 0.9382786811..., \quad \eta = f(0) = 1.1481267734...$$

are the best possible asymptotic bounds on $\rho(N)$. We have seen such oscillatory functions on numerous occasions elsewhere in number theory and combinatorics

[42, 43]. The quantities

$$\chi = \frac{1}{\Gamma(1+\lambda)} \prod_{p \text{ prime}} \left(1 - \frac{1}{p}\right)^{\lambda} \left(1 + \frac{\lambda}{p}\right) = 0.3495143728...,$$

$$\frac{\chi}{1-\ln(2)} \sqrt{\frac{\ln(2)}{2\pi}} = 0.3783186209... = \frac{\xi}{2.4801282017...} = \frac{\eta}{3.0348143331...}$$

also play an intermediate role [41], where $\lambda = \ln(2)^{-1}$.

1.5.1 *Practical Numbers*

A positive integer n is **practical** if every integer m with $1 \le m \le n$ can be written as a sum of distinct positive divisors of n [44–46]. No odd $n > 1$ can be practical; consider $m = 2$. Also, $n = 10$ is not practical; consider $m = 4$. Letting $P(N)$ denote the count of practical numbers $n \le N$, we have [47–52]

$$\lim_{N \to \infty} \frac{\ln(N)}{N} P(N) = \kappa,$$

where

$$\kappa = \frac{1}{1 - e^{-\gamma}} \sum_{\substack{n \\ \text{practical}}} \frac{1}{n} \left(\sum_{p \le \sigma(n)+1} \frac{\ln(p)}{p-1} - \ln(n) \right) \prod_{p \le \sigma(n)+1} \left(1 - \frac{1}{p}\right)$$

provably satisfies $1.311 < \kappa < 1.693$. This is consistent with an empirical estimate $\kappa \approx 1.341$ given in [49].

A positive integer n is φ-**practical** if the polynomial $t^n - 1$ has a divisor (with integer coefficients) of every degree up to n. Equivalently, n is φ-practical if every integer m with $1 \le m \le n$ is of the form $\varphi(d_1) + \varphi(d_2) + \cdots + \varphi(d_\ell)$ for some $\ell \ge 1$, where φ is the Euler totient function [53] and $d_1, d_2, \ldots, d_\ell$ are distinct positive divisors of n. For example, $n = 5$ is not φ-practical because $t^5 - 1$ has no divisor of degree 2; alternatively, sums involving $\varphi(1) = 1$ and $\varphi(5) = 4$ omit the value $m = 2$. By contrast, $n = 6$ is φ-practical because

$$t - 1, \quad t^2 + t + 1, \quad t^3 - 1, \quad t^4 + t^3 - t - 1, \quad t^5 + t^4 + t^3 + t^2 + t + 1$$

are divisors of $t^6 - 1$; alternatively, sums involving $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(3) = 2$ and $\varphi(6) = 2$ cover all values $1 \le m \le 6$. Letting $P_\varphi(N)$ denote the count of φ-practical numbers $n \le N$, we have [52, 54, 55]

$$\lim_{N \to \infty} \frac{\ln(N)}{N} P_\varphi(N) = \kappa_\varphi,$$

where κ_φ possesses a more complicated formula than κ and provably satisfies $0.945 < \kappa_\varphi < 0.967$. This is consistent with an empirical estimate $\kappa \approx 0.96$ given in [55].

Related work appears in [56–58], as well as the Erdős–Ford–Tenenbaum constant δ and an interesting equation

$$1 = \int_1^\infty \omega(y)(y+1)^{-1-x} dy$$

with solution $x = 0.433489...$, where ω is Buchstab's function [59].

Addendum Ford [60] proved that there exist positive constants $c < C$ such that

$$c\frac{N}{\ln(N)^\delta \ln(\ln(N))^{3/2}} \le M(N) \le C\frac{N}{\ln(N)^\delta \ln(\ln(N))^{3/2}}$$

for large N, and positive constants $c' < C'$ such that

$$c'\frac{N^{3/2}}{\ln(N)^\delta \ln(\ln(N))^{3/2}} \le R_1(N) \le C'\frac{N^{3/2}}{\ln(N)^\delta \ln(\ln(N))^{3/2}}$$

for large N. Thus, for the first time, the true order of magnitude of $M(N)$ and of $R_1(N)$ is known. See also [61] for an application to computer science.

A famous result is [62–66]

$$\limsup_{n\to\infty} \ln(d(n))\frac{\ln\ln n}{\ln n} = \ln 2,$$

but the analogous result for the iterated divisor

$$\begin{aligned}\limsup_{n\to\infty} \ln(d(d(n)))\frac{\ln\ln n}{\sqrt{\ln n}} &= \left(8\sum_{j=1}^\infty \ln\left(1+\frac{1}{j}\right)^2\right)^{1/2} \\ &= 2.7959802335...\end{aligned}$$

was proved comparatively recently [67, 68].

[1] P. Erdős, An asymptotic inequality in the theory of numbers (in Russian), *Vestnik Leningrad. Univ.*, v. 15 (1960) n. 13, 41–49; MR0126424 (23 #A3720).

[2] P. Erdős and A. Sárközy, On the number of prime factors in integers, *Acta Sci. Math. (Szeged)* 42 (1980) 237–246; MR0603312 (82c:10053).

[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A027384, A027424 and A074738.

[4] R. R. Hall and G. Tenenbaum, *Divisors*, Cambridge Univ. Press, 1988, pp. 27–34, 95–99, 207–211; MR0964687 (90a:11107).

[5] L. Babai, C. Pomerance, and P. Vértesi, The mathematics of Paul Erdős, *Notices Amer. Math. Soc.* 45 (1998) 19–31; MR1490536 (98m:01025).

[6] G. Tenenbaum, Sur deux fonctions de diviseurs, *J. London Math. Soc.* 14 (1976) 521–526; MR0432569 (55 #5557).

[7] G. Tenenbaum, 1105: first steps in a mysterious quest, *The Mathematics of Paul Erdős*, v. 1, ed. R. L. Graham and J. Nesetril, Springer-Verlag, 1997, pp. 268–275; MR1425191 (98d:11115).

[8] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A033676 and A033677.

[9] P. Erdős and R. R. Hall, The propinquity of divisors, *Bull. London Math. Soc*. 11 (1979) 304–307; MR0554398 (81m:10102).

[10] H. Maier and G. Tenenbaum, On the set of divisors of an integer, *Invent. Math.* 76 (1984) 121–128; MR0739628 (86b:11057).

[11] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002182 and A090196.

[12] S. Ramanujan, Some formulae in the analytic theory of numbers, *Messenger of Math.* 45 (1916) 81–84; also in *Collected Papers*, ed. G. H. Hardy, P. V. Seshu Aiyar, and B. M. Wilson, Cambridge Univ. Press, 1927, pp. 133–135, 339–340.

[13] S. R. Finch, Sierpinski's constant: Circle and divisor problems, *first volume*, pp. 123–125.

[14] S. R. Finch, Unitarism and infinitarism, *this volume*, §1.7.

[15] J.-M. Deshouillers, F. Dress, and G. Tenenbaum, Lois de répartition des diviseurs. I, *Acta Arith*. 34 (1979) 273–285 (loose errata); MR0543201 (81d:10043).

[16] V. Sita Ramaiah and D. Suryanarayana, Sums of reciprocals of some multiplicative functions, *Math. J. Okayama Univ.* 21 (1979) 155–164; MR0554306 (82d:10064a).

[17] S. R. Finch, Abundant numbers density constant, *first volume*, pp. 126–127.

[18] J.-M. DeKoninck and A. Ivić, *Topics in Arithmetical Functions: Asymptotic Formulae for Sums of Reciprocals of Arithmetical Functions and Related Fields*, North-Holland, 1980, pp. 3–13, 23; MR0589545 (82a:10047).

[19] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[20] G. Robin, Grandes valeurs de la fonction somme des diviseurs et hypothèse de Riemann, *J. Math. Pures Appl.* 63 (1984) 187–213; MR0774171 (86f:11069).

[21] J. C. Lagarias, An elementary problem equivalent to the Riemann hypothesis, *Amer. Math. Monthly* 109 (2002) 534–543; MR1908008 (2003d:11129).

[22] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.

[23] P. Erdős, On highly composite numbers, *J. London Math. Soc.* 19 (1944) 130–133; MR0013381 (7,145d).

[24] J.-L. Nicolas, Répartition des nombres hautement composés de Ramanujan, *Canad. J. Math.* 23 (1971) 116–130; MR0274400 (43 #165).

[25] J.-L. Nicolas, Nombres hautement composés, *Acta Arith.* 49 (1988) 395–412; MR0937935 (89c:11012).

[26] J.-L. Nicolas, On highly composite numbers, *Ramanujan Revisited*, Proc. 1987 Urbana conf., ed. G. E. Andrews, R. A. Askey, B. C. Berndt, K. G. Ramanathan, and R. A. Rankin, Academic Press, 1988, pp. 215–244; MR0938967 (89c:11011).

[27] P. Erdős and C. Pomerance, Matching the natural numbers up to n with distinct multiples in another interval, *Nederl. Akad. Wetensch. Indag. Math*. 42 (1980) 147–161; MR0577570 (81i:10053).

[28] M. Car, Polynômes de $F_q[X]$ ayant un diviseur de degré donné, *Acta Arith*. 43 (1984) 131–154; MR0736727 (85i:11103).

[29] R. R. Hall and G. Tenenbaum, On Behrend sequences, *Math. Proc. Cambridge Philos. Soc.* 112 (1992) 467–482; MR1177995 (93h:11026).

[30] M. Mkaouar, Sur les fractions continues des séries formelles quadratiques sur $\mathbb{F}_q(X)$, *Acta Arith.* 97 (2001) 241–251; MR1826003 (2002b:11015).

[31] K.-H. Indlekofer and N. M. Timofeev, Divisors of shifted primes, *Publ. Math. Debrecen* 60 (2002) 307–345; MR1898566 (2003b:11076).

[32] R. R. Hall, *Sets of Multiples*, Cambridge Univ. Press, 1996; MR1414678 (98d:11012).

[33] K. Ford, F. Luca and C. Pomerance, The image of Carmichael's λ-function, arXiv:1408.6506.

[34] S. Chow and C. Pomerance, Triangles with prime hypotenuse, arXiv:1703.10953.

[35] F. Roesler, How quadratic are the natural numbers? *Arch. Math. (Basel)* 73 (1999) 193–198; MR1705014 (2000h:11104).

[36] P. Fletcher, W. Lindgren, and C. Pomerance, Symmetric and asymmetric primes, *J. Number Theory* 58 (1996) 89–99; MR1387727 (97c:11007).

[37] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A090190 and A090191.

[38] S. R. Finch, Brun's constant, *first volume*, pp. 133–135.

[39] K. K. Norton, On the number of restricted prime factors of an integer. I, *Illinois J. Math.* 20 (1976) 681–705; MR0419382 (54 #7403).

[40] M. Balazard, J.-L. Nicolas, C. Pomerance, and G. Tenenbaum, Grandes déviations pour certaines fonctions arithmétiques, *J. Number Theory* 40 (1992) 146–164; MR1149734 (93c:11060).

[41] M. Deléglise and J.-L. Nicolas, Sur les entiers inférieurs à x ayant plus de $\log(x)$ diviseurs, *J. Théor. Nombres Bordeaux* 6 (1994) 327–357; MR1360649 (96m:11083).

[42] S. R. Finch, *first volume*, pp. 102, 146, 311, 340, 355, 437.

[43] S. R. Finch, Two asymptotic series, *this volume*, §1.4.

[44] A. K. Srinivasan, Practical numbers, *Current Sci.* 17 (1948) 179–180; MR0027799 (10,356e).

[45] A. A. Krishnaswami Ayyangar, Practical numbers, *Current Sci.* 19 (1950) 237–238.

[46] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005153 and A260653.

[47] M. Hausman and H. N. Shapiro, On practical numbers, *Comm. Pure Appl. Math.* 37 (1984) 705–713; MR0752596 (86a:11036).

[48] G. Tenenbaum, Sur un problème de crible et ses applications, *Annales Sci. École Norm. Sup.* 19 (1986) 1–30; MR0860809 (87m:11094).

[49] M. Margenstern, Les nombres pratiques: théorie, observations et conjectures, *J. Number Theory* 37 (1991) 1–36; MR1089787 (92a:11006).

[50] E. Saias, Entiers à diviseurs denses. I, *J. Number Theory* 62 (1997) 163–191; MR1430008 (98c:11096).

[51] A. Weingartner, Practical numbers and the distribution of divisors, *Quart. J. Math.* 66 (2015) 743–758; arXiv:1405.2585; MR3356847.

[52] A. Weingartner, On the constant factor in several related asymptotic estimates, arXiv:1705.06349.

[53] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.

[54] L. Thompson, Polynomials with divisors of every degree, *J. Number Theory* 132 (2012) 1038–1053; arXiv:1111.5401; MR2890525.

[55] C. Pomerance, L. Thompson and A. Weingartner, On integers n for which $X^n - 1$ has a divisor of every degree, *Acta Arith.* 175 (2016) 225–243; arXiv:1511.03357; MR3557122.

[56] P. Pollack and L. Thompson, On the degrees of divisors of $T^n - 1$, *New York J. Math.* 19 (2013) 91–116; arXiv:1206.2084; MR3028137.

[57] A. Weingartner, On the degrees of polynomial divisors over finite fields, *Math. Proc. Cambridge Philos. Soc.* 161 (2016) 469–487; arXiv:1507.01920; MR3569157.

[58] A. Weingartner, A sieve problem and its application, *Mathematika* 63 (2017) 213–229; arXiv:1605.05204; MR3610011.

[59] S. R. Finch, Golomb-Dickman constant, *first volume*, pp. 284–292.

[60] K. Ford, The distribution of integers with a divisor in a given interval, math.NT/0401223.

[61] N. Pippenger, The average amount of information lost in multiplication, *IEEE Trans. Inform. Theory* 51 (2005) 684–687; arXiv:math/0408043; MR2236076 (2007m:94073).

[62] S. Wigert, Sur l'ordre de grandeur du nombre des diviseurs d'un entier, *Ark. Mat. Astron. Fysik*, v. 3 (1906–1907) n. 18, 1–9.
[63] S. Ramanujan, Highly composite numbers, *Proc. London Math. Soc.* 14 (1915) 347–409; also in *Collected Papers*, ed. G. H. Hardy, P. V. Seshu Aiyar, and B. M. Wilson, Cambridge Univ. Press, 1927, pp. 78–128, 338–339.
[64] P. Shiu, The maximum orders of multiplicative functions, *Quart. J. Math.* 31 (1980) 247–252; MR0576341 (81i:10059).
[65] J.-L. Nicolas and G. Robin, Majorations explicites pour le nombre de diviseurs de *N*, *Canad. Math. Bull.* 26 (1983) 485–492; MR0716590 (85e:11006).
[66] A. G. Postnikov, *Introduction to Analytic Number Theory*, Amer. Math. Soc., 1988, pp. 13–14, 183–187; MR0932727 (89a:11001).
[67] Y. Buttkewitz, C. Elsholtz, K. Ford and J.-C. Schlage-Puchta, A problem of Ramanujan, Erdős, and Kátai on the iterated divisor function, *Internat. Math. Res. Notices* (2012) 4051–4061; arXiv:1108.1815; MR2972548.
[68] C. Elsholtz, M. Technau and N. Technau, The maximal order of iterated multiplicative functions, arXiv:1709.04799.

1.6 Discrepancy and Uniformity

Let $X=\{x_n\}_{n=1}^{\infty}$ be an infinite sequence of points in the interval $[0,1)$ and let X_N denote the finite subsequence $\{x_n\}_{n=1}^{N}$. Define, for each N, the **discrepancy**

$$D_N(X)=\sup_{0\leq a<b<1}\left|\frac{|X_N\cap[a,b)|}{N}-(b-a)\right|$$

and the **star discrepancy**

$$D_N^*(X)=\sup_{0\leq c<1}\left|\frac{|X_N\cap[0,c)|}{N}-c\right|.$$

It can be proved that $1/N\leq D_N\leq 1$ and $1/(2N)\leq D_N^*\leq D_N\leq 2D_N^*$. The sequence X is uniformly distributed in $[0,1)$ if and only if $\lim_{N\to\infty}D_N(X)=0$. We are interested in low-discrepancy sequences, that is, sequences X for which $D_N(X)$ is small for all N. The efficient construction of such X is essential in quasi-Monte Carlo algorithms used, for example, to approximate a multivariate integral or to simulate certain random processes [1–3].

On the one hand, Béjian [4, 5] showed that

$$S(X)=\limsup_{N\to\infty}\frac{N}{\ln(N)}D_N(X)\geq C,$$

$$S^*(X)=\limsup_{N\to\infty}\frac{N}{\ln(N)}D_N^*(X)\geq C^*$$

for all sequences X, where

$$C=\max_{r\geq 2}\frac{r-2}{4(r-1)\ln(r)}=0.120386...,\quad C^*=\frac{C}{2}=0.060193....$$

This is a consequence of work by van Aardenne-Ehrenfest [6, 7], Roth [8] and Schmidt [9] regarding the unavoidable irregularities that occur in any point distribution. Improvement is likely. On the other hand, there are special sequences X for which [10–12]

$$S(X) \leq \frac{23}{35\ln(6)} = 0.366758..., \quad S^*(X) \leq \frac{1919}{3454\ln(12)} = 0.223584...$$

and we shall discuss such examples shortly. The gap between lower and upper bounds is surprisingly wide: It would be good someday for these estimates to be tightened.

Students of statistics will recognize $D_N^*(X)$ as the **Kolmogorov–Smirnov one-sample statistic**, under the hypothesis that the sequence X_N is a random sample drawn from a Uniform $(0,1)$ distribution. Call this hypothesis H_0. We have the following asymptotic result [13–15]:

$$\lim_{N\to\infty} \mathrm{P}\left(\sqrt{N}D_N^*(X) \leq z \mid H_0 \text{ is true}\right) = 1 - 2\sum_{n=1}^{\infty}(-1)^{n-1}e^{-2n^2z^2}.$$

Call this expression $\theta(z)$. In an experimental data analysis setting, if $\sqrt{N}D_N^*(X)$ exceeds a sufficiently large threshold w (in the sense that the probability $1-\theta(w)$ is suitably small), then one must doubt the truth of H_0.

1.6.1 *Scrambled van der Corput Sequences*

Given an integer $n \geq 1$, write the base b representation for $n-1$ as

$$n - 1 = \sum_{k=0}^{\infty} m_k(n)b^k, \quad m_k \in B$$

where $B = \{0, 1, \ldots b-1\}$. Then, for any permutation $\sigma : B \to B$, the scrambled van der Corput sequence X_b^σ has n^{th} term given by

$$x_n = \sum_{j=0}^{\infty} \sigma(m_j(n))b^{-j-1}.$$

For example, if $b = 2$ and σ is the identity permutation ε, then

$$X_2^\varepsilon = \{0,\ 0.1,\ 0.01,\ 0.11,\ 0.001,\ 0.101,\ 0.011, ...\},$$

that is, x_n is simply the reflection of $n-1$ across the decimal point. Haber [16–18] computed that $S^*(X_2^\varepsilon) = 1/(3\ln(2)) = 0.480898....$ Faure [10] proved the more general result

$$S(X_b^\varepsilon) = S^*(X_b^\varepsilon) = \begin{cases} \dfrac{b-1}{4\ln(b)} & \text{if } b \text{ is odd,} \\[2ex] \dfrac{b^2}{4(b+1)\ln(b)} & \text{if } b \text{ is even.} \end{cases}$$

Introducing a non-identity permutation σ provides the smallest discrepancy currently known [12]:

$$S(X_{36}^{\sigma}) = 23/(35\ln(6)) = 0.366758...$$

where σ has cycle structure

$$(0)\ (1\ 25\ 33\ 16\ 23\ 4\ 31\ 30\ 10\ 34\ 24\ 21\ 28\ 2\ 17)\ (3\ 7)$$
$$(5\ 11\ 22\ 14\ 29\ 19\ 32\ 6\ 20\ 8\ 27\ 26\ 12)\ (9\ 13\ 15)\ (18)\ (35).$$

It is useful to generalize X_b^{σ} to X_b^{Σ}, where Σ is a sequence of permutations σ. If $\Sigma = \{\sigma_j\}_{j=0}^{\infty}$, then the n^{th} term of X_b^{Σ} is simply given by

$$x_n = \sum_{j=0}^{\infty} \sigma_j(m_j(n)) b^{-j-1}.$$

For example, if $j \ge 0$ is an integer, let h be the smallest integer $\ge \max\{1, \sqrt{j}\}$. If $h(h-1)+1 \le j \le h^2$, define σ_j to be the permutation

$$(0)\ (1\ 5)\ (2\ 9)\ (3)\ (4\ 7)\ (6\ 10)\ (8)\ (11);$$

otherwise define σ_j to be the permutation

$$(0\ 11)\ (1\ 6)\ (2)\ (3\ 8)\ (4)\ (5\ 10)\ (7)\ (9).$$

The alternating character of Σ plays a role in reducing the star discrepancy to [10]

$$S^*(X_{12}^{\Sigma}) = 1919/(3454\ln(12)) = 0.223584....$$

Again, this is the smallest such value currently known.

1.6.2 $\{n\alpha\}$-*Sequences*

Let $\alpha > 0$ be irrational. Define a sequence Y^{α} to have n^{th} term [19]

$$y_n = \{n\alpha\} = n\alpha \bmod 1,$$

that is, the fractional part of $n\alpha$. Let $D_N(\alpha) = D_N(Y^{\alpha})$ for convenience. It is known that $D_N(\alpha) \to 0$ as $N \to \infty$, just as for van der Corput sequences. The corresponding values of $S(\alpha)$ and $S^*(\alpha)$ are not as small as earlier, but are nevertheless interesting.

Dupain & Sós [20] proved that

$$\inf_{\alpha} S^*(\alpha) = S^*(\sqrt{2}) = \frac{1}{4\ln(1+\sqrt{2})} = 0.283648...$$

and Schoissengeier [21–24] expressed $S^*(\alpha)$ in terms of the continued fraction expansion of α. Baxa [25, 26] demonstrated that the image of the set of all irrational α under the map $\alpha \mapsto S^*(\alpha)$ is the ray $[S^*(\sqrt{2}), \infty]$, which contrasts sharply against the Lagrange and Markov spectra [27].

Ramshaw [28] proved that

$$S(\varphi) = \frac{1}{5\ln(\varphi)} = 0.415617...$$

where $\varphi = (1+\sqrt{5})/2$ is the Golden mean. A proof that

$$\inf_{\alpha} S(\alpha) = S(\varphi)$$

has never been published [29, 30]; a hole as such in the literature deserves to be filled.

1.6.3 *Self-Similar Sequences*

Terms $u(n) = \{(n-1)\varphi\}$ of the preceding sequence can be written as [31, 32]

$$u(n) = \begin{cases} 0 & \text{if } n = 1, \\ f_0\,(u(m)) & \text{if } n \geq 2 \text{ and } u(n) < \varphi - 1, \\ f_1\,(u(n-m)) & \text{if } n \geq 2 \text{ and } u(n) \geq \varphi - 1, \end{cases}$$

where

$$m = \#\,\{k : 1 \leq k \leq n \text{ and } 0 < u(k) \leq \varphi - 1\}$$

and f_0, f_1 are simple linear functions

$$f_0(x) = (\varphi - 1)(1 - x), \quad f_1(x) = (\varphi - 1) + (2 - \varphi)x$$

that map the interval $[0, 1)$ onto subintervals $I_0 = (0, \varphi - 1]$, $I_1 = [\varphi - 1, 1)$ respectively. For each $j \in \{0, 1\}$, the subsequence of terms $u(n_1), u(n_2), u(n_3), \ldots$ belonging to subinterval I_j with $n_1 < n_2 < n_3 < \ldots$ is the same as the image of the full sequence $u(1), u(2), u(3), \ldots$ under the function f_j. Equivalently, $u(n_\ell) = f_j(u(\ell))$ for all $\ell \geq 1$. Such *self-similar* sequences often possess low discrepancy; it is known that $S^*(U) = 3/(20\ln(\varphi)) = 0.311721...$ in this case [33]. Alternatively,

$$u(n) = \begin{cases} 0 & \text{if } n = 1, \\ f_0\left(u\left(n - \sum_{k=1}^{n} e_k\right)\right) & \text{if } n \geq 2 \text{ and } e_n = 0, \\ f_1\left(u\left(\sum_{k=1}^{n} e_k\right)\right) & \text{if } n \geq 2 \text{ and } e_n = 1, \end{cases}$$

where [34]

$$e_1 e_2 e_3 \ldots = 010100101001001010010100100101001001010010100100101001 0...$$

is the unique fixed point of the bit substitution $0 \mapsto 010$ and $1 \mapsto 10$ on infinite binary words. No periodicity in the digits

$$e_n = 1 - \lfloor (n-1)(\varphi - 1) \rfloor + \lfloor (n-2)(\varphi - 1) \rfloor$$

can be ascertained. By contrast, the van der Corput sequence X_2^{ξ} with cutoff $1/2$ (rather than $\varphi - 1$) has corresponding infinite word $\overline{01}$ (periodic with period 2).

A more complicated sequence is defined by

$$v(n) = \begin{cases} 0 & \text{if } n = 1, \\ g_0\left(v\left(n - \sum_{k=1}^{n} e_k\right)\right) & \text{if } n \geq 2 \text{ and } e_n = 0, \\ g_1\left(v\left(\sum_{k=1}^{n} e_k\right)\right) & \text{if } n \geq 2 \text{ and } e_n = 1 \end{cases}$$

where [31]

$$g_0(x) = \left(2 - \sqrt{2}\right)(1 - x), \quad g_1(x) = \left(2 - \sqrt{2}\right) + \left(-1 + \sqrt{2}\right)x$$

and here [35]

$$e_1 e_2 e_3 \ldots = 0101001010010101001010010101001010010100101010010100101...$$

is the fixed point of the substitution $0 \mapsto 01$ and $1 \mapsto 010$. Again, no periodicity in the digits

$$e_n = \left\lfloor (n+1)\left(\sqrt{2} - 1\right)\right\rfloor - \left\lfloor n\left(\sqrt{2} - 1\right)\right\rfloor$$

can be ascertained. A more straightforward representation of $v(n)$ does not seem to exist, although we mention

$$v(n) = P(n)\left(2 - \sqrt{2}\right) - Q(n) = \left\{P(n)\left(2 - \sqrt{2}\right)\right\}$$

where $P(1)$, $P(2)$, $P(3)$, ..., $Q(1)$, $Q(2)$, $Q(3)$, ... are integers satisfying [31]

$$\begin{pmatrix} P(n) \\ Q(n) \end{pmatrix} = \begin{cases} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \text{if } n = 1, \\ \begin{pmatrix} 1 + Q\left(n - \sum_{k=1}^{n} e_k\right) - 4P\left(n - \sum_{k=1}^{n} e_k\right) \\ -2P\left(n - \sum_{k=1}^{n} e_k\right) \end{pmatrix} & \text{if } n \geq 2 \text{ and } e_n = 0, \\ \begin{pmatrix} 1 + Q\left(\sum_{k=1}^{n} e_k\right) - 3P\left(\sum_{k=1}^{n} e_k\right) \\ Q\left(\sum_{k=1}^{n} e_k\right) - 2P\left(\sum_{k=1}^{n} e_k\right) \end{pmatrix} & \text{if } n \geq 2 \text{ and } e_n = 1; \end{cases}$$

$$\begin{aligned} \{P(n)\}_{n=2}^{\infty} = \{&1, -3, -2, 11, 7, 8, -37, 5, -23, -27, -26, 127, -16, -17, \\ &-19, 79, 93, \ldots\}, \end{aligned}$$

$$\begin{aligned} \{Q(n)\}_{n=2}^{\infty} = \{&0, -2, -2, 6, 4, 4, -22, 2, -14, -16, -16, 74, -10, -10, \\ &-12, 46, 54, \ldots\}. \end{aligned}$$

Borel [30, 36, 37] proved that $0.17451 < S^*(V) < 0.45696$ and conjectured that $S^*(V) = 3/\left(14\ln\left(1 + 1/\sqrt{2}\right)\right) = 0.400683....$ Owing to the elaborate details, resolving this issue may remain open for a long time. Other occurrences of recursive bit substitutions include [38–40].

1.6.4 *Erdős–Turán Inequality*

Erdős & Turán [41] proved that there exist constants c_1, c_2 such that

$$D_N(X) \le \frac{c_1}{K+1} + c_2 \sum_{k=1}^{K} \frac{1}{k} \left| \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i k x_n} \right|$$

for any positive integers N, K and any sequence $X \subseteq [0, 1)$. There is considerable flexibility in the choice of the two constants, as indicated here:

- $(c_1, c_2) = (6, \frac{4}{\pi})$ (Kuipers & Niederreiter [1])
- $(c_1, c_2) = (1, 3)$ (Baker [42] and Montgomery [43])
- $(c_1, c_2) = (1, 1)$ (Mauduit, Rivat & Sárközy [44]).

Rivat & Tenenbaum [45], building on the work of Vaaler [46], determined constants that are believed to be close to best for the Erdős–Turán inequality:

- $(c_1, c_2) = \left(1, \frac{2}{\pi}\gamma\right) = (1, 0.6527196578...)$
- $(c_1, c_2) = \left(1 + \xi, \frac{2}{\pi}\right) = (1.1434819845..., 0.6366197723...)$

where

$$f(t) = \sqrt{[\pi t(1-t)\cot(\pi t) + t]^2 + [\pi t(1-t)]^2}, \quad 0 \le t \le 1,$$

$$\gamma = \max_{0 \le t \le 1} f(t) = 1.0252896410...,$$

$$g(x, t) = \left(1 - 3x^2 + 3x^2 \left|\cos\left(\tfrac{\pi t}{3x}\right)\right|\right) f(t), \quad 0 \le x \le \tfrac{\sqrt{3}}{3},$$

and $\xi = 0.1434819845...$ is the smallest value of x for which $\max_{0 \le t \le 1} g(x, t) \le 1$. In fact, Rivat & Tenenbaum found a one-parameter *family* of admissible constants (c_1, c_2), but we have indicated only the endpoints of this family.

A similar set of formulas occur in the determination of close-to-best constants for the Berry–Esseen inequality. (This is a somewhat different version of the inequality from that discussed in [47].) Let F, G be two distribution functions with corresponding characteristic functions φ, ψ. Assume that G is differentiable and that $\sup_x |G'(x)| = M < \infty$. Then there exist constants c_1, c_2 such that

$$\sup_x |F(x) - G(x)| \le c_1 \frac{M}{T} + c_2 \int_{-T}^{T} \left| \frac{\varphi(t) - \psi(t)}{t} \right| dt$$

for all $T > 0$. Admissible values for these constants include

- $(c_1, c_2) = (\frac{24}{\pi}, \frac{1}{\pi})$ (Feller [48] and Loève [49])
- $(c_1, c_2) = (\pi, \frac{1}{2\pi}\gamma)$ (Vaaler [46] and Tenenbaum [50])

where $\gamma/(2\pi) = 0.1631799144...$ and γ is exactly as before. The new approach in [45] can perhaps be applied here too.

Addendum Larcher & Puchhammer [51, 52] proved that $S^*(X) \geq 0.065664679$ always while Ostromoukhov [53] proved that certain sequences X exist for which

$$S(X) \leq \frac{130}{83\ln(84)} = 0.353493..., \quad S^*(X) \leq \frac{32209}{35400\ln(60)} = 0.222223....$$

The gap between bounds remains large!

Let us reformulate matters while providing a sample two-dimensional result [54–56]. Given $p > 0$ and W a finite set of N points in the unit square, define

$$\Delta^*_{p,N}(W) = \left(\int_0^1 \int_0^1 \left| \frac{|W \cap \{[0,x) \times [0,y)\}|}{N} - x\,y \right|^p dx\,dy \right)^{1/p}.$$

The star discrepancy

$$\delta^*_{p,N} = \frac{N}{\sqrt{\ln(N)}} \cdot \inf_W \Delta^*_{p,N}(W)$$

here satisfies [57, 58]

$$\liminf_{N\to\infty} \delta^*_{2,N} \geq \frac{7}{216\sqrt{\ln(2)}} = 0.038925...,$$

$$\limsup_{N\to\infty} \delta^*_{1,N} \geq \frac{3}{64\sqrt{\ln(2)}} = 0.056302...,$$

$$\liminf_{N\to\infty} \delta^*_{1,N} \geq \frac{1}{1152\sqrt{e\ln(2)}} = 0.000632...$$

and $\limsup_{N\to\infty} \delta^*_{1,N}$ is conjectured to be strictly greater than $\liminf_{N\to\infty} \delta^*_{1,N}$. See [59, 60] for central behavior of $\Delta^*_{p,N}(W)$ rather than extreme behavior, where points in W are independent and uniformly distributed.

[1] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, 1974; MR0419394 (54 #7415).

[2] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, 1992; MR1172997 (93h:65008).

[3] M. Drmota and R. F. Tichy, *Sequences, Discrepancies and Applications*, Lect. Notes in Math. 1651, Springer-Verlag, 1997; MR1470456 (98j:11057).

[4] R. Béjian, Minoration de la discrépance à l'origine d'une suite quelconque, *Annales Faculté Sciences Toulouse Math.* 1 (1979) 201–213; MR0568146 (82b:10076).

[5] R. Béjian, Minoration de la discrépance d'une suite quelconque sur T, *Acta Arith.* 41 (1982) 185–202; MR0674832 (83k:10101).

[6] T. van Aardenne-Ehrenfest, Proof of the impossibility of a just distribution of an infinite sequence of points over an interval, *Proc. Konink. Nederl. Akad. Wetensch. Sci. Sect.* 48 (1945) 266–271; *Indag. Math.* 7 (1945) 71–76; MR0015143 (7,376l).

[7] T. van Aardenne-Ehrenfest, On the impossibility of a just distribution, *Proc. Konink. Nederl. Akad. Wetensch. Sci. Sect.* 52 (1949) 734–739; *Indag. Math.* 11 (1949) 264–269; MR0032717 (11,336d).
[8] K. F. Roth, On irregularities of distribution, *Mathematika* 1 (1954) 73–79; MR0066435 (16,575c).
[9] W. M. Schmidt, Irregularities of distribution. VII, *Acta Arith.* 21 (1972) 45–50; MR0319933 (47 #8474).
[10] H. Faure, Discrépances de suites associées à un système de numération (en dimension un), *Bull. Soc. Math. France* 109 (1981) 143–182; MR0623787 (82i:10069).
[11] A. Thomas, Discrépance en dimension un, *Annales Faculté Sciences Toulouse Math.* 10 (1989) 369–399; MR1425493 (97g:11083).
[12] H. Faure, Good permutations for extreme discrepancy, *J. Number Theory* 42 (1992) 47–56; MR1176419 (93j:11049).
[13] M. H. DeGroot, *Probability and Statistics*, Addison-Wesley, 1975, pp. 463–471.
[14] J. D. Gibbons, *Nonparametric Statistical Inference*, McGraw-Hill, 1971, pp. 75–85; MR0286223 (44 #3437).
[15] S. R. Finch, Variants of Brownian motion, *this volume*, §4.5.
[16] S. Haber, On a sequence of points of interest for numerical quadrature, *J. Res. Nat. Bur. Standards B* 70 (1966) 127–136; MR0203938 (34 #3785).
[17] R. Béjian and H. Faure, Discrépance de la suite de van der Corput, *C. R. Acad. Sci. Paris Sér. A-B* 285 (1977) A313–A316; MR0444600 (56 #2950).
[18] R. Béjian and H. Faure, Discrépance de la suite de van der Corput, *Séminaire Delange-Pisot-Poitou (Théorie des nombres)*, 19e année, 1977/78, n. 13, pp. 1–14; MR0520308 (80b:10078).
[19] S. R. Finch, Pisot-Vijayaraghavan-Salem constants, *first volume*, pp. 192–199.
[20] Y. Dupain and V. T. Sós, On the discrepancy of $(n\alpha)$ sequences, *Topics in Classical Number Theory*, v. I, Proc. 1981 Budapest conf., Colloq. Math. Soc. János Bolyai 34, North-Holland, 1984, pp. 355–387; MR0781147 (87c:11065).
[21] J. Schoissengeier, On the discrepancy of $(n\alpha)$, *Acta Arith.* 44 (1984) 241–279; MR0774103 (86c:11056).
[22] J. Schoissengeier, On the discrepancy of $(n\alpha)$. II, *J. Number Theory* 24 (1986) 54–64; MR0852190 (88d:11074).
[23] J. Schoissengeier, The discrepancy of $(n\alpha)_{n\geq 1}$, *Math. Annalen* 296 (1993) 529–545; MR1225989 (94e:11086).
[24] C. Baxa and J. Schoissengeier, Minimum and maximum order of magnitude of the discrepancy of $(n\alpha)$, *Acta Arith.* 68 (1994) 281–290; MR1308128 (95j:11073).
[25] C. Baxa, On the discrepancy of the sequence $(n\alpha)$, *J. Number Theory* 55 (1995) 94–107; MR1361562 (96i:11086).
[26] C. Baxa, On the distribution of the sequence $(n\alpha)$ with transcendental α, *Acta Arith.* 81 (1997) 357–363; MR1472815 (98m:11080).
[27] S. R. Finch, Freiman's constant, *first volume*, pp. 199–203.
[28] L. Ramshaw, On the discrepancy of the sequence formed by the multiples of an irrational number, *J. Number Theory* 13 (1981) 138–175; MR0612680 (82k:10071).
[29] C. Baxa, Comparing the distribution of $(n\alpha)$-sequences, *Acta Arith.* 94 (2000) 345–363; MR1779947 (2001e:11086).
[30] J.-P. Borel, Symbolic representation of piecewise linear functions on the unit interval and application to discrepancy, *Theoret. Comput. Sci.* 123 (1994) 61–87; MR1257204 (94m:11030).
[31] J.-P. Borel, *Contribution a l'etude de la distribution des suites*, Ph.D. thesis, Marseille-Luminy, 1989.

[32] J.-P. Borel, Self-similar measures and sequences, *J. Number Theory* 31 (1989) 208–241; MR987574 (90f:11064).
[33] Y. Dupain, Discrépance de la suite $(\{n\alpha\})$, $\alpha = \left(1+\sqrt{5}\right)/2$, *Annales Inst. Fourier (Grenoble)* 29 (1979) 81–106; MR0526778 (80f:10061).
[34] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A189661 and A189576 (the latter classifies several 01-sequences fixed by morphisms).
[35] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A159684 and A188037 (the former cites a relevant family of Sturmian words).
[36] J.-P. Borel, Sur une suite de fonctions liées à un procédé itératif. I, *Publ. Dépt. Math. Limoges* 12 (1990) 1–16.
[37] J.-P. Borel, Sur une suite de fonctions liées à un procédé itératif. II, *J. Théor. Nombres Bordeaux* 5 (1993) 235–261; MR1265904 (95h:11072).
[38] S. R. Finch, Prouhet-Thue-Morse constant, *first volume*, pp. 436–441.
[39] S. R. Finch, Molteni's composition constant, *this volume*, §1.40.
[40] S. R. Finch, Substitution dynamics, *this volume*, §4.24.
[41] P. Erdős and P. Turán, On a problem in the theory of uniform distribution. II, *Proc. Konink. Nederl. Akad. Wetensch. Sci. Sect.* 51 (1948) 1262–1269; *Indag. Math.* 10 (1948) 406–413; MR0027896 (10,372d).
[42] R. C. Baker, *Diophantine Inequalities*, Oxford Univ. Press, 1986, p. 19; MR0865981 (88f:11021).
[43] H. L. Montgomery, *Ten Lectures on the Interface between Analytic Number Theory and Harmonic Analysis*, Amer. Math. Soc., 1994, p. 8; MR1297543 (96i:11002).
[44] C. Mauduit, J. Rivat and A. Sárközy, On the pseudo-random properties of n^c, *Illinois J. Math.* 46 (2002) 185–197; MR1936084 (2003h:11088).
[45] J. Rivat and G. Tenenbaum, Constantes d'Erdős-Turán, *Ramanujan J.* 9 (2005) 111–121; MR2166382 (2006g:11158).
[46] J. D. Vaaler, Some extremal functions in Fourier analysis, *Bull. Amer. Math. Soc.* 12 (1985) 183–216; MR0776471 (86g:42005).
[47] S. R. Finch, Berry-Esseen constant, *first volume*, pp. 264–266.
[48] W. Feller, *An Introduction to Probability Theory and its Applications*, v. II., 2nd ed., Wiley, 1971, pp. 536–538, 542–546; MR0270403 (42 #5292).
[49] M. Loève, *Probability Theory. I*, 4th ed., Springer-Verlag, 1977, p. 297; MR0651017 (58 #31324a).
[50] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Univ. Press, 1995, pp. 240–241; MR1342300 (97e:11005b).
[51] G. Larcher, On the star discrepancy of sequences in the unit interval, *J. Complexity* 31 (2015) 474–485; arXiv:1407.2094; MR3325684.
[52] G. Larcher and F. Puchhammer, An improved bound for the star discrepancy of sequences in the unit interval, arXiv:1511.03869.
[53] V. Ostromoukhov, Recent progress in improvement of extreme discrepancy and star discrepancy of one-dimensional sequences, *Monte Carlo and Quasi-Monte Carlo Methods 2008*, Proc. Univ. of Montreal conf.,.ed. P. L'Ecuyer and A. B. Owen, Springer-Verlag, 2009, pp. 561–572; MR2743919 (2011k:65010).
[54] G. Halász, On Roth's method in the theory of irregularities of point distributions, *Recent Progress in Analytic Number Theory*, v. 2, Proc. 1979 Durham sympos., ed. H. Halberstam and C. Hooley, Academic Press, 1981, pp. 79–94; MR0637361 (83e:10072).
[55] J. Beck and W. W. L. Chen, *Irregularities of Distribution*, Cambridge Univ. Press, 1987, pp. 3–7, 20–23; MR0903025 (88m:11061).
[56] D. Bilyk, On Roth's orthogonal function method in discrepancy theory, *Unif. Distrib. Theory* 6 (2011) 143–184; MR2817765 (2012e:11141).

[57] A. Hinrichs and L. Markhasin, On lower bounds for the L_2-discrepancy, *J. Complexity* 27 (2011) 127–132; MR2776487 (2011m:60099).
[58] A. Vagharshakyan, Lower bounds for L_1 discrepancy, *Mathematika* 59 (2013) 365–379; arXiv:1209.2398; MR3081776.
[59] S. Steinerberger, The asymptotic behavior of the average L^p-discrepancies and a randomized discrepancy, *Elec. J. Combin.* 17 (2010) R106; MR2679560 (2011h:11082).
[60] A. Hinrichs and H. Weyhausen, Asymptotic behavior of average L_p-discrepancies, *J. Complexity* 28 (2012) 425–439; MR2925899.

1.7 Unitarism and Infinitarism

We will examine variations of four famous arithmetical functions. For a given function χ, let χ^* denote its unitary analog, $\tilde{\chi}$ its square-free analog, and χ' its unitary square-free analog. The meanings of these phrases will be made clear in each case. At the end, the infinitary analog χ_∞ will appear as well.

1.7.1 *Divisor Function*

If $d(n)$ is the number of distinct divisors of n, then

$$\sum_{n=1}^{N} d(n) = N\ln(N) + (2\gamma - 1)N + O(\sqrt{N})$$

as $N \to \infty$, where γ is the Euler–Mascheroni constant. Let us introduce a more refined notion of divisibility. A divisor k of n is **unitary** if k and n/k are coprime, that is, if $\gcd(k, n/k) = 1$. This condition is often written as $k||n$. The number $d^*(n)$ of unitary divisors of n is $2^{\omega(n)}$, where $\omega(n)$ is the number of distinct prime factors of n. This fact is easily seen to be true: If $p_1^{a_1} p_2^{a_2} \cdots p_r^{a_r}$ is the prime factorization of n, then the unitary divisors of n are of the form $p_1^{\varepsilon_1 a_1} p_2^{\varepsilon_2 a_2} \cdots p_r^{\varepsilon_r a_r}$, where each ε_s is either 0 or 1. There are 2^r possible choices for the r-tuple $(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r)$; hence the result follows. We have [1–5]

$$\sum_{n=1}^{N} d^*(n) = \frac{6}{\pi^2} N\ln(N) + \frac{6}{\pi^2}\left(2\gamma - 1 - \frac{12}{\pi^2}\zeta'(2)\right) N + O(\sqrt{N}),$$

where $\zeta(x)$ is the Riemann zeta function and $\zeta'(x)$ is its derivative.

A divisor k of n is **square-free** if k is divisible by no square exceeding 1. The number $\tilde{d}(n)$ of square-free divisors of n is also $2^{\omega(n)}$; the divisors in this case are of the form $p_1^{\varepsilon_1} p_2^{\varepsilon_2} \cdots p_r^{\varepsilon_r}$. Therefore the same asymptotics apply for $\tilde{d}(n)$, but the underlying sets of numbers overlap only somewhat [6].

Define $d'(n)$ to be the number of unitary square-free divisors of n. A more complicated asymptotic formula arises here [7, 8]:

$$\sum_{n=1}^{N} d'(n) = \frac{6\alpha}{\pi^2} N\ln(N) + \frac{6\alpha}{\pi^2}\left(2\gamma - 1 - \frac{12}{\pi^2}\zeta'(2) + X\right) N + O(\sqrt{N}\ln(N))$$

where

$$\alpha = \prod_p \left(1 - \frac{1}{p(p+1)}\right) = 0.7044422009...,$$

$$X = \sum_p \frac{(2p+1)\ln(p)}{(p+1)(p^2+p-1)} = 0.7483723334...$$

and we agree that the product and sum extend over all primes p. The constant α is the same as what is called $\pi^2 P/6$ in [9]. Calculation of X is found in [10].

We finally give corresponding reciprocal sums [11–15]:

$$\begin{aligned}\lim_{N\to\infty} \frac{\sqrt{\ln(N)}}{N} \sum_{n=1}^{N} \frac{1}{d(n)} &= \frac{1}{\sqrt{\pi}} \prod_p \sqrt{p(p-1)} \ln\left(\frac{p}{p-1}\right) \\ &= \frac{0.9692769438...}{\sqrt{\pi}} = 0.5468559552...,\end{aligned}$$

$$\begin{aligned}\lim_{N\to\infty} \frac{\sqrt{\ln(N)}}{N} \sum_{n=1}^{N} \frac{1}{d^*(n)} &= \frac{1}{\sqrt{\pi}} \prod_p \sqrt{1 + \frac{1}{4p(p-1)}} = \frac{1.0969831191...}{\sqrt{\pi}} \\ &= 0.6189064491....\end{aligned}$$

The former sum was mentioned in [16] with regard to the arcsine law for random divisors. It is not known what constant emerges for $1/d'(n)$; an analog of $d'(n)$, corresponding to unitary cube-free divisors of n, can be studied [8, 10].

1.7.2 *Sum-of-Divisors Function*

If $\sigma(n)$ is the sum of all distinct divisors of n, then

$$\sum_{n=1}^{N} \sigma(n) = \frac{\pi^2}{12} N^2 + O(N \ln(N))$$

as $N \to \infty$. Let $\sigma^*(n)$ be the sum of unitary divisors of n and $\tilde{\sigma}(n)$ be the sum of square-free divisors of n. Although $d^*(n) = \tilde{d}(n)$ always, it is usually false that $\sigma^*(n) = \tilde{\sigma}(n)$ [17]. We have [18–21]

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{n=1}^{N} \sigma^*(n) = \frac{\pi^2}{12\zeta(3)}, \quad \lim_{N\to\infty} \frac{1}{N^2} \sum_{n=1}^{N} \tilde{\sigma}(n) = \frac{1}{2}.$$

Further, if $\sigma'(n)$ is the sum of unitary square-free divisors of n, then [18]

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{n=1}^{N} \sigma'(n) = \frac{1}{2} \prod_p \left(1 - \frac{1}{p^2(p+1)}\right) = \frac{0.8815138397...}{2},$$

a constant which appeared in [22] and turns out to be connected with class number theory [23–25].

Corresponding reciprocal sums are [26, 27]

$$\sum_{n=1}^{N}\frac{1}{\sigma(n)}\sim Y_1\ln(N)+Y_1(\gamma+Y_2),\quad \sum_{n=1}^{N}\frac{1}{\sigma^*(n)}\sim Y_3\ln(N)+Y_3(\gamma+Y_4-Y_5)$$

where

$$\begin{aligned} Y_1 &= \prod_p f(p), \quad Y_2=\sum_p \frac{(p-1)^2 g(p)\ln(p)}{p f(p)}, \\ Y_3 &= \prod_p\left(1-\frac{1}{p}\right)\left(1+\sum_{k=1}^{\infty}\frac{1}{p^k+1}\right)=\prod_p h(p), \\ Y_4 &= \sum_p\left(\frac{(p-1)\ln(p)}{p\,h(p)}\sum_{j=1}^{\infty}\frac{j}{p^j(p^j+1)}\right) \end{aligned}$$

where

$$f(p)=1-\frac{(p-1)^2}{p}\sum_{j=1}^{\infty}\frac{1}{(p^j-1)(p^{j+1}-1)},\quad g(p)=\sum_{j=1}^{\infty}\frac{j}{(p^j-1)(p^{j+1}-1)},$$

$$h(p)=1-\frac{p-1}{p}\sum_{j=1}^{\infty}\frac{1}{p^j(p^j+1)}$$

and Y_5 is a similarly complicated expression. An estimate 0.6728... for Y_1 appears in [28]; other values of Y_i remain open. No one seems to have examined $1/\tilde{\sigma}(n)$ or $1/\sigma'(n)$ yet.

1.7.3 *Totient Function*

If $\varphi(n)$ is the number of positive integers $k\le n$ satisfying $\gcd(k,n)=1$, then [29, 30]

$$\sum_{n=1}^{N}\varphi(n)=\frac{3}{\pi^2}N^2+O(N\ln(N))$$

as $N\to\infty$. Define $\gcd_*(k,n)$ to be the greatest divisor of k that is also a unitary divisor of n. Let $\varphi^*(n)$ be the number of positive integers $k\le n$ satisfying $\gcd_*(k,n)=1$. Since $\gcd_*$ is never larger than gcd, it follows that φ^* is at least as large as φ. Also let $\tilde{\varphi}(n)$ be the number of positive square-free integers $k\le n$ satisfying $\gcd(k,n)=1$. We have [18, 31]

$$\lim_{N\to\infty}\frac{1}{N^2}\sum_{n=1}^{N}\varphi^*(n)=\frac{1}{2}\alpha,\quad \lim_{N\to\infty}\frac{1}{N^2}\sum_{n=1}^{N}\tilde{\varphi}(n)=\frac{3}{\pi^2}\alpha$$

where α is as defined earlier. The case for $\varphi'(n)$ remains open.

Corresponding reciprocal sums are [26, 27, 32]

$$\sum_{n=1}^{N} \frac{1}{\varphi(n)} \sim Z_1 \ln(N) + Z_1(\gamma - Z_2),$$
$$\sum_{n=1}^{N} \frac{1}{\varphi^*(n)} \sim Z_3 \ln(N) + Z_3(\gamma - Z_4 + Z_5 + Z_6)$$

where

$$Z_1 = \frac{315\zeta(3)}{2\pi^4}, \quad Z_2 = \sum_p \frac{\ln(p)}{p^2 - p + 1},$$

$$Z_3 = \prod_p \left(1 - \frac{1}{p}\right)\left(1 + \sum_{k=1}^{\infty} \frac{1}{p^k - 1}\right) = \prod_p u(p),$$

$$Z_4 = \sum_p \left(\frac{(p-1)\ln(p)}{p\,u(p)} \sum_{j=1}^{\infty} \frac{j}{p^j(p^j - 1)}\right)$$

where

$$u(p) = 1 + \frac{p-1}{p} \sum_{j=1}^{\infty} \frac{1}{p^j(p^j - 1)}, \quad v(p) = \sum_{j=1}^{\infty} \frac{1}{p^j(p^{j+1} - 1)}$$

and Z_5, Z_6 are similarly complicated expressions.

1.7.4 *Square-Free Core Function*

If $\tilde{\kappa}(n)$ is the maximal square-free divisor of n (also called [9] the square-free kernel of n), then [18, 20, 21, 33–35]

$$\sum_{n=1}^{N} \tilde{\kappa}(n) = \frac{\alpha}{2} N^2 + O\left(N^{3/2}\right)$$

as $N \to \infty$, where α is as before. Assuming the Riemann Hypothesis, the error term can be improved to $O(N^{7/5+\varepsilon})$ for any $\varepsilon > 0$. If $\kappa'(n)$ is the maximal unitary square-free divisor of n, then [34, 35]

$$\sum_{n=1}^{N} \kappa'(n) = \frac{\beta}{2} N^2 + O\left(N^{3/2}\right)$$

where

$$\beta = \prod_p \left(1 - \frac{p^2 + p - 1}{p^3(p+1)}\right) = 0.6496066993\ldots.$$

1.7.5 *Infinitary Arithmetic*

We continue refining the notion of divisibility [36, 37]. A divisor k of n is **biunitary** if the greatest common unitary divisor of k and n/k is 1, and **triunitary** if the greatest common biunitary divisor of k and n/k is 1. More generally, for any positive integer m, a divisor k of n is m**-ary** if the greatest common $(m-1)$**-ary** divisor of k and n/k is 1. We write $k|_m n$. Clearly $1|_m n$ and $n|_m n$.

When introducing infinitary divisors, it is best to start with prime powers. Let p be a prime, and let $x \geq 0$, $y \geq 1$ be integers. It can be proved that, for any $m \geq y-1$, $p^x|_m p^y$ if and only if $p^x|_{y-1} p^y$. Thus we define $p^x|_\infty p^y$ if $p^x|_{y-1} p^y$. For fixed y, the number of integers $0 \leq x \leq y$ satisfying $p^x|_\infty p^y$ is $2^{b(y)}$, where $b(y)$ is the number of ones in the binary expansion of y. Define as well $1|_\infty 1$. The sum $\sum_{y=0}^{z-1} 2^{b(y)}$ is approximately $z^{\ln(3)/\ln(2)}$ but is not well behaved asymptotically [38].

We now allow n to be arbitrary. A divisor k of n is **infinitary** if, for any prime p, the conditions $p^x||k$ and $p^y||n$ imply that $p^x|_\infty p^y$. We write $k|_\infty n$. Clearly $1|_\infty n$ and $n|_\infty n$. Each $n > 1$ has a unique factorization as a product of distinct elements from the set

$$I = \left\{ p^{2^j} : p \text{ is prime and } j \geq 0 \right\};$$

each element of I in this product is called an I-component of n. It follows that $k|_\infty n$ if and only if every I-component of k is also an I-component of n.

Assume that $n = P_1 P_2 \cdots P_t$, where $P_1 < P_2 < \cdots < P_t$ are the I-components of n. The infinitary analogs of the functions d and σ are defined by [39, 40]

$$d_\infty(n) = 2^t, \quad \sigma_\infty(n) = \prod_{i=1}^{t} (P_i + 1),$$

for $n > 1$; otherwise $d_\infty(1) = \sigma_\infty(1) = 1$. Two infinitary analogs of the function φ are known:

$$\varphi_\infty(n) = \text{ the number of positive integers } k \leq n \text{ satisfying } \gcd_\infty(k, n) = 1;$$

$$\hat{\varphi}_\infty(n) = \prod_{i=1}^{t} (P_i - 1) = n \prod_{i=1}^{t} \left(1 - \frac{1}{P_i}\right) \text{ for } n > 1, \quad \hat{\varphi}_\infty(1) = 1.$$

It is generally untrue that $\varphi_\infty(n) = \hat{\varphi}_\infty(n)$. No similar extension of the function $\tilde{\kappa}$ is known. Cohen & Hagis [39, 41] proved that

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{n=1}^{N} \sigma_\infty(n) = \frac{A}{2} = 0.7307182421...,$$

$$\lim_{N\to\infty} \frac{1}{N^2} \sum_{n=1}^{N} \hat{\varphi}_\infty(n) = \frac{B}{2} = 0.3289358388...,$$

$$\frac{1}{N^2} \sum_{n=1}^{N} d_\infty(n) \sim CN\ln(N) + DN \sim 2(0.3666252769...)N\ln(N)$$

where

$$A=\prod_{P\in I}\left(1+\frac{1}{P(P+1)}\right),\quad B=\prod_{P\in I}\left(1-\frac{1}{P(P+1)}\right),\quad C=\prod_{P\in I}\left(1-\frac{1}{(P+1)^2}\right)$$

but no such expression for D yet exists. It is known that $\varphi_\infty(n)=n^2/\sigma_\infty(n)+O(n^\varepsilon)$ for any $\varepsilon>0$; reciprocal sums involving d_∞, σ_∞ and $\hat{\varphi}_\infty$ also remain open. Alternative generalizations of unitary divisor have been given [42, 43] but will not be discussed here.

1.7.6 *Coprimality*

The probability that k randomly chosen integers are unitary coprime is [44]

$$\prod_p\left(1-\frac{(p-1)^k}{p^k(p^k-1)}\right).$$

The probability that they are *pairwise* unitary coprime is more complicated: for instance, it is

$$\zeta(2)\zeta(3)\prod_p\left(1-\frac{4}{p^2}+\frac{7}{p^3}-\frac{9}{p^4}+\frac{8}{p^5}-\frac{2}{p^6}-\frac{3}{p^7}+\frac{2}{p^8}\right)$$

when $k=3$ and

$$\zeta(2)^2\zeta(3)\zeta(4)\prod_p\left(1-\frac{8}{p^2}+\frac{3}{p^3}+\frac{27}{p^4}-\frac{24}{p^5}-\frac{14}{p^6}-\frac{3}{p^7}+\frac{37}{p^8}-\frac{30}{p^9}\right.$$
$$\left.+\frac{42}{p^{10}}-\frac{33}{p^{11}}-\frac{41}{p^{12}}+\frac{78}{p^{13}}-\frac{44}{p^{14}}+\frac{9}{p^{15}}\right)$$

when $k=4$. Expressions for arbitrary k appear in [44].

[1] F. Mertens, Über einige asymptotische Gesetze der Zahlentheorie, *J. Reine Angew. Math.* 77 (1874) 289–338.
[2] E. Cohen, The number of unitary divisors of an integer, *Amer. Math. Monthly* 67 (1960) 879–880; MR0122790 (23 #A124).
[3] A. A. Gioia and A. M. Vaidya, The number of squarefree divisors of an integer, *Duke Math. J.* 33 (1966) 797–799; MR0202678 (34 #2538).
[4] D. Suryanarayana and V. Siva Rama Prasad, The number of k-free divisors of an integer, *Acta Arith.* 17 (1970/71) 345–354; MR0327697 (48 #6039).
[5] R. C. Baker, The square-free divisor problem, *Quart. J. Math* 45 (1994) 269–277; part II, *Quart. J. Math* 47 (1996) 133–146; MR1295577 (95h:11098) and MR1397933 (97f:11080).
[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000005, A034444, and A056671.
[7] D. Suryanarayana, The number of unitary, square-free divisors of an integer. I, II, *Norske Vid. Selsk. Forh. (Trondheim)* 42 (1969) 6–21; MR0271048 (42 #5931).

[8] D. Suryanarayana and V. Siva Rama Prasad, The number of k-ary, $(k+1)$-free divisors of an integer, *J. Reine Angew. Math.* 276 (1975) 15–35; MR0376574 (51 #12749).
[9] S. R. Finch, Hafner-Sarnak-McCurley constant: Carefree couples, *first volume*, pp. 110–112.
[10] R. J. Mathar, Series of reciprocal powers of k-almost primes, arXiv:0803.0900.
[11] S. Ramanujan, Some formulae in the analytic theory of numbers, *Messenger of Math.* 45 (1916) 81–84; also in *Collected Papers*, ed. G. H. Hardy, P. V. Seshu Aiyar, and B. M. Wilson, Cambridge Univ. Press, 1927, pp. 133–135, 339–340.
[12] H. L. Abbott and M. V. Subbarao, On the distribution of the sequence $\{nd^*(n)\}$, *Canad. Math. Bull.* 32 (1989) 105–108; MR0996130 (90e:11136).
[13] J.-M. De Koninck and J. Grah, Arithmetic functions and weighted averages, *Colloq. Math.* 79 (1999) 249–272; MR1670205 (2000b:11111).
[14] P. Sebah, Approximation of several prime products, unpublished note (2004).
[15] M. Mazur and B. V. Petrenko, Representations of analytic functions as infinite products and their application to numerical computations, *Ramanujan J.* 34 (2014) 129–141; arXiv:1202.1335; MR3210260.
[16] S. R. Finch, Multiples and divisors, *this volume*, §1.5.
[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000203, A034448, A048250 and A092261.
[18] E. Cohen, Arithmetical functions associated with the unitary divisors of an integer, *Math. Z.* 74 (1960) 66–80; MR0112861 (22 #3707).
[19] R. Sitaramachandrarao and D. Suryanarayana, On $\sum_{n\leq x}\sigma^*(n)$ and $\sum_{n\leq x}\varphi^*(n)$, *Proc. Amer. Math. Soc.* 41 (1973) 61–66; MR0319922 (47 #8463).
[20] S. Wigert, Sur quelques formules asymptotiques de la théorie des nombres, *Ark. Mat. Astron. Fysik*, v. 22 (1931) n. 6, 1–6; v. 25 (1934) n. 3, 1–6.
[21] D. Suryanarayana, On the core of an integer, *Indian J. Math.* 14 (1972) 65–74; MR0330078 (48 #8417).
[22] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.
[23] M. R. Khan, A variation on the divisor functions $\sigma_a(n)$, unpublished note (2003).
[24] K. A. Broughan, Restricted divisor sums, *Acta Arith.* 101 (2002) 105–114; MR1880301 (2002k:11155).
[25] S. R. Finch, Class number theory, *this volume*, §1.12.
[26] V. Sita Ramaiah and D. Suryanarayana, Sums of reciprocals of some multiplicative functions, *Math. J. Okayama Univ.* 21 (1979) 155–164; MR0554306 (82d:10064a).
[27] V. Sita Ramaiah and D. Suryanarayana, Sums of reciprocals of some multiplicative functions. II, *Indian J. Pure Appl. Math*. 11 (1980) 1334–1355; MR0591410 (82d:10064b).
[28] L. G. Fel, Summatory multiplicative arithmetic functions: scaling and renormalization, arXiv:1108.0957.
[29] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000010, A047994, A073311 and A091813.
[30] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.
[31] E. Kowalski, Asymptotic mean of square-free totients, unpublished note (2004).
[32] V. Sita Ramaiah and M. V. Subbarao, Asymptotic formulae for sums of reciprocals of some multiplicative functions, *J. Indian Math. Soc.* 57 (1991) 153–167; MR1161332 (93b:11122).
[33] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007947 and A055231.
[34] E. Cohen, An elementary method in the asymptotic theory of numbers, *Duke Math. J.* 28 (1961) 183–192; MR0140496 (25 #3916).

[35] E. Cohen, Some asymptotic formulas in the theory of numbers, *Trans. Amer. Math. Soc.* 112 (1964) 214–227; MR0166181 (29 #3458).
[36] G. L. Cohen, On an integer's infinitary divisors, *Math. Comp.* 54 (1990) 395–411; MR0993927 (90e:11011).
[37] P. Hagis and G. L. Cohen, Infinitary harmonic numbers, *Bull. Austral. Math. Soc.* 41 (1990) 151–158; MR1043976 (91d:11001).
[38] S. R. Finch, Stolarsky-Harborth constant, *first volume*, pp. 145–151.
[39] G. L. Cohen and P. Hagis, Arithmetic functions associated with the infinitary divisors of an integer, *Internat. J. Math. Math. Sci.* 16 (1993) 373–383; MR1210848 (94c:11002).
[40] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A037445, A049417, A050376, A064380 and A091732.
[41] P. Sebah, Approximation of sums over $\{p^{2^j} : p \text{ prime}, j \geq 0\}$, unpublished note (2004).
[42] D. Suryanarayana, The number of k-ary divisors of an integer, *Monatsh. Math.* 72 (1968) 445–450; MR0236130 (38 #4428).
[43] K. Alladi, On arithmetic functions and divisors of higher order, *J. Austral. Math. Soc. Ser. A* 23 (1977) 9–27; MR0439721 (55 #12607).
[44] L. Tóth, Multiplicative arithmetic functions of several variables: a survey, *Mathematics Without Boundaries*, ed. T. M. Rassias and P. M. Pardalos, Springer-Verlag, 2014, pp. 483–514; arXiv:1310.7053; MR3330713.

1.8 Erdős' Minimum Overlap Problem

Let A, B be disjoint, complementary subsets of the set $\{1, 2, 3, \ldots, 2n\}$ with cardinality $|A| = |B| = n$. Let M_k denote the number of solutions of the equation $a_i - b_j = k$, where k is an integer between $-2n$ and $2n$. Define

$$M(n) = \min_{A,B} \max_{k} M_k.$$

We wish to estimate $M(n)$ as n grows large [1–3]. The work of Erdős, Scherk and others [4–6] implies that

$$\mu_L = \liminf_{n\to\infty} \frac{M(n)}{n} \geq \sqrt{4 - \sqrt{15}} > 0.35639$$

and specific examples [7] provide that

$$\mu_R = \limsup_{n\to\infty} \frac{M(n)}{n} \leq \frac{2}{5} = 0.4.$$

Haugland [6, 8] demonstrated that $\mu_L = \mu_R$ (meaning that the limit exists) and, using a theorem of Swinnerton-Dyer, obtained the improvement

$$0.35639 < \mu = \lim_{n\to\infty} \frac{M(n)}{n} < 0.38093.$$

No one has conjectured an exact value for this limiting ratio.

Observe that M_{-k} is the cardinality of the set $A_k \cap B$, where A_k is the translated set $\{a+k : a \in A\}$. Mycielski and Świerczkowski [4] considered a continuous analog of Erdős' problem. Let X, Y be disjoint, complementary measurable subsets of the interval $[0, 1]$ with Lebesgue measure $|X| = |Y| = 1/2$. It is not surprising that

$$\inf_{X,Y} \sup_t |X_t \cap Y| = \frac{\mu}{2}$$

where X_t is the translated set $\{x+t : x \in X\}$. Hence the bounds $0.17819 < \mu/2 < 0.19047$ carry over from before.

Moser and Murdeshwar [9–11] studied the following generalization. Let f, g be Lebesgue integrable functions on $\mathbb{R}$ satisfying

$$0 \le f(x) \le 1 \quad \text{for } 0 \le x \le 1, \quad f(x) = 0 \quad \text{otherwise};$$

$$0 \le g(x) \le 1 \quad \text{for } 0 \le x \le 1, \quad g(x) = 0 \quad \text{otherwise};$$

$$\int_0^1 f(x)\,dx = \frac{1}{2} = \int_0^1 g(x)\,dx.$$

(This scenario reduces to the preceding case by taking f to be the characteristic function of X and g to be the characteristic function of Y; clearly $f(x) + g(x) = 1$ for all $0 \le x \le 1$.) Define

$$\lambda = \inf_{f,g} \sup_t \int_0^1 f(x+t)\,g(x)\,dx.$$

It is known [11] that $0.136 \le \lambda \le 0.166$, but it is not presently known whether Swinnerton-Dyer's theorem [6] can be applied here (in some extended form) to improve these bounds.

Here is a related problem due to Czipszer [3, 12]. Let $\tilde{a}_1 < \tilde{a}_2 < \tilde{a}_3 < \cdots < \tilde{a}_n$ be arbitrary integers and define $\tilde{A}_k = \{\tilde{a}_j + k : 1 \le j \le n\}$ for each integer k. Let $\tilde{M}_k$ denote the cardinality $|\tilde{A}_k - \tilde{A}_0|$, that is, the number of elements of $\tilde{A}_k$ not in $\tilde{A}_0$. Define

$$\tilde{M}(n) = \min_{\tilde{A}} \max_{-n \le k \le n} \tilde{M}_k$$

and $\tilde{\mu}_L$, $\tilde{\mu}_R$ as earlier. It is known that $1/2 \le \tilde{M}(n)/n \le 2/3$ and, further, that $\tilde{M}(n)/n \ge 3/5$ for all $n \ge 26$ [13]. It is conjectured that $\tilde{\mu}_L = \tilde{\mu}_R = 2/3$. We give the corresponding functional version. Let $\tilde{f}$ be a Lebesgue integrable function on $\mathbb{R}$ satisfying

$$0 \le \tilde{f}(x) \le 1, \quad \int_{-\infty}^{\infty} \tilde{f}(x)\,dx = 1.$$

Define

$$\tilde{\lambda} = \inf_{\tilde{f}} \left\{ 1 - \inf_{-1 \le t \le 1} \int_{-\infty}^{\infty} \tilde{f}(x+t)\tilde{f}(x)\, dx \right\}.$$

It is known that $0.5892 \le \tilde{\lambda} \le 2/3$ [12]. As a corollary, if $\tilde{X}$ is a measurable subset of $\mathbb{R}$ with Lebesgue measure $|\tilde{X}| = 1$, then

$$0.5892 \le \inf_{\tilde{X}} \sup_{-1 \le t \le 1} \left| \tilde{X}_t - \tilde{X} \right| \le \frac{2}{3}.$$

The discrete and continuous analogs do not appear to be as closely linked as before. Again, we wonder whether current techniques [6] can be invoked to sharpen these bounds.

[1] P. Erdős, Some remarks on number theory (in Hebrew), *Riveon Lematematika* 9 (1955) 45–48; MR0073619 (17,460d).

[2] P. Erdős, Problems and results in additive number theory, *Colloque sur la Théorie des Nombres*, Proc. 1955 Bruxelles conf., Liège, G. Thone, 1956, pp. 127–137; MR0079027 (18,18a).

[3] R. K. Guy, The minimum overlap problem, *Unsolved Problems in Number Theory*, 2nd ed., Springer-Verlag, 1994, sect. C17; MR1299330 (96e:11002).

[4] S. Świerczkowski, On the intersection of a linear set with the translation of its complement, *Colloq. Math.* 5 (1958) 185–197; MR0103172 (21 #1955).

[5] L. Moser, On the minimal overlap problem of Erdős, *Acta Arith.* 5 (1959) 117–119; MR0106864 (21 #5594).

[6] J. K. Haugland, Advances in the minimum overlap problem, *J. Number Theory* 58 (1996) 71–78; MR1387725 (97c:11031).

[7] T. S. Motzkin, K. E. Ralston and J. L. Selfridge, Minimal overlapping under translation, *Bull. Amer. Math. Soc.* 62 (1956) 558.

[8] J. K. Haugland, The minimum overlap problem revisited, arXiv:1609.08000.

[9] W. O. J. Moser, A generalization of some results in additive number theory, *Math. Z.* 83 (1964) 304–313; MR0167475 (29 #4748).

[10] L. Moser and M. G. Murdeshwar, On the overlap of a function with the translation of its complement, *Colloq. Math.* 15 (1966) 93–97; MR0196023 (33 #4218).

[11] M. G. Murdeshwar, On the maximum correlation between functions, *Prace Mat.* 11 (1968) 193–197; MR0225955 (37 #1545).

[12] L. Moser and M. G. Murdeshwar, On the overlap of a function with its translates, *Nieuw Arch. Wisk.* 14 (1966) 15–18; MR0193201 (33 #1422).

[13] M. Katz and F. Schnitzer, On a problem of J. Czipszer, *Rend. Sem. Mat. Univ. Padova* 44 (1970) 85–90; MR0299492 (45 #8540).

1.9 Planar Graph Growth Constants

A **graph** of order n consists of a set of n **vertices** (points) together with a set of **edges** (unordered pairs of distinct points). Note that loops and multiple parallel edges are automatically disallowed. Two vertices joined by an edge are called

Figure 1.13 There exist 4 non-isomorphic graphs of order 3, that is, $g_3 = 4$.

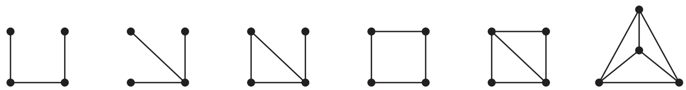

Figure 1.14 There exist 6 non-isomorphic connected graphs of order 4, that is, $c_4 = 6$.

adjacent. Two graphs X and Y are **isomorphic** if there is a one-to-one map from the vertices of X to the vertices of Y that preserves adjacency (see Figure 1.13). Diagrams for all non-isomorphic graphs of order ≤ 7 appear in [1].

A graph is **connected** if, for any two distinct vertices u and w, there is a sequence of adjacent vertices $v_0, v_1, \ldots, v_m$ such that $v_0 = u$ and $v_m = w$ (see Figure 1.14). The generating function for graphs [2]

$$\begin{aligned} g(x) &= \sum_{n=1}^{\infty} g_n x^n = x + 2x^2 + 4x^3 + 11x^4 + 34x^5 + 156x^6 + 1044x^7 \\ &\quad + 12346x^8 + 274668x^9 + \cdots, \end{aligned}$$

and the generating function for connected graphs

$$\begin{aligned} c(x) &= \sum_{n=1}^{\infty} c_n x^n = x + x^2 + 2x^3 + 6x^4 + 21x^5 + 112x^6 + 853x^7 \\ &\quad + 11117x^8 + 261080x^9 + \cdots \end{aligned}$$

are related via the Euler transform [3]

$$1 + g(x) = \exp\left(\sum_{k=1}^{\infty} \frac{c(x^k)}{k}\right).$$

If we agree that $g_0 = 1$, then the coefficients satisfy

$$g_n = \frac{1}{n} \sum_{k=1}^{n} \left(\sum_{d|k} d\, c_d\right) g_{n-k}, \quad n \geq 1.$$

Asymptotically, $g_n \sim 2^{n(n-1)/2}/n!$ as $n \to \infty$, or more precisely [4–6],

$$g_n \sim \frac{2^{n(n-1)/2}}{n!}\left(1 + 2\frac{n(n-1)}{2^n} + \frac{8}{3}\frac{n(n-1)(n-2)(3n-7)}{2^{2n}} + O\left(\frac{n^5}{2^{5n/2}}\right)\right).$$

A **separating set** or **vertex cut** of a graph X is a subset of the vertices of X, the removal of which disconnects X. Let k be a nonnegative integer. A graph

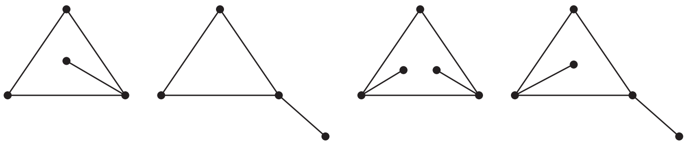

Figure 1.15 The left-hand pair of 1-connected graphs are isomorphic yet are distinct planar embeddings. The right-hand pair of 1-connected graphs are isomorphic yet are distinct spherical embeddings.

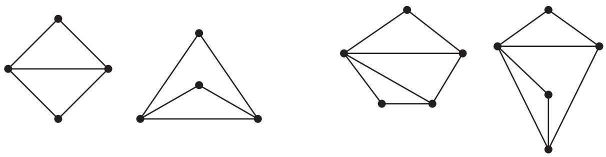

Figure 1.16 The left-hand pair of 2-connected graphs are isomorphic yet are distinct planar embeddings. The right-hand pair of 2-connected graphs are isomorphic yet are distinct spherical embeddings.

is k-**connected** if every vertex cut has at least k vertices. Clearly any graph is 0-connected and 1-connectedness is equivalent to connectedness. A 2-connected graph is often called **biconnected** or **nonseparable** and a 3-connected graph is often called **triconnected**. Observe that, when we count graphs, we do so abstractly; we are not counting embeddings in the plane or on the sphere (Figures 1.15 and 1.16).

If we **label** the vertices of a graph distinctly with the integers $1, 2, ..., n$, the corresponding enumeration problems often simplify; for example, there are exactly $2^{n(n-1)/2}$ labeled graphs. The generating function for labeled graphs

$$G(x) = \sum_{n=1}^{\infty} \frac{G_n}{n!} x^n = \sum_{n=1}^{\infty} \frac{2^{n(n-1)/2}}{n!} x^n$$

and the generating function for connected labeled graphs [7]

$$\begin{aligned} C(x) &= \sum_{n=1}^{\infty} \frac{C_n}{n!} x^n \\ &= x + \frac{1}{2!}x^2 + \frac{4}{3!}x^3 + \frac{38}{4!}x^4 + \frac{728}{5!}x^5 + \frac{26704}{6!}x^6 + \frac{1866256}{7!}x^7 + \cdots \end{aligned}$$

satisfy [3, 8]

$$1 + G(x) = \exp\left(C(x)\right), \quad G_n = \frac{1}{n} \sum_{k=1}^{n} \binom{n}{k} k G_{n-k} C_k,$$

where again we agree that $G_0 = 1$. In fact, $C_n \sim G_n$ as $n \to \infty$; consequently, almost all graphs are connected [6]. Likewise, almost all graphs are 2-connected.

A graph is **planar** if it can be embedded in the plane (as opposed to a **map**, which is a graph together with its embedding). In other words, a planar graph can be drawn so that no two edges meet except at a vertex at which both are incident. The first example of a nonplanar graph is the complete graph K_5 with 5 vertices and all 10 edges; a second well-known example is the complete bipartite graph $K_{3,3}$ with 6 vertices (3 houses and 3 utilities) and 9 edges (each house is adjacent to each utility). The generating function for planar graphs [9]

$$\begin{aligned}\bar{g}(x) &= \sum_{n=1}^{\infty} \bar{g}_n x^n \\ &= x+2x^2+4x^3+11x^4+33x^5+142x^6+822x^7+6966x^8+79853x^9+\cdots,\end{aligned}$$

the generating function for connected planar graphs

$$\begin{aligned}\bar{c}(x) &= \sum_{n=1}^{\infty} \bar{c}_n x^n \\ &= x+x^2+2x^3+6x^4+20x^5+99x^6+646x^7+5974x^8+71885x^9+\cdots,\end{aligned}$$

the generating function for 2-connected planar graphs (see Figure 1.17)

$$\begin{aligned}\bar{b}(x) &= \sum_{n=1}^{\infty} \bar{b}_n x^n \\ &= x^3+3x^4+9x^5+44x^6+294x^7+2893x^8+36496x^9+545808x^{10}+\cdots,\end{aligned}$$

and the generating function for 3-connected planar graphs (also called **polyhedra**)

$$\begin{aligned}\bar{a}(x) &= \sum_{n=1}^{\infty} \bar{a}_n x^n \\ &= x^4+2x^5+7x^6+34x^7+257x^8+2606x^9+32300x^{10}+\cdots,\end{aligned}$$

do not appear to be easily related. The growth rate of $\{\bar{g}_n\}_{n=1}^{\infty}$, defined as $\gamma_u = \lim_{n\to\infty} \bar{g}_n^{1/n}$, can be proved to exist and satisfies $\gamma_u \le 30.0606 = 2^{4.9098}$ [10–13]. We will discuss lower bounds on this constant shortly. Also, the asymptotics of $\{\bar{a}_n\}_{n=1}^{\infty}$ are precisely known [14–16]:

$$\bar{a}_n \sim \kappa\, n^{-7/2} \alpha^n$$

where

$$\alpha = \frac{16}{27}\left(17+7\sqrt{7}\right) = 21.0490424755... = (0.0475080992...)^{-1}$$

and κ is a constant (omitted).

The generating function for labeled planar graphs [17]

$$\begin{aligned}\bar{G}(x) &= \sum_{n=1}^{\infty} \frac{\bar{G}_n}{n!} x^n \\ &= x + \frac{2}{2!}x^2 + \frac{8}{3!}x^3 + \frac{64}{4!}x^4 + \frac{1023}{5!}x^5 + \frac{32071}{6!}x^6 + \frac{1823707}{7!}x^7 + \cdots ,\end{aligned}$$

the generating function for labeled connected planar graphs

$$\begin{aligned}\bar{C}(x) &= \sum_{n=1}^{\infty} \frac{\bar{C}_n}{n!} x^n \\ &= x + \frac{1}{2!}x^2 + \frac{4}{3!}x^3 + \frac{38}{4!}x^4 + \frac{727}{5!}x^5 + \frac{26013}{6!}x^6 + \frac{1597690}{7!}x^7 + \cdots ,\end{aligned}$$

and the generating function for labeled 2-connected planar graphs

$$\begin{aligned}\bar{B}(x) &= \sum_{n=1}^{\infty} \frac{\bar{B}_n}{n!} x^n \\ &= \frac{1}{3!}x^3 + \frac{10}{4!}x^4 + \frac{237}{5!}x^5 + \frac{10707}{6!}x^6 + \frac{774924}{7!}x^7 + \frac{78702536}{8!}x^8 + \cdots ,\end{aligned}$$

satisfy

$$1 + \bar{G}(x) = \exp\left(\bar{C}(x)\right), \quad \bar{C}'(x) = \exp\left(x + \bar{B}'(x\bar{C}'(x))\right)$$

where $\bar{C}'(x)$ and $\bar{B}'(x)$ denote the derivatives of $\bar{C}(x)$ and $\bar{B}(x)$. The growth rate of $\{\bar{G}_n\}_{n=1}^{\infty}$, defined as $\gamma_l = \lim_{n\to\infty}(\bar{G}_n/n!)^{1/n}$, can be proved to exist and satisfies $27.22685 < \gamma_l < 27.22688$ [18–20]. It is known that $\gamma_l < \gamma_u$, hence the lower bound for γ_l serves as a lower bound for γ_u. Further, the asymptotics of $\{\bar{B}_n\}_{n=1}^{\infty}$ are exactly known [21]:

$$\bar{B}_n \sim \lambda\, n^{-7/2} \beta^n n!$$

where

$$\beta = \frac{16\tau^3}{(1+3\tau)(1-\tau)^3} = 26.1841125556... = (0.0381910976...)^{-1},$$

τ is the unique solution of

$$\frac{1+2t}{(1+3t)(1-t)} \exp\left[-\frac{t^2(1-t)(18+36t+5t^2)}{2(3+t)(1+2t)(1+3t)^2}\right] - 2 = 0$$

and λ is a constant (again omitted). The growth constant for $\{\bar{C}_n\}_{n=1}^{\infty}$ is clearly the same as that for $\{\bar{G}_n\}_{n=1}^{\infty}$. In [20], it was asked: which of the following formulas:

$$\bar{C}_n \sim \mu\, n^{-5/2} \gamma_l^n n! \quad \text{or} \quad \bar{C}_n \sim \mu\, n^{-7/2} \gamma_l^n n!$$

is true? This appeared to be a difficult question. The answer is that $-7/2$ is the correct exponent (for $\bar{G}_n$ as well as $\bar{C}_n$): see the Addendum.

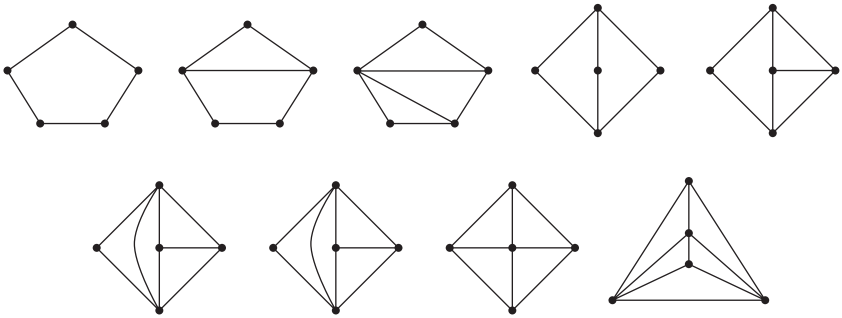

Figure 1.17 There exist 9 non-isomorphic 2-connected planar graphs of order 5.

A graph is **outerplanar** if it can be embedded in the plane so that all its vertices lie on the same face. This face, by convention, is usually chosen to be the unbounded exterior of the graph. Any tree is an outerplanar graph. Non-outerplanar graphs include the complete graph K_4 with 4 vertices and all 6 edges, and the complete bipartite graph $K_{2,3}$ with 5 vertices and 6 edges. The unlabeled case has not received much attention, except in the 2-connected case (the first three graphs in Figure 1.17, each pentagonal, constitute all possibilities with 5 vertices). The generating function for unlabeled 2-connected outerplanar graphs [22]

$$\begin{aligned}\hat{b}(x) &= \sum_{n=1}^{\infty} \hat{b}_n x^n \\ &= x^3 + 2x^4 + 3x^5 + 9x^6 + 20x^7 + 75x^8 + 262x^9 + 1117x^{10} + \cdots\end{aligned}$$

was obtained by Read [23–25], building on Motzkin [26]:

$$\begin{aligned}\hat{b}(x) = &\frac{\left(3x^2 - 2xf(x) + f(x)^2\right) - \left(2 + 2x + 7x^2 - 4xf(x) + 2f(x)^2\right)f(x^2) + 2f(x^2)^2}{4(2f(x^2) - 1)} \\ &+ \frac{1}{2}\sum_{k=3}^{\infty} \frac{1}{k}\left(\sum_{d|k} \varphi(d)\left(f(x^d)\right)^{k/d}\right)\end{aligned}$$

where

$$f(x) = x + \sum_{r=1}^{\infty} \frac{1}{r} \sum_{s=2}^{\infty} \binom{s-2}{r-1}\binom{r+s-1}{s} x^s$$

and φ is the Euler totient function. Counting such graphs is closely related to enumerating the number of dissections of the interior of a regular n-gon into smaller polygons by use of nonintersecting diagonals. Asymptotics for $\hat{b}_n$ are presently open. It would be good to learn more about $\hat{c}_n$ and $\hat{g}_n$ too.

By contrast, the generating function for labeled outerplanar graphs [22]

$$\begin{aligned}\hat{G}(x) &= \sum_{n=1}^{\infty}\frac{\hat{G}_n}{n!}x^n \\ &= x+\frac{2}{2!}x^2+\frac{8}{3!}x^3+\frac{63}{4!}x^4+\frac{893}{5!}x^5+\frac{19714}{6!}x^6+\frac{597510}{7!}x^7+\cdots,\end{aligned}$$

the generating function for labeled connected outerplanar graphs

$$\begin{aligned}\hat{C}(x) &= \sum_{n=1}^{\infty}\frac{\hat{C}_n}{n!}x^n \\ &= x+\frac{1}{2!}x^2+\frac{4}{3!}x^3+\frac{37}{4!}x^4+\frac{602}{5!}x^5+\frac{14436}{6!}x^6+\frac{458062}{7!}x^7+\cdots,\end{aligned}$$

and the generating function for labeled 2-connected outerplanar graphs

$$\begin{aligned}\hat{B}(x) &= \sum_{n=1}^{\infty}\frac{\hat{B}_n}{n!}x^n \\ &= \frac{1}{3!}x^3+\frac{9}{4!}x^4+\frac{132}{5!}x^5+\frac{2700}{6!}x^6+\frac{70920}{7!}x^7+\frac{2275560}{8!}x^8+\cdots,\end{aligned}$$

satisfy

$$1+\hat{G}(x)=\exp\left(\hat{C}(x)\right), \quad \hat{C}'(x)=\exp\left(x\hat{C}'(x)+\hat{B}'(x\hat{C}'(x))\right)$$

and, moreover,

$$\hat{B}'(x)=\frac{1}{8}\left(1+5x-\sqrt{1-6x+x^2}\right)-x.$$

In view of the algebraic nature of $\hat{B}'(x)$, it is not surprising that the growth constants possess closed-form expressions [27–29]:

$$\begin{aligned}\lim_{n\to\infty}\left(\frac{\hat{G}_n}{n!}\right)^{1/n} &= \frac{1}{\xi}\exp\left(\frac{1+5\xi-\sqrt{1-6\xi+\xi^2}}{8}\right) \\ &= 7.3209800548...=(0.1365937336...)^{-1}\end{aligned}$$

where $\xi=0.1707649868...$ has minimal polynomial $8-58x+70x^2-28x^3+3x^4$, and

$$\lim_{n\to\infty}\left(\frac{\hat{B}_n}{n!}\right)^{1/n}=3+2\sqrt{2}=5.8284271247....$$

Like before, the growth constant for $\{\hat{C}_n\}_{n=1}^{\infty}$ is the same as that for $\{\hat{G}_n\}_{n=1}^{\infty}$.

Addendum Giménez & Noy [30, 31] demonstrated that the growth constant for labeled planar graphs is $\gamma_l = 27.2268777685...$ and, further,

$$\bar{G}_n \sim \left(0.42609... \times 10^{-5}\right) n^{-7/2} (27.22687...)^n,$$

$$\bar{C}_n \sim \left(0.41043... \times 10^{-5}\right) n^{-7/2} (27.22687...)^n,$$

$$\bar{B}_n \sim \left(0.37044... \times 10^{-5}\right) n^{-7/2} (26.18411...)^n.$$

Bodirsky, Fusy, Kang & Vigerske [32] showed that

$$\hat{g}_n \sim (0.90994... \times 10^{-2}) n^{-5/2} (7.50360...)^n,$$

$$\hat{c}_n \sim (0.76047... \times 10^{-2}) n^{-5/2} (7.50360...)^n,$$

$$\hat{b}_n \sim (0.59602... \times 10^{-2}) n^{-5/2} \left(3 + 2\sqrt{2}\right)^n;$$

note that growth rates of unlabeled and labeled 2-connected outerplanar graphs coincide, whereas growth rates for the connected and general cases differ.

Upon introduction of randomness, we are quickly overwhelmed with various numerical results, far too many to summarize. Here are two examples [31]. Let P_n denote the number of edges in a uniform planar graph with n vertices. Then P_n is asymptotically normal with mean $(2.21326...)n$ and variance $(0.43034...)n$ as $n \to \infty$. Let Q_n denote the number of connected components in a uniform planar graph with n vertices. Then $Q_n - 1$ is asymptotically Poisson with parameter $\nu \approx 0.037439$. Consequently, the probability that a random planar graph is connected is $\exp(-\nu) \approx 0.963253$ and the expected number of components is $1 + \nu = 1.037439$ as $n \to \infty$. More such results appear in [33–42].

[1] R. C. Read and R. J. Wilson, *An Atlas of Graphs*, Oxford Univ. Press, 1998; MR1692656 (2000a:05001).

[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000088, A001349, A002218, and A006290.

[3] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences*, Academic Press, 1995, pp. 19–21; MR1327059 (96a:11001).

[4] W. Oberschelp, Kombinatorische Anzahlbestimmungen in Relationen, *Math. Annalen* 174 (1967) 53–78; MR0218255 (36 #1342).

[5] E. M. Wright, The number of graphs on many unlabelled nodes, *Math. Annalen* 183 (1969) 250–253; MR0250928 (40 #4160).

[6] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, 1973, pp. 195–208; MR0357214 (50 #9682).

[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001187 and A095983.

[8] H. S. Wilf, *generatingfunctionology*, 2nd ed., Academic Press, 1994, pp. 86–87; MR1277813 (95a:05002).

[9] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005470, A003094, A021103, and A000944.

[10] A. Denise, M. Vasconcellos and D. J. A. Welsh, The random planar graph, *Festschrift for C. St. J. A. Nash-Williams*, ed. A. J. W. Hilton, Congr. Numer. 113, Utilitas Math., 1996, pp. 61–79; MR1393702 (97e:05171).

[11] N. Bonichon, C. Gavoille and N. Hanusse, An information-theoretic upper bound of planar graphs using triangulation, *Proc. 2003 Sympos. on Theoretical Aspects of Computer Science (STACS)*, Berlin, ed. H. Alt and M. Habib, Lect. Notes in Comp. Sci. 2607, Springer-Verlag, 2003, pp. 499–510; MR2066778.

[12] N. Bonichon, C. Gavoille, N. Hanusse, D. Poulalhon and G. Schaeffer, Planar graphs, via well-orderly maps and trees, *Proc. 30th International Workshop on Graph-Theoretic Concepts in Computer Science (WG)*, Bad Honnef, ed. J. Hromkovic, M. Nagl and B. Westfechtel, Lect. Notes in Comp. Sci. 3353, Springer-Verlag, 2004, pp. 270–284; MR2158652 (2006d:05059).

[13] N. Bonichon, C. Gavoille, N. Hanusse, D. Poulalhon and G. Schaeffer, Planar graphs, via well-orderly maps and trees, *Graphs Combin.* 22 (2006) 185–202; MR2231990 (2007g:05085).

[14] E. A. Bender and L. B. Richmond, The asymptotic enumeration of rooted convex polyhedra, *J. Combin. Theory Ser. B* 36 (1984) 276–283; MR0753605 (85m:05004).

[15] E. A. Bender and N. C. Wormald, Almost all convex polyhedra are asymmetric, *Canad. J. Math.* 37 (1985) 854–871; MR0806644 (87b:52014).

[16] H. T. Croft, K. J. Falconer and R. K. Guy, Counting polyhedra, *Unsolved Problems in Geometry*, Springer-Verlag, 1991, sect. B15; MR1107516 (92c:52001).

[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A086789, A096332, A096331, and A096330.

[18] C. McDiarmid, A. Steger and D. J. A. Welsh, Random planar graphs, *J. Combin. Theory Ser. B* 93 (2005) 187–205; MR2117936 (2005k:05221).

[19] D. Osthus, H. J. Prömel and A. Taraz, On random planar graphs, the number of planar graphs and their triangulations, *J. Combin. Theory Ser. B* 88 (2003) 119–134; MR1973264 (2004a:05068).

[20] O. Giménez and M. Noy, Estimating the growth constant of labelled planar graphs, *Proc. 3rd Colloq. on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities*, Vienna, 2004, ed. M. Drmota, P. Flajolet, D. Gardy and B. Gittenberger, Birkhäuser, 2004, pp. 133–139; MR2090501 (2005f:05077).

[21] E. A. Bender, Z. Gao and N. C. Wormald, The number of labeled 2-connected planar graphs, *Elec. J. Combin.* 9 (2002) R43; MR1946145 (2003i:05071).

[22] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001004, A097998, A097999, A098000.

[23] R. C. Read, On general dissections of a polygon, *Aequationes Math*. 18 (1978) 370–388; MR0522523 (80e:05069).

[24] P. Lisonek, Closed forms for the number of polygon dissections, *J. Symbolic Comput.* 20 (1995) 595–601; MR1395415 (97j:52019).

[25] G. Hu, Group theory method for enumeration of outerplanar graphs, *Acta Math. Appl. Sinica* 14 (1998) 381–387; MR1651268 (99h:05061).

[26] T. Motzkin, Relations between hypersurface cross ratios, and a combinatorial formula for partitions of a polygon, for permanent preponderance, and for non-associative products, *Bull. Amer. Math. Soc*. 54 (1948) 352–360; MR0024411 (9,489d).

[27] M. Bodirsky, O. Giménez, M. Kang and M. Noy, On the number of series parallel and outerplanar graphs, *Proc. European Conf. on Combinatorics, Graph Theory, and Applications (EuroComb05)*, DMTCS Proc. AE (2005) 383–388.

[28] M. Bodirsky, O. Giménez, M. Kang and M. Noy, Enumeration and limit laws for series-parallel graphs, *European J. Combin.* 28 (2007) 2091–2105; math.CO/0512435; MR2351512 (2008i:05087).

[29] P. Flajolet and M. Noy, Analytic combinatorics of non-crossing configurations, *Discrete Math.*, 204 (1999) 203–229; MR1691870 (2000c:05012).
[30] O. Giménez and M. Noy, The number of planar graphs and properties of random planar graphs, *Internat. Conf. on Analysis of Algorithms (AofA'05)*, Barcelona, ed. C. Martínez, DMTCS Proc. AD (2005) 147–156; MR2193114.
[31] O. Giménez and M. Noy, Asymptotic enumeration and limit laws of planar graphs, *J. Amer. Math. Soc.* 22 (2009) 309–329; arXiv:math/0501269; MR2476775 (2010g:05031).
[32] M. Bodirsky, E. Fusy, M. Kang and S. Vigerske, Enumeration and asymptotic properties of unlabeled outerplanar graphs, *Elec. J. Combin.* 14 (2007) R66; arXiv:math/0511422; MR2350456 (2008j:05171).
[33] M. Drmota, O. Giménez and M. Noy, Degree distribution in random planar graphs, *Fifth Colloquium on Mathematics and Computer Science*, DMTCS Proc. AI, (2008) 163–1781; MR2508785 (2010g:05338).
[34] M. Drmota, O. Giménez and M. Noy, Degree distribution in random planar graphs, *J. Combin. Theory Ser. A* 118 (2011) 2102–2130; arXiv:0911.4331; MR2802191 (2012i:05258).
[35] M. Drmota, O. Giménez and M. Noy, Vertices of given degree in series-parallel graphs, *Random Structures Algorithms* 36 (2010) 273–314; MR2650053 (2011h:05226).
[36] M. Drmota, O. Giménez and M. Noy, The maximum degree of series-parallel graphs, *Combin. Probab. Comput.* 20 (2011) 529–570; arXiv:1008.5361; MR2805397 (2012e:05342).
[37] M. Drmota, O. Giménez, M. Noy, K. Panagiotou and A. Steger, The maximum degree of random planar graphs, *Proc. Lond. Math. Soc.* 109 (2014) 892–920; MR3273487.
[38] N. Fountoulakis and K. Panagiotou, 3-connected cores in random planar graphs, *Combin. Probab. Comput.* 20 (2011) 381–412; arXiv:0907.2326; MR2784634 (2012d:05351).
[39] M. Bodirsky, M. Kang, M. Löffler and C. McDiarmid, Random cubic planar graphs, *Random Structures Algorithms* 30 (2007) 78–94; MR2283223 (2007k:05193).
[40] S. Gerke, O. Giménez, M. Noy and A. Weißl, The number of graphs not containing $K_{3,3}$ as a minor, *Elec. J. Combin.* 15 (2008) R114; arXiv:0803.4418; MR2438586 (2009d:05109).
[41] O. Bernardi, M. Noy and D. Welsh, Growth constants of minor-closed classes of graphs, *J. Combin. Theory Ser. B* 100 (2010) 468–484; arXiv:0710.2995; MR2644234 (2011h:05245).
[42] O. Giménez, M. Noy and J. Rué, Graph classes with given 3-connected components: asymptotic enumeration and random graphs, *Random Structures Algorithms* 42 (2013) 438–479; arXiv:0907.0376; MR3068033.

1.10 Tauberian Constants

A series $\sum_{k=0}^{\infty} a_k$ of complex numbers is **Abel convergent** if

$$\lim_{x\to 1^-} \sum_{k=0}^{\infty} a_k x^k \quad \left(\text{equivalently, } \lim_{x\to 1^-} (1-x) \sum_{n=0}^{\infty} s_n x^n\right)$$

exists and **Cesàro convergent** if

$$\lim_{l\to\infty}\frac{1}{l+1}\sum_{n=0}^{l}s_n \quad \left(\text{equivalently, } \lim_{l\to\infty}\sum_{k=0}^{l}\left(1-\frac{k}{l+1}\right)a_k\right)$$

exists, where $s_n=\sum_{k=0}^{n}a_k$ for each $n\geq 0$. Ordinary convergence implies both Abel convergence and Cesàro convergence. Various converses of this theorem, in which ordinary convergence is deduced from a summability condition (as above) plus an additional condition (for example, $ka_k\to 0$ as $k\to\infty$), are called **Tauberian theorems** [1–3].

For notational convenience, when we use the symbol σ, we mean an arbitrary limit point of the partial sums $\{s_n\}_{n=0}^{\infty}$. By λ, we mean a limit point of the power series $\sum_{k=0}^{\infty}a_kx^k$ as $x\to 1^-$. By μ, we mean a limit point of the partial averages $\{m_l\}_{l=0}^{\infty}$, where $m_l=\sum_{n=0}^{l}s_n/(l+1)$ for each $l\geq 0$.

We start with a Tauberian theorem due to Hadwiger [4, 5] and Agnew [6–8]; it is quite general since no hypotheses are required! Constants C_1 and C_2 exist with the following properties:

- for each σ, there is a λ such that $|\lambda-\sigma|\leq C_1\,\limsup_{k\to\infty}|ka_k|$
- for each λ, there is a σ such that $|\lambda-\sigma|\leq C_2\,\limsup_{k\to\infty}|ka_k|$.

The least constant C_1 is known to be

$$\gamma+\ln(\ln(2))-2\,\mathrm{Ei}(-\ln(2))=0.9680448304...$$

where Ei is the exponential integral [9]. The least constant C_2 satisfies the inequality $0.4858\leq C_2\leq 0.7494386$, but its exact value is unknown. Likewise [8, 10, 11], constants C_3 and C_4 exist with the following properties:

- for each σ, there is a μ such that $|\mu-\sigma|\leq C_3\,\limsup_{k\to\infty}|ka_k|$
- for each μ, there is a σ such that $|\mu-\sigma|\leq C_4\,\limsup_{k\to\infty}|ka_k|$.

The least constant C_3 is known to be $\ln(2)=0.6931471805...$ and the least constant C_4 is the unique real solution y of the equation

$$y=e^{-(\pi/2)y},\quad \text{that is, } y=\frac{2}{\pi}W\left(\frac{\pi}{2}\right)=0.4745409995...,$$

where W is Lambert's function [12]. See a generalization by Rajagopal [13, 14].

Different constants emerge if we are more restrictive in our choices of σ, λ and μ. For example, the best constant $\tilde{C}_1$ such that [15–18]

$$\limsup_{n\to\infty}\left|\sum_{k=0}^{\infty}a_k\left(1-\frac{1}{n}\right)^k-\sum_{k=0}^{n}a_k\right|\leq \tilde{C}_1\limsup_{k\to\infty}|ka_k|$$

is

$$\gamma-2\mathrm{Ei}(-1)=1.0159835336...=1.7517424160...-2/e.$$

The constant $-\mathrm{Ei}(-1)=0.2193839343...$ is familiar: when multiplied by e, it gives the Euler–Gompertz constant [9]. In the definition of C_4, observe that the subsequence of $\{s_n\}_{n=0}^{\infty}$ with limit point σ may depend on the sequence $\{a_k\}_{k=0}^{\infty}$. If we deny any knowledge of $\{a_k\}_{k=0}^{\infty}$, then the required constant C_4' becomes larger. More precisely, there is a non-decreasing sequence $\{n_l\}_{l=0}^{\infty}$ independent of $\{a_k\}_{k=0}^{\infty}$ such that [8, 10, 11]

$$\limsup_{l\to\infty}|m_l-s_{n_l}|\le C_4'\limsup_{k\to\infty}|ka_k|$$

and $C_4'=\ln(2)$ is best possible; further, a simple such sequence is $n_l=\lfloor l/2\rfloor$. Here is a variation in which we permit knowledge of $\{a_k\}_{k=0}^{\infty}$ only to make a binary decision at each step. There exist two non-decreasing sequences $\{p_l\}_{l=0}^{\infty}$ and $\{q_l\}_{l=0}^{\infty}$ independent of $\{a_k\}_{k=0}^{\infty}$ such that

$$\limsup_{l\to\infty}|m_l-s_{n_l}|\le C_4''\limsup_{k\to\infty}|ka_k|$$

where n_l is, for each l, one of the two integers p_l and q_l, and the optimal C_4'' satisfies $C_4\le C_4''\le 0.56348$. The exact value of C_4'' is unknown, but it is believed to be close to its upper bound. This estimate comes from setting $p_l=\lfloor 3l/8\rfloor$, $q_l=\lfloor 5l/8\rfloor$ and choosing n_l appropriately.

Comparisons between different weighted averages of a sequence $\{a_k\}_{k=1}^{\infty}$ are important in prime number theory. Let $\mu(k)$ be the Möbius mu function and define

$$M(n)=\sum_{k\le n}\mu(k),\quad g(n)=\sum_{k\le n}\frac{\mu(k)}{k},\quad Q(n)=\sum_{k\le n}|\mu(k)|\,.$$

From [19, 20], we have

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n}a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2-\frac{6}{\pi^2}\sum_{k=1}^{n}a_k\frac{k}{n}\left\lfloor\frac{n}{k}\right\rfloor\right|\le C_5\limsup_{k\to\infty}|ka_k|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n}a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2-\frac{6}{\pi^2}\sum_{k=1}^{n}a_k\right|\le C_6\limsup_{k\to\infty}|ka_k|$$

where the best constants are [21, 22]

$$C_5=\int_1^\infty\left|\frac{Q(u)}{u^2}-\frac{6}{\pi^2}\frac{\lfloor u\rfloor}{u^2}\right|du=0.6945017...,$$

$$C_6=\int_1^\infty\left|\frac{Q(u)}{u^2}-\frac{6}{\pi^2 u}\right|du=0.4616041....$$

In the following, a non-classical expression appears on the right-hand side of the inequality, which allows for comparisons with 0 in the first three cases [23, 24]:

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{1}{n}\right)^{k}-0\right| \le C_7 \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{k-1}{n}\right)-0\right| \le C_8 \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2-0\right| \le C_9 \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2-\sum_{k=1}^{n} a_k\frac{k}{n}\left\lfloor\frac{n}{k}\right\rfloor\right| \le C_{10} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{1}{n}\right)^{k}-\sum_{k=1}^{n} a_k\frac{k}{n}\left\lfloor\frac{n}{k}\right\rfloor\right| \le C_{11} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{1}{n}\right)^{k}-\sum_{k=1}^{n} a_k\left(1-\frac{k-1}{n}\right)\right| \le C_{12} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\frac{k}{n}\left\lfloor\frac{n}{k}\right\rfloor-\sum_{k=1}^{n} a_k\left(1-\frac{k-1}{n}\right)\right| \le C_{13} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{1}{n}\right)^{k}-\sum_{k=1}^{n} a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2\right| \le C_{14} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|,$$

$$\limsup_{n\to\infty}\left|\sum_{k=1}^{n} a_k\left(1-\frac{k-1}{n}\right)-\sum_{k=1}^{n} a_k\frac{k}{n}\sum_{j\le n/k}\mu(j)^2\right| \le C_{15} \limsup_{n\to\infty}\left|\sum_{d|n} d a_d\right|$$

where the best constants are [21, 22]

$$C_7=\int_1^\infty \frac{|g(u)|}{u}du=1.09667..., \quad C_8=\int_1^\infty \frac{|u\,g(u)-M(u)|}{u^2}du=1.00004...,$$

$$C_9=2\int_1^\infty \frac{|M(u)|}{u^3}du=0.8921506905...,$$

$$C_{10}=2\int_1^\infty \frac{|M(u)-1|}{u^3}du=0.3921032696...,$$

$$C_{11}=\int\limits_1^\infty \frac{|u\,g(u)-1|}{u^2}du=0.483439..., \quad C_{12}=\int\limits_1^\infty \frac{|M(u)|}{u^2}du=1.01426...,$$

$$C_{13}=\int\limits_1^\infty \frac{|u\,g(u)-M(u)-1|}{u^2}du=0.613...,$$

$$C_{14}=\int\limits_1^\infty \frac{\left|M\left(\sqrt{u}\right)-u\,g(u)\right|}{u^2}du=0.49619...,$$

$$C_{15}=\int\limits_1^\infty \frac{\left|M\left(\sqrt{u}\right)-u\,g(u)+M(u)\right|}{u^2}du=1.00582....$$

The values of C_{14} and C_{15} here correct the values 0.486 and 0.994 given in [24].

Acknowledgment Tadej Kotnik kindly computed the constants C_5 to C_{15} at my request [21], publishing his detailed results later [22].

[1] R. L. Powell and S. M. Shah, *Summability Theory and Applications*, van Nostrand Reinhold, 1972.
[2] A. Peyerimhoff, *Lectures on Summability*, Lect. Notes in Math. 107, Springer-Verlag, 1969; MR0463744 (57 #3684).
[3] J. van de Lune, *An Introduction to Tauberian Theory: From Tauber to Wiener*, Centrum voor Wiskunde en Informatica, 1986; MR0882005 (88j:40011).
[4] H. Hadwiger, Über ein Distanz-theorem bei der A-Limitierung, *Comment. Math. Helv.* 16 (1944) 209–214; MR0010002 (5,236e).
[5] H. Hadwiger, Ueber eine Konstante Tauberscher Art, *Revista Mat. Hisp.-Amer.* 7 (1947) 65–69; MR0021602 (9,86b).
[6] R. P. Agnew, Abel transforms of Tauberian series, *Duke Math. J.* 12 (1945) 27–36; MR0012338 (7,12c).
[7] R. P. Agnew, Abel transforms and partial sums of Tauberian series, *Annals of Math.* 50 (1949) 110–117; MR0027351 (10,291i).
[8] R. P. Agnew, Tauberian relations among partial sums, Riesz transforms, and Abel transforms of series, *J. Reine Angew. Math.* 193 (1954) 94–118; MR0064159 (16,237a).
[9] S. R. Finch, Euler-Gompertz constant, *first volume*, pp. 423–428.
[10] R. P. Agnew, Arithmetic means of some Tauberian series and determination of a lower bound for a fundamental Tauberian constant, *Proc. London Math. Soc.* 2 (1952) 369–384; MR0049357 (14,160c).
[11] R. P. Agnew, Arithmetic means and the Tauberian constant .474541, *Acta Math.* 87 (1952) 347–359; MR0051331 (14,463e).
[12] S. R. Finch, Iterated exponential constants, *first volume*, pp. 448–452.
[13] C. T. Rajagopal, On an absolute constant in the theory of Tauberian series, *Proc. Indian Acad. Sci. Sect. A* 28 (1948) 537–544; postscript, 31 (1950) 60–61; MR0028443 (10,447g) and MR0035846 (12,21b).
[14] C. T. Rajagopal, On an absolute constant in the theory of Tauberian series. II, *Proc. Indian Acad. Sci. Sect. A* 39 (1954) 272–281; MR0063462 (16,125a).

[15] H. Hadwiger, Die Retardierungserscheinung bei Potenzreihen und Ermittlung zweier Konstanten Tauberscher Art, *Comment. Math. Helv.* 20 (1947) 319–332; MR0021605 (9,86e).
[16] A. Wintner, On Tauber's theorem, *Comment. Math. Helv.* 20 (1947) 216–222; MR0021604 (9,86d).
[17] P. Hartman, Tauber's theorem and absolute constants, *Amer. J. Math.* 69 (1947) 599–606; MR0021603 (9,86c).
[18] I. J. Maddox, Tauberian constants, *Bull. London Math. Soc.* 1 (1969) 193–200; MR0247328 (40 #596).
[19] K. A. Jukes, Equiconvergence of matrix transformations, *Proc. Amer. Math. Soc.* 69 (1978) 261–270; MR0477552 (57 #17071).
[20] H. W. Gretton and K. A. Jukes, Computations of Tauberian constants connected with equiconvergence and the quadratfrei numbers, *J. Comput. Appl. Math.* 5 (1979) 25–27; MR0525608 (80c:65079).
[21] T. Kotnik, Computation of eleven Tauberian constants, unpublished note (2005).
[22] T. Kotnik, Evaluation of some integrals of sums involving the Möbius function, *Internat. J. Comput. Math.* 84 (2007) 469–475; MR2326305.
[23] K. A. Jukes, Tauberian constants connected with the prime number theorem, *Math. Z.* 145 (1975) 187–193; MR0382896 (52 #3778).
[24] H. W. Gretton and K. A. Jukes, Tauberian constants connected with the prime number theorem. II. Computational aspects, *Math. Z.* 164 (1978) 25–30; MR0514605 (80a:10064).

1.11 Integer Partitions

Let L denote the positive octant of the regular d-dimensional cubic lattice. Each vertex $(j_1, j_2, \ldots, j_d)$ of L is adjacent to all vertices of the form $(j_1, j_2, \ldots, j_k + 1, \ldots, j_d)$, $1 \le k \le d$. A **d-partition** of a positive integer n is an assignment of nonnegative integers $n_{j_1, j_2, \ldots, j_d}$ to the vertices of L, subject to both an ordering condition

$$n_{j_1,j_2,\ldots,j_d} \ge \max_{1\le k\le d} n_{j_1,j_2,\ldots,j_k+1,\ldots,j_d}$$

and a summation condition $\sum n_{j_1,j_2,\ldots,j_d} = n$. The summands in the d-partition are thus nonincreasing in each of the d lattice directions. We agree to suppress all zero labels. A 1-partition is the same as an ordinary partition; a 2-partition is often called a **plane partition** and a 3-partition is often called a **solid partition**. Three sample plane partitions of $n = 26$ are

$$\begin{pmatrix} 8 \\ 9 \\ 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & & & \\ 1 & & & \\ 2 & 2 & 1 & \\ 4 & 2 & 1 & 1 \\ 5 & 3 & 2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 7 & 6 & 4 & 4 & 3 & 1 & 1 \end{pmatrix}.$$

Let $p_d(n)$ denote the number of d-partitions of n. The generating functions [1]

$$1+\sum_{n=1}^{\infty} p_1(n)x^n = 1+x+2x^2+3x^3+5x^4+7x^5+11x^6+15x^7+22x^8+\cdots$$
$$= \prod_{m=1}^{\infty}(1-x^m)^{-1},$$

$$1+\sum_{n=1}^{\infty} p_2(n)x^n = 1+x+3x^2+6x^3+13x^4+24x^5+48x^6+86x^7+160x^8+\cdots$$
$$= \prod_{m=1}^{\infty}(1-x^m)^{-m}$$

give rise to well-known asymptotics [2–5]:

$$p_1(n) \sim \frac{1}{4\sqrt{3}n}\exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$
$$\sim (0.1443375672...)n^{-1}\exp\left((2.5650996603...)n^{1/2}\right),$$

$$p_2(n) \sim \frac{\zeta(3)^{7/36}e^{\zeta'(-1)}}{2^{11/36}\sqrt{3\pi}n^{25/36}}\exp\left(3\zeta(3)^{1/3}\left(\frac{n}{2}\right)^{2/3}\right)$$
$$\sim (0.2315168134...)n^{-25/36}\exp\left((2.0094456608...)n^{2/3}\right)$$

as $n\to\infty$, where $\zeta(3)=1.2020569031...$ is Apéry's constant [6] and $\zeta'(-1)=-0.1654211437...=2(-0.0827105718...)=\ln(0.8475366941...)$ is closely related to the Glaisher–Kinkelin constant [7]. Although an infinite product expression for the generating function [1]

$$1+\sum_{n=1}^{\infty} p_3(n)x^n = 1+x+4x^2+10x^3+26x^4+59x^5+140x^6+307x^7+684x^8+\cdots$$

remains unknown, it is conjectured that [8, 9]

$$p_3(n) \sim \frac{C}{n^{61/96}}\exp\left(\frac{2^{7/4}\pi}{3^{5/4}5^{1/4}}n^{3/4}+\frac{\sqrt{15}\zeta(3)}{\sqrt{2}\pi^2}n^{1/2}-\frac{15^{5/4}\zeta(3)^2}{2^{7/4}\pi^5}n^{1/4}\right)$$
$$\sim Cn^{-61/96}\exp((1.7898156270...)n^{3/4}+(0.3335461354...)n^{1/2}$$
$$-(0.0414392867...)n^{1/4})$$

for some constant $C>0$. The evidence for this asymptotic formula includes exact enumerations (for $n\leq 68$) and Monte Carlo simulation. See [10–13] for more about planar partitions and [14–17] for more about solid partitions.

1.11.1 *Hardy–Ramanujan–Rademacher*

The Hardy–Ramanujan–Rademacher formula for $p_1(n)$ is a spectacular exact result [18–26]:

$$p_1(n) = \frac{\pi}{2^{5/4}3^{3/4}}\left(n - \frac{1}{24}\right)^{-3/4} \sum_{k=1}^{\infty} \frac{A_k(n)}{k} I_{3/2}\left(\sqrt{\frac{2}{3}}\frac{\pi}{k}\sqrt{n - \frac{1}{24}}\right)$$

where

$$I_{3/2}(x) = \sqrt{\frac{2x}{\pi}}\left(\frac{\cosh(x)}{x} - \frac{\sinh(x)}{x^2}\right)$$

is the modified Bessel function of order $3/2$,

$$A_k(n) = \sum_{\substack{\gcd(h,k)=1,\\ 1\le h<k}} \omega_{h,k} \exp\left(\frac{-2\pi i n h}{k}\right),$$

and $\omega_{h,k} = \exp(\pi i s(h,k))$ is the unique $24k^{\text{th}}$ root of unity with Dedekind sum

$$s(h,k) = \sum_{m=1}^{k-1}\left(\frac{m}{k} - \left\lfloor\frac{m}{k}\right\rfloor - \frac{1}{2}\right)\left(\frac{hm}{k} - \left\lfloor\frac{hm}{k}\right\rfloor - \frac{1}{2}\right).$$

For example,

$$A_1(n) = 1, \quad A_2(n) = (-1)^n, \quad A_3(n) = 2\cos\left(\frac{\pi(12n-1)}{18}\right),$$

$$A_4(n) = 2\cos\left(\frac{\pi(4n-1)}{8}\right), \quad A_5(n) = 2\cos\left(\frac{\pi(2n-1)}{5}\right) + 2\cos\left(\frac{4\pi n}{5}\right).$$

Defining

$$c = \sqrt{\frac{2}{3}}\pi, \quad \lambda(n) = \sqrt{n - \frac{1}{24}},$$
$$\mu(n) = c\lambda(n), \quad A_k^*(n) = A_k(n)/\sqrt{k},$$

we have the following variations:

$$\begin{aligned} p_1(n) &= \frac{1}{2^{1/2}\pi}\sum_{k=1}^{\infty} A_k(n) k^{1/2} \frac{d}{dn}\left[\frac{\sinh(c\lambda(n)/k)}{\lambda(n)}\right] \\ &= 2\frac{3^{1/2}}{24n-1}\sum_{k=1}^{\infty} A_k^*(n)\left[\left(1 - \frac{k}{\mu(n)}\right)\exp\left(\frac{\mu(n)}{k}\right) + \left(1 + \frac{k}{\mu(n)}\right)\exp\left(-\frac{\mu(n)}{k}\right)\right]. \end{aligned}$$

By contrast, the original Hardy–Ramanujan formula is only an asymptotic expansion:

$$\begin{aligned} p_1(n) &\sim \frac{1}{2^{3/2}\pi}\sum_{k=1}^{\infty} A_k(n) k^{1/2} \frac{d}{dn}\left[\frac{\exp(c\lambda(n)/k)}{\lambda(n)}\right] \\ &\sim 2\frac{3^{1/2}}{24n-1}\sum_{k=1}^{\infty} A_k^*(n)\left(1 - \frac{k}{\mu(n)}\right)\exp\left(\frac{\mu(n)}{k}\right), \end{aligned}$$

which was later proved to be divergent by Lehmer [27–29]. Therefore Rademacher's contribution was the identification of a small additional term that forces the original Hardy–Ramanujan series to converge.

A third formula for $p_1(n)$:

$$p_1(n) \sim \frac{\pi}{2^{5/4}3^{3/4}}\lambda(n)^{-3/2}\sum_{k=1}^{\infty}\frac{A_k(n)}{k}I_{-3/2}\left(\frac{c\lambda(n)}{k}\right)$$

appears in Almkvist [30, 31] and is a consequence of a more general theory (to be discussed shortly). The only difference between this formula and the Hardy–Ramanujan–Rademacher formula is that $I_{-3/2}$ appears rather than $I_{3/2}$. It is believed to be divergent, but this has not yet been proved. For practical purposes, using the modified Bessel function of order $-3/2$:

$$I_{-3/2}(x) = \sqrt{\frac{2x}{\pi}}\left(\frac{\sinh(x)}{x} - \frac{\cosh(x)}{x^2}\right)$$

gives only slightly different numerical results (for large $\sqrt{n}/k$).

Analogous series exist for plane partitions. The terms involve neither exponentials nor Bessel functions, but rather a new function

$$g(x,\gamma) = \sum_{\nu=0}^{\infty}\frac{x^{2\nu+\gamma-1}}{\nu!\Gamma(2\nu+\gamma)}$$

that satisfies the third-order differential equation

$$xg'''(x,\gamma) - (\gamma-3)g''(x,\gamma) - 2g(x,\gamma) = 0$$

(the derivatives are taken with respect to x) as well as

$$g'(x,\gamma) = g(x,\gamma-1), \quad 2g(x,\gamma+2) + (\gamma-1)g(x,\gamma) = xg(x,\gamma-1).$$

A heuristic argument in [30, 31] gives that

$$p_2(n) \sim \varphi_1(n) + \varphi_2(n) + \varphi_3(n) + \cdots$$

as $n \to \infty$, where

$$\varphi_1(n) = \zeta(3)^{13/24}e^{\zeta'(-1)}\sum_{k=0}^{\infty}a_{2k}\zeta(3)^k g\left(n\sqrt{\zeta(3)}, -\frac{1}{12} - 2k\right)$$

and a_{2k} is the coefficient of x^{2k} in the Maclaurin series of

$$h(x) = \exp\left(-\sum_{j=1}^{\infty}\frac{2(2j+1)!\zeta(2j)\zeta(2j+2)}{j(2\pi)^{4j+2}}x^{2j}\right),$$

$$\varphi_2(n) = (-1)^n 2^{-5/3}\zeta(3)^{7/12}e^{2\zeta'(-1)}\sum_{k=0}^{\infty}b_{2k}\left(\frac{\zeta(3)}{8}\right)^k g\left(n\sqrt{\frac{\zeta(3)}{8}}, -\frac{1}{6} - 2k\right)$$

and b_{2k} is the coefficient of y^{2k} in the Maclaurin series of

$$\frac{h(2y)^5}{h(y)h(4y)^2},$$

and so forth. The additional terms $\varphi_3(n)$, $\varphi_4(n)$ appear in [30] and $\varphi_5(n)$, $\varphi_6(n)$ appear in [31]. Taken together, these terms provide remarkably accurate estimates of $p_2(n)$. Govindarajan & Prabhakar [32] revisited Almkvist's results, using a modified function

$$\tilde{g}(x,\gamma) = \frac{1}{2}\sum_{\nu=0}^{\infty}\frac{x^\nu}{\nu!\Gamma((3-\gamma+\nu)/2)}$$

that seems better behaved than $g(x,\gamma)$ and evidently does for $p_2(n)$ akin to what Rademacher's modification of Hardy–Ramanujan did for $p_1(n)$.

Addendum Recent Monte Carlo work indicates that [33]

$$\lim_{n\to\infty} n^{-3/4}\ln\left(p_3(n)\right) \approx 1.822 > 1.789... = \frac{2^{7/4}\pi}{3^{5/4}5^{1/4}},$$

contradicting [8, 9]. The asymptotics of solid partitions appear to differ sharply from those of line and plane partitions; in addition to sub-leading terms of order $n^{1/2}$, $n^{1/4}$ and $\ln(n)$, there seems to be an oscillatory function at the $n^{-1/4}$ level. Theory lags far behind numerical experimentation here. Let

$$\begin{aligned} 1+\sum_{n=1}^{\infty} q(n)x^n &= 1+x+4x^2+10x^3+26x^4+59x^5+141x^6 \\ &\quad +310x^7+692x^8+\cdots \\ &= \prod_{m=1}^{\infty}(1-x^m)^{-m(m+1)/2}. \end{aligned}$$

Although the MacMahon conjecture is incorrect ($p_3(n) \neq q(n)$ for $n > 5$), there is still a possibility that $p_3(n) \sim q(n)$ as $n \to \infty$. The conjectured asymptotics for $p_3(n)$ given earlier are validated asymptotics for $q(n)$. In a recent breakthrough, Kotěšovec [34] deduced that the multiplicative constant C for $q(n)$ is

$$2^{-157/96}15^{-13/96}\exp\left(-\frac{\zeta(3)}{8\pi^2}+\frac{75\zeta(3)^3}{2\pi^8}+\frac{\zeta'(-1)}{2}\right)\pi^{1/24} = 0.2135951604...$$

and details are yet forthcoming.

Let us consider one of many possible variations on 1-partitions. Define $\hat{p}_1(n)$ to be the number of partitions of n into integers, each of which may occur only an odd number of times. It can be shown that [35]

$$\hat{p}_1(n) \sim \frac{B}{2\pi\, n}\exp\left(2B\sqrt{n}\right)$$

where

$$B^2 = \frac{\pi^2}{12} + \int_0^1 \frac{\ln(1 + x - x^2)}{x} dx = \frac{\pi^2}{12} + 2\ln(\varphi)^2$$
$$= \frac{\pi^2}{12} + 0.4631296411... = (1.1338415562...)^2$$

and $\varphi = (1 + \sqrt{5})/2$ is the Golden mean.

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000041, A000219, and A000293.

[2] G. H. Hardy and S. Ramanujan, Asymptotic formulae in combinatory analysis, *Proc. London Math. Soc.* 17 (1918) 75–115; also in *Collected Papers of S. Ramanujan*, ed. G. H. Hardy, P. V. Seshu Aiyar, and B. M. Wilson, Cambridge Univ. Press, 1927, pp. 276–309.

[3] J. V. Uspensky, Asymptotic formulae for numerical functions which occur in the theory of partitions (in Russian), *Izv. Akad. Nauk SSSR* 14 (1920) 199–218.

[4] E. M. Wright, Asymptotic partition formulae. I, Plane partitions, *Quart. J. Math.* 2 (1931) 177–189.

[5] L. Mutafchiev and E. Kamenov, Asymptotic formula for the number of plane partitions of positive integers, *C. R. Acad. Bulgare Sci.* 59 (2006) 361–366; arXiv:math.CO/0601253; MR2229373.

[6] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[7] S. R. Finch, Glaisher-Kinkelin constant, *first volume*, pp. 135–145.

[8] V. Mustonen and R. Rajesh, Numerical estimation of the asymptotic behaviour of solid partitions of an integer, *J. Phys. A* 36 (2003) 6651–6659; arXiv:cond-mat/0303607; MR2004301.

[9] S. Balakrishnan, S. Govindarajan and N. Prabhakar, On the asymptotics of higher dimensional partitions, *J. Phys. A* 45 (2012) 055001; arXiv:1105.6231; MR2874261 (2012m:11142).

[10] B. Gordon and L. Houten, Notes on plane partitions. III, *Duke Math. J.* 36 (1969) 801–824; MR0248104 (40 #1358).

[11] M. S. Cheema and W. E. Conway, Numerical investigation of certain asymptotic results in the theory of partitions, *Math. Comp.* 26 (1972) 999–1005; MR0314756 (47 #3308).

[12] C. Knessl, Asymptotic behavior of high-order differences of the plane partition function, *Discrete Math.* 126 (1994) 179–193; MR1264486 (95b:11098).

[13] S. Shlosman, Applications of the Wulff construction to the number theory, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* 292 (2002) 153–160, 180; arXiv:math-ph/0109027; MR1944089 (2003k:11156).

[14] A. O. L. Atkin, P. Bratley, I. G. Macdonald, and J. K. S. McKay, Some computations for m-dimensional partitions, *Proc. Cambridge Philos. Soc.* 63 (1967) 1097–1100; MR0217029 (36 #124).

[15] D. E. Knuth, A note on solid partitions, *Math. Comp.* 24 (1970) 955–961; MR0277401 (43 #3134).

[16] F. Y. Wu, The infinite-state Potts model and restricted multidimensional partitions of an integer, *Math. Comput. Modelling* 26 (1997) 269–274; MR1492510 (98i:82020).

[17] H. Y. Huang and F. Y. Wu, The infinite-state Potts model and solid partitions of an integer, *Internat. J. Modern Phys. B* 11 (1997) 121–126; MR1435061 (98c:11145).

[18] H. Rademacher, A convergent series for the partition function $p(n)$, *Proc. Nat. Acad. Sci. USA* 23 (1937) 78–84.
[19] H. Rademacher, On the partition function $p(n)$, *Proc. London Math. Soc.* 43 (1937) 241–254.
[20] H. Rademacher, On the expansion of the partition function in a series, *Annals of Math.* 44 (1943) 416–422; MR0008618 (5,35a).
[21] G. H. Hardy, *Ramanujan: Twelve Lectures on Subjects Suggested by His Life and Work*, Chelsea, 1940, pp. 83–100, 113–131; MR0106147 (21 #4881).
[22] R. Ayoub, *An Introduction to the Analytic Theory of Numbers*, Amer. Math. Soc., 1963, pp 135–205; MR0160743 (28 #3954).
[23] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, 1976, pp. 68–87; MR1634067 (99c:11126).
[24] T. M. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, 2$^{\text{nd}}$ ed., Springer-Verlag, 1990, pp. 94–112; MR1027834 (90j:11001).
[25] G. Almkvist and H. S. Wilf, On the coefficients in the Hardy-Ramanujan-Rademacher formula for $p(n)$, *J. Number Theory* 50 (1995) 329–334; MR1316828 (96e:11129).
[26] W. de Azevedo Pribitkin, Revisiting Rademacher's formula for the partition function $p(n)$, *Ramanujan J.* 4 (2000) 455–467; MR1811909 (2001m:11176).
[27] D. H. Lehmer, On the Hardy-Ramanujan series for the partition function, *J. London Math. Soc.* 12 (1937) 171–176.
[28] D. H. Lehmer, On the series for the partition function, *Trans. Amer. Math. Soc.* 43 (1938) 271–295.
[29] D. H. Lehmer, On the remainders and convergence of the series for the partition function, *Trans. Amer. Math. Soc.* 46 (1939) 362–373; MR0000410 (1,69c).
[30] G. Almkvist, A rather exact formula for the number of plane partitions, *A Tribute to Emil Grosswald: Number Theory and Related Analysis*, ed. M. Knopp and M. Sheingorn, Amer. Math. Soc., 1993, pp. 21–26; MR1210508 (94b:11103).
[31] G. Almkvist, Asymptotic formulas and generalized Dedekind sums, *Experim. Math.* 7 (1998) 343–359; MR1678083 (2000d:11126).
[32] S. Govindarajan and N. S. Prabhakar, A superasymptotic formula for the number of plane partitions, arXiv:1311.7227.
[33] N. Destainville and S. Govindarajan, Estimating the asymptotics of solid partitions, *J. Stat. Phys.* 158 (2015) 950–967; arXiv:1406.5605; MR3311486.
[34] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000294 and A255939.
[35] F. C. Auluck, K. S. Singwi and B. K. Agarwala, On a new type of partition, *Proc. Nat. Inst. Sci. India* 16 (1950) 147–156; MR0038365 (12,392f).

1.12 Class Number Theory

The problem of representing an integer as a sum of squares, or more generally as the value of a quadratic form, is very old and challenging [1–7]. We will barely scratch the surface of this enormous literature.

1.12.1 *Form Class Group*

A binary quadratic form $f(x, y) = ax^2 + bxy + cy^2$ with $a, b, c \in \mathbb{Z}$ is **primitive** if a, b, c are relatively prime and has **discriminant** $\delta_f = b^2 - 4ac$. The form f is **positive**

Table 1.1 *Interplay between m and D,* $-163 \leq D \leq 136$

m	-3	-1	-7	-2	-11	-15	-19	-5	-23	-6	-31	-35	$\ldots$	-163
D	-3	-4	-7	-8	-11	-15	-19	-20	-23	-24	-31	-35	$\ldots$	-163
m	5	2	3	13	17	21	6	7	29	33	37	10	$\ldots$	34
D	5	8	12	13	17	21	24	28	29	33	37	40	$\ldots$	136

definite if the matrix

$$\begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$$

is positive definite (meaning $a > 0$ and $\delta_f < 0$) and **indefinite** if $\delta_f > 0$. An integer d is a discriminant δ_f for some form f if and only if $d \equiv 0, 1 \bmod 4$. A discriminant $D \neq 0, 1$ is a **fundamental discriminant** assuming that

$$D = \begin{cases} m & \text{if } m \equiv 1 \bmod 4, \\ 4m & \text{if } m \equiv 2, 3 \bmod 4 \end{cases}$$

for some square-free integer m. Every nonsquare discriminant d can be uniquely expressed as De^2 where D is a fundamental discriminant and $e \geq 1$. A partial listing of fundamental discriminants appears in Table 1.1 and the correspondence $m \leftrightarrow D$ will be needed later [8].

Assume that D is a fundamental discriminant. Two quadratic forms f, g with $\delta_f = D = \delta_g$ are **properly equivalent** if there is a linear change of variables

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad ru - st = 1, \quad r, s, t, u \in \mathbb{Z}$$

for which $f(x, y) = g(x', y')$ always. We say that f, g are in the same **form class** and define the **form class number**

$$h^+(D) = \begin{cases} \text{the number of classes of primitive positive definite forms of discriminant } D & \text{if } D < 0, \\ \text{the number of classes of primitive indefinite forms of discriminant } D & \text{if } D > 0. \end{cases}$$

For example, $h^+(-4) = 1$ and $x^2 + y^2$ is a representative element of the unique form class of discriminant -4; $h^+(-20) = 2$ and $x^2 + 5y^2$, $2x^2 + 2xy + 3y^2$ are representative elements of the two corresponding classes of discriminant -20.

It is possible to endow the set of form classes, for fixed D, with the structure of an abelian group. We simply illustrate in the case $D = -4$:

$$(x_1^2 + y_1^2)(x_2^2 + y_2^2) = x_3^2 + y_3^2$$

where

$$x_3 = x_1 x_2 - y_1 y_2, \quad y_3 = x_1 y_2 + y_1 x_2;$$

and in the case $D = -20$:

$$(x_1^2 + 5y_1^2)(x_2^2 + 5y_2^2) = x_4^2 + 5y_4^2,$$

$$(x_1^2 + 5y_1^2)(2x_2^2 + 2x_2y_2 + 3y_2^2) = 2x_5^2 + 2x_5y_5 + 3y_5^2,$$

$$(2x_1^2 + 2x_1y_1 + 3y_1^2)(2x_2^2 + 2x_2y_2 + 3y_2^2) = x_6^2 + 5y_6^2$$

where

$$x_4 = x_1x_2 - 5y_1y_2, \quad y_4 = x_1y_2 + y_1x_2,$$

$$x_5 = x_1x_2 - y_1x_2 - 3y_1y_2, \quad y_5 = x_1y_2 + 2y_1x_2 + y_1y_2,$$

$$x_6 = 2x_1x_2 + x_1y_2 + y_1x_2 - 2y_1y_2, \quad y_6 = x_1y_2 + y_1x_2 + y_1y_2.$$

This multiplication is called **Gaussian composition** and is perhaps best understood via the following section.

We discuss two variations of the preceding. If the determinant of the linear transformation $(x, y) \mapsto (x', y')$ is allowed to be $ru - st = \pm 1$, then the corresponding number of equivalence classes is [9]

$$\hat{h}(D) = \frac{1}{2}\left(h^+(D) + 2^{\omega(D)-1}\right)$$

where $\omega(n)$ denotes the number of distinct prime factors of $|n|$. Rephrasing, $h^+(D)$ is the number of orbits under the action of the matrix group $\mathrm{SL}_2(\mathbb{Z})$ on the primitive binary quadratic forms of discriminant D, while $\hat{h}(D)$ is the same under the action of $\mathrm{GL}_2(\mathbb{Z})$. For instance, $h^+(-23) = 3 > 2 = \hat{h}(-23)$ and $h^+(136) = 4 > 3 = \hat{h}(136)$.

The second variation seems quite artificial but is actually important. Two quadratic forms f, g with $\delta_f = D = \delta_g$ are **vulgarly equivalent** if there is a linear change of variables

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad ru - st = \theta = \pm 1, \quad r, s, t, u \in \mathbb{Z}$$

for which $f(x, y) = \theta\, g(x', y')$ always. Note the factor θ in front of g. Define $h(D)$ to be the number of vulgar equivalence classes of primitive quadratic forms of discriminant D. Note here that forms are not assumed to be positive definite for $D < 0$. As an example, $h^+(12) = 2 > 1 = h(12)$ since the forms $-3x^2 + y^2$ and $-x^2 + 3y^2$ are not properly equivalent, but are vulgarly equivalent via the assignment $(x', y') = (y, x)$.

1.12.2 *Ideal Class Group*

Let $m \neq 0, 1$ be a square-free integer. The **quadratic number field**

$$\mathbb{Q}(\sqrt{m}) = \mathbb{Q} + \mathbb{Q}\sqrt{m} = \left\{u + v\sqrt{m} : u, v \in \mathbb{Q}\right\}$$

is the smallest subfield of $\mathbb{C}$ containing both $\mathbb{Q}$ and $\sqrt{m}$. An element $\alpha \in \mathbb{Q}(\sqrt{m})$ is an **algebraic integer** if it is a zero of a monic polynomial $z^2 + bz + c$ with $b, c \in \mathbb{Z}$. The set of algebraic integers of $\mathbb{Q}(\sqrt{m})$ is the subring

$$\mathcal{O}_m = \begin{cases} \mathbb{Z} + \mathbb{Z}\sqrt{m} & \text{if } m \equiv 2, 3 \bmod 4, \\ \mathbb{Z} + \mathbb{Z}\frac{1+\sqrt{m}}{2} & \text{if } m \equiv 1 \bmod 4 \end{cases}$$

of $\mathbb{Q}(\sqrt{m})$, often called the **maximal order** or simply the **integers**. Using the correspondence between the **radicand** m and the fundamental discriminant D, we have

$$\mathbb{Q}(\sqrt{m}) = \mathbb{Q}(\sqrt{D}), \quad \mathcal{O}_m = \mathbb{Z} + \mathbb{Z}\tfrac{D+\sqrt{D}}{2}.$$

For example, $\mathcal{O}_{-1}$ is the ring of Gaussian integers. In $\mathcal{O}_{-5}$, we have a surprising failure of unique factorization:

$$6 = 2 \cdot 3 = (1 + \sqrt{-5})(1 - \sqrt{-5}).$$

More will be said about this momentarily.

An **ideal** I of $\mathcal{O}_m$ is an additive subgroup of $\mathcal{O}_m$ with the property that, if $\alpha \in I$ and $\rho \in \mathcal{O}_m$, then $\rho\alpha \in I$. The set

$$(\alpha) = \{\rho\alpha : \rho \in \mathcal{O}_m\}$$

is the ideal of all multiples of a single element $\alpha \in \mathcal{O}_m$ and is called a **principal ideal**. The ideal

$$(\alpha_1, \alpha_2) = \{\rho_1\alpha_1 + \rho_2\alpha_2 : \rho_1, \rho_2 \in \mathcal{O}_m\}$$

is **nonprincipal** if $(\alpha_1, \alpha_2) \neq (\alpha_3)$ for any $\alpha_3 \in \mathcal{O}_m$. The **product** IJ of two ideals is the ideal of all finite sums of products of the form $\alpha\beta$ with $\alpha \in I$ and $\beta \in J$. In $\mathcal{O}_{-5}$, the principal ideal (6) can be written as

$$\begin{aligned} (6) &= (2)(3) = I_1^2 I_2 I_3 \\ &= \left(1 + \sqrt{-5}\right)\left(1 - \sqrt{-5}\right) = I_1 I_2 I_1 I_3 \end{aligned}$$

where

$$I_1 = \left(2, 1 + \sqrt{-5}\right) = \left(2, 1 - \sqrt{-5}\right),$$

$$I_2 = \left(3, 1 + \sqrt{-5}\right), \quad I_3 = \left(3, 1 - \sqrt{-5}\right).$$

Thus the two distinct factorizations of the number 6 in $\mathcal{O}_{-5}$ come from permuting I_1, I_2, I_3 in the factorization of the ideal (6).

Given $\alpha = u + v\sqrt{m} \in \mathbb{Q}(\sqrt{m})$, define its **conjugate** $\bar{\alpha} = u - v\sqrt{m}$ and its **norm** $N(\alpha) = \alpha\bar{\alpha} = u^2 - mv^2$. If $\alpha \in \mathcal{O}_m$, then clearly $\bar{\alpha} \in \mathcal{O}_m$ and $N(\alpha) \in \mathbb{Z}$. Given an ideal I of $\mathcal{O}_m$, define its conjugate $\bar{I} = \{\bar{\alpha} : \alpha \in I\}$ and its norm $N(I) = \gcd\{N(\alpha) : \alpha \in I\}$. For example, if I is the principal ideal (α), then $\bar{I} = (\bar{\alpha})$ and $N(I) = |N(\alpha)|$. If I and J are two ideals, then $N(IJ) = N(I)N(J)$; also $I\bar{I} = (N(I))$ is principal.

Two ideals I, J of $\mathcal{O}_m$ are **strictly equivalent** if there exist $\alpha, \beta \in \mathcal{O}_m$ such that

$$(\alpha)I = (\beta)J, \quad N(\alpha\beta) > 0.$$

We say that I, J are in the same **narrow ideal class** and define H_m^+ to be the finite abelian group of ideals modulo this relation. If the requirement that $N(\alpha\beta) > 0$ is removed, we instead say that I, J are in the same **wide ideal class** and define H_m analogously. H_m^+ is called the **narrow class group** and its cardinality h_m^+ is the **narrow class number**. The name for H_m is often abbreviated simply to **class group**. The **class number** h_m can be found in terms of h_m^+ via

$$h_m = \begin{cases} h_m^+ & \text{if } m < 0 \text{ or } (m > 0 \text{ and } N(\varepsilon) = -1)\,, \\ \frac{1}{2}h_m^+ & \text{if } m > 0 \text{ and } N(\varepsilon) = 1, \end{cases}$$

where ε is the **fundamental unit** of $\mathcal{O}_m$ (to be defined in the next section). Group-theoretic properties of H_m and the efficient computation of h_m have attracted much attention in recent years.

It turns out that the abelian group of classes of primitive binary quadratic forms of discriminant D is isomorphic to the narrow class group H_m^+, where the interplay $m \leftrightarrow D$ was described earlier. In particular, Gaussian composition of forms can be elegantly written using ideals and $h^+(D) = h_m^+$; see Tables 1.2 and 1.3 [10]. By the same reasoning, we have $h(D) = h_m$ but no interpretation of $\hat{h}(D)$ in ideal class theory seems to be useful. Our convention for treating the discriminant D as an argument and the radicand m as a subscript is perhaps new.

A maximal order $\mathcal{O}_m$ is a UFD (unique factorization domain) if and only if it is a PID (principal ideal domain), which is true if and only if $h_m = 1$. Also, $h_m \le 2$ if and only if any two decompositions of $\alpha \in \mathcal{O}_m$ into products of irreducible elements must possess the same number of factors [11–14]. Hence the class number measures, in a vague sense, how far $\mathcal{O}_m$ is from being a UFD.

Table 1.2 *Class numbers as functions of m,* $-163 \le m \le 34$

m	-1	-2	-3	-5	-6	-7	-10	-11	-13	-14	-15	-17	…	-163
h_m	1	1	1	2	2	1	2	1	2	4	2	4	…	1
$\hat{h}_m$	1	1	1	2	2	1	2	1	2	3	2	3	…	1
m	2	3	5	6	7	10	11	13	14	15	17	19	…	34
h_m^+	1	2	1	2	2	2	2	1	2	4	1	2	…	4
h_m	1	1	1	1	1	2	1	1	1	2	1	1	…	2
$\hat{h}_m$	1	2	1	2	2	2	2	1	2	4	1	2	…	3

Table 1.3 *Class numbers as functions of D*, $-163 \le D \le 136$

D	-3	-4	-7	-8	-11	-15	-19	-20	-23	-24	-31	-35	$\dots$	-163
$h(D)$	1	1	1	1	1	2	1	2	3	2	3	2	$\dots$	1
$\hat{h}(D)$	1	1	1	1	1	2	1	2	2	2	2	2	$\dots$	1
D	5	8	12	13	17	21	24	28	29	33	37	40	$\dots$	136
$h^+(D)$	1	1	2	1	1	2	2	2	1	2	1	2	$\dots$	4
$h(D)$	1	1	1	1	1	1	1	1	1	1	1	2	$\dots$	2
$\hat{h}(D)$	1	1	2	1	1	2	2	2	1	2	1	2	$\dots$	3

1.12.3 *Fundamental Unit*

Let $m > 1$ be square-free. A **unit** $\varepsilon \in \mathcal{O}_m$ satisfies $N(\varepsilon) = \pm 1$; it is the **fundamental unit** if $\varepsilon > 1$ and every other unit is of the form $\pm\varepsilon^n$, $n \in \mathbb{Z}$. Here is a conceptually simple algorithm for computing ε. If $m \equiv 2, 3 \bmod 4$, calculate mb^2 for $b = 1, 2, 3, \dots$ and stop at the first integer mb_0^2 that differs from a square a_0^2 by exactly ± 1; then $\varepsilon = a_0 + b_0\sqrt{m}$. If $m \equiv 1 \bmod 4$, stop instead at the first integer mb_0^2 that differs from a square a_0^2 by exactly ± 4; then $\varepsilon = (a_0 + b_0\sqrt{m})/2$. In both cases, we assume that $a_0 \ge 1$.

Two alternative algorithms involve continued fractions [15, 16]. For the first, define

$$\mu = \begin{cases} \frac{1+\sqrt{m}}{2} & \text{if } m \equiv 1 \bmod 4, \\ \sqrt{m} & \text{if } m \equiv 2, 3 \bmod 4 \end{cases} = \frac{P_0 + \sqrt{m}}{Q_0}$$

and let the (eventually periodic) continued fraction expansion of μ be

$$\mu = c_0 + \frac{1|}{|c_1} + \frac{1|}{|c_2} + \frac{1|}{|c_3} + \cdots.$$

Define

$$P_{j+1} = c_j Q_j - P_j, \quad Q_{j+1} = \frac{m - P_{j+1}^2}{Q_j}$$

for $j \ge 0$, so that

$$\frac{P_j + \sqrt{m}}{Q_j} = c_j + \frac{1|}{|c_{j+1}} + \frac{1|}{|c_{j+2}} + \frac{1|}{|c_{j+3}} + \cdots$$

and hence

$$\varepsilon = \prod_{j=1}^{\lambda} \frac{P_j + \sqrt{m}}{Q_j}$$

where λ is the period length of the continued fraction expansion for μ.

The second possesses a curiously ambiguous outcome. Let

$$\sqrt{m} = d_0 + \frac{1|}{|d_1} + \frac{1|}{|d_2} + \frac{1|}{|d_3} + \cdots$$

and define

$$A_0 = d_0, \quad A_1 = d_0 d_1 + 1, \quad B_0 = 1, \quad B_1 = d_1,$$

$$A_k = d_k A_{k-1} + A_{k-2}, \quad B_k = d_k B_{k-1} + B_{k-2},$$

for $k \geq 2$, so that

$$\frac{A_k}{B_k} = d_0 + \frac{1|}{|d_1} + \frac{1|}{|d_2} + \cdots + \frac{1|}{|d_k} = \text{the } k^{\text{th}} \text{ convergent of } \sqrt{m}.$$

Let l denote the period length of the continued fraction expansion for $\sqrt{m}$. It can be proved that, if $m \not\equiv 5 \bmod 8$, then $\varepsilon = A_{l-1} + B_{l-1}\sqrt{m}$. If $m \equiv 5 \bmod 8$, however, all we can conclude is that $A_{l-1} + B_{l-1}\sqrt{m}$ is either ε or ε^3. See Tables 1.4 and 1.5 [17].

A fast method to compute the set of square-free $m > 1$ for which $N(\varepsilon) = -1$ (equivalently, l is odd) is not known [18–23]. Likewise, the set of $m \equiv 5 \bmod 8$ for which $A_{l-1} + B_{l-1}\sqrt{m} = \varepsilon^3$ remains only partially understood [24–30]. Since ε can be exponentially large in m, the **regulator** $\ln(\varepsilon)$ is often used instead [31]. Hallgren [32, 33] gave a polynomial-time algorithm for computing $\ln(\varepsilon)$ that is based on a quantum Fourier transform period finding technique.

Another formula is $\varepsilon = (x + y\sqrt{D})/2$, where x, y are the smallest positive integer solutions of the Pell equation $x^2 - Dy^2 = \pm 4$. It follows immediately that $N(\varepsilon) = -1$ if and only $x^2 - Dy^2 = -4$. Let us define $\varepsilon^+ = (z + w\sqrt{D})/2$, where z, w are the smallest positive integer solutions of $z^2 - Dw^2 = 4$. Clearly $h^+(D)\ln(\varepsilon^+) = 2h(D)\ln(\varepsilon)$ for all $D > 0$; we will need ε^+ later.

Table 1.4 *Fundamental unit ε and norm $N(\varepsilon)$ as functions of m, $2 \leq m \leq 17$*

m	2	3	5	6	7	10	11	13	14	15	17
ε	$\frac{1+\sqrt{2}}{1}$	$\frac{2+\sqrt{3}}{1}$	$\frac{1+\sqrt{5}}{2}$	$\frac{5+2\sqrt{6}}{1}$	$\frac{8+3\sqrt{7}}{1}$	$\frac{3+\sqrt{10}}{1}$	$\frac{10+3\sqrt{11}}{1}$	$\frac{3+\sqrt{13}}{2}$	$\frac{15+4\sqrt{14}}{1}$	$\frac{4+\sqrt{15}}{1}$	$\frac{4+\sqrt{17}}{1}$
$N(\varepsilon)$	-1	$+1$	-1	$+1$	$+1$	-1	$+1$	-1	$+1$	$+1$	-1

Table 1.5 *Fundamental unit ε and norm $N(\varepsilon)$ as functions of D, $5 \leq D \leq 37$*

D	5	8	12	13	17	21	24	28	29	33	37
ε	$\frac{1+\sqrt{5}}{2}$	$\frac{1+\sqrt{2}}{1}$	$\frac{2+\sqrt{3}}{1}$	$\frac{3+\sqrt{13}}{2}$	$\frac{4+\sqrt{17}}{1}$	$\frac{5+\sqrt{21}}{2}$	$\frac{5+2\sqrt{6}}{1}$	$\frac{8+3\sqrt{7}}{1}$	$\frac{5+\sqrt{29}}{2}$	$\frac{23+4\sqrt{33}}{1}$	$\frac{6+\sqrt{37}}{1}$
$N(\varepsilon)$	-1	-1	$+1$	-1	-1	$+1$	$+1$	$+1$	-1	$+1$	-1

1.12.4 *Ideal Statistics over D*

The study of ideal class numbers as functions of fundamental discriminant D (equivalently, radicand m) has occupied mathematicians for centuries. Heegner [34], Stark [35–37], Baker [38], Deuring [39] and Siegel [40, 41] solved Gauss' class number one problem: $h(D)=1$ for $D=-3$, -4, -7, -8, -11, -19, -43, -67, -163 and for no other $D<-163$. See [42–51] for related work in the imaginary case. With respect to the real case, Gauss conjectured that $h(D)=1$ for infinitely many $D>0$, but a proof remains unknown.

Siegel [52–56] showed that

$$\ln(h(D)) \sim \ln(\sqrt{-D}) \quad \text{as } D \to -\infty,$$

$$\ln(h(D)\ln(\varepsilon)) \sim \ln(\sqrt{D}) \quad \text{as } D \to \infty$$

and the following mean value results apply [57–60]:

$$\sum_{0<-D<x} h(D) \sim \frac{c}{3\pi}x^{3/2}, \quad \sum_{0<D<x} h(D)\ln(\varepsilon) \sim \frac{c}{6}x^{3/2}$$

as $x \to \infty$, where [61]

$$c = \prod_p \left(1 - \frac{1}{p^2(p+1)}\right) = 0.8815138397...$$

and the infinite product is over all primes p. We may alternatively write

$$\lim_{x\to\infty} \mathrm{E}\left(\frac{h(D)}{\sqrt{-D}} \mid 0<-D<x\right) = \frac{\pi c}{6} = 0.4615595671...$$

$$\lim_{x\to\infty} \mathrm{E}\left(\frac{h(D)\ln(\varepsilon)}{\sqrt{D}} \mid 0<D<x\right) = \frac{\pi^2 c}{12} = 0.7250160726...$$

because $\sum_{0<-D<x} 1 \sim (3/\pi^2)x \sim \sum_{0<D<x} 1$ and since partial summation contributes an additional factor of $3/2$.

Taniguchi [62] conjectured a second-order analog

$$\sum_{0<-D<x} h(D)^2 \sim \frac{\pi^2 C'}{144}x^2, \quad \sum_{0<D<x} h(D)^2 \ln(\varepsilon)^2 \sim \frac{\pi^4 C'}{576}x^2$$

as $x \to \infty$, where [63]

$$C' = \prod_p \left(1 - \frac{3}{p^3} + \frac{2}{p^4} + \frac{1}{p^5} - \frac{1}{p^6}\right) = 0.6782344919....$$

With regard to extreme values, Granville & Soundararajan [64] suggested that perhaps

$$\max_{|D|<x} L(D) = e^{\gamma}(\ln\ln x + \ln\ln\ln x + c'' + o(1))$$

where γ is Euler's constant,

$$L(D) = \begin{cases} \dfrac{\pi h(D)}{\sqrt{-D}} & \text{if } D < -4, \\[2ex] \dfrac{2h(D)\ln(\varepsilon)}{\sqrt{D}} & \text{if } D > 4 \end{cases}$$

and

$$c'' = \int_0^1 \frac{\tanh(y)}{y} dy + \int_1^\infty \frac{\tanh(y) - 1}{y} dy = 0.8187801401...$$

Is it possible in any of these formulas, when $D > 0$, to somehow separate the class number and the regulator?

1.12.5 *Cohen-Lenstra Heuristics*

We merely state certain conjectures due to Cohen & Lenstra [65–70]. Define $\tilde{H}_m$ to be the odd part of the class group H_m, that is, $\tilde{H}_m$ is the subgroup of all elements in H_m of odd order. Let [71, 72]

$$C = \prod_{j=2}^{\infty} \zeta(j) = 2.2948565916...,$$

$$\Delta = \frac{\pi^2}{6} \prod_p \left(1 + \frac{1}{p^2(p-1)}\right) = 2.2038565964...$$

and, when q is prime,

$$\eta(q) = \prod_{k=1}^{\infty} \left(1 - \frac{1}{q^k}\right)$$

(which appeared in [73] for the special case $q = 2$). For random $m < 0$, it is believed that

- the probability that $\tilde{H}_m$ is cyclic is

$$\frac{\pi^2}{18} \frac{\zeta(3)}{\zeta(6)} \frac{1}{C\eta(2)} = 0.9775748102...$$

- if p is an odd prime, the probability that $p|h_m$ is

$$1 - \eta(p) = \begin{cases} 0.4398739220... & \text{if } p = 3, \\ 0.2396672041... & \text{if } p = 5, \\ 0.1632045929... & \text{if } p = 7 \end{cases}$$

and, likewise, for random $m > 0$,

- the probability that $\tilde{H}_m$ is cyclic is

$$\frac{3}{10} \frac{\Delta}{C\eta(2)} = 0.9976305717...$$

- if p is an odd prime, the probability that $p|h_m$ is

$$1-\left(1-\frac{1}{p}\right)^{-1}\eta(p)=\begin{cases}0.1598108831... & \text{if } p=3,\\ 0.0495840051... & \text{if } p=5,\\ 0.0237386917... & \text{if } p=7\end{cases}$$

- the probability that $h_m=1$, given that m itself is prime, is

$$\frac{1}{2C\eta(2)}=0.7544581722....$$

A proof of any of these conjectures would be a welcome breakthrough! See [74] for partial results concerning the prime $p=3$.

1.12.6 *Form Statistics over d*

Given a nonsquare discriminant d, define $h^+(d)$ and $\varepsilon^+(d)$ exactly as before (with D simply replaced by d). We had no need of such generalizations until now. See Table 1.6 [75].

Lipschitz [76], Mertens [77] and Siegel [78] proved that

$$\sum_{0<-d<x} h^+(d)\sim\frac{\pi}{18\zeta(3)}x^{3/2},\quad \sum_{0<d<x} h^+(d)\ln(\varepsilon^+)\sim\frac{\pi^2}{18\zeta(3)}x^{3/2}$$

as $x\to\infty$, where the sums are taken over all $d\equiv 0,1 \bmod 4$ that are not squares. Their efforts confirmed conjectures of Gauss [79–82]:

$$\sum_{\substack{0<-d<4x,\\ 4|d}} h^+(d)\sim\frac{4\pi}{21\zeta(3)}x^{3/2},\quad \sum_{\substack{0<d<4x,\\ 4|d}} h^+(d)\ln(\varepsilon^+)\sim\frac{4\pi^2}{21\zeta(3)}x^{3/2}.$$

When searching through the literature, it is helpful to be aware of Gauss's convention (that $d=4k$ or, equivalently, $f(x,y)=ax^2+2bxy+cy^2$) versus Eisenstein's convention (no parity requirement on the middle coefficient). We have adopted

Table 1.6 *Class number $h^+(d)$ for $-23\le d\le 32$; also $\varepsilon^+(d)$ for $5\le d\le 32$*

d	-3	-4	-7	-8	-11	-12	-15	-16	-19	-20	-23
$h^+(d)$	1	1	1	1	1	1	2	1	1	2	3
d	5	8	12	13	17	20	21	24	28	29	32
$h^+(d)$	1	1	2	1	1	1	2	2	2	1	2
$\varepsilon^+(d)$	$\frac{3+\sqrt{5}}{2}$	$\frac{3+2\sqrt{2}}{1}$	$\frac{2+\sqrt{3}}{1}$	$\frac{11+3\sqrt{13}}{2}$	$\frac{33+8\sqrt{17}}{1}$	$\frac{9+4\sqrt{5}}{1}$	$\frac{5+\sqrt{21}}{2}$	$\frac{5+2\sqrt{6}}{1}$	$\frac{8+3\sqrt{7}}{1}$	$\frac{27+5\sqrt{29}}{2}$	$\frac{3+2\sqrt{2}}{1}$

the latter, as do most contemporary authors. For example,

$$\begin{aligned}\lim_{x\to\infty} \mathrm{E}\left(\frac{h^+(d)}{\sqrt{-d}} \,|\, 0<-d<4x,\, d=4k\right) &= \frac{\pi}{7\zeta(3)} = 0.3733591557... \\ &= \frac{1.1729423808...}{\pi}\end{aligned}$$

in Gauss' scheme and

$$\begin{aligned}\lim_{x\to\infty} \mathrm{E}\left(\frac{h^+(d)\ln(\varepsilon^+)}{\sqrt{d}} \,|\, 0<d<x\right) &= \frac{\pi^2}{6\zeta(3)} = 1.3684327776... \\ &= 2(0.6842163888...)\end{aligned}$$

in Eisenstein's scheme. A second-moment analog of the latter is due to Barban [83–89]:

$$\begin{aligned}\lim_{x\to\infty} \mathrm{E}\left(\frac{h^+(d)^2\ln(\varepsilon^+)^2}{d} \,|\, 0<d<x\right) &= \prod_p \left(1+\frac{3p^2-1}{(p^2-1)p(p+1)}\right) \\ &= 2.5965362904.... = \frac{29}{18}(1.6116432147...).\end{aligned}$$

In fact, the probability distributions [90–95]

$$\lim_{x\to\infty} \mathrm{P}\left\{\ln\left(\frac{h^+(d)\ln(\varepsilon^+)}{\sqrt{d}}\right) \le s \,|\, 0<d<x\right\},$$

$$\lim_{x\to\infty} \mathrm{P}\left\{\ln\left(\frac{\pi h^+(d)}{\sqrt{-d}}\right) \le s \,|\, 0<-d<x\right\}$$

both coincide with the distribution of $S=\sum_p X_p$, an infinite sum of independent random variables, where

$$X_p = \begin{cases} 0 & \text{with probability } \frac{1}{p}, \\ -\ln\left(1-\frac{1}{p}\right) & \text{with probability } \frac{1}{2}\left(1-\frac{1}{p}\right), \\ -\ln\left(1+\frac{1}{p}\right) & \text{with probability } \frac{1}{2}\left(1-\frac{1}{p}\right) \end{cases}$$

for each prime number p.

We mention finally Hooley's conjecture [96]

$$\sum_{\substack{0<d<4x,\\ 4|d}} h^+(d) \sim \frac{25}{12\pi^2} x\ln(x)^2$$

and wonder if this (and other attempts to separate the class number and the regulator when $d>0$) someday can be verified.

Table 1.7 *Period length as a function of m,* $2 \le m \le 31$

m	2	3	5	6	7	10	11	13	14	15	17	19	21	22	23	26	29	30	31
l_m	1	2	1	2	4	1	2	5	4	2	1	6	6	6	4	1	5	2	8

1.12.7 *Continued Fraction Period Length*

Table 1.7 exhibits the period length l_m of the continued fraction expansion for $\sqrt{m}$, where $m > 1$ is square-free [97].

Very little can be said about the behavior of l_m. Podsypanin [98, 99] proved that

$$l_m = O\left(\sqrt{m}\ln(\ln(m))\right)$$

as $m \to \infty$, assuming the truth of the Extended Riemann Hypothesis. Williams [100, 101] gave evidence that the big O, on the one hand, can be replaced by

$$\frac{e^{\gamma}}{\ln(\varphi)} = 3.7012232975...$$

where φ is the Golden mean, or even

$$\frac{12e^{\gamma}\ln(2)}{\pi^2} = 1.5010271229....$$

It seems likely, on the other hand, that the values 1.05 or even 1.08 will *not* suffice. Pen & Skubenko [102] and Golubeva [103, 104] proved the inequality [105]

$$\frac{\ln(\varepsilon)}{\ln(4\sqrt{m})} < l_m < \frac{4\ln(\varepsilon)}{\ln(\varphi)} = 4(2.0780869212...)\ln(\varepsilon)$$

involving the fundamental unit ε of $\mathbb{Q}(\sqrt{m})$. This subject turns out to be related to what are called **Lévy constants** [106–109]:

$$\beta(\xi) = \lim_{k\to\infty} \frac{\ln(B_k)}{k}$$

where A_k/B_k is the k^{th} convergent of the quadratic irrational ξ. Let Σ denote the set of all such $\beta(\xi)$. It is known that $\Sigma \subseteq [\ln(\varphi), \infty)$ and that $\pi^2/(12\ln(2))$ is a limit point of Σ. It is also likely that Σ has a structure similar to the Markov spectrum [110] in the sense that a left-hand portion of Σ probably consists only of isolated points and a right-hand portion of Σ is much denser.

Let $3 < p \equiv 3 \bmod 4$ be prime and assume that $h_p = 1$. An astonishing formula due to Hirzebruch [111–114] states that

$$h_{-p} = \frac{1}{3}\sum_{j=1}^{l}(-1)^{l-j}d_j$$

where $d_1, d_2, \ldots, d_l$ is the sequence of denominators in one period of the continued fraction expansion for $\sqrt{p} - \lfloor\sqrt{p}\rfloor$. For example, $h_{23} = 1$ and $h_{-23} = (-1 + 3 - 1 + 8)/3 = 3$. Is an elementary proof of this theorem possible? What can be said if instead $p \equiv 1 \bmod 4$?

As an aside, there exist precisely 21 square-free integers m for which the pair $(\mathcal{O}_m, |N|)$ is a Euclidean domain, that is, for which $|N|$ is compatible with the division algorithm [16, 115–118]. Both $(\mathcal{O}_{14}, |N|)$ and $(\mathcal{O}_{69}, |N|)$ fail to be Euclidean, although $h_{14} = 1 = h_{69}$. An alternative function $N' : \mathcal{O}_{69} \to \mathbb{Z}$ can be constructed so that $(\mathcal{O}_{69}, |N'|)$ is Euclidean [119–123]; the proof turns out to be computer-assisted. Does such a construction exist for $\mathcal{O}_{14}$ [124, 125]?

As another aside, $h(j^2 + 4) > 1$ for odd $j > 17$ and $h(4k^2 + 1) > 1$ for $k > 13$. The arguments $j^2 + 4$ and $4k^2 + 1$ are assumed to be square-free. These two inequalities, known respectively as Yokoi's conjecture and Chowla's conjecture, were proved by Biró [126–130].

We have not discussed prime-producing polynomials [131], asymptotic $h(d)$-averages over subsets [132, 133], the theory of genera [1] or Dirichlet L-series, although our definition of $L(D)$ earlier provides some foreshadowing of the next essay.

[1] D. A. Cox, *Primes of the Form $x^2 + ny^2$. Fermat, Class Field Theory and Complex Multiplication*, Wiley, 1989; MR1028322 (90m:11016).

[2] D. A. Buell, *Binary Quadratic Forms. Classical Theory and Modern Computations*, Springer-Verlag, 1989; MR1012948 (92b:11021).

[3] P. Ribenboim, *My Numbers, My Friends. Popular Lectures on Number Theory*, Springer-Verlag, 2000, pp. 112–174; MR1761897 (2002d:11001).

[4] Z. I. Borevich and I. R. Shafarevich, *Number Theory*, Academic Press, 1966; MR0195803 (33 #4001).

[5] H. Cohn, *A Second Course in Number Theory*, Wiley, 1962; reissued as *Advanced Number Theory*, Dover, 1980, MR0133281 (24 #A3115).

[6] H. L. Keng, *Introduction to Number Theory*, Springer-Verlag, 1982; MR0665428 (83f:10001).

[7] R. A. Mollin, *Quadratics*, CRC Press, 1996; MR1383823 (97e:11135).

[8] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003657, A003658, and A037449.

[9] J. Jonasson, *Classes of Integral Binary Quadratic Forms*, Masters thesis, Göteborg University (2001).

[10] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000924, A003646, A003649, A003652, A006641, A104888, A106029, A106030, A106031 and A106032.

[11] L. Carlitz, A characterization of algebraic number fields with class number two, *Proc. Amer. Math. Soc.* 11 (1960) 391–392; MR0111741 (22 #2603).

[12] J. M. Masley, Where are number fields with small class number? *Number Theory, Carbondale 1979*, Proc. Southern Illinois conf., ed. M. B. Nathanson, Lect. Notes in Math. 751, Springer-Verlag, 1979, pp. 221–242; MR0564932 (81f:12004).

[13] A. Czogala, Arithmetic characterization of algebraic number fields with small class numbers, *Math. Z.* 176 (1981) 247–253; MR0607964 (82e:12006).
[14] F. Di Franco and F. Pace, Arithmetical characterization of rings of algebraic integers with class number three and four, *Boll. Un. Mat. Ital. D* 4 (1985) 63–69; MR0871911 (88c:11062).
[15] R. A. Mollin, K. Cheng and B. Goddard, The Diophantine equation $AX^2 - BY^2 = C$ solved via continued fractions, *Acta Math. Univ. Comenian.* 71 (2002) 121–138; MR1980374 (2004c:11031).
[16] W. Narkiewicz, *Elementary and Analytic Theory of Algebraic Numbers*, 3rd ed., Springer-Verlag, 2004, pp. 104–105, 115–118; MR2078267 (2005c:11131).
[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A014000, A014046, A048941, A048942, A053370, A053371, A053372, A053373, A053374 and A053375.
[18] B. D. Beach and H. C. Williams, A numerical investigation of the Diophantine equation $x^2 - dy^2 = -1$, *Proc. 3rd Southeastern Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1972, ed. F. Hoffman, R. B. Levow and R. S. D. Thomas, Congr. Numer. 6, Utilitas Math., 1972, pp. 37–68; MR0347729 (50 #231).
[19] P. Morton, On Rédei's theory of the Pell equation, *J. Reine Angew. Math.* 307/308 (1979) 373–398; MR0534233 (81f:12005).
[20] J. C. Lagarias, On the computational complexity of determining the solvability or unsolvability of the equation $X^2 - DY^2 = -1$, *Trans. Amer. Math. Soc.* 260 (1980) 485–508; MR0574794 (81g:10029).
[21] P. Stevenhagen, Frobenius distributions for real quadratic orders, *J. Théor. Nombres Bordeaux* 7 (1995) 121–132; MR1413571 (97g:11118).
[22] S. R. Finch, Pell-Stevenhagen constant, *first volume*, pp. 119–120.
[23] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003654, A003814, and A031396.
[24] P. Kaplan and K. S. Williams, Pell's equations $X^2 - mY^2 = -1, -4$ and continued fractions, *J. Number Theory* 23 (1986) 169–182; MR0845899 (87g:11035).
[25] A. J. Stephens and H. C. Williams, Some computational results on a problem of Eisenstein, *Théorie des nombres*, Proc. 1987 Québec conf., ed. J.-M. De Koninck and C. Levesque, de Gruyter, 1989, pp. 869–886; MR1024611 (91c:11066).
[26] H. C. Williams, Eisenstein's problem and continued fractions, *Utilitas Math*. 37 (1990) 145–157; MR1068514 (91h:11018).
[27] N. Ishii, P. Kaplan and K. S. Williams, On Eisenstein's problem, *Acta Arith*. 54 (1990) 323–345; MR1058895 (91f:11024).
[28] K. S. Williams and N. Buck, Comparison of the lengths of the continued fractions of $\sqrt{D}$ and $\frac{1}{2}(1+\sqrt{D})$, *Proc. Amer. Math. Soc.* 120 (1994) 995–1002; MR1169053 (94f:11062).
[29] P. Stevenhagen, On a problem of Eisenstein, *Acta Arith.* 74 (1996) 259–268; MR1373712 (97b:11138).
[30] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A107996, A107997, A107998, A107999 and A108160.
[31] M. J. Jacobson, R. F. Lukes and H. C. Williams, An investigation of bounds for the regulator of quadratic fields, *Experim. Math.* 4 (1995) 211–225; MR1387478 (97d:11173).
[32] S. Hallgren, Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem, *Proc. 34th ACM Symp. on Theory of Computing (STOC)*, Montreal, ACM, 2002, pp. 653–658.

[33] R. Jozsa, Notes on Hallgren's efficient quantum algorithm for solving Pell's equation, quant-ph/0302134.
[34] K. Heegner, Diophantische Analysis und Modulfunktionen, *Math. Z.* 56 (1952) 227–253; MR0053135 (14,725j).
[35] H. M. Stark, A complete determination of the complex quadratic fields of class-number one, *Michigan Math. J.* 14 (1967) 1–27; MR0222050 (36 #5102).
[36] H. M. Stark, On the "gap" in a theorem of Heegner, *J. Number Theory* 1 (1969) 16–27; MR0241384 (39 #2724).
[37] H. M. Stark, A historical note on complex quadratic fields with class-number one, *Proc. Amer. Math. Soc*. 21 (1969) 254–255; MR0237461 (38 #5743).
[38] A. Baker, Linear forms in the logarithms of algebraic numbers. I, *Mathematika* 13 (1966) 204–216; MR0220680 (36 #3732).
[39] M. Deuring, Imaginäre quadratische Zahlkörper mit der Klassenzahl Eins, *Invent. Math.* 5 (1968) 169–179; MR0228464 (37 #4044).
[40] C. L. Siegel, Zum Beweise des Starkschen Satzes, *Invent. Math.* 5 (1968) 180–191; *Gesammelte Abhandlungen*, v. 4, ed. K. Chandrasekharan and H. Maass, Springer-Verlag, 1966, pp. 41–52; MR0228465 (37 #4045).
[41] S. Chowla, The Heegner-Stark-Baker-Deuring-Siegel theorem, *J. Reine Angew. Math.* 241 (1970) 47–48; MR0258762 (41 #3408).
[42] A. Baker, Imaginary quadratic fields with class number 2, *Annals of Math.* 94 (1971) 139–152; MR0299583 (45 #8631).
[43] H. M. Stark, A transcendence theorem for class-number problems, *Annals of Math.* 94 (1971) 153–173; MR0297715 (45 #6767).
[44] H. M. Stark, On complex quadratic fields with class-number two, *Math. Comp.* 29 (1975) 289–302; MR0369313 (51 #5548).
[45] H. L. Montgomery and P. J. Weinberger, Notes on small class numbers, *Acta Arith.* 24 (1973/74) 529–542; MR0357373 (50 #9841).
[46] J. Oesterlé, Nombres de classes des corps quadratiques imaginaires, *Astérisque* 121–122 (1985) 309–323; MR0768967 (86k:11064).
[47] S. Arno, The imaginary quadratic fields of class number 4, *Acta Arith.* 60 (1992) 321–334; MR1159349 (93b:11144).
[48] C. Wagner, Class number 5, 6 and 7, *Math. Comp.* 65 (1996) 785–800; MR1333327 (96g:11135).
[49] S. Arno, M. L. Robinson and F. S. Wheeler, Imaginary quadratic fields with small odd class number, *Acta Arith*. 83 (1998) 295–330; MR1610549 (99a:11123).
[50] M. Watkins, Class numbers of imaginary quadratic fields, *Math. Comp*. 73 (2004) 907–938; MR2031415 (2005a:11175).
[51] D. Goldfeld, Gauss's class number problem for imaginary quadratic fields, *Bull. Amer. Math. Soc.* 13 (1985) 23–37; MR0788386 (86k:11065).
[52] C. L. Siegel, Über die Classenzahl quadratischer Zahlkörper, *Acta Arith.* 1 (1935) 83-86; *Gesammelte Abhandlungen*, v. I, ed. K. Chandrasekharan and H. Maass, Springer-Verlag, 1966, pp. 406–409.
[53] H. Heilbronn, On Dirichlet series which satisfy a certain functional equation, *Quart. J. Math*. 9 (1938) 194–195; *Collected Papers*, ed. E. J. Kani and R. A. Smith, Wiley, 1988, pp. 343–344.
[54] T. Estermann, On Dirichlet's *L* functions, *J. London Math. Soc.* 23 (1948) 275–279; MR0027797 (10,356c).
[55] S. Chowla, A new proof of a theorem of Siegel, *Annals of Math.* 51 (1950) 120–122; MR0033313 (11,420e).

[56] D. M. Goldfeld, A simple proof of Siegel's theorem, *Proc. Nat. Acad. Sci. U.S.A.* 71 (1974) 1055; MR0344222 (49 #8962).

[57] R. Ayoub, *An Introduction to the Analytic Theory of Numbers*, Amer. Math. Soc., 1963, pp. 320–323; MR0160743 (28 #3954).

[58] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993, pp. 232–234, 261–263, 290–293; MR1228206 (94i:11105).

[59] D. Goldfeld and J. Hoffstein, Eisenstein series of $1/2$-integral weight and the mean value of real Dirichlet L-series, *Invent. Math.* 80 (1985) 185–208; MR0788407 (86m:11029).

[60] B. A. Datskovsky, A mean-value theorem for class numbers of quadratic extensions, *A Tribute to Emil Grosswald: Number Theory and Related Analysis*, ed. M. Knopp and M. Sheingorn, Amer. Math. Soc., 1993, pp. 179–242; MR1210518 (94m:11137).

[61] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.

[62] T. Taniguchi, A mean value theorem for the square of class numbers of quadratic fields, math.NT/0410531.

[63] P. Sebah, Calculation of Taniguchi's constant, unpublished note (2005).

[64] A. Granville and K. Soundararajan, The distribution of values of $L(1, \chi_d)$, *Geom. Funct. Anal.* 13 (2003) 992–1028; math.NT/0206031; MR2024414 (2005d:11129).

[65] H. Cohen and H. W. Lenstra, Heuristics on class groups of number fields, *Number Theory, Noordwijkerhout 1983*, Proc. 13th Journées Arithmétiques, ed. H. Jager, Lect. Notes in Math. 1068, Springer-Verlag, 1984, pp. 33–62; MR0756082 (85j:11144).

[66] D. A. Buell, Class groups of quadratic fields. II, *Math. Comp.* 48 (1987) 85–93; MR0866100 (87m:11109).

[67] D. A. Buell, The last exhaustive computation of class groups of complex quadratic number fields, *Number Theory*, Proc. 1996 Canad. Number Theory Assoc. Ottawa conf., ed. R. Gupta and K. S. Williams, Amer. Math. Soc., 1999, pp. 35–53; MR1684589 (2000d:11156).

[68] L. C. Washington, Some remarks on Cohen-Lenstra heuristics, *Math. Comp.* 47 (1986) 741–747; MR0856717 (87j:11115).

[69] M. J. Jacobson, Experimental results on class groups of real quadratic fields (extended abstract), *Proc. 1998 Algorithmic Number Theory Sympos. (ANTS-III)*, Portland, ed. J. P. Buhler, Lect. Notes in Comp. Sci. 1423, Springer-Verlag, 1998, pp. 463–474; MR1726094 (2000h:11119).

[70] H. te Riele and H. Williams, New computations concerning the Cohen-Lenstra heuristics, *Experim. Math.* 12 (2003) 99–113; MR2002677 (2005d:11183).

[71] S. R. Finch, Abelian group enumeration constants, *first volume*, pp. 273–276.

[72] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.

[73] S. R. Finch, Digital search tree constants, *first volume*, pp. 354–361.

[74] H. Davenport and H. Heilbronn, On the density of discriminants of cubic fields. II, *Proc. Royal Soc. London Ser. A* 322 (1971) 405–420; *Collected Papers of Hans Arnold Heilbronn*, ed. E. J. Kani and R. A. Smith, Wiley, 1988, pp. 532–547; MR0491593 (58 #10816).

[75] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000037, A014600, A014601, A033313, A033317, A077425, A077428, A078355, A079896 and A087048.

[76] R. Lipschitz, Über die asymptotischen Gesetze von gewissen Gattungen zahlentheoretischer Functionen, *Monatsberichte der Königlichen Preussische Akademie der*

Wissenschaften zu Berlin. Sitzung der physikalisch-mathematischen Klasse (1865) 174–185.

[77] F. Mertens, Über einige asymptotische Gesetze der Zahlentheorie, *J. Reine Angew. Math*. 77 (1874) 289–338.

[78] C. L. Siegel, The average measure of quadratic forms with given determinant and signature, *Annals of Math*. 45 (1944) 667–685; *Gesammelte Abhandlungen*, v. II, ed. K. Chandrasekharan and H. Maass, Springer-Verlag, 1966, pp. 473–491; MR0012642 (7,51a).

[79] I. M. Vinogradov, On the mean value of the number of classes of properly primitive forms of negative discriminant (in Russian), *Soobchsheniya Kharkovskogo Matematicheskogo Obshestva* 16 (1918) 10–38; Engl. transl. in *Selected Works*, ed. L. D. Faddeev, R. V. Gamkrelidze, A. A. Karacuba, K. K. Mardzhanishvili and E. F. Miscenko, Springer-Verlag, 1985, pp. 28–52.

[80] T. Shintani, On zeta-functions associated with the vector space of quadratic forms, *J. Fac. Sci. Univ. Tokyo Sect. I A Math.* 22 (1975) 25–65; MR0384717 (52 #5590).

[81] J. Hoffstein and M. Rosen, Average values of L-series in function fields, *J. Reine Angew. Math*. 426 (1992) 117–150; MR1155750 (93c:11022).

[82] F. Chamizo and H. Iwaniec, On the Gauss mean-value formula for class number, *Nagoya Math. J.* 151 (1998) 199–208; MR1650293 (99h:11041).

[83] M. B. Barban, Linnik's "great sieve" and a limit theorem for the class number of ideals of an imaginary quadratic field (in Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* 26 (1962) 573–580; MR0151441 (27 #1426).

[84] M. B. Barban, The "large sieve" method and its application to number theory (in Russian), *Uspehi Mat. Nauk* 21 (1966) 51–102; Engl. transl. in *Russian Math. Surveys* 21 (1966) 49–103; MR0199171 (33 #7320).

[85] M. B. Barban and G. Gordover, On moments of the number of classes of purely radical quadratic forms with negative determinant (in Russian), *Dokl. Akad. Nauk SSSR* 167 (1966) 267–269; Engl. transl. in *Soviet Math. Dokl.* 7 (1966) 356–358; MR0197401 (33 #5566).

[86] D. Wolke, Moments of the number of classes of primitive quadratic forms with negative discriminant, *J. Number Theory* 1 (1969) 502–511; MR0252322 (40 #5543).

[87] A. F. Lavrik, The moments of the number of classes of primitive quadratic forms of negative determinant (in Russian), *Dokl. Akad. Nauk SSSR* 197 (1971) 32–35; Engl. transl. in *Soviet Math. Dokl.* 12 (1971) 399–403; MR0280446 (43 #6166).

[88] M. Jutila, On character sums and class numbers, *J. Number Theory* 5 (1973) 203–214; MR0335449 (49 #230).

[89] P. Sebah, Calculation of Barban's constant, unpublished note (2005).

[90] S. Chowla and P. Erdős, A theorem on the distribution of the values of L-functions, *J. Indian Math. Soc.* 15 (1951) 11–18; MR0044566 (13,439a).

[91] A. S. Fainleib, The limit theorem for the number of classes of primitive quadratic forms with negative determinant (in Russian), *Dokl. Akad. Nauk SSSR* 184 (1969) 1048–1049; Engl. transl. in *Soviet Math. Dokl.* 10 (1969) 206–207; MR0244157 (39 #5474).

[92] P. D. T. A. Elliott, The distribution of the quadratic class number, *Litovsk. Mat. Sb.* 10 (1970) 189–197; MR0285505 (44 #2723).

[93] P. D. T. A. Elliott, *Probabilistic Number Theory. II. Central Limit Theorems*, Springer-Verlag, 1980, pp. 313–329; MR0560507 (82h:10002b).

[94] E. Stankus, Elementary method for estimating the complex moments of $L(1, \chi_D)$ (in Russian), *Litovsk. Mat. Sb.* 22 (1982) 160–170; Engl. transl. in *Lithuanian Math. J.* 22 (1982) 170–177; MR0659029 (83i:10054).

[95] M. Peter, The distribution of class numbers of pure number fields, unpublished note (2002).
[96] C. Hooley, On the Pellian equation and the class number of indefinite binary quadratic forms, *J. Reine Angew. Math.* 353 (1984) 98–131; MR0765829 (86d:11032).
[97] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A035015 and A107356.
[98] E. V. Podsypanin, The length of the period of a quadratic irrationality (in Russian), Studies in Number Theory, 5, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov (LOMI)* 82 (1979) 95–99, 166; Engl. transl. in *J. Soviet Math.* 18 (1982) 919–923; MR0537024 (80h:12002).
[99] J. Robertson, Comment on a proof by Podsypanin, unpublished note (2006).
[100] H. C. Williams, A numerical investigation into the length of the period of the continued fraction expansion of $\sqrt{D}$, *Math. Comp.* 36 (1981) 593–601; MR0606518 (82f:10011).
[101] C. D. Patterson and H. C. Williams, Some periodic continued fractions with long periods, *Math. Comp.* 44 (1985) 523–532; MR0777283 (86h:11113).
[102] A. S. Pen and B. F. Skubenko, An upper bound of the period of a quadratic irrationality (in Russian), *Mat. Zametki* 5 (1969) 413–418; MR0245524 (39 #6830).
[103] E. P. Golubeva, The length of a period of quadratic irrationality (in Russian), *Mat. Sbornik* 123 (1984) 120–129; Engl. transl. in *Math. USSR Sbornik* 51 (1985) 119–128; MR0728933 (85f:11044).
[104] E. P. Golubeva, On the lengths of the periods of a continued fraction expansion of quadratic irrationalities and on the class numbers of real quadratic fields (in Russian), Anal. Teor. Chisel i Teor. Funktsii. 8, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 160 (1987) 72–81, 297–298; Engl. transl. in *J. Soviet Math.* 52 (1990) 3049–3056; MR0906845 (88j:11055).
[105] S. R. Finch, Porter-Hensley constants, *first volume*, pp. 156–160.
[106] C. Faivre, Distribution of Lévy constants for quadratic numbers, *Acta Arith*. 61 (1992) 13–34; MR1153919 (93c:11057).
[107] E. P. Golubeva, An estimate for the Lévy constant for $\sqrt{p}$, and a class one number criterion for $\mathbb{Q}(\sqrt{p})$ (in Russian), Anal. Teor. Chisel i Teor. Funkts. 14, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* 237 (1997) 21–30, 227; Engl. transl. in *J. Math. Sci.* 95 (1999) 2185–2191; MR1691280 (2000d:11129.
[108] E. P. Golubeva, The spectrum of Lévy constants for quadratic irrationalities (in Russian), Anal. Teor. Chisel i Teor. Funkts. 16, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* 263 (2000) 20–33, 237; Engl. transl. in *J. Math. Sci.* 110 (2002) 3040–3047; MR1756334 (2001b:11065).
[109] E. P. Golubeva, On the spectra of Lévy constants for quadratic irrationalities and class numbers of real quadratic fields (in Russian), Anal. Teor. Chisel i Teor. Funkts. 17, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov (POMI)* 276 (2001) 20–40, 349; Engl. transl. in *J. Math. Sci.* 118 (2003) 4740–4752; MR1850361 (2002k:11199).
[110] S. R. Finch, Freiman's constant, *first volume*, pp. 199–203.
[111] F. E. P. Hirzebruch, Hilbert modular surfaces, *Enseign. Math.* 19 (1973) 183–281; MR0393045 (52 #13856).
[112] P. Chowla and S. Chowla, On Hirzebruch sums and a theorem of Schinzel, *Acta Arith.* 24 (1973) 223–224; also in *Proc. Nat. Acad. Sci. U.S.A.* 69 (1972) 3745; MR0332717 (48 #11043) and MR0319942 (47 #8483).

[113] A. Schinzel, On two conjectures of P. Chowla and S. Chowla concerning continued fractions, *Annali Mat. Pura Appl.* 98 (1974) 111–117; MR0340187 (49 #4943).

[114] H. W. Lu, Hirzebruch sums and Hecke operators, *J. Number Theory* 38 (1991) 185–195; MR1111370 (92f:11153).

[115] K. Inkeri, Über den Euklidischen Algorithmus in quadratischen Zahlkörpern, *Annales Acad. Sci. Fennicae. Ser. A. I. Math.-Phys.* (1947) n. 41; MR0025498 (10,15g).

[116] H. Chatland and H. Davenport, Euclid's algorithm in real quadratic fields, *Canadian J. Math.* 2 (1950) 289–296; MR0041885 (13,15h).

[117] E. S. Barnes and H. P. F. Swinnerton-Dyer, The inhomogeneous minima of binary quadratic forms. I, *Acta Math.* 87 (1952) 259–323; MR0053162 (14,730a).

[118] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003174 and A003246.

[119] D. W. Dubois and A. Steger, A note on division algorithms in imaginary quadratric number fields, *Canad. J. Math.* 10 (1958) 285–286; MR0094325 (20 #844).

[120] H. W. Lenstra, On Artin's conjecture and Euclid's algorithm in global fields, *Invent. Math*. 42 (1977) 201–224; MR0480413 (58 #576).

[121] D. A. Clark, A quadratic field which is Euclidean but not norm-Euclidean, *Manuscripta Math.* 83 (1994) 327–330; MR1277533 (95f:11086).

[122] G. Niklasch, On the verification of Clark's example of a Euclidean but not norm-Euclidean number field, *Manuscripta Math.* 83 (1994) 443–446; MR1277541 (95f:11087).

[123] D. A. Clark, Non-Galois cubic fields which are Euclidean but not norm-Euclidean, *Math. Comp*. 65 (1996) 1675–1679; MR1355007 (97a:11169).

[124] E. Bedocchi, L'anneau $\mathbb{Z}[\sqrt{14}]$ et l'algorithme euclidien, *Manuscripta Math.* 53 (1985) 199–216; MR0807095 (87b:11102).

[125] E. Bedocchi, On the second minimum of a quadratic form and its applications (in Italian), *Riv. Mat. Univ. Parma* 15 (1989) 175–190; MR1064256 (91i:11074).

[126] M. Y. Zhang, On Yokoi's conjecture, *Math. Comp.* 64 (1995) 1675–1685; MR1308464 (95m:11117).

[127] F. Lu, A lower bound for the exceptional field in Yokoi's conjecture (in Chinese), *J. China Univ. Sci. Tech.* 26 (1996) 385–391; MR1434628 (97m:11136).

[128] J. Beck, Diophantine approximation and quadratic fields, *Number Theory*, Proc. 1996 Eger conf., ed. K. Györy, A. Pethö and V. T. Sos, de Gruyter, 1998, pp. 55–93; MR1628833 (99k:11108).

[129] A. Biró, Yokoi's conjecture, *Acta Arith.* 106 (2003) 85–104; MR1956977 (2003k:11162).

[130] A. Biró, Chowla's conjecture, *Acta Arith*. 107 (2003) 179–194; MR1970822 (2004a:11113).

[131] M. J. Jacobson and H. C. Williams, New quadratic polynomials with high densities of prime values, *Math. Comp.* 72 (2003) 499–519; MR1933834 (2003k:11146).

[132] P. C. Sarnak, Class numbers of indefinite binary quadratic forms. II, *J. Number Theory* 21 (1985) 333–346; MR0814010 (87h:11027).

[133] O. M. Fomenko, Class numbers of indefinite binary quadratic forms (in Russian), Anal. Teor. Chisel i Teor. Funkts. 17, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 276 (2001) 312–333, 354–355; Engl. transl. in *J. Math. Sci.* 118 (2003) 4918–4932; MR1850375 (2002d:11039).

1.13 Quadratic Dirichlet L-Series

Let $D=1$ or D be a fundamental discriminant [1]. The **Kronecker–Jacobi–Legendre symbol** (D/n) is a completely multiplicative function on the positive integers:

$$\left(\frac{D}{n}\right) = \begin{cases} \prod_{j=1}^{k} \left(\frac{D}{p_j}\right)^{e_j} & \text{if } n \geq 2, \\ 1 & \text{if } n = 1 \end{cases}$$

where $n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ is the prime factorization of n,

$$\left(\frac{D}{p}\right) = \begin{cases} 1 & \text{if } p \nmid D \text{ and } x^2 \equiv D \bmod p \text{ is solvable}, \\ -1 & \text{if } p \nmid D \text{ and } x^2 \equiv D \bmod p \text{ is not solvable}, \\ 0 & \text{if } p \mid D \end{cases}$$

assuming p is an odd prime, and

$$\left(\frac{D}{2}\right) = \begin{cases} 1 & \text{if } D \equiv 1, 7 \bmod 8, \\ -1 & \text{if } D \equiv 3, 5 \bmod 8, \\ 0 & \text{if } 2 \mid D. \end{cases}$$

The function $n \mapsto (D/n)$ is a real primitive Dirichlet character with modulus $|D|$. In particular, $(1/n) = 1$ always,

$$(-3/n)|_{n=1,2,3} = \{1, -1, 0\},$$

$$(-4/n)|_{n=1,2,3,4} = \{1, 0, -1, 0\},$$

$$(-7/n)|_{n=1,\ldots,7} = \{1, 1, -1, 1, -1, -1, 0\},$$

$$(-8/n)|_{n=1,\ldots,8} = \{1, 0, 1, 0, -1, 0, -1, 0\},$$

$$(5/n)|_{n=1,\ldots,5} = \{1, -1, -1, 1, 0\},$$

$$(8/n)|_{n=1,\ldots,8} = \{1, 0, -1, 0, -1, 0, 1, 0\},$$

$$(12/n)|_{n=1,\ldots,12} = \{1, 0, 0, 0, -1, 0, -1, 0, 0, 0, 1, 0\}.$$

Now define the **Dirichlet L-series associated to** (D/n):

$$L_D(z) = \sum_{n=1}^{\infty} \left(\frac{D}{n}\right) n^{-z}, \quad \mathrm{Re}(z) > 1$$

which can also be written as an infinite product over primes:

$$L_D(z) = \prod_{p} \left(1 - \left(\frac{D}{p}\right) p^{-z}\right)^{-1}, \quad \mathrm{Re}(z) > 1.$$

If $D=1$, then $L_1(z)=\zeta(z)$, which can be analytically continued over the whole complex plane except for a simple pole at $z=1$. For all other D, $L_D(z)$ can be made into an entire function with special values

$$L_D(1)=\begin{cases} \dfrac{\pi}{3\sqrt{3}} & \text{if } D=-3 \\ \dfrac{\pi}{4} & \text{if } D=-4 \\ \dfrac{\pi h(D)}{\sqrt{-D}} & \text{if } D<-4 \\ \dfrac{2h(D)\ln(\varepsilon)}{\sqrt{D}} & \text{if } D>1, \end{cases} \quad \text{(Dirichlet class number formula)}$$

where $h(D)$ is the ideal class number in the wide sense of the quadratic field $\mathbb{Q}(\sqrt{D})$, and ε is the fundamental unit of the integer subring $\mathbb{Z}+((D+\sqrt{D})/2)\mathbb{Z}$. It follows that

$$L_{-7}(1)=\frac{\pi}{\sqrt{7}}, \quad L_{-8}(1)=\frac{\pi}{2\sqrt{2}},$$

$$L_5(1)=\frac{2}{\sqrt{5}}\ln\left(\frac{1+\sqrt{5}}{2}\right), \quad L_8(1)=\frac{\ln\left(1+\sqrt{2}\right)}{\sqrt{2}}, \quad L_{12}(1)=\frac{\ln\left(2+\sqrt{3}\right)}{\sqrt{3}}.$$

The fact that $L_D(1)\neq 0$ leads to a proof of Dirichlet's theorem on arithmetic progressions r, $q+r$, $2q+r$, $\ldots$: There are infinitely many primes congruent to r modulo q if q, r are coprime [2].

A modification of an L-series $L_D(z)$, defined by [3]

$$L_D^*(z)=\begin{cases} (-D)^{z/2}\pi^{-z/2}\Gamma\left(\dfrac{z+1}{2}\right)L_D(z) & \text{if } D<0, \\ D^{z/2}\pi^{-z/2}\Gamma\left(\dfrac{z}{2}\right)L_D(z) & \text{if } D>0, \end{cases}$$

leads to the elegant functional equation $L_D^*(z)=L_D^*(1-z)$.

We turn attention to the points $z=2$, $z=3$ and $z=1/2$. If $D>0$, closed-form expressions for $L_D(2)$ are known:

$$L_1(2)=\frac{\pi^2}{6}, \quad L_5(2)=\frac{4\pi^2}{25\sqrt{5}},$$

$$L_8(2)=\frac{\pi^2}{8\sqrt{2}}, \quad L_{12}(2)=\frac{\pi^2}{6\sqrt{3}},$$

but if $D<0$, only numerical approximations apply:

$$L_{-3}(2)=0.7813024128\ldots \quad ([4]),$$

$$L_{-4}(2)=G=0.9159655941\ldots \quad \text{(Catalan's constant [5])},$$

$$L_{-7}(2)=1.1519254705\ldots, \quad L_{-8}(2)=1.0647341710\ldots.$$

There is an unproven conjecture that [6, 7]

$$L_{-7}(2) = \frac{24}{7\sqrt{7}} \int\limits_{\pi/3}^{\pi/2} \ln \left| \frac{\tan(t) + \sqrt{7}}{\tan(t) - \sqrt{7}} \right| dt$$

which has its origins in hyperbolic geometry and the Claussen function [8]. If $D < 0$, closed-form expressions for $L_D(3)$ are known:

$$L_{-3}(3) = \frac{4\pi^3}{81\sqrt{3}}, \quad L_{-4}(3) = \frac{\pi^3}{32},$$

$$L_{-7}(3) = \frac{32\pi^3}{343\sqrt{7}}, \quad L_{-8}(3) = \frac{3\pi^3}{64\sqrt{2}},$$

but if $D > 0$, only numerical approximations apply:

$$L_1(3) = \zeta(3) = 1.2020569031... \quad \text{(Apéry's constant [9])},$$

$$L_5(3) = 0.8548247666..., \quad L_8(3) = 0.9583804545...,$$

$$L_{12}(3) = 0.9900400194....$$

By way of contrast, virtually nothing is known about $L_D(1/2)$ (regardless of the sign of D):

$$L_1(1/2) = -1.4603545088... \text{ ([9, 10])},$$

$$L_{-3}(1/2) = 0.4808675576..., \quad L_{-4}(1/2) = 0.6676914571... \text{ ([10])}$$

$$L_{-7}(1/2) = 1.1465856669..., \quad L_{-8}(1/2) = 1.1004214095...,$$

$$L_5(1/2) = 0.2317509475......, \quad L_8(1/2) = 0.3736917129...,$$

$$L_{12}(1/2) = 0.4985570024....$$

It is expected that $L_D(1/2) \neq 0$ always [11]; the Generalized Riemann Hypothesis (GRH) states that all zeroes of $L_D(z)$ in the strip $0 \leq \mathrm{Re}(z) \leq 1$ must lie on the central line $\mathrm{Re}(z) = 1/2$. A deeper conjecture, known as the Grand Simplicity Hypothesis [12], asserts that the nonnegative imaginary parts of all such zeroes, taken as D varies across $1 \cup \{$fundamental discriminants$\}$, form a linearly independent set over $\mathbb{Q}$.

1.13.1 *Various Moments*

A discussion of the first and second moments of $L_D(1)$, over all fundamental discriminants $-x < D < 0$ and $0 < D < x$, appears in [1]. We will focus on $L_D(1/2)$ here. Many of the numerical results are due to Conrey, Farmer, Keating, Rubinstein & Snaith [13, 14].

Jutila [15, 16] proved that

$$\sum_{0<-D<x} L_D(1/2) \sim \frac{3}{\pi^2}\left(a_{1,1}\ln(x)+a_{1,0}^{-}\right)x$$
$$\sim (0.1070623764...)x\ln(x)+(0.0806503246...)x,$$

$$\sum_{0<D<x} L_D(1/2) \sim \frac{3}{\pi^2}\left(a_{1,1}\ln(x)+a_{1,0}^{+}\right)x$$
$$\sim (0.1070623764...)x\ln(x)-(0.2556960505...)x$$

as $x\to\infty$, where

$$P_1(s)=\prod_p\left(1-\frac{1}{(p+1)p^s}\right),$$

$$a_{1,1}=P_1(1)/2=(0.7044422009...)/2=0.3522211004...,$$

$$\begin{aligned} a_{1,0}^{-} &= \frac{P_1(1)}{2}\left(-1-\ln(\pi)+4\gamma+\frac{\Gamma'(3/4)}{\Gamma(3/4)}+4\frac{P_1'(1)}{P_1(1)}\right)=0.2653289331...\\ &= 0.6175500336...-a_{1,1}=1.2648891165...-(1+\ln(2\pi))a_{1,1}, \end{aligned}$$

$$\begin{aligned} a_{1,0}^{+} &= \frac{P_1(1)}{2}\left(-1-\ln(\pi)+4\gamma+\frac{\Gamma'(1/4)}{\Gamma(1/4)}+4\frac{P_1'(1)}{P_1(1)}\right)=-0.8412062886...\\ &= -0.4889851881...-a_{1,1}=0.1583538947...-(1+\ln(2\pi))a_{1,1}. \end{aligned}$$

The fact that $a_{1,1}>0$ confirms that $L_D(1/2)>0$ for infinitely many $D<0$ and for infinitely many $D>0$. Interestingly, the expression

$$\frac{P_1'(1)}{P_1(1)}=\sum_p\frac{\ln(p)}{p^2+p-1}=0.4187575787...$$

appears in [17] (concerning the Dedekind totient function).

Jutila [15] also proved that [13, 14]

$$\sum_{0<-D<x} L_D(1/2)^2 \sim \frac{3}{\pi^2}\left(a_{2,3}\ln(x)^3+a_{2,2}^{-}\ln(x)^2+a_{2,1}^{-}\ln(x)+a_{2,0}^{-}\right)x$$
$$\sim (0.0037642089...)x\ln(x)^3+(0.0436478230...)x\ln(x)^2$$
$$+(0.0239243562...)x\ln(x)-(0.0664474558...)x,$$

$$\sum_{0<D<x} L_D(1/2)^2 \sim \frac{3}{\pi^2}\left(a_{2,3}\ln(x)^3+a_{2,2}^{+}\ln(x)^2+a_{2,1}^{+}\ln(x)+a_{2,0}^{+}\right)x$$
$$\sim (0.0037642089...)x\ln(x)^3+(0.0081709895...)x\ln(x)^2$$
$$-(0.1388692446...)x\ln(x)+(0.4058928120...)x$$

as $x \to \infty$, where

$$P_2 = \prod_p \left(1 - \frac{4p^2 - 3p + 1}{(p+1)p^3}\right) = 0.2972100247,$$

$$a_{2,3} = P_2/24 = 0.0123837510...$$

($a_{2,2}^-, a_{2,2}^+, a_{2,1}^-, a_{2,1}^+$ formulas appear in the Addendum). The work of Soundararajan [11], Diaconu, Goldfeld & Hoffstein [18] and Zhang [19] gives rise to the conjecture [13, 14]:

$$\begin{aligned}\sum_{0<-D<x} L_D(1/2)^3 &\sim \frac{3}{\pi^2}\left(a_{3,6}\ln(x)^6 + \sum_{k=0}^{5} a_{3,k}^- \ln(x)^k\right)x + b^- x^{3/4} \\ &\sim (0.0000046457...)x\ln(x)^6 + (0.0002447286...)x\ln(x)^5 \\ &\quad + (0.0039480538...)x\ln(x)^4 + (0.0174395675...)x\ln(x)^3 \\ &\quad - (0.0110235234...)x\ln(x)^2 - (0.0487615392...)x\ln(x) \\ &\quad + (0.1926975162...)x - (0.07...)x^{3/4},\end{aligned}$$

$$\begin{aligned}\sum_{0<D<x} L_D(1/2)^3 &\sim \frac{3}{\pi^2}\left(a_{3,6}\ln(x)^6 + \sum_{k=0}^{5} a_{3,k}^+ \ln(x)^k\right)x + b^+ x^{3/4} \\ &\sim (0.0000046457...)x\ln(x)^6 + (0.0001571591...)x\ln(x)^5 \\ &\quad + (0.0007916339...)x\ln(x)^4 - (0.0094598480...)x\ln(x)^3 \\ &\quad + (0.0136781642...)x\ln(x)^2 + (0.1643132466...)x\ln(x) \\ &\quad - (0.5385378337...)x - (0.14...)x^{3/4}\end{aligned}$$

as $x \to \infty$, where

$$P_3 = \prod_p \left(1 - \frac{12p^5 - 23p^4 + 23p^3 - 15p^2 + 6p - 1}{(p+1)p^6}\right) = 0.0440172316...,$$

$$a_{3,6} = P_3/2880 = 0.0000152837...$$

($a_{3,5}^-$, $a_{3,5}^+$, $a_{3,4}^-$, $a_{3,4}^+$ formulas appear in the Addendum). The exceptional term $x^{3/4}$ has no analog in the first and second moment cases. It is believed that [19]

$$\begin{aligned}b^- + b^+ &= \frac{223\sqrt{2} - 253}{192}\left(\frac{\Gamma(1/8)^4}{\Gamma(3/8)^4} + \frac{\Gamma(1/8)\Gamma(5/8)^3}{\Gamma(3/8)\Gamma(7/8)^3}\right)\pi Q \\ &= \frac{4}{3}(-0.1615725999...) = -0.2154301332...\end{aligned}$$

where [20]

$$\begin{aligned} Q &= \left(\tfrac{\sqrt{2}-1}{\sqrt{2}}\right)^3 \zeta\left(\tfrac{1}{2}\right)^7 \\ &\quad \times \prod_{p>2}\left(1-\tfrac{14}{p^{3/2}}-\tfrac{1}{p^2}+\tfrac{78}{p^{5/2}}-\tfrac{84}{p^3}-\tfrac{58}{p^{7/2}}+\tfrac{154}{p^4}-\tfrac{70}{p^{9/2}}-\tfrac{49}{p^5}+\tfrac{64}{p^{11/2}}-\tfrac{22}{p^6}+\tfrac{1}{p^7}\right) \\ &= -0.0019314869... \end{aligned}$$

(might separate expressions for b^+ and b^- be possible?). For arbitrary $n \geq 1$, Conrey & Farmer [21] conjectured that

$$\sum_{|D|<x} L_D(1/2)^n \sim \frac{6}{\pi^2} a_{n,N}\, x \ln(x)^N$$

as $x \to \infty$, where $N = n(n+1)/2$ and

$$a_{n,N} = \prod_{j=1}^{n} \frac{j!}{(2j)!} \cdot \prod_{p} \frac{\left(1-\frac{1}{p}\right)^N}{1+\frac{1}{p}} \left\{ \frac{1}{2}\left(\left(1-\tfrac{1}{\sqrt{p}}\right)^{-n} + \left(1+\tfrac{1}{\sqrt{p}}\right)^{-n}\right) + \frac{1}{p} \right\}.$$

This is based, in part, on research in random matrix theory by Keating & Snaith [22, 23].

1.13.2 *Dedekind Zeta Function*

Given a fundamental discriminant D, define the **Dedekind zeta function** of $\mathbb{Q}(\sqrt{D})$ to be

$$\begin{aligned} \zeta_D(z) &= \zeta(z) \cdot L_D(z) \\ &= \prod_{\left(\frac{D}{p}\right)=1} (1-p^{-z})^{-2} \cdot \prod_{\left(\frac{D}{p}\right)=-1} (1-p^{-2z})^{-1} \cdot \prod_{\left(\frac{D}{p}\right)=0} (1-p^{-z})^{-1} \end{aligned}$$

(the latter formula is valid for $\mathrm{Re}(z) > 1$). For example, if $D = -4$ (which corresponds to the ring $\mathcal{O}_{-1}$ of Gaussian integers), we have

$$\zeta_{-4}(z) = \prod_{\substack{p\equiv 1 \\ \bmod 4}} (1-p^{-z})^{-2} \cdot \prod_{\substack{p\equiv 3 \\ \bmod 4}} (1-p^{-2z})^{-1} \cdot (1-2^{-z})^{-1}$$

and if $D = -3$ (which corresponds to the ring $\mathcal{O}_{-3}$ of Eisenstein–Jacobi integers), we have

$$\zeta_{-3}(z) = \prod_{\substack{p\equiv 1 \\ \bmod 3}} (1-p^{-z})^{-2} \cdot \prod_{\substack{p\equiv 2 \\ \bmod 3}} (1-p^{-2z})^{-1} \cdot (1-3^{-z})^{-1}.$$

Since $\zeta_D(z)$ has a simple pole at $z = 1$, its Laurent expansion at $z = 1$ is

$$\zeta_D(z) = c_{-1}(z-1)^{-1} + c_0 + c_1(z-1) + c_2(z-1)^2 + \cdots, \quad c_{-1} \neq 0.$$

Define the **Euler–Kronecker constant** of $\mathbb{Q}(\sqrt{D})$ to be

$$\gamma_D = \frac{c_0}{c_{-1}} = \gamma + \frac{L'_D(1)}{L_D(1)},$$

which generalizes Euler's constant $\gamma = 0.5772156649...$ [24]. (In the case $D = 1$, we merely have $\zeta_1(z) = \zeta(z)$ and thus $c_{-1} = 1$, $c_0 = \gamma$.) It follows that [25–28]

$$\gamma_{-4} = \ln\left(2\pi e^{2\gamma}\frac{\Gamma(\frac{3}{4})^2}{\Gamma(\frac{1}{4})^2}\right) = 0.8228252496... \quad \text{(Sierpinski's constant [29])},$$

$$\gamma_{-3} = \ln\left(2\pi e^{2\gamma}\frac{\Gamma(\frac{2}{3})^3}{\Gamma(\frac{1}{3})^3}\right) = 0.9454972808... \quad \text{([30])};$$

alternatively, by the Kronecker limit formula [31],

$$\gamma_{-4} = \frac{\pi}{3} - \ln(4) + 2\gamma - 4\sum_{k=1}^{\infty}\ln\left(1 - e^{-2\pi k}\right) = \tfrac{1}{2}(1.1870859072...)\ln(4),$$

$$\gamma_{-3} = \frac{\pi}{2\sqrt{3}} - \ln(3) + 2\gamma - 4\sum_{k=1}^{\infty}\ln\left|1 - e^{-2\pi i\omega k}\right| = \tfrac{1}{2}(1.7212574274...)\ln(3)$$

where $\omega = -(1 + i\sqrt{3})/2$ and i is the imaginary unit. Further, we have

$$\gamma_{-7} = \ln\left(2\pi e^{2\gamma}\frac{\Gamma(\frac{3}{7})\Gamma(\frac{5}{7})\Gamma(\frac{6}{7})}{\Gamma(\frac{1}{7})\Gamma(\frac{2}{7})\Gamma(\frac{4}{7})}\right) = 0.5928513548... = \tfrac{1}{2}(0.6093306571...)\ln(7),$$

$$\gamma_{-8} = \ln\left(2\pi e^{2\gamma}\frac{\Gamma(\frac{5}{8})\Gamma(\frac{7}{8})}{\Gamma(\frac{1}{8})\Gamma(\frac{3}{8})}\right) = 0.5565042591... = \tfrac{1}{2}(0.5352439565...)\ln(8).$$

In the event that $D > 0$, the only known formulas are [28, 32, 33]

$$\begin{aligned}\gamma_5 &= \ln(2\pi e^{2\gamma}) + \frac{R(\frac{1}{5}) - R(\frac{2}{5}) - R(\frac{3}{5}) + R(\frac{4}{5})}{2\ln\left(\frac{1+\sqrt{5}}{2}\right)} = 1.4048951416... \\ &= \tfrac{1}{2}(1.7458208617...)\ln(5),\end{aligned}$$

$$\begin{aligned}\gamma_8 &= \ln(2\pi e^{2\gamma}) + \frac{R(\frac{1}{8}) - R(\frac{3}{8}) - R(\frac{5}{8}) + R(\frac{7}{8})}{2\ln(1+\sqrt{2})} = 1.2093306309... \\ &= \tfrac{1}{2}(1.1631302027...)\ln(8),\end{aligned}$$

$$\begin{aligned}\gamma_{12} &= \ln(2\pi e^{2\gamma}) + \frac{R(\frac{1}{12}) - R(\frac{5}{12}) - R(\frac{7}{12}) + R(\frac{11}{12})}{2\ln(2+\sqrt{3})} = 1.0539656082... \\ &= \tfrac{1}{2}(0.8482939255...)\ln(12),\end{aligned}$$

where

$$R(x) = -\frac{\partial^2}{\partial z^2}\zeta(z,x)\bigg|_{z=0}$$

and $\zeta(z,x)$ is the second derivative of the **Hurwitz zeta function**, defined when $0 < x \le 1$ by $\zeta(z,x) = \sum_{n=0}^{\infty}(n+x)^{-z}$ for $\mathrm{Re}(z) > 1$ and by analytic continuation elsewhere. Of course, $\zeta(z,1) = \zeta(z)$, $\zeta(z,1/2) = (2^z - 1)\zeta(z)$, $\zeta'(0) = -\ln(2\pi)/2$ and

$$R(1) = -\zeta''(0) = -\tilde{\gamma} - \frac{1}{2}\gamma^2 + \frac{1}{24}\pi^2 + \frac{1}{2}\ln(2\pi)^2 = 2.0063564559...,$$

$$\lim_{x\to 0^+} R(x) = -\infty, \quad R\left(\frac{1}{2}\right) = \ln(2)\ln(2\pi) + \frac{1}{2}\ln(2)^2 = 1.5141458137...,$$

where $\tilde{\gamma} = -0.0728158454...$ is the first Stieltjes constant [34], but little else is known about special values of $R(x)$.

For small $|D|$, γ_D is positive. The first $D < 0$ for which γ_D is negative is $D = -47$, and the first $D > 0$ for which γ_D is negative is $D = 337$. It can be shown that $\lim_{|D|\to\infty} \gamma_D / \ln\sqrt{|D|} = 0$. For arbitrary number fields (finite algebraic extensions of $\mathbb{Q}$), a corresponding limit superior is also 0, assuming the truth of GRH. The corresponding limit inferior, however, appears to lie between -0.26049 and -0.17849, and its exact value is open [31, 35]. We wonder if similar optimization problems can be studied involving higher-order coefficients c_j in the Laurent expansion of $\zeta_D(z)$.

1.13.3 *Prime Products*

Formulas such as [36, 37]

$$\prod_p \frac{p^2+1}{p^2-1} = \frac{5}{2}, \quad \prod_p \frac{p^3+1}{p^3-1} = \frac{945\zeta(3)^2}{\pi^6},$$

$$\prod_{p\equiv 1 \bmod 4} \frac{p^2+1}{p^2-1} = \frac{12G}{\pi^2}, \quad \prod_{p\equiv 1 \bmod 4} \frac{p^3+1}{p^3-1} = \frac{105\zeta(3)}{4\pi^3},$$
$$\prod_{p\equiv 3 \bmod 4} \frac{p^2+1}{p^2-1} = \frac{\pi^2}{8G}, \quad \prod_{p\equiv 3 \bmod 4} \frac{p^3+1}{p^3-1} = \frac{28\zeta(3)}{\pi^3}$$

offer hope that prime products $\prod_{p\equiv k \bmod l}(p^m+1)/(p^m-1)$ might always be expressed via L-series values, where $m \ge 2$. Indeed, we have

$$\prod_{p\equiv 1 \bmod 3} \frac{p^2+1}{p^2-1} = \frac{27L_{-3}(2)}{2\pi^2}, \quad \prod_{p\equiv 1 \bmod 3} \frac{p^3+1}{p^3-1} = \frac{15\sqrt{3}\zeta(3)}{\pi^3},$$
$$\prod_{p\equiv 2 \bmod 3} \frac{p^2+1}{p^2-1} = \frac{4\pi^2}{27L_{-3}(2)}, \quad \prod_{p\equiv 2 \bmod 3} \frac{p^3+1}{p^3-1} = \frac{39\sqrt{3}\zeta(3)}{2\pi^3}.$$

More complicated examples include

$$\prod_{p\equiv 2 \text{ or } 3 \bmod 5} \frac{p^2+1}{p^2-1} = \sqrt{5}, \qquad \prod_{p\equiv 2 \text{ or } 3 \bmod 5} \frac{p^3+1}{p^3-1} = \frac{124\zeta(3)}{125L_5(3)},$$

$$\prod_{p\equiv 7 \bmod 8} \frac{p^2+1}{p^2-1} = \frac{\pi^2}{\sqrt{64\sqrt{2}G\,L_{-8}(2)}}, \qquad \prod_{p\equiv 7 \bmod 8} \frac{p^3+1}{p^3-1} = \frac{\sqrt{1792\sqrt{2}\zeta(3)L_8(3)}}{\sqrt{3}\pi^3}$$

and we wonder whether products over $p \equiv 2 \bmod 5$, or products over $p \equiv 3 \bmod 5$, dash the hope. Finally, series such as

$$\sum_{n=0}^{\infty} \frac{1}{(3n+1)^2} = \frac{1}{2}\left(\frac{4\pi^2}{27} + L_{-3}(2)\right), \quad \sum_{n=0}^{\infty} \frac{1}{(3n+1)^3} = \frac{2}{81\sqrt{3}}\pi^3 + \frac{13}{27}\zeta(3),$$

$$\sum_{n=0}^{\infty} \frac{1}{(3n+2)^2} = \frac{1}{2}\left(\frac{4\pi^2}{27} - L_{-3}(2)\right), \quad \sum_{n=0}^{\infty} \frac{1}{(3n+2)^3} = -\frac{2}{81\sqrt{3}}\pi^3 + \frac{13}{27}\zeta(3),$$

$$\sum_{n=0}^{\infty} \frac{1}{(5n+2)^2} + \sum_{n=0}^{\infty} \frac{1}{(5n+3)^2} = \frac{10-2\sqrt{5}}{125}\pi^2,$$

$$\sum_{n=0}^{\infty} \frac{1}{(5n+2)^3} + \sum_{n=0}^{\infty} \frac{1}{(5n+3)^3} = \frac{62}{125}\zeta(3) - \frac{1}{2}L_5(3),$$

$$\sum_{n=0}^{\infty} \frac{1}{(8n+7)^2} = \frac{1}{4}\left(\frac{1+\sqrt{2}}{8\sqrt{2}}\pi^2 - G - L_{-8}(2)\right),$$

$$\sum_{n=0}^{\infty} \frac{1}{(8n+7)^3} = -\frac{1}{4}\left(\frac{3+2\sqrt{2}}{64\sqrt{2}}\pi^3 - \frac{7}{8}\zeta(3) - L_8(3)\right)$$

raise similar issues.

1.13.4 *Primitive Characters*

Let $\mathbb{Z}_n^*$ denote the group (under multiplication modulo n) of integers relatively prime to n, and let $\mathbb{C}^*$ denote the group (under ordinary multiplication) of nonzero complex numbers. A **Dirichlet character modulo** n is a homomorphism $\chi : \mathbb{Z}_n^* \to \mathbb{C}^*$. It can be shown that $\chi(k)$ is a $\varphi(n)^{\text{th}}$ root of unity for any $k \in \mathbb{Z}_n^*$, where φ is the Euler totient function [38]. In particular, if χ is real-valued, then $\chi(k) = \pm 1$ for any k. We have [39–42]

$$\begin{array}{l}\text{\# complex Dirichlet characters} \\ \text{of modulus } \leq N\end{array} = \sum_{n\leq N} \varphi(n) \sim \frac{3}{\pi^2}N^2,$$

$$\begin{aligned}\begin{array}{l}\text{\# real Dirichlet characters} \\ \text{of modulus } \leq N\end{array} &= \sum_{\substack{n\leq N,\\ n\equiv 2,6 \bmod 8}} 2^{\omega(n)-1} + \sum_{\substack{n\leq N,\\ n\equiv 1,3,4,5,7 \bmod 8}} 2^{\omega(n)} + \sum_{\substack{n\leq N,\\ n\equiv 0 \bmod 8}} 2^{\omega(n)+1} \\ &\sim \frac{6}{\pi^2}N\cdot\ln(N)\end{aligned}$$

as $N \to \infty$, where $\omega(n)$ denotes the number of distinct prime factors of n. The constant $6/\pi^2$ appears in [43] as the probability that two randomly chosen integers are coprime; the above three-fold summation also counts the average number of solutions of $x^2 = 1$ in $\mathbb{Z}_n^*$. Why should a coprimality probability and square roots of unity mod n be at all related to real characters mod n?

Let m be a multiple of n. Extend the domain of χ to $\mathbb{Z}$ via the formula

$$\chi(k) = \begin{cases} \chi(j) & \text{if } \gcd(n,k) = 1 \text{ and } j \equiv k \bmod n,\ 1 \le j \le n \\ 0 & \text{otherwise} \end{cases}$$

and then define a new **induced character** mod m:

$$\hat{\chi}(k) = \begin{cases} \chi(k) & \text{if } \gcd(m,k) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

For example, if χ is the character mod 3 with $\chi(1) = 1$, $\chi(2) = -1$ and $\chi(3) = 0$, note that

$$\chi(k)|_{k=1,\ldots,6} = \{1,-1,0,1,-1,0\} \longmapsto \{1,0,0,0,-1,0\} = \hat{\chi}(k)|_{k=1,\ldots,6}\,.$$

As another example, if χ is the character mod 3 with $\chi(1) = 1$, $\chi(2) = 1$ and $\chi(3) = 0$, note that

$$\chi(k)|_{k=1,\ldots,6} = \{1,1,0,1,1,0\} \longmapsto \{1,0,0,0,1,0\} = \hat{\chi}(k)|_{k=1,\ldots,6}\,.$$

As a third and fourth example, if χ is the character mod 1 with $\chi(1) = 1$, note that

$$\chi(k)|_{k=1,2} = \{1,1\} \longmapsto \{1,0\} = \hat{\chi}(k)|_{k=1,2}\,,$$

$$\chi(k)|_{k=1,2,3,4} = \{1,1,1,1\} \longmapsto \{1,0,1,0\} = \hat{\chi}(k)|_{k=1,2,3,4}\,.$$

These are meant to prepare us for the following definition. A **primitive character mod** m is a character that is not induced by a character mod n for any divisor n of m other than m itself. The first two examples demonstrate that no primitive character mod 6 exists. Likewise, no primitive character mod 2 exists, but the mod 4 character χ with $\chi(1) = 1$, $\chi(2) = 0$, $\chi(3) = -1$ and $\chi(4) = 0$ is primitive.

Define a new multiplicative function

$$\psi(n) = \sum_{d|n} \varphi(d)\mu(n/d)$$

where μ is the Möbius mu function. Also, $\psi(p) = p - 2$ and $\psi(p^l) = p^{l-2}(p-1)^2$ for $l \ge 2$, for any prime p. We have [44, 45]

$$\begin{array}{l}\#\text{ complex primitive Dirichlet} \\ \text{characters of modulus } \le N\end{array} = \sum_{n \le N} \psi(n) \sim \frac{18}{\pi^4} N^2,$$

$$\begin{array}{l}\text{\# real primitive Dirichlet}\\ \text{characters of modulus } \leq N\end{array} = \sum_{|D| \leq N} 1 \sim \frac{6}{\pi^2} N$$

as $N \to \infty$, where D varies across the set $1 \cup \{\text{fundamental discriminants}\}$. For future convenience, the latter sum can be written more explicitly as $\sum_{n \leq N} \delta(n)$, where

$$\delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 1 & \text{if either } n \text{ or } -n \text{ is a fundamental discriminant (not both),} \\ 2 & \text{if } n \text{ and } -n \text{ are fundamental discriminants,} \\ 0 & \text{otherwise.} \end{cases}$$

In fact, $\delta(n)$ is multiplicative with $\delta(2) = 0$, $\delta(4) = 1$, $\delta(8) = 2$, $\delta(2^l) = 0$ for $l > 3$, $\delta(p) = 1$ for prime $p > 2$ and $\delta(p^l) = 0$ for $l > 1$; thus asymptotic techniques in [42] are applicable.

A less stringent version of primitiveness is also available. A **weakly primitive character mod** m is a character that does not coincide (as a function $\mathbb{Z} \to \mathbb{C}^*$) with a character mod n for any divisor n of m other than m itself. For example, $\{1, 0\}$ is weakly primitive as a character mod 2, but not as a character mod 4, since $\{1, 0, 1, 0\}$ is the same as $\{1, 0\}$ concatenated with itself. Both earlier-mentioned characters mod 6 are weakly primitive as well, but not mod 12.

Define another multiplicative function $\xi(n)$ with $\xi(p^l) = \psi(p^l)$ for $l \geq 2$, but $\xi(p) = p - 1$ instead. A Dirichlet convolution-type formula for $\xi(n)$ is also available:

$$\xi(n) = \sum_{d | \kappa'(n)} \psi(n/d)$$

where $\kappa'(n)$ is the product of primes that occur with multiplicity 1 when factoring n (using notation from [46]). We have [47, 48]

$$\begin{array}{l}\text{\# complex, weakly primitive Dirichlet}\\ \text{characters of modulus } \leq N\end{array} = \sum_{n \leq N} \xi(n) \sim \frac{1}{2} \rho N^2$$

where

$$\begin{aligned} \rho &= \prod_p \left(1 - \frac{p^2 + p - 1}{p^4}\right) = \frac{6}{\pi^2} \prod_p \left(1 + \frac{1}{p^3 + p^2 - 1}\right)^{-1} \\ &= \frac{6}{\pi^2} (1.1344121384...)^{-1} = 0.5358961538... \end{aligned}$$

as $N \to \infty$.

Define one last multiplicative function $\eta(n)$ with $\eta(2) = 1$, $\eta(4) = 1$, $\eta(8) = 2$, $\eta(2^l) = 0$ for $l > 3$, $\eta(p) = 2$ for prime $p > 2$ and $\eta(p^l) = 0$ for $l > 1$. A Dirichlet convolution-type formula for $\eta(n)$ is also available:

$$\eta(n) = \sum_{d | \kappa'(n)} \delta(n/d).$$

We have [42]

$$\begin{array}{l}\text{\# real, weakly primitive Dirichlet} \\ \text{characters of modulus } \le N\end{array} = \sum_{n \le N} \eta(n) \sim \sigma N \ln(N)$$

where

$$\sigma = \frac{6}{\pi^2} \prod_p \left(1 - \frac{2}{p(p+1)}\right) = \frac{36}{\pi^4} \prod_p \left(1 - \frac{1}{(p+1)^2}\right) = 0.2867474284...$$

as $N \to \infty$. The constant σ appears in [43] as the probability that three randomly chosen integers are pairwise coprime; it is also unexpectedly connected to the asymptotics of the average number of solutions of $x^3 = 0$ in $\mathbb{Z}_n$. Why should a coprimality probability and cubic roots of nullity mod n be at all related to weakly primitive characters mod n?

Addendum Writing exact expressions for $L_D(1/2)$ moments is difficult. We have, for example [49],

$$a_{2,2}^{\pm} = c^{\pm} - 3a_{2,3}, \quad a_{2,1}^{\pm} = d^{\pm} - 2c^{\pm} + 6a_{2,3}$$

where

$$\begin{aligned} c^- &= \frac{P_2}{4}\left(\frac{1}{2}\frac{\Gamma'(3/4)}{\Gamma(3/4)} + U\right) = 0.1807468351..., \\ c^+ &= \frac{P_2}{4}\left(\frac{1}{2}\frac{\Gamma'(1/4)}{\Gamma(1/4)} + U\right) = 0.0640327313..., \end{aligned}$$

$$\begin{aligned} d^- &= \frac{P_2}{2}\left[\left(\frac{4}{P_2}c^-\right)^2 - V\right] = 0.3658991414..., \\ d^+ &= \frac{P_2}{2}\left[\left(\frac{4}{P_2}c^+\right)^2 - V\right] = -0.4030985462... \end{aligned}$$

and

$$U = -\frac{1}{2}\ln(\pi) + 3\gamma + \sum_p \frac{5p^2 - 6p + 3}{(p-1)(p^3 + 2p^2 - 2p + 1)} \ln(p),$$

$$V = \gamma^2 + 2\tilde{\gamma} + \sum_p \frac{p\left(5p^5 - 5p^4 + 4p^3 + 4p^2 - 5p + 1\right)}{(p-1)^2 \left(p^3 + 2p^2 - 2p + 1\right)^2} \ln(p)^2.$$

To obtain $a_{2,0}^-$ or $a_{2,0}^+$ involves even more complicated formulas. As another example [49],

$$a_{3,5}^{\pm} = c^{\pm} - 6a_{3,6}, \quad a_{3,4}^{\pm} = d^{\pm} - 5c^{\pm} + 30a_{3,6}$$

where

$$c^- = \frac{P_3}{240}\left(\frac{1}{2}\frac{\Gamma'(3/4)}{\Gamma(3/4)} + U\right) = 0.0008968276...,$$
$$c^+ = \frac{P_3}{240}\left(\frac{1}{2}\frac{\Gamma'(1/4)}{\Gamma(1/4)} + U\right) = 0.0006087355...,$$

$$d^- = \frac{P_3}{48}\left[\left(\frac{240}{P_3}c^-\right)^2 - V\right] = 0.0170142017...,$$
$$d^+ = \frac{P_3}{48}\left[\left(\frac{240}{P_3}c^+\right)^2 - V\right] = 0.0051895362...$$

and

$$U = -\frac{1}{2}\ln(\pi) + 4\gamma + \sum_p \frac{4\left(3p^3 - 3p^2 + 3p - 1\right)}{(p-1)\left(p^4 + 4p^3 - 3p^2 + 3p - 1\right)}\ln(p),$$

$$V = \gamma^2 + 2\tilde{\gamma} + \sum_p \frac{p\left(10p^7 + 5p^5 + 17p^4 - 31p^3 + 20p^2 - 6p + 1\right)}{(p-1)^2\left(p^4 + 4p^3 - 3p^2 + 3p - 1\right)^2}\ln(p)^2.$$

Again, $a_{3,k}^-$ or $a_{3,k}^+$ are increasingly complicated for decreasing $k \le 3$.

For arbitrary $n \ge 1$, the rational function in p for the infinite series within U, needed to compute $c^\pm$ and $a_{n,N-1}^\pm$, is [49]

$$g_n(p) = \frac{n+1}{p-1} + \frac{-\left(\sqrt{p}-1\right)^{-n-1} + \left(\sqrt{p}+1\right)^{-n-1}}{\left(\sqrt{p}-1\right)^{-n} + \left(\sqrt{p}+1\right)^{-n} + 2p^{-n/2-1}}.$$

For arbitrary $n \ge 2$, the rational function in p for the infinite series within V, needed to compute $d^\pm$ and $a_{n,N-2}^\pm$, is

$$\frac{p}{(p-1)^2} + g_n(p)^2 - \frac{\left(\sqrt{p}-1\right)^{-n-2} + \left(\sqrt{p}+1\right)^{-n-2}}{\left(\sqrt{p}-1\right)^{-n} + \left(\sqrt{p}+1\right)^{-n} + 2p^{-n/2-1}}.$$

See also [50, 51].

The conjectured expression for $L_{-7}(2)$ is, in fact, a theorem due to Zagier [52]; other representations appear in [53, 54]. More on Euler–Kronecker constants is found in [55, 56]. We recommend Mathar's calculations [57, 58] for further study.

[1] S. R. Finch, Class number theory, *this volume*, §1.12.
[2] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976, pp. 146–156; MR0434929 (55 #7892).
[3] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993, pp. 232, 261; MR1228206 (94i:11105).

[4] S. R. Finch, Addendum, *first volume*, pp. 601–602.
[5] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[6] D. H. Bailey and J. M. Borwein, Experimental mathematics: examples, methods and implications, *Notices Amer. Math. Soc.* 52 (2005) 502–514; MR2140093.
[7] D. H. Bailey, J. M. Borwein, V. Kapoor and E. W. Weisstein, Ten problems in experimental mathematics, *Amer. Math. Monthly* 113 (2006) 481–509; MR2231135 (2007b:65001).
[8] S. R. Finch, Volumes of hyperbolic 3-manifolds, *this volume*, §5.3.
[9] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[10] S. R. Finch, Madelung's constant, *first volume*, pp. 76–81.
[11] K. Soundararajan, Nonvanishing of quadratic Dirichlet L-functions at $s = 1/2$, *Annals of Math.* 152 (2000) 447–488; MR1804529 (2001k:11164).
[12] M. Rubinstein and P. Sarnak, Chebyshev's bias, *Experim. Math.* 3 (1994) 173–197; MR1329368 (96d:11099).
[13] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Integral moments of L-functions, *Proc. London Math. Soc.* 91 (2005) 33–104; math.NT/0206018; MR2149530 (2006j:11120).
[14] M. O. Rubinstein, L-function calculator: a C++ class library and command line interface for computing zeros and values of L-functions.
[15] M. Jutila, On the mean value of $L(1/2, \chi)$ for real characters, *Analysis* 1 (1981) 149–161; MR0632705 (82m:10065).
[16] E. Stankus, The mean value of $L(1/2, \chi_d)$ for $d \equiv l \pmod{D}$ (in Russian), *Litovsk. Mat. Sb.* 23 (1983) 169–177; Engl. transl. in *Lithuanian Math. J.* 23 (1983) 215–221; MR0706019 (84m:10030).
[17] S. R. Finch, Series involving arithmetic functions, *this volume*, §1.20.
[18] A. Diaconu, D. Goldfeld and J. Hoffstein, Multiple Dirichlet series and moments of zeta and L-functions, *Compositio Math.* 139 (2003) 297–360; math.NT/0110092; MR2041614 (2005a:11124).
[19] Q. Zhang, On the cubic moment of quadratic Dirichlet L-functions, *Math. Res. Lett.* 12 (2005) 413–424; MR2150894 (2006h:11102).
[20] P. Sebah, Calculation of Zhang's constant, unpublished note (2005).
[21] J. B. Conrey and D. W. Farmer, Mean values of L-functions and symmetry, *Internat. Math. Res. Notices* (2000) 883–908; math.NT/9912107; MR1784410 (2001i:11111).
[22] J. P. Keating and N. C. Snaith, Random matrix theory and L-functions at $s = 1/2$, *Comm. Math. Phys.* 214 (2000) 91–110; MR1794267 (2002c:11108).
[23] J. P. Keating and N. C. Snaith, Random matrices and L-functions, *J. Phys. A* 36 (2003) 2859–2881; MR1986396 (2004d:11090).
[24] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[25] E. Landau, Über die zu einem algebraischen Zahlkörper gehörige Zetafunktion und die Ausdehnung der Tschebyschefschen Primzahlentheorie aus das Problem der Verteilung der Primideale, *J. Reine Angew. Math.* 125 (1903) 64–188; also in *Collected Works*, v. 1, ed. L. Mirsky, I. J. Schoenberg, W. Schwarz, and H. Wefelscheid, Thales Verlag, 1983, pp. 201–325.
[26] A. Selberg and S. Chowla, On Epstein's zeta-function, *J. Reine Angew. Math.* 227 (1967) 86–110; MR0215797 (35 #6632).
[27] J. Elstrodt, Note on the Selberg trace formula for the Picard group, *Abh. Math. Sem. Univ. Hamburg* 55 (1985) 207–209; MR0831529 (87g:11062).
[28] A. Feuerverger and G. Martin, Biases in the Shanks-Rényi prime number race, *Experim. Math.* 9 (2000) 535–570; MR1806291 (2002d:11111).
[29] S. R. Finch, Sierpinski's constant, *first volume*, pp. 122–125.

[30] S. R. Finch, Gauss' lemniscate constant, *first volume*, pp. 420–423.

[31] Y. Ihara, On the Euler-Kronecker constants of global fields and primes with small norms, *Algebraic Geometry and Number Theory*, ed. V. Ginzburg, Birkhäuser Boston, 2006, pp. 407–451; MR2263195 (2007h:11127).

[32] C. Deninger, On the analogue of the formula of Chowla and Selberg for real quadratic fields, *J. Reine Angew. Math.* 351 (1984) 171–191; MR0749681 (86f:11085).

[33] S. Kanemitsu, On evaluation of certain limits in closed form, *Théorie des nombres*, Proc. 1987 Québec conf., ed. J.-M. De Koninck and C. Levesque, de Gruyter, 1989, pp. 459–474; MR1024583 (90m:11127).

[34] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.

[35] M. Tsfasman, Asymptotic behaviour of the Euler-Kronecker constant, *Algebraic Geometry and Number Theory*, ed. V. Ginzburg, Birkhäuser Boston, 2006, pp. 453–458; math.NT/0503347; MR2263196 (2007h:11129).

[36] R. K. Guy, Products taken over primes, *Unsolved Problems in Number Theory*, 2nd ed., Springer-Verlag, 1994, sect. B48; MR1299330 (96e:11002).

[37] A. Walter, Dirichlet L-functions and prime products, unpublished note (2005).

[38] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.

[39] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000010 and A060594.

[40] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976, pp. 55–62, 138–139; MR0434929 (55 #7892).

[41] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Univ. Press, 1995, p. 257; MR1342300 (97e:11005b).

[42] S. Finch and P. Sebah, Squares and cubes modulo n, math.NT/0604465.

[43] S. R. Finch, Hafner-Sarnak-McCurley constant: Carefree couples, *first volume*, pp. 110–112.

[44] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007431 and A114643.

[45] H. Jager, Some remarks on primitive residue class characters, *Nieuw Arch. Wisk.* 21 (1973) 44–47; MR0323683 (48 #2039).

[46] S. R. Finch, Unitarism and infinitarism, *this volume*, §1.7.

[47] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A055231, A114810 and A114811.

[48] H. Jager, On the number of Dirichlet characters with modulus not exceeding x, *Nederl. Akad. Wetensch. Proc. Ser. A* 76 (1973) 452–455; *Indag. Math.* 35 (1973) 452–455; MR0327693 (48 #6035).

[49] I. P. Goulden, D. K. Huynh, Rishikesh and M. O. Rubinstein, Lower order terms for the moments of symplectic and orthogonal families of L-functions, *J. Number Theory* 133 (2013) 639–674; arXiv:1203.4647; MR2994379.

[50] M. W. Alderson and M. O. Rubinstein, Conjectures and experiments concerning the moments of $L(1/2, \chi_d)$, *Experim. Math.* 21 (2012) 307–328; arXiv:1110.0253; MR2988582.

[51] M. P. Young, The third moment of quadratic Dirichlet L-functions, *Selecta Math.* 19 (2013) 509–543; arXiv:1203.4457; MR3090236.

[52] D. H. Bailey, J. M. Borwein, D. Broadhurst and W. Zudilin, Experimental mathematics and mathematical physics, *Gems in Experimental Mathematics*, Proc. 2009 AMS Special Session on Experimental Mathematics, Washington, DC, ed. T. Amdeberhan, L. A. Medina and V. H. Moll, Amer. Math. Soc., 2010, pp. 41–58; arXiv:1005.0414; MR2731060 (2011m:33049).

[53] M. W. Coffey, Evaluation of a ln tan integral arising in quantum field theory, *J. Math. Phys.* 49 (2008) 093508; arXiv:0801.0272; MR2455847 (2009g:11166).
[54] M. W. Coffey, Alternative evaluation of a ln tan integral arising in quantum field theory, *Math. Phys. Anal. Geom.* 13 (2010) 191–204; arXiv:0810.5077; MR2654127 (2011d:33005).
[55] A. I. Badzyan, The Euler-Kronecker constant (in Russian), *Mat. Zametki* 87 (2010) 35–47; Engl. transl. in *Math. Notes* 87 (2010) 31–42; MR2730381.
[56] K. Ford, F. Luca and P. Moree, Values of the Euler φ-function not divisible by a given odd prime, and the distribution of Euler-Kronecker constants for cyclotomic fields, *Math. Comp.* 83 (2014) 1447–1476; arXiv:1108.3805; MR3167466.
[57] R. J. Mathar, Table of Dirichlet L-series and prime zeta modulo functions for small moduli, arXiv:1008.2547.
[58] R. J. Mathar, Survey of Dirichlet series of multiplicative arithmetic functions, arXiv:1106.4038.

1.14 Elliptic Curves over $\mathbb{Q}$

All polynomials and rational functions in this essay are assumed to have coefficients in $\mathbb{Q}$. Fix an integer $n \geq 1$. An **affine variety** is a simultaneous irreducible system of polynomial equations in n variables. The **$\mathbb{Q}$-points**, **$\mathbb{R}$-points** and **$\mathbb{C}$-points** of the affine variety are all solutions of the polynomial system in $\mathbb{Q}^n$, $\mathbb{R}^n$ and $\mathbb{C}^n$, respectively.

Rational projective n-space $\widetilde{\mathbb{Q}^n}$ is the set of lines through the origin in $\mathbb{Q}^{n+1}$. For example, the projective plane $\widetilde{\mathbb{Q}^2}$ is a quotient of the unit sphere in $\mathbb{Q}^3$ modulo the relation $(X, Y, Z) \sim (-X, -Y, -Z)$. We define $\widetilde{\mathbb{R}^n}$ and $\widetilde{\mathbb{C}^n}$ similarly. A **projective variety** is a simultaneous irreducible system of homogeneous polynomial equations in $n+1$ variables. The $\mathbb{Q}$-points, $\mathbb{R}$-points and $\mathbb{C}$-points of the projective variety are all solutions of the polynomial system in $\widetilde{\mathbb{Q}^n}$, $\widetilde{\mathbb{R}^n}$ and $\widetilde{\mathbb{C}^n}$, respectively; these are $(n+1)$-tuples, not n-tuples as before.

A **curve** is a projective variety corresponding to one homogeneous polynomial equation $p(X, Y, Z) = 0$. In particular, $n+1 = 3$; that is, $n = 2$. Such a curve is **smooth** or **non-singular** if there is no $\mathbb{C}$-point at which the partial derivatives p_X, p_Y, p_Z all vanish. For example, the conic

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

is expressed in homogeneous coordinates as

$$aX^2 + bXY + cY^2 + dXZ + eYZ + fZ^2 = 0;$$

irreducibility implies smoothness in this case. Triples (X, Y, Z) satisfying this equation with $Z = 0$ are called **points at infinity**.

Let F denote a smooth projective curve and $F(\mathbb{C})$ denote the $\mathbb{C}$-points of F. The **genus** g of F is defined topologically as the number of handles in the Riemann surface $F(\mathbb{C})$, and algebraically as $(m-1)(m-2)/2$, where m is the degree of the polynomial p. Lines and conics have genus 0 while smooth cubics have genus 1.

Any two smooth projective curves of genus 0 with a rational point must be **isomorphic** or **birationally equivalent** [1]. This means that the bijection between the curves, as well as its inverse, can be given locally by rational functions. For example, the circle $x^2+y^2=1$ is isomorphic to the hyperbola $x^2-y^2=1$ via the change of coordinates $(x,y)\mapsto(1/x,y/x)$. It is isomorphic to the line $y=0$ via the function $(x,y)\mapsto y/(x+1)$. The circle, moreover, is a commutative group under addition-of-angles, with identity element $(x,y)=(1,0)$. Its group of rational points is the direct sum of $\mathbb{Z}_4$, the cyclic group of order 4, and countably many copies of $\mathbb{Z}$ [2–6].

By contrast, there are (up to isomorphism) infinitely many smooth projective curves of genus 1 with a rational point. These are called **elliptic curves** (not to be confused with *ellipses*). Each such isomorphism class possesses a unique **Weierstrass minimal model** [7]

$$y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6,\quad a_j\in\mathbb{Z},\quad a_1,a_3\in\{0,1\},\quad a_2\in\{0,\pm1\},$$

for which $|\Delta|$ is minimized, where

$$\begin{aligned}\Delta = &-\left(a_1^2+4a_2\right)^2\left(a_1^2a_6-a_1a_3a_4+4a_2a_6+a_2a_3^2-a_4^2\right)\\ &-8\left(a_1a_3+2a_4\right)^3-27\left(a_3^2+4a_6\right)^2+9\left(a_1^2+4a_2\right)\left(a_1a_3+2a_4\right)\left(a_3^2+4a_6\right)\end{aligned}$$

is the **discriminant** of the cubic. For example, the Fermat cubic $x^3+y^3=1$ is isomorphic to the elliptic curve $y^2=x^3-432$ via the change of coordinates $(x,y)\mapsto((36-y)/(6x),(36+y)/(6x))$. Its minimal model is $y^2+y=x^3-7$, however, obtained via the additional transformation $(x,y)\mapsto(4x,8y+4)$. We will soon present a table of isomorphism classes, ordered according to increasing conductor N, along with several other associated constants.

An elliptic curve E is also a commutative group, with addition given by the familiar chord-and-tangent law, and with identity element the unique point at infinity $(X,Y,Z)=(0,1,0)$. It is a prototypal example of what is known as an *abelian variety*. Let $E(\mathbb{Q})$ denote the group of rational points of E. By Mordell's theorem,

$$E(\mathbb{Q})\approx\mathbb{Z}^r\oplus E_{\text{tors}}(\mathbb{Q})$$

where the **rank** r is a nonnegative integer and the **torsion subgroup** $E_{\text{tors}}(\mathbb{Q})$ is finite. Define t to be the order of $E_{\text{tors}}(\mathbb{Q})$, for convenience's sake. The overlap of geometry (E is a smooth curve) and algebra (E is an abelian variety) makes this subject rich and interesting.

1.14.1 *Naive Height*

If $x\in\mathbb{Q}$, write $x=a/b$, where a and b are coprime integers. Define $H(x)=\max\{|a|,|b|\}$. The set of $x\in\mathbb{Q}$ for which $H(x)\le k$ is clearly finite and [1, 8–11]

$$\lim_{k\to\infty}\frac{1}{k^2}\sum_{H(x)\le k} 1 = \frac{12}{\pi^2} = \frac{2}{\zeta(2)}.$$

Alternatively, if $x \in \widetilde{\mathbb{Q}}$, then x is represented by (a, b) in homogeneous coordinates and the same asymptotic result applies. The projective line is identical to the affine line in this regard.

Given a rational point (x, y) on the circle $x^2 + y^2 = 1$, define $H(x, y)$ to be simply $H(x)$. The set of such rational points for which $H(x, y) \le k$ is again finite and [12]

$$\lim_{k\to\infty}\frac{1}{k}\sum_{\substack{H(x)\le k,\\ x^2+y^2=1}} 1 = \frac{4}{\pi}.$$

Observe that it makes no sense to sum over all points $(x, y) \in \mathbb{Q}^2$ of bounded height. However, on the projective plane $\widetilde{\mathbb{Q}^2}$, choose an integer triple (a, b, c) representing a given rational point (x, y), where $\gcd(a, b, c) = 1$. Defining $H(x, y) = \max\{|a|, |b|, |c|\}$, we obtain [1, 9–11]

$$\lim_{k\to\infty}\frac{1}{k^3}\sum_{H(x,y)\le k} 1 = \frac{4}{\zeta(3)}$$

where $\zeta(3)$ is Apéry's constant [13]. For example, $c > 0$ may be taken to be the least common denominator of x and y, and thus $a = cx$ and $b = cy$.

Given a rational point (x, y, z) on the sphere $x^2 + y^2 + z^2 = 1$, define $H(x, y, z)$ to be the least common denominator of x, y and z. In this case, it is known that [14]

$$\lim_{k\to\infty}\frac{1}{k^2}\sum_{\substack{H(x,y,z)\le k,\\ x^2+y^2+z^2=1}} 1 = \frac{3}{2G}$$

where G is Catalan's constant [15]. An open frontier of asymptotic results like these, for higher-dimensional varieties and assorted height functions, awaits discovery.

Let us return to the affine plane. Consider an elliptic curve E and define the **naive height** $H(x, y) = H(x)$ for any rational point (x, y) on E (ignoring the vertical component, just as we did for the circle). The set of rational points for which $H(x, y) \le k$ can be proved to be finite and [1, 16–20]

$$\Theta = \lim_{k\to\infty}\frac{1}{\ln(k)^{r/2}}\sum_{\substack{H(x)\le k,\\ (x,y)\in E}} 1 = \frac{\pi^{r/2}}{\Gamma\left(1+\frac{r}{2}\right)}\frac{t}{\sqrt{R}},$$

where the integers $r \ge 0$ and $t \ge 1$ were defined previously and the real number $R > 0$ is the **regulator** of E. We will demonstrate how to compute R shortly (§1.14.2).

1.14.2 *Canonical Height*

Let $h(x,y) = \ln(H(x,y))$, the logarithm of the naive height on E. We also need the duplication formula, that is, the algorithm by which to calculate $2\cdot(x,y) = (x,y)+(x,y)$:

$$2\cdot(x,y) = (\nu, -(\lambda+a_1)\nu - \mu - a_3)$$

where

$$\lambda = \frac{3x^2+2a_2x+a_4-a_1y}{2y+a_1x+a_3}, \quad \mu = \frac{-x^3+a_4x+2a_6-a_3y}{2y+a_1x+a_3}$$

and $\nu = \lambda^2 + a_1\lambda - a_2 - 2x$. In fact, $y = \lambda x + \mu$ is the line L tangent to E at (x,y) and ν is the horizontal component of the other point of $L \cap E$. Clearly $2^n \cdot P = 2\cdot[2^{n-1}\cdot P]$ for all positive integers n, for any rational point P on E. Define the **canonical height** or **Néron–Tate height** of P to be [7, 21–24]

$$\hat{h}(P) = \frac{1}{2}\lim_{n\to\infty}\frac{h(2^n\cdot P)}{2^{2n}}.$$

For example, given the elliptic curve $y^2+y = x^3 - x$ and the point $P = (0,0)$, we have

$$2\cdot P = (1,0)\,, \quad 4\cdot P = (2,-3)\,, \quad 8\cdot P = \left(\tfrac{21}{25}, -\tfrac{69}{125}\right), \quad 16\cdot P = \left(\tfrac{480106}{4225}, \tfrac{332513754}{274625}\right)$$

and $\hat{h}(P) = 0.0255557041....$ It can be shown that $\hat{h}$ is a nonnegative definite quadratic form on $E(\mathbb{Q})$ that differs from $h/2$ by at most a constant. In particular, the **height pairing**

$$\langle\,,\,\rangle : E(\mathbb{Q}) \times E(\mathbb{Q}) \to \mathbb{R}$$

$$\langle P_i, P_j\rangle = \hat{h}(P_i+P_j) - \hat{h}(P_i) - \hat{h}(P_j)$$

is a symmetric bilinear form. The $r\times r$ determinant

$$R = \det\left(\langle P_i, P_j\rangle\right)_{\substack{1\le i\le r,\\ 1\le j\le r}}$$

is independent of the choice of basis $\{P_1, P_2, \ldots, P_r\}$ for $E(\mathbb{Q})/E_{\mathrm{tors}}(\mathbb{Q})$, and this defines the regulator. Continuing our example, we have $r = 1$ and $R = 2\hat{h}(P) = 0.0511114082....$ Since $t = 1$, it follows that the asymptotic growth constant $\Theta = 8.8464916552....$

Different variations on $\hat{h}$ and $\langle\,,\,\rangle$ abound, all involving factors of 2. Our conventions are consistent with the number-theoretic freeware PARI/GP [25–27].

Numerical algorithms exist for computing $\hat{h}$ to arbitrary precision [28–33]. Here is a curious approach, based on what is called an **elliptic divisibility sequence** [34–36]:

$$s_{2n+1} = s_{n+2}s_n^3 - s_{n-1}s_{n+1}^3, \quad s_{2n} = s_n\left(s_{n+2}s_{n-1}^2 - s_{n-2}s_{n+1}^2\right)$$

Table 1.8 *Regulator R, for four selected isomorphism classes of elliptic curves ($r > 0$)*

N	elliptic curve	r	t	P_i	R	Θ
37	$y^2 + y = x^3 - x$	1	1	$(0,0)$	0.0511114082...	8.8464916552...
43	$y^2 + y = x^3 + x^2$	1	1	$(0,0)$	0.0628165070...	7.9798201588...
389	$y^2 + y = x^3 + x^2 - 2x$	2	1	$(0,0), (1,0)$	0.1524601779...	8.0458449949...
5077	$y^2 + y = x^3 - 7x + 6$	3	1	$(0,2), (1,0), (2,0)$	0.4171435587...	6.4855354622...

with initial terms $s_0 = 0, s_1 = 1, s_2 = 1, s_3 = -1, s_4 = 1$. It can be proved that $s_n \mid s_m$ whenever $n \mid m$, that

$$s_{m-n}s_{m+n} = s_{m+1}s_{m-1}s_n^2 - s_{n+1}s_{n-1}s_m^2$$

for all $m \geq n \geq 0$, and that $\lim_{n\to\infty} n^{-2} \ln |s_n| = 0.0255557041....$ This is the same value $\hat{h}(P)$ obtained in our example.

Another example is the elliptic curve $y^2 + y = x^3 + x^2$; we compute $\hat{h}(P) = 0.0314082535...$ for the point $P = (0,0)$. These two cases constitute the two "simplest" rank-one elliptic curves. Table 1.8 summarizes these, as well as the "simplest" rank-two and rank-three elliptic curves [37, 38]. "Simplicity" means smallest possible conductor N; we will define this quantity later (§1.14.5).

1.14.3 *Real Period*

The complex torus $E(\mathbb{C})$ is isomorphic (as a Riemann surface) to $\mathbb{C}/\Lambda$, where Λ is a certain lattice $\omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ such that $\omega_1 > 0$ and $\mathrm{Im}(\omega_2) > 0$. Clearly the minimal model for E can be rewritten as

$$(2y + a_1x + a_3)^2 = 4x^3 + \left(a_1^2 + 4a_2\right)x^2 + 2\left(a_1a_3 + 2a_4\right)x + \left(a_3^2 + 4a_6\right);$$

let us denote the right-hand side of this equation by $f(x)$. Define the zeroes of $f(x)$ to be e_1, e_2, e_3 with the understanding that $e_1 < e_2 < e_3$ if $\Delta > 0$ and $e_1 \in \mathbb{R}$ uniquely if $\Delta < 0$. These two cases correspond to $E(\mathbb{R})$ being disconnected or connected, respectively ($\Delta \neq 0$ since otherwise E would be singular). Also define the **arithmetic-geometric mean** $M(u, v)$ of two numbers u, v to be the common limit as $n \to \infty$ of the sequences $\{u_n\}$, $\{v_n\}$, where

$$u_n = \frac{u_{n-1} + v_{n-1}}{2}, \quad v_n = \sqrt{u_{n-1}v_{n-1}}, \quad u_0 = u, \quad v_0 = v.$$

It follows that, if $\Delta > 0$,

$$\omega_1 = \int_{e_1}^{e_2} \frac{2\,dx}{\sqrt{f(x)}} = \int_{e_3}^{\infty} \frac{2\,dx}{\sqrt{f(x)}} = \frac{\pi}{M\left(\sqrt{e_3 - e_1}, \sqrt{e_3 - e_2}\right)},$$

Table 1.9 *Real period Ω, for seven selected isomorphism classes of elliptic curves*

N	elliptic curve	Δ	Ω	V
11	$y^2+y=x^3-x^2-10x-20$	-161051	1.2692093042...	1.8515436234...
14	$y^2+xy+y=x^3+4x-6$	-21952	1.9813419560...	2.6262514055...
15	$y^2+xy+y=x^3+x^2-10x-10$	50625	2.8012060846...	2.2357017126...
37	$y^2+y=x^3-x$	37	5.9869172924...	7.3381327407...
43	$y^2+y=x^3+x^2$	-43	5.4686895299...	7.4548214176...
389	$y^2+y=x^3+x^2-2x$	389	4.9804251217...	4.9100459911...
5077	$y^2+y=x^3-7x+6$	5077	4.1516879830...	3.0733872268...

$$\omega_2=\int_{-\infty}^{e_1}\frac{2\,dx}{\sqrt{f(x)}}=\int_{e_2}^{e_3}\frac{2\,dx}{\sqrt{f(x)}}=\frac{\pi i}{M\left(\sqrt{e_3-e_1},\sqrt{e_2-e_1}\right)}$$

and, if $\Delta<0$,

$$\omega_1=\int_{e_1}^{\infty}\frac{2\,dx}{\sqrt{f(x)}}=\frac{2\pi}{M\left(2\sqrt{\eta},\sqrt{2\eta+\xi}\right)},\quad \omega_2=-\frac{1}{2}\omega_1+\frac{\pi i}{M\left(2\sqrt{\eta},\sqrt{2\eta-\xi}\right)}$$

where [18, 39]

$$\xi=3e_3+\tfrac{1}{4}\left(a_1^2+4a_2\right),\quad \eta=\sqrt{3e_3^2+\tfrac{1}{2}\left(a_1^2+4a_2\right)e_3+\tfrac{1}{2}\left(a_1a_3+2a_4\right)}.$$

A path integral expression for ω_2 in the latter case also exists; the AGM sequences converge quadratically and are vastly preferred over numerical integration.

The **real period** Ω is $2\omega_1$ when $\Delta>0$ and ω_1 when $\Delta<0$, and the **real volume** V is $\omega_1\,\mathrm{Im}(\omega_2)$. Observe that ω_1 is the smallest positive real number contained in Λ and V is the area of the associated fundamental parallelogram. A related quantity is the **Faltings height** of E, defined to be the reciprocal of V.

Table 1.9 contains Ω and V for the elliptic curves given in Table 1.8, preceded by several rank-zero elliptic curves not mentioned earlier. In fact, there are three isomorphism classes of elliptic curves with conductor $N=11$, six classes with $N=14$ and eight classes with $N=15$. No examples with $N<11$ exist [40]. We use the notation of Cremona [37] to refer to certain elliptic curves. For instance, $11A1$ refers to the first curve in Table 1.9, while $11A2$ refers to

$$y^2+y=x^3-x^2-7820x-263580$$

with $\Delta=-11$, $\Omega_{11A2}=(1/5)\Omega_{11A1}$ and $V_{11A2}=(1/5)\,V_{11A1}$, and $11A3$ refers to

$$y^2+y=x^3-x^2$$

with $\Delta = -11, \Omega_{11A3} = 5\Omega_{11A1}$ and $V_{11A3} = 5V_{11A1}$. More generally, elliptic curves possessing the same conductor < 26 have the same real period and real volume, up to rational multiples.

Familiar numbers among the real periods include the lemniscate constants [41–44]

$$\Omega_{32A1} = \frac{1}{2\sqrt{2\pi}}\Gamma\left(\frac{1}{4}\right)^2 = 2.6220575542... = \frac{1}{2}(5.2441151085...),$$

$$\Omega_{432A1} = \frac{1}{4\pi}\Gamma\left(\frac{1}{3}\right)^3 = 1.5299540370... = \frac{1}{2}(3.0599080741...)$$

corresponding to the elliptic curves $y^2 = x^3 + 4x$ and $y^2 = x^3 - 16$, respectively. Note that $y^2 = x^3 - x$ and $y^2 = x^3 + x$ are related to the former:

$$\Omega_{32A2} = 2\Omega_{32A1}, \quad \Omega_{64A4} = \sqrt{2}\Omega_{32A1}$$

while $y^2 = x^3 - 1$ and $y^2 = x^3 + 1$ are related to the latter:

$$\Omega_{144A1} = \tfrac{4}{\sqrt[3]{16}}\Omega_{432A1}, \quad \Omega_{36A1} = \tfrac{4\sqrt{3}}{\sqrt[3]{16}}\Omega_{432A1}.$$

Such exact expressions in terms of gamma function values seem to be rare. General formulas for ω_1 and ω_2 in terms of hypergeometric function values are found in [45]. It would be good to better understand $\omega_1 = 2.9934586462...$ and $\omega_2 = (2.4513893819...)i$ for the special curve $37A1$, in particular [18, 46].

1.14.4 *Isogenies*

Let E and E' be two elliptic curves and denote the point at infinity by $\mathcal{O}$. Any isomorphism $E \to E'$ that maps $\mathcal{O}$ to itself induces an isomorphism $E(\mathbb{Q}) \to E'(\mathbb{Q})$ of groups. It is natural to attempt to classify all elliptic curves up to isomorphism; recall, for example, the three isomorphism classes $11A1$, $11A2$, $11A3$ with conductor 11. A weaker notion is as follows. Any homomorphism $E(\mathbb{Q}) \to E'(\mathbb{Q})$ that is not identically $\mathcal{O}$ is called an **isogeny**. It can be proved, in fact, that any isogeny is necessarily surjective. For example, an isogeny from $11A3$ to $11A1$ is given by [18, 47]

$$(x, y) \mapsto \left(x + \frac{1}{x^2} + \frac{2}{x-1} + \frac{1}{(x-1)^2}, y - (2y+1)\left(\frac{1}{x^3} + \frac{1}{(x-1)^3} + \frac{1}{(x-1)^2}\right)\right),$$

which clearly fails to be injective. We remarked earlier that every isomorphism class is represented uniquely by a minimal model; an algorithm for computing such representative curves is due to Tate [48, 49]. Isogeny classes encompass one or more isomorphism classes. The curves $11A1$, $11A2$, $11A3$ all fall in one isogeny class, which is written simply as $11A$. It can be proved that isogenic curves E and E' possess the same conductor N and the same L-series (see §1.14.5). The first

N for which two isogeny classes exist is 26; these are denoted $26A$ and $26B$. The first N for which three isogeny classes exist is 57; these are denoted $57A$, $57B$ and $57C$ [37, 50–52].

1.14.5 *L-Series*

For any prime p, let $\mathbb{Z}_p$ denote the field of integers modulo p. Starting with the minimal model for an elliptic curve E over $\mathbb{Q}$, define E_p to be its **reduction over** $\mathbb{Z}_p$:

$$y^2 + a_1xy + a_3y \equiv x^3 + a_2x^2 + a_4x + a_6 \mod p.$$

Let N_p denote the number of points $(x, y) \in \mathbb{Z}_p^2$ on E_p, plus one, and let [53]

$$b_p = \begin{cases} p + 1 - N_p & \text{if } p \nmid \Delta, \\ & \text{that is, } E \text{ has } \textbf{good reduction} \text{ at } p; \\ \pm 1 & \text{if } p \mid \Delta \text{ and } p \nmid \left(a_1^2 + 4a_2\right)^2 - 24\left(a_1a_3 + 2a_4\right), \\ & \text{that is, } E \text{ has } \textbf{multiplicative reduction} \text{ at } p; \\ 0 & \text{if } p \mid \Delta \text{ and } p \mid \left(a_1^2 + 4a_2\right)^2 - 24\left(a_1a_3 + 2a_4\right), \\ & \text{that is, } E \text{ has } \textbf{additive reduction} \text{ at } p. \end{cases}$$

The three cases correspond to when E_p is non-singular, has a node, or has a cusp, respectively. The last two cases, of course, correspond to when E has **bad reduction** at p. It remains for us to specify the sign of b_p in the nodal case. Does there exist a quadruple $(x_0, y_0, \alpha, \beta) \in \mathbb{Z}_p^4$ for which (x_0, y_0) is a singular point on E_p,

$$\begin{aligned} & y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 \\ \equiv\ & [(y - y_0) - \alpha(x - x_0)]\,[(y - y_0) - \beta(x - x_0)] - (x - x_0)^3 \mod p \end{aligned}$$

and $\alpha \neq \beta$? If yes, the reduction is said to be **split** at p and $b_p = 1$. If no, the reduction is **non-split** and $b_p = -1$. Finally, the **Hasse–Weil L-series of** E is defined to be

$$L_E(z) = \sum_{n=1}^{\infty} b_n n^{-z}, \quad \operatorname{Re}(z) > \tfrac{3}{2}$$

where $b_1 = 1$, $b_{p^k} = b_{p^{k-1}} b_p - p\, b_{p^{k-2}}$ for $k \geq 2$ and $b_{m\,n} = b_m b_n$ for coprime integers m, n. This can also be written as an infinite product:

$$L_E(z) = \prod_{p \mid \Delta} \frac{1}{1 - b_p p^{-z}} \cdot \prod_{p \nmid \Delta} \frac{1}{1 - b_p p^{-z} + p^{1-2z}}, \quad \operatorname{Re}(z) > \tfrac{3}{2}.$$

The combined efforts of Wiles [54], Taylor & Wiles [55] and others [56–59] yield that $L_E(z)$ can be analytically continued over the whole complex plane.

For example, the elliptic curve $11A3$ has bad reduction only at $p = 11$. It has split multiplicative reduction since $11 \nmid 16$ and since $(x_0, y_0, \alpha, \beta) = (-3, 5, 1, -1)$ satisfies the required equation; hence $b_{11} = 1$. As another example, $E = 37A1$ has bad reduction only at $p = 37$. It has non-split multiplicative reduction since

$37 \nmid 48$ and since $(x_0, y_0) = (5, 18)$ is the only singular point of E_{37} but no slopes $(\alpha, \beta) \in \mathbb{Z}_{37}^2$ work with this; hence $b_{37} = -1$. All other coefficients b_p are obtained easily. For the isogeny class $11A$, there is a miraculous q-expansion result [53, 59, 60]:

$$\sum_{n=1}^{\infty} b_n q^n = q \prod_{k=1}^{\infty} \left(1 - q^k\right)^2 \left(1 - q^{11k}\right)^2$$

(a weight 2 cusp form of level 11) and similarly for $14A$ and $15A$. Corresponding generating functions for $37A$ are much more complicated [46, 61].

Let us return to the entire function $L_E(z)$ and define the modification

$$\hat{L}_E(z) = \left(\frac{\sqrt{N}}{2\pi}\right)^z \Gamma(z) L_E(z),$$

where N is the conductor of E. Then the following functional equation

$$\hat{L}_E(z) = \varepsilon \cdot \hat{L}_E(2 - z)$$

is satisfied everywhere, where $\varepsilon = \pm 1$ is the **root number** of E. This equation serves to characterize N uniquely (the actual computation of N turns out to be difficult). The conductor N divides Δ and is divisible only by primes where E has bad reduction. It is conjectured that $\varepsilon = (-1)^r$, where r is the rank of E.

Consider the value of L_E and its derivatives at $z = 1$. Let m denote the smallest integer for which $L_E^{(m)}(1) \neq 0$. The famous **Birch–Swinnerton-Dyer conjecture** predicts that $m = r$ and that

$$\frac{L_E^{(r)}(1)}{r!} \frac{t^2}{\Omega R} \in \mathbb{Z}^+,$$

where t is the torsion order of E, Ω is the real period and R is the regulator (we take $R = 1$ in the event $r = 0$). More can be said if we introduce one additional quantity into the denominator – the **Tamagawa number** c – which cannot be defined here for reasons of space. The new ratio is then conjectured to be an integer square always (see Table 1.10). It is known exactly when $r = 0$ and approximately when $r > 0$ [37]. The first case for which the ratio equals 4 is the elliptic curve $66B3$; the first case for which it equals 9 is $182B3$. Associated with each elliptic curve E is the **Tate–Shafarevich group** $\text{III}(E)$ whose order is at issue. No effective procedure for computing $|\text{III}(E)|$ is known, short of assuming the truth of the BSD conjecture and numerically calculating m, $L_E^{(m)}(1)$, t, Ω, R and c. Gross & Zagier [62] and Kolyvagin [63, 64] proved that if $m = 0$, then $r = 0$; if $m = 1$, then $r = 1$; and that there exists an E with $m = r = 3$. (The curves $389A1$ and $5077A1$ provably satisfy $m = r = 2$ and $m \geq r = 3$, respectively.) We do not yet know an E with $m = r = 4$, or even an E with $r \geq 4$ and $L_E''(1) = 0$ [19, 29, 37, 38, 65–68].

Table 1.10 *BSD ratio, for nine selected isomorphism classes of elliptic curves*

elliptic curve	r	t	$L_E^{(r)}(1)/r!$	c	$\left(L_E^{(r)}(1)/r!\right)\left(t^2/(c\Omega R)\right)$
$11A1$	0	5	0.2538418608...	5	1
$14A1$	0	6	0.3302236593...	6	1
$15A1$	0	8	0.3501507605...	8	1
$37A1$	1	1	0.3059997738...	1	1.0
$43A1$	1	1	0.3435239746...	1	1.0
$66B3$	0	2	1.1021925301... $(=\Omega)$	1	4
$182B3$	0	1	1.9204065875... $(=9\Omega)$	1	9
$389A1$	2	1	0.7593165002...	1	1.0
$5077A1$	3	1	1.7318499001...	1	1.0

1.14.6 *Areas of Rational Right Triangles*

A square-free positive integer d is a **congruent number** if the set

$$\left\{(u,v)\in\mathbb{Q}^2 : \tfrac{1}{2}u\,v=d \text{ and } u^2+v^2=w^2 \text{ for some } w\in\mathbb{Q}\right\}$$

is nonempty [69, 70]. We wish to effectively distinguish congruent d from non-congruent d. Let E_d denote the elliptic curve $y^2=x^3-d^2x$; recall the special case $E_1=32A2$ from §1.14.3. It is known that d is congruent if and only if E_d has nonzero rank. By the (weak) BSD conjecture, the latter condition is equivalent to $L_{E_d}(1)=0$. Another miraculous q-expansion result holds for E_1:

$$\sum_{n=1}^{\infty} b_n q^n = q\prod_{k=1}^{\infty}\left(1-q^{4k}\right)^2\left(1-q^{8k}\right)^2$$

and this carries over to E_d via the **quadratic twist**

$$L_{E_d}(z)=\sum_{n=1}^{\infty}\left(\frac{d}{n}\right)b_n n^{-z}$$

of $L_{E_1}(z)$ by the Dirichlet character (d/n). For instance, $(1/n)=1$ always,

$$(2/n)|_{n=1,2}=\{1,0\},$$

$$(3/n)|_{n=1,2,3}=\{1,-1,0\}$$

and other examples appear in [71].

Define $(i,j)=(1,d)$ if d is odd and $(i,j)=(2,d/2)$ if d is even. In both cases, j is an odd square-free integer and $ij=d$. Define coefficients $c_{i,j}$ via the q-expansions

$$\begin{aligned}\sum_{n=1}^{\infty} c_{1,n}q^n &= q\prod_{k=1}^{\infty}\left(1-q^{8k}\right)\left(1-q^{16k}\right)\cdot\sum_{m=-\infty}^{\infty} q^{2m^2}\\ &= \sum_{\substack{(u,v,w)\in\mathbb{Z}^3,\\ v\equiv 1 \bmod 2}}\left(q^{2u^2+v^2+32w^2}-\tfrac{1}{2}q^{2u^2+v^2+8w^2}\right),\end{aligned}$$

$$\begin{aligned}\sum_{n=1}^{\infty} c_{2,n}q^n &= q\prod_{k=1}^{\infty}\left(1-q^{8k}\right)\left(1-q^{16k}\right)\cdot\sum_{m=-\infty}^{\infty} q^{4m^2}\\ &= \sum_{\substack{(u,v,w)\in\mathbb{Z}^3,\\ v\equiv 1 \bmod 2}}\left(q^{4u^2+v^2+32w^2}-\tfrac{1}{2}q^{4u^2+v^2+8w^2}\right).\end{aligned}$$

Tunnell [72–74] proved the following remarkable formula:

$$L_{E_d}(1)=\frac{1}{8\sqrt{2\pi}}\Gamma\left(\frac{1}{4}\right)^2\cdot c_{i,j}^2\sqrt{\frac{i}{j}}=\frac{1}{4}(2.6220575542...)\cdot c_{i,j}^2\sqrt{\frac{i}{j}}$$

which provides the required identification algorithm (flawed only in that it rests on the validity of an unproved conjecture). On the one hand, since $c_{1,1}=1$, $c_{2,1}=1$ and $c_{1,3}=2$, we have $L_{E_1}(1)=0.6555143885...$,

$$L_{E_2}(1)=\sqrt{2}L_{E_1}(1)=0.9270373386..., \quad L_{E_3}(1)=\tfrac{4}{\sqrt{3}}L_{E_1}(1)=1.5138456348....$$

On the other hand, since $c_{1,5}=c_{2,3}=c_{1,7}=0$, we deduce that $L_{E_5}(1)=L_{E_6}(1)=L_{E_7}(1)=0$. By the BSD conjecture, it can be concluded that 5, 6, 7 are congruent numbers and 1, 2, 3 are not. (These particular facts, however, are obtained via elementary means as well. We are merely illustrating the method.)

Observe that the change of variables $(x,y)\mapsto(x/d,(1/\sqrt{d})(y/d))$ maps E_1 to E_d. It is not an isomorphism over $\mathbb{Q}$ because of the presence of the irrationality $\sqrt{d}$; it is, rather, an isomorphism over $\mathbb{Q}(\sqrt{d})$. Other relevant papers on congruent numbers include [75–81]. A consequence of the BSD conjecture is that any square-free positive integer $\equiv 5,6,7 \bmod 8$ is a congruent number. Further, random matrix theory predicts that [82]

$$\#\{n\leq N: n\equiv 1,2,3 \bmod 8 \text{ is a conguent number}\}\sim C\,N^{3/4}\ln(N)^{11/8}$$

as $N\to\infty$, for some positive constant C.

Let us turn attention away from the curve $32A2$ and instead briefly to $E_1=11A3$. The L-series for E_1 was specified in §1.14.5; the L-series for the quadratic twist E_{-3} corresponds to the curve $99D1$ given by $y^2+y=x^3-3x-5$.

It is known that, for fundamental discriminants δ satisfying $0>\delta\equiv 2,6,7,8,$ $10 \bmod 11$, we have [83–85]

$$L_{E_\delta}(1)=\gamma\cdot c_{-\delta}^2\frac{1}{\sqrt{-\delta}}$$

where

$$\sum_{n=1}^{\infty}c_nq^n=\frac{1}{2}\sum_{\substack{(u,v,w)\in\mathbb{Z}^3,\\ u\equiv v \bmod 2}}q^{u^2+11v^2+11w^2}-\frac{1}{2}\sum_{\substack{(u,v,w)\in\mathbb{Z}^3,\\ u\equiv v \bmod 3,\\ v\equiv w \bmod 2}}q^{\left(u^2+11v^2+33w^2\right)/3}$$

and $\gamma=\sqrt{3}\Omega_{99D1}=2.9176332338....$ An expression for the real period of $99D1$ in terms of gamma function values seems not to be available. This formula for $L_{E_\delta}(1)$ is only the tip of a more general theory due to Shimura [86], Waldspurger [87] and Kohnen & Zagier [88]. See also [89–94].

The curve $E_1=144A1$ (mentioned in §1.14.3) has quadratic twist E_d given by $y^2=x^3-d^3$. We have, for example [95],

$$L_{E_d}(1)=\frac{1}{2\sqrt[3]{16}\pi}\Gamma\left(\frac{1}{3}\right)^3\cdot\frac{c_d^2}{\sqrt{d}}=\frac{2}{\sqrt[3]{16}}(1.5299540370...)\cdot\frac{c_d^2}{\sqrt{d}}$$

where $0<d\equiv 1 \bmod 24$ is square-free and

$$\sum_{n=1}^{\infty}c_nq^n=q\prod_{k=1}^{\infty}\left(1-q^{12k}\right)^2\sum_{m=-\infty}^{\infty}q^{m^2}.$$

The Fermat cubic $F_1=27A1$ (mentioned near the beginning) has quadratic twist F_d given by $y^2=x^3-432d^3$. Similar complicated formulas for $L_{F_d}(1)$ hold, depending again on the sign and modulus of d [96]. We will revisit F_1 shortly.

Here is an exercise that is vaguely similar to the congruent number problem [97]. Define

$$g(n)=\#\left\{(u,v)\in\mathbb{Z}^2: u\,v=n \text{ and } u+v=w^2 \text{ for some } w\in\mathbb{Z}\right\}.$$

It turns out, for square-free $d>0$, that $g(d)$ is a lower bound for 2^{r+2}, where r is the rank of the elliptic curve $y^2=x^3+dx$. No one knows whether $\limsup_{d\to\infty}g(d)=\infty$, which would imply that there exist elliptic curves of arbitrarily large rank. We do know, however, that $\limsup_{n\to\infty}g(n)=\infty$ and more precisely that [98]

$$\begin{aligned}\lim_{N\to\infty}N^{-3/4}\sum_{n=1}^{N}g(n) &= 2\int_0^1\sqrt{x+\frac{1}{x}}dx-\frac{4}{3}\\ &= \frac{4}{3}\left(\sqrt{2}-1\right)+\frac{1}{3\sqrt{\pi}}\Gamma\left(\frac{1}{4}\right)^2=3.0243843195....\end{aligned}$$

1.14.7 *Sums of Two Rational Cubes*

Let d be a cube-free positive integer and F_d denote the elliptic curve $y^2 = x^3 - 432d^2$; recall the special case $F_1 = 27A1$ from earlier. Note that the factor here is d^2 rather than d^3 as before. It is known that $2 < d = u^3 + v^3$ for $(u, v) \in \mathbb{Q}^2$ if and only if F_d has nonzero rank. (Reason: the group $F_d(\mathbb{Q})$ is torsion-free, hence $F_d(\mathbb{Q})$ contains infinitely many points if and only if $F_d(\mathbb{Q})$ contains at least one point [99–101].) By the (weak) BSD conjecture, the latter condition is equivalent to $L_{F_d}(1) = 0$. Yet another miraculous q-expansion result holds for F_1:

$$\sum_{n=1}^{\infty} b_n q^n = q \prod_{k=1}^{\infty} \left(1 - q^{3k}\right)^2 \left(1 - q^{9k}\right)^2,$$

but formulas for $L_{F_d}(z)$ via the **cubic twist** F_d of F_1 are considerably more complicated. We defer these until later [102]. No theorem analogous to Tunnell's is yet known. On the one hand, we have

$$L_{F_1}(1) = \frac{\sqrt{3}}{18\pi}\Gamma\left(\frac{1}{3}\right)^3 = 0.5888795834...,$$

$$L_{F_2}(1) = \tfrac{3}{2^{4/3}} L_{F_1}(1) = 0.7010910526..., \quad L_{F_3}(1) = 3^{2/3} L_{F_1}(1) = 1.2249188952...,$$

$$L_{F_4}(1) = \tfrac{3}{2^{2/3}} L_{F_1}(1) = 1.1129126745..., \quad L_{F_5}(1) = \tfrac{3}{5^{1/3}} L_{F_1}(1) = 1.0331366085....$$

On the other hand, $L_{F_6}(1) = L_{F_7}(1) = L_{F_9}(1) = 0$. By the BSD conjecture, it can be concluded that 6, 7, 9 are sums of two rational cubes and 3, 4, 5 are not. (Again, these facts are elementary – just for illustration – as are $1 = 0^3 + 1^3$ and $2 = 1^3 + 1^3$.)

Observe that the change of variables $(x, y) \mapsto (\sqrt[3]{d}x, \sqrt[3]{d}y)$ maps F_1 to F_d and is an isomorphism over $\mathbb{Q}(\sqrt[3]{d})$. The L-series arising in this case *differ* from the L-series of $x^3 + y^3 = 1$ twisted by cubic Dirichlet characters [103]; hence confusion is possible when surveying the literature. More on $x^3 + y^3 = d$ is found in [104–108]. A consequence of the BSD conjecture is that any square-free positive integer $\equiv 4, 6, 7, 8 \bmod 9$ is a sum of two rational cubes. Further, random matrix theory predicts that [109]

$$\begin{aligned}&\#\{n \le N : n \equiv 1, 2, 3, 5 \bmod 9 \text{ is square-free and is a sum of two rational cubes}\}\\ &\quad \sim C N^{5/6} \ln(N)^{\sqrt{3}/2 - 1/8}\end{aligned}$$

as $N \to \infty$, for some positive constant C. It would be more natural to express these asymptotics for cube-free integers, but apparently the result becomes less tractable.

1.14.8 *Lang's Conjecture*

Let E be an elliptic curve over $\mathbb{Q}$. Recall that the canonical height $\hat{h}: E(\mathbb{Q}) \to \mathbb{R}$ satisfies $\hat{h}(P) = 0$ if and only if P is a torsion point. We wonder whether [110–112]

$$\inf_{E} \inf_{\substack{\text{nontorsion} \\ P \in E(\mathbb{Q})}} \hat{h}(P) > 0.$$

The infimum is certainly small: taking E to be the minimal model $y^2 + xy + y = x^3 + x^2 - 125615x + 61201397$ and P to be the point $(7107, -602054)$, we obtain $\hat{h}(P) < 0.0045$.

Let Δ denote the discriminant of E and let $\mathcal{E}_D$ denote the set of all minimal models E satisfying $|\Delta| \geq D$. Lang [113] predicted that the aforementioned infimum is positive and further conjectured that

$$\inf_{D>0} \inf_{E \in \mathcal{E}_D} \inf_{\substack{\text{nontorsion} \\ P \in E(\mathbb{Q})}} \frac{\hat{h}(P)}{\ln |\Delta|} > 0.$$

Again, the infimum is small: for our earlier example, $\Delta = -1494018600480000000$ and thus $\hat{h}(P)/\ln|\Delta| < 1.07 \times 10^{-4}$. Elkies [114] found a different example with ratio less than 0.85×10^{-4}. Hindry & Silverman [115], however, demonstrated that Lang's conjecture would follow from a proof of the important Masser–Oesterlé *ABC* conjecture [116–118]. Another interesting constant is the value of

$$\lim_{D \to \infty} \inf_{E \in \mathcal{E}_D} \inf_{\substack{\text{nontorsion} \\ P \in E(\mathbb{Q})}} \frac{\hat{h}(P)}{\ln |\Delta|},$$

which may or may not exceed the preceding. Progress in resolving these issues is reported in [30, 31, 35, 110, 115, 119–121].

We conclude with a final glimpse at the height $\hat{h}(P) = 0.0255557041...$ of the point $P = (0, 0)$ on the elliptic curve $E = 37A1$. Consider the lattice $\omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$, where ω_1, ω_2 are given at the end of §1.14.3. Over all nonzero lattice points ω, define the **Weierstrass sigma function**

$$\sigma(z) = z \prod_{\omega \neq 0} \left(1 - \frac{z}{\omega}\right) \exp\left(\frac{z}{\omega} + \frac{z^2}{2\omega^2}\right)$$

as well as constants

$$\kappa = \int_1^{\infty} \frac{dx}{\sqrt{4x^3 - 4x + 1}} = 1.1342732156..., \quad \sigma(\kappa) = 1.1055557990....$$

It would be good someday to prove that $\hat{h}(P)$ is transcendental; one formula for achieving this might be [36, 122–124]

$$\hat{h}(P) = \frac{\kappa^2}{4\omega_1} \frac{\sigma'(\omega_1/2)}{\sigma(\omega_1/2)} - \frac{1}{4} \ln(\sigma(\kappa)).$$

Another helpful formula (a decomposition of $\hat{h}(P)$ into a sum of local heights over all primes p) appears in [31, 125]. No algebraic height $\hat{h}(P)$, for any curve E and nontorsion point P, has ever been found. But a transcendentality proof for even a single case escapes all known efforts.

In closing, we merely mention certain averages [126] without details; p and ℓ denote primes throughout. Concerning the value distribution of L-series coefficients b_p, we have a constant [127–129]

$$\prod_{\ell}\left(1-\frac{1}{(\ell-1)^2(\ell+1)}\right)=0.6151326573....$$

Concerning the growth of primes p such that N_p is prime, we have [130–132]

$$\prod_{\ell}\left(1-\frac{\ell^2-\ell-1}{(\ell-1)^3(\ell+1)}\right)=0.5051661682....$$

Concerning the growth of primes p such that the group E_p (together with a point at infinity) is cyclic, we have [133, 134]

$$\prod_{\ell}\left(1-\frac{1}{\ell(\ell-1)^2(\ell+1)}\right)=0.8137519061....$$

The constant $2C_{\text{twin}}/\pi^2=0.1337767531...$ appears in [135, 136]; recent progress on Lang's conjecture is reported in [137, 138].

[1] M. Hindry and J. H. Silverman, *Diophantine Geometry. An Introduction*, Springer-Verlag, 2000, pp. 12–17, 28–29, 67–81, 160–161, 210–215, 451–456, 464; MR1745599 (2001e:11058).

[2] E. J. Eckert, The group of primitive Pythagorean triangles, *Math. Mag.* 57 (1984) 22–27; MR0729035 (85a:51017).

[3] L. Tan, The group of rational points on the unit circle, *Math. Mag.* 69 (1996) 163–171.

[4] F. Lemmermeyer, Higher descent on Pell conics III. The first 2-descent, arXiv:math/0311310.

[5] F. Lemmermeyer, Conics - a poor man's elliptic curves, arXiv:math/0311306.

[6] B. Poonen, Elliptic curves, *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*, Proc. 2000 MSRI workshop, ed. J. P. Buhler and P. Stevenhagen, Cambridge Univ. Press, 2008, pp. 183–207; MR2467547 (2009k:11095).

[7] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer-Verlag, 1986, pp. 45–65, 146–170, 201–205, 227–240, 360–363; MR0817210 (87g:11070) and MR1329092 (95m:11054).

[8] J. H. Silverman and J. Tate, *Rational Points on Elliptic Curves*, Springer-Verlag, 1992, pp. 102–103; MR1171452 (93g:11003).

[9] S. Lang, *Fundamentals of Diophantine Geometry*, Springer-Verlag, 1983, pp. 70–71; MR0715605 (85j:11005).

[10] S. H. Schanuel, Heights in number fields, *Bull. Soc. Math. France* 107 (1979) 433–449; MR0557080 (81c:12025).

[11] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A018805, A071778.
[12] D. N. Lehmer, Asymptotic evaluation of certain totient sums, *Amer. J. Math.* 22 (1900) 293-335.
[13] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[14] W. Duke, Rational points on the sphere, *Ramanujan J.* 7 (2003) 235–239; MR2035804 (2005h:11094).
[15] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[16] D. Zagier, The Birch-Swinnerton-Dyer conjecture from a naive point of view, *Arithmetic Algebraic Geometry*, Proc. 1989 Texel Island conf., ed. G. van der Geer, F. Oort and J. Steenbrink, Birkhäuser, 1991, pp. 377–389; MR1085269 (92c:11063).
[17] D. Zagier, L-series of elliptic curves, the Birch-Swinnerton-Dyer conjecture, and the class number problem of Gauss, *Notices Amer. Math. Soc*. 31 (1984) 739–743; MR0765835 (86f:11047).
[18] A. W. Knapp, *Elliptic Curves*, Princeton Univ. Press, 1992, pp. 54–56, 106–107, 186–188, 308–309, 348–349, 363–365; MR1193029 (93j:11032).
[19] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer-Verlag, 1990, pp. 297–304, 343–358; MR1070716 (92e:11001).
[20] B. Mazur, Arithmetic on curves, *Bull. Amer. Math. Soc.* 14 (1986) 207–259; MR0828821 (88e:11050).
[21] S. Lang, Les formes bilinéaires de Néron et Tate, *1963/1964 Séminaire Bourbaki*, f. 3, e. 274, Secrétariat Mathématique, 1–11; *Séminaire Bourbaki*, v. 8, Soc. Math. France, 1995, pp. 435–445; MR0176984 (31 #1252) and MR1611559.
[22] A. Néron, Quasi-fonctions et hauteurs sur les variétés abéliennes, *Annals of Math*. 82 (1965) 249–331; MR0179173 (31 #3424).
[23] L. C. Washington, *Elliptic Curves. Number Theory and Cryptography*, Chapman & Hall/CRC, 2003, pp. 206–222; MR1989729 (2004e:11061).
[24] I. Connell, *Elliptic Curve Handbook*, unpublished manuscript (1996).
[25] H. Cohen, K, Belabas, et al., PARI/GP, downloadable software, Université Bordeaux I.
[26] J. Voight, Computing with elliptic curves: implementations, unpublished note (2000).
[27] T. Dokchitser, ComputeL, downloadable PARI script, Univ. of Bristol.
[28] J. Tate, unpublished letter to J.-P. Serre (1979).
[29] J. P. Buhler, B. H. Gross and D. B. Zagier, On the conjecture of Birch and Swinnerton-Dyer for an elliptic curve of rank 3, *Math. Comp.* 44 (1985) 473–481; MR0777279 (86g:11037).
[30] H. M. Tschöpe and H. G. Zimmer, Computation of the Néron-Tate height on elliptic curves, *Math. Comp.* 48 (1987) 351–370; MR0866121 (87m:14025).
[31] J. H. Silverman, Computing heights on elliptic curves, *Math. Comp.* 51 (1988) 339–358; MR0942161 (89d:11049).
[32] H. G. Zimmer, A limit formula for the canonical height of an elliptic curve and its application to height computations, *Number Theory*, Proc. 1988 Canad. Number Theory Assoc. Banff conf., ed. R. A. Mollin, de Gruyter, 1990, pp. 641–659; MR1106690 (93d:11060).
[33] J. H. Silverman, Computing canonical heights with little (or no) factorization, *Math. Comp.* 66 (1997) 787–805; MR1388892 (97f:11040).
[34] G. Everest, Elliptic divisibility sequences, unpublished note (2001).
[35] G. Everest and T. Ward, The canonical height of an algebraic point on an elliptic curve, *New York J. Math.* 6 (2000) 331–342; MR1800354 (2001j:11056).

[36] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006720, A006769, A028941.
[37] J. E. Cremona, *Algorithms for Modular Elliptic Curves*, 2nd ed., Cambridge Univ. Press, 1997; MR1628193 (99e:11068).
[38] W. A. Stein, Elementary Number Theory, lecture notes (2001).
[39] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993, pp. 380–405, 514–516; MR1228206 (94i:11105).
[40] M. K. Agrawal, J. H. Coates, D. C. Hunt and A. J. van der Poorten, Elliptic curves of conductor 11, *Math. Comp*. 35 (1980) 991–1002; MR0572871 (81e:10022).
[41] S. R. Finch, Gauss' lemniscate constant, *first volume*, pp. 420–423.
[42] B. J. Birch and H. P. F. Swinnerton-Dyer, Notes on elliptic curves. II, *J. Reine Angew. Math.* 218 (1965) 79–108; MR0179168 (31 #3419).
[43] N. M. Stephens, The diophantine equation $X^2 + Y^2 = DZ^2$ and the conjectures of Birch and Swinnerton-Dyer, *J. Reine Angew. Math.* 231 (1968) 121–162; MR0229651 (37 #5225).
[44] E. B. Burger and R. Tubbs, *Making Transcendence Transparent. An Intuitive Approach to Classical Transcendental Number Theory*, Springer-Verlag, 2004, pp. 193–221; MR2077395 (2005f:11145).
[45] D. Husemöller, *Elliptic Curves*, Springer-Verlag, 1987, pp. 162–182, 290–314; MR0868861 (88h:11039).
[46] D. Zagier, Modular points, modular curves, modular surfaces and modular forms, *Arbeitstagung Bonn 1984*, ed. F. Hirzebruch, J. Schwermer and S. Suter, Lect. Notes in Math. 1111, Springer-Verlag, 1985; pp. 225–248; MR0797423 (88j:11033).
[47] J. Coates, Elliptic curves – the crossroads of theory and computation, *Proc. 2002 Algorithmic Number Theory Sympos. (ANTS-V)*, Sydney, ed. C. Fieker and D. R. Kohel, Lect. Notes in Comp. Sci. 2369, Springer-Verlag, 2002, pp. 9–19; MR2041070 (2005a:11073).
[48] J. Tate, Algorithm for determining the type of a singular fiber in an elliptic pencil, *Modular Functions of One Variable. IV*, Proc. 1972 Antwerp conf., ed. B. J. Birch and W. Kuyk, Lect. Notes in Math. 476, Springer-Verlag, 1975, pp. 33–52; MR0393039 (52 #13850).
[49] M. Laska, An algorithm for finding a minimal Weierstrass equation for an elliptic curve, *Math. Comp.* 38 (1982) 257–260; MR0637305 (84e:14033).
[50] J. S. Milne, *Elliptic Curves*, BookSurge, 2006, pp. 76–79, 160–168, 193–229; MR2267743 (2007h:14044).
[51] E. Kowalski, Elliptic curves, rank in families and random matrices, *Ranks of Elliptic Curves and Random Matrix Theory*, ed. J. B. Conrey, D. W. Farmer, F. Mezzadri and N. C. Snaith, Cambridge Univ. Press, 2007, pp. 7–52; MR2322336 (2008j:11073).
[52] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005788, A052432, A060564, A110563, A110620.
[53] J. Tate, The arithmetic of elliptic curves, *Invent. Math*. 23 (1974) 179–206; MR0419359 (54 #7380).
[54] A. Wiles, Modular elliptic curves and Fermat's last theorem, *Annals of Math*. 141 (1995) 443–551; MR1333035 (96d:11071).
[55] R. Taylor and A. Wiles, Ring-theoretic properties of certain Hecke algebras, *Annals of Math*. 141 (1995) 553–572; MR1333036 (96d:11072).
[56] F. Diamond, On deformation rings and Hecke rings, *Annals of Math*. 144 (1996) 137–166; MR1405946 (97d:11172).
[57] B. Conrad, F. Diamond and R. Taylor, Modularity of certain potentially Barsotti-Tate Galois representations, *J. Amer. Math. Soc.* 12 (1999) 521–567; MR1639612 (99i:11037).

[58] C. Breuil, B. Conrad, F. Diamond and R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises, *J. Amer. Math. Soc.* 14 (2001) 843–939; MR1839918 (2002d:11058).
[59] H. Darmon, A proof of the full Shimura-Taniyama-Weil conjecture is announced, *Notices Amer. Math. Soc.* 46 (1999) 1397–1401; MR1723249 (2000j:11082).
[60] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000727, A006571, A007653, A030184, A030187, A030188.
[61] N. D. Elkies, Elliptic and modular curves over finite fields and related computational issues, *Computational Perspectives on Number Theory*, Proc. 1995 Chicago conf., ed. D. A. Buell and J. T. Teitelbaum, Amer. Math. Soc., 1998, pp. 21–76; MR1486831 (99a:11078).
[62] B. H. Gross and D. B. Zagier, Heegner points and derivatives of L-series, *Invent. Math*. 84 (1986) 225–320; MR0833192 (87j:11057)
[63] V. A. Kolyvagin, Finiteness of $E(\mathbb{Q})$ and $\text{III}(E,\mathbb{Q})$ for a subclass of Weil curves (in Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* 52 (1988) 522–540, 670–671; Engl. transl. in *Math. USSR-Izv*. 32 (1989) 523–541; MR0954295 (89m:11056).
[64] V. A. Kolyvagin, The Mordell-Weil and Shafarevich-Tate groups for Weil elliptic curves (in Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* 52 (1988) 1154–1180, 1327; Engl. transl. in *Math. USSR-Izv.* 33 (1989) 473–499; MR0984214 (90f:11035).
[65] A. Wiles, The Birch and Swinnerton-Dyer conjecture, Clay Math. Institute Millennium Problems.
[66] H. Darmon, *Rational Points on Modular Elliptic Curves*, Amer. Math. Soc., 2004, pp. 1–11; MR2020572 (2004k:11103).
[67] H. Darmon, Elliptic curves and p-adic uniformization, *CMS Notes*, v. 31 (1999) n. 2, 8–16.
[68] K. Rubin and A. Silverberg, Ranks of elliptic curves, *Bull. Amer. Math. Soc*. 39 (2002) 455–474; MR1920278 (2003f:11080).
[69] R. K. Guy, Congruent numbers, *Unsolved Problems in Number Theory*, 2$^{\text{nd}}$ ed., Springer-Verlag, 1994, sect. D27; MR1299330 (96e:11002).
[70] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003273, A006991, A072068–A072071.
[71] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[72] J. B. Tunnell, A classical Diophantine problem and modular forms of weight $3/2$, *Invent. Math.* 72 (1983) 323–334; MR0700775 (85d:11046).
[73] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, 1984, pp. 1–97, 212–222; MR0766911 (86c:11040).
[74] K. Ono, Tate-Shafarevich groups of the congruent number elliptic curves, *Acta Arith.* 81 (1997) 247–252; MR1460580 (98g:11064).
[75] G. Kramarz, All congruent numbers less than 2000, *Math. Annalen* 273 (1986) 337–340; MR0817886 (87e:11036).
[76] K. Noda and H. Wada, All congruent numbers less than 10000, *Proc. Japan Acad. Ser. A Math. Sci*. 69 (1993) 175–178; MR1232820 (94e:11064).
[77] F. R. Nemenzo, All congruent numbers less than 40000, *Proc. Japan Acad. Ser. A Math. Sci.* 74 (1998) 29–31; MR1617754 (99e:11070).
[78] D. R. Heath-Brown, The size of Selmer groups for the congruent number problem. II, *Invent. Math.* 118 (1994) 331–370; MR1292115 (95h:11064).
[79] H. Wada and M. Taira, Computations of the rank of elliptic curve $y^2 = x^3 - n^2x$, *Proc. Japan Acad. Ser. A Math. Sci.* 70 (1994) 154–157; MR1291172 (95g:11052).
[80] N. F. Rogers, Rank computations for the congruent number elliptic curves, *Experim. Math.* 9 (2000) 591–594; MR1806294 (2001k:11104).

[81] N. D. Elkies, Curves $Dy^2 = x^3 - x$ of odd analytic rank, *Proc. 2002 Algorithmic Number Theory Sympos. (ANTS-V)*, Sydney, ed. C. Fieker and D. R. Kohel, Lect. Notes in Comp. Sci. 2369, Springer-Verlag, 2002, pp. 244–251; arXiv:math/0208056; MR2041088 (2005a:11075).
[82] H. Cohen, *Number Theory*. v. 1, *Tools and Diophantine Equations*, Springer-Verlag, 2007, pp. 393–396, 450–456, 566–568; MR2312337 (2008e:11001).
[83] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Integral moments of L-functions, *Proc. London Math. Soc.* 91 (2005) 33–104; arXiv:math/0206018; MR2149530 (2006j:11120).
[84] A. Dabrowski and M. Wodzicki, Elliptic curves with large analytic order of $\text{III}(E)$, *Algebra, Arithmetic, and Geometry*, v. 1, ed. Y. Tschinkel and Y. Zarhin, Birkhäuser Boston, 2009, pp. 407–421; MR2641177 (2011m:11133).
[85] J. P. Keating, Lectures on Random Matrix Theory and Number Theory - II (2004).
[86] G. Shimura, On modular forms of half integral weight, *Annals of Math.* 97 (1973) 440–481; MR0332663 (48 #10989).
[87] J.-L. Waldspurger, Sur les coefficients de Fourier des formes modulaires de poids demi-entier, *J. Math. Pures Appl.* 60 (1981) 375–484; MR0646366 (83h: 10061).
[88] W. Kohnen and D. Zagier, Values of L-series of modular forms at the center of the critical strip, *Invent. Math.* 64 (1981) 175–198; MR0629468 (83b:10029).
[89] D. Goldfeld and C. Viola, Mean values of L-functions associated to elliptic, Fermat and other curves at the centre of the critical strip, *J. Number Theory* 11 (1979) 305–320; MR0544259 (80j:10046).
[90] J. B. Conrey, J. P. Keating, M. O. Rubinstein and N. C. Snaith, On the frequency of vanishing of quadratic twists of modular L-functions, *Number Theory for the Millennium*, v. 1, Proc. 2000 Urbana conf., ed. M. A. Bennett, B. C. Berndt, N. Boston, H. G. Diamond, A. J. Hildebrand, and W. Philipp, A. K. Peters, 2002, pp. 301–315; arXiv:math/0012043; MR1956231 (2003m:11141).
[91] J. B. Conrey, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Random matrix theory and the Fourier coefficients of half-integral weight forms, *Experim. Math.* 15 (2006) 67–82; arXiv:math/0412083; MR2229387 (2007i:11063).
[92] C. Delaunay, Note on the frequency of vanishing of L-functions of elliptic curves in a family of quadratic twists, *Ranks of Elliptic Curves and Random Matrix Theory*, ed. J. B. Conrey, D. W. Farmer, F. Mezzadri and N. C. Snaith, Cambridge Univ. Press, 2007, pp. 195–200; MR2322345 (2008e:11079).
[93] C. Delaunay, Heuristics on Tate-Shafarevitch groups of elliptic curves defined over $\mathbb{Q}$, *Experim. Math.* 10 (2001) 191–196; MR1837670 (2003a:11065).
[94] K. Ono, *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series*, Amer. Math. Soc., 2004, pp. 149–163; MR2020489 (2005c:11053).
[95] G. Frey, A relation between the value of the L-series of the curve: $y^2 = x^3 - k^3$ in $s = 1$ and its Selmer group, *Arch. Math. (Basel)* 45 (1985) 232–238; MR0807656 (87b:11045).
[96] J. Nekovár, Class numbers of quadratic fields and Shimura's correspondence, *Math. Annalen* 287 (1990) 577–594; MR1066816 (91k:11051).
[97] D. Clark, An arithmetical function associated with the rank of elliptic curves, *Canad. Math. Bull.* 34 (1991) 181–185; MR1113294 (92c:11059).
[98] J.-J. Lee, A remark on the ranks of twists of elliptic curves, *C. R. Math. Acad. Sci. Soc. Royale Canada* 26 (2004) 73–78; MR2083839 (2005f:11113).
[99] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A159843, A020898.
[100] D. Husemöller, *Elliptic Curves*, 2nd ed., Springer-Verlag, 2004, pp. 30–37; MR2024529 (2005a:11078).

[101] D. Zagier, Elliptic modular forms and their applications, *The 1-2-3 of Modular Forms*, Proc. Nordfjordeid, 2004 summer school, ed. K. Ranestad, Springer-Verlag, 2008, pp. 77, 97–99; MR2409678 (2010b:11047).
[102] S. R. Finch, Cubic and quartic characters, *this volume*, §1.33.
[103] C. David, J. Fearnley and H. Kisilevsky, On the vanishing of twisted L-functions of elliptic curves, *Experim. Math.* 13 (2004) 185–198; MR2068892 (2005e:11082).
[104] N. M. Stephens, The diophantine equation $X^3 + Y^3 = DZ^3$ and the conjectures of Birch and Swinnerton-Dyer, *J. Reine Angew. Math.* 231 (1968) 121–162; MR0229651 (37 #5225).
[105] D. Zagier and G. Kramarz, Numerical investigations related to the L-series of certain elliptic curves, *J. Indian Math. Soc*. 52 (1987) 51–69; MR0989230 (90d:11072).
[106] D. Lieman, Automorphic forms and cubic twists of elliptic curves, arXiv:math/9407202, Ph.D thesis, Brown Univ., 1992; MR2688630.
[107] M. Watkins, Rank distribution in a family of cubic twists, *Ranks of Elliptic Curves and Random Matrix Theory*, ed. J. B. Conrey, D. W. Farmer, F. Mezzadri and N. C. Snaith, Cambridge Univ. Press, 2007, pp. 237–246; arXiv:math/0412427; MR2322349 (2008k:11072).
[108] N. D. Elkies and N. F. Rogers, Elliptic curves $x^3 + y^3 = k$ of high rank, *Proc. 2004 Algorithmic Number Theory Sympos. (ANTS-VI)*, Burlington, VT, ed. D. Buell, Lect. Notes in Comp. Sci. 3076, Springer-Verlag, 2004, pp. 244–251; arXiv:math/0403116; MR2137353 (2006c:11064).
[109] H. Cohen, *Number Theory*. v. 1, *Tools and Diophantine Equations*, Springer-Verlag, 2007, pp. 376–385, 478–479, 568–570, 611; MR2312337 (2008e:11001).
[110] G. Everest, A. van der Poorten, I. Shparlinski and T. Ward, *Recurrence Sequences*, Amer. Math. Soc., 2003, pp. 163–175; MR1990179 (2004c:11015).
[111] W. A. Stein, Elliptic curves, the *ABC* conjecture, and points of small canonical height: from a seminar talk by M. Baker, unpublished note (2001).
[112] N. D. Elkies, Nontorsion points of low height on elliptic curves over $\mathbb{Q}$, unpublished note (2002).
[113] S. Lang, *Elliptic Curves: Diophantine Analysis*, Springer-Verlag, 1978, p. 92; MR0518817 (81b:10009).
[114] N. D. Elkies, Points of low canonical height on elliptic curves and surfaces, unpublished note (2001).
[115] M. Hindry and J. H. Silverman, The canonical height and integral points on elliptic curves, *Invent. Math.* 93 (1988) 419–450; MR0948108 (89k:11044).
[116] S. Lang, Old and new conjectured Diophantine inequalities, *Bull. Amer. Math. Soc.* 23 (1990) 37–75; MR1005184 (90k:11032).
[117] D. Goldfeld, Modular forms, elliptic curves and the *ABC*-conjecture, *A Panorama of Number Theory or the View from Baker's Garden*, Proc. 1999 Zürich conf., ed. G. Wüstholz, Cambridge Univ. Press, 2002, pp. 128–147; MR1975449 (2004f:11048).
[118] A. Nitaj, The *ABC* conjecture, unpublished note (2013).
[119] J. H. Silverman, Lower bound for the canonical height on elliptic curves, *Duke Math. J*. 48 (1981) 633–648; MR0630588 (82k:14043).
[120] S. David, Points de petite hauteur sur les courbes elliptiques, *J. Number Theory* 64 (1997) 104–129; MR1450488 (98k:11067).
[121] M. Krir, À propos de la conjecture de Lang sur la minoration de la hauteur de Néron-Tate pour les courbes elliptiques sur $\mathbb{Q}$, *Acta Arith.* 100 (2001) 1–16; MR1864622 (2002i:11062).
[122] A. N. W. Hone, Elliptic curves and quadratic recurrence sequences, *Bull. London Math. Soc*. 37 (2005) 161–171; corrigendum 38 (2006) 741–742; MR2119015 (2005h:11111) and MR2268357 (2007g:11060).

[123] H. W. Braden, V. Z. Enolskii and A. N. W. Hone, Bilinear recurrences and addition formulae for hyperelliptic sigma functions, *J. Nonlinear Math. Phys.* 12 (2005), suppl. 2, 46–62; arXiv:math/0501162; MR2217095 (2007a:14041).
[124] A. N. W. Hone, Algebraic curves, integer sequences and a discrete Painlevé transcendent, arXiv:0807.2538.
[125] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, 1994, pp. 454–480; MR1312368 (96b:11074).
[126] N. Jones, Averages of elliptic curve constants, *Math. Annalen* 345 (2009) 685–710; arXiv:0711.3484; MR2534114 (2010j:11090).
[127] S. Lang and H. Trotter, *Frobenius Distributions in* GL_2*-Extensions: Distribution of Frobenius Automorphisms in* GL_2*-Extensions of the Rational Numbers*, Lect. Notes in Math. 504, Springer-Verlag, 1976; MR0568299 (58 #27900).
[128] C. David and F. Pappalardi, Average Frobenius distributions of elliptic curves, *Internat. Math. Res. Notices* (1999) 165–183; MR1677267 (2000g:11045).
[129] S. R. Finch, Dedekind eta products, *this volume*, §1.19.
[130] N. Koblitz, Primality of the number of points on an elliptic curve over a finite field, *Pacific J. Math.* 131 (1988) 157–165; MR0917870 (89h:11023).
[131] D. Zywina, A refinement of Koblitz's conjecture, *Internat. J. Number Theory* 7 (2011) 739–769; arXiv:0909.5280; MR2805578 (2012e:11107).
[132] A. Balog, A-C. Cojocaru and C. David, Average twin prime conjecture for elliptic curves, *Amer. J. Math.* 133 (2011) 1179–1229; arXiv:0709.1461; MR2843097 (2012j:11118).
[133] J.-P. Serre, Résumé des cours de 1977–1978. *Annuaire du Collège de France 1978*, pp. 67–70; *Oeuvres*, v. III, Springer, 1986, pp. 465–468; MR0926691 (89h:01109c).
[134] W. D. Banks and I. E. Shparlinski, Sato-Tate, cyclicity, and divisibility statistics on average for elliptic curves of small height, *Israel J. Math.* 173 (2009) 253–277; arXiv:math/0609144; MR2570668 (2011a:11121).
[135] Y.-M. J. Chen and J. Yu, On a density problem for elliptic curves over finite fields, *Asian J. Math.* 4 (2000) 737–755; MR1870656 (2003a:11071).
[136] P. Moree, Artin's primitive root conjecture–a survey, *Integers* 12 (2012) 1305–1416; arXiv:math/0412262; MR3011564.
[137] P. Voutier and M. Yabuta, Lang's conjecture and sharp height estimates for the elliptic curves $y^2 = x^3 + ax$, *Internat. J. Number Theory* 9 (2013) 1141–1170; arXiv:1104.4645; MR3077706.
[138] P. Voutier and M. Yabuta, Lang's conjecture and sharp height estimates for the elliptic curves $y^2 = x^3 + b$, *Acta Arith.* 173 (2016) 197–224; arXiv:1305.6560; MR3512852.

1.15 Modular Forms on $\mathrm{SL}_2(\mathbb{Z})$

Let $k \in \mathbb{Z}$ and let $\mathrm{SL}_2(\mathbb{Z})$ denote the special linear group

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z} \text{ and } ad - bc = 1 \right\}.$$

A **modular form of weight** k is an analytic function f defined on the complex upper half plane $\mathbb{H} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ that transforms under the action of $\mathrm{SL}_2(\mathbb{Z})$

according to the relation [1]

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

and whose Fourier series $f(z) = \sum_{n=-\infty}^{\infty} \gamma_n e^{2\pi i n z}$ satisfies $\gamma_n = 0$ for all $n < 0$. In particular, we have

$$f(z+1) = f(z), \quad f(-1/z) = (-z)^k f(z).$$

If, additionally, we have $\gamma_0 = 0$, then f is a **cusp form of weight** k. Every nonconstant modular form has weight $k \geq 4$, where k is even, and every nonzero cusp form has weight $k \geq 12$. The set M_k of modular forms and the set S_k of cusp forms are finite-dimensional vector spaces over $\mathbb{C}$ with [2]

$$\dim(M_k) = \begin{cases} \left\lfloor \frac{k}{12} \right\rfloor & \text{if } k \equiv 2 \bmod 12, \\ \left\lfloor \frac{k}{12} \right\rfloor + 1 & \text{if } k \equiv 0, 4, 6, 8, 10 \bmod 12 \end{cases}$$

and $\dim(S_k) = \dim(M_k) - 1$ if $k \geq 12$. We will focus primarily on a specific basis element of S_{12}, leaving other aspects of this huge research area for later.

The **discriminant function** $\Delta : \mathbb{H} \to \mathbb{C}$, defined via

$$\Delta(z) = q \prod_{n=1}^{\infty} (1 - q^n)^{24} = \sum_{m=1}^{\infty} \tau(m) q^m$$

where $q = e^{2\pi i z}$ and $\tau : \mathbb{Z}^+ \to \mathbb{Z}$ is the **Ramanujan tau function** [3–7], can be proved to be a cusp form of weight 12. Nobody knows whether $\tau(m) \neq 0$ for all $m \geq 1$, but Mordell [8] proved that τ is a multiplicative function and Deligne [9–11] proved that $|\tau(p)| \leq 2p^{11/2}$ for any prime p. This implies that [12]

$$\tau(m) = O\left(m^{11/2+\varepsilon}\right)$$

as $m \to \infty$, for any $\varepsilon > 0$; further [13–17],

$$\liminf_{m\to\infty} m^{-11/2}\tau(m) = -\infty, \quad \limsup_{m\to\infty} m^{-11/2}\tau(m) = \infty.$$

Let the **Hecke L-series** be

$$L_\Delta(z) = \sum_{m=1}^{\infty} \tau(m) m^{-z} = \prod_p \frac{1}{1 - \tau(p)p^{-z} + p^{11-2z}}, \quad \mathrm{Re}(z) > \tfrac{13}{2},$$

and its modification be

$$L_\Delta^*(z) = (2\pi)^{-z}\Gamma(z) L_\Delta(z).$$

Then $L_\Delta(z)$ can be extended to an entire function and the functional equation $L_\Delta^*(z) = L_\Delta^*(12 - z)$ is satisfied everywhere. One can compute $L_\Delta(6) = 0.7921228386...$, for example, but it turns out that more can be said.

Define two constants [18–20]

$$\begin{aligned}\xi &= 30L_{\Delta}^{*}(6) = 0.0463463808...\\ &= 960(0.0000482774...) = 5(0.0092692761...),\end{aligned}$$

$$\begin{aligned}\eta &= 28L_{\Delta}^{*}(5) = 28L_{\Delta}^{*}(7) = 0.0457516089...\\ &= \tfrac{32}{15}(0.0214460667...) = \tfrac{2}{5}(0.1143790224...).\end{aligned}$$

It can be shown that the values of $L_{\Delta}^{*}(n)$ at even $2 \le n \le 10$ are rational multiples of ξ:

$$L_{\Delta}^{*}(4) = L_{\Delta}^{*}(8) = \tfrac{1}{24}\xi, \quad L_{\Delta}^{*}(2) = L_{\Delta}^{*}(10) = \tfrac{2}{25}\xi,$$

and that the values of $L_{\Delta}^{*}(n)$ at odd $1 \le n \le 11$ are rational multiples of η:

$$L_{\Delta}^{*}(3) = L_{\Delta}^{*}(9) = \tfrac{1}{18}\eta, \quad L_{\Delta}^{*}(1) = L_{\Delta}^{*}(11) = \tfrac{90}{691}\eta.$$

These can alternatively be written in terms of $L_{\Delta}(1)$ and $L_{\Delta}(2)$; see Table 1.11. Similar collapsing occurs at integer arguments $< k$ for the unique cusp forms of weight $k = 16$ and $k = 18$ [7]. An integral expression for $L_{\Delta}^{*}(n)$ is [21]

$$\begin{aligned}L_{\Delta}^{*}(n) = {} & \frac{1}{i^{n-1}\pi^{11}} \int_{0}^{1} \left(\int_{v}^{1} \frac{du}{\sqrt{u(u-1)(u-v)}} \right)^{n-1} \left(\int_{1}^{\infty} \frac{du}{\sqrt{u(u-1)(u-v)}} \right)^{11-n}\\ & \times v\,(1-v)\,dv\end{aligned}$$

where $n = 1, 2, ..., 11$ and i is the imaginary unit. The product $\xi\eta = 0.0021204214...$ also appears in the following [18, 19, 22–24]:

$$\begin{aligned}\lim_{x\to\infty} \frac{1}{x^{12}} \sum_{m \le x} \tau(m)^2 &= \frac{2^3\pi^{11}}{3^4 5^2 7\, 11}\xi\eta = 0.0320070045...\\ &= \frac{2^8\pi^{11}}{3^4 5^8 7\, 11}(1.0353620568...) = \frac{1}{12}(0.3840840544...),\end{aligned}$$

which is an interesting asymptotic mean square result. By contrast, we know that [25, 26]

$$\sum_{m \le x} \tau(m) = O\left(x^{35/6+\varepsilon}\right)$$

Table 1.11 *Values of $L_f(1)$, $L_f(2)$; f is the unique cusp form of weight $k = 12, 16, 18$*

k	12	16	18
$L_f(1)$	0.0374412812...	0.5870144080...	−3.5316483054...
$L_f(2)$	0.1463745420...	1.6654560382...	−8.6783515629...

as $x \to \infty$, for any $\varepsilon > 0$, and that [27, 28]

$$\liminf_{x\to\infty} x^{-23/4} \sum_{m\le x} \tau(m) = -\infty, \quad \limsup_{x\to\infty} x^{-23/4} \sum_{m\le x} \tau(m) = \infty,$$

but a more precise estimate of the mean apparently remains open. Moreover (§1.15.2),

$$\sum_{m\le x} |\tau(m)| = o\left(x^{13/2}\right)$$

as $x \to \infty$. See also [29–31].

1.15.1 *Congruence Subgroups*

Given N to be a positive integer, define the following subgroup of the full modular group $\mathrm{SL}_2(\mathbb{Z})$:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : c \equiv 0 \bmod N \right\}$$

and define a **weight** k **modular form of level** N exactly as before, with $\mathrm{SL}_2(\mathbb{Z})$ replaced by $\Gamma_0(N)$. Clearly the preceding discussion applies to the case $N = 1$ and k free; we focus henceforth on the case $k = 2$ and N free. The first nonzero weight 2 cusp form has level 11:

$$f(z) = q \prod_{n=1}^{\infty} (1 - q^n)^2 \left(1 - q^{11n}\right)^2$$

whose Fourier coefficients coincide [32] with those of the L-series for the elliptic curve isogeny class $11A$. The next two cusp forms have level 14 and 15, corresponding to $14A$ and $15A$. On the one hand, not all cusp forms are linked to elliptic curves: the first counterexamples have level 22 and 23. On the other hand, the Taniyama–Shimura conjecture (proved by Wiles, Taylor, Diamond, Conrad & Breuil [33]) asserts that every elliptic curve E is linked to a cusp form with level N equal to the conductor of E.

Let $S_2(N)$ denote the vector space of weight 2 cusp forms of level N. The dimension $\delta_0(N)$ of $S_2(N)$ over $\mathbb{C}$ possesses a more complicated formula than earlier [34–39]:

$$\delta_0(N) = 1 + \frac{\psi(N)}{12} - \frac{\nu_2(N)}{4} - \frac{\nu_3(N)}{3} - \frac{\chi(N)}{2}$$

where

$$\psi(N) = N \prod_{p|N} \left(1 + \frac{1}{p}\right), \quad \chi(N) = \sum_{d|N} \varphi\left(\gcd\left(d, \frac{N}{d}\right)\right),$$

$$\nu_2(N) = \begin{cases} 0 & \text{if } 4|N, \\ \prod_{p|N} \left(1 + \left(\frac{-4}{p}\right)\right) & \text{otherwise;} \end{cases} \quad \nu_3(N) = \begin{cases} 0 & \text{if } 9|N, \\ \prod_{p|N} \left(1 + \left(\frac{-3}{p}\right)\right) & \text{otherwise;} \end{cases}$$

$\varphi(N) = N\prod_{p|N}(1 - 1/p)$ is the Euler totient function [40], and $(-4/p)$, $(-3/p)$ are Kronecker–Jacobi–Legendre symbols [41]. We have asymptotic extreme results [36, 42]

$$\liminf_{N\to\infty} \frac{\delta_0(N)}{N} = \frac{1}{12}, \quad \limsup_{N\to\infty} \frac{\delta_0(N)}{N\ln(\ln(N))} = \frac{e^{\gamma}}{2\pi^2}$$

and average behavior

$$\sum_{N\leq y} \delta_0(N) = \frac{5}{8\pi^2}y^2 + o\left(y^2\right)$$

as $y\to\infty$. Similar dimension estimates can be found for the vector space $M_2(N)$ of weight 2, level N modular forms [43].

Define also the subgroup

$$\Gamma_1(N) = \left\{\begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : a\equiv d\equiv 1 \bmod N \text{ and } c\equiv 0 \bmod N\right\}$$

and the corresponding weight 2 cuspidal vector space dimension $\delta_1(N)$. An analogous formula for $\delta_1(N)$ is known [36, 37, 43], with extreme results

$$\liminf_{N\to\infty} \frac{\delta_1(N)}{N^2} = \frac{1}{4\pi^2} < \frac{1}{24} = \limsup_{N\to\infty} \frac{\delta_1(N)}{N^2}$$

and average behavior

$$\sum_{N\leq y} \delta_1(N) = \frac{1}{72\zeta(3)}y^3 + o\left(y^3\right)$$

as $y\to\infty$. Generalization to arbitrary integer weight k is also possible.

Let $D = 1$ or D be a fundamental discriminant [44]. A **level** N**, weight** k **modular form** $f: \mathbb{H}\to\mathbb{C}$ with **Nebentypus character** $(D/\cdot)$ transforms according to

$$f\left(\frac{az+b}{cz+d}\right) = \left(\frac{D}{d}\right)(cz+d)^k f(z) \qquad \text{for all } \begin{pmatrix} a & b \\ c & d\end{pmatrix} \in \Gamma_0(N).$$

The trivial case $D = 1$ reduces to the earlier definition. For example, we have

$$(-15/d)|_{d=1,2,\ldots,15} = \{1, 1, 0, 1, 0, 0, -1, 1, 0, 0, -1, 0, -1, -1, 0\}.$$

It turns out that the vector space of cusp forms corresponding to $(N, k, D) = (15, 3, -15)$ is two-dimensional, and that a certain basis element is given by [38, 45–47]

$$f(z) = q\prod_{n=1}^{\infty}\left(1 - q^{3n}\right)^3\left(1 - q^{5n}\right)^3 + q^2\prod_{n=1}^{\infty}\left(1 - q^{n}\right)^3\left(1 - q^{15n}\right)^3.$$

This will be useful later (§1.15.3). Also, the vector space of cusp forms corresponding to $(N, k, D) = (6, 4, 1)$ is one-dimensional with basis element

$$g(z) = q \prod_{n=1}^{\infty} (1-q^n)^2 \left(1-q^{2n}\right)^2 \left(1-q^{3n}\right)^2 \left(1-q^{6n}\right)^2,$$

which we likewise will see again.

1.15.2 *Ramanujan Tau Function*

Let us continue where we stopped earlier. It is conjectured that [48–52]

$$\sum_{m \le x} |\tau(m)| \sim A\, x^{13/2} (\ln(x))^{-1+8/(3\pi)}$$

as $x \to \infty$, for some constant $0 < A < \infty$, whereas it is known that [50, 53]

$$\sum_{m \le x} \tau(m)^4 \sim B\, x^{23} \ln(x)$$

for some constant $0 < B < \infty$. Improved numerical estimates of $A \approx 0.0996$ and $B \approx 0.0026$ [54] would be good to see someday. We cannot hope for similar accuracy in estimating $\sum_{m \le x} \tau(m)$ until the correct order of magnitude – conjectured to be $O\left(x^{23/4+\varepsilon}\right)$ – is established. Evidence that 23/4 is the best exponent includes the formula [55–62]

$$\frac{1}{x} \int_1^x \left(\sum_{m \le y} \tau(m) \right)^2 dy \sim C_\tau\, x^{23/2}$$

as $x \to \infty$, where [63, 64]

$$C_\tau = \frac{1}{50\pi^2} \sum_{k=1}^{\infty} \frac{\tau(k)^2}{k^{25/2}} = \frac{1.5882400955...}{50\pi^2}.$$

There are analogous formulas [56, 65–70] for the error terms in the divisor and circle problems [71]:

$$\frac{1}{x} \int_1^x \left(\sum_{m \le y} d(m) - y\ln(y) - (2\gamma - 1)y \right)^2 dy \sim C_d\, x^{1/2},$$

$$\frac{1}{x} \int_1^x \left(\sum_{m \le y} r(m) - \pi y \right)^2 dy \sim C_r\, x^{1/2}$$

where

$$C_d = \frac{1}{6\pi^2} \sum_{k=1}^{\infty} \frac{d(k)^2}{k^{3/2}} = \frac{\zeta(3/2)^4}{6\pi^2 \zeta(3)} = 0.6542839775...,$$

$$C_r = \frac{1}{3\pi^2}\sum_{k=1}^{\infty}\frac{r(k)^2}{k^{3/2}} = \frac{16\zeta(3/2)^2\beta(3/2)^2}{3\left(1+2^{-3/2}\right)\pi^2\zeta(3)} = 1.6939569917...$$

and $\zeta(z) = L_1(z)$, $\beta(z) = L_{-4}(z)$ denote the Riemann zeta and Dirichlet beta functions, respectively [72, 73].

Returning finally to the problem of estimating $\tau(m)$ itself, we ask about the values of constants c_+, c_- for which [17]

$$0 < \limsup_{m\to\infty} m^{-11/2}\exp\left(\frac{-c_+\ln(m)}{\ln(\ln(m))}\right)\tau(m) < \infty,$$

$$-\infty < \liminf_{m\to\infty} m^{-11/2}\exp\left(\frac{-c_-\ln(m)}{\ln(\ln(m))}\right)\tau(m) < 0.$$

Is there a reason to doubt that $c_+ = c_-$?

1.15.3 *Mahler's Measure*

Before beginning, we observe that the Laurent polynomial equation

$$1 + x + \frac{1}{x} + y + \frac{1}{y} = 0$$

is isomorphic to the elliptic curve $15A8$ via the change of coordinates [74, 75]

$$(x, y) \mapsto \left(\frac{y}{x}, \frac{x^3 - y^2 - xy}{xy}\right).$$

Similarly, the equation

$$1 + x + \frac{1}{x} + y + \frac{1}{y} + xy + \frac{1}{xy} = 0$$

is isomorphic to the curve $14A4$, and the equation

$$-1 + x + \frac{1}{x} + y + \frac{1}{y} + xy + \frac{1}{xy} = 0$$

is isomorphic to the curve $30A1$. Such representations of elliptic curves (as polynomials in x, x^{-1}, y, y^{-1}) are especially attractive when symmetric in x, y as shown.

The **(logarithmic) Mahler measure** of a Laurent polynomial $P(x_1, x_2, \ldots, x_n) \in \mathbb{Z}[x_1^{\pm 1}, x_2^{\pm 1}, \ldots, x_n^{\pm 1}]$ is defined to be

$$m(P) = \int_0^1\int_0^1 \cdots \int_0^1 \ln\left|P(e^{2\pi i\theta_1}, e^{2\pi i\theta_2}, ..., e^{2\pi i\theta_n})\right| \, d\theta_1 d\theta_2 \cdots d\theta_n.$$

We studied $\exp(m(P))$ for univariate P in [76]; our focus here will be on the case $n \geq 2$. Smyth [77, 78] proved that

$$\begin{aligned} m(1+x_1+x_2) &= L'_{-3}(-1) = \frac{3\sqrt{3}}{4\pi} L_{-3}(2) = 0.3230659472... \\ &= \ln(1.3813564445...), \end{aligned}$$

$$\begin{aligned} m(1+x_1+x_2+x_3) &= 14\zeta'(-2) = \frac{7}{2\pi^2}\zeta(3) = 0.4262783988... \\ &= \ln(1.5315470966...) \end{aligned}$$

and Rodriguez-Villegas [79–81] conjectured that

$$m(1+x_1+x_2+x_3+x_4) = -L'_f(-1) = \frac{675\sqrt{15}}{16\pi^5} L_f(4) = 0.5444125617...,$$

$$m(1+x_1+x_2+x_3+x_4+x_5) = -8L'_g(-1) = \frac{648}{\pi^6} L_g(5) = 0.6273170748...$$

where f, g are the cusp forms defined at the end of §1.15.1. Deninger [82] conjectured that

$$\begin{aligned} m\left(1+x+\frac{1}{x}+y+\frac{1}{y}\right) &= L'_{15A}(0) = \frac{15}{4\pi^2} L_{15A}(2) = 0.2513304337... \\ &= \ln(1.2857348642...) \end{aligned}$$

and Boyd [75] conjectured that

$$\begin{aligned} m\left(1+x+\frac{1}{x}+y+\frac{1}{y}+xy+\frac{1}{xy}\right) &= L'_{14A}(0) = \frac{7}{2\pi^2} L_{14A}(2) = 0.2274812230... \\ &= \ln(1.2554338662...). \end{aligned}$$

The latter is the smallest known measure of bivariate polynomials; the former is the second-smallest known. Both conjectures can be rephrased in completely explicit terms [75]: If

$$\sum_{n=1}^{\infty} a_n q^n = q \prod_{k=1}^{\infty} \left(1-q^k\right)\left(1-q^{3k}\right)\left(1-q^{5k}\right)\left(1-q^{15k}\right),$$

$$\sum_{n=1}^{\infty} b_n q^n = q \prod_{k=1}^{\infty} \left(1-q^k\right)\left(1-q^{2k}\right)\left(1-q^{7k}\right)\left(1-q^{14k}\right)$$

then

$$\int_0^{2\pi}\int_0^{2\pi} \ln|1+2\cos(s)+2\cos(t)|\,ds\,dt = 15\sum_{j=1}^{\infty}\frac{a_j}{j^2},$$

$$\int_0^{2\pi}\int_0^{2\pi} \ln|1+2\cos(s)+2\cos(t)+2\cos(s+t)|\,ds\,dt = 14\sum_{j=1}^{\infty}\frac{b_j}{j^2}.$$

These integrals bear some resemblance to certain constants in [83]. Rogers & Zudilin [84, 85] succeeded in proving Deninger's conjecture; Brunault [86] & Mellit [87] likewise proved Boyd's conjecture. Trivariate analogs of these two examples are [88–90]

$$m\left(1+x+\frac{1}{x}+y+\frac{1}{y}+z+\frac{1}{z}\right)=0.3703929298...=\ln(1.4483035845...),$$

$$\begin{aligned}&m\left(1+x+\frac{1}{x}+y+\frac{1}{y}+z+\frac{1}{z}+xy+\frac{1}{xy}+yz+\frac{1}{yz}+xyz+\frac{1}{xyz}\right)\\&\quad=0.4798982839...\end{aligned}$$

but no relation to special L-series values has yet been proposed. Other variations include [75, 90]

$$m\left(-1+x+\frac{1}{x}+y+\frac{1}{y}+xy+\frac{1}{xy}\right)=L'_{30A}(0)=\frac{15}{2\pi^2}L_{30A}(2)=0.6168709387...,$$

$$\begin{aligned}&m\left(-1+x+\frac{1}{x}+y+\frac{1}{y}+z+\frac{1}{z}+xy+\frac{1}{xy}+yz+\frac{1}{yz}+xyz+\frac{1}{xyz}\right)\\&\quad=0.8157244463....\end{aligned}$$

The third-smallest known measure of bivariate polynomials is [75, 89, 91]

$$m\left(-1+x+\frac{1}{x}-y-\frac{1}{y}+x^2y^2+\frac{1}{x^2y^2}\right)=0.2693386412...=\ln(1.3090983806...)$$

and the fourth-smallest known is [75, 89, 92]

$$\begin{aligned}&m\left(1+x^2+\frac{1}{x^2}+y^2+\frac{1}{y^2}+xy+\frac{1}{xy}+x^2y^2+\frac{1}{x^2y^2}+\frac{y}{x}+\frac{x}{y}\right)\\&\quad=0.2743632972...\\&\quad=\ln(1.3156927029...).\end{aligned}$$

We emphasize that Rodriguez-Villegas' conjectures and several other $m(P)$ formulas exhibited here still await rigorous proof.

1.15.4 *Klein's Modular Invariant*

The only modular form $f:\mathbb{H}\to\mathbb{C}$ of weight 0 is a constant. (Assume, as at the beginning, that f is of level 1 and has trivial character.) What happens if we weaken our hypotheses on f? A **modular function** f is an $\mathrm{SL}_2(\mathbb{Z})$-invariant meromorphic function on $\mathbb{H}$ whose Fourier series $f(z)=\sum_{n=-\infty}^{\infty}\gamma_n q^n$ has at most finitely many $\gamma_n \neq 0$ for $n < 0$. The set of modular functions can be proved to be a field, $\mathbb{C}(j)$, generated by Klein's **j-invariant** or **Hauptmodul** [1, 93–97]

$$j(z) = \frac{1}{Q}(1 + 256Q)^3 = \frac{1}{R}\left(1 + 250R + 3125R^2\right)^3 = \sum_{m=-1}^{\infty} c(m)q^m$$

where

$$Q = q\prod_{n=1}^{\infty}\left(\frac{1-q^{2n}}{1-q^n}\right)^{24} = \frac{\Delta(2z)}{\Delta(z)},$$

$$R = q\prod_{n=1}^{\infty}\left(\frac{1-q^{5n}}{1-q^n}\right)^{6} = \left(\frac{\Delta(5z)}{\Delta(z)}\right)^{1/4}$$

and $c(-1) = 1$, $c(0) = 744$, $c(1) = 196884$, $c(2) = 21493760$, ... Moreover, j is the unique modular function having a simple pole with residue 1 at $q = 0$. Closed-form expressions and asymptotics for $c(m)$ are known [98–100], akin to those for the number $p(m)$ of partitions of m [101]. Special values include

$$j(i) = 12^3, \quad j\left((1 + i\sqrt{3})/2\right) = 0, \quad j\left((1 + i\sqrt{163})/2\right) = (-640320)^3;$$

the latter, plus the fact that $j(z) \approx q^{-1} + 744$, is responsible for the surprising consequence that $e^{\pi\sqrt{163}}$ misses being an integer by less than 10^{-12}. More special values include

$$j\left((1 + i\sqrt{15})/2\right) = x, \quad j\left((1 + i\sqrt{23})/2\right) = y$$

where x, y have minimal polynomials $x^2 + 191025x - 121287375$ and $y^3 + 3491750y^2 - 5151296875y + 12771880859375$, respectively. (The class numbers $h_{-1} = h_{-3} = h_{-163} = 1$, $h_{-15} = 2$ and $h_{-23} = 3$ play a role here [44].) Schneider [102] proved that, if $j(z)$ is algebraic, then z is algebraic if and only if z is imaginary quadratic. It is also known that, if $q \in \mathbb{Q}$ is algebraic and $0 < |q| < 1$, then $j(z)$ is transcendental [103–105]. A connection between sporadic simple group theory and modular functions (on $\Gamma_0(N)$ and extensions) is beyond the scope of our study [106–108].

1.15.5 *Limits*

Here is a seemingly unrelated calculus problem. Let $f(x) = (\pi/4 - x)\ln(g(x))$ be integrable on $[0, \pi/4]$, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{\lfloor n/2\rfloor}\left(1 - \frac{2k}{n}\right)\ln\left[g\left(\frac{\pi k}{2n}\right)\right] = \frac{8}{\pi^2}\lim_{n\to\infty}\sum_{k=1}^{\lfloor n/2\rfloor} f\left(\frac{\pi k}{2n}\right)\left(\frac{\pi(k+1)}{2n} - \frac{\pi k}{2n}\right)$$

$$= \frac{8}{\pi^2}\int_0^{\pi/4} f(x)dx$$

(a limit of Riemann sums). As a simple example,

$$\lim_{n\to\infty}\prod_{k=1}^{\lfloor n/2\rfloor}\left(\frac{n}{2k}\right)^{\frac{1}{n}\left(1-\frac{2k}{n}\right)}=e^{\frac{3}{8}}$$

after setting $g(x)=\pi/(4x)$ and exponentiating. As a more complicated example,

$$\begin{aligned}\lim_{n\to\infty}\prod_{k=1}^{\lfloor n/2\rfloor}\cot\left(\frac{\pi k}{2n}\right)^{\frac{1}{n}\left(1-\frac{2k}{n}\right)} &= e^{\frac{7\zeta(3)}{2\pi^2}}=\exp(0.4262783988...)\\ &= \sqrt{2}\exp(0.0797048085...)\end{aligned}$$

after setting $g(x)=\cot(x)$. The latter appears in the asymptotics of what is called the Atiyah determinant from quantum physics [109].

[1] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, 1984, pp. 98–147, 164, 216–217; MR0766911 (86c:11040).

[2] T. M. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, 2^{nd} ed., Springer-Verlag, 1990, pp. 14, 20–22, 50–51, 113–140; MR1027834 (90j:11001).

[3] S. Ramanujan, On certain arithmetical functions, *Trans. Cambridge Philos. Soc.* 22 (1916) 159–184; also in *Collected Papers*, ed. G. H. Hardy, P. V. Seshu Aiyar, and B. M. Wilson, Cambridge Univ. Press, 1927, pp. 136–162, 340–341.

[4] G. H. Hardy, Note on Ramanujan's arithmetical function $\tau(n)$, *Proc. Cambridge Philos. Soc.* 23 (1927) 675–680; also in *Collected Papers*, v. 2, Oxford Univ. Press, 1967, pp. 358–363.

[5] G. H. Hardy, A further note on Ramanujan's arithmetical function $\tau(n)$, *Proc. Cambridge Philos. Soc.* 34 (1938) 309–315; also in *Collected Papers*, v. 2, Oxford Univ. Press, 1967, pp. 369–375.

[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000594, A027364, A037944.

[7] B. Conrey and D. Farmer, Degree 2, level 1 cuspforms, Database of L-functions, Modular Forms, and Related Objects (LMFDB).

[8] L. J. Mordell, On Mr. Ramanujan's empirical expansions of modular functions, *Proc. Cambridge Philos. Soc.* 19 (1917) 117–124.

[9] P. Deligne, Formes modulaires et représentations ℓ-adiques, *Séminaire Bourbaki. Vol. 1968/69: Exposés 347–363*, Lect. Notes in Math. 179, Springer-Verlag, 1971, pp. 139–172; MR0272579 (42 #7460).

[10] P. Deligne, La conjecture de Weil. I, *Inst. Hautes Études Sci. Publ. Math.* 43 (1974) 273–307; MR0340258 (49 #5013).

[11] N. M. Katz, An overview of Deligne's proof of the Riemann hypothesis for varieties over finite fields, *Mathematical Developments Arising From Hilbert Problems*, ed. F. E. Browder, Proc. Symp. Pure Math. 28, Amer. Math. Soc., 1976, pp. 275–305; MR0424822 (54 #12780).

[12] M. Ram Murty, The Ramanujan τ function, *Ramanujan Revisited*, Proc. 1987 Urbana conf., ed. G. E. Andrews, R. A. Askey, B. C. Berndt, K. G. Ramanathan, and R. A. Rankin, Academic Press, 1988, pp. 269–288; MR0938969 (89e:11027).

[13] R. A. Rankin, An Ω-result for the coefficients of cusp forms, *Math. Annalen* 203 (1973) 239–250; MR0321876 (48 #241).

[14] H. Joris, An Ω-result for the coefficients of cusp forms, *Mathematika* 22 (1975) 12–19; MR0419362 (54 #7383).

[15] R. Balasubramanian and M. Ram Murty, An Ω-theorem for Ramanujan's τ-function, *Invent. Math.* 68 (1982) 241–252; MR0666161 (84a:10023a).

[16] M. Ram Murty, Some Ω-results for Ramanujan's τ-function, *Number Theory*, Proc. 1981 Mysore conf., ed. K. Alladi, Lect. Notes in Math. 938, Springer-Verlag, 1982, pp. 123–137; MR0665444 (84a:10023b).

[17] M. Ram Murty, Oscillations of Fourier coefficients of modular forms, *Math. Annalen* 262 (1983) 431–446; MR0696516 (84g:10053).

[18] D. Zagier, Modular forms whose Fourier coefficients involve zeta-functions of quadratic fields, *Modular Functions of One Variable. VI*, Proc. 1976 Bonn conf., ed. J.-P. Serre and D. Zagier, Lect. Notes in Math. 627, Springer-Verlag, 1977, pp. 105–169; MR0485703 (58 #5525).

[19] W. Kohnen and D. Zagier, Values of L-series of modular forms at the center of the critical strip, *Invent. Math.* 64 (1981) 175–198; MR0629468 (83b:10029).

[20] D. Zagier, Periods of modular forms and Jacobi theta functions, *Invent. Math.* 104 (1991) 449–465; MR1106744 (92e:11052).

[21] M. Kontsevich and D. Zagier, Periods, *Mathematics Unlimited – 2001 and Beyond*, ed. B. Engquist and W. Schmid, Springer-Verlag, 2001, pp. 771–808; MR1852188 (2002i:11002).

[22] R. A. Rankin, Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions. I, The zeros of the function $\sum_{n=1}^{\infty} \tau(n)/n^s$ on the line $\mathrm{Re}(s) = 13/2$. II, The order of the Fourier coefficients of integral modular forms, *Proc. Cambridge Philos. Soc.* 35 (1939) 351–372; MR0000411 (1,69d).

[23] A. Selberg, On the estimation of Fourier coefficients of modular forms, *Theory of Numbers*, ed. A. L. Whiteman, Proc. Symp. Pure Math. 8, Amer. Math. Soc., 1965, pp. 1–15; MR0182610 (32 #93).

[24] D. H. Lehmer, Ramanujan's function $\tau(n)$, *Duke Math. J.* 10 (1943) 483–492; MR0008619 (5,35b).

[25] R. A. Rankin, Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions. III, A note on the sum function of the Fourier coefficients of integral modular forms, *Proc. Cambridge Philos. Soc.* 36 (1940) 150–151; MR0001249 (1,203d).

[26] G. H. Hardy, *Ramanujan: Twelve Lectures on Subjects Suggested by His Life and Work*, Chelsea, 1940, pp. 161–185; MR0106147 (21 #4881).

[27] W. B. Pennington, On the order of magnitude of Ramanujan's arithmetical function $\tau(n)$, *Proc. Cambridge Philos. Soc.* 47 (1951) 668–678; MR0043126 (13,209d).

[28] K. Chandrasekharan and R. Narasimhan, Hecke's functional equation and the average order of arithmetical functions, *Acta Arith.* 6 (1960/1961) 487–503; MR0126423 (23 #A3719).

[29] R. Spira, Calculation of the Ramanujan τ-Dirichlet series, *Math. Comp.* 27 (1973) 379–385; MR0326995 (48 #5337).

[30] H. Yoshida, On calculations of zeros of L-functions related with Ramanujan's discriminant function on the critical line, *J. Ramanujan Math. Soc.* 3 (1988) 87–95; MR0975839 (90b:11044).

[31] J. B. Keiper, On the zeros of the Ramanujan τ-Dirichlet series in the critical strip, *Math. Comp.* 65 (1996) 1613–1619; MR1344615 (97a:11073).

[32] S. R. Finch, Elliptic curves over $\mathbb{Q}$, *this volume*, §1.14.

[33] H. Darmon, A proof of the full Shimura-Taniyama-Weil conjecture is announced, *Notices Amer. Math. Soc.* 46 (1999) 1397–1401; MR1723249 (2000j:11082).

[34] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton Univ. Press, 1971, pp. 20–27; MR0314766 (47 #3318).
[35] H. Cohen and J. Oesterlé, Dimensions des espaces de formes modulaires, *Modular Functions of One Variable. VI*, Proc. 1997 Bonn conf., ed. J.-P. Serre and D. B. Zagier, Lect. Notes in Math. 627, Springer-Verlag, 1977, pp. 69–78; MR0472703 (57 #12396).
[36] G. Martin, Dimensions of the spaces of cusp forms and newforms on $\Gamma_0(N)$ and $\Gamma_1(N)$, *J. Number Theory* 112 (2005) 298–331; math.NT/0306128; MR2141534 (2005m:11069).
[37] W. Stein, *Modular Forms: A Computational Approach*, Amer. Math. Soc, 2007, ch. 6; MR2289048 (2008d:11037).
[38] W. Stein, Dimensions of spaces of cuspforms, The Modular Forms Database.
[39] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000086, A000089, A001615, A001616, A001617.
[40] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.
[41] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[42] J. A. Csirik, M. Zieve and J. Wetherell, On the genera of $X_0(N)$, math.NT/0006096.
[43] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A029937, A111248, A159050.
[44] S. R. Finch, Class number theory, *this volume*, §1.12.
[45] K. Ono, On a newform of level 15, weight 3 and character $[1, 2]$, unpublished note (2005).
[46] M. Somos, Eta expression for a newform and its similarity to Ramanujan's entry 10(v), unpublished note (2005).
[47] B.C. Berndt, *Ramanujan's Notebooks. Part III*, Springer-Verlag, 1991, pp. 34–38, 379–382; MR1117903 (92j:01069).
[48] P. D. T. A. Elliott, Mean value theorems for multiplicative functions bounded in mean α-power, $\alpha > 1$, *J. Austral. Math. Soc. Ser. A* 29 (1980) 177–205; MR0566285 (81i:10057).
[49] P. D. T. A. Elliott, Multiplicative functions and Ramanujan's τ-function, *J. Austral. Math. Soc. Ser. A* 30 (1980/81) 461–468; MR0621561 (82h:10058).
[50] R. A. Rankin, Sums of powers of cusp form coefficients, *Math. Annalen* 263 (1983) 227–236; MR0698005 (84h:10033).
[51] P. D. T. A. Elliott, C. J. Moreno and F. Shahidi, On the absolute value of Ramanujan's τ-function, *Math. Annalen* 266 (1984) 507–511; MR0735531 (85f:11030).
[52] R. A. Rankin, Sums of powers of cusp form coefficients. II, *Math. Annalen* 272 (1985) 593–600; MR0807293 (87d:11032).
[53] C. J. Moreno and F. Shahidi, The fourth moment of Ramanujan τ-function, *Math. Annalen* 266 (1983) 233–239; MR0724740 (85i:11039).
[54] L. G. Fel, Summatory multiplicative arithmetic functions: scaling and renormalization, arXiv:1108.0957.
[55] A. Walfisz, Über die Koeffizientensummen einiger Modulformen, *Math. Annalen* 108 (1933) 75-90.
[56] K. Chandrasekharan and R. Narasimhan, On the mean value of the error term for a class of arithmetical functions, *Acta Math.* 112 (1964) 41–67; MR0160765 (28 #3976).
[57] A. Ivic, Large values of certain number-theoretic error terms, *Acta Arith.* 56 (1990) 135–159; MR1075641 (91j:11078).
[58] Y. Cai, On the third and fourth power moments of Fourier coefficients of cusp forms, *Acta Math. Sinica* 13 (1997) 443–452; MR1612698 (99b:11048).

[59] Y.-K. Lau, On the mean square formula of the error term for a class of arithmetical functions, *Monatsh. Math.* 128 (1999) 111–129; MR1712484 (2000h:11107).
[60] W. Zhai, On higher-power moments of $\Delta(x)$, *Acta Arith.* 112 (2004) 367–395; MR2046947 (2005g:11188).
[61] W. Zhai, On higher-power moments of $\Delta(x)$. II, *Acta Arith.* 114 (2004) 35–54; MR2067871 (2005h:11216).
[62] W. Zhai, On higher-power moments of $\Delta(x)$. III, *Acta Arith.* 118 (2005) 263–281; MR2168766 (2006f:11121).
[63] T. Dokchitser, Calculating Rankin zeta function values using ComputeL, unpublished note (2005).
[64] T. Dokchitser, Computing special values of motivic L-functions, *Experim. Math.* 13 (2004) 137–149; MR2068888 (2005f:11128).
[65] H. Cramér, Über zwei Sätze des Herrn G. H. Hardy, *Math. Z.* 15 (1922) 201–210; also in *Collected Works*, v.1, ed. A. Martin-Löf, Springer-Verlag, pp. 236–245.
[66] E. Landau, Über die Gitterpunkte in einem Kreise (Vierte Mitteilung), *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Klasse* (1924) 58–65; also in *Collected Works*, v. 8, ed. L. Mirsky, I. J. Schoenberg, W. Schwarz, and H. Wefelscheid, Thales Verlag, 1983, pp. 59–66.
[67] E. Landau, *Vorlesungen über Zahlentheorie*, v. 2, Verlag von S. Hirzel, 1927, pp. 250–263.
[68] A. Walfisz, Teilerprobleme, *Math. Z.* 26 (1927) 66–88.
[69] K.-C. Tong, On divisor problems. I (in Chinese), *Acta Math. Sinica* 5 (1955) 313–324; MR0073632 (17,462c).
[70] K.-C. Tong, On divisor problems. II, III (in Chinese), *Acta Math. Sinica* 6 (1956) 139–152, 515–541; MR0098718 (20 #5173).
[71] S. R. Finch, Sierpinski's constant, *first volume*, pp. 122–125.
[72] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[73] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[74] J. W. S. Cassels, *Lectures on Elliptic Curves*, Cambridge Univ. Press, 1991, pp. 32–38; MR1144763 (92k:11058).
[75] D. W. Boyd, Mahler's measure and special values of L-functions, *Experim. Math.* 7 (1998) 37–82; MR1618282 (99d:11070).
[76] S. R. Finch, Kneser-Mahler polynomial constants, *first volume*, pp. 231–235.
[77] C. J. Smyth, On measures of polynomials in several variables, *Bull. Austral. Math. Soc.* 23 (1981) 49–63; G. Myerson and C. J. Smyth, corrigendum 26 (1982) 317–319; MR0615132 (82k:10074) and MR0683659 (84g:10088).
[78] C. J. Smyth, An explicit formula for the Mahler measure of a family of 3-variable polynomials, *J. Théor. Nombres Bordeaux* 14 (2002) 683–700; MR2040701 (2004k:11178).
[79] D. W. Boyd, D. Lind, F. Rodriguez-Villegas and C. Deninger, The many aspects of Mahler's measure, Proc. 2003 Banff workshop.
[80] D. W. Boyd, Explicit formulas for Mahler measure, *Bulletin du CRM*, v. 11 (2005) n. 1, pp. 14–15.
[81] F. Rodriguez-Villegas, R. Toledano and J. D. Vaaler, Estimates for Mahler's measure of a linear form, *Proc. Edinburgh Math. Soc.* 47 (2004) 473–494; MR2081066 (2005f:11237).
[82] C. Deninger, Deligne periods of mixed motives, K-theory and the entropy of certain $\mathbb{Z}^n$-actions, *J. Amer. Math. Soc.* 10 (1997) 259–281; MR1415320 (97k:11101).
[83] S. R. Finch, Monomer-dimer constants, *first volume*, pp. 406–412.
[84] M. Rogers and W. Zudilin, On the Mahler measure of $1+X+1/X+Y+1/Y$, *Internat. Math. Res. Notices* (2014) 2305–2326; arXiv:1102.1153; MR3207368.

[85] D. Borwein, J. M. Borwein, A. Straub and J. Wan, Log-sine evaluations of Mahler measures. II, *Integers* 12 (2012) 1179–1212; arXiv:1103.3035; MR3011556.

[86] F. Brunault, Version explicite du théorème de Beilinson pour la courbe modulaire X1(N), *C. R. Math. Acad. Sci. Paris* 343 (2006) 505–510; MR2267584 (2007g:11065).

[87] A. Mellit, Elliptic dilogarithms and parallel lines, arXiv:1207.4722.

[88] M. J. Bertin, Mesure de Mahler d'hypersurfaces $K3$, math.NT/0501153 (her value for $\exp(m(Q_1)) \approx 1.43517$, however, is incorrect).

[89] D. W. Boyd and M. J. Mossinghoff, Small limit points of Mahler's measure, *Experim. Math.* 14 (2005) 403–414; MR2193803 (2006g:11216).

[90] D. W. Boyd, Computations of Bertin's $m(Q_k)$ for $|k| \leq 250$, unpublished note (2005).

[91] D. S. Silver and S. G. Williams, Mahler measure of Alexander polynomials, *J. London Math. Soc.* 69 (2004) 767–782; math.GT/0105234.

[92] M. J. Mossinghoff, Several representations of $P_{3,5}(x,y)$, unpublished note (2005).

[93] T. M. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, 2nd ed., Springer-Verlag, 1990, pp. 15–22, 34–44, 74–91; MR1027834 (90j:11001).

[94] K. Ono, *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series*, Amer. Math. Soc., 2004, pp. 16–19; MR2020489 (2005c:11053).

[95] B. C. Berndt, H. H. Chan, J. Sohn and S. H. Son, Eisenstein series in Ramanujan's lost notebook, *Ramanujan J.* 4 (2000) 81–114; MR1754634 (2001j:11018).

[96] B. C. Berndt and A. J. Yee, Ramanujan's contributions to Eisenstein series, especially in his lost notebook, *Number Theoretic Methods - Future Trends*, Proc. 2001 Iizuka seminar, ed. S. Kanemitsu and C. Jia, Kluwer, 2002, pp. 31–53; MR1974133 (2004d:11088).

[97] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000521, A035099, A097340, A058496, A096563.

[98] H. Petersson, Über die Entwicklungskoeffizienten der automorphen formen, *Acta Math.* 58 (1932) 169–215.

[99] H. Rademacher, The Fourier coefficients of the modular invariant $J(\tau)$, *Amer. J. Math*. 60 (1938) 501–512.

[100] H. Rademacher, The Fourier series and the functional equation of the absolute modular invariant $J(\tau)$, *Amer. J. Math.* 61 (1939) 237–248; correction, 64 (1942) 456.

[101] S. R. Finch, Integer partitions, *this volume*, §1.11.

[102] T. Schneider, Arithmetische Untersuchungen elliptischer Integrale, *Math. Annalen* 113 (1936) 1–13.

[103] K. Barré-Sirieix, G. Diaz, F. Gramain and G. Philibert, Une preuve de la conjecture de Mahler-Manin, *Invent. Math.* 124 (1996) 1–9; MR1369409 (96j:11103).

[104] Y. V. Nesterenko and P. Philippon, *Introduction to Algebraic Independence Theory*, Lect. Notes in Math. 1752, Springer-Verlag, 2001; MR1837822 (2002g:11104).

[105] M. Waldschmidt, Elliptic functions and transcendence, *Surveys in Number Theory*, ed. K. Alladi, Springer-Verlag, 2008, pp. 143–188; MR2462949 (2010a:11142).

[106] R. E. Borcherds, What is moonshine? *International Congress of Mathematicians*, v. 1. Proc. 1998 Berlin conf., *Documenta Mathematica* Extra Volume I (1998) 607–615; MR1660657 (99j:17001).

[107] T. Gannon, Monstrous moonshine: The first twenty-five years, *Bull. London Math. Soc.* 38 (2006) 1–33; math.QA/0402345; MR2201600 (2006k:11071).

[108] T. Gannon, The algebraic meaning of genus-zero, math.NT/0512248.

[109] M. Mazur and B. V. Petrenko, On the conjectures of Atiyah and Sutcliffe, *Geom. Dedicata* 158 (2012) 329–342; arXiv:1102.4662; MR2922719.

1.16 Chebyshev's Bias

How do we quantify irregularities in the distribution of prime numbers? Define

$$\pi_{q,a}(n) = \#\{p \le n : p \equiv a \bmod q\}$$

where $\gcd(a,q) = 1$. A well-known result:

$$\lim_{n\to\infty} \frac{\ln(n)}{n}\pi_{q,a}(n) = \frac{1}{\varphi(q)}$$

informs us that primes are asymptotically equidistributed modulo q, where $\varphi(q)$ is the Euler totient. There is, however, unrest beneath the surface of such symmetry. For fixed $a_1, a_1, \ldots, a_r$ and q, define

$$S_N = \#\{n \le N : \pi_{q,a_1}(n) > \pi_{q,a_2}(n) > \ldots > \pi_{q,a_r}(n)\}$$

and

$$P(a_1 > a_2 > \ldots > a_r \bmod q) = \lim_{N\to\infty} \frac{1}{\ln(N)} \sum_{n\in S_N} \frac{1}{n}.$$

As the notation suggests, P is to be interpreted as a probability (via logarithmic measure). Rubinstein & Sarnak [1], assuming both the Generalized Riemann Hypothesis and the Grand Simplicity Hypothesis [2], succeeded in proving that

$$P(3 > 1 \bmod 4) = 0.9959280...,$$

$$P(2 > 1 \bmod 3) = 0.9990633....$$

Feuerverger & Martin [3] further proved that

$$P(3 > 5 > 7 \bmod 8) = P(7 > 5 > 3 \bmod 8) = 0.1928013...,$$

$$P(3 > 7 > 5 \bmod 8) = P(5 > 7 > 3 \bmod 8) = 0.1664263...,$$

$$P(5 > 3 > 7 \bmod 8) = P(7 > 3 > 5 \bmod 8) = 0.1407724...$$

and

$$P(5 > 7 > 11 \bmod 12) = P(11 > 7 > 5 \bmod 12) = 0.1984521...,$$

$$P(7 > 5 > 11 \bmod 12) = P(11 > 5 > 7 \bmod 12) = 0.1799849...,$$

$$P(5 > 11 > 7 \bmod 12) = P(7 > 11 > 5 \bmod 12) = 0.1215630...;$$

thus it is more probable that 5 will occupy the middle position for mod 8, and 7 will occupy the middle position for mod 12!

New constants do not always emerge: we have, for example,

$$P(1 > 4 \bmod 5) = P(2 > 3 \bmod 5) = \frac{1}{2}$$

which is due to 1, 4 being squares mod 5 and 2, 3 being nonsquares mod 5. Also

$$P(1>2>4 \bmod 7)=P(3>5>6 \bmod 7)=\frac{1}{6}$$

which is due to 1, 2, 4 being squares mod 7 and 3, 5, 6 being nonsquares mod 7. Examples with exact probabilities $1/r!$, where $r>3$, have not been found.

Define the logarithmic integral

$$\operatorname{li}(x)=\int_2^x \frac{1}{\ln(t)}\,dt$$

for $x \geq 2$ and

$$T_N=\#\left\{n \leq N: \pi_{1,0}(n)>\operatorname{li}(n)\right\}.$$

In another demonstration of their methods, Rubinstein & Sarnak [1] showed that

$$\lim_{N\to\infty}\frac{1}{\ln(N)}\sum_{n\in T_N}\frac{1}{n}=0.00000026...=1-0.99999973....$$

Further results have been obtained by Ng [4], as reported in [5]. Consider, for instance, the q-series coefficients $\{a_n\}_{n=1}^{\infty}$ of the modular form $\eta(z)\eta(23z)$ [6]. Letting

$$\chi_b(n)=\#\left\{p \leq n: a_p=b\right\}$$

for $b=2, 0, -1$, we have

$$\lim_{n\to\infty}\frac{\ln(n)}{n}\chi_b(n)=\begin{cases}1/6 & \text{if } b=2,\\ 1/2 & \text{if } b=0,\\ 1/3 & \text{if } b=-1\end{cases}$$

and

$$\lim_{N\to\infty}\frac{1}{\ln(N)}\sum_{n\in U_N}\frac{1}{n}=0.98309... \quad \text{where} \quad U_N=\#\left\{n \leq N: 2\chi_0(n)>6\chi_2(n)\right\},$$

$$\lim_{N\to\infty}\frac{1}{\ln(N)}\sum_{n\in V_N}\frac{1}{n}=0.72469... \quad \text{where} \quad V_N=\#\left\{n \leq N: 2\chi_0(n)>3\chi_{-1}(n)\right\},$$

$$\lim_{N\to\infty}\frac{1}{\ln(N)}\sum_{n\in W_N}\frac{1}{n}=0.97504... \quad \text{where} \quad W_N=\#\left\{n \leq N: 3\chi_{-1}(n)>6\chi_2(n)\right\}.$$

Let us return to the usual sense of probability (via uniform measure). Brent [7] conjectured that, for random $0<N<n$, we have

$$\lim_{n\to\infty}\mathrm{P}\left(\frac{\operatorname{li}(N)-\pi_{1,0}(N)}{\sqrt{N}/\ln(N)}<x\right)=F(x)$$

where the probability distribution F has mean $\mu = 1$ and variance $\sigma^2 \approx (0.21)^2$. If the Riemann Hypothesis is true, then it can be shown that [8]

$$\begin{aligned}\sigma^2 &= 2 - \ln(4\pi) + \gamma = (0.2149218879...)^2 \\ &= 0.0461914179... = 2(0.0230957089...)\end{aligned}$$

which we have seen elsewhere [9, 10]. An open question is whether F is the normal distribution; a density plot [1] and a time series graph [5] suggest that the answer might be yes. We also wonder about extensions of this probabilistic result to $\pi_{q,a}(n)$ for arbitrary a and q.

If, in the definition of $\pi_{q,a}(n)$, the symbol p is understood to encompass *semiprimes* (products of two primes) rather than primes, then with formulas for S_N and P exactly as before [11, 12],

$$P(3 > 1 \bmod 4) = 0.10572....$$

Hence the bias for semiprimes is reversed from that of primes, although it is less pronounced. The terms *2-almost prime* or *biprime* are often encountered; a less common term *quasi-prime* appears in [11].

[1] M. Rubinstein and P. Sarnak, Chebyshev's bias, *Experim. Math.* 3 (1994) 173–197; MR1329368 (96d:11099).

[2] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.

[3] A. Feuerverger and G. Martin, Biases in the Shanks-Rényi prime number race, *Experim. Math.* 9 (2000) 535–570; MR1806291 (2002d:11111).

[4] N. C. Ng, Limiting Distributions and Zeros of Artin L-functions, Ph.D. thesis, University of British Columbia, 2000.

[5] A. Granville and G. Martin, Prime number races, *Amer. Math. Monthly* 113 (2006) 1–33; Spanish version in *La Gaceta de la Real Sociedad Matematica Espanola* 8 (2005) 197–240; MR2158415 (2006f:11114).

[6] S. R. Finch, Dedekind eta products, *this volume*, §1.19.

[7] R. P. Brent, Irregularities in the distribution of primes and twin primes, *Math. Comp.* 29 (1975) 43–56; correction 30 (1976) 198; MR0369287 (51 #5522) and MR0396437 (53 #302).

[8] R. P. Brent, Comments on "Irregularities...", unpublished note (2006).

[9] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[10] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.

[11] K. Ford and J. Sneed, Chebyshev's bias for products of two primes, *Experim. Math.* 19 (2010) 385–398; arXiv:0908.0093; MR2778652 (2012a:11141).

[12] G. Martin and J. Scarfy, Comparative prime number theory: A survey, arXiv:1202.3408.

1.17 Pattern-Avoiding Permutations

Let $\sigma = \sigma_1\sigma_2\cdots\sigma_m$ be a permutation on $\{1, 2, \ldots, m\}$. Define a **pattern** $\tilde{\sigma}$ to be the string $\sigma_1\varepsilon_1\sigma_2\varepsilon_2\cdots\varepsilon_{m-1}\sigma_m$, where each ε_j is either the dash symbol - or the empty string. For example,

$$1\text{-}3\text{-}2, \quad 1\text{-}32, \quad 132$$

are three distinct patterns. The first is known as a **classical pattern** (dashes in all $m-1$ slots); the third is also known as a **consecutive pattern** (no dashes in any slots). Some authors call $\tilde{\sigma}$ a "generalized pattern" and use the word "pattern" exclusively for what we call "classical patterns".

Let $\tau = \tau_1\tau_2\cdots\tau_n$ be a permutation on $\{1,2,\ldots,n\}$, where $n \geq m$. We say that τ **contains** $\tilde{\sigma}$ if there exist $1 \leq i_1 < i_2 < \ldots < i_m \leq n$ such that

- for each $1 \leq j \leq m-1$, if ε_j is empty, then $i_{j+1} = i_j + 1$;
- for all $1 \leq k \leq m$, $1 \leq l \leq m$, we have $\tau_{i_k} < \tau_{i_l}$ if and only if $\sigma_k < \sigma_l$.

The string $\tau_{i_1}\tau_{i_2}\cdots\tau_{i_m}$ is called an **occurrence** of $\tilde{\sigma}$ in τ. If τ does not contain $\tilde{\sigma}$, then we say τ **avoids** $\tilde{\sigma}$ or that τ is $\tilde{\sigma}$**-avoiding**. For example,

24531 contains 1-3-2

because 253 has the same relative order as 132, but

42351 avoids 1-3-2.

As another example,

6725341 contains 4132

because 7253 has the same relative order as 4132 and consists of four consectutive elements, but

41352 avoids 4132.

As a final example,

3542716 contains 12-4-3

because 3576 has the same relative order as 1243 and its first two elements are consecutive, but

3542716 avoids 12-43.

Define $\alpha_n(\tilde{\sigma})$ to be the number of n-symbol, $\tilde{\sigma}$-avoiding permutations. We naturally wish to understand the rate of growth of $\alpha_n(\tilde{\sigma})$ with increasing n.

1.17.1 *Classical Patterns*

The Stanley–Wilf conjecture, proved by Marcus & Tardos [1], was rephrased by Arratia [2] as follows:

$$L(\tilde{\sigma}) = \lim_{n\to\infty} \left(\alpha_n(\sigma_1\text{-}\sigma_2\text{-}\cdots\text{-}\sigma_m)\right)^{1/n}$$

exists and is finite. We have [3–7]

$$L(\tilde{\sigma}) = 4 \quad \text{when } m = 3,$$

$$L(1\text{-}2\text{-}\cdots\text{-}m) = (m-1)^2 \quad \text{for all } m \geq 2,$$
$$L(1\text{-}3\text{-}4\text{-}2) = 8,$$
$$L(1\text{-}2\text{-}4\text{-}5\text{-}3) = \left(1+\sqrt{8}\right)^2 = 9 + 4\sqrt{2}.$$

A conjecture that $L(\tilde{\sigma}) \leq (m-1)^2$ was disproved [8]:

$$9.47 \leq L(1\text{-}3\text{-}2\text{-}4) \leq 288$$

and hence the maximum limiting value (as a function of m) remains open. We wonder if $L(\tilde{\sigma})$ is always necessarily an algebraic number. Also, the preceding bounds were improved [9–13]:

$$10.24 \leq L(1\text{-}3\text{-}2\text{-}4) \leq 13.5$$

and a nonrigorous estimate $L(1\text{-}3\text{-}2\text{-}4) \approx 11.6$ now exists [14, 15].

1.17.2 *Consecutive Patterns*

Elizalde & Noy [16, 17] examined the cases $m = 3$ and $m = 4$. The quantities $\alpha_n(123)$ and $\alpha_n(132)$ satisfy

$$\alpha_n(123) \sim \gamma_1 \cdot \rho_1^n \cdot n!, \quad \alpha_n(132) \sim \gamma_2 \cdot \rho_2^n \cdot n!$$

where

$$\rho_1 = 3\sqrt{3}/(2\pi) = 0.8269933431..., \quad \gamma_1 = \exp\left(\pi/(3\sqrt{3})\right) = 1.8305194665...,$$

$$\rho_2 = 1/\xi = 0.7839769312..., \quad \gamma_2 = \exp(\xi^2/2) = 2.2558142944...$$

and $\xi = 1.2755477364...$ is the unique positive solution of [18]

$$\int_0^x \exp(-t^2/2)\,dt = 1, \quad \text{that is,} \quad \sqrt{\frac{\pi}{2}}\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) = 1.$$

The quantities $\alpha_n(1342)$, $\alpha_n(1234)$ and $\alpha_n(1243)$ satisfy

$$\alpha_n(1342) \sim \gamma_1 \cdot \rho_1^n \cdot n!, \quad \alpha_n(1234) \sim \gamma_2 \cdot \rho_2^n \cdot n!, \quad \alpha_n(1243) \sim \gamma_3 \cdot \rho_3^n \cdot n!$$

where

$$\rho_1 = 1/\xi = 0.9546118344..., \quad \gamma_1 = 1.8305194...,$$
$$\rho_2 = 1/\eta = 0.9630055289..., \quad \gamma_2 = 2.2558142...,$$
$$\rho_3 = 1/\zeta = 0.9528914198..., \quad \gamma_3 = 1.6043282...;$$

ξ, η and ζ are the smallest positive solutions of

$$\int_0^x \exp(-t^3/6)\,dt = 1, \quad \cos(y) - \sin(y) + \exp(-y) = 0,$$

$$3^{1/2}\int_0^z \mathrm{Ai}(-s)\,ds+\int_0^z \mathrm{Bi}(-s)\,ds=\frac{3^{1/3}\Gamma(1/3)}{\pi},$$

respectively, where $\mathrm{Ai}(t)$ and $\mathrm{Bi}(t)$ are the Airy functions [19].

A permutation τ is **nonoverlapping** if it contains no permutation σ such that two copies of σ overlap in more than one entry [20]. For example, $\tau=214365$ contains both 2143 and 4365, both which follow the same pattern and overlap in two entries, hence τ is overlapping. Bóna [21] examined the probability p_n that a randomly selected n-permutation is nonoverlapping, showed that $\{p_n\}_{n=2}^{\infty}$ is strictly decreasing, and computed $\lim_{n\to\infty} p_n=0.36409....$

From the fact that $\alpha_n(123)>\alpha_n(132)$ and $\alpha_n(1234)>\alpha_n(1342)>\alpha_n(1243)$ for suitably large n, it is natural to speculate that $\alpha_n(123\ldots m)$ is asymptotically larger than $\alpha_n(\sigma)$ for any other m-permutation σ (except $m(m-1)\ldots 21$, which is equivalent by symmetry). This conjecture is now a theorem [22].

1.17.3 *Other Results*

Elizalde [23, 24] proved that

$$\lim_{n\to\infty}\left(\frac{\alpha_n(1\text{-}23\text{-}4)}{n!}\right)^{1/n}=0$$

and believed that the same applies to $\alpha_n(12\text{-}34)$, although a proof is not yet known. Ehrenborg, Kitaev & Perry [25] gave more detailed asymptotic expansions for $\alpha_n(123)$ and $\alpha_n(132)$; a similar "translation" of combinatorics into operator eigenvalue analysis was explored in [26]. The field is wide open for research.

Define $\sigma\le\tau$ if τ contains the classical pattern $\tilde{\sigma}$. A **permutation class** C is a set of permutations such that, if $\tau\in C$ and $\sigma\le\tau$, then $\sigma\in C$. Let C_n denote the permutations in C of length n. If $C=\{\text{all permutations}\}$, then $|C_n|=n!$; such behavior is regarded as degenerate and this case is excluded from now on. The Marcus–Tardos theorem implies that, for nondegenerate C,

$$L(C)=\limsup_{n\to\infty}|C_n|^{1/n}<\infty.$$

Consider the set R of all growth rates $L(C)$ and the derived set R' of all accumulation points of R. Vatter [27] proved that

$$\inf\{r\in R: r>2\}=2.0659948920...$$

which is the unique positive zero of $1+2x+x^2+x^3-x^4$, and

$$\inf\{s: s\text{ is an accumulation point of }R'\}=2.2055694304...$$

which is the unique positive zero of $1+2x^2-x^3$. Albert & Linton [28] proved that R is uncountable and thus contains transcendental numbers. Vatter [29] subsequently proved that

$$\inf\{t : R \text{ contains the interval } (t,\infty)\} \leq 2.4818728574...$$

which is the unique positive zero of $-1-2x-2x^2-2x^4+x^5$ and conjectured that $\leq$ can be replaced by $=$. The question of whether limsup in the definition of $L(C)$ can be replaced by lim is also unanswered.

[1] A. Marcus and G. Tardos, Excluded permutation matrices and the Stanley-Wilf conjecture, *J. Combin. Theory Ser. A* 107 (2004) 153–160; MR2063960 (2005b:05009).

[2] R. Arratia, On the Stanley-Wilf conjecture for the number of permutations avoiding a given pattern, *Elec. J. Combin.* 6 (1999) N1; MR1710623 (2000e:05005).

[3] A. Regev, Asymptotic values for degrees associated with strips of Young diagrams, *Adv. Math.* 41 (1981) 115–136; MR0625890 (82h:20015).

[4] M. Bóna, Exact enumeration of 1342-avoiding permutations: a close link with labeled trees and planar maps, *J. Combin. Theory Ser.* A 80 (1997) 257–272; math.CO/9702223; MR1485138 (98j:05003).

[5] M. Bóna, The limit of a Stanley-Wilf sequence is not always rational, and layered patterns beat monotone patterns, *J. Combin. Theory Ser. A* 110 (2005) 223–235; math.CO/0403502; MR2142176 (2005k:05005).

[6] M. Bóna, New records in Stanley-Wilf limits, *European J. Combin.* 28 (2007) 75–85; MR2261804 (2007f:05002).

[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A0022558, A005802, A061552.

[8] M. H. Albert, M. Elder, A. Rechnitzer, P. Westcott and M. Zabrocki, On the Stanley-Wilf limit of 4231-avoiding permutations and a conjecture of Arratia, *Adv. Appl. Math*. 36 (2006) 96–105; math.CO/0502504; MR2199982 (2006k:05003).

[9] A. Claesson, V. Jelínek and E. Steingrímsson, Upper bounds for the Stanley-Wilf limit of 1324 and other layered patterns, *J. Combin. Theory Ser. A* 119 (2012) 1680–1691; arXiv:1111.5736; MR2946382.

[10] M. Bóna, A new upper bound for 1324-avoiding permutations, *Combin. Probab. Comput.* 23 (2014) 717–724; arXiv:1207.2379; MR3249221.

[11] M. Bóna, A new record for 1324-avoiding permutations, *Europ. J. Math.* 1 (2015) 198–206; arXiv:1404.4033; MR3386234.

[12] D. Bevan, Permutations avoiding 1324 and patterns in Łukasiewicz paths, *J. Lond. Math. Soc.* 92 (2015) 105–122; MR3384507.

[13] D. Bevan, R. Brignall, A. E. Price and J. Pantone, Staircases, dominoes, and the growth rate of 1324-avoiders, *Elec. Notes Discrete Math.* 61 (2017) 123–129.

[14] N. Madras and H. Liu, Random pattern-avoiding permutations, *Algorithmic Probability and Combinatorics*, ed. M. E. Lladser, R. S. Maier, M. Mishna and A. Rechnitzer, Amer. Math. Soc., 2010, pp. 173–194; MR2681860 (2012a:05017).

[15] A. R. Conway and A. J. Guttmann, On 1324-avoiding permutations, *Adv. Appl. Math.* 64 (2015) 50–69; MR3300327.

[16] S. Elizalde and M. Noy, Consecutive patterns in permutations, *Adv. Appl. Math.* 30 (2003) 110–125; MR1979785 (2004d:05004).

[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A049774, A111004, A113228, A113229, A117156, A117158, A117226.

[18] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[19] S. R. Finch, Airy function zeroes, *this volume*, §2.5.

[20] A. Duane and J. Remmel, Minimal overlapping patterns in colored permutations, *Elec. J. Combin.* 18 (2011) R25; MR2861417 (2012m:05028).
[21] M. Bóna, Non-overlapping permutation patterns, *Pure Math. Appl. (PU.M.A.)* 22 (2011) 99–105; MR2924741.
[22] S. Elizalde, The most and the least avoided consecutive patterns, *Proc. London Math. Soc.* 106 (2013) 957–979; MR3066745.
[23] S. Elizalde, Asymptotic enumeration of permutations avoiding generalized patterns, *Adv. Appl. Math.* 36 (2006) 138–155; math.CO/0505254; MR2199985 (2006h:05004).
[24] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A113226, A113227.
[25] R. Ehrenborg, S. Kitaev and P. Perry, A spectral approach to pattern-avoiding permutations, *Formal Power Series and Algebraic Combinatorics (FPSAC)*, Proc. 2006 San Diego conf.
[26] R. Ehrenborg and Y. Farjoun, Asymptotics of the Euler number of bipartite graphs, *Adv. Appl. Math.* 44 (2010) 155–167; arXiv:0704.1782; MR2576845 (2011h:05017).
[27] V. Vatter, Small permutation classes, *Proc. Lond. Math. Soc.* 103 (2011) 879–921; arXiv:0712.4006; MR2852292 (2012j:05022).
[28] M. H. Albert and S. A. Linton, Growing at a perfect speed, *Combin. Probab. Comput.* 18 (2009) 301–308; MR2501429 (2010h:05010). .
[29] V. Vatter, Permutation classes of every growth rate above 2.48188..., *Mathematika* 56 (2010) 182–192; arXiv:0807.2815; MR2604993 (2011c:05012).

1.18 Cyclic Group Orders

Let $\mathbb{Z}_n$ denote the cyclic group (under addition) of integers modulo n. Given $m\in\mathbb{Z}^+$ and $x\in\mathbb{Z}_n$, define mx to be $\sum_{k=1}^{m} x$. The **order** of $x\in\mathbb{Z}_n$ is the least $m>0$ such that $mx=0$. Clearly $\operatorname{ord}(x)$ divides n and, for each divisor d of n, there are precisely $\varphi(d)$ elements in $\mathbb{Z}_n$ of order d. Define the **average order** in $\mathbb{Z}_n$ to be [1]

$$\alpha(n)=\frac{1}{n}\sum_{x\in\mathbb{Z}_n}\operatorname{ord}(x)=\frac{1}{n}\sum_{d|n}d\,\varphi(d).$$

Asymptotically, we have

$$\sum_{n\leq N}\alpha(n)\sim\frac{\zeta(3)}{2\zeta(2)}N^2=\frac{3\zeta(3)}{\pi^2}N^2=(0.3653814847...)N^2$$

as $N\to\infty$. Variations of this result include [1, 2]

$$\sum_{n\leq N}\frac{\alpha(n)}{n}\sim\frac{\zeta(3)}{\zeta(2)}N=\frac{6\zeta(3)}{\pi^2}N=(0.7307629694...)N,$$

$$\sum_{n\leq N}\frac{\alpha(n)}{\varphi(n)}\sim\frac{\zeta(3)\zeta(4)}{\zeta(8)}N=\frac{105\zeta(3)}{\pi^4}N=(1.2957309578...)N,$$

$$\sum_{n\leq N}\frac{n}{\alpha(n)}\sim C_1N,\quad\sum_{n\leq N}\frac{\varphi(n)}{\alpha(n)}\sim C_2N$$

where

$$C_1 = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \left(1 + \frac{1}{p}\right) \sum_{k=1}^{\infty} \frac{1}{p^k + p^{-k-1}}\right) = 1.4438675...,$$

$$C_2 = \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \left(1 - \frac{1}{p^2}\right) \sum_{k=1}^{\infty} \frac{1}{p^k + p^{-k-1}}\right) = 0.8014696934....$$

Let $\mathbb{F}_q^*$ denote the cyclic group (under multiplication) of nonzero elements of $\mathbb{F}_q$, the field of size q. It is well-known that q must be a prime power. The order of $x \in \mathbb{F}_q^*$ is the least $m > 0$ such that $x^m = 1$ and the average order in $\mathbb{F}_q^*$ is

$$\alpha(q-1) = \frac{1}{q-1} \sum_{x \in \mathbb{F}_q^*} \operatorname{ord}(x) = \frac{1}{q-1} \sum_{d|q-1} d\varphi(d).$$

We examine two cases: the first when q is actually a prime [2, 3]:

$$\sum_{q \le Q} \frac{\alpha(q-1)}{q-1} \sim C_3 \frac{Q}{\ln(Q)}, \quad \sum_{q \le Q} \frac{\alpha(q-1)}{\varphi(q-1)} \sim C_4 \frac{Q}{\ln(Q)}$$

where

$$C_3 = \prod_p \left(1 - \frac{p}{p^3 - 1}\right) = 0.5759599688...$$

is Stephens' constant [4, 5],

$$C_4 = \prod_p \left(1 + \frac{p+1}{(p-1)^2(p^2+p+1)}\right) = 1.5664205124...;$$

and the second when $q = 2^k$ for some $k \ge 1$ [2, 3]:

$$\sum_{k \le K} \frac{\alpha(2^k - 1)}{2^k - 1} \sim C_5 K, \quad \sum_{k \le K} \frac{\alpha(2^k - 1)}{\varphi(2^k - 1)} \sim C_6 K$$

where

$$C_5 = \sum_{\substack{n \ge 1, \\ n \text{ odd}}} \frac{f(n)}{t(n)} = 0.786125..., \quad C_6 = \sum_{\substack{n \ge 1, \\ n \text{ odd}}} \frac{g(n)}{t(n)} = 1.102488....$$

In the preceding formulas, f and g are multiplicative functions with

$$f(p^r) = -\frac{p-1}{p^{2r}}, \quad g(p^r) = \begin{cases} \dfrac{1}{p(p-1)} & \text{if } r = 1, \\ -\dfrac{1}{p^{2r-1}} & \text{if } r \ge 2 \end{cases}$$

and $t(n)$ is the order of the element 2 in $\mathbb{Z}_n^*$, the group (under multiplication) of integers relatively prime to n [6]. If we replace α by φ, the following emerge [1, 4]:

$$\sum_{q \le Q} \frac{\varphi(q-1)}{q-1} \sim C_7 \frac{Q}{\ln(Q)}, \quad \sum_{k \le K} \frac{\varphi(2^k-1)}{2^k-1} \sim C_8 K$$

where

$$C_7 = \prod_p \left(1 - \frac{1}{p(p-1)}\right) = 0.3739558136...$$

is Artin's constant [5],

$$C_8 = \sum_{\substack{n \ge 1, \\ n \text{ odd}}} \frac{\mu(n)}{n\, t(n)} = 0.73192...,$$

and μ is the Möbius mu function. Also, we have extreme results [1, 7]:

$$1 = \liminf_{n\to\infty} \frac{\alpha(n)}{\varphi(n)} < \limsup_{n\to\infty} \frac{\alpha(n)}{\varphi(n)} = \frac{\zeta(2)\zeta(3)}{\zeta(6)} = \frac{315}{2\pi^4}\zeta(3) = 1.9435964368....$$

The study of the average order $\xi(n)$ in $\mathbb{Z}_n^*$ was initiated in [8]. We have extreme results

$$\liminf_{n\to\infty} \frac{\xi(n)\ln(\ln(n))}{\lambda(n)} = \frac{e^{-\gamma}\pi^2}{6}, \quad \limsup_{n\to\infty} \frac{\xi(n)}{\lambda(n)} = 1$$

where $\lambda(n)$ is the **reduced totient** or **Carmichael function** [9]:

$$\lambda(n) = \begin{cases} \varphi(n) & \text{if } n = 1, 2, 4 \text{ or } q^j, \text{ where } q \text{ is an odd prime and } j \ge 1, \\ \varphi(n)/2 & \text{if } n = 2^k, \text{ where } k \ge 3, \\ \operatorname{lcm}\left\{\lambda(p_j^{e_j}) : 1 \le j \le l\right\} & \text{if } n = p_1^{e_1} p_2^{e_2} \cdots p_l^{e_l}, \text{ where } 2 \le p_1 < p_2 < \ldots \text{ and } l \ge 2. \end{cases}$$

Observe that $\lambda(n)$ is the size of the largest cyclic subgroup of $\mathbb{Z}_n^*$. A mean result [8, 9]:

$$\frac{1}{N}\sum_{n \le N} \xi(n) = \frac{N}{\ln(N)} \exp\left[\frac{C_9 \ln(\ln(N))}{\ln(\ln(\ln(N)))}(1 + o(1))\right]$$

holds as $N \to \infty$, where

$$C_9 = e^{-\gamma} \prod_p \left(1 - \frac{1}{(p-1)^2(p+1)}\right) = 0.3453720641....$$

There is a set S of positive integers of asymptotic density 1 such that, for $n \in S$,

$$\xi(n) = \frac{n}{(\ln(n))^{\ln(\ln(\ln(n))) + C_{10} + o(1)}}$$

and

$$C_{10} = -1 + \sum_p \frac{\ln(p)}{(p-1)^2} = 0.2269688056...;$$

it is not known whether $S = \mathbb{Z}^+$ is possible.

A different study of periodicity properties of $\{x^k\}_{k=0}^{\infty}$ for each $x \in \mathbb{Z}_n$ (including $\mathbb{Z}_n^*$ and more) has also been undertaken [10, 11]. The constants C_3 and C_9 moreover appear in theorems proved [12–14] assuming the Generalized Riemann Hypothesis.

[1] J. von zur Gathen, A. Knopfmacher, F. Luca, L. G. Lucht and I. E. Shparlinski, Average order in cyclic groups, *J. Théor. Nombres Bordeaux* 16 (2004) 107–123; MR2145575 (2006d:11111).
[2] K.-H. Indlekofer, S. Wehmeier and L. G. Lucht, Mean behaviour and distribution properties of multiplicative functions, *Comput. Math. Appl.* 48 (2004) 1947–1971; MR2116969 (2005m:11168).
[3] F. Luca, Some mean values related to average multiplicative orders of elements in finite fields, *Ramanujan J.* 9 (2005) 33–44; MR2166376 (2006i:11111).
[4] P. J. Stephens, An average result for Artin's conjecture, *Mathematika* 16 (1969) 178–188; MR0498449 (58 #16565).
[5] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.
[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002326.
[7] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.
[8] F. Luca and I. E. Shparlinski, Average multiplicative orders of elements modulo *n*, *Acta Arith.* 109 (2003) 387–411; MR2009051 (2004i:11113).
[9] P. Erdős, C. Pomerance and E. Schmutz, Carmichael's lambda function, *Acta Arith.* 58 (1991) 363–385; MR1121092 (92g:11093).
[10] S. Finch and P. Sebah, Squares and cubes modulo *n*, math.NT/0604465.
[11] S. Finch, Idempotents and nilpotents modulo *n*, math.NT/0605019.
[12] P. Kurlberg and C. Pomerance, On a problem of Arnold: The average multiplicative order of a given integer, *Algebra Number Theory* 7 (2013) 981–999; arXiv:1108.5209; MR3095233.
[13] S. Kim, Average results on the order of *a* modulo *p*, *J. Number Theory* 169 (2016) 353–368; arXiv:1509.01752; MR3531245.
[14] S. Kim, On the order of *a* modulo *n*, on average, *Internat. J. Number Theory* 12 (2016) 2073–2080; arXiv:1509.03768; MR3562014.

1.19 Dedekind Eta Products

For $\mathrm{Im}(z) > 0$, define the **Dedekind eta function**

$$\eta(z) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n) = \Delta(z)^{1/24}$$

where $q = e^{2\pi i z}$ and $\Delta(z)$ is the discriminant function studied earlier [1]. Euler's pentagonal-number theorem states that

$$\eta(24z) = q \prod_{n=1}^{\infty} \left(1 - q^{24n}\right) = \sum_{k=-\infty}^{\infty} (-1)^k q^{(6k+1)^2};$$

we also have

$$\eta(8z)^3 = q\prod_{n=1}^{\infty}\left(1-q^{8n}\right)^3 = \sum_{k=0}^{\infty}(-1)^k(2k+1)q^{(2k+1)^2}$$

via Jacobi's triple-product identity. The absence of a corresponding formula for

$$\eta(12z)^2 = q\prod_{n=1}^{\infty}\left(1-q^{12n}\right)^2$$

or for $\eta(8z)\eta(16z)$, $\eta(6z)\eta(18z)$, $\eta(4z)\eta(20z)$, $\eta(3z)\eta(21z)$, $\eta(2z)\eta(22z)$, $\eta(z)\eta(23z)$ is remarkable! At a minimum, we should be able to say something about the density of nonzero coefficients in the q-series expansion (on the right-hand side).

Given any eta product

$$\eta(b_1z)\eta(b_2z)\cdots\eta(b_mz) = \sum_{k=0}^{\infty}a_kq^k, \quad 1\le b_1\le b_2\le\ldots\le b_m,$$

define the counting function

$$M_{b_1,b_2,\ldots,b_m}(x) = \#\left\{k\le x : a_k\ne 0\right\}.$$

The eta product is said to be **lacunary** if

$$\lim_{x\to\infty}\frac{M_{b_1,b_2,\ldots,b_m}(x)}{x} = 0.$$

For example, it is clear that

$$M_{24}(x)\sim\frac{1}{3}\sqrt{x}, \quad M_{8,8,8}(x)\sim\frac{1}{2}\sqrt{x}$$

as $x\to\infty$. Serre [2–4] proved that

$$M_{12,12}(x)\sim\frac{c\,x}{(\ln x)^{3/4}}$$

where

$$\begin{aligned} c &= \left(\frac{\pi^6\ln(2+\sqrt{3})}{2\cdot 3^7}\right)^{1/4}\frac{1}{\Gamma(1/4)}\prod_{p\equiv 1 \bmod 12}\left(1-\frac{1}{p^2}\right)^{1/2} = 0.2015440949... \\ &= (2.4185291388...)/12 \end{aligned}$$

and Ng [5] proved that

$$M_{1,23}(x)\sim\frac{d\,x}{(\ln x)^{1/2}}$$

where

$$d = \left\{ \frac{3\sqrt{23}}{22} \prod_{p \in S} \left(1 - \frac{1}{p^2}\right)^{-1} \cdot \prod_{p \in T} \left(\frac{1 - 1/p^2}{1 - 1/p^3} \right)^2 \right\}^{1/2}.$$

The set S is defined as the set of all primes p with the property that the cubic polynomial $y^3 - y - 1$ has a single zero modulo p. This turns out to be the same as requiring that the Legendre symbol $(-23/p)$ be equal to -1. The set T is the set of all primes p with the property that $y^3 - y - 1$ has no zeroes modulo p (that is, it is irreducible over $\mathbb{Z}_p$). No equivalent condition involving the Legendre symbol is known [6].

It is also proved that $M_{6,6,6,6}(x)$, $M_{4,4,4,4,4,4}(x)$ and $M_{3,3,3,3,3,3,3,3}(x)$ correspond to lacunary eta products; further, each is asymptotically $Cx/\ln(x)^{1/2}$ for some constant C. In particular, $\eta(6z)^4$ is related to the L-series for the elliptic curve $36A1$:

$$v^2 = u^3 + 1$$

and thus it would be good to better understand the corresponding C.

By contrast, $\eta(2z)^{12}$, $\eta(z)^{24}$ and $\eta(z)^2\eta(11z)^2$ are *not* lacunary. It is conjectured that

$$M_{\underbrace{2,2,\ldots,2}_{12\text{ times}}}(x) \sim x, \quad M_{\underbrace{1,1,\ldots,1}_{24\text{ times}}}(x) \sim x,$$

and that

$$M_{1,1,11,11}(x) \sim \left(\frac{14}{15} \prod_{a_p = 0} \left(1 - \frac{1}{p+1}\right) \right) x = (0.84652\ldots)x.$$

In particular, $\eta(z)^2\eta(11z)^2$ is related to the L-series for the elliptic curve $11A3$:

$$v^2 + v = u^3 - u^2$$

and thus it would be good to compute the associated constant to higher precision.

We mention that the primes p satisfying $a_p = 0$ (as above) are called **supersingular primes**. This sequence of primes begins as 19, 29, 199, 569, 809, ... No explicit formula for a_p as a function of p, or for the n^{th} supersingular prime, is known [7–11].

Another related constant for $11A3$ is

$$\gamma_j = \lim_{x \to \infty} \frac{\#\{p \le x : a_p = j\}}{\sqrt{x}/\ln(x)}$$

for any integer j. If the Lang–Trotter conjecture were proved [7], then it would follow that $\gamma_0 = 23\pi/55 \approx 1.31375$,

$$\gamma_{-1} = \frac{1}{\pi} \frac{11^2}{2^3 \cdot 3^2} A \approx 0.49919, \quad \gamma_{-2} = \frac{1}{\pi} \frac{7 \cdot 11 \cdot 31}{2^4 \cdot 3^2 \cdot 5} A \approx 0.98478$$

where

$$A = \prod_{p \neq 2,5,11} \frac{p(p^2 - p - 1)}{(p-1)(p^2-1)} = \prod_{p \neq 2,5,11} \left(1 - \frac{1}{(p-1)(p^2-1)}\right)$$
$$= 0.9331892646...$$

Some doubt exists, however, whether assumptions underlying Lang–Trotter are justified. We refer the interested reader to [12], which is a work-in-progress addressed to both mathematicians and statisticians. See also [13] for a constant, similar to A, which arises in the study of the reduced totient or Carmichael function.

A recent preprint [14] is concerned not with the density of nonzero coefficients a_k, but instead with the asymptotic mean square of a_k (which perhaps is less difficult).

[1] S. R. Finch, Modular forms on $\mathrm{SL}_2(\mathbb{Z})$, *this volume*, §1.15.

[2] J.-P. Serre, Quelques applications du théorème de densité de Chebotarev, *Inst. Hautes Études Sci. Publ. Math.* n. 54 (1981) 323–401; MR0644559 (83k:12011).

[3] J.-P. Serre, Sur la lacunarité des puissances de η, *Glasgow Math. J.* 27 (1985) 203–221; MR0819840 (87e:11066).

[4] P. Sebah, An infinite product, unpublished note (2006).

[5] N. C. Ng, Limiting Distributions and Zeros of Artin L-functions, Ph.D. thesis, Univ. of British Columbia, 2000.

[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A033217, A086965, A106867.

[7] S. Lang and H. Trotter, *Frobenius Distributions in* GL_2*-Extensions: Distribution of Frobenius Automorphisms in* GL_2*-Extensions of the Rational Numbers*, Lect. Notes in Math. 504, Springer-Verlag, 1976; MR0568299 (58 #27900).

[8] G. Shimura, A reciprocity law in non-solvable extensions, *J. Reine Angew. Math.* 221 (1966) 209–220; MR0188198 (32 #5637).

[9] N. D. Elkies, The existence of infinitely many supersingular primes for every elliptic curve over $\mathbb{Q}$, *Invent. Math.* 89 (1987) 561–567; MR0903384 (88i:11034).

[10] J. H. Silverman, *A Friendly Introduction to Number Theory*, 3rd ed., 2006, Prentice-Hall, pp. 403–404.

[11] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002070, A006571, A006962.

[12] K. James and V. Kumar Murty, Frobenius distributions and Galois representations: Numerical evidence, unpublished manuscript (2005).

[13] S. Finch, Cyclic group orders, *this volume*, §1.18.

[14] S. Finch, Powers of Euler's q-series, math.NT/0701251.

1.20 Series involving Arithmetic Functions

We intend here to collect infinite series, each involving unusual combinations or variations of well-known arithmetic functions. For simplicity's sake, results are often quoted not with full generality but only to illustrate a special case.

Let $\sigma(n)$ denote the sum of all distinct divisors of n, $\kappa(n)$ denote the quotient of n with its greatest square divisor, and $\varphi(n)$ denote the number of positive integers $k \le n$ satisfying $\gcd(k,n)=1$. These multiplicative functions are called sum-of-divisors, square-free part, and Euler totient, respectively. It can be shown that the following series are convergent:

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{\sigma(n)\varphi(n)} &= \prod_p\left(1+\sum_{r=1}^{\infty}\frac{1}{p^{r-1}(p^{r+1}-1)}\right)\\ &= 1.7865764593...,\end{aligned}$$

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{\kappa(n)\varphi(n)} &= \prod_p\left(1+\frac{2p}{(p-1)(p^2-1)}\right)\\ &= \frac{\pi^2}{6}\prod_p\left(1+\frac{p+1}{p^2(p-1)}\right)\\ &= 3.9655568689... = A\end{aligned}$$

where the product is over all primes p. The former was considered by Silverman [1] while studying the number of generators possessing large order in the group $\mathbb{Z}_j^*$. With regard to the latter, more precise asymptotics can be given [2]:

$$\begin{aligned}\sum_{n\le N}\frac{1}{\kappa(n)\varphi(n)} &\sim A-\prod_p\left(1+\frac{\sqrt{p}+1}{p(p-1)}\right)\cdot\frac{1}{\sqrt{N}}\\ &\sim A-\prod_p\left(1+\frac{1}{p(\sqrt{p}-1)}\right)\cdot\frac{1}{\sqrt{N}}\\ &\sim A-\frac{4.9478356259...}{\sqrt{N}}.\end{aligned}$$

Let $d(n)$ denote the number of distinct divisors of n, and $\omega(n)$ denote the number of distinct prime factors of n. The divisor function $d(n)$ is multiplicative; by contrast, $\omega(n)$ is additive. It can be shown that [3, 4]

$$\sum_{n\le N} d(n)\omega(n) \sim 2N\ln(N)\ln(\ln(N)) + 2BN\ln(N)$$

where

$$\begin{aligned}B &= -\Gamma'(2)+\sum_p\left(\ln\left(1-\frac{1}{p}\right)+\frac{1}{2}\left(1-\frac{1}{p}\right)^2\sum_{k=1}^{\infty}\frac{k+1}{p^k}\right)\\ &= -(1-\gamma)+\sum_p\left(\ln\left(1-\frac{1}{p}\right)+\frac{1}{p}-\frac{1}{2p^2}\right)\\ &= M-1-\frac{1}{2}\sum_p\frac{1}{p^2} = -0.9646264971...\end{aligned}$$

where M is the Meissel–Mertens constant [5] and γ is the Euler–Mascheroni constant [6].

The mean of distinct divisors of n is clearly $\sigma(n)/d(n)$. It can be shown that [7, 8]

$$\sum_{n\le N}\frac{\sigma(n)}{d(n)}\sim\frac{C}{2\sqrt{\pi}}\frac{N^2}{\sqrt{\ln(N)}},\quad \#\left\{n:\frac{\sigma(n)}{d(n)}\le x\right\}\sim D\,x\ln(x)$$

where

$$\begin{aligned}C&=\prod_p\left(1+\sum_{k=1}^{\infty}\frac{1}{k+1}\left(\sum_{j=0}^{k}\frac{1}{p^j}\right)\frac{1}{p^k}\right)\left(1-\frac{1}{p}\right)^{1/2}\\&=\prod_p\left(1+\frac{1}{p-1}\sum_{k=1}^{\infty}\frac{1}{k+1}\frac{p^{k+1}-1}{p^{2k}}\right)\left(1-\frac{1}{p}\right)^{1/2}\\&=\prod_p\left(1-\frac{1}{p}\right)^{-1/2}p\ln\left(1+\frac{1}{p}\right)=1.2651951601...\\&=(0.7138099304...)\sqrt{\pi}=2(0.3569049652...)\sqrt{\pi},\end{aligned}$$

$$\begin{aligned}D&=\prod_p\left(1+\sum_{k=1}^{\infty}(k+1)\left(\sum_{j=0}^{k}p^j\right)^{-1}\right)\left(1-\frac{1}{p}\right)^2\\&=\prod_p\left(1+(p-1)\sum_{k=1}^{\infty}(k+1)\frac{1}{p^{k+1}-1}\right)\left(1-\frac{1}{p}\right)^2\\&=0.4950461958....\end{aligned}$$

A related series

$$\sum_{n\le N}\frac{\sigma(n)}{\varphi(n)}\sim(3.6174...)N$$

appears without comment in [9].

The lag-one autocorrelation of $d(n)$ is evident via [10]

$$\sum_{n\le N}d(n)d(n+1)\sim\frac{6}{\pi^2}N\ln(N)^2;$$

a variation of this includes [11]

$$\sum_{n\le N}d(n)^2d(n+1)\sim\frac{1}{\pi^2}\prod_p\left(1-\frac{1}{p}+\frac{1}{p}\left(1-\frac{1}{p}\right)^2\left(1+\frac{1}{p}\right)^{-1}\right)N\ln(N)^4.$$

Let $r(n)$ denote the number of representations of n as a sum of two squares, counting order and sign (note that $r(n)/4$ is multiplicative). We have [12]

$$\sum_{n\leq N} r(n)^2 d(n+1) \sim 6\prod_p \left(1-\frac{1}{p}+\frac{1}{p}\left(1-\frac{\chi(p)}{p}\right)^2\left(1+\frac{1}{p}\right)^{-1}\right) N\ln(N)^2$$

where $\chi(k)=(-4/k)$ is 0 when k is even and $(-1)^{(k-1)/2}$ when k is odd. Also, if $\tau(n)$ denotes the Ramanujan tau function [13], then [14–16]

$$\sum_{n\leq N} \tau(n)^2 d(n+1) \sim \prod_p \left(1-\frac{1}{p}+\frac{p^2-2p\cos(2\theta_p)+1}{p^2(p+1)}\right) N^{12}\ln(N)^2$$

where $2\cos(\theta_p)=\tau(p)p^{-11/2}$. Other autocorrelation results include [10]

$$\sum_{n\leq N} \sigma(n)\sigma(n+1) \sim \frac{5}{6}N^3,$$

$$\sum_{n\leq N} \varphi(n)\varphi(n+1) \sim \frac{1}{3}\prod_p\left(1-\frac{2}{p^2}\right)N^3 = \frac{0.3226340989...}{3}N^3$$

and the latter product is known as the Feller–Tornier constant [17]. The following series [18]

$$\begin{aligned}\sum_{n\leq N} \frac{d(n)}{d(n+1)} &\sim \frac{1}{\sqrt{\pi}}\prod_p\left(\frac{1}{\sqrt{p(p-1)}}+\sqrt{1-\frac{1}{p}}(p-1)\ln\left(\frac{p}{p-1}\right)\right)\cdot N\sqrt{\ln(N)}\\ &= (0.7578277106...)N\sqrt{\ln(N)}\end{aligned}$$

has a constant similar to that appearing in [19] for $\sum_{n\leq N} 1/d(n)$.

Logarithms of arithmetic functions provide some interesting constants [20–24]:

$$\frac{1}{\ln(2)}\sum_{n\leq N}\ln(d(n)) \sim N\ln(\ln(N)) + E_1 N,$$

$$\sum_{n\leq N}\ln(\varphi(n)) \sim N\ln(N)+E_2N, \quad \sum_{n\leq N}\ln(\sigma(n)) \sim N\ln(N)+E_3N,$$

$$\sum_{n\leq N}{}' \frac{\ln(\varphi(n))}{\ln(\sigma(n))} \sim N + E_4\frac{N}{\ln(N)},$$

$$\ln(2)\sum_{n\leq N}{}' \frac{1}{\ln(d(n))} \sim \frac{N}{\ln(\ln(N))}+E_5\frac{N}{\ln(\ln(N))^2},$$

$$\sum_{n\leq N}{}' \frac{1}{\ln(\varphi(n))} \sim \frac{N}{\ln(N)}+E_6\frac{N}{\ln(N)^2}, \quad \sum_{n\leq N}{}' \frac{1}{\ln(\sigma(n))} \sim \frac{N}{\ln(N)}+E_7\frac{N}{\ln(N)^2}$$

where [25–29]

$$\begin{aligned} E_1 &= \gamma + \sum_{k=2}^{\infty} \left(\frac{1}{\ln(2)} \ln\left(1+\frac{1}{k}\right) - \frac{1}{k} \right) \sum_p \frac{1}{p^k} \\ &= M + \frac{1}{\ln(2)} \sum_{k=2}^{\infty} \ln\left(1+\frac{1}{k}\right) \sum_p \frac{1}{p^k} = 0.6394076513..., \end{aligned}$$

$$\begin{aligned} E_2 &= -1 + \sum_p \frac{1}{p} \ln\left(1-\frac{1}{p}\right) = -1 + \ln(0.5598656169......) \\ &= -1.5800584938..., \end{aligned}$$

$$\begin{aligned} E_3 &= -1 + \sum_p \left(1-\frac{1}{p}\right) \sum_{k=1}^{\infty} \frac{1}{p^k} \ln\left(\frac{p^{k+1}-1}{p^k(p-1)}\right) \\ &= -1 + 0.4457089175... = 0.5542910824..., \end{aligned}$$

$$E_4 = \sum_p \left(1-\frac{1}{p}\right) \sum_{k=1}^{\infty} \left(2\ln\left(1-\frac{1}{p}\right) - \ln\left(1-\frac{1}{p^{k+1}}\right)\right) \frac{1}{p^k},$$

$E_5 = 1 - E_1$, $E_6 = -E_2$, $E_7 = -E_3$ (a sign error in [25] has been corrected to give E_5) and $\sum'$ is interpreted as summation over all n avoiding division by zero. The constant $\exp(1 + E_2)$ appeared in [30] as well.

Let $a(n)$ denote the number of non-isomorphic abelian groups of order n and $P(k)$ denote the number of unrestricted partitions of k. It can be shown that [31, 32]

$$\sum_{n\le N}' \frac{1}{\ln(a(n))} = N \int_{-\infty}^{0} \left(\prod_p \left(1 + \sum_{k=2}^{\infty} \frac{P(k)^t - P(k-1)^t}{p^k}\right) - \frac{6}{\pi^2} \right) dt.$$

Let $s(n)$ denote the number of non-isomorphic semisimple rings of order n and $Q(k)$ denote the number of unordered sets of integer pairs (r_j, m_j) for which $k = \sum_j r_j m_j^2$ and $r_j m_j^2 > 0$ for all j. Likewise, we have

$$\sum_{n\le N}' \frac{1}{\ln(s(n))} = N \int_{-\infty}^{0} \left(\prod_p \left(1 + \sum_{k=2}^{\infty} \frac{Q(k)^t - Q(k-1)^t}{p^k}\right) - \frac{6}{\pi^2} \right) dt.$$

If $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_r^{\alpha_r}$ is the prime factorization of n, define three additive functions

$$\beta(n) = \sum_{j=1}^{r} p_j, \quad B(n) = \sum_{j=1}^{r} \alpha_j p_j, \quad \hat{B}(n) = \sum_{j=1}^{r} p_j^{\alpha_j},$$

the first two of which contrast nicely with the better-known functions

$$\omega(n) = \sum_{j=1}^{r} 1, \quad \Omega(n) = \sum_{j=1}^{r} \alpha_j.$$

While [5]

$$\frac{1}{N}\sum_{n\le N}\omega(n) \sim \ln(\ln(N)) + M, \quad \frac{1}{N}\sum_{n\le N}\Omega(n) \sim \ln(\ln(N)) + M + \sum_{p}\frac{1}{p(p-1)}$$

we have [33–35]

$$\sum_{n\le N}\beta(n) \sim \sum_{n\le N}B(n) \sim \sum_{n\le N}\hat{B}(n) \sim \frac{\pi^2}{12}\frac{N^2}{\ln(N)}.$$

While [36, 37]

$$\begin{aligned}\sum_{n\le N}{}' \frac{1}{\Omega(n)-\omega(n)} &\sim N\int_0^1 \left(\prod_p \left(1+\sum_{k=2}^{\infty}\frac{t^{k-1}-t^{k-2}}{p^k}\right) - \frac{6}{\pi^2}\right)\frac{1}{t}dt \\ &\sim N\int_0^1 \left(\prod_p \left(1-\frac{1}{p}\right)\left(1-\frac{1}{t-p}\right) - \frac{6}{\pi^2}\right)\frac{1}{t}dt,\end{aligned}$$

we have [38, 39]

$$\begin{aligned}\sum_{n\le N}{}' \frac{1}{B(n)-\beta(n)} &\sim N\int_0^1 \left(\prod_p \left(1+\sum_{k=2}^{\infty}\frac{t^{(k-1)p}-t^{(k-2)p}}{p^k}\right) - \frac{6}{\pi^2}\right)\frac{1}{t}dt \\ &\sim N\int_0^1 \left(\prod_p \left(1-\frac{1}{p}\right)\left(1-\frac{1}{t^p-p}\right) - \frac{6}{\pi^2}\right)\frac{1}{t}dt.\end{aligned}$$

We also have [38, 40, 41],

$$\sum_{n\le N}{}' \frac{\Omega(n)}{\omega(n)} \sim \sum_{n\le N}{}' \frac{B(n)}{\beta(n)} \sim N,$$

$$\sum_{n\le N}{}' \frac{\hat{B}(n)}{\beta(n)} \sim e^{\gamma}N\ln(\ln(N)), \quad \sum_{n\le N}{}' \frac{\hat{B}(n)}{B(n)} \sim FN$$

where

$$F = \int_1^{\infty}\frac{1}{x}\sum_{j=0}^{\lfloor x\rfloor-1}\frac{\rho(x-\lfloor x\rfloor+j)}{\lfloor x\rfloor-j}dx = \sum_{k=1}^{\infty}\frac{1}{k}\int_0^{\infty}\frac{\rho(y)}{y+k}dy$$

and $\rho(z)$ is Dickman's function [42].

Other constants emerge when arithmetic functions are evaluated not at n, but at quadratic functions of n. For example [23, 43–48],

$$\sum_{n\le N} d(n^2+1) \sim \frac{3}{\pi} N\ln(N), \quad \sum_{n\le N} \sigma(n^2+1) \sim \frac{5G}{\pi^2} N^3,$$

$$\sum_{n\le N} r(n^2+1) \sim \frac{8}{\pi} N\ln(N), \quad \sum_{n\le N} \varphi(n^2+1) \sim \frac{H}{4} N^3$$

where G is Catalan's constant [49] and

$$H = \prod_{\substack{p\equiv 1 \\ \bmod 4}} \left(1 - \frac{2}{p^2}\right) = 0.8948412245...$$

is a modified Feller–Tornier constant that appeared in [50]. As another example [51–54],

$$\sum_{m,n\le N} d(m^2+n^2) \sim \frac{\pi}{2G} N^2 \ln(N), \quad \sum_{m,n\le N} \sigma(m^2+n^2) \sim I\,N^4$$

where

$$\begin{aligned} I &= \frac{2}{3}\sum_{j=1}^{\infty} \frac{\nu(j)}{j^3} \\ &= \frac{8}{9} \prod_{\substack{p\equiv 1 \\ \bmod 4}} \left(1 + \frac{2p+1}{(p+1)(p^2-1)}\right) \prod_{\substack{p\equiv 3 \\ \bmod 4}} \left(1 + \frac{1}{(p-1)(p^2+1)}\right) \\ &= 1.03666099... \end{aligned}$$

and $\nu(j)$ denotes the number of solutions of $x^2+y^2=0$ in $\mathbb{Z}_j$, counting order [55, 56].

The average prime factor of n may reasonably be defined in two ways: as an mean of distinct prime factors $\beta(n)/\omega(n)$ or as a mean of all prime factors $B(n)/\Omega(n)$ (with multiplicity). It can be shown that [57]

$$\sum_{n\le N} \frac{\beta(n)}{\omega(n)} \sim J \frac{N^2}{\ln(N)}, \quad \sum_{n\le N} \frac{B(n)}{\Omega(n)} \sim K \frac{N^2}{\ln(N)}$$

for constants $0 < K < J$. Infinite product expressions for J, K are possible but remain undiscovered (as far as is known).

The distance between consecutive distinct prime factors of $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_r^{\alpha_r}$ can be quantified in many ways: for example [58],

$$\frac{1}{r-1}\sum_{j=2}^{r}(p_j - p_{j-1}) = \frac{P^+(n) - P^-(n)}{\omega(n)-1}$$

(whose sum over $n \le N$ is $\sim \lambda N^2/\ln(N)$, where $2\lambda = \sum_{k=2}^{\infty} k^{-2}\omega(k)^{-1} = 0.59737...$) and

$$g(n) = \sum_{j=2}^{r} \frac{1}{p_j - p_{j-1}}$$

(which is comparatively artificial). Of course, $g(1) = 0 = g(p)$ for any prime p by the empty sum convention. It can be shown that [59]

$$\begin{aligned} \sum_{n \le N} g(n) &\sim N \sum_{p_L < p_R} \frac{1}{(p_R - p_L)p_L p_R} \prod_{p_L < p < p_R} \left(1 - \frac{1}{p}\right) \\ &\sim (0.299...)N \end{aligned}$$

where the sum is taken over all pairs of primes $p_L < p_R$ and the product is taken over all primes p strictly between the left prime p_L and the right prime p_R. If no such p exists, then the product is 1 by the empty product convention.

If $1 = \delta_1 < \delta_2 < \ldots < \delta_s = n$ are the consecutive distinct divisors of n, we might examine

$$\frac{1}{s-1} \sum_{j=2}^{s} (\delta_j - \delta_{j-1}) = \frac{n-1}{d(n)-1}$$

(whose sum over $n \le N$ is $\sim \mu N^2/\ln(N)^{1/2}$; the formula for $2\mu = (0.96927...)\pi^{-1/2}$ appears in [19, 20]) and

$$h(n) = \sum_{j=2}^{s} \frac{1}{\delta_j - \delta_{j-1}}.$$

If two positive integers $a < b$ are consecutive divisors of $c_{a,b} = \mathrm{lcm}(a, b)$, let

$$\Delta_{a,b} = \left\{ \frac{d}{\gcd(d, c_{a,b})} : a < d < b \right\}$$

and let $D_{a,b}$ be the largest subset of $\Delta_{a,b}$ such that no element of $D_{a,b}$ is a multiple of another element in $D_{a,b}$. (Clearly $1 \notin \Delta_{a,b}$.) Assuming $D_{a,b} = \{d_1, d_2, \ldots, d_t\}$, we denote by $T(a, b)$ the following expression:

$$1 - \sum_{1 \le i \le t} \tfrac{1}{d_i} + \sum_{1 \le i < j \le t} \tfrac{1}{\mathrm{lcm}(d_i, d_j)} - \sum_{1 \le i < j < k \le t} \tfrac{1}{\mathrm{lcm}(d_i, d_j, d_k)} + \cdots + (-1)^t \tfrac{1}{\mathrm{lcm}(d_1, d_2, \ldots, d_t)}.$$

It can be shown that [59]

$$\begin{aligned} \sum_{n \le N} h(n) &\sim N \sum_{a < b} \frac{1}{c_{a,b}(b-a)} T(a, b) \\ &\sim (1.77...)N \end{aligned}$$

where the sum is taken over all pairs of positive integers $a < b$ such that the consecutive divisor requirement is met by a, b.

1.20.1 *Subgroups of* $\mathbb{Z}_m \times \mathbb{Z}_n$

Let $\mathbb{Z}_m$ be the additive group of residue classes modulo m. The number of subgroups of $\mathbb{Z}_m$ is $d(m)$ and each subgroup is cyclic. The number $s(m,n)$ of subgroups of $\mathbb{Z}_m \times \mathbb{Z}_n$ satisfies [60–63]

$$s(m,n) = \sum_{a|m,b|n} \gcd(a,b),$$

$$\sum_{m,n\le x} s(m,n) \sim x^2 \left(A_3 \ln(x)^3 + A_2 \ln(x)^2 + A_1 \ln(x) + A_0\right)$$

where

$$A_3 = \frac{1}{3\zeta(2)} = \frac{2}{\pi^2}, \quad A_2 = \frac{1}{\zeta(2)}\left(3\gamma - 1 - \frac{\zeta'(2)}{\zeta(2)}\right),$$

$$A_1 = \frac{1}{\zeta(2)}\left(8\gamma^2 - 6\gamma - 2\gamma_1 + 1 - 2(3\gamma-1)\frac{\zeta'(2)}{\zeta(2)} + 2\left(\frac{\zeta'(2)}{\zeta(2)}\right)^2 - \frac{\zeta''(2)}{\zeta(2)}\right)$$

and the number $c(m,n)$ of cyclic subgroups of $\mathbb{Z}_m \times \mathbb{Z}_n$ satisfies

$$c(m,n) = \sum_{\substack{a|m,b|n,\\ \gcd\left(\frac{m}{a},\frac{n}{b}\right)=1}} \gcd(a,b),$$

$$\sum_{m,n\le x} c(m,n) \sim x^2 \left(B_3 \ln(x)^3 + B_2 \ln(x)^2 + B_1 \ln(x) + B_0\right)$$

where

$$B_3 = \frac{1}{3\zeta(2)^2} = \frac{12}{\pi^4}, \quad B_2 = \frac{1}{\zeta(2)^2}\left(3\gamma - 1 - 2\frac{\zeta'(2)}{\zeta(2)}\right),$$

$$B_1 = \frac{1}{\zeta(2)^2}\left(8\gamma^2 - 6\gamma - 2\gamma_1 + 1 - 4(3\gamma-1)\frac{\zeta'(2)}{\zeta(2)} + 6\left(\frac{\zeta'(2)}{\zeta(2)}\right)^2 - 2\frac{\zeta''(2)}{\zeta(2)}\right).$$

The expressions for A_0, B_0 are complicated and not helpful for numerical evaluation; γ_1 is the first Stieltjes constant [64]. In particular,

$$\sum_{n\le x} s(n,n) \sim \frac{5\pi^2}{24}x^2, \quad \sum_{n\le x} c(n,n) \sim \frac{5}{4}x^2;$$

analogously,

$$\sum_{n\le x} s(n,n,n) \sim \frac{1}{3}x^3 \left[H(3)\left(\ln(x) + 2\gamma - 1\right) + H'(3)\right]$$

where

$$H(z) = \zeta^2(z) \prod_p \left(1 + \frac{2}{p^{z-1}} + \frac{2}{p^z} + \frac{1}{p^{2z-1}}\right), \quad \mathrm{Re}(z) > 2.$$

Of related interest are series $\sum_{n\le x} t(n)$ and $\sum_{m,n\le x} t(m\,n)$, where $t(n)$ is the number of squares dividing n [65, 66]. More examples appear in [67, 68]; cases when the underlying Dirichlet series is a product of zeta function expressions give rise to asymptotic expansions with exact coefficients (found via residues).

1.20.2 *Dedekind Totient Constants*

The Dedekind totient ψ enjoys close parallels [69, 70] with the Euler totient φ:

$$\psi(n)=n\prod_{p|n}\left(1+\frac{1}{p}\right),\quad \varphi(n)=n\prod_{p|n}\left(1-\frac{1}{p}\right);$$

$$\sum_{n\le N}\psi(n)\sim\underbrace{\frac{1}{2}\prod_{p}\left(1+\frac{1}{p^2}\right)}_{15/(2\pi^2)}\cdot N^2,\quad \sum_{n\le N}\varphi(n)\sim\underbrace{\frac{1}{2}\prod_{p}\left(1-\frac{1}{p^2}\right)}_{3/\pi^2}\cdot N^2;$$

$$\sum_{n\le N}\frac{1}{\psi(n)}\sim\underbrace{\prod_{p}\left(1-\frac{1}{p(p+1)}\right)}_{C_{\text{carefree}}}\cdot\left(\ln(N)+\gamma+\sum_{p}\frac{\ln(p)}{p^2+p-1}\right),$$

$$\sum_{n\le N}\frac{1}{\varphi(n)}\sim\underbrace{\prod_{p}\left(1+\frac{1}{p(p-1)}\right)}_{315\zeta(3)/(2\pi^4)}\cdot\left(\ln(N)+\gamma-\sum_{p}\frac{\ln(p)}{p^2-p+1}\right).$$

Further results include [71]

$$\sum_{n\le N}\frac{\varphi(n)}{\psi(n)}\sim\prod_{p}\left(1-\frac{2}{p(p+1)}\right)\cdot N,$$

$$\sum_{n\le N}\psi(n)^2\sim\tfrac{1}{3}\prod_{p}\left(1+\tfrac{2}{p^2}+\tfrac{1}{p^3}\right)\cdot N^3,\quad \sum_{n\le N}\varphi(n)^2\sim\tfrac{1}{3}\prod_{p}\left(1-\tfrac{2}{p^2}+\tfrac{1}{p^3}\right)\cdot N^3.$$

The first of the three products appears in [72] with regard to cube roots of nullity mod n, and in [73] with regard to strongly carefree couples. Asymptotics for $\sum_{n\le N}\varphi(n)^{\ell}$ were found by Chowla [74], where ℓ is any positive integer. His formula naturally carries over to $\sum_{n\le N}\psi(n)^{\ell}$. It is known that the Riemann Hypothesis is true if and only if [75, 76]

$$\varphi\left(\prod_{k=1}^{n}p_k\right)<e^{-\gamma}\left(\prod_{k=1}^{n}p_k\right)/\ln\left(\ln\left(\prod_{k=1}^{n}p_k\right)\right),$$

$$\psi\left(\prod_{k=1}^{n}p_k\right)>\frac{6\,e^{\gamma}}{\pi^2}\left(\prod_{k=1}^{n}p_k\right)\cdot\ln\left(\ln\left(\prod_{k=1}^{n}p_k\right)\right)$$

for all $n\ge 3$, where $p_1=2,\ p_2=3,\ p_3=5,\ \ldots$ is the sequence of all primes. A related inequality, due to Robin, appears in [77]. Alternating series analogs, too numerous to include here, are found in [70].

1.20.3 Extreme Prime Factors

Let $P^+(n)$ denote the largest prime factor of n and $P^-(n)$ denote the smallest prime factor of n. Also let $P^+(1) = P^-(1) = 1$. It follows that [78–84]

$$\sum_{n\le N} P^+(n) \sim \frac{\pi^2}{12}\frac{N^2}{\ln(N)}, \quad \sum_{n\le N} P^-(n) \sim \frac{1}{2}\frac{N^2}{\ln(N)},$$

$$\sum_{n\le N} \frac{P^+(n)}{P^-(n)} \sim \frac{\pi^2}{12}\sum_p \left(\frac{1}{p^3}\prod_{q<p}\left(1-\frac{1}{q^2}\right)\right)\cdot\frac{N^2}{\ln(N)},$$

$$\sum_{n\le N} \frac{1}{P^+(n)} \sim N\int_2^N \rho\left(\frac{\ln(N)}{\ln(t)}\right)\frac{1}{t^2}dt,$$

$$\sum_{n\le N} \frac{P^-(n)}{P^+(n)} \sim \frac{N}{\ln(N)}, \quad \sum_{n\le N} \frac{1}{P^-(n)} \sim UN,$$

$$\sum_{n\le N} \frac{d(n)}{P^-(n)} \sim VN\ln(N), \quad \sum_{n\le N} \frac{\Omega(n)-\omega(n)}{P^-(n)} \sim WN,$$

$$\sum_{n\le N} \frac{\varphi(n)}{P^-(n)} \sim XN^2, \quad \sum_{n\le N} \frac{1}{n\ln(P^-(n))} \sim Y\ln(N)$$

where

$$U = \sum_p \frac{f(p)}{p^2}, \quad V = \sum_p \frac{(2p-1)f(p)^2}{p^3},$$

$$W = \sum_p \frac{f(p)}{p}\sum_{\alpha\ge 2}\frac{1}{p^\alpha} + \sum_p \frac{f(p)}{p^2}\sum_{q>p}\sum_{\alpha\ge 2}\frac{1}{q^\alpha},$$

$$X = \frac{3}{\pi^2}\sum_p \frac{1}{p\,(p+1)\tilde{f}(p)}, \quad Y = \sum_p \frac{f(p)}{p\ln(p)},$$

p and q are primes (of course), and

$$f(k) = \begin{cases} 1 & \text{if } k=2, \\ \prod_{p<k}\left(1-\frac{1}{p}\right) & \text{if } k>2, \end{cases} \quad \tilde{f}(k) = \begin{cases} 1 & \text{if } k=2, \\ \prod_{p<k}\left(1+\frac{1}{p}\right) & \text{if } k>2. \end{cases}$$

Mertens' formula implies that $\lim_{k\to\infty}\ln(k)f(k) = e^{-\gamma}$ and $\lim_{k\to\infty}\tilde{f}(k)/\ln(k) = 6\pi^{-2}e^{\gamma}$.

Variations of [85]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{\ln(n)}{\ln(P^+(n))} = e^{\gamma}$$

include [79]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n\le N}\frac{\ln(P^+(n))}{\ln(n)}=\Lambda=\lim_{N\to\infty}\frac{1}{N\ln(N)}\sum_{n\le N}\ln(P^+(n))$$

where $\Lambda=0.6243299885...$ is the Golomb–Dickman constant [42]. A simple, precise estimate of

$$\sum_{n\le N}\frac{1}{\ln(P^+(n))}$$

evidently has not yet been found.

Let $k(n)$ denote the smallest prime not dividing n and $\ell(n)$ denote the smallest integer >1 not dividing n. Their respective average values are [86–88]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{3\le n\le N}k(n)=\sum_{p}(p-1)/\prod_{q<p}q=2.9200509773...,$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{3\le n\le N}\ell(n)=\sum_{j\ge 2}\left(\frac{1}{\operatorname{lcm}\{1,2,\ldots,j-1\}}-\frac{1}{\operatorname{lcm}\{1,2,\ldots,j\}}\right)j=2.7877804561....$$

Compare these to the quadratic nonresidue constants at the end of [5].

Let $P_2^+(n)$ denote the second largest prime factor of n if it exists, otherwise set $P_2^+(n)=\infty$. The asymptotic behavior of $P_2^+(n)$ is completely different from that of $P^+(n)$ [89, 90]:

$$\begin{aligned}\sum_{n\le N}\frac{1}{P_2^+(n)}&\sim\left(\sum_{m=1}^{\infty}\frac{1}{m}\sum_{p\ge P^+(m)}\frac{1}{p^2}\right)\frac{N}{\ln(N)}\\&\sim\left(\sum_{p}\frac{1}{p^2}\prod_{q\le p}\left(1-\frac{1}{q}\right)^{-1}\right)\frac{N}{\ln(N)}\sim(1.254...)\,\frac{N}{\ln(N)}.\end{aligned}$$

Let $P_3^+(n)$ denote the third largest prime factor of n if it exists, otherwise set $P_3^+(n)=\infty$. Interestingly, the same constant occurs [89, 90]:

$$\sum_{n\le N}\frac{1}{P_3^+(n)}\sim(1.254...)\frac{N\ln(\ln(N))}{\ln(N)}$$

but the growth rate is faster. A well-known constant $\sum 1/p^2=0.4522474200...$ from [5] appears in [91], stemming (almost surely) from the reciprocal sum of a uniformly drawn prime factor of n, for each n. The growth rate $N/\ln(\ln(N))$ is faster still.

Here is a comparatively neglected topic: for a random integer n between 1 and N, since

$$\lim_{N\to\infty}\mathrm{P}\left(P^+(n)\le n^x\right)=\rho\left(\frac{1}{x}\right)$$

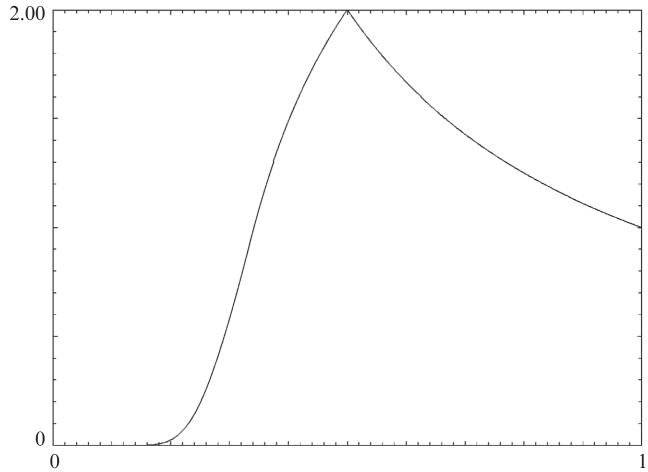

Figure 1.18 Plot of $d/dx\, \rho_1(1/x)$ when $0 < x < 1$; given random n, the density for x such that n^x is the largest prime factor of n. Image courtesy of David Broadhurst.

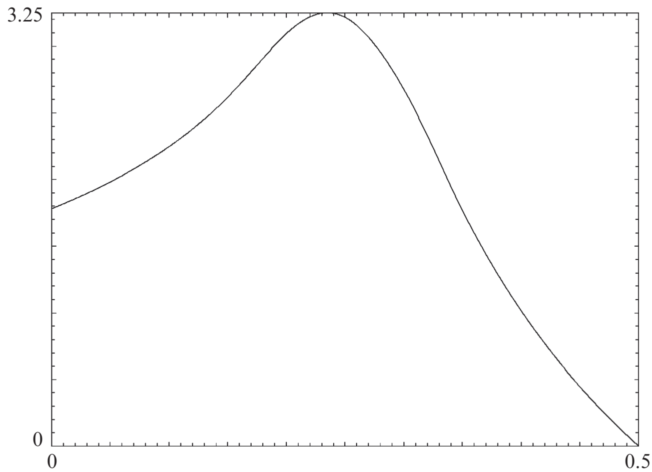

Figure 1.19 Plot of $d/dx\, \rho_2(1/x)$ when $0 < x < 1/2$; given random n, the density for x such that n^x is the second-largest prime factor of n. Image courtesy of David Broadhurst.

for $0 < x \leq 1$, the median value of x satisfies $\rho(1/x) = 1/2$, that is, $x = 1/\sqrt{e} = 0.6065306597....$ The mode (peak of density) is $1/2$; see Figure 1.18. Define the second-order Dickman function $\rho_2(x)$ by [89]

$$x\rho_2'(x) + \rho_2(x-1) = \rho(x-1) \text{ for } x > 1, \quad \rho_2(x) = 1 \text{ for } 0 \leq x \leq 1$$

then the corresponding median value satisfies $\rho_2(1/x) = 1/2$, that is, $x = 0.2117211464...$ [92]. An early approximation (0.24) appeared long ago [93]; medians are more robust estimators of centrality than means (being less sensitive to data outliers). The mode here is 0.2350396459...; see Figure 1.19. Likewise, the third-order Dickman function $\rho_3(x)$ is [89]

$$x\rho_3'(x) + \rho_3(x-1) = \rho_2(x-1) \text{ for } x > 1, \quad \rho_3(x) = 1 \text{ for } 0 \le x \le 1$$

and the corresponding median value satisfies $\rho_3(1/x) = 1/2$, that is, $x = 0.0758437231...$ [92]. A certain family of multiple integrals related to $\rho(x)$ is investigated in [94–96].

[1] P. Zimmermann, A peculiar sum, USENET sci.math.research newsgroup posting (1996).
[2] C. David and F. Pappalardi, Average Frobenius distributions of elliptic curves, *Internat. Math. Res. Notices* (1999) 165–183; MR1677267 (2000g:11045).
[3] J.-M. De Koninck and A. Mercier, Remarque sur un article "Identities for series of the type $\Sigma f(n)\mu(n)n^{-s}$" de T. M. Apostol, *Canad. Math. Bull.* 20 (1977) 77–88; MR0472733 (57 #12425).
[4] J.-M. De Koninck and A. Ivić, *Topics in Arithmetical Functions: Asymptotic Formulae for Sums of Reciprocals of Arithmetical Functions and Related Fields*, North-Holland, 1980, pp. 233–235; MR0589545 (82a:10047).
[5] S. R. Finch, Meissel-Merten constant, *first volume*, pp. 94–98.
[6] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[7] P. T. Bateman, P. Erdős, C. Pomerance and E. G. Straus, The arithmetic mean of the divisors of an integer, *Analytic Number Theory*, Proc. 1980 Philadelphia conf., ed. M. I. Knopp, Lect. Notes in Math. 899, Springer-Verlag, 1981, pp. 197–220; MR0654528 (84b:10066).
[8] M. Mazur and B. V. Petrenko, Representations of analytic functions as infinite products and their application to numerical computations, *Ramanujan J.* 34 (2014) 129–141; arXiv:1202.1335; MR3210260.
[9] L. G. Fel, Summatory multiplicative arithmetic functions: scaling and renormalization, arXiv:1108.0957.
[10] A. E. Ingham, Some asymptotic formulae in the theory of numbers, *J. London Math. Soc.* 2 (1927) 202–208.
[11] Y. Motohashi, An asymptotic formula in the theory of numbers, *Acta Arith.* 16 (1969/70) 255–264; MR0266884 (42 #1786).
[12] K.-H. Indlekofer, Eine asymptotische Formel in der Zahlentheorie, *Arch. Math. (Basel)* 23 (1972) 619–624; MR0318080 (47 #6629).
[13] S. R. Finch, Modular forms on $SL_2(\mathbb{Z})$, *this volume*, §1.15.
[14] D. Redmond, An asymptotic formula in the theory of numbers, *Math. Annalen* 224 (1976) 247–268; MR0419386 (54 #7407).
[15] D. Redmond, An asymptotic formula in the theory of numbers. II, *Math. Annalen* 234 (1978) 221–238; MR0480387 (58 #553).
[16] D. Redmond, An asymptotic formula in the theory of numbers. III, *Math. Annalen* 243 (1979) 143–151; MR0543724 (80h:10052).
[17] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.
[18] M. A. Korolev, On Karatsuba's problem concerning the divisor function, *Monatsh. Math.* 168 (2012) 403–441; arXiv:1011.1391; MR2993957.
[19] S. R. Finch, Unitarism and infinitarism, *this volume*, §1.6.
[20] B. M. Wilson, Proofs of some formulae enunciated by Ramanujan, *Proc. London Math. Soc.* 21 (1923) 235–255.
[21] D. R. Ward, Some series involving Euler's function, *J. London Math. Soc.* 2 (1927) 210-214.

[22] A. Mercier, Sommes de fonctions additives restreintes à une class de congruence, *Canad. Math. Bull.* 22 (1979) 59–73; MR0532271 (81a:10009).
[23] A. G. Postnikov, *Introduction to Analytic Number Theory*, Amer. Math. Soc., 1988, pp. 192–195; MR0932727 (89a:11001).
[24] De Koninck and Ivić, *Topics,* op. cit., pp. 106, 226.
[25] J.-M. De Koninck, On a class of arithmetical functions, *Duke Math. J.* 39 (1972) 807–818; MR0311598 (47 #160).
[26] J.-M. De Koninck and J. Galambos, Sums of reciprocals of additive functions, *Acta Arith.* 25 (1973/74) 159–164; MR0354598 (50 #7076).
[27] T. Cai, On a sum of Euler's totient function (in Chinese), *J. Shandong Univ. Nat. Sci. Ed.* 24 (1989) 106-110; Zbl 0684.10044.
[28] J. Bayless and D. Klyve, On the sum of reciprocals of amicable numbers, *Integers* 11 (2011) 315–332; arXiv:1101.0259; MR2988065.
[29] P. Sebah, Calculating the aliquot constant to 20 digits accuracy, unpublished note (2013).
[30] S. R. Finch, Alladi-Grinstead constant, *first volume*, pp. 120–122.
[31] De Koninck and Ivić, *Topics,* op. cit., p. 81–88.
[32] J.-M. De Koninck and A. Ivić, An asymptotic formula for reciprocals of logarithms of certain multiplicative functions, *Canad. Math. Bull.* 21 (1978) 409–413; MR0523581 (80g:10043).
[33] K. Alladi and P. Erdős, On an additive arithmetic function, *Pacific J. Math.* 71 (1977) 275–294; MR0447086 (56 #5401).
[34] De Koninck and Ivić, *Topics,* op. cit., p. 149–151, 171–172, 246–247.
[35] T.-Z. Xuan, On some sums of large additive number-theoretic functions (in Chinese), *Beijing Shifan Daxue Xuebao* (1984), n. 2, 11–18; MR0767509 (86i:11052).
[36] De Koninck and Ivić, *Topics,* op. cit., p. 132–133, 142–143.
[37] J.-M. De Koninck and A. Ivić, Sums of reciprocals of certain additive functions, *Manuscripta Math.* 30 (1979/80) 329–341; MR0567210 (81g:10061).
[38] J.-M. De Koninck, P. Erdős and A. Ivić, Reciprocals of certain large additive functions, *Canad. Math. Bull.* 24 (1981) 225–231; MR0619450 (82k:10053).
[39] De Koninck and Ivić, *Topics,* op. cit., p. 164–166.
[40] P. Erdős and A. Ivić, Estimates for sums involving the largest prime factor of an integer and certain related additive functions, *Studia Sci. Math. Hungar.* 15 (1980) 183–199; MR0681439 (84a:10046).
[41] T.-Z. Xuan, On a result of Erdős and Ivić, *Arch. Math. (Basel)* 62 (1994) 143–154; MR1255638 (94m:11109).
[42] S. R. Finch, Golomb-Dickman constant, *first volume*, pp. 284–292.
[43] H. N. Shapiro, *Introduction to the Theory of Numbers*, Wiley, 1983, pp. 175–185; MR0693458 (84f:10001).
[44] E. J. Scourfield, The divisors of a quadratic polynomial, *Proc. Glasgow Math. Assoc.* 5 (1961) 8–20; MR0144855 (26 #2396).
[45] C. Hooley, On the number of divisors of a quadratic polynomial, *Acta Math.* 110 (1963) 97–114; MR0153648 (27 #3610).
[46] J. McKee, On the average number of divisors of quadratic polynomials, *Math. Proc. Cambridge Philos. Soc.* 117 (1995) 389–392; MR1317484 (96e:11118).
[47] J. McKee, A note on the number of divisors of quadratic polynomials, *Sieve Methods, Exponential Sums, and their Applications in Number Theory*, ed. G. R. H. Greaves, G. Harman and M. N. Huxley, Proc. 1995 Cardiff conf., Cambridge Univ. Press, 1997, pp. 275–28; MR1635774 (99d:11106).

[48] J. McKee, The average number of divisors of an irreducible quadratic polynomial, *Math. Proc. Cambridge Philos. Soc.* 126 (1999) 17–22; MR1681650 (2000a:11053).
[49] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[50] S. R. Finch, Landau-Ramanujan constant, *first volume*, pp. 98–104.
[51] N. Gafurov, The sum of the number of divisors of a quadratic form (in Russian), *Dokl. Akad. Nauk Tadzhik. SSR* 28 (1985) 371–375; MR0819343 (87c:11038).
[52] N. Gafurov, Asymptotic formulas for the sum of powers of divisors of a quadratic form (in Russian), *Dokl. Akad. Nauk Tadzhik. SSR* 32 (1989) 427–431; MR1038632 (91c:11053).
[53] N. Gafurov, On the number of divisors of a quadratic form (in Russian), *Trudy Mat. Inst. Steklov.* 200 (1991) 124–135; Engl. transl. in *Proc. Steklov Inst. Math.* (1993), n. 2, 137–148; MR1143362 (93a:11079).
[54] G. Yu, On the number of divisors of the quadratic form $m^2 + n^2$, *Canad. Math. Bull.* 43 (2000) 239–256; MR1754029 (2001f:11162).
[55] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A086933.
[56] L. Tóth, Counting solutions of quadratic congruences in several variables revisited, *J. Integer Seq.* 17 (2014) 14.11.6; arXiv:1404.4214; MR3291084.
[57] J.-M. De Koninck and A. Ivić, The distribution of the average prime divisor of an integer, *Arch. Math. (Basel)* 43 (1984) 37–43; MR0758338 (85j:11116).
[58] J.-M. De Koninck, Sketch of proof giving a constant, unpublished note (2007).
[59] J.-M. De Koninck and A. Ivić, On the distance between consecutive divisors of an integer, *Canad. Math. Bull.* 29 (1986) 208–217; MR0844901 (87f:11074).
[60] L. Tóth, On the number of cyclic subgroups of a finite abelian group, *Bull. Math. Soc. Sci. Math. Roumanie* 55 (2012) 423–428; arXiv:1203.6201; MR2963406.
[61] M. Hampejs, N. Holighaus, L. Tóth and C. Wiesmeyr, On the subgroups of the group $\mathbb{Z}_m \times \mathbb{Z}_n$, arXiv:1211.1797.
[62] W. G. Nowak and L. Tóth, On the average number of subgroups of the group $\mathbb{Z}_m \times \mathbb{Z}_n$, arXiv:1307.1414.
[63] M. Hampejs and L. Tóth, On the subgroups of finite abelian groups of rank three, *Annales Univ. Sci. Budapest. Sect. Comput.* 39 (2013) 111–124; arXiv:1304.2961; MR3045601.
[64] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.
[65] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A046951.
[66] A. V. Lelechenko, Average number of squares dividing mn, arXiv:1407.1222.
[67] E. Krätzel, W. G. Nowak and L. Tóth, On certain arithmetic functions involving the greatest common divisor, *Central Europ. J. Math.* 10 (2012) 761–774; MR2886571.
[68] M. Kühleitner and W. G. Nowak, On a question of A. Schinzel: Omega estimates for a special type of arithmetic functions, *Central Europ. J. Math.* 11 (2013) 477–486; MR3016316.
[69] V. Sita Ramaiah and D. Suryanarayana, Sums of reciprocals of some multiplicative functions, *Math. J. Okayama Univ.* 21 (1979) 155–164; MR0554306 (82d:10064a).
[70] L. Tóth, Alternating sums concerning multiplicative arithmetic functions, *J. Integer Seq.* 20 (2017) 17.2.1; arXiv:1608.00795; MR3608534.
[71] D. Suryanarayana, On some asymptotic formulae of S. Wigert, *Indian J. Math.* 24 (1982) 81–98; MR0724328 (85d:11087).
[72] S. Finch and P. Sebah, Squares and cubes modulo n, math.NT/0604465.
[73] S. R. Finch, Hafner-Sarnak-McCurley constant, *first volume*, pp. 110–112.
[74] S. D. Chowla, An order result involving Euler's φ-function, *J. Indian Math. Soc.* 18 (1927) 138–141.

[75] J.-L. Nicolas, Petites valeurs de la fonction d'Euler, *J. Number Theory* 17 (1983) 375–388; MR0724536 (85h:11053).
[76] M. Planat and P. Solé, Extreme values of the Dedekind Ψ function, *J. Combinatorics and Number Theory* 3 (2011) 33–38; arXiv:1011.1825; MR2908180.
[77] S. R. Finch, Multiples and divisors, *this volume*, §1.5.
[78] A. E. Brouwer, Two number theoretic sums, *Afdeling Zuivere Wiskunde*, ZW 19/74 (1974); MR0345918 (49 #10647).
[79] J. van de Lune, Some sums involving the largest and smallest prime divisor of a natural number, *Afdeling Zuivere Wiskunde*, ZW 25/74 (1974).
[80] P. Erdős, A. Ivić and C. Pomerance, On sums involving reciprocals of the largest prime factor of an integer, Glasnik Mat. 21 (1986) 283–300; MR896810 (89a:11090).
[81] P. Erdős and J. H. van Lint, On the average ratio of the smallest and largest prime divisor of n, *Nederl. Akad. Wetensch. Indag. Math.* 44 (1982) 127–132; MR0662646 (83m:10075).
[82] W. P. Zhang, Average-value estimation of a class of number-theoretic functions (in Chinese), *Acta Math. Sinica* 32 (1989) 260–267; MR1025146 (90k:11124).
[83] Y. R. Zhang, Estimates for sums involving the smallest prime factor of an integer (in Chinese), *Acta Math. Sinica* 42 (1999) 997–1004; MR1756021 (2001c:11104).
[84] H.-Z. Cao, Sums involving the smallest prime factor of an integer, *Utilitas Math.* 45 (1994) 245–251; MR1284035 (95d:11126).
[85] J. B. van Rongen, On the largest prime divisor of an integer, *Nederl. Akad. Wetensch. Proc. Ser. A* 78 (1975) 70–76; *Indag. Math.* 37 (1975) 70–76; MR0376573 (51 #12748).
[86] I. Rivin, Geodesics with one self-intersection, and other stories, *Adv. Math.* 231 (2012) 2391–2412; arXiv:0901.2543; MR2970452.
[87] K. Bou-Rabee and D. B. McReynolds, Bertrand's postulate and subgroup growth, *J. Algebra* 324 (2010) 793–819; arXiv:0909.1343; MR2651569 (2011i:20035).
[88] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A053669, A007978 and A199968.
[89] J.-M. De Koninck, Sur les plus grands facteurs premiers d'un entier, *Monatsh. Math.* 116 (1993) 13–37; MR1239141 (94h:11088).
[90] J.-M. De Koninck and F. Luca, On the middle prime factor of an integer, *J. Integer Seq.* 16 (2013) 13.5.5; MR3065334.
[91] J.-M. De Koninck and J. Galambos, Some randomly selected arithmetical sums, *Acta Math. Hungar.* 52 (1988) 37–43; MR0956136 (89k:11068).
[92] D. Broadhurst, Higher-order Dickman functions, unpublished note (2014).
[93] M. C. Wunderlich and J. L. Selfridge, A design for a number theory package with an optimized trial division routine, *Commun. ACM*, v. 17 (1974) n. 5, 272–276.
[94] D. Broadhurst, Dickman polylogarithms and their constants, arXiv:1004.0519.
[95] K. Soundararajan, An asymptotic expansion related to the Dickman function, *Ramanujan J.* 29 (2012) 25–30; MR2994087.
[96] C. S. Franze, A note on an asymptotic expansion related to the Dickman function, arXiv:1706.06679.

1.21 Riemann Zeta Moments

The behavior of the Riemann zeta function $\zeta(z)$ on the critical line $\mathrm{Re}(z) = 1/2$ has been studied intensively for nearly 150 years. We start with a well-known asymptotic formula [1–6]:

$$\int_0^T |\zeta(1/2+i\,t)|^2\,dt \sim (\ln(T)+c)\,T$$

as $T \to \infty$, where $c = 2\gamma - 1 - \ln(2\pi)$ and γ is the Euler–Mascheroni constant [7]. This is often rewritten as

$$\frac{1}{T}\int_0^T |\zeta(1/2+i\,t)|^2\,dt \sim \int_0^T P_1\left(\ln(\tfrac{t}{2\pi})\right)dt$$

where $P_1(x) = x + 2\gamma$ is a polynomial of degree 1. More generally,

$$\frac{1}{T}\int_0^T |\zeta(1/2+i\,t)|^{2k}\,dt \sim \int_0^T P_k\left(\ln(\tfrac{t}{2\pi})\right)dt$$

where $P_k(x)$ is a polynomial of degree k^2. We are interested in the coefficients of $P_2(x)$, $P_3(x)$ and $P_4(x)$, but shall first assess the error term associated with $P_1(x)$. Observe that all moments examined here are of even order; the asymptotics of odd moments remain undiscovered [8].

1.21.1 *Error for k = 1*

Define

$$E(T) = \int_0^T |\zeta(1/2+i\,t)|^2\,dt - (\ln(T)+c)\,T.$$

Analogous to [9], we have a conjecture:

$$E(T) = O(T^{1/4+\varepsilon})$$

which is supported by the mean-square result [10, 11]:

$$\int_2^T E(t)^2 dt \sim C_2 T^{3/2}$$

where

$$C_2 = \frac{2}{3\sqrt{2\pi}}\sum_{n=1}^{\infty}\frac{d(n)^2}{n^{3/2}} = \frac{2\zeta(3/2)^4}{3\sqrt{2\pi}\,\zeta(3)}$$

and $d(n)$ is the number of divisors of n. Further supporting evidence includes [12–17]

$$\int_2^T E(t)^m dt \sim C_m T^{1+m/4}$$

where

$$C_3 = \frac{6}{7(2\pi)^{3/4}} \sum_{\sqrt{n_1}+\sqrt{n_2}=\sqrt{n_3}} \frac{d(n_1)d(n_2)d(n_3)}{(n_1 n_2 n_3)^{3/4}},$$

$$C_4 = \frac{3}{8\pi} \sum_{\sqrt{n_1}+\sqrt{n_2}=\sqrt{n_3}+\sqrt{n_4}} \frac{d(n_1)d(n_2)d(n_3)d(n_4)}{(n_1 n_2 n_3 n_4)^{3/4}},$$

$$\begin{aligned} C_5 &= \frac{10}{9(2\pi)^{5/4}} \sum_{\sqrt{n_1}+\sqrt{n_2}+\sqrt{n_3}=\sqrt{n_4}+\sqrt{n_5}} \frac{d(n_1)d(n_2)d(n_3)d(n_4)d(n_5)}{(n_1 n_2 n_3 n_4 n_5)^{3/4}} \\ &\quad - \frac{5}{9(2\pi)^{5/4}} \sum_{\sqrt{n_1}+\sqrt{n_2}+\sqrt{n_3}+\sqrt{n_4}=\sqrt{n_5}} \frac{d(n_1)d(n_2)d(n_3)d(n_4)d(n_5)}{(n_1 n_2 n_3 n_4 n_5)^{3/4}}. \end{aligned}$$

Numerical evaluation of such constants would be very challenging!

1.21.2 *Coefficients for $k \geq 2$*

Let F denote the Gauss hypergeometric function ${}_2F_1$ [18]. The leading coefficient $c_{k,0}$ of

$$P_k(x) = c_{k,0}x^{k^2} + c_{k,1}x^{k^2-1} + \cdots + c_{k,k^2-1}x + c_{k,k^2}$$

is conjectured to be [19]

$$c_{k,0} = \prod_p \left(\left(1 - \frac{1}{p}\right)^{k^2} F(k,k,1,1/p) \right) \cdot \prod_{j=0}^{k-1} \frac{j!}{(j+k)!}.$$

This is provably true for the cases

$$c_{1,0} = 1, \quad c_{2,0} = \frac{1}{12} \prod_p \left(1 - \frac{1}{p^2}\right) = \frac{1}{2\pi^2} = 0.0506605918....$$

Beyond these, the cases

$$c_{3,0} = \frac{1}{8640} \prod_p \left(1 - \frac{1}{p}\right)^4 \left(1 + \frac{4}{p} + \frac{1}{p^2}\right) = (5.708527...) \times 10^{-6},$$

$$c_{4,0} = \frac{1}{870912000} \prod_p \left(1 - \frac{1}{p}\right)^9 \left(1 + \frac{9}{p} + \frac{9}{p^2} + \frac{1}{p^3}\right) = (2.465018...) \times 10^{-13}$$

are conjectural only. For convenience, let

$$A(k) = \gamma + \sum_p \left[\frac{1}{p-1} - \frac{F(k+1,k+1,2,1/p)}{p\,F(k,k,1,1/p)} \right] \ln(p),$$

$$\begin{aligned} B(k) &= \sum_p \left[\frac{p}{(p-1)^2} + 2k^2 \frac{F(k+1,k+1,2,1/p)^2}{p^2 F(k,k,1,1/p)^2} \right. \\ &\quad \left. - \frac{k(k+1)}{2} \frac{F(k+2,k+2,3,1/p)}{p^2 F(k,k,1,1/p)} - \frac{F(k+1,k+1,1,1/p)}{p\,F(k,k,1,1/p)} \right] \ln(p)^2. \end{aligned}$$

The next coefficient $c_{k,1}$ is conjectured to be

$$c_{k,1} = 2c_{k,0}k^3 A(k)$$

which is provably true for $c_{1,1} = 2\gamma = 1.1544313298....$ Beyond this,

$$\begin{aligned} c_{2,1} &= \frac{8}{\pi^2} \left(\gamma + \frac{1}{2} \sum_p \frac{\ln(p)}{p^2-1} \right) \\ &= \frac{8}{\pi^4} \left(\gamma\pi^2 - 3\zeta'(2) \right) = 0.6988698848..., \end{aligned}$$

$$\begin{aligned} c_{3,1} &= 54c_{3,0} \left(\gamma + \frac{2}{3} \sum_p \frac{(3p+1)\ln(p)}{(p-1)(p^2+4p+1)} \right) \\ &= 0.0004050213... \end{aligned}$$

are conjectural only. The next coefficient

$$c_{k,2} = c_{k,0}k^2(k^2-1)\left(2k^2A(k)^2 - B(k) - \gamma^2 - 2\gamma_1\right)$$

gives rise to [19, 20]

$$\begin{aligned} c_{2,2} &= \frac{6}{\pi^2} \left(\frac{8}{\pi^4}(\gamma\pi^2 - 3\zeta'(2))^2 - 2\sum_p \frac{p^2\ln(p)^2}{(p^2-1)^2} - \gamma^2 - 2\gamma_1 \right) \\ &= \frac{6}{\pi^6}\left(-48\gamma\zeta'(2)\pi^2 - 12\zeta''(2)\pi^2 + 7\gamma^2\pi^4 + 144\zeta'(2)^2 - 2\,\gamma_1\pi^4\right) \\ &= 2.4259621988..., \end{aligned}$$

$$\begin{aligned} c_{3,2} &= 72c_{3,0} \left(18A(3)^2 - \sum_p \frac{p^2(7p^2+12p+7)\ln(p)^2}{(p-1)^2(p^2+4p+1)^2} - \gamma^2 - 2\gamma_1 \right) \\ &= 0.0110724552... \end{aligned}$$

where γ_m is the m^{th} Stieltjes constant [21] (for example, $\gamma_1 = -0.0728158454...$). Such values are conjectural, as well as [20]

$$\begin{aligned} c_{2,3} &= \frac{12}{\pi^8}\left(6\gamma^3\pi^6 - 84\gamma^2\zeta'(2)\pi^4 + 24\gamma_1\zeta'(2)\pi^4 - 1728\zeta'(2)^3 + 576\gamma\zeta'(2)^2\pi^2 \right. \\ &\quad \left. +288\zeta'(2)\zeta''(2)\pi^2 - 8\zeta'''(2)\pi^4 - 10\gamma_1\gamma\pi^6 - \gamma_2\pi^6 - 48\gamma\zeta''(2)\pi^4\right) \\ &= 3.2279079649..., \end{aligned}$$

$$\begin{aligned}
c_{2,4} &= \frac{4}{\pi^{10}}\big(-12\zeta''''(2)\pi^6 + 36\gamma_2\zeta'(2)\pi^6 + 9\gamma^4\pi^8 + 21\gamma_1^2\pi^8 + 432\zeta''(2)^2\pi^4 \\
&\quad +3456\gamma\zeta'(2)\zeta''(2)\pi^4 + 3024\gamma^2\zeta'(2)^2\pi^4 - 36\gamma^2\gamma_1\pi^8 - 252\gamma^2\zeta''(2)\pi^6 \\
&\quad +3\gamma\gamma_2\pi^8 + 72\gamma_1\zeta''(2)\pi^6 + 360\gamma_1\gamma\zeta'(2)\pi^6 - 216\gamma^3\zeta'(2)\pi^6 \\
&\quad -864\gamma_1\zeta'(2)^2\pi^4 + 5\gamma_3\pi^8 + 576\zeta'(2)\zeta'''(2)\pi^4 - 20736\gamma\zeta'(2)^3\pi^2 \\
&\quad -15552\zeta''(2)\zeta'(2)^2\pi^2 - 96\gamma\zeta'''(2)\pi^6 + 62208\zeta'(2)^4\big) \\
&= 1.3124243859...,
\end{aligned}$$

$$c_{3,3} = 0.1484007308..., \quad c_{3,4} = 1.0459251779...,$$

$$c_{3,5} = 3.9843850948..., \quad c_{3,6} = 8.6073191457...,$$

$$c_{3,7} = 10.2743308307..., \quad c_{3,8} = 6.5939130206...,$$

$$c_{3,9} = 0.9165155076....$$

Why are such calculations important? Since the conjectures originate in random matrix theory and appear to agree with empirical evaluations of the zeta moments, it would follow that RMT acts as a "model" for arithmetical L-function value distributions.

1.21.3 *Additive Divisor Problems*

Estermann [22–25] solved the following binary additive divisor problem:

$$\sum_{n\le N} d_2(n)d_2(n+1) \sim \frac{6}{\pi^2}N\ln(N)^2 + \alpha\, N\ln(N) + \beta\, N,$$

where $d_\ell(n)$ is the number of sequences $x_1, x_2, ..., x_\ell$ of positive integers such that $n = x_1x_2\cdots x_\ell$, and

$$\alpha = \frac{12}{\pi^4}\left(\pi^2(2\gamma-1) - 12\zeta'(2)\right) = 1.5737449203...,$$

$$\begin{aligned}
\beta &= \frac{6}{\pi^6}\left(\pi^4\left[(2\gamma-1)^2+1\right] - 24\pi^2(2\gamma-1)\zeta'(2) + 288\zeta'(2)^2 - 24\pi^2\zeta''(2)\right) \\
&= -0.5243838319....
\end{aligned}$$

For $\ell \ge 3$, it is conjectured that [26–28]

$$\sum_{n\le N} d_\ell(n)d_\ell(n+1) \sim N\, Q_\ell(\ln(N))$$

where $Q_\ell(x)$ is a polynomial of degree $2(\ell-1)$, but even the leading coefficient of $Q_3(x)$ is not known. Describing the connection between ternary additive divisors as such and the sixth moment of $\zeta(1/2+i\,t)$ would take us too far afield.

Another conjecture is [29]

$$\sum_{n \le N} d_2(n-1)d_2(n)d_2(n+1) \sim \frac{11}{8}\kappa\, N \ln(N)^3$$

where

$$\kappa = \prod_p \left(1-\frac{1}{p}\right)^2 \left(1+\frac{2}{p}\right) = 0.2867474284...$$

is the strongly carefree constant [30]. Discussion of generalizations and supporting evidence again would take us too far afield.

[1] G. H. Hardy and J. E. Littlewood, Contributions to the theory of the Riemann zeta-function and the theory of the distribution of primes, *Acta Math.* 41 (1918) 119–196; also in *Collected Papers of G. H. Hardy*, v. 2, Oxford Univ. Press, 1967, pp. 20–99.
[2] J. E. Littlewood, Researches in the theory of the Riemann ζ-function, *Proc. London Math. Soc.* 20 (1922) xxii–xxviii.
[3] A. E. Ingham, Mean-value theorems in the theory of the Riemann zeta-function, *Proc. London Math. Soc.* 27 (1926) 273–300.
[4] E. C. Titchmarsh, On van der Corput's method and the zeta-function of Riemann. V, *Quart. J. Math.* 5 (1934) 195–210.
[5] F. V. Atkinson, The mean-value of the Riemann zeta function, *Acta Math.* 81 (1949) 353–376; MR0031963 (11,234d).
[6] M. Jutila, On a formula of Atkinson, *Topics in Classical Number Theory*, v. I, Proc. 1981 Budapest conf., Colloq. Math. Soc. János Bolyai 34, North-Holland, 1984, pp. 807–823; MR0781164 (86e:11068).
[7] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[8] J. B. Conrey, *L*-functions and random matrices, *Mathematics Unlimited - 2001 and Beyond*, ed. B. Engquist and W. Schmid, Springer-Verlag, 2001, pp. 331–352; math.NT/0005300; MR1852163 (2002g:11134).
[9] S. R. Finch, Modular forms on $\mathrm{SL}_2(\mathbb{Z})$, *this volume*, §1.15.
[10] D. R. Heath-Brown, The mean value theorem for the Riemann zeta-function, *Mathematika* 25 (1978) 177–184; MR0533124 (80g:10035).
[11] T. Meurman, On the mean square of the Riemann zeta-function, *Quart. J. Math.* 38 (1987) 337–343; MR0907241 (88j:11054).
[12] K. M. Tsang, Higher-power moments of $\Delta(x)$, $E(t)$ and $P(x)$, *Proc. London Math. Soc.* 65 (1992) 65–84; MR1162488 (93c:11082).
[13] A. Ivić, On some problems involving the mean square of $\zeta(1/2+it)$, *Bull. Cl. Sci. Math. Nat. Sci. Math.* 23 (1998) 71–76; MR1744092 (2000k:11098).
[14] W. Zhai, On higher-power moments of $E(t)$, *Acta Arith.* 115 (2004) 329–348; MR2099830 (2005g:11155).
[15] W. Zhai, On higher-power moments of $\Delta(x)$, *Acta Arith.* 112 (2004) 367–395; MR2046947 (2005g:11188).
[16] W. Zhai, On higher-power moments of $\Delta(x)$. II, *Acta Arith.* 114 (2004) 35–54; MR2067871 (2005h:11216).
[17] W. Zhai, On higher-power moments of $\Delta(x)$. III, *Acta Arith.* 118 (2005) 263–281; MR2168766 (2006f:11121).
[18] S. R. Finch, Notation: Hypergeometric function, *this volume*, p. xi.
[19] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Lower order terms in the full moment conjecture for the Riemann zeta function, *J. Number Theory* 128 (2008) 1516–1554; math.NT/0612843; MR2419176 (2009b:11139).

[20] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Integral moments of L-functions, *Proc. London Math. Soc.* 91 (2005) 33–104; math.NT/0206018; MR2149530 (2006j:11120).
[21] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.
[22] T. Estermann, Über die Darstellungen einer Zahl als Differenz von zwei Produkten, *J. Reine Angew. Math.* 164 (1931) 173-182.
[23] J.-M. Deshouillers and H. Iwaniec, An additive divisor problem, *J. London Math. Soc.* 26 (1982) 1–14; MR0667238 (84b:10067).
[24] Y. Motohashi, An asymptotic series for an additive divisor problem, *Math. Z.* 170 (1980) 43–63; MR0558887 (81j:10067).
[25] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000005, A007425, A007426, A092517.
[26] A. I. Vinogradov, SL_n techniques and the density conjecture (in Russian), Anal. Teor. Chisel i Teor. Funktsii. 9, *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)* 168 (1988), 5–10, 187; Engl. transl. in *J. Soviet Math.* 53 (1991) 225–228; MR0982478 (90e:11124).
[27] A. Ivić, On the ternary additive divisor problem and the sixth moment of the zeta-function, *Sieve Methods, Exponential Sums, and their Applications in Number Theory*, ed. G. R. H. Greaves, G. Harman and M. N. Huxley, Proc. 1995 Cardiff conf., Cambridge Univ. Press, 1997, pp. 205–243; MR1635762 (99k:11129).
[28] J. B. Conrey and S. M. Gonek, High moments of the Riemann zeta-function, *Duke Math. J.* 107 (2001) 577–604; math.NT/9902162; MR1828303 (2002b:11112).
[29] T. D. Browning, The divisor problem for binary cubic forms, *J. Théor. Nombres Bordeaux* 23 (2011) 579–602; arXiv:1006.3476; MR2861076 (2012m:11131).
[30] S. R. Finch, Hafner-Sarnak-McCurley constant: Carefree couples, *first volume*, pp. 110–112.

1.22 Central Binomial Coefficients

The largest coefficient of the polynomial $(1+x)^n$ is [1]

$$A(n) = \binom{n}{\lfloor n/2 \rfloor} = \binom{n}{\lceil n/2 \rceil}.$$

It possesses recursion

$$\left\lceil \frac{n+1}{2} \right\rceil A(n+1) = (n+1)A(n), \quad A(0) = 1$$

and asymptotics

$$A(n) \sim \sqrt{\frac{2}{\pi}} n^{-1/2} 2^n$$

as $n \to \infty$. Another interpretation of $A(n)$ is as the number of sign choices + and − such that

$$\begin{cases} \pm 1 \pm 1 \pm 1 \pm \cdots \pm 1 = 0 & \text{if } n \text{ is even,} \\ \underbrace{\pm 1 \pm 1 \pm 1 \pm \cdots \pm 1}_{n} = 1 & \text{if } n \text{ is odd.} \end{cases}$$

The latter is an especially attractive characterization of the n^{th} central binomial coefficient.

Contrast this with the n^{th} central trinomial coefficient, $B(n)$, defined to be the largest coefficient of the polynomial $(1+x+x^2)^n$. There is no simple closed-form expression for $B(n)$ [2]. It possesses recursion

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

and asymptotics

$$B(n) \sim \sqrt{\frac{3}{4\pi}}n^{-1/2}3^n.$$

Here, $B(n)$ can be interpreted as the number of solutions of

$$\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \cdots + \varepsilon_n = 0$$

where each $\varepsilon_j \in \{-1, 0, 1\}$. Easy proofs of the asymptotics of $A(n)$ and $B(n)$ can be based on such additive representations, coupled with the Central Limit Theorem [3].

1.22.1 *Divisibility*

Let $\omega(n,k)$ denote the number of distinct prime factors of $\binom{n}{k}$. Erdős [4, 5] proved that

$$\omega(2n,n) \sim 2\ln(2)\frac{n}{\ln(n)}$$

as $n \to \infty$ and wondered what else could be said about the prime factors. Let

$$f(n) = \sum_{\substack{p \le n, \\ p \nmid \binom{2n}{n}}} \frac{1}{p},$$

then [6, 7]

$$c = \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N} f(n) = \sum_{k=2}^{\infty}\frac{\ln(k)}{2^k} = 0.5078339228...,$$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}(f(n)-c)^2 = 0.$$

These two facts together express that $f(n) \to c$ for almost all integers n, hence $\binom{2n}{n}$ is almost always divisible by high powers of small primes. Let $g(n)$ be the smallest odd prime factor of $\binom{2n}{n}$. Whether $f(n)$ or $g(n)$ are bounded remains an open question.

Sárközy [8] and others [9–11] proved that $\binom{2n}{n}$ is not square-free for any $n > 4$. The largest n for which $\binom{2n}{n}$ is not divisible by p^2 for any odd prime p is $n = 786$.

We turn attention to $\binom{n}{k}$, the $(k+1)^{\text{st}}$ element in the n^{th} row of Pascal's triangle. For each $k \geq 1$, the sequence of integers n such that $\binom{n}{k}$ is square-free has asymptotic density c_k, where

$$c_1 = \frac{6}{\pi^2} = 0.6079271018..., \quad c_2 = \frac{3}{4}\prod_{p\geq 3}\left(1-\frac{2}{p^2}\right) = 0.4839511484...$$

(the latter is related to the Feller–Tornier constant [12]). More generally, write k in base p:

$$k = a_0 + a_1 p + a_2 p^2 + \cdots + a_\ell p^\ell, \quad 0 \leq a_j < p \text{ for all } 0 \leq j \leq \ell, \quad a_{\ell+1} = 0,$$

and define

$$c_{k,p} = \begin{cases} \prod\limits_{i=0}^{\ell}\left(1-\dfrac{a_i}{p}\right)\cdot\left(1+\sum\limits_{j=0}^{\ell}\dfrac{a_j(p-1-a_{j+1})}{(p-a_j)(p-a_{j+1})}\right) & \text{if } p \leq k, \\ 1-\dfrac{k}{p^2} & \text{if } p > k. \end{cases}$$

Then c_k is equal to $\prod_p c_{k,p}$, where the product is taken over all primes p. We have $c_3 = 0.251...$, $c_4 = 0.360...$, $c_5 = 0.191...$, $c_6 = 0.189...$, $c_7 = 0.062...$ and

$$0 < c_k = \exp\left[-(\alpha + o(1))\sqrt{k}/\ln(k)\right]$$

as $k \to \infty$, where

$$\begin{aligned} \alpha &= \sum_{j=1}^{\infty}\frac{1}{j(j+1)}\int_0^\infty \{x\}^j x^{-3/2}dx \\ &= \sum_{j=1}^{\infty}\binom{2j}{j}\zeta(j+1/2)\frac{1}{2^{2j-1}}\left(1-j\sum_{i>j}\frac{1}{i^2}\right) \\ &= 1.825108.... \end{aligned}$$

Integrals involving $\{x\} = x - \lfloor x \rfloor$ as such also appear in [13, 14]. It follows that there are $\sim \tau N$ square-free binomial coefficients $\binom{n}{k}$ with $0 \leq k < n \leq N$, where

$$\tau = 2\sum_{k=0}^{\infty} c_k = 2(5.3275...) = 10.655....$$

In words, each row of Pascal's triangle possesses approximately $10\frac{2}{3}$ square-free entries (on average).

1.22.2 *Relevant Sums*

Let φ denote the Golden mean $(1+\sqrt{5})/2$. We have [15–19]

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}}=\frac{1}{3}+\frac{2\sqrt{3}\pi}{27},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}}=\frac{1}{5}+\frac{4\sqrt{5}\ln(\varphi)}{25},$$

$$\sum_{n=1}^{\infty}\frac{n}{\binom{2n}{n}}=\frac{2}{3}+\frac{2\sqrt{3}\pi}{27},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}n}{\binom{2n}{n}}=\frac{6}{25}+\frac{4\sqrt{5}\ln(\varphi)}{125},$$

$$\sum_{n=1}^{\infty}\frac{n^2}{\binom{2n}{n}}=\frac{4}{3}+\frac{10\sqrt{3}\pi}{81},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}n^2}{\binom{2n}{n}}=\frac{4}{25}-\frac{4\sqrt{5}\ln(\varphi)}{125},$$

$$\sum_{n=1}^{\infty}\frac{n^3}{\binom{2n}{n}}=\frac{10}{3}+\frac{74\sqrt{3}\pi}{243},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}n^3}{\binom{2n}{n}}=-\frac{2}{125}-\frac{28\sqrt{5}\ln(\varphi)}{625}$$

and, more generally [20],

$$\sum_{n=1}^{\infty}\frac{n^k}{\binom{2n}{n}}=p_k+q_k\sqrt{3}\pi,\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}n^k}{\binom{2n}{n}}=r_k+s_k\sqrt{5}\ln(\varphi)$$

for appropriate rationals p_k, q_k, r_k, s_k. Let L_D denote the Dirichlet L-series with character $(D/\cdot)$ and Li_k denote the k^{th} polylogarithm function [14]. The following are more difficult [15–19]:

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}n}=\frac{\sqrt{3}\pi}{9},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}n}=\frac{2\sqrt{5}\ln(\varphi)}{5},$$

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}n^2}=\frac{\pi^2}{18},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}n^2}=2\ln(\varphi)^2,$$

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}n^3}=\frac{\sqrt{3}\pi}{2}L_{-3}(2)-\frac{4\zeta(3)}{3},\quad \sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}n^3}=\frac{2\zeta(3)}{5},$$

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}n^4}=\frac{17\pi^4}{3240},$$

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}n^4} &= 8\,\mathrm{Li}_4\left(\frac{1}{\varphi}\right)+8\ln(\varphi)\,\mathrm{Li}_3\left(\frac{1}{\varphi}\right)-\frac{1}{2}\,\mathrm{Li}_4\left(\frac{1}{\varphi^2}\right)\\ &\quad+\frac{7\pi^2\ln(\varphi)^2}{15}-\frac{13\ln(\varphi)^4}{6}-\frac{4\zeta(3)\ln(\varphi)}{5}-\frac{7\pi^4}{90},\end{aligned}$$

$$\sum_{n=1}^{\infty}\frac{1}{\binom{2n}{n}n^5}=\frac{9\sqrt{3}\pi}{8}L_{-3}(4)+\frac{\pi^2\zeta(3)}{9}-\frac{19\zeta(5)}{3},$$

$$\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{\binom{2n}{n}n^5} = \frac{5}{2}\,\mathrm{Li}_5\left(\frac{1}{\varphi^2}\right)+5\ln(\varphi)\,\mathrm{Li}_4\left(\frac{1}{\varphi^2}\right) +4\zeta(3)\ln(\varphi)^2-\frac{4\pi^2\ln(\varphi)^3}{9}+\frac{4\ln(\varphi)^5}{3}-2\zeta(5).$$

Let $G = L_{-4}(2)$ denote Catalan's constant. Other series include

$$\sum_{n=0}^{\infty}\frac{1}{\binom{2n}{n}(2n+1)}=\frac{2\sqrt{3}\pi}{9},\quad \sum_{n=0}^{\infty}\frac{(-1)^n}{\binom{2n}{n}(2n+1)}=\frac{4\sqrt{5}\ln(\varphi)}{5},$$

$$\sum_{n=0}^{\infty}\frac{1}{\binom{2n}{n}(2n+1)^2}=\frac{8G}{3}-\frac{\pi\ln(2+\sqrt{3})}{3},\quad \sum_{n=0}^{\infty}\frac{(-1)^n}{\binom{2n}{n}(2n+1)^2}=\frac{\pi^2}{6}-3\ln(\varphi)^2$$

and

$$\sum_{n=0}^{\infty}\frac{2^n}{\binom{2n}{n}(2n+1)}=\frac{\pi}{2},\quad \sum_{n=0}^{\infty}\frac{(-1)^n 2^n}{\binom{2n}{n}(2n+1)}=\frac{2}{\sqrt{3}}\ln\left(\frac{1+\sqrt{3}}{\sqrt{2}}\right),$$

$$\sum_{n=0}^{\infty}\frac{2^n}{\binom{2n}{n}(2n+1)^2}=2L_{-8}(2)-\frac{\sqrt{2}\pi}{4}\ln(1+\sqrt{2}),$$

$$\sum_{n=0}^{\infty}\frac{2^{2n}}{\binom{2n}{n}(2n+1)^2}=2G,\quad \sum_{n=0}^{\infty}\frac{(-1)^n 2^{2n}}{\binom{2n}{n}(2n+1)^2}=\frac{\pi^2}{8}-\frac{1}{2}\ln(1+\sqrt{2})^2,$$

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{2^{2n}}{\binom{2n}{n}(2n+1)^3} &= 2i\left[\mathrm{Li}_3\left(\frac{1-i}{2}\right)-\mathrm{Li}_3\left(\frac{1+i}{2}\right)\right]-\frac{\pi\ln(2)^2}{8}-\frac{\pi^3}{32}\\ &= 2i\left[\mathrm{Li}_3(1+i)-\mathrm{Li}_3(1-i)\right]+\frac{\pi\ln(2)^2}{4}+\frac{3\pi^3}{16}.\end{aligned}$$

The latter sum is due to Gosper [21]. Batir [22, 23] proved that

$$\sum_{n=1}^{\infty}\frac{2^{4n}}{\binom{2n}{n}^2 n^3}=8\pi G-14\zeta(3),\quad \sum_{n=0}^{\infty}\frac{2^{4n+2}}{\binom{2n}{n}^2(2n+1)^3}=14\zeta(3)-4\pi G$$

and also derived a complicated formula for $\sum_{n=1}^{\infty} 1/\binom{3n}{n}$. We will barely mention cases for which $\binom{2n}{n}$ is in the numerator, for example [15, 17, 24],

$$\sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{2n}(2n+1)}=\frac{\pi}{2},$$

$$\sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{2n}(2n+1)^2}=\frac{\pi\ln(2)}{2},\quad \sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{3n}(2n+1)^2}=\frac{\sqrt{2}}{8}\left(\pi\ln(2)+4G\right),$$

$$\sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{4n}(2n+1)^2}=\frac{3\sqrt{3}}{4}L_{-3}(2),$$

$$\sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{4n}(2n+1)^3}=\frac{7\pi^3}{216},\quad \sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{2^{4n}(2n+1)^4}=\frac{\pi\zeta(3)}{12}+\frac{27\sqrt{3}}{64}L_{-3}(4),$$

$$\sum_{n=0}^{\infty}\binom{2n}{n}^2\frac{1}{2^{4n}(2n+1)}=\frac{4G}{\pi}.$$

Deninger's conjecture [25]

$$\sum_{n=0}^{\infty}\binom{2n}{n}^2\frac{1}{2^{8n}(2n+1)}=\frac{15}{\pi^2}L_{15A}(2)$$

was proved by Rogers & Zudilin [26], where L_{15A} is the L-series for the elliptic curve isogeny class 15A. See [27] for a sampling of other conjectures and [28] for other techniques.

1.22.3 *Middle Stirling Numbers*

Asymptotic results for middle Stirling numbers are more complicated than those for central binomial coefficients. Let $s_{2n,n}$ denote the number of permutations on $2n$ symbols possessing exactly n cycles; let $S_{2n,n}$ denote the number of partitions of a $(2n)$-element set possessing exactly n blocks. We have [29–32]

$$\frac{n!}{(2n)!}s_{2n,n}\sim\kappa_1\frac{\lambda_1^n}{\sqrt{n}},\quad \frac{n!}{(2n)!}S_{2n,n}\sim\kappa_2\frac{\lambda_2^n}{\sqrt{n}},$$

where

$$\lambda_1=\frac{\xi}{[1-\exp(-\xi)]^2}=2.4554074822...,\quad \lambda_2=\frac{\exp(\eta)-1}{\eta^2}=1.5441386523...$$

and ξ, η are unique positive solutions of the equations

$$\frac{\exp(\xi)-1}{\xi}=2,\quad \frac{\eta}{1-\exp(-\eta)}=2.$$

The latter is a Lambert W function value: $\eta=2+W(-2e^{-2})=1.5936242600...$ [33] while the former satisfies $2\xi/(2\xi+1)=0.7153318629....$ Generalizations of such results appear in [34–37].

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001405, A002426, A126869.
[2] M. Petkovsek, H. S. Wilf and D. Zeilberger, $A=B$, A. K. Peters, 1996, pp. 160, 165; MR1379802 (97j:05001).
[3] S. R. Finch, Signum equations and extremal coefficients, *this volume*, §1.30.
[4] P. Erdős, Über die Anzahl der Primfaktoren von $\binom{n}{k}$, *Arch. Math. (Basel)* 24 (1973) 53–56; MR0319881 (47 #8422).
[5] P. Goetgheluck, On prime divisors of binomial coefficients, *Math. Comp.* 51 (1988) 325–329; MR0942159 (89f:11033).

[6] P. Erdős, R. L. Graham, I. Z. Ruzsa and E. G. Straus, On the prime factors of $\binom{2n}{n}$, *Math. Comp.* 29 (1975) 83–92. MR0369288 (51 #5523).

[7] J. W. Sander, On primes not dividing binomial coefficients, *Math. Proc. Cambridge Philos. Soc.* 113 (1993) 225–232; MR1198408 (93m:11099).

[8] A. Sárközy, On divisors of binomial coefficients. I, *J. Number Theory* 20 (1985) 70–80; MR0777971 (86c:11002).

[9] G. Velammal, Is the binomial coefficient $\binom{2n}{n}$ squarefree? *Hardy-Ramanujan J.* 18 (1995) 23–45; MR1300106 (95j:11016).

[10] A. Granville and O. Ramaré, Explicit bounds on exponential sums and the scarcity of squarefree binomial coefficients, *Mathematika* 43 (1996) 73–107; MR1401709 (97m:11023).

[11] P. A. Cutter, Finding prime pairs with particular gaps, *Math. Comp.* 70 (2001) 1737–1744; MR1836931 (2002c:11174).

[12] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.

[13] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.

[14] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[15] I. J. Zucker, On the series $\sum_{k=1}^{\infty}\binom{2k}{k}^{-1}k^{-n}$ and related sums, *J. Number Theory* 20 (1985) 92–102; MR0777973 (86e:11018).

[16] D. H. Lehmer, Interesting series involving the central binomial coefficient, *Amer. Math. Monthly* 92 (1985) 449–457; MR0801217 (87c:40002).

[17] B. C. Berndt, *Ramanujan's Notebooks. Part I*, Springer-Verlag, 1985, pp. 264–265, 289–290; MR0781125 (86c:01062).

[18] J. M. Borwein, D. J. Broadhurst and J. Kamnitzer, Central binomial sums, multiple Clausen values, and zeta values, *Experim. Math.* 10 (2001) 25–34; MR1821569 (2002k:11105).

[19] T. Sherman, Summation of Glaisher- and Apéry-like series, unpublished manuscript (2000).

[20] J. M. Borwein and R. Girgensohn, Evaluations of binomial series, *Aequationes Math.* 70 (2005) 25–36; MR2167981 (2006e:05006).

[21] R. W. Gosper, Several central binomial sums, unpublished note (2008).

[22] N. Batir, Integral representations of some series involving $\binom{2k}{k}^{-1}k^{-n}$ and some related series, *Appl. Math. Comput.* 147 (2004) 645–667; MR2011078 (2004h:33009).

[23] N. Batir, On the series $\sum\binom{3k}{k}^{-1}k^{-n}x^{k}$, *Proc. Indian Acad. Sci. Math. Sci.* 115 (2005) 371–381; math.AC/0512310; MR2184197 (2006j:11170).

[24] S. Ruscheweyh, Einige neue Darstellungen des Di- und Trilogarithmus, *Math. Nachr.* 57 (1973) 237–244; MR0326012 (48 #4358).

[25] M. Rogers and B. Yuttanan, Modular equations and lattice sums, *Computational and Analytical Mathematics*, Proc. 2001 Burnaby conf., ed. D. H. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. D. Vanderwerff and H. Wolkowicz, Springer-Verlag, 2013, pp. 667–680; arXiv:1001.4496; MR3108448.

[26] M. Rogers and W. Zudilin, On the Mahler measure of $1+X+1/X+Y+1/Y$, *Internat. Math. Res. Notices* (2014) 2305–2326; arXiv:1102.1153; MR3207368.

[27] Z.-W. Sun, List of conjectural series for powers of π and other constants, arXiv:1102.5649.

[28] S. Weinzierl, Expansion around half-integer values, binomial sums, and inverse binomial sums, *J. Math. Phys.* 45 (2004) 2656–2673; arXiv:hep-ph/0402131; MR2067580 (2005f:33042).

[29] L. Moser and M. Wyman, Asymptotic development of the Stirling numbers of the first kind, *J. London Math. Soc.* 33 (1958) 133–146; MR0098202 (20 #4664).

[30] L. Moser and M. Wyman, Stirling numbers of the second kind, *Duke Math. J.* 25 (1957) 29–43; MR0091964 (19,1039b).
[31] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009, pp. 588–589; MR2483235 (2010h:05005).
[32] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007820 and A187646.
[33] S. R. Finch, Iterated exponential constants, *first volume*, pp. 448–452.
[34] E. G. Tsylova, Probabilistic methods for obtaining asymptotic formulas for generalized Stirling numbers (in Russian), *Statisticheskie Metody Otsenivaniyi i Poverki Gipotez*, Proc 1990 Perm conf., Gos. Univ., 1991, pp. 165–178; Engl. transl. in *J. Math. Sci.* 75 (1995) 1607–1614; MR1253585.
[35] A. N. Timashov, On asymptotic expansions of Stirling numbers of the first and second kinds (in Russian), *Diskret. Mat.*, v. 10 (1998) n. 3, 148–159; Engl. transl. in *Discrete Math. Appl.* 8 (1998) 533–544; MR1673690 (99m:11021).
[36] G. Louchard, Asymptotics of the Stirling numbers of the first kind revisited: a saddle point approach,. *Discrete Math. Theor. Comput. Sci.* 12 (2010) 167–184; MR2676669 (2012c:05031).
[37] G. Louchard, Asymptotics of the Stirling numbers of the second kind revisited, *Appl. Anal. Discrete Math.* 7 (2013) 193–210; MR3113398.

1.23 Fractional Parts of Bernoulli Numbers

The Bernoulli numbers B_0, B_1, B_2, ... are defined via [1–3]

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}$$

and satisfy $B_0 = 1$, $B_1 = -1/2$, $(-1)^{k+1} B_{2k} > 0$ and $B_{2k+1} = 0$ for $k \geq 1$. It can be shown that $|B_{2k}|$ is strictly increasing after its minimum at $B_6 = 1/42$, and

$$|B_{2k}| \sim \frac{2(2k)!}{(2\pi)^{2k}} \sim 4\sqrt{\pi k}\left(\frac{k}{e\pi}\right)^{2k}$$

as $k \to \infty$. Let $\{x\} = x - \lfloor x \rfloor$ denote the fractional part of a real number x; for example,

$$\{B_2\} = \{\tfrac{1}{6}\} = \tfrac{1}{6}, \quad \{B_4\} = \{-\tfrac{1}{30}\} = \tfrac{29}{30},$$
$$\{B_{14}\} = \{\tfrac{7}{6}\} = \tfrac{1}{6}, \quad \{B_{16}\} = \{-\tfrac{3617}{510}\} = \tfrac{463}{510}.$$

The sequence $\{B_2\}$, $\{B_4\}$, $\{B_6\}$, ... is dense in the unit interval $[0, 1]$, but it is not uniformly distributed [4]. Certain rational numbers appear with positive probability: $1/6$ is most likely with probability 0.151..., $29/30$ is next with probability 0.064... [5]. In fact, the limiting distribution F is piecewise linear with countably many jump discontinuities: F increases only when jumping (see Figure 1.20). We wonder, in particular, about the moments of F. By the von Staudt–Clausen theorem, the mean fractional part is [6]

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left\{ - \sum_{(p-1)|2n} \frac{1}{p} \right\} = 0.5486...$$

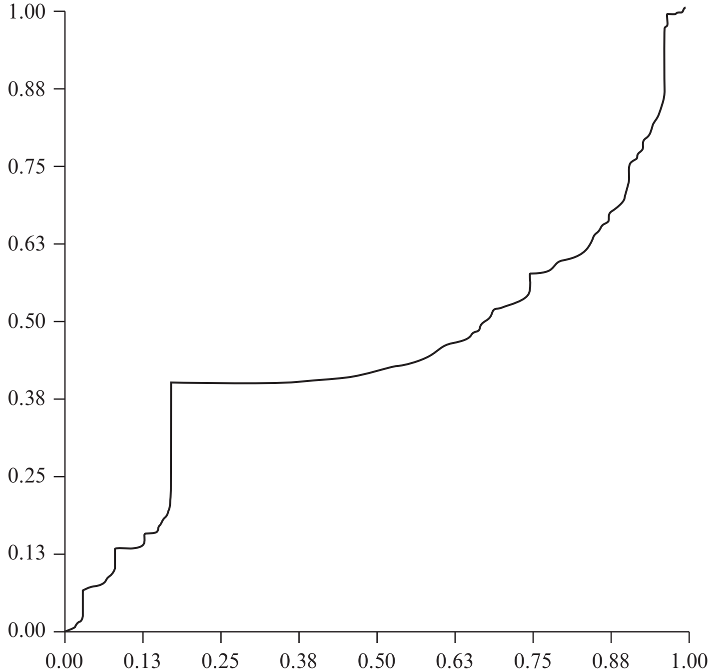

Figure 1.20 Bernoulli numbers fractional parts distribution

and the mean fractional part squared is

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\left\{-\sum_{(p-1)|2n}\frac{1}{p}\right\}^2=0.4396....$$

The inner sum is over all primes p such that $p-1$ divides $2n$. No analytic simplification of such formulas is known.

A proof of the equality [7]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\sum_{p|n}\frac{1}{p}=\sum_{p}\frac{1}{p^2}=0.4522474200...$$

will be given shortly. If the sum $\sum 1/p$ is replaced by the reciprocal of the least prime factor $P^{-}(n)$ of n, then interestingly [8, 9]

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\frac{1}{P^{-}(n)}=\sum_{p}\frac{1}{p^2}\prod_{q<p}\left(1-\frac{1}{q}\right)$$

where the inner product is over all primes q less than p. In principle, this latter expression can be evaluated to high precision. A similar replacement for the average of $\{B_{2n}\}$ is not clear. Observe that $p=2$ and $p=3$ both satisfy $(p-1)|2n$

automatically for any $n \geq 1$. The issue is thus determining the smallest such prime exceeding 3 for each n (if one exists) and this may be awkward.

The promised proof starts by letting $S_N = \sum_{n\leq N}\sum_{p|n} 1/p$. It is clear that

$$S_N = \sum_{p\leq N}\frac{1}{p}\sum_{\substack{n\leq N,\\ p|n}} 1 = \sum_{p\leq N}\frac{1}{p}\sum_{m\leq N/p} 1 = \sum_{p\leq N}\frac{\lfloor N/p\rfloor}{p}$$

and, since $N/p - 1 < \lfloor N/p\rfloor \leq N/p$,

$$\sum_{p\leq N}\frac{1}{p^2} - \frac{1}{N}\sum_{p\leq N}\frac{1}{p} < \frac{1}{N}S_N \leq \sum_{p\leq N}\frac{1}{p^2}.$$

The result follows because $\sum_{p\leq N} 1/p = O(\ln\ln N)$.

A famous conjecture, due to Siegel [10–13], is as follows. An odd prime p is **regular** if it does not divide the numerator of any of the Bernoulli numbers B_2, B_4, B_6, ..., B_{p-3}; otherwise p is **irregular**. It seems to be true that

$$\lim_{N\to\infty}\frac{\sum\limits_{\substack{p\leq N,\\ p\text{ irregular}}} 1}{\sum\limits_{\substack{p\leq N,\\ p\text{ regular}}} 1} = e^{1/2} - 1 = 0.6487212707...$$

but a proof is not known. Equivalently, we have

$$\lim_{N\to\infty}\frac{\ln(N)}{N}\sum_{\substack{p\leq N,\\ p\text{ irregular}}} 1 = 1 - e^{-1/2} = 0.3934693402...,$$

$$\lim_{N\to\infty}\frac{\ln(N)}{N}\sum_{\substack{p\leq N,\\ p\text{ regular}}} 1 = e^{-1/2} = 0.6065306597....$$

In 1851, Kummer proved that Fermat's Last Theorem holds when the exponent is a regular prime. Although FLT was proved by Wiles in 1995, we still do not know whether there exist infinitely many regular primes.

See also [14, 15] for the asymptotics for $\prod_{k\leq K}|B_{2k}|$.

[1] K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, 2nd ed., Springer-Verlag, 1990, pp. 228–248, 264–266; MR1070716 (92e:11001).

[2] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976, pp. 264–266; MR0434929 (55 #7892).

[3] C. J. Moreno and S. S. Wagstaff, *Sums of Squares of Integers*, Chapman & Hall/CRC, 2006, pp. 61–116; MR2189437 (2006m:11048).

[4] P. Erdős and S. S. Wagstaff, The fractional parts of the Bernoulli numbers, *Illinois J. Math.* 24 (1980) 104–112; MR0550654 (81c:10064).

[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A045979, A046886, A051222, A051225, A051226, A067513.
[6] S. Plouffe, Computations involving Bernoulli denominators, unpublished note (2007).
[7] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.
[8] W. P. Zhang, Average-value estimation of a class of number-theoretic functions (in Chinese), *Acta Math. Sinica* 32 (1989) 260–267; MR1025146 (90k:11124).
[9] S. R. Finch, Series involving arithmetic functions, *this volume*, §1.20.
[10] J. L. Selfridge, C. A. Nicol and H. S. Vandiver, Proof of Fermat's last theorem for all prime exponents less than 4002, *Proc. Nat. Acad. Sci. U.S.A.* 41 (1955) 970–973; MR0072892 (17,348a).
[11] C. L. Siegel, Zu zwei Bemerkungen Kummers, *Nachr. Akad. Wiss. Göttingen, Math.-Phys. Klasse* (1964) 51–57; also in *Gesammelte Abhandlungen*, v. 3, ed. K. Chandrasekharan and H. Maass, Springer-Verlag, 1966, pp. 436–442; MR0163899 (29 #1198).
[12] W. Johnson, Irregular primes and cyclotomic invariants, *Math. Comp.* 29 (1975) 113–120; MR0376606 (51 #12781).
[13] S. S. Wagstaff, The irregular primes to 125000, *Math. Comp.* 32 (1978) 583–591; MR0491465 (58 #10711).
[14] B. Kellner, On asymptotic constants related to products of Bernoulli numbers and factorials, *Integers* 9 (2009) 83–106; math.NT/0604505; MR2506140 (2010d:11027).
[15] S. R. Finch, Minkowski-Siegel mass constants, *this volume*, §5.6.

1.24 Products of Consecutive-Integer Ratios

Consider the random product

$$P(N) = \prod_{n=1}^{N} \left(\frac{n}{n+1}\right)^{\varepsilon_n} = \prod_{n=1}^{N} \left(1 + \frac{1}{n}\right)^{-\varepsilon_n}$$

where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ are independent variables satisfying $\mathrm{P}(\varepsilon_n = 1) = \mathrm{P}(\varepsilon_n = -1) = 1/2$ for each n. The maximum value of $P(N)$ is $N+1$, which occurs if and only if all ε_n are -1. The minimum value of $P(N)$ is $1/(N+1)$, which occurs if and only if all ε_n are 1. We are interested in the average behavior of $P(N)$ and it makes sense to examine $\ln(P(N))$ henceforth (with extreme values $-\ln(N+1)$ and $\ln(N+1)$ symmetric about the origin).

Before continuing, let us mention the random sum

$$S(N) = \sum_{n=1}^{N} \frac{\varepsilon_n}{n}$$

which converges almost surely [1, 2]. The maximum value of $S(N)$ diverges to ∞ as $N \to \infty$ and the minimum value of $S(N)$ diverges to $-\infty$. Clearly $\mathrm{E}(S(N)) = 0$ and

$$\mathrm{Var}(S(N)) = \sum_{n=1}^{N} \frac{1}{n^2} \to \frac{\pi^2}{6}$$

as $N \to \infty$. It is perhaps surprising that $\mathrm{Var}(S(N))$ is finite. Define $\theta_n = -1$ if $n \equiv 0 \bmod 3$ and $\theta_n = 1$ otherwise; define $\omega_n = -1$ if $n \equiv 2, 3 \bmod 4$ and $\omega_n = 1$ otherwise. On the one hand [3],

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln(2), \quad \sum_{n=1}^{\infty} \frac{\omega_n}{n} = \frac{\pi}{4} - \frac{1}{2}\ln(2);$$

on the other hand [4],

$$\sum_{n=1}^{N} \frac{1}{n} \sim \ln(N) + \gamma, \quad \sum_{n=1}^{N} \frac{\theta_n}{n} \sim \frac{1}{3}\ln(N) + \frac{2}{3}\ln(3) + \frac{1}{3}\gamma$$

where γ is the Euler–Mascheroni constant [5].

Returning to the product $P(N)$, we have $\mathrm{E}(\ln(P(N))) = 0$ and

$$\mathrm{Var}(\ln(P(N))) = \sum_{n=1}^{N} \ln\left(\frac{n}{n+1}\right)^2 \to 0.977189...$$

as $N \to \infty$. No closed-form expression for this expression is known. Again, it is perhaps surprising that $\mathrm{Var}(\ln(P(N)))$ is finite. By Wallis' formula [6, 7], we have

$$\prod_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{(-1)^{n+1}} = \frac{1}{2}\frac{3}{2}\frac{3}{4}\frac{5}{4}\frac{5}{6}\frac{7}{6}\frac{7}{8}\cdots = \frac{2}{\pi}$$

but as before an unbalanced distribution of $+1$ and -1 exponents leads to divergence (to either ∞ or 0).

Here is a far more difficult problem. Let $a(N)$ and $b(N)$ denote the numerator and denominator of $P(N)$, expressed in lowest terms. Rather than maximizing $P(N)$ for fixed N as previously, consider instead maximizing $a(N)$. Note that, by changing each ε_k to $-\varepsilon_k$, the maximum value of $b(N)$ is equal to the maximum value of $a(N)$. Hence we lose nothing by studying only numerators in the following.

Let $A(N)$ denote the maximum value of $a(N)$. See Table 1.12 for sample values [8]. For example, when $N = 6$,

$$\text{the numerator of } \left(\tfrac{1}{2}\right)^{-1}\tfrac{2}{3}\left(\tfrac{3}{4}\right)^{-1}\tfrac{4}{5}\left(\tfrac{5}{6}\right)^{-1}\tfrac{6}{7} \text{ is } 2^8$$

whereas

$$\text{the numerator of } \left(\tfrac{1}{2}\right)^{-1}\left(\tfrac{2}{3}\right)^{-1}\left(\tfrac{3}{4}\right)^{-1}\tfrac{4}{5}\left(\tfrac{5}{6}\right)^{-1}\tfrac{6}{7} \text{ is } 2^6 3^2;$$

hence $A(6) = 576$. Nicolas [9] and de la Bretèche, Pomerance & Tenenbaum [10] proved that

$$0.107 < \liminf_{N\to\infty} \frac{1}{N\ln(N)} \ln(A(N)) \le \limsup_{N\to\infty} \frac{1}{N\ln(N)} \ln(A(N)) \le \frac{2}{3} < 0.667.$$

At the end of [10], the lower bound was improved to 0.112 (due to Fouvry). We wonder whether the limit supremum is equal to the limit infimum and, if so, what the limiting value might be.

Table 1.12 *Sample values of maximum numerator* $A(N)$ *and* $\ln(A(N))/(N\ln(N))$

N	1	2	3	4	5	6
$A(N)$	2	4	16	64	128	576
$\dfrac{\ln(A(N))}{N\ln(N)}$		1.0000	0.8407	0.7500	0.6031	0.5909
N	7	8	9	10	11	12
$A(N)$	4608	16384	64000	640000	2560000	10240000
$\dfrac{\ln(A(N))}{N\ln(N)}$	0.6195	0.5833	0.5596	0.5806	0.5592	0.5414

1.24.1 *Highly Composite Numbers*

A positive integer n is **highly composite** if, for all $m < n$, we have $d(m) < d(n)$, where $d(k)$ denotes the number of distinct divisors of k. The integer n is also called a d-**champion**. It is known that

$$|\{n \leq N : n \text{ is highly composite}\}| = O\left(\ln(N)^{1.71}\right)$$

as $N \to \infty$, and conjectured that 1.71 can be replaced by any constant $c > \ln(30)/\ln(16) = 1.2267...$ [11].

A positive integer n is **superior highly composite** if there exists $\delta > 0$ such that, for all positive integers m, we have $d(m)/m^{\delta} \leq d(n)/n^{\delta}$. It is known that

$$|\{n \leq N : n \text{ is superior highly composite}\}| \sim \ln(N)$$

as $N \to \infty$. While these asymptotics are well-understood, those for the quotient of two consecutive highly composite numbers are not.

Define

$$\lambda = \limsup_{N\to\infty} \frac{1}{N\ln(N)} \ln(A(N))$$

where $A(N)$ is as before. If M is a sufficiently large superior highly composite number and M' is the highly composite number following M, then [11]

$$\frac{M'}{M} \geq 1 + \frac{1}{\ln(M)^{\kappa}}$$

for any constant $\kappa > \lambda/\ln(2)$. Since we know $\lambda \leq 2/3$, it follows that the exponent $2/(3\ln(2)) = 0.961796...$ works. A sharper upper bound on λ (for example, $\lambda \leq 3/5$ or even $\lambda \leq 1/2$) would be very helpful.

[1] K. E. Morrison, Cosine products, Fourier transforms, and random sums, *Amer. Math. Monthly* 102 (1995) 716–724; MR1357488 (96h:42001).

[2] B. Schmuland, Random harmonic series, *Amer. Math. Monthly* 110 (2003) 407–416; MR2040884 (2005a:60024).

[3] L. B. W. Jolley, *Summation of Series*, 2nd ed., Dover, 1961, pp. 14–15; MR0134458 (24 #B511).
[4] J. M. Hyslop, *Infinite Series*, Oliver and Boyd; Interscience; 1942, p. 67; MR0007792 (4,193j).
[5] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[6] T. J. Osler, The union of Vieta's and Wallis's products for pi, *Amer. Math. Monthly* 106 (1999) 774–776.
[7] S. R. Finch, Archimedes' constant, *first volume*, pp. 17–28.
[8] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002182, A002201, A127588.
[9] J.-L. Nicolas, Nombres hautement composés, *Acta Arith.* 49 (1988) 395–412; MR0937935 (89c:11012).
[10] R. de la Bretèche, C. Pomerance and G. Tenenbaum, Products of ratios of consecutive integers, *Ramanujan J.* 9 (2005) 131–138; MR2166384 (2006e:11147).
[11] J.-L. Nicolas, Some open questions, *Ramanujan J.* 9 (2005) 251–264; MR2166389 (2006e:11148).

1.25 Prime Number Theorem

Let $\pi(x)=\sum_{p\le x}1$, the number of primes p not exceeding x. Gauss and Legendre conjectured an asymptotic expression for $\pi(x)$. Define the Möbius mu function

$$\mu(n)=\begin{cases}1 & \text{if } n=1,\\ (-1)^r & \text{if } n \text{ is a product of } r \text{ distinct primes},\\ 0 & \text{if } n \text{ is divisible by a square } >1;\end{cases}$$

the von Mangoldt function

$$\Lambda(n)=\begin{cases}\ln(p) & \text{if } n=p^m \text{ for some prime } p \text{ and integer } m\ge 1,\\ 0 & \text{otherwise;}\end{cases}$$

and the Chebyshev functions

$$\theta(x)=\sum_{p\le x}\ln(p),$$

$$\psi(x)=\sum_{\substack{p^m\le x,\\ m\ge 1}}\ln(p)=\sum_{n\le x}\Lambda(n)=\ln(\operatorname{lcm}\{1,2,\ldots,\lfloor x\rfloor\}).$$

Hadamard and de la Vallée Poussin proved the Gauss–Legendre conjecture, namely,

$$\pi(x)\sim\frac{x}{\ln(x)},\quad \theta(x)\sim x,\quad \psi(x)\sim x$$

as $x\to\infty$. These three formulas are equivalent to each other and also to

$$\sum_{n=1}^{\infty}\frac{\mu(n)}{n}=0.$$

The Riemann zeta function clearly plays a role here since, for $\mathrm{Re}(s) > 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s), \quad \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}.$$

Of many aspects of the Prime Number Theorem, we focus on the following error estimates [1–4]:

$$\begin{aligned} 1 + \int_1^{\infty} \frac{\theta(x) - x}{x^2} dx &= \lim_{N \to \infty} \left(\sum_{p \le N} \frac{\ln(p)}{p} - \ln(N) \right) \\ &= -\gamma - \sum_{p} \frac{\ln(p)}{p(p-1)} = -1.3325822757..., \end{aligned}$$

$$1 + \int_1^{\infty} \frac{\psi(x) - x}{x^2} dx = \lim_{N \to \infty} \left(\sum_{n \le N} \frac{\Lambda(n)}{n} - \ln(N) \right) = -\gamma = -0.5772156649...$$

where γ is the Euler–Mascheroni constant [5, 6]. The latter implies that

$$\sum_{n=1}^{\infty} \frac{\Lambda(n) - 1}{n} = -2\gamma.$$

What can be said about analogous expressions connected with the Prime Number Theorem for arithmetic progressions $3k + 1$ and $4k + 1$?

Nevanlinna [7, 8] gave a straightforward generalization:

$$\sum_{n \equiv 1 \bmod 3} \frac{\mu(n)}{n} = \frac{1}{2} \frac{1}{L_{-3}(1)} = \frac{3\sqrt{3}}{2\pi},$$

$$\sum_{n \equiv 1 \bmod 4} \frac{\mu(n)}{n} = \frac{1}{2} \frac{1}{L_{-4}(1)} = \frac{2}{\pi};$$

$$\begin{aligned} \lim_{N \to \infty} \left(\sum_{\substack{n \equiv 1 \bmod 3, \\ n \le N}} \frac{2\Lambda(n)}{n} - \ln(N) \right) &= -\gamma - \frac{\ln(3)}{2} - \frac{L'_{-3}(1)}{L_{-3}(1)} \\ &= -\gamma - \frac{\ln(3)}{2} - \ln\left(2\pi e^{\gamma} \frac{\Gamma(\frac{2}{3})^3}{\Gamma(\frac{1}{3})^3} \right), \end{aligned}$$

$$\begin{aligned} \lim_{N \to \infty} \left(\sum_{\substack{n \equiv 1 \bmod 4, \\ n \le N}} \frac{2\Lambda(n)}{n} - \ln(N) \right) &= -\gamma - \ln(2) - \frac{L'_{-4}(1)}{L_{-4}(1)} \\ &= -\gamma - \ln(2) - \ln\left(2\pi e^{\gamma} \frac{\Gamma(\frac{3}{4})^2}{\Gamma(\frac{1}{4})^2} \right), \end{aligned}$$

which imply that [5]

$$\sum_{n\equiv 1 \bmod 3} \frac{2\Lambda(n)-3}{n} = -3\gamma + \frac{\ln(3)}{2} - \frac{\sqrt{3}\pi}{6} - 4\ln(2\pi) + 6\ln(\Gamma(1/3)),$$

$$\sum_{n\equiv 1 \bmod 4} \frac{2\Lambda(n)-4}{n} = -3\gamma - \ln(2) - \frac{\pi}{2} - 3\ln(2\pi) + 4\ln(\Gamma(1/4)).$$

Here is a more complicated generalization. Define

$$\Lambda_{1,3}(n) = \begin{cases} \ln(p) & \text{if } n = p^m \text{ for some prime } p \equiv 1 \bmod 3 \text{ and integer } m \geq 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$\Lambda_{1,4}(n) = \begin{cases} \ln(p) & \text{if } n = p^m \text{ for some prime } p \equiv 1 \bmod 4 \text{ and integer } m \geq 1, \\ 0 & \text{otherwise;} \end{cases}$$

$$\theta_{1,3}(x) = \sum_{\substack{p\leq x,\\ p\equiv 1 \bmod 3}} \ln(p), \quad \theta_{1,4}(x) = \sum_{\substack{p\leq x,\\ p\equiv 1 \bmod 4}} \ln(p);$$

$$\psi_{1,3}(x) = \sum_{n\leq x} \Lambda_{1,3}(n), \quad \psi_{1,4}(x) = \sum_{n\leq x} \Lambda_{1,4}(n).$$

Just as [1]

$$\sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s} = -\frac{\zeta'(s)}{\zeta(s)} \sim \frac{1}{s-1} - \gamma \sim \zeta(s) - 2\gamma,$$

we have [9]

$$\begin{aligned} 2\sum_{n=1}^{\infty} \frac{\Lambda_{1,3}(n)}{n^s} &= -\frac{\zeta'(s)}{\zeta(s)} - \frac{L'_{-3}(s)}{L_{-3}(s)} - \frac{\ln(3)}{3^s-1} - 2\sum_{p\equiv 2 \bmod 3} \frac{\ln(p)}{p^{2s}-1} \\ &\sim \zeta(s) - 2\gamma - \frac{L'_{-3}(s)}{L_{-3}(s)} - \frac{\ln(3)}{3^s-1} - 2\sum_{p\equiv 2 \bmod 3} \frac{\ln(p)}{p^{2s}-1} \end{aligned}$$

as $s \to 1$. On the one hand,

$$\sum_{n=1}^{\infty} \frac{2\Lambda_{1,3}(n)-1}{n} = -2\gamma - \frac{L'_{-3}(1)}{L_{-3}(1)} - \frac{\ln(3)}{2} - 2\sum_{p\equiv 2 \bmod 3} \frac{\ln(p)}{p^2-1}$$

but on the other hand,

$$\begin{aligned} \sum_{n\leq N} \frac{2\Lambda_{1,3}(n)-1}{n} &\sim 2\sum_{\substack{p\leq N,\\ p\equiv 1 \bmod 3}} \frac{\ln(p)}{p} + 2\sum_{\substack{p\leq N,\\ m\geq 2,\\ p\equiv 1 \bmod 3}} \frac{\ln(p)}{p^m} - \sum_{n\leq N} \frac{1}{n} \\ &\sim \ln(N) + c_{1,3} + 2\sum_{p\equiv 1 \bmod 3} \frac{\ln(p)}{p(p-1)} - \ln(N) - \gamma \\ &\sim -\gamma + c_{1,3} + 2\sum_{p\equiv 1 \bmod 3} \frac{\ln(p)}{p(p-1)} \end{aligned}$$

as $N \to \infty$. It follows that

$$\begin{aligned}
1+\int_1^\infty \frac{2\theta_{1,3}(x)-x}{x^2}dx &= \lim_{N\to\infty}\left(2\sum_{\substack{p\le N,\\ p\equiv 1 \bmod 3}}\frac{\ln(p)}{p}-\ln(N)\right)=c_{1,3}\\
&= -\gamma-\frac{L_{-3}'(1)}{L_{-3}(1)}-\frac{\ln(3)}{2}-2\sum_{p\equiv 2 \bmod 3}\frac{\ln(p)}{p^2-1}\\
&\quad -2\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p-1)}\\
&= -2.3754945198....
\end{aligned}$$

Similarly,

$$\begin{aligned}
1+\int_1^\infty \frac{2\theta_{1,4}(x)-x}{x^2}dx &= \lim_{N\to\infty}\left(2\sum_{\substack{p\le N,\\ p\equiv 1 \bmod 4}}\frac{\ln(p)}{p}-\ln(N)\right)=c_{1,4}\\
&= -\gamma-\frac{L_{-4}'(1)}{L_{-4}(1)}-\ln(2)-2\sum_{p\equiv 3 \bmod 4}\frac{\ln(p)}{p^2-1}\\
&\quad -2\sum_{p\equiv 1 \bmod 4}\frac{\ln(p)}{p(p-1)}\\
&= -2.2248371388....
\end{aligned}$$

A simple series acceleration technique [10] arises from the identity

$$\frac{1}{p(p-1)}-\frac{1}{p^2-1}=\frac{1}{p(p^2-1)};$$

hence

$$\begin{aligned}
\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p-1)} &= \sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p^2-1)}+\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p^2-1}\\
&= \sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p^2-1)}+\left(\sum_{p}\frac{\ln(p)}{p^2-1}-\sum_{p\equiv 2 \bmod 3}\frac{\ln(p)}{p^2-1}-\frac{\ln(3)}{8}\right);
\end{aligned}$$

hence

$$\sum_{p\equiv 2 \bmod 3}\frac{\ln(p)}{p^2-1}+\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p-1)}=\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p^2-1)}-\frac{\zeta'(2)}{\zeta(2)}-\frac{\ln(3)}{8};$$

hence

$$c_{1,3}=-2\gamma-4\log(2\pi)+\frac{9\log(3)}{8}+6\log(\Gamma(1/3))+\frac{\zeta'(2)}{\zeta(2)}-2\sum_{p\equiv 1 \bmod 3}\frac{\ln(p)}{p(p^2-1)}.$$

Similarly,

$$c_{1,4} = -2\gamma - 3\log(2\pi) + \frac{\log(2)}{3} + 4\log(\Gamma(1/4)) + \frac{\zeta'(2)}{\zeta(2)} - 2\sum_{p\equiv 1 \bmod 4} \frac{\ln(p)}{p(p^2-1)}.$$

More complex acceleration techniques yield [9]

$$\sum_{p\equiv 2 \bmod 3} \frac{\ln(p)}{p^2-1} = 0.3516478132..., \qquad \sum_{p\equiv 3 \bmod 4} \frac{\ln(p)}{p^2-1} = 0.2287363531...,$$

which permit numerical evaluations such as

$$\begin{aligned} 1 + \int_1^\infty \frac{2\psi_{1,3}(x) - x}{x^2}dx &= \lim_{N\to\infty}\left(\sum_{n\le N}\frac{2\Lambda_{1,3}(n)}{n} - \ln(N)\right) \\ &= -\gamma - \frac{L'_{-3}(1)}{L_{-3}(1)} - \frac{\ln(3)}{2} - 2\sum_{p\equiv 2 \bmod 3}\frac{\ln(p)}{p^2-1} \\ &= -2(1.0990495258...), \end{aligned}$$

$$\begin{aligned} 1 + \int_1^\infty \frac{2\psi_{1,4}(x) - x}{x^2}dx &= \lim_{N\to\infty}\left(\sum_{n\le N}\frac{2\Lambda_{1,4}(n)}{n} - \ln(N)\right) \\ &= -\gamma - \frac{L'_{-4}(1)}{L_{-4}(1)} - \ln(2) - 2\sum_{p\equiv 3 \bmod 4}\frac{\ln(p)}{p^2-1} \\ &= -2(0.9867225683...) \end{aligned}$$

and

$$\sum_{p\equiv 1 \bmod 3} \frac{\ln(p)}{p(p-1)} = 0.0886977340..., \qquad \sum_{p\equiv 1 \bmod 4} \frac{\ln(p)}{p(p-1)} = 0.1256960010....$$

The estimates $-2.375...$ and $-2.224...$ for the theta function integrals are also found in [11, 12]. A parallel analysis of integrals involving

$$\theta_{2,3}(x) = \sum_{\substack{p\le x,\\ p\equiv 2 \bmod 3}} \ln(p), \quad \theta_{3,4}(x) = \sum_{\substack{p\le x,\\ p\equiv 3 \bmod 4}} \ln(p)$$

could be done as well.

Another type of error estimate was provided by McCurley [13]. The maximum value of $\theta_{2,3}(x)/x$ occurs at $x = 1619$ and, further, $\theta_{2,3}(x) < 0.50933118\,x$ for all x. This result is essentially best possible. By contrast, the maximum value of $\theta_{1,3}(x)/x$ is not known! (For $x \le 10^8$, it occurs at $x = 52553329$.) It can be shown that $\theta_{1,3}(x) < 0.5040354\,x$ for all x, but improvement is likely. Sharp analyses of $\theta_{3,4}(x)$ and $\theta_{1,4}(x)$ as such seem still to be open.

The maximum value of $\psi(x)/x$ occurs at $x=113$ and $\psi(x)<1.03882058\,x$ always [2]. Montgomery [14] conjectured that

$$\liminf_{x\to\infty}\frac{\psi(x)-x}{\sqrt{x}\ln(\ln(\ln(x)))^2}=-\frac{1}{2\pi},\quad \limsup_{x\to\infty}\frac{\psi(x)-x}{\sqrt{x}\ln(\ln(\ln(x)))^2}=\frac{1}{2\pi}.$$

Let $M(x)=\sum_{n\le x}\mu(x)$; Odlyzko & te Riele [15] proved that

$$\liminf_{x\to\infty}\frac{M(x)}{\sqrt{x}}<-1.009,\quad \limsup_{x\to\infty}\frac{M(x)}{\sqrt{x}}>1.06.$$

The precise growth rate of $M(x)$ has been the subject of speculation [16–18]. Gonek and Ng [19, 20] independently conjectured that

$$\liminf_{x\to\infty}\frac{M(x)}{\sqrt{x}\ln(\ln(\ln(x)))^{5/4}}=-C,\quad \limsup_{x\to\infty}\frac{M(x)}{\sqrt{x}\ln(\ln(\ln(x)))^{5/4}}=C$$

for some positive constant C. A proof of this or of Montgomery's conjecture would be sensational!

The second-order Landau–Ramanujan constant for counting integers of the form a^2+3b^2 is [21]

$$\frac{1}{2}\left(1-\frac{\gamma}{2}-\frac{1}{2}\frac{L'_{-3}(1)}{L_{-3}(1)}+\frac{\ln(3)}{4}+\sum_{p\equiv 2\bmod 3}\frac{\ln(p)}{p^2-1}\right)=0.5767761224...$$

and the (classical) second-order Landau–Ramanujan constant for counting integers of the form a^2+b^2 is

$$\frac{1}{2}\left(1-\frac{\gamma}{2}-\frac{1}{2}\frac{L'_{-4}(1)}{L_{-4}(1)}+\frac{\ln(2)}{2}+\sum_{p\equiv 3\bmod 4}\frac{\ln(p)}{p^2-1}\right)=0.5819486593....$$

The fact that $0.576...<0.581...$ resolves a question raised by Shanks & Schmid [22, 23]. Further, the second-order LR constant corresponding to a^2+2b^2 is [24]

$$\frac{1}{2}\left(1-\frac{\gamma}{2}-\frac{1}{2}\frac{L'_{-8}(1)}{L_{-8}(1)}+\frac{\ln(2)}{2}+\sum_{p\equiv 5,7\bmod 8}\frac{\ln(p)}{p^2-1}\right)=0.6093010224...$$

and the second-order LR constant corresponding to a^2-2b^2 is

$$\frac{1}{2}\left(1-\frac{\gamma}{2}-\frac{1}{2}\frac{L'_{8}(1)}{L_{8}(1)}+\frac{\ln(2)}{2}+\sum_{p\equiv 3,5\bmod 8}\frac{\ln(p)}{p^2-1}\right)=0.5045371359....$$

The fact that $0.609...>0.581...>0.504...$ verifies an assertion in [22]; we used the Selberg–Delange method and formulas in [25] to deduce the preceding expressions for $a^2\pm 2b^2$. See also [26].

[1] E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, 2nd ed., Chelsea, 1953, pp. 197–203; MR0068565 (16,904d).
[2] J. B. Rosser and L. Schoenfeld, Approximate formulas for some functions of prime numbers, *Illinois J. Math.* 6 (1962) 64-94; MR0137689 (25 #1139).
[3] W. Ellison and F. Ellison, *Prime Numbers*, Wiley, 1985, p. 111; MR0814687 (87a:11082).
[4] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003418, A014963, A034386.
[5] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[6] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.
[7] V. Nevanlinna, On constants connected with the prime number theorem for arithmetic progressions, *Annales Academiae Scientiarum Fennicae Ser. A I* (1973) n. 539, 1–11; MR0342473 (49 #7219).
[8] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[9] P. Moree, Chebyshev's bias for composite numbers with restricted prime divisors, *Math. Comp.* 73 (2004) 425–449; MR2034131 (2005b:11154).
[10] X. Gourdon and P. Sebah, Some constants from number theory, Numbers, Constants and Computation.
[11] M. P. Young, Lower-order terms of the 1-level density of families of elliptic curves, *Int. Math. Res. Notices* (2005) 587–633; math.NT/0408359; MR2147004 (2006c:11076).
[12] S. J. Miller, Lower order terms in the 1-level density for families of holomorphic cuspidal newforms, *Acta Arith.* 137 (2009) 51–98; arXiv:0704.0924; MR2481982 (2010f:11146).
[13] K. S. McCurley, Explicit estimates for $\theta(x; 3, l)$ and $\psi(x; 3, l)$, *Math. Comp.* 42 (1984) 287–296; MR0726005 (85g:11085).
[14] H. L. Montgomery, The zeta function and prime numbers, *Queen's Papers in Pure and Applied Mathematics*, n. 54, Proc. 1979 Queen's Number Theory Conf., ed. P. Ribenboim, 1980, pp. 1–31; MR0634679 (82k:10047).
[15] A. M. Odlyzko and H. J. J. te Riele, Disproof of the Mertens conjecture, *J. Reine Angew. Math.* 357 (1985) 138–160; MR0783538 (86m:11070).
[16] I. J. Good and R. F. Churchhouse, The Riemann hypothesis and pseudorandom features of the Möbius sequence, *Math. Comp.* 22 (1968) 857–861; MR0240062 (39 #1416).
[17] B. Saffari, Sur la fausseté de la conjecture de Mertens, *C. R. Acad. Sci. Paris Sér. A-B* 271 (1970) A1097–A1101; MR0280447 (43 #6167).
[18] T. Kotnik and J. van de Lune, On the order of the Mertens function, *Experim. Math.* 13 (2004) 473–481; MR2118272 (2005i:11139).
[19] N. Ng, The distribution factor of values of the summatory function of the Möbius function, *CMS Notes*, v. 34 (2002) n. 5, 5–8.
[20] N. Ng, The distribution of the summatory function of the Möbius function, *Proc. London Math. Soc.* 89 (2004) 361–389; MR2078705 (2005f:11215).
[21] S. R. Finch, Landau-Ramanujan constant, *first volume*, pp. 98–104.
[22] D. Shanks and L. P. Schmid, Variations on a theorem of Landau. I, *Math. Comp.* 20 (1966) 551–569; MR0210678 (35 #1564).
[23] P. Moree and H. J. J. te Riele, The hexagonal versus the square lattice, *Math. Comp.* 73 (2004) 451–473; arXiv:math/0204332; MR2034132 (2005b:11155).
[24] P. Sebah, Two modulo 8 prime sums, unpublished note (2009).
[25] S. Finch, Quartic and octic characters modulo *n*, arXiv:0907.4894.
[26] S. R. Finch, Mertens' formula, *this volume*, §1.26.

1.26 Mertens' Formula

An elegant generalization of Mertens' famous formula appears in [1]:

$$\lim_{x\to\infty} \ln(x)^{1/\varphi(k)} \prod_{\substack{p\le x,\\ p\equiv \ell \bmod k}} \left(1-\frac{1}{p}\right) = \left[e^{-\gamma} \prod_{p} \left(1-\frac{1}{p}\right)^{\alpha(p;k,\ell)}\right]^{1/\varphi(k)}$$

where φ is the Euler totient function, γ is the Euler–Mascheroni constant [2], and $\alpha(p;k,\ell)$ is equal to $\varphi(k)-1$ if $p\equiv \ell \bmod k$ and is -1 otherwise. This constitutes a vast simplification of earlier such formulas [3, 4]. Computing the constant $e^{-\gamma}\Lambda_{k,\ell} = (0.5614594835...)\Lambda_{k,\ell}$ inside the square brackets, as well as the related limit [5]:

$$M_{k,\ell} = \lim_{x\to\infty}\left(\sum_{\substack{p\le x\\ p\equiv \ell \bmod k}} \frac{1}{p} - \frac{1}{\varphi(k)}\ln(\ln(x))\right)$$

will occupy us for the remainder of this essay.

Let $\zeta(s)$ denote the Riemann zeta function and

$$P_{k,\ell}(s) = \sum_{p\equiv \ell \bmod k} \frac{1}{p^s}$$

denote the $(k,\ell)^{\text{th}}$ **prime zeta function** for $\mathrm{Re}(s)>1$. Clearly $\Lambda_{1,0}=1$; to efficiently compute $M_{1,0}$, we utilize the series [6–8]

$$P_{1,0}(s) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \ln\left(\zeta(ns)\right).$$

The numerical evaluation of other $P_{k,\ell}(s)$ will be discussed momentarily. For now, note that

$$\frac{-\gamma+\ln(\Lambda_{k,\ell})}{\varphi(k)} = \lim_{x\to\infty}\left(\sum_{\substack{p\le x\\ p\equiv \ell \bmod k}} \ln\left(1-\frac{1}{p}\right) + \frac{1}{\varphi(k)}\ln(\ln(x))\right);$$

hence

$$\begin{aligned} M_{k,\ell} + \frac{\ln(\Lambda_{k,\ell})-\gamma}{\varphi(k)} &= \sum_{p\equiv \ell \bmod k}\left(\ln\left(1-\frac{1}{p}\right)+\frac{1}{p}\right) \\ &= -\sum_{p\equiv \ell \bmod k}\left(\frac{1}{2p^2}+\frac{1}{3p^3}+\frac{1}{4p^4}+\cdots\right) \\ &= -\sum_{n=2}^{\infty}\frac{P_{k,\ell}(n)}{n}; \end{aligned}$$

hence

$$M_{1,0} = \gamma - \sum_{n=2}^{\infty} \frac{P_{1,0}(n)}{n} = 0.2614972128....$$

In the following two sections, we will discuss the cases $k=3$ and $k=4$. In both cases, $\varphi(k)=2$, which implies that

$$\begin{aligned}
\Lambda_{k,1} &= \frac{k}{2} \prod_{p\equiv 1 \bmod k} \left(1-\frac{1}{p}\right) \cdot \prod_{p\equiv -1 \bmod k} \left(1-\frac{1}{p}\right)^{-1} \\
&= \frac{k}{2} \frac{1}{L_{-k}(1)} \prod_{p\equiv -1 \bmod k} \left(1+\frac{1}{p}\right)^{-1} \cdot \prod_{p\equiv -1 \bmod k} \left(1-\frac{1}{p}\right)^{-1} \\
&= \frac{k}{2} \frac{1}{L_{-k}(1)} \prod_{p\equiv -1 \bmod k} \left(1-\frac{1}{p^2}\right)^{-1}
\end{aligned}$$

where L_{-k} is Dirichlet's L-series associated to $(-k/\cdot)$. The infinite product can be evaluated via the $(k,-1)^{\text{th}}$ prime zeta function since

$$\begin{aligned}
\ln\left(\prod_{p\equiv \ell \bmod k} \left(1-\frac{1}{p^2}\right)\right) &= \sum_{p\equiv \ell \bmod k} \left(\ln\left(1+\frac{1}{p}\right) + \ln\left(1-\frac{1}{p}\right)\right) \\
&= -\sum_{p\equiv \ell \bmod k} \left(\frac{1}{p^2} + \frac{1}{2p^4} + \frac{1}{3p^6} + \cdots\right) \\
&= -\sum_{n=1}^{\infty} \frac{P_{k,\ell}(2n)}{n}.
\end{aligned}$$

Thus we first compute $\Lambda_{3,1}$ and $\Lambda_{4,1}$, and then $M_{3,1}$ and $M_{4,1}$.

Let χ_0 denote the principal character $\bmod\, k$ and χ_1 denote the nonprincipal character $\bmod\, k$ (χ_1 is unique since $k=3$ or $k=4$). In order to evaluate $P_{k,1}(s)$ and $P_{k,-1}(s)$, the associated Dirichlet L-series:

$$L_{\chi_j}(s) = \sum_{n=1}^{\infty} \frac{\chi_j(n)}{n^s} = \frac{1}{k^s}\left(\chi_j(1)\zeta\left(s,\tfrac{1}{k}\right) + \chi_j(-1)\zeta\left(s,1-\tfrac{1}{k}\right)\right), \quad j=0,1$$

are required, where $\zeta(s,a)$ is the Hurwitz zeta-function. It can be shown that [9]

$$P_{k,-1}(s) = \frac{1}{2}\sum_{n=0}^{\infty} \frac{\mu(2n+1)}{2n+1} \ln\left(\frac{L_{\chi_0}((2n+1)s)}{L_{\chi_1}((2n+1)s)}\right),$$

$$P_{k,1}(s) = \frac{1}{2}\sum_{n=0}^{\infty} \frac{\mu(2n+1)}{2n+1} \ln\left(\frac{L_{\chi_0}((2n+1)s)L_{\chi_1}((2n+1)s)}{L_{\chi_0}((4n+2)s)}\right).$$

We will additionally exhibit $\prod_{p\equiv 1 \bmod k}\left(1-p^{-2}\right)$ and $\prod_{p\equiv -1 \bmod k}\left(1-p^{-2}\right)$, since these are also easily available.

1.26.1 *Residue Classes Mod 3*

The two characters modulo 3 are

$$\chi_0(n)|_{n=1,2,3} = \{1, 1, 0\}, \quad \chi_1(n)|_{n=1,2,3} = \{1, -1, 0\};$$

thus $L_{\chi_0}(s) = \zeta(s)(1 - 1/3^s)$ and $L_{\chi_1}(s) = L_{-3}(s)$. We have [10]

$$\prod_{p\equiv 1 \bmod 3} \left(1 - \frac{1}{p^2}\right) = 0.9671040753... = \frac{9\sqrt{3}}{8}(0.7044984335...)^2 = \frac{27\sqrt{3}}{2\pi^2}K_3^2,$$

$$\prod_{p\equiv 2 \bmod 3} \left(1 - \frac{1}{p^2}\right) = 0.7071813747... = \frac{9\sqrt{3}}{2}(0.3012165544...)^2 = \frac{\sqrt{3}}{6}\frac{1}{K_3^2}$$

where $K_3 = 0.6389094054...$ is the Landau–Ramanujan constant for counting integers of the form $a^2 + 3b^2$ [11]. Also $\Lambda_{3,1} = 27K_3^2/\pi$ and therefore

$$\lim_{x\to\infty} \sqrt{\ln(x)} \prod_{\substack{p\le x \\ p\equiv 1 \bmod 3}} \left(1 - \frac{1}{p}\right) = 3\sqrt{\frac{3}{\pi}}e^{-\gamma/2}K_3 = 1.4034774468...,$$

$$\lim_{x\to\infty} \sqrt{\ln(x)} \prod_{\substack{p\le x \\ p\equiv 2 \bmod 3}} \left(1 - \frac{1}{p}\right) = \frac{1}{2}\sqrt{\frac{\pi}{3}}e^{-\gamma/2}\frac{1}{K_3} = 0.6000732161...,$$

$$M_{3,1} = \frac{\gamma}{2} - \ln\left(3\sqrt{\frac{3}{\pi}}K_3\right) + \sum_{p\equiv 1 \bmod 3}\left[\ln\left(1 - \frac{1}{p}\right) + \frac{1}{p}\right] = -0.3568904795...,$$

$$M_{3,2} = \frac{\gamma}{2} - \ln\left(\frac{1}{2}\sqrt{\frac{\pi}{3}}\frac{1}{K_3}\right) + \sum_{p\equiv 2 \bmod 3}\left[\ln\left(1 - \frac{1}{p}\right) + \frac{1}{p}\right] = 0.2850543590....$$

1.26.2 *Residue Classes Mod 4*

An alternative approach is given in [12]. The two characters modulo 4 are

$$\chi_0(n)|_{n=1,2,3,4} = \{1, 0, 1, 0\}, \quad \chi_1(n)|_{n=1,2,3,4} = \{1, 0, -1, 0\};$$

thus $L_{\chi_0}(s) = \zeta(s)(1 - 1/2^s)$ and $L_{\chi_1}(s) = L_{-4}(s)$. We have [10]

$$\prod_{p\equiv 1 \bmod 4} \left(1 - \frac{1}{p^2}\right) = 0.9468064071... = 4(0.4865198883...)^2 = \frac{16}{\pi^2}K_1^2,$$

$$\prod_{p\equiv 3 \bmod 4} \left(1 - \frac{1}{p^2}\right) = 0.8561089817... = 8\,(0.3271293669...)^2 = \frac{1}{2}\frac{1}{K_1^2}$$

where $K_1 = 0.7642236535...$ is the (classical) Landau–Ramanujan constant for counting integers of the form $a^2 + b^2$ [11]. Also $\Lambda_{4,1} = 16K_1^2/\pi$ and therefore

$$\lim_{x\to\infty} \sqrt{\ln(x)} \prod_{\substack{p\leq x \\ p\equiv 1 \bmod 4}} \left(1 - \frac{1}{p}\right) = \frac{4}{\sqrt{\pi}} e^{-\gamma/2} K_1 = 1.2923041571...,$$

$$\lim_{x\to\infty} \sqrt{\ln(x)} \prod_{\substack{p\leq x \\ p\equiv 3 \bmod 4}} \left(1 - \frac{1}{p}\right) = \frac{\sqrt{\pi}}{2} e^{-\gamma/2} \frac{1}{K_1} = 0.8689277682...,$$

$$M_{4,1} = \frac{\gamma}{2} - \ln\left(\frac{4}{\sqrt{\pi}} K_1\right) + \sum_{p\equiv 1 \bmod 4} \left[\ln\left(1 - \frac{1}{p}\right) + \frac{1}{p}\right] = -0.2867420562...,$$

$$M_{4,3} = \frac{\gamma}{2} - \ln\left(\frac{\sqrt{\pi}}{2} \frac{1}{K_1}\right) + \sum_{p\equiv 3 \bmod 4} \left[\ln\left(1 - \frac{1}{p}\right) + \frac{1}{p}\right] = 0.0482392690....$$

Some low-precision results are known [13] for the residue classes mod 6 and 8; it would be good to repeat these calculations (using the prime zeta function, as above) to high accuracy.

Addendum Languasco & Zaccagnini [14–16] proved new formulas and greatly extended the preceding calculations, confirming our values for $M_{k,\ell}$ and for $(e^{-\gamma}\Lambda_{k,\ell})^{1/\varphi(k)}$ (what they call $C_{k,\ell}$) when $k = 3$ and $k = 4$.

[1] A. Languasco and A. Zaccagnini, A note on Mertens' formula for arithmetic progressions, *J. Number Theory* 127 (2007) 37–46; MR2351662 (2009g:11136).

[2] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.

[3] K. S. Williams, Mertens' theorem for arithmetic progressions, *J. Number Theory* 6 (1974) 353–359; MR0364137 (51 #392).

[4] E. A. Vasil'kovskaja, Mertens' formula for an arithmetic progression (in Russian), *Voprosy Mat. Sbornik Naucn. Trudy Taskent. Gos. Univ.* (1977) n. 548, 14–17, 139-140; MR0565984 (58 #27848).

[5] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.

[6] C. -E. Fröberg, On the prime zeta function, *Nordisk Tidskr. Informationsbehandling (BIT)* 8 (1968) 187–202; MR0236123 (38 #4421).

[7] H. Cohen, High precision computation of Hardy-Littlewood constants, unpublished note (1999).

[8] X. Gourdon and P. Sebah, Some constants from number theory, Numbers, Constants and Computation.

[9] S. Finch and P. Sebah, Squares and cubes modulo *n*, math.NT/0604465.

[10] P. Moree, Chebyshev's bias for composite numbers with restricted prime divisors, *Math. Comp.* 73 (2004) 425–449; MR2034131 (2005b:11154).

[11] S. R. Finch, Landau-Ramanujan constant, *first volume*, pp. 98–104.

[12] S. Uchiyama, On some products involving primes, *Proc. Amer. Math. Soc.* 28 (1971) 629–630; MR0277494 (43 #3227).

[13] E. Grosswald, Some number theoretical products, *Rev. Colombiana Mat.* 21 (1987) 231–242; MR0988131 (90e:11129).
[14] A. Languasco and A. Zaccagnini, On the constant in the Mertens product for arithmetic progressions. I, Identities, *Funct. Approx. Comment. Math.* 42 (2010) 17–27; arXiv:0706.2807; MR2640766 (2011b:11127)
[15] A. Languasco and A. Zaccagnini, On the constant in the Mertens product for arithmetic progressions. II, Numerical values, *Math. Comp.* 78 (2009) 315–326; arXiv:0712.1665; MR2448709 (2010g:11164)
[16] A. Languasco and A. Zaccagnini, Computing the Mertens and Meissel-Mertens constants for sums over arithmetic progressions, *Experim. Math.* 19 (2010) 279–284; arXiv:0906.2132; MR2743571 (2011j:11247).

1.27 Cyclotomic Polynomials

Let

$$\prod_{\gcd(j,n)=1}\left(x-e^{2\pi j\,i/n}\right)=\sum_{k=0}^{\varphi(n)}a_n(k)x^k$$

denote the n^{th} **cyclotomic polynomial**, where $\varphi(n)$ is Euler's totient function [1], i denotes the imaginary unit, and the product is taken over all integers $1\leq j\leq n$ coprime with n. The coefficients $a_n(k)$ are always integers. Define

$$A(n)=\max_k|a_n(k)|,$$

the largest coefficient of the n^{th} polynomial in absolute value; and

$$B(k)=\max_n|a_n(k)|,$$

the largest k^{th} coefficient in absolute value (taken over all polynomials). A simple argument gives $B(k)\leq p(k)$, where $p(k)$ is the number of integer partitions of k, hence $B(k)$ is finite.

Vaughan [2–4] proved that

$$\limsup_{n\to\infty}\frac{\ln(\ln(A(n)))}{\ln(n)/\ln(\ln(n))}=\ln(2)$$

(a maximal order) and Bachman [5] proved that

$$\lim_{k\to\infty}\frac{\ln(k)^{1/4}}{\sqrt{k}}\ln(B(k))=C=1.5394450081...$$

(an asymptotic result). The constant C is related to the solution of an interesting optimization problem involving L-series [6]. Define $\iota(D)=2$ if $D>0$ and $\iota(D)=1$ if $D<0$. Over all fundamental discriminants D, it can be proved that $D=12$ maximizes the quantity

$$\sqrt{\frac{\iota(D)}{\pi\,\varphi(D)}}\frac{L_D(2)}{\sqrt{L_D(1)}}\prod_{(D/p)=-1}\left(1-\frac{1}{p^2}\right)^{1/2},$$

where the product is taken over all primes p satisfying a negativity condition on the Legendre symbol. When $D = 12$, this quantity simplifies to

$$\sqrt{\frac{1}{2\pi}} \frac{\pi^2/(6\sqrt{3})}{\sqrt{\ln(2+\sqrt{3})/\sqrt{3}}} \prod_{\substack{p\equiv 5 \text{ or } 7 \\ \text{mod } 12}} \left(1 - \frac{1}{p^2}\right)^{1/2} = 0.4189414873...$$

$$= 2^{-5/2}C^2.$$

It is hoped that other statistics (for example, means and variances) summarizing the coefficient array $a_n(k)$ might be feasible. See [7, 8] for work in this area. The precise estimate of C is due to Sebah [9].

[1] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.
[2] R. C. Vaughan, Bounds for the coefficients of cyclotomic polynomials, *Michigan Math. J.* 21 (1974) 289–295; MR0364141 (51 #396).
[3] P. T. Bateman, C. Pomerance and R. C. Vaughan, On the size of the coefficients of the cyclotomic polynomial, *Topics in Classical Number Theory*, v. I, Proc. 1981 Budapest conf., Colloq. Math. Soc. János Bolyai 34, North-Holland, 1984, pp. 171–202; MR0781138 (86e:11089).
[4] C. Pomerance and N. C. Ryan, Maximal height of divisors of $x^n - 1$, *Illinois J. Math.* 51 (2007) 597–604; MR2342677 (2008j:12012).
[5] G. Bachman, *On the Coefficients of Cyclotomic Polynomials*, Amer. Math. Soc., 1993, pp. 1–10, 74–80; MR1172916 (94d:11068).
[6] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[7] H. Möller, Über die i-ten Koeffizienten der Kreisteilungspolynome, *Math. Annalen* 188 (1970) 26–38; MR0266899 (42 #1801).
[8] Y. Gallot, P. Moree and H. Hommersom, Value distribution of cyclotomic polynomial coefficients, *Unif. Distrib. Theory* 6 (2011) 177–206; arXiv:0803.2483; MR2904047.
[9] P. Sebah, An infinite product, unpublished note (2008).

1.28 Minkowski–Alkauskas Constant

In addition to examining [1]

$$?\left(0 + \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots\right) = \sum_{k=1}^{\infty} (-1)^{k-1} 2^{-(a_1+a_2+\cdots+a_k-1)},$$

we study [2]

$$F\left(a_0 + \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots\right) = \sum_{k=1}^{\infty} (-1)^{k-1} 2^{-(a_0+a_1+a_2+\cdots+a_k)}.$$

The former is the original Minkowski question mark function, a self-map of $[0,1]$; the latter is defined on the nonnegative real line with $2F(x) = ?(x)$ for all $x \in [0,1]$. In particular,

$$F(0) = 0, \quad F(\tfrac{1}{2}) = \tfrac{1}{4}, \quad F(1) = \tfrac{1}{2}, \quad F(\sqrt{2}) = \tfrac{3}{5},$$

$$F(\tfrac{1+\sqrt{5}}{2}) = \tfrac{2}{3}, \quad F(2) = \tfrac{3}{4}, \quad F(3) = \tfrac{7}{8}, \quad \lim_{x\to\infty} F(x) = 1^-.$$

The distribution F is continuous, strictly increasing, singular, and uniquely determined by the functional equation

$$2F(x) = \begin{cases} F(x-1)+1 & \text{if } x \geq 1, \\ F\left(\dfrac{x}{1-x}\right) & \text{if } 0 \leq x < 1. \end{cases}$$

Define moments

$$M_\ell = \int\limits_0^\infty x^\ell dF(x), \quad m_\ell = \int\limits_0^1 x^\ell d?(x)$$

then $m_1 = M_1 - 1 = 1/2$ follows easily. Similar closed-form expressions for

$$m_2 = M_2 - 4 = 0.2909264764...,$$

$$m_4 = M_4 - 24m_2 - 100 = 0.1269922584...$$

presently do not exist, although progress has recently been made [3]. It is known that

$$2m_3 = 3m_2 - 1/2 = 2(0.1863897146...),$$

$$2M_3 = 9m_2 + 69/2, \quad 2m_5 = 5m_4 - 5m_2 + 1$$

and analogous relations hold for higher-order moments. Hence calculating m_2, $m_4, \ldots$ to high precision is important for understanding $m_3, m_5, \ldots$.

Alkauskas [4, 5] proved the following asymptotic formula:

$$\begin{aligned} m_\ell &\sim \sqrt[4]{4\pi^2\ln(2)} \cdot c \cdot \left(e^{-2\sqrt{\ln(2)}}\right)^{\sqrt{\ell}} \ell^{1/4} \\ &\sim (2.3562298899...)(0.1891699952...)^{\sqrt{\ell}}\ell^{1/4} \end{aligned}$$

as $\ell \to \infty$, where

$$c = \int\limits_0^1 2^x(1-F(x))dx = 1.0301995633... = \frac{1.4281598455...}{2\ln(2)}.$$

This is a fascinating result, especially because $m_2, m_4, \ldots$ remain so mysterious! One would not have expected an asymptotic formula for m_ℓ as such to be possible.

An infinite series for m_ℓ that does not explicitly involve continued fractions was unveiled in [6]:

$$\frac{1}{(\ell-1)!}\sum_{n=0}^{\infty} \int \cdots \int_{[0,\infty)^{n+1}} x_0^\ell \cdot \frac{(x_0x_n)^{-1/2} \cdot \prod\limits_{j=0}^{n-1} I_1\left(2\sqrt{x_jx_{j+1}}\right)}{\prod\limits_{j=0}^{n} e^{x_j}\left(2e^{x_j}-1\right)} dx_0 \cdots dx_n$$

where $I_1(z)$ is the modified Bessel function of the first kind. Unfortunately this does not improve upon numerical accuracy found in [3]. Does a simpler formula exist (even if only for $\ell = 2$ or $\ell = 4$)?

Integrals of the form

$$\int_0^1 \cos(2\pi k x)\, d?(x)$$

are evaluated to high precision in [7]; another sample calculation is

$$\pi \int_0^1 (?(x) - x)\cot(\pi x)\, dx = -0.4559592037...,$$

which corresponds to the value of an associated zeta function at unity.

[1] S. R. Finch, Minkowski-Bower constant, *first volume*, pp. 441–443.

[2] G. Alkauskas, The moments of Minkowski question mark function: the dyadic period function, *Glasgow Math. J.* 52 (2010) 41–64; arXiv:0801.0051; MR2587817 (2011d:11007).

[3] G. Alkauskas, The Minkowski question mark function: explicit series for the dyadic period function and moments, *Math. Comp.* 79 (2010) 383–418; addenda/corrigenda 80 (2011) 2445–2454; arXiv:0805.1717; MR2552232 (2010k:11006) and MR2813370 (2012d:11009).

[4] G. Alkauskas, An asymptotic formula for the moments of the Minkowski question mark function in the interval $[0, 1]$, *Lithuanian Math. J.* 48 (2008) 357–367; arXiv:0802.2721; MR2470798 (2009i:11115).

[5] R. Bacher, The Stern sequence and moments of Minkowski's question mark function, arXiv:1703.07268.

[6] G. Alkauskas, Semi-regular continued fractions and an exact formula for the moments of the Minkowski question mark function, *Ramanujan J.* 25 (2011) 359–367; arXiv:0912.1039; MR2819722 (2012f:11008).

[7] G. Alkauskas, Fourier-Stieltjes coefficients of the Minkowski question mark function, *Analytic and Probabilistic Methods in Number Theory*, Proc. 2011 Palanga conf., ed. A. Laurinčikas, E. Manstavičius and G. Stepanauskas, TEV, 2012, pp. 19–33; arXiv:1008.4014; MR3025455.

1.29 Two-Colorings of Positive Integers

Let $f : \{1, 2, 3, \ldots\} \to \{-1, 1\}$ be an arbitrary function. Given a threshold $M > 0$, we ask two questions:

- Do there exist integers $a > 0$, $b \geq 0$, $\ell > 0$ such that

$$|f(a+b) + f(2a+b) + f(3a+b) + \cdots + f(\ell a + b)| > M?$$

- Do there exist integers $a > 0$, $\ell > 0$ such that

$$|f(a) + f(2a) + f(3a) + \cdots + f(\ell a)| > M?$$

The answer to the first question is yes. In words, every two-coloring of the positive integers has unbounded discrepancy, taken over the family of arithmetic progressions. Restricting attention to the subset $\{1, 2, 3, \ldots, n\}$, we have [1–5]

$$c\,n^{1/4} \le P(n) = \min_{f} \max_{\substack{a,b,\ell \\ \ell\, a+b \le n}} \left| \sum_{k=1}^{\ell} f(k\,a+b) \right| \le C\,n^{1/4}$$

for all n, with constants $c \ge 1/20$ and $C < \infty$. The lower bound on c was improved to $1/14$ in [6]; no finite upper bound on C is known. It is natural to wonder about the numerical values of

$$\liminf_{n\to\infty} n^{-1/4} P(n), \quad \limsup_{n\to\infty} n^{-1/4} P(n).$$

The second question, due to Erdős [7–9] and Chudakov [10, 11], was answered affirmatively only recently by Tao [12, 13]. It is remarkable that, upon mere constraint to homogeneity ($b = 0$), the problem becomes unimaginably difficult. The existence of near-counterexamples (four are given in [12]) serve to isolate the key difficulty of the problem. More on the buildup to a solution appears shortly.

If we expand the family under consideration, the problem simplifies. For almost all real numbers $\alpha \ge 1$, there exists $\ell > 0$ such that [14–16]

$$|f(\lfloor\alpha\rfloor) + f(\lfloor 2\alpha\rfloor) + f(\lfloor 3\alpha\rfloor) + \cdots + f(\lfloor \ell\,\alpha\rfloor)| > M.$$

Such **quasi-arithmetic progressions** collapse to homogeneous arithmetic progressions when α is an integer. Even though the set S of counterexamples α has measure zero, we definitely know (thanks to Tao) that S avoids all integers. Further, for any $\varepsilon > 0$,

$$d\,n^{1/6} \le Q(n) = \min_{f} \max_{\substack{\alpha,\ell \\ \lfloor \ell\,\alpha\rfloor \le n}} \left| \sum_{k=1}^{\ell} f(\lfloor k\,\alpha\rfloor) \right| \le D\,n^{1/3+\varepsilon}$$

where $d \ge 1/50$ and we speculate whether bounds on $Q(n)$ might someday be significantly improved.

The expression [17, 18]

$$R(n) = \min_{f} \max_{\substack{a,b,\ell \\ a<b \\ \ell+b \le n}} \left| \sum_{k=1}^{\ell} f(k+a) f(k+b) \right|$$

is also interesting and we wonder about the numerical values of

$$\liminf_{n\to\infty} n^{-1/2} R(n), \quad \limsup_{n\to\infty} n^{-1/2} R(n).$$

1.29.1 *Erdős–Chudakov–Tao*

If f is random (independently taking values ± 1 with probability $1/2$ at each integer $1 \le k \le n$), then asymptotically [19]:

$$\mathrm{E}\left(|f(1)+f(2)+f(3)+\cdots+f(n)|\right) \sim \sqrt{\frac{2n}{\pi}}$$

as $n \to \infty$. The use of an average is somewhat deceiving because, for almost all such f,

$$|f(1)+f(2)+f(3)+\cdots+f(n)| \sim \sqrt{2n\ln(\ln(n))}$$

by the law of the iterated logarithm. In words, for typical f, sums are larger than expected. Hence solving the Erdős–Chudakov problem requires an understanding of atypical f, for which sums remain small.

Here are two relevant results obtained prior to Tao's groundbreaking work.

Nikolov & Talwar [20], building on Alon & Kalai [21], showed that the following statement is true for infinitely many positive integers n. There is a set $W \subseteq \{1, \ldots, n\}$ of square-free integers such that, for any $f\colon W \to \{-1, 1\}$, there exists a positive integer a so that

$$\left|\sum_{w \in W,\, a|w} f(w)\right| = n^{1/O(\ln(\ln(n)))}$$

as $n \to \infty$. (If we were permitted to define $f = 0$ outside of W, then the Erdős–Chudakov problem would be solved. The values of f, however, are restricted to ± 1, disallowing such a construction.)

Konev & Lisitsa [22, 23], assisted by computer, exhibited a length 1160 sequence whose discrepancy is bounded by $M = 2$, but proved that such cannot be true for any sequence of length ≥ 1161. Hence the Erdős–Chudakov conjecture (for infinite sequences) is true for $M = 2$. Twenty years earlier, Mathias [11] showed likewise for $M = 1$ via elementary means. A length 13000 sequence whose discrepancy is bounded by $M = 3$ is known; what is the shortest length $L > 13000$ beyond which this cannot be true?

Discussion of completely multiplicative functions and k-regular sequences would take us too far afield [24–27]. We mention the important role played here by the Polymath wiki – which documents massively collaborative online mathematical projects – and highlight the summaries provided in [28, 29].

[1] K. F. Roth, Remark concerning integer sequences, *Acta Arith.* 9 (1964) 257–260; MR0168545 (29 #5806).

[2] J. Beck, Roth's estimate of the discrepancy of integer sequences is nearly sharp, *Combinatorica* 1 (1981) 319–325; MR0647981 (83i:05040).

[3] J. Beck and V. T. Sós, Discrepancy theory, *Handbook of Combinatorics*, v. 2, ed. R. L. Graham, M. Grötschel and L. Lovász, MIT Press, 1995, pp. 1405–1446; MR1373682 (96m:11060).

[4] J. Matoušek and J. Spencer, Discrepancy in arithmetic progressions, *J. Amer. Math. Soc.* 9 (1996) 195–204; MR1311824 (96c:11089).

[5] C. Dartyge, K. Gyarmati and A. Sárközy, On irregularities of distribution of binary sequences relative to arithmetic progressions, I (General results), *Unif. Distrib. Theory* 12 (2017) 55–67; MR3666533.

[6] L. Lovász, Integer sequences and semidefinite programming, *Publ. Math. Debrecen* 56 (2000) 475–479; MR1765994 (2001k:11020).

[7] P. Erdős, Quelques problèmes de théorie des nombres, *Monographies de L'Enseignement Mathématique,* v. 6, L'Enseignement Mathématique, Université, Geneva, 1963, pp. 81–135; MR0158847 (28 #2070).

[8] P. Erdős, Some unsolved problems, *Michigan Math. J.* 4 (1957) 291–300; MR0098702 (20 #5157).

[9] P. Erdős, On the combinatorial problems which I would most like to see solved, *Combinatorica* 1 (1981) 25–42; MR0602413 (82k:05001).

[10] N. G. Tchudakoff, Theory of the characters of number semigroups, *J. Indian Math. Soc.* 20 (1956) 11–15; MR0083515 (18,719e).

[11] A. R. D. Mathias, On a conjecture of Erdős and Čudakov, *Combinatorics, Geometry and Probability*, Proc. 1993 Cambridge conf., ed. B. Bollobás and A. Thomason, Cambridge Univ. Press, 1997, pp. 487–488; MR1476466 (98i:11058).

[12] T. Tao, The Erdős discrepancy problem, *Discrete Analysis* (2016) N1; arXiv:1509.05363; MR3533300.

[13] K. Soundararajan, Tao's resolution of the Erdős discrepancy problem, *Bull. Amer. Math. Soc.* 55 (2018) 81–92; MR3737211.

[14] J. Beck, On irregularities of ± 1-sequences, *Österreichische Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Klasse. Sitzungsberichte. Abteilung II* 195 (1986) 13–23; MR0881327 (88e:11065).

[15] R. Hochberg, Large discrepancy in homogeneous quasi-arithmetic progressions, *Combinatorica* 26 (2006) 47–64; MR2201283 (2006i:11012).

[16] S. Vijay, On the discrepancy of quasi-progressions, *Elec. J. Combin.* 15 (2008) R104; math.CO/0604511; MR2438576 (2009i:05170).

[17] J. Cassaigne, C. Mauduit and A. Sárközy, On finite pseudorandom binary sequences. VII, The measures of pseudorandomness, *Acta Arith.* 103 (2002) 97–118; MR1904866 (2004c:11139).

[18] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira and V. Rödl, Measures of pseudorandomness for finite sequences: minimal values, *Combin. Probab. Comput.* 15 (2006) 1–29; MR2195573 (2006j:60007).

[19] S. R. Finch, Moments of sums, *this volume*, §4.2.

[20] A. Nikolov and K. Talwar, On the hereditary discrepancy of homogeneous arithmetic progressions, *Proc. Amer. Math. Soc.* 143 (2015) 2857–2863; arXiv:1309.6034; MR3336610.

[21] G. Kalai, Erdős discrepancy problem, EDP 22, Gower's Weblog posting (2012).

[22] B. Konev and A. Lisitsa, A SAT attack on the Erdős discrepancy conjecture, *Theory and Applications of Satisfiability Testing – SAT 2014*, Proc. Vienna conf., ed. C. Sinz and U. Egl, Lect. Notes in Comput. Sci., 8561, Springer-Verlag, 2014, pp. 219–226; arXiv:1402.2184; MR3252248.

[23] B. Konev and A. Lisitsa, Computer-aided proof of Erdős discrepancy properties, *Artificial Intelligence* 224 (2015) 103–118; arXiv:1405.3097; MR3332786.

[24] P. Borwein, S. K. K. Choi and M. Coons, Completely multiplicative functions taking values in $\{-1, 1\}$, *Trans. Amer. Math. Soc.* 362 (2010) 6279–6291; arXiv:0809.1691; MR2678974 (2011f:11123).
[25] M. Coons, On the multiplicative Erdős discrepancy problem, arXiv:1003.5388.
[26] J. Vandehey, On multiplicative functions with bounded partial sums, *Integers* 12 (2012) 741–755; arXiv:1108.1487; MR2988544.
[27] J. P. Bell, M. Coons and K. G. Hare, The minimal growth of a k-regular sequence, *Bull. Aust. Math. Soc.* 90 (2014) 195–203; arXiv:1410.5517; MR3252000.
[28] The Erdős discrepancy problem, Polymath 5 online wiki (2010–2012).
[29] W. T. Gowers, Erdős and arithmetic progressions, *Erdős Centennial*, ed. L. Lovász, I. Z. Ruzsa and V. T. Sós, János Bolyai Math. Soc., 2013, pp. 265–287; arXiv:1509.03421; MR3203599.

1.30 Signum Equations and Extremal Coefficients

Let $a(n)$ denote the number of sign choices $+$ and $-$ such that

$$\pm 1 \pm 2 \pm 3 \pm \cdots \pm n = 0$$

and $b(n)$ denote the number of solutions of

$$\varepsilon_1 \cdot 1 + \varepsilon_2 \cdot 2 + \varepsilon_3 \cdot 3 + \cdots + \varepsilon_n \cdot n = 0$$

where each $\varepsilon_j \in \{-1, 0, 1\}$. It can be proved that [1, 2]

$$a(n) \text{ is the coefficient of } x^{n(n+1)/2} \text{ in the polynomial } \prod_{k=1}^{n} \left(1 + x^{2k}\right),$$

$$b(n) \text{ is the coefficient of } x^{n(n+1)/2} \text{ in the polynomial } \prod_{k=1}^{n} \left(1 + x^{k} + x^{2k}\right).$$

Clearly $a(n) = 0$ when $n \equiv 1, 2 \bmod 4$. If we think of sign choices as independent random variables with equal weight on $\{-1, 1\}$, then

$$\mathrm{E}\left(\sum_{k=1}^{n} \pm k\right) = 0, \quad \mathrm{Var}\left(\sum_{k=1}^{n} \pm k\right) = \frac{n(n+1)(2n+1)}{6} \sim \frac{n^3}{3}$$

as $n \to \infty$. By the Central Limit Theorem,

$$\mathrm{P}\left(\sqrt{3}n^{-3/2} \sum_{k=1}^{n} \pm k \le x\right) \sim \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$$

which implies that [3, 4]

$$\mathrm{P}\left(\sum_{k=1}^{n} \pm k = 0\right) \sim s\sqrt{\frac{3}{2\pi}} n^{-3/2} \exp\left(-\frac{x^2}{2}\right)\Bigg|_{x=0}$$

where $s = 1 - (-1) = 2$ is the span of the distribution of $\pm$; hence [5, 6]

$$a(n) \sim \sqrt{\frac{6}{\pi}} n^{-3/2} 2^n.$$

In the same way,

$$b(n) \sim \frac{1}{2\sqrt{\pi}} n^{-3/2} 3^{n+1}.$$

Let $c(n)$ denote the number of sign choices such that

$$\pm 1 \pm 2 \pm 3 \pm \cdots \pm n = \pm 1 \pm 2 \pm 3 \pm \cdots \pm n.$$

Here [7]

$$c(n) \text{ is the coefficient of } x^{n(n+1)/2} \text{ in the polynomial } \prod_{k=1}^{n} \left(1 + x^k\right)^2$$

and [8–11]

$$c(n) \sim \sqrt{\frac{3}{\pi}} n^{-3/2} 2^{2n}.$$

Define [12]

$$\alpha(n) \text{ to be the maximal coefficient in the polynomial } \prod_{k=1}^{n} \left(1 + x^{2k}\right),$$
$$\beta(n) \text{ to be the maximal coefficient in the polynomial } \prod_{k=1}^{n} \left(1 + x^k + x^{2k}\right),$$
$$\gamma(n) \text{ to be the maximal coefficient in the polynomial } \prod_{k=1}^{n} \left(1 + x^k\right)^2.$$

The first of these has an immediate combinatorial interpretation: $\alpha(n)$ is the number of sign choices such that

$$\pm 1 \pm 2 \pm 3 \pm \cdots \pm n \text{ is } 0 \text{ or } 1.$$

While $\beta(n)$ seems not to have such a representation, the last sequence satisfies trivially $\gamma(n) = c(n)$ always.

We look at several more examples. Define [13]

$$\lambda_{\max}(n) \text{ to be the maximal coefficient in } \prod_{k=1}^{n} \left(1 - x^{2k}\right)$$
$$\text{and } -\lambda_{\min}(n) \text{ to be the corresponding minimal coefficient;}$$

$$\mu_{\max}(n) \text{ to be the maximal coefficient in } (-1)^n \prod_{k=1}^{n} \left(1 - x^k\right)^2$$
$$\text{and } -\mu_{\min}(n) \text{ to be the corresponding minimal coefficient.}$$

Only the third one possesses a clear simplification:

$$\mu_{\max}(n) \text{ is the coefficient of } x^{n(n+1)/2} \text{ in } (-1)^n \prod_{k=1}^{n} \left(1 - x^k\right)^2$$

and the asymptotics

$$\mu_{\max}(n)^{1/n} \sim 1.48... \sim 2\, e^{-0.29...}$$

are of interest [14, 15]. Greater understanding of the other sequences is desired.

1.30.1 *Number Partitioning*

What is the number of ways to partition the set $\{1, 2, \ldots, n\}$ into two subsets whose sums are as nearly equal as possible? If $n \equiv 0, 3 \bmod 4$, the answer is $\alpha(n)$; if $n \equiv 1, 2 \bmod 4$, the answer is $\alpha(n)/2$. In the former case, the subsets have the same sum; in the latter, the subsets have sums that differ by 1 [16, 17]. Partitioning arbitrary sets of n integers, each typically of order 2^m, is an NP-complete problem. The ratio m/n characterises the difficulty in searching for a perfect partition (one in which subset sums differ by at most 1). A phase transition exists for this problem (at $m/n = 1$, in fact) and perhaps similarly for all NP problems [17–19].

As an aside, we observe that

$$\lambda_{\max}(n) \text{ is the coefficient of } x^{n(n+1)/2} \text{ in the polynomial } \prod_{k=1}^{n} \left(1 - x^{2k}\right)$$

for $n \equiv 0 \bmod 4$, but this fails elsewhere (a conjectural relation involving $x^{(n+1)^2/2}$ coefficients for $n \equiv 3 \bmod 4$ falls apart when $n = 27$). It seems to be true that

$$\lambda_{\max}(n)^{1/n} \sim 1.21... \sim 2\, e^{-0.50...}$$

as $n \to \infty$ via multiples of 4.

As another aside, if $d(n)$ is the number of solutions of

$$\varepsilon_1 \cdot 1 + \varepsilon_2 \cdot 2 + \varepsilon_3 \cdot 3 + \cdots + \varepsilon_n \cdot n = \varepsilon_{-1} \cdot 1 + \varepsilon_{-2} \cdot 2 + \varepsilon_{-3} \cdot 3 + \cdots + \varepsilon_{-n} \cdot n,$$

then [20]

$$d(n) \text{ is the coefficient of } x^{n(n+1)} \text{ in the polynomial } \prod_{k=1}^{n} \left(1 + x^k + x^{2k}\right)^2$$

(in fact, it is the maximal such coefficient)

and

$$d(n) \sim \frac{1}{2\sqrt{2\pi}} n^{-3/2} 3^{2n+1}.$$

This grows more quickly than $b(n)$, of course. We wonder what else can be said in both cases. For example, what is the mean percentage of 0s in $\{\varepsilon_j\}$ taken over all solutions, as $n \to \infty$? It may well be $1/3$ for both, but it may be $> 1/3$ for one or the other.

Addendum Define a function $G: (0,1) \to \mathbb{R}$ by

$$G(x) = \int_0^1 \ln(\sin(\pi x t))\, dt.$$

There is a unique point $x_0 = 0.7912265710...$ at which G attains its maximum value $G(x_0) = -0.4945295653....$ Let

$$r = \exp(2G(x_0)) = 0.3719264606... = \frac{1}{4}(1.4877058426...),$$

$$C = \frac{4\sin(\pi x_0)}{x_0}\sqrt{\frac{\pi}{-G''(x_0)}} = 2.4057458393...$$

then [21]

$$\mu_{\max}(n) \sim C\frac{(4r)^n}{\sqrt{n}}$$

as $n \to \infty$, making impressively precise our earlier conjecture. An analogous formula for $\lambda_{\max}(n)$ for $n \equiv 0 \bmod 4$ remains open.

[1] H. S. Wilf, *generatingfunctionology*, 2nd ed., Academic Press, 1994, pp. 27–28, 201–202; MR1277813 (95a:05002).

[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007576 and A063865.

[3] B. Fristedt and L. Gray, *A Modern Approach to Probability Theory*, Birkhäuser, 1997, pp. 282–288; MR1422917 (98e:60002).

[4] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 2nd ed., Oxford Univ. Press, 1992, pp. 174–193; MR1199812 (93m:60002).

[5] D. Andrica and I. Tomescu, On an integer sequence related to a product of trigonometric functions, and its combinatorial relevance, *J. Integer Seq.* 5 (2002) 02.2.4; MR1938223 (2003j:05005).

[6] H.-K. Hwang, Review of "On an integer sequence...", MR1938223 (2003j:05005).

[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000980 and A047653.

[8] J. H. van Lint, Representation of 0 as $\sum_{k=-N}^{N} \varepsilon_k k$, *Proc. Amer. Math. Soc.* 18 (1967) 182–184; MR0205964 (34 #5789).

[9] R. C. Entringer, Representation of m as $\sum_{k=-n}^{n} \varepsilon_k k$, *Canad. Math. Bull.* 11 (1968) 289–293; MR0229564 (37 #5138).

[10] L. Clark, On the representation of m as $\sum_{k=-n}^{n} \varepsilon_k k$, *Internat. J. Math. Math. Sci.* 23 (2000) 77–80; MR1741328 (2001c:11011).

[11] G. Louchard and H. Prodinger, Representations of numbers as $\sum_{k=-n}^{n} \varepsilon_k k$: A saddle point approach, *Infinity in Logic and Computation*, Proc. 2007 Cape Town conf., ed. M. Archibald, V. Brattka, V. Goranko and B. Löwe, Lect. Notes in Comp. Sci. 5489, Springer-Verlag, 2009, pp. 87–96; MR2737664 (2012d:05052).

[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A025591, A039826, and A047653.
[13] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A086376, A086394, A133871, and A156082.
[14] S. Jaidee, S. Stevens and T. Ward, Mertens' theorem for toral automorphisms, *Proc. Amer. Math. Soc.* 139 (2011) 1819–1824; arXiv:0801.2082; MR2763768 (2011m:37029).
[15] R. B. Israel and D. W. Cantrell, Re: Bounds for an integral arising in dynamical Mertens, unpublished note (2008).
[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A069918.
[17] B. Hayes, The easiest hard problem, *Amer. Scientist*, v. 90 (2002) n. 2, 113–117.
[18] S. Mertens, A physicist's approach to number partitioning, *Theoret. Comput. Sci.* 265 (2001) 79–108; MR1848213 (2003j:68068).
[19] C. Borgs, J. Chayes and B. Pittel, Phase transition and finite-size scaling for the integer partitioning problem, *Random Structures Algorithms* 19 (2001) 247–288; MR1871556 (2002j:90061).
[20] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A156181.
[21] J. Gaither, G. Louchard, S. Wagner and M. D. Ward, Resolution of T. Ward's question and the Israel-Finch conjecture: Precise analysis of an integer sequence arising in dynamics. *Combin. Probab. Comput.* 24 (2015) 195–215; MR3318044.

1.31 Monoids of Natural Numbers

Let $\mathbb{N}$ denote the set of nonnegative integers. If $A = \{a_1, a_2, \ldots, a_m\}$ is a set of positive integers satisfying $\gcd(a_1, a_2, \ldots, a_m) = 1$, then

$$\langle a_1, a_2, \ldots, a_m \rangle = \left\{ \sum_{j=1}^{m} x_j a_j : x_j \in \mathbb{N} \text{ for each } 1 \le j \le m \right\}$$

is the subset of $\mathbb{N}$ **generated by** A. For example,

$$\langle a, a+1, a+2, a+3, \ldots, 2a-1 \rangle = \{0\} \cup \{a, a+1, a+2, a+3, \ldots\}$$

and

$$\langle 2, b \rangle = \{0, 2, 4, \ldots, b-3\} \cup \{b-1, b, b+1, b+2, b+3, \ldots\}$$

when $b \ge 3$ is odd.

A **numerical monoid** S is a subset of $\mathbb{N}$ that is closed under addition, contains 0, and has finite complement in $\mathbb{N}$. (Most authors use the phrase "numerical semigroup", but semigroups by definition need not contain 0, hence the usage is puzzling.) The **Frobenius number** f of S is the maximum element in the set $\mathbb{N} - S$, and the **genus** g of S is the cardinality of $\mathbb{N} - S$. Therefore

$$f(\langle a, a+1, a+2, a+3, \ldots, 2a-1 \rangle) = a-1, \quad f(\langle 2, b \rangle) = b-2,$$

Table 1.13 *Numerical monoids with small Frobenius number or genus*

$f=1$	$f=2$	$f=3$	$f=4$	$g=1$	$g=2$	$g=3$	$g=4$
$\langle 2,3\rangle$	$\langle 3,4,5\rangle$	$\langle 4,5,6,7\rangle$	$\langle 5,6,7,8,9\rangle$	$\langle 2,3\rangle$	$\langle 3,4,5\rangle$	$\langle 4,5,6,7\rangle$	$\langle 5,6,7,8,9\rangle$
		$\langle 2,5\rangle$	$\langle 3,5,7\rangle$		$\langle 2,5\rangle$	$\langle 3,5,7\rangle$	$\langle 4,6,7,9\rangle$
						$\langle 3,4\rangle$	$\langle 3,7,8\rangle$
						$\langle 2,7\rangle$	$\langle 4,5,7\rangle$
							$\langle 4,5,6\rangle$
							$\langle 3,5\rangle$
							$\langle 2,9\rangle$

$$g\left(\langle a, a+1, a+2, a+3, \ldots, 2a-1\rangle\right) = a-1, \quad g\left(\langle 2, b\rangle\right) = (b-1)/2$$

and, more generally [1],

$$f(\langle a, b\rangle) = (a-1)(b-1)-1, \quad g\left(\langle a, b\rangle\right) = (a-1)(b-1)/2$$

when $\gcd(a, b) = 1$. It is known that $f+1 \le 2g$ always [2, 3]. Table 1.13 gives all monoids S with $1 \le f \le 4$ or $1 \le g \le 4$.

Define sequences [4–7]

$$\{F_n\}_{n=1}^{\infty} = \{1, 1, 2, 2, 5, 4, 11, 10, \ldots\},$$

$$\{G_n\}_{n=1}^{\infty} = \{1, 2, 4, 7, 12, 23, 39, 67, \ldots\}$$

by

$$F_n = (\text{the number of monoids } S \text{ with } f(S) = n),$$

$$G_n = (\text{the number of monoids } S \text{ with } g(S) = n).$$

Backelin [8] showed that

$$0 < \liminf_{n\to\infty} 2^{-n/2} F_n < \limsup_{n\to\infty} 2^{-n/2} F_n < \infty,$$

$$\tfrac{1}{2}(2.47) < \lim_{\substack{n\to\infty \\ n\equiv 0 \bmod 2}} 2^{-n/2} F_n < \tfrac{1}{2}(3.3), \quad \tfrac{1}{\sqrt{2}}(2.5) < \lim_{\substack{n\to\infty \\ n\equiv 1 \bmod 2}} 2^{-n/2} F_n < \tfrac{1}{\sqrt{2}}(3.32)$$

and the work of others [5, 9–11] culminated with a theorem by Zhao [12] and Zhai [13]:

$$\lim_{n\to\infty} \frac{G(n)}{\varphi^n} \quad \text{exists, is finite, and is at least } 3.78$$

where $\varphi = (1+\sqrt{5})/2 = 1.6180339887...$ is the Golden mean. See also [14, 15]. Tighter bounds are needed for F_n asymptotics; it has not even been proved that G_n is increasing.

What can be said about rates of growth of $F_{n,k}$ and $G_{n,k}$, the counts of monoids when the number of generators is fixed to be k?

A monoid is **irreducible** if it cannot be written as the intersection of two monoids properly containing it [16]. A monoid S is irreducible if and only if S is maximal (with respect to set inclusion) in the collection of all monoids with Frobenius number $f(S)$. Irreducible monoids with odd f are the same as **symmetric** monoids (for which $f=2g-1$ always); irreducible monoids with even f are the same as **pseudo-symmetric** monoids (for which $f=2(g-1)$ always). As an example, $\langle 3,4\rangle$ and $\langle 2,7\rangle$ are the two symmetric monoids with Frobenius number 5; $\langle 4,5,7\rangle$ is the unique pseudo-symmetric monoid with Frobenius number 6. Another characterization of symmetry and pseudo-symmetry will be given shortly. Define [4, 17]

$$\{H_n\}_{n=1}^{\infty}=\{1,1,1,1,2,1,3,2,3,3,6,2,8,\ldots\}$$

by

$$H_n=(\text{the number of irreducible monoids } S \text{ with } f(S)=n)\,;$$

it follows that [8]

$$0<\liminf_{n\to\infty}2^{-n/6}H_n<\limsup_{n\to\infty}2^{-n/6}H_n<\infty,$$

$$\tfrac{1}{2}(9.36)<\lim_{\substack{n\to\infty\\ n\equiv 0 \bmod 6}}2^{-n/6}H_n=\tfrac{1}{\sqrt{2}}\lim_{\substack{n\to\infty\\ n\equiv 3 \bmod 6}}2^{-n/6}H_n<c.$$

No finite value c (as an upper bound for H_n asymptotics) has been rigorously proved.

1.31.1 *Sets without Closure*

A **numerical set** S is a subset of $\mathbb{N}$ that contains 0 and has finite complement in $\mathbb{N}$. The **Frobenius number** of S is, as before, the maximum element in the set $\mathbb{N}-S$. Nothing has been assumed about additivity so far. Every numerical set S has an associated **atom monoid** $A(S)$ defined by

$$A(S)=\{n\in\mathbb{Z}:n+S\subseteq S\}\,.$$

Clearly $A(S)\subseteq S$; also $A(S)=S$ if and only if S is itself a numerical monoid. The Frobenius number of $A(S)$ is the same as the Frobenius number of S; thus there is no possible ambiguity when speaking about $f(S)$. Let

$$\mathbb{N}_n=\langle n+1,n+2,n+3,\ldots,2n+1\rangle=\{0\}\cup\{n+1,n+2,n+3,\ldots\},$$

which we already know has Frobenius number n. Given n, which sets S have $A(S)=\mathbb{N}_n$? Table 1.14 answers the question for $1\le n\le 5$. For brevity, we give only T, where $S=T\cup\mathbb{N}_n$ is a disjoint union.

Define [18]

$$\{P_n\}_{n=1}^{\infty}=\{1,2,3,6,10,20,37,74,\ldots\}$$

Table 1.14 *Numerical sets* $T \cup \mathbb{N}_n$ *with atom monoid* $\mathbb{N}_n$

$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$\emptyset^*$	$\emptyset$	$\emptyset$	$\emptyset$	$\emptyset$
	$\{1\}$	$\{1\}^*$	$\{1\}$	$\{1\}$
		$\{1,2\}$	$\{2\}$	$\{2\}$
			$\{1,2\}$	$\{1,2\}^*$
			$\{1,3\}$	$\{1,3\}^*$
			$\{1,2,3\}$	$\{1,4\}$
				$\{2,3\}$
				$\{1,2,3\}$
				$\{1,2,4\}$
				$\{1,2,3,4\}$

by

$$P_n = (\text{the number of sets } S \text{ with } A(S) = \mathbb{N}_n)\,;$$

Marzuola & Miller [19] showed that

$$\lim_{n\to\infty} \frac{P_n}{2^{n-1}} \approx 0.484451 \pm 0.005.$$

Also, a numerical set S with Frobenius number n satisfying

$$x \in S \text{ if and only if } n - x \notin S$$

is **symmetric** if n is odd and **pseudo-symmetric** if n is even and $n/2 \notin S$ (we agree to exclude $x = n/2$ from consideration). The symmetric cases in Table 1.14 are marked by $*$. Define [18]

$$\{Q_k\}_{k=1}^{\infty} = \{1, 1, 2, 3, 6, 10, 20, 37, 73, \ldots\}$$

by

$$Q_k = (\text{the number of symmetric sets } S \text{ with } A(S) = \mathbb{N}_{2k-1})$$

then [19]

$$\lim_{k\to\infty} \frac{Q_k}{2^{k-1}} \approx 0.230653 \pm 0.006.$$

It is interesting that Q_{k+2} is the number of additive 2-bases for $\{0, 1, 2, \ldots, k\}$, meaning sets Σ that satisfy

$$\Sigma \subseteq \{0, 1, 2, \ldots, k\} \subseteq \Sigma + \Sigma.$$

The asymptotics for the corresponding "anti-atom" problem for pseudo-symmetric sets are identical to the preceding.

1.31.2 *Frobenius Numbers with Three Arguments*

Given $\gcd(a,b,c)=1$, let $\tilde{f}(a,b,c)=f(\langle a,b,c\rangle)+a+b+c$ denote the **modified Frobenius number.** Ustinov [20–22] proved that, on average, $\tilde{f}(a,b,c)$ is asymptotic to $(8/\pi)\sqrt{abc}$. The following probability density function

$$p(t)=\begin{cases} \dfrac{12}{\pi}\left(\dfrac{t}{\sqrt{3}}-\sqrt{4-t^2}\right) & \text{for } \sqrt{3}\le t\le 2 \\ \dfrac{12}{\pi^2}\left[\sqrt{3}\,t\arccos\left(\dfrac{t+3\sqrt{t^2-4}}{4\sqrt{t^2-3}}\right)+\dfrac{3}{2}\sqrt{t^2-4}\ln\left(\dfrac{t^2-4}{t^2-3}\right)\right] & \text{for } t>2 \end{cases}$$

describes more fully the behavior of $\tilde{f}(a,b,c)/\sqrt{abc}$ as $\max\{a,b,c\}\to\infty$; in particular, the distribution has a sharp peak at mode 2 and has mean

$$\int\limits_{\sqrt{3}}^{\infty} t\,p(t)dt=\frac{8}{\pi}.$$

In words, $\tilde{f}(a,b,c)$ is the largest positive integer not representable as $xa+yb+zc$, $x>0$, $y>0$, $z>0$. This is more convenient for the analysis because it is multiplicative in two arguments: if $d\ge 1$ is a divisor of both b and c, then

$$\tilde{f}(a,b,c)=d\tilde{f}\left(a,\frac{b}{d},\frac{c}{d}\right).$$

The proof is based on continued fraction theory; for example, Porter's constant [23] appears in [20]. Properties of the original $f(\langle a,b,c\rangle)$ appear in [24], along with discussion of the coin exchange or money changing problem [25, 26].

The **modified genus**

$$\tilde{g}(a,b,c)=g(\langle a,b,c\rangle)+\frac{a+b+c-1}{2}$$

does not possess as simple an interpretation as $\tilde{f}(a,b,c)$ (recall that $g(\langle a,b,c\rangle)$ denotes the cardinality of all positive integers not representable as $xa+yb+zc$, $x\ge 0$, $y\ge 0$, $z\ge 0$). Again, multiplicativity and continued fractions play a role. Vorobev [27] proved that, on average, $\tilde{g}(a,b,c)/\sqrt{abc}$ approaches $64/(5\pi^2)$ as $\max\{a,b,c\}\to\infty$. A corresponding density function $q(t)$ remains open, although its support is known to be the interval $[5\sqrt{3}/9,\infty)$ and its mode is 1 on empirical grounds [28].

1.31.3 *Missing Sums and Differences*

A **more sums than differences** (MSTD) set is a finite subset S of $\mathbb{N}$ satisfying $|S+S|>|S-S|$. The probability that a uniform random subset of

$\{0, 1, ..., n-1\}$ is an MSTD set is provably > 0.000428 and conjecturally ≈ 0.00045, as $n \to \infty$. Underlying solution techniques [29, 30] resemble those in [12]; the problem itself reminds us of [31].

The probability mass function of $M_n = 2n - 1 - |S+S|$ for arbitrary $S \subseteq \{0, 1, ..., n-1\}$ appears in [32] as well as moments

$$\lim_{n\to\infty} \mathrm{E}(M_n) = 10, \quad \lim_{n\to\infty} \mathrm{Var}(M_n) = 35.9658....$$

A closed-form expression for the variance is not known. What is the analog of this result when sums are replaced by differences?

[1] R. Fröberg, C. Gottlieb and R. Häggkvist, On numerical semigroups, *Semigroup Forum* 35 (1987) 63–83; MR0880351 (88d:20092).
[2] M. Bras-Amorós and A. de Mier, Representation of numerical semigroups by Dyck paths, *Semigroup Forum* 75 (2007) 677–682; arXiv:math/0612634; MR2353289 (2008g:20132).
[3] M. Bras-Amorós and S. Bulygin, Towards a better understanding of the semigroup tree, *Semigroup Forum* 79 (2009) 561–574; arXiv:0810.1619; MR2564064 (2010m:20096).
[4] J. C. Rosales, P. A. García-Sánchez, J. I. García-García and J. A. Jiménez Madrid, Fundamental gaps in numerical semigroups, *J. Pure Appl. Algebra* 189 (2004), 301–313; MR2038577 (2004j:20116).
[5] M. Bras-Amorós, Fibonacci-like behavior of the number of numerical semigroups of a given genus, *Semigroup Forum* 76 (2008) 379–384; MR2377597 (2009c:20110).
[6] N. Medeiros, Listing of Numerical Semigroups up to Genus 12, unpublished table (2002).
[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007323 and A124506.
[8] J. Backelin, On the number of semigroups of natural numbers, *Math. Scand.* 66 (1990) 197–215; MR1075137 (91k:11084).
[9] M. Bras-Amorós, Bounds on the number of numerical semigroups of a given genus, *J. Pure Appl. Algebra* 213 (2009) 997–1001; arXiv:0802.2175; MR2498791 (2010e:05008).
[10] V. Blanco, P. A. García-Sánchez and J. Puerto, Counting numerical semigroups with short generating functions, *Internat. J. Algebra Comput.* 21 (2011) 1217–1235; arXiv:0901.1228; MR2863434 (2012m:20101).
[11] S. Elizalde, Improved bounds on the number of numerical semigroups of a given genus, *J. Pure Appl. Algebra* 214 (2010) 1862–1873; arXiv:0905.0489; MR2608114 (2011i:20088).
[12] Y. Zhao, Constructing numerical semigroups of a given genus, *Semigroup Forum* 80 (2010), no. 2, 242–254; arXiv:0910.2075; MR2601762 (2011h:20125).
[13] A. Zhai, Fibonacci-like growth of numerical semigroups of a given genus, *Semigroup Forum* 86 (2013) 634–662; arXiv:1111.3142; MR3053785.
[14] J. Fromentin and F. Hivert, Exploring the tree of numerical semigroups, *Math. Comp.* 85 (2016) 2553–2568; arXiv:1305.3831; MR3511292.
[15] M. Bernardini and F. Torres, Counting numerical semigroups by genus and even gaps, *Discrete Math.* 340 (2017) 2853–2863; arXiv:1612.01212; MR3698073.
[16] J. C. Rosales and M. B. Branco, Irreducible numerical semigroups, *Pacific J. Math.* 209 (2003) 131–143; MR1973937 (2004b:20091).

[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A158206, A158278, and A158279.
[18] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A066062, A008929, and A158291.
[19] J. Marzuola and A. Miller, Counting numerical sets with no small atoms, *J. Combin. Theory Ser. A* 117 (2010) 650–667; arXiv:0805.3493; MR2645183 (2011e: 20086).
[20] A. V. Ustinov, Solution of the Arnold problem on weak asymptotics for Frobenius numbers with three arguments (in Russian), *Mat. Sb.*, v. 200 (2009) n. 4, 131–160; Engl. transl. in *Sb. Math.* 200 (2009) 597–627; MR2531884 (2010e:11091).
[21] A. V. Ustinov, On the distribution of Frobenius numbers with three arguments (in Russian), *Izv. Ross. Akad. Nauk Ser. Mat.*, v. 74 (2010) n. 5, 145–170; Engl. transl. in *Izv. Math.* 74 (2010) 1023–1049; MR2757903 (2011m:11197).
[22] D. A. Frolenkov, The mean value of Frobenius numbers with three arguments (in Russian), *Izv. Ross. Akad. Nauk Ser. Mat.*, v. 76 (2012) n. 4, 125–184; Engl. transl. in *Izv. Math.* 76 (2012) 760–819; arXiv:1103.5427; MR3013273.
[23] S. R. Finch, Porter-Hensley constants, *first volume*, pp. 156–160.
[24] A. Tripathi, Formulae for the Frobenius number in three variables, *J. Number Theory* 170 (2017) 368–389; MR3541713.
[25] M. Beck and S. Robins, The coin-exchange problem of Frobenius, *Computing the Continuous Discretely. Integer-Point Enumeration in Polyhedra*, 2nd ed., Springer-Verlag, 2015, pp. 3–26; MR3410115.
[26] R. K. Guy, The money-changing problem, *Unsolved Problems in Number Theory*, 2nd ed., Springer-Verlag, 1994, sect. C7; MR1299330 (96e:11002).
[27] I. S. Vorobev, On the Frobenius problem with three arguments (in Russian), *Mat. Sb.*, v. 207 (2016) n. 6, 53–78; Engl. transl. in *Sb. Math.* 207 (2016) 816–840; MR3507503.
[28] I. S. Vorobev, Experimental investigation of the Frobenius problem for three arguments (in Russian), *Dalnevost. Mat. Zh.* 11 (2011) 3–9; MR2851920 (2012j:11069).
[29] G. Martin and K. O'Bryant, Many sets have more sums than differences, *Additive Combinatorics*, Proc. 2006 Montréal conf., ed. A. Granville, M. B. Nathanson and J. Solymosi, Amer. Math. Soc., 2007, pp. 287–305; arXiv:math/0608131; MR2359479 (2008i:11038).
[30] Y. Zhao, Sets characterized by missing sums and differences, *J. Number Theory* 131 (2011) 2107–2134; arXiv:0911.2292; MR2825117 (2012g:11023).
[31] S. R. Finch, Cameron's sum-free set constants, *first volume*, pp. 180–183.
[32] O. Lazarev, S. J. Miller and K. O'Bryant, Distribution of missing sums in sumsets, *Experim. Math.* 22 (2013) 132–156; arXiv:1109.4700; MR3047909.

1.32 Primitive Cusp Form

Let $S_k(N)$ denote the vector space of weight k cusp forms on $\Gamma_0(N)$ with trivial character; see [1] for background. There are two circumstances under which $f \in S_k(N)$ might fail to be **primitive** [2]:

- $f \in S_k(N/d)$ for some divisor $d > 1$ of N
- $f(z) = g(dz)$ and $g \in S_k(N/d)$ for some divisor $d > 1$ of N.

For example, let f_{11A} denote the (unique) level 11 weight 2 cusp form, then both $f_{11A}(z)$ and $f_{11A}(2z)$ are level 22 cusp forms. Similarly, both $f_{14A}(z)$ and $f_{14A}(2z)$

are level 28 cusp forms, and both $f_{15A}(z)$ and $f_{15A}(2z)$ are level 30 cusp forms. None of these are "new" at $N=22$, 28 or 30 since they arise from lower levels.

Define $S_k^{\#}(N)$ to be the vector space of weight k primitive cusp forms (or **Hecke newforms**) on $\Gamma_0(N)$ with trivial character. We restrict attention to the case $k=2$ henceforth. The dimension $\delta_0^{\#}(N)$ of $S_2^{\#}(N)$ over $\mathbb{C}$ possesses the following formula [3–5]:

$$\delta_0^{\#}(N)=\mu(N)+\frac{\lambda(N)}{12}-\frac{\omega_2(N)}{4}-\frac{\omega_3(N)}{3}-\frac{\kappa(N)}{2}$$

where λ, κ, ω_2, ω_3 are multiplicative functions with

$$\lambda\left(p^e\right)=\begin{cases} p-1 & \text{if } e=1,\\ p^2-p-1 & \text{if } e=2,\\ p^{e-3}(p+1)(p-1)^2 & \text{if } e\geq 3,\end{cases}$$

$$\kappa\left(p^e\right)=\begin{cases} 0 & \text{if } e\equiv 1 \bmod 2,\\ p-2 & \text{if } e=2,\\ p^{e/2-2}(p-1)^2 & \text{if } 4\leq e\equiv 0 \bmod 2,\end{cases}$$

$$\omega_2\left(p^e\right)=\begin{cases} -1 & \text{if } p=2 \text{ and } e\leq 2,\\ 1 & \text{if } p=2 \text{ and } e=3,\\ 0 & \text{if } p=2 \text{ and } e\geq 4,\\ \left(\frac{-4}{p}\right)-1 & \text{if } p\neq 2 \text{ and } e=1,\\ -\left(\frac{-4}{p}\right) & \text{if } p\neq 2 \text{ and } e=2,\\ 0 & \text{if } p\neq 2 \text{ and } e\geq 3,\end{cases}$$

$$\omega_3\left(p^e\right)=\begin{cases} -1 & \text{if } p=3 \text{ and } e\leq 2,\\ 1 & \text{if } p=3 \text{ and } e=3,\\ 0 & \text{if } p=3 \text{ and } e\geq 4,\\ \left(\frac{-3}{p}\right)-1 & \text{if } p\neq 3 \text{ and } e=1,\\ -\left(\frac{-3}{p}\right) & \text{if } p\neq 3 \text{ and } e=2,\\ 0 & \text{if } p\neq 3 \text{ and } e\geq 3,\end{cases}$$

$\mu(N)$ is the Möbius mu function [6], and $(-4/p)$, $(-3/p)$ are Kronecker–Jacobi–Legendre symbols [7]. We have asymptotic extreme results [4]

$$\tfrac{1}{12}(0.3739558136...)=\tfrac{1}{12}\prod_p\left(1-\tfrac{1}{p(p-1)}\right)=\liminf_{N\to\infty}\frac{\delta_0^{\#}(N)}{\varphi(N)}<\limsup_{N\to\infty}\frac{\delta_0^{\#}(N)}{\varphi(N)}=\tfrac{1}{12}$$

and average behavior

$$\sum_{N\leq x}\delta_0^{\#}(N)=\frac{45}{2\pi^6}x^2+o\left(x^2\right)$$

as $x \to \infty$, where $\varphi(N)$ is the Euler totient function [8] and the infinite product is Artin's constant [9].

For concreteness' sake, here is a list of basis elements of $S_2^{\#}(N)$ for $1 \le N \le 32$ [10–13]:

$$f_{11A}(z) = \eta(z)^2\eta(11z)^2,$$

$$f_{14A}(z) = \eta(z)\eta(2z)\eta(7z)\eta(14z),$$

$$f_{15A}(z) = \eta(z)\eta(3z)\eta(5z)\eta(15z),$$

$$f_{17A}(z) = \frac{\eta(z)\eta(4z)^2\eta(34z)^5}{\eta(2z)\eta(17z)\eta(68z)^2} - \frac{\eta(2z)^5\eta(17z)\eta(68z)^2}{\eta(z)\eta(4z)^2\eta(34z)},$$

$$f_{19A}(z) = \left(\frac{\eta(8z)^2\eta(76z)^5}{\eta(4z)\eta(38z)^2\eta(152z)^2} - \frac{\eta(2z)^2\eta(38z)^2}{\eta(z)\eta(19z)} + \frac{\eta(4z)^5\eta(152z)^2}{\eta(2z)^2\eta(8z)^2\eta(76z)}\right)^2,$$

$$f_{20A}(z) = \eta(2z)^2\eta(10z)^2,$$

$$\begin{aligned} f_{21A}(z) \;=\; & \frac{\eta(7z)\left[3\eta(z)^2\eta(7z)^2\eta(9z)^3 - \eta(3z)^5\eta(7z)\eta(21z) + 7\eta(z)\eta(3z)^2\eta(21z)^4\right]}{2\eta(z)^2\eta(3z)\eta(21z)} \\ & + \frac{3\eta(7z)\eta(63z)\left[\eta(z)^2\eta(7z)\eta(9z)^3 - \eta(3z)^5\eta(21z)\right]}{2\eta(z)\eta(3z)\eta(9z)\eta(21z)} \\ & + \frac{3\eta(z)^2\eta(7z)\eta(9z)\eta(63z)^2}{2\eta(3z)\eta(21z)}, \end{aligned}$$

$$f_{23A}(z) = q - \tfrac{1-\sqrt{5}}{2}q^2 - \sqrt{5}q^3 - \tfrac{1+\sqrt{5}}{2}q^4 - (1-\sqrt{5})q^5 - \tfrac{5-\sqrt{5}}{2}q^6 + \cdots,$$

$$f_{23B}(z) = q - \tfrac{1+\sqrt{5}}{2}q^2 + \sqrt{5}q^3 - \tfrac{1-\sqrt{5}}{2}q^4 - (1+\sqrt{5})q^5 - \tfrac{5+\sqrt{5}}{2}q^6 + \cdots,$$

$$f_{24A}(z) = \eta(2z)\eta(4z)\eta(6z)\eta(12z),$$

$$f_{26A}(z) = q - q^2 + q^3 + q^4 - 3q^5 - q^6 - q^7 - q^8 - 2q^9 + 3q^{10} + 6q^{11} + q^{12} + \cdots,$$

$$f_{26B}(z) = q + q^2 - 3q^3 + q^4 - q^5 - 3q^6 + q^7 + q^8 + 6q^9 - q^{10} - 2q^{11} - 3q^{12} + \cdots,$$

$$f_{27A}(z) = \eta(3z)^2\eta(9z)^2,$$

$$f_{29A}(z) = q - (1-\sqrt{2})q^2 + (1-\sqrt{2})q^3 + (1-2\sqrt{2})q^4 - q^5 - (3-2\sqrt{2})q^6 + \cdots,$$

$$f_{29B}(z) = q - (1+\sqrt{2})q^2 + (1+\sqrt{2})q^3 + (1+2\sqrt{2})q^4 - q^5 - (3+2\sqrt{2})q^6 + \cdots,$$

$$f_{30A}(z) = \eta(3z)\eta(5z)\eta(6z)\eta(10z) - \eta(z)\eta(2z)\eta(15z)\eta(30z),$$

$$f_{31A}(z) = q + \tfrac{1-\sqrt{5}}{2}q^2 - (1-\sqrt{5})q^3 - \tfrac{1+\sqrt{5}}{2}q^4 + q^5 - (3-\sqrt{5})q^6 + \cdots,$$

$$f_{31B}(z) = q + \tfrac{1+\sqrt{5}}{2}q^2 - (1+\sqrt{5})q^3 - \tfrac{1-\sqrt{5}}{2}q^4 + q^5 - (3+\sqrt{5})q^6 + \cdots,$$

$$f_{32A}(z) = \eta(4z)^2\eta(8z)^2$$

where $\eta(z) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n)$ is the Dedekind eta function and $q = e^{2\pi i z}$ [14]. It is natural to ask whether basis elements possessing integer coefficients necessarily have an eta expression. Counterexamples might include $f_{26A}(z)$ and $f_{26B}(z)$. Another counterexample might be $f_{49A}(z)$, which evidently can be represented via Ramanujan's two-variable theta function [15]. We know that [16]

$$\frac{1}{\sqrt{5}} f_{23A}(z) - \frac{1}{\sqrt{5}} f_{23B}(z) = \eta(z)^2 \eta(23z)^2$$

but no analogous simple expressions exist for $N = 29$ or $N = 31$ ($N = 26$ remains open).

What can be said about the relative number of newforms to cusp forms in $\Gamma_0(N)$? Martin [4] proved that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{N\le n} \frac{\delta_0^{\#}(N)}{\delta_0(N)} = \prod_p \left(1 + \frac{1}{p}\right)^{-1} \left(1 - \frac{1}{p}\right) \left(1 + \frac{2}{p} - \frac{1}{p^4} - \frac{1}{p^5}\right) = 0.444301....$$

A parallel theory can be developed for weight 2 primitive cusp forms on $\Gamma_1(N)$ with trivial character [5]. The answer to the same question over $\Gamma_1(N)$ is [4]

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \sum_{N\le n} \frac{\delta_1^{\#}(N)}{\delta_1(N)} &= \prod_p \left(1 + \frac{1}{p}\right)^{-1} \left(1 + \frac{1}{p} - \frac{2}{p^3} - \frac{2}{p^4} - \frac{2}{p^5} + \frac{1}{p^6} + \frac{1}{p^7} + \frac{1}{p^8}\right) \\ &= 0.652036.... \end{aligned}$$

Given a weight k primitive cusp form $f(z) = \sum_{m=1}^{\infty} a_m q^m$ on $\Gamma_0(N)$, define

$$L_f(z) = \sum_{m=1}^{\infty} a_m m^{-z}, \quad \mathrm{Re}(z) > (k+1)/2.$$

This admits analytic continuation to all of $\mathbb{C}$. What can be said about L-series moments over all such f at $z = 1/2$? Conrey [17] proved that, for $k = 2$,

$$\frac{1}{\delta_0^{\#}(N)} \sum_{f\in S_2^{\#}(N)} L_f(1/2) \sim \zeta(2),$$

$$\frac{1}{\delta_0^{\#}(N)} \sum_{f\in S_2^{\#}(N)} L_f^2(1/2) \sim 2\zeta(2)^2 \prod_p \left(1 + \frac{1}{p^2}\right) \cdot \frac{\ln\left(\sqrt{N}\right)}{1!},$$

$$\frac{1}{\delta_0^{\#}(N)} \sum_{f\in S_2^{\#}(N)} L_f^3(1/2) \sim 8\zeta(2)^3 \prod_p \left(1 - \frac{1}{p}\right) \left(1 + \frac{1}{p} + \frac{4}{p^2} + \frac{1}{p^3} + \frac{1}{p^4}\right) \cdot \frac{\ln\left(\sqrt{N}\right)^3}{3!},$$

$$\begin{aligned} \frac{1}{\delta_0^{\#}(N)} \sum_{f\in S_2^{\#}(N)} L_f^4(1/2) \sim{}& 128\zeta(2)^4 \prod_p \left(1 - \frac{1}{p}\right)^3 \left(1 + \frac{3}{p} + \frac{11}{p^2} + \frac{10}{p^3} + \frac{11}{p^4} + \frac{3}{p^5} + \frac{1}{p^6}\right) \\ &\cdot \frac{\ln\left(\sqrt{N}\right)^6}{6!} \end{aligned}$$

as $N \to \infty$ passes through the prime numbers. The expression in p can be verified for each exponent $1 \le \ell \le 4$ by use of a double summation [18]

$$\left(1-\frac{1}{p}\right)^{\ell(\ell+1)/2} \sum_{r=0}^{\infty} \frac{1}{p^{2r}} \sum_{j=r}^{\infty} \binom{\ell+2j-1}{2j} \frac{1}{p^{j-r}(1+1/p)^{2j}} \left(\binom{2j}{j-r} - \binom{2j}{j-r-1}\right).$$

We wonder about the $\Gamma_1(N)$-analog of the four moments, as well as any connection between such results and others given in [19].

1.32.1 *Half-Integer Weights*

Let $k \ge 1$ be an odd integer and $N \ge 4$ be a multiple of 4. A **modular form of weight** $k/2$ **and level** N is an analytic function f defined on the complex upper half plane that transforms under the action of $\Gamma_0(N)$ according to [2, 20, 21]

$$f\left(\frac{az+b}{cz+d}\right) = \left(\frac{c}{d}\right)^k \varepsilon_d^{-k}(cz+d)^{k/2} f(z) \quad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$$

and whose Fourier series $f(z) = \sum_{n=-\infty}^{\infty} \gamma_n e^{2\pi i n z}$ satisfies $\gamma_n = 0$ for all $n < 0$. For the preceding relation, define

$$\varepsilon_d = \begin{cases} 1 & \text{if } d \equiv 1 \bmod 4, \\ i & \text{if } d \equiv 3 \bmod 4. \end{cases}$$

Note that d must be odd since otherwise $ad - bc$ would be divisible by 2, contradicting $ad - bc = 1$. For negative odd d or zero c, let

$$\left(\frac{c}{d}\right) = \begin{cases} \left(\frac{c}{|d|}\right) & \text{if } d < 0 \text{ and } c > 0, \\ -\left(\frac{c}{|d|}\right) & \text{if } d < 0 \text{ and } c < 0, \\ 1 & \text{if } d = \pm 1 \text{ and } c = 0. \end{cases}$$

If, additionally, we have $\gamma_0 = 0$, then f is a **cusp form** of weight $k/2$ and level N. The space $M_{k/2}(N)$ of modular forms and the space $S_{k/2}(N)$ of cusp forms satisfy

$$\dim(M_{k/2}(4)) = \left\lfloor \frac{k}{4} \right\rfloor + 1$$

and $\dim(S_{k/2}(4)) = \dim(M_{k/2}(4)) - 2$ if $k \ge 9$. Straightforward formulas for $\dim(S_{1/2}(N))$ and $\dim(S_{3/2}(N))$ have not yet been found, but we know that [22–25]

$$\dim(S_{5/2}(N)) = \frac{1}{8}\psi(N) - \frac{1}{2\,\alpha(N)}\beta(N)\chi(N)$$

where

$$\psi(N) = N \prod_{p|N} \left(1 + \frac{1}{p}\right), \quad \chi(N) = \sum_{d|N} \varphi\left(\gcd\left(d, \frac{N}{d}\right)\right),$$

$$\alpha(N) = \begin{cases} 3 \cdot 2^{r_2(N)/2-1} & \text{if } r_2(N) \text{ is even,} \\ 2^{(r_2(N)+1)/2} & \text{if } r_2(N) \text{ is odd,} \end{cases}$$

$r_p(N)$ is the largest exponent e such that p^e divides N for prime p, and

$$\beta(N) = \begin{cases} \alpha(N) & \text{if } r_2(N) \geq 4, \\ 3 & \text{if } r_2(N) = 3, \\ 2 & \begin{array}{l}\text{if } r_2(N) = 2 \text{ and there exists } p \equiv 3 \bmod 4 \\ \text{such that } p|N \text{ and } r_p(N) \text{ is odd,}\end{array} \\ 3/2 & \text{otherwise.} \end{cases}$$

There are slightly different formulas for $\dim(S_{k/2}(N))$ for larger k as well. The proof, due to Cohen & Oesterlé [22], has never been published.

In the following, we will need one of the two basis elements of $M_2(4)$:

$$F(z) = \sum_{n=0}^{\infty} \sigma(2n+1) q^{2n+1}$$

where $\sigma(m)$ is the sum of all divisors of m. It can be shown that [2, 23]

$$F(z) = \frac{\eta(4z)^8}{\eta(2z)^4}.$$

The simplest half-integer weight modular form has weight $1/2$ and level 4:

$$\theta(z) = \sum_{n=-\infty}^{\infty} q^{n^2} = \frac{\eta(2z)^5}{\eta(z)^2 \eta(4z)^2}.$$

(It turns out that $\theta(z)^4$ is the other basis element of $M_2(4)$.) Let us focus on cusp forms henceforth [26]. The first nonzero cusp form of weight $1/2$ occurs at level 1728:

$$\frac{1}{2} \sum_{n=-\infty}^{\infty} \left(\frac{12}{n}\right) q^{3n^2} = \eta(72z)$$

and the first nonzero cusp form of weight $3/2$ occurs at level 28:

$$\frac{\eta(z)\eta(4z)\eta(14z)^4}{\eta(2z)\eta(7z)\eta(28z)}.$$

The first nonzero cusp form of level 4 has weight $9/2$:

$$\theta(z) F(z) \left(\theta(z)^4 - 16 F(z)\right) = \frac{\eta(2z)^{12}}{\theta(z)^3};$$

the first nonzero cusp form of level 8 has weight $7/2$:

$$\frac{\eta(z)^2 \eta(4z)^6}{\eta(2z)};$$

the first nonzero cusp form of level 12 has weight $5/2$:

$$\frac{\eta(2z)^3\eta(6z)^3}{\theta(3z)}.$$

A prominent example is one of the two basis elements of $S_{13/2}(4)$:

$$\theta(z)F(z)\left(\theta(z)^4-16F(z)\right)\left(\theta(z)^4-2F(z)\right),$$

which is the image of $\Delta(z)\in S_{12}(1)$ under what is called the *Shimura correspondence* [2, 27]. Further discussion of this topic, with application to Tunnell's solution of the congruent number problem, is beyond our scope. We have not mentioned newforms of half-integer weight – in fact, two distinct definitions are commonly used, one due to Serre & Stark [28] and the other due to Kohnen [29] – but we must cease here.

1.32.2 *Complex Multiplication*

A cusp form $f(z)=\sum_{n=1}^{\infty}\gamma_n q^n\in S_k(N)$ has **complex multiplication (CM)** by a nontrivial Dirichlet character ξ if [30]

$$f(z)=\sum_{n=1}^{\infty}\xi(n)\gamma_n q^n;$$

equivalently, $\xi(p)=1$ or $\gamma_p=0$ for each prime p. It can be shown that ξ is necessarily a quadratic character, thus we often refer to CM by the corresponding quadratic field. There is a one-to-one correspondence between imaginary quadratic fields of class number one [31]:

$$\mathbb{Q}(\sqrt{-1}),\ \mathbb{Q}(\sqrt{-2}),\ \mathbb{Q}(\sqrt{-3}),\ \mathbb{Q}(\sqrt{-7}),\ \mathbb{Q}(\sqrt{-11}),$$
$$\mathbb{Q}(\sqrt{-19}),\ \mathbb{Q}(\sqrt{-43}),\ \mathbb{Q}(\sqrt{-67}),\ \mathbb{Q}(\sqrt{-163})$$

and CM-newforms of weight 2 (elliptic curves with CM) up to twisting [32]:

$$64A4,\ 256A1,\ 27A3,\ 49A1,\ 121B1,\ 361A1,\ 1849A1,\ 4489A1,\ 26569A1$$

with rational coefficients. Schütt [33] classified similarly CM-newforms of weight 3 and 4.

1.32.3 *Singular K3 Surfaces*

We merely mention a class of projective varieties, called **K3 surfaces**, that are two-dimensional analogs of elliptic curves [34]. The name K3 is given in honor of Kummer, Kahler & Kodaira and also refers to the mountain K2 [35]. Existence of rational points is one theme; canonical heights of such points can be computed [36, 37] as with elliptic curves.

A K3 surface over $\mathbb{Q}$ is *not* modular, in general [38]. If we restrict attention to what are called **singular** (or **extremal**) K3 surfaces, however, then modularity holds with associated newform of weight 3 and possibly nontrivial Nebentypus character [39–41]. Further, the newform is CM.

For example, the Fermat quartic surface in $\widetilde{\mathbb{C}^3}$:

$$Z_0^4 + Z_1^4 + Z_2^4 + Z_3^4 = 0$$

has corresponding unique CM-newform of weight 3 and level 16 [34]:

$$\eta(4z)^6$$

which has character $(-4/\cdot)$. There are unique CM-newforms of weight 3 and levels 7, 8, 11 and 15 [33, 42, 43]:

$$\eta(z)^3\eta(7z)^3,$$

$$\eta(z)^2\eta(2z)\eta(4z)\eta(8z)^2,$$

$$\left(G(z)^2 + 4G(2z)^2 + 8G(4z)^2\right) G(z)^2/G(2z),$$

$$\eta(3z)^3\eta(5z)^3 - \eta(z)^3\eta(15z)^3$$

where

$$G(z) = \eta(z)\eta(11z)$$

and we wonder if algebraic expressions for geometric realizations of these (for example, as intersections of varieties) can be found.

Acknowledgments I am grateful to Michael Somos for his computations of eta expressions and Hung Bui, Jeremy Rouse, Matthias Schütt, Kevin James, Gonzalo Tornaria & William Stein for answers to many questions.

[1] S. R. Finch, Modular forms on $\mathrm{SL}_2(\mathbb{Z})$, *this volume*, §1.15.
[2] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, 1984, pp. 145–146, 173–174, 176–184, 213, 216, 238; MR0766911 (86c:11040).
[3] E. Halberstadt and A. Kraus, Courbes de Fermat: résultats et problèmes, *J. Reine Angew. Math.* 548 (2002) 167–234; MR1915212 (2003h:11068).
[4] G. Martin, Dimensions of the spaces of cusp forms and newforms on $\Gamma_0(N)$ and $\Gamma_1(N)$, *J. Number Theory* 112 (2005) 298–331; math.NT/0306128; MR2141534 (2005m:11069).
[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A127788 and A159046.
[6] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.
[7] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[8] S. R. Finch, Euler totient constants, *first volume*, pp. 115–118.
[9] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.
[10] Y. Martin and K. Ono, Eta-quotients and elliptic curves, *Proc. Amer. Math. Soc.* 125 (1997) 3169–3176; MR1401749 (97m:11057).
[11] W. Stein, Table of q-expansions of normalized new eigenforms, The Modular Forms Database.
[12] M. Somos, Eta expressions for newforms $17A$, $19A$, $21A$ and $30A$, unpublished note (2011).

[13] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006571, A030187, A030184, A128263, A187096, A030205, A187846, A030188, A247198, A251913, A030206, A122777, A002171.

[14] S. R. Finch, Dedekind eta products, *this volume*, §1.19.

[15] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A140686.

[16] A. Arnold-Roksandich, K. James and R. Keaton, Counting eta-quotients of prime level, arXiv:1611.08969.

[17] J. B. Conrey, L-functions and random matrices, *Mathematics Unlimited - 2001 and Beyond*, ed. B. Engquist and W. Schmid, Springer-Verlag, 2001, pp. 331–352; math.NT/0005300; MR1852163 (2002g:11134).

[18] H. M. Bui and J. P. Keating, On the mean values of L-functions in orthogonal and symplectic families, *Proc. Lond. Math. Soc.* 96 (2008) 335–366; arXiv:math/0612348; MR2396123 (2009a:11172).

[19] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein and N. C. Snaith, Integral moments of L-functions, *Proc. London Math. Soc.* 91 (2005) 33–104; math.NT/0206018; MR2149530 (2006j:11120).

[20] G. Shimura, Modular forms of half integral weight, *Modular Functions of One Variable. I*, Proc. 1972 Antwerp conf., Lect. Notes in Math. 320, ed. W. Kuyk, Springer-Verlag, 1973, pp. 57–74; correction, *Modular Functions of One Variable. IV*, Lect. Notes in Math. 476, p. 145; MR0480346 (58 #519).

[21] G. Shimura, On modular forms of half integral weight, *Annals of Math.* 97 (1973) 440–481; MR0332663 (48 #10989).

[22] H. Cohen and J. Oesterlé, Dimensions des espaces de formes modulaires, *Modular Functions of One Variable. VI*, Proc. 1976 Bonn conf., ed. J.-P. Serre and D. B. Zagier, Lect. Notes in Math. 627, Springer-Verlag, 1977, pp. 69–78; MR0472703 (57 #12396).

[23] K. Ono, *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series*, Amer. Math. Soc., 2004, pp. 10–16, 19; MR2020489 (2005c:11053).

[24] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A159630–A159636.

[25] P. Humphries, A conjecture related to the Cohen-Oesterlé dimension formula of spaces of modular forms for half-integer weights, MathOverflow question 161948 (2014).

[26] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000122, A159813, A029769, A159814, A159584 and A054891.

[27] W. Kohnen and D. Zagier, Values of L-series of modular forms at the center of the critical strip, *Invent. Math.* 64 (1981) 175–198; MR0629468 (83b:10029).

[28] J.-P. Serre and H. M. Stark, Modular forms of weight $1/2$, *Modular Functions of One Variable. VI*, Proc. 1976 Bonn conf., ed. J.-P. Serre and D. B. Zagier, Lect. Notes in Math. 627, Springer-Verlag, 1977, pp. 27–67; MR0472707 (57 #12400).

[29] W. Kohnen, Newforms of half-integral weight, *J. Reine Angew. Math.* 333 (1982) 32–72; MR0660784 (84b:10038).

[30] K. A. Ribet, Galois representations attached to eigenforms with Nebentypus, *Modular Functions of One Variable. V*, Proc. 1976 Bonn conf., ed. J.-P. Serre and D. B. Zagier, Lect. Notes in Math. 601, Springer-Verlag, 1977, pp. 17–51; MR0453647 (56 #11907).

[31] S. R. Finch, Class number theory, *this volume*, §1.12.

[32] J. H. Silverman, *Advanced Topics in the Arithmetic of Elliptic Curves*, Springer-Verlag, 1994, pp. 101–102, 109–111, 140–143, 483; MR1312368 (96b:11074).

[33] M. Schütt, CM newforms with rational coefficients, *Ramanujan J.* 19 (2009) 187–205; arXiv:math/0511228; MR2511671 (2010c:11052).

[34] M. Schütt, Arithmetic of K3 surfaces, *Jahresbericht. Deutsch. Math.-Verein.* 111 (2009) 23–41; arXiv:0808.1061; MR2508556 (2010f:14042).
[35] W. Barth, C. Peters and A. Van de Ven, *Compact Complex Surfaces*, Springer-Verlag, 1984, pp. 238–288; MR0749574 (86c:32026).
[36] J. H. Silverman, Rational points on $K3$ surfaces: a new canonical height, *Invent. Math.* 105 (1991) 347–373; MR1115546 (92k:14025).
[37] G. S. Call and J. H. Silverman, Computing the canonical height on $K3$ surfaces, *Math. Comp.* 65 (1996) 259–290; MR1322885 (96g:11067).
[38] N. D. Elkies and M. Schütt, Modular forms and $K3$ surfaces, *Adv. Math.* 240 (2013) 106–131; arXiv:0809.0830; MR3046305.
[39] T. Shioda and H. Inose, On singular $K3$ surfaces, *Complex Analysis and Algebraic Geometry*, ed. W. L. Baily and T. Shioda, Cambridge Univ. Press, 1977, pp. 119–136; MR0441982 (56 #371).
[40] R. Livné, Motivic orthogonal two-dimensional representations of $\mathrm{Gal}(\overline{Q}/Q)$, *Israel J. Math.* 92 (1995) 149–156; MR1357749 (96m:11050).
[41] N. Yui, Arithmetic of Calabi-Yau varieties, *Mathematisches Institut, Georg-August-Universität Göttingen: Seminars Summer Term 2004*, ed. Y. Tschinkel, Universitätsdrucke Göttingen, 2004, pp. 9–29; MR2183123 (2007h:11050).
[42] M. Somos, Eta expression for CM-newforms of weight 3 and levels 11 & 15, unpublished note (2007).
[43] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002656, A030207, A129522, A136549 and A000729.

1.33 Cubic and Quartic Characters

In this essay, we revisit Dirichlet characters [1], but focus here on non-real cases (that is, of order exceeding 2).

Let $\mathbb{Z}_n^*$ denote the group (under multiplication modulo n) of integers relatively prime to n, and let $\mathbb{C}^*$ denote the group (under ordinary multiplication) of nonzero complex numbers. We wish to examine homomorphisms $\chi : \mathbb{Z}_n^* \to \mathbb{C}^*$ satisfying certain requirements. A Dirichlet character χ is **quadratic** if $\chi(k)^2 = 1$ for every k in $\mathbb{Z}_n^*$. It is well-known that, if $\chi \neq 1$ is a primitive quadratic character modulo n, then $D = \chi(-1)n$ is a fundamental discriminant and

$$\chi(k) = \left(\frac{D}{k}\right) \quad \text{for all } k \in \mathbb{Z}_n^*$$

where (D/k) is the Kronecker–Jacobi–Legendre symbol. A character χ is real if and only if it is quadratic. By the correspondence with $(D/.)$, quadratic characters can be said to be completely understood.

A Dirichlet character χ is **cubic** if $\chi(k)^3 = 1$ for every k in $\mathbb{Z}_n^*$. Let $\omega = (-1 + i\sqrt{3})/2$ where i is the imaginary unit. Let $a + b\omega$ be a prime in the ring $\mathbb{Z}[\omega]$ of Eisenstein–Jacobi integers with norm $a^2 - ab + b^2 \neq 3$. For any positive integer n in $\mathbb{Z}$, define the cubic residue symbol [2, 3]

$$\left(\frac{n}{a + b\omega}\right)_3$$

to be 0 if n is divisible by $a+b\omega$; otherwise it is the unique power ω^j for $0 \le j \le 2$ such that

$$n^{(a^2-ab+b^2-1)/3} \equiv \omega^j \bmod (a+b\omega).$$

The only prime divisor of 9 is $1-\omega$, which has norm 3. Hence we will need an alternative way of representing characters:

$$f_q(n,k) = \begin{cases} \omega^e & \text{if } n \equiv k^e \bmod q, \\ 0 & \text{otherwise,} \end{cases}$$

especially in the case $q=9$. The first several cubic characters are

$$f_7(n,5) = \left(\tfrac{n}{2+3\omega}\right)_3\Big|_{n=1,\ldots,7} = \{1,\omega,\omega^2,\omega^2,\omega,1,0\},$$

$$f_7(n,3) = \left(\tfrac{n}{-1-3\omega}\right)_3\Big|_{n=1,\ldots,7} = \{1,\omega^2,\omega,\omega,\omega^2,1,0\},$$

$$f_9(n,2)|_{n=1,\ldots,9} = \{1,\omega,0,\omega^2,\omega^2,0,\omega,1,0\},$$

$$f_9(n,5)|_{n=1,\ldots,9} = \{1,\omega^2,0,\omega,\omega,0,\omega^2,1,0\},$$

$$f_{13}(n,2) = \left(\tfrac{n}{-4-3\omega}\right)_3\Big|_{n=1,\ldots,13} = \{1,\omega,\omega,\omega^2,1,\omega^2,\omega^2,1,\omega^2,\omega,\omega,1,0\},$$

$$f_{13}(n,6) = \left(\tfrac{n}{-1+3\omega}\right)_3\Big|_{n=1,\ldots,13} = \{1,\omega^2,\omega^2,\omega,1,\omega,\omega,1,\omega,\omega^2,\omega^2,1,0\},$$

$$\begin{aligned} f_{19}(n,2) &= \left(\tfrac{n}{2-3\omega}\right)_3\Big|_{n=1,\ldots,19} = \{1,\omega,\omega,\omega^2,\omega,\omega^2,1,1,\omega^2,\omega^2,1,1,\omega^2,\omega, \\ &\qquad \omega^2,\omega,\omega,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{19}(n,10) &= \left(\tfrac{n}{5+3\omega}\right)_3\Big|_{n=1,\ldots,19} = \{1,\omega^2,\omega^2,\omega,\omega^2,\omega,1,1,\omega,\omega,1,1,\omega,\omega^2,\omega, \\ &\qquad \omega^2,\omega^2,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{31}(n,3) &= \left(\tfrac{n}{5+6\omega}\right)_3\Big|_{n=1,\ldots,31} \\ &= \{1,1,\omega,1,\omega^2,\omega,\omega,1,\omega^2,\omega^2,\omega^2,\omega,\omega^2,\omega,1,1,\omega,\omega^2,\omega,\omega^2, \\ &\quad \omega^2,\omega^2,1,\omega,\omega,\omega^2,1,\omega,1,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{31}(n,11) &= \left(\tfrac{n}{-1-6\omega}\right)_3\Big|_{n=1,\ldots,31} \\ &= \{1,1,\omega^2,1,\omega,\omega^2,\omega^2,1,\omega,\omega,\omega,\omega^2,\omega,\omega^2,1,1,\omega^2,\omega,\omega^2,\omega, \\ &\quad \omega,\omega,1,\omega^2,\omega^2,\omega,1,\omega^2,1,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{37}(n,2) &= \left(\tfrac{n}{-4+3\omega}\right)_3\Big|_{n=1,\ldots,37} \\ &= \{1,\omega,\omega^2,\omega^2,\omega^2,1,\omega^2,1,\omega,1,1,\omega,\omega^2,1,\omega,\omega,\omega,\omega^2,\omega^2,\omega, \\ &\quad \omega,\omega,1,\omega^2,\omega,1,1,\omega,1,\omega^2,1,\omega^2,\omega^2,\omega^2,\omega,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{37}(n,5) &= \left(\tfrac{n}{-7-3\omega}\right)_3\Big|_{n=1,\ldots,37} \\ &= \{1,\omega^2,\omega,\omega,\omega,1,\omega,1,\omega^2,1,1,\omega^2,\omega,1,\omega^2,\omega^2,\omega^2,\omega,\omega,\omega^2, \\ &\quad \omega^2,\omega^2,1,\omega,\omega^2,1,1,\omega^2,1,\omega,1,\omega,\omega,\omega,\omega^2,1,0\}. \end{aligned}$$

A Dirichlet character χ is **quartic (biquadratic)** if $\chi(k)^4 = 1$ for every k in $\mathbb{Z}_n^*$. Let $a+bi$ be a prime in the ring $\mathbb{Z}[i]$ of Gaussian integers with norm $a^2+b^2 \neq 2$. For any positive integer n in $\mathbb{Z}$, define the quartic (biquadratic) residue symbol [2, 3]

$$\left(\frac{n}{a+bi}\right)_4$$

to be 0 if n is divisible by $a+bi$; otherwise it is the unique power i^j for $0 \le j \le 3$ such that

$$n^{(a^2+b^2-1)/4} \equiv i^j \bmod (a+bi).$$

The only prime divisor of 16 is $1+i$, which has norm 2. We will again need alternative ways of representing characters:

$$f_q(n,k) = \begin{cases} i^e & \text{if } n \equiv k^e \bmod q, \\ 0 & \text{otherwise,} \end{cases}$$

$$g_q(n,k) = \begin{cases} i^e & \text{if } n \equiv k^e \bmod q \text{ or } q-n \equiv k^e \bmod q, \\ 0 & \text{otherwise,} \end{cases}$$

$$h_q(n,k,\ell,m) = \begin{cases} i^e & \text{if } n \equiv k^e \bmod q \text{ or } n \equiv \ell^e \bmod q, \\ (-1)^{e+1} & \text{if } q-n \equiv m^e \bmod q, \\ 0 & \text{otherwise,} \end{cases}$$

especially in the cases $q = 15, 16, 20$ and 35. The first several non-real quartic characters are

$$f_5(n,2) = \left(\tfrac{n}{-1-2i}\right)_4\Big|_{n=1,\ldots,5} = \{1,i,-i,-1,0\},$$

$$f_5(n,3) = \left(\tfrac{n}{-1+2i}\right)_4\Big|_{n=1,\ldots,5} = \{1,-i,i,-1,0\},$$

$$f_{13}(n,2) = \left(\tfrac{n}{3-2i}\right)_4\Big|_{n=1,\ldots,13} = \{1,i,1,-1,i,i,-i,-i,1,-1,-i,-1,0\},$$

$$f_{13}(n,7) = \left(\tfrac{n}{3+2i}\right)_4\Big|_{n=1,\ldots,13} = \{1,-i,1,-1,-i,-i,i,i,1,-1,i,-1,0\},$$

$$g_{15}(n,2)|_{n=1,\ldots,15} = \{1,i,0,-1,0,0,-i,-i,0,0,-1,0,i,1,0\},$$
$$g_{15}(n,8)|_{n=1,\ldots,15} = \{1,-i,0,-1,0,0,i,i,0,0,-1,0,-i,1,0\},$$
$$g_{16}(n,3)|_{n=1,\ldots,16} = \{1,0,i,0,-i,0,-1,0,-1,0,-i,0,i,0,1,0\},$$
$$g_{16}(n,5)|_{n=1,\ldots,16} = \{1,0,-i,0,i,0,-1,0,-1,0,i,0,-i,0,1,0\},$$
$$h_{16}(n,3,5,9)|_{n=1,\ldots,16} = \{1,0,i,0,i,0,1,0,-1,0,-i,0,-i,0,-1,0\},$$
$$h_{16}(n,11,13,9)|_{n=1,\ldots,16} = \{1,0,-i,0,-i,0,1,0,-1,0,i,0,i,0,-1,0\},$$
$$f_{17}(n,3) = \left(\tfrac{n}{1-4i}\right)_4\Big|_{n=1,\ldots,17} = \{1,-1,i,1,i,-i,-i,-1,-1,-i,-i,i,1,i,-1,1,0\},$$
$$f_{17}(n,6) = \left(\tfrac{n}{1+4i}\right)_4\Big|_{n=1,\ldots,17} = \{1,-1,-i,1,-i,i,i,-1,-1,i,i,-i,1,-i,-1,1,0\},$$
$$g_{20}(n,3)|_{n=1,\ldots,20} = \{1,0,i,0,0,0,-i,0,-1,0,-1,0,-i,0,0,0,i,0,1,0\},$$
$$g_{20}(n,7)|_{n=1,\ldots,20} = \{1,0,-i,0,0,0,i,0,-1,0,-1,0,i,0,0,0,-i,0,1,0\},$$

$$\begin{aligned} f_{29}(n,2) &= \left(\tfrac{n}{-5-2i}\right)_4\Big|_{n=1,\ldots,29} \\ &= \{1,i,i,-1,-1,-1,1,-i,-1,-i,i,-i,-1,i,-i,1,i,-i,i,1,\\ &\quad i,-1,1,1,1,-i,-i,-1,0\}, \end{aligned}$$

$$\begin{aligned} f_{29}(n,8) &= \left(\tfrac{n}{-5+2i}\right)_4\Big|_{n=1,\ldots,29} \\ &= \{1,-i,-i,-1,-1,-1,1,i,-1,i,-i,i,-1,-i,i,1,-i,i,-i,1,\\ &\quad -i,-1,1,1,1,i,i,-1,0\}, \end{aligned}$$

$$\begin{aligned} g_{35}(n,2)|_{n=1,\ldots,35} &= \{1,i,i,-1,0,-1,0,-i,-1,0,1,-i,i,0,0,1,-i,-i,1,0,\\ &\quad 0,i,-i,1,0,-1,-i,0,-1,0,-1,i,i,1,0\}, \end{aligned}$$

$$\begin{aligned} g_{35}(n,18)|_{n=1,\ldots,35} &= \{1,-i,-i,-1,0,-1,0,i,-1,0,1,i,-i,0,0,1,i,i,1,0,\\ &\quad 0,-i,i,1,0,-1,i,0,-1,0,-1,-i,-i,1,0\}, \end{aligned}$$

$$\begin{aligned} f_{37}(n,2) &= \left(\tfrac{n}{-1+6i}\right)_4\Big|_{n=1,\ldots,37} \\ &= \{1,i,-1,-1,-i,-i,1,-i,1,1,-1,1,-i,i,i,1,-i,i,-i,i,\\ &\quad -1,-i,-i,i,-1,1,-1,-1,i,-1,i,i,1,1,-i,-1,0\}, \end{aligned}$$

$$\begin{aligned} f_{37}(n,5) &= \left(\tfrac{n}{-1-6i}\right)_4\Big|_{n=1,\ldots,37} \\ &= \{1,-i,-1,-1,i,i,1,i,1,1,-1,1,i,-i,-i,1,i,-i,i,-i,\\ &\quad -1,i,i,-i,-1,1,-1,-1,-i,-1,-i,-i,1,1,i,-1,0\}. \end{aligned}$$

We mention that [4]

$$\begin{array}{l}\text{\# Dirichlet characters of}\\ \text{order } \ell \text{ and modulus } n\end{array} = \begin{array}{l}\text{\# solutions } x \text{ in } \mathbb{Z}_n^* \text{ of}\\ \text{the equation } x^\ell = 1\end{array}$$

and thus, by Möbius inversion,

$$\begin{array}{l}\text{\# primitive quadratic Dirichlet}\\ \text{characters of modulus } \le N\end{array} \sim \frac{6}{\pi^2}N,$$

$$\begin{array}{l}\text{\# primitive cubic Dirichlet}\\ \text{characters of modulus } \le N\end{array} \sim A\,N,$$

$$\begin{array}{l}\text{\# primitive quartic Dirichlet}\\ \text{characters of modulus } \le N\end{array} \sim B\,N\ln(N),$$

as $N \to \infty$, where [5–7]

$$A = \frac{11\sqrt{3}}{18\pi} \prod_{p \equiv 1 \bmod 3} \left(1 - \frac{2}{p(p+1)}\right) = 0.3170565167...,$$

$$B = \frac{7}{\pi} \frac{1}{16K^2} \prod_{p \equiv 1 \bmod 4} \left(1 - \frac{5p-3}{p^2(p+1)}\right) = 0.1908767211...$$

and K is the Landau–Ramanujan constant [8]. No one appears to have examined B before.

Now define the **Dirichlet L-series associated to** $\chi \neq 1$:

$$L_\chi(z) = \sum_{n=1}^{\infty} \chi(n) n^{-z} = \prod_p \left(1 - \chi(p) p^{-z}\right)^{-1}, \quad \mathrm{Re}(z) > 1,$$

which can be made into an entire function. Special values are more complicated for cubic/quartic characters than for quadratic characters [1]. For example, if $\chi = (\cdot/(2+3\omega))_3$, then

$$L_\chi(1) = 7^{-2/3}(-2-3\omega)^{1/3}\left(\omega^2 \ln(y_1) + \omega \ln(y_2) + \ln(y_3)\right)$$

where $y_1 < y_2 < y_3$ are the (real) zeroes of $y^3 - 7y^2 + 14y - 7$; if $\chi = f_9(\cdot, 2)$, then

$$L_\chi(1) = -\tfrac{2}{3}\omega^{1/3}\left(\omega^2 \ln\left(\sin\left(\tfrac{2\pi}{9}\right)\right) + \omega \ln\left(\cos\left(\tfrac{\pi}{18}\right)\right) + \ln\left(\sin\left(\tfrac{\pi}{9}\right)\right)\right).$$

As more examples, if $\chi = (\cdot/(-1-2i))_4$, then

$$L_\chi(1) = 2^{1/2} 5^{-5/4} (3+4i)^{1/4} \pi;$$

if $\chi = g_{16}(\cdot, 3)$, then

$$L_\chi(1) = -\tfrac{1}{2} i^{1/4}\left(i \ln\left(\cot\left(\tfrac{3\pi}{16}\right)\right) + \ln\left(\tan\left(\tfrac{\pi}{16}\right)\right)\right);$$

if $\chi = h_{16}(\cdot, 3, 5, 9)$, then

$$L_\chi(1) = 8^{-1/2} i^{1/4} \pi.$$

See a general treatment of quartic cases in [9].

The elaborate formulas for moments of $L_\chi(1/2)$ over primitive quadratic characters χ do not yet appear to have precise analogs for primitive cubic characters. Baier & Young [10] proved that

$$\sum_{q \le Q} \sum_{\chi} |L_\chi(1/2)|^2 = O\left(Q^{6/5+\varepsilon}\right)$$

as $Q \to \infty$, for any $\varepsilon > 0$, where the big-O constant depends on ε. The inner summation is over all primitive cubic characters modulo q. As a consequence, $L_\chi(1/2) \ne 0$ for infinitely many such χ.

1.33.1 *Cubic Twists*

Given an elliptic curve E over $\mathbb{Q}$ with L-series

$$L_E(z) = \sum_{n=1}^{\infty} c_n n^{-z},$$

the L-series obtained via twisting $L_E(z)$ by a cubic character χ is

$$L_{E,\chi}(z) = \sum_{n=1}^{\infty} \chi(n) c_n n^{-z}.$$

Of course, while each $c_n \in \mathbb{Z}$, the coefficients $\chi(n) c_n \in \mathbb{Z}[\omega]$ need not be real. This generalizes the sense of quadratic twists discussed in [11]; we refer to a paper of David, Fearnley & Kisilevsky [6] for more information on such L-series.

There is a different sense of cubic twists that interests us – it is important for the study of the family of elliptic curves F_d given by $x^3 + y^3 = d$ – and features the cubic residue symbol $(d/\cdot)_3$ in an intriguing way. We mentioned the problem of evaluating $L_{F_d}(1)$ for cube-free $d > 2$ in [11] but did not give details. By definition [12],

$$\begin{aligned} L_{F_d}(z) &= \sum_{\substack{a,b \in \mathbb{Z} \\ a \equiv 1 \bmod 3 \\ b \equiv 0 \bmod 3}} (a + b\omega^2) \left(\frac{d}{a + b\omega}\right)_3 (a^2 - ab + b^2)^{-z} \\ &= \sum_{\substack{a,b \in \mathbb{Z} \\ a \equiv 1 \bmod 3 \\ b \equiv 0 \bmod 3}} (a + b\omega) \left(\frac{d}{a + b\omega^2}\right)_3 (a^2 - ab + b^2)^{-z} \\ &= \prod_{p \equiv 2 \bmod 3} (1 + p^{1-2z})^{-1} \cdot \prod_{p \equiv 1 \bmod 3} (1 - c_p p^{-z} + p^{1-2z})^{-1} \end{aligned}$$

where

$$c_p = (h + k\omega^2)\left(\frac{d}{h+k\omega}\right)_3 + (h + k\omega)\left(\frac{d}{h+k\omega^2}\right)_3$$

and $p = (h + k\omega)(h + k\omega^2)$, $h \equiv 1 \bmod 3$, $k \equiv 0 \bmod 3$. To extend to composite indices, use the usual recurrence $c_{p^j} = c_{p^{j-1}} c_p - p\, c_{p^{j-2}}$ for $j \geq 2$, $c_1 = 1$ and $c_{mn} = c_m c_n$ for coprime integers m, n.

For $d = 1$ and $p \equiv 1 \bmod 3$, it is known that $c_p = \gamma_p$, where γ_p is the unique integer $\alpha \equiv 2 \bmod 3$ such that $\alpha^2 + 3\beta^2 = 4p$ for some integer $\beta \equiv 0 \bmod 3$. Now, for $d > 1$ and $p \equiv 1 \bmod 3$, $p \nmid d$, it can be shown that c_p is the unique integer $\alpha \equiv 2 \bmod 3$ such that three conditions:

- $\alpha^2 + 3\beta^2 = 4p$ for some integer β
- $\alpha \equiv d^{(p-1)/3}\gamma_p \bmod p$
- $|\alpha| < 2\sqrt{p}$

are simultaneously satisfied [13].

Sextic twists are required to study Bachet's equation $y^2 = x^3 + n$ for arbitrary n (the Fermat cubic problem is a special case with $n = -432d^2$ and d cube-free). Such residue symbols are beyond us. Here is a formula for L-series coefficients c_p in this more general setting: when $p = 3$, $p|n$ or $p \equiv 2 \bmod 3$, we have $c_p = 0$; otherwise [14]

$$c_p = \left(\frac{n}{p}\right) \cdot \begin{cases} 2a - b & \text{if } (4n)^{(p-1)/3} \equiv 1 \bmod p, \\ -a - b & \text{if } (4n)^{(p-1)/3} b \equiv -a \bmod p, \\ 2b - a & \text{if } (4n)^{(p-1)/3} a \equiv -b \bmod p, \end{cases}$$

where $p = a^2 - ab + b^2$ with $a \equiv 1 \bmod 3$, $b \equiv 0 \bmod 3$ and $(\cdot/\cdot)$ is the Kronecker–Jacobi–Legendre symbol. The sequence of integers for which $y^2 = x^3 + n$ has zero rank [15]:

$$\ldots, -12, -10, -9, -8, -6, -5, -3, -1, 1, 4, 6, 7, 13, 14, 16, 20, \ldots$$

deserves close attention!

1.33.2 *Quartic Twists*

Quartic twists are required to study $y^2 = x^3 - n\,x$ for arbitrary n (the congruent number problem is a special case with $n = d^2$ and d square-free [11]). Analogous to the expression for $L_{F_d}(z)$,

$$L_{\mathrm{E}_n}(z) = \sum_{\substack{a,b\in\mathbb{Z}\\ a\equiv 1 \bmod 4\\ b\equiv 0 \bmod 2}} (a-bi)\left(\frac{-n}{a+bi}\right)_4 (a^2+b^2)^{-z}$$
$$= \sum_{\substack{a,b\in\mathbb{Z}\\ a\equiv 1 \bmod 4\\ b\equiv 0 \bmod 2}} (a+bi)\left(\frac{-n}{a-bi}\right)_4 (a^2+b^2)^{-z}.$$

Here also is the corresponding formula for L-series coefficients c_p: when $p=2$, $p|n$ or $p\equiv 3 \bmod 4$, we have $c_p=0$; otherwise [14]

$$c_p = 2\left(\frac{2}{p}\right)\cdot \begin{cases} -a & \text{if } n^{(p-1)/4}\equiv 1 \bmod p, \\ a & \text{if } n^{(p-1)/4}\equiv -1 \bmod p, \\ -b & \text{if } n^{(p-1)/4}b\equiv -a \bmod p, \\ b & \text{if } n^{(p-1)/4}b\equiv a \bmod p, \end{cases}$$

where $p=a^2+b^2$ with $a\equiv 3 \bmod 4$, $b\equiv 0 \bmod 2$. Again, the sequence of integers for which $y^2=x^3-n\,x$ has zero rank [15]:

$$\ldots, -12, -11, -10, -7, -6, -4, -2, -1, 1, 3, 4, 8, 9, 11, 13, 18, \ldots$$

is worthy of deeper study.

As a quintic follow-on to [5, 7], we merely mention [16].

[1] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[2] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976, pp. 55–62, 138–139, 229; MR0434929 (55 #7892).
[3] B. C. Berndt, R. J. Evans and K. S. Williams, *Gauss and Jacobi Sums*, Wiley, 1998, pp. 234–251; MR1625181 (99d:11092).
[4] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A060594, A114643, A060839, A160498, A073103 and A160499.
[5] S. Finch and P. Sebah, Squares and cubes modulo n, arXiv:math/0604465.
[6] C. David, J. Fearnley and H. Kisilevsky, On the vanishing of twisted L-functions of elliptic curves, *Experim. Math.* 13 (2004) 185–198; MR2068892 (2005e:11082).
[7] S. Finch, Quartic and octic characters modulo n, arXiv:0907.4894.
[8] S. R. Finch, Landau-Ramanujan constant, *first volume*, pp. 98–104.
[9] K. Hardy and K. S. Williams, Evaluation of the infinite series $\sum_{\substack{n=1\\(n/p)=1}}^{\infty}\left(\frac{n}{p}\right)_4 n^{-1}$, *Proc. Amer. Math. Soc.* 109 (1990) 597–603; MR1019275 (90k:11114).
[10] S. Baier and M. P. Young, Mean values with cubic characters, *J. Number Theory* 130 (2010) 879–903; arXiv:0804.2233; MR2600408 (2011c:11138).
[11] S. R. Finch, Elliptic curves over $\mathbb{Q}$, *this volume*, §1.14.
[12] N. M. Stephens, The diophantine equation $X^2+Y^2=DZ^2$ and the conjectures of Birch and Swinnerton-Dyer, *J. Reine Angew. Math.* 231 (1968) 121–162; MR0229651 (37 #5225).

[13] D. Zagier and G. Kramarz, Numerical investigations related to the L-series of certain elliptic curves, *J. Indian Math. Soc.* 52 (1987) 51–69; MR0989230 (90d:11072).

[14] H. Cohen, *Number Theory*. v. 1, *Tools and Diophantine Equations*, Springer-Verlag, 2007, pp. 566–570; MR2312337 (2008e:11001).

[15] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002150, A002151, A002156 and A002158.

[16] S. Finch and P. Sebah, Residue of a mod 5 Euler product, arXiv:0912.3677.

1.34 Distribution of Error Terms

Let ν_N be a random integer chosen uniformly in $[1, N]$. Let $\varphi(n)$ denote the number of positive integers $m \leq n$ satisfying $\gcd(m, n) = 1$ and $\sigma(n)$ denote the sum of all divisors of n. The limiting probability distributions of $\varphi(\nu_N)/\nu_N$ and $\sigma(\nu_N)/\nu_N$, as $N \to \infty$, are continuous but singular in the sense that

$$F_\varphi(x) = \lim_{N\to\infty} \frac{\#\{n \leq N : \varphi(n)/n \leq x\}}{N}, \quad F_\sigma(x) = \lim_{N\to\infty} \frac{\#\{n \leq N : \sigma(n)/n \leq x\}}{N}$$

satisfy $F'_\varphi = 0 = F'_\sigma$ almost everywhere [1–3]. Considerable effort is needed, for example, to compute that $1 - F_\sigma(2) = 0.247...$, the density of abundant numbers relative to the set of positive integers [4]. See [5–9] for recent work concerning F_φ and F_σ.

Starting from

$$\lim_{N\to\infty} \mathrm{E}\left(\frac{\varphi(\nu_N)}{\nu_N}\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} \frac{\varphi(n)}{n} = \frac{6}{\pi^2},$$

$$\lim_{N\to\infty} \mathrm{E}\left(\frac{\sigma(\nu_N)}{\nu_N}\right) = \lim_{N\to\infty} \frac{1}{N} \sum_{n\leq N} \frac{\sigma(n)}{n} = \frac{\pi^2}{6},$$

we examine distributions that perhaps are more open to analysis. Define error terms

$$H(n) = \sum_{m\leq n} \frac{\varphi(m)}{m} - \frac{6}{\pi^2} n,$$

$$K(n) = \sum_{m\leq n} \frac{\sigma(m)}{m} - \frac{\pi^2}{6} n + \frac{1}{2} \ln(n) + \frac{\gamma + \ln(2\pi)}{2},$$

then it can be shown that [10–19]

$$\lim_{N\to\infty} \mathrm{E}\left(H(\nu_N)\right) = \frac{3}{\pi^2}, \quad \lim_{N\to\infty} \mathrm{Var}\left(H(\nu_N)\right) = \frac{1}{2\pi^2} - \frac{3}{\pi^4},$$

$$\lim_{N\to\infty} \mathrm{E}\left(K(\nu_N)\right) = \frac{\pi^2}{12}, \quad \lim_{N\to\infty} \mathrm{Var}\left(K(\nu_N)\right) = \frac{5\pi^2}{144} - \frac{\pi^4}{432}.$$

Further, it is known that the limiting distributions corresponding to $H(\nu_N) - 3/\pi^2$ and $K(\nu_N) - \pi^2/12$ are symmetric and all corresponding odd moments vanish. In particular, the skewness coefficients of both quantities are zero. What is *not* precisely known are the kurtosis excesses:

$$\frac{\mathrm{E}\left[\left(H(\nu_N) - \mathrm{E}\left(H(\nu_N)\right)\right)^4\right]}{\mathrm{Var}\left(H(\nu_N)\right)^2} - 3 = -0.93...,$$

$$\frac{\mathrm{E}\left[\left(K(\nu_N) - \mathrm{E}\left(K(\nu_N)\right)\right)^4\right]}{\mathrm{Var}\left(K(\nu_N)\right)^2} - 3 = 0.10...,$$

which would imply that tails are thin for $H(\nu_N)$ and tails are fat for $K(\nu_N)$. This may be a consequence of the simple fact that the support of the distribution for $\varphi(\nu_N)/\nu_N$ is $[0,1]$ whereas the support of the distribution for $\sigma(\nu_N)/\nu_N$ is $[0,\infty)$.

Exact formulas for all even moments would allow us to accurately construct the distributions corresponding to $H(\nu_N)$ and $K(\nu_N)$. Evaluating the fourth moments, however, seems to be hard. Related material includes [20–23].

[1] I. Schoenberg, Über die asymptotische Verteilung reeller Zahlen mod 1, *Math. Z.* 28 (1918) 171–199.
[2] P. Erdős, On the smoothness of the asymptotic distribution of additive arithmetical functions, *Amer. J. Math.* 61 (1939) 722–725; MR0000248 (1,41a).
[3] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Univ. Press, 1995, pp. 281–296; MR1342300 (97e:11005b).
[4] S. R. Finch, Abundant numbers density constant, *first volume*, pp. 126–127.
[5] V. Toulmonde, Module de continuité de la fonction de répartition de $\phi(n)/n$, *Acta Arith.* 121 (2006) 367–402; MR2224402 (2007e:11090).
[6] V. Toulmonde, Sur les variations de la fonction de répartition de $\phi(n)/n$, *J. Number Theory* 120 (2006) 1–12; MR2256793 (2007g:11123).
[7] G. Tenenbaum and V. Toulmonde, Sur le comportement local de la répartition de l'indicatrice d'Euler, *Funct. Approx. Comment. Math.* 35 (2006) 321–338; MR2271621 (2007g:11126).
[8] A. Weingartner, The distribution functions of $\sigma(n)/n$ and $n/\phi(n)$, *Proc. Amer. Math. Soc.* 135 (2007) 2677–2681; MR2317939 (2009c:11156).
[9] A. Weingartner, The limiting distribution of the divisor function, *J. Number Theory* 129 (2009) 1432–1442; MR2521483 (2010i:11127).
[10] A. Walfisz, Teilerprobleme. Zweite Abhandlung, *Math. Z.* 34 (1932) 448–472.
[11] S. S. Pillai and S. D. Chowla, On the error terms in some asymptotic formulae in the theory of numbers (I), *J. London Math. Soc.* 5 (1930) 95–101.
[12] S. D. Chowla and S. S. Pillai, On the error terms in some asymptotic formulae in the theory of numbers (II), *J. Indian Math. Soc.* 18 (1930) 181–184.
[13] S. D. Chowla, Contributions to the analytic theory of numbers, *Math. Z.* 35 (1932) 279–299.
[14] P. Erdős and H. N. Shapiro, On the changes of sign of a certain error function, *Canad. J. Math.* 3 (1951) 375–385; MR0045144 (13,535i).
[15] P. Erdős and H. N. Shapiro, The existence of a distribution function for an error term related to the Euler function, *Canad. J. Math.* 7 (1955) 63–75; MR0065580 (16,448f).
[16] D. Suryanarayana and R. Sitaramachandra Rao, On the average order of the function $E(x) = \sum_{n \le x} \phi(n) - 3x^2/\pi^2$, *Ark. Mat.* 10 (1972) 99–106; MR0309883 (46 #8988).

[17] D. Suryanarayana, On the average order of the function $E(x) = \sum_{n \le x} \phi(n) - 3x^2/\pi^2$. II, *J. Indian Math. Soc.* 42 (1978) 179–195; MR0558994 (81a:10059).

[18] Y.-F. S. Pétermann, An Ω-theorem for an error term related to the sum-of-divisors function, *Monatsh. Math.* 103 (1987) 145–157; addendum, 105 (1988) 193–194; MR0881720 (88e:11086) and MR0939941 (89g:11087).

[19] Y.-F. S. Pétermann, On the distribution of values of an error term related to the Euler function, *Théorie des nombres*, Proc. 1987 Québec conf., ed. J.-M. De Koninck and C. Levesque, de Gruyter, 1989, pp. 785–797; MR1024603 (91d:11117).

[20] P. Codecà, On the B^λ almost periodic behaviour of certain arithmetical convolutions, *Rend. Sem. Mat. Univ. Padova* 72 (1984) 373–387; MR0778352 (86e:42021).

[21] P. Codecà, On the properties of oscillation and almost periodicity of certain convolutions, *Rend. Sem. Mat. Univ. Padova* 71 (1984) 103–119; MR0769430 (87d:11069).

[22] J. Liu, On an error term of Chowla. I, II, III, *J. Number Theory* 64 (1997) 20–35, 36–50, 51–58; MR1450484 (98d:11113).

[23] P. J. Almeida, Sign changes of error terms related to arithmetical functions, *J. Théor. Nombres Bordeaux* 19 (2007) 1–25; MR2332050 (2008d:11115).

1.35 Cilleruelo's LCM Constants

Let a, b be coprime integers such that $a \ge 1$, $a + b \ge 1$. The Prime Number Theorem for Arithmetic Progressions implies that

$$\ln\left(\operatorname*{lcm}_{1\le k\le n}\{a\,k+b\}\right) \sim A\,n$$

as $n \to \infty$, where the constant A is

$$A = \frac{a}{\varphi(a)} \sum_{\substack{1\le j\le a,\\ \gcd(j,a)=1}} \frac{1}{j}$$

(independent of b) and φ is the Euler totient function [1, 2]. What happens if we replace the linear polynomial $ax + b$ by a quadratic polynomial $ax^2 + bx + c$? On the one hand, if the quadratic is reducible over the integers, then there is not much change (the growth rate is still $A\,n$ for some new rational number A). On the other hand, if the quadratic is irreducible over the integers, then there is a more interesting outcome [3]:

$$\ln\left(\operatorname*{lcm}_{1\le k\le n}\left\{a\,k^2+b\,k+c\right\}\right) = n\ln(n) + B\,n + o(n)$$

as $n \to \infty$, where the constant B will occupy our attention for the remainder of this essay.

Henceforth we set $a=1$, $b=0$, $c\in\{1,2,-2\}$. It follows that the fundamental discriminant $d\in\{-4,-8,8\}$. The constant B for our three special cases is

$$B = \gamma - 1 - \frac{1}{2}\ln(2) - \sum_{k=1}^{\infty}\left(\frac{\zeta'(2^k)}{\zeta(2^k)} - \frac{L_d'(2^k)}{L_d(2^k)} + \frac{\ln(2)}{2^{2^k}-1}\right) + \frac{L_d'(1)}{L_d(1)}$$

$$= \begin{cases} -0.0662756342... & \text{if } c=1, \\ -0.4895081630... & \text{if } c=2, \\ 0.3970903472... & \text{if } c=-2. \end{cases}$$

As an example, if $c=1$, we have [4]

$$\frac{L_{-4}'(1)}{L_{-4}(1)} = \ln\left(2\pi e^{\gamma}\frac{\Gamma(\frac{3}{4})^2}{\Gamma(\frac{1}{4})^2}\right) = \ln\left(\frac{\pi^2 e^{\gamma}}{2\Lambda^2}\right)$$

where Λ is Gauss' lemniscate constant [5]; it can be shown here that

$$B = -3 - \frac{3}{2}\ln(2) + 2\gamma + 4\tilde{C}$$

where $\tilde{C} = 0.7047534517...$ is the second-order constant corresponding to non-hypotenuse numbers [6, 7]. Similar relationships with second-order constants listed in [8] can be found.

Cilleruelo [3] further noted that, in the general case,

$$B = C_0 + C_d + C(f)$$

where

$$C_0 = \gamma - 1 - 2\ln(2) - \sum_{k=1}^{\infty}\frac{\zeta'(2^k)}{\zeta(2^k)} = -1.1725471674...$$

is universal,

$$C_d = \sum_{k=0}^{\infty}\frac{L_d'(2^k)}{L_d(2^k)} - \sum_{p|d}\sum_{k=1}^{\infty}\frac{\ln(p)}{p^{2^k}-1}$$

depends only on d, and $C(f)$ is too complicated to reproduce (but is equal to $(3/2)\ln(2)$ for our three special cases). Although other irreducible quadratics are examined in [3], we note the absence of $x^2 \pm 3$ and wonder what can be deduced here. See also [9–12].

[1] P. Bateman, J. Kalb and A. Stenger, A limit involving least common multiples, *Amer. Math. Monthly* 109 (2002) 393–394.
[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003418, A193181.
[3] J. Cilleruelo, The least common multiple of a quadratic sequence, *Compos. Math.* 147 (2011) 1129–1150; arXiv:1001.3438; MR2822864 (2012f:11185).
[4] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[5] S. R. Finch, Gauss' lemniscate constant, *first volume*, pp. 420–423.

[6] P. Moree, Counting numbers in multiplicative sets: Landau versus Ramanujan, *Math. Newsl.* 21 (2011) 73–81; arXiv:1110.0708; MR3012680.
[7] S. R. Finch, Landau-Ramanujan constant, *first volume*, pp. 98–104.
[8] S. R. Finch, Prime number theorem, *this volume*, §1.25.
[9] J. Rué, P. Šarka and A. Zumalacárregui, On the error term of the logarithm of the lcm of a quadratic sequence, *J. Théor. Nombres Bordeaux* 25 (2013) 457–470; arXiv:1110.0939; MR3228315.
[10] G. Qian, Q. Tan and S. Hong, The least common multiple of consecutive terms in a quadratic progression, *Bull. Aust. Math. Soc.* 86 (2012) 389–404; arXiv:1208.5119; MR2995891.
[11] B. Farhi, Nontrivial lower bounds for the least common multiple of some finite sequences of integers, *J. Number Theory* 125 (2007) 393–411; MR2332595 (2008i:11001).
[12] S.-M. Oon, Note on the lower bound of least common multiple, *Abstr. Appl. Anal.* 2013, Art. ID 218125; MR3034956.

1.36 Amicable Pairs and Aliquot Sequences

If n is a positive integer, let $s(n)$ denote the sum of all positive divisors of n that are strictly less than n. Then n is said to be **perfect** or **1-sociable** if $s(n) = n$. We mentioned perfect numbers in [1], asking whether infinitely many exist, but did not report their reciprocal sum [2]

$$\frac{1}{6} + \frac{1}{28} + \frac{1}{496} + \frac{1}{8128} + \frac{1}{33550336} + \frac{1}{8589869056} + \cdots = 0.2045201428....$$

This constant can, in fact, be rigorously calculated to 149 digits (and probably much higher accuracy if needed).

Define $s^k(n)$ to be the k^{th} iterate of s with starting value n. The integer n is **amicable** or **2-sociable** if $s^2(n) = n$ but $s(n) \neq n$. Such phrasing is based on older terminology [3]: two distinct integers m, n are said to form an "amicable pair" if $s(m) = n$ and $s(n) = m$. The (infinite?) sequence of amicable numbers possesses zero asymptotic density [4] and, further, has reciprocal sum [5–8]

$$\begin{aligned} &\frac{1}{220} + \frac{1}{284} + \frac{1}{1184} + \frac{1}{1210} + \frac{1}{2620} + \frac{1}{2924} + \frac{1}{5020} + \frac{1}{5564} \\ &+ \frac{1}{6232} + \frac{1}{6368} + \cdots = 0.0119841556.... \end{aligned}$$

In contrast with the preceding, *none* of the digits are provably correct. The best rigorous upper bound for this constant is 222; deeper understanding of the behavior of amicable numbers will be required to improve upon this poor estimate.

Fix $k \geq 3$. An integer n is k**-sociable** if $s^k(n) = n$ but $s^\ell(n) \neq n$ for all $1 \leq \ell < k$. No examples of 3-sociable numbers are known [9, 10]; the first example for $4 \leq k < 28$ is the 5-cycle $\{12496, 14288, 15472, 14536, 14264\}$ and the next example is the 4-cycle $\{1264460, 1547860, 1727636, 1305184\}$. Let S_k denote the sequence of all k-sociable numbers and S be the union of S_k over all k. It is conjectured that the (infinite?) sequence S possesses zero asymptotic density and progress toward confirming this appears in [11]. No one is ready to compute the reciprocal sum of S; a proof of convergence would seem to be faraway.

As an aside, we mention the sequence of **prime-indexed primes**, which is clearly infinite and has reciprocal sum [12]

$$\frac{1}{3}+\frac{1}{5}+\frac{1}{11}+\frac{1}{17}+\frac{1}{31}+\frac{1}{41}+\frac{1}{59}+\frac{1}{67}+\frac{1}{83}+\frac{1}{109}+\frac{1}{127}$$
$$+\frac{1}{157}+\frac{1}{179}+\frac{1}{191}+\cdots=1.0432015....$$

Again, this is conjectural only. The best rigorous lower/upper bounds for this constant are 1.04299 and 1.04365 [2]. Such bounds are tighter than those (1.83408 and 2.34676) for the reciprocal sum of twin primes [13].

A positive integer n is **deficient** if $s(n)<n$. A **primitive nondeficient number** is nondeficient yet all its proper divisors are deficient. As another aside, we mention the reciprocal sum of such numbers [8]:

$$\frac{1}{6}+\frac{1}{20}+\frac{1}{28}+\frac{1}{70}+\frac{1}{88}+\frac{1}{104}+\frac{1}{272}+\frac{1}{304}$$
$$+\frac{1}{368}+\frac{1}{464}+\frac{1}{496}+\frac{1}{550}+\cdots=0.3481648657...$$

and note that the best rigorous upper bound for this constant is 13.7.

Our main interest is in the "aliquot sequence" $\{s^k(n)\}_{k=1}^{\infty}$, where we assume without loss of generality that n is even. For example, if $n=12$, the sequence $\{16,15,9,4,3,1\}$ is finite (terminates at 1). From earlier, we know that infinite cyclic behavior is possible. Does an infinite *unbounded* aliquot sequence exist? On the one hand, starting with $n=276$, extensive computation has yielded 1769 terms with no end in sight [14–18]; probabilistic arguments in [19, 20], based on the arithmetic mean of $s(2n)/(2n)$, also support a belief that most sequences grow without bound.

On the other hand, the geometric mean of $s(2n)/(2n)$:

$$\sqrt[N]{\prod_{n=1}^{N}\frac{s(2n)}{2n}}=\exp\left(\frac{1}{N}\sum_{n=1}^{N}\ln\left(\frac{s(2n)}{2n}\right)\right)$$

(which seems a more appropriate tool than a simple average) predicts the opposite. Bosma & Kane [21] proved that

$$\begin{aligned}\lambda &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\ln\left(\frac{s(2n)}{2n}\right)=2\alpha(2)+\sum_{p\geq 3}\alpha(p)-\sum_{j\geq 1}\left((2\beta_j(2)-1)\prod_{p\geq 3}\beta_j(p)\right)\frac{1}{j}\\ &= -0.0332594808...<0,\end{aligned}$$

which implies that the geometric mean $\mu=\exp(\lambda)=0.9672875344...<1$. The indicated numerical estimates are due to Sebah [22]. Sums and products over

p are restricted to primes; further,

$$\alpha(p) = \left(1 - \frac{1}{p}\right) \sum_{m=1}^{\infty} \frac{1}{p^m} \ln\left(\frac{p^{m+1} - 1}{p^m(p-1)}\right),$$

$$\beta_j(p) = \left(1 - \frac{1}{p}\right) \sum_{m=0}^{\infty} \frac{1}{p^m} \left(\frac{p^{m+1} - 1}{p^m(p-1)}\right)^{-j}.$$

The fact that $\mu < 1$ suggests that aliquot sequences tend to decrease ultimately, evidence in favor of the Catalan–Dickson conjecture. It would be good to compute other related constants, appearing in [23], to similar levels of precision.

From [1, 24], the probability that $s(n)$ exceeds n, for arbitrary n, is

$$\lim_{n\to\infty} \frac{1}{n} \cdot \left| \left\{ i \le n : \frac{s(i)}{i} > 1 \right\} \right| = 0.2476...$$

(what was called $A(2)$). The fact that these odds are significantly less than $1/2$ again suggests that unboundedness is a rare event, if it occurs at all.

[1] S. R. Finch, Abundant numbers density constant, *first volume*, pp. 126–127.
[2] J. Bayless and D. Klyve, Reciprocal sums as a knowledge metric: theory, computation, and perfect numbers, *Amer. Math. Monthly* 120 (2013) 822–831; MR3115443.
[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000396, A006039, A006450, A063990, and A122726.
[4] P. Erdős, On amicable numbers, *Publ. Math. Debrecen* 4 (1955) 108–111; MR0069198 (16,998h).
[5] C. Pomerance, On the distribution of amicable numbers. II, *J. Reine Angew. Math.* 325 (1981) 183–188; MR0618552 (82m:10012).
[6] J. Bayless and D. Klyve, On the sum of reciprocals of amicable numbers, *Integers* 11 (2011) 315–332; arXiv:1101.0259; MR2988065.
[7] H. M. Nguyen and C. Pomerance, The reciprocal sum of the amicable numbers, submitted.
[8] J. D. Lichtman, The reciprocal sum of primitive nondeficient numbers, arXiv:1801.01925.
[9] D. Moews, A list of aliquot cycles of length greater than 2, unpublished table (2016).
[10] J. M. Pedersen, Tables of aliquot cycles, unpublished table (2007).
[11] M. Kobayashi, P. Pollack and C. Pomerance, On the distribution of sociable numbers, *J. Number Theory* 129 (2009) 1990–2009; MR2522719 (2010h:11162).
[12] J. Bayless, D. Klyve, and T. Oliveira e Silva, New bounds and computations on prime-indexed primes, *Integers* 13 (2013) 613–633; MR3097157.
[13] S. R. Finch, Brun's constant, *first volume*, pp. 133–135.
[14] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A008892, A014360, A014361, A014362, and A014363.
[15] M. Benito, W. Creyaufmüller, J. L. Varona and P. Zimmermann, Aliquot sequence 3630 ends after reaching 100 digits, *Experim. Math.* 11 (2002) 201–206; MR1959263 (2003j:11150).
[16] W. Creyaufmüller, Aliquot sequences, unpublished table (2016).
[17] P. Zimmermann, Aliquot sequences, unpublished table (2009).

[18] M. Tervooren, Aliquot sequences, The Factoring Database.
[19] M. C. Wunderlich, The aliquot project: an application of job chaining in number theoretic computing, *Proc. 3rd ACM Symp. on Symbolic and Algebraic Computation (SYMSAC)*, ed. R. D. Jenks, ACM, 1976, pp. 276–284.
[20] R. K. Guy and J. L. Selfridge, What drives an aliquot sequence? *Math. Comp.* 29 (1975) 101–107; corrigendum 34 (1980) 319–321; MR0384669 (52 #5542) and MR0551309 (81f:10008).
[21] W. Bosma and B. Kane, The aliquot constant, *Quart. J. Math.* 63 (2012) 309–323; arXiv:0912.3660; MR2925292.
[22] P. Sebah, Calculating the aliquot constant to 20 digits accuracy, unpublished note (2013).
[23] S. R. Finch, Series involving arithmetic functions, *this volume*, §1.20.
[24] S. R. Finch, Distribution of error terms, *this volume*, §1.34.

1.37 Fermat Numbers and Elite Primes

The Fermat numbers $F_n = 2^{2^n} + 1$ satisfy a quadratic recurrence [1]

$$F_{n+1} = (F_n - 1)^2 + 1, \quad n \geq 0$$

and are pairwise coprime. It is conjectured that F_n are always square-free and that, beyond F_4, they are never prime. The latter would imply that there are exactly 31 regular polygons with an odd number G_m of sides that can be constructed by straightedge and compass [2]. The values $G_1, G_2, \ldots, G_{31}$ encompass all divisors of $2^{32} - 1$ except unity [3]. Let $G_0 = 1$. If we scan each row of Pascal's triangle modulo 2 as a binary integer, then the numbers G_m (listed in ascending order) are naturally extended without bound. The reciprocal sum [4]

$$\sum_{m=0}^{\infty} \frac{1}{G_m} = \prod_{n=0}^{\infty} \left(1 + \frac{1}{F_n}\right) = 1.7007354952...$$

is irrational [2]; by contrast,

$$\sum_{m=0}^{\infty} \frac{(-1)^{t_m}}{G_m} = \frac{1}{2}$$

is rational, where $\{t_m\}$ is the Thue–Morse sequence $\{0, 1, 1, 0, 1, 0, 0, 1, 1, 0, ...\}$ [5]. Golomb [6] proved that

$$\sum_{n=0}^{\infty} \frac{1}{F_n} = 0.5960631721...$$

is irrational and Duverney [7] proved that it is transcendental; there is evidence that Mahler possessed these results far earlier [8].

Let P denote the set of all primes p for which there exists n such that p divides F_n. Křížek, Luca & Somer [9] proved that

$$\sum_{p \in P} \frac{1}{p} = 0.5976404758...$$

is convergent, answering a question raised in [10]. The series $\sum_{d\in D} 1/d$ likewise converges, where D is the set of all divisors $d > 1$ (prime or composite) for which there exists n such that d divides F_n. The smallest element of D not in P is F_5 itself [11, 12].

A prime p is called **elite** [13] if there exists m for which all F_n with $n > m$ are quadratic non-residues of p, that is, the equation

$$x^2 \equiv F_n \bmod p$$

has no solutions x for $n > m$. Let E denote the (infinite?) set of all elite primes. The series [14–17]

$$\sum_{p\in E} \frac{1}{p} = 0.7007640115...$$

is convergent [9]. This numerical evaluation, as well as that for the series over $p \in P$, is non-rigorous. For our calculation over $p \in E$ to be valid, we would need

$$\#\{p \in E : p \leq q\} = O(\ln(q))$$

as $q \to \infty$; the best current bound is $O\left(q/\ln(q)^2\right)$, hence improvement in our knowledge of E will be required. Generalization to the numbers $F_{b,n} = b^{2^n} + 1$, for fixed integer $b \geq 2$, is found in [18].

We conclude with the fact that

$$\sum_{n=0}^{\infty} \frac{1}{2^{2^n}} = 0.8164215090...$$

is transcendental, proved by Kempner [19] and revisited in [20].

[1] S. R. Finch, Quadratic recurrence constants, *first volume*, pp. 443–448.
[2] M. Křížek, F. Luca and L. Somer, *17 Lectures on Fermat Numbers: From Number Theory to Geometry*, Springer-Verlag, 2001, pp. 33–35, 44, 75–77,104–113; MR1866957 (2002i:11001).
[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000215, A001317, A004729 and A045544.
[4] V. Shevelev, On Stephan's conjectures concerning Pascal triangle modulo 2, arXiv:1011.6083.
[5] S. R. Finch, Prouhet-Thue-Morse constant, *first volume*, pp. 436–441.
[6] S. W. Golomb, On the sum of the reciprocals of the Fermat numbers and related irrationalities, *Canad. J. Math.* 15 (1963) 475–478; MR0150102 (27 #105).
[7] D. Duverney, Transcendence of a fast converging series of rational numbers, *Math. Proc. Cambridge Philos. Soc.* 130 (2001) 193–207; MR1806772 (2002e:11089).
[8] M. Coons, On the rational approximation of the sum of the reciprocals of the Fermat numbers, *Ramanujan J.* 30 (2013) 39–65; arXiv:1112.5072; MR3010463.
[9] M. Křížek, F. Luca and L. Somer, On the convergence of series of reciprocals of primes related to the Fermat numbers, *J. Number Theory* 97 (2002) 95–112; MR1939138 (2003i:11015).

[10] S. W. Golomb, Sets of primes with intermediate density, *Math. Scand.* 3 (1955) 264–274; MR0075981 (17,828d).
[11] W. Keller, Prime factors $k \cdot 2^n + 1$ of Fermat numbers F_m and complete factoring status, unpublished table (2016).
[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A023394 and A102742.
[13] A. Aigner, Über Primzahlen, nach denen (fast) alle Fermatschen Zahlen quadratische Nichtreste sind, *Monatsh. Math.* 101 (1986) 85–93; MR0843294 (87g:11010).
[14] A. Chaumont and T. Müller, All elite primes up to 250 billion, *J. Integer Seq.* 9 (2006) 06.3.8; MR2240858 (2007e:11005).
[15] T. Müller, Searching for large elite primes, *Experim. Math.* 15 (2006) 183–186; MR2253004 (2008a:11008).
[16] A. Chaumont, J. Leicht, T. Müller and A. Reinhart, The continuing search for large elite primes, *Internat. J. Number Theory* 5 (2009) 209–218; MR2502805 (2010d:11003).
[17] D. R. Martin, Elite prime search complete to 1E14, unpublished table (2010).
[18] T. Müller, A generalization of a theorem by Křížek, Luca, and Somer on elite primes, *Analysis* 28 (2008) 375–382; MR2477353 (2010a:11013).
[19] A. J. Kempner, On transcendental numbers, *Trans. Amer. Math. Soc.* 17 (1916) 476–482.
[20] B. Adamczewski, The many faces of the Kempner number, *J. Integer Seq.* 16 (2013) 13.2.15; arXiv:1303.1685; MR3032398.

1.38 Average Least Nonresidues

Fifty years separate two computations: the mean value of a certain function $f(p)$ over primes p, mentioned in [1], and the mean value of $f(m)$ over all positive integers m. We anticipate that the overlap between number theory and probability will only deepen with time.

1.38.1 *Quadratic*

Let $f(m)$ be the smallest positive quadratic nonresidue modulo $m > 2$. Erdős [2] proved that

$$\lim_{x\to\infty}\left(\sum_{2<p\le x} 1\right)^{-1} \sum_{2<p\le x} f(p) = \sum_{k=1}^{\infty} \frac{p_k}{2^k} = 3.6746439660...$$

where $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, … is the sequence of prime numbers. Pollack [3, 4] extended this result to

$$\lim_{x\to\infty}\left(\sum_{2<m\le x} 1\right)^{-1} \sum_{2<m\le x} f(m) = \sum_{k=1}^{\infty} \frac{p_k - 1}{p_1 p_2 \cdots p_{k-1}} = 2.9200509773....$$

In words, the right-hand side is the average value of the least prime not dividing m.

1.38.2 Character

Given a fundamental discriminant D, let $F(D)$ be the least positive integer n for which $(D/n) \notin \{0,1\}$. The set of all real primitive Dirichlet characters χ, except the principal character χ_0, is encompassed by $(D/.)$ as D runs over all fundamental discriminants [5]. It can be shown that [3, 6]

$$\lim_{x\to\infty}\left(\sum_{|D|\le x} 1\right)^{-1}\sum_{|D|\le x} F(D) = \sum_{q}\frac{q^2}{2(q+1)}\prod_{p<q}\frac{p+2}{2(p+1)} = 4.9809473396...$$

where p, q are primes.

What is the corresponding result for the set of all complex nonprincipal Dirichlet characters χ? Given an integer $m>2$, let

$$F'(m) = \sum_{\substack{\chi\ (\mathrm{mod}\, m),\\ \chi\neq\chi_0}} \left(\text{the least positive integer } n \text{ for which } \chi(n)\notin\{0,1\}\right),$$

noting that $F'(8) = F(8) + F(4) + F(-8) = 3+3+5 = 11$, for example [7], and $\sum_\chi 1 = \varphi(m)$ where φ is the Euler totient function. Martin & Pollack [8] proved that

$$\begin{aligned}\lim_{x\to\infty}\left(\sum_{2<m\le x}(\varphi(m)-1)\right)^{-1}\sum_{2<m\le x} F'(m) &= \sum_{k=1}^{\infty}\frac{p_k^2}{(p_1+1)(p_2+1)\cdots(p_k+1)}\\ &= 2.5350541804....\end{aligned}$$

What is the corresponding result for the set of all complex primitive Dirichlet characters χ? Given an integer $m>2$, let

$$F''(m) = \sum_{\substack{\chi\ (\mathrm{mod}\, m),\\ \chi\ \text{primitive}}} \left(\text{the least positive integer } n \text{ for which } \chi(n)\notin\{0,1\}\right),$$

noting that $F''(8) = F(8) + F(-8) = 8$ and $\sum_\chi 1 = \psi(m)$ where ψ is given by [5]

$$\psi(m) = \sum_{d|m}\varphi(d)\mu(m/d)$$

and μ is the Möbius mu function. We may use the fact that χ is primitive if and only if the Gauss sum [9]

$$\sum_{k=1}^{m}\chi(k)\exp\left(\frac{2\pi i k n}{m}\right) = 0 \quad \text{whenever } \gcd(n,m)>1.$$

It can be shown that [8]

$$\lim_{x\to\infty}\left(\sum_{2<m\le x}\psi(m)\right)^{-1}\sum_{2<m\le x}F''(m) = \sum_{q}\frac{q^4}{(q+1)^2(q-1)}\prod_{p<q}\frac{p^2-p-1}{(p+1)^2(p-1)}$$
$$= 2.1514351057....$$

1.38.3 *Variations*

Let $G(m)$ denote the least q such that the primes $\le q$ generate $\mathbb{Z}_m^*$, the multiplicative group modulo m. Also let $G'(m)$ denote the unique index k satisfying $p_k = q$. The latter function was first examined experimentally in [11]. For prime arguments, assuming that the Generalized Riemann Hypothesis is true, it follows that [3, 10]

$$\lim_{x\to\infty}\left(\sum_{2<p\le x}1\right)^{-1}\sum_{2<p\le x}G(p) = 3.9748384704...,$$

$$\lim_{x\to\infty}\left(\sum_{2<p\le x}1\right)^{-1}\sum_{2<p\le x}G'(p) = 2.2060828940...$$

but the infinite series expressions for these constants are too elaborate to present here. For arbitrary integer arguments, Bach [12, 13] proved that

$$\left(\sum_{2<m\le x}1\right)^{-1}\sum_{2<m\le x}G(m) \ge (1+o(1))\ln\ln x\ln\ln\ln x$$

as $x\to\infty$ and conjectured that the reverse inequality is valid too. The connection between $G(m)$ and least character nonresidues is [14]

$$G(m) = \max_{\substack{\chi\,(\mathrm{mod}\,m),\\ \chi\ne\chi_0}}\left(\text{the least positive integer } n \text{ for which } \chi(n)\notin\{0,1\}\right).$$

Previously we examined a sum $F'(m)$; here we examine a maximum.

Another interesting connection is that $f(p)$ is the least positive integer n for which $(n/p)\notin\{0,1\}$.

Let $h(m)$ be the least prime p for which $(m/p)\notin\{0,1\}$. Let $h'(m)$ be the least prime q for which $(m/q)\ne 1$. Since $p\ge q$, it is not surprising that [15]

$$C = \lim_{x\to\infty}\frac{1}{x}\sum_{m\le x}h(m) = \sum_{j=1}^{\infty}\frac{p_j-1}{2^j}\prod_{i=1}^{j-1}\left(1+\frac{1}{p_i}\right) = 5.6043245854...$$

is greater than

$$\lim_{x\to\infty}\frac{1}{x}\sum_{m\leq x}h'(m)=\sum_{j=1}^{\infty}\frac{p_j+1}{2^j}\prod_{i=1}^{j-1}\left(1-\frac{1}{p_i}\right)=2.5738775742....$$

The first (larger) average was examined by Elliott [16], but the second expression in p_i, p_j mistakenly appeared as the outcome.

Let $k(m)$ be the least prime p such that m is a quadratic nonresidue modulo p. It is easy to see that $k(m)=h(m)$ except when $h(m)=2$, in which case $k(m)>h(m)$. We have finally

$$\lim_{x\to\infty}\frac{1}{x}\sum_{m\leq x}k(m)=\sum_{j=2}^{\infty}\frac{p_j-1}{2^{j-1}}\prod_{i=2}^{j-1}\left(1+\frac{1}{p_i}\right)=\frac{4}{3}\left(C-\frac{1}{2}\right)=6.8057661139...$$

and wonder whether mean square analogs of these results are within reach.

Acknowledgments I thank Eric Bach for his extensive computations involving $G(p)$ and Greg Martin for theoretical help regarding $h(m)$, $h'(m)$ and $k(m)$.

[1] S. R. Finch, Meissel-Mertens constants: Quadratic residues, *first volume*, pp. 96–98.
[2] P. Erdős, Remarks on number theory. I (in Hungarian), *Mat. Lapok* 12 (1961) 10–17; MR0144869 (26 #2410).
[3] P. Pollack, The average least quadratic nonresidue modulo m and other variations on a theme of Erdős, *J. Number Theory* 132 (2012) 1185–1202; MR2899801.
[4] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A020649 and A053760.
[5] S. R. Finch, Quadratic Dirichlet L-series, *this volume*, §1.13.
[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A232929, A232930, A232931 and A232932.
[7] R. J. Mathar, Table of Dirichlet L-series and prime zeta modulo functions for small moduli, arXiv:1008.2547.
[8] G. Martin and P. Pollack, The average least character non-residue and further variations on a theme of Erdős, *J. London Math. Soc.* 87 (2013) 22–42; MR3022705.
[9] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976, pp. 138–139, 165–172; MR0434929 (55 #7892).
[10] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A232927 and A232928.
[11] H. Brown and H. Zassenhaus, Some empirical observations on primitive roots, *J. Number Theory* 3 (1971) 306–309; MR0288072 (44 #5270).
[12] E. Bach, Explicit bounds for primality testing and related problems, *Math. Comp.* 55 (1990) 355–380; MR1023756 (91m:11096).
[13] E. Bach and L. Huelsbergen, Statistical evidence for small generating sets, *Math. Comp.* 61 (1993) 69–82; MR1195432 (93k:11089).
[14] R, J. Burthe, Upper bounds for least witnesses and generating sets, *Acta Arith.* 80 (1997) 311–326; MR1450926 (98h:11117).
[15] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A092419 and A144294.
[16] P. D. T. A. Elliott, On the mean value of $f(p)$, *Proc. London Math. Soc.* 21 (1970) 28–96; MR0266881 (42 #1783).

1.39 Apollonian Circles with Integer Curvatures

Given four mutually tangent circles (one of them internally tangent to the other three), we can inscribe into each of the remaining curvilinear triangles a unique circle. Continuing iteratively in this manner, we obtain what is known as an **Apollonian circle packing**. If the initial four circles possess integer curvatures (reciprocal radii), then all of the circles in the packing possess integer curvatures. Some introductory accounts of this subject include [1–4]. We examine just two examples, the first starting with curvatures $\{-1, 2, 2, 3\}$ (Figure 1.21) and the second starting with curvatures $\{-11, 21, 24, 28\}$ (Figure 1.22). The outer circle is given negative curvature – indicating that the other circles are in its interior – and it is the unique circle with this property.

How are the integer curvatures obtained for each example? Define four 4×4 matrices

$$S_1 = \begin{pmatrix} -1 & 2 & 2 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad S_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & -1 & 2 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$S_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 2 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad S_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 2 & 2 & -1 \end{pmatrix}$$

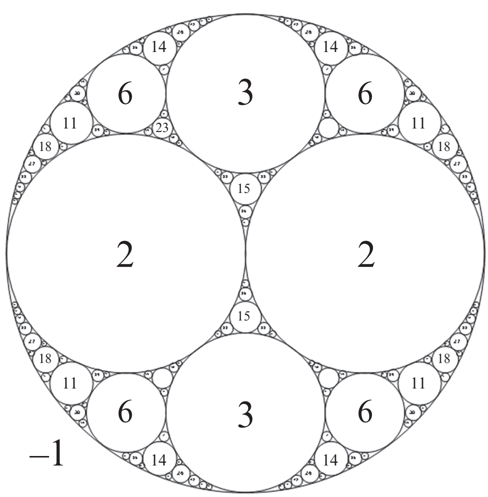

Figure 1.21 Bugeye circle packing.

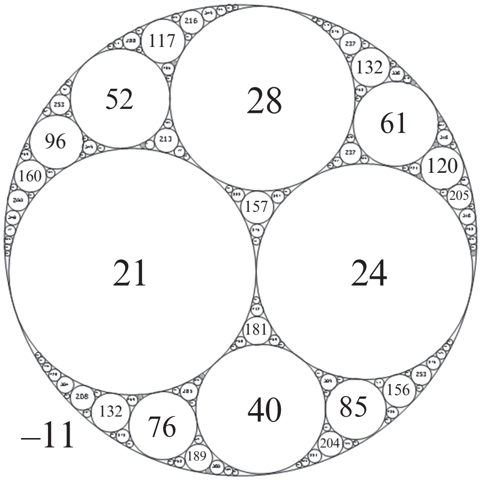

Figure 1.22 Nickel-dime-quarter packing.

and consider products $S_{j_1} S_{j_2} \cdots S_{j_n}$ with each $j_k \in \{1,2,3,4\}$ and $j_k \neq j_{k+1}$ for any k. The second generation of circles has curvatures

$$\begin{aligned}(S_4 w)_4 &= 3,\\ (S_3 w)_3 &= 6,\\ (S_2 w)_2 &= 6,\\ (S_1 w)_1 &= 15\end{aligned}$$

when $w = (-1, 2, 2, 3)$ (the bugeye circle packing) and

$$\begin{aligned}(S_4 w)_4 &= 40,\\ (S_3 w)_3 &= 52,\\ (S_2 w)_2 &= 61,\\ (S_1 w)_1 &= 157\end{aligned}$$

when $w = (-11, 21, 24, 28)$ (the nickel-dime-quarter packing). The third generation of circles has curvatures

$$\begin{array}{lll}(S_1 S_4 w)_1 = 15, & (S_2 S_4 w)_2 = 6, & (S_3 S_4 w)_3 = 6,\\ (S_1 S_3 w)_1 = 23, & (S_2 S_3 w)_2 = 14, & (S_4 S_3 w)_4 = 11,\\ (S_1 S_2 w)_1 = 23, & (S_3 S_2 w)_3 = 14, & (S_4 S_2 w)_4 = 11,\\ (S_2 S_1 w)_2 = 38, & (S_3 S_1 w)_3 = 38, & (S_4 S_1 w)_4 = 35\end{array}$$

when $w = (-1, 2, 2, 3)$ and

$$\begin{array}{lll}(S_1 S_4 w)_1 = 181, & (S_2 S_4 w)_2 = 85, & (S_3 S_4 w)_3 = 76,\\ (S_1 S_3 w)_1 = 213, & (S_2 S_3 w)_2 = 117, & (S_4 S_3 w)_4 = 96,\\ (S_1 S_2 w)_1 = 237, & (S_3 S_2 w)_3 = 132, & (S_4 S_2 w)_4 = 120,\\ (S_2 S_1 w)_2 = 397, & (S_3 S_1 w)_3 = 388, & (S_4 S_1 w)_4 = 376\end{array}$$

when $w = (-11, 21, 24, 28)$. The fourth generation of circles for the latter starts with $(S_4S_3S_4w)_4 = 132$, which is the first duplicate; the next two terms are $(S_4S_2S_4w)_4 = 156$ and $(S_3S_4S_3w)_3 = 160$. Arranging all the curvatures in order (with multiplicities), we have [5]

$$2, 2, 3, 3, 6, 6, 6, 6, 11, 11, 11, 11, 14, 14, 14, 14, 15, 15, 18, 18, 18, 18, 23,$$
$$23, 23, 23, 26, 26, 26, 26, 27, 27, 27, 27, 30, 30, 30, 30, 35, 35, 35, 35, 35,$$
$$35, 38, 38, 38, 38, 38, 38, 38, 38, 39, 39, 39, 39, 42, 42, 42, 42, 47, 47, 47,$$
$$47, 50, 50, 50, 50, 51, 51, 51, 51, 54, 54, 54, 54, 59, 59, 59, 59, 59, 59, 59, 59, \ldots$$

when $w = (-1, 2, 2, 3)$ and

$$21, 24, 28, 40, 52, 61, 76, 85, 96, 117, 120, 132, 132, 156, 157, 160, 181, 189,$$
$$204, 205, 208, 213, 216, 237, 237, 244, 253, 253, 285, 288, 304, 309, 316, 316, \ldots$$

when $w = (-11, 21, 24, 28)$. A theorem due to Kontorovich & Oh [6] provides the growth rate for these sequences:

$$\nu(x) \sim c \cdot x^{\delta}$$

as $x \to \infty$, where $\nu(x)$ is the number of circles in the packing with curvature less than x, the exponent $\delta = 1.3056867280...$ has been discussed [7, 8], and the coefficients

$$c = \begin{cases} 0.402... & \text{if } w = (-1, 2, 2, 3), \\ 0.0176... & \text{if } w = (-11, 21, 24, 28) \end{cases}$$

were estimated by Fuchs & Sanden [9]. (The values 0.201... in [2] and 0.0458... in [3] are apparently mistaken.) Expressions for c exist [10–12], but are not suitably practical to allow numerical calculations.

Rather than counting all circles with curvature $< x$, we might instead restrict attention to the n^{th} generation (which has $4 \cdot 3^{n-2}$ members) and determine the average curvature as a function of n. Most circles born at a large generation n possess curvature $\sim \exp(\gamma\, n)$, where $\gamma = 0.9149...$ is the Lyapunov exponent associated with random products $S_{j_1} S_{j_2} \cdots S_{j_n}$. The logarithm of curvature, divided by n, is asymptotically normal with mean γ and variance $\sim \alpha/n$, where $\alpha = 0.065....$ This alternative approach would be worth further study [3, 13], but we must stop here.

1.39.1 *Kissing Primes*

The primes appearing in the preceding sequences (curvatures with multiplicities) are [5]

$$2, 2, 3, 3, 11, 11, 11, 11, 23, 23, 23, 23, 47, 47, 47, 47, 59, 59, 59, 59, 59, 59, 59, 59,$$
$$71, 71, 71, 71, 83, 83, 83, 83, 83, 83, 83, 83, 107, 107, 107, 107, 107, 107, 107, 107,$$
$$131, 131, 131, 131, 131, 131, 131, 131, 131, 131, 131, 131,$$
$$167, 167, 167, 167, 167, 167, 167, 167, 167, 167, 167, 167, ...$$

when $w = (-1, 2, 2, 3)$ and

$$61, 157, 181, 349, 373, 397, 421, 541, 661, 709, 733, 829, 853, 877, ...$$

when $w = (-11, 21, 24, 28)$. Each term corresponds to a circle C of prime curvature $a(C)$. Define a weighted prime count

$$\psi(x) = \sum_{\substack{a(C)<x,\\ a(C)\text{ prime}}} \ln(a(C))$$

then it is conjectured that

$$\psi(x) \sim G \cdot \nu(x)$$

as $x \to \infty$, where the coefficient $G = 0.9159655941...$ is Catalan's constant [14]. It is remarkable that the coefficient is independent of the packing.

Assume that an unordered pair of tangent circles C, C' are both of prime curvature p, p'. The two primes are said to be **kissing primes** (for the packing under consideration). We have pairs (with multiplicities)

$$\begin{gathered}(2,2), (2,3), (2,3), (2,3), (2,3), (2,11), (2,11), (2,11), (2,11), (2,23),\\ (2,23), (2,23), (2,23), (3,23), (3,23), (3,23), (3,23), (3,47), (3,47),\\ (3,47), (3,47), (2,59), (2,59), (2,59), (2,59), ...\end{gathered}$$

when $w = (-1, 2, 2, 3)$ and

$$(157, 397), (61, 421), (61, 1069), (157, 1093), (181, 1213), ...$$

when $w = (-11, 21, 24, 28)$. Define a weighted prime count

$$\psi^{(2)}(x) = \sum_{\substack{a(C),a(C')<x,\\ C,C' \text{ tangent},\\ a(C),a(C')\text{ prime}}} \ln(a(C)) \cdot \ln(a(C'))$$

then it is conjectured that

$$\psi^{(2)}(x) \sim H \cdot \nu(x)$$

as $x \to \infty$, where the coefficient

$$H = G^2 \cdot 2 \prod_{p\equiv 3 \bmod 4} \left(1 - \frac{2}{p(p-1)^2}\right) = G^2(1.6493376890...) = 3(0.4612609086...)$$

is again independent of the packing. These estimates improve upon the values 1.646... in [9] and 0.460... in [3].

The number of circles of prime curvature $< x$ is asymptotically $\psi(x)/\ln(x)$, hence $\sim G \cdot \nu(x)/\ln(x)$ by the first-order conjecture. For the number of kissing prime circles both with curvatures $< x$, the relationship with $\psi^{(2)}(x)/\ln(x)^2$ is less clear. This would be good to clarify someday. Interestingly, Catalan's constant also appears in [1], although in an unrelated manner.

Recent progress on this subject is described in [15–20].

[1] R. L. Graham, J. C. Lagarias, C. L. Mallows, A. R. Wilks and C. H. Yan, Apollonian circle packings: number theory, *J. Number Theory* 100 (2003) 1–45; arXiv:math/0009113; MR1971245 (2004d:11055).

[2] D. Mackenzie, A tisket, a tasket, an Apollonian gasket, *American Scientist*, v. 98 (2010) n. 1, 10–14.

[3] P. Sarnak, Integral Apollonian packings, *Amer. Math. Monthly* 118 (2011) 291–306; MR2800340 (2012e:52047).

[4] E. Pegg, The Circles of Descartes, Wolfram Demonstrations project (2008).

[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A042944, A042946, A189226 and A189227.

[6] A. Kontorovich and H. Oh, Apollonian circle packings and closed horospheres on hyperbolic 3-manifolds, *J. Amer. Math. Soc.* 24 (2011) 603–648; arXiv:0811.2236; MR2784325.

[7] S. R. Finch, Apollonian packing constant, *first volume*, pp. 537–539.

[8] Z.-Q. Bai and S. R. Finch, Precise calculation of Hausdorff dimension of Apollonian gasket, Fractals 26 (2018) 1850050.

[9] E. Fuchs and K. Sanden, Some experiments with integral Apollonian circle packings, *Experim. Math.* 20 (2011) 380–399; arXiv:1001.1406; MR2859897 (2012j:52039).

[10] H. Oh and N. Shah, The asymptotic distribution of circles in the orbits of Kleinian groups, *Invent. Math.* 187 (2012) 1–35; arXiv:1004.2130; MR2874933 (2012k:37011).

[11] M. Lee and H. Oh, Effective circle count for Apollonian packings and closed horospheres, *Geom. Funct. Anal.* 23 (2013) 580–621; arXiv:1202.1067; MR3053757.

[12] I. Vinogradov, Effective bisector estimate with application to Apollonian circle packings, *Internat. Math. Res. Notices* (2014) 3217–3262; arXiv:1204.5498; MR3217660.

[13] E. Fuchs, *Arithmetic Properties of Apollonian Circle Packings*, Ph.D. thesis, Princeton Univ., 2010.

[14] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.

[15] J. Bourgain and E. Fuchs, A proof of the positive density conjecture for integer Apollonian circle packings, *J. Amer. Math. Soc.* 24 (2011) 945–967; arXiv:1001.3894; MR2813334 (2012d:11072).

[16] J. Bourgain, Integral Apollonian circle packings and prime curvatures, *J. Anal. Math.* 118 (2012) 221–249; arXiv:1105.5127; MR2993027.

[17] E. Fuchs, Counting problems in Apollonian packings, *Bull. Amer. Math. Soc.* 50 (2013) 229–266; MR3020827.

[18] A. Kontorovich, From Apollonius to Zaremba: local-global phenomena in thin orbits, *Bull. Amer. Math. Soc.* 50 (2013) 187–228; arXiv:1208.5460; MR3020826.

[19] H. Oh, Apollonian circle packings: dynamics and number theory, *Japanese J. Math.* 9 (2014) 69–97; arXiv:1312.1383; MR3173439.

[20] J. Bourgain and A. Kontorovich, On the local-global conjecture for integral Apollonian gaskets, *Invent. Math.* 196 (2014) 589–650; arXiv:1205.4416; MR3211042.

1.40 Molteni's Composition Constant

This essay continues where we left off in [1]: the number of (unordered) partitions of 2^{k-1} as a sum of k powers of 2 is well-understood [2–6]. What can be said about the number $w(k)$ of (ordered) compositions of 2^{k-1} as a sum of k powers of 2? Clearly $w(1) = w(2) = 1$; $w(3) = 3$ since there are three ways to sort $\{1, 1, 2\}$ and $w(4) = 13$ since there are twelve ways to sort $\{1, 1, 2, 4\}$ plus $8 = 2 + 2 + 2 + 2$. A few more terms of $\{w(k)\}$ appear in [7, 8] but a pattern is far from clear.

The following doubly-indexed recursive formula [9]

$$m_{k,\ell} = \begin{cases} 0 & \text{if } \ell \geq k, \\ 1 & \text{if } k > 1 \text{ and } \ell = k - 1, \\ \displaystyle\sum_{j=1}^{2\ell} \binom{k + \ell - 1}{2\ell - j} m_{k-\ell,j} & \text{if } 1 \leq \ell < k - 1, \end{cases}$$

coupled with $w_k = m_{k,1}$, $k > 1$, makes efficient calculation of many more terms possible. It further allowed Molteni [10] to deduce the asymptotic behavior of $\{w(k)\}$:

$$\lim_{k\to\infty} \left(\frac{w(k)}{k!}\right)^{1/k} = 1.1926743412...$$

– a remarkable achievement! – but an exact formula for this constant seems to be unavailable. The same constant appears in a more general setting when 2^{k-1} is replaced by, for instance, a sum of two distinct powers of 2. As an example, $w'(3) = 6$ since $10 = 2 + 8$, there are three ways to sort $\{1, 1, 8\}$ plus three ways to sort $\{2, 4, 4\}$, and such a portfolio is maximal. Replacing w by w' in the limiting expression does not change the constant.

1.40.1 *Euler Binary Partitions*

Given $d \geq 2$ and $n \geq 0$, let $b_d(n)$ denote the number of integer sequences x_1, x_2, x_3, … satisfying $0 \leq x_i \leq d - 1$ for all i for which $n = \sum_{i=0}^{\infty} x_i 2^i$. Clearly $b_2(n) = 1$ for all n, $\{b_3(n)\}$ is related to Stern's sequence [11], and $b_4(n) = \lfloor n/2 \rfloor + 1$ for all n. Define

$$\kappa_d = \liminf_{n\to\infty} \frac{\ln(b_d(n))}{\ln(n)}, \quad \lambda_d = \limsup_{n\to\infty} \frac{\ln(b_d(n))}{\ln(n)}.$$

The most interesting asymptotics occur for odd d and we list several results here [12–16]:

$$2^{\kappa_3} = 1, \quad 2^{\lambda_3} = \varphi = \left(1 + \sqrt{5}\right)/2 = 1.6180339887...;$$

$$2^{\kappa_5} = 1 + \sqrt{2} = 2.4142135623..., \quad 2^{\lambda_5} = 2.5386157635...$$

has minimal polynomial $z^4 - 2z^3 - 2z^2 + 2z - 1$;

$$2^{\kappa_7} = 3.4918910516..., \quad 2^{\lambda_7} = 3.5115471416...$$

have minimal polynomials $z^5 - z^4 - 7z^3 - 5z^2 - 3z - 1$ and $z^3 - 4z^2 + 2z - 1$, respectively; and

$$2^{\kappa_9} = 4.4944928370..., \quad 2^{\lambda_9} = 4.5030994219...$$

have minimal polynomials $z^3 - 4z^2 - 2z - 1$ and $z^8 - 3z^7 - 9z^6 + 9z^5 + 5z^4 - z^3 - z^2 - z + 1$, respectively.

1.40.2 *Joint Spectral Radius*

The joint spectral radius [17] of two real 2×2 matrices A, B is the maximum possible exponential rate of growth of long products of As and Bs. The set $\{A, B\}$ is said to have the finiteness property if there exists a periodic product that attains this maximal rate of growth. At one point, it was believed that every set $\{A, B\}$ satisfies the finiteness property. This was eventually disproved; the first explicit counterexample was given in [18]. It takes the form

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = c\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

where the constant c requires elaboration. Define

$$e_{n+1} = e_n e_{n-1} - e_{n-2}, \quad e_0 = 1, \quad e_1 = 2, \quad e_2 = 2$$

and

$$f_{n+1} = f_n + f_{n-1}, \quad f_0 = 0, \quad f_1 = 1$$

(the latter is the Fibonacci sequence). It follows that

$$\begin{aligned} c &= \lim_{n\to\infty}\left(\frac{e_n^{f_{n+1}}}{e_{n+1}^{f_n}}\right)^{(-1)^n} = \prod_{n=1}^{\infty}\left(1 - \frac{e_{n-1}}{e_{n+1}e_n}\right)^{(-1)^n f_{n+1}} \\ &= 0.7493265463... \end{aligned}$$

converges unconditionally. No uniqueness claims have been made about c; we are simply attracted by its intricate construction. The authors of [18] wondered whether c is irrational, tying it to the Fibonacci substitution $0 \to 01$, $1 \to 0$ [19] and to the quantity $1/\varphi^2 = (3 - \sqrt{5})/2$. They conjectured that $\tilde{c}$ is irrational, where $\tilde{c}$ (unspecified but distinct from c) is tied to the substitution $0 \to 001$, $1 \to 0$ and to the quantity $1 - 1/\sqrt{2}$. It would be good to understand more about $\tilde{c}$ someday.

[1] S. R. Finch, Kalmár's composition constant, *first volume*, pp. 292–295.

[2] D. W. Boyd, The asymptotic number of solutions of a Diophantine equation from coding theory, *J. Combin. Theory Ser. A* 18 (1975) 210–215; MR0360437 (50 #12887).

[3] J. Komlós, W. Moser and T. Nemetz, On the asymptotic number of prefix codes, *Mitteilungen aus dem Math. Seminar Giessen*, Heft 165, *Coxeter–Festschrift*, Teil III, 1984, pp. 35–48.
[4] P. Flajolet and H. Prodinger, Level number sequences for trees, *Discrete Math.* 65 (1987) 149–156; MR0893076 (88e:05030).
[5] S. Lehr, J. Shallit and J. Tromp, On the vector space of the automatic reals, *Theoret. Comput. Sci.* 163 (1996) 193–210; MR1407021 (97i:03037).
[6] C. Elsholtz, C. Heuberger and H. Prodinger, The number of Huffman codes, compact trees, and sums of unit fractions, *IEEE Trans. Inform. Theory* 59 (2013) 1065–1075; arXiv:1108.5964; MR3015716.
[7] G. Molteni, Cancellation in a short exponential sum, *J. Number Theory* 130 (2010) 2011–2027; MR2653211 (2011e:11135).
[8] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002572, A007178, A007657, and A022405.
[9] A. Giorgilli and G. Molteni, Representation of a 2-power as sum of k 2-powers: a recursive formula, *J. Number Theory* 133 (2013) 1251–1261; MR3003997.
[10] G. Molteni, Representation of a 2-power as sum of k 2-powers: the asymptotic behavior, *Internat. J. Number Theory* 8 (2012) 1923–1963; MR2978848.
[11] S. R. Finch, Lyapunov exponents IV, *this volume*, §4.16.
[12] B. Reznick, Some binary partition functions, *Analytic Number Theory*, Proc. 1989 Allerton Park conf., ed. B.C. Berndt. H.G. Diamond, H. Halberstam, and A. Hildebrand, Birkhäuser, 1990, pp. 451–477; MR1084197 (91k:11092).
[13] V. Yu. Protasov, Asymptotics of the partition function (in Russian), *Mat. Sbornik*, v. 191 (2000) n. 3, 65–98; Engl. trans. in *Sbornik Math.* 191 (2000) 381–414; MR1773255 (2001h:11134).
[14] V. Yu. Protasov, On the problem of the asymptotics of the partition function (in Russian), *Mat. Zametki*, v. 76 (2004) n. 1, 151–156; Engl. transl. in *Math. Notes* 76 (2004) 144–149; MR2099853 (2006a:11132).
[15] V. Yu. Protasov, R. M. Jungers and V. D. Blondel, Joint spectral characteristics of matrices: a conic programming approach, *SIAM J. Matrix Anal. Appl.* 31 (2009/10) 2146–2162; MR2678961 (2011g:15020).
[16] N. Guglielmi and V. Yu. Protasov, Exact computation of joint spectral characteristics of linear operators, *Found. Comput. Math.* 13 (2013) 37–97; MR3009529.
[17] R. Jungers, *The Joint Spectral Radius. Theory and Applications*, Lect. Notes in Control and Info. Sci. 385, Springer-Verlag, 2009; MR2507938 (2011c:15001).
[18] K. G. Hare, I. D. Morris, N. Sidorov and J. Theys, An explicit counterexample to the Lagarias-Wang finiteness conjecture, *Adv. Math.* 226 (2011) 4667–4701; MR2775881 (2012b:15016).
[19] S. R. Finch, Substitution dynamics, *this volume*, §4.24.

1.41 Boolean Decision Functions

Let $f\colon \{0,1\}^n \to \{0,1\}$ be the Boolean function that decides whether a given $(n+1)$-bit odd integer is square-free. More precisely,

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } 2\xi + 1 \text{ is square-free,} \\ 0 & \text{otherwise} \end{cases}$$

where the string $x_1x_2, \ldots x_n$ is the integer ξ written in binary (with leading zeroes added as necessary). Let x denote the vector $(x_1, x_2, \ldots, x_n)$. There are many ways of characterizing the computational complexity of f; we focus on a single combinatorial method related to what is called the *average sensitivity* of f. The **influence** of x_i on f, denoted by $I_i(f)$, is the probability that flipping the i^{th} component of the input vector, selected at random from $\{0,1\}^n$, will flip the output. That is,

$$I_i(f) = 2^{-n} \sum_{x \in \{0,1\}^n} \left| f(x) - f\left(x^{(i)}\right) \right|$$

where $x^{(i)} = (x_1, x_2, \ldots, x_i + 1, \ldots, x_n)$ modulo 2. Bernasconi, Damm & Shparlinski [1, 2] proved that

$$I_i(f) = 2\gamma_{\text{int}} + o(n)$$

as $n \to \infty$, where

$$\gamma_{\text{int}} = \frac{8}{\pi^2} - 2\prod_p \left(1 - \frac{2}{p^2}\right) = 0.1653012713... = \frac{0.3306025426...}{2}.$$

In words, an odd integer changes from square-free to square-full or vice versa with probability $\approx 33\%$ if one of its bits is flipped. The infinite product is familiar – called the Feller–Tornier constant in [3] – and its appearance here is quite interesting.

We turn attention from integers to polynomials with coefficients in the finite field $\mathbb{Z}_2$. Let $g: \{0,1\}^n \to \{0,1\}$ decide whether a given binary polynomial with constant coefficient unity

$$\eta(x) = y_n x^n + y_{n-1} x^{n-1} + \cdots + y_1 x + 1$$

is square-free. More precisely,

$$g(y_1, y_2, \ldots, y_n) = \begin{cases} 1 & \text{if } \eta(x) \text{ is square-free,} \\ 0 & \text{otherwise} \end{cases}$$

and we again abbreviate the vector as y. The influence $I_i(g)$ of y_i on g is defined similarly. Clearly the polynomial corresponding to the vector $y^{(i)}$ is $\eta(x) + x^i$ modulo 2. Allender, Bernasconi, Damm, von zur Gathen, Saks & Shparlinski [4] proved that

$$I_i(g) = 2\gamma_{\text{poly}} + O\left(2^{-n/4}\right)$$

as $n \to \infty$, where

$$\gamma_{\text{poly}} = \frac{2}{3} - 2\prod_{k=1}^{\infty} \left(1 - \frac{1}{2^{2k-1}}\right)^{a_k} = 0.2735795624... = \frac{0.5471591248...}{2}.$$

The sequence $\{a_k\}_{k=1}^{\infty} = \{2, 1, 2, 3, 6, 9, 18, 30, \ldots\}$ counts all irreducible polynomials over $\mathbb{Z}_2$ of degree k and satisfies [5]

$$2^k = \sum_{d \mid k} d\, a_k;$$

equivalently,

$$a_k = \frac{1}{k} \sum_{d \mid k} \mu\left(\frac{k}{d}\right) 2^d$$

where μ is the Möbius mu function [6]. Note that the error term is tighter for $I_i(g)$ than that for $I_i(f)$.

A fascinating unanswered question arises if we replace square-freeness by primality (for odd integers) and irreducibility (for binary polynomials). What are the influence I_i asymptotics in this new scenario? Formulas analogous to the preceding would be good to see someday.

With regard to integers, a positive proportion of primes become composite when *any* one of their bits is changed [7–9]. As a consequence, it is not possible to establish whether an arbitrary integer is prime without examining all of its bits. With regard to polynomials, it is curious that [10]

$$\prod_{k=1}^{\infty} \left(1 - \frac{1}{2^{2k}}\right)^{a_k} = \frac{1}{2}$$

is trivial while a slight modification yields the unrecognizable constant γ_{poly}.

[1] A. Bernasconi, C. Damm and I. Shparlinski, On the average sensitivity of testing square-free numbers, *Computing and Combinatorics (COCOON)*, Proc. 1999 Tokyo conf., ed. T. Asano, H. Imai, D. T. Lee, S.-I. Nakano and T. Tokuyama, Lect. Notes in Comp. Sci. 1627, Springer-Verlag, 1999, pp. 291–299; MR1730345 (2000i:11192).
[2] A. Bernasconi, C. Damm and I. Shparlinski, The average sensitivity of square-freeness, *Comput. Complexity* 9 (2000) 39–51; MR1791089 (2001k:11179).
[3] S. R. Finch, Meissel-Mertens constants, *first volume*, pp. 94–98.
[4] E. Allender, A. Bernasconi, C. Damm, J. von zur Gathen, M. Saks and I. Shparlinski, Complexity of some arithmetic problems for binary polynomials, *Comput. Complexity* 12 (2003) 23–47; MR2054893 (2005b:68119).
[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001037, A059966, and A060477.
[6] S. R. Finch, Artin's constant, *first volume*, pp. 104–109.
[7] F. Cohen and J. L. Selfridge, Not every number is the sum or difference of two prime powers, *Math. Comp.* 29 (1975) 79–81; MR0376583 (51 #12758).
[8] Z.-W. Sun, On integers not of the form $\pm p^a \pm q^b$, *Proc. Amer. Math. Soc.* 128 (2000) 997–1002; MR1695111 (2000i:11157).
[9] T. Tao, A remark on primality testing and decimal expansions, *J. Aust. Math. Soc.* 91 (2011) 405–413; arXiv:0802.3361; MR2900615.
[10] E. R. Berlekamp, *Algebraic Coding Theory*, rev. ed., Aegean Park Press, 1984, pp. 70–86; MR0238597 (38 #6873).

1.42 Map Asymptotics Constant

A **map** on a compact surface S without boundary is an embedding of a graph G into S such that all components of $S - G$ are simply connected [1]. These components are thus homeomorphic to open disks and are called **faces**. The graph G is allowed to have both loops and multiple parallel edges (unlike those in [2]). A map is **rooted** when an edge, a direction along that edge, and a side of the edge, are distinguished. The edge is called the **root edge**, and the face on the distinguished side is the **root face**. Two rooted maps are **equivalent** if there is a homeomorphism between the underlying surfaces that preserves all graph incidences and rootedness.

In the case when S is orientable, two rooted maps are equivalent if and only they are related by an orientation-preserving homeomorphism that (merely) preserves all graph incidences. Such thinking does not apply, of course, when S is non-orientable. For orientable surfaces, the **genus** g is 0 for the sphere, 1 for the torus, 2 for the connected sum of two tori, and so forth. For non-orientable surfaces, the **type** h is $1/2$ for the projective plane, 1 for the Klein bottle, $3/2$ for the connected sum of three projective planes, and so forth.

The requirement that faces be simply connected implies that the graph G itself must be connected [3]. Proof: if G were to possess two components, then a curve drawn around one of the components could not be contracted to a point (because the other component would present an obstacle), which is a contradiction. The converse is true if the surface S is a sphere, but is false if S is a torus. Reason: consider the figure-eight graph G consisting of one vertex and two edges (orthogonal loops that together generate the torus). While $S - G$ is simply connected, this is not true for any proper subgraph of G.

Let $T_g(n)$ denote the number of rooted maps with n edges on an orientable surface of genus g. Let $P_h(n)$ denote the number of rooted maps with n edges on a non-orientable surface of type h. (T stands for "torus" and P stands for "projective plane".) It is known that $T_0(n)$ is the coefficient of x^n in the Maclaurin series expansion [1, 4, 5]

$$\begin{aligned}\frac{4(1+2r)}{3(1+r)^2} &= 1 + 2x + 9x^2 + 54x^3 + 378x^4 + 2916x^5 + 24057x^6 \\ &\quad + 208494x^7 + 1876446x^8 + 17399772x^9 + 165297834x^{10} + \cdots,\end{aligned}$$

$T_1(n)$ is the coefficient of x^n in the expansion [6, 7]

$$\begin{aligned}\frac{(-1+r)^2}{12r^2(2+r)} &= x^2 + 20x^3 + 307x^4 + 4280x^5 + 56914x^6 \\ &\quad + 736568x^7 + 9370183x^8 + 117822512x^9 + 1469283166x^{10} + \cdots,\end{aligned}$$

$P_{1/2}(n)$ is the coefficient of x^n in the expansion [1, 6]

$$\begin{aligned}\frac{-q}{(-1+r)(1+r)} &= x + 10x^2 + 98x^3 + 982x^4 + 10062x^5 + 105024x^6 \\ &\quad + 1112757x^7 + 11934910x^8 + 129307100x^9 \\ &\quad + 1412855500x^{10} + \cdots\end{aligned}$$

and $P_1(n)$ is the coefficient of x^n in the expansion [8–10]

$$\begin{aligned}\frac{(1+r)q}{2r^2(2+r)} &= 4x^2 + 84x^3 + 1340x^4 + 19280x^5 + 263284x^6 \\ &\quad + 3486224x^7 + 45247084x^8 + 579150012x^9 \\ &\quad + 7338291224x^{10} + \cdots\end{aligned}$$

where $r = \sqrt{1-12x}$ and $q = 2 + 4r - 2\sqrt{3}\sqrt{r(2+r)}$ throughout. Moreover [11],

$$T_g(n) \sim t_g n^{5(g-1)/2} 12^n, \quad P_h(n) \sim p_h n^{5(h-1)/2} 12^n$$

as $n \to \infty$, where t_g is the **orientable map asymptotics constant**:

$$t_0 = \frac{2}{\sqrt{\pi}}, \quad t_1 = \frac{1}{24}, \quad t_2 = \frac{7}{4320\sqrt{\pi}}, \quad t_3 = \frac{245}{15925248}, \quad t_4 = \frac{37079}{96074035200\sqrt{\pi}}$$

and p_h is the **non-orientable map asymptotics constant**:

$$p_{1/2} = \frac{\sqrt{3}}{2\pi}\Gamma(1/4) = -\frac{2\sqrt{6}}{\Gamma(-1/4)}, \quad p_1 = \frac{1}{2}, \quad p_{3/2} = \frac{\sqrt{6}}{3\Gamma(1/4)} = \frac{5}{8\sqrt{6}\Gamma(9/4)}.$$

Since the status of t_g is quite different from the status of p_h, we shall treat them separately.

For many years, the values of t_g for $g > 2$ were unknown, owing to difficulties in their formulation. Impressive progress has been made recently. Define a sequence

$$u_0 = 1, \quad u_n = \frac{25(n-1)^2 - 1}{48} u_{n-1} - \frac{1}{2}\sum_{k=1}^{n-1} u_k u_{n-k} \quad \text{for } n \geq 1,$$

then provably

$$t_g = -\frac{1}{2^{g-2}\Gamma\left((5g-1)/2\right)} u_g$$

for all integers $g \geq 0$. The formal power series $u(z) = \sum_{n=0}^{\infty} u_n z^{-(5n-1)/2}$ satisfies the Painlevé I differential equation

$$u''(z) = 6u(z)^2 - 6z$$

which makes possible the following asymptotics:

$$t_g \sim \frac{40\sin(\pi/5)K}{\sqrt{2\pi}} \left(\frac{1440g}{e}\right)^{-g/2}$$

as $g \to \infty$ and

$$K = \sqrt{\frac{3}{5}} \frac{\Gamma(1/5)\Gamma(4/5)}{4\pi^2} = 0.1048689877....$$

We explain further: Bender, Gao & Richmond [12] discovered the preceding approximation for t_g but with only a rough numerical estimate 0.1034 for K. The connection with Painlevé I, streamlined u_n recursion and exact K expression are due to Garoufalidis, Lê & Mariño [13]. A (somewhat different) full asymptotic series is also possible. We give the first term only:

$$u_n \sim -\frac{1}{2\pi} \frac{3^{1/4}}{\sqrt{\pi}} \left(\frac{8\sqrt{3}}{5} \right)^{-2n+\frac{1}{2}} \Gamma\left(2n - \frac{1}{2}\right)$$

as $n \to \infty$, quoting [14]. This is reminiscent of other quadratic recurrence studies [15, 16].

Likewise, the path to understanding p_h for $h > 2$ is fraught with peril. Define a sequence

$$v_0 = -\sqrt{3}, \quad v_n = \frac{1}{2\sqrt{3}} \left(-3u_{n/2} + \frac{5n-6}{2} v_{n-1} + \sum_{k=1}^{n-1} v_k v_{n-k} \right) \quad \text{for } n \geq 1$$

(the dependence of v_n on $u_{n/2}$ from before is striking: if n is odd, let $u_{n/2} = 0$). Conjecturally, we have [14]

$$p_h = \frac{1}{2^{h-2}\Gamma\left((5h-3)/2\right)} v_{2h-1}$$

for all integers/half-integers $h \geq 1/2$. Evidence for this equality comes from quantum physics. As consequences,

$$p_2 = \frac{5}{36\sqrt{\pi}}, \quad p_{5/2} = \frac{1033}{1024\sqrt{6}\Gamma(19/4)}, \quad p_3 = \frac{3149}{442368}, \quad p_{7/2} = \frac{1599895}{294912\sqrt{6}\Gamma(29/4)}.$$

The formal power series $v(z) = \sum_{n=0}^{\infty} v_n z^{-(5n-1)/4}$ satisfies the differential equation

$$2v'(z) = v(z)^2 - 3u(z)$$

and a full asymptotic series is again possible. We give the first term only:

$$v_n \sim \frac{C}{2\pi} \left(\frac{4\sqrt{3}}{5} \right)^{-n} \Gamma(n)$$

as $n \to \infty$, where the Stokes constant C is conjectured to be $\sqrt{6}$. See [17, 18] for a bivariate analog of the preceding theory.

[1] E. R. Canfield, Calculating the number of rooted maps on a surface, *Proc. 21st Southeastern Conf. on Combinatorics, Graph Theory, and Computing*, Boca Raton, 1990, ed. F. Hoffman, R. C. Mullin, R. G. Stanton and K. Brooks Reid, Congr. Numer. 76, Utilitas Math., 1990, pp. 21–34; MR1140414 (92i:05109).

[2] S. R. Finch, Planar graph growth constants, *this volume*, §1.9.
[3] T. R. Walsh, Counting maps on doughnuts, *Theoret. Comput. Sci.* 502 (2013) 4–15; MR3101687.
[4] W. T. Tutte, A census of planar maps, *Canad. J. Math.* 15 (1963) 249–271; MR0146823 (26 #4343).
[5] W. T. Tutte, The enumerative theory of planar maps, *A Survey of Combinatorial Theory*, Proc. 1971 Colorado State Univ. conf., ed. J. N. Srivastava, North-Holland, 1973, pp. 437–448; MR0364004 (51 #259).
[6] E. A. Bender, E. R. Canfield and R. W. Robinson, The enumeration of maps on the torus and the projective plane, *Canad. Math. Bull.* 31 (1988) 257–271; MR0956356 (89g:05056).
[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000168, A006300, A006301, and A104742.
[8] D. Arquès and J.-F. Béraud, Énumération des cartes pointées sur la bouteille de Klein, *RAIRO Inform. Théor. Appl.* 31 (1997) 385–409; MR1605788 (99a:05063).
[9] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007137, A006344, A118451, and A214334.
[10] D. Arquès and A. Giorgetti, Counting rooted maps on a surface, *Theoret. Comput. Sci.* 234 (2000) 255–272; MR1745078 (2001f:05078).
[11] E. A. Bender and E. R. Canfield, The asymptotic number of rooted maps on a surface, *J. Combin. Theory Ser. A* 43 (1986) 244–257; MR0867650 (88a:05080).
[12] E. A. Bender, Z. Gao and L. B. Richmond, The map asymptotics constant t_g, *Elec. J. Combin.* 15 (2008) R51; MR2398843 (2009a:05096).
[13] S. Garoufalidis, T. T. Q. Lê and M. Mariño, Analyticity of the free energy of a closed 3-manifold, *SIGMA Symmetry Integrability Geom. Methods Appl.* 4 (2008) P080, MR2470516 (2010e:57015).
[14] S. Garoufalidis and M. Mariño, Universality and asymptotics of graph counting problems in non-orientable surfaces, *J. Combin. Theory Ser. A* 117 (2010) 715–740; MR2645187 (2011i:05098).
[15] S. R. Finch, Quadratic recurrence constants, *first volume*, pp. 443–448.
[16] S. R. Finch, Otter's tree enumeration constants, *first volume*, Cambridge Univ. Press, 2003, pp. 295–316.
[17] E. A. Bender, E. R. Canfield and L. B. Richmond, The asymptotic number of rooted maps on a surface. II, Enumeration by vertices and faces, *J. Combin. Theory Ser. A* 63 (1993) 318–329; MR1223687 (95c:05068).
[18] Z. Gao, A formula for the bivariate map asymptotics constants in terms of the univariate map asymptotics constants, *Elec. J. Combin.* 17 (2010) R155; MR2745708 (2012c:05089).

1.43 Injections, Surjections and More

Let $I_{m,n}$ denote the set of all injections $\{1,\ldots,m\} \to \{1,\ldots,n\}$ where $m \leq n$. An element of $I_{m,n}$ can be thought of as a permutation on n symbols taken m at a time. We define $I_{0,n}$ to possess one element (the empty permutation) for convenience; therefore [1–3]

$$\# I_{m,n} = \frac{n!}{(n-m)!}$$

and

$$\#\bigcup_{0\le m\le n} I_{m,n}=\sum_{k=0}^{n}\frac{n!}{k!}=\begin{cases}\lfloor n!e\rfloor & \text{if } n>0,\\ 1 & \text{if } n=0\end{cases}$$

where e is the natural logarithmic base [4]. In counting all injections, we treat extensions as distinct; for example, the function $f:\{1,2\}\to\{1,2\}$ with $f(x)=x$ is not the same as the function $g:\{1,2\}\to\{1,2,3\}$ with $g(x)=x$, nor is it the same as the function $h:\{1,2,3\}\to\{1,2,3\}$ with $h(x)=x$.

Let $J_{n,m}$ denote the set of all surjections $\{1,\ldots,n\}\to\{1,\ldots,m\}$ where $n\ge m$. An element of $J_{n,m}$ can be thought of as an ordered m-tuple consisting of preimage blocks (m disjoint nonempty sets that cover n symbols). We define $J_{0,0}$ to possess one element (the empty tuple) for convenience; therefore [5–7]

$$\# J_{n,m}=\sum_{j=0}^{m}(-1)^j\binom{m}{j}(m-j)^n=m!S_{n,m}$$

and

$$\#\bigcup_{0\le m\le n} J_{n,m}=\frac{1}{2}\sum_{k=0}^{\infty}\frac{k^n}{2^k}\sim\frac{n!}{2}\left(\frac{1}{\ln(2)}\right)^{n+1}\sim\frac{n!}{2\ln(2)}(1.4426950408...)^n$$

as $n\to\infty$, where $S_{n,m}$ is a Stirling number of the second kind [8]. In counting all surjections, we treat extensions as distinct; for example, the preceding function f is not the same as the function $g:\{1,2,3\}\to\{1,2\}$ with $g(x)=x \bmod 2$, nor is it the same as the preceding function h.

Various refinements of surjections are available. An **ℓ-surjection** has the property that every value in the range $\{1,\ldots,m\}$ is taken with multiplicity at least ℓ. (The phrase "double surjection" was used in [6], while "2-surjection" meant something different.) Asymptotic counting results for 2-surjections, 3-surjections and 4-surjections are

$$\frac{n!}{(1+r)r}(0.8724532496...)^n \quad \text{where} \quad r=1.1461932206... \text{ solves } e^r=2+r,$$

$$\frac{n!}{\left(1+\frac{1}{2}r^2\right)r}(0.6377063010...)^n \quad \text{where} \quad r=1.5681199923... \text{ solves } 2e^r=4+2r+r^2,$$

$$\frac{n!}{\left(1+\frac{1}{6}r^3\right)r}(0.5060319662...)^n \quad \text{where} \quad r=1.9761597421... \text{ solves } 6e^r=12+6r+3r^2+r^3,$$

respectively (the numerical value within parentheses is $1/r$). The formulas for $\ell=3$ and 4 are due to Kotěšovec [5].

Another way of imagining a surjection is as a **labeled clique**, that is, a hierarchy on $\{1,\ldots,n\}$ in which vertical ordering is important but horizontal ordering is not. We illustrate $\#\, J_{3,1} = 1$, $\#\, J_{3,2} = 6$, $\#\, J_{3,3} = 6$ here:

$$\left|1,2,3\right|,$$

$$\left|\begin{array}{c}1\\2,3\end{array}\right|,\ \left|\begin{array}{c}2\\1,3\end{array}\right|,\ \left|\begin{array}{c}3\\1,2\end{array}\right|,\ \left|\begin{array}{c}1,2\\3\end{array}\right|,\ \left|\begin{array}{c}1,3\\2\end{array}\right|,\ \left|\begin{array}{c}2,3\\1\end{array}\right|,$$

$$\left|\begin{array}{c}1\\2\\3\end{array}\right|,\ \left|\begin{array}{c}1\\3\\2\end{array}\right|,\ \left|\begin{array}{c}2\\3\\1\end{array}\right|,\ \left|\begin{array}{c}2\\1\\3\end{array}\right|,\ \left|\begin{array}{c}3\\1\\2\end{array}\right|,\ \left|\begin{array}{c}3\\2\\1\end{array}\right|.$$

If we remove labels, then just 4 hierarchies emerge:

$$\left|*,*,*\right|,\ \left|\begin{array}{c}*\\*,*\end{array}\right|,\ \left|\begin{array}{c}*,*\\*\end{array}\right|,\ \left|\begin{array}{c}*\\*\\*\end{array}\right|.$$

More generally [9], the number of **unlabeled cliques** on n integers is 2^n.

A **labeled society** on $\{1,\ldots,n\}$ is created by distributing the elements into cliques. The ordering of the cliques is not important. Let S_n denote the number of such societies and s_n denote the unlabeled analog. The cliques are visually separated by bars and (as before) hierarchy within a clique is indicated by the vertical arrangement. We illustrate $S_3 = 23$ and $s_3 = 7$, omitting the 13 one-clique cases for the former and the 4 one-clique cases for the latter (which were already given):

$$\left|\begin{array}{c|c}1,2 & 3\end{array}\right|,\ \left|\begin{array}{c|c}1,3 & 2\end{array}\right|,\ \left|\begin{array}{c|c}2,3 & 1\end{array}\right|,$$

$$\left|\begin{array}{c|c}1 & \\ 2 & 3\end{array}\right|,\ \left|\begin{array}{c|c}1 & \\ 3 & 2\end{array}\right|,\ \left|\begin{array}{c|c}2 & \\ 3 & 1\end{array}\right|,$$

$$\left|\begin{array}{c|c}2 & \\ 1 & 3\end{array}\right|,\ \left|\begin{array}{c|c}3 & \\ 1 & 2\end{array}\right|,\ \left|\begin{array}{c|c}3 & \\ 2 & 1\end{array}\right|,$$

$$\left|\begin{array}{c|c|c}1 & 2 & 3\end{array}\right|;$$

$$\left|\begin{array}{c|c}*,* & *\end{array}\right|,\ \left|\begin{array}{c|c}* & \\ * & *\end{array}\right|,\ \left|\begin{array}{c|cc}* & * & *\end{array}\right|.$$

More generally [10–12],

$$S_n = \frac{d^n}{dx^n}\exp\left(\frac{1}{2-e^x} - 1\right)\Bigg|_{x=0}, \quad s_n = \frac{1}{n!}\,\frac{d^n}{dx^n}\prod_{k=1}^{\infty}\frac{1}{(1-x^k)^{2^{k-1}}}\Bigg|_{x=0}$$

and

$$S_n \sim C\frac{e^{\sqrt{2n/\ln(2)}}}{n^{3/4}\ln(2)^n}n!, \quad s_n \sim \frac{c}{\sqrt{2\pi}}\frac{e^{\sqrt{2n}-1/4}}{n^{3/4}}2^{n-3/4}$$

as $n \to \infty$, where

$$C = \frac{1}{4\sqrt{\pi}}\left(\frac{2}{e}\right)^{3/4}\left(\frac{e^{1/\ln(2)}}{\ln(2)}\right)^{1/4} = (1038.9726974426...)^{-1/4},$$

$$c = \exp\left(\sum_{j=2}^{\infty}\frac{1}{j(2^j-1)}\right) = 1.3976490050....$$

The constant c, overlooked in [10], was subsequently determined in [13].

Let us focus entirely on the labeled scenario henceforth. A clique is **elitist** if, given any two adjacent levels, the number of elements in the higher level never exceeds the number of elements in the lower level. Define R_n to be the number of elitist cliques on $\{1,\ldots,n\}$. Clearly $R_2 = 3$ and $R_3 = 10$. More generally [9, 12, 14],

$$R_n = \left.\frac{d^n}{dx^n}\prod_{k=1}^{\infty}\left(1-\frac{x^k}{k!}\right)^{-1}\right|_{x=0}$$

and $R_n \sim B\,n!$ as $n \to \infty$, where

$$B = \prod_{k=2}^{\infty}\left(1-\frac{1}{k!}\right)^{-1} = 2.5294774720....$$

Another interpretation involves multinomial coefficients [15]: for suitably large m,

$$(x_1+x_2+\cdots+x_m)^2 = \sum_i x_i^2 + 2\sum_{i<j}x_ix_j,$$

$$(x_1+x_2+\cdots+x_m)^3 = \sum_i x_i^3 + 3\sum_{i\neq j}x_ix_j^2 + 6\sum_{i<j<k}x_ix_jx_k,$$

$$\begin{aligned}(x_1+x_2+\cdots+x_m)^4 &= \sum_i x_i^4 + 4\sum_{i\neq j}x_ix_j^3 + 6\sum_{i<j}x_i^2x_j^2\\ &\quad + 12\sum_{\substack{i<j,\\ i\neq k,\, j\neq k}}x_ix_jx_k^2 + 24\sum_{i<j<k<\ell}x_ix_jx_kx_\ell,\end{aligned}$$

hence $R_2 = 1+2$, $R_3 = 1+3+6$ and $R_4 = 1+4+6+12+24$.

Finally, a society is elitist if all of its cliques are elitist. Define Q_n to be the number of elitist societies on $\{1,\ldots,n\}$. Clearly $Q_2 = 4$ and $Q_3 = 20$. More generally,

$$Q_n = \left.\frac{d^n}{dx^n}\exp\left(\prod_{k=1}^{\infty}\left(1-\frac{x^k}{k!}\right)^{-1}-1\right)\right|_{x=0},$$

but an asymptotic expression for Q_n appears to be open.

In closing, we give a sequence [9, 16]

$$p_n = \frac{1}{n!}\,\frac{d^n}{dx^n}\left(2-\prod_{k=1}^{\infty}\left(1-x^k\right)^{-1}\right)^{-1}\Bigg|_{x=0} = \frac{1}{n!}\,\frac{d^n}{dx^n}\frac{1}{f(x)}\bigg|_{x=0},$$

which arises from unlabeled cliques on *set partitions* rather than integers. It is quite similar to the sequence 2^n mentioned earlier. We illustrate $p_3 = 8$ here:

$$\overline{\underline{|\{*,*,*\}|}},\quad \overline{\underline{|\{*,*\},\{*\}|}},\quad \overline{\underline{|\{*\},\{*\},\{*\}|}},\quad \overline{\underline{\left|\begin{array}{c}\{*,*\}\\ \{*\}\end{array}\right|}},$$

$$\overline{\underline{\left|\begin{array}{c}\{*\}\\ \{*,*\}\end{array}\right|}},\quad \overline{\underline{\left|\begin{array}{c}\{*\}\\ \{*\},\{*\}\end{array}\right|}},\quad \overline{\underline{\left|\begin{array}{c}\{*\},\{*\}\\ \{*\}\end{array}\right|}},\quad \overline{\underline{\left|\begin{array}{c}\{*\}\\ \{*\}\\ \{*\}\end{array}\right|}}.$$

It is easily shown that $p_n \sim a\,b^n$ where $b = 2.6983291064...$ is the unique positive solution of the equation $f(1/y) = 0$ and

$$a = \frac{-b}{f'(1/b)} = 0.4141137931....$$

The fit is excellent. Moreover, the occurrence of the Dedekind eta function [17] is unexpected. Replacing $f(x)$ by $f(x) - 1$ spawns another (alternating in sign) integer sequence [16]; we wonder whether this perturbation possesses a combinatorial interpretation. Societies (labeled or not, elitist or not) can also be imposed in the new partitional framework and more asymptotic results await discovery.

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000522 and A007526.
[2] L. Halbeisen and N. Hungerbühler, Number theoretic aspects of a combinatorial function. *Notes Number Theory Discrete Math.,* v. 5 (1999) n. 4, 138–150; MR1764301 (2001c:11029).
[3] M. Hassani, Counting and computing by *e*, arXiv:math/0606613.
[4] S. R. Finch, Natural logarithmic base, *first volume*, pp. 12–17.
[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000670, A032032, A102233, and A232475.
[6] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009, pp. 106–110, 244–245, 259–261, 335; MR2483235 (2010h:05005).
[7] H. S. Wilf, *generatingfunctionology*, 2nd ed., Academic Press, 1994, pp. 175–176; MR1277813 (95a:05002).
[8] S. R. Finch, Lengyel's constant: Stirling partition numbers, *first volume*, pp. 316–317.
[9] T. Wieder, The number of certain rankings and hierarchies formed from labeled or unlabeled elements and sets, *Appl. Math. Sci. (Ruse)* 3 (2009) 2707–2724; MR2563113 (2010m:05025).
[10] N. J. A. Sloane and T. Wieder, The number of hierarchical orderings, *Order* 21 (2004) 83–89; arXiv:math/0307064; MR2128036 (2005k:06012).
[11] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge Univ. Press, 2009, p. 571; MR2483235 (2010h:05005).

[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005651, A034691, A034899, A075729, and A143463.

[13] V. Kotěšovec, Asymptotics of sequence A034691, unpublished note (2014).

[14] A. Knopfmacher, A. M. Odlyzko, B. Pittel, L. B. Richmond, D. Stark, G. Szekeres and N. C. Wormald, The asymptotic number of set partitions with unequal block sizes, *Elec. J. Combin.* 6 (1999) R2; MR1663723 (2000b:05018).

[15] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Reidel, 1974, p. 126; MR0460128 (57 #124).

[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A055887 and A082531.

[17] S. R. Finch, Dedekind eta products, *this volume*, §1.19.

2

Inequalities and Approximation

2.1 Hardy–Littlewood Maximal Inequalities

The operators M and N defined here were first introduced by Hardy & Littlewood [1]. These tools are useful in several areas, e.g., harmonic analysis [2], but we disregard the applications entirely and focus rather on properties of M and N in themselves.

2.1.1 *One Dimension, Uncentered*

For a locally integrable function $f : \mathbb{R} \to \mathbb{R}$, define

$$(Mf)(x) = \sup_{\substack{a<x \\ b>x}} \frac{1}{b-a} \int_a^b |f(t)|\, dt.$$

In the Banach space $L_p(\mathbb{R})$, $1 \leq p < \infty$, with norm

$$||f||_p = \left(\int_{-\infty}^{\infty} |f(t)|^p\, dt \right)^{\frac{1}{p}},$$

we examine the inequality

$$||Mf||_p \leq c_p \cdot ||f||_p$$

and ask for the best constant c_p. (By "best", we mean that c_p is the smallest positive constant for which the inequality holds for all f.) It is known, for $1 < p < \infty$, that c_p is the unique positive solution of [3]

$$(p-1)x^p - px^{p-1} - 1 = 0;$$

hence, for example, we have $c_2 = 1 + \sqrt{2}$ and $\lim_{p\to\infty} c_p = 1$.

For $p = 1$, we examine instead the weak type $(1, 1)$ inequality

$$|\{x : (Mf)(x) > \lambda\}| \leq C \cdot \frac{1}{\lambda} \cdot ||f||_1$$

where $|S|$ denotes the Lebesgue measure of a measurable set $S \subseteq \mathbb{R}$ and $\lambda > 0$. In this case, it is comparatively simple to prove that $C = 2$ is the best constant [4], valid for all f and all λ.

2.1.2 *One Dimension, Centered*

For a locally integrable function $f: \mathbb{R} \to \mathbb{R}$, define

$$(Nf)(x) = \sup_{h>0} \frac{1}{2h} \int_{x-h}^{x+h} |f(t)|\, dt.$$

Soria & Carbery [5–7] conjectured that $C = 3/2$ is the best constant for the weak type $(1, 1)$ inequality

$$|\{x : (Nf)(x) > \lambda\}| \le C \cdot \frac{1}{\lambda} \cdot ||f||_1.$$

Aldaz [8] refuted this conjecture and showed that $37/24 \le C \le (9 + \sqrt{41})/8$. Further progress was made in [9, 10] before Melas [4] established that

$$C = \frac{11 + \sqrt{61}}{12} = 1.5675208063....$$

The impressive proof underlying this formula is far more complicated than the corresponding uncentered result (§2.1.1).

For the strong type (p, p) inequality with $p > 1$, Dror, Ganguli & Strichartz [7] conjectured that the best constant c_p is given by

$$c_p = \frac{(y+1)^{\frac{p-1}{p}} + (y-1)^{\frac{p-1}{p}}}{2y^{\frac{p-1}{p}}}$$

where $y > 1$ uniquely satisfies

$$\left(1 - \frac{y}{p}\right)^p (y+1) - \left(1 + \frac{y}{p}\right)^p (y-1) = 0;$$

hence, for example, $c_2 = \sqrt[4]{27}/\sqrt{2}$ and $\lim_{p\to\infty} c_p = 1$. Grafakos, Montgomery-Smith & Motrunich [11] confirmed the truth of this formula for a special class of "bell-shaped" functions, but expressed doubt that it holds for all $f \in L_p(\mathbb{R})$. The problem remains unsolved.

2.1.3 *n Dimensions, Uncentered*

Let $n \ge 2$. For a locally integrable function $f: \mathbb{R}^n \to \mathbb{R}$, define

$$(M_n f)(x) = \sup_Q \frac{1}{|Q|} \int_Q |f(t)|\, dt,$$

where the supremum is taken over all compact cubes Q with sides parallel to the coordinate axes, subject only to $x \in Q$. For fixed $1 < p < \infty$, the best constant $c_{p,n}$ must grow at least exponentially as $n \to \infty$ [3]. This result is also true if we replace cubes by balls.

2.1.4 *n Dimensions, Centered*

Let $n \geq 2$. Define similarly

$$(N_n f)(x) = \sup_Q \frac{1}{|Q|} \int\limits_Q |f(t)|\, dt,$$

where we insist not only that $x \in Q$, but additionally that each cube Q is centered at x. For the weak type $(1,1)$ inequality, we have lower bounds on the best constants C_n, for example [12]

$$C_2 \geq \frac{3 + \sqrt{2}(2\sqrt{3} - 1)}{4},$$

$$\liminf_{n\to\infty} C_n \geq \frac{47\sqrt{2}}{36}.$$

It would be good someday to know the exact values of these constants. Moreover, we have $C_1 < C_2$ and $C_n \leq C_{n+1}$ for all n [13]. Stein & Strömberg [14] demonstrated that C_n grows at most like $O(n \ln(n))$ and like $O(n)$ if we replace cubes by balls.

Let us return finally to the strong type (p, p) setting. There exists a constant K for which [14]

$$c_{p,n} \leq K \cdot \frac{p}{p-1} \cdot n$$

for all p and n. If we replace cubes by balls, then n can be further replaced by $\sqrt{n}$. Also, it is possible to write

$$c_{p,n} \leq F(p)$$

for all n, in the case of balls (but the expression $F(p)$ may have to grow more rapidly than $p/(p-1)$ as $p \to 1^+$). Thus, for fixed $1 < p < \infty$, $c_{p,n}$ is bounded as $n \to \infty$. This result contrasts strikingly with the uncentered case (§2.1.3).

[1] G. H. Hardy and J. E. Littlewood, A maximal theorem with function-theoretic applications, *Acta Math.* 54 (1930) 81–116; also in *Collected Papers of G. H. Hardy*, v. 2, Oxford Univ. Press, 1967, pp. 509–545.

[2] A. Torchinsky, *Real-Variable Methods in Harmonic Analysis*, Academic Press, 1986, pp. 76–84; MR0869816 (88e:42001).

[3] L. Grafakos and S. Montgomery-Smith, Best constants for uncentred maximal functions, *Bull. London Math. Soc.* 29 (1997) 60–64; math.FA/9412218; MR1416408 (98b:42031).

[4] A. D. Melas, The best constant for the centered Hardy-Littlewood maximal inequality, *Annals of Math.* 157 (2003) 647–688; MR1973058 (2004e:42032).
[5] D. A. Brannan and W. K. Hayman, Research problems in complex analysis, *Bull. London Math. Soc*. 21 (1989) 1–35; MR0967787 (89m:30001).
[6] M. Trinidad Menarguez and F. Soria, Weak type $(1,1)$ inequalities of maximal convolution operators, *Rend. Circ. Mat. Palermo* 41 (1992) 342–352; MR1230582 (94i:42025).
[7] R. Dror, S. Ganguli, and R. S. Strichartz, A search for best constants in the Hardy-Littlewood maximal theorem, *J. Fourier Anal. Appl.* 2 (1996) 473–486; MR1412064 (98b:42030).
[8] J. M. Aldaz, Remarks on the Hardy-Littlewood maximal function, *Proc. Roy. Soc. Edinburgh Sect. A* 128 (1998) 1–9; MR1606325 (99b:42020).
[9] A. D. Melas, On the centered Hardy-Littlewood maximal operator, *Trans. Amer. Math. Soc.* 354 (2002) 3263–3273; MR1897399 (2003c:42021).
[10] A. D. Melas, On a covering problem related to the centered Hardy-Littlewood maximal inequality, *Ark. Mat.* 41 (2003) 341–361; MR2011925 (2004h:42022).
[11] L. Grafakos, S. Montgomery-Smith, and O. Motrunich, A sharp estimate for the Hardy-Littlewood maximal function, *Studia Math.* 134 (1999) 57–67; math.FA/9704211; MR1688215 (2000b:42015).
[12] J. M. Aldaz, A remark on the centered n-dimensional Hardy-Littlewood maximal function, *Czechoslovak Math. J.* 50 (2000) 103–112; MR1745465 (2002b:42024).
[13] J. M. Aldaz and J. L. Varona, Singular measures and convolution operators, *Acta Math. Sinica (Engl. Ser.)* 23 (2007) 487–490; math.CA/0207285; MR2292694.
[14] E. M. Stein and J.-O. Strömberg, Behavior of maximal functions in $\mathbb{R}^n$ for large n, *Ark. Mat.* 21 (1983) 259–269; MR0727348 (86a:42027).

2.2 Bessel Function Zeroes

The Bessel function $J_\nu(x)$ of the first kind

$$J_\nu(x)=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!\Gamma(\nu+k+1)}\left(\frac{x}{2}\right)^{\nu+2k}, \quad \nu>-1$$

has infinitely many positive zeros

$$0<j_{\nu,1}<j_{\nu,2}<j_{\nu,3}<\cdots,$$

as does its derivative $J'_\nu(x)$:

$$0<j'_{\nu,1}<j'_{\nu,2}<j'_{\nu,3}<\cdots, \quad \nu>0,$$

$$0=j'_{0,1}<j'_{0,2}<j'_{0,3}<j'_{0,4}<\cdots, \quad \nu=0.$$

See Tables 2.1 & 2.2 for the cases $\nu=0,1,2$ and Tables 2.3 & 2.4 for the cases $\nu=1/2,3/2,5/2$. These appear in many physical applications that we cannot hope to survey in entirety. We will state only a few properties and several important inequalities. A starting point for research is Watson's monumental treatise [1].

Table 2.1 *Zeroes of J_ν for $s = 1, 2, 3$ and integer ν*

$j_{0,s}$	$j_{1,s}$	$j_{2,s}$
2.4048255576...	3.8317059702...	5.1356223018...
5.5200781102...	7.0155866698...	8.4172441403...
8.6537279129...	10.1734681350...	11.6198411721...

Table 2.2 *Zeroes of J'_ν for $s = 1, 2, 3$ and integer ν*

$j'_{0,s}$	$j'_{1,s}$	$j'_{2,s}$
0	1.8411837813...	3.0542369282...
3.8317059702...	5.3314427735...	6.7061331941...
7.0155866698...	8.5363163663...	9.9694678230...

Table 2.3 *Zeroes of J_ν for $s = 1, 2, 3$ and half-integer ν*

$j_{1/2,s}$	$j_{3/2,s}$	$j_{5/2,s}$
π	4.4934094579...	5.7634591968...
2π	7.7252518369...	9.0950113304...
3π	10.9041216594...	12.3229409705...

Table 2.4 *Zeroes of J'_ν for $s = 1, 2, 3$ and half-integer ν*

$j'_{1/2,s}$	$j'_{3/2,s}$	$j'_{5/2,s}$
1.1655611852...	2.4605355721...	3.6327973198...
4.6042167772...	6.0292923816...	7.3670089715...
7.7898837511...	9.2614019262...	10.6635613904...

Clearly $j_{\nu,s} \to \infty$ as $s \to \infty$ with ν fixed; in fact, $j_{\nu,s+1} - j_{\nu,s} \to \pi$. For $\nu \geq 0$, here is a straightforward lower bound [2, 3]:

$$j_{\nu,s} > \sqrt{\left(s - \tfrac{1}{4}\right)^2 \pi^2 + \nu^2}$$

and, for $\nu > 0$, here are more complicated bounds [4–6]:

$$\nu + \alpha_s \nu^{1/3} < j_{\nu,s} < \nu + \alpha_s \nu^{1/3} + \frac{3\alpha_s^2}{10} \nu^{-1/3},$$

where $\alpha_s = 2^{-1/3} a_s$ and a_s is the s^{th} positive root of the equation

$$J_{\frac{1}{3}}\left(\tfrac{2}{3}x^{3/2}\right) + J_{-\frac{1}{3}}\left(\tfrac{2}{3}x^{3/2}\right) = 0.$$

For example, $a_1 = 2.3381074104...$ and thus the coefficients of $\nu^{1/3}$ and $\nu^{-1/3}$ for $s = 1$ are 1.8557570814... and 1.0331503036..., respectively. (The left-hand side of the equation is the same as $3\,\mathrm{Ai}(-x)/\sqrt{x}$, where Ai is the Airy function.) These bounds are asymptotically precise; more terms in the asymptotic expansion of $j_{\nu,s}$ as $\nu \to \infty$, for any fixed s, can be obtained [7–10]. Related work includes [11–15].

Similarly we have

$$\nu + \alpha_s' \nu^{1/3} < j_{\nu,s}' < \nu + \alpha_s' \nu^{1/3} + \frac{3\alpha_s'^3 - 1}{10\alpha_s'} \nu^{-1/3},$$

where $\alpha_s' = 2^{-1/3} a_s'$ and a_s' is the s^{th} positive root of the equation

$$J_{\frac{2}{3}}\left(\tfrac{2}{3}x^{3/2}\right) - J_{-\frac{2}{3}}\left(\tfrac{2}{3}x^{3/2}\right) = 0.$$

For example, $a_1' = 1.0187929716...$ and thus the coefficients of $\nu^{1/3}$ and $\nu^{-1/3}$ for $s = 1$ are 0.8086165174... and 0.0724901862..., respectively. (The left-hand side of the equation is the same as $3\,\mathrm{Ai}'(-x)/x$.) The zeroes of J_ν and J_ν' are interlaced:

$$... < j_{\nu,s}' < j_{\nu,s} < j_{\nu,s+1}' < j_{\nu,s+1} < ...$$

and further satisfy [16]

$$j_{\nu,s+1}' > \sqrt{j_{\nu,s} j_{\nu,s+1}}.$$

Let $n \geq 0$ be an integer. Every Bessel function $J_{n+1/2}(x)$ is elementary; for example, $\sqrt{x} J_{1/2}(x)$ can be simplified to $\sqrt{2/\pi}\sin(x)$. Consequently $j_{3/2,s}$ is the s^{th} positive root of the equation

$$\sin(x) - x\cos(x) = 0, \text{ that is, } \tan(x) = x,$$

and $j_{1/2,s}'$ is the s^{th} positive root of the equation

$$\sin(x) - 2x\cos(x) = 0, \text{ that is, } \tan(x) = 2x.$$

Siegel [1, 17, 18] proved that $J_\nu(\xi)$ is transcendental whenever ν is rational and ξ is algebraic. It follows immediately that every zero $j_{\nu,s}$ is transcendental. Further, if μ is rational and $\nu - \mu \neq 0$ is an integer, then $J_\nu(x)$ and $J_\mu(x)$ can never have common zeroes (other than $x = 0$) [19–22].

Series of the form [1, 23]

$$\sum_{s=1}^{\infty} \frac{1}{j_{\nu,s}^2} = \frac{1}{4(\nu+1)}, \quad \sum_{s=1}^{\infty} \frac{1}{j_{\nu,s}^4} = \frac{1}{16(\nu+1)^2(\nu+2)}$$

possess well-known special cases. If $\nu = 1/2$, then $j_{\nu,s} = \pi s$ and

$$\sum_{s=1}^{\infty} \frac{1}{s^2} = \frac{\pi^2}{6}, \quad \sum_{s=1}^{\infty} \frac{1}{s^4} = \frac{\pi^4}{90}$$

as given in [24]. We also have

$$\sum_{s=1}^{\infty} \frac{1}{j_{0,s}^2} = \frac{1}{4}, \quad \sum_{s=1}^{\infty} \frac{1}{j_{3/2,s}^2} = \frac{1}{10}$$

and the latter series appears in [25]. Other identities can be found in [26, 27].

We need three more tables before continuing. Define

$$\begin{aligned} P_\nu(x) &= \frac{d}{dx}\left(x^{1-\nu} J_\nu(x)\right) = x^{-\nu}\left((1-\nu)J_\nu(x) + xJ'_\nu(x)\right), \\ Q_\nu(x) &= J_\nu(x) I_{\nu+1}(x) + I_\nu(x) J_{\nu+1}(x) \end{aligned}$$

where $I_\nu(x)$ is the modified Bessel function of the first kind:

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{k!\Gamma(\nu+k+1)} \left(\frac{x}{2}\right)^{\nu+2k} = i^{-\nu} J_\nu(ix).$$

Let $p_{\nu,s}$ and $q_{\nu,1}$ denote the s^{th} smallest positive zeroes of $P_\nu(x)$ and $Q_\nu(x)$. It is clear that $p_{1,s} = j'_{1,s}$ for all s. (See Tables 2.5 & 2.6.)

Finally, we offer an application. Table 2.7 gives the vibration modes of an idealized timpani (or kettledrum). By contrast, the frequency ratios for overtones of an idealized guitar string are all integers [28].

Table 2.5 *Zeroes of P_ν for $s = 1, 2, 3$*

$p_{1,s}$	$p_{3/2,s}$	$p_{2,s}$
1.8411837813...	2.0815759778...	2.2999103302...
5.3314427735...	5.9403699905...	6.5414028262...
8.5363163663...	9.2058401429...	9.8647278383...

Table 2.6 *Zeroes of Q_ν for $s = 1, 2, 3$*

$q_{0,s}$	$q_{1/2,s}$	$q_{1,s}$
3.1962206165...	3.9266023120...	4.6108998790...
6.3064370476...	7.0685827456...	7.7992738008...
9.4394991378...	10.2101761228...	10.958067191...

Table 2.7 *Frequency ratios for the first five overtones of a fixed circular membrane*

ν	s	$j_{\nu,s}/\pi$	$j_{\nu,s}/j_{0,1}$
0	1	0.7654797495...	1
1	1	1.2196698912...	1.5933405056...
2	1	1.6347193503...	2.1355487866...
0	2	1.7570954350...	2.2954172674...
3	1	2.0308686069...	2.6530664045...
1	2	2.2331305943...	2.9172954551...

2.2.1 *Membrane and Plate Inequalities*

Let $n \geq 2$. Let $\Omega \subseteq \mathbb{R}^n$ be a connected bounded open set of volume $|\Omega|$, and assume that its boundary $\partial\Omega$ is smooth. Define the **Laplacian** and **bi-Laplacian (biharmonic)** operators

$$\triangle f = \sum_{k=1}^{n} \frac{\partial^2 f}{\partial^2 x_k}, \quad \triangle^2 f = \triangle(\triangle f)$$

for smooth functions $f: \Omega \to \mathbb{R}$. We will briefly consider four famous eigenvalue problems (i.e., isoperimetric inequalities) that occur in structural dynamics for which Bessel function zeroes play a role [29, 30].

The **fixed (fastened) membrane** problem involves the Laplacian with **Dirichlet** boundary conditions:

$$-\triangle u = \lambda u \quad \text{in } \Omega,$$

$$u = 0 \quad \text{on } \partial\Omega.$$

We seek the smallest eigenvalue $\lambda_1(\Omega)$, that is, the fundamental frequency of vibration. When is $\lambda_1(\Omega)$ minimal? The **Rayleigh–Faber–Krahn** inequality provides that [31]

$$\lambda_1(\Omega) \geq \left(\frac{\omega_n}{|\Omega|}\right)^{2/n} j^2_{\frac{n}{2}-1,1}$$

with equality if and only if Ω is a ball. Here $\omega_n = \pi^{n/2}/\Gamma(n/2+1)$ is the volume of the unit ball in $\mathbb{R}^n$. Only the case $n = 2$ was mentioned in [32]. For example, $j^2_{0,1} = 5.7831859629...$

The **free membrane** problem involves the Laplacian with **Neumann** boundary conditions:

$$-\triangle v = \mu v \quad \text{in } \Omega,$$

$$\frac{\partial v}{\partial n} = 0 \quad \text{on } \partial\Omega,$$

where $\partial v/\partial n$ denotes the outward normal derivative of v. Since $\mu_1(\Omega)=0$, we seek the next-to-smallest eigenvalue $\mu_2(\Omega)$. When is $\mu_2(\Omega)$ maximal? The **Szegö–Weinberger** inequality provides that [33–36]

$$\mu_2(\Omega) \leq \left(\frac{\omega_n}{|\Omega|}\right)^{2/n} p^2_{\frac{n}{2},1}$$

with equality if and only if Ω is a ball.

The **clamped plate** problem involves the bi-Laplacian with the following boundary conditions:

$$\triangle^2 w = \Lambda w \quad \text{in } \Omega,$$

$$w = \frac{\partial w}{\partial n} = 0 \quad \text{on } \partial\Omega.$$

We seek the smallest eigenvalue $\Lambda_1(\Omega)$. When is $\Lambda_1(\Omega)$ minimal? The **Nadirashvili–Ashbaugh-Benguria** inequality provides that [37–39]

$$\Lambda_1(\Omega) \geq \left(\frac{\omega_n}{|\Omega|}\right)^{4/n} q^4_{\frac{n}{2}-1,1}$$

with equality if and only if Ω is a ball. This has been rigorously proved only for $2 \leq n \leq 3$, but it is known to be true for $n \geq 4$ up to a constant factor $\to 1$ as $n \to \infty$. Only the case $n=2$ was mentioned in [32].

The **buckling load** problem involves both the Laplacian and bi-Laplacian with the following boundary conditions:

$$\triangle^2 z = -M \triangle z \quad \text{in } \Omega,$$

$$z = \frac{\partial z}{\partial n} = 0 \quad \text{on } \partial\Omega.$$

We seek the smallest eigenvalue $M_1(\Omega)$. When is $M_1(\Omega)$ minimal? Pólya & Szegö [39, 40] conjectured that

$$M_1(\Omega) \geq \left(\frac{\omega_n}{|\Omega|}\right)^{2/n} j^2_{\frac{n}{2},1}$$

with equality if and only if Ω is a ball, but this is only known to be true up to a constant factor $\to 1$ as $n \to \infty$.

We return to the original Dirichlet problem to state one more idea. The **Payne–Pólya–Weinberger** conjecture, proved by Ashbaugh & Benguria [41–43], involves the maximal ratio of the two smallest eigenvalues $\lambda_1(\Omega)$ and $\lambda_2(\Omega)$:

$$\frac{\lambda_2(\Omega)}{\lambda_1(\Omega)} \leq \frac{j^2_{\frac{n}{2},1}}{j^2_{\frac{n}{2}-1,1}}$$

with equality if and only if Ω is a ball. For example, when $n=2$, the right-hand side is 2.5387339670... What can be said about the maximal ratios of two arbitrary eigenvalues [44]?

2.2.2 *Other Best Constants*

Bessel function zeroes occur in best constants associated with Nash's inequality [45], uncertainty inequalities [46], and with an improved version of Hardy's inequality [47–51].

We close with remarks about the multiplicities of the zeroes. It appears that, for fixed $\nu > 0$, the positive zeroes $j''_{\nu,s}$ of the second derivative $J''_{\nu}(x)$ are all simple, like those of $J_{\nu}(x)$ and $J'_{\nu}(x)$. This is no longer true when considering positive zeroes $j'''_{\nu,s}$ of the third derivative $J'''_{\nu}(x)$: there exists a value $\nu_0 = 0.755378...$ for which J'''_{ν_0} has a double zero $x_0 = 0.959621...$ [52, 53]. Related papers include [54–64] the latter of which are more concerned with the strictly increasing behavior of $j''_{\nu,s}$ as a function of ν for fixed s (rather than of s for fixed ν).

[1] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge Univ. Press, 1944; MR1349110 (96i:33010).

[2] H. W. Hethcote, Bounds for zeros of some special functions, *Proc. Amer. Math. Soc.* 25 (1970) 72–74; MR0255909 (41 #569).

[3] R. C. McCann, Lower bounds for the zeros of Bessel functions, *Proc. Amer. Math. Soc*. 64 (1977) 101–103; MR0442316 (56 #700).

[4] T. Lang and R. Wong, "Best possible" upper bounds for the first two positive zeros of the Bessel function $J_\nu(x)$: the infinite case, *J. Comput. Appl. Math.* 71 (1996) 311–329; MR1399899 (97h:33016).

[5] L. Lorch and R. Uberti, "Best possible" upper bounds for the first positive zeros of Bessel functions—the finite part, *J. Comput. Appl. Math.* 75 (1996) 249–258; MR1426265 (98b:33010).

[6] C. K. Qu and R. Wong, "Best possible" upper and lower bounds for the zeros of the Bessel function $J_\nu(x)$, *Trans. Amer. Math. Soc.* 351 (1999) 2833–2859; MR1466955 (99j:33006).

[7] F. Tricomi, Sulle funzioni di Bellel di ordine e argomento pressochè uguali, *Atti Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat.* 83 (1949) 3–20; MR0034478 (11,594a).

[8] F. W. J. Olver, A further method for the evaluation of zeros of Bessel functions and some new asymptotic expansions for zeros of functions of large order, *Proc. Cambridge Philos. Soc.* 47 (1951) 699–712; MR0043551 (13,283c).

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, pp. 371, 447; MR1225604 (94b:00012).

[10] Á. Elbert and A. Laforgia, Asymptotic expansion for zeros of Bessel functions and of their derivatives for large order, *Atti Sem. Mat. Fis. Univ. Modena Suppl.* 46 (1998) 685–695; MR1645747 (99h:33012).

[11] L. G. Chambers, An upper bound for the first zero of Bessel functions, *Math. Comp.* 38 (1982) 589–591; MR0645673 (83h:33011).

[12] R. Piessens, A series expansion for the first positive zero of the Bessel functions, *Math. Comp.* 42 (1984) 195–197; MR0725995 (84m:33014).

[13] M. E. H. Ismail, and M. E. Muldoon, On the variation with respect to a parameter of zeros of Bessel and q-Bessel functions, *J. Math. Anal. Appl.* 135 (1988) 187–207; MR0960813 (89i:33011).

[14] L. Lorch, Some inequalities for the first positive zeros of Bessel functions, *SIAM J. Math. Anal.* 24 (1993) 814–823; MR1215440 (95a:33010).

[15] Á. Elbert and P. D. Siafarikas, On the square of the first zero of the Bessel function $J_n(z)$, *Canad. Math. Bull.* 42 (1999) 56–67; MR1695874 (2000f:33004).
[16] J. Segura, On a conjecture regarding the extrema of Bessel functions and its generalization, *J. Math. Anal. Appl.* 280 (2003) 54–62; MR1972191 (2004c:33013).
[17] C. L. Siegel, Über einige Anwendungen diophantischer Approximationen, *Abh. Preuss. Akad. Wiss., Phys.-Math. Klasse* (1929) n. 1, 1–70; also in *Gesammelte Abhandlungen*, v. 1, ed. K. Chandrasekharan and H. Maass, Springer-Verlag, 1966, pp. 209–266.
[18] C. L. Siegel, *Transcendental Numbers*, Chelsea 1965, pp. 59, 71–72; MR0032684 (11,330c).
[19] T. C. Benton and H. D. Knoble, Common zeros of two Bessel functions, *Math. Comp.* 32 (1978) 533–535; MR0481160 (58 #1303).
[20] T. C. Benton, Common zeros of two Bessel functions. II. Approximations and tables, *Math. Comp*. 41 (1983) 203–217; MR0701635 (85a:33010).
[21] E. N. Petropoulou, P. D. Siafarikas and I. D. Stabolas, On the common zeros of Bessel functions, *J. Comput. Appl. Math.* 153 (2003) 387–393; MR1985709 (2004g:33010).
[22] J. Haletky, The spacing of zeros of Bessel functions, unpublished note (1999).
[23] I. N. Sneddon, On some infinite series involving the zeros of Bessel functions of the first kind, *Proc. Glasgow Math. Assoc.* 4 (1960) 144–156; MR0120400 (22 #11154).
[24] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[25] S. R. Finch, Du Bois Reymond's constants, *first volume*, pp. 237–240.
[26] X. A. Lin and O. P. Agrawal, A new identity involving positive roots of Bessel functions of the first kind, *J. Franklin Inst. B* 332 (1995) 333–336; MR1368354 (97a:33005).
[27] E. A. Skelton, A new identity for the infinite product of zeros of Bessel functions of the first kind or their derivatives, *J. Math. Anal. Appl.* 267 (2002) 338–344; MR1886832 (2003g:33007).
[28] P. M. Morse, *Vibration and Sound*, 2nd ed., McGraw-Hill, 1948.
[29] M. S. Ashbaugh, Isoperimetric and universal inequalities for eigenvalues, *Spectral Theory and Geometry*, Proc. 1998 Edinburgh conf., ed. E. B. Davies and Yu. Safarov, Cambridge Univ. Press, 1999, pp. 95–139; math.SP/0008087; MR1736867 (2001a:35131).
[30] G. Talenti, On isoperimetric theorems of mathematical physics, *Handbook of Convex Geometry*, ed. P. M. Gruber and J. M. Wills, Elsevier, 1993, pp. 1131–1147; MR1243005 (94i:49002).
[31] E. Krahn, Über Minimaleigenschaften der Kugel in drei und mehr Dimensionen, *Acta et Commentationes Universitatis Tartuensis (Dorpatensis)* A9 (1926) 1–44; Engl. transl. in *Edgar Krahn 1894-1961: A Centenary Volume*, ed. U. Lumiste and J. Peetre, IOS Press, 1994, pp. 139–174.
[32] S. R. Finch, Sobolev isoperimetric constants, *first volume*, pp. 219–225.
[33] G. Szegö, Inequalities for certain eigenvalues of a membrane of given area, *J. Rational Mech. Anal.* 3 (1954) 343–356; MR0061749 (15,877c).
[34] H. F. Weinberger, An isoperimetric inequality for the N-dimensional free membrane problem, *J. Rational Mech. Anal.* 5 (1956) 633–636; MR0079286 (18,63c).
[35] M. S. Ashbaugh and R. D. Benguria, Universal bounds for the low eigenvalues of Neumann Laplacians in n dimensions, *SIAM J. Math. Anal.* 24 (1993) 557–570; MR1215424 (94b:35191).

[36] L. Lorch and P. Szego, Bounds and monotonicities for the zeros of derivatives of ultraspherical Bessel functions, *SIAM J. Math. Anal.* 25 (1994) 549–554; MR1266576 (95b:33034).
[37] N. S. Nadirashvili, Rayleigh's conjecture on the principal frequency of the clamped plate, *Arch. Rational Mech. Anal.* 129 (1995) 1–10; MR1328469 (97j:35113).
[38] M. S. Ashbaugh and R. D. Benguria, On Rayleigh's conjecture for the clamped plate and its generalization to three dimensions, *Duke Math. J.* 78 (1995) 1–17; MR1328749 (97j:35111).
[39] M. S. Ashbaugh and R. S. Laugesen, Fundamental tones and buckling loads of clamped plates, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* 23 (1996) 383–402; MR1433428 (97j:35112).
[40] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Princeton Univ. Press, 1951; MR0043486 (13,270d).
[41] M. S. Ashbaugh and R. D. Benguria, Proof of the Payne-Pólya-Weinberger conjecture, *Bull. Amer. Math. Soc.* 25 (1991) 19–29; MR1085824 (91m:35173).
[42] M. S. Ashbaugh and R. D. Benguria, A sharp bound for the ratio of the first two eigenvalues of Dirichlet Laplacians and extensions, *Annals of Math.* 135 (1992) 601–628; MR1166646 (93d:35105).
[43] M. S. Ashbaugh and R. D. Benguria, A second proof of the Payne-Pólya-Weinberger conjecture, *Comm. Math. Phys.* 147 (1992) 181–190; MR1171765 (93k:33002).
[44] M. Levitin and R. Yagudin, Range of the first three eigenvalues of the planar Dirichlet Laplacian, *London Math. Soc. J. Comput. Math.* 6 (2003) 1–17; math.SP/0203231; MR1971489 (2004c:35305).
[45] S. R. Finch, Nash's inequality, *this volume*, §2.3.
[46] S. R. Finch, Uncertainty inequalities, *this volume*, §2.4.
[47] S. R. Finch, Copson-de Bruijn constant, *first volume*, pp. 217–219.
[48] H. Brezis and J. L. Vázquez, Blow-up solutions of some nonlinear elliptic problems, *Rev. Mat. Univ. Complut. Madrid* 10 (1997) 443–469; MR1605678 (99a:35081).
[49] S. Filippas and A. Tertikas, Optimizing improved Hardy inequalities, *J. Funct. Anal.* 192 (2002) 186–233; MR1918494 (2003f:46045).
[50] F. Gazzola, H.-C. Grunau and E. Mitidieri, Hardy inequalities with optimal constants and remainder terms, *Trans. Amer. Math. Soc.* 356 (2004) 2149–2168; MR2048513 (2005c:26031).
[51] G. Barbatis, S. Filippas and A. Tertikas, A unified approach to improved L^p Hardy inequalities with best constants, *Trans. Amer. Math. Soc.* 356 (2004) 2169–2196; math.AP/0302326; MR2048514 (2005a:26016).
[52] L. Lorch and P. Szego, Monotonicity of the zeros of the third derivative of Bessel functions, *Methods Appl. Anal.* 2 (1995) 103–111; MR1337456 (96g:33006).
[53] L. Lorch, The zeros of the third derivative of Bessel functions of order less than one, *Methods Appl. Anal.* 2 (1995) 147–159; MR1350893 (96k:33005).
[54] M. K. Kerimov and S. L. Skorokhodov, Calculation of multiple zeros of derivatives of cylindrical Bessel functions $J_\nu(z)$ and $Y_\nu(z)$ (in Russian), *Zh. Vychisl. Mat. Mat. Fiz.* 25 (1985) 1749–1760, 1918; Engl. transl. in *USSR Comput. Math. Math. Phys.*, v. 25 (1985) n. 6, 101–107; MR0821711 (87f:33018).
[55] M. K. Kerimov and S. L. Skorokhodov, Multiple zeros of derivatives of cylindrical Bessel functions (in Russian), *Doklady Akad. Nauk SSSR* 288 (1986) 285–288; Engl. transl. in *Soviet Math. Dokl.* 33 (1986) 650–653; MR0843438 (87j:33011).
[56] M. K. Kerimov and S. L. Skorokhodov, Some asymptotic formulas for cylindrical Bessel functions (in Russian), *Zh. Vychisl. Mat. Mat. Fiz.* 30 (1990) 1775–1784; Engl. transl. in *USSR Comput. Math. Math. Phys.*, v. 30 (1990) n. 6, 126–133; MR1099143 (92f:33006).

[57] L. Lorch and P. Szego, On the points of inflection of Bessel functions of positive order. I, *Canad. J. Math.* 42 (1990) 933–948; corrigenda 42 (1990) 1132; MR1081004 (92i:33002a) and MR1099462 (92i:33002b).
[58] R. Wong and T. Lang, On the points of inflection of Bessel functions of positive order. II, *Canad. J. Math.* 43 (1991) 628–651; MR1118013 (92f:33008).
[59] A. McD. Mercer, The zeros of $az^2J_\nu''(z)+bzJ_\nu'(z)+cJ_\nu(z)$ as functions of order, *Internat. J. Math. Math. Sci.* 15 (1992) 319–322; MR1155524 (93a:33010).
[60] L. Lorch and P. Szego, Further on the points of inflection of Bessel functions, *Canad. Math. Bull.* 39 (1996) 216–218; MR1390358 (97i:33002).
[61] C. G. Kokologiannaki and P. D. Siafarikas, An alternative proof of the monotonicity of $j_{\nu,1}''$, *Boll. Unione Mat. Ital. A* 7 (1993) 373–376; errata corrige 9 (1995) 415; MR1249113 (94j:33001) and MR1336247.
[62] E. K. Ifantis, C. G. Kokologiannaki and C. B. Kouris, On the positive zeros of the second derivative of Bessel functions, *J. Comput. Appl. Math.* 34 (1991) 21–31; MR1095193 (92i:33001).
[63] L. Lorch, M. E. Muldoon and P. Szego, Inflection points of Bessel functions of negative order, *Canad. J. Math.* 43 (1991) 1309–1322; MR1145591 (93a:33009).
[64] R. Wong and T. Lang, Asymptotic behaviour of the inflection points of Bessel functions, *Proc. Roy. Soc. London Ser. A* 431 (1990) 509–518; MR1086356 (92i:33003).

2.3 Nash's Inequality

Consider all smooth, compactly supported, s-integrable functions $f:\mathbb{R}^n\to\mathbb{R}$ with the property that the Euclidean norm of the gradient $\nabla f:\mathbb{R}^n\to\mathbb{R}^n$ is q-integrable:

$$||f||_s=\left(\int\limits_{\mathbb{R}^n}|f(x)|^s\,dx\right)^{\frac{1}{s}}<\infty,\quad s\geq 1;$$

$$||\nabla f||_q=\left(\int\limits_{\mathbb{R}^n}|\nabla f(x)|^q\,dx\right)^{\frac{1}{q}}<\infty,\quad q\geq 1.$$

For example, let $q=2$ and $s=1$. Nash's inequality [1]

$$||f||_2^{2+\frac{4}{n}}\leq A_n\cdot||\nabla f||_2^2\cdot||f||_1^{\frac{4}{n}},$$

that is,

$$\left(\int\limits_{\mathbb{R}^n}|f(x)|^2\,dx\right)^{1+\frac{2}{n}}\leq A_n\left(\int\limits_{\mathbb{R}^n}|\nabla f(x)|^2\,dx\right)\left(\int\limits_{\mathbb{R}^n}|f(x)|\,dx\right)^{\frac{4}{n}},$$

is useful in the study of nonlinear partial differential equations (PDEs). Best constants A_n were proved by Carlen & Loss [2] to be

$$A_n=\left(1+\frac{2}{n}\right)\Gamma\left(\frac{n}{2}+1\right)^{\frac{2}{n}}\frac{1}{\pi j_{n/2,1}^2}$$

where $j_{n/2,1}$ is the smallest positive zero [3] of the Bessel function $J_{n/2}(x)$. Hence

$$A_1 = \frac{27}{16\pi^2} = 0.1709794973..., \quad A_2 = 0.0867212975..., \quad A_3 = 0.0585146159...$$

and $A_n \sim 2/(\pi e n)$ as $n \to \infty$. This asymptotic result is due to Beckner [4–7].

As another example, let $q = 2$ and $s = 2$. Best constants for Moser's inequality [8, 9]

$$||f||_{2+\frac{4}{n}}^{2+\frac{4}{n}} \le B_n \cdot ||\nabla f||_2^2 \cdot ||f||_2^{\frac{4}{n}},$$

that is,

$$\int\limits_{\mathbb{R}^n} |f(x)|^{2+\frac{4}{n}}\, dx \le B_n \left(\int\limits_{\mathbb{R}^n} |\nabla f(x)|^2\, dx \right) \left(\int\limits_{\mathbb{R}^n} |f(x)|^2\, dx \right)^{\frac{2}{n}},$$

are known exactly only for $n = 1$ [10]:

$$B_1 = \frac{4}{\pi^2} = 0.4052847345...$$

When $n = 2$, we have a numerical estimate $B_2 = 0.170927... = (5.85043...)^{-1}$; more will be said about this constant shortly. Here too it is known that $B_n \sim 2/(\pi e n)$ as $n \to \infty$ [5].

2.3.1 *Gagliardo–Nirenberg*

A generalization of Nash's inequality is [11–13]

$$||f||_r \le \kappa_n(q, r, s) \cdot ||\nabla f||_q^\theta \cdot ||f||_s^{1-\theta},$$

where $1 < q < n$, $s \ge 1$, $0 \le \theta \le 1$ and

$$\frac{1}{r} = \left(\frac{1}{q} - \frac{1}{n} \right) \theta + \frac{1}{s}(1 - \theta).$$

These conditions force $r \ge 1$. Note that the Gagliardo–Nirenberg inequality trivially encompasses the p-Sobolev inequality when $q = p$ and $\theta = 1$ (details appear in §2.3.3). We have already examined best constants for one case:

$$\kappa_n(2, 2, 1) = A_n^{\theta/2}, \quad \theta = \frac{n}{n+2}$$

and wonder about any other nontrivial cases possessing explicit formulas for all n. Del Pino & Dolbeault discovered two one-parameter families that assist in answering the question [14–17]:

$$\kappa_n\left(q, q\frac{s-1}{q-1}, s\right) = \left(\frac{s-q}{q\sqrt{\pi}}\right)^{\theta}\left(\frac{qs}{n(s-q)}\right)^{\theta/q}\left(\frac{\delta}{qs}\right)^{1/r} \times \left(\frac{\Gamma\left(s\frac{q-1}{s-q}\right)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{q-1}{q}\frac{\delta}{s-q}\right)\Gamma\left(n\frac{q-1}{q}+1\right)}\right)^{\theta/n}$$

for all $1 < q < s$, where $q(s-1) = r(q-1)$ and $\delta = nq - s(n-q) \geq q$, and

$$\kappa_n\left(q, r, q\frac{r-1}{q-1}\right) = \left(\frac{q-r}{q\sqrt{\pi}}\right)^{\theta}\left(\frac{qr}{n(q-r)}\right)^{\theta/q}\left(\frac{qr}{\delta}\right)^{(1-\theta)/s} \times \left(\frac{\Gamma\left(\frac{q-1}{q}\frac{\delta}{q-r}+1\right)\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(r\frac{q-1}{q-r}+1\right)\Gamma\left(n\frac{q-1}{q}+1\right)}\right)^{\theta/n}$$

for all $1 < r < q$, where $q(r-1) = s(q-1)$ and $\delta = nq - r(n-q) > 0$. Gunson [18] stated the first result (in which q, s are free and $r = q(s-1)/(q-1)$), but without proof.

Most cases, however, are like

$$\kappa_n\left(2, 2+\tfrac{4}{n}, 2\right) = B_n^{\theta/2}, \quad \theta = \frac{n}{n+2}$$

in the sense that explicit expressions are presently unavailable for all n. For example [19–23],

$$\kappa_2(2,3,2) = \frac{1}{1.379427\ldots}, \quad \theta = \frac{1}{3};$$

$$\kappa_2(2,4,2) = B_2^{1/4} = \sqrt[4]{\frac{1}{\pi \cdot 1.86225\ldots}} = \frac{1}{1.555239\ldots}, \quad \theta = \frac{1}{2};$$

$$\kappa_2(2,6,2) = \sqrt[3]{\frac{1}{4.5981\ldots}} = \frac{1}{1.663066\ldots}, \quad \theta = \frac{2}{3};$$

$$\kappa_3(2,4,2) = \frac{1}{2.2258\ldots}, \quad \theta = \frac{3}{4}.$$

As a prelude to the next section, define

$$C_n(\sigma) = \kappa_n(2, 2\sigma+2, 2)$$

for $\sigma > 0$; this two-parameter family includes the four constants just listed.

2.3.2 *Schrödinger*

Let $\triangle$ denote the Laplacian operator. A space function $f(x)$ is **radial** if f is a function of $|x|$ alone. Also, a time function $g(t)$ is **global** if it is finite for all t, that is, no blow ups occur in finite time.

Here is an alternative characterization [20] of $C_n(\sigma)$ for $0 < \sigma < 2/(n-2)$:

$$C_n(\sigma) = \left(\frac{\sigma + 1}{||\psi||_2^{2\sigma}}\right)^{\frac{1}{2\sigma+2}}$$

where $\psi : \mathbb{R}^n \to \mathbb{R}$ is a smooth, positive, radial solution of the nonlinear PDE

$$\frac{n\sigma}{2}\triangle\psi - \frac{2\sigma + 2 - n\sigma}{2}\psi + \psi^{2\sigma+1} = 0$$

of minimal norm $||\psi||_2$ (the ground state). If $n = 2$, such a function $\psi(x)$ can be proved to be unique; further,

$$||\psi||_2^2 = (2\pi)(1.86225...)$$

when $\sigma = 1$. This gives rise to our numerical estimate of $C_2(1) = B_2^{1/4}$. It is known (among many things) that the cubic Schrödinger PDE in $\mathbb{R}^2$:

$$2i\frac{\partial\varphi}{\partial t} + \triangle\varphi + |\varphi|^2\varphi = 0$$

with initial conditions

$$\varphi(x, 0) = \varphi_0(x)$$

possesses a global solution $\varphi : \mathbb{R}^2 \times \mathbb{R}^+ \to \mathbb{C}$ if $||\varphi_0||_2 < ||\psi||_2$. The latter inequality is sharp in a certain technical sense involving instability. Solutions $\varphi(x, t)$ of the Schrödinger equation find application in optics and plasma physics [24].

The constant B_2 also appears in the study of intersection local times for planar random walks and planar Brownian motion [25–27].

2.3.3 *Sobolev*

Let $\omega_n = \pi^{n/2}/\Gamma(n/2 + 1)$ denote the volume enclosed by the unit sphere in $\mathbb{R}^n$; consequently $\tilde{\omega}_{n-1} = n\omega_n$ is its surface area. For any $1 \le p < n$, let $p^* = np/(n-p)$. The classical p-Sobolev inequality is as follows:

$$\left(\int\limits_{\mathbb{R}^n} |f(x)|^{p^*}\, dx\right)^{\frac{1}{p^*}} \le K \left(\int\limits_{\mathbb{R}^n} |\nabla f(x)|^p\, dx\right)^{\frac{1}{p}}$$

and the best constant

$$K(n, p) = \kappa_n(p, p^*, s) \quad (s \text{ is immaterial since } \theta = 1)$$

was independently determined by Aubin [28, 29] and Talenti [30]:

$$K(n,p) = \begin{cases} \frac{1}{n}\left(\frac{n}{\tilde{\omega}_{n-1}}\right)^{\frac{1}{n}} & \text{if } p=1 \\ n^{-\frac{1}{p}}\left(\frac{p-1}{n-p}\right)^{1-\frac{1}{p}}\left(\frac{\Gamma(n+1)}{\Gamma\left(\frac{n}{p}\right)\Gamma\left(n+1-\frac{n}{p}\right)\tilde{\omega}_{n-1}}\right)^{\frac{1}{n}} & \text{if } 1<p<n, \end{cases}$$

$$= \begin{cases} \frac{1}{n}\left(\frac{1}{\omega_n}\right)^{\frac{1}{n}} & \text{if } p=1 \\ n^{-\frac{1}{p}}\left(\frac{p-1}{n-p}\right)^{1-\frac{1}{p}}\left(\frac{\Gamma(n)}{\Gamma\left(\frac{n}{p}\right)\Gamma\left(n+1-\frac{n}{p}\right)\omega_n}\right)^{\frac{1}{n}} & \text{if } 1<p<n. \end{cases}$$

Note the special case

$$K(n,2) = \sqrt{\frac{4}{n(n-2)\tilde{\omega}_n^{2/n}}} = (\pi n(n-2))^{-\frac{1}{2}}\left(\frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)}\right)^{\frac{1}{n}},$$

which arises frequently in applications [31, 32]. Only the case $p=1$ was discussed in [33].

As an aside, let $p^{\#} = pn/(n-2p)$. The best constant in the second-order Sobolev inequality

$$\left(\int_{\mathbb{R}^n} |f(x)|^{2^{\#}}\,dx\right)^{\frac{1}{2^{\#}}} \le M\left(\int_{\mathbb{R}^n} |\triangle f(x)|^2\,dx\right)^{\frac{1}{2}}$$

is known to be [34, 35]

$$M(n) = \sqrt{\frac{16}{n(n-4)(n^2-4)\tilde{\omega}_{n+1}^{4/n}}} = \left(\pi^2 n(n-4)(n^2-4)\right)^{-\frac{1}{2}}\left(\frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)}\right)^{\frac{2}{n}}.$$

The similarity between $K(n,2,)$ and $M(n)$ is interesting: The former involves ∇f while the latter involves $\triangle f$. We wonder about the p-generalization of the latter.

2.3.4 *Trudinger–Moser*

A limiting scenario (as $p \to n^-$) of the Sobolev inequality is as follows. Let D denote a bounded open domain with smooth boundary in $\mathbb{R}^n$; for example, let D be an open ball. Let $|D|$ denote the Lebesgue measure of D. Consider all smooth, compactly supported functions $f: D \to \mathbb{R}$ with the property that ∇f is n-integrable and

$$\int_D |\nabla f(x)|^n\,dx \le 1.$$

Then there exists a constant c_n depending only on n (and not on D) such that [36, 37]

$$\frac{1}{|D|}\int_D \exp\left(\alpha\cdot|f(x)|^{n/(n-1)}\right)\,dx \le c_n$$

for any value $\alpha \le n\tilde{\omega}_{n-1}^{1/(n-1)}$. Further, if α exceeds the indicated threshold, then the left-hand side can be made arbitrarily large by appropriate choice of $f(x)$.

Carleson & Chang [38] obtained that $c_2 = 4.3556...$ (with computational help by Gamelin). In principle, accurate estimates of c_n are possible, but no one appears to have done this. Variations and elaborations of the fascinating Trudinger–Moser inequality are found in [39–44].

[1] J. Nash, Continuity of solutions of parabolic and elliptic equations, *Amer. J. Math.* 80 (1958) 931–954; MR0100158 (20 #6592).

[2] E. A. Carlen and M. Loss, Sharp constant in Nash's inequality, *Internat. Math. Res. Notices* (1993) 213–215; MR1230297 (95f:26022).

[3] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.

[4] W. Beckner, Geometric proof of Nash's inequality, *Internat. Math. Res. Notices* (1998) 67–71; MR1604804 (99c:35024).

[5] W. Beckner, Asymptotic estimates for Gagliardo-Nirenberg embedding constants, *Potential Anal.* 17 (2002) 253–266; MR1917809 (2003f:35009).

[6] O. Druet, E. Hebey and M. Vaugon, Optimal Nash's inequalities on Riemannian manifolds: the influence of geometry, *Internat. Math. Res. Notices* (1999) 735–779; MR1704184 (2000j:53044).

[7] E. Humbert, Best constants in the L^2-Nash inequality, *Proc. Roy. Soc. Edinburgh Sect. A* 131 (2001) 621–646; MR1838504 (2002i:58050).

[8] J. Moser, On Harnack's theorem for elliptic differential equations, *Comm. Pure Appl. Math.* 14 (1961) 577–591; MR0159138 (28 #2356).

[9] J. Moser, A Harnack inequality for parabolic differential equations, *Comm. Pure Appl. Math.* 17 (1964) 101–134; correction 20 (1967) 231–236; MR0159139 (28 #2357).

[10] B. v. Sz. Nagy, Über Integralungleichungen zwischen einer Funktion und ihrer Ableitung, *Acta Univ. Szeged. Sect. Sci. Math.* 10 (1941) 64–74; MR0004277 (2,351b).

[11] E. Gagliardo, Proprietà di alcune classi di funzioni in più variabili, *Ricerche Mat.* 7 (1958) 102–137; MR0102740 (21 #1526).

[12] L. Nirenberg, On elliptic partial differential equations, *Annali Scuola Norm. Sup. Pisa* 13 (1959) 115–162; MR0109940 (22 #823).

[13] C. Brouttelande, The best-constant problem for a family of Gagliardo-Nirenberg inequalities on a compact Riemannian manifold, *Proc. Edinburgh Math. Soc.* 46 (2003) 117–146; MR1961176 (2004b:58025).

[14] M. Del Pino and J. Dolbeault, Best constants for Gagliardo-Nirenberg inequalities and applications to nonlinear diffusions, *J. Math. Pures Appl.* 81 (2002) 847–875; MR1940370 (2003h:35051).

[15] M. Del Pino and J. Dolbeault, The optimal Euclidean L^p-Sobolev logarithmic inequality, *J. Funct. Anal.* 197 (2003) 151–161; MR1957678 (2004a:35091).

[16] D. Cordero-Erausquin, B. Nazaret and C. Villani, A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities, *Adv. Math.* 182 (2004) 307–332; MR2032031 (2005b:26023).

[17] M. Agueh, N. Ghoussoub and X. Kang, Geometric inequalities via a general comparison principle for interacting gases, *Geom. Funct. Anal.* 14 (2004) 215–244; math.AP/0310234; MR2053603 (2005c:82050).

[18] J. Gunson, Inequalities in mathematical physics, *Inequalities: Fifty years on from Hardy, Littlewood and Pólya*, Proc. 1987 Birmingham conf., ed. W. N. Everitt, Dekker, 1991, pp. 53–79; MR1112572 (92h:82002).

[19] H. A. Levine, An estimate for the best constant in a Sobolev inequality involving three integral norms, *Annali Mat. Pura Appl.* 124 (1980) 181–197; MR0591555 (81m:26009).

[20] M. I. Weinstein, Nonlinear Schrödinger equations and sharp interpolation estimates, *Comm. Math. Phys.* 87 (1982/83) 567–576; MR0691044 (84d:35140).

[21] S. M. Nasibov, Optimal constants in some Sobolev inequalities and their applications to the nonlinear Schrödinger equation (in Russian), *Doklady Akad. Nauk SSSR* 307 (1989) 538–542; Engl. transl. in *Soviet Math. Dokl.* 40 (1990) 110–115; MR1020379 (90m:35051).

[22] S. M. Nasibov, On the exact constant in a Sobolev-Nirenberg inequality and its application to a lower bound for the blow-up time of the nonlinear evolution Schrödinger equation with a critical power (in Russian), *Doklady Akad. Nauk* 382 (2002) 750–753; Engl. transl. in *Doklady Math.* 65 (2002) 135–138; MR1932214 (2003g:35210).

[23] E. J. M. Veling, Lower bounds for the infimum of the spectrum of the Schrödinger operator in $\mathbb{R}^N$ and the Sobolev inequalities, *J. Inequal. Pure Appl. Math.* 3 (2002) article 63; corrigendum 4 (2003) article 109; MR1923362 (2003g:35039).

[24] P.-L. Sulem, C. Sulem and A. Patera, Numerical simulation of singular solutions to the two-dimensional cubic Schrödinger equation, *Comm. Pure Appl. Math.* 37 (1984) 755–778; MR0762872 (86e:65145).

[25] X. Chen, Exponential asymptotics and law of the iterated logarithm for intersection local times of random walks, *Annals of Probab.* 32 (2004) 3248–3300; MR2094445 (2005m:60174).

[26] R. F. Bass and X. Chen, Self-intersection local time: critical exponent, large deviations, and laws of the iterated logarithm, *Annals of Probab.* 32 (2004) 3221–3247; math.PR/0503592; MR2094444 (2005i:60149).

[27] J.-F. Le Gall, Some properties of planar Brownian motion, *École d'Été de Probabilités de Saint-Flour XX—1990*, Lecture Notes in Math. 1527, Springer-Verlag, 1992, pp. 111–235; MR1229519 (94g:60156).

[28] T. Aubin, Problèmes isopérimétriques et espaces de Sobolev, *J. Differential Geometry* 11 (1976) 573–598; MR0448404 (56 #6711).

[29] T. Aubin, *Nonlinear Analysis on Manifolds: Monge-Ampère Equations*, Springer-Verlag, 1982, pp. 39–40; MR0681859 (85j:58002).

[30] G. Talenti, Best constant in Sobolev inequality, *Annali Mat. Pura Appl.* 110 (1976) 353–372; MR0463908 (57 #3846).

[31] O. Druet, The best constants problem in Sobolev inequalities, *Math. Annalen* 314 (1999) 327–346; MR1697448 (2000d:58033).

[32] E. Lutwak, D. Yang and G. Zhang, Sharp affine L_p Sobolev inequalities, *J. Differential Geom.* 62 (2002) 17–38; MR1987375 (2004d:46039).

[33] S. R. Finch, Sobolev isoperimetric constants, *first volume*, pp. 219–225.

[34] D. E. Edmunds, D. Fortunato and E. Jannelli, Critical exponents, critical dimensions and the biharmonic operator, *Arch. Rational Mech. Anal.* 112 (1990) 269–289; MR1076074 (91k:35191).

[35] E. H. Lieb, Sharp constants in the Hardy-Littlewood-Sobolev and related inequalities, *Annals of Math*. 118 (1983) 349–374; MR0717827 (86i:42010).
[36] N. S. Trudinger, On imbeddings into Orlicz spaces and some applications, *J. Math. Mech*. 17 (1967) 473–483; MR0216286 (35 #7121).
[37] J. Moser, A sharp form of an inequality by N. Trudinger, *Indiana Univ. Math. J*. 20 (1970/71) 1077–1092; MR0301504 (46 #662).
[38] L. Carleson and S.-Y. A. Chang, On the existence of an extremal function for an inequality of J. Moser, *Bull. Sci. Math*. 110 (1986) 113–127; MR0878016 (88f:46070).
[39] D. R. Adams, A sharp inequality of J. Moser for higher order derivatives, *Annals of Math*. 128 (1988) 385–398; MR0960950 (89i:46034).
[40] K.-C. Lin, Extremal functions for Moser's inequality, *Trans. Amer. Math. Soc*. 348 (1996) 2663–2671; MR1333394 (96i:58043).
[41] B. Ruf, On a result by Carleson-Chang concerning the Trudinger-Moser inequality, *Nonlinear Anal*. 47 (2001) 6041–6051; MR1970776 (2004b:46040).
[42] D. G. de Figueiredo, J. M. do Ó and B. Ruf, On an inequality by N. Trudinger and J. Moser and related elliptic equations, *Comm. Pure Appl. Math.* 55 (2002) 135–152; MR1865413 (2002j:35104).
[43] M. Jodeit, An inequality for the indefinite integral of a function in L^q, *Studia Math*. 44 (1972) 545–554; MR0341055 (49 #5805).
[44] S. Hudson and M. Leckband, Extremals for a Moser-Jodeit exponential inequality, *Pacific J. Math*. 206 (2002) 113–128; MR1924821 (2003g:46033).

2.4 Uncertainty Inequalities

If an integrable function $f: \mathbb{R}^n \to \mathbb{R}$ is thought of as the amplitude of a time signal or space image, then the Fourier transform $\hat{f}$ of f:

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-2\pi i\xi\cdot x} f(x)\, dx$$

conveys information on how $f(x)$ is built from sine waves of different frequencies. Assume that $f \in L_r(\mathbb{R}^n)$ for some $r \geq 1$; equivalently, $|f(x)|^r$ is integrable and decays rapidly enough as $|x| \to \infty$ so that

$$||f||_r = \left(\int_{\mathbb{R}^n} |f(x)|^r\, dx \right)^{\frac{1}{r}} < \infty.$$

Define $P_p f$ and $Q_q f$ to be the functions

$$(P_p f)(x) = |x|^p f(x), \quad (Q_q f)(\xi) = |\xi|^q \hat{f}(\xi).$$

Heisenberg's famous inequality arises from the case when $p = q = 1$ and $r = 2$ [1]:

$$||P_1 f||_2 \cdot ||Q_1 f||_2 \geq \frac{n}{4\pi} ||f||_2^2.$$

In words, if $f(x)$ is concentrated close to 0 (having a small variance), then $\hat{f}(\xi)$ must be relatively spread out (having a large variance) unless $f(x)$ is zero almost everywhere. The constant $n/(4\pi)$ is best possible if $n=1$: consider functions of the form $a\exp(-bx^2)$ for some $b>0$ [2].

When f is smooth, it follows that $||\nabla f||_2 = 2\pi||Q_1 f||_2$ where ∇f is the gradient of f and $|\nabla f|$ is its Euclidean norm. Therefore Heisenberg's inequality is an uncertainty principle in the same sense as expressed in [3].

Here are two sample variations [4, 5]. Let $f \in L_1(\mathbb{R}^n) \cap L_2(\mathbb{R}^n)$ and recall that J_ν is the Bessel function of the first kind [6]. For $r>0$, define

$$J(r) = r^{1-\frac{n}{2}} J_{\frac{n}{2}-1}(r)$$

and, for $y>0$,

$$g_y(x) = \begin{cases} J(|x|) - J(y) + \dfrac{J'(y)}{2y}\left(y^2 - |x|^2\right) & \text{if } |x| < y, \\ 0 & \text{if } |x| \geq y. \end{cases}$$

The best constant μ_n in the inequality

$$||P_2 f||_1 \cdot ||Q_1 f||_2^2 \geq \frac{\mu_n}{4\pi^2}||f||_1 ||f||_2^2$$

is achieved when $f = g_c$, where c is the smallest positive root of the equation

$$||g_y||_2 = ||\nabla g_y||_2.$$

In particular, if $n=1$, the equation simplifies to

$$y(5-2y^2)\tan(y)^2 + 5(3-2y^2)\tan(y) - 15y = 0$$

and hence $c = 1.7502456171...$ and $\mu_n = 0.4283683675... = \frac{1}{2}(0.8567367350...) = \frac{M}{2}$. The constant M will be useful to us later.

Also, the best constant μ_n in the inequality

$$||P_2 f||_1^{\frac{2}{n+6}} \cdot ||Q_1 f||_2^{\frac{n+4}{n+6}} \geq \mu_n ||f||_2$$

is achieved when $f = g_c$, where c is the smallest positive root of the equation

$$\sqrt{n+4}\,||g_y||_2 = \sqrt{n+6}\,||\nabla g_y||_2, \quad \text{that is,} \quad (y^2 - 2n)J'(y) = 2yJ(y).$$

In particular, if $n=1$ (and thus the two exponents are $2/7$ and $5/7$), the equation simplifies to

$$(2-y^2)\tan(y) = 2y$$

and hence $c = 2.0815759778$ and $\mu_n^{-1} = 4.1731026567....$ Closed-form expressions do not seem to be possible here! This formulation is, in fact, only a special case of a considerably broader theorem [5].

2.4.1 *Positive Definite Probability Densities*

A probability density function $f\colon \mathbb{R}^n \to \mathbb{R}$ is **positive definite** if [7, 8]

$$\sum_{j=1}^{m}\sum_{k=1}^{m} f(x_k - x_j) z_j \bar{z}_k \geq 0$$

for all $x_j \in \mathbb{R}^n$, for all $z_j \in \mathbb{C}$ $(j = 1, \ldots, n)$ and for each $m \geq 1$, where $\bar{z}$ denotes the complex conjugate of z. Clearly $f(-x) = f(x) < f(0)$ for all $x \neq 0$. Let F_n denote the class of all continuous, positive definite probability density functions on $\mathbb{R}^n$. If $f \in F_n$, then $\hat{f}$ is nonnegative and integrable over $\mathbb{R}^n$; in fact, $\hat{f}/f(0)$ is itself a probability density.

Fix, for now, $n = 1$. Among the well-known members of F_1 are the normal, t, and logistic densities. Define a product of variances

$$\lambda(f) = 4\pi^2 \frac{||P_2 f||_1 \cdot ||Q_2 f||_1}{\hat{f}(0) \cdot f(0)}$$

and a greatest lower bound, called **Laue's constant** [8]:

$$\Lambda = \inf_{f \in F_1} \lambda(f).$$

An immediate consequence of Laeng & Morpurgo's work [4], for example, is that $\Lambda \leq M < 0.85674$. Estimating Λ has occupied several researchers over several years [9–12]:

$$0.543 < \Lambda < 0.85024$$

yet a determination of its exact value still seems far away.

For $n \geq 1$, choose an arbitrary unit vector $u \in \mathbb{R}^n$. If X is a random n-vector with density $f \in F_n$, let $f_u \in F_1$ denote the density for the one-dimensional projection $u \cdot X$ of X onto u. Then define [11]

$$\Lambda_n = \inf_{f \in F_n} \sup_{||u||=1} \lambda(f_u).$$

Clearly $\Lambda_1 = \Lambda$ and $\Lambda_{n+1} \geq \Lambda_n$ for all n. We have the following estimates [12]:

$$\begin{aligned} \Lambda_n &\leq \frac{1}{2} \frac{9 + 4\sqrt{5}}{(1+\sqrt{5})^2} < 0.856763... \quad && \text{if } n \leq 7, \\ 1 - \frac{3}{n} \leq \Lambda_n &\leq 1 - \frac{n-5}{2(n-4)} \frac{3}{n} && \text{if } n \geq 8, \end{aligned}$$

which demonstrate that $\lim_{n \to \infty} \Lambda_n = 1$.

2.4.2 *Fourier Optimization*

We mention the optimization problem [13–15]:

$$C = \sup_{0 \neq f \in \mathcal{E}_1} \frac{|f(0)|}{||f||_1}$$

where $\mathcal{E}_r$ is the set of all continuous $f : \mathbb{R} \to \mathbb{R}$ with $f \in L_r(\mathbb{R})$ and support $(\hat{f}) \subseteq [-1, 1]$. Solving such problems with band-limited functions is difficult; we know that there exists an even $g \in \mathcal{E}_1$ with $g(0) = 1$ that maximizes the ratio and that $1.08185 \leq C \leq 1.09769$. If f was further assumed to be nonnegative, then the problem would simplify and the Fejér kernel

$$g(x) = \left(\frac{\sin(\pi x)}{\pi x} \right)^2$$

would emerge, giving $C_{\text{nonneg}} = 1$. Analogous issues for 1-periodic trigonometric polynomials are studied in [14, 16, 17]. For example, if $f : [-1/2, 1/2] \to \mathbb{R}$ is the first-order expression

$$f(x) = a_{-1}\, e^{-2\pi i x} + a_0 + a_1\, e^{2\pi i x},$$

then $C_1 = \pi/(2\omega) = 2.1253252923...$, where $\omega = 0.7390851332...$ is the unique root of $\cos(\omega) = \omega$. An exact C_2 formula is not known for the second-order expression

$$f(x) = a_{-2}\, e^{-4\pi i x} + a_{-1}\, e^{-2\pi i x} + a_0 + a_1\, e^{2\pi i x} + a_2\, e^{4\pi i x},$$

nor for higher ℓ^{th} orders; however, $\lim_{\ell \to \infty} C_\ell/\ell = C$. We also note [18, 19]

$$D = \sup_{0 \neq f \in \mathcal{E}_2} \frac{||f||_4}{||f||_2} = \left(\frac{0.6869812930...}{\pi} \right)^{1/4}$$

and again the absence of familiar functions. These constants deserve to be better known!

[1] G. B. Folland and A. Sitaram, The uncertainty principle: a mathematical survey, *J. Fourier Anal. Appl.* 3 (1997) 207–238; MR1448337 (98f:42006).

[2] H. Dym and H. P. McKean, *Fourier Series and Integrals*, Academic Press, 1972, pp. 116–121; MR0442564 (56 #945).

[3] S. R. Finch, Sobolev isoperimetric constants, *first volume*, pp. 219–225.

[4] E. Laeng and C. Morpurgo, An uncertainty inequality involving L^1 norms, *Proc. Amer. Math. Soc.* 127 (1999) 3565–3572; MR1621969 (2000b:42006).

[5] C. Morpurgo, Extremals of some uncertainty inequalities, *Bull. London Math. Soc.* 33 (2001) 52–58; MR1798575 (2001m:26062).

[6] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.

[7] J. Stewart, Positive definite functions and generalizations, an historical survey, *Rocky Mountain J. Math.* 6 (1976) 409–434; MR0430674 (55 #3679).

[8] H.-J. Rossberg, Positive definite probability densities and probability distributions, *J. Math. Sci.* 76 (1995) 2181–2197; MR1356657 (96h:60027).

[9] I. Dreier, On the uncertainty principle for positive definite densities, *Z. Anal. Anwendungen* 15 (1996) 1015–1023; MR1422654 (98c:60014).

[10] T. Gneiting, On the uncertainty relation for positive definite probability densities, *Statistics* 31 (1998) 83–88; MR1625593 (99b:60019).

[11] W. Ehm, T. Gneiting and D. Richards, On the uncertainty relation for positive-definite probability densities. II, *Statistics* 33 (1999) 267–286; MR1750252 (2000m:60011).
[12] I. Dreier, W. Ehm, T. Gneiting and D. Richards, Improved bounds for Laue's constant and multivariate extensions, *Math. Nachr*. 228 (2001) 109–122. MR1845909 (2002e:60021).
[13] N. N. Andreev, S. V. Konyagin and A. Y. Popov, Extremal problems for functions with small support (in Russian), *Mat. Zametki* 60 (1996) 323–332, 478; Engl. transl. in *Math. Notes* 60 (1996) 241–247; MR1428847 (98g:42002).
[14] D. V. Gorbachev, An integral problem of Konyagin and the (C,L)-constants of Nikolskii, *Proc. Steklov Inst. Math.* 2005*, Function Theory*, suppl. 2, S117–S138; MR2200225 (2006k:41062).
[15] E. Carneiro, M. B. Milinovich and K. Soundararajan, Fourier optimization and prime gaps, arXiv:1708.04122.
[16] L. V. Taĭkov, A group of extremal problems for trigonometric polynomials, *Uspehi Mat. Nauk* 20 (1965) 205–211; MR0185361 (32 #2829).
[17] D. V. Gorbachev, Sharpening the Taĭkov lower bound in the inequality between the *C*- and *L*-norms for trigonometric polynomials (in Russian); *Mat. Zametki* 74 (2003) 132–134; Engl. transl. in *Math. Notes* 74 (2003) 123–126; MR2010684 (2004h:41010).
[18] A. Garsia, E. Rodemich and H. Rumsey, On some extremal positive definite functions, *J. Math. Mech.* 18 (1968/1969) 805–834; MR0251454 (40 #4682).
[19] M. I. Ganzburg and S. Y. Tikhonov, On sharp constants in Bernstein-Nikolskii inequalities, *Constr. Approx.* 45 (2017) 449–466; MR3640642.

2.5 Airy Function Zeroes

On the negative real axis ($x<0$), the Airy function

$$\mathrm{Ai}(x)=\frac{1}{3}(-x)^{1/2}\left[J_{-\frac{1}{3}}\left(\tfrac{2}{3}(-x)^{3/2}\right)+J_{\frac{1}{3}}\left(\tfrac{2}{3}(-x)^{3/2}\right)\right]$$

has an oscillatory behavior similar to that of the Bessel function $J_\nu(x)$ [1]. Note the special values [2]

$$\mathrm{Ai}(0)=\frac{1}{3^{\frac{2}{3}}\Gamma\left(\frac{2}{3}\right)}=0.3550280538..., \quad \mathrm{Ai}'(0)=-\frac{1}{3^{\frac{1}{3}}\Gamma\left(\frac{1}{3}\right)}=-0.2588194037...$$

and the integral representations

$$\mathrm{Ai}(x)=\frac{1}{\pi}\int_0^\infty\cos\left(\frac{1}{3}t^3+x\,t\right)dt, \quad \mathrm{Ai}'(x)=-\frac{1}{\pi}\int_0^\infty t\sin\left(\frac{1}{3}t^3+x\,t\right)dt.$$

Let $0<a_1<a_2<\ldots$ be the zeroes of $\mathrm{Ai}(-x)$ and $0<a_1'<a_2'<\ldots$ be the zeroes of $\mathrm{Ai}'(-x)$. See Table 2.8 for the first several terms of both sequences. We saw these values when bounding the zeroes of $J_\nu(x)$ [1] and we will see them again when estimating the L_1-norm of Brownian motion [3]. In the present essay, our focus is on two applications to physics.

Table 2.8 *Negatives of zeroes of* Ai *and* Ai′ *for* $n = 1, 2, 3, 4, 5$

a_n	a'_n
2.3381074104...	1.0187929716...
4.0879494441...	3.2481975821...
5.5205598280...	4.8200992111...
6.7867080900...	6.1633073556...
7.9441335871...	7.3721772550...

2.5.1 *Quantum Mechanics of Falling*

Consider a quantum mechanical (QM) particle in free fall, that is, on the positive x-axis with linear potential x. The time-independent Schrödinger equation becomes

$$\frac{d^2f}{dx^2} + (\lambda - x)f = 0, \quad \lim_{x\to\infty} f(x) = 0.$$

If a Dirichlet condition $f(0) = 0$ is imposed (elastic reflection), then the eigenvalues λ are the Airy function zeroes $\{a_n\}_{n=1}^{\infty}$ [4–9]. If instead a Neumann condition $f'(0) = 0$ is imposed, then the eigenvalues λ are the derivative zeroes $\{a'_n\}_{n=1}^{\infty}$ [10, 11].

What is the physical significance of these results? The eigenfunctions f contain information about the behavior of the particle, for example, the probability densities of position and momentum. Admissible solutions to the time-independent Schrödinger equation exist only if the total energy of the particle is quantized, that is, restricted to a discrete set of eigenvalues λ. (This counterintuitive fact is akin to Bohr's model of the hydrogen atom possessing discrete shells for the electron to occupy, as indicated by spectroscopy.) Different boundary conditions or different potentials, of course, lead to different allowed energy levels.

Consider rather a QM particle on the whole x-axis with the potential $|x|$. Then the eigenvalues corresponding to even eigenfunctions come from $\{a'_n\}$ and the eigenvalues corresponding to odd eigenfunctions come from $\{a_n\}$ [10, 11]. A listing of the eigenvalues λ consists of the interlaced zeroes of Ai′ and Ai. It is remarkable that the Airy function zeroes occur here, in the QM analog of the simplest of all classical physics problems.

2.5.2 *Van der Pol's Equation*

For constant $\mu > 0$, all solutions of van der Pol's equation

$$\frac{d^2g}{dt^2} + \mu(g^2 - 1)\frac{dg}{dt} + g = 0,$$

other than the trivial solution $g=0$, tend to a unique periodic limit cycle as $t\to\infty$. The proof of this theorem is due to Liénard [12]. We are interested in how the magnitude $A(\mu)$ and the period $T(\mu)$ of the limit cycle vary with increasing μ.

Let $\alpha=a_1=2.3381074104...$ for convenience. The work of Haag [13], Dorodnicyn [14] and others [15–20] gives

$$\begin{aligned}A(\mu) &= 2+\frac{1}{3}\alpha\mu^{-4/3}-\frac{16}{27}\mu^{-2}\ln(\mu)\\&\quad+\frac{1}{9}\left(3\beta+2\ln(2)-8\ln(3)-1\right)\mu^{-2}+O\left(\mu^{-8/3}\right),\end{aligned}$$

$$\begin{aligned}T(\mu) &= (3-2\ln(2))\mu+3\alpha\mu^{-1/3}-\frac{2}{3}\mu^{-1}\ln(\mu)\\&\quad+\left(3\beta+\ln(2)-\ln(3\pi)-2\ln\left(\mathrm{Ai}'(-\alpha)\right)-1\right)\mu^{-1}+O\left(\mu^{-4/3}\ln(\mu)\right)\end{aligned}$$

as $\mu\to\infty$, where $\beta=0.17234...$ is defined as follows. The function $-\mathrm{Ai}'(x)/\mathrm{Ai}(x)$ maps the interval $(-\alpha,\infty)$ onto $(-\infty,\infty)$ in a one-to-one fashion; let $z(x)$ denote its inverse. Define $Q(x)=x^2-z(x)$ and

$$P(x)=\exp\left(-\int_0^x\frac{1}{Q(u)^2}du\right).$$

Then the expression

$$\frac{1}{P(x)}\int_x^\infty P(v)\left\{\frac{v}{Q(v)}-\frac{v^3}{3Q(v)^2}-\frac{2v}{3(v^2+\alpha/2)}+\frac{\ln(v^2+\alpha/2)}{3Q(v)^2}\right\}dv$$

approaches β as $x\to-\infty$. Hence, for example, we have the asymptotic expression

$$\begin{aligned}T(\mu) &\sim (1.613705...)\mu+(7.014322...)\mu^{-1/3}-(0.666666...)\mu^{-1}\ln(\mu)\\&\quad-(1.3232...)\mu^{-1}.\end{aligned}$$

The final coefficient for $T(\mu)$ is sometimes written as $3\beta+3\ln(2)-\ln(3)-1-2\iota$ or as $\beta+3\ln(2)-\ln(3)-3/2-2\delta$, where

$$\begin{aligned}&\iota=\ln(2)+\tfrac{1}{2}\ln(\pi)+\ln\left(\mathrm{Ai}'(-\alpha)\right)=0.9105654320...,\\&\delta=-\beta+\iota-\tfrac{1}{4}=0.4882....\end{aligned}$$

Two additional terms in the series for $A(\mu)$ were determined by Bavinck & Grasman [20, 21]; we omit these for reasons of space. Early textbooks [22, 23] often repeat errors originating in [14]; the final two coefficients for $T(\mu)$ are mistakenly given as $-22/9$ and $+0.0087$.

A relevant theory of special functions arose in [24–26]. For example, the Haag function $\mathrm{Hg}(x)$ is defined to be what we call $-z(-x)$; thus $\mathrm{Hg}(0)=a_1'$,

$\lim_{x\to\infty} \mathrm{Hg}(x) = a_1$ and

$$\frac{d}{dx}\,\mathrm{Hg}(x) = \frac{1}{x^2 + \mathrm{Hg}(x)}, \quad \lim_{x\to-\infty} \frac{\mathrm{Hg}(x)}{x^2} = -1.$$

The Dorodnicyn function $\mathrm{Dn}(x)$ satisfies

$$\frac{d}{dx}\,\mathrm{Dn}(x) = -\frac{\mathrm{Dn}(x)}{(x^2 + \mathrm{Hg}(x))^2} + \frac{x}{x^2 + \mathrm{Hg}(x)}, \quad \lim_{x\to-\infty} \mathrm{Dn}(x) = -\frac{1}{2}$$

as well as

$$\lim_{x\to\infty} (\mathrm{Dn}(x) - \ln(x)) = -\frac{3}{2}\beta - \frac{1}{4} = -0.50851....$$

Clearly Hg has a unique zero at $-3^{1/3}\Gamma(2/3)/\Gamma(1/3) = -0.7290111329...$; a similar exact expression for the unique zero 0.8452... of Dn is not known.

[1] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.
[2] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, pp. 446–450; 475–478; MR1225604 (94b:00012).
[3] S. R. Finch, Variants of Brownian motion, *this volume*, §4.5.
[4] S. Flügge, *Practical Quantum Mechanics*, Springer-Verlag, 1974, pp. 101–105; MR0366248 (51 #2496).
[5] R. G. Winter, *Quantum Physics*, 2nd ed., Faculty Publishing, 1986, pp. 92–97; MR0620164 (82i:81003).
[6] P. W. Langhoff, Schrödinger particle in a gravitational well, *Amer. J. Phys.* 39 (1971) 954–957.
[7] R. L. Gibbs, The quantum bouncer, *Amer. J. Phys.* 43 (1975) 25–28; comment 51 (1983) 82–84.
[8] V. C. Aguilera-Navarro, H. Iwamoto, E. Ley-Koo and A. H. Zimerman, Quantum bouncer in a closed court, *Amer. J. Phys.* 49 (1981) 648–651; comment 51 (1983) 84–85.
[9] J. Gea-Banacloche, A quantum bouncing ball, *Amer. J. Phys.* 67 (1999) 776–782; comments 68 (2000) 672–673 and 68 (2000) 866–868.
[10] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, 1978, pp. 28, 521–522; MR0538168 (80d:00030).
[11] M. S. Ashbaugh and J. D. Morgan, Remarks on Turschner's eigenvalue formula, *J. Phys. A* 14 (1981) 809–819; MR0609826 (82d:81036).
[12] G. F. Simmons, *Differential Equations with Applications and Historical Notes*, McGraw-Hill, 1972, pp. 338–244, 349–352; MR0499365 (58 #17258).
[13] J. Haag, Exemples concrets d'étude asymptotique d'oscillations de relaxation, *Annales Sci. École Norm. Sup.* 61 (1944) 73–117; MR0014539 (7,299d).
[14] A. A. Dorodnicyn, Asymptotic solution of van der Pol's equation (in Russian), *Akad. Nauk SSSR. Prikl. Mat. Mech.* 11 (1947) 313–328; Engl. transl. in *Amer. Math. Soc. Transl.* 88 (1953) 1–24; MR0022011 (9,144g).
[15] M. L. Cartwright, Van der Pol's equation for relaxation oscillations, *Contributions to the Theory of Nonlinear Oscillations*, v. II, Princeton Univ. Press, 1952, pp. 3–18; MR0052617 (14,647a).

[16] W. S. Krogdahl, Numerical solutions of the Van der Pol equation, *Z. Angew. Math. Phys.* 11 (1960) 59–63; MR0111149 (22 #2013).

[17] M. Urabe, Periodic solutions of van der Pol's equation with damping coefficient $\lambda = 0 \sim 10$, *IEEE Trans. Circuit Theory* CT-7 (1960) 382–386; MR0120095 (22 #10852).

[18] P. J. Ponzo and N. Wax, On the periodic solution of the van der Pol equation, *IEEE Trans. Circuit Theory* CT-12 (1965) 135–136.

[19] J. A. Zonneveld, Periodic solutions of the Van der Pol equation, *Nederl. Akad. Wetensch. Proc. Ser. A* 69 (1966) 620–622; *Indag. Math*. 28 (1966) 620–622; MR0226864 (37 #2450).

[20] H. Bavinck and J. Grasman, The method of matched asymptotic expansions for the periodic solution of the van der Pol equation, *Internat. J. Nonlinear Mechanics* 9 (1974) 421–434; *Toegepaste Wiskunde* TN 73/73.

[21] J. Grasman, *Asymptotic Methods for Relaxation Oscillations and Applications*, Springer-Verlag, 1987, pp. 55–72; MR0884527 (88i:34001).

[22] J. J. Stokes, *Nonlinear Vibrations in Mechanical and Electrical Systems*, Interscience, 1950, pp. 140–141.

[23] H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, Dover, 1962, pp. 358–368; MR0181773 (31 #6000).

[24] M. I. Zharov, E. F. Mishchenko and N. K. Rozov, Some special functions and constants that arise in the theory of relaxation oscillations (in Russian), *Dokl. Akad. Nauk SSSR* 261 (1981) 1292–1296; Engl. transl. in *Soviet Math. Dokl.* 24 (1981) 672–675; MR0640839 (83e:34080).

[25] M. K. Kerimov, In memory of Anatoliĭ Alekseevich Dorodnitsyn (in Russian), *Zh. Vychisl. Mat. i Mat. Fiz.* 35 (1995) 819–842; Engl. transl. in *Comput. Math. Math. Phys.* 35 (1995) 649–666; MR1342522 (96g:01039).

[26] M. K. Kerimov, Special functions that arise in the theory of nonlinear oscillations (in Russian), *Zh. Vychisl. Mat. i Mat. Fiz.* 36 (1996) 57–72; Engl. transl. in *Comput. Math. Math. Phys.* 36 (1996) 1027–1039 (1997); MR1407727 (98g:34048).

2.6 Projections of Minimal Norm

Let X be a real Banach space and Y be a closed subspace of X. A continuous linear operator $P : X \to Y$ is a **projection** if $P(y) = y$ for all $y \in Y$. The **norm** of P is defined by

$$\|P\| = \sup_{\|x\| \le 1} \|P(x)\| = \sup_{x \ne 0} \frac{\|P(x)\|}{\|x\|}.$$

Out of all such projections (for fixed X and Y), which ones have the smallest possible norm? [1–4] We will answer this question for the special scenario when

$$X = \{\text{admissible functions } x : [-1,1] \to \mathbb{R}\},$$

$$Y = \{\text{real polynomials of degree} \le n\},$$

$$\|x\| = \left(\int_{-1}^{1} |x(t)|^q dt \right)^{1/q},$$

and $n \in \{0,1,2,3,4,5\}$, $q \in \{1,\infty\}$. We understand the word "admissible" to mean "continuous" if $q = \infty$ and "Lebesgue integrable" if $q = 1$. The minimal norm $\|P\|$, considered over all $P: X \to Y$, will be denoted by $\pi(n,q)$ and is called the **relative projection constant** of $Y \subseteq X$.

It is known that $\pi(0,\infty) = \pi(1,\infty) = \pi(0,1) = 1$, which is demonstrated by taking $(Px)(t)$ to be [5]

$$x(0), \quad \tfrac{1}{2}\left((1-t)\,x(-1) + (1+t)\,x(1)\right), \quad \int_{-1}^{1} x(s)\,ds,$$

respectively. The other cases are far more difficult. Franchetti & Cheney [6] proved that $\pi(1,1) = 1.2204049171... = 1 - \varphi^2 + \varphi$, where $\varphi = 0.3279677853...$ satisfies the equation

$$2\varphi\left(1 - \varphi^2 + \varphi\right)\ln(\varphi) + 1 - \varphi^2 = 0.$$

The corresponding projection is unique:

$$(Px)(t) = \int_{-1}^{1} x(s)\,u_1(s)\,ds + t\int_{-1}^{1} x(s)\,u_2(s)\,ds,$$

where

$$u_2(t) = \frac{1 - \varphi^2 + \varphi}{2\left(1 + \lambda^2 t^2 - \lambda\,t\sqrt{\lambda^2 t^2 + 1}\right)}, \quad u_1(t) = u_2(t)\left(-\lambda\,t + \sqrt{\lambda^2 t^2 + 1}\right)$$

for $0 \le t \le 1$ and $u_2(t) = -u_2(-t)$, $u_1(t) = u_1(-t)$ for $-1 \le t < 0$. The constant $\lambda = 1.3605560846...$ is defined to be $(1 - \varphi^2)/(2\varphi)$ or, equivalently, $-(1 - \varphi^2 + \varphi)\ln(\varphi)$.

Chalmers & Metcalf [7] proved that $\pi(2,\infty) = 1.2201730642...$ and a corresponding projection is

$$\begin{aligned}(Px)(t) \;=\; &\left(A - C\,t + D\,t^2\right)x(-1) + B\left(1 - t^2\right)x(0) + \left(A + C\,t + D\,t^2\right)x(1)\\ &+ \sum_{k=1}^{2}\left(\int_{-s_{k,2}}^{-s_{k,1}} + \int_{s_{k,1}}^{s_{k,2}}\right)\frac{(b_k + a_k|s|) + (c_k s)\,t + (-b_k + d_k|s|)\,t^2}{(1 + w_k|s|)^3}\,x(s)\,ds.\end{aligned}$$

Whether this is the unique such projection remains open. It turns out that $\pi(2,\infty) = 1 - 4A$; hence it remains to define all the parameters in the formula for P. Given $0 \le t_0 \le \sqrt{2} - 1 \le t_1 \le 1$ and $\theta < 0$, let

$$t_c = \frac{2t_1 t_0^2 + (2t_0 - 1)(1 + t_1^2) + (1 - t_1^2)(1 - t_0)\sqrt{1 - 2t_0 - t_0^2}}{t_0^2(1 + t_1^2) + 2t_1(2t_0 - 1)},$$

$$\beta = \frac{2\left(t_1 t_0(t_0 + t_1) + t_0 - t_1\right)}{(1 - t_1)^2}, \quad \delta = -\beta - t_0^2,$$

$$\kappa = \frac{t_0(2-t_0\theta^{-1})-1}{2t_0^2-\delta}+1, \quad w_1 = \frac{\kappa-\theta^{-1}}{2}, \quad w_2 = -w_1,$$

$$s_{1,1} = \frac{t_1-1}{\kappa+t_1\theta^{-1}}, \quad s_{1,2} = \frac{t_c-1}{\kappa+t_c\theta^{-1}}, \quad s_{2,1} = \frac{t_c+1}{\kappa-t_c\theta^{-1}}, \quad s_{2,2} = \frac{t_0+1}{\kappa-t_0\theta^{-1}},$$

$$I_1(\sigma,\tau) = \frac{8\theta^2(\sigma-\tau)(\tau\sigma-1)}{(\kappa\theta+1)^2(\tau+1)^2(\sigma+1)^2}, \quad I_{1,1} = I_1(t_c,t_1), \quad I_{1,2} = I_1(-t_0,-t_c),$$

$$D = \frac{1}{2(1+\delta-2t_0^2)}, \quad d_1 = \frac{-D\,t_0^2}{(1+\kappa\theta)I_{1,1}}, \quad d_2 = \frac{D\,\delta}{(1+\kappa\theta)I_{1,2}},$$

$$A = -D\,t_0^2, \quad B = \frac{2D(1-2t_0-t_0^2)}{(1-t_1)^2}, \quad C = 2D\,t_0,$$

$$b_1 = \theta\,d_1, \quad b_2 = -\theta\,d_2, \quad a_1 = \kappa\,b_1, \quad a_2 = -\kappa\,b_2,$$

$$c_1 = -(a_1+d_1), \quad c_2 = a_2+d_2,$$

$$\alpha_1 = \frac{-1}{(a_1-d_1)^3}, \quad \beta_1 = \frac{a_1}{2(a_1-d_1)}, \quad \gamma_1 = \frac{d_1}{(a_1-d_1)^2}, \quad \nu(\xi) = \frac{1}{a_1+d_1\xi},$$

$$L(\sigma,\tau) = \ln\left(\frac{1+\sigma}{1+\tau}\frac{\nu(\sigma)}{\nu(\tau)}\right), \quad L_1 = L(t_c,t_1), \quad L_2 = L(-t_0,-t_c),$$

$$V(\sigma,\tau) = b_1^2(a_1+d_1)^2\left\{\alpha_1 L(\sigma,\tau) + \frac{\nu(\tau)-\nu(\sigma)}{d_1}\left[\beta_1\left(\nu(\tau)+\nu(\sigma)\right)+\gamma_1\right]\right\},$$

$$\mu_c = \frac{-(a_1+d_1)(s_{2,1}^2-s_{1,2}^2)}{2(1-t_c^2)}, \quad \mu_1 = \frac{-(a_1+d_1)s_{1,1}^2}{2(1-t_1^2)},$$

$$\tilde{V} = V(t_c,t_1) - V(-t_0,-t_c), \quad \varepsilon = 2(b_1 s_{2,2} - t_c\mu_c - t_1\mu_1 - \tilde{V}),$$

$$I_0(\sigma,\tau) = \left(\frac{(1+\kappa\theta)(2+\tau+\sigma)}{2\,(\tau\sigma-1)}+1\right)\theta^{-1}I_1(\sigma,\tau),$$

$$I_{0,1} = I_0(t_c,t_1), \quad I_{2,1} = \left(\frac{L_1}{w_1} - I_{0,1} - 2w_1 I_{1,1}\right)w_1^{-2},$$

$$I_{0,2} = -I_0(-t_0,-t_c), \quad I_{2,2} = \left(\frac{L_2}{w_2} - I_{0,2} - 2w_2 I_{1,2}\right)w_2^{-2}.$$

The following three equations in (t_0,t_1,θ):

$$\begin{aligned}\frac{1}{2}\left(1-\frac{3}{w_1}\right) \;=\; & A + D - \frac{3C}{w_1} + (1+\kappa\theta)\frac{d_1(s_{1,2}-s_{1,1}) - d_2(s_{2,2}-s_{2,1})}{w_1^3} \\ & + \frac{(1+\kappa\theta)B + 2(A-\kappa\theta D) - 1}{2\theta\,w_1^3} + \frac{3(t_0^2-\delta)D}{w_1^2},\end{aligned}$$

$$\frac{1}{2} = C + D\left(t_0^2\frac{I_{2,1}}{I_{1,1}} + \delta\frac{I_{2,2}}{I_{1,2}}\right),$$

$$a_1(1-s_{2,2}^2) + \varepsilon\,(1-t_0) = 2b_1 t_0(1-s_{2,2})$$

give $t_0 = 0.3762232453...$, $t_1 = 0.6849260549...$ and $\theta = -0.2884707066...$, from which all other parameters are computed.

It is also known that $\pi(2,1) = 1.35948...$, $\pi(3,1) = 1.46184...$, $\pi(4,1) = 1.54874...$ and $\pi(5,1) = 1.61031...$ [9–11]. The values of $\pi(3,\infty)$, $\pi(4,\infty)$ and $\pi(5,\infty)$, however, are unknown. Helzel & Petras [8] determined that $1.3539 < \pi(3,\infty) < 1.3577$, $1.4524 < \pi(4,\infty) < 1.4611$, $1.5254 < \pi(5,\infty) < 1.5427$ and remarked that the upper bounds might be more accurate than the lower bounds.

A simpler scenario is when $Y = \mathbb{R}^n$, equipped with the Euclidean norm. The **(absolute) projection constant** of Y is known to be [12, 13]

$$\rho(n) = \sup_X \inf_P \|P\| = \frac{2}{\sqrt{\pi}} \frac{\Gamma((n+2)/2)}{\Gamma((n+1)/2)} = \begin{cases} \dfrac{n+1}{2^n} \dbinom{n}{(n-1)/2} & \text{if } n \text{ is odd,} \\ \dfrac{2^{n+1}}{\pi} \dbinom{n}{n/2}^{-1} & \text{if } n \text{ is even.} \end{cases}$$

More generally, we may examine the q-norm

$$\|x\| = \left(\sum_{k=1}^{n} |x_k|^q \right)^{1/q}$$

for $1 \le q \le \infty$. The case $q = 1$ appears, however obliquely, in [14].

[1] E. W. Cheney and K. H. Price, Minimal projections, *Approximation Theory*, Proc. 1969 Lancaster sympos., ed. A. Talbot, Academic Press, 1970, pp. 261–289; MR0265842 (42 #751).

[2] H. König, D. R. Lewis and P. K. Lin, Finite-dimensional projection constants, *Studia Math.* 75 (1983) 341–358; MR0722257 (84k:46014).

[3] C. Franchetti and U. Tiberio, A numerical evaluation of projection constants, *Linear Algebra Appl.* 109 (1988) 179–196; MR0961577 (90g:46026).

[4] K. Petras, On the minimal norms of polynomial projections, *J. Approx. Theory* 62 (1990) 206–212; MR1068429 (92b:41014).

[5] W. Light, A brief glance at the research of Ward Cheney, *Approximation Theory VIII*, v. 1, *Approximation and Interpolation*, Proc. 1995 College Station conf., ed. C. K. Chui and L. L. Schumaker, World Sci. Publishing, 1995, pp. 343–370; MR1471748 (98g:01066).

[6] C. Franchetti and E. W. Cheney, Minimal projections in L_1-spaces, *Duke Math. J.* 43 (1976) 501–510; MR0423061 (54 #11044).

[7] B. L. Chalmers and F. T. Metcalf, Determination of a minimal projection from $C[-1,1]$ onto the quadratics, *Numer. Funct. Anal. Optim.* 11 (1990) 1–10; MR1058773 (91d:46027).

[8] C. Helzel and K. Petras, Numerical estimation of projection constants, *Numer. Funct. Anal. Optim.* 18 (1997) 555–566; MR1467662 (98h:41004).

[9] B. L. Chalmers and F. T. Metcalf, The determination of minimal projections and extensions in L^1, *Trans. Amer. Math. Soc.* 329 (1992) 289–305. MR1034660 (92e:41017).

[10] B. L. Chalmers and F. T. Metcalf, The minimal projection from L^1 onto π_n, *Stochastic Processes and Functional Analysis*, Proc. 1994 Riverside conf., ed. J. A. Goldstein, N. E. Gretsky and J. J. Uhl, Dekker, 1997, pp. 61–69; MR1440415 (98k:41046).

[11] B. L. Chalmers, F. T. Metcalf and B. Shekhtman, On the computation of minimal projections: millennium report, *Applied Mathematics Reviews*, v. 1, ed. G. A. Anastassiou, World Sci. Publishing, 2000, pp. 119–156; MR1782730 (2001j:41044).

[12] B. Grünbaum, Projection constants, *Trans. Amer. Math. Soc.* 95 (1960) 451–465; MR0114110 (22 #4937).

[13] H. König, C. Schütt and N. Tomczak-Jaegermann, Projection constants of symmetric spaces and variants of Khintchine's inequality, *J. Reine Angew. Math.* 511 (1999) 1–42; MR1695788 (2000i:46014).

[14] S. R. Finch, Moments of sums, *this volume*, §4.2.

2.7 Bohr's Inequality

This essay complements an earlier one [1] on uncertainty inequalities. Let $B_{n,r}$ denote the open n-dimensional ball of radius r centered at the origin. Assume that $f:\mathbb{R}^n \to \mathbb{R}$ is integrable and that its Fourier transform:

$$\hat{f}(\xi) = \int\limits_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x)\, dx$$

satisfies $\hat{f}(\xi) = 0$ for all $\xi \in B_{n,r}$. Note that, to be consistent with the partial differential equations literature, we omit the factor 2π from the exponent (compare with [1]). Assume also that both f and its gradient ∇f are continuous and bounded on $\mathbb{R}^n$. In the case $n = 1$, Bohr [2–4] proved that

$$r \sup_{x\in\mathbb{R}} |f(x)| \le \frac{\pi}{2} \sup_{x\in\mathbb{R}} |f'(x)|.$$

The constant $\pi/2$ is clearly best possible, for examine the periodic function $f(x) = -r|x| + \pi/2$ with $|x| \le \pi/r$ (of period $2\pi/r$). See §2.7.1 for more discussion of this example. In the case $n = 2$, Rüssmann [5] and Hörmander & Bernhardsson [6] calculated that the best constant in the inequality

$$r \sup_{x\in\mathbb{R}^2} |f(x)| \le C \sup_{x\in\mathbb{R}^2} \|\nabla f(x)\|$$

is $C = 2.9038872827\ldots$ (the indicated vector norm is Euclidean). They succeeded in reducing the computation of C to the following one-dimensional optimization problem:

$$C = \min \int\limits_0^\infty |g(y)|\, dy$$

where the minimum is taken over all integrable functions $g:\mathbb{R}\to\mathbb{R}$ satisfying $g(0) = 1$, $g(y) = g(-y)$ for all y, and $\hat{g}(\eta) = 0$ for all $|\eta| \ge 1$. In fact, g can be extended to an entire analytic function of exponential type 1 on the complex plane; the zeroes of g are all real and simple.

The constants $\pi/2$ and C also appear in connection with solving the linear operator equation $PX - XQ = Y$, where $P: \mathbb{H} \to \mathbb{H}$, $Q: \mathbb{K} \to \mathbb{K}$ and $Y: \mathbb{H} \to \mathbb{K}$ are bounded operators on Hilbert spaces $\mathbb{H}$ and $\mathbb{K}$. If the spectra $\sigma(P)$, $\sigma(Q)$ of P, Q are disjoint subsets of $\mathbb{C}$, then the equation $PX - XQ = Y$ possesses a unique solution X. Let $\delta = \inf_{\lambda \in \overline{\sigma(P)}, \mu \in \overline{\sigma(Q)}} |\lambda - \mu|$, the separation between closed sets containing the spectra. The norm of the transformation $Y \mapsto X$ can be bounded by $(\pi/2)/\delta$ if P, Q are self-adjoint and C/δ if instead P, Q are normal. Bhatia, Davis & Koosis [7–9] wrote that there is "no substantial evidence" for expecting these two constants to be best possible here, but added that they cannot be far off. A related problem involves perturbation bounds for spectral subspaces.

What can be said if higher-order derivatives of f are continuous and bounded on $\mathbb{R}^n$? In the case $n = 1$, Favard [10] proved that

$$r^m \sup_{x \in \mathbb{R}} |f(x)| \le K_m \sup_{x \in \mathbb{R}} |f^{(m)}(x)|$$

for each positive integer m, where the constants [11]

$$1 = K_0 < K_2 = \frac{\pi^2}{8} < K_4 < \ldots < \frac{4}{\pi} < \ldots < K_5 < K_3 = \frac{\pi^3}{24} < K_1 = \frac{\pi}{2}$$

are all best possible. This is called the **Bohr-Favard inequality**. The case $n \ge 2$ remains open.

What can be said if we assume instead that $\hat{f}(\xi) = 0$ for all $\xi \notin \bar{B}_{n,r}$? (That is, we assume the support of f is completely contained within the closed r-ball, the opposite of before.) In the case $n = 1$, Bernstein [12, 13] proved that

$$\sup_{x \in \mathbb{R}} |f^{(m)}(x)| \le r^m \sup_{x \in \mathbb{R}} |f(x)|$$

for each positive integer m, where the constant 1 is best possible. Such functions f are said to be **band-limited** and, like g, can be extended to an entire function of exponential type r. The generalization

$$\sup_{x \in \mathbb{R}^n} \|\nabla f(x)\| \le r \sup_{x \in \mathbb{R}^n} |f(x)|$$

for $n \ge 2$ (when $m = 1$) and higher-order analogs (when $m > 1$) were apparently first found by Nikolskii [14, 15].

2.7.1 *Tempered Distributions*

Let f denote the periodic triangular wave function mentioned earlier. It is not true that f is integrable on $\mathbb{R}$: strictly speaking, its Fourier transform is undefined (although signal processing engineers would describe $\hat{f}$ as a weighted sequence of equidistant Dirac impulses at $\xi = \pm r, \pm 2r, \pm 3r, \ldots$). We can circumvent this difficulty by defining a family of rapidly decreasing test functions

$$\varphi_k(x) = e^{-x^2/k^2}, \quad k = 1, 2, 3, \ldots$$

and then taking

$$\hat{f}(\xi) = \lim_{k\to\infty} \int_{-\infty}^{\infty} e^{-i\xi x}\varphi_k(x) f(x)\, dx.$$

What allows us, however, to conclude that $\hat{f}$ is independent of the choice of test functions $\{\varphi_k\}_{k=1}^{\infty}$, $\varphi_k \to 1$ as $k \to \infty$?

Here is a little background. The space of all infinitely differentiable functions φ such that $\varphi^{(j)}(x) = O(|x|^{-n})$ as $x \to \pm\infty$, for any $j \geq 0$ and $n \geq 1$, is called the **Schwarz space** $\mathcal{S}$. A **tempered distribution** is a continuous linear functional T on $\mathcal{S}$ and its (generalized) Fourier transform is defined by

$$\hat{T}(\varphi) = T(\hat{\varphi});$$

this induces an automorphism $\mathcal{S}' \to \mathcal{S}'$ of the dual space $\mathcal{S}'$ of $\mathcal{S}$. Consider now the example

$$F(\varphi) = \int_{-\infty}^{\infty} \varphi(x) f(x)\, dx,$$

where f is the periodic triangular wave, and let $f_k = \varphi_k f$. Clearly

$$\hat{F}(\varphi) = \int_{-\infty}^{\infty} \hat{\varphi}(x) f(x)\, dx = \int_{-\infty}^{\infty} \hat{\varphi}(x) \left(\lim_{k\to\infty} f_k(x)\right) dx,$$

while

$$\lim_{k\to\infty} \int_{-\infty}^{\infty} \varphi(\xi) \hat{f}_k(\xi)\, d\xi = \lim_{k\to\infty} \int_{-\infty}^{\infty} \hat{\varphi}(x) f_k(x)\, dx$$

follows by interchanging the order of integration. Since $|\hat{\varphi} f_k| \leq |\hat{\varphi} f|$ and $\hat{\varphi} f$ is integrable on $\mathbb{R}$, the limit may be brought inside the integral by Lebesgue's dominated convergence theorem. Hence, just as f and F are regarded as the same, we may identify $\hat{f}$ and $\hat{F}$.

[1] S. R. Finch, Uncertainty inequalities, *this volume*, §2.4.

[2] H. Bohr, Ein allgemeiner Satz über die Integration eines trigonometrischen Polynoms, *Prace Matematyczno-Fizyczne* 43 (1935) 273–288; *Collected Mathematical Works*, v. II, ed. E. Følner and B. Jessen, Dansk Matematisk Forening, 1952, C36, pp. 1–16.

[3] B. Sz.-Nagy, Über die Ungleichung von H. Bohr, *Math. Nachr.* 9 (1953) 255–259; MR0054765 (14,976d).

[4] L. Hörmander, A new proof and a generalization of an inequality of Bohr, *Math. Scand.* 2 (1954). 33–45; MR0064906 (16,354e).

[5] H. Rüssmann, On an inequality for trigonometric polynomials in several variables, *Analysis, et cetera*, ed. P. H. Rabinowitz and E. Zehnder, Academic Press, 1990, pp. 545–562; MR1039361 (91f:42001).

[6] L. Hörmander and B. Bernhardsson, An extension of Bohr's inequality, *Boundary Value Problems for Partial Differential Equations and Applications*, ed. J.-L. Lions and C. Baiocchi, Masson, 1993, pp. 179–194; MR1260445 (95e:46052).
[7] R. Bhatia, C. Davis and P. Koosis, An extremal problem in Fourier analysis with applications to operator theory, *J. Funct. Anal.* 82 (1989) 138–150; MR0976316 (91a:42006).
[8] R. Bhatia, *Matrix Analysis*, Springer-Verlag, 1997, pp. 163, 203–225, 229, 250; MR1477662 (98i:15003).
[9] R. Bhatia and P. Rosenthal, How and why to solve the operator equation $AX-XB=Y$, *Bull. London Math. Soc.* 29 (1997) 1–21; MR1416400 (97k:47016).
[10] J. Favard, Application de la formule sommatoire d'Euler à la démonstration de quelques propriétés extrémales des intégrale des fonctions périodiques ou presquepériodiques, *Matematisk Tidsskrift B (Kobenhavn)* (1936) 81–94.
[11] S. R. Finch, Achieser-Krein-Favard constants, *first volume*, pp. 255–257.
[12] S. N. Bernstein, Sur une propri'et'e des fonctions enti'eres, *C. R. Acad. Sci. Paris* 176 (1923) 1603–1605.
[13] R. P. Boas, *Entire Functions*, Academic Press, 1954, pp. 206–207; MR0068627 (16,914f).
[14] S. M. Nikolskii, Inequalities for entire functions of finite order and their application in the theory of differentiable functions of several variables (in Russian), *Trudy Mat. Inst. Steklov.*, v. 38, Izdat. Akad. Nauk SSSR, 1951, pp. 244–278; MR0048565 (14,32e).
[15] S. M. Nikolskii, *Approximation of Functions of Several Variables and Imbedding Theorems* (in Russian), Izdat. Nauka, 1969; Engl. transl., Springer-Verlag, 1975; 2nd ed., Nauka, 1977; MR0310616 (46 #9714), MR0374877 (51 #11073) and MR0506247 (81f:46046).

2.8 Moduli of Continuity

2.8.1 *Bernstein Polynomials*

Bernstein's proof of the Weierstrass approximation theorem makes use of the operator

$$B_nf(x)=\sum_{j=0}^{n}\binom{n}{j}x^j(1-x)^{n-j}f\left(\frac{j}{n}\right),$$

given any continuous function $f\colon[0,1]\to\mathbb{R}$. To demonstrate that

$$\lim_{n\to\infty}B_nf(x)=f(x)$$

uniformly on $[0,1]$ requires a bound of the form

$$\sup_{0\le x\le 1}|B_nf(x)-f(x)|\le c\cdot\omega(f,n^{-1/2}),$$

where $\omega(f,\delta)$ is the **first modulus of continuity**

$$\omega(f,\delta)=\sup_{|u-v|<\delta}|f(u)-f(v)|$$

Table 2.9 *Best constants c_n: exact expressions and decimal approximations*

n	Exact	Decimal	n	Exact	Decimal
1	1	1	5	$\frac{21-7\sqrt{5}}{5}$	1.0695048315...
2	$\frac{5-2\sqrt{2}}{2}$	1.0857864376...	6	$\frac{4306+837\sqrt{6}}{5832}$	1.0898873310...
3	$\frac{27-10\sqrt{3}}{9}$	1.0754991027...	7	$\frac{35442+33754\sqrt{7}}{117649}$	1.0603293674...
4	$\frac{17}{16}$	1.0625	8	$\frac{3865512\sqrt{8}-1937991}{8388608}$	1.0723266591...

and $0 \le \delta \le 1$. What is the best possible constant c that works for all $n \ge 1$? Starting from [1, 2], Sikkema [3–5] proved that

$$\sup_{n\ge1}\sup_{f}\sup_{0\le x\le1}\frac{|B_nf(x)-f(x)|}{\omega(f,n^{-1/2})}=\frac{4306+837\sqrt{6}}{5832}=1.0898873310...$$

and this value is attained only for $n=6$. Table 2.9 lists the best possible constants c_n that work for specified $n=1,2,\ldots,8$.

Esseen [6–9] examined the limiting behavior of c_n as n grows without bound:

$$\begin{aligned}\limsup_{n\to\infty}\sup_{f}\sup_{0\le x\le1}\frac{|B_nf(x)-f(x)|}{\omega(f,n^{-1/2})} &= 2\sum_{m=0}^{\infty}(m+1)\left(\Phi(2m+2)-\Phi(2m)\right)\\ &= 1.0455636083...\end{aligned}$$

where $\Phi(x)$ is the standard normal distribution function [10]. Of course, we understand to omit constant functions f from the supremum (for which $\omega=0$).

Define the **second modulus of continuity**

$$\tilde{\omega}(f,\delta)=\sup_{|u-v|<\delta}\left|f(u)-2f\left(\frac{u+v}{2}\right)+f(v)\right|.$$

In contrast with the preceding results, the best constant and best asymptotic constant here coincide:

$$\sup_{f}\sup_{0\le x\le1}\frac{|B_nf(x)-f(x)|}{\tilde{\omega}(f,n^{-1/2})}=1$$

for each $n\ge1$. This was proved by Paltanea [11], building on earlier results [12–17]. Here, of course, we understand to omit linear functions f from the supremum (for which $\tilde{\omega}=0$).

Let us return to the first modulus ω for the remainder of this essay. Define Ω to be the set of all continuous functions $g:[0,1]\to\mathbb{R}$ that vanish at zero, are nondecreasing and subadditive (meaning $g(x+y) \le g(x) + g(y)$ always). Each member g of Ω satisfies $g(x) = \omega(g, x)$ and thus is itself a modulus of continuity. Define Ω^* to be the subset of Ω whose elements g are such that $x \mapsto x^{-1}g(x)$ is nonincreasing on $(0, 1]$. Then [18, 19]

$$\sup_{n\ge1}\sup_{0<x\le1}\sup_{g\in\Omega}\frac{B_n g(x)}{g(x)} = 2 > 1.1855905950... = \alpha = \sup_{0<x\le1}\sup_{n\ge1}\sup_{g\in\Omega^*}\frac{B_n g(x)}{g(x)},$$

where

$$\alpha = \sup_{k\ge0}\sup_{k\le x\le k+1} 1 + e^{-x}\left(\frac{x^k}{k!} - 1\right) = 1 + \frac{\xi^2}{2}e^{-\xi}$$

and $\xi = 3.4920333011...$ is the unique real zero of the cubic equation $x^3 - 3x^2 - 6 = 0$.

A seemingly related problem involves the ratio of moduli [20, 21]

$$\rho_1(n) = \sup_{0<\delta\le1}\sup_{f}\frac{\omega(B_n f, \delta)}{\omega(f, \delta)} = 2$$

for each $n \ge 1$. There are interesting multivariate versions of this result. Consider the operator

$$B_n f(x, y) = \sum_{i=0}^{n}\sum_{j=0}^{n}\binom{n}{i}x^i(1-x)^{n-i}\binom{n}{j}y^j(1-y)^{n-j}f\left(\frac{i}{n}, \frac{j}{n}\right),$$

given any continuous function $f\colon [0, 1] \times [0, 1] \to \mathbb{R}$. This is also called the bivariate **tensor product** Bernstein polynomial on the unit square. De La Cal, Cárcamo & Valle [22, 23] proved that, in this two-dimensional case, the ratio

$$\rho_2(n) = \sup_{0<\delta\le1}\sup_{f}\frac{\omega(B_n f, \delta)}{\omega(f, \delta)}$$

depends on n and

$$\begin{aligned}\sup_{n\ge1}\rho_2(n) &= 1 - \frac{1}{e^2} + \sum_{t=0}^{\infty}\left[1 - \frac{1}{e^2}\left(\sum_{s=0}^{t}\frac{1}{s!}\right)^2\right] \\ &= 2.3884423285... = 1 - e^{-2} + \beta,\end{aligned}$$

where $\beta = 1.5237776118...$ is the mean of the maximum of two independent Poisson(1) random variables. One would expect the k-dimensional case, $k \ge 3$, to be even more complicated. In fact, $\rho_k(n) = k$ for all $n \ge 1$. Hence only the bivariate case gives n-dependent behavior as well as a new constant, which is quite surprising.

2.8.2 Müntz–Jackson theorem

Müntz's theorem gives that the power functions

$$\left\{x^{\lambda_j} : 0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots\right\}$$

generate a dense subspace of the space of all continuous functions on [0, 1] if and only if $\sum_{j=0}^{\infty} 1/\lambda_j = \infty$. Jackson's theorem is in the spirit of other results in this essay: It provides bounds on the error in approximating a continuous function f by polynomials in terms of ω. Newman [24, 25] combined the two theorems in the following way. Define

$$\Lambda = \{\lambda_j : 0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n\}$$

and generalized polynomials

$$Q_\Lambda = \left\{ \sum_{j=0}^{n} a_j x^{\lambda_j} : a_j \in \mathbb{R} \text{ for all } 0 \le j \le n \right\}.$$

Then

$$\inf_{q \in Q_\Lambda} \sup_{0 \le x \le 1} |q(x) - f(x)| \le C \cdot \omega(f, \varepsilon_\Lambda),$$

where C is a constant independent of f and Λ, and

$$\varepsilon_\Lambda = \sup_{\mathrm{Re}(z)=1} \left| \frac{1}{z} \frac{z-\lambda_1}{z+\lambda_1} \frac{z-\lambda_2}{z+\lambda_2} \cdots \frac{z-\lambda_n}{z+\lambda_n} \right|.$$

Newman [24, 25] demonstrated that $1/50 < C < 368$ and Odogwu [26] improved the upper bound to 66. Over and beyond the value of C, the Blaschke product formula for ε_Λ is intriguing. Special cases (when consecutive λs are at least 2 apart, or when consecutive λs are at most 2 apart) with simpler formulas also exist.

An L_p-generalization of ω can be defined; the constants in this essay correspond only to the case $p = \infty$. It would be good to see their L_p-analogs for $p < \infty$. Clearly $\lim_{\delta \to 0} \omega(f, \delta) \cdot \delta^{-1} = 0$ implies that f is constant. Consequences of the weaker condition $\lim_{\delta \to 0} \omega(f, \delta) \cdot \ln(\delta) = 0$ are mentioned in [27].

[1] T. Popoviciu, Sur l'approximation des fonctions convexes d'ordre supérieur, *Mathematica (Cluj)* 10 (1935) 49–54.

[2] G. G. Lorentz, *Bernstein Polynomials*, 2nd ed., Chelsea, 1986; pp. 20–21, 51; MR0864976 (88a:41006).

[3] P. C. Sikkema, Über den Grad der Approximation mit Bernstein-Polynomen, *Numer. Math.* 1 (1959) 221–239; MR0110178 (22 #1060).

[4] P. C. Sikkema, Der Wert einiger Konstanten in der Theorie der Approximation mit Bernstein-Polynomen, *Numer. Math.* 3 (1961) 107–116; MR0123128 (23 #A459).

[5] H. van Iperen and P. C. Sikkema, Determination of a class of best constants in the approximation by powers of generalized Bernstein operators, *Nederl. Akad. Wetensch. Proc. Ser. A* 71 (1968) 336–352; *Indag. Math.* 30 (1968) 336–352; MR0241855 (39 #3192).

[6] C. G. Esseen, Über die asymptotisch beste Approximation stetiger Funktionen mit Hilfe von Bernstein-Polynomen, *Numer. Math.* 2 (1960) 206–213; MR0132948 (24 #A2784).

[7] H. Walk, Probabilistic methods in the approximation by linear positive operators, *Nederl. Akad. Wetensch. Indag. Math*. 42 (1980) 445–455; MR0598003 (82c:41024).

[8] Y. I. Volkov, Multiple sequences of multidimensional linear positive operators (in Russian), *Ukrain. Mat. Zh.* 36 (1984) 286–291; Engl. transl. in *Ukrainian Math. J.* 36 (1984) 257–262; MR0749716 (86d:41028).

[9] X. Xiang, The best asymptotic constant of a class of approximation operators, *J. Approx. Theory* 70 (1992) 348–357; MR1178378 (93f:41037).

[10] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[11] R. Păltănea, Optimal constant in approximation by Bernstein operators, *J. Comput. Anal. Appl.* 5 (2003) 195–235; MR1980393 (2004b:41029).

[12] Y. A. Brudnyi, On a certain method for approximation of bounded functions given on a segment (in Russian), *Studies of Contemporary Problems in the Constructive Theory of Functions*, Proc. Second All-Union conf., Baku, 1962, ed. I. I. Ibragimov, Izdat. Akad. Nauk Azerbaidzan. SSR, pp. 40–45; MR0199614 (33 #7757).

[13] J.-D. Cao, On linear approximation methods (in Chinese), *Acta Sci. Natur. Univ. Fudan* 9 (1964) 43–52.

[14] H. H. Gonska, Two problems on best constants in direct estimates, *Second Edmonton Conference on Approximation Theory*, 1982, ed. by Z. Ditzian, A. Meir, S. D. Riemenschneider and A. Sharma, Canad. Math. Soc., 1983, p. 394; MR0729317 (84k:41002).

[15] H. H. Gonska and R. K. Kovacheva, The second order modulus revisited: remarks, applications, problems, *Conferenze del Seminario di Matematica dell'Universita di Bari* 257 (1994) 1–32; MR1366531 (96k:41023).

[16] R. Păltănea, Best constants in estimates with second order moduli of continuity, *Proc. 1995 Int. Dortmund Meeting on Approximation Theory (IDoMAT)*, Witten, ed. M. W. Müller, M. Felten and D. H. Mache, Wiley, 1995, pp. 251–275; MR1377075 (97h:41042).

[17] R. Păltănea, On an optimal constant in approximation by Bernstein operators, *Rend. Circ. Mat. Palermo Suppl.* n. 52, v. II (1998) 663–686; MR1644582 (2000a:41008).

[18] Z. Li, Bernstein polynomials and modulus of continuity, *J. Approx. Theory* 102 (2000) 171–174; MR1736050 (2000j:41029).

[19] J. de La Cal and J. Cárcamo, On certain best constants for Bernstein-type operators, *J. Approx. Theory* 113 (2001) 189–206; MR1876322 (2002i:41036).

[20] G. A. Anastassiou, C. Cottin and H. H. Gonska, Global smoothness of approximating functions, *Analysis* 11 (1991) 43–57; MR1113067 (92h:41047).

[21] J. A. Adell and A. Pérez-Palomares, Best constants in preservation inequalities concerning the first modulus and Lipschitz classes for Bernstein-type operators, *J. Approx. Theory* 93 (1998) 128–139; MR1612798 (99a:41038).

[22] J. de La Cal and A. M. Valle, Best constants for tensor products of Bernstein, Szász and Baskakov operators, *Bull. Austral. Math. Soc*. 62 (2000) 211–220; MR1786203 (2001j:41029).

[23] J. de La Cal, J. Cárcamo and A. M. Valle, A best constant for bivariate Bernstein and Szász-Mirakyan operators, *J. Approx. Theory* 123 (2003) 117–124; MR1985019 (2004f:41029).

[24] D. J. Newman, A general Müntz-Jackson theorem, *Amer. J. Math.* 96 (1974) 340–345; MR0352804 (50 #5290).
[25] R. P. Feinerman and D. J. Newman, *Polynomial Approximation*, Williams & Wilkins, 1974, pp. 13–15, 121–137; MR0499910 (58 #17657).
[26] H. N. Odogwu, An improved constant for the Müntz-Jackson theorem, *Publ. Inst. Math. (Beograd)* 45(59) (1989) 103–108; MR1021924 (90m:41015).
[27] S. R. Finch, Lebesgue constants, *first volume*, pp. 250–255.

2.9 Quinn–Rand–Strogatz Constant

We present two problems: one is easy (for the sake of comparison) and the other is difficult. The unique solution $s > 0$ of the algebraic equation

$$0 = \sum_{j=1}^{n} \left[1 - 3s^2 \left(1 - 2\tfrac{j-1}{n-1}\right)^2\right]$$

is

$$s = \sqrt{\tfrac{n-1}{n+1}} \sim 1 - \tfrac{1}{n} + \tfrac{1}{2}\tfrac{1}{n^2} - \tfrac{1}{2}\tfrac{1}{n^3} + \tfrac{3}{8}\tfrac{1}{n^4} - \tfrac{3}{8}\tfrac{1}{n^5} + - \cdots$$

as $n \to \infty$. Define $s_n = 1 - 1/n$, the first-order approximation, and a certain partial sum

$$f_n(x) = \sum_{j=1}^{n} \left[1 - s_n^2 \left(1 - 2\tfrac{j-1}{n-1}\right)^2\right]^{-x}$$

for $x > 0$. It follows that

$$\lim_{n\to\infty} \frac{f_n(1)}{n\ln(n)} = \frac{1}{2}, \quad \lim_{n\to\infty} \frac{f_n(2)}{n^2} = \frac{\pi^2}{16}$$

and such formulas for other values of x are possible.

The unique solution $s > 0$ of the algebraic equation

$$0 = \sum_{j=1}^{n} \left[2\sqrt{1 - s^2\left(1 - 2\tfrac{j-1}{n-1}\right)^2} - \frac{1}{\sqrt{1 - s^2\left(1 - 2\tfrac{j-1}{n-1}\right)^2}}\right]$$

satisfies [1–3]

$$s \sim 1 - \tfrac{c_1}{n} - \tfrac{c_2}{n^2} - \tfrac{c_3}{n^3} - \tfrac{c_4}{n^4} - \cdots$$

as $n \to \infty$, where

$$c_1 = 0.6054436571..., \quad c_2 = -0.1046854594...,$$

$$c_3 = 0.1263143361..., \quad c_4 = -0.0159376251....$$

Bailey, Borwein & Crandall [2] proved that c_1 is the unique solution $y \in (0, 2)$ of the transcendental equation

$$\zeta(1/2, y/2) = 0$$

where

$$\zeta(z, a) = \sum_{\substack{k=0 \\ k+a \neq 0}}^{\infty} \frac{1}{(k+a)^z}$$

is the Hurwitz zeta function (with analytic continuation). Further,

$$c_2 = c_1 - c_1^2 - 30\frac{\zeta(-1/2, c_1/2)}{\zeta(3/2, c_1/2)}$$

but exact expressions for c_3, c_4 remain open [3]. Define $s_n = 1 - c_1/n$, the first-order approximation, and a partial sum $f_n(x)$ exactly as before. It follows that

$$\lim_{n\to\infty} \frac{f_n(3/2)}{n^{3/2}} = \frac{1}{4}\zeta\left(\frac{3}{2}, \frac{c_1}{2}\right) = 2.0381693797....$$

It is believed that analogous formulas involving Hurwitz zeta function values should exist for other choices of x.

2.9.1 *Self-Synchronization*

We briefly discuss a model underlying coherent phenomena in biology such as flashing fireflies and cardiac pacemaker cells [4]. Let $0 \leq \lambda < 1$. Consider a population of n oscillators with natural frequencies ω_i chosen at random from a symmetric unimodal density $g(\omega)$ on the interval $[1-\lambda, 1+\lambda]$. Assume that the mean of $g(\omega)$ is equal to 1. The **Winfree model** [5] is a system of differential equations

$$\frac{d\theta_i}{dt} = \omega_i - \frac{\kappa}{n}\sin(\theta_i)\sum_{j=1}^{n}\left[1 + \cos(\theta_j)\right], \quad 1 \leq i \leq n,$$

where $\theta_i(t)$ is the phase of the i^{th} oscillator at time t and $\kappa \geq 0$ is the (constant) coupling strength. To study system dynamics for large n, identify oscillators by their frequency ω_i instead of their index i. Defining $\Theta(t, \nu)$ to be $\theta_{1+\lambda\nu}(t)$ for $-1 \leq \nu \leq 1$, we obtain the following integro-differential equation [6, 7]:

$$\frac{\partial\Theta}{\partial t}(t, \nu) = 1 + \lambda\nu - \kappa\sin(\Theta(t, \nu))\int_{-1}^{1}\left[1 + \cos(\Theta(t, \mu))\right]h(\mu)d\mu$$

in the limit as $n \to \infty$, where the density h is simply g translated to $[-1, 1]$ and normalized. Such a result permits a linear stability analysis necessary to determine system behavior and bifurcation curves as a function of (λ, κ). When κ is small relative to the spread λ of natural frequencies, each oscillator behaves independently of the others; when κ is larger than a critical value, some of the oscillators spontaneously synchronize to a common frequency while others remain adrift;

when κ exceeds another (greater) threshold, all of the oscillators are in line. For example, the prescription of a certain saddle-node condition is determined via [1]

$$0=\int\limits_{-1}^{1}\frac{\cos(2\varphi(\nu))}{\cos(\varphi(\nu))}h(\nu)d\nu=\int\limits_{-1}^{1}\left[2\cos(\varphi(\nu))-\frac{1}{\cos(\varphi(\nu))}\right]h(\nu)d\nu$$

by a double-angle formula, where $\nu=\sin(\varphi)/s$. In the special case of a discrete uniform density

$$h(\nu)=\frac{1}{n}\sum_{j=1}^{n}\delta\left[\nu-\left(1-2\tfrac{j-1}{n-1}\right)\right]$$

where δ is the Dirac delta function, we have nonzero contributions precisely when

$$\sin(\varphi)=s\left(1-2\tfrac{j-1}{n-1}\right),\quad \cos(\varphi)=\sqrt{1-s^2\left(1-2\tfrac{j-1}{n-1}\right)^2}$$

and thus the second (difficult) algebraic equation emerges. The asymptotic expansion for s gives insight concerning possible singularities in the bifurcation curve separating partially and fully phase-locked states. Recent associated work includes [8–11].

[1] D. D. Quinn, R. H. Rand and S. H. Strogatz, Singular unlocking transition in the Winfree model of coupled oscillators, *Phys. Rev. E* 75 (2007) 036218; MR2358543.

[2] D. H. Bailey, R. E. Crandall and J. M. Borwein, Resolution of the Quinn-Rand-Strogatz constant of nonlinear physics, *Experim. Math.* 18 (2009) 107–116; MR2548991 (2010k:11192).

[3] N. J. Durgin, S. M. Garcia, T. Flournoy and D. H. Bailey, 'Syncing' up with the Quinn-Rand-Strogatz constant: Hurwitz-zeta functions in non-linear physics, unpublished note (2007).

[4] A. Pikovsky, M. Rosenblum and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Science,* Cambridge Univ. Press, 2001.

[5] A. T. Winfree, Biological rhythms and the behavior of populations of coupled oscillators, *J. Theoret. Biology* 16 (1967) 15–42.

[6] J. T. Ariaratnam and S. H. Strogatz, Phase diagram for the Winfree model of coupled nonlinear oscillators, *Phys. Rev. Lett.* 86 (2001) 4278–4281; arXiv:nlin/0012017.

[7] J. T. Ariaratnam, *Collective Dynamics of the Winfree Model of Coupled Nonlinear Oscillators*, Ph.D. thesis, Cornell Univ., 2002.

[8] S.-Y. Ha, J. Park and S. W. Ryoo, Emergence of phase-locked states for the Winfree model in a large coupling regime, *Discrete Contin. Dyn. Syst.* 35 (2015) 3417–3436; MR3320132.

[9] S.-Y. Ha, D. Ko, J. Park and S. W. Ryoo, Emergent dynamics of Winfree oscillators on locally coupled networks, *J. Differential Equations* 260 (2016) 4203–4236; MR3437585.

[10] S.-Y. Ha, D. Ko, J. Park and S. W. Ryoo, Emergence of partial locking states from the ensemble of Winfree oscillators, *Quart. Appl. Math.* 75 (2017) 39–68; MR3580095.

[11] W. Oukil, A. Kessi and P. Thieullen, Synchronization hypothesis in the Winfree model, *Dynam. Systems* 32 (2017) 326–339; arXiv:1507.06061; MR3669804.

2.10 Tsirelson's Constant

All infinite-dimensional, separable, complex Hilbert spaces are isometrically isomorphic [1]. Fix such a space X for consideration. Let P, Q be self-adjoint linear operators on X that satisfy the canonical commutation relations

$$PQ - QP = -iI,$$

where i is the imaginary unit and I is the identity operator. Such unbounded operators P, Q are each defined only on a dense linear subspace of X, and the intersection of two dense linear subspaces generally need not be dense. The commutation relations ensure, however, that

$$R = -(P + Q)$$

is well-defined and is a self-adjoint linear operator. Hence we have three operators P, Q, R such that $P + Q + R = 0$ and

$$PQ - QP = QR - RQ = RP - PR = -iI.$$

Let us define a **sign function** for operators [2]. First, the scalar sign function is given by

$$\operatorname{sgn}(z) = \begin{cases} 1 & \text{if } \operatorname{Re}(z) > 0, \\ -1 & \text{if } \operatorname{Re}(z) < 0 \end{cases}$$

for $z \in \mathbb{C}$ lying off the imaginary axis. Next, the matrix sign function is given by

$$\operatorname{sgn}(M) = U \operatorname{sgn}(\Lambda)\, U^{-1},$$

where $M \in \mathbb{C}^{n \times n}$ is a Hermitian matrix with no eigenvalues on the imaginary axis. The unitary $n \times n$ matrix U has column vectors equal to the orthonormal eigenvector basis of $\mathbb{C}^n$ determined by M, and the diagonal $n \times n$ matrix Λ has components equal to the (real) eigenvalues of M:

$$M = U \Lambda\, U^{-1}.$$

By $\operatorname{sgn}(\Lambda)$ is meant the diagonal $n \times n$ matrix with sgn applied component-wise to Λ. Finally, the operator sign function can be defined similarly by use of the spectral theorem for unbounded operators (upon which we do not elaborate).

It is remarkable that the operator norm [3]

$$c = \|\operatorname{sgn}(P) + \operatorname{sgn}(Q) + \operatorname{sgn}(R)\| \approx 1.2$$

is independent of the choice of P, Q, R.[1] It is a nontrivial constant and a more precise estimate would be good to see. We will provide a limiting expression for c shortly.

[1] The addendum clarifies the meaning of $PQ - QP = -iI$ and the well-definition of c.

2.10.1 *Schrödinger Representation*

Let $X = L_2(\mathbb{R})$ and, for wave functions $\psi \in L_2(\mathbb{R})$,

$$(P\psi)(x) = -i\frac{d}{dx}\psi(x), \quad (Q\psi)(x) = x\,\psi(x).$$

These are the momentum and position (or coordinate) operators that arise in quantum mechanics. Further, the time-independent Schrödinger ODE for the quantum harmonic oscillator [4–7]:

$$\frac{d^2\psi}{dx^2} + \left(\lambda - x^2\right)\psi = 0$$

(in natural units) can be written as

$$(P^2 + Q^2)\psi = \lambda\,\psi$$

with eigenvalues $\lambda_n = 2n + 1$ for $n = 0, 1, 2, \ldots$ and orthonormal eigenfunctions

$$\psi_n(x) = \left(\sqrt{\pi}n!2^n\right)^{-1/2} e^{-x^2/2} H_n(x).$$

The Hermite polynomials $H_n(x)$ satisfy Rodrigues' formula

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}\left(e^{-x^2}\right)$$

as well as the recurrence

$$H_{n+1}(x) = 2x\,H_n(x) - 2n\,H_{n-1}(x), \quad H_0(x) = 1, \quad H_1(x) = 2x.$$

It is well-known that

$$\int_{-\infty}^{\infty} \psi_n(x)^2 dx = 1$$

and $\psi_n(x)^2$ is the probability density for location of a particle in the n^{th} energy state of a harmonic oscillator. Corresponding to any observable physical quantity, there is a self-adjoint linear operator T, and its expected value for the same particle is

$$\mathrm{E}_n(T) = \int_{-\infty}^{\infty} \psi_n(x)(T\psi_n)(x)dx.$$

For example,

$$\sqrt{\mathrm{Var}_n(P)}\sqrt{\mathrm{Var}_n(Q)} = n + \frac{1}{2} \geq \frac{1}{2},$$

which constitutes the Heisenberg uncertainty principle for a quantum harmonic oscillator (in dimensionless variables). The fact that the product of uncertainties is bounded away from zero can be proved under much more general circumstances.

In the following section, the Laguerre polynomials

$$L_n(x) = \frac{1}{n!} e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right)$$

are essential. These obey the recurrence

$$(n+1)L_{n+1}(x) = (2n+1-x)\,L_n(x) - n\,L_{n-1}(x), \quad L_0(x)=1, \quad L_1(x)=1-x$$

and are orthogonal with respect to the exponential distribution Exp(1), just as the Hermite polynomials are orthogonal with respect to the normal distribution N(0, 1/2).

2.10.2 *Wigner Function*

One might believe that, to estimate c, all we must do is to find $n \times n$ matrices P, Q satisfying the commutation relations for arbitrarily large n. Unfortunately no such matrices exist since otherwise we would have

$$0 = \mathrm{tr}(P\,Q) - \mathrm{tr}(Q\,P) = \mathrm{tr}(P\,Q - Q\,P) = \mathrm{tr}(-i\,I) = -i\,n,$$

a contradiction. A different approach must be found.

The **Wigner function** (or **quasi-distribution**) offers a way to compute c. All we require are its values on the Hermite eigenfunction basis of $L_2(\mathbb{R}) \times L_2(\mathbb{R})$:

$$\begin{aligned} w_{m,n}(x,y) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi_m\left(x+\frac{\xi}{2}\right) e^{i\,\xi\,y}\psi_n\left(x-\frac{\xi}{2}\right) d\xi \\ &= \begin{cases} \dfrac{(-1)^m}{\pi}\sqrt{\dfrac{m!}{n!}}(2\bar{z})^{n-m}e^{-2|z|^2}L_m^{(n-m)}\left(4|z|^2\right) & \text{if } m \le n, \\ w_{n,m}(x,y) & \text{if } m > n, \end{cases} \end{aligned}$$

where $z = (x+i\,y)/\sqrt{2}$ and $\bar{z} = (x-i\,y)/\sqrt{2}$. See [8–11] for details. The generalized Laguerre polynomials are related to the (ordinary) Laguerre polynomials via

$$L_m^{(k)}(x) = (-1)^k \frac{d^k}{dx^k} L_{m+k}(x).$$

The n^{th} expected value of any physical quantity $f(Q,P)$ can alternatively be calculated via

$$\mathrm{E}_n(f(Q,P)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,w_{n,n}(x,y)\,dy\,dx.$$

For example,

$$\mathrm{Var}_n(P) = \frac{(-1)^n}{\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 e^{-(x^2+y^2)} L_n(2(x^2+y^2))dy\,dx = n + \frac{1}{2}$$

and $\mathrm{Var}_n(Q)$ likewise, confirming Heisenberg's principle. Note that $w_{1,1}(0,0) = -1/\pi$, for instance, and thus the Wigner function is not a probability density in the usual sense (because it may take negative values).

2.10.3 *Operator Norm*

The $(m,n)^{\text{th}}$ element in the matrix representation of the operator $T = \mathrm{sgn}(P)$ relative to the Hermite eigenfunction basis of $L_2(\mathbb{R})$ is

$$\begin{aligned}\int_{-\infty}^{\infty} \psi_m(x)(T\psi_n)(x)dx &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{sgn}(y) w_{m,n}(x,y)\,dy\,dx \\ &= \int_{-\infty}^{\infty}\int_{0}^{\infty} w_{m,n}(x,y)\,dy\,dx - \int_{-\infty}^{\infty}\int_{-\infty}^{0} w_{m,n}(x,y)\,dy\,dx\end{aligned}$$

for integers $m \geq 0$, $n \geq 0$. Changing to polar coordinates

$$x = r\cos(\theta), \quad y = r\sin(\theta)$$

in the upper half plane, we obtain

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{0}^{\infty} w_{m,n}(x,y)\,dy\,dx &= \int_{0}^{\pi}\int_{0}^{\infty} w_{m,n}(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta \\ &= \int_{0}^{\pi} e^{i(m-n)\theta} d\theta \int_{0}^{\infty} w_{m,n}(r,0)\,r\,dr,\end{aligned}$$

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{0} w_{m,n}(x,y)\,dy\,dx &= \int_{\pi}^{2\pi}\int_{0}^{\infty} w_{m,n}(r\cos(\theta), r\sin(\theta))\,r\,dr\,d\theta \\ &= \int_{\pi}^{2\pi} e^{i(m-n)\theta} d\theta \int_{0}^{\infty} w_{m,n}(r,0)\,r\,dr.\end{aligned}$$

When $m \leq n$,

$$\int_{0}^{\infty} w_{m,n}(r,0)\,r\,dr = \frac{(-1)^m}{\pi}\sqrt{\frac{m!}{n!}}\int_{0}^{\infty}\left(\sqrt{2}r\right)^{n-m} e^{-r^2} L_m^{(n-m)}\left(2r^2\right) r\,dr$$

and thus the $(m,n)^{\text{th}}$ matrix element simplifies to [12]

$$\gamma_{m-n}\frac{(-1)^{m+n}}{\pi}\sqrt{m!\,n!}\sum_{k=\max\{m,n\}}^{m+n}(-1)^k 2^{k-(m+n)/2-1}\frac{\Gamma(k-(m+n)/2+1)}{(m+n-k)!\,(k-m)!\,(k-n)!}$$

where

$$\gamma_j = \int_0^{\pi} e^{ij\theta}\, d\theta - \int_{\pi}^{2\pi} e^{ij\theta}\, d\theta = \begin{cases} 0 & \text{if } j \equiv 0 \bmod 2, \\ 4i/j & \text{if } j \equiv 1 \bmod 2. \end{cases}$$

The norm $\|\mathrm{sgn}(P)\|$ of the infinite matrix is found by numerically evaluating the largest eigenvalue of the upper left $N \times N$ submatrix of $\mathrm{sgn}(P)$ and letting $N \to \infty$.

The matrices $\mathrm{sgn}(Q)$ and $\mathrm{sgn}(R)$ are obtained similarly, with γ_{m-n} replaced by δ_{m-n} and ε_{m-n} respectively, where

$$\delta_j = \int_{-\pi/2}^{\pi/2} e^{ij\theta}\, d\theta - \int_{\pi/2}^{3\pi/2} e^{ij\theta}\, d\theta = \begin{cases} 0 & \text{if } j \equiv 0 \bmod 2, \\ 4/j & \text{if } j \equiv 1 \bmod 4, \\ -4/j & \text{if } j \equiv 3 \bmod 4 \end{cases}$$

and

$$\varepsilon_j = \int_{-5\pi/4}^{-\pi/4} e^{ij\theta}\, d\theta - \int_{-\pi/4}^{3\pi/4} e^{ij\theta}\, d\theta = \begin{cases} 0 & \text{if } j \equiv 0 \bmod 2, \\ 2\sqrt{2}(-1-i)/j & \text{if } j \equiv 1 \bmod 8, \\ 2\sqrt{2}(-1+i)/j & \text{if } j \equiv 3 \bmod 8, \\ 2\sqrt{2}(1+i)/j & \text{if } j \equiv 5 \bmod 8, \\ 2\sqrt{2}(1-i)/j & \text{if } j \equiv 7 \bmod 8. \end{cases}$$

Adding the three matrices and taking the largest eigenvalue, we obtain a limiting value ≈ 1.2 for the operator norm.

2.10.4 *Quantum Probability*

For convenience, define the **indicator function**

$$\mathrm{ind}(\xi) = \begin{cases} 1 & \text{if } \xi > 0, \\ 0 & \text{if } \xi < 0 \end{cases} = \frac{1}{2} + \frac{1}{2}\,\mathrm{sgn}(\xi).$$

Let $q\cos(t) + p\sin(t)$ denote the coordinate of a classical harmonic oscillator at time t, where q, p are the initial coordinate and momentum, and the period is 2π. Choose $\tau \in \{0, 2\pi/3, 4\pi/3\}$ at random. What is the probability that $q\cos(\tau) + p\sin(\tau) > 0$? Clearly this depends on the initial state and is given by

$$\begin{aligned} &\sum_{k=0}^{2} \mathrm{P}\left(q\cos(\tau) + p\sin(\tau) > 0 \,|\, \tau = \tfrac{2\pi k}{3}\right) \mathrm{P}\left(\tau = \tfrac{2\pi k}{3}\right) \\ &= \tfrac{1}{3}\left(\mathrm{ind}\,(q) + \mathrm{ind}\left(-\tfrac{1}{2}q + \tfrac{\sqrt{3}}{2}p\right) + \mathrm{ind}\left(-\tfrac{1}{2}q - \tfrac{\sqrt{3}}{2}p\right)\right) \\ &= \begin{cases} \tfrac{2}{3} & \text{if } \tfrac{\pi}{6} < \theta < \tfrac{\pi}{2} \text{ or } \tfrac{5\pi}{6} < \theta < \tfrac{7\pi}{6} \text{ or } -\tfrac{\pi}{2} < \theta < -\tfrac{\pi}{6}, \\ \tfrac{1}{3} & \text{if } -\tfrac{\pi}{6} < \theta < \tfrac{\pi}{6} \text{ or } \tfrac{\pi}{2} < \theta < \tfrac{5\pi}{6} \text{ or } -\tfrac{5\pi}{6} < \theta < -\tfrac{\pi}{2} \end{cases} \end{aligned}$$

where θ is the polar angle of (q, p) in the plane. Thus the solution is $\frac{1}{2} \pm \frac{1}{6}$.

Consider now the quantum harmonic oscillator $Q\cos(t)+P\sin(t)$ [13]. Answering the same question reduces to evaluating the spectral bounds of the operator

$$\tfrac{1}{2}I+\tfrac{1}{6}\left(\operatorname{sgn}(Q)+\operatorname{sgn}\left(-\tfrac{1}{2}Q+\tfrac{\sqrt{3}}{2}P\right)+\operatorname{sgn}\left(-\tfrac{1}{2}Q-\tfrac{\sqrt{3}}{2}P\right)\right),$$

which turn out to be

$$\tfrac{1}{2}\pm\tfrac{1}{6}c\approx\tfrac{1}{2}\pm 0.21.$$

The maximum probability ≈ 0.71 is calculated in [12] and is rigorously proved to be <1. We wonder if there are other such fascinating numbers in the intersection between functional analysis and quantum mechanics.

2.10.5 *Generalized Oscillator*

The Schrödinger ODE for the anharmonic oscillator:

$$\frac{d^2\psi}{dx^2}+\left(\lambda-x^4\right)\psi=0$$

with quartic potential cannot be solved in closed-form (unlike the harmonic oscillator). It is worth mentioning that the smallest eigenvalue is

$$\lambda_0=1.0603620904...$$

and this constant is now known to more than 1000 digits [14–17]. The corresponding eigenvalues for the sextic and octic potentials are 1.1448024537... and 1.2258201138... [18, 19]. See [20] for mention of the linear potential case.

Addendum Tsirelson [3] warned readers that he uses $PQ-QP=-iI$ merely as shorthand for the Weyl relations

$$\exp(i\alpha P)\exp(i\beta Q)=\exp(i\alpha\beta)\exp(i\beta Q)\exp(i\alpha P)\quad\text{for all }\alpha,\beta\in\mathbb{R}.$$

The consequential independence of $\|\operatorname{sgn}(P)+\operatorname{sgn}(Q)+\operatorname{sgn}(R)\|$ of the choice of P, Q, R follows from von Neumann's theorem [1].

He also offered the following explanation for §2.10.2: "The operator norm is the supremum of the corresponding quadratic form over the unit sphere. We may choose an increasing sequence of finite-dimensional subspaces whose union is dense, and consider the corresponding finite-dimensional suprema; they increase to the infinite-dimensional supremum. Thus the operator norm is the limit of an increasing sequence of matrix norms. A good choice of a basis (in the Hilbert space) simplifies the calculation of the matrices. We use the basis of eigenvectors of the Hamiltonian (of the oscillator). The calculation of the matrices may be made via the Wigner function."

[1] M. Reed and B. Simon, *Methods of Modern Mathematical Physics. I. Functional Analysis*, 2nd ed., Academic Press, 1980, pp. 47, 71, 249–250, 270–275, 302–305; MR0751959 (85e:46002).

[2] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, 2008, pp. 39–43, 107–132; MR2396439 (2009b:15001).

[3] B. Tsirelson, Automorphisms of the type II_1 Arveson system of Warren's noise, arXiv:math.OA/0612303.

[4] G. F. Simmons, *Differential Equations with Applications and Historical Notes*, McGraw-Hill, 1972, pp. 194–196; MR0499365 (58 #17258).

[5] C. A. Tracy, Lectures on Ordinary Differential Equations, unpublished notes (2007).

[6] S. A. Ponomarenko, Quantum harmonic oscillator revisited: A Fourier transform approach, *Amer. J. Phys.* 72 (2004) 1259–1260.

[7] A. Engel, Comment on 'Quantum harmonic oscillator revisited: A Fourier transform approach', *Amer. J. Phys.* 74 (2006) 837.

[8] H. J. Groenewold, On the principles of elementary quantum mechanics, *Physica* 12 (1946) 405–460; MR0018562 (8,301a).

[9] M. S. Bartlett and J. E. Moyal, The exact transition probabilities of quantum-mechanical oscillators calculated by the phase-space method, *Proc. Cambridge Philos. Soc.* 45 (1949) 545–553; MR0034709 (11,631e).

[10] T. Takabayasi, The formulation of quantum mechanics in terms of ensemble in phase space, *Progress Theoret. Physics* 11 (1954) 341–373; MR0064670 (16,319b).

[11] T. Curtright, T. Uematsu and C. Zachos, Generating all Wigner functions, *J. Math. Phys.* 42 (2001) 2396–2415; MR1821868 (2002f:81060).

[12] B. Tsirelson, How often is the coordinate of a harmonic oscillator positive? arXiv:quant-ph/0611147.

[13] P. Nardone, Heisenberg picture in quantum mechanics and linear evolutionary systems, *Amer. J. Phys.* 61 (1993) 232–237.

[14] M. Trott, High-precision value for the quartic anharmonic oscillator ground state, arXiv:quant-ph/0012147.

[15] H. A. Alhendi and E. I. Lashin, Spectrum of one-dimensional anharmonic oscillators, *Canad. J. Phys.* 83 (2005) 541–550; arXiv:quant-ph/0305128.

[16] P. Kościk and A. Okopińska, Application of the Frobenius method to the Schrödinger equation for a spherically symmetric potential: an anharmonic oscillator, *J. Phys. A* 38 (2005) 7743–7755; MR2169486 (2006j:81051).

[17] A. J. Zakrzewski, Highly precise solutions of the one-dimensional Schrödinger equation with an arbitrary potential, *Comput. Phys. Comm.* 175 (2006) 397–403; MR2260324 (2007d:81056).

[18] F. Vinette and J. Čížek, Upper and lower bounds of the ground state energy of anharmonic oscillators using renormalized inner projection, *J. Math. Phys.* 32 (1991) 3392–3404; MR1137393 (93m:81024).

[19] E. J. Weniger, J. Čížek, and F. Vinette, The summation of the ordinary and renormalized perturbation series for the ground state energy of the quartic, sextic, and octic anharmonic oscillators using nonlinear sequence transformations, *J. Math. Phys.* 34 (1993) 571–609; MR1201021 (93j:81018).

[20] S. R. Finch, Airy function zeroes, *this volume*, §2.5.

2.11 Mathieu Eigenvalues

Consider the differential equation [1–3]

$$y''(x) + (\lambda - 2\mu\cos(2x))\,y(x) = 0,$$

which admits periodic solutions of (least) period π and 2π for four countably infinite sets of eigenvalues, for each value of μ.

2.11.1 *Even Solutions of Period* π

Given boundary conditions $y'(0) = y'(\pi/2) = 0$, the eigenvalues $\lambda = \alpha_{2k}$ for $k \geq 0$ satisfy the infinite tridiagonal determinant equation [4]

$$\begin{vmatrix} 0^2-\lambda & \sqrt{2}\mu & 0 & 0 & 0 & \\ \sqrt{2}\mu & 2^2-\lambda & \mu & 0 & 0 & \\ 0 & \mu & 4^2-\lambda & \mu & 0 & \\ 0 & 0 & \mu & 6^2-\lambda & \mu & \\ 0 & 0 & 0 & \mu & 8^2-\lambda & \ddots \\ & & & & \ddots & \ddots \end{vmatrix} = 0$$

as well as the continued fraction equation [5]

$$-\frac{\lambda}{2} = \frac{\mu^2|}{|2^2-\lambda} - \frac{\mu^2|}{|4^2-\lambda} - \frac{\mu^2|}{|6^2-\lambda} - \frac{\mu^2|}{|8^2-\lambda} - \frac{\mu^2|}{|10^2-\lambda} - \cdots.$$

For example, if $\mu = 1$, then [6] $\alpha_0 = -0.4551386041...$ and $\alpha_2 = 4.3713009827....$ The corresponding eigenfunctions are written as $\mathrm{ce}_{2k}(x)$. Only for complex μ can the equality $\alpha_0 = \alpha_2$ occur; the first such example [7–11] happens when $\mu = (1.4687686137...)i$, at which $\alpha_0 = \alpha_2 = 2.0886989027....$

2.11.2 *Odd Solutions of Period* π

Given boundary conditions $y(0) = y(\pi/2) = 0$, the eigenvalues $\lambda = \beta_{2k+2}$ for $k \geq 0$ satisfy the infinite tridiagonal determinant equation

$$\begin{vmatrix} 2^2-\lambda & \mu & 0 & 0 & 0 & \\ \mu & 4^2-\lambda & \mu & 0 & 0 & \\ 0 & \mu & 6^2-\lambda & \mu & 0 & \\ 0 & 0 & \mu & 8^2-\lambda & \mu & \\ 0 & 0 & 0 & \mu & 10^2-\lambda & \ddots \\ 0 & 0 & 0 & 0 & \ddots & \ddots \end{vmatrix} = 0$$

as well as the continued fraction equation

$$4-\lambda = \frac{\mu^2|}{|4^2-\lambda} - \frac{\mu^2|}{|6^2-\lambda} - \frac{\mu^2|}{|8^2-\lambda} - \frac{\mu^2|}{|10^2-\lambda} - \frac{\mu^2|}{|12^2-\lambda} - \cdots.$$

For example, if $\mu = 1$, then $\beta_2 = 3.9170247729\ldots$ and $\beta_4 = 16.0329700814\ldots$. The corresponding eigenfunctions are written as $\mathrm{se}_{2k+2}(x)$. Only for complex μ can the equality $\beta_2 = \beta_4$ occur; the first such example [9–12] happens when $\mu = (6.9289547587\ldots)i$, at which $\beta_2 = \beta_4 = 11.1904735991\ldots$.

2.11.3 *Even Solutions of Period 2π*

Given boundary conditions $y'(0) = y(\pi/2) = 0$, the eigenvalues $\lambda = \alpha_{2k+1}$ for $k \geq 0$ satisfy the infinite tridiagonal determinant equation

$$\begin{vmatrix} 1+\mu-\lambda & \mu & 0 & 0 & 0 & \\ \mu & 3^2-\lambda & \mu & 0 & 0 & \\ 0 & \mu & 5^2-\lambda & \mu & 0 & \\ 0 & 0 & \mu & 7^2-\lambda & \mu & \\ 0 & 0 & 0 & \mu & 9^2-\lambda & \ddots \\ 0 & 0 & 0 & 0 & \ddots & \ddots \end{vmatrix} = 0$$

as well as the continued fraction equation

$$1+\mu-\lambda = \frac{\mu^2 \,|}{|3^2-\lambda} - \frac{\mu^2 \,|}{|5^2-\lambda} - \frac{\mu^2 \,|}{|7^2-\lambda} - \frac{\mu^2 \,|}{|9^2-\lambda} - \frac{\mu^2 \,|}{|11^2-\lambda} - \cdots.$$

For example, if $\mu = 1$, then $\alpha_1 = 1.8591080725\ldots$ and $\alpha_3 = 9.0783688472\ldots$. The corresponding eigenfunctions are written as $\mathrm{ce}_{2k+1}(x)$. Only for complex μ can the equality $\alpha_1 = \alpha_3$ occur; the first such example [9–11, 13] happens when

$$\mu = 1.93139250\ldots + (3.23763841\ldots)i = (3.7699574940\ldots)e^{i\theta},$$

$$\theta = \arccos(0.51231148\ldots) \approx 59.182°$$

at which

$$\alpha_1 = \alpha_3 = 6.17649\ldots + (1.23174\ldots)i.$$

2.11.4 *Odd Solutions of Period 2π*

Given boundary conditions $y(0) = y'(\pi/2) = 0$, the eigenvalues $\lambda = \beta_{2k+1}$ for $k \geq 0$ satisfy the infinite tridiagonal determinant equation

$$\begin{vmatrix} 1-\mu-\lambda & \mu & 0 & 0 & 0 & \\ \mu & 3^2-\lambda & \mu & 0 & 0 & \\ 0 & \mu & 5^2-\lambda & \mu & 0 & \\ 0 & 0 & \mu & 7^2-\lambda & \mu & \\ 0 & 0 & 0 & \mu & 9^2-\lambda & \ddots \\ 0 & 0 & 0 & 0 & \ddots & \ddots \end{vmatrix} = 0$$

as well as the continued fraction equation

$$1-\mu-\lambda=\frac{\mu^2}{|3^2-\lambda}-\frac{\mu^2}{|5^2-\lambda}-\frac{\mu^2}{|7^2-\lambda}-\frac{\mu^2}{|9^2-\lambda}-\frac{\mu^2}{|11^2-\lambda}-\cdots.$$

For example, if $\mu=1$, then $\beta_1=-0.1102488169...$ and $\beta_3=9.0477392598....$ The corresponding eigenfunctions are written as $\mathrm{se}_{2k+1}(x)$. No new constants emerge in connection with $\beta_1=\beta_3$ because $\beta_1(\mu)=\alpha_1(-\mu)$ and $\beta_3(\mu)=\alpha_3(-\mu)$; hence this case reduces to the preceding.

2.11.5 *Double Points*

The values $|\mu|=1.468...$, $6.928...$, $3.769...$ are first terms of the three sequences [10, 11]

- $\{a_k\}$, where $a_k=|\mu|$ and μ is the complex point closest to 0 satisfying $\alpha_{2k}(\mu)=\alpha_{2k+2}(\mu)$
- $\{b_k\}$, where $b_k=|\mu|$ and μ is the complex point closest to 0 satisfying $\beta_{2k+2}(\mu)=\beta_{2k+4}(\mu)$
- $\{c_k\}$, where $c_k=|\mu|$ and μ is the complex point closest to 0 satisfying $\alpha_{2k+1}(\mu)=\alpha_{2k+3}(\mu)$ if k is even and $\beta_{2k+1}(\mu)=\beta_{2k+3}(\mu)$ if k is odd.

It is conjectured (among other things) that

$$a_k\sim b_k\sim c_k$$

asymptotically as $k\to\infty$ and $a_k\approx(2.042)k^2$ for large k. Conceivably $\pi^{-1/4}e=2.04177...$ could be an exact expression for the leading coefficient [11]: no one knows.

2.11.6 *Hill and Ince*

Let n be a positive integer. Hill's equation is the following generalization [14]

$$y''(x)+\left(\lambda-2\sum_{j=1}^{n}\mu_j\cos(2jx)\right)y(x)=0$$

of Mathieu's equation (for which $n=1$ was assumed). A special case of Hill's equation is Ince's equation [4, 15]

$$y''(x)+c\sin(2x)y'(x)+(\lambda-\mu\,c\cos(2x))\,y(x)=0$$

after a suitable transformation (assuming here that $n=2$). Let λ denote the leftmost eigenvalue of the above. We merely mention that the derivatives $\lambda'(0)$ and $\lambda''(0)$ of the function $\mu\mapsto\lambda(\mu)$, for fixed c, play an interesting role in [16]. By contrast, $\alpha_0'(0)=0$ and $\alpha_0''(0)=-1$ for Mathieu's equation, which are comparatively straightforward.

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, pp. 721–750; MR1225604 (94b:00012).
[2] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions*, v. III, based on notes left by H. Bateman, McGraw-Hill, 1955, pp. 97–166; MR0066496 (16,586c).
[3] C. Hunter and B. Guerrieri, The eigenvalues of Mathieu's equation and their branch points, *Stud. Appl. Math.* 64 (1981) 113–141; MR0608595 (82c:34030).
[4] H. Volkmer, *The Ince and Lamé Differential Equations*, unpublished manuscript (2004).
[5] J. Meixner and F. W. Schäfke, *Mathieusche Funktionen und Sphäroidfunktionen mit Anwendungen auf physikalische und technische Probleme*, Springer-Verlag, 1954, pp. 117–121; MR0066500 (16,586g).
[6] T. Tamir, Characteristic exponents of Mathieu functions, *Math. Comp.* 16 (1962) 100–106; MR0135739 (24 #B1782).
[7] H. P. Mulholland and S. Goldstein, The characteristic numbers of the Mathieu equation with purely imaginary numbers, *Philos. Mag.* 8 (1929) 834–840.
[8] C. J. Bouwkamp, A note on Mathieu functions, *Nederl. Akad. Wetensch. Proc.* 51 (1948) 891–893; *Indag. Math.* 10 (1948) 319–321; MR0029008 (10,533b).
[9] G. Blanch and D. S. Clemm, The double points of Mathieu's differential equation, *Math. Comp*. 23 (1969) 97–108; MR0239727 (39 #1084).
[10] F. W. Schäfke and H. Groh, Zur Berechnung der Eigenwerte der Mathieuschen Differentialgleichung, *Numer. Math.* 4 (1962) 64–67; MR0150358 (27 #359).
[11] J. Meixner, F. W. Schäfke and G. Wolf, *Mathieu Functions and Spheroidal Functions and their Mathematical Foundations. Further Studies*, Springer-Verlag, Lect. Notes in Math. 837, 1980, pp. 85–88; MR0606934 (83b:33013).
[12] P. N. Shivakumar and J. Xue, On the double points of a Mathieu equation, *J. Comput. Appl. Math.* 107 (1999) 111–125; MR1698481 (2000c:34065).
[13] C. Hunter and B. Guerrieri, Deducing the properties of singularities of functions from their Taylor series coefficients, *SIAM J. Appl. Math.* 39 (1980) 248–263; erratum 41 (1981) 203; MR0588498 (82b:65015a) and MR0622882 (82b:65015b).
[14] F. M. Arscott, *Periodic Differential Equations*, Macmillan, 1964, pp. 141–152, 270–272; MR0173798 (30 #4006).
[15] E. L. Ince, A linear differential equation with periodic coefficients, *Proc. London Math. Soc.* 23 (1925) 56–74.
[16] S. R. Finch, CLT variance associated with Baxendale's SDE, arXiv:0809.5274.

2.12 Thomas–Fermi Model

The boundary value problem

$$y''(x) = x^{-1/2}y(x)^{3/2}, \quad y(0) = 1, \quad \lim_{x\to\infty} y(x) = 0$$

is an important model in atomic physics [1–4]. Two well-known series expansions for $y(x)$ are

$$y(x) = \sum_{k=0}^{\infty} p_k x^{k/2}, \quad x \approx 0$$

$$p_0 = 1, \quad p_1 = 0, \quad p_2 = -\xi, \quad p_3 = 4/3, \quad p_4 = 0, \quad \ldots$$

due to Baker [5] and

$$y(x) = \frac{144}{x^3} \sum_{k=0}^{\infty} q_k \eta^k x^{-\lambda k}, \quad x \approx \infty$$

$$q_0 = 1, \quad q_1 = -1, \quad \ldots$$

due to Coulson & March [6], where $\lambda = (-7 + \sqrt{73})/2$. The coefficient p_k is a polynomial in ξ; the coefficient $q_k \eta^k$ is (even more clearly) a polynomial in η. Hence it is important to compute

$$\xi = \lim_{x \to 0^+} \frac{1 - y(x)}{x} = -y'(0), \quad \eta = \lim_{x \to \infty} x^\lambda \left(1 - \frac{x^3}{144} y(x)\right)$$

as accurately as feasible.

More precisely, we have recursive formulas [7–9]

$$\begin{aligned} p_k = \frac{1}{(k-3)\left[(k-1)^2 - 1\right]} \Bigg\{ & \frac{3}{2} \sum_{j=1}^{k-4} (j+1) \left[(k-j-2)^2 - 1\right] p_{j+1} p_{k-j-1} \\ & - \sum_{j=0}^{k-6} (j+1) \left[(j+3)^2 - 1\right] p_{j+4} p_{k-j-4} \Bigg\} \end{aligned}$$

for $k \geq 5$ and

$$\begin{aligned} q_k = \frac{1}{(k-1)k(\lambda^2 k + 6)} \sum_{j=0}^{k-2} (j+1) \Bigg\{ & \frac{3}{2} \left[\lambda^2 (k-j-1)(k-j-2)\right. \\ & \left. + 6(k-j-1) + 12\right] - \lambda^2 j(j+1) - 6(j+1) - 12 \Bigg\} q_{j+1} q_{k-j-1} \end{aligned}$$

for $k \geq 2$. Special values include

$$p_5 = -\frac{2}{5}\xi, \quad p_6 = \frac{1}{3}, \quad p_7 = \frac{3}{70}\xi^2,$$

$$p_8 = -\frac{2}{15}\xi, \quad p_9 = \frac{4}{63}\left(\frac{7}{6} + \frac{1}{16}\xi^3\right), \quad p_{10} = \frac{1}{175}\xi^2$$

and

$$q_2 = \frac{201 + 21\sqrt{73}}{608}, \quad q_3 = -\frac{15377 + 1813\sqrt{73}}{98496}.$$

Such information, however, does not lead easily to numerical estimates of ξ or η. Various attempts to do this include [10–29]. We mention that the solution $y(x)$

minimizes the integral

$$I(\varphi) = \int_0^\infty \left(\frac{1}{2}\varphi'(x)^2 + \frac{2}{5}\frac{\varphi(x)^{5/2}}{x^{1/2}} \right) dx$$

subject to the constraints $\varphi(0) = 1$ and $\lim_{x\to\infty} \varphi(x) = 0$, and maximizes the integral

$$J(\psi) = -\int_0^\infty \left(\frac{1}{2}\psi'(x)^2 + \frac{3}{5}\frac{(x^{1/2}\psi''(x))^{5/3}}{x^{1/2}} \right) dx - \psi'(0)$$

with no essential constraints [30–32]. The extreme values of I and J agree:

$$J(\psi) \le J(y) = I(y) \le I(\varphi)$$

and thus the difference $I(z) - J(z)$ serves to measure how close a candidate function $z(x)$ is to $y(x)$.

2.12.1 *Majorana Transformation*

The following derivation of ξ, η was discovered in 1928 but remained unknown until 2008 [33, 34]. Write

$$t = 144^{-1/6} x^{1/2} y(x)^{1/6},$$

$$u = -\left(\frac{16}{3}\right)^{1/3} y(x)^{-4/3} y'(x),$$

then

$$\dot{u}(t) = 8\frac{t\,u(t)^2 - 1}{1 - t^2 u(t)}, \quad u(0) = \left(\frac{16}{3}\right)^{1/3} \xi, \quad u(1) = 1$$

and hence

$$u(t) = \sum_{m=0}^{\infty} a_m (1-t)^m,$$

where $a_0 = 1$, $a_1 = 9 - \sqrt{73}$ and

$$a_m = \frac{1}{2(m+8) - (m+1)a_1} \left\{ \sum_{n=1}^{m-2} \left[(n+1)a_{n+1} - 2(n+4)a_n + (n+7)a_{n-1} \right] a_{m-n} \right.$$
$$\left. + \left[(m+7) - 2(m+3)a_1 \right] a_{m-1} + (m+6)a_1 a_{m-2} \right\}.$$

It follows that

$$\xi = \left(\frac{3}{16}\right)^{1/3} \sum_{m=0}^{\infty} a_m = 1.5880710226....$$

We have

$$y(x) = \frac{144}{x^3} t^6,$$

$$x(t) = 144^{1/3} t^2 \exp\left[2 \int_0^t \frac{s\,u(s)}{1 - s^2 u(s)} ds\right]$$

and, further,

$$\int_0^t \frac{s\,u(s)}{1 - s^2 u(s)} ds = \int_{1-t}^1 \frac{\sum_{m=0}^\infty b_m \tau^m}{\sum_{m=0}^\infty c_m \tau^m} d\tau$$

where $b_0 = 1$, $c_0 = 0$,

$$b_m = a_m - a_{m-1}, \quad c_m = b_{m-1} - b_m \quad \text{for } m \geq 1.$$

It follows that

$$\eta = \lim_{t \to 1^-} x(t)^\lambda \left(1 - t^6\right) = 13.2709738480....$$

These numerical computations appear to be more straightforward than any other technique invented over the past eighty years!

A starting point for theory underlying the Thomas–Fermi equation can be found in [35–38]; see also [39] for a connection with counting lattice points within a planar closed curve.

[1] H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, Dover, 1962, pp. 405–407; MR0181773 (31 #6000).

[2] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, McGraw-Hill, 1978, pp. 167–170, 199–200; MR0538168 (80d:00030).

[3] I. M. Torrens, *Interatomic Potentials*, Academic Press, 1972, pp. 17–34, 223–226.

[4] B.-G. Englert, *Semiclassical Theory of Atoms*, Lect. Notes in Phys. 300 Springer-Verlag, 1988, pp. 37–39, 46–66.

[5] E. B. Baker, The application of the Fermi-Thomas statistical model to the calculation of potential distribution in positive ions, *Phys. Rev.* 36 (1930) 630–647.

[6] C. A. Coulson and N. H. March, Momenta in atoms using the Thomas-Fermi method, *Proc. Phys. Soc. London Sect. A* 63 (1950) 367–374.

[7] F. M. Fernández and J. F. Ogilvie, Approximate solutions to the Thomas-Fermi equation, *Phys. Rev. A* 42 (1990) 149–154.

[8] J. E. Enstrom, Convergence of the series expansion for the Thomas-Fermi function, *J. Chem. Phys.* 46 (1967) 1213–1214.

[9] P. J. Rijnierse, *Algebraic Solutions of the Thomas-Fermi Equation for Atoms*, Ph.D. thesis, Univ. of St. Andrews Scotland, 1968.

[10] V. Bush and S. H. Caldwell, Thomas-Fermi equation solution by the differential analyzer, *Phys. Rev.* 38 (1931) 1898–1902.

[11] C. Miranda, Teoremi e metodi per l'integrazione numerica della equazione differenziale di Fermi, *Memorie della classe di Scienze Fisiche, Matematiche e Naturali, Reale Accademia d'Italia*, v. 5 (1934) n. 7, 285–322.

[12] J. C. Slater and H. M. Krutter, The Thomas-Fermi method for metals, *Phys. Rev.* 47 (1935) 559–568.
[13] R. P. Feynman, N. Metropolis, and E. Teller, Equations of state of elements based on the generalized Fermi-Thomas theory, *Phys. Rev.* 75 (1949) 1561–1573.
[14] S. Kobayashi, T. Matsukuma, S. Nagai and K. Umeda, Accurate value of the initial slope of the ordinary TF function, *J. Phys. Soc. Japan* 10 (1955) 759–762.
[15] H. Krutter, Numerical integration of the Thomas-Fermi equation from zero to infinity, *J. Comput. Phys.* 47 (1982) 308–312.
[16] G. I. Plindov and S. K. Pogrebnya, The analytical solution of the Thomas-Fermi equation for a neutral atom, *J. Phys. B* 20 (1987) L547–L550.
[17] C. M. Bender, K. A. Milton, S. S. Pinsky and L. M. Simmons, A new perturbative approach to nonlinear problems, *J. Math. Phys.* 30 (1989) 1447–1455; MR1002247 (90i:34085).
[18] B. J. Laurenzi, An analytic solution to the Thomas-Fermi equation, *J. Math. Phys.* 31 (1990) 2535–2537; MR1072970 (92g:81033).
[19] K. Tu, Analytic solution to the Thomas-Fermi and Thomas-Fermi-Dirac-Weizsäcker equations, *J. Math. Phys.* 32 (1991) 2250–2253; MR1123619.
[20] A. J. MacLeod, Chebyshev series solution of the Thomas-Fermi equation, *Comp. Phys. Comm.* 67 (1992) 389–391.
[21] A. Cedillo, A perturbative approach to the Thomas-Fermi equation in terms of the density, *J. Math. Phys.* 34 (1993) 2713–2717; MR1224172 (94h:81028).
[22] S. N. Venkatarangan and K. Rajalakshmi, Modification of Adomian's decomposition method to solve equations containing radicals, *Comput. Math. Appl.* 29 (1995) 75–80; MR1317861.
[23] L. N. Epele, H. Fanchiotti, C. A. García Canal and J. A. Ponciano, Padé approximant approach to the Thomas-Fermi problem, *Phys. Rev. A* 60 (1999) 280–283.
[24] A.-M. Wazwaz, The modified decomposition method and Padé approximants for solving the Thomas-Fermi equation, *Appl. Math. Comput.* 105 (1999) 11–19; MR1706059 (2000d:65120).
[25] S. Liao, An explicit analytic solution to the Thomas-Fermi equation, *Appl. Math. Comput.* 144 (2003) 495–506; MR1994087 (2004e:34002).
[26] B. K. Datta, Approximate solution for the Thomas-Fermi equation, *Nuovo Cimento B* 121 (2006) 501–504.
[27] H. Khan and H. Xu, Series solution to the Thomas-Fermi equation, *Phys. Lett. A* 365 (2007) 111–115.
[28] F. M. Fernández, Comment on: "Series solution to the Thomas-Fermi equation", *Phys. Lett. A* 372 (2008) 5258–5260.
[29] B. Yao, A series solution to the Thomas-Fermi equation, *Appl. Math. Comput.* 203 (2008) 396–401.
[30] T. Ikebe and T. Kato, Application of variational method to the Thomas-Fermi equation, *J. Phys. Soc. Japan* 12 (1957) 201–203.
[31] R. Bellman, Dynamic programming and the variational solution of the Thomas-Fermi equation, *J. Phys. Soc. Japan* 12 (1957) 1049.
[32] N. Anderson and A. M. Arthurs, Variational solutions of the Thomas-Fermi equation, *Quart. Appl. Math.* 39 (1981/82) 127–129; MR0613956 (82g:81009).
[33] S. Esposito, Majorana solution of the Thomas-Fermi equation, *Amer. J. Phys*. 70 (2002) 852–856; arXiv:physics/0111167.
[34] F. Guerra and N. Robotti, Ettore Majorana's forgotten publication on the Thomas-Fermi model, *Phys. Perspect.* 10 (2008) 56–76; arXiv:physics/0511222.

[35] A. Mambriani, Su un teorema relativo alle equazioni differenziali ordinarie del 2° ordine, *Rendiconti della Reale Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali* 9 (1929) 620–622.
[36] G. Scorza-Dragoni, Su un'equazione differenziale particolare, *Rendiconti della Reale Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali* 9 (1929) 623–625.
[37] E. Hille, On the Thomas-Fermi equation, *Proc. Nat. Acad. Sci. U.S.A.* 62 (1969) 7–10; MR0243145 (39 #4469).
[38] E. Hille, Some aspects of the Thomas-Fermi equation, *J. d'Analyse Math.* 23 (1970) 147–170; MR0279376 (43 #5098).
[39] A. Córdoba, C. L. Fefferman and L. A. Seco, A number-theoretic estimate for the Thomas-Fermi density, *Comm. Partial Differential Equations* 21 (1996) 1087–1102; MR1399191 (97f:81181).

2.13 Prandtl–Blasius Flow

The boundary value problem

$$y'''(x) + y''(x)y(x) = 0, \quad y(0) = 0, \quad y'(0) = 0, \quad \lim_{x\to\infty} y'(x) = 1$$

arises in the study of two-dimensional incompressible viscous flow past a thin semi-infinite flat plate [1–4]. Such an equation is similar to the Thomas–Fermi equation [5], but is even more difficult to solve (because it is of higher order).

A well-known series for $y(x)$ is

$$y(x) = \sum_{k=0}^{\infty} (-1)^k \frac{p_k \xi^{k+1}}{(3k+2)!} x^{3k+2}, \quad x \approx 0,$$

where $p_0 = 1$ and [6]

$$p_k = \sum_{j=0}^{k-1} \binom{3k-1}{3j} p_j p_{k-j-1}, \quad k \geq 1.$$

Hence it is important to compute

$$\xi = \lim_{x\to 0^+} \frac{y(x) - 0}{x^2/2} = y''(0)$$

as accurately as feasible. Blasius' series has only a finite radius of convergence [7–11]:

$$\rho = \lim_{k\to\infty} \left(\frac{(3k)(3k+1)(3k+2)p_{k-1}}{p_k \xi} \right)^{1/3} = 4.0234644935...$$

(in fact, the associated singularities are at $x = -\rho$ and $\rho \exp(\pm i\pi/3)$). Unlike the Thomas–Fermi equation, an efficient transformation for the Blasius equation is not yet known that permits high-precision estimates of ξ. A Runge–Kutta numerical ODE solver gives $\xi = 0.4695999883...$, as well as [2, 3, 10, 11]

$$\eta = \lim_{x\to\infty} (x - y(x)) = 1.2167806216....$$

The fluid dynamics literature is somewhat bewildering because of (small) variations in the presentation of Blasius' equation. Let us generalize our discussion to clear up any confusion. Consider

$$z'''(x) + a\,z''(x)z(x) = 0, \quad z(0) = 0, \quad z'(0) = 0, \quad \lim_{x\to\infty} z'(x) = b$$

where $a > 0$, $b > 0$. Let $c = z''(0)$; it can be easily shown that $c = a^{1/2}b^{3/2}\xi$ and thus

$$c(a = 1/2, b = 1) = \xi/\sqrt{2} = 0.3320573362...,$$

$$c(a = 1, b = 2) = 2\sqrt{2}\xi = 1.3282293448... = 2(0.6641146724...),$$

$$b(a = 1, c = 1) = \xi^{-2/3} = 1.6551903602...,$$

$$b(a = 1/2, c = 1) = 2^{1/3}\xi^{-2/3} = 2.0854091764....$$

From formulas for the radius of convergence

$$R = \frac{\rho}{(a\,b)^{1/2}} = \left(\frac{\xi}{a\,c}\right)^{1/3} \rho$$

and for the limit

$$L = \lim_{x\to\infty} (b\,x - z(x)) = \left(\frac{b}{a}\right)^{1/2} \eta,$$

we obtain

$$R(a = 1/2, b = 1) = \sqrt{2}\rho = 5.6900380545...,$$

$$R(a = 1, c = 1) = \xi^{1/3}\rho = 3.1273479155...,$$

$$R(a = 2, c = 1) = (\xi/2)^{1/3}\rho = 2.4821776854...$$

(long ago Weyl [12, 13] gave bounds 2.08 and 3.11 for the latter) and

$$L(a = 1, b = 2) = L(a = 1/2, b = 1) = \sqrt{2}\eta = 1.7207876575...,$$

$$L(a = 2, b = 1) = L(a = 1, b = 1/2) = \eta/\sqrt{2} = 0.8603938287....$$

When moving fluid encounters a solid, a layer is formed adjacent to the boundary of the solid. Strong frictional effects exist inside this layer; on the outside, by contrast, the flow essentially displays no viscosity [2, 14]. For the case of a thin plate, the fluid velocity changes rapidly from zero (along the plate) to its original value (beyond the boundary layer). Three relevant quantities in this physical model are the *displacement thickness*

$$\delta_1 = \int_0^\infty (1 - y'(x))\,dx = \eta = 1.2167806216...,$$

the *momentum thickness*

$$\delta_2 = \int_0^\infty y'(x)\,(1 - y'(x))\,dx = \xi = 0.4695999883...$$

and the *energy thickness*

$$\delta_3 = \int_0^\infty y'(x)\left(1 - y'(x)^2\right)dx = 2\int_0^\infty y''(x)y'(x)y(x)dx = 0.73848498....$$

It is also known that [2, 10]

$$y''(x) \sim \kappa \exp\left[-(x-\eta)^2/2\right]$$

as $x \to \infty$, where $\kappa = 0.3305407719... = (0.2337276212...)\sqrt{2}$. We wonder whether δ_3 and κ are closely related. The literature associated with $y(x)$ is massive [15–54].

2.13.1 *Falkner–Skan Equation*

Consider

$$y'''(x) + y''(x)y(x) + \lambda\left(1 - y'(x)^2\right) = 0, \quad y(0) = 0, \quad y'(0) = 0, \quad \lim_{x\to\infty} y'(x) = 1,$$

which arises in the study of viscous flow past a wedge of angle $\lambda\pi$, $0 \le \lambda \le 1$. The special case $\lambda = 0$ is Blasius' equation, in which the wedge reduces to a flat plate. The special case $\lambda = 1/2$ is called Homann's equation; we here have [2, 3, 55–57]

$$y''(0) = 0.92768003... = (1.31193769...)/\sqrt{2}, \quad \lim_{x\to\infty}(x - y(x)) = 0.804548....$$

The special case $\lambda = 1$ is called Hiemenz's equation (corresponding to stagnation flow, for example, past a large disk); we here have [2, 3, 58, 59]

$$y''(0) = 1.23258765..., \quad \lim_{x\to\infty}(x - y(x)) = 0.647900....$$

It is known that a smooth solution $y(x)$ exists and is unique [12, 13, 60–65] for each λ, $0 \le \lambda \le 1$. An especially simple proof for $\lambda = 0$, due to Serrin, appears in [66, 67].

Physically relevant solutions also exist for negative λ, more precisely, in the range $-0.19883768... = \mu \le \lambda < 0$. (Positive λ corresponds to flow toward the wedge; negative λ corresponds to flow away from the wedge.) By "physically relevant", we mean that a solution $y(x)$ further satisfies

$$0 < y'(x) < 1 \quad \text{for all } x > 0,$$

$$1 - y'(x) = O\left(e^{-\gamma x}\right) \quad \text{as } x \to \infty$$

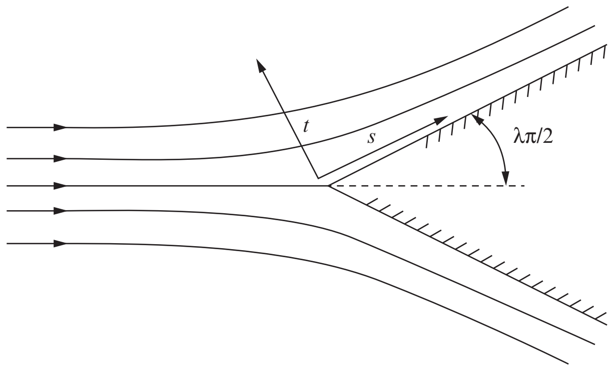

Figure 2.1 Falkner–Skan flow past a wedge.

for some $\gamma > 0$. It follows that $y''(0) > 0$ when $\lambda > \mu$ and $y''(0) = 0$ when $\lambda = \mu$. A deeper understanding of the constant μ is desired [17, 58, 68–79]. Again, the associated literature is massive [18–21, 80–102].

2.13.2 *Streamlines*

At each point in the first quadrant of (s, t)-space, define a velocity vector (u, v) by

$$u(s, t) = s^m y'(\theta),$$

$$v(s, t) = -\sqrt{\frac{s^{m-1}}{2(2-\lambda)}}\,((\lambda - 1)\theta\, y'(\theta) + y(\theta))$$

where

$$m = \frac{\lambda}{2-\lambda}, \quad \theta = t\sqrt{(m+1)s^{m-1}}.$$

The vector field $(s, t) \mapsto (u, v)$ determines the streamlines for laminar boundary-layer fluid flow past a wedge, as suggested in Figure 2.1, for a specified viscosity coefficient [3, 83].

Recent papers devoted to the Blasius ODE include [103–106]; related examples (with numerical estimates) are discussed in [107, 108].

[1] H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, Dover, 1962, pp. 400–404; MR0181773 (31 #6000).
[2] L. Rosenhead, *Laminar Boundary Layers*, Oxford Univ. Press, 1963, pp. 198–292, 419–420; MR0155499 (27 #5433).
[3] H. Schlichting, *Boundary-Layer Theory*, 7th ed., McGraw-Hill, 1979, pp. 95–101, 127–149, 163–170; MR0122222 (22 #12948).
[4] F. M. White, *Viscous Fluid Flow*, 2nd ed., McGraw-Hill, 1991, pp. 233–253, 504–520.
[5] S. R. Finch, Thomas-Fermi model, *this volume*, §2.12.

[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A018893.
[7] A. Ostrowski, Sur le rayon de convergence de la série de Blasius, *C. R. Acad. Sci. Paris* 227 (1948) 580–582; MR0027406 (10,300c).
[8] D. Shanks, The Blasius and Weyl constants in boundary-layer theory, *Phys. Rev.* 90 (1953) 377.
[9] B. Punnis, Zur Differentialgleichung der Plattengrenzschicht von Blasius, *Arch. Math.* 7 (1956) 165–171; abstract, *Z. Angew. Math. Mech.* 36 (1956) S26; MR0079450 (18,88a).
[10] J. P. Boyd, The Blasius function in the complex plane, *Experim. Math.* 8 (1999) 381–394; MR1737233 (2000i:34100).
[11] J. P. Boyd, The Blasius function: computations before computers, the value of tricks, undergraduate projects, and open research problems, *SIAM Rev.* 50 (2008) 791–804; MR2460804 (2009i:65237).
[12] H. Weyl, Concerning the differential equations of some boundary layer problems, *Proc. Nat. Acad. Sci. U. S. A.* 27 (1941) 578–583; 28 (1942) 100–102; MR0005849 (3,220c) and MR0005850 (3,221a).
[13] H. Weyl, On the differential equations of the simplest boundary-layer problems. *Annals of Math.* 43 (1942) 381–407; MR0006294 (3,284d).
[14] J. D. Anderson, Ludwig Prandtl's boundary layer, *Physics Today* 58 (2005) 42–48.
[15] K. Töpfer, Bemerkung zu dem Aufsatz von H. Blasius "Grenzschichten in Flüssigkeiten mit kleiner Reibung", *Z. Math. Phys.* 60 (1912) 397–398.
[16] S. Goldstein, Concerning some solutions of the boundary-layer equations in hydrodynamics, *Proc. Camb. Philos. Soc.* 26 (1930) 1–30.
[17] D. R. Hartree, On an equation occurring in Falkner and Skan's approximate treatment of the equations of the boundary layer, *Proc. Cambridge Philos. Soc.* 33 (1937) 223–239.
[18] L. Howarth, On the solution of the laminar boundary layer equations, *Proc. Royal Soc. London Ser. A* 164 (1938) 547–579.
[19] A. M. O. Smith, Rapid laminar boundary-layer calculations by piecewise application of similar solutions, *J. Aeronaut. Sci.* 23 (1956) 901–912.
[20] D. Meksyn, Integration of the boundary-layer equations, *Proc. Roy. Soc. London. Ser. A* 237 (1956) 543–559; MR0083317 (18,693f).
[21] D. Meksyn, *New Methods in Laminar Boundary-Layer Theory*, Pergamon Press, 1961, pp. 40–46, 52–96.
[22] A. J. Callegari and M. B. Friedman, An analytical solution of a nonlinear, singular boundary value problem in the theory of viscous fluids, *J. Math. Anal. Appl.* 21 (1968) 510–529; MR0224331 (36 #7375).
[23] S. Richardson, On Blasius's equation governing flow in the boundary layer on a flat plate, *Proc. Cambridge Philos. Soc.* 74 (1973) 179–184; MR0386437 (52 #7291).
[24] J. Y. Parlange, R. D. Braddock and G. Sander, Analytical approximations to the solution of the Blasius equation, *Acta Mechanica* 38 (1981) 119–125.
[25] C. M. Bender, K. A. Milton, S. S. Pinsky and L. M. Simmons, A new perturbative approach to nonlinear problems, *J. Math. Phys.* 30 (1989) 1447–1455; MR1002247 (90i:34085).
[26] P. L. Sachdev, *Nonlinear Ordinary Differential Equations and their Applications*, Dekker, 1991, pp. 56–57, 100–101, 218–220; MR1109918 (92i:34010).
[27] R. Fazio, The Blasius problem formulated as a free boundary value problem, *Acta Mechanica* 95 (1992) 1–7.
[28] J. P. Boyd, Padé approximant algorithm for solving nonlinear ordinary differential equation boundary value problems on an unbounded domain, *Computers in Physics* 11 (1997) 299–303.

[29] L.-T.Yu and C.-K. Chen, The solution of the Blasius equation by the differential transformation method, *Math. Comput. Modelling* 28 (1998) 101–111; MR1638616 (99d:34017).
[30] S.-J. Liao, A kind of approximate solution technique which does not depend upon small parameters. II, An application in fluid mechanics, *Internat. J. Nonlinear Mech.* 32 (1997) 815–822.
[31] S.-J. Liao, A uniformly valid analytic solution of two-dimensional viscous flow over a semi-infinite flat plate, *J. Fluid Mech.* 385 (1999) 101–128; MR1690937 (2000a:76057).
[32] S.-J. Liao, An explicit, totally analytic approximate solution for Blasius viscous flow problems, *Internat. J. Nonlinear Mech.* 34 (1999) 759–778.
[33] J. Lin, A new approximate iteration solution of Blasius equation, *Commun. Nonlinear Sci. Numer. Simul.* 4 (1999) 91–99; MR1723812.
[34] L. Roman-Miller and P. Broadbridge, Exact integration of reduced Fisher's equation, reduced Blasius equation, and the Lorenz model, *J. Math. Anal. Appl.* 251 (2000) 65–83; MR1790398 (2001f:34008).
[35] G. R. Liu and T. Y. Wu, Application of generalized differential quadrature rule in Blasius and Onsager equations, *Internat. J. Numer. Methods Engrg.* 52 (2001) 1013–1027.
[36] J.-H. He, A simple perturbation approach to Blasius equation, *Appl. Math. Comput.* 140 (2003) 217–222; MR1953895 (2003m:34039).
[37] B. K. Datta, Analytic solution for the Blasius equation, *Indian J. Pure Appl. Math.* 34 (2003) 237–240; MR1964524 (2004a:34028).
[38] L. Wang, A new algorithm for solving classical Blasius equation, *Appl. Math. Comput.* 157 (2004) 1–9; MR2085519 (2005h:65113).
[39] I. Hashim, Comments on: "A new algorithm for solving classical Blasius equation", *Appl. Math. Comput.* 176 (2006) 700–703; MR2232061.
[40] R. Cortell, Numerical solutions of the classical Blasius flat-plate problem, *Appl. Math. Comput.* 170 (2005) 706–710; MR2177573 (2006d:76033).
[41] D. E. Panayotounakos, N. B. Sotiropoulos, A. B. Sotiropoulou and N. D. Panayotounakou, Exact analytic solutions of nonlinear boundary value problems in fluid mechanics (Blasius equations), *J. Math. Phys.* 46 (2005) 033101; MR2125571 (2005j:76031).
[42] R. Iacono, Comment on: "Exact analytic solutions of nonlinear boundary value problems in fluid mechanics", *J. Math. Phys.* 47 (2006) 034101; MR2219808 (2006m:76046).
[43] T. Fang, F. Guo, C. F. Lee, A note on the extended Blasius equation, *Appl. Math. Lett.* 19 (2006) 613-617; MR2224413.
[44] A.-M. Wazwaz, The variational iteration method for solving two forms of Blasius equation on a half-infinite domain, *Appl. Math. Comput.* 188 (2007) 485–491; MR2327376.
[45] F. Ahmad, Application of Crocco-Wang equation to the Blasius problem, *Technical Acoustics*, 2007, n. 2.
[46] F. Ahmad, Degeneracy in the Blasius problem, *Elec. J. Differential Equations* 2007, N79; MR2308879 (2008d:34044).
[47] F. Ahmad and W. H. Al-Barakati, An approximate analytic solution of the Blasius problem, *Commun. Nonlinear Sci. Numer. Simul.* 14 (2009) 1021–1024; MR2468935.

[48] S. Abbasbandy, A numerical solution of Blasius equation by Adomians decomposition method and comparison with homotopy perturbation method, *Chaos Solitons Fractals* 31 (2007) 257–260.

[49] D. Bichsel and P. Wittwer, Stationary flow past a semi-infinite flat plate: analytical and numerical evidence for a symmetry-breaking solution, *J. Stat. Phys.* 127 (2007) 133–170; MR2313064 (2008d:76022).

[50] M. Benlahsen, M. Guedda and R. Kersner, The generalized Blasius equation revisited, *Math. Comput. Modelling* 47 (2008) 1063–1076; MR2413736 (2009e:76055).

[51] T. Tajvidi, M. Razzaghi and M. Dehghan, Modified rational Legendre approach to laminar viscous flow over a semi-infinite flat plate, *Chaos Solitons Fractals* 35 (2008) 59–66; MR2355238.

[52] A. Alizadeh-Pahlavan and S. Borjian-Boroujen, On the analytical solution of viscous fluid flow past a flat plate, *Phys. Lett. A* 372 (2008) 3678–3682.

[53] C.-W. Chang, J.-R. Chang and C.-S. Liu, The Lie-group shooting method for solving classical Blasius flat-plate problem, *Computers Materials & Continua* 7 (2008) 139–153; MR2437465 (2009h:76076).

[54] R. Fazio, Transformation methods for the Blasius problem and its recent variants, *Proc. World Congress on Engineering 2008*, v. II, 2008, pp. 938–943.

[55] H. Nasr, I. A. Hassanien and H. M. El-Hawary, Chebyshev solution of laminar boundary layer flow, *Internat. J. Comput. Math.* 33 (1990) 127–132.

[56] D. F. Rogers, *Laminar Flow Analysis*, Cambridge Univ. Press, 1992, pp. 53–104.

[57] R. Fazio, A novel approach to the numerical solution of boundary value problems on infinite intervals, *SIAM J. Numer. Anal.* 33 (1996) 1473–1483; MR1403554 (97e:65069).

[58] H. L. Evans, *Laminar Boundary-Layer Theory*, Addison-Wesley, 1968, pp. 20–26, 52–73.

[59] C. A. Forbrich, Improved solutions to the Falkner-Skan boundary-layer equation, *AIAA J.* 20 (1982) 1306–1307.

[60] W. A. Coppel, On a differential equation of boundary-layer theory, *Philos. Trans. Roy. Soc. London. Ser. A* 253 (1960) 101–136; MR0114966 (22 #5775).

[61] P. Hartman, *Ordinary Differential Equations*, Wiley, 1964, pp. 519–537; MR0171038 (30 #1270).

[62] A. H. Craven and L. A. Peletier, On the uniqueness of solutions of the Falkner-Skan equation, *Mathematika* 19 (1972), 129–133. MR0326054 (48 #4400a).

[63] E. F. F. Botta, F. J. Hut and A. E. P. Veldman, The role of periodic solutions in the Falkner-Skan problem for $\lambda > 0$, *J. Engineering Math.* 20 (1986) 81–93; MR0835374 (87d:76076).

[64] W. R. Utz, Existence of solutions of a generalized Blasius equation, *J. Math. Anal. Appl.* 66 (1978) 55–59; MR0513485 (81g:34030).

[65] W. R. Utz, The existence of solutions of a third-order boundary value problem, *Nonlinear Anal.* 5 (1981) 445–448; MR0611655 (82d:34024).

[66] R. E. Meyer, *Introduction to Mathematical Fluid Dynamics*, Wiley, 1971, pp. 104–106; MR0366189 (51 #2437).

[67] H. K. Wilson, *Ordinary Differential Equations*, Addison-Wesley, 1971, pp. 23–24, 254–255; MR0280764 (43 #6483).

[68] K. K. Chen and P. A. Libby, Boundary layers with small departures from the Falkner-Skan profile, *J. Fluid Mech.* 33 (1968) 273–282.

[69] K. Stewartson, Further solutions of the Falkner-Skan equation, *Proc. Cambridge Philos. Soc.* 50 (1954); 454–465; MR0062572 (15,999b).
[70] R. Iglisch and F. Kemnitz, Über die in der Grenzschichttheorie auftretende Differentialgleichung $f''' + ff'' + \beta(1 - f'^2) = 0$ für $\beta < 0$ bei gewissen Absauge- und Ausblatzgesetzen, *50 Jahre Grenzschichtforschung*, ed. H. Görtler und W. Tollmien, Vieweg, 1955, pp. 34–46.
[71] P. Hartman, On the asymptotic behavior of solutions of a differential equation in boundary layer theory, *Z. Angew. Math. Mech.* 44 (1964) 123–128; MR0175437 (30 #5621).
[72] S. N. Brown and K. Stewartson, On the reversed flow solutions of the Falkner-Skan equation, *Mathematika* 13 (1966) 1–6; MR0204011 (34 #3858).
[73] P. A. Libby and T. M. Liu, Further solutions of the Falkner-Skan equation, *AIAA J.* 5 (1967) 1040–1042.
[74] N. Riley and P. D. Weidman, Multiple solutions of the Falkner-Skan equation for flow past a stretching boundary, *SIAM J. Appl. Math.* 49 (1989) 1350–1358; MR1015066 (90g:76051).
[75] S. P. Hastings, Reversed flow solutions of the Falkner-Skan equation, *SIAM J. Appl. Math.* 22 (1972) 329–334; MR0306632 (46 #5754).
[76] B. Oskam and A. E. P. Veldman, Branching of the Falkner-Skan solutions for $\lambda < 0$, *J. Engineering Math.* 16 (1982) 295–308; MR0684010 (83m:76032).
[77] G. C. Yang and K. Q. Lan, The velocity and shear stress functions of the Falkner-Skan equation arising in boundary layer theory, *J. Math. Anal. Appl.* 328 (2007) 1297–1308; MR2290054 (2008b:76051).
[78] G. C. Yang and J. C. Hu, New numerical solutions of the Falkner-Skan equation using equivalent equation, *Nonlinear Anal. Forum* 12 (2007) 235–240; MR2404200 (2009j:65158).
[79] K. Q. Lan and G. C. Yang, Positive solutions of the Falkner-Skan equation arising in the boundary layer theory, *Canad. Math. Bull.* 51 (2008) 386–398; MR2436929 (2009e:34051).
[80] A. Aziz and T. Y. Na, New approach to the solution of Falkner-Skan equation, *AIAA J.* 19 (1981) 1242–1244.
[81] N. Afzal, Improved series solutions of Falkner-Skan equation, *AIAA J.* 23 (1985) 969–971; comment and reply 24 (1986) 863–864; MR0789856 (86h:76030) and MR0838167 (87h:76051).
[82] L. Yu and L. Yili, New series expansion method for the solution of the Falkner-Skan equation, *AIAA J.* 27 (1989) 1453–1455.
[83] I. Sher and A. Yakhot, New approach to solution of the Falkner-Skan equation, *AIAA J.* 39 (2001) 965–967; comments 40 (2002) 1475 and 42 (2004) 2620–2621.
[84] T. Cebeci and H. B. Keller, Shooting and parallel shooting methods for solving the Falkner-Skan boundary-layer equation, *J. Comput. Phys.* 7 (1971) 289–300.
[85] T. Y. Na, *Computational Methods in Engineering Boundary Value Problems*, Academic Press, 1979, pp. 138–148, 249–252, 261–264, 280–286, 303–304; MR0561710 (82g:65001).
[86] A. R. Wadia and F. R. Payne, A boundary value technique for the analysis of laminar boundary layer flows, *Internat. J. Comput. Math.* 9 (1981) 163–172; MR0607258 (82b:65073).
[87] P. M. Beckett, Finite difference solutions of boundary-layer type equations, *Internat. J. Comput. Math.* 14 (1983) 183–190; MR0723557 (84k:65082).
[88] H. T. Sharp and W. L. Harris, A pseudo-spectral method and parametric differentiation applied to the Falkner-Skan equation, *J. Comput. Phys.* 55 (1984) 377–386.

[89] C. Laine and L. Reinhart, Further numerical methods for the Falkner-Skan equations: shooting and continuation techniques, *Internat. J. Numer. Methods Fluids* 4 (1984) 833–852; MR0760001 (85k:76004).
[90] Z. U. A. Warsi and G. P. Koomullil, Numerical solution of the Falkner-Skan equation – a fresh look, *Ganita* 41 (1990) 49–60 (1991); MR1155951.
[91] T. M. El-Gindy, H. M. El-Hawary and H. S. Hussien, An optimization technique for the Falkner-Skan equation, *Optimization* 35 (1995) 357–366; MR1396230 (97e:65068).
[92] A. Asaithambi, A finite-difference method for the Falkner-Skan equation, *Appl. Math. Comput.* 92 (1998) 135–141; MR1625529.
[93] A. Asaithambi, A second-order finite-difference method for the Falkner-Skan equation, *Appl. Math. Comput*. 156 (2004) 779–786; MR2088620.
[94] A. Asaithambi, Numerical solution of the Falkner-Skan equation using piecewise linear functions, *Appl. Math. Comput.* 159 (2004) 267–273; MR2094970 (2005g:65164).
[95] A. Asaithambi, Solution of the Falkner-Skan equation by recursive evaluation of Taylor coefficients, *J. Comput. Appl. Math.* 176 (2005) 203–214; MR2108109 (2005k:34108).
[96] M. B. Zaturska and W. H. H. Banks, A new solution branch of the Falkner-Skan equation, *Acta Mechanica* 152 (2001) 197–201.
[97] A. A. Salama, Higher-order method for solving free boundary-value problems, *Numer. Heat. Transfer Part B* 45 (2004) 385–394.
[98] W. A. Albarakati, An analytic solution of the stagnation point flow problem, *Technical Acoustics*, 2007, n. 21.
[99] C.-S. Liu and J.-R. Chang, The Lie-group shooting method for multiple-solutions of Falkner-Skan equation under suction-injection conditions, *Internat. J. Nonlinear Mech.* 43 (2008) 844–851.
[100] T. Fang and J. Zang, An exact analytical solution of the Falkner-Skan equation with mass transfer and wall stretching, *Internat. J. Nonlinear Mech.* 43 (2008) 1000–1006.
[101] N. S. Elgazery, Numerical solution for the Falkner-Skan equation, *Chaos Solitons & Fractals* 35 (2008) 738–746.
[102] P. L. Sachdev, R. B. Kudenatti and N. M. Bujurke, Exact analytic solution of a boundary value problem for the Falkner-Skan equation, *Stud. Appl. Math.* 120 (2008) 1–16; MR2382924 (2008k:34106).
[103] F. Ahmad and W. A. Albarakati, Application of Padé approximation to solve the Blasius problem, *Proc. Pakistan Acad. Sci.* 44 (2007) 17–19; MR2299721 (2007m:41012).
[104] B. Brighi, A. Fruchard and T. Sari, On the Blasius problem, *Adv. Differential Equations* 13 (2008) 509–600; MR2482397 (2009k:34030).
[105] V. P. Varin, A solution of the Blasius problem, *Comput. Math. Math. Phys.* 54 (2014) 1025–1036; MR3217273.
[106] V. P. Varin, Asymptotic expansion of Crocco solution and Blasius constant, Keldysh Institute Appl. Math. preprint (2016).
[107] P. Amore and F. M. Fernandez, Rational approximation for two-point boundary value problems, arXiv:0705.3862.
[108] S. Abbasbandy and C. Bervillier, Analytic continuation of Taylor series and the boundary value problems of some nonlinear ordinary differential equations, *Appl. Math. Comput.* 218 (2011) 2178–2199; arXiv:1104.5073; MR2831493.

2.14 Lane–Ritter–Emden Constants

The Lane–Emden equation of index p:

$$y''(x)+\frac{2}{x}y'(x)+y(x)^p=0, \quad y(0)=1, \quad y'(0)=0$$

is useful in astrophysics for computing the structure of interiors of polytropic stars [1–3]. A well-known series for $y(x)$ is [4]

$$y(x)=\sum_{k=0}^{\infty}a_k x^{2k}, \quad x\approx 0,$$

where $a_0=1$, $a_1=-1/6$ and

$$a_k=\frac{1}{(k-1)k(2k+1)}\sum_{j=1}^{k-1}(jp+j-k+1)(k-j)(2k-2j+1)a_j a_{k-j}, \quad k\geq 2.$$

This series has radius of convergence [5–10]

$$\gamma=\lim_{k\to\infty}\left|\frac{a_k}{a_{k+1}}\right|^{1/2}=\begin{cases}\infty & \text{if } p=0 \text{ or } p=1,\\ 3.6537537362...=\sqrt{13.3499163649...} & \text{if } p=3/2,\\ 3.9645856345...=\sqrt{15.7179392534...} & \text{if } p=2,\\ 2.5748367419...=\sqrt{6.6297842476...} & \text{if } p=3,\\ 2.0348941557...=\sqrt{4.1407942251...} & \text{if } p=4,\\ 1.7320508075...=\sqrt{3} & \text{if } p=5\end{cases}$$

and details on relevant calculations will appear momentarily. The *dimensionless radius* of a polytropic star is the smallest positive x_0 for which $y(x_0)=0$: [1, 11–15]

$$x_0=\begin{cases}2.4494897427...=\sqrt{6} & \text{if } p=0,\\ 3.1415926535...=\pi & \text{if } p=1,\\ 3.6537537362...=\sqrt{13.3499163649...} & \text{if } p=3/2,\\ 4.3528745959...=\sqrt{18.9475172480...} & \text{if } p=2,\\ 6.8968486193...=\sqrt{47.5665208786...} & \text{if } p=3,\\ 14.9715463488...=\sqrt{224.1472000754...} & \text{if } p=4,\\ \infty & \text{if } p=5\end{cases}$$

and the *dimensionless mass* of a polytropic star is x_0^2 multiplied by $-y'(x_0)$:

$$\mu=-x_0^2 y'(x_0)=\begin{cases}4.8989794855...=2\sqrt{6} & \text{if } p=0,\\ 3.1415926535...=\pi & \text{if } p=1,\\ 2.7140551201... & \text{if } p=3/2,\\ 2.4110460120... & \text{if } p=2,\\ 2.0182359509... & \text{if } p=3,\\ 1.7972299144... & \text{if } p=4,\\ 1.7320508075...=\sqrt{3} & \text{if } p=5.\end{cases}$$

No closed-form expressions for constants associated with the range $1<p<5$ are known. The functions $\gamma(p)$ and $x_0(p)$ are initially equal for $p>1$, but they separate at $p\approx 1.9121$ [7]. The function $\mu(p)$ initially decreases, but encounters a minimum at $p\approx 4.823$ and increases henceforth [16, 17].

A simpler formula for the coefficients $\{a_k\}$ is valid for $p=2$: [18]

$$a_k=\frac{-1}{(2k)(2k+1)}\sum_{j=0}^{k-1}a_ja_{k-j},\quad k\geq 1,$$

which makes the alternating character of the series obvious. Is there an analogous formula for $p=3$ or $p=4$?

Let us explain how $\gamma(p)$ is computed for $2\leq p\leq 5$. Write $t=-x^2$ and $u=y^{-1/p}$, then

$$-6p\,u\frac{du}{dt}+4p(p+1)t\left(\frac{du}{dt}\right)^2-4p\,t\,u\frac{d^2u}{dt^2}=u^{-p^2+p+2}$$

$$u(0)=1,\quad \frac{du}{dt}(0)=-\frac{1}{6p}.$$

For example, supposing $p=2$, we find $u(10)=0.312...$ and $\frac{du}{dt}(10)=-0.058....$ By the Inverse Function Theorem,

$$-6p\,u\left(\frac{dt}{du}\right)^2+4p(p+1)t\frac{dt}{du}+4p\,t\,u\frac{d^2t}{du^2}=u^{-p^2+p+2}\left(\frac{dt}{du}\right)^3.$$

In the case $p=2$, initial conditions $t(0.312...)=10$ and $\frac{dt}{du}(0.312...)=\frac{1}{-0.058...}$ clearly hold. We find $t(0)=15.717...$, thus $x=(\pm 3.964...)i$ correspond to where $y=u^{-p}$ explodes [19]. This technique works because 10 is large enough that $u(10)$ is small, making the computation of $t(0)$ feasible.

See also [26–62]; the challenge of more fully understanding $x_0=4.3528745959...$ for $p=2$ was featured in [63, 64]. Other works include [65–70].

2.14.1 *Polytropic and Isothermal Spheres*

A generalization of the Lane–Emden equation is [17, 20–24]

$$y''(x)+\frac{N}{x}y'(x)+y(x)^p=0,\quad y(0)=1,\quad y'(0)=0$$

corresponding to N-dimensional polytropic spheres in $\mathbb{R}^{N+1}$. The case $N=2$ was discussed earlier. For $N=1$ (polytropic cylinders), we have

$$x_0 = \begin{cases} 2 & \text{if } p=0, \\ 2.4048255576... = z & \text{if } p=1, \\ 2.6477767662... & \text{if } p=3/2, \\ 2.9213207237... & \text{if } p=2, \\ 3.5739009819... & \text{if } p=3, \end{cases}$$

where z is the smallest positive zero of the Bessel function J_0 and [25]

$$\mu = -x_0 y'(x_0) = \begin{cases} 2 & \text{if } p=0, \\ 1.2484591696... = z\,J_1(z) & \text{if } p=1, \\ 1.0611147888... & \text{if } p=3/2, \\ 0.9253532703... & \text{if } p=2, \\ 0.7401221205... & \text{if } p=3. \end{cases}$$

No closed-form expressions for constants associated with $p>1$ are known. By contrast, for $N=0$ (polytropic slabs),

$$x_0 = \left(\frac{\pi}{2(p+1)}\right)^{1/2} \frac{\Gamma\left(\frac{1}{p+1}\right)}{\Gamma\left(\frac{p+3}{2(p+1)}\right)} = \begin{cases} 1.4142135623... = \sqrt{2} & \text{if } p=0, \\ 1.5707963267... = \pi/2 & \text{if } p=1, \\ 1.6453408471... & \text{if } p=3/2, \\ 1.7173153422... & \text{if } p=2, \\ 1.8540746773... & \text{if } p=3 \end{cases}$$

and

$$\mu = -y'(x_0) = \left(\frac{2}{p+1}\right)^{1/2} = \begin{cases} 1.4142135623... = \sqrt{2} & \text{if } p=0, \\ 1 & \text{if } p=1, \\ 0.8944271909... & \text{if } p=3/2, \\ 0.8164965809... & \text{if } p=2, \\ 0.7071067811... = 1/\sqrt{2} & \text{if } p=3. \end{cases}$$

A different generalization involves the limit as $p \to \infty$:

$$y''(x) + \frac{2}{x}y'(x) = e^{-y(x)}, \quad y(0) = y'(0) = 0.$$

This corresponds to 2-dimensional isothermal spheres in $\mathbb{R}^3$ and has the following series expansion:

$$y(x) = \sum_{k=1}^{\infty} b_k x^{2k}, \quad x \approx 0,$$

where $b_1 = 1/6$ and

$$b_k = \frac{-1}{(k-1)k(2k+1)} \sum_{j=1}^{k-1} j(k-j)(2k-2j+1) b_j b_{k-j}, \quad k \geq 2.$$

The radius of convergence, squared, is [7, 8, 71]

$$\lim_{k\to\infty}\left|\frac{b_k}{b_{k+1}}\right| = 10.7170288238... = 2(5.3585144119...).$$

This is computed by writing $t = -x^2$ and $u = e^{y/2}$, then applying the Inverse Function Theorem to

$$-12u\frac{du}{dt} + 8t\left(\frac{du}{dt}\right)^2 - 8t\,u\frac{d^2u}{dt^2} = 1,$$

$$u(0) = 1, \quad \frac{du}{dt}(0) = -\frac{1}{12}.$$

It is also known that

$$-y(x) \sim \ln\left(\frac{2}{x^2}\right) + \frac{C}{\sqrt{x}}\cos\left(\frac{\sqrt{7}}{2}\ln(x) - c\right)$$

as $x\to\infty$ for certain unspecified constants C and c. A more precise statement of this asymptotic formula, with expressions for C and c, would be good to see. Related materials include [72–77].

[1] S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Dover, 1957, pp. 84–182; MR0092663 (19,1142b).

[2] M. H. Wrubel, Stellar interiors, *Encyclopedia of Physics*, v. 51, *Astrophysics II: Stellar Structure*, ed. S. Flügge, Springer-Verlag, 1958, pp. 1–74.

[3] H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, Dover, 1962, pp. 371–394, 408; MR0181773 (31 #6000).

[4] Z. F. Seidov and R. K. Kuzakhmedov, Solution of the Lane-Emden problem in series (in Russian), *Astronomicheskii Zh.* 54 (1977) 707–708; Engl. transl. in *Soviet Astronomy* 21 (1977) 399–400; MR0537843 (80j:85007).

[5] K. A. Thernöe, Numerische Untersuchungen über die Emden-Funktionen. I, *Z. Astrophys.* 22 (1943) 122–155; MR0009675 (5,182f).

[6] I. W. Roxburgh and L. M. Stockman, Power series solutions of the polytrope equation, *Monthly Notices Royal Astron. Soc.* 303 (1999) 466–470.

[7] C. Hunter, Series solutions for polytropes and the isothermal sphere, *Monthly Notices Royal Astron. Soc.* 328 (2001) 839–847.

[8] A. Aslanov, Determination of convergence intervals of the series solutions of Emden-Fowler equations using polytropes and isothermal spheres, *Phys. Lett. A* 372 (2008) 3555–3561; MR2413857 (2009e:35015).

[9] M. I. Nouh, Accelerated power series solution of polytropic and isothermal gas spheres, *New Astronomy* 9 (2004) 467–473.

[10] A. S. Saad, Symbolic analytical solution of N-dimensional radially symmetric polytropes, *Astron. Nachr.* 325 (2004) 733–739.

[11] D. H. Sadler, J. C. P. Miller and J. R. Airey, *Mathematical Tables*, v. II, *Emden Functions*, British Assoc. Adv. Sci., 1932.

[12] C. Mohan and A. R. A1-Bayaty, Power series solutions of the Lane-Emden equation, *Astrophys. Space Sci.* 73 (1980) 227–239.

[13] J. R. Jabbar, The boundary conditions for polytropic gas spheres, *Astrophys. Space Sci.* 100 (1984) 447–449.
[14] G. P. Horedt, Seven-digit tables of Lane-Emden functions, *Astrophys. Space Sci.* 126 (1986) 357–408.
[15] Z. F. Seidov, Lane-Emden equation: perturbation method, arXiv:astro-ph/0402130.
[16] Z. F. Seidov and R. K. Kuzakhmedov, New solutions of the Lane-Emden equation (in Russian), *Astronomicheskii Zh.* 55 (1978) 1250–1255; Engl. transl. in *Soviet Astronomy* 22 (1978) 711–714; MR0537843 (80j:85007).
[17] G. P. Horedt, Exact solutions of the Lane-Emden equation in N-dimensional space, *Astron. Astrophys.* 160 (1986) 148–156.
[18] Z. F. Seidov, The power series as solution of the Lane-Emden equation with index two (in Russian), *Dokl. Akad. Nauk AzSSR* 35 (1979) 21–24.
[19] D. Rusin and R. B. Israel, Convergence radius for a nonlinear ODE solution, unpublished note (2000).
[20] G. P. Horedt, Approximate analytical solutions of the Lane-Emden equation in N-dimensional space, *Astron. Astrophys.* 172 (1987) 359–367; MR0878810 (88a:85004).
[21] J. Ostriker, The equilibrium of polytropic and isothermal cylinders, *Astrophys. J.* 140 (1964) 1056–1066; MR0189788 (32 #7208).
[22] J. Ostriker, Cylindrical Emden and associated functions, *Astrophys. J. Suppl.* 11 (1965) 167–184.
[23] T. F. Tascione, Plane-symmetric polytropic functions, *Astrophys. J. Suppl.* 24 (1972) 479–492; MR0314224 (47 #2776).
[24] E. R. Harrison and R. G. Lake, Polytropic and isothermal plane-symmetric configurations, *Astrophys. J.* 171 (1972) 323–330.
[25] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.
[26] R. V. Ramnath, On a class of nonlinear differential equations of astrophysics, *J. Math. Anal. Appl.* 35 (1971) 27–47; MR0277830 (43 #3563).
[27] Y. Viala and G. P. Horedt, Polytropic sheets, cylinders and spheres with negative index, *Astron. Astrophys.* 33 (1974) 195–202.
[28] Y. Viala and G. P. Horedt, Emden's functions for polytropes with negative index, *Astronom. Astrophys. Suppl.* 16 (1974) 173–206; MR0353631 (50 #6114).
[29] A. P. Linnell, Approximation functions to the Emden and associated Emden functions near the first zero, *Astrophys. Space Sci.* 37 (1975) 57–72.
[30] A. P. Linnell, Approximation functions to the Emden and associated Emden functions near the first zero. II, *Astrophys. Space Sci.* 37 (1975) 401–426.
[31] P. Pascual, Lane-Emden equation and Padé approximants, *Astron. Astrophys.* 60 (1977) 161–163.
[32] A. T. Service, Concise approximation formulae for the Lane-Emden functions, *Astrophys. J.* 211 (1977) 908–909.
[33] P. G. L. Leach, First integrals for the modified Emden equation $\ddot{q} + \alpha(t)\dot{q} + q^n = 0$, *J. Math. Phys.* 26 (1985) 2510–2514; MR0803793 (86k:34001).
[34] P. B. Burt, Nonperturbative solutions of the Lane-Emden equation, *Phys. Lett. A* 109 (1985) 133–135; MR0788102 (86h:85001).
[35] P. B. Burt, Nonperturbative solution of nonlinear field equations, *Nuovo Cimento B* 100 (1987) 43–52; MR0926436 (89h:81021).
[36] M. Beech, An approximate solution for the polytrope $n = 3$, *Astrophys. Space Sci.* 132 (1987) 393–396.
[37] C. M. Bender, K. A. Milton, S. S. Pinsky and L. M. Simmons, A new perturbative approach to nonlinear problems, *J. Math. Phys.* 30 (1989) 1447–1455; MR1002247 (90i:34085).

[38] M. M. Chawla, M. K. Jain and R. Subramanian, The application of explicit Nyström methods to singular second order differential equations, *Comput. Math. Appl.* 19 (1990) 47–51; MR1044440 (91a:65172).

[39] G. P. Horedt, Analytical and numerical values of Emden-Chandrasekhar associated functions, *Astrophys. J.* 357 (1990) 560–565.

[40] N. T. Shawagfeh, Nonperturbative approximate solution for Lane-Emden equation, *J. Math. Phys.* 34 (1993) 4364–4369; MR1233277 (94e:85001).

[41] G. Adomian, R. Rach and N. T. Shawagfeh, On the analytic solution of Lane-Emden equation, *Found. Phys. Lett.* 8 (1995) 161–181.

[42] B. K. Datta, Analytic solution to the Lane-Emden equation, *Nuovo Cimento B* 111 (1996) 1385–1388; MR1433976 (97j:34019).

[43] H. Goenner and P. Havas, Exact solutions of the generalized Lane-Emden equation, *J. Math. Phys.* 41 (2000) 7029–7042; MR1781422 (2001d:34004).

[44] V. B. Mandelzweig and F. Tabakin, Quasilinearization approach to nonlinear problems in physics with application to nonlinear ODEs, *Comput. Phys. Comm.* 141 (2001) 268–281; arXiv:physics/0102041; MR1872036 (2002j:34016).

[45] H. Goenner, Symmetry transformations for the generalized Lane-Emden equation, *General Relativity and Gravitation* 33 (2001) 833–841.

[46] Z. F. Seidov, Lane-Emden equation: Picard vs Pade, arXiv:astro-ph/0107395.

[47] D. C. Biles, M. P. Robinson and J. S. Spraker, A generalization of the Lane-Emden equation, *J. Math. Anal. Appl.* 273 (2002) 654–666; MR1932513 (2003g:34007).

[48] A.-M. Wazwaz, A new method for solving singular initial value problems in the second-order ordinary differential equations, *Appl. Math. Comput.* 128 (2002) 45–57; MR1891179 (2002m:34011).

[49] S. Liao, A new analytic algorithm of Lane-Emden type equations, *Appl. Math. Comput.* 142 (2003) 1–16; MR1978243 (2004d:85001).

[50] J.-H. He, Variational approach to the Lane-Emden equation, *Appl. Math. Comput.* 143 (2003) 539–541; MR1981715.

[51] J. I. Ramos, Piecewise quasilinearization techniques for singular boundary-value problems, *Comput. Phys. Comm.* 158 (2004) 12–25; MR2105397 (2005f:65091).

[52] J. I. Ramos, Linearization techniques for singular initial-value problems of ordinary differential equations, *Appl. Math. Comput.* 161 (2005) 525–542; MR2112422 (2005h:65101).

[53] S. A. Yousefi, Legendre wavelets method for solving differential equations of Lane-Emden type, *Appl. Math. Comput.* 181 (2006) 1417–1422; MR2270772.

[54] V. S. Ertürk, Differential transformation method for solving differential equations of Lane-Emden type, *Math. Comput. Appl.* 12 (2007) 135–139; MR2396843 (2009c:34009).

[55] N. Guzel, M. Bayram and C. Guler, On the numerical solution of Lane-Emden differential equations with index-*m*, *Far East J. Appl. Math.* 29 (2007) 217–232; MR2372061 (2009g:65139).

[56] A. Yildirim, A. and T. Özis, Solutions of singular IVPs of Lane-Emden type by homotopy perturbation method, *Phys. Lett. A* 369 (2007) 70–76.

[57] C. M. Khalique and P. Ntsime, Exact solutions of the Lane-Emden-type equation, *New Astronomy* 13 (2008) 476–480.

[58] H. R. Marzban, H. R. Tabrizidooz and M. Razzaghi, Hybrid functions for nonlinear initial-value problems with applications to Lane-Emden type equations, *Phys. Lett. A* 372 (2008) 5883–5886.

[59] J. I. Ramos, Series approach to the Lane-Emden equation and comparison with the homotopy perturbation method, *Chaos Solitons Fractals* 38 (2008) 400–408; MR2415942 (2009e:65117).

[60] M. Dehghan and F. Shakeri, Approximate solution of a differential equation arising in astrophysics using the variational iteration method, *New Astronomy* 13 (2008) 53–59.
[61] R. A. Van Gorder and K. Vajravelu, Analytic and numerical solutions to the Lane-Emden equation, *Phys. Lett. A* 372 (2008) 6060–6065; MR2457579.
[62] M. S. H. Chowdhury and I. Hashim, Solutions of Emden-Fowler equations by homotopy-perturbation method, *Nonlinear Analysis: Real World Appl.* 10 (2009) 104–115.
[63] J. M. Borwein and R. E. Crandall, Closed forms: What they are and why they matter (Some reflections on practical complexity theory), *Workshop on Parameterized Complexity – Not About Graphs*, Charles Darwin Univ., 2011.
[64] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A254037.
[65] S. Srivastava, A new solution of the Lane-Emden equation of index $n = 5$, *Astrophys. J.* 136 (1962) 680–681.
[66] V. D. Sharma, A non-elliptic solution of the Lane-Emden equation for index $n = 5$, *Phys. Lett. A* 60 (1977) 381–382; MR0452446 (56 #10726).
[67] J. O. Murphy, A finite radius solution for the polytrope index 5, *Proc. Astron. Soc. Austral.* 4 (1980) 37–39.
[68] J. O. Murphy, On the *F*-type and *M*-type solutions of the Lane-Emden equation, *Proc. Astron. Soc. Austral.* 4 (1980) 41–43.
[69] J. O. Murphy, A sequence of *E*-type composite analytical solutions of the Lane-Emden equation, *Proc. Astron. Soc. Austral.* 4 (1982) 376–378.
[70] J. O. Murphy, Composite analytical solutions of the Lane-Emden equation with polytropic indices $n = 1$ and $n = 5$, *Proc. Astron. Soc. Austral.* 5 (1983) 175–179.
[71] M. A. Lampert and R. U. Martinelli, Solution of the non-linear Poisson-Boltzmann equation in the interior of charged, spherical and cylindrical vesicles. I, The high-charge limit, *Chem. Phys.* 88 (1984) 399–413.
[72] S. Chandrasekhar and G. W. Wares, The isothermal function, *Astrophys. J.* 109 (1949) 551–554.
[73] F. K. Liu, Polytropic gas spheres: an approximate analytic solution of the Lane-Emden equation, *Monthly Notices Royal Astron. Soc.* 281 (1996) 1197–1205; arXiv:astro-ph/9512061.
[74] P. Natarajan and D. Lynden-Bell, An analytic approximation to the isothermal sphere, *Monthly Notices Royal Astron. Soc.* 286 (1997) 268–270; arXiv:astro-ph/9604084v1.
[75] A.-M. Wazwaz, A new algorithm for solving differential equations of Lane-Emden type, *Appl. Math. Comput.* 118 (2001) 287–310; MR1812960 (2001k:65107).
[76] E. Momoniat and C. Harley, Approximate implicit solution of a Lane-Emden equation, *New Astronomy* 11(2006) 520–526.
[77] N. Guzel and M. T. Oylumlu, Numerical solution of isothermal gas sphere problem, *Internat. J. Comput. Methods* 4 (2007) 163–171; MR2348108 (2008g:65095).

2.15 Radiative Transfer Equations

Modeling the passage of light through an absorbing and scattering medium (for example, a planetary atmosphere) is a difficult challenge. Its solution is applicable to neutron diffusion in nuclear reactor theory. We can hope only to present a few important integral equations and associated constants [1–12].

2.15.1 *Schwarzschild–Milne*

Let $s \geq 0$. For a homogeneous semi-infinite plane-parallel atmosphere with isotropic scattering, the Milne equation [13, 14]

$$f(s) = \frac{\omega}{2} \int_0^\infty f(t) E_1\left(|s-t|\right) dt, \quad f(0) = 1$$

arises, where $0 < \omega \leq 1$ is a constant (albedo) and

$$E_n(x) = \int_1^\infty \frac{e^{-xy}}{y^n} dy$$

for $n \geq 1$, which is $-\operatorname{Ei}(-x)$ if $n = 1$. Define

$$Z(\mu) = (1 - \omega\, \mu \operatorname{arctanh}(\mu))^2 + \frac{1}{4}\pi^2\omega^2\mu^2$$

and $H(\mu)$ exactly as later; we suppress the dependence on ω. In the special case when $\omega = 1$ (conservative case), the solution is given by [2]

$$f(s) = \sqrt{3}\,(s + q(s)),$$

where

$$q(s) = \frac{1}{\sqrt{3}} + \frac{1}{2\sqrt{3}} \int_0^1 \frac{1 - e^{-s/\mu}}{H(\mu)Z(\mu)} d\mu = q_\infty - \frac{1}{2\sqrt{3}} \int_0^1 \frac{e^{-s/\mu}}{H(\mu)Z(\mu)} d\mu$$

and q_∞ is **Hopf's constant** [1, 6, 10, 15–24]:

$$\begin{aligned} q_\infty &= \frac{1}{\pi} \int_0^{\pi/2} \left(\frac{3}{\sin(\theta)^2} - \frac{1}{1 - \theta \cot(\theta)} \right) d\theta \\ &= \frac{6}{\pi^2} + \frac{1}{\pi} \int_0^{\pi/2} \left(\frac{3}{\theta^2} - \frac{1}{1 - \theta \cot(\theta)} \right) d\theta \\ &= \frac{6}{\pi^2} - \frac{1}{\pi} \sum_{n=1}^\infty \frac{b_{n+1}}{2n-1} \left(\frac{\pi}{2}\right)^{2n-1} = 0.7104460895.... \end{aligned}$$

The series coefficients $b_2, b_3, b_4, \ldots$ are defined recursively via

$$\sum_{k=1}^n a_k b_{n-k+1} = 0, \quad b_1 = 3,$$

where

$$a_k = \frac{(-1)^{k-1} 2^{2k} B_{2k}}{(2k)!}$$

and $B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, $\ldots$ are the Bernoulli numbers.

Mark [25] was the first to solve Milne's equation, building on work by Wiener & Hopf [26] and Placzek & Seidel [15]. An integral equation for $q(s)$ directly is [27, 28]

$$q(s) = \frac{1}{2}E_3(s) + \frac{1}{2}\int_0^\infty q(t)E_1\left(|s-t|\right)dt$$

and a related formula for Hopf's constant is

$$q_\infty = \frac{3}{8} + \frac{3}{2}\int_0^\infty q(t)E_3\left(t\right)dt.$$

2.15.2 *Ambarzumian–Chandrasekhar*

Let $0 \le \mu \le 1$ and $0 < \omega \le 1$. The equation [4, 5, 29–33]

$$H(\mu) = 1 + \frac{1}{2}\omega\,\mu\int_0^1 \frac{H(\mu)H(\lambda)}{\mu+\lambda}d\lambda$$

possesses a continuous solution; further, it is unique if $\omega = 1$. A better definition of $H(\mu)$ for arbitrary ω avoids ambiguity [2, 34]:

$$H(\mu) = f(0,\mu)$$

where

$$f(s,\mu) = e^{-s/\mu} + \frac{\omega}{2}\int_0^\infty f(t,\mu)E_1\left(|s-t|\right)dt.$$

Halpern, Lueneburg & Clark [35] and Fock [36] proved that [10, 37]

$$\begin{aligned} H(\mu) &= \exp\left[-\frac{\mu}{\pi}\int_0^\infty \ln\left(1-\omega\frac{\arctan(\lambda)}{\lambda}\right)\frac{d\lambda}{1+\mu^2\lambda^2}\right] \\ &= \exp\left[-\frac{\mu}{\pi}\int_0^{\pi/2}\frac{\ln\left(1-\omega\,\theta\cot(\theta)\right)}{\cos(\theta)^2+\mu^2\sin(\theta)^2}d\theta\right] \end{aligned}$$

and it is clear that $H(0) = 1$ and H increases with μ. Define moments

$$\alpha_n = \int_0^1 H(\mu)\mu^n d\mu,$$

then for $\omega = 1$ we have [11, 38]

$$\alpha_0 = 2, \quad \alpha_1 = 2/\sqrt{3} = 1.1547005383...,$$

$$\alpha_2 = \frac{2}{\sqrt{3}} q_\infty = \frac{2}{\sqrt{3}}(1 - \eta_0) = 0.8203524821...,$$

$$\alpha_3 = \left(\frac{1}{5} + \frac{1}{3} q_\infty^2\right)\sqrt{3} = 0.6378182680...,$$

$$\alpha_4 = \frac{2}{\sqrt{3}}\left(\frac{1}{3} - \eta_2 + \frac{3}{10} q_\infty + \frac{1}{6} q_\infty^3\right) = 0.5222273037...,$$

where

$$\eta_j = \int_0^1 \xi^j \left[\frac{1}{2} - \frac{1}{\pi}\arctan\left(\frac{2(1 - \xi\,\mathrm{arctanh}(\xi))}{\pi\,\xi}\right)\right] d\xi$$

for $j \geq 0$. See also [39–55].

Let us examine a generalization to finite atmospheres. Let $\tau > 0$. The coupled equations

$$X(\mu) = 1 + \frac{1}{2}\omega\,\mu \int_0^1 \frac{X(\mu)X(\lambda) - Y(\mu)Y(\lambda)}{\mu + \lambda} d\lambda,$$

$$Y(\mu) = e^{-\tau/\mu} + \frac{1}{2}\omega\,\mu \int_0^1 \frac{Y(\mu)X(\lambda) - X(\mu)Y(\lambda)}{\mu - \lambda} d\lambda$$

give solutions related by

$$Y(\mu) = e^{-\tau/\mu} X(-\mu), \quad X(\mu) = e^{-\tau/\mu} Y(-\mu)$$

(appropriately extended for $\mu < 0$), but the solutions are non-unique if $\omega = 1$. A better definition is

$$X(\mu) = f(0, \mu), \quad Y(\mu) = f(\tau, \mu)$$

where

$$f(s, \mu) = e^{-s/\mu} + \frac{\omega}{2} \int_0^\tau f(t, \mu) E_1(|s - t|)\, dt.$$

The only difference from before is that the upper limit of integration here is $\tau < \infty$; in fact,

$$\lim_{\tau\to\infty} X(\mu) = H(\mu), \quad \lim_{\tau\to\infty} Y(\mu) = 0.$$

Integral expressions for X, Y analogous to H are not known. Clearly $X(0) = 1$, $Y(0) = 0$ and both X, Y increase with μ. Define moments

$$\alpha_n = \int_0^1 X(\mu)\mu^n d\mu, \quad \beta_n = \int_0^1 Y(\mu)\mu^n d\mu,$$

then for $\omega = 1$ the following hold [11]:

$$\alpha_0 + \beta_0 = 2, \quad \alpha_1 - \beta_1 = \tau\,\beta_0, \quad \alpha_2 + \beta_2 = \frac{2}{3\beta_0} - \frac{\tau}{2}\,(\alpha_1 + \beta_1)$$

for any τ. When $\tau = 1/10$, we have [24]

$$\alpha_0 = 1.1420220619..., \ \alpha_1 = 0.5765390018..., \ \alpha_2 = 0.3851978742...,$$

$$\beta_0 = 0.8579779380..., \ \beta_1 = 0.4907412080..., \ \beta_2 = 0.3384588719...$$

and when $\tau = 5$, we have

$$\alpha_0 = 1.8201574310..., \ \alpha_1 = 1.0269371382..., \ \alpha_2 = 0.7210212649...,$$

$$\beta_0 = 0.1798425689..., \ \beta_1 = 0.1277242933..., \ \beta_2 = 0.0992710166....$$

No exact formulas for α_n or β_n are known for $\tau < \infty$. The solutions $X(\mu),\ Y(\mu)$ for the conservative case are not the same as the "standard solutions"

$$\tilde{X}(\mu) = X(\mu) + \frac{\beta_0 \mu}{\alpha_1 + \beta_1}\,(X(\mu) + Y(\mu)), \quad \tilde{Y}(\mu) = Y(\mu) - \frac{\beta_0 \mu}{\alpha_1 + \beta_1}\,(X(\mu) + Y(\mu))$$

described by Chandrasekhar [56, 57], which satisfy $\tilde{\alpha}_0 = 2$ and $\tilde{\beta}_0 = 0$ (rather than the non-homogenous Milne equation for f). It is known that

$$\tilde{\alpha}_1^2 - \tilde{\beta}_1^2 = \frac{4}{3}$$

and, further, that pairwise moment sums are invariant [24]:

$$\alpha_1 + \beta_1 = 1.0672802099... = \tilde{\alpha}_1 + \tilde{\beta}_1,$$

$$\alpha_2 + \beta_2 = 0.7236567462... = \tilde{\alpha}_2 + \tilde{\beta}_2$$

when $\tau = 1/10$, and

$$\alpha_1 + \beta_1 = 1.1546614315... = \tilde{\alpha}_1 + \tilde{\beta}_1,$$

$$\alpha_2 + \beta_2 = 0.8202922816... = \tilde{\alpha}_2 + \tilde{\beta}_2$$

when $\tau = 5$. Conceivably an integral expression might exist for $\alpha_n + \beta_n$ but not for either α_n or β_n. Note that, in general [2],

$$\lim_{\mu \to \infty} H(\mu) = (1 - \omega)^{-1/2},$$

whereas

$$\lim_{\mu \to \infty} X(\mu) = \left[1 - \frac{\omega}{2}\,(\alpha_0 - \beta_0)\right]^{-1} = \lim_{\mu \to \infty} Y(\mu)$$

assuming $\tau < \infty$. See also [58–72]. Much territory remains for exploration.

[1] E. Hopf, *Mathematical Problems of Radiative Equilibrium*, Cambridge Univ. Press, 1934, pp. 21–31, 52–54, 85.
[2] I. W. Busbridge, *The Mathematics of Radiative Transfer*, Cambridge Univ. Press, 1960, pp. 1–103.
[3] E. A. Milne, Thermodynamics of the stars, *Handbuch der Astrophysik*, v. 3, n. 1, Springer-Verlag, 1930, pp. 65–255; *Selected Papers on the Transfer of Radiation*, ed. D. H. Menzel, Dover, 1966, pp. 77–269.
[4] S. Chandrasekhar, The transfer of radiation in stellar atmospheres, *Bull. Amer. Math. Soc.* 53 (1947) 641–711; *Selected Papers*. v. II. *Radiative Transfer and Negative Ion of Hydrogen*, Univ. of Chicago Press, 1989, pp. 439–509; MR0022303 (9,189e).
[5] S. Chandrasekhar, *Radiative Transfer*, Oxford Univ. Press, 1950, pp. 14–17, 105–126, 183–207, 346–350; MR0042603 (13,136a).
[6] V. Kourganoff, *Basic Methods in Transfer Problems. Radiative Equilibrium and Neutron Diffusion*, Oxford Univ. Press, 1952, pp 153–193; MR0054158 (14,879b).
[7] B. Davison, *Neutron Transport Theory*, Oxford Univ. Press, 1957, pp. 67–86; MR0095716 (20 #2217).
[8] V. A. Ambartsumyan, *Theoretical Astrophysics*, Pergamon Press, 1958, pp. 550–572.
[9] V. V. Sobolev, *A Treatise on Radiative Transfer*, Van Nostrand, 1963, pp. 38–96.
[10] V. V. Ivanov, *Transfer of Radiation in Spectral Lines*, Nat. Bur. Standards, Spec. Publ. 385, 1973, pp. 125–141.
[11] H. C. van de Hulst, *Multiple Light Scattering*, v. 1, Academic Press, 1980, pp. 52, 149–171, 191, 224–235.
[12] S. N. Shore, Blue sky and hot piles: the evolution of radiative transfer theory from atmospheres to nuclear reactors, *Historia Math.* 29 (2002) 463–489; MR1936801 (2004d:85003).
[13] K. Schwarzschild, Über Diffusion und Absorption in der Sonnenatmosphäre, *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin* 47 (1914) 1183–1200; Engl. transl. in *Selected Papers on the Transfer of Radiation*, ed. D. H. Menzel, Dover, 1966, pp. 35–51.
[14] E. A. Milne, Radiative equilibrium in the outer layers of a star, *Monthly Notices Royal Astron. Soc.* 81 (1921) 361–375.
[15] G. Placzek and W. Seidel, Milne's problem in transport theory, *Phys. Rev.* 72 (1947) 550-555; MR0022029 (9,147g).
[16] J. King, The Hopf q-function simply and precisely evaluated, *Astrophys. J.* 132 (1960) 509–511.
[17] J. I. F. King, R. V. Sillars and R. H. Harrison, Hopf q-function evaluated to eight-digit accuracy, *Astrophys. J.* 142 (1965) 1655–1659.
[18] H. C. van de Hulst, Asymptotic fitting, a method for solving anisotropic transfer problems in thick layers, *J. Comput. Phys.* 3 (1968) 291–306 (computation due to K. Grossman).
[19] G. R. Bond and C. E. Siewert, On the Milne-problem extrapolation distance, *Nuclear Sci. Eng.* 36 (1969) 447.
[20] W. C. DeMarcus and R. A. Stokes, An accurate determination of the value of the Hopf q-function at infinity, *Nuclear Sci. Eng.* 42 (1970) 230–231.
[21] E. R. Cohen, A numerical calculation of the Milne problem extrapolation limit, *Nuclear Sci. Eng.* 42 (1970) 231–232.
[22] S. K. Loyalka and S. Naz, An approximate method in transport theory. II, *Physics of Fluids* 18 (2006) 117106.
[23] S. K. Loyalka and S. Naz, Milne's half-space problem: A numerical solution of the related integral equation, *Annals Nuclear Energy* 35 (2008) 1900–1902.

[24] T. Viik, An efficient method to calculate Ambarzumian-Chandrasekhar's and Hopf's functions, *Astrophys. Space Sci.* 127 (1986) 285–307.

[25] C. Mark, The neutron density near a plane surface, *Phys. Rev.* 72 (1947) 558–564; MR0022031 (9,147i).

[26] N. Wiener and E. Hopf, Über eine Klasse singulärer Integralgleichungen, *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin, Phys.-Math. Klasse* 30–32 (1931) 696–706; *Collected Works*, v. 3, ed. P. Masani, MIT Press, 1981, pp. 33–53.

[27] J. LeCaine, Application of a variational method to Milne's problem, *Phys. Rev.* 72 (1947) 564–566; MR0022032 (9,147j).

[28] S.-S. Huang, The variational method for problems of neutron diffusion and of radiative transfer, *Phys. Rev.* 88 (1952) 50–52.

[29] V. A. Ambarzumian, Diffuse reflection of light by a foggy medium, *C. R. (Doklady) Acad. Sci. URSS* 38 (1943) 229–232; MR0008739 (5,52g).

[30] V. A. Ambarzumian, The scattering of light by planetary atmospheres (in Russian), *Astr. J. Soviet Union [Astr. Zhurnal]* v. 19 (1942) n. 5, 30–41; Engl. transl. in *A Life in Astrophysics*, Allerton Press, 1998, pp. 67–82; MR0009729 (5,192e).

[31] V. A. Ambarzumian, The problem of diffuse reflection of light by a turbid medium, *Dokl. Akad. Nauk SSSR* 38 (1943) 257–261; *A Life in Astrophysics*, Allerton Press, 1998, pp. 83–87.

[32] V. A. Ambarzumian, On the problem of the diffuse reflection of light, *Acad. Sci. URSS. J. Phys.* 8 (1944) 65–75; *A Life in Astrophysics*, Allerton Press, 1998, pp. 89–106; MR0013024 (7,97e).

[33] S. Chandrasekhar, On the radiative equilibrium of a stellar atmosphere. XIV, XVII, XX, XXII, *Astrophys. J.* 105 (1947) 164–203, 105 (1947) 441–460, 106 (1947) 145–151, 107 (1948) 48–72, 188–215; *Selected Papers*. v. II. *Radiative Transfer and Negative Ion of Hydrogen*, Univ. of Chicago Press, 1989, pp. 219–338; MR0019836 (8,467g), MR0021233 (9,41a), MR0024048 (9,444c), MR0025075 (9,593f).

[34] D. F. Mayers, Calculation of Chandrasekhar's X- and Y- functions for isotropic scattering, *Monthly Notices Royal Astron. Soc.* 123 (1962) 471–484.

[35] O. Halpern, R. Lueneburg and O. Clark, On multiple scattering of neutrons. I, Theory of the albedo of a plane boundary, *Phys. Rev.* 53 (1938) 173–183.

[36] V. Fock, Sur certaines équations intégrales de physique mathématique (in Russian), *Rec. Math. [Mat. Sbornik]* 14 [56] (1944) 3–50; MR0012190 (6,272b).

[37] D. W. N. Stibbs and R. E. Weir, On the H-functions for isotropic scattering, *Monthly Notices Royal Astron. Soc.* 119 (1959) 512–525.

[38] V. V. Sobolev, Some functions in the theory of light scattering (in Russian), *Astrofizika* 3 (1967) 433–445; *Astrophysics* 3 (1967) 205–209; MR0307649 (46 #6769).

[39] G. Placzek, The angular distribution of neutrons emerging from a plane surface, *Phys. Rev.* 72 (1947) 556–558; MR0022030 (9,147h).

[40] M. M. Crum, On an integral equation of Chandrasekhar, *Quart. J. Math.* 18 (1947) 244–252; MR0023437 (9,356c).

[41] D. L. Harris, Diffuse reflection from planetary atmospheres, *Astrophys. J.* 126 (1957) 408–412.

[42] C. Fox, A solution of Chandrasekhar's integral equation, *Trans. Amer. Math. Soc.* 99 (1961) 285–291; MR0123900 (23 #A1221).

[43] K. D. Abhyankar and A. L. Fymat, On the solution $H_1(\mu)$ of Ambarzumian-Chandrasekhar H-equation, *J. Quant. Spectrosc. Radiat. Transfer* 9 (1969) 1563–1566.

[44] K. D. Abhyankar and A. L. Fymat, Tables of auxiliary functions for the nonconservative Rayleigh phase matrix in semi-infinite atmospheres, *Astrophys. J. Suppl.* 23 (1971) 35–101.
[45] S. R. Das Gupta, On H-functions of radiative transfer, *Astrophys. Space Sci.* 30 (1974) 327–342; MR0356447 (50 #8917).
[46] S. R. Das Gupta, A new representation of the H-functions of radiative transfer, *Astrophys. Space Sci.* 50 (1977) 187–203; MR0452373 (56 #10653).
[47] R. L. Bowden, A note on the iteration of the Chandrasekhar nonlinear H-equation, *J. Math. Phys.* 20 (1979) 608–610; MR0529724 (80i:82005).
[48] C. T. Kelley, Solution of the Chandrasekhar H-equation by Newton's method, *J. Math. Phys.* 21 (1980) 1625–1628; MR0575595 (81j:85003).
[49] P. B. Bosma and W. A. de Rooij, Efficient methods to calculate Chandrasekhar's H-function, *Astron. Astrophys.* 126 (1983) 283–292.
[50] D. M. Sedrakyan, K. M. Shakhabasyan and A. G. Movsisyan, A high accuracy approximation of Ambartsumian's function (in Russian), *Astrofizika* 19 (1983) 175–177.
[51] A. Holubec and J. C. McConnell, An approximation of Chandrasekhar's H-function for conservative isotropic scattering, *Astron. Astrophys.* 227 (1990) 649–652; MR1036846 (90k:85003).
[52] E. Y. Deeba and S. A. Khuri, The decomposition method applied to Chandrasekhar H-equation, *Appl. Math. Comput.* 77 (1996) 67–78; MR1386691 (96m:65121).
[53] A. Natarajan and N. Mohankumar, A note on the numerical evaluation of the H function, *Comput. Phys. Comm.* 107 (1997) 54–60; MR1488792.
[54] N. Mohankumar and A. Natarajan, On the evaluation of H and X functions, *Comput. Phys. Comm.* 176 (2007) 266–270.
[55] R. N. Das, Solution of a Riemann-Hilbert problem for a new expression of Chandrasekhar's H-function in radiative transfer, *Astrophys. Space Sci.* 317 (2008) 119–126.
[56] S. Chandrasekhar, D. Elbert, and A. Franklin, The X- and Y-functions for isotropic scattering. I, *Astrophys. J.* 115 (1952) 244–268.
[57] S. Chandrasekhar and D. Elbert, The X- and Y-functions for isotropic scattering. II, *Astrophys. J.* 115 (1952) 269–278.
[58] I. W. Busbridge, On the X- and Y-functions of S. Chandrasekhar, *Astrophys. J.* 122 (1955) 327–348; MR0073069 (17,377d).
[59] T. W. Mullikin, A complete solution of the X- and Y-equations of Chandrasekhar, *Astrophys. J.* 136 (1962) 627–635; MR0153467 (27 #3433).
[60] T. W. Mullikin, Chandrasekhar's X and Y equations, *Trans. Amer. Math. Soc.* 113 (1964) 316–332; MR0166577 (29 #3850).
[61] J. L. Carlstedt and T. W. Mullikin, Chandrasekhar's X- and Y-functions, *Astrophys. J. Suppl.* 12 (1966) 449–586.
[62] Y. Sobouti, Chandrasekhar's X-, Y-, and related functions, *Astrophys. J. Suppl.* 7 (1963) 411–560.
[63] R. Bellman, H. Kagiwada, R. Kalaba and S. Ueno, Numerical results for Chandrasekhar's X and Y functions of radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer* 6 (1966) 479–500.
[64] R. Bellman, P. T. Y. Poon, and. S. Ueno, Numerical results for finite order X and Y functions of radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer* 13 (1973) 1273–1284.
[65] H. Cohen, A new computational method for the X and Y functions of radiative transfer, *J. Quant. Spectrosc. Radiat. Transfer* 9 (1969) 931–942.

[66] J. Caldwell, The X- and Y-functions of Chandrasekhar, *Astrophys. J.* 163 (1971) 111–120; MR0282555 (43 #8264).
[67] J. Caldwell and A. J. Perks, Chandrasekhar's X- and Y-functions, *Astrophys. J.* 249 (1981) 258–262.
[68] V. M. Loskutov, Calculation of radiation field in the case of monochromatic isotropic scattering. II, Ambartsumyan's functions $\varphi(\eta, \tau_0)$, $\psi(\eta, \tau_0)$ and their moments (in Russian), *Astrofizika* 9 (1973) 361–370; Engl. transl. in *Astrophysics* 9 (1973) 205–210; MR0334865 (48 #13183).
[69] W. L. Dunn, D. M. Garcia and C. E. Siewert, Concise and accurate solutions for Chandrasekhar's X and Y functions, *Astrophys. J.* 260 (1982) 849–854.
[70] M. H. Haggag, H. M. Machali, S. El-Labany and M. T. Attia, Chandrasekhar's X- and Y-functions, *Astrophys. Space Sci.* 149 (1988) 13–18.
[71] L. Chevallier and B. Rutily, Exact solution of the standard transfer problem in a stellar atmosphere, *J. Quant. Spectrosc. Radiat. Transfer* 91 (2005) 373–391; arXiv:astro-ph/0504049.
[72] B. Rutily, L. Chevallier and J. Pelkowski, K. Schwarzschild's problem in radiation transfer theory, *J. Quant. Spectrosc. Radiat. Transfer* 98 (2006) 290–307; arXiv:astro-ph/0601340.

2.16 Carleman's Inequality

The inequality

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} < e \sum_{k=1}^{\infty} a_k$$

relates the geometric and arithmetic means of an infinite sequence a_1, a_2, ..., where $a_k \geq 0$ for all k and $a_\ell > 0$ for at least one ℓ. The constant e is best possible [1–5].

A number of refined versions of Carleman's original inequality have appeared including [6, 7]

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} < e \sum_{k=1}^{\infty} \left[1 - \frac{1}{2(k+1)}\right] a_k$$

and a generalization exists [8–13]:

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} < e \sum_{k=1}^{\infty} \left[1 - \sum_{j=1}^{m} \frac{b_j}{(k+1)^j}\right] a_k,$$

where m is any positive integer and $b_1 = 1/2$, $b_2 = 1/24$, $b_3 = 1/48$, $b_4 = 73/5760$, $b_5 = 11/128$, $b_6 = 3625/580608$, ... are generated via

$$b_j = -\frac{1}{j} \sum_{i=1}^{j} \frac{b_{j-i}}{i+1}, \quad b_0 = -1.$$

In different directions, we have

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} \le e \sum_{k=1}^{\infty} \left[1 - \frac{1 - 2/e}{k}\right] a_k,$$

$$\sum_{k=1}^{\infty} (a_1 a_2 \cdots a_k)^{1/k} \le e \sum_{k=1}^{\infty} \left[1 + \frac{1}{k}\right]^{1-1/\ln(2)} a_k$$

and a common extension of these also exists [14, 15].

Our interest is in the n^{th} finite section of Carleman's inequality:

$$\sum_{k=1}^{n} (a_1 a_2 \cdots a_k)^{1/k} < C_n \sum_{k=1}^{n} a_k.$$

It is known that the best constant C_n satisfies [16, 17]

$$C_n = e - 2\pi^2 e \frac{1}{\ln(n)^2} + O\left(\frac{1}{\ln(n)^3}\right)$$

asymptotically as $n \to \infty$. The rate at which C_n approaches e is quite slow. What can be said for small values of n?

It is not difficult to show that

$$C_2 = \frac{1}{2}\left(1 + \sqrt{2}\right), \quad C_3 = \frac{4}{3}$$

via direct minimization of $\sum_{k=1}^{n} (a_1 a_2 \cdots a_k)^{1/k}$ subject to the constraint $\sum_{k=1}^{n} a_k = 1$. A symbolic technique in [18] gives that $C_4 = 1.4208443854...$ is algebraic of degree 24 with minimal polynomial

$$\begin{aligned}
&1090491731185059590307 84x^{24} - 654295038711035754184704x^{23} \\
&\quad + 147216383709983044691 5584x^{22} - 138734781356321470100 2752x^{21} \\
&\quad + 220843507713085418766336x^{20} + 361130725214496730644480x^{19} \\
&\quad + 18738444188050884919296x^{18} - 149735761790067869220864x^{17} \\
&\quad - 20033038006659651207168x^{16} + 14417509185682352898048x^{15} \\
&\quad + 1690553030369369024102 4x^{14} - 2098418839125516877824x^{13} \\
&\quad - 198705178996352483328x^{12} + 42744743365616389324 8x^{11} \\
&\quad + 41447678188009291776x^{10} - 2629784260986273792x^{9} \\
&\quad + 660475521813381120x^{8} + 342213608420278272x^{7} \\
&\quad + 42624005978423296x^{6} - 201976270848000x^{5} \\
&\quad + 274965186525696x^{4} + 12841816536576x^{3} \\
&\quad + 373658292864x^{2} + 22039921152x \\
&\quad + 387420489;
\end{aligned}$$

also we have $C_5 = 1.4863532289...$ and $C_6 = 1.5379375565...$ by numeric means. The minimal polynomials of C_5 and C_6 are presently unknown.

[1] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, 1934, pp. 249–250; MR0944909 (89d:26016).
[2] D. S. Mitrinović, J. E. Pečarić and A. M. Fink, *Inequalities Involving Functions and Their Integrals and Derivatives*, Kluwer, 1991, pp. 143–186; MR1190927 (93m:26036).
[3] J. Pečarić and K. B. Stolarsky, Carleman's inequality: history and new generalizations, *Aequationes Math.* 61 (2001) 49–62; MR1820809 (2002m:26029).
[4] J. Duncan and C. M. McGregor, Carleman's inequality, *Amer. Math. Monthly* 110 (2003) 424–431; MR2040885 (2004k:26027).
[5] M. Johansson, L.-E. Persson and A. Wedestig, Carleman's inequality - history, proofs and some new generalizations, *J. Inequal. Pure Appl. Math.* 4 (2003) A 53; MR2044402 (2005a:26027).
[6] B. Yang and L. Debnath, Some inequalities involving the constant *e*, and an application to Carleman's inequality, *J. Math. Anal. Appl.* 223 (1998) 347–353; MR1627296 (99h:26026).
[7] J.-L. Li, Notes on an inequality involving the constant *e*, *J. Math. Anal. Appl.* 250 (2000) 722–725; MR1786093 (2001h:26032).
[8] X. Yang, On Carleman's inequality, *J. Math. Anal. Appl.* 253 (2001) 691–694; MR1808160 (2001h:26036).
[9] X. Yang, Approximations for constant *e* and their applications, *J. Math. Anal. Appl.* 262 (2001) 651–659; MR1859330 (2002g:26035).
[10] M. Gyllenberg and Y. Ping, On a conjecture by Yang, *J. Math. Anal. Appl.* 264 (2001) 687–690; MR1876758 (2002j:26019).
[11] H. Chen, On an infinite series for $(1 + 1/x)^x$ and its application, *Internat. J. Math. Math. Sci.* 29 (2002) 675–680; MR1902838 (2003c:26026).
[12] H. Yue, A strengthened Carleman's inequality, *Commun. Math. Anal.* 1 (2006) 115–119; MR2301825 (2008b:26034).
[13] X. Yang, Y.-I. Kim and K. Lo, Approximation for constant *e* and its applications, *Appl. Math. Comput.* 206 (2008) 50–55; MR2474956.
[14] B.-Q. Yuan, Refinements of Carleman's inequality, *J. Inequal. Pure Appl. Math.* 2 (2001) A 21; MR1873861 (2002i:26020).
[15] D.-Y. Hwang, Some refinements and generalizations of Carleman's inequality, *Internat. J. Math. Math. Sci.* 2004, 2171–2180; MR2100097 (2005i:26043).
[16] N. G. de Bruijn, Carleman's inequality for finite series, *Nederl. Akad. Wetensch. Proc. Ser. A* 66 = *Indag, Math.* 25 (1963) 505–514; MR0156141 (27 #6073).
[17] H. S. Wilf, *Finite Sections of Some Classical Inequalities*, Springer-Verlag, 1970, pp. 6-7, 70–78; MR0271762 (42 #6643).
[18] Y.-D. Wu, Z. H. Zhang and Z. G. Wang, The best constant for Carleman's inequality of finite type, *Acta Math. Acad. Paedagog. Nyházi.* 24 (2008) 235–241; MR2461634 (2009i:26043).

2.17 Golay–Littlewood Problem

Two independent streams of investigation, one from digital communications engineering and the other from complex analysis on the unit circle, come together in this essay [1–5].

2.17.1 *Merit Factor of Binary Sequences*

Given a sequence a_0, a_1, a_2, ..., a_n where each $a_j = \pm 1$, define the k^{th} **acyclic autocorrelation** to be

$$c_k = \sum_{j=0}^{n-k} a_j a_{j+k} \quad \text{for } 0 \le k \le n; \quad c_k = c_{-k} \quad \text{for } -n \le k < 0$$

and the **merit factor** to be the ratio

$$F = \frac{c_0^2}{\sum_{k \neq 0} c_k^2} = \frac{(n+1)^2}{2\sum_{k=1}^{n} c_k^2}.$$

Identifying binary sequences $\{a_j\}$ whose autocorrelations $\{c_k\}$ are jointly as small as possible, for fixed n, is important for engineering design purposes. The "best" sequences are those with the largest merit factor F. As an example, the sequence $1, -1, 1, -1, 1, 1, -1, -1, 1, 1, 1, 1, 1$ has the largest F value $169/12 = 14.0833...$ among all such with $n = 12$. As another example, the sequence $1, -1, 1, 1, -1, 1, 1, 1, -1, -1, -1$ has the largest F value $121/10 = 12.1$ among all such with $n = 10$. No other merit factor exceeding 10 is known for any n; a proof that $169/12$ and $121/10$ are the maximum possible values for F is still open.

2.17.2 *L_4 Norm of Polynomials on Unit Circle*

Given a polynomial of complex variable z:

$$f(z) = \sum_{j=0}^{n} a_j z^j,$$

the L_p norm of f over the unit circle for $p \ge 1$ is

$$\|f\|_p = \left[\frac{1}{2\pi}\int_0^{2\pi} \left|f\left(e^{i\theta}\right)\right|^p d\theta\right]^{1/p}.$$

Since the complex conjugate $\bar{z}$ is equal to $1/z$ and all polynomial coefficients a_j are real, we have $\overline{f(z)} = f(\bar{z}) = f(1/z)$. Hence

$$|f(z)|^2 = f(z) f\left(\frac{1}{z}\right) = c_0 + \sum_{k \neq 0} c_k z^k$$

and, after integrating, $\|f\|_2^2 = c_0 = n + 1$ because each $a_j = \pm 1$. Also, we have

$$|f(z)|^4 = f(z)^2 f\left(\frac{1}{z}\right)^2 = \sum_k c_k^2 + \sum_{k+\ell \neq 0} c_k c_\ell z^{k+\ell}$$

and, after integrating, $\|f\|_4^4 = \sum c_k^2 = (n+1)^2(1+1/F)$. Thus Littlewood's question [6, 7] about how closely the ratio $\|f\|_4 \,/\, \|f\|_2$ can approach 1 as $n \to \infty$ translates into Golay's question [8–13] about the limit supremum of F.

2.17.3 *Bounds on Asymptotic Behavior*

On the one hand, let $\xi = 1.157677...$ denote the smallest zero of $27x^3 - 498x^2 + 1164x - 722$. Jedwab, Katz & Schmidt [14] proved that there is a Littlewood polynomial sequence $\{f_n\}$ such that $\deg(f_n) \to \infty$ and

$$\frac{\|f_n\|_4}{\|f_n\|_2} \to \sqrt[4]{\xi} = 1.037282...$$

as $n \to \infty$. As a consequence,

$$\limsup_{n\to\infty} F_n \geq \eta = \frac{1}{\xi - 1} = 6.342061....$$

The preceding best result, namely $\xi = 7/6 = 1.16...$ ($\eta = 6$), had remained in place for more than twenty years [15, 16]. Recent numerical computations indicate that $\xi = 1.1553...$ ($\eta = 6.4382...$) is feasible. We might have to wait a long time for rigorous verification of this result because, in the words of [17], "inclusion of the steep descent algorithm ... would seem to make a proof much more difficult". Theory lags considerably behind experiment here: there is good evidence that $\eta > 8$ or even $\eta > 8.5$. Merit factors exceeding 9 are not uncommon for sequence lengths ≈ 100, but it is difficult to project whether such extremities will continue to grow slowly or level off [18, 19].

On the other hand, no one has proved that the limit supremum of F is necessarily *finite*. (An argument in [11, 20] that it is approximately 12.32 is only heuristic.) This would be good to see someday.

Imagine the set of all sequences of length $n+1$, endowed with the uniform distribution. Draw one such sequence and compute F. The mean value of $1/F$ is exactly [21, 22]

$$\mathrm{E}\left(\frac{1}{F}\right) = \frac{n}{n+1} \to 1$$

as $n \to \infty$. An exact expression for $\mathrm{Var}(1/F)$ is not available, but it is $O(1/n)$ according to [4]. Thus most sequences should have merit factor close to 1 [23]. What else can be said about the distribution of $1/F$ or, indeed, of F itself?

Relevant material is covered in [24, 25]. The survey [4] mentions Mahler's measure and Lehmer's conjecture surrounding a certain polynomial of degree 10 (with largest zero 1.1762808182...) [26]. Related problems involving ± 1 sequences appear in [27–29].

Addendum Choi [30] supplemented the result $\mathrm{E}(\|f\|_4^4) = (n+1)(2n+1)$ with a new one:

$$\mathrm{Var}\left(\|f\|_4^4\right) = \frac{8}{3}(n+1)\left(2n^2 - 2n + 3\right) - 8\left\lfloor\frac{n^2+2n+2}{2}\right\rfloor$$

giving a formula for $\mathrm{Var}(1/F)$ as a corollary. Golay's constant is, to higher precision,

$$12.3247958363... = \frac{2y^2}{2y - \ln(2y+1)},$$

where y is the unique positive solution of the equation $(y+1)\ln(2y+1) = 2(1 + \ln(2))y$ [20].

[1] J. Jedwab, A survey of the merit factor problem for binary sequences, *Sequences and Their Applications - SETA 2004*, ed. T. Helleseth, D. Sarwate, H.-Y. Song and K. Yang, Lect. Notes in Comp. Sci. 3486, Springer-Verlag, 2005, pp. 30–55.

[2] P. Borwein, *Computational Excursions in Analysis and Number Theory*, Springer-Verlag, 2002, pp. 109–132, 181–202; MR1912495 (2003m:11045).

[3] T. Hoholdt, The merit factor problem for binary sequences, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes - AAECC-16*, ed. M. P. C. Fossorier, H. Imai, S. Lin and A. Poli, Lect. Notes in Comp. Sci. 3857, Springer-Verlag, 2006, pp. 51–59.

[4] K.-K. S. Choi, Extremal problems about norms of Littlewood polynomials, unpublished manuscript (2006).

[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A091386 and A102780.

[6] J. E. Littlewood, On polynomials $\sum^n \pm z^m$, $\sum^n e^{\alpha_m i} z_m$, $z = e^{\theta i}$, *J. London Math. Soc.* 41 (1966) 367–376; MR0196043 (33 #4237).

[7] J. E. Littlewood, *Some Problems in Real and Complex Analysis*, D. C. Heath and Co., 1968, pp. 25–32; MR0244463 (39 #5777).

[8] M. J. E. Golay, A class of finite binary sequences with alternate autocorrelation values equal to zero, *IEEE Trans. Inform. Theory* IT-18 (1972) 449–450.

[9] M. J. E. Golay, Hybrid low autocorrelation sequences, *IEEE Trans. Inform. Theory* IT-21 (1975) 460–462.

[10] M. J. E. Golay, Sieves for low autocorrelation binary sequences, *IEEE Trans. Inform. Theory* IT-23 (1977) 43–51.

[11] M. J. E. Golay, The merit factor of long low autocorrelation binary sequences, *IEEE Trans. Inform. Theory* IT-28 (1982) 543–549.

[12] M. J. E. Golay, The merit factor of Legendre sequences, *IEEE Trans. Inform. Theory* IT-29 (1983) 934–936.

[13] M. J. E. Golay and D. B. Harris, A new search for skewsymmetric binary sequences with optimal merit factors, *IEEE Trans. Inform. Theory* 36 (1990) 1163–1166.

[14] J. Jedwab, D. J. Katz and K.-U. Schmidt, Littlewood polynomials with small L^4 norm, *Adv. Math.* 241 (2013) 127–136; arXiv:1205.0260; MR3053707.

[15] T. Hoholdt and H. E. Jensen, Determination of the merit factor of Legendre sequences, *IEEE Trans. Inform. Theory* 34 (1988) 161–164; Zbl 0652.40006.

[16] P. Borwein, K.-K. S. Choi and J. Jedwab, Binary sequences with merit factor greater than 6.34, *IEEE Trans. Inform. Theory* 50 (2004) 3234–3249; MR2103494.

[17] J. M. Baden, Efficient optimization of the merit factor of long binary sequences, *IEEE Trans. Inform. Theory* 57 (2011) 8084–8094; MR2895382 (2012m:94243).

[18] S. Mertens, Exhaustive search for low-autocorrelation binary sequences, *J. Phys. A* 29 (1996) L473–L481; MR1419192 (97i:82050).
[19] P. Borwein, R. Ferguson and J. Knauer, The merit factor problem, *Number Theory and Polynomials*, ed. J. McKee and C. Smyth, Cambridge Univ. Press, 2008, pp. 52–70; MR2428515 (2010b:11030).
[20] J. Bernasconi, Low autocorrelation binary sequences: statistical mechanics and configuration state analysis, *J. Physique* 48 (1987) 559–567.
[21] D. V. Sarwate, Mean-square correlation of shift-register sequences, *Communications, Radar and Signal Processing. IEE Proceedings F*, v. 131 (1984) n. 2, 101–106.
[22] D. J. Newman and J. S. Byrnes, The L^4 norm of a polynomial with coefficients ± 1, *Amer. Math. Monthly 97* (1990) 42–45; MR1034349 (91d:30006).
[23] K.-U. Schmidt, On random binary sequences, *Sequences and Their Applications - SETA 2012*, ed. T. Helleseth and J. Jedwab, Lect. Notes in Comp. Sci. 7280, Springer-Verlag, 2012, pp. 303–314; MR3015481.
[24] D. J. Katz, Asymptotic L^4 norm of polynomials derived from characters, *Pacific J. Math.* 263 (2013) 373–398; arXiv:1205.1069; MR3068549.
[25] J. Jedwab, D. J. Katz and K.-U. Schmidt, Advances in the merit factor problem for binary sequences, *J. Combin. Theory Ser. A* 120 (2013) 882–906; arXiv:1205.0626; MR3022619.
[26] S. R. Finch, Pisot-Vijayaraghavan-Salem constants, *first volume*, pp. 192–199.
[27] S. R. Finch, Power series with restricted coefficients, *this volume*, §3.9.
[28] S. R. Finch, Two-colorings of positive integers, *this volume*, §1.29.
[29] S. R. Finch, Moments of sums, *this volume*, §4.2.
[30] K.-K. S. Choi, Variance of L_4-norm of the polynomials with ± 1 coefficients, unpublished note (1997).

2.18 Online Matching Coins

The first game we discuss originated in [1, 2], although we mostly follow [3] in our exposition. The second and third games appear in [4].

Nature tosses a fair coin repeatedly and independently, yielding an infinite sequence N_0, N_1, N_2, … of 1s and 2s. Just prior to each toss, Alice and Bob simultaneously declare their guess A and B for the resulting N. They win the toss if both guessed correctly. Their goal is to maximize the probability of winning. They are permitted to strategize no later than one hour beforehand; after the game starts, any communication between them is only via the As, Bs and Ns. With no further information, if they agree beforehand to always both guess 1 (for example) then the probability of winning is $1/2$. No improvement is possible.

Suppose now that, during their strategizing, Alice and Bob are told that Alice will be given the full sequence N_0, N_1, N_2, … one minute before the game! To improve their odds, Alice must pass relevant information she knows to Bob in an agreed-upon manner. Setting $A = N$ always does not help their cause! At toss 0, Alice might declare

$$A_0 = N_2$$

(sacrificing her knowledge of N_0) so that Bob understands to declare $B_1 = B_2 = A_0$. They will win toss 2 since Alice will declare $A_2 = N_2$. At toss 1, Alice might declare

$$A_1 = N_4$$

(sacrificing her knowledge of N_1) so that Bob understands to declare $B_3 = B_4 = A_1$. They will win toss 4 since Alice will declare $A_4 = N_4$. At toss 3, Alice might declare

$$A_3 = N_6$$

(sacrificing her knowledge of N_3) so that Bob understands to declare $B_5 = B_6 = A_3$. They will win toss 6 since Alice will declare $A_6 = N_6$, and so forth. In summary, Alice and Bob will score one win out of two whenever $\{N_{2t+1}, N_{2t+2}\} = \{1, 2\}$ or $\{2, 1\}$. When $\{N_{2t+1}, N_{2t+2}\} = \{1, 1\}$ or $\{2, 2\}$, they will score one win out of two half the time and two out of two the remaining half, giving odds of

$$\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{\frac{1}{2} + 1}{2} = \frac{5}{8} = 0.625.$$

Instead of partitioning time into blocks modulo 2, let us do so modulo 3. Define the mode M_t of $\{N_{3t+1}, N_{3t+2}, N_{3t+3}\}$ to be the most common element in the set. At toss 0, Alice might declare

$$A_0 = M_0$$

(sacrificing her knowledge of N_0) so that Bob understands to declare $B_1 = B_2 = B_3 = M_0$. Assume that indices $1 \le i, j, k \le 3$ are distinct. Alice's next three declarations might be

$$A_i = A_j = M_0 \text{ and } A_k = M_1 \quad \text{if } N_k \ne M_0$$

and

$$A_1 = A_2 = M_0 \text{ and } A_3 = M_1 \quad \text{if } N_1 = N_2 = N_3 = M_0$$

(sacrificing her knowledge of N_3 for the latter) so that Bob understands to declare $B_4 = B_5 = B_6 = M_1$. In summary, Alice and Bob will score two wins out of three whenever $\{N_1, N_2, N_3\}$ contains two 1s and one 2, or two 2s and one 1. When $\{N_1, N_2, N_3\}$ contains all 1s or all 2s, they will score two wins out of three half the time and three out of three the remaining half, giving odds of

$$\frac{3}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{\frac{2}{3} + 1}{2} = \frac{17}{24} = 0.7083....$$

A more sophisticated strategy allows the win probability to approach $x = 0.8107103750...$ as closely as desired, where x is the unique solution of the equation [1]

$$-x \ln(x) - (1 - x) \ln(1 - x) + (1 - x) \ln(3) = \ln(2).$$

No further improvement is possible beyond this point.

2.18.1 *Symmetric Online Matching Coins*

The preceding game is asymmetric – Alice knows everything and Bob knows nothing – for the following game, information is distributed equally among the players and they will both need to send signals to each other. Imagine here that a fair coin has four equally-likely sides, not two. (A regular tetrahedral die would be a better metaphor.) Also define

$$f(N)=\begin{cases}1 & \text{if } N=1 \text{ or } 3,\\ 0 & \text{if } N=2 \text{ or } 4,\end{cases} \quad g(N)=\begin{cases}1 & \text{if } N=1 \text{ or } 2,\\ 0 & \text{if } N=3 \text{ or } 4\end{cases}$$

for convenience, that is, $f(N)$ anwers the question "Is N odd?" and $g(N)$ answers the question "Is $N \leq 2$?"

Nature tosses a fair coin repeatedly and independently, yielding an infinite sequence N_1, N_2, N_3, … of 1s, 2s, 3s and 4s. Just prior to each toss, Alice and Bob simultaneously declare their guess A and B for the resulting N. They win the toss if both guessed correctly. Their goal is to maximize the probability of winning. They are permitted to strategize no later than one hour beforehand; after the game starts, any communication between them is only via the As, Bs and Ns. With no further information, if they agree beforehand to always both guess 1 (for example) then the probability of winning is $1/4$. No improvement is possible.

Suppose now that, during their strategizing, Alice and Bob are told that Alice will be given the sequence $f(N_1)$, $f(N_2)$, $f(N_3)$, … and Bob will be given the sequence $g(N_1)$, $g(N_2)$, $g(N_3)$, … one minute before the game! At toss 1, Alice might declare

$$A_1=\begin{cases}1 & \text{if } f(N_1)=1 \text{ and } f(N_2)=0,\\ 2 & \text{if } f(N_1)=0 \text{ and } f(N_2)=1,\\ 3 & \text{if } f(N_1)=1 \text{ and } f(N_2)=1,\\ 4 & \text{if } f(N_1)=0 \text{ and } f(N_2)=0\end{cases}$$

and Bob might declare

$$B_1=\begin{cases}1 & \text{if } g(N_1)=1 \text{ and } g(N_2)=0,\\ 2 & \text{if } g(N_1)=0 \text{ and } g(N_2)=1,\\ 3 & \text{if } g(N_1)=1 \text{ and } g(N_2)=1,\\ 4 & \text{if } g(N_1)=0 \text{ and } g(N_2)=0\end{cases}$$

so that they will win toss 2. The odds here are

$$\frac{1}{2}\left(\frac{1}{2}\cdot\frac{1}{2}+1\right)=\frac{5}{8}=0.625.$$

Instead of devoting resources to guessing N_1, let us shift emphasis entirely to signaling ahead for N_2 and N_3. At toss 1, Alice might declare

$$A_1 = \begin{cases} 1 & \text{if } f(N_2) = 1 \text{ and } f(N_3) = 0, \\ 2 & \text{if } f(N_2) = 0 \text{ and } f(N_3) = 1, \\ 3 & \text{if } f(N_2) = 1 \text{ and } f(N_3) = 1, \\ 4 & \text{if } f(N_2) = 0 \text{ and } f(N_3) = 0 \end{cases}$$

and Bob might declare

$$B_1 = \begin{cases} 1 & \text{if } g(N_2) = 1 \text{ and } g(N_3) = 0, \\ 2 & \text{if } g(N_2) = 0 \text{ and } g(N_3) = 1, \\ 3 & \text{if } g(N_2) = 1 \text{ and } g(N_3) = 1, \\ 4 & \text{if } g(N_2) = 0 \text{ and } g(N_3) = 0 \end{cases}$$

(both sacrificing their partial knowledge of N_1) so that they will win tosses 2 and 3. The odds here are

$$\frac{1}{3}\left(\frac{1}{4} \cdot \frac{1}{4} + 1 + 1\right) = \frac{33}{48} = 0.6875.$$

A more sophisticated strategy allows the win probability to approach $\kappa = 0.7337221510...$ as closely as desired [4]. The formulas underlying this constant are more elaborate than before. Define a hyperplanar region in $\mathbb{R}^8$:

$$\Delta(8) = \left\{ (x_1, x_2, \ldots, x_8) : \sum_{\ell=1}^{8} x_\ell = 1 \text{ and } x_\ell \geq 0 \text{ for all } \ell \right\}$$

and a real-valued function on $\Delta(8)$:

$$h(x) = -\frac{1}{\ln(2)} \sum_{\ell=1}^{8} x_\ell \ln(x_\ell)$$

with the convention that $0 \cdot \ln(0) = 0$. Let $\varphi : [0, 3] \to \mathbb{R}$ be given by

$$\varphi(r) = \max \left\{ \sum_{\ell=1}^{4} x_\ell^2 : x \in \Delta(8) \text{ and } h(x) \geq r \right\}$$

and let $\psi : [0, 3] \to \mathbb{R}$ be the minimal concave function $\geq \varphi$. The desired probability κ is $\psi(1)$, which numerically appears to be equal to $\varphi(1)$. No further improvement is possible beyond this point. It also appears that the minimizing vector x can be taken such that $x_1 = x_2 = x_3$ and $x_5 = x_6 = x_7$, which would simplify our presentation.

2.18.2 *Cross Over Matching Coins*

Here the game is symmetric, as for the preceding, but Nature instead tosses a *pair* of distinguishable coins (two sides apiece). Thus we have two infinite sequences N_0^α, N_1^α, N_2^α, ... and N_0^β, N_1^β, N_2^β, ... of 1s and 2s. Just prior to each toss, Alice and Bob simultaneously declare their guesses A and B for the resulting N^α and N^β, respectively. During their one-hour prior strategizing, they learn that Alice will be given the sequence N_0^β, N_1^β, N_2^β, ... and Bob will be given the sequence N_0^α, N_1^α, N_2^α, ... at one-minute prior! Their goal is to maximize the average of (the probability of Alice winning) and (the probability of Bob winning). Communication between them, via the As, Bs, N^αs and N^βs, is again critical to their success.

The optimal win probability here is $\lambda = 0.8041565330...$ [4]. Define a line segment in $\mathbb{R}^2$:

$$\Delta(2) = \left\{ (x_1, x_2) : \sum_{\ell=1}^{2} x_\ell = 1 \text{ and } x_\ell \geq 0 \text{ for all } \ell \right\}$$

and a real-valued function on $\Delta(2)$:

$$h(x) = -\frac{1}{\ln(2)} \sum_{\ell=1}^{2} x_\ell \ln(x_\ell).$$

Let $\varphi : [0, 1] \to \mathbb{R}$ be given by

$$\varphi(r) = \max \left\{ \sum_{\ell=1}^{2} x_\ell^2 : x \in \Delta(2) \text{ and } h(x) \geq r \right\}$$

and let $\psi : [0, 1] \to \mathbb{R}$ be the minimal concave function $\geq \varphi$. The desired probability λ is $\psi(1/2)$, which is (in this case) provably equal to $\varphi(1/2)$.

A simpler presentation is hence clear: $\lambda = y^2 + (1-y)^2$ where y is either of the two reals satisfying

$$-2y \ln(y) - 2(1-y) \ln(1-y) = \ln(2).$$

No closed-form expression for this constant (or for other constants in this essay) seems to be available.

It is possible to generalize the symmetric online game to an arbitrary number m of players and a single n^m-sided coin. The real-valued function h on $\Delta(n^{m+1})$ gives rise to a φ (maximum sum of m^{th} powers, indices from 1 to n^m) and a minimal concave $\psi \geq \varphi$. For $m > 2$ or $n > 2$, however, $\psi(\ln(n)/\ln(2))$ is strictly greater than $\varphi(\ln(n)/\ln(2))$. This complicates the numerical calculation of a optimal win probability in the general setting.

It is also possible to generalize the cross over matching game to a pair of n-sided coins. The real-valued function h on $\Delta(n)$ gives rise to a φ (maximum sum of n

squares) and a minimal concave $\psi \geq \varphi$. For $n > 2$, however, $\psi(\ln(n)/(2\ln(2)))$ is strictly greater than $\varphi(\ln(n)/(2\ln(2)))$. This again complicates calculations in general.

Related ideas appear in [5] (best strategies) and [6] (maximal convex function $\leq \varphi$).

[1] O. Gossner, P. Hernández and A. Neyman, Online matching pennies, Center for the Study of Rationality, Hebrew Univ., discussion paper, 2003.
[2] O. Gossner, P. Hernández and A. Neyman, Optimal use of communication resources, *Econometrica* 74 (2006) 1603–1636; MR2268411 (2007e:91112).
[3] P. Winkler, *Mathematical Mind-Benders*, A. K. Peters, 2007, pp. 11, 17–18; MR2334790 (2008f:00002).
[4] A. Shapira, *Communication Games with Asymmetric Information*, Ph.D. thesis, Hebrew Univ., 2008.
[5] S. R. Finch, Optimal stopping constants, *first volume*, pp. 361–363.
[6] S. R. Finch, Shapiro-Drinfeld constant, *first volume*, pp. 208–211.

2.19 Toothpicks and Live Cells

We understand a **toothpick** to be a compact unit subinterval of the real line. At time 1, place a toothpick in the xy-plane with endpoints at $(0, \pm 1/2)$. Both endpoints are *exposed* and must be *covered* at time 2. This is done by simultaneously placing a new toothpick with endpoints at $(\pm 1/2, 1/2)$ and a new toothpick with endpoints at $(\pm 1/2, -1/2)$. New toothpicks at odd times are always vertical; new toothpicks at even times are always horizontal. Any old endpoint is exposed if it is neither the endpoint nor the midpoint of any other existing toothpick. If exposed, it must be covered by the midpoint of a new toothpick without delay. At time 3, four new toothpicks are needed; likewise for times 4 and 5. At time 6, eight new toothpicks are required (see Figure 2.2); at time 7, twelve are required. No toothpicks are ever removed [1].

Let $T(n)$ denote the total number of toothpicks at time n. For $k \geq 0$, we have the following recursion:

$$T\left(2^k + j\right) = \begin{cases} \frac{1}{3}\left(2^{2k+1} + 1\right) & \text{if } j = 0, \\ T\left(2^k\right) + 2T(j) + T(j+1) - 1 & \text{if } 1 \leq j \leq 2^k - 1. \end{cases}$$

No simple formula for $T(n)$ is known; it is not well behaved asymptotically in the sense that [2]

$$0.4513058284... = c = \liminf_{n\to\infty} \frac{T(n)}{n^2} < \limsup_{n\to\infty} \frac{T(n)}{n^2} = \frac{2}{3}.$$

We understand **cells** to be the basis elements of an infinite planar square lattice. Neighbors of each cell are defined to be the four squares that share an edge with it (see Figure 2.3). At time 1, a single cell is alive. At time $n > 1$, a cell newly comes

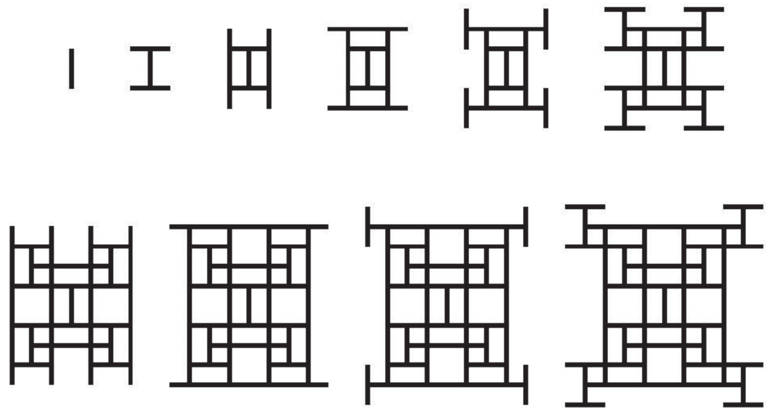

Figure 2.2 Toothpicks $\{T(n)\}_{n=1}^{10} = \{1, 3, 7, 11, 15, 23, 35, 43, 47, 55\}$, from [1].

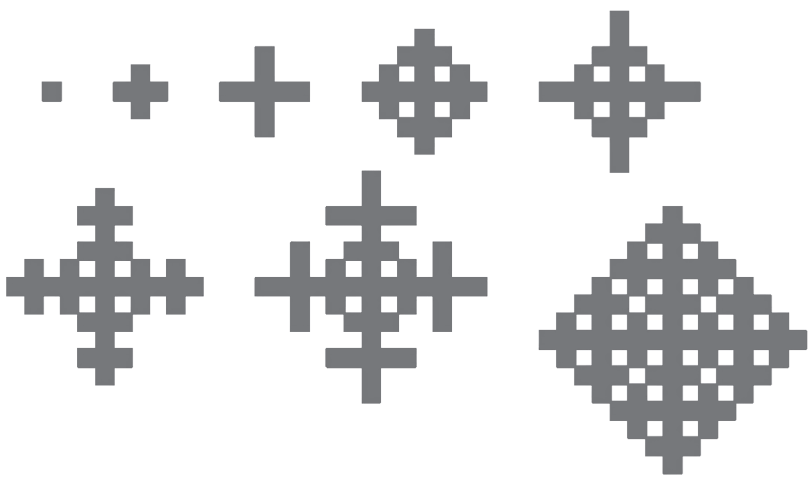

Figure 2.3 Live cells $\{U(n)\}_{n=1}^{8} = \{1, 5, 9, 21, 25, 37, 49, 85\}$, from [9].

to life if and only if exactly one of its neighbors is alive and older (that is, alive at time $n-1$). Once a cell is alive, it remains alive forever [3–9].

Let $U(n)$ denote the total number of live cells at time n. For $n \geq 1$, a simple formula applies:

$$U(n) = \frac{1}{3}\left(4\sum_{m=0}^{n-1} 3^{b(m)} - 1\right),$$

where $b(m)$ is the number of ones in the binary expansion of m. We have seen such exponential sums of digital sums before [10] and find [11]

$$0.9026116569... = \liminf_{n\to\infty} \frac{U(n)}{n^2} < \limsup_{n\to\infty} \frac{U(n)}{n^2} = \frac{4}{3}.$$

The fact that $4/3$ is the limit superior has been known for years [4]; by contrast, no one seems to have studied the limit inferior until now. Is this quantity equal to $2c$? Why should the toothpick and Ulam–Warburton automata be so closely related? Sloane [12] provided an overview of associated ideas.

Acknowledgment I am thankful to Robert Price for numerically confirming that the lower limit of $U(n)/n^2$ is twice that of $T(n)/n^2$ to high precision.

[1] D. Applegate, O. E. Pol and N. J. A. Sloane, The toothpick sequence and other sequences from cellular automata, *Congr. Numer.* 206 (2010) 157–191; arXiv:1004.3036; MR2762248 (2011k:37017).
[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A139250, A170927, and A195853.
[3] S. M. Ulam, On some mathematical problems connected with patterns of growth of figures, *Mathematical Problems in the Biological Sciences*, ed. R. Bellman, Proc. Symp. Applied Math., v. 14, Amer. Math. Soc., 1962, pp. 215–224; MR0157851 (28 #1080).
[4] R. P. Stanley and R. J. Chapman, A tree in the integer lattice, *Amer. Math. Monthly* 105 (1998) 769–771.
[5] S. Wolfram, *A New Kind of Science*, Wolfram Media, 2002, pp. 928–929; MR1920418 (2003i:37002).
[6] M. Warburton, One-edge connections, *M500* 188 (Oct 2002) 11.
[7] D. Singmaster, On the cellular automaton of Ulam and Warburton, *M500* 195 (Dec 2003) 2–7.
[8] T. Khovanova, E. Nie and A. Puranik, The Sierpinski triangle and the Ulam-Warburton automaton, arXiv:1408.5937.
[9] A. Fink, A. S. Fraenkel and C. Santos, LIM is not slim, *Internat. J. Game Theory* 43 (2014) 269–281; MR3197293.
[10] S. R. Finch, Stolarsky-Harborth constant, *first volume*, pp. 145–151.
[11] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A147562, A260239, and A261313.
[12] N. J. A. Sloane, On the number of ON cells in cellular automata, arXiv:1503.01168.

2.20 Virial Coefficients

A fluid is a large collection of small particles. The simplest model for fluids in D-dimensional space gives rise to the ideal gas law

$$\frac{P}{\kappa T} = \rho,$$

where P is pressure, T is temperature, ρ is density and κ is Boltzmann's constant. A more general model takes interparticle interactions of all orders into consideration. It features the virial series expansion

$$\frac{P}{\kappa T} = \rho + \sum_{n=2}^{\infty} B_{n,D}\, \rho^n$$

where coefficients $B_{n,D}$ depend on the choice of potential function. We will focus on the hard core potential

$$\begin{cases} \infty & \text{if } r \le 1, \\ 0 & \text{if } r > 1, \end{cases}$$

which implies that two particles have no interaction if their distance > 1 and they are prohibited from approaching a distance ≤ 1. The particles are called hard rods if $D = 1$, hard disks if $D = 2$ and hard spheres if $D = 3$. A more realistic potential

$$\begin{cases} \infty & \text{if } r \le 1, \\ -\varepsilon & \text{if } 1 < r \le 1 + \delta, \\ 0 & \text{if } r > 1 + \delta \end{cases}$$

includes a region of attraction as well as a repulsive hard core; this is called the square-well potential. Other choices exist.

If $D = 1$, then [1–3]

$$\frac{P}{\kappa T} = \frac{\rho}{1 - \rho},$$

that is, $B_{n,1} = 1$ for all $n \ge 1$, corresponding to a fluid of hard rods. For $D \ge 2$, we need to discuss nonseparable graphs on n vertices, building on material covered in [4, 5]. The number of such graphs is 1, 1, 3, 10 for $2 \le n \le 5$. Figure 2.4 exhibits the 15 graphs so far mentioned and symbols representing each [6, 7]. English letters correspond to the number of vertices; integers correspond to the number of

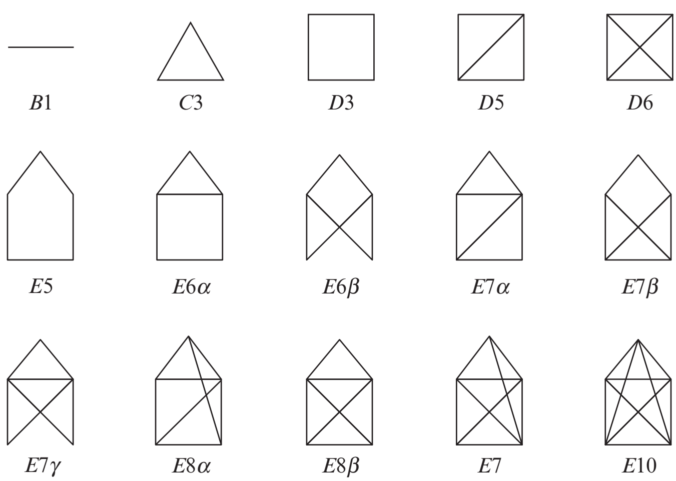

Figure 2.4 15 unlabeled nonseparable graphs on ≤ 5 vertices.

edges; Greek letters will be explained shortly. The number of *labeled* nonseparable graphs is 1, 1, 10, 238 for $2 \le n \le 5$. Our interest is in the labeled case. For $n=4$, there are 3 graphs of type $D4$, 6 graphs of type $D5$ and 1 graph of type $D6$. For $n=5$, there are 12 graphs of type $E5$, 70 graphs of type $E6$, 100 graphs of type $E7$, 45 graphs of type $E8$, 10 graphs of type $E9$ and 1 graph of type $E10$. Further refinement is needed for three cases:

$$\begin{aligned} 70\ E6 \text{ graphs} &= 60\ E6\alpha \text{ graphs} + 10\ E6\beta \text{ graphs}, \\ 100\ E7 \text{ graphs} &= 60\ E7\alpha \text{ graphs} + 30\ E7\beta \text{ graphs} + 10\ E7\gamma \text{ graphs}, \\ 45\ E8 \text{ graphs} &= 15\ E8\alpha \text{ graphs} + 30\ E8\beta \text{ graphs}. \end{aligned}$$

Let us now illustrate what is called the Mayer formalism for representing virial coefficients $B_{n,D}$ for $2 \le n \le 5$ and $D \ge 2$. Given n points $\vec{r}_1, \vec{r}_2, \vec{r}_3, \ldots, \vec{r}_n$ in $\mathbb{R}^D$ with $\vec{r}_1 = \vec{0}$ by convention, define $r_{ij} = |\vec{r}_i - \vec{r}_j|$ and

$$f(r) = \begin{cases} -1 & \text{if } r \le 1, \\ 0 & \text{if } r > 1. \end{cases}$$

We abuse notation and allow graph symbols to serve as shorthand for certain integrals:

$$B_{2,D} = -\frac{1}{2} \int_{\mathbb{R}^D} f(r_{12}) d\vec{r}_2 = -\frac{1}{2} \frac{1}{1!} B1,$$

$$B_{3,D} = -\frac{1}{3} \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} f(r_{12}) f(r_{23}) f(r_{31}) d\vec{r}_2\, d\vec{r}_3 = -\frac{2}{3} \frac{1}{2!} C3,$$

$$B_{4,D} = -\frac{3}{4} \frac{1}{3!} \left(3D4 + 6D5 + D6\right),$$

where

$$D4 = \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} f(r_{12}) f(r_{23}) f(r_{34}) f(r_{41}) d\vec{r}_2\, d\vec{r}_3\, d\vec{r}_4,$$

$$D5 = \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} f(r_{12}) f(r_{23}) f(r_{34}) f(r_{41}) f(r_{13}) d\vec{r}_2\, d\vec{r}_3\, d\vec{r}_4,$$

$$D6 = \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} \int_{\mathbb{R}^D} f(r_{12}) f(r_{23}) f(r_{34}) f(r_{41}) f(r_{13}) f(r_{24}) d\vec{r}_2\, d\vec{r}_3\, d\vec{r}_4.$$

Continuing,

$$\begin{aligned} B_{5,D} = -\frac{4}{5} \frac{1}{4!} &\left(12E5 + 60E6\alpha + 10E6\beta + 60E7\alpha + 30E7\beta + 10E7\gamma \right. \\ &\left. + 15E8\alpha + 30E8\beta + 10E9 + E10\right), \end{aligned}$$

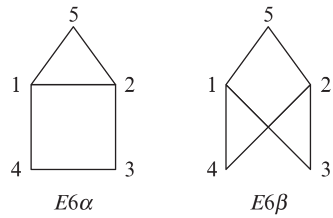

Figure 2.5 Selected labeled nonseparable graphs on 5 vertices.

where, for example,

$$E6\alpha = \int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D} f(r_{12})f(r_{14})f(r_{15})f(r_{23})f(r_{25})f(r_{34})d\vec{r}_2\, d\vec{r}_3\, d\vec{r}_4\, d\vec{r}_5,$$

$$E6\beta = \int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D}\int\limits_{\mathbb{R}^D} f(r_{13})f(r_{14})f(r_{15})f(r_{23})f(r_{24})f(r_{25})d\vec{r}_2\, d\vec{r}_3\, d\vec{r}_4\, d\vec{r}_5$$

and we have used the helpful labels in Figure 2.5.

From these formulas, we deduce that [8–17]

$$B_{2,D} = \frac{\pi^{D/2}}{2\Gamma\left(1 + D/2\right)} = \begin{cases} \pi/2 & \text{if } D = 2, \\ 2\pi/3 & \text{if } D = 3; \end{cases}$$

$$\frac{B_{3,D}}{B_{2,D}^2} = \frac{4\Gamma\left(1 + D/2\right)}{\sqrt{\pi}\Gamma\left((1 + D)/2\right)} \int\limits_0^{\pi/3} \sin(\theta)^D d\theta = \begin{cases} 4/3 - \sqrt{3}/\pi & \text{if } D = 2, \\ 5/8 & \text{if } D = 3; \end{cases}$$

$$\frac{B_{4,D}}{B_{2,D}^3} = \begin{cases} 2 - (9/2)\left(\sqrt{3}/\pi\right) + 10/\pi^2 & \text{if } D = 2, \\ 2707/4480 + (219/2240)\left(\sqrt{2}/\pi\right) - (4131/4480)\,(\operatorname{arcsec}(3)/\pi) & \text{if } D = 3; \end{cases}$$

$$\frac{B_{5,D}}{B_{2,D}^4} = \begin{cases} 0.33355604... & \text{if } D = 2, \\ 0.110252... & \text{if } D = 3. \end{cases}$$

Elaborating on $B_{5,D}$ for $D = 3$:

$$\frac{E5}{B_2^4} = -\frac{40949}{10752}, \quad \frac{E6\alpha}{B_2^4} = \frac{68419}{26880}, \quad \frac{E6\beta}{B_2^4} = \frac{82}{35},$$

$$\frac{E7\alpha}{B_2^4} = -\frac{34133}{17920}, \quad \frac{E7\beta}{B_2^4} = -\frac{18583}{5376} + \frac{33291}{9800}\frac{\sqrt{3}}{\pi}, \quad \frac{E7\gamma}{B_2^4} = -\frac{73491}{35840},$$

$$\frac{E8\beta}{B_2^4} = -\frac{35731}{6720} + \frac{1458339}{627200}\frac{\sqrt{2}}{\pi} - \frac{33291}{9800}\frac{\sqrt{3}}{\pi} + \frac{683559}{35840}\frac{\operatorname{arcsec}(3)}{\pi},$$

but exact expressions for

$$\frac{E8\alpha}{B_2^4} \approx 2(0.56965), \quad \frac{E9}{B_2^4} \approx 3(-0.30490) \quad \frac{E10}{B_2^4} \approx 30(0.02369)$$

remain open. Even less is known about $B_{5,D}$ for $D=2$:

$$\frac{E6\beta}{B_2^4}=16-\frac{116}{\pi^2},\quad \frac{E7\gamma}{B_2^4}=-16+\frac{16\sqrt{3}}{\pi}+\frac{196}{3\pi^2}-\frac{117\sqrt{3}}{2\pi^3}.$$

Numerical integration is evidently required for the remaining subcases. For example [14, 16],

$$\begin{aligned} E6\alpha &= 4\pi^2\int_0^1\int_0^{1-r} A(r)A(s)r\,s\,ds\,dr \\ &\quad +4\pi\int_0^1\int_{1-r}^{1+r} A(r)A(s)\arccos\left(\frac{r^2+s^2-1}{2r\,s}\right)r\,s\,ds\,dr \\ &\approx (4.46966949)B_2^4\approx\frac{1}{2}(8.93933899)B_2^4, \end{aligned}$$

$$\begin{aligned} E7\alpha &= -4\pi^2\int_0^1\int_0^{1-r} A(r)A(s)r\,s\,ds\,dr \\ &\quad -4\pi\int_0^1\int_{1-r}^{1} A(r)A(s)\arccos\left(\frac{r^2+s^2-1}{2r\,s}\right)r\,s\,ds\,dr \\ &\approx (-3.61831477)B_2^4\approx\frac{1}{2}(-7.23662954)B_2^4, \end{aligned}$$

$$\begin{aligned} E5 &= -E6\alpha-4\pi\int_1^2\int_{-1+r}^{2} A(r)A(s)\arccos\left(\frac{r^2+s^2-1}{2r\,s}\right)r\,s\,ds\,dr \\ &\approx (-5.97307832)B_2^4\approx\frac{5}{2}(-2.38923133)B_2^4, \end{aligned}$$

where

$$A(r)=2\arccos\left(\frac{r}{2}\right)-\frac{r}{2}\sqrt{4-r^2}$$

is the area of the intersection of two overlapping disks, each of unit radius, with distance r between their centers. Other symbols require evaluation of trivariate integrals or worse; computational difficulty seems to increase with the number of edges in the graph. A remarkable breakthrough was achieved recently [18, 19], giving $E10$ for $D=2$ solely in terms of bivariate integrals and hence to high accuracy:

$$\frac{E10}{B_2^4}=1.8090652427...=5(0.3618130485...)=30(0.0603021747...).$$

Details of this computation are still forthcoming. Analogous estimates for the other unsolved contributions to $B_{5,2}$ are unavailable; the corresponding difficulties for $B_{5,3}$ are insurmountable.

A different normalization for virial coefficients often appears:

$$\tilde{B}_{n,D} = \frac{B_{n,D}}{(\omega_D/2^D)^{n-1}}$$

where $\omega_D = \pi^{D/2}/\Gamma(1 + D/2)$, the volume enclosed by the unit sphere in $\mathbb{R}^D$. Thus $\tilde{B}_{2,2} = 2$, $\tilde{B}_{2,3} = 4$, $\tilde{B}_{3,2} = 16/3 - 4\sqrt{3}/\pi$ and $\tilde{B}_{3,3} = 10$. We merely mention challenging research for $n > 5$ and $D > 3$, which is beyond the scope of his essay [20–29].

Addendum An expression for the area of the intersection I of three overlapping disks, each of unit radius, is found in [30]. Let the centers be $(-r/2, 0)$, $(r/2, 0)$ and (x, y), where $0 < r < 2$ and the third point is assumed to be inside the intersection J of the first two disks. Assume further that a nonempty arc of ∂J lies outside of the third circle, that is, I is nondegenerate. Let

$$d_{12} = r, \quad d_{13} = \sqrt{(x + r/2)^2 + y^2}, \quad d_{23} = \sqrt{(x - r/2)^2 + y^2},$$

$$x_{12} = d_{12}/2, \quad x'_{13} = d_{13}/2, \quad x''_{23} = d_{23}/2,$$

$$y_{12} = \sqrt{1 - d_{12}^2/4}, \quad y'_{13} = -\sqrt{1 - d_{13}^2/4}, \quad y''_{23} = \sqrt{1 - d_{23}^2/4},$$

$$\lambda' = \frac{d_{12}^2 + d_{13}^2 - d_{23}^2}{2d_{12}d_{13}}, \quad \mu' = \sqrt{1 - \lambda'^2}, \quad \lambda'' = -\frac{d_{12}^2 + d_{23}^2 - d_{13}^2}{2d_{12}d_{23}}, \quad \mu'' = \sqrt{1 - \lambda''^2},$$

$$x_{13} = x'_{13}\lambda' - y'_{13}\mu', \quad y_{13} = x'_{13}\mu' + y'_{13}\lambda',$$

$$x_{23} = x''_{23}\lambda'' - y''_{23}\mu'' + d_{12}, \quad y_{23} = x''_{23}\mu'' + y''_{23}\lambda'',$$

$$c_1 = \sqrt{(x_{12} - x_{13})^2 + (y_{12} - y_{13})^2}, \quad c_2 = \sqrt{(x_{12} - x_{23})^2 + (y_{12} - y_{23})^2},$$

$$c_3 = \sqrt{(x_{13} - x_{23})^2 + (y_{13} - y_{23})^2}.$$

Then the desired area is

$$\begin{aligned} \aleph(x, y, r) \;=\; & \frac{1}{4}\sqrt{(c_1 + c_2 + c_3)(-c_1 + c_2 + c_3)(c_1 - c_2 + c_3)(c_1 + c_2 - c_3)} \\ & + \sum_{k=1}^{3}\left[\arcsin\left(\frac{c_k}{2}\right) - \frac{c_k}{4}\sqrt{4 - c_k^2}\right]. \end{aligned}$$

Define also

$$u(x, r) = \sqrt{1 - x^2} - \sqrt{1 - r^2/4}, \quad v(x, r) = \sqrt{1 - (x + r/2)^2},$$

$$w(r) = \frac{1}{4}\left(-r + \sqrt{3}\sqrt{4 - r^2}\right);$$

exact formulas for

$$\theta(r) = A(r) \int_0^{r/2} \int_0^{u(x,r)} dy\, dx,$$

$$\varphi(r) = A(r) \int_0^{w(r)} \int_0^{u(x,r)} dy\, dx, \quad \psi(r) = A(r) \int_{w(r)}^{1-r/2} \int_0^{v(x,r)} dy\, dx$$

exist but are omitted for brevity's sake. Two additional symbols for $D = 2$ are therefore [14]

$$\begin{aligned} E8\beta &= 8\pi \left[\int_0^1 \theta(r) A(r)\, r\, dr \right. \\ &\quad + \int_0^1 \int_{r/2}^{w(r)} \int_0^{-u(x,r)} A\left(d_{13}\right) A(r)\, r\, dy\, dx\, dr + \int_0^1 \int_{w(r)}^{1-r/2} \int_0^{v(x,r)} A\left(d_{13}\right) A(r)\, r\, dy\, dx\, dr \\ &\quad \left. + \int_0^1 \int_0^{r/2} \int_{u(x,r)}^{v(x,r)} \aleph(x,y,r) A(r)\, r\, dy\, dx\, dr + \int_0^1 \int_{r/2}^{w(r)} \int_{-u(x,r)}^{v(x,r)} \aleph(x,y,r) A(r)\, r\, dy\, dx\, dr \right] \\ &\approx (2.810839) B_2^4, \end{aligned}$$

$$\begin{aligned} E7\beta &= -E8\beta - 2\pi \int_{\sqrt{3}}^{2} A(r)^3 r\, dr \\ &\quad -8\pi \left[\int_1^{\sqrt{3}} \varphi(r) A(r)\, r\, dr + \int_1^{\sqrt{3}} \psi(r) A(r)\, r\, dr + \int_1^{\sqrt{3}} \int_0^{w(r)} \int_{u(x,r)}^{v(x,r)} \aleph(x,y,r) A(r)\, r\, dy\, dx\, dr \right] \\ &\approx (-3.202747) B_2^4. \end{aligned}$$

We have not attempted to independently evaluate [16]

$$\frac{E8\alpha}{B_2^4} \approx 2.529628 \approx 2(1.264814), \quad \frac{E9}{B_2^4} \approx -2.160499 \approx 3(-0.720166)$$

except to verify that a certain identity

$$E6\beta + E7\gamma + 3(E7\beta + E8\alpha + E8\beta) + 4E9 + E10 = 0$$

is satisfied.

[1] M. Bishop, Virial coefficients for one-dimensional hard rods, *Amer. J. Phys*. 51 (1983) 1151–1152.

[2] M. Baus and J. L. Colot, Thermodynamics and structure of a fluid of hard rods, disks, spheres, or hyperspheres from rescaled virial expansions, *Phys. Rev. A* 36 (1987) 3912–3925.

[3] G. Labelle, P. Leroux and M. G. Ducharme, Graph weights arising from Mayer's theory of cluster integrals, *Sém. Lothar. Combin.* 54 (2005/07) B54m; MR2341745 (2009b:82003).

[4] S. R. Finch, Planar graph growth constants, *this volume*, §1.9.

[5] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002218, A013922 and A123534.

[6] R. J. Riddell and G. E. Uhlenbeck, On the theory of the virial development of the equation of state of mono-atomic gases, *J. Chem. Phys.* 21 (1953) 2056–2064; MR0059207 (15,491g).

[7] J. S. Rowlinson, The fifth virial coefficient of a fluid of hard spheres, *Proc. Royal Soc. London A* 279 (1964) 147–160.

[8] I. Lyberg, *Statistical Mechanics of Hard Spheres and the Two Dimensional Ising Lattice*, Ph.D. thesis, State Univ. of New York at Stony Brook, 2007.

[9] I. Lyberg, The fourth virial coefficient of a fluid of hard spheres in odd dimensions, *J. Stat. Phys.* 119 (2005) 747–764; arXiv:cond-mat/0410080; MR2151221 (2006c:82061).

[10] N. Clisby, *Negative Virial Coefficients for Hard Spheres*, Ph.D. thesis, State Univ. of New York at Stony Brook, 2004.

[11] N. Clisby and B. M. McCoy, Analytic calculation of B_4 for hard spheres in even dimensions, *J. Stat. Phys.* 114 (2004) 1343–1360; arXiv:cond-mat/0303098; MR2039480 (2004m:82040).

[12] N. Clisby and B. M. McCoy, Ninth and tenth order virial coefficients for hard spheres in D dimensions, *J. Stat. Phys.* 122 (2006) 15–57; arXiv:cond-mat/0503525; MR2203780 (2007c:82017).

[13] P. C. Hemmer, Virial coefficients for the hard-core gas in two dimensions, *J. Chem. Phys.* 42 (1965) 1116–1118; MR0186278 (32 #3738).

[14] A. P.-H. Yu, *The Fifth Virial Coefficient of a Hard Sphere Gas and a Hard Disk Gas*, Ph.D. thesis, Rice Univ., 1967.

[15] S. Kim and D. Henderson, Exact values of two cluster integrals in the fifth virial coefficient for hard spheres, *Phys. Lett. A* 27 (1968) 378–379.

[16] K. W. Kratky, Fifth virial coefficient for a system of hard disks, *Physica A* 85 (1976) 607–615.

[17] K. W. Kratky, A new graph expansion of virial coefficients, *J. Stat. Phys.* 27 (1982) 533–551; MR0659808 (83h:82003).

[18] N. Clisby, The fifth virial coefficient for hard discs, 54^{th} Annual Meeting of the Australian Mathematical Society, 2010, Univ. of Queensland.

[19] N. Clisby, Connections between graph theory and the virial expansion, Discrete Mathematics Research Group meeting, 2012, Monash Univ..

[20] F. H. Ree and W. G. Hoover, Fifth and sixth virial coefficients for hard spheres and hard discs, *J. Chem. Phys.* 40 (1964) 939–950.

[21] F. H. Ree and W. G. Hoover, Seventh virial coefficients for hard spheres and hard discs, *J. Chem. Phys.* 46 (1967) 4181–4196.

[22] K. W. Kratky, Fifth to tenth virial coefficients of a hard-sphere fluid, *Physica A* 87 (1977) 584–600.

[23] K. W. Kratky, Overlap graph representation of B_6 and B_7, *J. Stat. Phys.* 29 (1982) 129–138; MR0676934 (84f:82010).
[24] E. J. Janse van Rensburg, Virial coefficients for hard discs and hard spheres, *J. Phys. A* 26 (1993) 4805–4818.
[25] E. J. Janse van Rensburg and G. M. Torrie, Estimation of multidimensional integrals: Is Monte Carlo the best method? *J. Phys. A* 26 (1993) 943–953; MR1211087 (93m:65008).
[26] A. Yu. Vlasov, X.-M. You and A. J. Masters, Monte-Carlo integration for virial coefficients re-visited: Hard convex bodies, spheres with a square-well potential and mixtures of hard spheres, *Molecular Phys.* 100 (2002) 3313–3324.
[27] S. Labik, J. Kolafa and A. Malijevsky, Virial coefficients of hard spheres and hard disks up to the ninth, *Phys. Rev. E* 71 (2005) 021105.
[28] C. Zhang and B. M. Pettitt, Computation of high-order virial coefficients in high-dimensional hard-sphere fluids by Mayer sampling, *Molecular Phys.* 112 (2014) 1427–1447.
[29] A. J. Schultz and D. A. Kofke, Fifth to eleventh virial coefficients of hard spheres, *Phys. Rev. E* 90 (2014) 023301.
[30] M. P. Fewell, Area of common overlap of three circles, Technical Note DSTO-TN-0722, Australian Dept. of Defence, 2006.

2.21 Strong Triangle Inequality

Let a, b, c denote the sides of a triangle, h denote the altitude to side c, and γ denote the angle opposite c. It is known that the inequality [1, 2]

$$a + b > c + h$$

is true for all triangles with $\gamma < \pi - 4\arctan(1/2) = 1.2870022175... \approx 73.74°$ but is false for all triangles with $\gamma \geq \pi/2$. For the intermediate range of angles, there are several ways to express the percentage of triangles satisfying the inequality. Certain authors [3] assumed that the angles α, β opposite sides a, b are uniformly distributed on the region

$$0 < \alpha < \pi, \quad 0 < \beta < \pi, \quad \alpha + \beta < \pi.$$

Let

$$K = \int_0^{\pi/2} \left[2\arctan\left(1 - \tan\left(\frac{x}{2}\right)\right) - \left(\frac{\pi}{2} - x\right)\right] dx = 0.2922839193...$$

for convenience. Supposing $0 < \gamma < \pi$, the probability that a random triangle satisfies the inequality is

$$1 - \frac{2}{\pi^2}\left(\frac{\pi^2}{8} + K\right) = 1 - \frac{1}{4} - \frac{2K}{\pi^2} = 0.690770....$$

Supposing instead $0 < \gamma < \pi/2$, the probability is

$$1 - \frac{8K}{3\pi^2} = 0.921027....$$

(This is why $a+b>c+h$ is said to hold for "most" triangles with acute γ.) Supposing instead $\pi-4\arctan(1/2)<\gamma<\pi/2$, the probability is

$$1-\frac{8K}{64\arctan(1/2)^2-\pi^2}=0.398657....$$

We wonder about the odds corresponding to a *fixed* angle γ in the intermediate range. This is found by integrating the joint (α,β)-density

$$\begin{cases} \frac{2}{\pi^2} & \text{if } 0<x<\pi,\ 0<y<\pi \text{ and } x+y<\pi, \\ 0 & \text{otherwise} \end{cases}$$

to obtain a marginal density

$$f(x)=\int_0^{\pi-x}\frac{2}{\pi^2}dy=\frac{2}{\pi^2}(\pi-x);$$

the desired probability is hence

$$\begin{aligned} 1-\frac{1}{f(\gamma)}\int_z^w\frac{2}{\pi^2}dx &= \frac{2z}{\pi-\gamma} \\ &= \begin{cases} 1 & \text{if } \gamma=\pi-4\arctan(1/2), \\ 0.770368... & \text{if } \gamma=5\pi/12=75^\circ, \\ 0.335397... & \text{if } \gamma=11\pi/24=82.5^\circ, \\ 0.166040... & \text{if } \gamma=23\pi/48=86.25^\circ, \\ 0 & \text{if } \gamma=\pi/2, \end{cases} \end{aligned}$$

where z is the smallest positive solution of the equation

$$\tan\left(\frac{z}{2}\right)+\cot\left(\frac{\gamma+z}{2}\right)=1$$

and $w=\pi-\gamma-z$.

We additionally wonder about the odds corresponding to a different choice of distribution for α, β. If the triangle vertices are independent random Gaussian points in two dimensions, all of which have mean vector zero and covariance matrix identity, then we have joint (α,β)-density [4, 5]

$$\begin{cases} \frac{6}{\pi}\frac{\sin(x)\sin(y)\sin(x+y)}{\left(\sin(x)^2+\sin(y)^2+\sin(x+y)^2\right)^2} & \text{if } 0<x<\pi,\ 0<y<\pi \text{ and } x+y<\pi, \\ 0 & \text{otherwise.} \end{cases}$$

Integrating with respect to y over $[0,\pi-x]$, a marginal density [4, 6]

$$g(x)=\frac{3}{\pi}\frac{\cos(x)}{\left(4-\cos(x)^2\right)^{3/2}}\left(\frac{\pi}{2}+\arcsin\left(\frac{\cos(x)}{2}\right)\right)+\frac{3}{\pi}\frac{1}{4-\cos(x)^2}$$

emerges. The desired probability becomes

$$1-\frac{1}{g(\gamma)}\int\limits_{z}^{w}\frac{6}{\pi}\frac{\sin(\gamma)\sin(x)\sin(\gamma+x)}{\left(\sin(\gamma)^2+\sin(x)^2+\sin(\gamma+x)^2\right)^2}dx$$

$$=\begin{cases}1 & \text{if } \gamma=\pi-4\arctan(1/2),\\ 0.662855... & \text{if } \gamma=5\pi/12=75^\circ,\\ 0.141612... & \text{if } \gamma=11\pi/24=82.5^\circ,\\ 0.034758... & \text{if } \gamma=23\pi/48=86.25^\circ,\\ 0 & \text{if } \gamma=\pi/2\end{cases}$$

where z, w are exactly as before.

A benefit of working with 2D Gaussian triangles is that the joint density for sides a, b, c is available [4, 7]:

$$\begin{cases}\dfrac{2}{3\pi}\dfrac{a\,b\,c}{\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}}\exp\left(-\dfrac{1}{6}\left(a^2+b^2+c^2\right)\right) \\ \qquad \text{if } |a-b|<c<a+b,\\ 0 \qquad \text{otherwise.}\end{cases}$$

The (ordinary) triangle inequality gives rise to an expected difference

$$\mathrm{E}(a+b-c)=\sqrt{\pi}=1.7724538509...$$

and an expected ratio

$$\mathrm{E}\left(\frac{a+b}{c}\right)\approx 2.94.$$

For the strong triangle inequality, we utilize a variation of the density function

$$\frac{1}{3\pi}a\,b\exp\left[-\frac{1}{3}\left(a^2-a\,b\cos(\gamma)+b^2\right)\right]$$

over $a>0$, $b>0$, $0<\gamma<\pi$ to compute the expected difference

$$\begin{aligned}\mathrm{E}(a+b-c-h) &= \mathrm{E}\left(a+b-\sqrt{a^2-2a\,b\cos(\gamma)+b^2}-\frac{a\,b\sin(\gamma)}{\sqrt{a^2-2a\,b\cos(\gamma)+b^2}}\right)\\ &\approx 0.79\end{aligned}$$

and the expected ratio

$$\begin{aligned}\mathrm{E}\left(\frac{a+b}{c+h}\right) &= \mathrm{E}\left(\frac{a+b}{\sqrt{a^2-2a\,b\cos(\gamma)+b^2}+\frac{a\,b\sin(\gamma)}{\sqrt{a^2-2a\,b\cos(\gamma)+b^2}}}\right)\\ &\approx 1.44.\end{aligned}$$

Other possible models to consider are 3D Gaussian triangles [4] and broken L triangles of unit perimeter [8].

Let us turn attention away from a Euclidean setting and toward the hyperbolic plane. The strong triangle inequality holds for any hyperbolic triangle if $\gamma < \xi$ where $\xi = 1.1496525950... \approx 65.87^\circ$ is the smallest positive solution of the equation [9]

$$-1 - \cos(\xi) + \sin(\xi) + \sin\left(\frac{\xi}{2}\right)\sin(\xi) = 0.$$

Analogous probabilistic results for uniform angles are uncovered in [10]. An unusual feature of the latter paper is its careful analysis – numerical results here can be computed to arbitrary precision and the error can be bounded – we wonder if such rigour can be feasibly carried over to the Gaussian case.

[1] H. R. Bailey and R. Bannister, A stronger triangle inequality, *College Math. J.* 28 (1997) 182–186; MR1444005 (98b:51021).
[2] M. S. Klamkin, A sharp triangle inequality, *College Math. J.* 29 (1998) 33.
[3] V. Faiziev, R. C. Powers and P. K. Sahoo, When can one expect a stronger triangle inequality? *College Math. J.* 44 (2013) 24–31; MR3005907.
[4] S. R. Finch, Random triangles, *this volume*, §5.10.
[5] K. V. Mardia, Recent directional distributions with applications, *Statistical Distributions in Scientific Work*, v. 6, *Applications in Physical, Social, and Life Sciences*, Proc. 1980 NATO Advanced Study Institute, Univ. Trieste, ed. C. Taillie, G. P. Patil and B. A. Baldessari, Reidel, 1981, pp. 1–19; MR0656354 (84b:62071).
[6] W. S. Kendall, *Some Problems in Probability Theory*, Ph.D. thesis, Oxford Univ., 1979.
[7] P. Clifford and N. J. B. Green, Distances in Gaussian point sets, *Math. Proc. Cambridge Philos. Soc.* 97 (1985) 515–524; MR0778687 (86i:62091).
[8] S. R. Finch, Random triangles. VI, *this volume*, §5.15.
[9] M. Baker and R. C. Powers, A stronger triangle inequality for neutral geometry, *Forum Geom.* 7 (2007) 25–29; MR2282228.
[10] C. Biró and R. C. Powers, A strong triangle inequality in hyperbolic geometry, arXiv:1507.04033.

3

Real and Complex Analysis

3.1 Radii in Geometric Function Theory

First, we talk about geometry. A region $R \subseteq \mathbb{C}$ is **convex** if, for any two points $p, q \in R$, the line segment $pq \subseteq R$. A region $R \subseteq \mathbb{C}$ is **starlike** with respect to the origin if $0 \in R$ and if, for any point $p \in R$, the line segment $0p \subseteq R$.

Next, we talk about functions. A complex analytic function f defined on an open region is **univalent** (or **schlicht**) if f is one-to-one; that is, $f(z) = f(w)$ if and only if $z = w$. Let

$$D = \{z : |z| < 1\} \quad \text{(the open disk of radius 1)},$$

$$E = \{z : 0 < |z| < 1\} \quad \text{(the open punctured disk)},$$

$$S = \left\{ \text{univalent } f \text{ on } D \text{ with } f(z) = z + \sum_{n=2}^{\infty} a_n z^n \right\},$$

$$\Sigma = \left\{ \text{univalent } f \text{ on } E \text{ with } f(z) = \frac{1}{z} + \sum_{n=0}^{\infty} b_n z^n \right\}.$$

Geometry and functions now come together. The various subclasses of S include

$$\begin{aligned} CV &= \{f \in S : f(D) \text{ is convex}\} \\ &= \left\{ f \in S : \operatorname{Re}\left(1 + z\frac{f''(z)}{f'(z)}\right) > 0 \text{ for all } z \in D \right\}, \end{aligned}$$

the class of convex functions on D, and

$$\begin{aligned} ST &= \{f \in S : f(D) \text{ is starlike with respect to } 0\} \\ &= \left\{ f \in S : \operatorname{Re}\left(z\frac{f'(z)}{f(z)}\right) > 0 \text{ for all } z \in D \right\}, \end{aligned}$$

the class of starlike functions on D. We will mostly discuss S (the analytic case), but will mention Σ (the meromorphic case) occasionally in the following [1–5].

3.1.1 *Radius of Convexity*

Define $D_r = \{z : |z| < r\}$, the open disk of radius r, for each $r > 0$. For each $f \in S$, let $r(f)$ be the supremum of all numbers r such that $f(D_r)$ is convex. The **radius of convexity** for S is [1]

$$\rho_{cv}(S) = \inf_{f \in S} r(f) = 2 - \sqrt{3} = 0.2679491924...$$

and is achieved by the Koebe function $f(z) = z(1 - z)^{-2}$. This fact was first proved by Nevanlinna [6]. Generalization of ρ_{cv} to any subclass of S gives rise to some interesting optimization problems. Trivially we have

$$\rho_{cv}(CV) = 1, \quad \rho_{cv}(ST) = 2 - \sqrt{3}$$

(the latter follows since the Koebe function is starlike). Define, however, the special class of starlike functions of order α:

$$S_\alpha^* = \left\{ f \in S : \mathrm{Re}\left(z\frac{f'(z)}{f(z)} \right) > \alpha \text{ for all } z \in D \right\}.$$

Zmorovic [7], extending work in [8–10], proved that

$$\rho_{cv}(S_\alpha^*) = \begin{cases} \dfrac{1}{2 - 3\alpha + \sqrt{(1-\alpha)(3-5\alpha)}} & \text{if } 0 \le \alpha < \alpha_0, \\ \left(\dfrac{5\alpha - 1}{4\alpha^2 - \alpha + 1 + 4\alpha\sqrt{\alpha^2 - 3\alpha + 2}} \right)^{\frac{1}{2}} & \text{if } \alpha_0 \le \alpha < 1, \end{cases}$$

where $\alpha_0 = 0.3349596751...$ is the smallest positive zero of $20\alpha^4 - 52\alpha^3 + 15\alpha^2 + 12\alpha - 4$. Note that $\rho_{cv}(S_0^*) = 2 - \sqrt{3}$, as expected.

We turn attention to the class Σ. Define $E_r = \{z : 0 < |z| < r\}$ and, for $f \in \Sigma$, let $r(f)$ be the supremum of all numbers r such that the complement of $f(E_r)$ in $\mathbb{C}$ is convex. Goluzin [5, 11] proved that

$$\rho_{cv}(\Sigma) = \inf_{f \in \Sigma} r(f) = x = 0.5600798519...,$$

where x is the unique positive solution of the equation

$$\frac{E(x)}{K(x)} + \frac{x^2}{8} - \frac{7}{8} = 0$$

and $K(x)$, $E(x)$ are complete elliptic integrals of the first and second kind [12]. Letting

$$\Sigma_\beta^* = \left\{ f \in \Sigma : \mathrm{Re}\left(z\frac{f'(z)}{f(z)} \right) < -\beta \text{ for all } z \in E \right\},$$

we also have [7, 9, 11, 13, 14]

$$\rho_{cv}(\Sigma_\beta^*) = \begin{cases} \left(\dfrac{4\beta - 5 + 4\sqrt{\beta^2 - \beta + 1}}{8\beta - 3} \right)^{\frac{1}{2}} & \text{if } 0 \le \beta < \beta_0, \\ \dfrac{1}{\beta + \sqrt{(1-\beta)(3\beta - 1)}} & \text{if } \beta_0 \le \beta < 1, \end{cases}$$

where $\beta_0 = 0.8673407553...$ is the largest positive zero of $12\beta^4 - 28\beta^3 + 33\beta^2 - 20\beta + 4$. Note here that $\rho_{cv}(\Sigma_0^*) = 1/\sqrt{3} = 0.577... > 0.560... = x$. In this case, the extremal function is not starlike, which accounts for the strict inequality.

3.1.2 *Radius of Starlikeness*

For each $f \in S$, let $r(f)$ be the supremum of all numbers r such that $f(D_r)$ is starlike with respect to the origin. The **radius of starlikeness** for S is [1]

$$\rho_{st}(S) = \inf_{f \in S} r(f) = \frac{1 - e^{-\pi/2}}{1 + e^{-\pi/2}} = \tanh\left(\frac{\pi}{4}\right) = 0.6557942026...$$

and this fact was first discovered by Grunsky [15].

Goluzin [5, 16] found several interesting generalizations. Define a region $R \subseteq \mathbb{C}$ to be ***n*-starlike** with respect to the origin if $0 \in R$ and if every point of R can be connected with 0 by a piecewise linear curve that lies entirely in R and that consists of no more than n line segments. Let δ_n be the supremum of all r such that an arbitrary $f \in S$ maps D_r onto an n-starlike region with respect to 0. Then

$$\tanh\left(\frac{\pi}{4}\right) = \delta_1 \le \delta_2 \le \delta_3 \le \cdots, \quad \delta_n \ge \tanh\left(\frac{n\pi}{4}\right),$$

but values for δ_n, $n \ge 2$, are unknown. See also [17, 18].

Likewise, let ϵ_n be the supremum of all r such that an arbitrary $f \in \Sigma$ maps E_r onto a region, the complement of which is n-starlike with respect to 0. Then

$$0.85 < \epsilon_1, \quad 1 - 1.11 \exp\left(\frac{-n\pi}{2}\right) < \epsilon_n \quad \text{for all } n > 1.$$

An exact expression for ϵ_1 would be good to see someday.

3.1.3 *Radius of Close-to-Convexity*

A region $R \subseteq \mathbb{C}$ is **close-to-convex** (or **linearly accessible**) if its complement is a union of closed half-lines such that the corresponding open half-lines are pairwise disjoint. Any starlike region is close-to-convex. A half-annulus is also close-to-convex, but this property fails for any larger subsection of an annulus.

An analytic function $f: D \to \mathbb{C}$ is close-to-convex if $f(D)$ is close-to-convex. Equivalently, f is close-to-convex if there is a convex function $g: D \to \mathbb{C}$ such that $\mathrm{Re}(f'(z)/g'(z)) > 0$ for all $z \in D$ [1, 19–25]. It can be shown that every close-to-convex function is univalent.

Define

$$CC = \{f \in S : f(D) \text{ is close-to-convex}\}$$
$$= \left\{ f \in S : \begin{array}{l} \displaystyle\int_{\theta_1}^{\theta_2} \operatorname{Re}\left(1 + z\frac{f''(z)}{f'(z)}\right) d\theta > -\pi, \text{ where } z = re^{i\theta}, \\ \text{for each } 0 < r < 1 \text{ and each pair } 0 < \theta_1 < \theta_2 < 2\pi \end{array} \right\}.$$

Let $\rho_{cc}(S)$ be the supremum of all r such that an arbitrary $f \in S$ maps D_r onto a close-to-convex region. Krzyz [26] determined that

$$\rho_{cc}(S) = y = 0.8098139153...,$$

where y is the unique real solution of the equation

$$2\arctan\left(\frac{\kappa(y)}{\lambda(y)}\right) + \ln\left(1 + \lambda(y)^2\right) - 2\ln\left(\frac{2y}{1 - y^2}\right) = 0$$

in the interval $0 < y < 1$, $\kappa(y) = (1 + y^2)/(1 - y^2)$, and $\lambda = \lambda(y)$ is the unique real solution of the equation

$$\lambda^3 - \kappa(y)\lambda^2 + \kappa(y)^2\lambda - \kappa(y) = 0.$$

Sizuk [27] extended this result to the class of close-to-convex functions of order γ.

3.1.4 *Radius of Convexity in One Direction*

A region $R \subseteq \mathbb{C}$ is **convex in the direction of the imaginary axis** if, for every vertical line L, the set $L \cap R$ is either empty or connected. Any region that is convex in one direction can be rotated so that it is convex in the imaginary direction [3, 28, 29].

Define

$$CD = \{f \in S : f(D) \text{ is convex in the imaginary direction}\}$$

and let $\rho_{cd}(S)$ be the supremum of all numbers r such that an arbitrary $f \in S$ maps D_r onto a region that is convex in the imaginary direction. Umezawa [30] and Goodman & Saff [31] proved that

$$0.394... = 4 - \sqrt{13} \le \rho_{cd}(S) \le \sqrt{2} - 1 = 0.414....$$

The exact value of this constant is unknown.

A subclass of CD was considered by Hengartner & Schober [32]:

$$\{f \in S : \operatorname{Re}\left((1 - z^2)f'(z)\right) \ge 0 \text{ for all } z \in D\}$$

but we omit details. See also [33, 34].

3.1.5 *Radius of Majorization*

Let $f: D \to \mathbb{C}$ be analytic with $f(0) = 0$ and $f'(0) \geq 0$. Let $F \in S$. The function f is **subordinate** to F, written $f \preceq F$, if $f(D_r) \subseteq F(D_r)$ for all $0 < r < 1$ [1, 35].

Shah [36, 37], verifying conjectures of Goluzin [5, 38], proved that if $f \preceq F$, then

$$|f(z)| \leq |F(z)| \quad \text{for all } |z| \leq \tfrac{1}{2}(3 - \sqrt{5}) = 0.3819660112...,$$

$$|f'(z)| \leq |F'(z)| \quad \text{for all } |z| \leq 3 - 2\sqrt{2} = 0.1715728752....$$

Both of these radii are best possible. If we further assume that f is univalent and $f'(0) > 0$, then [5, 39]

$$|f(z)| \leq |F(z)| \quad \text{for all } |z| \leq u = 0.3908507887...,$$

where u is the unique real solution of

$$\ln\left(\frac{1+u}{1-u}\right) + 2\arctan(u) = \frac{\pi}{2}.$$

Again, this radius of majorization is best possible. Problems as such (subordination implies majorization) were first examined by Biernacki [40].

Converse problems (majorization implies subordination) were studied by Lewandowski [41]. Under the same conditions as earlier, if $|f(z)| \leq |F(z)|$ for all $z \in D$ and f is not necessarily univalent, then $f \preceq F$ in the disk D_v, where $0.21 < v < 0.29$. The exact value of v is unknown. If f is assumed to be univalent, then the constant $u = 0.390...$ arises again [42, 43].

3.1.6 *Radius of Zeroness*

Let $\rho_N(\Sigma)$ be the supremum of all numbers r such that an arbitrary $f \in \Sigma$ never vanishes on the punctured disk E_r. Goluzin [16] proved that $0.86 < \rho_N(\Sigma) \leq \sqrt{3}/2 < 0.867$, but a subsequent theorem of his [5, 44] implies that $\rho_N(\Sigma) = \xi = 0.8649789576...$, where ξ is the unique positive solution of the equation

$$\frac{E(\xi)}{K(\xi)} + \frac{\xi^2}{4} - \frac{3}{4} = 0.$$

This is quite similar to the equation prescribed earlier for the radius of convexity $\rho_{cv}(\Sigma)$.

Given an analytic function f, we may likewise define $\rho_N(f)$ to be the supremum of all numbers r such that f, when restricted to E_r, is never zero. For example,

$$\rho_N(f) = 2|z_0| \quad \text{for } f(z) = z - \frac{1}{2z_0}z^2 \text{ (a quadratic function)}$$

and

$$\rho_N(f) = 2\pi \quad \text{for } f(z) = \exp(z) - 1 \text{ (the exponential function).}$$

3.1.7 *Radius of Univalence*

Given an analytic function f, define the **radius of univalence** of f to be the supremum of all numbers r such that f, when restricted to the disk D_r, is univalent. Let us first consider the case of polynomials. We clearly have

$$\rho_s(f) = |z_0| \quad \text{for } f(z) = z - \frac{1}{2z_0}z^2$$

in the quadratic case. Kakeya's theorem [45–47] provides that

$$\sin\left(\frac{\pi}{n}\right) \leq \frac{\rho_s(f)}{|z_0|} \leq 1 \quad \text{for } f(z) = z + \sum_{k=2}^{n} a_k z^k$$

in the general case, where $n \geq 2$ and $z_0 \neq 0$ is the zero of $f'(z)$ of smallest modulus. These bounds are sharp.

Now, let us consider the case of transcendental functions. We have

$$\rho_s(f) = \pi \quad \text{for } f(z) = \exp(z) - 1,$$

as is well-known (although $f'(z)$ never vanishes); [48]

$$\rho_s(f) = 1.5748375891... \quad \text{for } f(z) = \operatorname{erf}(z),$$

corresponding to the smallest modulus, of points z not on the x-axis, for which $\operatorname{erf}(z)$ is real; [49, 50]

$$\rho_s(f) = 0.9241388730... \quad \text{for } f(z) = \exp(z^2)\operatorname{erf}(z),$$

corresponding to the unique positive solution of $\sqrt{\pi}y\,\operatorname{Im}(f(iy)) = 1$; [51–53]

$$\rho_s(f) = p_{\nu,1} \quad \text{for } f(z) = z^{1-\nu}J_\nu(z),\ \nu > -1,$$

corresponding to the smallest positive zero of $f'(z)$; [54]

$$\rho_s(f) = 0.5040830082... \quad \text{for } f(z) = 1/\Gamma(z),$$

corresponding to the smallest positive zero of $\Gamma'(-z)$; and [55]

$$\rho_s(f) = 0.4616321449... \quad \text{for } f(z) = \Gamma(z+1),$$

corresponding to the smallest positive zero of $\Gamma'(z+1)$. See also [56].

We digress briefly to other radii. For $f(z) = \exp(z) - 1$, it is known that [57, 58]

$$\rho_{cv}(f) = 1, \quad \rho_{st}(f) = 2.8329700604...$$

and the latter corresponds to $\sqrt{1+\eta^2}$, where η is the smallest positive solution of the equation

$$\eta\sin(\eta) + \cos(\eta) = \frac{1}{e}.$$

See also [59, 60].

3.1.8 *Sums and Products*

Here are two procedures for combining univalent functions:

$$S+S=\{h:h(z)=tf(z)+(1-t)g(z) \text{ for some } f,g\in S \text{ and } 0\le t\le 1\},$$

$$S\cdot S=\left\{h:h(z)=f(z)^t g(z)^{1-t} \text{ for some } f,g\in S \text{ and } 0\le t\le 1\right\}.$$

On the one hand, MacGregor [61] demonstrated that

$$\rho_s(S+S)=\sin\left(\frac{\pi}{8}\right)=\frac{1}{2}\sqrt{2-\sqrt{2}}=0.3826834323...$$

$$\rho_s(CV+CV)=\frac{\sqrt{2}}{2}=0.7071067811...$$

and Robertson [62] showed that

$$\rho_s(ST+ST)=\chi=0.4035150049...$$

where χ is the unique positive zero of $\chi^6+5\chi^4+79\chi^2-13$. Further results appear in [63–65]. On the other hand, we have [3]

$$CV\cdot CV\subseteq ST\cdot ST\subseteq ST, \quad CV\cdot CV\nsubseteq CV$$

but virtually nothing is known about the class $S\cdot S$.

3.1.9 *Derivatives and Integrals*

Define the following classes of functions:

$$T=\left\{f:f(z)=\frac{1}{2}\frac{d}{dz}(zg(z)) \text{ for some } g\in S\right\},$$

$$U_\alpha=\left\{f:f(z)=\int_0^z\left(\frac{g(w)}{w}\right)^\alpha dw \text{ for some } g\in S\right\},$$

$$V_\beta=\left\{f:f(z)=\int_0^z g'(w)^\beta dw \text{ for some } g\in S\right\},$$

where α, β are complex numbers and hence the logarithmic branch is selected so that $f'(0)=1$. Barnard [66, 67] and Pearce [68], building on Robinson [69], proved that

$$0.49<\rho_s(T)\le\tfrac{1}{2}, \quad 0.435<\rho_{st}(T)<0.445.$$

In particular, these two constants must be distinct.

Biernacki [70] claimed that $\rho_s(U_1)=1$, but this was disproved by Krzyz & Lewandowski [71]. It was later shown [72] that $0.91<\rho_s(U_1)\le\tanh(\pi)<0.9963$.

Let A denote the set of all complex numbers α for which $U_\alpha \subseteq S$. Kim & Merkes [73] proved that $D_{1/4} \subseteq A \subseteq D_{1/2}$; we wonder whether $D_r \subseteq A$ for some $r > 1/4$.

Trivially $\rho_s(V_1) = 1$. Let B denote the set of all complex numbers β for which $V_\beta \subseteq S$. Royster [74] and Pfaltzgraff [75] proved that $D_{1/4} \subseteq B \subseteq D_{1/3} \cup \{1\}$; we again wonder whether $D_r \subseteq B$ for some $r > 1/4$. See also [76, 77].

[1] P. L. Duren, *Univalent Functions*, Springer-Verlag, 1983, pp. 40–51, 95–98, 202–209; MR0708494 (85j:30034).

[2] A. W. Goodman, *Univalent Functions*, v. 1, Mariner, 1983, pp. 119–121, 193–197; MR0704183 (85j:30035a).

[3] A. W. Goodman, *Univalent Functions*, v. 2, Mariner, 1983, pp. 84–172; MR0704184 (85j:30035b).

[4] W. K. Hayman, *Multivalent Functions*, 2nd ed., Cambridge Univ, Press, 1994, pp. 226–227; MR1310776 (96f:30003).

[5] G. M. Goluzin, *Geometric Theory of Functions of a Complex Variable* (in Russian), 2nd ed., Izdat. Akad. Nauk SSSR, 1966; Engl. transl. in *Amer. Math. Soc. Transl.*, v. 26, 1969, pp. 136–140, 165–170, 375–379, 599, 643–644; MR0219714 (36 #2793) and MR0247039 (40 #308).

[6] R. Nevanlinna, Über die Eigenschaften einer analytischen Funktion in der Umgebung einer singulären Stelle oder Linie, *Acta Soc. Sci. Fenn.,* v. 50 (1922) n. 5, 1–46.

[7] V. A. Zmorovic, On bounds of convexity for starlike functions of order α in the circle $|z| < 1$ and in the circular region $0 < |z| < 1$ (in Russian), *Mat. Sbornik* 68 (1965) 518–526; Engl. transl. in *Amer. Math. Soc. Transl.*, v. 80, 1969, pp. 203–213; MR0197712 (33 #5875).

[8] T. H. MacGregor, The radius of convexity for starlike functions of order $\frac{1}{2}$, *Proc. Amer. Math. Soc*. 14 (1963) 71–76; MR0150282 (27 #283).

[9] M. S. Robertson, Some radius of convexity problems, *Michigan Math. J*. 10 (1963) 231–236; MR0151596 (27 #1580).

[10] A. Schild, On starlike functions of order α, *Amer. J. Math.* 87 (1965) 65–70; MR0174729 (30 #4929).

[11] M. S. Robertson, Extremal problems for analytic functions with positive real part and applications, *Trans. Amer. Math. Soc.* 106 (1963) 236–253; MR0142756 (26 #325).

[12] S. R. Finch, Random triangles, *this volume*, §5.10.

[13] K. S. Padmanabhan, On the radius of convexity of a certain class of meromorphically starlike functions in the unit circle, *Math. Z*. 91 (1966) 308–313; MR0190319 (32 #7732).

[14] V. Singh and R. M. Goel, On radii of convexity and starlikeness of some classes of functions, *J. Math. Soc. Japan* 23 (1971) 323–339; MR0281903 (43 #7617).

[15] H. Grunsky, Zwei Bemerkungen zur konformen Abbildung, *Jahresbericht Deutsch. Math.-Verein.* 43 (1934) 140–143.

[16] G. M. Goluzin, Zur Theorie der schlichten konformen Abbildungen (in Russian), *Mat. Sbornik* 42 (1935) 169–190.

[17] S. Gelfer, Sur les bornes de l'étoilement et de la convexité des fonctions p-valentes (in Russian), *Mat. Sbornik* 16 (1945) 81–86; MR0012662 (7,55a).

[18] I. A. Aleksandrov, Domains of definition of some functionals on the class of functions univalent and regular in a circle (in Russian), *Issledovaniia Po Sovremennym Problemam Teorii Funkcii Kompleksnovo Peremennovo*, Gosudarstv. Izdat. Fiz.-Mat. Lit., 1960, pp. 39–45; MR0114928 (22 #5744).

[19] M. Biernacki, Sur la représentation conforme des domaines linéairement accessibles, *Prace Matematyczno-Fizyczne* 44 (1936) 293–314.

[20] W. Kaplan, Close-to-convex schlicht functions, *Michigan Math. J.* 1 (1952) 169–185 (1953); MR0054711 (14,966e).

[21] Z. Lewandowski, Sur l'identité de certaines classes de fonctions univalentes. I, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 12 (1958) 131–146; MR0130354 (24 #A217).

[22] Z. Lewandowski, Sur l'identité de certaines classes de fonctions univalentes. II, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 14 (1960) 19–46; MR0156958 (28 #200).

[23] A. Bielecki and Z. Lewandowski, Sur un théorème concernant les fonctions univalentes linéairement accessibles de M. Biernacki, *Ann. Polon. Math.* 12 (1962) 61–63; MR0147636 (26 #5151).

[24] V. A. Pohilevic, The equivalence of two classes of univalent functions (in Russian), *Teor. Funkcii Funkcional Anal. i Prilozhen* 8 (1969) 57–62; MR0262482 (41 #7088).

[25] W. Koepf, On close-to-convex functions and linearly accessible domains, *Complex Variables Theory Appl.* 11 (1989) 269–279; MR1007662 (91f:30010).

[26] J. Krzyz, The radius of close-to-convexivity within the family of univalent functions, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* 10 (1962) 201–204; MR0148887 (26 #6384).

[27] P. I. Sizuk, The radius of almost convexity of order α in a class of schlicht functions (in Russian), *Mat. Zametki* 20 (1976) 105–112; Engl. transl. in *Math. Notes* 20 (1976) 616–620; MR0425103 (54 #13061).

[28] L. Fejér, Neue Eigenschaften der Mittelwerte bei den Fourierreihen, *J. London Math. Soc.* 8 (19133) 53–62.

[29] M. S. Robertson, On the theory of univalent functions, *Annals of Math.* 37 (1936) 374–408.

[30] T. Umezawa, Analytic functions convex in one direction, *J. Math. Soc. Japan* 4 (1952) 194–202; MR0051313 (14,461b).

[31] A. W. Goodman and E. B. Saff, On univalent functions convex in one direction, *Proc. Amer. Math. Soc*. 73 (1979) 183–187; MR0516461 (80e:30006).

[32] W. Hengartner and G. Schober, Analytic functions close to mappings convex in one direction, *Proc. Amer. Math. Soc*. 28 (1971) 519–524; MR0277704 (43 #3437).

[33] R. M. Goel, On a class of analytic functions, *Indian J. Math.* 17 (1975) 9–20; MR0603705 (82i:30011).

[34] M. Dorff, I. Naraniecka and J. Szynal, Doubly close-to-convex functions, *J. Math. Anal. Appl.* 290 (2004) 55–62; MR2032224 (2004k:30024).

[35] D. M. Campbell, Majorization-subordination theorems for locally univalent functions, *Bull. Amer. Math. Soc.* 78 (1972) 535–538; MR0299769 (45 #8817).

[36] T.-S. Shah, Goluzin's number $(3-\sqrt{5})/2$ is the radius of superiority in subordination, *Sci. Record* 1 (1957) 219–222; MR0100094 (20 #6530).

[37] T.-S. Shah, On the radius of superiority in subordination, *Sci. Record* 1 (1957) 329–333; MR0100095 (20 #6531).

[38] G. M. Goluzin, On majorants of subordinate analytic functions. I (in Russian), *Mat. Sbornik* 29 (1951) 209–224; MR0043209 (13,223b).

[39] G. M. Goluzin, On majoration of subordinate analytic functions. II (in Russian), *Mat. Sbornik* 29 (1951) 593–602; MR0044644 (13,454d).

[40] M. Biernacki, Sur les fonctions univalentes, *Mathematica (Cluj)* 12 (1936) 49–64.

[41] Z. Lewandowski, Sur les majorantes des fonctions holomorphes dans le cercle $|z| < 1$, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 15 (1961) 5–11; MR0142745 (26 #314).

[42] Z. Lewandowski, Starlike majorants and subordination, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 15 (1961) 79–84; MR0138743 (25 #2186).

[43] A. Bielecki and Z. Lewandowski, Sur certaines majorantes des fonctions holomorphes dans le cercle unité, *Colloq. Math.* 9 (1962) 299–303; MR0152650 (27 #2625).

[44] G. M. Goluzin, Zur Theorie der schlichten Funktionen (in Russian), *Mat. Sbornik* 12 (1943) 48–55; MR0009053 (5,93f).

[45] S. Kakeya, On zeros of a polynomial and its derivatives, *Tôhuku Math. J.* 11 (1917) 5–16.

[46] L. N. Cakalov, On domains of univalence of certain classes of analytic functions (in Russian), *Dokl. Akad. Nauk SSSR* 132 (1960) 1277–1279; Engl. transl. in *Soviet Math. Dokl*. 1 (1960) 781–783; MR0122965 (23 #A296).

[47] A. Gluchoff and F. Hartmann, On a "much underestimated" paper of Alexander, *Arch. Hist. Exact Sci*. 55 (2000) 1–41; MR1780441 (2001i:30001).

[48] E. Kreyszig and J. Todd, The radius of univalence of the error function, *Numer. Math*. 1 (1959) 78–89; MR0101918 (21 #724).

[49] E. Kreyszig and J. Todd, On the radius of univalence of the function $\exp z^2 \int_0^z \exp(-t^2)dt$, *Pacific J. Math*. 9 (1959) 123–127; MR0107032 (21 #5759).

[50] Y. L. Luke, The radius of univalence of the function $\exp z^2 \int_0^z \exp(-t^2)dt$, *Numer. Math*. 3 (1961) 76–78; MR0118844 (22 #9613).

[51] E. Kreyszig and J. Todd, The radius of univalence of Bessel functions. I, *Illinois J. Math*. 4 (1960) 143–149; MR0110827 (22 #1695).

[52] R. K. Brown, Univalence of Bessel functions, *Proc. Amer. Math. Soc*. 11 (1960) 278–283; MR0111846 (22 #2706).

[53] H. S. Wilf, The radius of univalence of certain entire functions, *Illinois J. Math*. 6 (1962) 242–244; MR0138734 (25 #2177).

[54] E. P. Merkes, M. S. Robertson and W. T. Scott, On products of starlike functions, *Proc. Amer. Math. Soc*. 13 (1962) 960–964; MR0142747 (26 #316).

[55] E. P. Merkes and M. Salmassi, On univalence of certain infinite products, *Complex Variables Theory Appl*. 33 (1997) 207–215; MR1624939 (99b:30023).

[56] G. V. Kuzmina, Numerical determination of radii of univalence of analytic functions (in Russian), *Trudy Mat. Inst. Steklov.* 53 (1959) 192–235; MR0118841 (22 #9610).

[57] P. T. Mocanu, Sur le rayon de stellarité des fonctions univalentes (in Romanian), *Acad. R. P. Romêne Fil. Cluj Stud. Cerc. Mat.* 11 (1960) 337–341; MR0146371 (26 #3893).

[58] P. T. Mocanu, About the radius of starlikeness of the exponential function, *Studia Univ. Babes-Bolyai Ser. Math.-Phys*. 14 (1969) 35–40; MR0257339 (41 #1990).

[59] D. Coman, The radius of starlikeness for the error function, *Studia Univ. Babes-Bolyai Math*. 36 (1991) 13–16; MR1280904.

[60] D. Coman and I. Tigan, The close-to-convexity radius of some functions, *Studia Univ. Babes-Bolyai Math.* 36 (1991) 61–65; MR1281977.

[61] T. H. MacGregor, The univalence of a linear combination of convex mappings, *J. London Math. Soc*. 44 (1969) 210–212; MR0236369 (38 #4665).

[62] M. S. Robertson, The sum of univalent functions, *Duke Math. J.* 37 (1970) 411–419; MR0264052 (41 #8649).

[63] G. Labelle and Q. I. Rahman, Remarque sur la moyenne arithmétique de fonctions univalentes convexes, *Canad. J. Math.* 21 (1969) 977–981; MR0245775 (39 #7081).

[64] D. M. Campbell, The radius of convexity of a linear combination of functions in K, $CV_k(\beta)$, S or U_α, *Canad. J. Math*. 25 (1973), 982–985; MR0328045 (48 #6387).

[65] D. H. Hamilton and P. D. Tuan, Radius of starlikeness of convex combinations of univalent starlike functions, *Proc. Amer. Math. Soc.* 78 (1980) 56–58; MR0548084 (81f:30007).
[66] R. W. Barnard, On the radius of starlikeness of $(zf)'$ for f univalent, *Proc. Amer. Math. Soc.* 53 (1975) 385–390; MR0382615 (52 #3497).
[67] R. W. Barnard, On Robinson's $\frac{1}{2}$ conjecture, *Proc. Amer. Math. Soc.* 72 (1978) 135–139; MR0503547 (80j:30014).
[68] K. Pearce, A note on a problem of Robinson, *Proc. Amer. Math. Soc*. 89 (1983) 623–627; MR0718985 (85h:30014).
[69] R. M. Robinson, Univalent majorants, *Trans. Amer. Math. Soc*. 61 (1947) 1–35; MR0019114 (8,370e).
[70] M. Biernacki, Sur l'intégrale des fonctions univalentes, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astr. Phys*. 8 (1960) 29–34; MR0117341.
[71] J. Krzyz and Z. Lewandowski, On the integral of univalent functions, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* 11 (1963) 447–448; MR0153830 (27 #3791).
[72] Z. Lewandowski, On a problem of M. Biernacki, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 17 (1963) 39–41 (1965); MR0196055 (33 #4249).
[73] Y. J. Kim and E. P. Merkes, On an integral of powers of a spirallike function, *Kyungpook Math. J.* 12 (1972) 249–252; MR0320295 (47 #8834).
[74] W. C. Royster, On the univalence of a certain integral, *Michigan Math. J.* 12 (1965) 385–387; MR0183866 (32 #1342).
[75] J. A. Pfaltzgraff, Univalence of the integral of $f'(z)^{\lambda}$, *Bull. London Math. Soc.* 7 (1975) 254–256; MR0387570 (52 #8410).
[76] D. V. Prokhorov and J. Szynal, On the radius of univalence for the integral of $f'(z)^{\alpha}$, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 33 (1979) 157–163; MR0689595 (84c:30034).
[77] M. Dorff and J. Szynal, Linear invariance and integral operators of univalent functions, *Demonstratio Math.* 38 (2005) 47–57; MR2123720 (2005k:30016).

3.2 Numerical Radii of Linear Operators

Let $A : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator. The **numerical radius** $w(A)$ of A is defined by

$$w(A) = \sup_{||x||=1} |x^* A\, x|$$

where x^* denotes the conjugate transpose of $x \in \mathbb{C}^n$. For example, if A is self-adjoint or Hermitian (meaning $A = A^*$), then the eigenvalues $\{\lambda_j\}_{j=1}^n$ of A are all real and $w(A)$ coincides with both the **operator norm** $||A||$ of A:

$$||A|| = \sup_{||x||=1} ||A\, x||$$

and the **spectral radius** $r(A)$ of A:

$$r(A) = \lim_{k\to\infty} ||A^k||^{1/k} = \max_{1\le j\le n} |\lambda_j|.$$

In general, however, these three quantities are not equal [1].

Let A and B be linear operators satisfying $AB = BA$. What is the smallest constant γ such that

$$w(A\,B) \leq \gamma \cdot w(A) \cdot ||B||$$

always? On the one hand, Crabb [2] proved that

$$\gamma \leq \frac{1}{2}\sqrt{2+2\sqrt{3}} = 1.1687....$$

On the other hand, Müller [3], Davidson & Holbrook [4] and Chkliar [5] constructed explicit examples to show that $\gamma > 1.066$. There is interest not only in tightening the bounds on γ, but also in tailoring the sizes of the matrices involved.

We have restricted attention to operators on $\mathbb{C}^n$ for the sake of simplicity only. Given a bounded linear operator A on an arbitrary complex Banach space X, its numerical radius and operator norm are defined by formulation exactly as before. Here, however, x^* is to be interpreted as the bounded linear functional $X \to \mathbb{C}$ that maps x to 1 and that maps $y + \beta\,x$ to β, where β is a scalar and $y \neq \alpha\,x$ for any scalar α. This definition extends what we discussed earlier [6]. It is natural to consider as well the **numerical index** $i(X)$ of the space X:

$$i(X) = \inf_{||A||=1} w(A).$$

For example, $i(\mathbb{C}^n) = 1/2$ for $n > 1$ and $0.3678... = 1/e \leq i(X) \leq 1$ always. The constants $1/e$ and 1 are best possible [7–10]. Of course, the Euclidean l_2 norm is in effect for $\mathbb{C}^n$. Computing the numerical index for $\mathbb{C}^n$ equipped with the l_p norm, where $1 < p < \infty$, $p \neq 2$, is more complicated and remains an open issue [11].

In closing, here is an unrelated problem. Given subsets $\Lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, $M = \{\mu_1, \mu_1, \ldots, \mu_n\}$ of $\mathbb{C}$, define the **optimal matching distance**

$$d(\Lambda, M) = \min_{\pi} \max_{1 \leq j \leq n} |\lambda_j - \mu_{\pi(j)}|,$$

where the minimum is taken over all permutations π on n symbols. Also, let $\sigma(A)$ denote the set (with multiplicities) of all eigenvalues of the $n \times n$ complex matrix A. What can be said about the distance between the eigenvalues of two matrices A and B in terms of the operator norm of their difference? The smallest constant c such that

$$d(\sigma(A), \sigma(B)) \leq c \cdot (||A|| + ||B||)^{1-1/n} \cdot ||A - B||^{1/n}$$

always is known to satisfy $2 \leq c < 16\sqrt{3}/9 = 3.0792...$ [12–15]. Variations on the problems raised here suggest themselves.

Addendum We have discussed only the case of complex spaces; somewhat more is known for real spaces. For example, $i(\mathbb{R}^n) = 0$ for $n > 1$ and $0 \leq i(X) \leq 1$ always.

Further, the numerical index for $\mathbb{R}^2$ equipped with the l_p norm with $1<p<\infty$ satisfies [16–18]

$$\max\left\{2^{-1/p},2^{-1/q}\right\}M_p \le i(\mathbb{R}_p^2)\le M_p,$$

where $1/p+1/q=1$ and

$$M_p=\sup_{0\le t\le 1}\frac{\left|t^{p-1}-t\right|}{t^p+1}.$$

The lower bound is not sharp if $p\neq 2$.

Here is another unrelated problem. A **Minkowski plane** is a real two-dimensional normed linear space X; an example is $\mathbb{R}_p^2$ for $1\le p\le\infty$. Let S denote the unit circle of X. If $x,y\in S$, then clearly $-y\in S$ and the average of distances from x to $\pm y$ is an interesting quantity for study. Letting

$$A_1(X)=\inf_{x\in S}\sup_{y\in S}\left(\frac{||x-y||+||x+y||}{2}\right),\quad A_2(X)=\sup_{x\in S}\sup_{y\in S}\left(\frac{||x-y||+||x+y||}{2}\right)$$

we have [19, 20]

$$\frac{2.5275...}{2}=1.2637...=\frac{3+\sqrt{21}}{6}\le A_1(X)\le\frac{-3+7\sqrt{3}}{6}=1.5207...=\frac{3.0414...}{2}$$

but we do not know whether such bounds are sharp. A Minkowski plane X' exists for which $A_1(X')<1.28405=(2.56811)/2$; hence the lower bound is close. By contrast, $\sqrt{2}\le A_2(X)\le 2$ with equality on the left for $\mathbb{R}_2^2$ and equality on the right for both $\mathbb{R}_1^2$ and $\mathbb{R}_\infty^2$.

[1] K. E. Gustafson and D. K. M. Rao, *Numerical Range. The Field of Values of Linear Operators and Matrices*, Springer-Verlag, 1997, pp. 1–41; MR1417493 (98b:47008).

[2] K. Okubo and T. Ando, Operator radii of commuting products, *Proc. Amer. Math. Soc.* 56 (1976) 203–210; MR0405132 (53 #8927).

[3] V. Müller, The numerical radius of a commuting product, *Michigan Math. J.* 35 (1988) 255–260; MR0959272 (89k:47008).

[4] K. R. Davidson and J. A. R. Holbrook, Numerical radii of zero-one matrices, *Michigan Math. J.* 35 (1988) 261–267; MR0959273 (89k:47007).

[5] V. Chkliar, Numerical radii of simple powers, *Linear Algebra Appl.* 265 (1997) 119–121; MR1466894 (98i:47005).

[6] M. Martín, A survey on the numerical index of a Banach space, *Extracta Math.* 15 (2000) 265–276; MR1823892 (2002b:46027).

[7] G. Lumer, Semi-inner-product spaces, *Trans. Amer. Math. Soc.* 100 (1961) 29–43; MR0133024 (24 #A2860).

[8] H. F. Bohnenblust and S. Karlin, Geometrical properties of the unit sphere of Banach algebras, *Annals of Math.* 62 (1955) 217–229; MR0071733 (17,177a).

[9] B. W. Glickfeld, On an inequality of Banach algebra geometry and semi-inner product space theory, *Illinois J. Math.* 14 (1970) 76–81; MR0253024 (40 #6239).

[10] J. Duncan, C. M. McGregor, J. D. Pryce and A. J. White, The numerical index of a normed space, *J. London Math. Soc.* 2 (1970) 481–488; MR0264371 (41 #8967).

[11] E. Ed-Dari and M. A. Khamsi, The numerical index of the L_p space, *Proc. Amer. Math. Soc.* 134 (2006) 2019–2025; MR2215771 (2006m:46019).
[12] R. Bhatia, *Matrix Analysis*, Springer-Verlag, 1997, pp. 229, 250; MR1477662 (98i:15003).
[13] D. Phillips, Improving spectral-variation bounds with Chebyshev polynomials, *Linear Algebra Appl*. 133 (1990) 165–173; MR1058112 (91d:15035).
[14] R. Bhatia, L. Elsner and G. Krause, Bounds for the variation of the roots of a polynomial and the eigenvalues of a matrix, *Linear Algebra Appl.* 142 (1990) 195–209; MR1077985 (92i:12001).
[15] G. M. Krause, Bounds for the variation of matrix eigenvalues and polynomial roots, *Linear Algebra Appl.* 208/209 (1994) 73–82; MR1287341 (95j:15012).
[16] E. Ed-Dari, On the numerical index of Banach spaces, *Linear Algebra Appl.* 403 (2005) 86–96; MR2140274 (2006f:47005).
[17] M. Martín and J. Merí, A note on the numerical index of the L_p space of dimension two, *Linear Multilinear Algebra* 57 (2009) 201–204; MR2492102 (2010b:46030).
[18] M. Martín, J. Merí and M. Popov, On the numerical index of real $L_p(\mu)$-spaces, *Israel J. Math.* 184 (2011) 183–192; MR2823974 (2012g:46032).
[19] M. Baronti, E. Casini and P. L. Papini, Triangles inscribed in a semicircle, in Minkowski planes, and in normed spaces, *J. Math. Anal. Appl.* 252 (2000) 124–146; MR1797848 (2002a:46007).
[20] M. Baronti, E. Casini and C. Franchetti, The size of a Minkowski ellipse that contains the unit ball, *J. Math. Anal. Appl.* 385 (2012) 278–284; MR2834254 (2012g:46024).

3.3 Coefficient Estimates for Univalent Functions

A complex analytic function f defined on the open unit disk D is **univalent** (or **schlicht**) if f is one-to-one; that is, $f(z)=f(w)$ if and only if $z=w$. We are interested in estimating the coefficients $\{a_n\}_{n=0}^{\infty}$ of the Maclaurin series expansion $\sum_{n=0}^{\infty} a_n z^n$ of $f(z)$. Define a set

$$S=\{f\colon D\to\mathbb{C} : f \text{ is univalent}, f(0)=0 \text{ and } f'(0)=1\}$$

and subsets

$$S_{\mathbb{R}}=\{f\in S : a_n\in\mathbb{R} \text{ for all } n\geq 2\}\,,$$

$$S_{\text{odd}}=\{f\in S : f(z)=-f(-z) \text{ for all } z\in D\}\,,$$

$$S_M=\{f\in S : |f(z)|<M \text{ for all } z\in D\}\,,$$

where $M>1$. On the one hand, the Koebe function

$$\kappa(z)=\frac{z}{(1-z)^2}=\sum_{n=1}^{\infty} n\,z^n$$

is a member of $S_{\mathbb{R}}$ but not of $S_{\text{odd}}\cup S_M$. On the other hand, the Pick function

$$P_M(z)=M\,\kappa^{-1}\left(\frac{\kappa(z)}{M}\right),\quad \text{where } \kappa^{-1}(w)=\frac{2w+1-\sqrt{4w+1}}{2w},$$

is a member of $S_{\mathbb{R}} \cap S_M$ but not of S_{odd}. De Branges [1, 2] proved Bieberbach's famous conjecture [3]:

$$\max_{f \in S} |a_n| = n,$$

which occurs if and only if f is a rotation of κ; equivalently, $f(z) = e^{-i\theta}\kappa(e^{i\theta})$ for some $\theta \in \mathbb{R}$. Actually, he proved something even more subtle: Milin's conjecture, which involves not the coefficients $\{a_n\}$ but rather the **logarithmic coefficients** $\{b_n\}$, where

$$\ln\left(\frac{f(z)}{z}\right) = 2\sum_{n=1}^{\infty} b_n z^n.$$

It is surprising how much material here remains unresolved, even twenty years after de Branges' achievement!

3.3.1 *Bombieri's Conjecture*

While proving a local version of Bieberbach's conjecture, Bombieri [4] speculated about a formula for

$$\sigma_{m,n} = \liminf_{\substack{f \to \kappa \\ f \in S}} \frac{n - \operatorname{Re}(a_n)}{m - \operatorname{Re}(a_m)},$$

where $m \geq 2$, $n \geq 2$ and where $f \to \kappa$ means locally uniform convergence on D (uniform on every compact subset of D). He determined, for example, that

$$\liminf_{\substack{f \to \kappa \\ f \in S}} \frac{3 - \operatorname{Re}(a_3)}{(2 - \operatorname{Re}(a_2))^{3/2}} = \frac{8}{3}$$

and hence $\sigma_{2,3} = 0$. Likewise, $\sigma_{4,3} = 0$. Bshouty & Hengartner [5] proved Bombieri's conjecture for f with real coefficients:

$$\liminf_{\substack{f \to \kappa \\ f \in S_{\mathbb{R}}}} \frac{n - a_n}{m - a_m} = \min_{0 \leq \theta < 2\pi} \frac{n\sin(\theta) - \sin(n\theta)}{m\sin(\theta) - \sin(m\theta)} = \beta_{m,n},$$

but the case of f with complex coefficients was left open. The first counterexample to Bombieri's conjecture was found by Greiner & Roth [6]:

$$\sigma_{3,2} = \frac{e-1}{4e} = 0.1580301397\ldots < 0.25 = \beta_{3,2} = \frac{1}{4}.$$

Prokhorov & Vasil'ev [7] gave additional counterexamples:

$$\sigma_{4,2} = 0.050057\ldots < 0.1 = 1/10 = \beta_{4,2},$$

$$\sigma_{2,4} = 0.969556\ldots < 1 = \beta_{2,4},$$

$$\sigma_{3,4} = 0.791557\ldots < 0.828427\ldots < 2\left(\sqrt{2} - 1\right) = \beta_{3,4}.$$

Their interesting work involves Löwner's differential equation, Pontryagin's maximum principle and the numerical solution of an optimal control system.

3.3.2 *Fekete–Szegö Theorem*

Littlewood & Paley [8] proved that the coefficients in S_{odd} are bounded; that is, there exists $A > 0$ for which $|a_{2n+1}| \le A$ for all $f \in S_{\text{odd}}$ and all $n \ge 1$. In a footnote to their paper, they wrote "No doubt the true bound is given by $A = 1$." It is clearly true that $\max_{f\in S_{\text{odd}}} |a_3| = 1$. Fekete & Szegö [9, 10], however, disproved the Littlewood–Paley conjecture for the next coefficient:

$$\alpha = \max_{f\in S_{\text{odd}}} |a_5| = \frac{1}{2} + e^{-2/3} = 1.0134171190....$$

Schaeffer & Spencer [11] exhibited explicitly the unique extremal function f and noted that $f \in S_{\mathbb{R}}$ as well. They demonstrated that

$$\max_{f\in S_{\text{odd}}\cap S_{\mathbb{R}}} |a_{2n+1}| > 1$$

for each $n \ge 2$. Leeman [12] studied the case $n = 3$:

$$\max_{f\in S_{\text{odd}}\cap S_{\mathbb{R}}} |a_7| = \frac{1090}{1083} = 1.0064635272...$$

and such extremal functions f must additionally satisfy $a_3 = \pm 18/19$ and $a_5 = 351/261$. The occurrence of rational numbers here is quite surprising. The best general estimate is due to Hu Ke [13], improving upon [8, 14–18]:

$$\max_{f\in S_{\text{odd}}} |a_{2n+1}| \le 1.1305....$$

Ke's proof is based on Milin's conjecture (now de Branges' theorem), which we will discuss shortly.

3.3.3 *Tammi's Conjecture*

The following estimates hold for the bounded univalent function scenario:

$$\max_{f\in S_M} |a_2| = 2\left(1 - M^{-1}\right),$$

$$\max_{f\in S_M} |a_3| = \begin{cases} 1 - M^{-2} & \text{if } 1 < M < e, \\ 1 - M^{-2} + 2\left(\lambda - M^{-1}\right)^2 & \text{if } M \ge e, \end{cases}$$

$$\max_{f\in S_M} |a_4| = \begin{cases} \frac{2}{3}\left(1 - M^{-3}\right) & \text{if } 1 < M \le \frac{34}{19}, \\ 2\left(2 - 10M^{-1} + 15M^{-2} - 7M^{-3}\right) & \text{if } M \ge \mu, \end{cases}$$

where the parameter λ is the largest of the two real solutions of $\lambda \ln(\lambda) + M^{-1} = 0$ and the constant μ is the smallest for which the formula holds (to be ascertained). The first estimate dates back to Pick [19]; the second is due to Löwner [20], Schaeffer & Spencer [21] and Janowski [22]; and the third comes from Schiffer & Tammi [23], who computed that $\mu \leq 100/3$. It turns out, for large M, that $\max_{f \in S_M} |a_{2n}|$ is the $(2n)^{\text{th}}$ coefficient in the Maclaurin series expansion of the Pick function $P_M(z)$, for any $n \geq 1$ [24, 25].

Note the sizable gap in the formula for $\max_{f \in S_M} |a_4|$. Tammi [26] determined, when f has real coefficients and $M \geq 11$, that

$$\max_{f \in S_{\mathbb{R}} \cap S_M} |a_4| = 2\left(2 - 10M^{-1} + 15M^{-2} - 7M^{-3}\right).$$

The formula fails for $M < 11$. Hence it was natural for him to conjecture [27] that $\mu = 11$ for f with complex coefficients as well. Prokhorov & Vasil'ev [7] disproved this conjecture, showing that $\mu = 22.9569...$, again using a numerical optimal control-based approach.

3.3.4 *Greiner–Roth Theorem*

Elaborate expressions built from series coefficients can also be optimized. Greiner & Roth [28], starting from [29, 30], proved that the function $f \in S$ maximizing

$$\operatorname{Re}\left(a_3 + \frac{p-3}{3} a_2^2\right) + \frac{p+1}{3}|a_2|^2, \quad p \in \mathbb{R} \text{ fixed},$$

is

$$f(z) = \begin{cases} \pm i\, K(\mp i\, z) & \text{if } p \leq \dfrac{3}{4\ln(2)} - \dfrac{1}{2} = 0.5820212806..., \\[2ex] \pm K(\pm z) & \text{if } p \geq \dfrac{1}{2}\dfrac{2e^3+1}{e^3-1} = 1.0785935447.... \end{cases}$$

In the gap, f cannot be a rotation of the Koebe function. Starting from [31], they also proved that the function $f \in S$ maximizing

$$\operatorname{Re}\left(a_3 - q\, a_2^2\right) + q|a_2|^2, \quad q \in \mathbb{R} \text{ fixed},$$

is

$$f(z) = \begin{cases} \pm K(\pm z) & \text{if } q \leq \dfrac{1}{2} = 0.5, \\[2ex] \pm i\, K(\mp i\, z) & \text{if } q \geq \dfrac{1}{2}\dfrac{e}{e-1} = 0.7909883534.... \end{cases}$$

Again, in the gap, f can be proved not to be a rotation of the Koebe function. Such expressions serve to generalize those used to obtain the Fekete–Szegö constant α

mentioned earlier. Explicit formulas for the gap extremals are not available, but these functions can be found numerically via optimal control.

3.3.5 *Milin's Constant*

Define

$$\delta = \sup_{n \geq 1} \sup_{f \in S} \sum_{k=1}^{n} \left(k|b_k|^2 - \frac{1}{k} \right),$$

then it can be shown [10, 16, 17] that $0.0266 < 2\ln(\alpha) < \delta < 0.3119 < \mathrm{Ei}(\ln(2))/2 - \gamma - \ln(\ln(2))$, where Ei is the exponential integral and γ is Euler's constant. A more precise estimate of Milin's constant δ would be good to see, as well as the corresponding extremal functions. Note that, if $f = \kappa$, then the logarithmic coefficients $b_n = 1/n$ for all n; hence the Koebe function is far from optimal in this setting. It is known that $\max_{f \in S_{\mathrm{odd}}} |a_{2n+1}| < e^{\delta/2}$ (which gave, at one time, the best general estimate 1.17 of the odd coefficents); if it were true that $\delta = 0$, then the Littlewood–Paley conjecture would follow.

By contrast, we have

$$\hat{\delta} = \sup_{n \geq 1} \sup_{f \in S} \sum_{m=1}^{n} \sum_{k=1}^{m} \left(k|b_k|^2 - \frac{1}{k} \right) = 0,$$

which is Milin's conjecture (now proved, as stated earlier). Here, of course, the Koebe function is optimal. For $f \neq \kappa$, the lower order contributions to the sum evidently tend to be negative, forcing $\hat{\delta} < \delta$. See also [32, 33].

3.3.6 *Bieberbach-Eilbenberg Functions*

Define a new set

$$\tilde{S} = \{ f : D \to \mathbb{C} : f \text{ is univalent}, f(0) = 0 \text{ and } f(z)f(w) \neq 1 \text{ for any } z, w \in D \}.$$

Note that nothing is assumed about a_1. In fact,

$$\max_{f \in \tilde{S}} |a_1| = 1,$$

which occurs if and only if $f(z) = e^{i\theta} z$ for some $\theta \in \mathbb{R}$. Nehari [34] and Aharonov [35] proved that

$$\frac{e^{-1/2}}{\sqrt{n}} \leq \max_{f \in \tilde{S}} |a_n| < \frac{e^{-\gamma/2}}{\sqrt{n-1}}$$

for all $n \geq 2$; in particular, $|a_2|$ is less than $e^{-\gamma/2} < 0.74931$. Hummel & Schiffer [36, 37] obtained the estimate

$$\max_{f \in \tilde{S}} |a_2| = \frac{1}{2}\eta = 0.5811002808...,$$

where $\eta = 1.1622005617...$ is the unique real solution of the equation

$$\int_0^1 \left(\frac{1-t}{\eta^2+t^2}\right)^{1/2} dt = \frac{1}{\sqrt{2}} \int_0^{\pi/2} \left[\left(1+\eta^2\sin(\theta)^2\right)^{1/2} - 1\right]^{1/2} d\theta.$$

Another interesting result is the estimate

$$\max_{f\in\tilde{S}} |a_1 a_2| \le \frac{8}{27}\eta^2 = 0.4002104135...$$

and we wonder about higher order coefficients of such functions.

3.3.7 *Krzyz's Conjecture*

Define two new sets

$$U = \{f : D \to \mathbb{C} : f \text{ is analytic, } 0 < |f(z)| < 1\}, \quad V = \{f \in U : f \text{ is univalent}\}.$$

Obviously $U \cap S = \emptyset$ and $0 < |a_0| < 1$ for every $f \in U$. For $n \ge 1$, Krzyz [38] conjectured that

$$\max_{f\in U} |a_n| = \frac{2}{e} = 0.7357588823...,$$

which occurs if and only if $f(z) = e^{(z^n+1)/(z^n-1)}$ or a rotation of this. Note that f is not univalent. Krzyz's conjecture has been proved only for $n \le 5$ [39–50]. A general estimate also applies [51–53]:

$$\max_{f\in U} |a_n| < 0.99918...$$

For univalent functions, Prokhorov & Szynal [37] demonstrated that

$$\max_{f\in V} |a_1| = 12 - 8\sqrt{2} = 0.6862915010...,$$

$$\max_{f\in V} |a_2| = \frac{8\xi(1-\xi)(1-2\xi-\xi^2)}{(1+\xi)^3} = 0.4553841384...,$$

where $\xi = 0.1414780159...$ has minimal polynomial $\xi^4 + 4\xi^3 + 6\xi^2 - 8\xi + 1$.

We close with one more problem. Grinshpan [54, 55], improving upon [56–58], showed that $-2.97 < |a_{n+1}| - |a_n| < 3.61$ for all $f \in S$ and all $n \ge 1$. It is further known that the constants on the left and right cannot be replaced by -1 and 1, respectively, even if we restrict discussion to $f \in S_{\text{odd}}$ [59]. See other related problems in [60, 61].

Addendum Michel [62] claimed that $\max_{f\in S_{\text{odd}}} |a_7|$ is at least 1.006763... and that Milin's constant δ is at least 0.034856.... He further conjectured that δ is equal to this lower bound. A precise expression for the latter is

$$2y^4 e^{-4y} + \left(3y^2 + 2y + 1\right) e^{-2y} - 1 = 0.0348561121...,$$

where $y = 0.3900456802\ldots$ is the unique real solution of the equation

$$4x^2(1 - x)e^{-2x} + (1 - 3x) = 0.$$

No precise expression for the former, analogous to that for $\max_{f \in S_{\mathrm{odd}}} |a_5|$, appears to be known.

[1] L. de Branges, A proof of the Bieberbach conjecture, *Acta Math.* 154 (1985) 137–152; MR0772434 (86h:30026).
[2] A. Z. Grinshpan, The Bieberbach conjecture and Milin's functionals, *Amer. Math. Monthly* 106 (1999) 203–214; MR1682341 (2000b:30027).
[3] L. Bieberbach, Über die Koeffizienten derjenigen Potenzreihen welche eine schlichte Abbildungen des Einheitskreises vermitteln, *Sitzungsber. Preuss. Akad. Wiss.* (1916) 940–955.
[4] E. Bombieri, On the local maximum property of the Koebe function, *Invent. Math.* 4 (1967) 26–67; MR0218549 (36 #1635).
[5] D. Bshouty and W. Hengartner, A variation of the Koebe mapping in a dense subset of S, *Canad. J. Math.* 39 (1987) 54–73; MR0889106 (88j:30048).
[6] R. Greiner and O. Roth, On support points of univalent functions and a disproof of a conjecture of Bombieri, *Proc. Amer. Math. Soc.* 129 (2001) 3657–3664; MR1860500 (2002g:30016).
[7] D. Prokhorov and A. Vasil'ev, Optimal control in Bombieri's and Tammi's conjectures, *Georgian Math. J.* 12 (2005) 743–761; math.CV/0410578; MR2197195 (2006j:30034).
[8] J. E. Littlewood and R. E. A. C. Paley, A proof that an odd schlicht function has bounded coefficients, *J. London Math. Soc.* 7 (1932) 167–169.
[9] M. Fekete and G. Szegö, Eine Bemerkung über ungerade schlichte Funktionen, *J. London Math. Soc.* 8 (1933) 85–89.
[10] P. L. Duren, *Univalent Functions*, Springer-Verlag, 1983, pp. 103–110, 151–156; MR0708494 (85j:30034).
[11] A. C. Schaeffer and D. C. Spencer, The coefficients of schlicht functions, *Duke Math. J*. 10 (1943) 611–635; MR0009631 (5,175i).
[12] G. B. Leeman, The seventh coefficient of odd symmetric univalent functions, *Duke Math. J.* 43 (1976) 301–307; MR0409793 (53 #13545).
[13] H. Ke, Coefficients of odd univalent functions, *Proc. Amer. Math. Soc.* 96 (1986) 183–186; MR0813835 (87a:30029).
[14] V. I. Levin, Some remarks on the coefficients of schlicht functions, *Proc. London Math. Soc.* 39 (1935) 467–480.
[15] K. Sun, Contributions to the theory of schlicht functions. II: The coefficient problem, *Sci. Sinica* 4 (1955) 359–373.
[16] I. M. Milin, The coefficients of schlicht functions (in Russian), *Dokl. Akad. Nauk SSSR* 176 (1967) 1015–1018; Engl. transl in *Soviet Math. Dokl.* 8 (1967) 1255–1258; MR0222276 (36 #5328).
[17] I. M. Milin, *Univalent Functions and Orthonormal Systems* (in Russian), Izdat. Nauka., 1971; Engl. transl. and errata in Amer. Math. Soc. Transl., v. 49, Amer. Math. Soc., 1977; MR0369684 (51 #5916), MR0427620 (55 #651) and MR0450541 (56 #8835).
[18] V. I. Milin, Estimate of the coefficients of odd univalent functions (in Russian), *Metric Questions of the Theory of Functions*, Naukova Dumka, 1980, pp. 78–86, 160; MR0598843 (82i:30025).

[19] G. Pick, Über die konforme Abbildung eines Kreises auf ein schlichtes und zugleich beschränktes Gebiet, *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-Naturwissenschaftliche Klasse. Abteilung* II a 126 (1917) 247–263.

[20] K. Löwner, Untersuchungen über schlichte konforme Abbildungen des Einheitskreises. I, *Math. Annalen* 89 (1923) 103–121.

[21] A. C. Schaeffer and D. C. Spencer, The coefficients of schlicht functions. II, *Duke Math. J.* 12 (1945) 107–125; MR0011721 (6,206c).

[22] W. Janowski, Le maximum des coefficients A_2 et A_3 des fonctions univalentes bornées, *Annales Polonici Math.* 2 (1955) 145–160; MR0074510 (17,598f).

[23] M. Schiffer and O. Tammi, On the fourth coefficient of bounded univalent functions, *Trans. Amer. Math. Soc.* 119 (1965) 67–78; MR0185109 (32 #2579).

[24] D. V. Prokhorov, Even coefficient estimates for bounded univalent functions, *Annales Polonici Math.* 58 (1993) 267–273; MR1244398 (94k:30039).

[25] D. V. Prokhorov, Bounded univalent functions, *Handbook of Complex Analysis*, v. 1, *Geometric Function Theory*, ed. R. Kühnau, North-Holland, 2002, pp. 207–228; MR1966195 (2004a:30017).

[26] O. Tammi, *Extremum Problems for Bounded Univalent Functions. II*, Lect. Notes in Math. 913, Springer-Verlag, 1982, p. 163; MR0653811 (84a:30032).

[27] O. Tammi, *Extremum Problems for Bounded Univalent Functions*, Lect. Notes in Math. 646, Springer-Verlag, 1978, p. 210; MR0492191 (58 #11337).

[28] R. Greiner and O. Roth, A coefficient problem for univalent functions related to two-point distortion theorems, *Rocky Mountain J. Math.* 31 (2001) 261–283; MR1821380 (2002a:30026).

[29] C. Blatter, Ein Verzerrungssatz für schlichte Funktionen, *Comment. Math. Helv.* 53 (1978) 651–659; MR0511855 (80d:30010).

[30] S.-A. Kim and D. Minda, Two-point distortion theorems for univalent functions, *Pacific J. Math.* 163 (1994) 137–157; MR1256180 (94m:30042).

[31] Z. J. Jakubowski and K. Zyskowska, On an estimate of a functional in the class of holomorphic univalent functions, *Math. Bohemica* 118 (1993) 281–296; MR1239123 (94i:30016).

[32] S. M. Zemyan, Estimates of logarithmic coefficients of univalent functions, *Internat. J. Math. Math. Sci.* 16 (1993) 311–318; MR1210840 (94h:30024).

[33] L.-J. Lin, On the logarithmic coefficients of univalent functions, *Math. Japon.* 42 (1995) 165–168; MR1344645 (96i:30013).

[34] Z. Nehari, On the coefficients of Bieberbach-Eilenberg functions, *J. d'Analyse Math.* 23 (1970) 297–303; MR0273003 (42 #7884).

[35] D. Aharonov, On Bieberbach Eilenberg functions, *Bull. Amer. Math. Soc.* 76 (1970) 101–104; MR0257343 (41 #1994).

[36] J. A. Hummel and M. M. Schiffer, Variational methods for Bieberbach-Eilenberg functions and for pairs, *Annales Acad. Sci. Fenn. Ser. A I Math.* 3 (1977) 3–42; MR0584969 (58 #28473).

[37] D. V. Prokhorov and J. Szynal, Coefficient estimates for bounded nonvanishing functions, *Bull. Acad. Polon. Sci. Sér. Sci. Math.* 29 (1981) 223–230; MR0640466 (82m:30024).

[38] J. Krzyz, Coefficient problem for non-vanishing functions, *Annales Polonici Math.* 20 (1968) 314.

[39] V. Levin, W. Fenchel and E. Reissner, Lösing der Aufgabe 163, *Jahresbericht Deutsch. Math.-Verein.*, v. 44 (1934) n. 2, 80–83.

[40] H. S. Shapiro, Bounds for the coefficient a_1 in certain series expansions, *Amer. Math. Monthly* 60 (1953) 131–132.

[41] J. A. Hummel, S. Scheinberg and L. Zalcman, A coefficient problem for bounded nonvanishing functions, *J. d'Analyse Math.* 31 (1977) 169–190; MR0492212 (58 #11358).

[42] D. L. Tan, Estimates of coefficients of bounded nonvanishing analytic functions (in Chinese), *Chinese Annals Math. Ser. A* 4 (1983) 97–104; Engl. summary in *Chinese Annals Math. Ser. B* 4 (1983) 131–132; MR0718010 (85b:30026).

[43] J. E. Brown, Iteration of functions subordinate to schlicht functions, *Complex Var. Theory Appl.* 9 (1987) 143–152; MR0923215 (89a:30019).

[44] R. Peretz, Applications of subordination theory to the class of bounded nonvanishing functions, *Complex Var. Theory Appl.* 17 (1992) 213–222; MR1147052 (93c:30050).

[45] W. Koepf and D. Schmersau, Bounded nonvanishing functions and Bateman functions, *Complex Var. Theory Appl.* 25 (1994) 237–259; MR1314519 (96a:30016).

[46] W. Szapiel, A new approach to the Krzyz conjecture, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 48 (1994) 169–192; MR1346570 (96m:30021).

[47] Z. Lewandowski and J. Szynal, On the Krzyz conjecture and related problems, *XVI*th *Rolf Nevanlinna Colloquium*, Proc. 1995 Joensuu conf., ed. I. Laine and O. Martio, de Gruyter, 1996, pp. 257–268; MR1427091 (97k:30018).

[48] Z. Lewandowski and J. Szynal, On the Krzyz conjecture and related problems. II, *Ann. Univ. Mariae Curie-Sklodowska Sect. A* 52 (1998) 73–82; MR1665531 (2000a:30031).

[49] S. V. Romanova, Estimates for functionals of bounded nonvanishing functions (in Russian), *Sibirsk. Mat. Zh.* 44 (2003) 1113–1123; Engl. transl. in *Siberian Math. J.* 44 (2003) 869–876; MR2019565 (2004j:30025).

[50] N. Samaris, A proof of Krzyz's conjecture for the fifth coefficient, *Complex Var. Theory Appl.* 48 (2003) 753–766; MR2002237 (2004g:30027)

[51] C. Horowitz, Coefficients of nonvanishing functions in H^{∞}, *Israel J. Math.* 30 (1978) 285–291; MR0508270 (80a:30035).

[52] R. Ermers, *Coefficient Estimates for Bounded Nonvanishing Functions*, Ph.D. thesis, Katholieke Universiteit Nijmegen, 1990.

[53] D. V. Prokhorov, Coefficients of holomorphic functions (in Russian), *Itogi Nauki Tekh. Ser. Sovrem. Mat. Prilozh. Temat. Obz.*, v. 71, *Complex Analysis and Representation Theory*-2, 2000; Engl. transl. in *J. Math. Sci.* 106 (2001) 3518–3544; MR1871193 (2003a:30021).

[54] A. Z. Grinshpan, The sharpening of the difference of the moduli of adjacent coefficients of schlicht functions (in Russian), *Some Problems in Modern Function Theory*, Proc. 1976 Novosibirsk conf., Akad. Nauk SSSR Sibirsk. Otdel. Inst. Mat., 1976, pp. 41–45; MR0507758 (58 #22520).

[55] W. K. Hayman, *Multivalent Functions*, 2nd ed., Cambridge Univ. Press, 1994, pp. 165–196; MR1310776 (96f:30003).

[56] I. M. Milin, Adjacent coefficients of univalent functions (in Russian), *Dokl. Akad. Nauk SSSR* 180 (1968) 1294–1297; Engl. transl. in *Soviet Math. Dokl.* 9 (1968) 762–765; MR0231992 (38 #318).

[57] L. P. Il'ina, On the relative growth of nearby coefficients of schlicht functions (in Russian), *Mat. Zametki* 4 (1968) 715–722, Engl. transl. in *Math. Notes* 4 (1968) 918–922; MR0240295 (39 #1644).

[58] M. Alves, Successive coefficients of univalent functions, *Bol. Soc. Brasil. Mat.* 12 (1981) 1–10; MR0688183 (84i:30019).

[59] P. L. Duren, Successive coefficients of univalent functions, *J. London Math. Soc.* 19 (1979) 448–450; MR0540059 (80i:30030).

[60] S. R. Finch, Littlewood-Clunie-Pommerenke constants, *first volume*, pp. 470–473.
[61] S. R. Finch, Radii in geometric function theory, *this volume*, §3.1.
[62] E. Michel, A Löwner variational method in the theory of schlicht functions, arXiv:1112.5187.

3.4 Planar Harmonic Mappings

Let D denote the open unit disk. A function $f: D \to \mathbb{C}$ is **planar harmonic** if it can be written as $f(z) = u(x, y) + i\, v(x, y)$ where $z = x + i\, y$ and where $u: D \to \mathbb{R}$, $v: D \to \mathbb{R}$ are harmonic, that is, are twice continuously differentiable and obey

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

It can be shown that f is planar harmonic if and only if $f = g + \bar{h}$, where g, h are analytic on D and the overbar indicates complex conjugation ($\bar{z} = x - i\, y$).

Of course, a planar harmonic function f is analytic if and only if u and v are harmonic conjugates, that is, the Cauchy–Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

are satisfied. We are interested, in this essay, in functions f whose real and imaginary parts are not necessarily conjugate [1].

It turns out that f may be written as a twice continuously differentiable function of z and $\bar{z}$; we abuse notation and use the same letter f to represent the new function. The Cauchy–Riemann equations become a single concise equation:

$$\frac{\partial f}{\partial \bar{z}} = 0$$

and the condition that Laplacians vanish becomes

$$4\frac{\partial^2 f}{\partial z\, \partial \bar{z}} = 0.$$

Thus the expression f is independent of $\bar{z}$ for analytic functions f, and the expression $\partial f / \partial z$ is independent of $\bar{z}$ for planar harmonic functions f.

A planar harmonic function $f: D \to \mathbb{C}$ is a **mapping** if it is one-to-one. Hence the class of planar harmonic mappings includes the subclass of univalent functions we have studied elsewhere [2–5]. Define also the **dilatation** of f

$$\omega = \overline{\frac{\partial f}{\partial \bar{z}}} \Big/ \frac{\partial f}{\partial z},$$

which will be needed later.

3.4.1 *Heinz's Inequality*

We consider here planar harmonic mappings f that map D onto D, with the property that $f(0)=0$. Heinz [6] proved that

$$\left|\frac{\partial f}{\partial z}(0,0)\right|^2+\left|\frac{\partial f}{\partial \bar{z}}(0,0)\right|^2\geq c$$

for some constant $c\geq 0.1788=0.3576/2$. The lower bound was improved to $0.32=0.64/2$ by Nitsche [7, 8], $0.4345=0.8691/2$ by de Vries [9], $0.4476=0.8952/2$ by Nitsche [10], $0.6411=1.2822/2$ by de Vries [11], and $0.6584=1.3168/2$ by Wegmann [12]. The conjecture that

$$c=\frac{27}{4\pi^2}=0.6839179895...=\frac{1}{2}(1.3678359791...),$$

mentioned by Wegmann [12], seems to have been anticipated by Hopf [13]. A proof of this conjecture was first given by Hall [1, 14]; the extremal function is achieved via approximations $D\to D$ of a mapping $D\to T$, where T is an inscribed equilateral triangle, with dilatation $\omega(z)=z$.

Hall's proof involves the Fourier coeffients of homeomorphisms $C\to C$ of the unit circle C. Some related problems are given in [14]; one of these has been solved [15]. Heinz [16] also proved the inequality

$$\left|\frac{\partial f}{\partial z}(z,\bar{z})\right|^2+\left|\frac{\partial f}{\partial \bar{z}}(z,\bar{z})\right|^2\geq\frac{1}{\pi^2},$$

which is valid for all $z\in D$; improvements in special cases appear in [17, 18].

3.4.2 *Minimal Surfaces*

Consider a minimal surface over the unit disk D of the form

$$\{(x,y,z)\in\mathbb{R}^3:z=F(x,y),\ (x,y)\in D\}$$

and let κ denote its Gaussian curvature at the origin. In words, the surface is locally area-minimizing: Each suitable small piece of it has the least possible area for any surface spanning the boundary of that piece. By the calculus of variations, we have the nonlinear PDE

$$\left[1+\left(\frac{\partial F}{\partial y}\right)^2\right]\frac{\partial^2 F}{\partial x^2}-2\frac{\partial F}{\partial x}\frac{\partial F}{\partial y}\frac{\partial^2 F}{\partial x\,\partial y}+\left[1+\left(\frac{\partial F}{\partial x}\right)^2\right]\frac{\partial^2 F}{\partial y^2}=0;$$

hence the mean curvature of the surface is everywhere zero. A precise determination of F is difficult – this is called **Plateau's problem** – but nature solves it effortlessly, as can be demonstrated by dipping a bent wire loop in a soap solution [19, 20]. We will revisit this topic in greater detail [21]; see especially the "Matlab help" example near the end.

A consequence of Heinz's inequality [6] is that

$$|\kappa| \leq \frac{4}{c} = \frac{16\pi^2}{27} = 5.8486544599...$$

by Hall's theorem [14], but this is not sharp. In fact, it is conjectured that [1]

$$|\kappa| \leq \frac{\pi^2}{2} = 4.9348022005;$$

this has however been proved only in the special case that the minimal surface has a horizontal tangent plane at the origin [22]. A general proof could be obtained utilizing the following.

Consider planar harmonic mappings f that map D onto D, with the two properties that $f(0) = 0$ and ω is the square of an analytic function. (Note that this final requirement is not met by $\omega(z) = z$.) Hall [23] computed that

$$\left|\frac{\partial f}{\partial z}(0,0)\right|^2 + \left|\frac{\partial f}{\partial \bar{z}}(0,0)\right|^2 \geq \tilde{c}$$

for some constant $\tilde{c} > c + 10^{-5}/2$. It is conjectured that $\tilde{c} = 8/\pi^2$ (from which $4/\tilde{c} = \pi^2/2$ would proceed immediately). The expected extremal function is a mapping $D \to S$, where S is an inscribed square, with dilatation $\omega(z) = z^2$. A proof that $\tilde{c} = 8/\pi^2$ would be a major step forward in understanding minimal surfaces. See [24] for more open questions.

3.4.3 *Soap Films*

As an aside, we give an elementary problem [25, 26]. Consider the catenoid-shaped soap film formed between two parallel rings centered at $(-\xi, 0, 0)$ and $(\xi, 0, 0)$ and of unit radius, where $\xi > 0$ is suitably small. If the rings are slowly pulled apart (that is, if ξ increases), there is a certain threshold at which the minimal surface becomes unstable and is likely to collapse to a disjoint union of two disks. More precisely, if $\xi < \xi_0 = 0.5276973969...$, then the catenoid corresponds to the global minimum for surface area while the two-disk configuration corresponds to only a local minimum. Here ξ_0 and $a = 0.8255174536...$ are solutions of the simultaneous equations

$$\begin{cases} a\cosh\left(\dfrac{\xi_0}{a}\right) = 1, \\[2ex] 2\pi a^2 \sinh\left(\dfrac{\xi_0}{a}\right)\cosh\left(\dfrac{\xi_0}{a}\right) + 2\pi a\,\xi_0 = 2\pi. \end{cases}$$

If $\xi > \xi_0$, then the two-disk configuration corresponds to the global minimum while the catenoid corresponds to only a local minimum for $\xi < \xi_1 = 0.6627434193...$; no such catenoid exists for $\xi > \xi_1$. Here ξ_1 and $b = 0.5524341245...$ are solutions of the simultaneous equations

$$\begin{cases} b\cosh\left(\dfrac{\xi_1}{b}\right) = 1, \\ \cosh\left(\dfrac{\xi_1}{b}\right) - \dfrac{\xi_1}{b}\sinh\left(\dfrac{\xi_1}{b}\right) = 0. \end{cases}$$

Interestingly, we have seen the value for ξ_1 before: In [27], it arose in a different context altogether and was called the *Laplace limit constant*.

[1] P. Duren, *Harmonic Mappings in the Plane*, Cambridge Univ. Press, 2004, pp. 1–7, 18–23, 66–72, 182–186; MR2048384 (2005d:31001).

[2] S. R. Finch, Bloch-Landau constants, *first volume*, pp. 456–459.

[3] S. R. Finch, Littlewood-Clunie-Pommerenke constants, *first volume*, pp. 470–473.

[4] S. R. Finch, Radii in geometric function theory, *this volume*, §3.1.

[5] S. R. Finch, Coefficient estimates for univalent functions, *this volume*, §3.3.

[6] E. Heinz, Über die Lösungen der Minimalflächengleichung, *Nachr. Akad. Wiss. Göttingen. Math.-Phys. Kl. Math.-Phys.-Chem. Abt.* (1952) 51–56; MR0054182 (14,885c).

[7] J. C. C. Nitsche, On harmonic mappings, *Proc. Amer. Math. Soc.* 9 (1958) 268–271; MR0104233 (21 #2990a).

[8] J. C. C. Nitsche, On the constant of E. Heinz, *Rend. Circ. Mat. Palermo* 8 (1959) 178–181; MR0113054 (22 #3895).

[9] H. L. de Vries, A remark concerning a lemma of Heinz on harmonic mappings, *J. Math. Mech.* 11 (1962) 469–471; MR0142769 (26 #338).

[10] J. C. C. Nitsche, Zum Heinzschen Lemma über harmonische Abbildungen, *Arch. Math.* 14 (1963) 407–410; MR0160919 (28 #4128).

[11] H. L. de Vries, Über Koeffizientenprobleme bei Eilinien und über die Heinzsche Konstante, *Math. Z.* 112 (1969) 101–106; MR0254742 (40 #7949).

[12] R. Wegmann, Extremalfiguren für eine Klasse von isoperimetrischen Problemen, *Math. Nachr.* 69 (1975) 173–190; MR0428198 (55 #1223).

[13] E. Hopf, On an inequality for minimal surfaces $z = z(x, y)$, *J. Rational Mech. Anal.* 2 (1953) 519–522; correction 2 (1953) 801–802; MR0055735 (14,1119d) and MR0058268 (15,348a).

[14] R. R. Hall, On an inequality of E. Heinz, *J. d'Analyse Math.* 42 (1982/83) 185–198; MR0729409 (85k:42011).

[15] A. Weitsman, On the Fourier coefficients of homeomorphisms of the circle, *Math. Res. Lett.* 5 (1998) 383–390; MR1637852 (99h:31003).

[16] E. Heinz, On one-to-one harmonic mappings, *Pacific J. Math.* 9 (1959) 101–105; MR0104933 (21 #3683).

[17] D. Partyka and K. Sakan, On Heinz's inequality, *Bull. Soc. Sci. Lett. Lodz Sér. Rech. Déform.* 37 (2002) 27–34; MR2055781 (2005e:31003).

[18] D. Partyka and K. Sakan, On an asymptotically sharp variant of Heinz's inequality, *Annales Acad. Sci. Fenn. Math.* 30 (2005) 167–182; MR2140305 (2006g:30023).

[19] J. C. C. Nitsche, *Vorlesungen über Minimalflächen*, Springer-Verlag, 1975; revised Engl. transl., *Lectures on Minimal Surfaces*, v. I, Cambridge Univ. Press, 1989; MR0448224 (56 #6533) and MR1015936 (90m:49031).
[20] U. Dierkes, S. Hildebrandt, A. Küster and O. Wohlrab, *Minimal Surfaces*, v. I and II, Springer-Verlag, 1992; MR1215267 (94c:49001a) and MR1215268 (94c:49001b).
[21] S. R. Finch, Soap film experiments, *this volume*, §5.19.
[22] R. Finn and R. Osserman, On the Gauss curvature of non-parametric minimal surfaces, *J. d'Analyse Math.* 12 (1964) 351–364; MR0166694 (29 #3967).
[23] R. R. Hall, The Gaussian curvature of minimal surfaces and Heinz' constant, *J. Reine Angew. Math.* 502 (1998) 19–28; MR1647626 (99j:53009).
[24] J. Clunie and T. Sheil-Small, Harmonic univalent functions, *Annales Acad. Sci. Fenn. Ser. A I Math.* 9 (1984) 3–25; MR0752388 (85i:30014).
[25] G. A. Bliss, *Calculus of Variations*, Math. Assoc. Amer., 1925, pp. 5–8, 85–127.
[26] J. Oprea, *The Mathematics of Soap Films: Explorations with Maple®*, Amer. Math. Soc., 2000, pp. 11–12, 71–72, 181–191; MR1779498 (2002c:53013).
[27] S. R. Finch, Laplace limit constant, *first volume*, pp. 266–268.

3.5 Constant of Interpolation

A bounded entire function is necessarily constant (by Liouville's theorem). For our purposes, let us therefore restrict attention to function f analytic on the upper half plane $\mathrm{Im}(z) > 0$. Define the H^∞-norm of f to be

$$||f||_\infty = \sup_{y>0} |f(x+iy)|.$$

Also, given a finite or infinite sequences $W = \{w_j\}$ of complex numbers, define its l^∞-norm by

$$||W||_\infty = \sup_{j\geq 1} |w_j|.$$

We say that a sequence $Z = \{z_j\}$ of distinct complex numbers in the upper half plane is an **interpolating sequence** if there exists an analytic function f for which $||f||_\infty < \infty$ and

$$f(z_j) = w_j, \quad j = 1, 2, 3, ...$$

for each sequence W with $||W||_\infty < \infty$. In words, Z has the property that, for any bounded W, there must be a bounded analytic interpolant f taking z_j to w_j for all j. There may be many such f. We wish to be as efficient as possible and define $M(Z)$ to be the smallest constant C such that

$$||f||_\infty \leq C \cdot ||W||_\infty$$

always; if Z is not an interpolating sequence, define instead $M(Z) = \infty$. Carleson [1–4] proved that $M(Z) < \infty$ if and only if a uniform separation criterion

$$\delta = \inf_{k\geq 1} \prod_{j\neq k} \left| \frac{z_j - z_k}{z_j - \bar{z}_k} \right| > 0$$

is met.

Define the **Blaschke product** corresponding to Z by [4]

$$B(z) = \prod_{n \geq 1} \frac{|z_n^2 + 1|}{z_n^2 + 1} \frac{z - z_n}{z - \bar{z}_n}$$

with the understanding that, if $z = i$ (the imaginary unit), then the left-hand factor is to be interpreted as 1. If Z is an interpolating sequence, then B is uniformly convergent on compact subsets of the upper half plane and hence represents an analytic function. Further, $||B||_\infty = 1$ and B vanishes only at the points z_n. Let

$$B_k(z) = \frac{z - \bar{z}_k}{z - z_k} B(z)$$

so that we may write $\delta = \inf_{k \geq 1} |B_k(z_k)|$. Also let $z_j = x_j + i\, y_j$.

Beurling [5], Jones [6] and Havin [7] examined the problem of exhibiting an explicit formula for f. Nicolau, Ortega-Cerdà & Seip [8] used this work as a basis for estimating $M(Z)$. Define

$$\Phi(Z) = \sup_{k \geq 1} \sum_{y_j \leq y_k} \frac{4y_j(y_j + y_k)}{|z_j - \bar{z}_k|^2} \frac{1}{|B_j(z_j)|},$$

$$\Psi(Z) = \sup_{k \geq 1} \sum_{n \geq 1} \frac{4y_k y_n}{|z_k - \bar{z}_n|^2} \frac{1}{|B_n(z_n)|}.$$

Then, for every interpolating sequence Z in the upper half plane, we have

$$\frac{1}{2} \leq \frac{M(Z)}{\Phi(Z)} \leq \kappa, \quad 1 \leq \frac{M(Z)}{\Psi(Z)} \leq \lambda$$

for constants κ and λ satisfying

$$2.2661... = \frac{\pi}{2\ln(2)} \leq \kappa \leq e = 2.7182...,$$

$$1.5707... = \frac{\pi}{2} \leq \lambda \leq 2e = 5.4365....$$

Can these bounds be improved? Also, can simpler expressions than Φ or Ψ for the denominators be found?

An alternative definition of $M(Z)$ is related to Nevanlinna–Pick theory [4, 9, 10]. Let $M_n(Z)$ be the smallest constant C_n such that the matrix $A = (a_{j,k})$ with

$$a_{j,k} = \frac{1 - \bar{w}_j w_k}{z_j - \bar{z}_k}, \quad j = 1, 2, ..., n, \quad k = 1, 2, ..., n,$$

is nonnegative definite whenever $||W||_\infty < 1/C_n$. The constant of interpolation $M(Z)$ is thus $M_n(Z)$ if Z consists of exactly n points and $\lim_{n \to \infty} M_n(Z)$ if Z is infinite [8].

We could alternatively restrict attention to functions f analytic on the unit disk $|z| < 1$. Some relevant formulas in this new setting are

$$\delta = \inf_{k \geq 1} \prod_{j \neq k} \left| \frac{z_j - z_k}{\bar{z}_j z_k - 1} \right|,$$

$$B(z) = \prod_{n \geq 1} \frac{|z_n|}{z_n} \frac{z - z_n}{\bar{z}_n z - 1},$$

$$a_{j,k} = \frac{1 - \bar{w}_j w_k}{1 - \bar{z}_j z_k}, \quad j = 1, 2, ..., n, \quad k = 1, 2, ..., n.$$

Similar interpolation questions can be asked for the H^p-norm on the unit disk (for example):

$$||f||_p = \sup_{0<r<1} \left(\frac{1}{2\pi} \int_0^{2\pi} \left| f(r\, e^{i\,\theta}) \right|^p d\theta \right)^{1/p}$$

where $1 < p < \infty$ [4, 11]. It would be good to see results paralleling those in [8] for $p = 2$ and $p = 1$.

[1] L. Carleson, An interpolation problem for bounded analytic functions, *Amer. J. Math.* 80 (1958) 921–930; MR0117349 (22 #8129).
[2] J. P. Earl, On the interpolation of bounded sequences by bounded functions, *J. London Math. Soc.* 2 (1970) 544–548; MR0284588 (44 #1813).
[3] J. P. Earl, A note on bounded interpolation in the unit disc, *J. London Math. Soc.* 13 (1976) 419–423; MR0419773 (54 #7791).
[4] J. B. Garnett, *Bounded Analytic Functions*, Academic Press, 1981, pp. 1–10, 50–57, 284–317; MR0628971 (83g:30037).
[5] L. Carleson, Interpolations by bounded analytic functions and the Corona problem, *International Congress of Mathematicians*, Proc. 1962 Stockholm conf., Inst. Mittag-Leffler, 1963, pp. 314–316; MR0176274 (31 #549).
[6] P. W. Jones, L^∞ estimates for the $\bar{\partial}$ problem in a half-plane, *Acta Math.* 150 (1983) 137–152; MR0697611 (84g:35135).
[7] V. P. Havin, Jones' interpolation formula, in P. Koosis, *Introduction to H_p Spaces*, 2nd ed., Cambridge Univ. Press, 1998, pp. 263–270; MR1669574 (2000b: 30052)
[8] A. Nicolau, J. Ortega-Cerdà and K. Seip, The constant of interpolation, *Pacific J. Math.* 213 (2004) 389–398; math.CV/0301334; MR2036925 (2004m:30061).
[9] J. E. McCarthy, Pick's theorem – what's the big deal? *Amer. Math. Monthly* 110 (2003) 36–45; MR1952746 (2003k:30055).
[10] R. A. Kortram, A new characterization of the unit ball of H^2, *Proc. Amer. Math. Soc.* 132 (2004) 127–133; MR2021255 (2004i:30031).
[11] H. S. Shapiro and A. L. Shields, On some interpolation problems for analytic functions, *Amer. J. Math.* 83 (1961) 513–532; MR0133446 (24 #A3280).

3.6 Dirichlet Integral

Consider the class of complex analytic functions f on the open unit disk Δ with $f(0) = 0$ and finite Dirichlet integral:

$$D(f) = \frac{1}{\pi} \int\limits_{\Delta} |f'(z)|^2 dx\, dy < \infty.$$

Clearly $\pi\, D(f)$ is the area of the region $f(\Delta)$ in $\mathbb{C}$, counting multiplicities [1].

Chang & Marshall [2–4] proved that there exists a constant $C > 0$ such that $D(f) \leq 1$ implies

$$\frac{1}{2\pi} \int\limits_{0}^{2\pi} \exp\left(|f(e^{i\theta})|^2\right) d\theta \leq C.$$

Andreev & Matheson [5–7] conjectured that the best constant C is $e = 2.7182818284...$, corresponding to the identity function $f(z) = z$. The mere existence of an extremal function, however, remains open [8]. Interestingly, extremal functions provably exist for the closely-related Trudinger–Moser inequality [9].

In the following, we distinguish the unit disk Δ in z-space from the unit disk in w-space (where $w = f(z)$) by writing $\tilde{\Delta}$ for the latter. Define, for $s > 0$,

$$\Omega(s) = \{z \in \Delta : |f(z)| < s\}$$

and let

$$A(s) = \int\limits_{\Omega(s)} |f'(z)|^2 dx\, dy.$$

Obviously $\Omega(\infty) = \Delta$ and $A(\infty) = \pi\, D(f)$. Marshall [3] asked whether there exists a constant $r > 0$ such that, for any $s > 0$, $A(s) \leq \pi\, s^2$ implies $f(r\,\Delta) \subseteq s\,\tilde{\Delta}$. In words, the constant r is so small that, for any radius s, if

$$\begin{pmatrix} \text{the area of the portion} \\ \text{of } f(\Delta) \text{ lying within } s\,\tilde{\Delta} \end{pmatrix} \text{ is strictly less than } \left(\text{the area of } s\,\tilde{\Delta}\right),$$

then f must map $r\,\Delta$ into $s\,\tilde{\Delta}$ itself.

Poggi-Corradini [10] demonstrated that r exists. Solynin [11] further proved that the best constant r is at least $r_0 = 0.03949....$ In fact, r_0 is best possible for the larger class of analytic functions f that omit two values of a doubly-sheeted Riemann surface corresponding to $z \mapsto \sqrt{z}$. It is given exactly by

$$r_0 = \frac{L\left(\sqrt{\sqrt{2}-1}\right) - K\left(\sqrt{\sqrt{2}-1}\right)}{L\left(\sqrt{\sqrt{2}-1}\right) + K\left(\sqrt{\sqrt{2}-1}\right)} = 0.0394929227...,$$

where $K(x)$ denotes the complete elliptic integral of the first kind [12] and $L(x) = K\left(\sqrt{1-x^2}\right)$. Unfortunately r_0 is not sharp for Marshall's original class of analytic functions: identifying r here remains open, as is the problem of describing extremal functions.

Marshall [3] pointed out that, if f is univalent, then the associated best value of r is at least $1/16 = 0.0625$. Solynin [11] indicated that the sharp r here is exactly $3 - 2\sqrt{2} = 0.1715728752...$, corresponding to rotations of the Koebe function $f(z) = z/(1-z)^2$.

[1] W. T. Ross, The classical Dirichlet space, *Recent Advances in Operator-Related Function Theory*, Proc. 2004 Dublin conf., ed. A. L. Matheson, M. I. Stessin and R. M. Timoney, Amer. Math. Soc., 2006, pp. 171–197; MR2198379 (2006k:31007).

[2] S.-Y. A. Chang and D. E. Marshall, On a sharp inequality concerning the Dirichlet integral, *Amer. J. Math.* 107 (1985) 1015–1033; MR0805803 (87a:30055).

[3] D. E. Marshall, A new proof of a sharp inequality concerning the Dirichlet integral, *Ark. Mat.* 27 (1989) 131–137; MR1004727 (90h:30097).

[4] M. Essén, Sharp estimates of uniform harmonic majorants in the plane, *Ark. Mat.* 25 (1987) 15–28; MR0918377 (89b:30024).

[5] V. V. Andreev and A. Matheson, Extremal functions and the Chang-Marshall inequality, *Pacific J. Math.* 162 (1994) 233–246; MR1251899 (95f:30051).

[6] J. Cima and A. Matheson, A nonlinear functional on the Dirichlet space, *J. Math. Anal. Appl.* 191 (1995) 380–401; MR1324020 (96g:46015).

[7] V. V. Andreev, An extremal function for the Chang-Marshall inequality over the Beurling functions, *Proc. Amer. Math. Soc.* 133 (2005) 2069–2076; MR2137873 (2006g:30084).

[8] A. Matheson and A. R. Pruss, Properties of extremal functions for some nonlinear functionals on Dirichlet spaces, *Trans. Amer. Math. Soc.* 348 (1996) 2901–2930; MR1357401 (96j:30003).

[9] S. R. Finch, Nash's inequality, *this volume*, §2.3.

[10] P. Poggi-Corradini, Mapping properties of analytic functions on the disk, *Proc. Amer. Math. Soc.* 135 (2007) 2893–2898; math.CV/0601080; MR2317966 (2008d:30035).

[11] A. Y. Solynin, Mapping properties of analytic functions on the unit disk, *Proc. Amer. Math. Soc.* 136 (2008) 577–585; MR2358498 (2009f:30050).

[12] S. R. Finch, Random triangles, *this volume*, §5.10.

3.7 Brachistochrone Problem

Think of a continuously differentiable curve as a frictionless wire in a vertical plane, with positive x-axis extending to the right and positive y-axis extending downward. Of all curves $y(x)$ joining the origin and a fixed point (p, q) in the first quadrant, which possesses the minimum descent time

$$T = \frac{1}{\sqrt{2g}} \int_0^p \sqrt{\frac{1 + y'(x)^2}{y(x)}}\, dx$$

from $(0,0)$ to (p,q)? In words, $y(x)$ is the wire configuration along which a bead will slide, starting from rest, in the shortest possible time T. For simplicity, we take the gravitational acceleration constant g to be $1/2$, so that the coefficient of the integral defining T is 1.

It is well-known that this calculus-of-variations problem reduces to solving the boundary value problem [1–7]

$$y(x)\left(1+y'(x)^2\right)=c, \quad y(0)=0, \quad y(p)=q,$$

where c is an arbitrary constant, and that $y(x)$ is represented parametrically by

$$x=\frac{c}{2}(t-\sin(t)), \quad y=\frac{c}{2}(1-\cos(t)).$$

Let $0<\theta<2\pi$ be the unique value satisfying

$$\frac{\theta-\sin(\theta)}{1-\cos(\theta)}=\frac{p}{q},$$

then

$$T=2\sqrt{q}\frac{\theta/2}{\sin(\theta/2)}.$$

For example, if $p/q=1$, then $T/\sqrt{q}=2.5819045128...$ and if $p/q=\pi/2$, then $T/\sqrt{q}=\pi$. While the latter result is simple, no closed-form expression is known for the former [8, 9].

Interestingly, we have $y'(p)>0$ when $p/q=1$, whereas $y'(p)=0$ when $p/q=\pi/2$. The sliding bead reaches the endpoint with *zero slope* in the latter case.

With this in mind, we introduce a revision of the brachistochrone problem. Let the starting point be $(0,b)$ where $b\geq 0$ is fixed and let the initial speed of the bead along the wire be $\sqrt{2gb}$. Let the endpoint be (p,q), where $p>0$ is fixed but $q>b$ is free to vary, subject to the constraint that the trajectory slope is zero at (p,q). Of all curves $y(x)$ joining $(0,b)$ and (p,q) satisfying these conditions, which possesses the minimum descent time? [10, 11]

In this revised setting, the boundary value problem is

$$y(x)\left(1+y'(x)^2\right)=c, \quad y(0)=b, \quad y'(p)=0$$

and the solution $y(x)$ is represented parametrically by

$$x=-\frac{c}{2}(t+\sin(t))+p, \quad y=\frac{c}{2}(1+\cos(t)).$$

Clearly $q=c$ upon setting $t=0$. When setting $x=0$ instead, we obtain

$$p=\frac{c}{2}(t+\sin(t)), \quad b=\frac{c}{2}(1+\cos(t)),$$

hence

$$t=\arccos\left(\frac{2}{c}b-1\right)$$

hence

$$\sqrt{(c-b)b}+\frac{c}{2}\arccos\left(\frac{2}{c}b-1\right)=p$$

hence

$$\sqrt{\left(\frac{q}{p}-\frac{b}{p}\right)\frac{b}{p}}+\frac{1}{2}\frac{q}{p}\arccos\left(2\frac{b/p}{q/p}-1\right)=1.$$

Our interest is in the value of q/p, given $b/p=0$, 1, 2 or 3. If $b/p=0$, it follows that $q/p=2/\pi=0.6366197723...$, consistent with before. If $b/p=1$, 2 or 3, then

$$q/p=1.2184055294..., \quad 2.1201938103..., \quad 3.0818460494...,$$

respectively. The latter value appears in [12, 13], obtained via completely different means. Other revisions of the brachistochrone problem can be found in [14, 15].

[1] G. F. Simmons, *Differential Equations with Applications and Historical Notes*, McGraw-Hill, 1972, pp. 25–31, 353–363; MR0499365 (58 #17258).
[2] R. Weinstock, *Calculus of Variations with Applications to Physics and Engineering*, Dover, 1974, pp. 19, 28–29, 38, 70–71; MR0443487 (56 #1856).
[3] W. Hrusa and J. L. Troutman, Elementary characterization of classical minima, *Amer. Math. Monthly* 88 (1981) 321–327; MR0611388 (82f:49001).
[4] D. C. Benson, An elementary solution of the brachistochrone problem, *Amer. Math. Monthly* 76 (1969) 890–894; MR0251608 (40 #4835).
[5] E. Balder, The brachistochrone problem made elementary, unpublished note (2002).
[6] G. Brookfield, Yet another elementary solution of the brachistochrone problem, *Math. Magazine* 83 (2010) 59–63.
[7] R. Coleman, A detailed analysis of the brachistochrone problem, arXiv:1001.2181.
[8] D. S. Shafer, The brachistochrone: Historical gateway to the calculus of variations, *Materials Matemàtics* (2007).
[9] M. Desaix, D. Anderson and M. Lisak, The brachistochrone problem – an introduction to variational calculus for undergraduate students, *Europ. J. Phys.* 26 (2005) 857–864.
[10] T. M. Atanackovic, The brachistrochrone for a material point with arbitrary initial velocity, *Amer. J. Phys.* 46 (1978) 1274–1275.
[11] G. Lawlor, A new minimization proof for the brachistochrone, *Amer. Math. Monthly* 103 (1996) 242–249; MR1376179.
[12] T. Y. Na, *Computational Methods in Engineering Boundary Value Problems*, Academic Press, 1979, pp. 149–152; MR0561710 (82g:65001).
[13] R. Fazio, A noniterative transformation method applied to two-point boundary-value problems, *Appl. Math. Comput.* 39 (1990) 79–87; MR1067571 (91i:34019).
[14] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*, rev. ed., Hemisphere, 1975, pp. 119–121; MR0446628 (56 #4953).
[15] F. Y. M. Wan, *Introduction to the Calculus of Variations and its Applications*, 2nd ed., Chapman & Hall, 1995, pp. 4, 22–26, 323–326, 345–350; MR1369576 (97a:49001).

3.8 Unconditional Basis Constants

Define the **Haar functions** $h_n : [0,1) \to \mathbb{R}$ by

$$h_1(t) = 1; \quad h_2(t) = \begin{cases} 1 & 0 \le t < 1/2, \\ -1 & 1/2 \le t < 1; \end{cases}$$

$$h_3(t) = \begin{cases} 1 & 0 \le t < 1/4, \\ -1 & 1/4 \le t < 1/2, \\ 0 & 1/2 \le t < 1; \end{cases} \quad h_4(t) = \begin{cases} 0 & 0 \le t < 1/2, \\ 1 & 1/2 \le t < 3/4, \\ -1 & 3/4 \le t < 1; \end{cases}$$

$$h_5(t) = \begin{cases} 1 & 0 \le t < 1/8, \\ -1 & 1/8 \le t < 1/4, \\ 0 & 1/4 \le t < 1; \end{cases} \quad h_6(t) = \begin{cases} 0 & 0 \le t < 1/4, \\ 1 & 1/4 \le t < 3/8, \\ -1 & 3/8 \le t < 1/2, \\ 0 & 1/2 \le t < 1; \end{cases}$$

$$h_7(t) = \begin{cases} 0 & 0 \le t < 1/2, \\ 1 & 1/2 \le t < 5/8, \\ -1 & 5/8 \le t < 3/4, \\ 0 & 3/4 \le t < 1; \end{cases} \quad h_8(t) = \begin{cases} 0 & 0 \le t < 3/4, \\ 1 & 3/4 \le t < 7/8, \\ -1 & 7/8 \le t < 1 \end{cases}$$

and so on. Schauder [1–3] proved that $\{h_n\}_{n\ge 1}$ form a basis of the classical Banach space $L_p[0,1]$, $1 \le p < \infty$, that is, for every function $f \in L_p[0,1]$, there exists a unique sequence $\{a_n\}_{n\ge 1}$ of real numbers satisfying

$$\lim_{n\to\infty} \left\| f - \sum_{k=1}^{n} a_k h_k \right\|_p = 0.$$

Let $1 < p < \infty$ and $1/p + 1/q = 1$. Define a **sign sequence** to consist entirely of elements in $\{+1, -1\}$ and a **bit sequence** to consist entirely of elements in $\{0, 1\}$. Work by Paley [4], Marinkiewicz [5] and Burkholder [6, 7] leads to

$$\left\| \sum_{k=1}^{\infty} \varepsilon_k a_k h_k \right\|_p \le (p^* - 1) \left\| \sum_{k=1}^{\infty} a_k h_k \right\|_p$$

for any real sequence $\{a_k\}$ and any sign sequence $\{\varepsilon_k\}$, where $p^* = \max\{p, q\}$ and the constant

$$p^* - 1 = \begin{cases} 1/(p-1) & \text{if } 1 < p \le 2, \\ p - 1 & \text{if } 2 < p < \infty \end{cases}$$

is best possible. Work by Choi [8, 9] leads to a similar inequality corresponding to bit sequences $\{\varepsilon_k\}$, but the best constant c_p here is more complicated. The **unconditional basis constant** c_p captures extreme behavior of the L_p-norm of a series $\sum a_k h_k$ when we discard some of the terms. It is known that $c_2 = 1$. By duality, $c_p = c_q$ and hence it suffices to determine c_p for $2 < p < \infty$.

Let $p_0 = 2.5455457214...$ be the unique solution of the equation

$$p - 2 = \left[\frac{(p-1)(p-2)}{-p^2 + 5p - 5}\right]^{p-1}, \quad 2 < p < 3.$$

For $-1 \le t \le 1$, define functions

$$E(t) = \begin{cases} t^{p-1} - (p-1)t + p - 2 & \text{if } t \ge 0, \\ -(-t)^{p-1} - (p-1)t + p - 2 & \text{if } t < 0; \end{cases}$$

$$A(t) = (p-1)(1-t)^2 - [(p-2) - p\,t]\,E(t);$$

$$D(t) = (p-1)(1-t)^2 + t\,E(t);$$

$$B(t) = (p-1)(1-t)^2 E(t) - t\,A(t) = [(p-1) - p\,t]\,E(t) - t\,D(t)$$

and subintervals of the real line

$$I_p = \begin{cases} \left(\dfrac{p-3}{2p}, 0\right] & \text{if } 2 < p \le p_0, \\ \left(\max\left\{0, \dfrac{p-3}{p-1}\right\}, \dfrac{p-2}{p}\right) & \text{if } p_0 < p < \infty. \end{cases}$$

For $2 < p < \infty$, there exists a unique solution $t_p \in I_p$ of the equation

$$[(p-2) - (p-1)t]\,A(t)^{p-1} = B(t)^{p-1}$$

and it follows that

$$c_p = \left[\frac{A(t_p)}{(p-1)(1-t_p)^2}\right]^{1/p} \frac{D(t_p)}{A(t_p)}.$$

As an example, $t_{p_0} = 0$ and $c_{p_0} = (p_0 - 2)^{(2-p_0)/p_0} = 1.1386774769...$, which is greater than $1 = c_2$. More examples include $c_3 = 1.3291719357...$ and $c_4 = 1.7919250903....$ As $p \to \infty$, we have

$$c_p = \frac{p}{2} + \frac{1}{2}\ln\left(\frac{1+e^{-2}}{2}\right) + \frac{\alpha}{p} + \cdots$$

where

$$\alpha = \frac{1}{4}\ln\left(\frac{1+e^{-2}}{2}\right)^2 + \frac{1}{2}\ln\left(\frac{1+e^{-2}}{2}\right) - 2\left(\frac{e^{-2}}{1+e^{-2}}\right).$$

The numbers t_3 and t_4 are algebraic of degrees 4 and 5, respectively, but this fact does not help us determine closed-form expressions for c_3 or c_4.

3.8.1 *Dyadic Martingales*

A sequence of random variables $\{f_n\}_{n\geq 1}$ is a **martingale** if, for all n, $\mathrm{E}\,(|f_n|) < \infty$ and [10]

$$\mathrm{E}\,(f_{n+1}\ |f_n, f_{n-1}, \ldots, f_1) = f_n.$$

Thinking of f_n as the fortune of a gambler at trial n of a game, the conditional equality states that the game is "fair" in the sense that the expected fortune at trial $n+1$, given knowledge of all past trials, is the same as the fortune at trial n.

To go further, we re-index and normalize what we called h_n earlier (in a manner often consistent with the wavelets literature). For integers $k \geq 1$ and $1 \leq j \leq 2^{k-1}$, define the **Haar functions** $\chi_k^j : [0,1) \to \mathbb{R}$ by

$$\chi_k^j(t) = \begin{cases} 2^{(k-1)/2} & \text{if } \dfrac{j-1}{2^{k-1}} \leq t < \dfrac{j-1/2}{2^{k-1}}, \\ -2^{(k-1)/2} & \dfrac{j-1/2}{2^{k-1}} \leq t < \dfrac{j}{2^{k-1}}, \\ 0 & \text{otherwise.} \end{cases}$$

In words, for each **dyadic subinterval** I of $[0,1)$ of length $2^{-(k-1)}$, we have a function equal to $2^{(k-1)/2}$ on the left half of I and $-2^{(k-1)/2}$ on the right half of I.

Let X and Y be real Banach spaces. A **dyadic martingale** is a set $\{f_n\}_{n=1}^{\infty}$ where each $f_n : [0,1) \to X$ is a linear combination of Haar functions:

$$f_n(t) = \sum_{k=1}^{n} \sum_{j=1}^{2^{k-1}} \chi_k^j(t) x_k^j$$

and each $x_k^j \in X$ is independent of n. Let $f_0 = 0$ and denote by $d_k = f_k - f_{k-1}$ the martingale differences. Given an operator $T : X \to Y$, the n^{th} **dyadic UMD constant** $\mu_n(T)$ is the least quantity $c \geq 0$ such that

$$\left\| \sum_{k=1}^{n} \varepsilon_k T d_k \right\|_2 \leq c \left\| \sum_{k=1}^{n} d_k \right\|_2$$

for all martingale differences $d_1, \ldots, d_n$ and all sequences $\varepsilon_1, \ldots, \varepsilon_n$ of signs. The norm on the right-hand side is the L_2-norm on measurable X-valued functions, the norm on the left-hand side is the L_2-norm on measurable Y-valued functions, and the acronym UMD stands for "unconditional martingale differences".

We are interested in the case when $X = \ell_1^m$ and $Y = \ell_\infty^m$, sequence spaces of m dimensions, and T is the finite summation operator

$$T_m(\xi_1, \ldots, \xi_m) = \left(\xi_1, \xi_1 + \xi_2, \xi_1 + \xi_2 + \xi_3, \ldots, \sum_{i=1}^{m} \xi_i \right)$$

where $m = 2^n$ for notational convenience. It is known that there exist constants $a > 0$, $b > 0$ such that

$$\sqrt{n} \le a\,\mu_n(T_m), \quad b\,\mu_n(T_m) \le n$$

independent of n. What, however, is the true asymptotic behavior of $\mu_n(T_m)$?

Wenzel [11, 12] proved that the growth rate of $\mu_n(T_m)$ is the same as the growth rate of

$$\theta_n = \sup_{\pi} \frac{1}{2^n} \sum_{i=0}^{2^n-1} \sup_{0\le k<2^n} \left| \sum_{j:\pi(j)\le k} (-2)^{-\kappa(i\oplus j)} \right|,$$

where the outer summation is taken over all permutations π of the set $\{0,\ldots,2^n-1\}$, $i \oplus j$ denotes the bitwise XOR sum of i and j (addition modulo two without carries [13]), and $\kappa(n) = 1 + \lfloor \ln(n)/\ln(2) \rfloor$ if $n > 0$, $\kappa(0) = 2$. He computed that

$$\theta_3 \approx 0.5937, \quad \theta_4 \approx 0.6718, \quad \theta_5 \approx 0.7509, \quad \theta_6 \approx 0.8203$$

and therefore conjectured that $\sqrt{n}$ is the correct growth rate. In fact, his calculations suggest that $\theta_n \sim (0.3...)\sqrt{n}$ as $n \to \infty$, and we wonder if the corresponding constant for $\mu_n(T_m)/\sqrt{n}$ will ever be known [14].

[1] J. Schauder, Eine Eigenschaft des Haarschen Orthogonalsystems, *Math. Z.* 28 (1928) 317–320.

[2] I. Singer, *Bases in Banach spaces. I*, Springer-Verlag, 1970, pp. 13–17, 396–432; MR0298399 (45 #7451).

[3] R. C. James, Bases in Banach spaces, *Amer. Math. Monthly* 89 (1982) 625–640; MR0678808 (84h:46019).

[4] R. E. A. C. Paley, A remarkable series of orthogonal functions. I, *Proc. London Math. Soc.* 34 (1932) 241–264.

[5] J. Marcinkiewicz, Quelques théorèmes sur les séries orthogonales, *Ann. Soc. Polon. Math.* 16 (1937) 84–96.

[6] D. L. Burkholder, Boundary value problems and sharp inequalities for martingale transforms, *Annals of Probab.* 12 (1984) 647–702; MR0744226 (86b:60080).

[7] D. L. Burkholder, An elementary proof of an inequality of R. E. A. C. Paley, *Bull. London Math. Soc.* 17 (1985) 474–478; MR0806015 (87a:42041).

[8] K. P. Choi, A sharp inequality for martingale transforms and the unconditional basis constant of a monotone basis in $L^p(0,1)$, *Trans. Amer. Math. Soc.* 330 (1992) 509–529; MR1034661 (92f:60073).

[9] K. P. Choi, *Some Sharp Inequalities for Martingale Transforms*, Ph.D. thesis, Univ. of Illinois at Urbana-Champaign, 1987.

[10] R. Durrett, *Essentials of Stochastic Processes*, Springer-Verlag, 1999, pp. 102–110; MR1703359 (2002i:60001).

[11] J. Wenzel, The UMD constants of the summation operators, *Quaest. Math.* 27 (2004) 111–136; arXiv:math/0407481; MR2091691 (2005h:46019).

[12] J. Wenzel, *Haar Functions, Martingales and Geometry of Banach Spaces*, Habilitation thesis, Friedrich-Schiller-Universität Jena, 2003.

[13] S. R. Finch, Plouffe's constant, *first volume*, pp. 430–433.
[14] S. Geiss, S. Montgomery-Smith and E. Saksman, On singular integral and martingale transforms, *Trans. Amer. Math. Soc.* 362 (2010) 553–575; arXiv:math/0701516; MR2551497 (2011a:60168).

3.9 Power Series with Restricted Coefficients

Define a family of functions

$$\mathcal{F} = \left\{ 1 + \sum_{n=1}^{\infty} a_n x^n : a_n \in \{-1, 0, 1\} \right\}$$

and three closed subsets of the open interval $(0, 1)$:

$$\Omega_2 = \{x : \exists f \in \mathcal{F} \text{ for which } f(x) = f'(x) = 0\},$$

$$\Omega_3 = \{x : \exists f \in \mathcal{F} \text{ for which } f(x) = f'(x) = f''(x) = 0\},$$

$$\Omega_4 = \{x : \exists f \in \mathcal{F} \text{ for which } f(x) = f'(x) = f''(x) = f'''(x) = 0\}.$$

Elements of Ω_2 are called **double zeroes**, those of Ω_3 **triple zeroes** and those of Ω_4 **quadruple zeroes**. For each $k = 2, 3, 4$, define [1]

$$\alpha_k = \min \Omega_k, \quad \tilde{\alpha}_k = \sup \Omega_k^c,$$

where Ω_k^c is the complement of Ω_k in $(0, 1)$. The structure of Ω_k is very complicated – it appears to possess infinitely many connected components – but provably $\alpha_2 = 0.6684756...$ and conjecturally

$$\tilde{\alpha}_2 = 0.669..., \quad \alpha_3 = 0.743..., \quad \tilde{\alpha}_3 \approx 0.75....$$

No one has yet examined α_4 or $\tilde{\alpha}_4$ numerically, as far as is known. Elements of Ω_2^c are said to satisfy a certain **tranversality condition**, in the sense that $y \in \Omega_2^c$ and $f(y) = 0$ imply that $f'(y) \neq 0$ for all $f \in \mathcal{F}$. Such a property is useful in [2] for a seemingly unrelated analysis of fractals.

Define instead

$$\hat{\mathcal{F}} = \left\{ 1 + \sum_{n=1}^{\infty} a_n x^n : a_n \in \{-2, -1, 0, 1, 2\} \right\}$$

and $\hat{\Omega}_2$ to be the corresponding set of double zeroes in $(0, 1)$. In this case, $\min \hat{\Omega}_2$ is precisely $1/2$ and is an isolated point of $\hat{\Omega}_2$. Removing $1/2$ from $\hat{\Omega}_2$ appears to give a connected set (that is, an interval) and the minimum of this set is conjectured to be ≈ 0.5437. The fact that Ω_2 and $\hat{\Omega}_2$ are so distinct topologically is very striking [1].

A different family of functions, studied earlier in [3, 4], is

$$\mathcal{G} = \left\{ 1 + \sum_{n=1}^{\infty} b_n x^n : b_n \in [-1, 1] \right\}.$$

Let β_k denote the associated minimum zero of order k (at least) of g, taken over all $g \in \mathcal{G}$. It turns out that β_k is always algebraic: $\beta_2 = 0.6491378608...$ has minimal polynomial

$$2z^5 - 8z^2 + 11z - 4,$$

$\beta_3 = 0.7278832326...$ has minimal polynomial

$$10z^{12} - 14z^{11} + 14z^6 - 10z^5 - 80z^3 + 185z^2 - 147z + 40,$$

and $\beta_4 = 0.7773295434...$ has minimal polynomial

$$\begin{aligned}&126z^{22} - 296z^{21} + 176z^{20} + 44z^{12} - 104z^{11} + 54z^{10} + 96z^7\\ &-146z^6 + 56z^5 - 684z^4 + 2236z^3 - 2797z^2 + 1584z - 342.\end{aligned}$$

Of course, $\beta_1 = 1/2$, which corresponds to $g(x) = 1 - \sum_{n=1}^{\infty} x^n$. The following least squares approximation

$$\beta_k \approx 1 - \frac{1}{(1.23909318...) + (0.81255949...)k}$$

was obtained in [4] and is based on data up to $k = 27$. We wonder if more precise asymptotics are feasible. Additional relevant references include [5–7].

[1] P. Shmerkin and B. Solomyak, Zeros of $\{-1, 0, 1\}$ power series and connectedness loci for self-affine sets, *Experim. Math.* 15 (2006) 499–511; arXiv:math/0504545; MR2293600 (2007k:30003).

[2] I. Benjamini and B. Solomyak, Spacings and pair correlations for finite Bernoulli convolutions, *Nonlinearity* 22 (2009) 381–393; arXiv:0808.1568; MR2475552 (2009m:28035).

[3] B. Solomyak, On the random series $\sum \pm\lambda^n$ (an Erdős problem), *Annals of Math.* 142 (1995) 611–625; MR1356783 (97d:11125).

[4] F. Beaucoup, P. Borwein, D. W. Boyd and C. Pinner, Multiple roots of $[-1, 1]$ power series, *J. London Math. Soc.* 57 (1998) 135–147; MR1624809 (99c:30005).

[5] A. M. Odlyzko and B. Poonen, Zeros of polynomials with 0, 1 coefficients, *Enseign. Math.* 39 (1993) 317–348; MR1252071 (95b:11026).

[6] P. Borwein, T. Erdélyi and G. Kós, Littlewood-type problems on [0, 1], *Proc. London Math. Soc.* 79 (1999) 22–46; MR1687555 (2000c:11111).

[7] P. Borwein, T. Erdélyi and G. Kós, The multiplicity of the zero at 1 of polynomials with constrained coefficients, *Acta Arith.* 159 (2013) 387–395; MR3080800.

3.10 Hankel and Toeplitz Determinants

The most famous Hankel matrix is the Hilbert matrix

$$H_n = \left(\frac{1}{i+j-1}\right)_{1 \le i,j \le n},$$

which has determinant equal to a ratio of Barnes G-function values:

$$\det(H_n) = \frac{\prod_{k=1}^{n-1}(k!)^4}{\prod_{\ell=1}^{2n-1}\ell!} = \frac{G(n+1)^4}{G(2n+1)} \to 0$$

as $n \to \infty$. More precisely [1],

$$\frac{\det(H_n)}{4^{-n^2}(2\pi)^n n^{-1/4}} \to 2^{1/12}e^{1/4}A^{-3} = 0.6450024485...$$

where A denotes the Glaisher–Kinkelin constant [2]. Such Hankel determinants are important in random matrix theory and applications [3], but we shall forsake all this, giving instead only a few examples [4–6]. Another interesting fact is that $\det(H_n)$ is always the reciprocal of a positive integer [7].

The Hankel determinant of Euler numbers [8] is, in absolute value,

$$\begin{aligned} |E_{i+j}|_{0\le i,j\le n-1} &= \prod_{k=1}^{n-1}(k!)^2 = G(n+1)^2 \\ &\sim \frac{e^{\frac{1}{6}}}{A^2}e^{-\frac{3}{2}n^2}(2\pi)^n n^{n^2-\frac{1}{6}} \end{aligned}$$

as $n \to \infty$. The simplicity of this result contrasts with the following. The Hankel determinant of Bernoulli numbers [9] is, in absolute value,

$$\begin{aligned} |B_{i+j}|_{0\le i,j\le n-1} &= \prod_{k=1}^{n-1}\frac{(k!)^6}{(2k)!(2k+1)!} \\ &= \frac{2^{\frac{1}{12}}e^{\frac{1}{4}}}{A^3}4^{-n^2}(2\pi)^n\frac{G(n+1)^4}{G(n+1/2)G(n+3/2)} \\ &\sim \frac{2^{\frac{1}{12}}e^{\frac{5}{12}}}{A^5}4^{-n^2}e^{-\frac{3}{2}n^2}(2\pi)^{2n}n^{n^2-\frac{5}{12}} \end{aligned}$$

as $n \to \infty$. We mention three formulas of Krattenthaler [10]:

$$\left|\frac{B_{2i+2j+2}}{(2i+2j+2)!}\right|_{0\le i,j\le n-1} = 4^{-n^2}\prod_{k=1}^{2n-1}(2k+1)^{-2n+k},$$

$$\left|\frac{B_{2i+2j+4}}{(2i+2j+4)!}\right|_{0\le i,j\le n-1} = 4^{-n^2-n}9^{-n}\prod_{k=1}^{2n-1}(2k+3)^{-2n+k},$$

$$\left|\frac{B_{2i+2j+6}}{(2i+2j+6)!}\right|_{0\le i,j\le n-1} = (n+1)(2n+3)4^{-n^2-2n}\prod_{k=1}^{2n+1}(2k+1)^{-2n-2+k},$$

which are always reciprocals of integers (unlike $|E_{i+j}|$ and $|B_{i+j}|$). The asymptotics of these three sequences remain open.

More difficult are determinants of Riemann zeta function values:

$$a_n^{(0)} = |\zeta(i+j)|_{1\le i,j\le n}, \quad a_n^{(1)} = |\zeta(i+j+1)|_{1\le i,j\le n},$$

which evidently satisfy

$$a_n^{(0)} \sim C\cdot\left(\frac{2n+1}{e^{3/2}}\right)^{-(n+1/2)^2}, \quad a_n^{(1)} \sim \frac{e^{9/8}}{\sqrt{6}}C\cdot\left(\frac{2n}{e^{3/2}}\right)^{-n^2+3/4}$$

thanks to numerical experiments by Zagier [11]. No closed-form expression for the constant $C = 0.351466738331...$ is known.

A famous Toeplitz matrix, called the alternating Hilbert matrix in [12], is

$$\tilde{H}_n = \left(\frac{1}{i-j}\right)_{1\le i,j\le n},$$

where we understand the diagonal elements to be 0. Schur [13] proved long ago that the maximum eigenvalue (in modulus) of both H_n and $\tilde{H}_n$ is less than π and approaches π as $n\to\infty$. The determinant is, of course, the product of all eigenvalues. When n is odd, $\det(\tilde{H}_n)=0$. When n is even, a closed-form expression for $\det(\tilde{H}_n)$ seems to be unavailable, despite the existence of a combinatorial approach [14]. Note that the "symbol" associated with $\tilde{H}_n$ is

$$\sum_{r=1}^{\infty}\frac{e^{ir\theta}}{-r} + \sum_{r=1}^{\infty}\frac{e^{-ir\theta}}{r} = i(\theta-\pi)$$

for $0<\theta<2\pi$, hence a theorem due to Grenander & Szegő [15] gives

$$\lim_{\substack{n\to\infty \\ n \text{ even}}}\frac{1}{n}\ln\left(\det(\tilde{H}_n)\right) = \frac{1}{2\pi}\int_0^{2\pi}\ln\left[i(\theta-\pi)\right]d\theta = -1+\ln(\pi) = 0.1447298858....$$

A refined estimate shown subsequently in [15], potentially governing the value of

$$\lim_{n\to\infty}\det(\tilde{H}_n)\cdot\left(\frac{\pi}{e}\right)^n,$$

has conditions that must be verified.

Consider finally another Toeplitz matrix

$$K_n = \left(\frac{1}{1+|i-j|}\right)_{1\le i,j\le n}$$

for which little is known. The "symbol" here is

$$\sum_{r=0}^{\infty}\frac{e^{ir\theta}}{1+r} + \sum_{r=1}^{\infty}\frac{e^{-ir\theta}}{1+r} = -1 - e^{i\theta}\ln\left(1-e^{-i\theta}\right) - e^{-i\theta}\ln\left(1-e^{i\theta}\right)$$

for $0 < \theta < 2\pi$, hence

$$\lim_{n\to\infty} \frac{1}{n} \ln\left(\det(K_n)\right) = \frac{1}{2\pi} \int_0^{2\pi} \ln\left[-1 - e^{i\theta} \ln\left(1 - e^{-i\theta}\right) - e^{-i\theta} \ln\left(1 - e^{i\theta}\right)\right] d\theta$$
$$= -0.3100863233....$$

An exact formula for this constant is desired; might, at least, the integral be simplified in some way?

3.10.1 *Combinatorial Approach*

Assume that n is even. Let S denote the set of all $(n/2)$-tuples of ordered pairs:

$$(p_k, q_k)_{k=1}^{n/2}$$

of positive integers $p_k < q_k$ satisfying

$$\bigcup_{k=1}^{n/2} \{p_k, q_k\} = \{1, 2, \ldots, n\}$$

and $p_1 < p_2 < \ldots < p_{n/2}$. Note that the qs need not be in ascending order. Let us verify a formula in [14]:

$$\det(\tilde{H}_n) = \sum_{(p_k,q_k)_{k=1}^{n/2} \in S} \prod_{k=1}^{n/2} \frac{1}{(q_k - p_k)^2}$$

for $n = 4$. Three such 2-tuples exist:

$$p_1 = 1 < p_2 = 2 < q_1 = 3 < q_2 = 4,$$

$$p_1 = 1 < p_2 = 2 < q_2 = 3 < q_1 = 4,$$

$$p_1 = 1 < q_1 = 2 < p_2 = 3 < q_2 = 4$$

yielding

$$\frac{1}{(3-1)^2(4-2)^2} + \frac{1}{(4-1)^2(3-2)^2} + \frac{1}{(2-1)^2(4-3)^2} = \frac{169}{144} = \det(\tilde{H}_4).$$

The case $\det(\tilde{H}_2) = 1$ is trivial; the case $\det(\tilde{H}_6) = 6723649/4665600$ will require some effort. We wonder if a simple method for computing the size of S, as a function of n, can be found. An analogous approach for $\det(K_n)$ would also be good to see.

Acknowledgments I am grateful to Olivier Lévêque for evaluating $\det(K_n)$ asymptotics and Hartmut Monien for helpful correspondence.

[1] Wikipedia contributors, Hilbert matrix, *Wikipedia, The Free Encyclopedia* (2014).
[2] S. R. Finch, Glaisher-Kinkelin constant, *first volume*, pp. 135–145.
[3] E. L. Basor, Y. Chen and H. Widom, Determinants of Hankel matrices, *J. Funct. Anal.* 179 (2001) 214–234; MR1807258 (2001m:47043).
[4] A. Junod, Hankel determinants and orthogonal polynomials, *Expo. Math.* 21 (2003) 63–74; MR1955218 (2004a:15009).
[5] C. Krattenthaler, Advanced determinant calculus, *Sémin. Lothar. Combin.* 42 (1999) B42q; arXiv:math/9902004; MR1701596 (2002i:05013).
[6] C. Krattenthaler, Advanced determinant calculus: a complement, *Linear Algebra Appl.* 411 (2005) 68–166; arXiv:math/0503507; MR2178686 (2006g:05022).
[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005249 and A067689.
[8] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[9] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[10] S. Fukuhara and Y. Yang, Period polynomials and explicit formulas for Hecke operators on $\Gamma_0(2)$, *Math. Proc. Cambridge Philos. Soc.* 146 (2009) 321–350; arXiv:math/0608372; MR2475970 (2009m:11062).
[11] H. Monien, Hankel determinants of Dirichlet series, arXiv:0901.1883.
[12] N. Kurokawa and H. Ochiai, Spectra of alternating Hilbert operators, *Spectral Analysis in Geometry and Number Theory*, Proc. 2007 Nagoya conf., ed. M. Kotani, H. Naito and T. Tate, Amer. Math. Soc., 2009, pp. 89–101; arXiv:0709.2675; MR1500140 (2010c:11113).
[13] I. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, *J. Reine Angew. Math.* 140 (1911) 1–28.
[14] E. Preissmann and O. Lévêque, On generalized weighted Hilbert matrices, *Pacific J. Math.* 265 (2013) 199–219; arXiv:1303.1085; MR3095119.
[15] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*, Chelsea, 1984, pp. 62–80; MR0890515 (88b:42031).

3.11 Goľdberg's Zero-One Constants

Let F be the set of all functions f that are analytic on some ring $\{z : \rho(f) < |z| < 1\}$ and omit the values of both 0 and 1 there. Each function is defined in its own distinct ring. By *omit*, it is meant that $f(z) \notin \{0, 1\}$ for all z. We assume $\rho(f)$ to be as small as possible. Let $G \subseteq F$ consist of all functions that are analytic on the open unit disk D. Thus, for $f \in G$, we have

$$\rho(f) = \begin{cases} 0 & \text{if } f \text{ is never 0 or 1,} \\ \sup\{|z| : f(z) \in \{0,1\}\} & \text{otherwise.} \end{cases}$$

Given a real number a, the a**-points** of f are the points z for which $f(z) = a$. Of course, 0-points are more commonly referred to as *zeroes*.

Consider the circle σ defined by

$$\left\{z : |z| = \sqrt{\rho(f)}\right\}$$

with counterclockwise orientation, and let γ_f be the image of σ under f. The **index** or **winding number** of γ_f with respect to the point a is

$$n(\gamma_f, a) = \frac{1}{2\pi i} \int\limits_{\gamma_f} \frac{dz}{z-a}.$$

Our interest is in the scenario when $n(\gamma_f, 0)$, $n(\gamma_f, 1)$ are nonzero and distinct; without loss of generality, we assume that $n(\gamma_f, 0) > n(\gamma_f, 1)$. Let $F(N_0, N_1) \subseteq F$ consist of all functions f with $n(\gamma_f, 0) = N_0$ and $n(\gamma_f, 1) = N_1$. Let $G(M_0, M_1) \subseteq G$ consist of all functions g with exactly one 0-point [of multiplicity M_0] and exactly one 1-point [of multiplicity M_1]. Again, we focus on $M_0 \neq 0$, $M_1 \neq 0$ and $M_0 \neq M_1$; without loss of generality, assume that $M_0 > M_1$.

Goľdberg [1] studied constants similar to

$$A(N_0, N_1) = \inf \{\rho(f) : f \in F(N_0, N_1)\},$$

$$B(M_0, M_1) = \inf \{\rho(g) : g \in G(M_0, M_1)\}.$$

Bergweiler & Eremenko [2] discovered closed-form expressions:

$$A(2,1) = \nu = \exp\left(-\frac{\pi^2}{\ln\left(3+2\sqrt{2}\right)}\right) = 0.0037015991...,$$

$$A(3,1) = A(3,2) = \exp\left(-\frac{\pi^2}{\ln\left(5+2\sqrt{6}\right)}\right) = 0.0134968456...,$$

$$A(4,1) = A(4,3) = \exp\left(-\frac{\pi^2}{\ln\left(7+4\sqrt{3}\right)}\right) = 0.0235855221...$$

and moreover proved that

$$A = \inf \{\rho(f) : f \in F \text{ and } N_0 > N_1 \geq 1\} = \nu.$$

(Goľdberg's original bounds for A were strengthened by Jenkins [3].) The numerical computation of

$$B(2,1) = \mu = 0.0252896...,$$

$$B(3,1) = 0.084924..., \quad B(3,2) = 0.227417...,$$

$$B(4,1) = 0.140571..., \quad B(4,3) = 0.290697...$$

is more difficult – no precise formulas are known – and it is merely conjectured that

$$B = \inf \{\rho(g) : g \in G \text{ and } M_0 > M_1 \geq 1\} = \mu.$$

(The best lower bound 0.00587 for B in [2], improving on [4–6], is still far off.) An elaborate construction of a certain transcendental analytic function on D possessing exactly one 0-point at $-\mu$ [with $M_0 = 2$] and exactly one 1-point at μ [with $M_1 = 1$] occupies much of the discussion in [2]. It shows that $B \leq \mu$. A proof that $B \geq \mu$ remains open.

3.11.1 *Belgian Chocolate Problem*

Here the difficulties of construction are overwhelming. What is the smallest $\tau > 0$ for which there exists an analytic function on D possessing exactly one 0-point at 0 [of multiplicity 1] and exactly two 1-points at $\pm\tau$ [each of multiplicity 1]? The current best bounds are [2, 7]

$$0.01450779 < \tau < 0.10913022.$$

Blondel's question [8, 9] is often phrased as follows. Let $a(z) = z^2 - 2\delta z + 1$ and $b(z) = z^2 - 1$. What is the largest $\delta > 0$ for which there exist stable real polynomials p and q with $\deg(p) \geq \deg(q)$ such that $ap + bq$ is stable? (A polynomial is called **stable** if all its zeroes are in the left half plane.) The numbers τ and δ are related by

$$\tau = \sqrt{\frac{1-\delta}{1+\delta}}, \quad \delta = \frac{1-\tau^2}{1+\tau^2}$$

and the current best bounds are

$$0.97646152 < \delta < 0.99957913.$$

Incremental progress in specifying such constraints is found in [4, 10–14].

3.11.2 *Landau's Theorem with Explicit Bound*

If an analytic function g on D omits the values of both 0 and 1, then [15–17]

$$|g(0)| \leq 2\,|g'(0)|\,(|\ln|g(0)|| + K),$$

where the constant

$$K = \frac{1}{4\pi^2}\Gamma\left(\frac{1}{4}\right)^4 = 4.3768792304...$$

is best possible. Other occurrences of K are similar to results appearing in [18]. If analytic g satisfies $g(0) = 0$ and $g'(0) = 1$, then $g(D)$ covers a segment of each line passing through the origin; further, each segment has length at least $2/K = 0.4569465810...$ and this is sharp [19–21]. If analytic g satisfies $g(-z) = -g(z)$ for all $z \in D$ and $g'(0) = 1$, then $g(D)$ covers a disk with center at the origin and radius $1/K = 0.2284732905...$; again, this is sharp [8, 22]. The presence of the elliptic modular function

$$J(z) = 16\exp(\pi i z)\prod_{n=1}^{\infty}\left(\frac{1+\exp(2n\pi i z)}{1+\exp((2n-1)\pi i z)}\right)^{8}, \quad \mathrm{Im}(z) > 0,$$

$$\frac{1}{J'(i)} = \frac{4}{K}i$$

is keenly felt here.

[1] A. A. Goľdberg, A certain theorem of Landau type (in Russian), *Teor. Funkciĭ Funkcional. Anal. i Priložen.* 17 (1973) 200–206, 246; MR0352464 (50 #4951).
[2] W. Bergweiler and A. Eremenko, Goľdberg's constants, *J. d'Analyse Math.* 119 (2013) 365–402; arXiv:1111.2296; MR3043157.
[3] J. A. Jenkins, On a problem of A. A. Goľdberg, *Annales Univ. Mariae Curie-Skłodowska Sect. A* 36/37 (1982/83) 83–86; MR0808435 (87m:30011).
[4] V. D. Blondel, R. Rupp and H. S. Shapiro, On zero and one points of analytic functions, *Complex Variables Theory Appl.* 28 (1995) 189–192; MR1700083 (2000d:30008).
[5] P. Batra, On small circles containing zeros and ones of analytic functions, *Complex Var. Theory Appl.* 49 (2004) 787–791; MR2097217 (2005h:30010).
[6] P. Batra, On Goľdberg's constant A_2, *Comput. Methods Funct. Theory* 7 (2007) 33–41; MR2321799 (2008f:30013).
[7] N. Boston, On the Belgian chocolate problem and output feedback stabilization: Efficacy of algebraic methods, *50th Annual Allerton Conference on Communication, Control, and Computing*, 2012, pp. 869–873.
[8] V. D. Blondel, *Simultaneous Stabilization of Linear Systems*, Lect. Notes in Control and Info. Sci. 191, Springer-Verlag, 1994, pp. 118–124, 132–144, 149–150, 160–168; MR1254658 (95m:93001).
[9] V. D. Blondel, Simultaneous stabilization of linear systems and interpolation with rational functions, *Open Problems in Mathematical Systems and Control Theory*, ed. V. D. Blondel, E. D. Sontag, M. Vidyasagar and J. C. Willems, Springer-Verlag, 1999, pp. 53–59; MR1727924 (2000g:93003).
[10] R. Rupp, A covering theorem for a composite class of analytic functions, *Complex Variables Theory Appl.* 25 (1994) 35–41; MR1310853 (96c:30006).
[11] R. Rupp, A note on covering theorems for composite classes of analytic functions, *Mitt. Math. Sem. Giessen* 223 (1995) 65–73; MR1368244 (97b:30005).
[12] V. V. Patel, G. Deodhare and T. Viswanath, Some applications of randomized algorithms for control system design, *Automatica* 38 (2002) 2085–2092; MR2134873.
[13] YJ Chang and N. V. Sahinidis, Global optimization in stabilizing controller design, *J. Global Optim.* 38 (2007) 509–526; MR2335309 (2008d:93055).
[14] J. V. Burke, D. Henrion, A. S. Lewis and M. L. Overton, Stabilization via nonsmooth, nonconvex optimization, *IEEE Trans. Automat. Control* 51 (2006) 1760–1769; MR2265983 (2007h:93082).
[15] W. T. Lai, The exact value of Hayman's constant in Landau's theorem, *Sci. Sinica*, v. 22 (1979) n. 2, 129–134; MR0532327 (80k:30002).
[16] J. A. Hempel, The Poincaré metric on the twice punctured plane and the theorems of Landau and Schottky, *J. London Math. Soc.* 20 (1979) 435–445; MR0561135 (81c:30025).

[17] J. A. Jenkins, On explicit bounds in Landau's theorem. II, *Canad. J. Math.* 33 (1981) 559–562; MR0627642 (83a:30026).
[18] S. R. Finch, Bloch-Landau constants, *first volume*, pp. 456–459.
[19] G. M. Goluzin, *Geometric Theory of Functions of a Complex Variable*, Amer. Math. Soc., pp. 72–75, 86–89; MR0247039 (40 #308).
[20] A. Bermant, Dilatation d'une fonction modulaire et problèmes de recouvrement (in Russian), *Rec. Math. [Mat. Sbornik]* 15 [57] (1944) 285–324; MR0013415 (7,150d).
[21] A. Yu. Solynin and M. Vuorinen, Estimates for the hyperbolic metric of the punctured plane and applications, *Israel J. Math.* 124 (2001) 29–60; MR1856503 (2002j:30071).
[22] Z. Nehari, *Conformal Mapping*, Dover Publ., 1975, pp. 318–332; MR0377031 (51 #13206).

3.12 Electrical Capacitance

We mentioned **logarithmic capacity** or **transfinite diameter** in [1]. Given a compact set A in $\mathbb{R}^2$, the measure

$$\gamma_0(A) = \lim_{n\to\infty} \max_{\xi_1,\ldots,\xi_n\in A} \left(\prod_{j<k} |\xi_j - \xi_k| \right)^{\frac{2}{n(n-1)}}$$

is invariant under rigid motions and continuous, but fails to be additive since $\gamma_0(A) = \gamma_0(\partial A)$ [2–4]. The unit interval has logarithmic capacity $1/4$; the unit disk, square and equilateral triangle have logarithmic capacities

$$1, \quad \frac{1}{4\pi^{3/2}}\Gamma\left(\frac{1}{4}\right)^2 = 0.5901702995..., \quad \frac{\sqrt{3}}{8\pi^2}\Gamma\left(\frac{1}{3}\right)^3 = 0.4217539346...,$$

respectively. Discussion of the geometric mean (of all pairs of points) often seems to be restricted to planar sets; we now turn to the harmonic mean and subsequently to the arithmetic mean.

Given a compact set A in $\mathbb{R}^3$, define [5, 6]

$$\gamma_{-1}(A) = \lim_{n\to\infty} \max_{\xi_1,\ldots,\xi_n\in A} \left(\frac{2}{n(n-1)} \sum_{j<k} \frac{1}{|\xi_j - \xi_k|} \right)^{-1}$$

to be the **Newtonian capacity** or **electrical capacitance** or **generalized transfinite diameter** of order -1. This is also the reciprocal of what is known as the **optimal Riesz 1-energy** [7]. The unit interval and unit circle both have electrical capacitance 0; one way to see the latter is to notice the inequality [8]

$$\sum_{j<k} \frac{1}{|\xi_j - \xi_k|} \geq \frac{n}{4} \sum_{\ell=1}^{n-1} \csc\left(\frac{\ell\pi}{n}\right)$$

(for which equality holds when $\xi_1, \ldots, \xi_n$ are n^{th} roots of unity). The unit disk has capacitance $2/\pi$ [9] If A is the closure of a bounded, open, connected set in

$\mathbb{R}^3$, then $\gamma_{-1}(A) = \gamma_{-1}(\partial A)$ [10]. The unit ball (and hence the unit sphere) has capacitance 1. Another way to see this is to invoke a formula for s-energy of the d-sphere [11] with $s = 1$, $d = 2$.

Interesting constants arise here. For example, let A be the solid formed by revolving a disk of radius 1 about a tangent line (a "torus without hole"). It follows that [12]

$$\gamma_{-1}(A) = \frac{4}{\pi}\int_0^\infty \frac{1}{I_0(t)^2}\,dt = 4\,(0.4353450662...)$$

where $I_0(t)$ is the zeroth modified Bessel function. More generally, consider the surface formed by revolving an arc of a circle about its chord (a "spindle"). A definite integral involving Legendre functions of complex degree, parametrized by the included angle, is found [13]. As another example, consider the (disconnected) set consisting of two congruent parallel line segments. Its capacitance is obtained via a transcendental equation that involves elliptic integrals [14–16]. See [10, 17–20] for more examples.

Seemingly simple sets present formidably difficult challenges [21]. The unit cube C has attracted enormous attention [22–41] and the best numerical estimate is [2, 9, 42]

$$\gamma_{-1}(C) = 0.6606781540... = \frac{1}{2}(1.3213563081...).$$

A conjectured exact expression for $\gamma_{-1}(C)$ in [43, 44] is evidently incorrect. For the unit square S and the unit equilateral triangle T, we have less precision:

$$\gamma_{-1}(S) = 0.3667880... = \frac{1}{2}(0.7335760...) = \frac{2}{\pi}(0.5761492...),$$

$$\gamma_{-1}(T) = 0.2508... = \frac{2}{\pi}(0.3940...).$$

It would be good someday to see improvements of these estimates, as well as $0.3565... = (1.7465...)/\sqrt{24}$ for the unit regular tetrahedron. We wonder if formulation in [45, 46] might assist in accomplishing this.

The preceding results are dimensionless, of course. Certain authors chose to express their estimates in the following manner:

$$\gamma_{-1}(C) \approx \frac{1}{4\pi\varepsilon_0}(73.51036),$$

$$\gamma_{-1}(S) \approx \frac{1}{4\pi\varepsilon_0}(40.811) \approx \frac{1}{\sqrt{2}}\frac{1}{4\pi\varepsilon_0}(57.715),$$

$$\gamma_{-1}(T) \approx \frac{1}{4\pi\varepsilon_0}(27.91) \approx \frac{1}{\sqrt{3}}\frac{1}{4\pi\varepsilon_0}(48.33),$$

where $4\pi\varepsilon_0 \approx 111.265006$ picofarads/meter and ε_0 is the permittivity constant of free space. Such decisions are a little unfortunate for us, since the value of ε_0 is based on physical experimentation and thus the normalization has changed somewhat with the passage of time.

Moving back to geometry, define the **generalized transfinite diameter** of order 1 or **optimal Riesz** (-1)**-energy**

$$\gamma_1(A) = \lim_{n\to\infty} \max_{\xi_1,\ldots,\xi_n\in A} \left(\frac{2}{n(n-1)} \sum_{j<k} |\xi_j - \xi_k| \right)$$

where A is a compact set in $\mathbb{R}^3$ [5, 7]. For lack of a convenient phrase ("sums of distances" is vague), we call $\gamma_1(A)$ the **Euclidean capacity** of A. The unit interval has Euclidean capacity $1/2$. The unit disk (and hence the unit circle) has Euclidean capacity $4/\pi$; notice the inequality [8]

$$\sum_{j<k} |\xi_j - \xi_k| \le n \cot\left(\frac{\pi}{2n}\right)$$

(for which equality holds when $\xi_1, \ldots, \xi_n$ are n^{th} roots of unity). The unit ball (and hence the unit sphere) has Euclidean capacity $4/3$; set $s=-1$, $d=2$ in the formula for s-energy of the d-sphere [11]. We wrote $2/3$ in [47] since sums were divided by n^2 rather than $2/(n(n-1))$. Higher order asymptotics for the latter are conjectured in [48].

It is remarkable that no numerical results for Euclidean capacity (akin to those for Newtonian capacity) of the unit cube, square, equilateral triangle or regular tetrahedron appear yet to exist. A starting point for a literature search might be [49–54].

[1] S. R. Finch, Integer Chebyshev constant, *first volume*, pp. 268–272.

[2] T. Ransford, Computation of logarithmic capacity, *Comput. Methods Funct. Theory* 10 (2010) 555–578; MR2791324 (2012c:65047).

[3] W. K. Hayman, *Subharmonic Functions*. v. II, Academic Press, 1989, pp. 429–436; MR1049148 (91f:31001).

[4] P. Pyrih, Logarithmic capacity is not subadditive – a fine topology approach, *Comment. Math. Univ. Carolin.* 33 (1992) 67–72; MR1173748 (93g:31002).

[5] G. Pólya and G. Szegö, Über den transfiniten Durchmesser (Kapazitätskonstante) von ebenen und räumlichen Punktmengen, *J. Reine Angew. Math.* 165 (1931) 4–49.

[6] L. Carleson, *Selected Problems on Exceptional Sets*, Van Nostrand Co., 1967, pp. 37–39; MR0225986 (37 #1576).

[7] J. S. Brauchart, Optimal discrete Riesz energy and discrepancy, *Unif. Distrib. Theory* 6 (2011) 207–220; MR2904048.

[8] L. Fejes Tóth, On the sum of distances determined by a pointset, *Acta Math. Acad. Sci. Hungar.* 7 (1956) 397–401; MR0107212 (21 #5937).

[9] Q. Rajon, T. Ransford and J. Rostand, Computation of capacity via quadratic programming, *J. Math. Pures Appl.* 94 (2010) 398–413; MR2719894 (2011h:31011).

[10] N. S. Landkof, *Foundations of Modern Potential Theory*, Springer-Verlag, 1972, pp. 160–177; MR0350027 (50 #2520).

[11] J. S. Brauchart and P. J. Grabner, Distributing many points on spheres: minimal energy and designs, *J. Complexity* 31 (2015) 293–326; arXiv:1407.8282; MR3325677.

[12] C. J. Bouwkamp, A simple method of calculating electrostatic capacity, *Physica* 24 (1958) 538–542; MR0074301 (17,563c).

[13] V. P. Pyati, Capacitance of a spindle, *Proc. IEEE* 61 (1973) 505–506.

[14] I. B. Simonenko and A. A. Čekulaeva, The capacity of a condenser that consists of infinite bands (in Russian), *Izvestija Vysših Učebnyh Zavedeniĭ. Èlektromehanika* 4 (1972) 362–365; MR0304684 (46 #3816).

[15] D. Karp, Capacity of a condenser whose plates are circular arcs, *Complex Var. Theory Appl.* 50 (2005) 103–122; arXiv:math/0604033; MR2122748 (2005j:31001).

[16] V. N. Dubinin and D. Karp, Capacities of certain plane condensers and sets under simple geometric transformations, *Complex Var. Elliptic Equ.* 53 (2008) 607–622; MR2421820 (2010e:31001).

[17] E. J. Routh, *A Treatise on Analytical Statics*, v. II, Cambridge Univ. Press, 1922, pp. 219–224.

[18] W. R. Smythe, *Static and Dynamic Electricity*, McGraw-Hill, 1950, pp. 111–123.

[19] C. Snow, *Formulas for Computing Capacitance and Inductance*, National Bureau of Standards circular #544, US Government Printing Office, 1954.

[20] Y. Y. Jossel, E. S. Kochanov and M. G. Strunskij, *The Calculation of Electrical Capacitance*, Energija, 1969 (in Russian); Engl. transl. in Tech. Report AD0727198, US Defense Technical Information Center.

[21] G. Pólya, Estimating electrostatic capacity, *Amer. Math. Monthly* 54 (1947) 201–206; MR0020175 (8,514a).

[22] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Princeton Univ. Press, 1951, pp. 76–78; MR0043486 (13,270d).

[23] D. K. Reitan and T. J. Higgins, Calculation of the electrical capacitance of a cube, *J. Appl. Phys.* 22 (1951) 223–226; MR0040505 (12,704g).

[24] D. K. Reitan and T. J. Higgins, Accurate determination of the capacitance of a thin rectangular plate, *AIEE Trans. Part I. Communication and Electronics* 75 (1957) 761–766.

[25] B. Noble, The numerical solution of the singular integral equation for the charge distribution on a flat rectangular lamina, *Symposium on the Numerical Treatment of Ordinary Differential Equations, Integral and Integro-Differential Equations*, Rome, 1960, Birkhäuser, 1960, pp. 530–543; MR0129149 (23 #B2186).

[26] B. Noble, Some applications of the numerical solution of integral equations to boundary value problems, *Conference on Applications of Numerical Analysis*, Dundee, 1971, Lect. Notes in Math. 228, Springer-Verlag, 1971, pp. 137–154; MR0353711 (50 #6194).

[27] D. Greenspan and E. Silverman, Approximate solution of the exterior Dirichlet problem and the calculation of electrostatic capacity, *Riv. Mat. Univ. Parma* 7 (1966) 305–314; MR0233522 (38 #1843).

[28] D. Greenspan. Resolution of classical capacity problems by means of a digital computer, *Canad. J. Phys.* 44 (1966) 2605–2614.

[29] E. E. Okon and R. F. Harrington, A method of computing the capacitance of flat discs of arbitrary shape, *Electromagnetics* 1 (1981) 229–241.

[30] E. F. Kuester, Explicit approximation for the static capacitance of a microstrip patch of arbitrary shape, *J. Electromagnetic Waves Appl.* 2 (1988) 103–135.

[31] C. S. Brown, Capacity of the regular polyhedra, *Comput. Math. Appl.* 20 (1990) 43–56; MR1051753 (91f:65198).
[32] H. J. Wintle, The capacitance of the regular tetrahdron and equilateral triangle, *J. Electrostatics* 26 (1991) 115–200.
[33] H. J. Wintle, Maxwell and the boundary element method: A historical puzzle, *IEEE Electrical Insulation Magazine* 14 (1998) 23–25.
[34] H. J. Wintle, The capacitance of the cube and square plate by random walk methods, *J. Electrostatics* 62 (2004) 51–62.
[35] E. Goto, Y. Shi and N. Yoshida, Extrapolated surface charge method for capacity calculation of polygons and polyhedra, *J. Comput. Phys.* 100 (1992) 105–115.
[36] F. H. Read, Improved extrapolation technique in the boundary element method to find the capacitances of the unit square and cube, *J. Comput. Phys.* 133 (1997) 1–5.
[37] F. H. Read, Capacitances and singularities of the unit triangle, square, tetrahedron and cube, *COMPEL* 23 (2004) 572–578.
[38] E.-W. Bai and K. E. Lonngren, On the capacitance of a cube, *Computers and Electrical Engineering* 28 (2002) 317–321.
[39] M. Mascagni and N. A. Simonov, The random walk on the boundary method for calculating capacitance, *J. Comput. Phys.* 195 (2004) 465–473.
[40] S. Ghosh and A. Chakrabarty, Estimation of capacitance of different conducting bodies by the method of rectangular subareas, *J. Electrostatics* 66 (2008) 142–146.
[41] C.-O. Hwang, M. Mascagni, and T. Won, Monte Carlo methods for computing the capacitance of the unit cube, *Math. Comput. Simulation* 80 (2010) 1089–1095; MR2610070.
[42] J. Helsing and K.-M. Perfekt, On the polarizability and capacitance of the cube, *Appl. Comput. Harmon. Anal.* 34 (2013) 445–468; arXiv:1203.5997; MR3027912.
[43] J. B. Hubbard and J. F. Douglas, Hydrodynamic friction of arbitrarily shaped Brownian particles, *Phys. Rev. E* (1993) R2983–R2986.
[44] J. F. Douglas, H.-X. Zhou and J. B. Hubbard, Hydrodynamic friction and the capacitance of arbitrarily shaped objects, *Phys. Rev. E* (1994) 5319–5331.
[45] S. S. Vinogradov, P. D. Smith and E. D. Vinogradova, *Canonical Problems in Scattering and Potential Theory.* v. I, *Canonical Structures in Potential Theory*, Chapman & Hall/CRC, 2001, pp. 291–309; MR1851593 (2002f:78009).
[46] A. G. Ramm, A variational principle and its application to estimating the electrical capacitance of a perfect conductor, *Amer. Math. Monthly* 120 (2013) 747–751; MR3096484.
[47] S. R. Finch, Tammes' constants, *first volume*, pp. 508–511.
[48] J. S. Brauchart, D. P. Hardin and E. B. Saff, The next-order term for optimal Riesz and logarithmic energy asymptotics on the sphere, *Recent Advances in Orthogonal Polynomials, Special Functions, and Their Applications*, ed. J. Arvesú and G. López Lagomasino, Amer. Math. Soc., 2012, pp. 31–61; arXiv:1202.4037; MR2964138.
[49] E. Hille, Some geometric extremal problems, *J. Austral. Math. Soc.* 6 (1966) 122–128; MR0198349 (33 #6507).
[50] R. Alexander and K. B. Stolarsky, Extremal problems of distance geometry related to energy integrals, *Trans. Amer. Math. Soc.* 193 (1974) 1–31; MR0350629 (50 #3121).
[51] B. Meyer, Inequalities for the generalized transfinite diameter, *Rocky Mountain J. Math.* 11 (1981) 267–270; MR0619675 (82g:31011).
[52] E. Alarcon and K. B. Stolarsky, Local diameters of compact planar sets, *J. Geom.* 55 (1996) 5–22; MR1377433 (96m:52003).

[53] X. Hou and J. Shao, Spherical distribution of 5 points with maximal distance sum, *Discrete Comput. Geom.* 46 (2011) 156–174; arXiv:0906.0937; MR2794362 (2012e:52005).

[54] R. Nerattini, J. S. Brauchart and M. K.-H. Kiessling, Optimal N-point configurations on the sphere: "magic" numbers and Smale's 7th problem, *J. Stat. Phys.* 157 (2014) 1138–1206; arXiv:1307.2834; MR3277763.

3.13 Aissen's Convex Set Function

Let D be a bounded open convex set in the plane and let C denote the boundary of D. For each $p \in D$ and $q \in C$, let h_{pq} be the Euclidean distance from p to the support line (tangent line) to D at q. Let ds_q denote the line element at q. It is known that [1, 2]

$$\text{arclength of } C = \int_C ds_q,$$

$$\text{area of } D = \frac{1}{2}\int_C h_{pq}\, ds_q \quad \text{(independent of } p\text{)},$$

$$r(D) = \text{inradius of } D = \max_{p\in D}\min_{q\in C} h_{pq}$$

where r is the radius of the largest disk contained by D [3]. The boundary of such a disk is called an incircle; its center is called an incenter. Aissen [1, 2] studied the function

$$B(D) = \min_{p\in D}\int_C h_{pq}^{-1}\, ds_q$$

and deduced that the optimizing point p corresponds to an incenter of D if D is a triangle, parallelogram, regular polygon or ellipse. (We are careful to say "an incenter" rather than "the incenter": a suitably elongated parallelogram has infinitely incircles, all of the same radius. By contrast, the incenter for an arbitrary triangle is unique.) This is a remarkable feature of B. It is natural to wonder whether the same is true for an arbitrary convex set.

The simplest counterexample is a trapezoid with vertices $(\pm 1, 1)$, $(\pm 3, -1)$, for which the optimizing point p has x-coordinate 0 (by symmetry) but y-coordinate > 0. More generally, examine the trapezoid with vertices $\left(\pm\left(\sqrt{2}-1+t\right), 1\right)$, $\left(\pm\left(\sqrt{2}+1+t\right), -1\right)$ where $t \geq 0$ is fixed. The integral within B becomes a sum of four ratios:

$$2\left(\frac{\sqrt{2}-1+t}{1-y} + \frac{\sqrt{2}+1+t}{1+y} + \frac{2}{\sqrt{2}+t+x-y} + \frac{2}{\sqrt{2}+t-x-y}\right)$$

each of the form sidelength/distance. As an instance, the rightmost side has equation

$$v - \frac{1}{\sqrt{2}} = -u + \left(\frac{1}{\sqrt{2}} + t\right)$$

in the uv-plane, that is, $u + v - \sqrt{2} - t = 0$. The distance from the point (x, y) to the line is

$$\frac{\left|x + y - \sqrt{2} - t\right|}{\sqrt{1^2 + 1^2}} = \frac{\sqrt{2} + t - x - y}{\sqrt{2}}$$

and the sidelength is $\sqrt{2^2 + 2^2} = 2\sqrt{2}$. Forming a ratio gives the final term in the sum. Differentiating the sum with respect to x, we see that $x = 0$ is necessary for minimization. The derivative with respect to y is more complicated. In the special case $t = 0$, each of the trapezoidal sides is tangent to the unit circle, thus $y = 0$. If instead $t = 2 - \sqrt{2}$, then the inradius is still 1 but $y \approx 0.116257$ is the unique positive zero of the quartic $y^4 + 8y^3 - 25y^2 + 20y - 2$. If instead $t = 3 - \sqrt{2}$, we have $y \approx 0.130385$ (increasing). If instead $t = 4 - \sqrt{2}$, we have $y \approx 0.110399$ (decreasing). As $t \to \infty$, we have $y \to 0^+$. Aissen's optimizing point appears not to be associated with the trapezoidal incenter except at the extremes $t = 0$, $t = \infty$.

Another counterexample – the half-disk $0 \le v \le \sqrt{1 - u^2}$ – comes from [1, 2]. Again $x = 0$ follows by symmetry. The integral within B here becomes

$$\frac{2}{y} + \frac{2\arcsin(y) + \pi}{\sqrt{1 - y^2}}$$

and is minimized when $y = 0.5432763603... > 1/2$. The value of B itself is $8.7915361561...$. Such values play a role in estimating hard physical quantities like torsional rigidity P in terms of area A [4]. For the half-disk, P turns out to be known exactly and the lower bound [5]

$$0.2975567820... = \frac{\pi}{2} - \frac{4}{\pi} = P \ge A^2 B^{-1} = \frac{(\pi/2)^2}{8.7915361561...} \approx 0.280$$

is excellent.

Returning to geometry, let d_{pq} simply be the Euclidean distance from p to q. Clearly

$$R(D) = \text{circumradius of } D = \min_{p \in D} \max_{q \in C} d_{pq}$$

where R is the radius of the smallest disk containing D [3]. The boundary of such a disk is called a circumcircle; its center is called a circumcenter. The circumcenter for an arbitrary convex set is unique. We wonder if a "dual" to Aissen's function can be defined and what its interplay with the circumcenter for various D might be.

[1] M. I. Aissen, *On the Estimation and Computation of Torsional Rigidity*, Ph.D. thesis, Stanford Univ., 1952.
[2] M. I. Aissen, A set function defined for convex plane domaines, *Pacific J. Math.* 8 (1958) 383–399 MR0123968 (23 #A1289).
[3] S. R. Finch, Circumradius-inradius constants, *first volume*, pp. 534–537.
[4] S. R. Finch, Expected lifetimes and inradii, *this volume*, §4.7.
[5] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, 1951, pp. 14–15, 71, 98–100, 254–255, 272; MR0043486 (13,270d).

3.14 Condition Numbers of Matrices

Let A be a real $n \times n$ matrix and let

$$\|A\| = \sup_{\|x\|=1} \|A\,x\|$$

denote its Euclidean operator norm (often called the 2-norm). If A is nonsingular, then its **condition number** $\kappa(A)$ is defined by

$$\kappa(A) = \|A\| \left\|A^{-1}\right\| = \frac{\sigma_1(A)}{\sigma_n(A)}$$

where $\sigma_1 \geq \sigma_1 \geq \ldots \geq \sigma_n \geq 0$ are the singular values of A. The σs constitute lengths of the semi-axes of the hyperellipsoid $E = \{A\,x : \|x\| = 1\}$ in n-dimensional space; thus κ measures elongation of E at its extreme [1]. The role that κ plays in numerical analysis cannot be overstated: real matrices with large κ are called **ill-conditioned** whereas matrices with small κ are called **well-conditioned**. In a nutshell, κ quantifies the sensitivity of x to pertubations in A and b when solving the linear system $A\,x = b$.

It remains to understand the meaning of "large" versus "small" in this context. Let the entries of A be independent normally distributed random variables with mean 0 and variance 1. Edelman [2] proved that the condition number κ_n satisfies

$$\mathrm{E}\left(\ln(\kappa_n)\right) = \ln(n) + c + o(1)$$

as $n \to \infty$, where

$$c = -\frac{1}{2}\tilde{c} + \ln(2) = 1.5370894353... = \ln(4.6510334182...),$$

$$\begin{aligned}
\tilde{c} &= \int_0^\infty \ln(x)\frac{1+\sqrt{x}}{2\sqrt{x}} \exp\left(-\frac{x}{2} - \sqrt{x}\right) dx \\
&= -2\gamma - 2e^{1/2} \int_1^\infty \frac{1}{y+1} \exp\left(-\frac{1}{2}y^2\right) dy \\
&= -1.6878845096...
\end{aligned}$$

and γ is the Euler–Mascheroni constant [3]. Therefore random dense matrices are well-conditioned, in the sense that κ_n grows only linearly with n.

Let A be the same as before except all superdiagonal entries are zero and all diagonal elements are one. That is, A is a unit lower triangular matrix, all of whose subdiagonal entries are independent $N(0,1)$. Viswanath & Trefethen [4] proved that

$$\begin{aligned}\sqrt[n]{\kappa_n} &\to \exp\left[\frac{1}{2\sqrt{2\pi}}\int_{-\infty}^{\infty}\ln\left(1+x^2\right)\exp\left(-\frac{1}{2}x^2\right)dx\right]\\ &= 1.3056834105...\end{aligned}$$

almost surely as $n\to\infty$. Therefore random unit lower triangular matrices are ill-conditioned, in the sense that κ_n grows exponentially with n. Such behavior is in striking contrast to the linear growth for random dense matrices.

Similar conclusions follow if we replace the normal distribution by, say, the Cauchy distribution with density function

$$\frac{1}{\pi}\frac{1}{1+x^2}$$

for $-\infty<x<\infty$. An exact limiting expression for $\mathrm{E}\left(\ln(\kappa_n/n)\right)$ analogous to that in [2] is unknown, although Monte Carlo simulation suggests that a constant c indeed exists and is close to 7.0. For random unit lower triangular matrices, we have [4]

$$\begin{aligned}\sqrt[n]{\kappa_n} &\to \exp\left[\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\ln\left(1+|x|\right)}{1+x^2}dx\right]\\ &= \exp\left(\frac{\ln(2)}{2}+\frac{2G}{\pi}\right)=2.5337372794...\end{aligned}$$

almost surely as $n\to\infty$, where G is Catalan's constant [5]. An interesting variation arises if we allow the diagonal entries of the latter to be independent Cauchy as well (rather than fixed at unity):

$$\begin{aligned}\sqrt[n]{\kappa_n} &\to \exp\left[\frac{2}{\pi^2}\int_{-\infty}^{\infty}\frac{\ln\left(1+|x|\right)\ln\left(|x|\right)}{x^2-1}dx\right]\\ &= \exp\left(\ln(2)+\frac{7\zeta(3)}{2\pi^2}\right)=3.0630941933...\end{aligned}$$

almost surely as $n\to\infty$, where $\zeta(3)$ is Apéry's constant [6].

We can extend our discussion to complex matrices. Let the real and imaginary parts of entries of A be independent normally distributed random variables with mean 0 and variance 1. From [2], we have

$$\mathrm{E}\left(\ln(\kappa_n)\right) = \ln(n) + d + o(1)$$

as $n \to \infty$, where

$$d = -\frac{1}{2}\tilde{d} + \frac{3}{2}\ln(2) = 0.9817550130... = \ln(2.6691365030...),$$

$$\tilde{d} = \int_0^\infty \ln(x)\frac{1}{2}\exp\left(-\frac{x}{2}\right)dx = \ln(2) - \gamma = 0.1159315156....$$

If we replace the normal distribution by the Cauchy distribution, then simulation suggests that d indeed exists and is close to 6.4.

Finally, let real/imaginary parts of entries of unit lower triangular A be independent normal with mean 0 and variance $1/2$ (different scaling than previously). From [4], we have

$$\begin{aligned}\sqrt[n]{\kappa_n} &\to \exp\left[\frac{1}{4}\int_0^\infty \ln\left(1+\frac{x}{2}\right)\exp\left(-\frac{x}{2}\right)dx\right]\\ &= \exp\left(-\frac{e}{2}\,\mathrm{Ei}(-1)\right) = 1.3473957848...\end{aligned}$$

almost surely as $n \to \infty$, where Ei is the exponential integral [7]. Numerical values when replacing the normal distribution here by the Cauchy distribution (for some choice of scaling) remain open. Other choices of densities are possible (symmetric strictly stable distributions, for example) and corresponding constants would be good to see someday.

[1] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins Univ. Press, 1996, pp. 69–72, 80–82; MR1417720 (97g:65006).

[2] A. Edelman, Eigenvalues and condition numbers of random matrices, *SIAM J. Matrix Anal. Appl.* 9 (1988) 543–560; MR964668 (89j:15039).

[3] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.

[4] D. Viswanath and L. N. Trefethen, Condition numbers of random triangular matrices, *SIAM J. Matrix Anal. Appl.* 19 (1998) 564–581; MR1614019 (99b:65061).

[5] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.

[6] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[7] S. R. Finch, Euler-Gompertz constant, *first volume*, pp. 423–428.

3.15 Goddard's Rocket Problem

A rocket lifts off vertically at time $t = 0$. Let $m(t)$ be the mass of the rocket (payload and fuel) and $s(t)$ be the altitude. We wish to choose the thrust $u(t)$ and a final time t_∞ such that the altitude $s(t_\infty)$ is maximized [1]. It is assumed that

$m(t_\infty)/m(0) = 3/5$, that is, 60% of the rocket is payload and 40% is fuel. For convenience, let $m_0 = m(0)$, $m_\infty = m(t_\infty)$ and $s_\infty = s(t_\infty)$. It is further assumed that the Earth is flat and the atmosphere is negligible, thus

$$\frac{d^2s}{dt^2} = \frac{1}{m(t)}u(t) - g, \quad s(0) = 0, \quad \left.\frac{ds}{dt}\right|_{t=0} = 0,$$

$$\frac{dm}{dt} = -\frac{1}{c}u(t),$$

where g is the (constant) acceleration due to gravity. Normalize g to be 1. Fuel consumption is proportional to thrust; set $c = 1/2$. Integrating

$$\frac{d^2s}{dt^2} = -\frac{c}{m(t)}\frac{dm}{dt} - g,$$

we obtain

$$\frac{ds}{dt} = c\ln\left(\frac{m_0}{m(t)}\right) - g\,t,$$

hence

$$0 = c\ln\left(\frac{m_0}{m_\infty}\right) - g\,t_\infty$$

and

$$t_\infty = \frac{c}{g}\ln\left(\frac{m_0}{m_\infty}\right) = 0.255412....$$

Integrating again, we have [2, 3]

$$\begin{aligned} s_\infty &= \int_0^{t_\infty}\left[c\ln\left(\frac{m_0}{m(t)}\right) - g\,t\right]dt \\ &= c\ln\left(\frac{m_0}{m_\infty}\right)t_\infty - \frac{1}{2}g\,t_\infty^2 = \frac{c^2}{2g}\ln\left(\frac{m_0}{m_\infty}\right)^2 = 0.032617... \end{aligned}$$

since, by the calculus of variations, it is optimal to select

$$m(t) = \begin{cases} m_0 & \text{if } t = 0, \\ m_\infty & \text{if } 0 < t \le t_\infty. \end{cases}$$

In words, the rocket will reach maximum altitude if the thrust $u(t)$ is an impulse at $t = 0$ (a special case of a *bang-bang control*). All fuel is used instantaneously; the rocket achieves maximum velocity immediately. For consistency with [4, 5], define

$$v(t) = \frac{1}{c}\frac{ds}{dt} = \begin{cases} 0 & \text{if } t = 0, \\ \ln\left(\frac{m_0}{m_\infty}\right) - \frac{g}{c}t & \text{if } 0 < t \le t_\infty \end{cases}$$

and

$$v_0 = \lim_{t\to 0^+} v(t) = \ln\left(\frac{m_0}{m_\infty}\right) = 0.510825....$$

If there is non-negligible aerodynamic drag, then an interesting tradeoff occurs. High velocity achieved at low altitudes (by an impulsive start) will confront great resistance. It appears that a better strategy would be to save some fuel for intermediate altitudes, but determination of exactly how to execute this is nontrivial.

Replace the first ODE by

$$\frac{d^2 s}{d t^2} = \frac{1}{m(t)}\left[u(t) - W\exp\left(-\alpha\, s(t)\right)\left(\frac{ds}{dt}\right)^2\right] - g,$$

where $W = 310 = (1/2)(620)$ and $\alpha = 500$. In words, air density decreases exponentially with altitude but drag increases quadratically with velocity. Although we cannot solve this nonlinear equation in the same manner as previously, it is remarkable that closed-form expressions for certain quantities even exist. Note that $W/m_\infty = 310/(3m_0/5) = (1550/3)/m_0$. The following discussion is due to Tsien & Evans [4], with follow-on work by Leitmann [5–7].

Let t_1 be the burnout time, that is, the end of powered flight. The optimal t_1 is 0 for travel in a vacuum; $t_1 > 0$ if there is significant drag. The rocket continues to coast upward, without fuel, until time t_∞. Of course $m(t_1) = m_\infty$. Let $s_1 = s(t_1)$ and $v_1 = v(t_1)$. Let

$$\beta = \frac{g}{\alpha\, c^2},\quad \gamma = \sqrt{(1-\beta)^2 + 8\beta},\quad f(x) = \mathrm{Ei}\left(-2\beta\frac{W c^2}{m_\infty g}\exp(-\alpha\, x)\right),$$

$$p(x) = \frac{2x + (1-\beta) - \gamma}{2x + (1-\beta) + \gamma},\qquad q(x,y) = \frac{x^2 + (1-\beta)x - 2\beta}{y^2 + (1-\beta)y - 2\beta},$$

$$r(x) = \frac{x+2}{x^2 + (1-\beta)x - 2\beta},$$

where Ei is the exponential integral [8]. Here is a system of five simultaneous equations, arising from the calculus of variations, that enable us to solve for t_1, s_1, s_∞, v_0, v_1:

$$v_1^2 = -2\beta\exp\left(2\beta\frac{W c^2}{m_\infty g}\exp(-\alpha\, s_1)\right)\left[f(s_\infty) - f(s_1)\right],$$

$$\frac{W c^2}{m_\infty g} v_1^2\,(1 + v_1) = \exp(\alpha\, s_1),$$

$$\alpha\, s_1 = v_1 - v_0 + \frac{\gamma}{2}\ln\left(\frac{p(v_1)}{p(v_0)}\right) + \frac{3+\beta}{2}\ln\left(q(v_1, v_0)\right),$$

$$\frac{g\, t_1}{c} = \ln\left(\frac{v_0}{v_1}\right) + \frac{\gamma}{2}\ln\left(\frac{p(v_1)}{p(v_0)}\right) + \frac{1+\beta}{2}\ln\left(q(v_1, v_0)\right),$$

$$\frac{m_0}{m_\infty} = \beta \frac{W c^2}{m_\infty g} v_0 \left[v_0^2 + (1-\beta) v_0 - 2\beta \right] \left[r(v_0) - r(v_1) \right] \exp(v_0) + \exp\left(v_1 + \frac{g\, t_1}{c} \right).$$

Given the prescribed parameter values, we obtain $t_1 = 0.062642...$, $s_1 = 0.005085...$, $s_\infty = 0.013579...$, $v_0 = 0.102753...$ and $v_1 = 0.277580...$. In particular, s_∞ is smaller than the final altitude $0.032617...$ computed for a vacuum and v_1 is considerably larger than v_0. Finding the thrust at $t = 0$ is equivalent to computing

$$\frac{1}{m_\infty} \lim_{t \to 0^+} m(t) = \frac{m_0}{m_\infty} \exp(-v_0) = 1.503915...,$$

that is, approximately 9.8% of the rocket mass is expended at the start. Mass at any time $0 < t < t_1$ can be found via replacing $r(v_0)$ in the right-hand side of the fifth equation by $r(v(t))$, and then multiplying the whole by $m_\infty \exp(-v(t) - g\,t/c)$. Finding velocity, given $0 < t < t_1$, is done by substituting t_1, v_1 everywhere in the fourth equation by t, v and then solving for v. The trickiest part is calculating t_∞, for which no analogous equation seems to be available. By call to a numerical ODE solver:

$$\frac{d^2 s}{d t^2} + \frac{W}{m_\infty} \exp(-\alpha\, s(t)) \left(\frac{ds}{dt} \right)^2 + g = 0, \quad s(t_1) = s_1, \quad \left. \frac{ds}{dt} \right|_{t = t_1} = c\, v_1,$$

we obtain $t_\infty = 0.192021...$ at which ds/dt vanishes. This, again, is smaller than the final time $0.255412...$ computed for a vacuum. We also confirm numerically that $s(t_\infty) = s_\infty$.

The Earth is, in fact, round – let its radius be 1 – therefore a distance $h(t) = s(t) + 1$ separates the rocket and Earth's center. Replace the first ODE by

$$\frac{d^2 s}{d t^2} = \frac{1}{m(t)} \left[u(t) - W \exp(-\alpha\, s(t)) \left(\frac{ds}{dt} \right)^2 \right] - \frac{g}{(s+1)^2}$$

where W and α are as before. Suppose that $m_0 = 1$. Additional realistic constraints on thrust and dynamic pressure

$$0 \le u(t) \le \frac{7}{2}, \quad q(t) = \frac{1}{2} \rho_0 \exp(-\alpha\, s(t)) \left(\frac{ds}{dt} \right)^2 \le 10$$

make the optimization more difficult, where the parameter $\rho_0 = 12400$ is air density at sea level. A substantial literature exists on the numerical solution of this problem [9–17]; the optimal final time is $t_\infty = 0.204055...$ and the optimal final distance is $h_\infty = s_\infty + 1 = 1.012717...$. Figure 3.1 constitutes relevant Matlab graphical output [18, 19], where ε is a penalty parameter. The phase between initial thrust $= 3.5$ and final thrust $= 0$ is known as the *singular arc* [20]. See also [21] for informal history and [22–24] for more examples and techniques. We mention finally control problems involving a missile moving obliquely in a vertical plane, maximizing the horizontal range covered [25] or a spacecraft attempting to make a soft landing on the moon, minimizing fuel consumption [26].

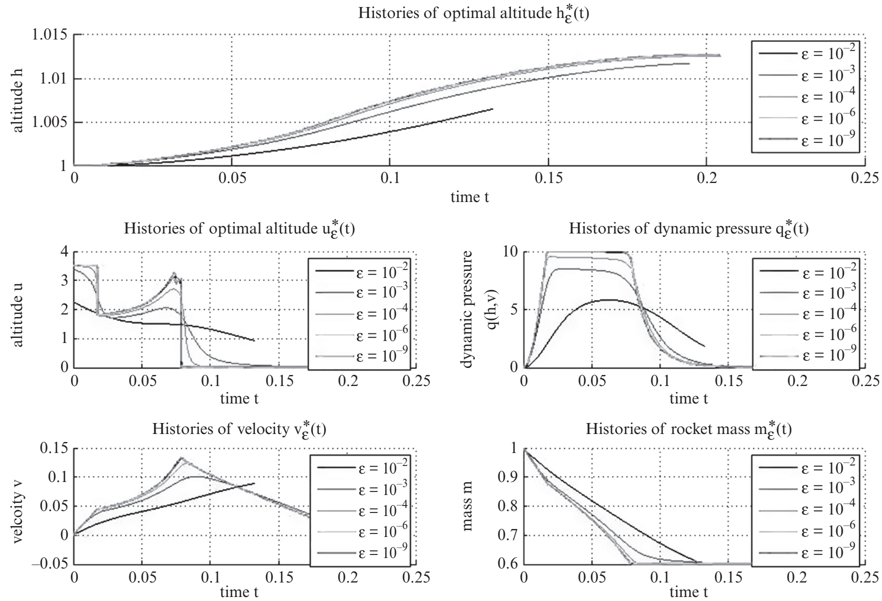

Figure 3.1 Histories of optimal flight characteristics for decreasing values of ε.

Acknowledgment I am thankful to Paul Malisani for his Matlab code [18, 19], which provided Figure 3.1, and for his kind correspondence.

[1] R. H. Goddard, *A Method of Reaching Extreme Altitudes*, Smithsonian Instit., 1919, pp. 5–11.

[2] F. J. Malina and A. M. O. Smith, Flight analysis of the sounding rocket, *J. Aeronautical Sciences*, v. 5 (1938) n. 5, 199–202.

[3] F. Y. M. Wan, *Introduction to the Calculus of Variations and its Applications*, 2^{nd} ed., Chapman & Hall, 1995, pp. 53–55; MR1369576 (97a:49001).

[4] H. S. Tsien, and R. C. Evans, Optimum thrust programming for a sounding rocket, *J. Amer. Rocket Society*, v. 21 (1951) n. 5, 99–107.

[5] G. Leitmann, Optimum thrust programming for high-altitude rockets, *Aeronautical Engineering Rev.*, v. 16 (1957) n. 6, 63–66.

[6] G. Leitmann, A calculus of variations solution of Goddard's problem, *Astronautica Acta* 2 (1956) 55–62; MR0083259 (18,685h).

[7] G. Leitmann, A note on Goddard's problem. *Astronautica Acta* 3 (1957) 237–240.

[8] S. R. Finch, Euler-Gompertz constant, *first volume*, pp. 423–428.

[9] H. J. Oberle, Numerical computation of singular control functions in trajectory optimization problems, *J. Guidance, Control, and Dynamics* 13 (1990) 153–159.

[10] P. Tsiotras and H. J. Kelley, Drag-law effects in the Goddard problem, *Automatica J. IFAC* 27 (1991) 481–490; MR1104508.

[11] P. Tsiotras and H. J. Kelley, Goddard problem with constrained time of flight, *J. Guidance, Control, and Dynamics* 15 (1992) 289–296.

[12] H. Seywald and E. M. Cliff, Goddard problem in presence of a dynamic pressure limit, *J. Guidance, Control, and Dynamics* 16 (1993) 776–781.

[13] K. Graichen and N. Petit, Solving the Goddard problem with thrust and dynamic pressure constraints using saturation functions, *Proc. 17th World Congress, International Federation of Automatic Control*, Seoul, 2008, ed. M. J. Chung and P. Misra, pp. 14301–14306.
[14] E. Cristiani and P. Martinon, Initialization of the shooting method via the Hamilton-Jacobi-Bellman approach, *J. Optim. Theory Appl.* 146 (2010) 321–346; arXiv:0910.0521; MR2679665 (2011g:49050).
[15] M. S. Aronna, J. F. Bonnans and P. Martinon, A shooting algorithm for optimal control problems with singular arcs, *J. Optim. Theory Appl.* 158 (2013) 419–459; arXiv:1206.0839; MR3084385.
[16] P. Malisani, F. Chaplais and N. Petit, An interior penalty method for optimal control problems with state and input constraints of nonlinear systems, *Optim. Control Appl. Methods* 37 (2016) 3–33; MR3457202.
[17] M. Gerdts, *Optimal Control of ODEs and DAEs*, de Gruyter, 2012, pp. 4–5; MR2883534.
[18] P. Malisani, Goddard problem Matlab code, 2012.
[19] P. Malisani, *Pilotage dynamique de l'énergie du bâtiment par commande optimale sous contraintes utilisant la pénalisation intérieure*, Ph.D. thesis, Ecole Nationale Supérieure des Mines de Paris, 2012.
[20] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*, rev. ed., Hemisphere, 1975, pp. 253–257; MR0446628 (56 #4953).
[21] R. D. Rugescu, Goddard's 85 years optimal ascent problem finally solved, *Proc. 57th Internat. Astronautical Congress*, Valencia, 2006, pp. 10073–10084; *History of Rocketry and Astronautics,* Proc. 40th History Sympos., Internat. Academy of Astronautics, ed. M. Freeman, 2012, pp. 337–355.
[22] H. Maurer, Numerical solution of singular control problems using multiple shooting techniques, *J. Optim. Theory Appl.* 18 (1976) 235–257; MR0408246 (53 #12011).
[23] G. Fraser-Andrews, Numerical methods for singular optimal control, *J. Optim. Theory Appl.* 61 (1989) 377–401; MR1000312.
[24] G. Vossen, Switching time optimization for bang-bang and singular controls, *J. Optim. Theory Appl.* 144 (2010) 409–429; MR2581113 (2011a:49075).
[25] G. Leitmann, *The Calculus of Variations and Optimal Control: An Introduction*, Plenum Press, 1981, pp. 3–5, 125–130, 150–155, 161–164, 226–237, 247–249; MR0641031 (84m:49002).
[26] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, 1975, pp. 21–22, 28–33, 35–37; MR0454768 (56 #13016).

3.16 Swing-Up Control of a Pendulum

A pendulum is a bob of mass m, attached to a frictionless pivot point via a massless rod of length ℓ. The bob is free to swing from side to side in a vertical plane. Let g denote the acceleration due to gravity. Let θ denote the angle between the rod and a vertical axis. The pendulum has two equilibrium positions, a stable one at $\theta = 0$ (bottom) and an unstable one at $\theta = \pi$ (top). Assume further that we apply a torque τ to the pendulum, increasing θ (counterclockwise motion) when $\tau > 0$. Let τ be constrained by $|\tau| \leq \tau_0$. The angular equation of motion is [1, 2]

$$I \frac{d^2\theta}{ds^2} + m\, g\, \ell \sin(\theta) = \tau,$$

where $I = m\,\ell^2$ is the moment of inertia and s is time. Define non-dimensional parameters

$$t = \sqrt{\frac{m\,g\,\ell}{I}}\,s, \quad u = \frac{\tau}{\tau_0}, \quad \kappa = \frac{\tau_0}{m\,g\,\ell},$$

so that

$$\frac{d\theta}{dt} = \frac{d\theta}{ds}\frac{ds}{dt} = \sqrt{\frac{I}{m\,g\,\ell}}\frac{d\theta}{ds}, \quad \frac{d^2\theta}{dt^2} = \sqrt{\frac{I}{m\,g\,\ell}}\frac{d^2\theta}{ds^2}\frac{ds}{dt} = \frac{I}{m\,g\,\ell}\frac{d^2\theta}{ds^2}, \quad \frac{\tau}{m\,g\,\ell} = \kappa\,u$$

and hence

$$\frac{d^2\theta}{dt^2} + \sin(\theta) = \kappa\,u$$

subject to $|u| \le 1$. We shall first solve a simple problem with $u = 0$ before allowing more complicated controls in our study. For simplicity, let $\omega = d\theta/dt$.

Let $\kappa = 1$ for now. Let $\theta = \pi/2$ and $\omega = 0$ at $t = 0$. Under these initial conditions and the assumption that $u = 0$ for all t, the pendulum swings down due to gravity alone. What is the angular velocity when $\theta = 0$? Here an exact formula exists:

$$\theta(t) = -2\arcsin\left(\frac{1}{\sqrt{2}}\,\mathrm{sn}\left(t - K\left(\frac{1}{\sqrt{2}}\right), \frac{1}{\sqrt{2}}\right)\right),$$

where $K(x)$ is the complete elliptic integral of the first kind and $\mathrm{sn}(x, y)$ is one of the Jacobi elliptic functions [3]. Solving $\theta(t) = 0$ gives [4]

$$t = K\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{4\sqrt{\pi}}\Gamma\left(\frac{1}{4}\right)^2 = 1.8540746773...$$

and substituting this value into $\omega(t)$ gives $-\sqrt{2} = -1.4142135623...$ [5]. A simple outcome as such is possible only because $u = 0$.

Assume either that $u = 1$ for all t or that $u = -1$ for all t. Given initial conditions $\theta(t_0) = \theta_0$ and $\omega(t_0) = \omega_0$, we have

$$\omega\frac{d\omega}{d\theta} = \frac{d^2\theta}{dt^2} = -\sin(\theta) \pm \kappa,$$

hence

$$\frac{1}{2}\omega^2 = \cos(\theta) \pm \kappa\,\theta + c, \quad c = \frac{1}{2}\omega_0^2 - \cos(\theta_0) \mp \kappa\,\theta_0$$

hence

$$\left(\frac{d\theta}{dt}\right)^2 = \omega^2 = \omega_0^2 + 2\left[\cos(\theta) - \cos(\theta_0) \pm \kappa\,\theta \mp \kappa\,\theta_0\right]$$

hence

$$\frac{|d\theta|}{\sqrt{\omega_0^2 + 2\left[\cos(\theta) - \cos(\theta_0) \pm \kappa\,\theta \mp \kappa\,\theta_0\right]}} = dt.$$

Define

$$T_+(\theta_0,\omega_0,t_0,\kappa;\theta)=t_0+\left|\int\limits_{\theta_0}^{\theta}\frac{d\varphi}{\sqrt{\omega_0^2+2\left[\cos(\varphi)-\cos(\theta_0)+\kappa\,\varphi-\kappa\,\theta_0\right]}}\right|$$

to be the time to reach θ, corresponding to $u=1$ and

$$T_-(\theta_0,\omega_0,t_0,\kappa;\theta)=t_0+\left|\int\limits_{\theta_0}^{\theta}\frac{d\psi}{\sqrt{\omega_0^2+2\left[\cos(\psi)-\cos(\theta_0)-\kappa\,\psi+\kappa\,\theta_0\right]}}\right|$$

to be the time to reach θ, corresponding to $u=-1$. For example, $T_-(\pi/2,0,0,1;0)=1.2794771227...$ is the time required for the pendulum to swing down due to both gravity and a clockwise unit torque. This is unsurprisingly less than the time 1.854... calculated for gravity alone. As another example, $T_+(0,0,0,1;\pi/2)=2.1000505566...$ is the time required for the pendulum to swing halfway up due to a counterclockwise unit torque. This is greater than the preceding since here we are working against gravity. These constants are unrecognizable, as are the associated velocities $\omega_-=-2.2675080272...$ and $\omega_+=1.0684533932...$ obtained using a nonlinear ODE solver.

A more challenging problem is as follows [6–8]. Given $(\theta_0,\omega_0)=(0,0)$, what is the unique strategy to drive the pendulum to $(\theta,\omega)=(\pi,0)$ via a bang-bang control $u=\pm1$ with one switching? The solution is to initially apply $u=1$ until the precise time t_1 when

$$(\theta,\omega)=\left(\frac{\pi}{2}+\frac{1}{\kappa},\sqrt{\kappa\,\pi-2\sin\left(\frac{1}{\kappa}\right)}\right)$$

and subsequently apply $u=-1$ until the precise time t_∞ when $(\theta,\omega)=(\pi,0)$. See Figure 3.2. For example, if $\kappa=1$, then

$$t_1=T_+\left(0,0,0,\kappa;\frac{\pi}{2}+\frac{1}{\kappa}\right)=3.0063538276...,$$

$$t_\infty=T_-\left(\frac{\pi}{2}+\frac{1}{\kappa},\sqrt{\kappa\,\pi-2\sin\left(\frac{1}{\kappa}\right)},t_1,\kappa;\pi\right)=4.0300186879...,$$

but this is valid only since $\cos(\varphi)-1+\kappa\,\varphi>0$ for all $0<\varphi<\pi$. The minimizing value $\varphi_{\min}$ on the left-hand side of the inequality is $\pi-\arcsin(\kappa)$. After substituting $\varphi_{\min}$ into the expression, we solve

$$1-\pi\,\kappa+\sqrt{1-\kappa^2}+\kappa\arcsin(\kappa)=0$$

and obtain $\kappa=0.7246113537...$ as the smallest number for which t_1 is well-defined. Both this number and a related quantity $\pi-\arcsin(\kappa)=2.3311223704...$ appear in [9] in connection not with *swing-up* control, but

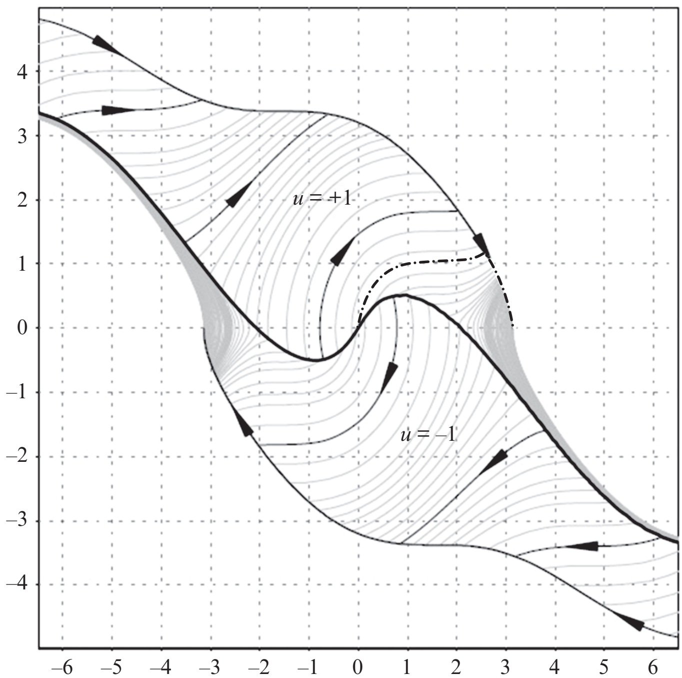

Figure 3.2 Phase portrait (θ on horizontal axis, ω on vertical axis) for $\kappa = 1$ from [6]; start at $(0,0)$, switching at $(2.570..., 1.207...)$, end at $(\pi, 0)$.

rather with *damping* (from unstable equilibrium position to stable). By contrast, the inequality $\cos(\psi) + 1 - \kappa\,\psi + \kappa\,\pi > 0$ does not impose any additional restrictions on κ.

If $\kappa = 1/2$, then we need to consider bang-bang controls $u = \pm 1$ with two switchings. Infinitely many strategies exist by which $u = -1$ is applied for $0 < t < t_1$, $u = 1$ is applied for $t_1 < t < t_2$, $u = -1$ is applied for $t_2 < t < t_\infty$ and required initial/terminal conditions for (θ, ω) are satisfied. Of these, there is a unique strategy with minimal t_∞; see Figure 3.3. It is remarkable that optimality is achieved by first allowing $\omega < 0$ (clockwise motion), seemingly out of the way, before simultaneously reversing torque and exploiting gravity to push $\omega > 1.3$. Omitting the first stage would lead to the pendulum falling far short of $(\theta, \omega) = (\pi, 0)$.

If $\kappa = 3/4$, then both a one-switching strategy and a minimal two-switching strategy exist. For the former, the required time is $t_{\infty,1} = 6.5690173615...$; for the latter, it is $t_{\infty,2} = 5.8397....$ The motion with two switchings is faster:

$$\frac{t_{\infty,1} - t_{\infty,2}}{t_{\infty,2}} \approx 12.5\%,$$

but a motion with three (or more) switchings cannot improve upon $t_{\infty,2}$. We write $N_{3/4}(0,0) = 2$, where $(0,0)$ is the initial point and it is understood that the

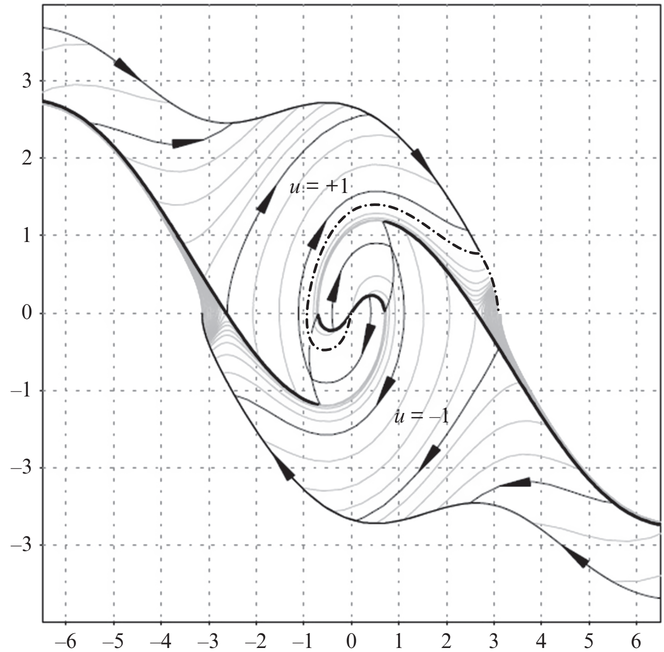

Figure 3.3 Phase portrait (θ on horizontal axis, ω on vertical axis) for $\kappa = 1/2$ from [6]; start at $(0,0)$, switchings at $(-0.877..., -0.394...)$ & $(2.693..., 0.803...)$, end at $(\pi, 0)$.

terminal point is $(\pi, 0)$. In the same way, $N_1(0,0) = 1$ and $N_{1/2}(0,0) = 2$. Define N_κ to be the supremum of $N_\kappa(\theta_0, \omega_0)$ over all θ_0 and ω_0 in the phase space.

Pontryagin's principle guarantees that the optimal control, for any choice of κ, must be of bang-bang type. The complexity of such a control can be characterized by the optimal switching number $N_\varkappa$. Greater knowledge of the function $\kappa \mapsto N_\varkappa$ is therefore desirable. Numerical computations suggest that [10]

$$\inf_{N_\kappa = 1} \kappa \approx 0.80, \quad \inf_{N_\kappa = 2} \kappa \approx 0.44,$$

which are bifurcation values of the parameter κ (analogous to bifurcation values of the parameter a discussed in [11] with regard to quadratic iterates and period doubling). More precise estimates of these values would be good to see someday.

3.16.1 *Damping Control*

This scenario is dual to that for swing-up [8, 9]. Given $(\theta_0, \omega_0) = (\pi, 0)$, what is the unique strategy to drive the pendulum to $(\theta, \omega) = (0,0)$ via a bang-bang control $u = \pm 1$ with one switching? The solution is to initially apply $u = -1$ until the precise time $\tilde{t}_1$ when

$$(\theta,\omega)=\left(\frac{\pi}{2}+\frac{1}{\kappa},-\sqrt{\kappa\pi-2\sin\left(\frac{1}{\kappa}\right)}\right)$$

and subsequently apply $u=1$ until the precise time $\tilde{t}_\infty$ when $(\theta,\omega)=(0,0)$. For example, if $\kappa=1$, then

$$\tilde{t}_1=T_-\left(\pi,0,0,\kappa;\frac{\pi}{2}+\frac{1}{\kappa}\right)=1.0236648603...=t_\infty-t_1,$$

$$\tilde{t}_\infty=T_+\left(\frac{\pi}{2}+\frac{1}{\kappa},-\sqrt{\kappa\pi-2\sin\left(\frac{1}{\kappa}\right)},\tilde{t}_1,\kappa;0\right)=4.0300186879...=t_\infty,$$

but again this is valid only since $\cos(\varphi)-1+\kappa\varphi>0$ for all $0<\varphi<\pi$.

We write $\tilde{N}_1(\pi,0)=1$, where $(\pi,0)$ is the initial point and it is understood that the terminal point is $(0,0)$. One might expect that $\tilde{N}_1(\theta_0,\omega_0)$ to be 1 always, but this is false. By an example given in [12], $\tilde{N}_1(-100,14.16)=2$ and the time improvement is 0.27% (less dramatic than before). The principal bifurcation value here is [13]

$$\inf_{\tilde{N}_\kappa=1}\kappa\approx 1.04$$

and we have asymptotics [14]

$$\inf_{\tilde{N}_\kappa=n}\kappa\sim\frac{1}{n}\frac{G}{2}=\frac{0.9259685259...}{n}$$

as $n\to\infty$, where the constant

$$G=\int\limits_0^\pi\frac{\sin(z)}{z}\,dz=\sum_{j=0}^{\infty}\frac{(-1)^j\pi^{2j+1}}{(2j+1)(2j+1)!}=1.8519370519...$$

is well-known from approximation theory [15]. Although the theory in [14] is devoted to damping, which differs substantially from swing-up, the asymptotic constant $G/2$ evidently remains the same.

More references appear in [16], including mention of a double pendulum and chaos. Time optimal control of such appears to be difficult [17].

Addendum The formula $\theta_1=\pi/2+1/\kappa$ corresponding to one switching has a complicated analog for two switchings [18]. Define

$$\xi(\rho)=\frac{\rho}{2}+\frac{\cos(\rho)-1}{2\kappa},\quad \eta(\rho)=\xi(\rho)+\frac{\pi}{2}+\frac{1}{\kappa},$$

$$F(u,\rho)=\sqrt{-2\kappa(u-\rho)-2\left[\cos(u)-\cos(\rho)\right]}$$

and solve for ρ via the following equation:

$$\frac{1}{F(\xi(\rho),\rho)}+\frac{1}{F(\eta(\rho),\rho)}+\int\limits_{\xi(\rho)}^{\rho}\frac{-\sin(u)+\sin(\rho)}{F(u,\rho)^3}du+\int\limits_{\eta(\rho)}^{\rho}\frac{-\sin(v)+\sin(\rho)}{F(v,\rho)^3}dv=0.$$

In the event $\kappa=1/2$, we obtain $\rho=-0.937739...$ and hence $\theta_1=\xi=-0.877...$, $\theta_2=\eta=2.693....$ In the event $\kappa=3/4$, we obtain $\rho=-0.521237...$ and hence $\theta_1=-0.349...$, $\theta_2=2.554....$ To compute ω_1 and ω_2 involves $F(\xi,\rho)$ and $F(\eta,\rho)$, respectively.

Acknowledgment I am thankful to Sergey Reshmin for employing his software to accurately calculate the cases $\kappa=1/2$ and $\kappa=3/4$, and for his helpful correspondence.

[1] H. Goldstein, C. Poole and J. Safko, *Classical Mechanics*, 3rd ed., Addison-Wesley, 2002, pp. 265-271; MR0575343 (81j:70001).
[2] G. L. Baker and J. A. Blackburn, *The Pendulum. A Case Study in Physics*, Oxford Univ. Press, 2005, pp. 121–152; MR2155144 (2006a:70001).
[3] S. R. Finch, Archimedes' constant, *first volume*, pp. 17–28.
[4] S. R. Finch, Gauss' lemniscate constant, *first volume*, pp. 420–423.
[5] S. R. Finch, Pythagoras' constant, *first volume*, pp. 1–5.
[6] F. L. Chernousko and S. A. Reshmin, Time-optimal swing-up feedback control of a pendulum, *Nonlinear Dynam.* 47 (2007) 65–73; MR2289530 (2007i:70044).
[7] S. A. Reshmin and F. L. Chernousko, Time-optimal control of an inverted pendulum in the form of synthesis (in Russian), *Izv. Ross. Akad. Nauk Teor. Sist. Upr.* (2006) n. 3, 51–62; Engl. transl. in *J. Comput. Syst. Sci. Int.* 45 (2006) 383–394; MR2300991 (2007k:49047).
[8] F. L. Chernousko, I. M. Ananievski and S. A. Reshmin, *Control of Nonlinear Dynamical Systems*, Springer-Verlag, 2008, pp. 367–387.
[9] S. A. Reshmin and F. L. Chernousko, A time-optimal synthesis of a control of a nonlinear pendulum (in Russian), *Izv. Ross. Akad. Nauk Teor. Sist. Upr.* (2007) n. 1, 13–22; Engl. transl. in *J. Comput. Syst. Sci. Int.* 46 (2007) 9–18; MR2321690 (2008d:49025).
[10] S. A. Reshmin, Bifurcation in the time-optimality problem for a second-order nonlinear system (in Russian), *Prikl. Mat. Mekh.* 73 (2009) 562–572; Engl. transl. in *J. Appl. Math. Mech.* 73 (2009) 403–410; MR2584526 (2010m:49004).
[11] S. R. Finch, Feigenbaum-Coullet-Tresser constants, *first volume*, pp. 65–76.
[12] S. A. Reshmin, Properties of the time-optimal control for Lagrangian single-degree-of-freedom systems, *IEEE Trans Automatic Control* 60 (2015) 3350–3355.
[13] S. A. Reshmin, Finding the principal bifurcation value of the maximum control torque in the problem of the synthesis of an optimal control for a pendulum (in Russian), *Izv. Ross. Akad. Nauk Teor. Sist. Upr.* (2008) n. 2, 5–20; Engl. transl. in *J. Comput. Syst. Sci. Int.* 47 (2008) 163–178; MR2456160 (2009k:70030).
[14] A. Ovseevich, Complexity of the minimum-time damping of a physical pendulum, *SIAM J. Control Optim.* 52 (2014) 82–96; MR3148641.
[15] S. R. Finch, Gibbs-Wilbraham constant, *first volume*, pp. 248–250.
[16] S. R. Finch, Lyapunov exponents. II, *this volume*, §4.14.

[17] S. A. Reshmin, Decomposition method in the problem of controlling an inverted double pendulum with the use of one control moment, *Izv. Ross. Akad. Nauk Teor. Sist. Upr.* (2005) n. 6, 28–45; Engl. transl. in *J. Comput. Syst. Sci. Int.* 44 (2005) 861–877; MR2300991 (2007k:49047).
[18] S. A. Reshmin and F. L. Chernousko, Properties of the time-optimal feedback control for a pendulum-like system, *J. Optim. Theory Appl.* 163 (2014) 230–252; MR3260983.

3.17 Zermelo's Navigation Problem

A river is assumed to be of unit width and infinite length. Place the origin on one bank. Let the positive x-axis point upstream (rightward) and the positive y-axis point across the river (upward). The current is characterized by a velocity $-v(y)e_x$, where $e_x = (1, 0)$ and $v(y) \geq 0$ for all $0 \leq y \leq 1$. Hence the water moves from right to left, with possible speed variation depending on the distance from the original bank. By contrast, an ocean is planar and infinite in all directions. While the current in the ocean has velocity depending only on y, there is no sign restriction on $v(y)$.

Our interest is primarily in minimizing travel time, subject to constraints. We will be more specific soon. The shortest path from origin to target is usually not the optimal solution! Our examples are special cases of the work of Zermelo [1–4]; we closely follow [5] for the initial section and [6] for the final.

3.17.1 *Canoe on a River*

Assume that a canoe moves at a constant speed 1 relative to the water. The goal is to reach the point $(0, 1)$ directly across the river from $(0, 0)$. Suppose first the existence of a uniform current, that is, $v(y) = c$ where $0 < c < 1$. A naive strategy is for the canoe's orientation to always be directed at the target. Under this strategy, the canoe is swept downstream somewhat before it overcomes the current and heads upstream. The resulting path (as viewed by a stationary observer from above) is [5, 7, 8]

$$x(y) = \frac{1}{2}\left[(1-y)^{1+c} - (1-y)^{1-c}\right]$$

and the travel time is $1/\left(1 - c^2\right)$. It is better, however, to point the canoe at a constant angle $\arccos(c)$ relative to the x-axis. Such a strategy is optimal: the resulting path is simply $x_c(y) = y$ and the travel time is $1/\sqrt{1 - c^2}$.

Suppose instead the existence of a linear current, that is, $v(y) = b\,y$ where $b > 0$. The naive strategy gives rise to a path [5, 7]

$$x(y) = \frac{1}{2}\left[(1-y)^{1+b}\exp(b\,y) - (1-y)^{1-b}\exp(-b\,y)\right]$$

and a travel time

$$-\frac{1}{2}\left[b^{-b-1}\exp(b)\Gamma(1+b,b)+(-b)^{b-1}\exp(-b)\Gamma(1-b,-b)+\right.$$
$$\left.-(-b)^{b}\exp(-b)\Gamma(-b,0)-b^{-b}\exp(b)\Gamma(b,0)\right],$$

where

$$\Gamma(z,w)=\int_{w}^{\infty}s^{z-1}\exp(-s)ds$$

is the incomplete gamma function. For example, if $b=1/2$, then the travel time is ≈ 1.13546; if $b=9/10$, then the travel time is ≈ 2.61567; as $b\to 1^{-}$, the travel time $\to\infty$. Again, it is better to point the canoe at some other angle $\theta(y)$ relative to the x-axis (here the angle is time-varying). The optimal strategy gives rise to a path [5]

$$\begin{aligned} x_{\ell}(y) \;=\; -\frac{1}{2b}\Bigg\{&\lambda\sqrt{\lambda^2-1}+(b\,y-\lambda)\sqrt{(b\,y+\lambda)^2-1}\\ &+\ln\left[\frac{\lambda+\sqrt{\lambda^2-1}}{b\,y+\lambda+\sqrt{(b\,y+\lambda)^2-1}}\right]\Bigg\}\end{aligned}$$

and a travel time

$$\tau=\frac{1}{b}\left(\sqrt{(b+\lambda)^2-1}-\sqrt{\lambda^2-1}\right),$$

where the parameter $\lambda=\lambda(b)$ is chosen so that $x_{\ell}(1)=0$. For example, if $b=1/2$, then the parameter value is ≈ 3.76109 and the travel time is ≈ 1.03275; if $b=1$, then $\lambda\approx 1.60647$ and $\tau\approx 1.14973$; if $b=3/2$, then $\lambda\approx 1.02830$ and $\tau\approx 1.38836$. Figure 3.4 resembles a graph in [9], reflected across the vertical axis. Define $\beta=1.6626273716...$ to be the largest quantity b for which $\lambda(b)\geq 1$, that is, the solution of the equation

$$(1-b)\sqrt{b(2+b)}+\ln\left(1+b+\sqrt{b(2+b)}\right)=0.$$

The corresponding travel time is $\sqrt{1+2/\beta}=1.4842221390....$ If $b>\beta$, then the canoe cannot overcome the current to reach the target (directly opposite the origin); it necessarily will be swept downstream a finite nonzero distance. For $0<b\leq\beta$, the optimizing angle is given by

$$\theta(y)=\arccos\left(\frac{1}{b\,y+\lambda}\right)$$

and thus, for example, $\theta(0)\approx 51.5°$ and $\theta(1)\approx 67.4°$ if $b=1$.

Suppose instead the existence of a "reverse" linear current, that is, $v(y)=b(1-y)$ where $b>0$. The naive strategy gives rise to a path [7]

$$x(y)=-(1-y)\sinh(b\,y)$$

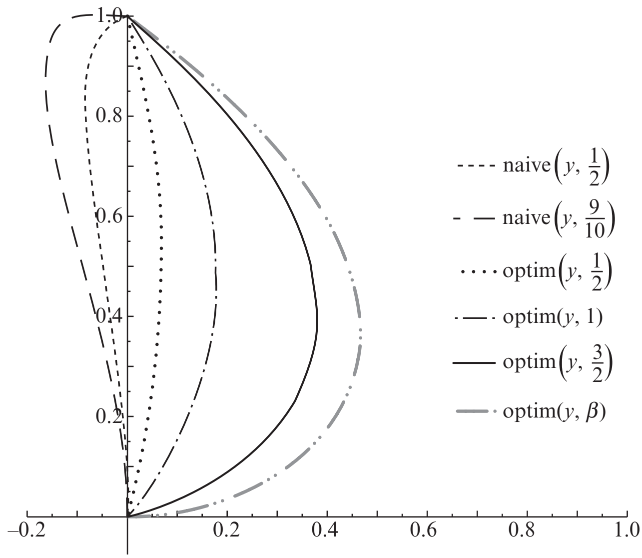

Figure 3.4 Path of a canoe traveling from $(0,0)$ to $(0,1)$ assuming a right-to-left linear current with coefficient b. Curves in $x<0$ show paths for which the canoe's heading is always at the target. Curves in $x>0$ show solutions to Zermelo's navigation problem.

and a travel time $\sinh(b)/b$, which is well-defined for all finite b. The optimal path is $-x_\ell(1-y)$, where x_ℓ was prescribed earlier for a "forward" linear current. The same travel time τ and threshold β apply here as before. Evidently the Euler–Lagrange approach does not work for large b. We do not know an optimal path to the target when $b>\beta$, yet the naive strategy provides a perfectly admissible path *whatever* the current. Resolving this issue seems to be open.

Suppose finally the existence of a parabolic current, that is, $v(y)=4a\,y(1-y)$ where $a>0$. The naive strategy gives rise to a path [5]

$$x(y)=-(1-y)\sinh\left(2a\,y^2\right)$$

and a travel time

$$\frac{1}{4}\sqrt{\frac{\pi}{2a}}\left(\operatorname{erf}\left(\sqrt{2a}\right)+\operatorname{erfi}\left(\sqrt{2a}\right)\right),$$

where $\operatorname{erfi}(s)=-i\,\operatorname{erf}(i\,s)$ is the imaginary error function [10]. Again, this expression is well-defined for all finite a. Let $\alpha=1.148590538...$ be the solution of the equation

$$\int_0^1 \frac{1-4a\,y(1-y)\,[4a\,y(1-y)+1]}{\sqrt{[4a\,y(1-y)+1]^2-1}}=0.$$

It is possible to compute the optimal path $x_p(y)$ and travel time for small a. As before, however, we do not know an optimal path to the target when $a > \alpha$, even though the naive path is admissible for all large a.

3.17.2 *Ship on an Ocean*

For the sake of consistency with [6], let the current be $-y\, e_x$ for all $-\infty < y < \infty$. Hence the water moves from left to right in the lower half plane, is motionless on the horizontal axis, and moves from right to left in the upper half plane. Assume that a ship moves at a constant speed 1 relative to the water. The goal is to reach the point $(0,0)$, given that the ship starts at (x_0, y_0). Suppose that $x_0 = 0$ and $y_0 = -1$. This scenario is essentially the same as the reverse linear current discussed earlier, except the direction of flow is opposite to before and now there is no restriction against the ship venturing beyond the target. There is no shoreline to block passage. Figure 3.5 depicts the optimal path (as viewed by a stationary observer from above), along with initial angle 180° – 67.4° and final angle 180° – 51.5° relative to the x-axis. The optimal time is the same as before.

Suppose instead that $x_0 = 0$ and $y_0 = -1.86$. The powerful current at the onset sweeps the ship considerably farther to the right than in the preceding example; see Figure 3.6. It is optimal for the ship to venture slightly into the region $y > 0$,

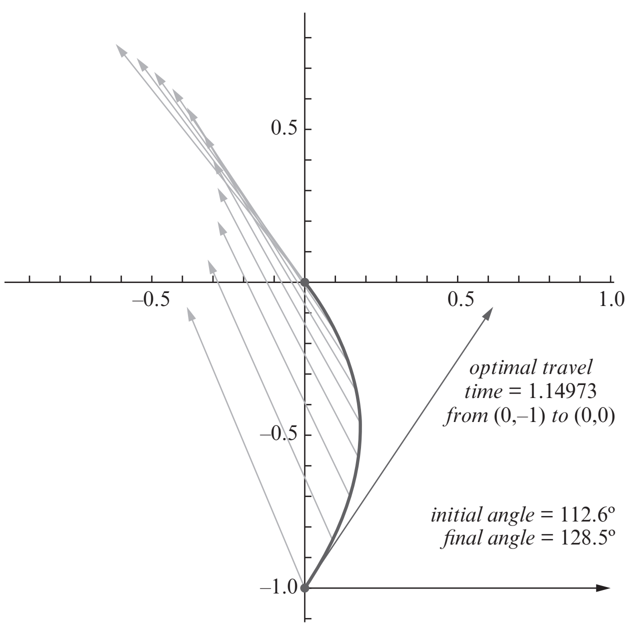

Figure 3.5 Optimal path of a ship traveling from $(0,-1)$ to $(0,0)$ assuming a left-to-right current with coefficient 1. The trajectory is the curve (in bold) whereas the heading vectors point to the northwest (roughly). The current vector at $(0,-1)$ points to the east; the tangent vector at $(0,-1)$ is the sum of the two arrows.

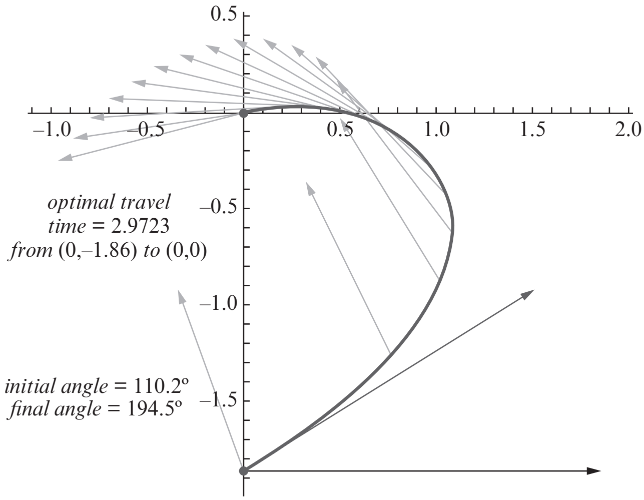

Figure 3.6 Optimal path of a ship traveling from $(0,-1.86)$ to $(0,0)$ assuming a left-to-right current with coefficient 1.

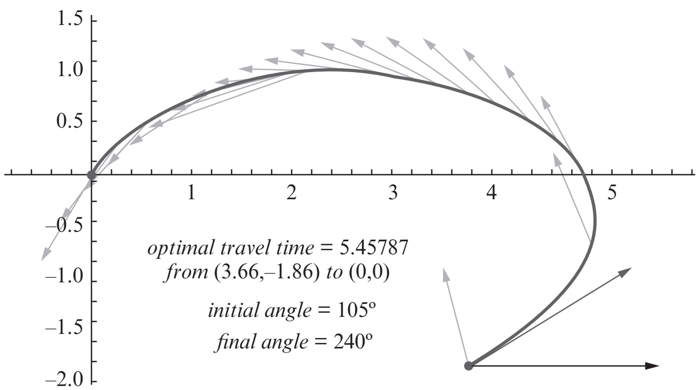

Figure 3.7 Optimal path of a ship traveling from $(3.66,-1.86)$ to $(0,0)$ assuming a left-to-right current with coefficient 1.

taking advantage of the leftward current to bring it back to $x=0$. The optimal time is ≈ 2.9723.

Suppose finally that $x_0 = 3.66$ and $y_0 = -1.86$. A rough approximation of the travel time is found by summing the preceding $\tau \approx 2.9723$ and the time 3.66 for unimpeded travel along the horizontal axis, yielding ≈ 6.6323. Substantial improvement is possible. Figure 3.7 illustrates that the optimal strategy is to penetrate deeply into the region $y > 0$, circling back with an optimal time ≈ 5.45787.

Let us give equations underlying the preceding oceanic results. It is convenient to use θ as the independent variable here. Define θ_0 to be the initial angle and θ_1 to be the final angle. Define

$$f(\theta,\theta_1) = -\frac{1}{2}\left\{\sec(\theta_1)\left(\tan(\theta_1)-\tan(\theta)\right)-\tan(\theta)\left(\sec(\theta_1)-\sec(\theta)\right) + \ln\left[\frac{\sec(\theta_1)+\tan(\theta_1)}{\sec(\theta)+\tan(\theta)}\right]\right\},$$

$$g(\theta,\theta_1) = -\left(\sec(\theta_1)-\sec(\theta)\right).$$

Then the equations $x_0 = f(\theta_0,\theta_1)$, $y_0 = g(\theta_0,\theta_1)$ jointly determine values for θ_0, θ_1 and, more generally, $x = f(\theta,\theta_1)$, $y = g(\theta,\theta_1)$ parametrically represent the optimal path in the plane. Further, the travel time is computed via $\tau = \tan(\theta_1) - \tan(\theta_0)$. Such examples also appear in [11–13]. In the case of a non-smooth wind/water field, a purely numerical approach in [13] suffices to obtain curvilinear plots and optimal travel times.

3.17.3 *Details*

We return to the canoe on a river. Let the velocity u of the canoe relative to the water be $u = \sqrt{1-q^2}\,e_x + q\,e_y$ (which is possible since $|u| = 1$). The canoe's absolute velocity is hence $u - v\,e_x$ (where v is the water speed, a function of distance y alone). We wish to minimize travel time [3, 5]

$$\int_0^1 \frac{1}{q(y)}dy$$

subject to the constraint

$$\int_0^1 \frac{\sqrt{1-q(y)^2}-v(y)}{q(y)}dy = 0$$

(because there is no net lateral displacement). To do this, we seek a stationary point of

$$\int_0^1 \left(\frac{1}{q(y)} - \frac{1}{\lambda}\frac{\sqrt{1-q(y)^2}-v(y)}{q(y)}\right)dy = \int_0^1 h(q,y)dy,$$

where λ is an undetermined multiplier. The Euler–Lagrange equation in this case is simply $dh/dq = 0$, that is,

$$-\frac{1}{q^2} + \frac{1}{\lambda}\left(\frac{\sqrt{1-q^2}-v}{q^2} + \frac{1}{\sqrt{1-q^2}}\right) = 0,$$

that is,

$$q(y) = \sqrt{1 - \frac{1}{(v(y) + \lambda)^2}}.$$

Substituting back into the constraint yields

$$\int_0^1 \frac{1 - v(y)\,(v(y) + \lambda)}{\sqrt{(v(y) + \lambda)^2 - 1}}\,dy = 0.$$

Once λ is known, the optimal path is computed via

$$x(y) = \int_0^y \frac{\sqrt{1 - q(r)^2} - v(r)}{q(r)}\,dr = \int_0^y \frac{1 - v(r)\,(v(r) + \lambda)}{\sqrt{(v(r) + \lambda)^2 - 1}}\,dr$$

with associated angle function

$$\theta(y) = \arccos\left(\sqrt{1 - q(y)^2}\right) = \arccos\left(\frac{1}{v(y) + \lambda}\right).$$

For the uniform current ($v = c$) we have $\lambda = -c + 1/c$, thus $\theta = \arccos(c)$ identically [5]. For the linear and parabolic currents, no closed-form expression for λ is available. Figure 3.8 depicts the analog of Figure 3.4 corresponding to $v(y) = 4a\,y(1 - y)$. In order to obtain this, we had no choice but to numerically evaluate the definite integral underlying $x(y)$.

It is easier to compute the naive path [5]. The vertical component of the vector difference $(0, 1) - (x, y)$, normalized to have unit length, is

$$q(y) = \frac{1 - y}{\sqrt{x^2 + (1 - y)^2}},$$

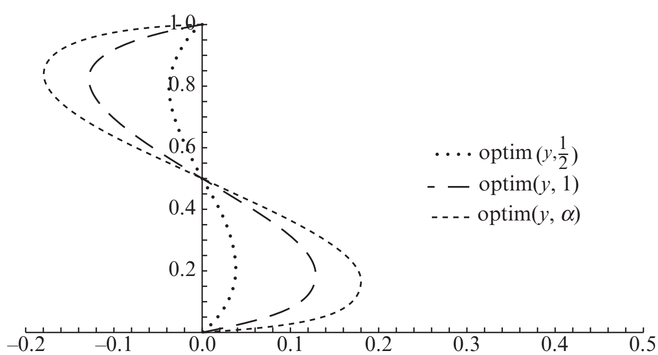

Figure 3.8 Path of a canoe traveling from $(0, 0)$ to $(0, 1)$ assuming a right-to-left parabolic current with coefficient $4a$.

which leads to a differential equation

$$\frac{dx}{dy} = \frac{\sqrt{1-q(y)^2} - v(y)}{q(y)} = \frac{-x - \sqrt{x^2 + (1-y)^2}\, v(y)}{1-y}$$

assuming that $x < 0$ and $0 < y < 1$.

We return to the ship on an ocean. A special case $d\theta/dt = -\cos(\theta)^2 dv/dy$ of Zermelo's general equation for optimality [3, 6] implies that

$$\frac{dt}{d\theta} = \sec(\theta)^2,$$

since $v(y) = -y$, so $t_1 - t = \tan(\theta_1) - \tan(\theta)$. From $dy/dt = \sin(\theta)$, it follows that

$$\frac{dy}{d\theta}\cos(\theta)^2 = \frac{dy}{d\theta}\frac{d\theta}{dt} = \sin(\theta)$$

giving $dy/d\theta = \sec(\theta)\tan(\theta)$ and so $y = -\sec(\theta_1) + \sec(\theta)$. From $dx/dt = \cos(\theta) - y$, it follows that

$$\frac{dx}{d\theta}\cos(\theta)^2 = \frac{dx}{d\theta}\frac{d\theta}{dt} = \cos(\theta) - y = \cos(\theta) + \sec(\theta_1) - \sec(\theta)$$

giving

$$\frac{dx}{d\theta} = \sec(\theta) + \sec(\theta_1)\sec(\theta)^2 - \sec(\theta)^3.$$

We have

$$\int \left(\sec(\theta) + C\sec(\theta)^2 - \sec(\theta)^3\right) d\theta \;=\; C\tan(\theta) - \frac{1}{2}\sec(\theta)\tan(\theta) + \frac{1}{2}\ln\left(\sec(\theta) + \tan(\theta)\right)$$

and so the desired expression for x is true.

Fraser [3] remarked that "the canonical problems of the calculus of variations – the isoperimetric problem, the hanging chain, the brachistochrone – go back centuries and appear at an early stage in the history of the subject." Zermelo's navigation problem (like Goddard's rocket problem [14]) is "somewhat unusual in providing a simple and signature example of very recent vintage, arising from technological developments of the twentieth century."

Addendum Define $\omega = -9/10$. Let the ship be in an ocean with purely rotational current $(x, y) \mapsto (-\omega\, y, \omega\, x)$, starting at location $(x_0, y_0) = (\sqrt{3}/2, 1/2)$ and ending at location $(x_1, y_1) = (0, 1)$. The ship could simply travel (at constant speed 1) with the flow, requiring a travel time $10\pi/3$, but this is unnecessarily lengthy. The optimal path is provably of the form [15]

$$x(t, \psi_0) = x_0\cos(\omega\, t) - y_0\sin(\omega\, t) + t\cos(\psi_0 + \omega\, t),$$

$$y(t, \psi_0) = x_0\sin(\omega\, t) + y_0\cos(\omega\, t) + t\sin(\psi_0 + \omega\, t)$$

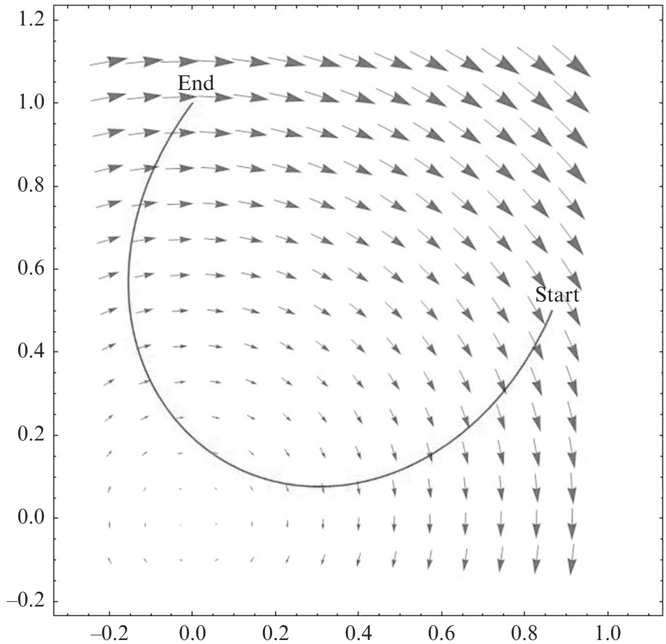

Figure 3.9 Optimal path of a ship traveling from $(\sqrt{3}/2, 1/2)$ to $(0, 1)$ assuming a purely rotational current with coefficient $-9/10$.

and requires a travel time $t_1 \approx 1.974938$, where t_1 and the initial heading $\psi_0 \approx 3.506716 \approx 200.9°$ jointly satisfy the equations $x(t_1, \psi_0) = x_1, y(t_1, \psi_0) = y_1$. Figure 3.9 (like Figure 3.7) indicates that an optimal path may contain a subarc where the distance between ship and target temporarily increases.

The examples discussed here have all been in the plane. Zermelo's solution can be extended to the surface of a sphere [4, 16, 17] and arises in the study of Riemannian manifolds [18–20]. Some generalized problems appear in [21–24]. An especially far-flung application is the implementation of information processing tasks in controlled quantum systems [25–28].

Acknowledgment I am thankful to Marian Mureoşan [12] for writing and sharing helpful Mathematica® code that provided Figures 3.5, 3.6 and 3.7.

[1] E. Zermelo, Über die Navigation in der Luft als Problem der Variationsrechnung, *Jahresbericht. Deutsch. Math.-Verein.* 39 (1930) 44–48.
[2] E. Zermelo, Über das Navigationsproblem bei ruhender oder veränderlicher Windverteilung, *Z. Angew. Math. Mech.* 11 (1931) 114–124.

[3] E. Zermelo, *Collected Works*, v. II., ed. H.-D. Ebbinghaus and A. Kanamori, Springer-Verlag, 2013, pp. 672–721 (introductory note by C. G. Fraser); MR3137671.

[4] H.-D. Ebbinghaus, *Ernst Zermelo. An Approach to His Life and Work*, 2nd ed., Springer-Verlag, 2015; pp. 159–161; MR3381798.

[5] J. Talbot, Going against the flow: finding the optimal path, *Europ. J. Phys.* 31 (2010) 205–213.

[6] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control*, rev. ed., Hemisphere, 1975, pp. 77–80, 97–99; MR0446628 (56 #4953).

[7] M. J. O'Shea, Crossing a river in a canoe – how complicated can it get? *Europ. J. Phys.* 31 (2010) 857–862.

[8] O. Bolina, Difficult problems having easy solutions, *Phys. Educ.* 34 (1999) 148–149.

[9] G. C. Hays, A. Christensen, S. Fossette, G. Schofield, J. Talbot and P. Mariani, Route optimization and solving Zermelo's navigation problem during long distance migration in cross flows, *Ecology Letters* 17 (2014) 127–260.

[10] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[11] A. Enes and W. Book, Blended shared control of Zermelo's navigation problem, *Proc. American Control Conference (ACC)*, 2010, Baltimore, IEEE, pp. 4307-4312.

[12] M. Mureşan, On Zermelo's navigation problem with Mathematica, *J. Appl. Funct. Anal.* 9 (2014) 349–355; MR3156218.

[13] H. J. Oberle and R. Rosendahl, On singular arcs in nonsmooth optimal control, *Control Cybernet.* 37 (2008) 429–450; MR2472884 (2009j:49043).

[14] S. R. Finch, Goddard's rocket problem, *this volume*, §3.15.

[15] L. Techy, Optimal navigation in planar time-varying flow: Zermelo's problem revisited, *Intelligent Service Robotics* 4 (2011) 271–283.

[16] K. J. Arrow, On the use of winds in flight planning, *J. Meteorology* 6 (1949) 150–159; *Collected Papers of Kenneth J. Arrow*, v. 5: *Production and Capital*, Harvard Univ. Press, 1985, pp. 1-24.

[17] M. R. Jardin and A. E. Bryson, Methods for computing minimum-time paths in strong winds, *J. Guidance, Control, and Dynamics* 35 (2012) 165–171.

[18] D. Bao, C. Robles and Z. Shen, Zermelo navigation on Riemannian manifolds, *J. Differential Geom.* 66 (2004) 377–435; arXiv:math/0311233; MR2106471 (2005k:58023).

[19] S.-S. Chern and Z. Shen, *Riemann-Finsler Geometry*, World Scientific, 2005, pp. 20–25; MR2169595 (2006d:53094).

[20] U. Serres, On the curvature of two-dimensional optimal control systems and Zermelo's navigation problem (in Russian), *Sovrem. Mat. Prilozh*, v. 21, *Geom. Zadachi Teor. Upr.* (2004) 134–153; Engl. transl. in *J. Math. Sci.* 135 (2006) 3224–3243; MR2157923 (2006f:93036).

[21] E. D. Poppleton, A generalization of Zermelo's condition for the minimum-time flight path, *Z. Angew. Math. Mech.* 49 (1969) 363–365.

[22] S. Zlobec, Partly convex programming and Zermelo's navigation problems, *J. Global Optim.* 7 (1995) 229–259; MR1355655 (96g:90055).

[23] E. Bakolas and P. Tsiotras, Optimal synthesis of the Zermelo-Markov-Dubins problem in a constant drift field, *J. Optim. Theory Appl.* 156 (2013) 469–492; MR3019324.

[24] B. Li, C. Xu, K. L. Teo and J. Chu, Time optimal Zermelo's navigation problem with moving and fixed obstacles, *Appl. Math. Comput.* 224 (2013) 866–875; MR3127670.

[25] B. Russell and S. Stepney, Zermelo navigation and a speed limit to quantum information processing, *Phys. Rev. A* 90 (2014) 012303; arXiv:1310.6731.

[26] B. Russell and S. Stepney, Zermelo navigation in the quantum brachistochrone, *J. Phys. A* 48 (2015) 115303; arXiv:1409.2055; MR3317198.

[27] D. C. Brody and D. M. Meier, Solution to the quantum Zermelo navigation problem, *Phys. Rev. Lett.* 114 (2015) 100502; arXiv:1409.3204.

[28] D. C. Brody and D. M. Meier, Elementary solution to the time-independent quantum navigation problem, *J. Phys. A* 48 (2015) 055302; arXiv:1409.6669; MR3314700.

4

Probability and Stochastic Processes

4.1 Hammersley's Path Process

The following is a generalization of a process introduced by Hammersley [1, 2]. Fix three parameters $\lambda > 0$, $\alpha_+ \geq 0$ and $\alpha_- \geq 0$. Let $P(\mu)$ denote a Poisson random variable with mean μ. Baik & Rains [3] constructed a set of points S in the unit square $[0, 1] \times [0, 1]$ according to three rules:

- $P(\lambda^2)$ points are selected uniformly inside $(0, 1) \times (0, 1)$
- $P(\alpha_+\lambda)$ points are selected uniformly on the open bottom edge $(0, 1) \times \{0\}$
- $P(\alpha_-\lambda)$ points are selected uniformly on the open left edge $\{0\} \times (0, 1)$.

These rules are independently executed. No points are selected from the closed top and right edges, nor is the origin $(0, 0)$ allowed.

Consider any sequence of distinct points of the form

$$(0, 0),\ (s_1, t_1),\ (s_2, t_2),\ \ldots\ ,\ (s_n, t_n),\ (1, 1)$$

where each $(s_k, t_k) \in S$, $1 \leq k \leq n$, and n is arbitrary. For convenience, define $(s_0, t_0) = (0, 0)$ and $(s_{n+1}, t_{n+1}) = (1, 1)$. Define such a point sequence to be an **up/right path** if, for any $k \geq 1$, we have $s_{k-1} \leq s_k$ and $t_{k-1} \leq t_k$. Hence an up/right path joins points of S in a continuous, piecewise linear manner with line segments of slope m_k, $0 \leq m_k \leq \infty$, attaching (s_{k-1}, t_{k-1}) and (s_k, t_k) for all k.

Of all up/right paths determined by S, there is (at least) one with a maximum number n of points. Call this number N_λ. (This is usually referred to as a *length* in the literature. Of course, it also depends implicitly on α_+ and α_-.) What can be said about the probability distribution of N_λ as $\lambda \to \infty$?

A special case of the above is the longest increasing subsequence problem [4], achieved when $\alpha_+ = \alpha_- = 0$. Its solution will be folded into the formulas we give shortly for the general problem. This turns out to be related to the polynuclear growth (PNG) model in physics due to Prähofer & Spohn [5–7], but we cannot discuss such topics now.

When $0 \le \alpha_+ \le 1$ and $0 \le \alpha_- \le 1$ are fixed, the following formulas hold [3]:

$$\lim_{\lambda\to\infty} \mathrm{P}\left(\frac{N_\lambda - 2\lambda}{\lambda^{1/2}} \le x\right) = \begin{cases} F_{\mathrm{GUE}}(x) & \text{if } \alpha_+ < 1 \text{ and } \alpha_- < 1, \\ F_{\mathrm{GOE}}(x)^2 & \text{if } \alpha_+ = 1,\ \alpha_- < 1 \text{ or } \alpha_+ < 1,\ \alpha_- = 1, \\ F_0(x) & \text{if } \alpha_+ = 1 \text{ and } \alpha_- = 1, \end{cases}$$

where the distribution functions $F_{\mathrm{GUE}}(x)$, $F_{\mathrm{GOE}}(x)$ and $F_0(x)$ will be defined shortly. Also, when $\alpha_+ > 1$ or $\alpha_- > 1$, we have

$$\lim_{\lambda\to\infty} \mathrm{P}\left(\frac{N_\lambda - (\alpha+\alpha^{-1})\lambda}{\sqrt{\alpha-\alpha^{-1}}\,\lambda^{1/2}} \le x\right) = \begin{cases} \Phi(x) & \text{if } \alpha_+ \ne \alpha_-, \\ \Phi(x)^2 & \text{if } \alpha_+ = \alpha_-, \end{cases}$$

where $\alpha = \max\{\alpha_+, \alpha_-\}$ and $\Phi(x)$ is the standard normal distribution function [8]. We provide moments corresponding to these distributions (and more) in Tables 4.1 and 4.2; computations were performed by Prähofer [9]. The functions $F_{\mathrm{GUE}}(x)$, $F_{\mathrm{GOE}}(x)$ and $F_{\mathrm{GSE}}(x)$ were first discovered by Tracy & Widom [10–12],

Table 4.1 *Moments of GUE, GOE, GUE² and GOE²*

	F_{GUE}
mean	$-1.7710868074...$
variance	$0.8131947928... = (0.9017731382...)^2$
skewness	$0.2240842036...$
kurtosis	$0.0934480876...$
	F_{GOE}
mean	$-1.2065335745... = 2^{2/3}(-0.7600685240...)$
variance	$1.6077810345... = (1.2679830576...)^2 = 2^{4/3}(0.6380483264...)$
skewness	$0.2934645240...$
kurtosis	$0.1652429384...$
	F^2_{GUE}
mean	$-1.2633181526...$
variance	$0.6066887541... = (0.7789022750...)^2$
skewness	$0.3290093382...$
kurtosis	$0.2254319482...$
	F^2_{GOE}
mean	$-0.4936399332...$
variance	$1.2320144032... = (1.1099614422...)^2$
skewness	$0.3917246784...$
kurtosis	$0.3086329720...$

Table 4.2 *Moments of GSE and other distributions*

	F_{GSE}	
mean	$-2.3068848932... = \frac{1}{\sqrt{2}}(-3.2624279028...)$	
variance	$0.5177237207... = (0.7195302083...)^2 = \frac{1}{2}(1.0354474415...) = \frac{1}{2}(1.0175693792...)^2$	
skewness	0.1655094943	
kurtosis	0.0491951565	
	F_0	
mean	0	
variance	$1.1503944782... = 2^{2/3}(0.7247031094...) = (0.8104567006...)^{-2/3}$	
skewness	$0.3594116897...$	
kurtosis	$0.2891570248...$	
	Φ	Φ^2
mean	0	$\frac{1}{\sqrt{\pi}} = 0.5641895835...$
variance	1	$1 - \frac{1}{\pi} = 0.6816901138... = (0.8256452711...)^2$
skewness	0	$\frac{4-\pi}{2(\pi-1)^{3/2}} = 0.1369487673...$
kurtosis	0	$\frac{2(\pi-3)}{(\pi-1)^2} = 0.0617443154...$

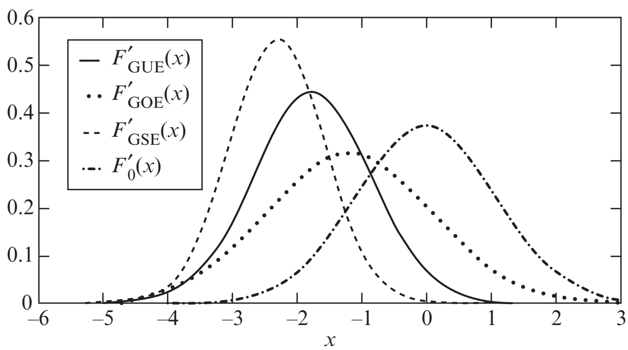

Figure 4.1 The Tracy–Widom density functions, as well as $F_0'(x)$.

whereas $F_0(x)$ arose more recently [3]. See Figure 4.1 for the associated density plots.

Let $u(x)$ be the solution of the Painlevé II differential equation:

$$u''(x) = 2u(x)^3 + xu(x), \quad u(x) \sim -\frac{1}{2\sqrt{\pi}} x^{-1/4} \exp\left(-\frac{2}{3}x^{3/2}\right) \text{ as } x \to \infty,$$

and define

$$U(x) = -\int_x^\infty u(r)\,dr, \quad V(x) = -\int_x^\infty v(r)\,dr$$

where

$$v(x) = -\int_x^\infty u(r)^2\,dr.$$

The largest eigenvalue of a random complex Hermitian matrix, when generated according to the Gaussian Unitary Ensemble (GUE) probability law and properly normalized, has distribution function

$$F_{\mathrm{GUE}}(x) = \exp(-V(x)) \quad \text{(often denoted as the case } \beta = 2\text{)}.$$

More details appear in §4.1.1. Replacing Hermitian matrices by real symmetric matrices, we obtain the Gaussian Orthogonal Ensemble (GOE) and corresponding distribution function

$$F_{\mathrm{GOE}}(x) = \exp\left(-\frac{U(x) + V(x)}{2}\right) \quad \text{(often denoted as the case } \beta = 1\text{)}.$$

Likewise, for the Gaussian Symplectic Ensemble (GSE), we have

$$F_{\mathrm{GSE}}(x) = \cosh\left(\frac{U(x)}{2}\right)\exp\left(-\frac{V(x)}{2}\right) \quad \text{(the case } \beta = 4\text{)}.$$

Define also

$$F_0(x) = \left[1 - \left(x + 2u'(x) + 2u(x)^2\right) v(x)\right] \exp\left(-2U(x) - V(x)\right),$$

which does not yet seem to possess a random matrix interpretation. These formulas serve as the computational basis for Tables 4.1 and 4.2, where skewness and kurtotis of a random variable Y are given as

$$\mathrm{Skew}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^3\right]}{\mathrm{Var}(Y)^{3/2}}, \quad \mathrm{Kurt}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^4\right]}{\mathrm{Var}(Y)^2} - 3.$$

For example, if $\alpha_+ < 1$ and $\alpha_- < 1$, it follows that

$$\lim_{\lambda\to\infty} \lambda^{-1/3}(\mathrm{E}(N_\lambda) - 2\lambda) = -1.7710868074...,$$
$$\lim_{\lambda\to\infty} \lambda^{-2/3}\,\mathrm{Var}(N_\lambda) = 0.8131947928...,$$

which generalize results given earlier by Tracy & Widom and Baik, Deift & Johansson [4]. If instead $\alpha_+ = 1$ and $\alpha_- = 1$, we have

$$\lim_{\lambda\to\infty} \lambda^{-1/3}\,\mathrm{Var}(N_\lambda) = 1.1503944782...$$

which is called the Baik–Rains constant in [7].

4.1.1 *GUE/GOE/GSE*

A random complex Hermitian $N \times N$ matrix X belongs to GUE if its (real) diagonal elements x_{jj} and (complex) upper triangular elements $x_{jk} = \xi_{jk} + i\eta_{jk}$ are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ for $1 \le j \le N$ and $\mathrm{Var}(\xi_{jk}) = \mathrm{Var}(\eta_{jk}) = 1$ for $1 \le j < k \le N$. Let λ denote the largest (real) eigenvalue of X and define the normalization [12]

$$\tilde{\lambda} = \frac{N^{1/6}(\lambda - 2\sigma\sqrt{N})}{\sigma}$$

where $\sigma = \sqrt{\mathrm{Var}(x_{jk})} = \sqrt{2}$. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GUE in Table 4.1. A related discussion, involving spacings between adjacent eigenvalues of X and featuring connections to the Riemann Hypothesis, appears in [13].

A random real symmetric $N \times N$ matrix X belongs to GOE if its diagonal elements x_{jj} and upper triangular elements x_{jk} are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ and $\mathrm{Var}(x_{jk}) = 1$. Let $\tilde{\lambda}$ denote the largest (real) eigenvalue of X, normalized as before with $\sigma = 1$ in this case. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GOE in Table 4.1.

A complex Hermitian $2N \times 2N$ matrix is said to be **real quaternionic** [14] if, when viewed as an $N \times N$ matrix X consisting of 2×2 blocks, the diagonal blocks X_{jj} look like

$$X_{jj} = \begin{pmatrix} x_{jj} & 0 \\ 0 & x_{jj} \end{pmatrix}, \quad x_{jj} \in \mathbb{R}$$

and the upper triangular blocks X_{jk} look like

$$X_{jk} = \begin{pmatrix} \xi_{jk} + i\eta_{jk} & \xi'_{jk} + i\eta'_{jk} \\ -\xi'_{jk} + i\eta'_{jk} & \xi_{jk} - i\eta_{jk} \end{pmatrix}.$$

A random real quaternionic matrix X belongs to GSE if the nonzero distinct elements of its diagonal and upper triangular blocks are independently chosen from zero-mean Gaussian distributions with $\mathrm{Var}(x_{jj}) = 2$ and $\mathrm{Var}(\xi_{jk}) = \mathrm{Var}(\eta_{jk}) = \mathrm{Var}(\xi'_{jk}) = \mathrm{Var}(\eta'_{jk}) = 1$. Let $\tilde{\lambda}$ denote the largest (real) eigenvalue of X, normalized as before with $\sigma = 2$ in this case. Then the distribution of $\tilde{\lambda}$ has the moments indicated for GSE in Table 4.2.

Here is an occurrence of $F_{\mathrm{GUE}}(x)^2$: Define a **signed permutation** π to be a bijection from $\{-n, -n+1, \ldots, -2, -1, 1, 2, \ldots, n-1, n\}$ onto itself which satisfies $\pi(-k) = -\pi(k)$ for all k. Tracy & Widom [15, 16] proved that the length L_{2n} of the longest increasing subsequence of a random signed permutation π satisfies

$$\lim_{n \to \infty} \mathrm{P}\left(\frac{L_{2n} - 2\sqrt{2n}}{2^{2/3}(2n)^{1/6}} \le x \right) = F_{\mathrm{GUE}}(x)^2.$$

A nice combinatorial application involving $F_{\mathrm{GSE}}(x)$ or $F_{\mathrm{GSE}}(x)^2$, especially one as simple as this, would be good to find.

Other applications appear in [12, 17, 18]. A d-dimensional analog of Hammersley's original process (with $\alpha_+ = \alpha_- = 0$) appears in [19]: Let S denote a set of $P(\lambda^d)$ points selected uniformly inside the d-dimensional unit cube and N_λ denote the number of points in a maximal chain (totally ordered subset) of S. Define c_d to be $\limsup_{\lambda\to\infty} \mathrm{E}(N_\lambda)/\lambda$. Then it is known that $c_2 = 2$ and $c_\infty = e$, but $2.363 \le c_3 \le 2.366$, $2.514 \le c_4 \le 2.521$, $2.583 \le c_5 \le 2.589$ and $2.607 \le c_6 \le 2.617$. We draw attention finally to the obvious identity [2]:

$$F_{\mathrm{GSE}}(x) = \frac{1}{2}\left(F_{\mathrm{GOE}}(x) + \frac{F_{\mathrm{GUE}}(x)}{F_{\mathrm{GOE}}(x)}\right)$$

and wonder whether a similar identity relating F_0 to other distributions can ever be found.

4.1.2 *Positive Definite/Indefinite*

Among many possible questions, we ask for the probability that a random $N \times N$ matrix, distributed according to GOE, is positive definite. Since

$$\begin{aligned} \mathrm{P}\,(\text{indefinite}) &= 1 - \mathrm{P}\,(\text{positive definite}) - \mathrm{P}\,(\text{negative definite}) \\ &= 1 - 2\mathrm{P}\,(\text{positive definite}), \end{aligned}$$

the answer for indefinite matrices is clear once it is found for positive definite matrices. The joint density for the N unordered (real) eigenvalues of a GOE matrix is [20]

$$\frac{1}{C_N} \prod_{1\le i<j\le N} |\lambda_i - \lambda_j| \cdot \exp\left(-\frac{1}{4}\sum_{k=1}^{N} \lambda_k^2\right)$$

where

$$C_N = N!(2\pi)^{N/2} 2^{N(N+1)/4} \prod_{\ell=1}^{N} \frac{\Gamma(\ell/2)}{\Gamma(1/2)}.$$

A complicated formula associated with the density for the smallest eigenvalue follows, as do the results in Table 4.3 for small N.

Table 4.3 *Probabilities that an $N \times N$ GOE matrix is positive definite/indefinite*

N	positive definite	indefinite
1	$1/2 = 0.5$	0
2	$1/2 - \sqrt{2}/4 \approx 0.1464$	$\sqrt{2}/2 \approx 0.7071$
3	$1/4 - \left(\sqrt{2}/2\right)\pi^{-1} \approx 0.0249$	$1/2 + \sqrt{2}\pi^{-1} \approx 0.9502$
4	$1/4 - \sqrt{2}/16 - (1/2)\pi^{-1} \approx 0.0025$	$1/2 + \sqrt{2}/8 + \pi^{-1} \approx 0.9951$
5	$1/8 - \left(1/3 + \sqrt{2}/24\right)\pi^{-1} \approx 0.0001$	$3/4 + \left(2/3 + \sqrt{2}/12\right)\pi^{-1} \approx 0.9997$

Consider now the quadratic form

$$Q(x_1, x_2, \ldots, x_N) = \sum_{1 \le i \le j \le N} m_{ij} x_i x_j,$$

where the coefficients m_{ij} form the upper triangular portion of a GOE matrix M. Another way of saying M is indefinite is that $Q = 0$ possesses a nonzero solution in $\mathbb{R}^N$. If we constrain the m_{ij} to be integers, what is the probability that $Q = 0$ possesses a nonzero solution in $\mathbb{Z}^N$? The answer is 0 for $1 \le N \le 3$, is the same as the real indefinite case for $N \ge 5$, but is miraculously [21, 22]

$$\left(\frac{1}{2} + \frac{\sqrt{2}}{8} + \frac{1}{\pi}\right) \prod_p \left(1 - \frac{p^3}{4(p+1)^2\,(p^4 + p^3 + p^2 + p + 1)}\right) = 0.9825845607...$$

for $N = 4$. If we replace the GOE distribution by, say, a uniform distribution on $[-1/2, 1/2]$ for each m_{ij}, then the probability becomes 0.97... instead. The structure of the formula – leading coefficient multiplied by prime product – is similar. While the prime product 0.9874362482... remains identical, no closed-form expression is known for the leading coefficient.

[1] J. M. Hammersley, A few seedlings of research, *Proc. Sixth Berkeley Symp. Math. Stat. Probab.*, v. 1, ed. L. M. Le Cam, J. Neyman and E. L. Scott, Univ. of Calif. Press, 1972, pp. 345–394; MR0405665 (53 #9457).

[2] P. J. Forrester, Growth models, random matrices and Painlevé transcendents, *Nonlinearity* 16 (2003) R27–R49; math-ph/0309058; MR2012845 (2004k:82071).

[3] J. Baik and E. M. Rains, Limiting distributions for a polynuclear growth model with external sources, *J. Stat. Phys.* 100 (2000) 523–541; math.PR/0003130; MR1788477 (2001h:82067).

[4] S. R. Finch, Longest subsequence constants, *first volume*, pp. 382–387.

[5] M. Prähofer and H. Spohn, Statistical self-similarity of one-dimensional growth processes, *Physica A* 279 (2000) 342–352; cond-mat/9910273; MR1797145 (2001j:82098).

[6] M. Prähofer and H. Spohn, Universal distributions for growth processes in $1 + 1$ dimensions and random matrices, *Phys. Rev. Lett.* 84 (2000) 4882–4885; cond-mat/9912264.

[7] M. Prähofer and H. Spohn, Exact scaling functions for one-dimensional stationary KPZ growth, *J. Stat. Phys.* 115 (2004) 255–279; cond-mat/0212519; MR2070096 (2005d:82097).

[8] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[9] M. Prähofer, High precision moments corresponding to GUE/GOE/GSE, unpublished note (2004).

[10] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Comm. Math. Phys.* 159 (1994) 151–174; hep-th/9211141; MR1257246 (95e:82003).

[11] C. A. Tracy and H. Widom, On orthogonal and symplectic matrix ensembles, *Comm. Math. Phys.* 177 (1996) 727–754; solv-int/9509007; MR1385083 (97a:82055).

[12] C. A. Tracy and H. Widom, Distribution functions for largest eigenvalues and their applications, *International Congress of Mathematicians*, v. 1, Proc. 2002 Beijing conf., Higher Ed. Press, pp. 587–596; math-ph/0210034; MR1989209 (2004f:82034).

[13] S. R. Finch, Glaisher-Kinkelin constant: GUE hypothesis, *first volume*, pp. 138–145.

[14] P. J. Forrester, *Log-Gases and Random Matrices*, Princeton University Press, 2010; MR2641363 (2011d:82001).

[15] C. A. Tracy and H. Widom, Random unitary matrices, permutations and Painlevé, *Comm. Math. Phys.* 207 (1999) 665–685; MR1727236 (2001h:15019).

[16] A. Borodin, Longest increasing subsequences of random colored permutations, *Elec. J. Combin.* 6 (1999) R13; MR1667453 (2000a:05014).

[17] J. Baik and E. M. Rains, Symmetrized random permutations, *Random Matrix Models and their Applications*, ed. P. Bleher and A. Its, Cambridge Univ. Press, 2001, pp. 1–19; math.CO/9910019; MR1842780 (2002i:82038).

[18] P. L. Ferrari, Random matrices and determinantal processes, *Note di Matematica e Fisica (CERFIM)* 12 (2003) 67–83.

[19] E. Breimer, M. Goldberg, B. Kolstad and M. Magdon-Ismail, On the height of a random set of points in a d-dimensional unit cube, *Experim. Math.* 10 (2001) 583–597; *Algorithm Engineering and Experimentation (ALENEX)*, Proc. 2001 Washington DC conf., ed. A. L. Buchsbaum and J. Snoeyink, Lect. Notes in Comp. Sci. 2153, Springer-Verlag, 2001, pp. 161–171; MR1881759 (2002j:60013).

[20] G. W. Anderson, A. Guionnet and O. Zeitouni, *An Introduction to Random Matrices*, Cambridge Univ. Press, 2010, pp. 51–53; MR2760897 (2011m:60016).

[21] M. Bhargava, J. E. Cremona, T. A. Fisher, N. G. Jones and J. P. Keating, What is the probability that a random integral quadratic form in n variables has an integral zero?, *Internat. Math. Research Notices* (2016) 3828–3848; arXiv:1502.05992; MR3544620.

[22] P. Sebah, Another infinite product over primes, unpublished note (2015).

4.2 Moments of Sums

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables. A huge amount of work has been done on estimating the L_p-norm of the sum of the Xs:

$$\left\| \sum_{k=1}^{n} X_k \right\|_p = \left\{ \mathrm{E}\left(\left| \sum_{k=1}^{n} X_k \right|^p \right) \right\}^{1/p}, \quad p > 0.$$

We first discuss Khintchine's inequality [1], which deals with the Rademacher sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, where

$$\mathrm{P}(\varepsilon_k = 1) = \mathrm{P}(\varepsilon_k = -1) = 1/2 \quad \text{(symmetric Bernoulli distribution)}$$

for each k. It is known that there exist constants A_p, B_p such that the bounds

$$A_p \left(\sum_{k=1}^{n} c_k^2 \right)^{1/2} \le \left\| \sum_{k=1}^{n} c_k \varepsilon_k \right\|_p \le B_p \left(\sum_{k=1}^{n} c_k^2 \right)^{1/2}$$

hold for arbitrary $c_1, c_2, \ldots, c_n \in \mathbb{R}$ and $n \geq 1$. Szarek [2] and Haagerup [3], building on [4–9], proved that the best such constants are

$$A_p = \begin{cases} \|W\|_p & \text{if } 0 < p \leq p_0 \\ \|Z\|_p & \text{if } p_0 < p < 2 \\ 1 & \text{if } 2 \leq p < \infty \end{cases} = \begin{cases} 2^{1/2-1/p} & \text{if } 0 < p \leq p_0 \\ 2^{1/2}\left(\dfrac{\Gamma((p+1)/2)}{\sqrt{\pi}}\right)^{1/p} & \text{if } p_0 < p < 2 \\ 1 & \text{if } 2 \leq p < \infty, \end{cases}$$

$$B_p = \begin{cases} 1 & \text{if } 0 < p \leq 2 \\ \|Z\|_p & \text{if } 2 < p < \infty \end{cases} = \begin{cases} 1 & \text{if } 0 < p \leq 2 \\ 2^{1/2}\left(\dfrac{\Gamma((p+1)/2)}{\sqrt{\pi}}\right)^{1/p} & \text{if } 2 < p < \infty, \end{cases}$$

where $W = 2^{-1/2}(\varepsilon_1 + \varepsilon_2)$, Z is Normal(0, 1), and $p_0 = 1.8474163360\ldots$ is the unique solution of the equation

$$\Gamma\left(\frac{p+1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

in the interval $0 < p < 2$. In words, if $\sum_{k=1}^n c_k^2 = 1$, then $A_1 = 2^{-1/2}$ and $B_1 = 1$ encompass the average of $|\pm c_1 \pm c_2 \pm \cdots \pm c_n|$ taken over all 2^n possible choices of signs. See also [10–15].

A complex analog of Khintchine's inequality deals with the Steinhaus sequence $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, where ε_k is uniformly distributed on the unit circle $\{z : |z| = 1\}$ for each k. We keep notation identical to before, except that we allow $c_1, c_2, \ldots, c_n \in \mathbb{C}$. The best constants A_p, B_p in the inequality

$$A_p\left(\sum_{k=1}^n |c_k|^2\right)^{1/2} \leq \left\|\sum_{k=1}^n c_k\varepsilon_k\right\|_p \leq B_p\left(\sum_{k=1}^n |c_k|^2\right)^{1/2}$$

were conjectured by Haagerup [16] to be

$$A_p = \begin{cases} \|W\|_p & \text{if } 0 < p \leq p_0 \\ \|Z\|_p & \text{if } p_0 < p < 2 \\ 1 & \text{if } 2 \leq p < \infty \end{cases} = \begin{cases} 2^{1/2}\left(\dfrac{\Gamma((p+1)/2)}{\sqrt{\pi}\,\Gamma((p+2)/2)}\right)^{1/p} & \text{if } 0 < p \leq p_0 \\ (\Gamma((p+2)/2))^{1/p} & \text{if } p_0 < p < 2 \\ 1 & \text{if } 2 \leq p < \infty, \end{cases}$$

$$B_p = \begin{cases} 1 & \text{if } 0 < p \leq 2 \\ \|Z\|_p & \text{if } 2 < p < \infty \end{cases} = \begin{cases} 1 & \text{if } 0 < p \leq 2 \\ (\Gamma((p+2)/2))^{1/p} & \text{if } 2 < p < \infty, \end{cases}$$

where $W = 2^{-1/2}(\varepsilon_1 + \varepsilon_2)$, $Z = 2^{-1/2}(U + iV)$ with U, V independent and Normal(0, 1), and $p_0 = 0.4756170089\ldots$ is the unique solution of the equation

$$2^{p/2}\Gamma\left(\frac{p+1}{2}\right) = \sqrt{\pi}\left(\Gamma\left(\frac{p+2}{2}\right)\right)^2$$

in the interval $0 < p < 2$. Here, if $\sum_{k=1}^n |c_k|^2 = 1$, then $A_1 = \sqrt{\pi}/2$ and $B_1 = 1$ encompass an average taken over all "complex signs" rather than only "real signs" as earlier. Sawa [17] announced that he could verify significant portions of Haagerup's conjecture, but only the case $p \approx 1$ was published. See also [14, 15, 18, 19]. We mention as well the following result [20, 21] for which $p = 1$ and n is the parameter of interest:

$$\mathrm{E}\left(\left|\sum_{k=1}^{n} \varepsilon_k\right|\right) = \begin{cases} \dfrac{2}{\pi}\displaystyle\int_0^\infty \dfrac{1-\cos(t)^n}{t^2}\,dt & \text{for the real case} \\ \displaystyle\int_0^\infty \dfrac{1-J_0(t)^n}{t^2}\,dt & \text{for the complex case} \end{cases}$$

where $J_0(t)$ is the zeroth Bessel function of the first kind. On the one hand, we have

$$\frac{2}{\pi}\int_0^\infty \frac{1-\cos(t)^n}{t^2}\,dt = \frac{n!}{2^{n-1} m!(n-m-1)!} \sim \sqrt{\frac{2n}{\pi}}$$

for the real case, where $m = \lfloor (n-1)/2 \rfloor$. On the other hand, the Bessel integral takes on the values 1, $4/\pi$, 1.57459723... and 1.79909248... for $n = 1$, 2, 3 and 4. Keane [22] determined that the third value in this list has the following closed-form expression:

$$\frac{1}{8\pi^3}\Gamma\left(\frac{1}{6}\right)^2 \Gamma\left(\frac{1}{3}\right)^2 + 48\pi\Gamma\left(\frac{1}{6}\right)^{-2} \Gamma\left(\frac{1}{3}\right)^{-2} = 1.5745972375...$$

but the fourth value still remains open.

We next discuss Rosenthal's inequalities [23]:

$$\left\|\sum_{k=1}^{n} X_k\right\|_p \le C_p \cdot \max\left\{\left(\sum_{k=1}^{n} \|X_k\|_p^p\right)^{1/p}, \left\|\sum_{k=1}^{n} X_k\right\|_1\right\}, \quad p \ge 1$$

for nonnegative random variables and

$$\left\|\sum_{k=1}^{n} X_k\right\|_p \le D_p \cdot \max\left\{\left(\sum_{k=1}^{n} \|X_k\|_p^p\right)^{1/p}, \left\|\sum_{k=1}^{n} X_k\right\|_2\right\}, \quad p \ge 2$$

for symmetric random variables (meaning that the distribution of $-X$ is the same as the distribution of X). A variation of the latter inequality arises if we loosen the restrictive hypothesis "symmetric" to "zero mean"; the constant is then denoted E_p rather than D_p. Johnson, Schechtman & Zinn [24] showed that the growth rate of the best constants C_p, D_p, E_p is $p/\ln(p)$ as $p \to \infty$; by contrast, the growth rate for B_p is only $\sqrt{p}$. Subsequent work [25–28] yielded that

$$C_p = \begin{cases} 1 & \text{if } p = 1 \\ 2^{1/p} & \text{if } 1 < p < 2 \\ \|Q\|_p & \text{if } 2 \le p < \infty \end{cases}, \quad D_p = \begin{cases} 1 & \text{if } p = 2 \\ \left(1 + \|Z\|_p^p\right)^{1/p} & \text{if } 2 < p < 4 \\ \|R - S\|_p & \text{if } 4 \le p < \infty \end{cases}$$

where Q is Poisson(1), Z is Normal(0, 1), and R, S are independent Poisson(1/2) variables. It is known that $\|Q\|_m^m = \alpha_m$ and $\|R-S\|_{2m}^{2m} = \beta_m$ for integer m, where $\{\alpha_m\}_{m=1}^{\infty} = \{1, 2, 5, 15, 52, 203, \ldots\}$ is the sequence of Bell numbers [29, 30]

$$\alpha_m = \frac{1}{e}\sum_{j=0}^{\infty}\frac{j^m}{j!} = \frac{d^m}{dx^m}\exp(\exp(x)-1)\bigg|_{x=0}$$

and $\{\beta_m\}_{m=1}^{\infty} = \{1, 4, 31, 379, \ldots\}$ is the sequence

$$\beta_m = \frac{2}{e}\sum_{k=1}^{\infty}\sum_{j=0}^{\infty}\frac{k^{2m}}{j!(j+k)!2^{2j+k}} = \frac{d^{2m}}{dx^{2m}}\exp(\cosh(x)-1)\bigg|_{x=0}.$$

Ibragimov & Sharakhmetov [31] conjectured that

$$E_p = \begin{cases} \left(1+\|Z\|_p^p\right)^{1/p} & \text{if } 2 < p < 4 \\ \|Q-1\|_p & \text{if } 4 \le p < \infty \end{cases}$$

and proved that this is true when $p = 2m$; further, $\|Q-1\|_{2m}^{2m} = \gamma_m$ and $\{\gamma_m\}_{m=1}^{\infty} = \{1, 4, 41, 715, \ldots\}$ is the sequence

$$\gamma_m = \frac{1}{e}\sum_{j=0}^{\infty}\frac{(j-1)^{2m}}{j!} = \frac{d^{2m}}{dx^{2m}}\exp(\exp(x)-x-1)\bigg|_{x=0}.$$

Combinatorial interpretations apply for each of the three sequences: α_n is the number of partitions of an n-element set into blocks; β_n is the number of partitions of a $2n$-element set into blocks, each containing an even number of elements; and γ_n is the number of partitions of a $2n$-element set into blocks, each containing more than one element [30].

Define the following Orlicz-type norm:

$$[\Xi]_p = \inf\left\{\lambda > 0 : \prod_{k=1}^{\infty}\mathrm{E}\left(\left|1+\frac{X_k}{\lambda}\right|^p\right) \le e^p\right\}$$

for an arbitrary sequence $\Xi = \{X_k\}_{k=1}^{\infty}$ of independent random variables, for any $p > 0$. We mention Latała's inequality [32]:

$$\frac{e-1}{2e^2}\cdot[\Xi]_p \le \left\|\sum_{k=1}^{\infty}X_k\right\|_p \le e\cdot[\Xi]_p$$

which holds either if all the Xs are nonnegative and $p \ge 1$, or if all the Xs are symmetric and $p \ge 2$. Observe here that the bounds do not depend on p, unlike the earlier inequalities. For the nonnegative case, Hitczenko & Montgomery-Smith [33] improved the left-hand constant $(e-1)/(2e^2) = 0.116272...$ to

$\xi = 0.154906...$, where ξ is the unique positive solution of the equation

$$\sum_{k=0}^{\infty} \frac{(2k+1)^k}{k!} x^k = e.$$

It is not known if this improvement carries over to the symmetric case, nor whether a calculation of best constants is feasible at present.

Assuming $\sum_{k=1}^{n} c_k^2 = 1$, it is conjectured that the Rademacher sequence satisfies [34–38]

$$P_n = \mathrm{P}\left(\left| \sum_{k=1}^{n} c_k \varepsilon_k \leq 1 \right| \right) \geq \frac{1}{2}$$

always. This inequality is provably true if $1/2$ is replaced by $3/8$ [35] or if all cs are equal [37]. For the latter scenario, we deduce that

$$\lim_{n \to \infty} P_n = \operatorname{erf}\left(1/\sqrt{2}\right) = 0.6826894921...$$

by the normal approximation [39] to the binomial distribution. This constant also appears in [40] with regard to a continued fraction expansion. Related work includes [41, 42].

[1] A. Khintchine, Über dyadische Brüche, *Math. Z.* 18 (1923) 109–116.

[2] S. J. Szarek, On the best constants in the Khinchin inequality, *Studia Math.* 58 (1976) 197–208; MR0430667 (55 #3672).

[3] U. Haagerup, The best constants in the Khintchine inequality, *Studia Math.* 70 (1981) 231–283; MR0654838 (83m:60031).

[4] J. E. Littlewood, On bounded bilinear forms in an infinite number of variables, *Quart. J. Math.* 1 (1930) 164–174.

[5] R. R. Hall, On a conjecture of Littlewood, *Math. Proc. Cambridge Philos. Soc.* 78 (1975) 443–445; MR0404151 (53 #7954).

[6] P. Whittle, Bounds for the moments of linear and quadratic forms in independent random variables (in Russian), *Teor. Verojatnost. i Primenen.* 5 (1960) 331–335; Engl. transl. in *Theory Probab. Appl.* 5 (1960) 302–305; MR0133849 (24 #A3673).

[7] S. B. Steckin, On the best lacunary system of functions (in Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* 25 (1961) 357–366; MR0131097 (24 #A951).

[8] M. L. Eaton, A note on symmetric Bernoulli random variables, *Annals Math. Statist.* 41 (1970) 1223–1226; MR0268930 (42 #3827).

[9] R. M. G. Young, On the best possible constants in the Khintchine inequality, *J. London Math. Soc.* 14 (1976) 496–504; MR0438089 (55 #11008).

[10] B. Tomaszewski, A simple and elementary proof of the Khintchine inequality with the best constant, *Bull. Sci. Math.* 111 (1987) 103–109; MR0886960 (89b:42004).

[11] R. Komorowski, On the best possible constants in the Khintchine inequality for $p \geq 3$, *Bull. London Math. Soc.* 20 (1988) 73–75; MR0916079 (89e:60037).

[12] F. L. Nazarov and A. N. Podkorytov, Ball, Haagerup, and distribution functions, *Complex Analysis, Operators, and Related Topics: S. A. Vinogradov Memorial Volume*, ed. V. P. Havin and N. K. Nikolski, Birkhäuser Verlag, 2000, pp. 247–267; MR1771767 (2001i:26022).

[13] J. Wissel and J. Turian, On the best constants in the Khintchine inequality, unpublished note (1997).

[14] G. Peshkir and A. N. Shiryaev, Khinchin inequalities and a martingale extension of the sphere of their action (in Russian), *Uspekhi Mat. Nauk,* v. 50 (1995) n. 5, 3–62; Engl. transl. in *Russian Math. Surveys* 50 (1995) 849–904; MR1365047 (96k: 60038).

[15] A. Baernstein and R. C. Culverhouse, Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions, *Studia Math*. 152 (2002) 231–248; MR1916226 (2003d:60040).

[16] A. Pełczyński, Norms of classical operators in function spaces, *Astérisque* 131 (1985) 137–162; MR0816744 (87b:47036).

[17] J. Sawa, The best constant in the Khintchine inequality for complex Steinhaus variables, the case $p = 1$, *Studia Math.* 81 (1985) 107–126; MR0818175 (87d:26024).

[18] G. Peškir, Best constants in Kahane-Khintchine inequalities for complex Steinhaus functions, *Proc. Amer. Math. Soc.* 123 (1995) 3101–3111; MR1283561 (95m: 46037).

[19] H. König and S. Kwapień, Best Khintchine type inequalities for sums of independent, rotationally invariant random vectors, *Positivity* 5 (2001) 115–152; MR1825172 (2002a:60023).

[20] B. Grünbaum, Projection constants, *Trans. Amer. Math. Soc.* 95 (1960) 451–465; MR0114110 (22 #4937).

[21] H. König, C. Schütt and N. Tomczak-Jaegermann, Projection constants of symmetric spaces and variants of Khintchine's inequality, *J. Reine Angew. Math.* 511 (1999) 1–42; MR1695788 (2000i:46014).

[22] J. Keane, Exact value of an integral involving Bessel J_0, unpublished note (2004).

[23] H. P. Rosenthal, On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables, *Israel J. Math.* 8 (1970) 273–303; MR0271721 (42 #6602).

[24] W. B. Johnson, G. Schechtman and J. Zinn, Best constants in moment inequalities for linear combinations of independent and exchangeable random variables, *Annals of Probab.* 13 (1985) 234–253; MR0770640 (86i:60054).

[25] I. F. Pinelis and S. A. Utev, Estimates of the moments of sums of independent random variables (in Russian), *Teor. Veroyatnost. i Primenen.* 29 (1984) 554–557; Engl. transl. in *Theory Probab. Appl.* 29 (1984) 574–577; MR0761144 (85m:60034).

[26] S. A. Utev, Extremal problems in moment inequalities (in Russian), *Limit Theorems of Probability Theory*, ed. A. A. Borovkov, Trudy Inst. Mat. 5, Nauka Sibirsk. Otdel., 1985, pp. 56–75, 175; MR0821753 (87d:60021).

[27] T. Figiel, P. Hitczenko, W. B. Johnson, G. Schechtman and J. Zinn, Extremal properties of Rademacher functions with applications to the Khintchine and Rosenthal inequalities, *Trans. Amer. Math. Soc.* 349 (1997) 997–1027; MR1390980 (97j: 60035).

[28] R. Ibragimov and S. Sharakhmetov, On an exact constant for the Rosenthal inequality (in Russian), *Teor. Veroyatnost. i Primenen.* 42 (1997) 341–350; Engl. transl. in *Theory Probab. Appl.* 42 (1997) 294–302; MR1474714 (98j:60026).

[29] S. R. Finch, Lengyel's constant: Stirling partition numbers, *first volume*, pp. 316–317.

[30] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000110, A000296 and A005046.

[31] R. Ibragimov and S. Sharakhmetov, The exact constant in the Rosenthal inequality for random variables with mean zero (in Russian), *Teor. Veroyatnost. i Primenen.* 46 (2001) 134–138; Engl. transl. in *Theory Probab. Appl.* 46 (2002) 127–131; MR1968709 (2004b:60054a).

[32] R. Latała, Estimation of moments of sums of independent real random variables, *Annals of Probab*. 25 (1997) 1502–1513; MR1457628 (98h:60021).
[33] P. Hitczenko and S. Montgomery-Smith, A note on sums of independent random variables, *Advances in Stochastic Inequalities*, Proc. 1997 Atlanta conf., ed. T. P. Hill and C. Houdré, Amer. Math. Soc., 1999, pp. 69–73; math.PR/9804068; MR1694763 (2000d:60080).
[34] R. K. Guy, Any answers anent these analytical enigmas? *Amer. Math. Monthly* 93 (1986) 279–281; MR1540838.
[35] R. Holzman and D. J. Kleitman, On the product of sign vectors and unit vectors, *Combinatorica* 12 (1992) 303–316; MR1226134 (94g:05065).
[36] I. Pinelis, An asymptotically Gaussian bound on the Rademacher tails, *Elec. J. Probab.* 17 (2012) N35, arXiv:1007.2137; MR2924368.
[37] M. C. A. van Zuijlen, On a conjecture concerning the sum of independent Rademacher random variables, arXiv:1112.4988.
[38] H. Hendriks and M. C. A. van Zuijlen, Sharp concentration inequalities for deviations from the mean for sums of independent Rademacher random variables, *Annals Combin.* 21 (2017) 281–291; arXiv:1210.1059; MR3654899.
[39] S. R. Finch, Notation: Error function, *this volume*, p. xi.
[40] S. R. Finch, Euler-Gompertz constant, *first volume*, pp. 423–428.
[41] H. König, On the best constants in the Khintchine inequality for Steinhaus variables, *Israel J. Math.* 203 (2014) 23–57; MR3273431.
[42] A. Eskenazis, P. Nayar and T. Tkocz, Gaussian mixtures: entropy and geometric inequalities, arXiv:1611.04921.

4.3 Ornstein–Uhlenbeck Process

We first define several words. A stochastic process $\{Y_t : t \geq 0\}$ is

- **stationary** if, for all $t_1 < t_2 < \ldots < t_n$ and $h > 0$, the random n-vectors $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})$ and $(Y_{t_1+h}, Y_{t_2+h}, \ldots, Y_{t_n+h})$ are identically distributed; that is, time shifts leave joint probabilities unchanged
- **Gaussian** if, for all $t_1 < t_2 < \ldots < t_n$, the n-vector $(Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n})$ is multivariate normally distributed
- **Markovian** if, for all $t_1 < t_2 < \ldots < t_n$, $\mathrm{P}(Y_{t_n} \leq y \mid Y_{t_1}, Y_{t_2}, \ldots, Y_{t_{n-1}}) = \mathrm{P}(Y_{t_n} \leq y \mid Y_{t_{n-1}})$; that is, the future is determined only by the present and not the past.

Also, a process $\{Y_t : t \geq 0\}$ is said to have **independent increments** if, for all $t_0 < t_1 < \ldots < t_n$, the n random variables $Y_{t_1} - Y_{t_0}$, $Y_{t_2} - Y_{t_1}$, ..., $Y_{t_n} - Y_{t_{n-1}}$ are independent. This condition implies that $\{Y_t : t \geq 0\}$ is Markovian, but not conversely. The increments are further said to be **stationary** if, for any $t > s$ and $h > 0$, the distribution of $Y_{t+h} - Y_{s+h}$ is the same as the distribution of $Y_t - Y_s$. This additional provision is needed for the following definition.

A stochastic process $\{W_t : t \geq 0\}$ is a **Wiener–Lévy process** or **Brownian motion** if it has stationary independent increments, if W_t is normally distributed and $\mathrm{E}(W_t) = 0$ for each $t > 0$, and if $W_0 = 0$. It follows immediately that $\{W_t : t > 0\}$ is Gaussian and that $\mathrm{Cov}(W_s, W_t) = \theta^2 \min\{s, t\}$, where the variance parameter θ^2 is a positive constant. For concreteness' sake, we henceforth assume that

$\theta = 1$. Almost all sample paths of Brownian motion are everywhere continuous but nowhere differentiable.

One technical stipulation is required for the following. A stochastic process $\{Y_t : t \geq 0\}$ is **continuous in probability** if, for all $u \in \mathbb{R}^+$ and $\varepsilon > 0$, $\mathrm{P}(|Y_v - Y_u| \geq \varepsilon) \to 0$ as $v \to u$. This holds if $\mathrm{Cov}(Y_s, Y_t)$ is continuous over $\mathbb{R}^+ \times \mathbb{R}^+$. Note that this is a statement about distributions, not sample paths.

Having dispensed with preliminaries, we turn to the central topic. A stochastic process $\{X_t : t \geq 0\}$ is an **Ornstein–Uhlenbeck process** or a **Gauss–Markov process** if it is stationary, Gaussian, Markovian, and continuous in probability [1, 2]. A fundamental theorem, due to Doob [3–5], ensures that $\{X_t : t \geq 0\}$ necessarily satisfies the following linear stochastic differential equation:

$$dX_t = -\rho(X_t - \mu)dt + \sigma\, dW_t,$$

where $\{W_t : t \geq 0\}$ is Brownian motion with unit variance parameter and μ, ρ, σ are constants. We have moments

$$\mathrm{E}(X_t) = \mu, \quad \mathrm{Cov}(X_s, X_t) = \frac{\sigma^2}{2\rho} e^{-\rho|s-t|}$$

in the unconditional (strictly stationary) case and

$$\mathrm{E}(X_t \mid X_0 = c) = \mu + (c - \mu)e^{-\rho t},$$

$$\mathrm{Cov}(X_s, X_t \mid X_0 = c) = \frac{\sigma^2}{2\rho}\left(e^{-\rho|s-t|} - e^{-\rho(s+t)}\right)$$

in the conditional (asymptotically stationary) case, where X_0 is initially constant. The latter case encompasses Brownian motion when $\mu = c = 0$, $\sigma = 1$ and $\rho \to 0^+$. The former case encompasses idealized **white noise** $\{dW_t/dt : t \geq 0\}$ when $\mu = 0$, $\sigma = \rho$ and $\rho \to \infty$.

Before proceeding, we note the following simple algorithm for generating a sample path of the Ornstein–Uhlenbeck process (also known as **colored noise**) over the time interval $[0, T]$. Let N be a large integer and let z_0, z_1, ..., z_N be independent random numbers generated from a normal distribution with mean 0 and variance $\sigma^2/(2\rho)$. Define $x_0 = \mu + z_0$ for the unconditional case and $x_0 = c$ for the conditional case. Then define recursively

$$x_n = \mu + \kappa_N(x_{n-1} - \mu) + \sqrt{1 - \kappa_N^2}\, z_n$$

for $1 \leq n \leq N$, where $\kappa_N = \exp(-\rho T/N)$. The sequence x_0, x_1, ..., x_N is called a first-order autoregressive sequence (a discrete analog of the OU process) with lag-one correlation coefficient κ_N. Finally, interpolate linearly the values $X(nT/N) = x_n$ for $0 \leq n \leq N$ to obtain the desired path [6–8]. More sophisticated simulation methods are found in [9–11].

For concreteness' sake, we henceforth assume that $\mu = 0$, $\rho = 1$ and $\sigma^2 = 2$. (Some authors take $\sigma^2 = 1$ instead; the decision becomes apparent in any paper by seeing whether $\mathrm{Cov}(X_s, X_t)$ is $e^{-|s-t|}$ or $e^{-|s-t|}/2$.) The conditional probability

$$\mathrm{P}(X_t \le x \mid X_0 = c) = \frac{1}{\sqrt{2\pi(1-e^{-2t})}} \int_{-\infty}^{x} \exp\left(-\frac{(\xi - ce^{-t})^2}{2(1-e^{-2t})}\right) d\xi$$

tends to the standard normal distribution, of course, as $t \to \infty$ (meaning that transients die out with time and do not affect long-term behavior). Likewise, $\mathrm{P}(X_s \le x \text{ and } X_t \le y \mid X_0 = c)$ can be evaluated. One might believe that the solution of any problem involving the OU process would be similarly straightforward; the following sections serve, however, to eliminate such ideas [12, 13].

4.3.1 *First-Passage Times*

For $a \in \mathbb{R}$, we wish to find the length of time required for an OU process to cross the level $x = a$, given that it started at $x = c$. Define the **first-passage time** or **hitting time** $T_{a,c}$ by $T_{a,c} = \inf\{t \ge 0 : X_t = a \mid X_0 = c\}$. The random variable $T_{a,c}$ is 0 if and only if $a = c$. Let $f_{a,c}(t)$ denote the density function of $T_{a,c}$. In the special case when $a = 0$, it is known that [2, 12, 14, 15]

$$f_{0,c}(t) = \sqrt{\frac{2}{\pi}} \frac{|c|e^{-t}}{(1-e^{-2t})^{3/2}} \exp\left(-\frac{c^2 e^{-2t}}{2(1-e^{-2t})}\right)$$

but for $a \ne 0$, the formulas for $f_{a,c}(t)$ are more complicated (as we shall soon see). For $a > 0$ and $c > 0$, Thomas [16] and Ricciardi & Sato [17, 18] demonstrated that [19]

$$\mathrm{E}(T_{a,0}) = \sqrt{\frac{\pi}{2}} \int_{0}^{a} \left(1 + \mathrm{erf}\left(\frac{t}{\sqrt{2}}\right)\right) \exp\left(\frac{t^2}{2}\right) dt = \frac{1}{2} \sum_{k=1}^{\infty} \frac{\left(\sqrt{2}a\right)^k}{k!} \Gamma\left(\frac{k}{2}\right),$$

$$\mathrm{E}(T_{0,c}) = \sqrt{\frac{\pi}{2}} \int_{-c}^{0} \left(1 + \mathrm{erf}\left(\frac{t}{\sqrt{2}}\right)\right) \exp\left(\frac{t^2}{2}\right) dt = \frac{1}{2} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\left(\sqrt{2}c\right)^k}{k!} \Gamma\left(\frac{k}{2}\right)$$

and, for example,

$$\mathrm{E}(T_{1,0}) = 2.0934066496..., \quad \mathrm{E}(T_{0,1}) = 0.9019080126...,$$
$$\mathrm{E}(T_{2,0}) = 10.4284093979..., \quad \mathrm{E}(T_{0,2}) = 1.4252045655....$$

The asymmetry in going from 0 to x, versus going from x to 0, is unsurprising: The process has mean 0, hence it tends to arrive at 0 more often than it departs from 0. For $a > 0$ and $c > 0$, we have [17, 18]

$$\begin{aligned}\operatorname{Var}(T_{a,0}) &= \sqrt{2\pi}\int_0^a \int_{-\infty}^t \int_s^a \left(1+\operatorname{erf}\left(\frac{r}{\sqrt{2}}\right)\right)\exp\left(\frac{r^2+t^2-s^2}{2}\right)dr\,ds\,dt - \mathrm{E}(T_{a,0})^2 \\ &= -\frac{1}{2}\sum_{k=1}^{\infty}\frac{\left(\sqrt{2}a\right)^k}{k!}\Gamma\left(\frac{k}{2}\right)\Psi\left(\frac{k}{2}\right) + \mathrm{E}(T_{a,0})^2,\end{aligned}$$

$$\begin{aligned}\operatorname{Var}(T_{0,c}) &= \sqrt{2\pi}\int_{-c}^0 \int_{-\infty}^t \int_s^0 \left(1+\operatorname{erf}\left(\frac{r}{\sqrt{2}}\right)\right)\exp\left(\frac{r^2+t^2-s^2}{2}\right)dr\,ds\,dt - \mathrm{E}(T_{0,c})^2 \\ &= -\frac{1}{2}\sum_{k=1}^{\infty}(-1)^k\frac{\left(\sqrt{2}c\right)^k}{k!}\Gamma\left(\frac{k}{2}\right)\Psi\left(\frac{k}{2}\right) - \mathrm{E}(T_{0,c})^2\end{aligned}$$

where $\Psi(x) = \psi(x) - \psi(1)$ and $\psi(x)$ is the digamma function [20]. In particular, $\Psi(1) = 0$ and

$$\Psi(x) = \begin{cases} \displaystyle\sum_{j=1}^{x-1}\frac{1}{j} & \text{if } x \text{ is an integer} > 1 \\ \displaystyle -2\ln(2) + 2\sum_{j=1}^{x-1/2}\frac{1}{2j-1} & \text{if } x \text{ is a half-integer} > 0. \end{cases}$$

For example,

$$\begin{aligned} &\operatorname{Var}(T_{1,0}) = 5.8420278024\ldots, && \operatorname{Var}(T_{0,1}) = 0.8510837032\ldots, \\ &\operatorname{Var}(T_{2,0}) = 105.2752035488\ldots, && \operatorname{Var}(T_{0,2}) = 1.0669454393\ldots. \end{aligned}$$

To compute $f_{a,c}(t)$ exactly for arbitrary a and c, we would need to invert the following (Laplace transform) identity due to Darling & Siegert [21–24]:

$$\mathrm{E}(e^{-\lambda T_{a,c}}) = \int_0^{\infty} f_{a,c}(t)e^{-\lambda t}dt = \begin{cases} \dfrac{D_{-\lambda}(-c)}{D_{-\lambda}(-a)}\exp\left(\dfrac{c^2-a^2}{4}\right) & \text{if } c < a \\ \dfrac{D_{-\lambda}(c)}{D_{-\lambda}(a)}\exp\left(\dfrac{c^2-a^2}{4}\right) & \text{if } c > a, \end{cases}$$

where $D_\nu(x)$ is the **parabolic cylinder function** or **Weber function** [25]:

$$D_\nu(x) = \begin{cases} \sqrt{\dfrac{2}{\pi}}\exp\left(\dfrac{x^2}{4}\right)\displaystyle\int_0^{\infty} t^\nu \exp\left(-\frac{t^2}{2}\right)\cos\left(xt - \frac{\nu\pi}{2}\right)dt & \text{if } \nu > -1 \\ \dfrac{1}{\Gamma(-\nu)}\exp\left(-\dfrac{x^2}{4}\right)\displaystyle\int_0^{\infty} t^{-\nu-1}\exp\left(-\frac{t^2}{2} - xt\right)dt & \text{if } \nu < 0. \end{cases}$$

The two branches of this formula agree for $-1<\nu<0$. A differential equation

$$\frac{d^2y}{dx^2}-\left(\frac{x^2}{4}-\nu-\frac{1}{2}\right)y(x)=0$$

is satisfied by $D_\nu(x)$ and, if ν is not an integer, independently by $D_{-\nu}(x)$. A series representation in terms of confluent hypergeometric functions (§4.3.4) is also useful. Unfortunately a closed-form expression for the inverse Laplace transform seems not to be possible; only a numerical approach is feasible at present. Keilson & Ross [26] tabulated the distribution of $T_{a,c}$ for a number of values a and c. For example, the median time for an OU process X_t to reach $a=1$, given that $X_0=c=0$, is 1.1892.... This corresponds to the 50$^{\text{th}}$ percentile of the distribution of $T_{1,0}$. The median of $T_{2,0}$, by contrast, is 7.2521....

We turn to a more complicated problem involving two (absorbing) boundaries rather than just one. Given $a<c<b$, what is the length of time required for the process to escape the interval (a,b), given that it started at $x=c$? Define $T_{a,b,c}=\inf\{t\geq 0: X_t=a \text{ or } X_t=b\,|\,X_0=c\}$ and let $f_{a,b,c}(t)$ denote the density function of $T_{a,b,c}$. Efforts have focused on the scenario in which $-a=b>0$. The Laplace transform of $f_{-b,b,c}(t)$ satisfies [23]

$$\mathrm{E}(e^{-\lambda T_{-b,b,c}})=\frac{D_{-\lambda}(c)+D_{-\lambda}(-c)}{D_{-\lambda}(b)+D_{-\lambda}(-b)}\exp\left(\frac{c^2-b^2}{4}\right)$$

assuming $-b<c<b$. From another table in [26], the median of $T_{-1,1,0}$ is found to be 0.4449.... The reason that this is less than 1.1892... is clear: Each direction of travel leads to a potential crossing. The median of $T_{-2,2,0}$ is 3.2439....

Keilson & Ross' approach to evaluating such probabilities was based on finding zeroes and residues in the complex plane of the parabolic cylinder functions. Alternative approaches for numerically computing $f_{a,c}(t)$ and $f_{-b,b,c}(t)$ include [27–30]. We report on some related asymptotics in §4.3.4.

There is an obvious connection between first-passage times and extreme values of a process (in the conditional case). We simply summarize:

$$\left.\begin{array}{ll}\mathrm{P}\left(\max\limits_{0\leq t\leq T}X_t\leq a\,\middle|\,X_0=c\right) & \text{if } c<a\\ \mathrm{P}\left(\min\limits_{0\leq t\leq T}X_t\geq a\,\middle|\,X_0=c\right) & \text{if } c>a\end{array}\right\}=\mathrm{P}(T_{a,c}>T)=1-F_{a,c}(T)$$

and, if $a<c<b$,

$$\mathrm{P}\left(a\leq\min_{0\leq t\leq T}X_t\leq\max_{0\leq t\leq T}X_t\leq b\,\middle|\,X_0=c\right)=\mathrm{P}(T_{a,b,c}>T)=1-F_{a,b,c}(T)$$

where $F_{a,c}(t)$, $F_{a,b,c}(t)$ are the cumulative distribution functions of $T_{a,c}$, $T_{a,b,c}$. In the special case when $-a=b>0$, the latter formula becomes a statement about

$\max_{0\le t\le T}|X_t|$, given $X_0=c$. Also, the **range** of the process satisfies [23]

$$\mathrm{P}\left(\max_{0\le t\le T}X_t-\min_{0\le t\le T}X_t\le r\,\middle|\,X_0=c\right)=\int\limits_0^r\int\limits_{c-q}^c\frac{\partial^2}{\partial a\,\partial b}F_{a,b,c}(T)\bigg|_{b=a+q}da\,dq$$

but no one apparently has calculated this probability.

4.3.2 *Historical Maximums*

If the condition $X_0=c$ is discarded, what then can be said about $\max_{0\le t\le T}X_t$ or $\max_{0\le t\le T}|X_t|$? We focus solely on the former expression and write $M_T=\max_{0\le t\le T}X_t$. It can be shown that [31–33]

$$\mathrm{P}(M_T\le 0)=\frac{1}{\pi}\arcsin\left(e^{-T}\right)$$

which is a beautiful (but isolated) result. More generally [33],

$$\int\limits_0^\infty \mathrm{P}(M_t\le y)e^{-\lambda t}dt=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^y\frac{1}{\lambda}\left(1-\frac{D_{-\lambda}(-x)}{D_{-\lambda}(-y)}\exp\left(\frac{x^2-y^2}{4}\right)\right)\exp\left(-\frac{x^2}{2}\right)dx$$

for arbitrary y, or

$$\int\limits_0^\infty g_t(y)e^{-\lambda t}dt=\frac{1}{\sqrt{2\pi}}\frac{D_{-\lambda-1}(-y)^2}{D_{-\lambda}(-y)^2}\exp\left(\frac{-y^2}{2}\right)$$

where $g_t(y)$ is the density function of M_t. For example, the median value of M_1 is 1.0393... and the median value of M_{10} is 2.2202.... It can be inferred from §4.3.3 that the median of M_T is asymptotically $\sqrt{2\ln(T)}$ as $T\to\infty$.

An alternative approach for numerically computing $\mathrm{P}(M_t\le y)$ via the Mellin transform is due to DeLong [34–36]. An interesting application to computer science, involving the maximum size reached by a dynamic data structure over a long span of time, is described in [37].

4.3.3 *Pickands' Constants*

Assume that $\{Y_t:t\ge 0\}$ is a stationary Gaussian process with zero mean, unit variance and covariance function of the form

$$r(|s-t|)=\mathrm{Cov}(Y_s,Y_t)=1-C|s-t|^\alpha+o(|s-t|^\alpha)$$

as $|s-t|\to 0$, where $0<\alpha\le 2$ and $C>0$ are constants. Assume further that $r(\tau)\ln(\tau)\to 0$ as $\tau\to\infty$. Pickands [38–42] demonstrated that $M_T=\max_{0\le t\le T}Y_t$ has the Gumbel limiting distribution [43]

$$\lim_{T\to\infty}\mathrm{P}\left(\sqrt{2\ln(T)}\,(M_T-k_T)\le x\right)=\exp(-e^{-x}),$$

where

$$k_T = \sqrt{2\ln(T)} + \frac{1}{\sqrt{2\ln(T)}}\left\{\frac{2-\alpha}{2\alpha}\ln(\ln(T)) + \ln\left((2\pi)^{-\frac{1}{2}}2^{\frac{2-\alpha}{2\alpha}}C^{\frac{1}{\alpha}}H_\alpha\right)\right\}$$

and H_α is a positive constant independent of C. It is known that $H_1 = 1$ (corresponding to the OU process) and $H_2 = 1/\sqrt{\pi}$. No other exact values for H_α are known. An alternative characterization of H_α is

$$H_\alpha = \lim_{T\to\infty}\int_0^\infty \mathrm{P}(\tilde{M}_T > y)e^y dy,$$

where $\{\tilde{Y}_t : t \geq 0\}$ is a nonstationary Gaussian process with

$$\mathrm{E}(\tilde{Y}_t) = -|t|^\alpha, \quad \mathrm{Cov}(\tilde{Y}_s, \tilde{Y}_t) = |s|^\alpha + |t|^\alpha - |s-t|^\alpha$$

but this does not seem to help. Shao [44] and Debicki, Michna & Rolski [45] gave bounds on H_α; for example,

$$0.009 \leq H_{1/2} \leq 715.94, \quad 0.208 \leq H_{3/2} \leq 3.04.$$

A conjecture that $H_\alpha = 1/\Gamma(1/\alpha)$ remains unproved. There is also a connection with the Gaussian correlation conjecture and with estimating small ball probabilities [46].

4.3.4 *Upper Tail Asymptotics*

We revisit the single-boundary first-passage time distribution and ask about the limiting value

$$\lambda(a) = \lim_{t\to\infty}\frac{1}{t}\ln\{\mathrm{P}(T_{a,0} > t)\}$$

as a function of $a > 0$. In words, what can be said about the upper tail of the distribution of the first hitting time $T_{a,0}$ for an OU process X_t across the level $x = a$, given that $X_0 = 0$? Mandl [47, 48] and Beekman [49] demonstrated that $-1 < \lambda(a) < 0$ and that $\lambda(a)$ is the zero of $D_{-\lambda}(-a)$ closest to 0. Sample values include [17, 50, 51]

$$\lim_{a\to 0^+}\lambda(a) = -1, \quad \lim_{a\to\infty}\lambda(a)\cdot\frac{\exp(a^2/2)}{a} = \frac{-1}{\sqrt{2\pi}},$$

$$\lambda(0.7649508673...) = -\tfrac{1}{2},$$

$$\lambda(1) = -0.3882382947... = 2(-0.1941191473...),$$

$$\lambda(2) = -0.0972745958... = 2(-0.0486372979...).$$

For the symmetric double-boundary first-passage time distribution, we examine

$$\lambda(-b, b) = \lim_{t\to\infty}\frac{1}{t}\ln\{\mathrm{P}(T_{-b,b,0} > t)\}$$

as a function of $b>0$. Breiman [52] proved that $-\infty<\lambda(-b,b)<0$ and that $\lambda(-b,b)$ is the zero of $\Phi(\lambda/2,1/2,b^2/2)$ closest to 0, where

$$\Phi(u,v,w)=1+\sum_{k=1}^{\infty}\frac{u(u+1)(u+2)\cdots(u+k-1)}{v(v+1)(v+2)\cdots(v+k-1)}\frac{w^k}{k!}$$

is the **confluent hypergeometric function of the first kind**. For simplicity, define $\mu(b)=\lambda(-b,b)$. Sample values include [51–53]

$$\lim_{b\to0^+}\mu(b)=-\infty,\quad \lim_{b\to\infty}\mu(b)\cdot\frac{\exp(b^2/2)}{b}=\frac{-1}{\sqrt{2\pi}},$$

$$\mu(1)=-2,\quad \mu(1.3069297277...)=-1,\quad \mu(1.6438001904...)=-\tfrac{1}{2},$$

$$\mu\left(\sqrt{3-\sqrt{6}}\right)=\mu(0.7419637843...)=-4,$$

$$\mu(2)=-0.2429928807...,\quad \mu(3)=-0.0239463006...,$$

$$\mu\left(\sqrt{2}\right)=-0.7984598320...,\quad \mu\left(2\sqrt{2}\right)=-0.0374612092....$$

The latter two values come from [53], where a different time scaling was chosen. Also, the constant $(3-6^{1/2})^{1/2}$ appears in [54–56] with regard to stopping rules in statistical sequential analysis.

For completeness' sake, here is the expression for $D_{-\lambda}(x)$ in terms of confluent hypergeometric functions:

$$\begin{aligned}D_{-\lambda}(x) &= \frac{\sqrt{\pi}2^{-\lambda/2}}{\Gamma((1+\lambda)/2)}e^{-x^2/4}\Phi\left(\frac{\lambda}{2},\frac{1}{2},\frac{x^2}{2}\right)\\ &\quad -2\frac{\sqrt{\pi}2^{-(1+\lambda)/2}}{\Gamma(\lambda/2)}xe^{-x^2/4}\Phi\left(\frac{1+\lambda}{2},\frac{3}{2},\frac{x^2}{2}\right)\end{aligned}$$

which gives rise to the values $\lambda(1)$, $\lambda(2)$ and $\lambda^{-1}(-1/2)$ listed earlier. The constant $\mu^{-1}(-1)$ is important in the study of sample path behavior of Brownian motion [51, 57, 58] and first appeared in [55], as far as is known. Some higher-dimensional results are given in [51, 59]. Csáki [60, 61] outlined the distributional asymptotics of the maximum M_T, but we cannot discuss this topic further.

Addendum New numerical transform inversion algorithms [62–64] make enhancement of the tables in [26, 33] possible. Also, the distribution of the L_2-norm of X_t on $[0,T]$ can be inferred from closed-form expressions in [65, 66]. We wonder about corresponding results for L_1 and L_∞-norms. The conjectured formula for H_α in terms of the gamma function is probably false [67–70]; simulation-based point estimates $H_{3/2}\approx0.77$ and confidence bounds $0.768\leq H_{3/2}\leq0.786$ do not carry over well to $H_{1/2}$ since the underlying algorithm becomes unreliable for $0<\alpha<1$.

[1] G. E. Uhlenbeck and L. S. Ornstein, On the theory of Brownian motion, *Phys. Rev.* 36 (1930) 823–841.

[2] M. C. Wang and G. E. Uhlenbeck, On the theory of Brownian motion. II, *Rev. Modern Phys.* 17 (1945) 323–342; MR0013266 (7,130e).

[3] J. L. Doob, The Brownian movement and stochastic equations, *Annals of Math.* 43 (1942) 351–369; also in *Selected Papers on Noise and Stochastic Processes*, ed. N. Wax, Dover, 1954, pp. 319–337; MR0006634 (4,17d).

[4] L. Breiman, *Probability*, Addison-Wesley, 1968, pp. 347–351; MR0229267 (37 #4841).

[5] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1988, p. 358; MR1121940 (92h:60127).

[6] W. Feller, *An Introduction to Probability Theory and Its Applications*, v. 2, 2nd ed., Wiley, 1971, pp. 96–99, 335–337; MR0270403 (42 #5292).

[7] I. I. Gringorten, Estimating finite-time maxima and minima of a stationary Gaussian Ornstein-Uhlenbeck process by Monte Carlo simulation, *J. Amer. Statist. Assoc*. 63 (1968) 1517–1521.

[8] L. Bartosch, Generation of colored noise, *Int. J. Mod. Phys. C* 12 (2001) 851–855.

[9] R. L. Honeycutt, Stochastic Runge-Kutta algorithms. I, White noise, *Phys. Rev. A* 45 (1992) 600–603; II, Colored noise, *Phys. Rev. A* 45 (1992) 604–610.

[10] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992; MR1214374 (94b:60069) and MR1260431 (95a:60084).

[11] D. J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations, *SIAM Rev*. 43 (2001) 525–546; MR1872387 (2002j:65008).

[12] I. F. Blake and W. C. Lindsey, Level-crossing problems for random processes, *IEEE Trans. Inform. Theory* IT-19 (1973) 295–315; MR0370729 (51 #6955).

[13] J. Abrahams, A survey of recent progress on level-crossing problems for random processes, *Communications and Networks: A Survey of Recent Advances*, ed. I. F. Blake and H. V. Poor, Springer-Verlag, 1986, pp. 6–25.

[14] A. Göing-Jaeschke and M. Yor, A clarification note about hitting times densities for Ornstein-Uhlenbeck processes, *Finance Stochast.* 7 (2003) 413–415.

[15] L. Alili, P. Patie and J. L. Pedersen, Representations of the first hitting time density of an Ornstein-Uhlenbeck process, *Stoch. Models* 21 (2005) 967–980; MR2179308 (2006i:60111).

[16] M. U. Thomas, Some mean first-passage time approximations for the Ornstein-Uhlenbeck process, *J. Appl. Probab.* 12 (1975) 600–604; MR0388558 (52 #9394).

[17] S. Sato, Evaluation of the first-passage time probability to a square root boundary for the Wiener process, *J. Appl. Probab.* 14 (1977) 850–856; MR0471056 (57 #10797).

[18] L. M. Ricciardi and S. Sato, First-passage-time density and moments of the Ornstein-Uhlenbeck process, J. *Appl. Probab*. 25 (1988) 43–57; MR0929503 (89b:60189).

[19] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[20] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.

[21] A. J. F. Siegert, On the first passage time probability problem, *Phys. Rev.* 81 (1951) 617–623; MR0054192 (14,887b).

[22] R. Bellman and T. Harris, Recurrence times for the Ehrenfest model, *Pacific J. Math.* 1 (1951) 179–193; MR0045323 (13,566b).

[23] D. A. Darling and A. J. F. Siegert, The first passage problem for a continuous Markov process, *Annals Math. Statist*. 24 (1953) 624–639; MR0058908 (15,449f).

[24] J. A. Beekman, Gaussian-Markov processes and a boundary value problem, *Trans. Amer. Math. Soc.* 126 (1967) 29–42; MR0203793 (34 #3641).

[25] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, *Higher Transcendental Functions*, v. II, based on notes left by H. Bateman, McGraw-Hill, 1953, pp. 116–126; MR0058756 (15,419i).

[26] J. Keilson and H. F. Ross, Passage time distributions for Gaussian Markov (Ornstein-Uhlenbeck) statistical processes, *Selected Tables in Mathematical Statistics*, v. III, Amer. Math. Soc., 1975, pp. 233–327; MR0408175 (53 #11941).

[27] C. Park and F. J. Schuurmann, Evaluations of absorption probabilities for the Wiener process on large intervals, *J. Appl. Probab.* 17 (1980) 363–372; MR0568947 (81f:60111).

[28] K. Pötzelberger and L. Wang, Boundary crossing probability for Brownian motion, *J. Appl. Probab.* 38 (2001) 152–164; MR1816120 (2002a:60138).

[29] E. Di Nardo, A. G. Nobile, E. Pirozzi and L. M. Ricciardi, A computational approach to first-passage-time problems for Gauss-Markov processes, *Adv. Appl. Probab*. 33 (2001) 453–482; MR1842303 (2002e:60053).

[30] U. Sumita, J.-Y. Gotoh and H. Jin, Numerical exploration of dynamic behavior of Ornstein-Uhlenbeck processes via Ehrenfest process approximation, *J. Oper. Res. Soc. Japan* 49 (2006) 256–278; MR2267891 (2008g:60108).

[31] S. O. Rice, Distribution of the duration of fades in radio transmission: Gaussian noise model, *Bell System Tech. J.* 37 (1958) 581–635; MR0094269 (20 #788).

[32] D. Slepian, The one-sided barrier problem for Gaussian noise, *Bell System Tech. J.* 41 (1962) 463–501; MR0133183 (24 #A3017).

[33] J. Keilson and H. F. Ross, *Gaussian Markov Related Variates for Meteorological Planning*, AFGL-TR-79-0282, ADA081382 (1979), Univ. of Rochester; available through the National Technical Information Service, Springfield VA 22161.

[34] D. M. DeLong, Crossing probabilities for a square root boundary by a Bessel process, *Comm. Statist. A–Theory Methods* 10 (1981) 2197–2213; erratum 12 (1983) 1699; MR0629897 (82i:62119) and MR0711257 (84i:62106).

[35] J.-M. Azaïs and M. Wschebor, The distribution of the maximum of a Gaussian process: Rice method revisited, *In and Out of Equilibrium: Probability with a Physics Flavor,* Proc. 2000 Mambucaba conf., ed. V. Sidoravicius, Birkhäuser Boston, 2002, pp. 321–348; MR1901961 (2003d:60075).

[36] A. Estrella, Critical values and p values of Bessel process distributions: Computation and application to structural break tests, *Econometric Theory* 19 (2003) 1128–1143.

[37] G. Louchard, C. Kenyon and R. Schott, Data structures' maxima, *SIAM J. Comput.* 26 (1997) 1006–1042; MR1460714 (98m:68045).

[38] J. Pickands, Upcrossing probabilities for stationary Gaussian processes, *Trans. Amer. Math. Soc*. 145 (1969) 51–73; MR0250367 (40 #3606).

[39] J. Pickands, Asymptotic properties of the maximum in a stationary Gaussian process, *Trans. Amer. Math. Soc*. 145 (1969) 75–86; MR0250368 (40 #3607).

[40] S. M. Berman, Maxima and high level excursions of stationary Gaussian processes, *Trans. Amer. Math. Soc.* 160 (1971) 65–85; MR0290449 (44 #7630).

[41] C. Qualls and H. Watanabe, Asymptotic properties of Gaussian processes, *Annals Math. Statist.* 43 (1972) 580–596; MR0307318 (46 #6438).

[42] M. R. Leadbetter, G. Lindgren and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, 1983, pp. 216–237; MR0691492 (84h:60050).

[43] S. R. Finch, Extreme value constants, *first volume*, pp. 363–367.

[44] Q.-M. Shao, Bounds and estimators of a basic constant in extreme value theory of Gaussian processes, *Statist. Sinica* 6 (1996) 245–257; MR1379060 (97e:60048).

[45] K. Debicki, Z. Michna and T. Rolski, Simulation of the asymptotic constant in some fluid models, *Stoch. Models* 19 (2003) 407–423; MR1993949 (2004h:60059).

[46] W. V. Li and Q.-M. Shao, Gaussian processes: Inequalities, small ball probabilities and applications, *Stochastic Processes: Theory and Methods*, ed. D. N. Shanbhag and C. R. Rao, North-Holland, 2001, pp. 533–597; MR1861734.

[47] P. Mandl, The distribution of time needed for the Uhlenbeck process to reach a border (in Czech), *Aplikace Mat.* 7 (1962) 141–148; Zbl. 0138.11303.

[48] P. Mandl, *Analytical Treatment of One-Dimensional Markov Processes*, Springer-Verlag, 1968, p. 145; MR0247667 (40 #930).

[49] J. A. Beekman, Asymptotic distributions for the Orstein-Uhlenbeck process, *J. Appl. Probab*. 12 (1975) 107–114; MR0388521 (52 #9357).

[50] K. Uchiyama, Brownian first exit from and sojourn over one-sided moving boundary and application, *Z. Wahrsch. Verw. Gebiete* 54 (1980) 75–116; MR0595482 (82c:60143).

[51] E. Perkins, On the Hausdorff dimension of the Brownian slow points, *Z. Wahrsch. Verw. Gebiete* 64 (1983) 369–399; MR0716493 (85g:60085b).

[52] L. Breiman, First exit times from a square root boundary, *Proc. Fifth Berkeley Symp. Math. Stat. Probab.*, vol. II, part 2, ed. L. M. Le Cam and J. Neyman, Univ. of Calif. Press, 1967, pp. 9–16; MR0212865 (35 #3730).

[53] K. Lindenberg, K. E. Shuler, J. Freeman and T. J. Lie, First passage time and extremum properties of Markov and independent processes, *J. Statist. Phys.* 12 (1975) 217–251; MR0458604 (56 #16804).

[54] Y. S. Chow and H. Teicher, On second moments of stopping rules, *Annals Math. Statist.* 37 (1966) 388–392; MR0192528 (33 #753).

[55] L. A. Shepp, A first passage problem for the Wiener process, *Annals Math. Statist*. 38 (1967) 1912–1914; MR0217879 (36 #968).

[56] B. M. Brown, Moments of a stopping rule related to the central limit theorem, *Annals Math. Statist*. 40 (1969) 1236–1249; MR0243689 (39 #5010).

[57] B. Davis, On Brownian slow points, *Z. Wahrsch. Verw. Gebiete* 64 (1983) 359–367; MR0716492 (85g:60085a).

[58] P. Greenwood and E. Perkins, A conditioned limit theorem for random walk and Brownian local time on square root boundaries, *Annals of Probab.* 11 (1983) 227–261; MR0690126 (85d:60068).

[59] F. Haake, J. W. Haus and R. Glauber, Passage-time statistics for the decay of unstable equilibrium states, *Phys. Rev. A* 23 (1981) 3255–3271; MR0615182 (82f:81039).

[60] E. Csáki and Z. Shi, Some lim inf results for two-parameter processes, *Stochastic Process. Appl.* 78 (1998) 27–46; MR1653288 (99i:60064).

[61] E. Csáki and Y. Hu, On the ranked excursion heights of a Kiefer process, *J. Theoret. Probab*. 17 (2004) 145–163; MR2054583 (2005f:60092).

[62] P. Valkó and J. Abate, Numerical Laplace Inversion, Wolfram MathSource posting 4738 (2002).

[63] J. Abate and P. Valkó, Numerical Inversion of Laplace Transform with Multiple Precision Using the Complex Domain, Wolfram MathSource posting 5026 (2003).

[64] J. Abate and P. Valkó, Multi-precision Laplace transform inversion, *Internat. J. Numer. Methods Engineering.* 60 (2004) 979–993.

[65] T. Dankel, On the distribution of the integrated square of the Ornstein-Uhlenbeck process, *SIAM J. Appl. Math.* 51 (1991) 568–574; MR1095035 (92c:60056).

[66] F. Gao, J. Hannig, T.-Y. Lee and F. Torcaso, Laplace transforms via Hadamard factorization, *Elec. J. Probab.* 8 (2003) N13; MR1998764 (2005h:60110).
[67] K. Burnecki and Z. Michna, Simulation of Pickands constants, *Probab. Math. Statist.* 22 (2002) 193–199; MR1944151.
[68] K. Debicki and P. Kisowski, A note on upper estimates for Pickands constants, *Statist. Probab. Lett.* 78 (2008) 2046–2051; MR2458013 (2009k:60089).
[69] A. J. Harper, Bounds on the suprema of Gaussian processes, and omega results for the sum of a random multiplicative function, *Annals Appl. Probab.* 23 (2013) 584–616; MR3059269.
[70] A. B. Dieker and B. Yakir, On asymptotic constants in the theory of extremes for Gaussian processes, *Bernoulli* 20 (2014) 1600–1619; arXiv:1206.5840; MR3217455.

4.4 Zero Crossings

In this essay, we presuppose basic knowledge of stochastic processes [1]. Let $\{X_t : t \geq 0\}$ be a zero mean, unit variance, stationary Gaussian process with twice differentiable correlation function $r(|s-t|) = \mathrm{Cov}(X_s, X_t)$. We wish to study the distribution of lengths of intervals between zeroes of X_t. There are two cases: the first in which $r(\tau)$ is analytic (implying differentiability up to all orders) and the second in which the third derivative of $r(\tau)$ possesses a jump discontinuity at $\tau = 0$.

Define $f_m(\tau)$ to be the probability density associated with the interval length τ between an arbitrary zero t_0 and the $(m+1)^{\text{st}}$ later zero t_{m+1}. In particular, $f_0(\tau)$ is the probability density for differences between successive zeroes t_0 and t_1. We will focus on the limiting behavior of $f_m(\tau)$ as $\tau \to 0^+$.

When $r(\tau)$ is analytic, it is clear that

$$r(\tau) = 1 + \frac{r''(0)}{2!}\tau^2 + \frac{r^{(4)}(0)}{4!}\tau^4 + O(\tau^6)$$

since $r(\tau)$ must be an even function. It is known, in this case, that [2]

$$f_m(\tau) = O\left(\tau^{\frac{1}{2}(m+2)(m+3)-2}\right)$$

as $\tau \to 0^+$. Further, the big O coefficient is known. We merely give an example: If $r(\tau) = \exp(-\alpha\tau^2)$ for $\alpha > 0$, then

$$\lim_{\tau\to 0^+} \frac{f_0(\tau)}{\tau} = \frac{1}{2}\alpha, \quad \lim_{\tau\to 0^+} \frac{f_1(\tau)}{\tau^4} = \frac{\sqrt{6}}{27\pi}\alpha^{5/2}.$$

The more interesting case is when $r(\tau)$ has a singularity at the origin. If

$$r(\tau) = 1 - \frac{1}{2}\tau^2 + \alpha|\tau|^3 + o(|\tau|^3),$$

then $f_m(\tau) \to C_m\alpha$ as $\tau \to 0^+$, where $C_m > 0$ is a constant (independent of α). Longuet-Higgins [3] determined the following bounds

$$1.1556 < C_0 < 1.158, \quad 0.1971 < C_1 < 0.198, \quad 0.0491 < C_2 < 0.0556,$$

but it remained for someone else to find a specific process $\{X_t\}$, and its corresponding α, for which $f_m(\tau)$ could be computed.

Wong [4–7], building upon McKean [8], examined the process

$$X_t = \sqrt{3}\exp\left(-\sqrt{3}t\right) \int_0^{\exp\left(2t/\sqrt{3}\right)} W_s\, ds$$

where W_s is standard Brownian motion ("standard" meaning that its variance parameter is 1). The correlation function for Wong's process is

$$r(\tau) = \frac{3}{2}\exp\left(-\frac{|\tau|}{\sqrt{3}}\right)\left(1 - \frac{1}{3}\exp\left(-\frac{2|\tau|}{\sqrt{3}}\right)\right)$$

and hence $\alpha = 2\sqrt{3}/9$. It turns out that $f_0(\tau)$ can be written in terms of complete elliptic integrals, and a more complicated integral expression applies for $f_m(\tau)$, $m \geq 1$. This is sufficient to deduce that

$$C_0 = \frac{37}{32} = 1.15625, \quad C_1 = \frac{47}{64} - \frac{108}{64\pi} = 0.1972270670\ldots,$$

$$C_2 = \frac{121}{128} - \frac{81}{32\pi} - \frac{27}{32\pi^2} = 0.0541008518\ldots.$$

In fact,

$$C_m = \frac{27}{4\pi^2}\int_0^\infty \frac{x^3 - 1}{x^3 + 1}\frac{x^m \ln(x)}{(x^2 + 1)^{m+1}}dx,$$

which can be evaluated exactly via residue calculus. The limiting behavior of $f_m(\tau)$ as $\tau \to 0^+$ is thus solved for all m. No one has found another stationary Gaussian process that permits exact analysis as this. Wong [4] also proved that $f_0(\tau) \to 0$ as $\tau \to \infty$ and, moreover,

$$\lim_{\tau \to \infty}\exp\left(\frac{\tau}{2\sqrt{3}}\right)f_0(\tau) = \frac{L}{\sqrt{2}} = K\left(\frac{1}{\sqrt{2}}\right) = \frac{1}{4\sqrt{\pi}}\Gamma\left(\frac{1}{4}\right)^2 = 1.8540746773\ldots,$$

where L is Gauss' lemniscate constant [9] and $K(x)$ denotes the complete elliptic integral of the first kind [10]. For $m \geq 1$, such precise asymptotics for $f_m(\tau)$ as $\tau \to \infty$ remain open. See [11–15] as well.

We shift attention to counting zeroes in an interval of prescribed length 1. Again, $\{X_t\}$ is assumed to be a zero mean, unit variance, stationary Gaussian process with twice differentiable correlation function $r(\tau)$. Let N denote the number of zeroes of X_t per unit time. The expected value of N is [16–19]

$$\mathrm{E}(N) = \frac{1}{\pi}\sqrt{-r''(0)}$$

and the variance of N is [20–26]

$$\mathrm{Var}(N) = \mathrm{E}(N) - \mathrm{E}(N)^2 + \frac{2}{\pi^2}\int_0^1 (1-\tau)F(\tau)\,d\tau,$$

where

$$F(\tau) = \left(1 - r(\tau)^2\right)^{-1} G(\tau)\left(1 + H(\tau)\arctan(H(\tau))\right),$$

$$G(\tau) = \sqrt{k_1(\tau)k_2(\tau)}, \quad H(\tau) = \frac{k_3(\tau)}{\sqrt{(1 - r(\tau)^2)\,k_1(\tau)k_2(\tau)}},$$

$$k_1(\tau) = (1 + r(\tau))\,(r''(0) - r''(\tau)) + r'(\tau)^2,$$
$$k_2(\tau) = (1 - r(\tau))\,(r''(0) + r''(\tau)) + r'(\tau)^2,$$

$$k_3(\tau) = \left(1 - r(\tau)^2\right) r''(\tau) + r(\tau)r'(\tau)^2.$$

Needless to say, an exact evaluation of $\mathrm{Var}(N)$ is generally impossible. In the case when $r(\tau)$ is analytic, we have [27]

$$\lim_{\tau\to 0^+} \frac{2}{\pi}\left(\frac{1}{H(\tau)} + \arctan\left(H(\tau)\right)\right) = 1.$$

By contrast, in the case when $r(\tau)$ has a singularity at the origin (as before),

$$\lim_{\tau\to 0^+} \frac{2}{\pi}\left(\frac{1}{H(\tau)} + \arctan\left(H(\tau)\right)\right) = \frac{2\sqrt{3}}{\pi} + \frac{1}{3} = 1.4359911241...,$$

which is an interesting occurrence of the first Lebesgue constant [28]. For Wong's process, $\mathrm{E}(N) = 1/\pi$ and [26]

$$\mathrm{Var}(N) = \frac{4}{3\pi} - \frac{1}{12} + \frac{3}{\pi^2}\left\{\arcsin\left(\frac{1}{2}\exp\left(-\frac{1}{\sqrt{3}}\right)\right)\right\}^2.$$

Only a few other stationary Gaussian processes are known to possess a closed-form expression for this variance; for example, those with correlation functions [29–31]

$$r(\tau) = \frac{1}{2} + \frac{1}{2}\cos\left(\sqrt{2}\tau\right) \quad \text{or} \quad r(\tau) = 1 - \frac{1}{2}\tau^2 + \frac{1}{6\sqrt{3}}|\tau|^3.$$

See also [32–38].

4.4.1 *Integrated Brownian Motion*

Wong's process involves an integral of standard Brownian motion. We briefly examine a simpler integral [39]:

$$Z_t = \int_0^t W_s \, ds,$$

which is zero mean Gaussian with covariance function

$$\mathrm{Cov}(Z_u, Z_v) = \int_0^u \int_0^v \min\{x, y\} \, dx \, dy = \begin{cases} \frac{1}{6}u^2(3v - u) & \text{if } v \ge u \ge 0 \\ \frac{1}{6}v^2(3u - v) & \text{if } u \ge v \ge 0. \end{cases}$$

One unsolved problem is concerned with the asymptotics of the maximum of $|Z_t|$ over the unit interval [40–43]:

$$\lim_{\varepsilon \to 0^+} \varepsilon^{2/3} \ln \left\{ \mathrm{P}\left(\max_{0 \le t \le 1} |Z_t| < \varepsilon \right) \right\} = \kappa,$$

where the constant κ is known to satisfy

$$\frac{3}{8} \le \kappa \le (2\pi)^{2/3} \frac{3}{8}.$$

These are the sharpest known bounds. Another unsolved problem is concerned with the probability that the integrated Wiener process is currently at its maximum value [44, 45]:

$$\lambda = \mathrm{P}\left(Z_t = \max_{0 \le s \le t} Z_s \right),$$

which is known to be independent of t. Since integration has the effect of smoothing W_s, it is reasonable to conjecture for Z_t that λ is positive. Two terms of a complicated infinite series were used in [44] to give an approximation $\lambda = 0.372...$, but a more accurate estimation procedure apparently has not been attempted.

4.4.2 *Random Polynomials*

Let $q(x)$ be a random polynomial of degree n, with real coefficients independently chosen from a standard Gaussian distribution. Asymptotic properties of the expected number of real zeroes of $q(x)$ were summarized in [46]; associated probabilities are more difficult to study. The probability that $q(x)$ does *not* have any zeroes in $\mathbb{R}$ is $n^{-b+o(1)}$ as $n \to \infty$ through even integers, where [47]

$$b = -4 \lim_{T \to \infty} \frac{1}{T} \ln \left(\mathrm{P}\left(\sup_{0 \le t \le T} Y(t) \le 0 \right) \right)$$

and $Y(t)$ is a zero mean, unit variance, stationary Gaussian process with correlation function $r(\tau) = \mathrm{sech}(\tau/2)$. It is known [48–50] that $0.5 < b < 1.0$ and, via simulation, $b \approx 0.76$. An exact value for b would be sensational! The statistics of real zeroes of $q(x)$ turn out to be identical in the four subintervals $(-\infty,-1)$, $[-1,0]$, $[0,1]$, $(1,\infty)$ of $\mathbb{R}$; hence the probability that $q(x)$ does not have zeroes in $[0,1]$ is $n^{-b/4+o(1)} \approx n^{-0.19}$ [51, 52]. A related topic is the capture time in the random pursuit problem for fractional Brownian particles [48–50].

[1] S. R. Finch, Ornstein-Uhlenbeck process, *this volume*, §4.3.

[2] M. S. Longuet-Higgins, The distribution of intervals between zeros of a stationary random function, *Philos. Trans. Roy. Soc. London Ser. A* 254 (1961/1962) 557–599; MR0158431 (28 #1654).

[3] M. S. Longuet-Higgins, Bounding approximations to the distribution of intervals between zeros of a stationary Gaussian process, *Proc. Sympos. Time Series Analysis*, Brown Univ., ed. M. Rosenblatt, Wiley, 1963, pp. 63–88; MR0148124 (26 #5633).

[4] E. Wong, Some results concerning the zero-crossings of Gaussian noise, *SIAM J. Appl. Math*. 14 (1966) 1246–1254; MR0207059 (34 #6875).

[5] E. Wong, The distribution of intervals between zeros for a stationary Gaussian process, *SIAM J. Appl. Math.* 18 (1970) 67–73; MR0256456 (41 #1112).

[6] I. F. Blake and W. C. Lindsey, Level-crossing problems for random processes, *IEEE Trans. Inform. Theory* IT-19 (1973) 295–315; MR0370729 (51 #6955).

[7] J. Abrahams, A survey of recent progress on level-crossing problems for random processes, *Communications and Networks: A Survey of Recent Advances*, ed. I. F. Blake and H. V. Poor, Springer-Verlag, 1986, pp. 6–25.

[8] H. P. McKean, A winding problem for a resonator driven by a white noise, *J. Math. Kyoto Univ.* 2 (1963) 227–235; MR0156389 (27 #6312).

[9] S. R. Finch, Gauss' lemniscate constant, *first volume*, pp. 420–423.

[10] S. R. Finch, Random triangles, *this volume*, §5.10.

[11] D. S. Palmer, Properties of random functions, *Proc. Cambridge Philos. Soc.* 52 (1956) 672–686; corrigenda 53 (1957) 266; MR0080398 (18,241e).

[12] A. J. Rainal, Zero-crossing intervals of Gaussian processes, *IEEE Trans. Inform. Theory* IT-8 (1962) 372–378; MR0146011 (26 #3537).

[13] D. Slepian, On the zeros of Gaussian noise, *Proc. Sympos. Time Series Analysis*, Brown Univ., ed. M. Rosenblatt, Wiley, 1963, pp. 104–115; MR0148128 (26 #5636).

[14] S. M. Cobb, The distribution of intervals between zero crossings of sine wave plus random noise and allied topics, *IEEE Trans. Inform. Theory* IT-11 (1965) 220–233; MR0186460 (32 #3920).

[15] J. Abrahams, The zero-crossing problem for some nonstationary Gaussian processes, *IEEE Trans. Inform. Theory* IT-28 (1982) 677–678.

[16] S. O. Rice, Mathematical analysis of random noise, *Bell System Tech. J.* 23 (1944) 282–332; 24 (1945) 46–156; also in *Selected Papers on Noise and Stochastic Processes*, ed. N. Wax, Dover, 1954, pp. 133–294; MR0010932 (6,89b) and MR0011918 (6,233i).

[17] A. J. Rainal, Origin of Rice's formula, *IEEE Trans. Inform. Theory* 34 (1988) 1383–1387; MR0993433 (90f:60001).

[18] K. Itô, The expected number of zeros of continuous stationary Gaussian processes, *J. Math. Kyoto Univ*. 3 (1963/1964) 207–216; MR0166824 (29 #4097).

[19] N. D. Ylvisarer, The expected number of zeros of a stationary Gaussian process, *Annals Math. Statist*. 36 (1965) 1043–1046; MR0177458 (31 #1721).

[20] H. Steinberg, P. M. Schultheiss, C. A. Wogrin and F. Zwieg, Short-time frequency measurement of narrow-band random signals by means of a zero counting process, *J. Appl. Phys.* 26 (1955) 195–201.

[21] I. Miller and J. E. Freund, Some results on the analysis of random signals by means of a cut-counting process, *J. Appl. Phys.* 27 (1958) 1290–1293.

[22] J. S. Bendat, *Principles and Applications of Random Noise Theory*, Wiley, 1958, pp. 124–128; 385–414; MR0105753 (21 #4489).

[23] V. A. Volkonskiĭ and Ju. A. Rozanov, Some limit theorems for random functions. II (in Russian), *Teor. Verojatnost. i Primenen.* 6 (1961) 202–215; Engl. transl. in *Theory Probab. Appl.* 6 (1961) 186–198; MR0137141 (25 #597).

[24] M. R. Leadbetter and J. D. Cryer, The variance of the number of zeros of a stationary normal process, *Bull. Amer. Math. Soc.* 71 (1965) 561–563; MR0174093 (30 #4300).

[25] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes: Sample Function Properties and their Applications*, Wiley, 1967, pp. 190–218; MR0217860 (36 #949).

[26] R. N. Miroshin, On the variance of number of zeros of Gaussian stationary processes, *Vestnik St. Petersburg Univ. Math.* 34 (2001) 30–35; MR1903454 (2003h:60051).

[27] J. A. McFadden, The axis-crossing intervals of random functions, *IEEE Trans. Inform. Theory* IT-2 (1956) 146–150; IT-4 (1958) 14–24; MR0098436 (20 #4895) and MR0098437 (20 #4896).

[28] S. R. Finch, Lebesgue constants, *first volume*, pp. 250–255.

[29] R. N. Miroshin, Moments of the number of zeros of stationary Gaussian locally Markovian processes of the first order on the small time interval, *Vestnik St. Petersburg Univ. Math.* 35 (2002) 18–23; MR1961951 (2004d:60090).

[30] R. N. Miroshin, Variance of the number of zeros of some stationary Gaussian processes: slight deviations from simple solutions, *Vestnik St. Petersburg Univ. Math.* 39 (2006) 38–44; MR2302636.

[31] R. N. Miroshin, On bounds for the variance of the number of zeros of a differentiable Gaussian stationary process, *Vestnik St. Petersburg Univ. Math.* 48 (2015) 82–88; MR3354011.

[32] C. W. Helstrom, The distribution of the number of crossings of a Gaussian stochastic process, *IEEE Trans. Inform. Theory* IT-3 (1957) 232–237.

[33] T. T. Soong, *Random Differential Equations in Science and Engineering*, Academic Press, 1973, pp. 296–319; MR0451405 (56 #9691).

[34] R. N. Miroshin, Markov and reciprocal stationary Gaussian processes of second order (in Russian), *Teor. Veroyatnost. i Primenen*. 24 (1979) 847–853; Engl. transl. in *Theory Probab. Appl.* 24 (1979) 845–852; MR0550542 (81k:60047).

[35] A. J. Rainal, Passage times of Gaussian noise crossing a time-varying boundary, *IEEE Trans. Inform. Theory* 36 (1990) 1179–1183.

[36] J. T. Barnett and B. Kedem, Zero-crossing rates of functions of Gaussian processes, *IEEE Trans. Inform. Theory* 37 (1991) 1188–1194; comment by G. L. Wise, 38 (1992) 213; MR1111822 (92b:60038).

[37] R. Illsley, The moments of the number of exits from a simply connected region, *Adv. Appl. Probab*. 30 (1998) 167–180; MR1618829 (2000f:60076).

[38] I. Rychlik, On some reliability applications of Rice's formula for the intensity of level crossings, *Extremes* 3 (2000) 331–348; MR1870462 (2002i:60083).

[39] A. Lachal, Application de la théorie des excursions à l'intégrale du mouvement brownien, *Séminaire de Probabilités XXXVII*, ed., J. Azema, M. Emery, M. Ledoux and M. Yor, Lect. Notes in Math. 1832, Springer-Verlag, 2003, pp. 109–195; MR2053045 (2005g:60128).

[40] D. Khoshnevisan and Z. Shi, Chung's law for integrated Brownian motion, *Trans. Amer. Math. Soc*. 350 (1998) 4253–4264; MR1443196 (98m:60056).

[41] X. Chen and W. V. Li, Quadratic functionals and small ball probabilities for the m-fold integrated Brownian motion, *Annals of Probab.* 31 (2003) 1052–1077; MR1964958 (2004j:60078).
[42] F. Gao, J. Hannig and F. Torcaso, Integrated Brownian motions and exact L_2-small balls, *Annals of Probab*. 31 (2003) 1320–1337; MR1989435 (2004k:60104).
[43] W. V. Li and Q.-M. Shao, Gaussian processes: Inequalities, small ball probabilities and applications, *Stochastic Processes: Theory and Methods*, ed. D. N. Shanbhag and C. R. Rao, North-Holland, 2001, pp. 533–597; MR1861734.
[44] M. Goldman, On the first passage of the integrated Wiener process, *Annals Math. Statist*. 42 (1971) 2150–2155; MR0297017 (45 #6075).
[45] P. Groeneboom, G. Jongbloed and J. A. Wellner, Integrated Brownian motion, conditioned to be positive, *Annals of Probab*. 27 (1999) 1283–1303; MR1733148 (2000i:60092).
[46] S. R. Finch, Glaisher-Kinkelin constant: GUE hypothesis, *first volume*, pp. 135–145.
[47] A. Dembo, B. Poonen, Q.-M. Shao and O. Zeitouni, Random polynomials having few or no real zeros, *J. Amer. Math. Soc.* 15 (2002) 857–892; MR1915821 (2003f:60092).
[48] W. V. Li and Q.-M. Shao, A normal comparison inequality and its applications, *Probab. Theory Related Fields* 122 (2002) 494–508; MR1902188 (2003b:60034).
[49] W. V. Li and Q.-M. Shao, Lower tail probabilities for Gaussian processes, *Annals of Probab.* 32 (2004) 216–242; MR2040781 (2005f:60094).
[50] W. V. Li and Q.-M. Shao, Recent developments on lower tail probabilities for Gaussian processes, *Cosmos* 1 (2005) 95–106; MR2329259.
[51] G. Schehr and S. N. Majumdar, Statistics of the number of zero crossings: from random polynomials to the diffusion equation, *Phys. Rev. Lett.* 99 (2007) 060603; arXiv:0705.2648.
[52] G. Schehr and S. N. Majumdar, Real roots of random polynomials and zero crossing properties of diffusion equation, *J. Stat. Phys.* 132 (2008) 235–273; arXiv:0803.4396; MR2415102 (2010a:81078).

4.5 Variants of Brownian Motion

We defined standard Brownian motion $\{W_t : t \geq 0\}$ in [1]. An alternative characterization of the Wiener process involves the limit of random walks. Let ε_1, $\varepsilon_2, \ldots, \varepsilon_n$ be a sequence of independent identically distributed random variables, each possessing mean 0 and variance 1. Let

$$S_0 = 0, \quad S_1 = \varepsilon_1, \quad S_2 = \varepsilon_1 + \varepsilon_2, \quad \ldots, \quad S_n = \sum_{k=1}^{n} \varepsilon_k.$$

Then the random walk $\{S_k\}_{k=1}^n$ approaches Brownian motion on the unit interval in the sense that

$$\frac{S_{\lfloor nt \rfloor}}{\sqrt{n}} \to W_t, \quad 0 \leq t \leq 1$$

as $n \to \infty$, via the functional central limit theorem of Donsker [2, 3]. We are interested in the L_p-norm of Brownian motion

$$\|W\|_p = \begin{cases} \left(\int\limits_0^1 |W_t|^p \, dt \right)^{1/p} & \text{if } 0 < p < \infty, \\ \max\limits_{0 \le t \le 1} |W_t| & \text{if } p = \infty \end{cases}$$

for a number of reasons [4, 5]. Note that $\|W\|_p$ is itself a random variable. A distributional statement about $\|W\|_p$ hence translates into an asymptotic distributional statement about the l_p-norm of the random walk:

$$\mathrm{P}\left(\|W\|_p \le x\right) = \begin{cases} \lim\limits_{n\to\infty} \mathrm{P}\left(\left(\sum\limits_{k=1}^n |S_k|^p \right)^{1/p} \le n^{\frac{1}{2}+\frac{1}{p}} x \right) & \text{if } 0 < p < \infty, \\ \lim\limits_{n\to\infty} \mathrm{P}\left(\max\left\{|S_1|, |S_2|, \ldots, |S_n|\right\} \le n^{1/2} x\right) & \text{if } p = \infty. \end{cases}$$

In the following sections, we will discuss the cases $p = \infty$, 1 and 2 for several variants of Brownian motion. Corresponding problems for all other values of $p > 0$ remain unsolved.

Some preliminary definitions include

$$\delta_m = \frac{\Gamma(m + \frac{1}{2})}{\sqrt{\pi} m!} = \begin{cases} \dfrac{1 \cdot 3 \cdot 5 \cdots (2m-1)}{2 \cdot 4 \cdot 6 \cdots (2m)} & \text{if } m \ge 1, \\ 1 & \text{if } m = 0, \end{cases}$$

$$\mathrm{Ai}(x) = \begin{cases} \frac{1}{3}(-x)^{1/2} \left[J_{-\frac{1}{3}}\left(\frac{2}{3}(-x)^{3/2}\right) + J_{\frac{1}{3}}\left(\frac{2}{3}(-x)^{3/2}\right) \right] & \text{if } x < 0, \\ \frac{1}{3}x^{1/2} \left[I_{-\frac{1}{3}}\left(\frac{2}{3}x^{3/2}\right) - I_{\frac{1}{3}}\left(\frac{2}{3}x^{3/2}\right) \right] & \text{if } x \ge 0, \end{cases}$$

$$K_{\frac{1}{4}}(x) = \frac{\pi}{\sqrt{2}} \left[I_{-\frac{1}{4}}(x) - I_{\frac{1}{4}}(x) \right]$$

where $J_\nu(x)$ and $I_\nu(x)$ are the well-known Bessel functions. Also, for $x > 0$ and $0 < a < b$, let

$$U(a, b, x) = \frac{1}{\Gamma(a)} \int\limits_0^\infty e^{-tx} t^{a-1} (1+t)^{b-a-1} dt.$$

This is called the **confluent hypergeometric function of the second kind** (in contrast to [1]). Finally, define the **Riemann xi function**

$$\xi(z) = \tfrac{1}{2} z(z-1)\pi^{-z/2}\Gamma(\tfrac{1}{2}z)\zeta(z), \quad \mathrm{Re}(z) > 1,$$

which serves as a tantalizing link between Brownian motion and number theory [6]. This can be analytically continued to an entire function via functional equation $\xi(z) = \xi(1-z)$.

4.5.1 *Bridge*

A Brownian bridge $\{X_t : 0 \le t \le 1\}$ has the same distribution as $\{W_t : 0 \le t \le 1\}$, conditioned on $W_1 = 0$. The maximum of $|X_t|$ turns out to be closely allied with the Kolmogorov–Smirnov goodness-of-fit test [7–16]:

$$\mathrm{P}\left(\|X\|_\infty \le x\right) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2x^2} = \frac{\sqrt{2\pi}}{x} \sum_{k=0}^{\infty} e^{-\pi^2(2k+1)^2/(8x^2)}$$

(and the right-hand equality follows via Poisson summation). This distribution has moments

$$\mathrm{E}\left(\|X\|_\infty\right) = \sqrt{\frac{\pi}{2}}\ln(2), \quad \mathrm{E}\left(\|X\|_\infty^2\right) = \frac{\pi^2}{12}$$

and median 0.8275735551.... It also satisfies [17, 18]

$$\mathrm{E}\left(\|X\|_\infty^z\right) = 2\frac{1-2^{1-z}}{z-1}\left(\frac{\pi}{2}\right)^{z/2}\xi(z)$$

for all complex z.

Takács [19, 20], building on Cifarelli [21], Shepp [22], Rice [23] and Johnson & Killeen [24], computed that

$$\mathrm{P}\left(\|X\|_1 \le x\right) = \frac{\sqrt{\pi}}{18^{1/6}x} \sum_{j=1}^{\infty} e^{-u_j} u_j^{-1/3} \operatorname{Ai}\left((3u_j/2)^{2/3}\right)$$

for $x > 0$, where $u_j = (a_j')^3/(27x^2)$ and $0 < a_1' < a_2' < \ldots$ are the zeroes [25] of $\operatorname{Ai}'(-x)$. This distribution has moments

$$\mathrm{E}\left(\|X\|_1\right) = \frac{1}{4}\sqrt{\frac{\pi}{2}}, \quad \mathrm{E}\left(\|X\|_1^2\right) = \frac{7}{60}$$

and median 0.2817802658....

Anderson & Darling [26–29], building on Smirnov [30], obtained that

$$\mathrm{P}\left(\|X\|_2^2 \le x\right) = \frac{1}{\pi\sqrt{x}} \sum_{j=0}^{\infty} \sqrt{4j+1}\, e^{-(4j+1)^2/(16x)} \delta_j K_{1/4}\left((4j+1)^2/(16x)\right),$$

which has moments

$$\mathrm{E}\left(\|X\|_2^2\right) = \frac{1}{6}, \quad \mathrm{E}\left(\|X\|_2^4\right) = \frac{1}{20}$$

and median 0.1188795509.... The L_2-norm, squared, of X_t turns out to be closely allied with the Cramér–von Mises goodness-of-fit test [31–33].

4.5.2 *Excursion*

A Brownian excursion $\{Y_t : 0 \le t \le 1\}$ has the same distribution as $\{W_t : 0 \le t \le 1\}$, conditioned on $W_t > 0$ for all $0 < t < 1$ and $W_1 = 0$.

Chung [34, 35], Kennedy [36] and Durrett & Iglehart [37, 38] showed that

$$\mathrm{P}\left(\|Y\|_\infty \le x\right) = \sum_{k=-\infty}^{\infty}(1-4k^2x^2)e^{-2k^2x^2} = \frac{\sqrt{2}\pi^{5/2}}{x^3}\sum_{k=1}^{\infty}k^2e^{-\pi^2k^2/(2x^2)},$$

which has moments

$$\mathrm{E}\left(\|Y\|_\infty\right) = \sqrt{\frac{\pi}{2}}, \quad \mathrm{E}\left(\|Y\|_\infty^2\right) = \frac{\pi^2}{6},$$

median 1.2234880197..., and also satisfies [17, 18]

$$\mathrm{E}\left(\|Y\|_\infty^z\right) = 2\left(\frac{\pi}{2}\right)^{z/2}\xi(z)$$

for all complex z.

Takács [19, 39], building on Getoor & Sharpe [40], Darling [41], Louchard [42, 43] and Groenboom [44], obtained that

$$\mathrm{P}\left(\|Y\|_1 \le x\right) = \frac{\sqrt{6}}{x}\sum_{j=1}^{\infty}e^{-v_j}v_j^{2/3}U\left(\tfrac{1}{6},\tfrac{4}{3},v_j\right)$$

for $x > 0$, where $v_j = 2a_j^3/(27x^2)$ and $0 < a_1 < a_2 < \ldots$ are the zeroes [25] of $\mathrm{Ai}(-x)$. This distribution has moments

$$\mathrm{E}\left(\|Y\|_1\right) = \sqrt{\frac{\pi}{8}}, \quad \mathrm{E}\left(\|Y\|_1^2\right) = \frac{5}{12}$$

and median 0.6070363869....

The L_2 case seems to be open for Brownian excursion.

4.5.3 *Meander*

A Brownian meander $\{Z_t : 0 \le t \le 1\}$ has the same distribution as $\{W_t : 0 \le t \le 1\}$, conditioned on $W_t > 0$ for all $0 < t < 1$. Note that Z_1 need not be zero.

Durrett & Iglehart [37, 38] computed that

$$\mathrm{P}\left(\|Z\|_\infty \le x\right) = \sum_{k=-\infty}^{\infty}(-1)^k e^{-k^2x^2/2} = \frac{2^{3/2}\sqrt{\pi}}{x}\sum_{k=0}^{\infty}e^{-\pi^2(2k+1)^2/(2x^2)}.$$

Observe that the distribution of $\|Z\|_\infty$ is the same as the distribution of $2\|X\|_\infty$. Hence it has moments

$$\mathrm{E}\left(\|Z\|_\infty\right) = \sqrt{2\pi}\ln(2), \quad \mathrm{E}\left(\|Z\|_\infty^2\right) = \frac{\pi^2}{3}$$

and median 1.6551471103...

Takács [45] proved that

$$\mathrm{P}\left(\|Z\|_1 \le x\right) = \frac{\sqrt{\pi}}{18^{1/6}x}\sum_{j=1}^{\infty} b_j e^{-\tilde{v}_j}\tilde{v}_j^{-1/3}\,\mathrm{Ai}\left((3\tilde{v}_j/2)^{2/3}\right)$$

for $x > 0$, where $\tilde{v}_j = v_j/2$ and v_j, a_j are as before, and where

$$b_j = \frac{a_j}{3\,\mathrm{Ai}'(-a_j)}\left(1 + 3\int_0^{a_j}\mathrm{Ai}(-s)\,ds\right).$$

This distribution has moments

$$\mathrm{E}\left(\|Z\|_1\right) = \frac{3}{4}\sqrt{\frac{\pi}{2}}, \quad \mathrm{E}\left(\|Z\|_1^2\right) = \frac{59}{60}$$

and median 0.8900420723....

The L_2 case seems to be open for Brownian meander.

4.5.4 *Motion*

We return to Brownian motion. Erdős & Kac [46–49] computed that [50]

$$\begin{aligned}
\mathrm{P}\left(\|W\|_\infty \le x\right) &= \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{2k+1}e^{-\pi^2(2k+1)^2/(8x^2)} \\
&= \frac{1}{2}\sum_{k=-\infty}^{\infty}(-1)^k\left[\operatorname{erf}\left(\tfrac{(2k+1)x}{\sqrt{2}}\right) - \operatorname{erf}\left(\tfrac{(2k-1)x}{\sqrt{2}}\right)\right] \\
&= -1 + \sum_{k=-\infty}^{\infty}\left[\operatorname{erf}\left(\tfrac{(4k+1)x}{\sqrt{2}}\right) - \operatorname{erf}\left(\tfrac{(4k-1)x}{\sqrt{2}}\right)\right],
\end{aligned}$$

which has moments

$$\mathrm{E}\left(\|W\|_\infty\right) = \sqrt{\frac{\pi}{2}}, \quad \mathrm{E}\left(\|W\|_\infty^2\right) = 2G$$

and median 1.1489732581.... This is a remarkable appearance of Catalan's constant G!

Takács [51, 52], building on Kac [53] and Schwinger [54], found that

$$\mathrm{P}\left(\|W\|_1 \le y\right) = \sqrt{\frac{3}{\pi}}\int_0^y \frac{1}{x}\sum_{j=1}^{\infty} c_j e^{-\tilde{u}_j}\tilde{u}_j^{2/3}\,U\left(\tfrac{1}{6}, \tfrac{4}{3}, \tilde{u}_j\right)dx$$

for $y > 0$, where $\tilde{u}_j = 2u_j$ and u_j, a_j' are as before, and where

$$c_j = \frac{1}{3a_j'\,\mathrm{Ai}(-a_j')}\left(1 + 3\int_0^{a_j'}\mathrm{Ai}(-s)\,ds\right).$$

This distribution has moments

$$\mathrm{E}\left(\|W\|_1\right) = \frac{4}{3}\frac{1}{\sqrt{2\pi}}, \quad \mathrm{E}\left(\|W\|_1^2\right) = \frac{3}{8}$$

and median 0.4510953819.... We wonder whether the integral for $\mathrm{P}(\|W\|_1 \leq y)$ can be termwise integrated.

Cameron & Martin [46, 55–58] proved that

$$\mathrm{P}\left(\|W\|_2^2 \leq x\right) = \sqrt{2}\sum_{j=0}^{\infty}(-1)^j \delta_j \,\mathrm{erfc}\left(\tfrac{4j+1}{2\sqrt{2x}}\right),$$

which has moments

$$\mathrm{E}\left(\|W\|_2^2\right) = \frac{1}{2}, \quad \mathrm{E}\left(\|W\|_2^4\right) = \frac{7}{12},$$

median 0.2904760595... and Laplace transform

$$\mathrm{E}\left(\exp(-\lambda\,\|W\|_2^2)\right) = \sqrt{\sec\left(\sqrt{-2\lambda}\right)}.$$

We close with several unanswered questions. Define the positive part of W_t to be $W_t^+ = \max\{W_t, 0\}$. Perman & Wellner [59, 60] studied the 1-norm of W_t^+ and found the following double Laplace transform:

$$\int_0^\infty e^{-\mu\lambda}\mathrm{E}\left\{\exp\left(-\sqrt{2}\lambda^{3/2}\,\|W^+\|_1\right)\right\}d\lambda = \frac{\mu^{-1/2}\,\mathrm{Ai}(\mu) + \frac{1}{3} - \int_0^\mu \mathrm{Ai}(s)\,ds}{\sqrt{\mu}\,\mathrm{Ai}(\mu) - \mathrm{Ai}'(\mu)}$$

as well as moments:

$$\mathrm{E}\left(\|W^+\|_1\right) = \frac{2}{3}\frac{1}{\sqrt{2\pi}}, \quad \mathrm{E}\left(\|W^+\|_1^2\right) = \frac{17}{96}.$$

Does an explicit formula for $\mathrm{P}\left(\|W^+\|_1 \leq x\right)$ exist? What can be said for other values of $p > 0$?

Brownian motion with drift (of linear type $W_t + \alpha t$ or parabolic type $W_t - \beta t^2$) would be interesting to report on [61–74]. Of all possible issues, we examine just two. When analysing $W_t + \alpha t$ for $\alpha > 0$, is the formula [64]

$$\int_0^{\pi/2} \frac{\exp(-x\cot(x))\sin(x)}{1+\exp(-\pi\cot(x))}\,dx = \int_0^\infty \left[\frac{1}{2} - \exp(-y\coth(y))\sinh(y)\right]dy$$

valid? Numerics suggest yes – both sides are approximately equal to 0.457524 – a rigorous proof would be good to see someday. The expected maximum value of $W_t - (1/2)t^2$ is [73]

$$\frac{2^{-1/3}}{2\pi i}\int_{-\infty}^{\infty}\frac{z}{\mathrm{Ai}(iz)^2}\,dz = 0.9961930199...$$

(among several integral expressions) and we wonder if similar formulas exist for higher moments.

[1] S. R. Finch, Ornstein-Uhlenbeck process, *this volume*, §4.3.

[2] M. D. Donsker, An invariance principle for certain probability limit theorems, *Memoirs Amer. Math. Soc.* v. 6 (1951) n. 4, 1–12; MR0040613 (12,723a).

[3] P. Billingsley, *Convergence of Probability Measures*, Wiley, 1968, pp. 68–80; MR0233396 (38 #1718).

[4] S. R. Finch, Discrepancy and uniformity, *this volume*, §1.6.

[5] S. R. Finch, Shapes of binary trees, *this volume*, §4.6.

[6] P. Biane, J. Pitman and M. Yor, Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions, *Bull. Amer. Math. Soc.* 38 (2001) 435–465; MR1848256 (2003b:11083).

[7] A. N. Kolmogorov, Sulla determinazione empirica di una legge di distribuzione, *Giornale dell'Istituto Italiano degli Attuari* 4 (1933) 83–91; Engl. transl. in *Selected Works, II. Probability Theory and Mathematical Statistics*, ed. A. N. Shiryayev, Kluwer, 1992, pp. 139–146.

[8] A. N. Kolmogorov, Confidence limits for an unknown distribution function, *Annals Math. Statist*. 12 (1941) 461–463; MR0006684 (4,25f).

[9] N. V. Smirnov, On the deviation of the empirical distribution function (in Russian), *Mat. Sbornik* 6 (1939) 3–26; MR0001483 (1,246b).

[10] N. V. Smirnov, On the estimation of the discrepancy between empirical curves of distribution for two independent samples (in Russian), *Vestnik Moskov. Univ.* 2 (1939) 3–16; MR0002062 (1,345f).

[11] N. V. Smirnov, Table for estimating the goodness of fit of empirical distributions, *Annals Math. Statist.* 19 (1948) 279–281; MR0025109 (9,599j).

[12] W. Feller, On the Kolmogorov-Smirnov limit theorems for empirical distributions, *Annals Math. Statist.* 19 (1948) 177–189; MR0025108 (9,599i).

[13] J. L. Doob, Heuristic approach to the Kolmogorov-Smirnov theorems, *Annals Math. Statist.* 20 (1949) 393–403; MR0030732 (11,43a).

[14] M. D. Donsker, Justification and extension of Doob's heuristic approach to the Komogorov-Smirnov theorems, *Annals Math. Statist.* 23 (1952) 277–281; MR0047288 (13,853n).

[15] J. D. Gibbons, *Nonparametric Statistical Inference*, McGraw-Hill, 1971, pp. 75–85; MR0286223 (44 #3437).

[16] G. Marsaglia, W. W. Tsang and J. Wang, Evaluating Kolmogorov's distribution, *J. Statist. Software* 8 (2003) i. 18.

[17] P. Biane and M. Yor, Valeurs principales associées aux temps locaux browniens, *Bull. Sci. Math.* 111 (1987) 23–101; MR0886959 (88g:60188).

[18] J. Pitman and M. Yor, The law of the maximum of a Bessel bridge, *Elec. J. Probab*. 4 (1999) N15; MR1701890 (2000j:60101).

[19] L. Takács, Random walk processes and their applications in order statistics, *Annals Appl. Probab.* 2 (1992) 435–459; MR1161061 (93h:60114).

[20] L. Tolmatz, Asymptotics of the distribution of the integral of the absolute value of the Brownian bridge for large arguments, *Annals Probab*. 28 (2000) 132–139; MR1756000 (2002a:60068).

[21] D. M. Cifarelli, Contributi intorno ad un test per l'omogeneità tra due campioni, *Giornale degli Economisti e Annali di Economia* 34 (1975) 233–249; MR0433704 (55 #6676).

[22] L. A. Shepp, On the integral of the absolute value of the pinned Wiener process, *Annals Probab*. 10 (1982) 234–239; acknowledgment of priority, 19 (1991) 1397; MR0637389 (82m:60100) and MR1112423 (92e:60159).

[23] S. O. Rice, The integral of the absolute value of the pinned Wiener process—calculation of its probability density by numerical integration, *Annals Probab.* 10 (1982) 240–243; MR0637390 (82m:60099).

[24] B. McK. Johnson and T. Killeen, An explicit formula for the C.D.F. of the L_1 norm of the Brownian bridge, *Annals Probab.* 11 (1983) 807–808; MR0704570 (85b:60038).

[25] S. R. Finch, Airy function zeroes, *this volume*, §2.5.

[26] T. W. Anderson and D. A. Darling, Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, *Annals Math. Statist.* 23 (1952) 193–212; MR0050238 (14,298h).

[27] L. Tolmatz, On the distribution of the square integral of the Brownian bridge, *Annals Probab.* 30 (2002) 253–269; addenda 31 (2003) 530–532; MR1894107 (2003g:60064).

[28] R. Ghomrasni, On distributions associated with the generalized Lévy's stochastic area formula, *Studia Sci. Math. Hungar.* 41 (2004) 93–100; MR2082064 (2005h:60111).

[29] G. Marsaglia and J. Marsaglia, Evaluating the Anderson-Darling distribution, *J. Statist. Software* 9 (2004) i. 2.

[30] N. V. Smirnov, On the distribution of the ω^2-criterion of von Mises (in Russian), *Mat. Sbornik* 2 (1937) 973–993; French transl. abbrev. in *C. R. Séances Acad. Sci.* 202 (1936) 449–452.

[31] M. Knott, The distribution of the Cramér-von Mises statistic for small sample sizes, *J. Royal Statist. Soc. Ser B* 36 (1974) 430–438; MR0365883 (51 #2135).

[32] G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics*, Wiley, 1986, pp. 142–150; MR0838963 (88e:60002).

[33] R. B. D'Agostino and M. A. Stephens (eds.), *Goodness-of-Fit Techniques*, 1986, pp. 111–122; MR0874534 (88c:62075).

[34] K. L. Chung, Maxima in Brownian excursions, *Bull. Amer. Math. Soc.* 81 (1975) 742–745; MR0373035 (51 #9237).

[35] K. L. Chung, Excursions in Brownian motion, *Ark. Mat.* 14 (1976) 155–177; MR0467948 (57 #7791).

[36] D. P. Kennedy, The distribution of the maximum Brownian excursion, *J. Appl. Probab.* 13 (1976) 371–376; MR0402955 (53 #6769).

[37] R. T. Durrett, D. L. Iglehart and D. R. Miller, Weak convergence to Brownian meander and Brownian excursion, *Annals Probab.* 5 (1977) 117–129; MR0436353 (55 #9300).

[38] R. T. Durrett and D. L. Iglehart, Functionals of Brownian meander and Brownian excursion, *Annals Probab.* 5 (1977) 130–135; MR0436354 (55 #9301).

[39] L. Takács, A Bernoulli excursion and its various applications, *Adv. Appl. Probab.* 23 (1991) 557–585; MR1122875 (92m:60057).

[40] R. K. Getoor and M. J. Sharpe, Excursions of Brownian motion and Bessel processes, *Z. Wahrsch. Verw. Gebiete* 47 (1979) 83–106; MR0521534 (80b:60104).

[41] D. A. Darling, On the supremum of a certain Gaussian process, *Annals Probab*. 11 (1983) 803–806; MR0704564 (84f:60058).

[42] G. Louchard, Kac's formula, Levy's local time and Brownian excursion, *J. Appl. Probab.* 21 (1984) 479–499; MR0752014 (86f:60100).

[43] G. Louchard, The Brownian excursion area: a numerical analysis, *Comput. Math. Appl.* 10 (1984) 413–417; erratum, *A*12 (1986) 375; MR0783514 (87b:60123a) and MR0837285 (87b:60123b).

[44] P. Groeneboom, Brownian motion with a parabolic drift and Airy functions, *Probab. Theory Related Fields* 81 (1989) 79–109; MR0981568 (90c:60052).

[45] L. Takács, Limit distributions for the Bernoulli meander, *J. Appl. Probab.* 32 (1995) 375–395; MR1334893 (96c:60087).

[46] P. Erdős and M. Kac, On certain limit theorems of the theory of probability, *Bull. Amer. Math. Soc*. 52 (1946) 292–302; MR0015705 (7,459b).

[47] K. L. Chung, On the maximum partial sums of sequences of independent random variables, *Trans. Amer. Math. Soc.* 64 (1948) 205–233; MR0026274 (10,132b).

[48] M. Kac, On deviations between theoretical and empirical distributions, *Proc. Nat. Acad. Sci. USA* 35 (1949) 252–257; MR0029490 (10,614b).

[49] A. Rényi, On the distribution function $L(z)$ (in Hungarian), *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 2 (1957) 43–50; Engl. transl. in *Selected Transl. Math. Statist. and Probab.,* v. 4, Inst. Math. Stat. and Amer. Math. Soc., 1963, pp. 219–224; MR0099710 (20 #6148).

[50] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[51] L. Takács, On the distribution of the integral of the absolute value of the Brownian motion, *Annals Appl. Probab.* 3 (1993) 186–197; MR1202522 (93i:60154).

[52] L. Tolmatz, Asymptotics of the distribution of the integral of the absolute value of the Brownian bridge for large arguments, *Annals Probab.* 28 (2000) 132–139; MR1756000 (2002a:60068).

[53] M. Kac, On the average of a certain Wiener functional and a related limit theorem in calculus of probability, *Trans. Amer. Math. Soc.* 59 (1946) 401–414; MR0016570 (8,37b).

[54] M. Kac, *Probability, Number Theory, and Statistical Physics. Selected Papers*, ed. K. Baclawski and M. D. Donsker, MIT Press, 1979, pp. ix–xxiii; MR0521078 (80i:01016).

[55] R. H. Cameron and W. T. Martin, The Wiener measure of Hilbert neighborhoods in the space of real continuous functions, *J. Math. Phys. Mass. Inst. Tech.* 23 (1944) 195–209; MR0011174 (6,132a).

[56] R. Fortet, Quelques travaux récents sur le mouvement brownien, *Annales de l'Institut Henri Poincaré* 11 (1949) 175–226; MR0035411 (11,731e).

[57] P. Lévy, Wiener's random function, and other Laplacian random functions, *Proc. Second Berkeley Symp. Math. Stat. Probab.*, ed. J. Neyman, Univ. of Calif. Press, 1951, pp. 171–187; MR0044774 (13,476b).

[58] A. N. Borodin and P. Salminen, *Handbook of Brownian motion—Facts and Formulae*, 2nd ed., Birkhäuser Verlag, 2002, pp. 168–169, 333, 343–344; MR1912205 (2003g:60001).

[59] M. Perman and J. A. Wellner, On the distribution of Brownian areas, *Annals Appl. Probab*. 6 (1996) 1091–1111; MR1422979 (97m:60119).

[60] L. Tolmatz, Asymptotics of the distribution of the integral of the positive part of the Brownian bridge for large arguments, *J. Math. Anal. Appl.* 304 (2005) 668–682; MR2126559 (2006b:60179).

[61] L. Takács, On a generalization of the arc-sine law, *Annals Appl. Probab.* 6 (1996) 1035–1040; MR1410128 (97k:60195).

[62] R. A. Doney and M. Yor, On a formula of Takács for Brownian motion with drift, *J. Appl. Probab.* 35 (1998) 272–280; MR1641856 (99f:60143).

[63] E. Buffet, On the time of the maximum of Brownian motion with drift, *J. Appl. Math. Stochastic Anal.* 16 (2003) 201–207; MR2010509 (2004k:60224).

[64] M. Magdon-Ismail, A. F. Atiya, A. Pratap and Y. S. Abu-Mostafa, On the maximum drawdown of a Brownian motion, *J. Appl. Probab.* 41 (2004) 147–161; MR2036278.

[65] E. Tanré and P. Vallois, Range of Brownian motion with drift, *J. Theoret. Probab.* 19 (2006) 45–69; MR2256479 (2007j:60135).

[66] P. Salminen and P. Vallois, On maximum increase and decrease of Brownian motion, *Annales de l'Institut Henri Poincaré Sect. B* 43 (2007) 655–676; arXiv:math/0512440; MR3252425.

[67] O. Hadjiliadis and J. Vecer, Drawdowns preceding rallies in the Brownian motion model, *Quantitative Finance* 6 (2006) 403–409; MR2261219 (2007m:60247).

[68] L. Pospisil, J. Vecer and O. Hadjiliadis, Formulas for stopped diffusion processes with stopping times based on drawdowns and drawups, *Stochastic Process. Appl.* 119 (2009) 2563–2578; MR2532216 (2010j:60100).

[69] C. Vardar-Acar, C. L. Zirbel and G. J. Székely, On the correlation of the supremum and the infimum and of maximum gain and maximum loss of Brownian motion with drift, *J. Comput. Appl. Math.* 248 (2013), 61–75; MR3029217.

[70] P. Groeneboom, The maximum of Brownian motion minus a parabola, *Elec. J. Probab.* 15 (2010) 1930–1937; arXiv:1011.0022; MR2738343 (2011k:60274).

[71] P. Groeneboom and N. Temme, The tail of the maximum of Brownian motion minus a parabola, *Elec. Commun. Probab.* 16 (2011) 458–466; arXiv:1011.3972; MR2831084 (2012h:60247).

[72] P. Groeneboom, S. Lalley and N. Temme, Chernoff's distribution and differential equations of parabolic and Airy type, *J. Math. Anal. Appl.* 423 (2015) 1804–1824; arXiv:1305.6053; MR3278229.

[73] S. Janson, G. Louchard and A. Martin-Löf, The maximum of Brownian motion with parabolic drift, *Elec. J. Probab.* 15 (2010) 1893–1929; MR2738342 (2012b:60271).

[74] S. Janson, Moments of the location of the maximum of Brownian motion with parabolic drift, *Elec. Commun. Probab.* 18 (2013) N15; MR3037213.

4.6 Shapes of Binary Trees

This is a sequel to our treatment of various attributes of trees [1], expressed in the language of probability. Let $\{Y_t : 0 \le t \le 1\}$ be standard Brownian excursion. Define the L_p-norm

$$\|Y\|_p = \begin{cases} \left(\int\limits_0^1 |Y_t|^p \, dt \right)^{1/p} & \text{if } 0 < p < \infty, \\ \max\limits_{0 \le t \le 1} |Y_t| & \text{if } p = \infty \end{cases}$$

and a (new) seminorm

$$\langle Y\rangle_p = \begin{cases} \left(\int_0^1 \int_0^v \left| Y_u + Y_v - 2\min_{u\le t\le v} Y_t\right|^p du\,dv\right)^{1/p} & \text{if } 0<p<\infty, \\ \max_{0\le u<v\le 1} \left| Y_u + Y_v - 2\min_{u\le t\le v} Y_t\right| & \text{if } p=\infty. \end{cases}$$

We examined $\| Y\|_p$ earlier [2]; $\langle Y\rangle_p$ is a less familiar random variable but nevertheless important in the study of trees. Note that $\langle Y\rangle_p$ is not a norm since, for any constant c, $\langle c\rangle_p = 0$ even if $c \neq 0$.

Let T be an ordered (strongly) binary tree with $N = 2n+1$ vertices. The **distance** between two vertices of T is the number of edges in the shortest path connecting them. The **height** of a vertex is the number of edges in the shortest path connecting the vertex and the root.

The **Wiener index** $d_1(T)$ is the sum of all $\binom{N}{2}$ distances between pairs of distinct vertices of T, and the **diameter** $d_\infty(T)$ is the maximum such distance. If $\delta(v,w)$ denotes the distance between vertices v and w, then

$$d_\lambda(T) = \left(\frac{1}{2}\sum_{v,w} \delta(v,w)^\lambda\right)^{1/\lambda}, \quad \lambda>0,$$

includes both the Wiener index and diameter as special cases.

The **internal path length** $h_1(T)$ of a tree is the sum of all N heights of vertices of T, and the **height** $h_\infty(T)$ is the maximum such height. Let o denote the root of T. The generalization

$$h_\lambda(T) = \left(\sum_{v} \delta(v,o)^\lambda\right)^{1/\lambda}, \quad \lambda>0,$$

includes both the internal path length and height as special cases. If we restrict attention to only those $n+1$ vertices $\hat{v}_k$ that are leaves (terminal nodes) of T, listed from left to right, then a sequence $\delta(\hat{v}_1,o), \delta(\hat{v}_2,o), \ldots, \delta(\hat{v}_{n+1},o)$ emerges. This is called the **contour** of T.

The **width** $w_\infty(T)$ of a tree is the maximum of $\zeta_l(T)$ over all $l \ge 0$, where $\zeta_l(T)$ is the number of vertices of height l in T. Note that

$$w_\lambda(T) = \left(\sum_{l=0}^{h_\infty(T)} \zeta_l(T)^\lambda\right)^{1/\lambda}, \quad \lambda>0,$$

includes the trivial case $w_1(T) = N$. The sequence $\zeta_0(T), \zeta_1(T), \ldots, \zeta_{h_\infty}(T)$ is known as the **profile** of T.

4.6.1 *Uniform Combinatorial Model*

In this model, the $\binom{2n}{n}/(n+1)$ ordered binary trees are weighted with equal probability, where $N=2n+1$ is fixed.

Janson [3] determined the joint distribution of internal path length and Wiener index:

$$\left(\frac{h_1(T)}{2N^{3/2}}, \frac{d_1(T)}{2N^{5/2}}\right) \to (\|Y\|_1, \langle Y\rangle_1)$$

as $N\to\infty$. The marginal distribution of $h_1(T)$ was obtained earlier by Takács [4–6]; the result for $d_1(T)$ is apparently new. No explicit formula for $\mathrm{P}(\langle Y\rangle_1 \le x)$ is known; see [2] for the corresponding result for $\|Y\|_1$. We have expected values

$$\mathrm{E}(\|Y\|_1) = \frac{1}{2}\sqrt{\frac{\pi}{2}}, \quad \mathrm{E}(\langle Y\rangle_1) = \frac{1}{4}\sqrt{\frac{\pi}{2}}$$

and correlation coefficient

$$\frac{\mathrm{Cov}(\|Y\|_1, \langle Y\rangle_1)}{\sqrt{\mathrm{Var}(\|Y\|_1)}\sqrt{\mathrm{Var}(\langle Y\rangle_1)}} = \sqrt{\frac{48-15\pi}{50-15\pi}} = 0.5519206030...$$

As an aside, we mention that $\|Y\|_1 - \langle Y\rangle_1 \ge 0$ always. Underlying the joint moment [3]

$$\mathrm{E}\left(\|Y\|_1^k (\|Y\|_1 - \langle Y\rangle_1)^l\right) = \frac{k!l!\sqrt{\pi}}{2^{(7k+9l-4)/2}\Gamma((3k+5l-1)/2)} a_{k,l}$$

is the following interesting quadratic recurrence [7–12]:

$$\begin{aligned} a_{k,l} &= 2(3k+5l-4)a_{k-1,l} + 2(3k+5l-6)(3k+5l-4)a_{k,l-1} \\ &\quad + \sum_{0<i+j<k+l} a_{i,j}a_{k-i,l-j} \end{aligned}$$

with $a_{0,0}=-1/2$, $a_{1,0}=1=a_{0,1}$ and $a_{k,l}=0$ when $k<0$ or $l<0$. All $a_{k,l}$ but $a_{0,0}$ are positive integers when $k\ge 0$ and $l\ge 0$. Applications include the enumeration of connected graphs with n vertices and $n+m$ edges. We have asymptotics [3, 13]

$$a_{k,0} \sim \frac{1}{2\pi}6^k(k-1)!, \quad a_{0,l} \sim C\cdot 50^l\left((l-1)!\right)^2,$$

where the precise identity of the constant

$$C = \frac{\sqrt{15}}{20\pi^2} = 0.0196207628... = \frac{1}{50}(0.9810381421...) = \frac{1}{50.9664179720...}$$

remained masked until its recent unveiling by Kotěšovec [12].

Chassaing, Marckert & Yor [14] determined the joint distribution of height and width:

$$\left(\frac{h_\infty(T)}{N^{1/2}}, \frac{w_\infty(T)}{N^{1/2}}\right) \to \left(\int_0^1 \frac{dt}{Y_t}, \|Y\|_\infty\right)$$

as $N \to \infty$. The marginal distribution of height was obtained earlier by Rényi & Szekeres and Stepanov [15–24]; earlier works on width include [25–30]. It turns out that the marginal distributions are identical (up to a factor of 2) and that this is the first of several theta distributions [31] we will see here:

$$\mathrm{P}\left(\frac{1}{2}\int_0^1 \frac{dt}{Y_t} \le x\right) = \mathrm{P}\left(\|Y\|_\infty \le x\right) = \frac{\sqrt{2}\pi^{5/2}}{x^3}\sum_{k=1}^{\infty} k^2 e^{-\pi^2 k^2/(2x^2)}.$$

The expected values are thus equal:

$$\mathrm{E}\left(\frac{1}{2}\int_0^1 \frac{dt}{Y_t}\right) = \mathrm{E}\left(\|Y\|_\infty\right) = \sqrt{\frac{\pi}{2}}.$$

Rényi & Szekeres also computed the location of the maximum of the probability density [15]:

$$\mathrm{mode}\left(\|Y\|_\infty\right) = \frac{1}{2}(2.3151543618...) = \frac{1}{2}\sqrt{\frac{2}{0.3731385248...}}.$$

Returning to the joint distribution formula, it is clear that $h_\infty(T)$ and $w_\infty(T)$ are negatively correlated. A numerical estimate for the correlation coefficient was open until recently [14, 32] ; Janson [33] computed that

$$\begin{aligned}
\mathrm{E}\left(\int_0^1 (1/Y_t)\,dt \cdot \|Y\|_\infty\right) &= 1 + \sum_{m=1}^{\infty} \frac{\ln\left[m(m+1)\right]}{m(m+1)} \\
&= 1 + 2.0462774528... \\
&= \pi - 0.0953152007...,
\end{aligned}$$

$$\begin{aligned}
\frac{\mathrm{Cov}\left(\int_0^1 (1/Y_t)\,dt, \|Y\|_\infty\right)}{\sqrt{\mathrm{Var}\left(\int_0^1 (1/Y_t)\,dt\right)}\sqrt{\mathrm{Var}\left(\|Y\|_\infty\right)}} &= \frac{3(3.0462774528... - \pi)}{\pi(\pi - 3)} \\
&= -0.6428251027...
\end{aligned}$$

and the infinite series [34, 35] is a Lüroth analog of Lévy's constant $\pi^2/(6\ln(2))$. Why is the joint distribution of height and width of trees related to the ergodic theory of numbers? Such a coincidence does not happen without a reason.

For the generalized height and diameter parameters, we have marginal distributions [3, 14, 36–38]:

$$\frac{h_\lambda(T)}{2N^{(\lambda+2)/(2\lambda)}} \to \|Y\|_\lambda, \quad \frac{d_\lambda(T)}{2N^{(\lambda+4)/(2\lambda)}} \to \langle Y\rangle_\lambda$$

as $N \to \infty$. The latter includes the special cases of Wiener index ($\lambda = 1$, as mentioned before) and diameter ($\lambda = \infty$):

$$\mathrm{P}(\langle Y\rangle_\infty \le x) = \frac{1024\sqrt{2}\pi^{5/2}}{3x^9}\sum_{k=1}^{\infty} k^2\left[\left(3+\pi^2k^2\right)x^4 - 36\pi^2k^2x^2 + 64\pi^4k^4\right] \times e^{-8\pi^2k^2/x^2},$$

which possesses expected value

$$\mathrm{E}\left(\langle Y\rangle_\infty\right) = \frac{4}{3}\sqrt{2\pi}$$

and maximum location [36]

$$\mathrm{mode}\left(\langle Y\rangle_\infty\right) = 3.2015131492... = \sqrt{\frac{8}{0.7805116813...}}.$$

Nothing is known for other values of λ (even $\lambda = 2$ seems to have been neglected). It would also be good to learn the value of the correlation coefficient of $d_\infty(T)$ and $h_\infty(T)$, or of $d_\infty(T)$ and $w_\infty(T)$.

Consider finally the minimum height $\eta(T)$ of a leaf, that is,

$$\eta(T) = \min_{1\le k\le n+1} \delta(\hat{v}_k, o),$$

and the height $\delta(\hat{v}_{\lceil n/2\rceil}, o)$ of the central leaf. It is known that

$$\mathrm{E}(\eta) \to \sum_{k=1}^{\infty} 2^{k+1-2^k} = 1.5629882961...$$

as $N \to \infty$ [39, 40]. It is also known that [41–44]

$$\sqrt{n}\mathrm{P}\left(\frac{\delta(\hat{v}_{\lceil n/2\rceil}, o)}{\sqrt{n}} \le x\right) \to \frac{1}{2\sqrt{\pi}}\int_0^x t^2e^{-t^2/4}dt = \mathrm{P}\left(\sqrt{X_1^2 + X_2^2 + X_3^2} \le \frac{x}{\sqrt{2}}\right),$$

the Maxwell distribution from thermodynamics, where X_1, X_2, X_3 are independent standard normal variables. (This can also be written in terms of the chi square distribution with 3 degrees of freedom.) Can these results be related to Brownian excursion in some way? More on the properties of leaves of T would be good to see.

4.6.2 *Critical Galton–Watson Model*

In this model, the size $N = 2n + 1$ is free to vary: All ordered binary trees are included but with weighting 2^{-N}. (We omit subcritical and supercritical cases for reasons of space.)

Let T be a random tree. The probability that T has precisely N vertices is clearly [45]

$$\frac{1}{n+1}\binom{2n}{n}2^{-N} \sim \sqrt{\frac{2}{\pi}}N^{-3/2};$$

hence the expected number of vertices of T is infinite. We examine this result in another way. If

$$\nu_l = \sum_{k=0}^{l} \zeta_k$$

where ζ_k is the number of vertices of height k in T, then $\mathrm{E}(\nu_l) = l+1$ and $\mathrm{Var}(\nu_l) = (2l+1)(l+1)l/6$, both which $\to \infty$ as $l \to \infty$. More complicated conditional distributions are due to Pakes [46, 47]:

$$\lim_{l\to\infty} \mathrm{P}\left(\frac{\nu_l}{l^2} \le x \,|\, \zeta_l > 0\right) = \int_0^x f(t)\, dt,$$

$$\lim_{l\to\infty} \mathrm{P}\left(\frac{\nu_l}{l^2} \le x \,|\, \zeta_m > 0 \text{ for all positive integers } m\right) = \int_0^x g(t)\, dt,$$

where the first density function is given by

$$f(t) = \frac{2}{\sqrt{2\pi}t^{3/2}} \sum_{k=0}^{\infty} \left(\frac{(2k+1)^2}{t} - 1\right) \exp\left(-\frac{(2k+1)^2}{2t}\right)$$

with mean $1/3$, variance $2/45$, and Laplace transform

$$\int_0^\infty e^{-st} f(t)\, dt = \sqrt{2s}\,\operatorname{csch}\left(\sqrt{2s}\right).$$

The second density function is not explicitly known, but has mean $1/2$, variance $1/12$ and satisfies

$$\int_0^\infty e^{-st} g(t)\, dt = \operatorname{sech}^2\left(\sqrt{\frac{s}{2}}\right).$$

Consequently $g(t)$ is the convolution of $\tilde{g}(t)$ with itself, where

$$\tilde{g}(t) = \frac{1}{\sqrt{2\pi}t^{3/2}} \sum_{k=0}^{\infty} (-1)^k (2k+1) \exp\left(-\frac{(2k+1)^2}{8t}\right),$$

but this appears to be as far as we can go.

Define T_l to be the subtree of T consisting of all ν_l vertices up to and including height l. We have the parameters $d_\lambda(T_l)$, $h_\lambda(T_l)$ and $w_\lambda(T_l)$ available for study, but little seems to be known. Of course, $w_1(T_l) = \nu_l$. Athreya [48], building on [49–51], proved that $\mathrm{E}\left(w_\infty(T_l)\right) \sim \ln(l)$ as $l \to \infty$, which contrasts nicely with the fact that $\mathrm{P}\left(\zeta_k = 0\right) \to 1$ as $k \to \infty$. See also [52–58]. Kesten, Ney & Spitzer [59–61] demonstrated that $\mathrm{P}\left(h_\infty(T_l) = j\right) \sim 2/j^2$ as $j \to \infty$; further references include

[62–64]. Can exact distributional results be found? What about other values of λ? Is anything known about diameter for Galton–Watson trees?

Just as the limit behavior for the uniform model is related to Brownian excursion, the limit behavior for the critical GW model is related to what is known as the two-sided three-dimensional Bessel process $\{B_t : -\infty < t < \infty\}$. That is, $\{B_t : t \geq 0\}$ and $\{B_{-t} : t \geq 0\}$ are independent copies of standard 3D radial Brownian motion $\sqrt{W_{1,t}^2 + W_{2,t}^2 + W_{3,t}^2}$, each starting from zero [38, 65]. It would be good to learn more about the concrete distributional results arising from this correspondence.

4.6.3 *Leaves of Maximum Height*

Our closing remarks are concerned not with binary trees, but instead with labeled rooted trees. Choose such a tree T with N vertices uniformly out of the N^{N-2} possibilities (we agree that the root is labeled 1). Out of all possible parameters (suitably generalized), we mention only the minimum height $\eta(T)$ of a leaf. Meir & Moon [40] computed that

$$\mathrm{E}(\eta) \to 9 \sum_{k=1}^{\infty} \frac{1}{4^k(1+2\cdot 4^{-k})^2} = 1.6229713847...$$

as $N \to \infty$. A more difficult problem involves counting the leaves $\hat{v}_k$ at prescribed distance from the root. Kesten & Pittel [66] proved, for leaves of maximum height, that there exists a probability distribution q_l such that

$$\lim_{N\to\infty} \mathrm{P}\left(\zeta_{h_\infty}(T) = l\right) = q_l, \quad l \geq 1.$$

Further, q_l is the unique nonnegative solution of the system of equations

$$l!e^l q_l = \sum_{k=1}^{\infty} k^l q_k, \quad \sum_{k=1}^{\infty} q_k = 1$$

and thus $q_1 = 0.602...$, $q_2 = 0.248...$, $q_3 = 0.094...$, $q_4 = 0.035...$ with mean 1.636... and standard deviation 0.995.... No exact expressions for these quantities are known. What is the corresponding distribution for the uniform ordered binary tree case?

[1] S. R. Finch, Otter's tree enumeration constants, *first volume*, pp. 295–316.
[2] S. R. Finch, Variants of Brownian motion, *this volume*, §4.5.
[3] S. Janson, The Wiener index of simply generated random trees, *Random Structures Algorithms* 22 (2003) 337–358; MR1980963 (2004b:05186).
[4] L. Takács, On the total heights of random rooted trees, *J. Appl. Probab.* 29 (1992) 543–556; MR1174430 (93g:05127).
[5] L. Takács, The asymptotic distribution of the total heights of random rooted trees, *Acta Sci. Math. (Szeged)* 57 (1993) 613–625; MR1243312 (94i:05086).

[6] L. Takács, On the total heights of random rooted binary trees, *J. Combin. Theory Ser. B* 61 (1994) 155–166; MR1280604 (95e:05035).

[7] E. M. Wright, The number of connected sparsely edged graphs, *J. Graph Theory* 1 (1977) 317–330; MR0463026 (57 #2990).

[8] V. A. Voblyĭ, Wright and Stepanov-Wright coefficients (in Russian), *Mat. Zametki* 42 (1987) 854–862, 911; Engl. transl. in *Math. Notes* 42 (1987) 969–974; MR0934817 (89d:05103).

[9] J. Spencer, Enumerating graphs and Brownian motion, *Comm. Pure Appl. Math.* 50 (1997) 291–294; MR1431811 (97k:05105).

[10] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent, *Annals of Probab.* 25 (1997) 812–854; MR1434128 (98d:60019).

[11] P. Flajolet and G. Louchard, Analytic variations on the Airy distribution, *Algorithmica* 31 (2001) 361–377; MR1855255 (2002j:68115).

[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A062980 and A094199.

[13] S. Janson and P. Chassaing, The center of mass of the ISE and the Wiener index of trees, *Elec. Comm. Probab.* 9 (2004) 178–187; arXiv:math.PR/0309284; MR2108865 (2005k:60303).

[14] P. Chassaing, J. F. Marckert and M. Yor, The height and width of simple trees, *Mathematics and Computer Science. Algorithms, Trees, Combinatorics and Probabilities*, Proc. 2000 Versailles colloq., ed. D. Gardy and A. Mokkadem, Birkhäuser, 2000, pp. 17–30; MR1798284 (2001j:68095).

[15] A. Rényi and G. Szekeres, On the height of trees, *J. Austral. Math. Soc.* 7 (1967) 497–507; MR0219440 (36 #2522).

[16] V. E. Stepanov, The distribution of the number of vertices in strata of a random tree (in Russian), *Teor. Verojatnost. i Primenen.* 14 (1969) 64–77; Engl. transl. in *Theory Probab. Appl.* 14 (1969) 65–78; MR0246344 (39 #7648).

[17] N. G. de Bruijn, D. E. Knuth and S. O. Rice, The average height of planted plane trees, *Graph Theory and Computing*, ed. R. C. Read, Academic Press, 1972, pp. 15–22; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 215–223; MR0505710 (58 #21737).

[18] V. F. Kolchin, The extinction time of a branching process and the height of a random tree (in Russian), *Mat. Zametki* 24 (1978) 859–870, 894; Engl. transl. in *Math. Notes* 24 (1978) 954–961; MR0522419 (80i:60118).

[19] V. F. Kolchin, *Random Mappings*, Optimization Software Inc., 1986, pp. 135–146; MR0865130 (88a:60022).

[20] P. Flajolet and A. Odlyzko, The average height of binary trees and other simple trees, *J. Comput. System Sci.* 25 (1982) 171–213; MR0680517 (84a:68056).

[21] G. G. Brown and B. O. Shubert, On random binary trees, *Math. Oper. Res.* 9 (1984) 43–65; MR0736638 (86c:05099).

[22] R. Durrett, H. Kesten and E. Waymire, On weighted heights of random trees, *J. Theoret. Probab.* 4 (1991) 223–237; MR1088403 (91m:60021).

[23] P. Flajolet, Z.-C. Gao, A. Odlyzko and B. Richmond, The distribution of heights of binary trees and other simple trees, *Combin. Probab. Comput*. 2 (1993) 145–156; MR1249127 (94k:05061).

[24] T. Luczak, A greedy algorithm estimating the height of random trees, *SIAM J. Discrete Math.* 11 (1998) 318–329; MR1617161 (99d:68117).

[25] A. M. Odlyzko and H. S. Wilf, Bandwidths and profiles of trees, *J. Combin. Theory Ser. B* 42 (1987) 348–370; MR0888687 (89b:05077).

[26] L. Takács, Limit distributions for queues and random rooted trees, *J. Appl. Math. Stochastic Anal.* 6 (1993) 189–216; MR1238599 (94m:60194).

[27] M. Drmota and B. Gittenberger, On the profile of random trees, *Random Structures Algorithms* 10 (1997) 421–451; MR1608230 (99c:05176).

[28] G. Kersting, On the height profile of a conditioned Galton-Watson tree, arXiv:1101.3656.

[29] P. Chassaing and J.-F. Marckert, Parking functions, empirical processes, and the width of rooted labeled trees, *Elec. J. Combin.* 8 (2001) R14; MR1814521 (2002b:05037).

[30] M. Drmota and B. Gittenberger, The width of Galton-Watson trees conditioned by the size, *Discrete Math. Theor. Comput. Sci.* 6 (2004) 387–400; MR2081482 (2005f:60180).

[31] L. Devroye, Simulating theta random variates, *Statist. Probab. Lett.* 31 (1997) 275–279; MR1439946 (97j:65009).

[32] C. Donati-Martin, Some remarks about the identity in law for the Bessel bridge $\int_0^1 \frac{ds}{r(s)} \overset{(law)}{=} 2\sup_{s\le 1} r(s)$, *Studia Sci. Math. Hungar.* 37 (2001) 131–144; MR1834327 (2002j:60141).

[33] S. Janson, On the asymptotic joint distribution of height and width in random trees, *Studia Sci. Math. Hungar.* 45 (2008) 451–467; MR2641443 (2011d:60239).

[34] S. R. Finch, Khintchine-Lévy constants, *first volume*, pp. 59–65.

[35] W. Bosma, K. Dajani and C. Kraaikamp, Entropy and counting correct digits, Dept. of Math. Report 9925 (1999), Univ. of Nijmegen.

[36] G. Szekeres, Distribution of labelled trees by diameter, *Combinatorial Mathematics X*, Proc. 1982 Adelaide conf., ed. L. R. A. Casse, Lect. Notes in Math. 1036, Springer-Verlag, 1983, pp. 392–397; MR0731595 (85b:05069).

[37] T. Luczak, The number of trees with large diameter, *J. Austral. Math. Soc. Ser. A* 58 (1995) 298–311; MR1329866 (96c:05086).

[38] D. Aldous, The continuum random tree. II. An overview, *Stochastic Analysis*, Proc. 1990 Durham conf., ed. M.T. Barlow and N.H. Bingham, Cambridge Univ. Press, 1991, pp. 23–70; MR1166406 (93f:60010).

[39] G. R. Grimmett, Random labelled trees and their branching networks, *J. Austral. Math. Soc. Ser. A* 30 (1980/81) 229–237; MR0607933 (82g:05042).

[40] A. Meir and J. W. Moon, Terminal path numbers for certain families of trees, *J. Austral. Math. Soc. Ser. A* 31 (1981) 429–438; MR0638271 (82m:05039).

[41] W. Gutjahr and G. C. Pflug, The asymptotic distribution of leaf heights in binary trees, *Graphs Combin.* 8 (1992) 243–251; MR1185403 (93i:05049).

[42] W. Gutjahr and G. C. Pflug, The limiting common distribution of two leaf heights in a random binary tree, *RAIRO Inform. Théor. Appl.* 26 (1992) 1–18; MR1155342 (93b:68032).

[43] W. Gutjahr and G. C. Pflug, The asymptotic contour process of a binary tree is a Brownian excursion, *Stochastic Process. Appl.* 41 (1992) 69–89; MR1162719 (93i:60017).

[44] M. Drmota, Distribution of the height of leaves of rooted trees (in Russian), *Diskret. Mat.* 6 (1994) 67–82; Engl. transl. in *Discrete Math. Appl.* 4 (1994) 45–58; MR1273234 (95c:05045).

[45] R. Otter, The multiplicative process, *Annals of Math. Statist.* 20 (1949) 206–224; MR0030716 (11,41a).

[46] A. G. Pakes, Some limit theorems for the total progeny of a branching process, *Adv. Appl. Probab.* 3 (1971) 176–192; MR0283892 (44 #1122).

[47] A. G. Pakes, The laws of some random series of independent summands, *Advances in the Theory and Practice of Statistics. A Volume in Honor of Samuel Kotz*, ed. N. L. Johnson and N. Balakrishnan, Wiley, 1997, pp. 499–516; MR1481190 (99g:60027).

[48] K. B. Athreya, On the maximum sequence in a critical branching process, *Annals of Probab.* 16 (1988) 502–507; MR0929060 (89f:60098).

[49] H. Weiner, Moments of the maximum in a critical branching process, *J. Appl. Probab.* 21 (1984) 920–923; MR0766828 (86a:60112).

[50] K. Kämmerle and H.-J. Schuh, The maximum in critical Galton-Watson and birth and death processes, *J. Appl. Probab.* 3 (1986) 601–613; MR0855368 (87m:60188).

[51] A. G. Pakes, Remarks on the maxima of a martingale sequence with application to the simple critical branching process, *J. Appl. Probab.* 24 (1987) 768–772; MR0900316 (88g:60113).

[52] J. Bishir, Maximum population size in a branching process, *Biometrics* 18 (1962) 394–403; MR0145637 (26 #3166).

[53] S. R. Adke, The maximum population size in the first N generations of a branching process, *Biometrics* 20 (1964) 649–651; MR0181013 (31 #5243).

[54] T. Lindvall, On the maximum of a branching process, *Scand. J. Statist*. 3 (1976) 209–214; MR0426188 (54 #14134).

[55] O. Nerman, On the maximal generation size of a non-critical Galton-Watson process, *Scand. J. Statist.* 4 (1977) 131–135; MR0474531 (57 #14169).

[56] A. Spataru, Equations concerning maxima of critical Galton-Watson processes, *Rev. Roumaine Math. Pures Appl.* 37 (1992) 935–937; MR1195094 (94f:60110).

[57] V. A. Vatutin and V. A. Topchii, The maximum of critical Galton-Watson processes, and left-continuous random walks (in Russian), *Teor. Veroyatnost. i Primenen.* 42 (1997) 21–34; Engl. transl. in *Theory Probab. Appl.* 42 (1997) 17–27; MR1453327 (98g:60154).

[58] E. M. Bondarenko and V. A. Topchii, Estimates for the expectation of the maximum of a critical Galton-Watson process on a finite interval (in Russian), *Sibirsk. Mat. Zh.* 42 (2001) 249–257; Engl. transl. in *Siberian Math. J.* 42 (2001) 209–216; MR1833155 (2002b:60145).

[59] H. Kesten, P. Ney and F. Spitzer, The Galton-Watson process with mean one and finite variance, *Teor. Verojatnost. i Primenen.* 11 (1966) 579–611; Engl. transl. in *Theory Probab. Appl*. 11 (1966) 513–540; MR0207052 (34 #6868).

[60] E. Seneta, The Galton-Watson process with mean one, *J. Appl. Probab.* 4 (1967) 489–495; MR0228075 (37 #3659).

[61] K. B. Athreya and P. Ney, *Branching Processes*, Springer-Verlag, 1972, pp. 19–23; MR0373040 (51 #9242).

[62] T. E. Harris, Branching processes, *Annals of Math. Statist.* 19 (1948) 474–494; MR0027465 (10,311b).

[63] L. Castoldi, Sulla distribuzione dei tempi di estinzione nelle discendenze biologiche, *Boll. Un. Mat. Ital.* 11 (1956) 158–167; MR0079380 (18,77d).

[64] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications*, McGraw-Hill, 1960, pp. 26–27; MR0112177 (22 #3032).

[65] B. Gittenberger, Convergence of branching processes to the local time of a Bessel process, *Random Structures Algorithms* 13 (1998) 423–438; MR1662793 (99k:60213).

[66] H. Kesten and B. Pittel, A local limit theorem for the number of nodes, the height, and the number of final leaves in a critical branching process tree, *Random Structures Algorithms* 8 (1996) 243–299; MR1603252 (98m:60137).

4.7 Expected Lifetimes and Inradii

In earlier essays [1, 2], we examined 1-dimensional Brownian motion starting at 0; here, we generalize. A d-dimensional stochastic process $\{W_t : t \geq 0\}$ is a **Brownian motion** with *arbitrary* starting point W_0 if the component processes

$$W_{t,1} - W_{0,1},\ W_{t,2} - W_{0,2},\ \ldots,\ W_{t,d} - W_{0,d}$$

are independent 1-dimensional Brownian motions starting at 0 and, further, are independent of $W_{0,1}$, $W_{0,2}$, . . ., $W_{0,d}$.

It is remarkable that d-dimensional Brownian motion can be used to represent the solution of the heat PDE [3, 4]:

$$\begin{cases} \dfrac{\partial u}{\partial t} = \dfrac{1}{2}\triangle u, & t \geq 0,\ \xi \in \mathbb{R}^d, \\ u(0,\xi) = f(\xi), & f : \mathbb{R}^d \to \mathbb{R} \text{ piecewise continuous} \end{cases}$$

in the following sense:

$$\begin{aligned} u(t,\xi) &= \mathrm{E}\left(f(W_t) \mid W_0 = \xi\right) \\ &= \frac{1}{(2\pi t)^{d/2}} \int\limits_{\mathbb{R}^d} f(\omega) \exp\left(-\frac{|\xi - \omega|^2}{2t}\right) d\omega. \end{aligned}$$

As a corollary, if f is the Dirac impulse at 0, then u simplifies to

$$u(t,\xi) = \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|\xi|^2}{2t}\right);$$

that is, the *heat kernel* coincides with the *Brownian transition density* starting at 0.

Also, let D denote an open, simply connected domain in $\mathbb{R}^d$ with piecewise smooth, closed, orientable boundary C. The solution of the Laplace PDE (Dirichlet boundary value problem):

$$\begin{cases} \triangle v = 0, & \xi \in D, \\ v(\xi) = g(\xi), & \xi \in C,\ g : C \to \mathbb{R} \text{ piecewise continuous} \end{cases}$$

can be written as

$$v(\xi) = \mathrm{E}\left(g(W_\tau) \mid W_0 = \xi\right),$$

where τ is the **lifetime** or **first exit time** of Brownian motion in D:

$$\tau = \inf\{t > 0 : W_t \notin D\}.$$

Consequently, if $C = C_0 \cup C_1$, $C_0 \cap C_1 = \emptyset$ and $g(\xi) = k$ for $\xi \in C_k$, then $v(\xi)$ is the probability that a Brownian particle which starts at $\xi \in D$ stops at some point $\eta \in C_1$.

These two examples are special cases of a more general principle that solutions of any parabolic or elliptic PDE can be represented as expectations of certain stochastic functionals. (A hyperbolic PDE such as the wave equation $\partial^2 u/dt^2 = (1/2)\triangle u$ apparently cannot be solved in this manner.)

So far we have seen how probability is a servant of analysis. An example of how analysis serves probability is that the expected lifetime $v(\xi) = \mathrm{E}\,(\tau \mid W_0 = \xi)$ satisfies the Poisson PDE

$$\begin{cases} \triangle v = -2, & \xi \in D, \\ v(\xi) = 0, & \xi \in C. \end{cases}$$

For instance, if D is the ball of radius r in $\mathbb{R}^d$ centered at 0, then $v_D(\xi) = (r^2 - |\xi|^2)/d$. In the remainder of this essay, let $d = 2$. If T is the equilateral triangular region in $\mathbb{R}^2$ with vertices $(0, 2a/3)$, $(\pm a/\sqrt{3}, -a/3)$, then

$$v_T(x,y) = \frac{1}{2a}\left(y - \sqrt{3}x - \frac{2}{3}a\right)\left(y + \sqrt{3}x - \frac{2}{3}a\right)\left(y + \frac{1}{3}a\right).$$

If S is the square region in $\mathbb{R}^2$ with vertices $(\pm b, \pm b)$, then [5]

$$\begin{aligned} v_S(x,y) &= \frac{32b^2}{\pi^3}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)^3}\left[1 - \operatorname{sech}\left(\frac{(2k+1)\pi}{2}\right)\cosh\left(\frac{(2k+1)\pi y}{2b}\right)\right] \\ &\quad \times \cos\left(\frac{(2k+1)\pi x}{2b}\right). \end{aligned}$$

The lifetime functions $v_D(x,y)$, $v_T(x,y)$ and $v_S(x,y)$ are each maximized when $x = y = 0$. Define, for $b = 1/2$,

$$\gamma = v_S(0,0) = \frac{8}{\pi^3}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)^3}\left[1 - \operatorname{sech}\left(\frac{(2k+1)\pi}{2}\right)\right] = 0.1473427065....$$

This constant will be useful in the following; we wonder whether it has a closed-form expression.

When $r = 1/\sqrt{\pi}$, $a = \sqrt[4]{3}$ and $b = 1/2$, each of D, T and S have area 1 and

$$v_D(0,0) = \tfrac{1}{2\pi} = 0.159... > v_S(0,0) = \gamma = 0.147... > v_T(0,0) = \tfrac{2\sqrt{3}}{27} = 0.128....$$

In fact, among all planar regions of fixed area, the disk possesses the longest lifetime [6]. No such region with shortest lifetime exists, for consider the $c \times (1/c)$ finite strip as $c \to \infty$.

When $r = 1$, $a = 3$ and $b = 1$, each of D, T and S have inradius 1 (meaning the radius of the largest inscribed disk is unity) and

$$v_D(0,0) = \tfrac{1}{2} = 0.5 < v_S(0,0) = 4\gamma = 0.589... < v_T(0,0) = \tfrac{2}{3} = 0.666....$$

Clearly, among all planar regions of fixed inradius, the disk possesses the shortest lifetime. By way of contrast with the preceding, finding such a region with longest lifetime is an unsolved problem. Let

$$K = \sup_{D}\ \sup_{(x,y)\in D} \mathrm{E}\,(\tau \mid W_0 = (x,y)),$$

where the outer supremum is over all simply connected domains D in $\mathbb{R}^2$ of unit inradius; thus $K \geq 2/3$. The $2 \times \infty$ infinite strip improves this inequality to $K \geq 1$ and is the best such convex domain [7, 8]. Bañuelos & Carroll [9, 10] demonstrated that $1.584 < K < 3.228$; they speculated that the associated nonconvex domain D is extremal for certain other optimization problems as well.

4.7.1 *Fundamental Drum Frequency*

The bass tone of a kettledrum, whose head shape is a simply connected domain D in $\mathbb{R}^2$, is the square root of the smallest eigenvalue λ of [11, 12]

$$\begin{cases} \triangle u = -\lambda\, u, & \xi \in D, \\ u(\xi) = 0, & \xi \in C. \end{cases}$$

For instance, if D is the disk of radius r centered at $(0,0)$, then the first eigenfunction/eigenvalue pair is

$$u_D(x,y) = J_0\left(\frac{j_0\sqrt{x^2+y^2}}{r}\right), \quad \lambda_D = \left(\frac{j_0}{r}\right)^2$$

where $J_0(z)$ is the zeroth Bessel function of the first kind and $j_0 = 2.4048255576...$ is its smallest positive zero. If T is the equilateral triangular region of height a centered at $(0, a/6)$, then [13, 14]

$$\begin{aligned} u_T(x,y) &= \sin\left(\frac{\pi}{a}\left(y - \sqrt{3}x - \frac{2}{3}a\right)\right) + \sin\left(\frac{\pi}{a}\left(y + \sqrt{3}x - \frac{2}{3}a\right)\right) \\ &\quad - \sin\left(\frac{2\pi}{a}\left(y + \frac{1}{3}a\right)\right), \end{aligned}$$

$$\lambda_T = \frac{4\pi^2}{a^2}.$$

If S is the square region of side $2b$ centered at $(0,0)$, then

$$u_S(x,y) = \cos\left(\frac{\pi x}{2b}\right)\cos\left(\frac{\pi y}{2b}\right), \quad \lambda_S = \frac{\pi^2}{2b^2}.$$

When D, T and S each have area 1,

$$\lambda_D = \pi j_0^2 = 18.168... < \lambda_S = 2\pi^2 = 19.739... < \lambda_T = \tfrac{4\pi^2}{\sqrt{3}} = 22.792....$$

The Faber–Krahn inequality states that, among all planar regions of fixed area, the disk possesses the lowest bass tone. No such region with highest bass tone exists, for consider the $c \times (1/c)$ finite strip as $c \to \infty$.

When D, T and S each have inradius 1,

$$\lambda_D = j_0^2 = 5.783... > \lambda_S = \tfrac{\pi^2}{2} = 4.934... > \lambda_T = \tfrac{4\pi^2}{9} = 4.386....$$

Clearly, among all planar regions of fixed inradius, the disk possesses the highest bass tone. Finding such a region with lowest bass tone is an unsolved problem. Let

$$\Lambda = \inf_D \lambda_D$$

where the infimum is over all simply connected domains D in $\mathbb{R}^2$ of unit inradius; thus $\Lambda \le 4\pi^2/9$. The $2 \times \infty$ infinite strip improves this inequality to $\Lambda \le \pi^2/4 = 2.467...$ and is the best such convex domain [15–17]. In the other direction, Makai [18–22] proved that $\Lambda \ge 1/4$. The best bounds currently known [9] are $0.6197 < \Lambda < 2.1292$ and the associated nonconvex domain D is conjectured to be the same as before.

What does this have to do with Brownian motion? We give just one (of several) formulas [10, 23]:

$$\Lambda_D = 2\ \sup\left\{ c \ge 0 : \sup_{(x,y)\in D} \mathrm{E}\left(e^{c\tau} \mid W_0 = (x,y)\right) < \infty \right\}$$

for bounded, simply connected D. In words, the fact that $\lambda_D \ge \Lambda/\rho^2 > 0$ for D of inradius ρ means that if a drum produces an arbitrarily low bass tone, then it must contain an arbitrarily large circular subdrum.

4.7.2 *Torsional Rigidity*

Let us return to the expected lifetime function $v(x,y)$ and evaluate not its maximum value in the domain D, but rather twice its average value

$$\mu = \frac{2}{\mathrm{area}(D)} \int_D \mathrm{E}\left(\tau \mid W_0 = (x,y)\right) dx\, dy.$$

For instance, if D is the disk of radius r centered at $(0,0)$, then $\mu_D = r^2/2$. If T is the equilateral triangular region of height a centered at $(0, a/6)$, then $\mu_T = a^2/15$. If S is the square region of side $2b$ centered at $(0,0)$, then [5]

$$\begin{aligned}\mu_S &= \frac{4b^2}{3}\left[1 - \frac{192}{\pi^5}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^5}\tanh\left(\frac{(2k+1)\pi}{2}\right)\right] \\ &= \frac{1}{4}b^2(2.2492322392...) = b^2(0.5623080598...) = 4b^2(0.1405770149...).\end{aligned}$$

Again, we wonder about the possibility of closed-form evaluation.

When $r = 1/\sqrt{\pi}$, $a = \sqrt[4]{3}$ and $b = 1/2$,

$$\mu_D = \tfrac{1}{2\pi} = 0.159... > \mu_S = 0.140... > \mu_T = \tfrac{\sqrt{3}}{15} = 0.115....$$

This can be expressed in the language of elasticity theory. Pólya [24–27] proved Saint Venant's conjecture that, among all cylindrical beams of prescribed cross-sectional area, the circular beam has the highest *torsional rigidity*. No such beam with lowest torsional rigidity exists, for consider the $c \times (1/c)$ rectangle as $c \to \infty$.

When $r = 1$, $a = 3$ and $b = 1$,

$$\mu_D = \tfrac{1}{2} = 0.5 < \mu_S = 0.562... < \mu_T = \tfrac{3}{5} = 0.6.$$

Among all cylindrical beams of prescribed cross-sectional inradius, the circular beam has the lowest normalized torsional rigidity (normalized by area, as defined earlier). Finding such a beam with highest normalized torsional rigidity is an unsolved problem. Let

$$M = \sup_D \mu_D$$

where the supremum is over all simply connected domains D in $\mathbb{R}^2$ of unit inradius; thus $M \geq 3/5$. The $2 \times c$ rectangle improves this inequality, as $c \to \infty$, to $M \geq 4/3$ and is the best such convex domain [28]. For nonconvex domains, we have the upper bound 6.456 [9], but little else is known about this problem.

4.7.3 *Conformal Mapping*

If E is an open, simply connected region in $\mathbb{C}$, define $\rho(E)$ to be the inradius of E. The **univalent Bloch–Landau constant** Θ is given by [29]

$$\Theta = \inf_f \rho(f(D))$$

where the infimum is over all one-to-one analytic functions f defined on the open unit disk D satisfying $f(0) = 1, f'(0) = 1$. Let g denote the conformal mapping of D onto the infinite strip $-\pi/4 < \mathrm{Im}(z) < \pi/4$:

$$g(z) = \frac{1}{2}\ln\left(\frac{1+z}{1-z}\right) = \sum_{k=0}^{\infty} \frac{z^{2k+1}}{2k+1},$$

hence $\Theta \geq \pi/4$. Szegö [30, 31] further proved that, if $f(D)$ is convex, then $\rho(f(D)) \leq \rho(g(D))$. For the nonconvex scenario, the best bounds currently known [9, 32, 33] are $0.57088 < \Theta < 0.65642$ and the associated nonconvex region $f(D)$ is conjectured to be the same as the nonconvex domain for the constants K and Λ.

Addendum The constant γ indeed has a closed-form expression [34, 35]:

$$\gamma = 4\frac{{}_4F_3\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, \frac{1}{2}; \frac{5}{4}, \frac{5}{4}, 1; 1\right)}{B\left(\frac{1}{4}, \frac{1}{2}\right)^2} = 0.1473427065... = \frac{1}{2}(0.2946854131...),$$

where ${}_pF_q$ is the generalized hypergeometric function [36] and B is the Euler beta function ($B(x,y)=I(1,x,y)$ in [37]). An interesting double series representation:

$$\gamma=\frac{32}{\pi^4}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n}}{(2m-1)(2n-1)\left[(2m-1)^2+(2n-1)^2\right]}$$

follows from a formula in [38] which, in turn, was corrected in [39]. See also [40].

Both λ and μ can be defined via the calculus of variations [26]. It is more customary to take area$(D)\mu$ as torsional rigidity and this is equal to [41, 42]

$$\frac{1}{12}-\frac{16}{\pi^5}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^5}\coth\left(\frac{(2k+1)\pi}{2}\right)=0.0260896517...$$

for an isosceles right triangle with sides 1, 1, $\sqrt{2}$ and [43, 44]

$$\begin{aligned}9&\left[\frac{17\sqrt{3}}{192}-\frac{1}{\pi^5}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^5}\left\{2\tanh\left(\frac{(2k+1)\pi\sqrt{3}}{2}\right)-9\tanh\left(\frac{(2k+1)\pi}{2\sqrt{3}}\right)\right.\right.\\&\left.\left.\quad+(-1)^k9\sqrt{3}\,\mathrm{sech}\left(\frac{(2k+1)\pi}{2\sqrt{3}}\right)+27\sqrt{3}\sin\left(\frac{(2k+1)\pi}{3}\right)\right\}\right]\\&=0.0044516625...=\frac{9}{16}(0.0079140667...)\end{aligned}$$

for a 30°-60°-90° triangle with sides $1/2$, $\sqrt{3}/2$ and 1. The corresponding value for a regular hexagon of unit side has attracted considerable attention [45–48] – see history in [42] – a complicated formula in [49] gives ≈ 1.035459, as reported in [50], and verifies an unpublished calculation in [51].

[1] S. R. Finch, Ornstein-Uhlenbeck process, *this volume*, §4.3.
[2] S. R. Finch, Variants of Brownian motion, *this volume*, §4.5.
[3] R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth, 1984, pp. 36–39, 219–270; MR0750829 (87a:60054).
[4] R. Durrett, *Stochastic Calculus: A Practical Introduction*, CRC Press, 1996, pp. 125–166; MR1398879 (97k:60148).
[5] S. Timoshenko and J. N. Goodier, *Theory of Elasticity*, 3rd ed, McGraw-Hill, 1970, pp. 297–301, 309–315; MR0045547 (13,599i).
[6] R. Bañuelos, Queen Dido's property for eigenvalues and Brownian motion, unpublished lecture (2001).
[7] L. E. Payne, Bounds for the maximum stress in the Saint Venant torsion problem, *Indian J. Mech. Math.,* special issue (1968/69), part I, 51–59; MR0351225 (50 #3714).
[8] R. P. Sperb, *Maximum Principles and their Applications*, Academic Press, 1981, pp. 85–89; MR0615561 (84a:35033).
[9] R. Bañuelos and T. Carroll, Brownian motion and the fundamental frequency of a drum, *Duke Math. J.* 75 (1994) 575–602; addendum 82 (1996) 227; MR1291697 (96m:31003) and MR1387227 (97f:31004).
[10] T. Carroll, Old and new on the bass note, the torsion function and the hyperbolic metric, *Irish Math. Soc. Bull.* 47 (2001) 41–65; MR1880191 (2002k:58072).

[11] S. R. Finch, Sobolev isoperimetric constants, *first volume*, pp. 219–225.

[12] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.

[13] M. A. Pinsky, The eigenvalues of an equilateral triangle, *SIAM J. Math. Anal.* 11 (1980) 819-827; 16 (1985) 848-851; MR0586910 (82d:35077) and MR0793926 (86k:35115).

[14] L. M. Cureton and J. R. Kuttler, Eigenvalues of the Laplacian on regular polygons and polygons resulting from their dissection, *J. Sound Vibration* 220 (1999) 83-98; MR1672258 (99j:35025).

[15] J. Hersch, Sur la fréquence fondamentale d'une membrane vibrante: évaluations par défaut et principe de maximum, *Z. Angew. Math. Phys.* 11 (1960) 387–413; MR0125319 (23 #A2622).

[16] R. Bañuelos, R. Latala and P. J. Méndez-Hernández, A Brascamp-Lieb-Luttinger-type inequality and applications to symmetric stable processes, *Proc. Amer. Math. Soc.* 129 (2001) 2997–3008; MR1840105 (2002c:60125).

[17] P. J. Méndez-Hernández, Brascamp-Lieb-Luttinger inequalities for convex domains of finite inradius, *Duke Math. J.* 113 (2002) 93–131; MR1905393 (2003b:31005).

[18] E. Makai, A lower estimation of the principal frequencies of simply connected membranes, *Acta Math. Acad. Sci. Hungar.* 16 (1965) 319–323; MR0185263 (32 #2732).

[19] W. K. Hayman, Some bounds for principal frequency, *Applicable Anal.* 7 (1977/78) 247–254; MR0492339 (58 #11468).

[20] R. Osserman, A note on Hayman's theorem on the bass note of a drum, *Comment. Math. Helv.* 52 (1977) 545–555; MR0459099 (56 #17297).

[21] M. E. Taylor, Estimate on the fundamental frequency of a drum, *Duke Math. J.* 46 (1979) 447–453; MR0534061 (81g:35097).

[22] C. B. Croke, The first eigenvalue of the Laplacian for plane domains, *Proc. Amer. Math. Soc.* 81 (1981) 304–305; MR0593476 (82e:35061).

[23] S. E. Graversen and M. Rao, Brownian motion and eigenvalues for the Dirichlet Laplacian, *Math. Z.* 203 (1990) 699–708; MR1044073 (91f:35197).

[24] G. Pólya, Torsional rigidity, principal frequency, electrostatic capacity and symmetrization, *Quart. Appl. Math.* 6 (1948) 267–277; MR0026817 (10,206b).

[25] G. Pólya and A. Weinstein, On the torsional rigidity of multiply connected cross-sections, *Annals of Math.* 52 (1950) 154–163; MR0040159 (12,652e).

[26] G. Pólya and G. Szegö, *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, 1951, pp. 87–100, 121, 251–261; MR0043486 (13,270d).

[27] C. Bandle, *Isoperimetric Inequalities and Applications*, Pitman, 1980, pp. 63–67, 158–159; MR0572958 (81e:35095).

[28] E. Makai, On the principal frequency of a membrane and the torsional rigidity of a beam, *Studies in Mathematical Analysis and Related Topics: Essays in Honor of George Pólya,* ed. G. Szegö, C. Loewner, S. Bergman, M. M. Schiffer, J. Neyman, D. Gilbarg, and H. Solomon, Stanford Univ. Press, 1962, pp. 227–231; MR0167004 (29 #4277).

[29] S. R. Finch, Bloch-Landau constants, *first volume*, pp. 456–459.

[30] G. Szegö, Über eine Extremalaufgabe aus der Theorie der schlichten Abbildungen, *Sitzungsberichte der Berliner Mathematische Gesellschaft* 22 (1923) 38–47; corrigendum 23 (1924) 64; *Collected Papers*, v. 1, ed. R. Askey, Birkhäuser Verlag, 1982, pp. 607-618.

[31] D. Minda, Lower bounds for the hyperbolic metric in convex regions, *Rocky Mountain J. Math.* 13 (1983) 61–69; MR0692577 (84j:30039).

[32] E. Beller and J. A. Hummel, On the univalent Bloch constant, *Complex Variables Theory Appl.* 4 (1985) 243-252; MR0801641 (86i:30005).

[33] S. Y. Zhang, On the schlicht Bloch constant, *Beijing Daxue Xuebao* 25 (1989) 537–540 (in Chinese); MR1045870 (91a:30014).
[34] G. Markowsky, On the expected exit time of planar Brownian motion from simply connected domains, *Elec. Commun. Probab.* 16 (2011) 652–663; arXiv:1108.1188; MR2853103 (2012j:60228).
[35] G. Markowsky, A method for deriving hypergeometric and related identities from the H^2 Hardy norm of conformal maps, *Integral Transforms Spec. Funct.* 24 (2013) 302–313; arXiv:1205.2458; MR3037253.
[36] S. R. Finch, Notation: Hypergeometric function, *this volume*, p. xi.
[37] S. R. Finch, Gambler's ruin, *this volume*, §4.20.
[38] F. B. Knight, *Essentials of Brownian Motion and Diffusion*, Amer. Math. Soc. 1981, pp. 70–71; MR0613983 (82m:60098).
[39] K. Helmes, S. Röhl and R. H. Stockbridge, Computing moments of the exit time distribution for Markov processes by linear programming, *Oper. Res.* 49 (2001) 516–530; MR1852358.
[40] M. W. Coffey, Expected exit times of Brownian motion from planar domains: Complements to a paper of Markowsky, arXiv:1203.5142.
[41] B. G. Galerkin, Torsion d'un prisme triangulaire, *Bulletin de l'Académie des Sciences de Russie* 13 (1919) 111–118.
[42] T. J. Higgins, A comprehensive review of Saint-Venant's torsion problem, *Amer. J. Phys.* 10 (1942) 248–259.
[43] G. E. Hay, The method of images applied to the problem of torsion, *Proc. London Math. Soc.* 45 (1939) 382–397; MR0000116 (1,21c).
[44] L.-C. Pan, Equilibrium, buckling and vibration of a 30°-60°-90°-triangular plate simply supported at the edges, *Sci. Sinica* 6 (1957) 347–379; MR0091635 (19,998c).
[45] T. Tsumura, On the torsion of prismatic cylinder with a regular polygonal section, *Journal of the Society of Mechanical Engineers (Tokyo)* 36 (1933) 60–63.
[46] J. L. Synge, *The Hypercircle in Mathematical Physics: A Method for the Approximate Solution of Boundary Value Problems*, Cambridge Univ. Press, 1957, pp. 214–270; MR0097605 (20 #4073).
[47] E. Romanelli, M. J. Maurizi and L. A. Patricio, Torsion of bars of regular polygonal cross section, *Journal of the Engineering Mechanics Division (ASCE)* 98 (1972) 415–424.
[48] W. C. Hassenpflug, Torsion of uniform bars with polygon cross-section, *Comput. Math. Appl.* 46 (2003) 313–392; MR2015439 (2004i:74069).
[49] B. R. Seth. Torsion of beams whose cross-section is a regular polygon of n sides, *Proc. Camb. Phil. Soc.* 30 (1934) 139–149.
[50] E. E. Jones, The torsional rigidity of anisotropic prisms, *J. Engineering Math.* 9 (1975) 39–51; MR0411328 (53 #15064).
[51] M. I. Aissen, *On the Estimation and Computation of Torsional Rigidity*, Ph.D. thesis, Stanford Univ., 1952.

4.8 Subcritical Galton–Watson Trees

Fix a probability $0 < p < 1$. For any 2-vector $u = (v, w)$, write $u_L = v$ and $u_R = w$. A **Galton–Watson tree** is an ordered, strongly binary tree $T = \tau(p)_L$ that is defined recursively in terms of left and right subtrees of the root as follows:

$$\tau(p) = \left(\begin{cases} \tau(p) & \text{if } X \le p, \\ \emptyset & \text{otherwise} \end{cases}, \begin{cases} \tau(p) & \text{if } X \le p, \\ \emptyset & \text{otherwise} \end{cases} \right).$$

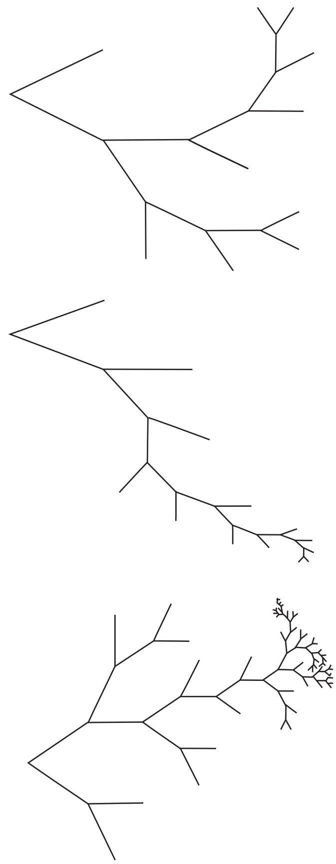

Figure 4.2 Three sample binary trees, grown via the Galton–Watson process.

Each variable X is a new, independent Uniform $[0, 1]$ random number. For example, $T = \emptyset$ with probability $1 - p$, $T = (\emptyset, \emptyset)$ with probability $p(1 - p)^2$, and $T = ((\emptyset, \emptyset), \emptyset)$ with probability $p^2(1 - p)^3$ (Figure 4.2).

The **number of vertices** N is equal to twice the number of left parentheses (parents) in the expression for T, plus one. Equivalently, N is twice the number of $\emptyset s$ (leaves), minus one. It can be shown that N is finite with probability 1 if $p \leq 1/2$ and $1/p - 1$ if $p > 1/2$. We will focus on the **subcritical case** $p < 1/2$ for the remainder of this essay.

Let N_k denote the number of vertices at distance k from the root, that is, the size of the k^{th} generation. Clearly $N_0 = 1$ and $N < \infty$ if and only if $N_k = 0$ for all sufficiently large k. Define

$$\text{the **height** } H \text{ of } T \text{ to be } \max_{N_k > 0} k,$$

$$\text{the **width** } W \text{ of } T \text{ to be } \max_{k \geq 0} N_k.$$

We wish to evaluate the joint distribution of (N, H, W) as a function of p. Some partial results (mostly of a numerical nature) are all we can report now.

The sequence N_0, N_1, ..., N_H is called the **profile** of T. Dual to this is the sequence of $(N+1)/2$ leaf distances from the root, ordered from left to right, called the **contour** of T. It would be good someday to better understand joint profile and contour distributions as well.

4.8.1 *Number of Vertices*

The probability that T has at least 3 vertices is p. Let $m \geq 1$. From the conditional relation:

$$\begin{aligned}&\mathrm{P}(T \text{ has } 2m+1 \text{ vertices})\\ &= \sum_{j=1}^{2m-1} \mathrm{P}(T_L \text{ has } 2m-j \text{ vertices } \wedge\ T_R \text{ has } j \text{ vertices} \mid T \text{ has at least 3 vertices}) \cdot p,\end{aligned}$$

we deduce that

$$\mathrm{P}(N=2m+1) = p \sum_{j=1}^{2m-1} \mathrm{P}(N=2m-j)\mathrm{P}(N=j)$$

and hence

$$\mathrm{P}(N=n) = \begin{cases} 0 & \text{if } n=2m, \\ \dfrac{1}{m+1}\dbinom{2m}{m} p^m (1-p)^{m+1} & \text{if } n=2m+1. \end{cases}$$

Well-known asymptotics for the Catalan numbers

$$\frac{1}{m+1}\binom{2m}{m} \sim \frac{1}{\sqrt{\pi}} \frac{2^{2m}}{m^{3/2}}$$

give a sense of the rate at which $\mathrm{P}(N=n) \to 0$ as $n \to \infty$, n odd. More precisely [1, 2],

$$\mathrm{P}(N=n) \sim \left(\sqrt{\frac{2}{\pi}} n^{-3/2} + c\, n^{-5/2} + d\, n^{-7/2} + \cdots\right) (2p)^m (2(1-p))^{m+1}.$$

We also have moments [3–5]

$$\mathrm{E}(N) = \frac{1}{1-2p}, \quad \mathrm{Var}(N) = \frac{4p(1-p)}{(1-2p)^3}.$$

4.8.2 *Height*

Let a_k denote the probability that $N_k = 0$, equivalently, the probability that $H < k$. The conditional distribution of N_k, given $N_1 = j$, is the same as the sum of

j independent random copies of N_{k-1} [6]. Of course, $j=0$ and $j=2$ are the only possible values for N_1; thus we have

$$\mathrm{P}(N_k=0)=\underbrace{\mathrm{P}(N_k=0\,|\,N_1=0)}_{1}\overbrace{\mathrm{P}(N_1=0)}^{1-p}+\underbrace{\mathrm{P}(N_k=0\,|\,N_1=2)}_{\mathrm{P}(N_{k-1}=0)^2}\overbrace{\mathrm{P}(N_1=2)}^{p}$$

and hence [7]

$$a_0=0, \qquad a_k=(1-p)+p\,a_{k-1}^2 \quad \text{for } k\geq 1, \qquad \lim_{k\to\infty} a_k=1.$$

Let us prove that the convergence rate of $\{a_k\}$ is exponential, that is,

$$0<\lim_{k\to\infty}\frac{1-a_k}{(2p)^k}<1.$$

First, note that $0\leq a_k<1$ for all k by induction ($a_k\geq 1-p>0$ is obvious; supposing $0\leq a_{k-1}<1$, we obtain $a_k<(1-p)+p=1$). Now, writing $b_k=1-a_k$, we have $b_0=1$, $0<b_k\leq 1$ and

$$\begin{aligned} b_k &= p\left(1-a_{k-1}^2\right)=p(1-a_{k-1})(1+a_{k-1}) \\ &= p\,b_{k-1}(2-b_{k-1}) \\ &< 2p\,b_{k-1}<(2p)^2 b_{k-2}<(2p)^3 b_{k-3} \end{aligned}$$

thus $b_k<(2p)^k$ for all k. Observe that

$$\begin{aligned} b_k &= 2p\,b_{k-1}\left(1-\frac{b_{k-1}}{2}\right) \\ &= (2p)^2 b_{k-2}\left(1-\frac{b_{k-2}}{2}\right)\left(1-\frac{b_{k-1}}{2}\right) \\ &= (2p)^3 b_{k-3}\left(1-\frac{b_{k-3}}{2}\right)\left(1-\frac{b_{k-2}}{2}\right)\left(1-\frac{b_{k-1}}{2}\right) \\ &= (2p)^k\prod_{j=0}^{k-1}\left(1-\frac{b_j}{2}\right) \end{aligned}$$

hence

$$\begin{aligned} C &= \lim_{k\to\infty}\frac{1-a_k}{(2p)^k}=\lim_{k\to\infty}\frac{b_k}{(2p)^k}=\prod_{j=0}^{\infty}\left(1-\frac{b_j}{2}\right) \\ &= \prod_{j=0}^{\infty}\left(1-\frac{1-a_j}{2}\right)=\prod_{j=0}^{\infty}\frac{1+a_j}{2} \end{aligned}$$

exists and is nonzero since

$$\sum_{j=0}^{\infty}\frac{b_j}{2}<\frac{1}{2}\sum_{j=0}^{\infty}(2p)^j$$

Table 4.4 *Height-related parameters*

p	C	$\mathrm{E}(H)$	$\mathrm{Var}(H)$
0.2	0.4238945378...	0.3179675669...	0.6053027749...
0.25	0.3929068527...	0.4610125877...	1.0724312517...
0.3	0.3539671772...	0.6568327963...	1.9336638291...
0.35	0.3039572818...	0.9422336526...	3.7158517879...
0.4	0.2376466589...	1.4045313857...	8.2383270278...

converges. This completes the proof. The expression for C as an infinite product turns out to be useful for high precision estimates of C, given p (see Table 4.4).

The algorithm for $\{b_k\}$:

$$b_0 = 1, \qquad b_k = p\, b_{k-1}(2 - b_{k-1}) \quad \text{for } k \geq 1, \qquad \lim_{k\to\infty} b_k = 0$$

is helpful from a numerical perspective. While formulas in a_ℓ are easily converted into formulas in b_ℓ and vice versa:

$$\mathrm{P}(H = k) = a_{k+1} - a_k = b_k - b_{k+1},$$

$$\mathrm{E}(e^{tH}) = \sum_{k=0}^{\infty} e^{tk}(a_{k+1} - a_k) = \sum_{k=0}^{\infty} e^{tk}(b_k - b_{k+1})$$

the difference $a_{k+1} - a_k$ is harder to calculate than $b_k - b_{k+1}$. (Reason: the subtraction of nearly equal quantities, each approaching 1, leads to a loss of floating point precision.)

Since the series for the moment generating function is telescoping:

$$\mathrm{E}(e^{tH}) = b_0 + \sum_{k=1}^{\infty}(e^{tk} - e^{t(k-1)})b_k$$

we obtain

$$\mathrm{E}(H) = \sum_{k=1}^{\infty} b_k, \quad \mathrm{E}(H^2) = \sum_{k=1}^{\infty}(2k-1)b_k$$

upon differentiation. No closed-form expressions for the quantities in Table 4.4 are known.

We mention an interesting result for the **critical case** $p = 1/2$. The recurrence

$$a_0 = 0, \qquad a_k = \tfrac{1}{2}\left(1 + a_{k-1}^2\right) \quad \text{for } k \geq 1, \qquad \lim_{k\to\infty} a_k = 1$$

satisfies [8–10]

$$1 - a_k \sim \frac{2}{k + \ln(k) + 1.76799378...}.$$

It is clear, therefore, that $\mathrm{E}(H) = \infty$. The relevance of [8, 9] to Galton–Watson trees seems not to have been noticed before.

4.8.3 *Height via Markov*

The sequence $N_0, N_1, N_2, \ldots$ is a time-homogeneous Markov chain with transition probability matrix Q, where

$$q_{i,j} = \mathrm{P}(N_1 = j \mid N_0 = i) = \begin{cases} \binom{i}{j/2} p^{j/2}(1-p)^{i-j/2} & \text{if } 2 \le j \le 2i \text{ is even,} \\ 0 & \text{otherwise} \end{cases}$$

is the $(i,j)^{\text{th}}$ element of Q and $i \ge 1, j \ge 1$. Observe that

$$\mathrm{P}(H = 0 \mid N_0 = i) = (1-p)^i$$

and

$$\mathrm{P}(H = k \mid N_0 = i) = \sum_{j \ge 1} \underbrace{\mathrm{P}(H = k \mid N_1 = j)}_{\mathrm{P}(H = k-1 \mid N_0 = j)} \overbrace{\mathrm{P}(N_1 = j \mid N_0 = i)}^{q_{i,j}}$$

for $k \ge 1$. We will use these formulas to derive an alternative matrix expression for $\mathrm{E}(H)$, as outlined in [11]. Let $\mu_i = \mathrm{E}(H \mid N_0 = i)$ and $\nu_i = 1 - (1-p)^i$. From

$$\begin{aligned} \mu_i &= \sum_{k \ge 0} k\mathrm{P}(H = k \mid N_0 = i) \\ &= \sum_{k \ge 1} k \sum_{j \ge 1} q_{i,j} \mathrm{P}(H = k-1 \mid N_0 = j) \\ &= \sum_{j \ge 1} q_{i,j} \sum_{k \ge 1} k\mathrm{P}(H = k-1 \mid N_0 = j) \\ &= \sum_{j \ge 1} q_{i,j} \left(1 + \sum_{k \ge 0} k\mathrm{P}(H = k \mid N_0 = j) \right) \\ &= \nu_i + \sum_{j \ge 1} q_{i,j} \mu_j, \end{aligned}$$

it follows that $(I - Q)\mu = \nu$ and thus $\mu = (I - Q)^{-1}\nu$. Only the first component of μ is desired since $\mathrm{E}(H) = \mathrm{E}(H \mid N_0 = 1)$. Of course, we must restrict $i \le \ell, j \le \ell$ when evaluating μ_1, where ℓ is large. As $\ell \to \infty$, indeed $\mu_1 \to \mathrm{E}(H)$ numerically as found in the previous section.

4.8.4 *Width*

Clearly $\mathrm{P}(W=0)=0$ since $N_0=1$ and $\mathrm{P}(W=1)=\mathrm{P}(N_1=0)=1-p$. An elementary expression $f(p)$ for $\mathrm{P}(W=2)$ arises from

$$\begin{aligned} f(p) &= p(1-p)^2+2p^2(1-p)^3+4p^3(1-p)^4+8p^4(1-p)^5+\cdots \\ &= p(1-p)^2\left(1+2p(1-p)\left(1+2p(1-p)\left(1+2p(1-p)\cdots\right)\right)\right) \\ &= p(1-p)^2\left(1+2f(p)/(1-p)\right); \end{aligned}$$

hence

$$(1-2p(1-p))f(p)=p(1-p)^2;$$

hence

$$\mathrm{P}(W=2)=\frac{p(1-p)^2}{1-2p(1-p)}=\frac{p(1-p)^2}{2p^2-2p+1}.$$

An analogous argument leading to $\mathrm{P}(W=4)$ does not seem to work. We turn therefore to the alternative approach.

4.8.5 *Width via Markov*

Define the matrix Q exactly as before with $i\geq 1, j\geq 1$. Observe that [12, 13]

$$\mathrm{P}(W\leq 0\,|\,N_0=i)=0$$

and

$$\mathrm{P}(W\leq m\,|\,N_0=i)=(1-p)^i+\sum_{j=1}^{m}\underbrace{\mathrm{P}(W\leq m\,|\,N_1=j)}_{\mathrm{P}(W\leq m\,|\,N_0=j)}\overbrace{\mathrm{P}(N_1=j\,|\,N_0=i)}^{q_{i,j}}$$

for $m\geq 1$. Let $\alpha_i(m)=\mathrm{P}(W>m\,|\,N_0=i)$, $\beta_0=1$, $\beta_m=\alpha_1(m)$ and

$$\gamma_i(m)=\begin{cases}\displaystyle\sum_{j=m+1}^{2i} q_{i,j} & \text{if } m+1\leq 2i,\\ 0 & \text{otherwise.}\end{cases}$$

It follows that

$$\begin{aligned} \alpha_i(m) &= 1-(1-p)^i-\sum_{j=1}^{m} q_{i,j}(1-\alpha_j(m)) \\ &= \sum_{j=1}^{2i} q_{i,j}-\sum_{j=1}^{m} q_{i,j}+\sum_{j=1}^{m} q_{i,j}\alpha_j(m) \\ &= \gamma_i(m)+\sum_{j=1}^{m} q_{i,j}\alpha_j(m) \end{aligned}$$

Table 4.5 *Width-related parameters*

p	$D(p)$	$\mathrm{E}(W)$	$\mathrm{Var}(W)$
0.2	0.8...	1.2243696655...	0.2507547512...
0.25	1.2...	1.3038399841...	0.3903119417...
0.3	1.7...	1.4072057242...	0.6311389283...
0.35	2.6...	1.5526227137...	1.1020414724...
0.4	4.3...	1.7823528114...	2.2389987484...

and thus $\alpha(m)=(I-Q)^{-1}\gamma(m)$. Only the first component of $\alpha(m)$ is desired since $\mathrm{P}(W>m)=\mathrm{P}(W>m\mid N_0=1)$. A theorem in [14, 15] leads to a conjecture that

$$D=\lim_{\substack{m\to\infty\\ m\text{ even}}} m\left(\frac{1}{p}-1\right)^{m}\beta_m=\lim_{\substack{m\to\infty\\ m\text{ odd}}} m\left(\frac{1}{p}-1\right)^{m-1}\beta_m$$

exists and is nonzero. We have

$$\mathrm{P}(W=m)=\beta_{m-1}-\beta_m,$$

$$\mathrm{E}(e^{tW})=\sum_{m=1}^{\infty}e^{tm}(\beta_{m-1}-\beta_m)=e^t\beta_0+\sum_{m=1}^{\infty}(e^{t(m+1)}-e^{tm})\beta_m$$

and hence

$$\mathrm{P}(W=4)=\frac{p^3(1-p)^4(2p^2-2p-1)}{(2p^2-2p+1)(8p^6-24p^5+30p^4-20p^3+4p^2+2p-1)},$$

$$\mathrm{E}(W)=\sum_{m=0}^{\infty}\beta_m,\quad \mathrm{E}(W^2)=\sum_{m=0}^{\infty}(2m+1)\beta_m.$$

No closed-form expressions for the quantities in Table 4.5 are known.

For the critical case, it can be proved [16, 17] that $\mathrm{E}(W)=\infty$ and, in fact, $\mathrm{E}(\max_{0\le k\le\ell}N_k)\sim\ln(\ell)$ as $\ell\to\infty$.

4.8.6 *Cross-Correlation*

Lacking any better methods to study association, we generated 10^6 Galton–Watson trees for each probability $p=0.2,\ \ldots,\ 0.4$. The cross-correlation coefficients between N, H and W are each large, but we observe that roughly

$$0.95\approx\rho(N,H)>\rho(N,W)>\rho(H,W)\approx0.85.$$

No clear pattern in these, as functions of p, are yet evident. Clearly this is an area for further research [18].

Acknowledgments I am grateful to Robert Israel and Anthony Quas for the simple proof that $C(p)$ exists, and to Frank Purcell for his algorithm that plots a binary tree (given as a nested list).

[1] R. Otter, The multiplicative process, *Annals of Math. Statistics* 20 (1949) 206–224; MR0030716 (11,41a).
[2] N. Minami, On the number of vertices with a given degree in a Galton-Watson tree, *Adv. Appl. Probab*. 37 (2005) 229–264; MR2135161 (2006a:60158).
[3] D. Hawkins and S. Ulam, Theory of multiplicative processes, Los Alamos classified report LA-171, 1944; in *Analogies between Analogies*, ed. A. R. Bednarek and F. Ulam, Univ. of California Press, 1990, pp. 1–15; MR1123485 (92i:01033).
[4] I. J. Good, The number of individuals in a cascade process, *Proc. Cambridge Philos. Soc.* 45 (1949) 360–363; MR0029489 (10,614a).
[5] A. G. Pakes, Some limit theorems for the total progeny of a branching process, *Adv. Appl. Probab.* 3 (1971) 176–192; MR0283892 (44 #1122).
[6] W. Feller, *An Introduction to Probability Theory and Its Applications*, v. 1, 3rd ed., Wiley, 1968, pp. 293–302; MR0228020 (37 #3604).
[7] S. R. Finch, Otter's tree enumeration constants, *first volume*, pp. 295–316.
[8] L. Moser, On a problem of Cayley, *Scripta Math.* 22 (1956) 289–292.
[9] I. Guttman, On a problem of L. Moser, *Canad. Math. Bull.* 3 (1960) 35–39; MR0129027 (23 #B2064).
[10] S. R. Finch, Optimal stopping constants, *first volume*, pp. 361–363.
[11] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications*, McGraw-Hill, 1960, pp. 26–27; MR0112177 (22 #3032).
[12] J. Bishir, Maximum population size in a branching process, *Biometrics* 18 (1962) 394–403; MR0145637 (26 #3166).
[13] S. R. Adke, The maximum population size in the first N generations of a branching process, *Biometrics* 20 (1964) 649–651; MR0181013 (31 #5243).
[14] T. Lindvall, On the maximum of a branching process, *Scand. J. Statist*. 3 (1976) 209–214; MR0426188 (54 #14134).
[15] O. Nerman, On the maximal generation size of a non-critical Galton-Watson process, *Scand. J. Statist.* 4 (1977) 131–135; MR0474531 (57 #14169).
[16] K. B. Athreya, On the maximum sequence in a critical branching process, *Annals of Probab.* 16 (1988) 502–507; MR0929060 (89f:60098).
[17] E. M. Bondarenko and V. A. Topchii, Estimates for the expectation of the maximum of a critical Galton-Watson process on a finite interval (in Russian), *Sibirsk. Mat. Zh.* 42 (2001) 249–257; Engl. transl. in *Siberian Math. J.* 42 (2001) 209–216; MR1833155 (2002b:60145).
[18] S. R. Finch, Shapes of binary trees, *this volume*, §4.6.

4.9 Continued Fraction Transformation

We are interested in iterates of the **continued fraction transformation** $T\colon [0,1] \to [0,1]$ defined by [1]

$$T(x) = \begin{cases} \left\{ \dfrac{1}{x} \right\} & \text{if } 0 < x \le 1, \\ 0 & \text{if } x = 0 \end{cases}$$

where $\{\xi\} = \xi - \lfloor \xi \rfloor$ denotes the fractional part of ξ. For example,

$$\begin{aligned}
\pi - 3 &= 0.141592..., & \left\lfloor \tfrac{1}{\pi-3} \right\rfloor &= 7, \\
T(\pi - 3) &= 0.062513..., & \left\lfloor \tfrac{1}{T(\pi-3)} \right\rfloor &= 15, \\
T^2(\pi - 3) &= 0.996594..., & \left\lfloor \tfrac{1}{T^2(\pi-3)} \right\rfloor &= 1, \\
T^3(\pi - 3) &= 0.003417..., & \left\lfloor \tfrac{1}{T^3(\pi-3)} \right\rfloor &= 292, \\
T^4(\pi - 3) &= 0.634591..., & \left\lfloor \tfrac{1}{T^4(\pi-3)} \right\rfloor &= 1
\end{aligned}$$

and

$$\pi = 3 + \frac{1|}{|7} + \frac{1|}{|15} + \frac{1|}{|1} + \frac{1|}{|292} + \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|2} + \frac{1|}{|1} + \frac{1|}{|3} + \cdots$$

is the regular continued fraction expansion for π. In words, T discards the first "digit" in any expansion, that is,

$$T\left(\frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots\right) = \frac{1|}{|a_2} + \frac{1|}{|a_3} + \frac{1|}{|a_4} + \cdots.$$

What can be said about the moments of $T^j X$ and of $\ln(T^j X)$, where X is a random variable in $[0, 1]$? There are two cases: the first when X follows the uniform distribution, and the second when X follows the **Gauss–Kuzmin distribution**:

$$\mathrm{P}(X \le x) = \frac{\ln(x+1)}{\ln(2)}.$$

We will later study the partial convergents to x, for example,

$$\frac{p_1}{q_1} = \frac{3}{1}, \ \frac{p_2}{q_2} = \frac{22}{7}, \ \frac{p_3}{q_3} = \frac{333}{106}, \ \frac{p_4}{q_4} = \frac{355}{113}, \ \frac{p_5}{q_5} = \frac{103993}{33102}, \ ...$$

when $x = \pi$. The asymptotic distribution of denominators Q_n, corresponding to uniformly distributed X as $n \to \infty$, turns out to be related to our earlier work on $\ln(T^j X)$ statistics.

4.9.1 *Uniform Distribution*

Let γ denote the Euler–Mascheroni constant [2], ζ denote the Riemann zeta function and Li_k denote the k^{th} polylogarithm function [3]. If X is a random variable following the uniform distribution on $[0, 1]$, then

$$\mathrm{E}(X) = \int_0^1 x\,dx = \frac{1}{2}, \quad \mathrm{E}(X^2) = \int_0^1 x^2 dx = \frac{1}{3},$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = \frac{1}{12}$$

and, via the substitution $y = 1/x$,

$$\begin{aligned}\mathrm{E}(TX) &= \int_0^1 \left\{\frac{1}{x}\right\} dx = \int_1^\infty \frac{\{y\}}{y^2} dy = \sum_{n=1}^\infty \int_n^{n+1} \frac{y-n}{y^2} dy \\ &= \sum_{n=1}^\infty \left(\ln\left(\frac{n+1}{n}\right) - \frac{1}{n+1} \right) = 1 - \gamma = 0.4227843351...\end{aligned}$$

(which is related to de la Vallée Poussin's theorem [2, 4]),

$$\mathrm{E}((TX)^2) = \ln(2\pi) - \gamma - 1,$$

$$\mathrm{Var}(TX) = \ln(2\pi) - \gamma^2 + \gamma - 2 = 0.0819148075... = (0.2862076300...)^2,$$

$$\mathrm{E}(X \cdot TX) = 1 - \frac{\pi^2}{12},$$

$$\mathrm{Cov}(X, TX) = \mathrm{E}(X \cdot TX) - \mathrm{E}(X)\mathrm{E}(TX) = \frac{1}{12}\left(6 - \pi^2 + 6\gamma\right),$$

$$\begin{aligned}\rho(X, TX) &= \frac{\mathrm{Cov}(X, TX)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(TX)}} = \frac{6 - \pi^2 + 6\gamma}{\sqrt{12}\sqrt{\ln(2\pi) - \gamma^2 + \gamma - 2}} \\ &= -0.4098133678...\end{aligned}$$

where ρ denotes cross-correlation. Likewise,

$$\mathrm{E}(\ln(X)) = -1, \quad \mathrm{E}(\ln(X)^2) = 2, \quad \mathrm{Var}(\ln(X)) = 1,$$

and, via the substitutions $y = 1/x$ and $z = y - n$,

$$\begin{aligned}\mathrm{E}(\ln(TX)) &= \int_0^1 \ln\left\{\frac{1}{x}\right\} dx = \int_1^\infty \frac{\ln\{y\}}{y^2} dy = \sum_{n=1}^\infty \int_n^{n+1} \frac{\ln(y-n)}{y^2} dy \\ &= \sum_{n=1}^\infty \int_0^1 \frac{\ln(z)}{(z+n)^2} dz = -\sum_{n=1}^\infty \frac{1}{n} \ln\left(\frac{n+1}{n}\right) \\ &= -\left(\ln(2) + \sum_{k=2}^\infty (-1)^k \frac{\zeta(k) - 1}{k-1} \right) = -1.2577468869...\end{aligned}$$

(this constant appears elsewhere [5, 6]),

$$\mathrm{E}(\ln(TX)^2) = -2\sum_{n=1}^\infty \frac{1}{n} \mathrm{Li}_2\left(-\frac{1}{n}\right) = \zeta(2) - 2\sum_{k=1}^\infty (-1)^k \frac{\zeta(k+1) - 1}{k^2},$$

$$\mathrm{Var}(\ln(TX)) = 1.2665694005... = (1.1254196552...)^2,$$

$$\begin{aligned}
\mathrm{E}(\ln(X)\cdot\ln(TX)) &= \sum_{n=1}^{\infty}\frac{1}{n}\left[\ln\left(\frac{n+1}{n}\right)(1+\ln(n)) - \mathrm{Li}_2\left(\frac{1}{n+1}\right)\right] \\
&= -\zeta(2) + \sum_{k=2}^{\infty}\left[\left(\zeta(2)-\sum_{\ell=1}^{k-1}\frac{1}{\ell^2}\right)(\zeta(k)-1) - \left(1+\frac{(-1)^k}{k-1}\right)\zeta'(k)\right],
\end{aligned}$$

$$\rho(\ln(X),\ln(TX)) = -0.2275522084....$$

The cumulative distribution for TX can be expressed in terms of the digamma function:

$$F(x) = \mathrm{P}(TX \le x) = \sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+x}\right) = \gamma + \psi(x+1),$$

and its density in terms of the trigamma function:

$$f(x) = \sum_{n=1}^{\infty}\frac{1}{(n+x)^2} = \psi'(x+1).$$

For example, the median of TX is $F^{-1}(1/2) = 0.3846747346....$ The cumulative distribution for T^2X is

$$\begin{aligned}
G(x) &= \mathrm{P}(T^2X \le x) = \sum_{n=1}^{\infty}\left(F\left(\frac{1}{n}\right) - F\left(\frac{1}{n+x}\right)\right) \\
&= \sum_{n=1}^{\infty}\left(\psi\left(\frac{1}{n}+1\right) - \psi\left(\frac{1}{n+x}+1\right)\right),
\end{aligned}$$

its density is

$$g(x) = \sum_{n=1}^{\infty}\psi'\left(\frac{1}{n+x}+1\right)\frac{1}{(n+x)^2},$$

and its median is $G^{-1}(1/2) = 0.42278....$ It is certainly inconvenient that $F \neq G$!

4.9.2 *Gauss–Kuzmin Distribution*

If X is a random variable following the Gauss–Kuzmin distribution on $[0, 1]$, then

$$\mathrm{E}(X) = \frac{1}{\ln(2)} - 1 = 0.4426950408... = \mathrm{E}(TX),$$

$$\mathrm{E}(X^2) = 1 - \frac{1}{2\ln(2)} = \mathrm{E}((TX)^2),$$

$$\mathrm{Var}(X) = \frac{(3/2)\ln(2) - 1}{\ln(2)^2} = 0.0826735803... = (0.2875301381...)^2 = \mathrm{Var}(TX)$$

by invariance under T, and

$$\mathrm{E}(X\cdot TX) = 1 - \frac{\gamma}{\ln(2)}, \quad \mathrm{Cov}(X, TX) = \frac{(2-\gamma)\ln(2) - 1}{\ln(2)^2},$$

$$\rho(X, TX) = \frac{(2-\gamma)\ln(2) - 1}{(3/2)\ln(2) - 1} = -0.3474517057....$$

Likewise,

$$\mathrm{E}(\ln(X)) = -\frac{\pi^2}{12\ln(2)} = -1.1865691104... = \mathrm{E}(\ln(TX)),$$

$$\mathrm{E}(\ln(X)^2) = \frac{3\zeta(3)}{2\ln(2)} = \mathrm{E}(\ln(TX)^2),$$

$$\begin{aligned}\mathrm{Var}(\ln(X)) &= \frac{216\ln(2)\zeta(3) - \pi^4}{144\ln(2)^2} = 1.1933560457... \\ &= (1.0924083695...)^2 = \mathrm{Var}(\ln(TX)),\end{aligned}$$

$$\begin{aligned}\mathrm{E}(\ln(X)\cdot\ln(TX)) &= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\left[\frac{1}{2}\ln\left(\frac{n+1}{n}\right)^2\ln((n+1)n) + \ln(n)\,\mathrm{Li}_2\left(-\frac{1}{n}\right)\right. \\ &\quad - \ln(n+1)\,\mathrm{Li}_2\left(-\frac{1}{n+1}\right) + \ln(n+1)\,\mathrm{Li}_2\left(\frac{1}{(n+1)^2}\right) \\ &\quad \left. + 2\,\mathrm{Li}_3\left(-\frac{1}{n}\right) - 2\,\mathrm{Li}_3\left(-\frac{1}{n+1}\right) + \mathrm{Li}_3\left(\frac{1}{(n+1)^2}\right)\right] \\ &= \frac{1}{\ln(2)}\left[-\frac{3\zeta(3)}{2} + \sum_{k=1}^{\infty}\left(\frac{\zeta(2k)-1}{k^3} - \frac{\zeta'(2k)}{k^2} + \frac{\zeta''(2k)}{2k}\right)\right],\end{aligned}$$

$$\rho(\ln(X), \ln(TX)) = -0.1858801270... = r_1.$$

The median of $T^j X$ is $\sqrt{2} - 1 = 0.4142135623...$ for every j. We wish to understand the decay rate of $\rho(X, T^j X)$ and $\rho(\ln(X), \ln(T^j X))$ as j increases, but this appears to be a difficult problem.

4.9.3 *Variance of Sample Mean*

Let us consider the sample mean

$$\hat{\mu}_n(X) = -\frac{1}{n}\sum_{0\le j<n}\ln(T^j X),$$

that is, the average of the time series $\ln(X), \ln(TX), \ldots, \ln(T^{n-1}X)$ built from iterates of T evaluated at X. (The negative sign will simplify subsequent formulation.) It can be proved that

$$\lim_{n\to\infty}\mathrm{E}\left(\hat{\mu}_n(X)\right) = \frac{\pi^2}{12\ln(2)} = 1.1865691104... = \mu,$$

$$\lim_{n\to\infty} n\,\mathrm{Var}\left(\hat{\mu}_n(X)\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{\substack{0\le j<n,\\ 0\le k<n}} \mathrm{Cov}(\ln(T^j X), \ln(T^k X)) = \sigma^2$$

$$\approx \frac{216\ln(2)\zeta(3) - \pi^4}{144\ln(2)^2}\left(1 + \frac{2r_1}{1-r_1}\right) \approx 0.8$$

for a wide variety of initial distributions for X on $[0, 1]$. The latter is a poor numerical estimate (since it presumes that the lag-ℓ correlation r_ℓ is approximately r_1^ℓ, which is not true). It is inspired, in part, by Salamin [7]. A more precise estimate will be given shortly.

4.9.4 *Partial Convergents*

The denominator $Q_n(X)$ of the n^{th} partial convergent to X is connected to our exposition via the formula

$$\underbrace{\ln(Q_n(X))}_{A_n} = \underbrace{-\sum_{0\le j<n} \ln(T^j X) + \varepsilon_n}_{B_n}$$

where $|\varepsilon_n| < c$ for all n, for some constant c. It is clear that

$$\lim_{n\to\infty} \frac{\mathrm{E}(A_n)}{n} = \lim_{n\to\infty} \frac{\mathrm{E}(B_n)}{n} = \mu$$

and further known [8] that

$$0 < \lim_{n\to\infty} \frac{\mathrm{Var}(A_n)}{n} < \infty.$$

We wish to prove that

$$\lim_{n\to\infty} \frac{\mathrm{Var}(A_n)}{n} = \lim_{n\to\infty} \frac{\mathrm{Var}(B_n)}{n}.$$

From $B_n = A_n - \varepsilon_n$, deduce that

$$\mathrm{Var}(B_n) = \mathrm{Var}(A_n) - 2\,\mathrm{Cov}(A_n, \varepsilon_n) + \mathrm{Var}(\varepsilon_n);$$

hence

$$\begin{aligned}
|\mathrm{Var}(A_n) - \mathrm{Var}(B_n)| &\le 2\,|\mathrm{Cov}(A_n, \varepsilon_n)| + \mathrm{Var}(\varepsilon_n) \\
&\le 2\sqrt{\mathrm{Var}(A_n)\,\mathrm{Var}(\varepsilon_n)} + \mathrm{Var}(\varepsilon_n) \\
&\le 2\sqrt{\mathrm{Var}(A_n)\mathrm{E}(\varepsilon_n^2)} + \mathrm{E}(\varepsilon_n^2) \\
&\le 2c\sqrt{\mathrm{Var}(A_n)} + c^2;
\end{aligned}$$

hence

$$\left|\frac{\mathrm{Var}(A_n)}{n} - \frac{\mathrm{Var}(B_n)}{n}\right| \leq 2c\sqrt{\frac{\mathrm{Var}(A_n)}{n^2}} + \frac{c^2}{n} \to 0$$

as $n \to \infty$. In particular,

$$\mathrm{Var}(\ln(Q_n(X))) \sim \sigma^2 n$$

and the importance of computing σ^2 (as attempted using iterates of T) becomes evident.

In fact, the existence of σ^2 (in connection with the denominators Q_n) has been known for a long time. Ibragimov [9], Philipp [10–12] and others [13–19] proved the following Central Limit Theorem:

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\frac{1}{n}\ln(Q_n(X)) - \mu}{\frac{\sigma}{\sqrt{n}}} \leq t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du.$$

No numerical estimate of σ^2 appeared until Flajolet & Vallée [8, 20, 21] computed that

$$\begin{aligned}\sigma^2 &= \lambda_1''(2) - \lambda_1'(2)^2 = 0.8621470373... = (0.9285187329...)^2 \\ &= \frac{1}{4}(9.0803731646...) - \mu^2 = (0.5160624088...) \cdot \mu^3,\end{aligned}$$

where $\lambda_1(s)$ is the dominant eigenvalue of a family of linear operators (indexed by s) on a certain infinite-dimensional function space. Lhote [22, 23] proved that σ^2 is polynomial-time computable and obtained higher accuracy. An elementary expression for σ^2 seems to be impossible. The quantities $4\lambda_1''(2)$ or σ^2/μ^3 are often called **Hensley's constant**.

We close with Loch's theorem [1, 24, 25]:

$$\lim_{n\to\infty} \frac{m(n,x)}{n} = \frac{6\ln(2)\ln(10)}{\pi^2} = 0.9702701143... = (1.0306408341...)^{-1} = \alpha$$

for almost all real x, where $m(n,x)$ is the number of partial denominators of x correctly predicted by the first n decimal digits of x. A corresponding Central Limit Theorem was proved by Faivre [26, 27]:

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\frac{m(n,X)}{n} - \alpha}{\frac{\theta}{\sqrt{n}}} \leq t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du$$

where

$$\begin{aligned}\theta^2 &= \frac{\alpha\,\sigma^2}{\mu^2} = \frac{864\ln(2)^3\ln(10)}{\pi^6}\sigma^2 \\ &= 0.5941388048... = (0.7708039990...)^2.\end{aligned}$$

For example, the first 10000 decimal digits of π give 9757 partial denominators, consistent with the value of α. A similar empirical confirmation of the value of θ would be good to see.

Acknowledgments I thank Eugene Salamin, William Gosper, Philippe Flajolet and Brigitte Vallée for helpful discussions in 1999. Regrettably, in early printings of [28], the formula for σ^2/μ^3 is wrong by a factor of π^6.

[1] K. Dajani and C. Kraaikamp, *Ergodic Theory of Numbers*, Math. Assoc. Amer., 2002, pp. 20–26, 80–88, 172–175; MR1917322 (2003f:37014).
[2] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[3] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[4] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, v. 1, Springer-Verlag 1998, problems 42, 43; MR0580154 (81e:00002).
[5] S. R. Finch, Alladi-Grinstead constant, *first volume*, pp. 120–122.
[6] P. Erdős, S. W. Graham, A. Ivić, and C. Pomerance, On the number of divisors of $n!$, *Analytic Number Theory*, Proc. 1995 Allerton Park conf., v. 1, ed. B. C. Berndt, H. G. Diamond, and A. J. Hildebrand, Birkhäuser, 1996, pp. 337–355; MR1399347 (97d:11142).
[7] W. Gosper and E. Salamin, Lévy's (1936) limit and zeta(3), unpublished note (1999).
[8] S. R. Finch, Gauss-Kuzmin-Wirsing constant, *first volume*, pp. 151–156.
[9] I. A. Ibragimov, A theorem from the metric theory of continued fractions (in Russian), *Vestnik Leningrad. Univ.* v. 16 (1961) n. 1, 13–24; MR0133619 (24 #A3445).
[10] W. Philipp, Ein zentraler Grenzwertsatz mit Anwendungen auf die Zahlentheorie, *Z. Wahrsch. Verw. Gebiete* 8 (1967) 185–203; MR0215360 (35 #6201).
[11] W. Philipp and O. P. Stackelberg, Zwei Grenzwertsätze für Kettenbrüche, *Math. Annalen* 181 (1969) 152–156; MR0244186 (39 #5503).
[12] W. Philipp, Some metrical theorems in number theory. II, *Duke Math. J.* 37 (1970) 447–458; errata 37 (1970) 788; MR0272739 (42 #7620) and MR0274412 (43 #177).
[13] M. I. Gordin and M. H. Reznik, The law of the iterated logarithm for the denominators of continued fractions (in Russian), *Vestnik Leningrad. Univ.* v. 25 (1970) n. 13, 28–33; MR0276191 (43 #1939).
[14] M. I. Gordin, The behavior of the dispersion of sums of random variables that generate a stationary process (in Russian), *Teor. Verojatnost. i Primenen.* 16 (1971) 484–494; Engl. transl. in *Theor. Probability Appl.* 10 (1971) 474–484; MR0287606 (44 #4809).
[15] G. A. Misjavicus, Evaluation of remainders in limit theorems for functions of elements of continued fractions (in Russian), *Litovsk. Mat. Sb.* 10 (1970) 293–308; MR0296041 (45 #5102).
[16] G. A. Misjavicus, Evaluation of remainders in limit theorems for denominators of continued fractions (in Russian), *Litovsk. Mat. Sb.* 21 (1981) 63–74; Engl. transl. in *Lithuanian Math. J.* 21 (1981) 245–253; MR0637846 (83d:10063).
[17] T. Morita, Local limit theorem and distribution of periodic orbits of Lasota-Yorke transformations with infinite Markov partition, *J. Math. Soc. Japan* 46 (1994) 309–343; correction 47 (1995) 191–192; MR1264944 (95h:58079) and MR1304197 (95k:58095).
[18] B. Vallée, Opérateurs de Ruelle-Mayer généralisés et analyse en moyenne des algorithmes d'Euclide et de Gauss, *Acta Arith.* 81 (1997) 101–144; MR1456238 (98g:11091).

[19] P. Flajolet and B. Vallée, Continued fraction algorithms, functional operators, and structure constants, *Theoret. Comput. Sci.* 194 (1998) 1–34; MR1491644 (98j:11061).
[20] P. Flajolet and B. Vallée, Hensley's constant, unpublished note (1999).
[21] P. Flajolet and B. Vallée, Continued fractions, comparison algorithms, and fine structure constants, *Constructive, Experimental, and Nonlinear Analysis*, Proc. 1999 Limoges conf., ed. M. Théra, Amer. Math. Soc., 2000, pp. 53-82; INRIA preprint RR4072; MR1777617 (2001h:11161).
[22] L. Lhote, Modélisation et approximation de sources complexes, Masters thesis, University of Caen, 2002.
[23] L. Lhote, Computation of a class of continued fraction constants, *Analytic Algorithmics and Combinatorics (ANALCO)*, Proc. 2004 New Orleans workshop.
[24] S. R. Finch, Khintchine-Lévy constants, *first volume*, pp. 59–65.
[25] G. Lochs, Gustav Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch, *Abh. Math. Sem. Univ. Hamburg* 27 (1964) 142–144; MR0162753 (29 #57).
[26] C. Faivre, A central limit theorem related to decimal and continued fraction expansion, *Arch. Math. (Basel)* 70 (1998) 455–463; MR1621982 (99m:11088).
[27] C. Faivre, On calculating a continued fraction expansion from a decimal expansion, *Acta Sci. Math. (Szeged)* 67 (2001) 505–519; MR1876450 (2002j:11088).
[28] S. R. Finch, Porter-Hensley constants, *first volume*, pp. 156–160.

4.10 Continued Fraction Transformation. II

As in our earlier essay [1], define $T: [0,1] \to [0,1]$ by

$$T(x) = \begin{cases} \left\{ \dfrac{1}{x} \right\} & \text{if } 0 < x \le 1, \\ 0 & \text{if } x = 0 \end{cases}$$

where $\{\xi\} = \xi - \lfloor \xi \rfloor$ denotes the fractional part of ξ. Previously, we examined the moments of $T^j X$ and of $\ln(T^j X)$, where X is a random variable in $[0,1]$. The distribution of X was assumed to be either uniform or Gauss–Kuzmin.

What can be said about the moments of $\lfloor 1/T^j X \rfloor$ and of $\ln \lfloor 1/T^j X \rfloor$? An answer to this question helps in determining the asymptotic distribution of the first n continued fraction "digits", corresponding to uniformly distributed X as $n \to \infty$.

4.10.1 *Uniform Distribution*

Let γ denote the Euler–Mascheroni constant, ψ denote the, and ζ denote the Riemann zeta function. If X is a random variable following the uniform distribution on $[0,1]$, then

$$\mathrm{E}\left\lfloor \frac{1}{X} \right\rfloor = \int_1^\infty \frac{\lfloor y \rfloor}{y^2} dy \sim \sum_{n \le N} \int_n^{n+1} \frac{n}{y^2} dy \sim \sum_{n \le N} n \left(\frac{1}{n} - \frac{1}{n+1} \right) \sim \sum_{n \le N} \frac{1}{n+1} \sim \ln(N),$$

$$\begin{aligned}
\mathrm{E}\left\lfloor\frac{1}{TX}\right\rfloor &= \int_{1}^{\infty}\left\lfloor\frac{1}{\{y\}}\right\rfloor\frac{dy}{y^2}\sim\sum_{n\le N}\int_{n}^{n+1}\left\lfloor\frac{1}{y-n}\right\rfloor\frac{dy}{y^2}\\
&\sim\sum_{n\le N}\int_{0}^{1}\left\lfloor\frac{1}{z}\right\rfloor\frac{dy}{(z+n)^2}\sim\sum_{n\le N}\int_{1}^{\infty}\frac{\lfloor w\rfloor}{(1+n\,w)^2}dw\\
&\sim\sum_{n\le N}\sum_{m\le N}\int_{m}^{m+1}\frac{m}{(1+n\,w)^2}dw\sim\sum_{n\le N}\sum_{m\le N}\frac{m}{n}\left(\frac{1}{1+n\,m}-\frac{1}{1+n(m+1)}\right)\\
&\sim\sum_{n\le N}\sum_{m\le N}\frac{1}{n(1+n\,m)}\sim\sum_{m\le N}\left(\psi\left(1+\frac{1}{m}\right)+\gamma\right)\sim\frac{\pi^2}{6}\ln(N)
\end{aligned}$$

as $N\to\infty$, via the substitutions $y=1/x$, $z=y-n$ and $w=1/z$. Hence both expected values are infinite. By contrast,

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\right) &= \int_{1}^{\infty}\frac{\ln\lfloor y\rfloor}{y^2}dy=\sum_{n=1}^{\infty}\int_{n}^{n+1}\frac{\ln(n)}{y^2}dy=\sum_{n=1}^{\infty}\frac{\ln(n)}{n(n+1)}\\
&= -\sum_{k=2}^{\infty}(-1)^k\zeta'(k)=0.7885305659...
\end{aligned}$$

(Lüroth analog of Khintchine's constant [2]),

$$\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor^2\right)=\sum_{n=1}^{\infty}\frac{\ln(n)^2}{n(n+1)}=\sum_{k=2}^{\infty}(-1)^k\zeta''(k),$$

$$\mathrm{Var}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\right)=1.1759638742...=(1.0844186803...)^2,$$

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor\right) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\ln(m)}{n}\left(\frac{1}{1+n\,m}-\frac{1}{1+n(m+1)}\right)\\
&= \sum_{n=1}^{\infty}\sum_{m=2}^{\infty}\frac{\ln(m)-\ln(m-1)}{n(1+n\,m)}\\
&= \sum_{m=2}^{\infty}(\ln(m)-\ln(m-1))\left(\psi\left(1+\frac{1}{m}\right)+\gamma\right)\\
&= \sum_{k=2}^{\infty}(-1)^k\zeta(k)\sum_{j=1}^{\infty}\binom{1-k}{j}\zeta'(j+k-1)\\
&= 1.06479...,
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor^2\right) &= \sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\ln(m)^2}{n}\left(\frac{1}{1+nm}-\frac{1}{1+n(m+1)}\right)\\
&= \sum_{m=2}^{\infty}\left(\ln(m)^2-\ln(m-1)^2\right)\left(\psi\left(1+\frac{1}{m}\right)+\gamma\right)\\
&= -\sum_{k=2}^{\infty}(-1)^k\zeta(k)\sum_{j=1}^{\infty}\binom{1-k}{j}\zeta''(j+k-1),
\end{aligned}$$

$$\mathrm{Var}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor\right) = 1.49522... = (1.22279...)^2.$$

We shall not attempt to compute the cross-moments

$$\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\cdot\ln\left\lfloor\frac{1}{TX}\right\rfloor\right)\quad\text{or}\quad\rho\left(\ln\left\lfloor\frac{1}{X}\right\rfloor,\ln\left\lfloor\frac{1}{TX}\right\rfloor\right)$$

and leave these as open problems.

4.10.2 *Gauss–Kuzmin Distribution*

If X is a random variable following the Gauss–Kuzmin distribution on $[0, 1]$, then

$$\begin{aligned}
\mathrm{E}\left\lfloor\frac{1}{X}\right\rfloor &= \frac{1}{\ln(2)}\int_1^{\infty}\frac{\lfloor y\rfloor}{y(y+1)}dy \sim \frac{1}{\ln(2)}\sum_{n\le N}\int_n^{n+1}\frac{n}{y(y+1)}dy\\
&\sim \frac{1}{\ln(2)}\sum_{n\le N} n\ln\left(1+\frac{1}{n(n+2)}\right) \sim \frac{1}{\ln(2)}\ln(N) \sim \mathrm{E}\left\lfloor\frac{1}{TX}\right\rfloor
\end{aligned}$$

as $N\to\infty$. Hence both expected values are infinite. By contrast,

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\right) &= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\ln(n)\ln\left(1+\frac{1}{n(n+2)}\right)\\
&= \frac{1}{\ln(2)}\sum_{j=2}^{\infty}(-1)^j\frac{2\zeta'(j)-2^j\left(\zeta'(j)+\frac{\ln(2)}{2^j}+\frac{\ln(3)}{3^j}\right)}{j}\\
&\qquad +(1-\ln(2))+\frac{\ln(3)}{\ln(2)}\left(\frac{2}{3}-\ln\left(\frac{5}{3}\right)\right)\\
&= 0.9878490568... = \ln(K) = \mathrm{E}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor\right)
\end{aligned}$$

(Khintchine's constant [2]),

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor^2\right) &= \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\ln(n)^2\ln\left(1+\frac{1}{n(n+2)}\right)\\
&= -\frac{1}{\ln(2)}\sum_{j=2}^{\infty}(-1)^j\frac{2\zeta''(j)-2^j\left(\zeta''(j)-\frac{\ln(2)^2}{2^j}-\frac{\ln(3)^2}{3^j}\right)}{j}\\
&\quad+\ln(2)\left(1-\ln(2)\right)+\frac{\ln(3)^2}{\ln(2)}\left(\frac{2}{3}-\ln\left(\frac{5}{3}\right)\right)\\
&= \mathrm{E}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor^2\right),
\end{aligned}$$

$$\mathrm{Var}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\right)=1.4094310970...=(1.1871946331...)^2=\mathrm{Var}\left(\ln\left\lfloor\frac{1}{TX}\right\rfloor\right).$$

The joint expectation

$$\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\cdot\ln\left\lfloor\frac{1}{TX}\right\rfloor\right)$$

simplifies to

$$\frac{1}{\ln(2)}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\ln(n)\ln(m)\ln\left(1+\frac{1}{(1+(n+1)m)(1+n(m+1))}\right)$$

and can be numerically evaluated via suitable generalization of Kummer's method [3]. It follows that the cross-correlation is

$$\rho\left(\ln\left\lfloor\frac{1}{X}\right\rfloor,\ln\left\lfloor\frac{1}{TX}\right\rfloor\right)=-0.0876526887...=r_1.$$

4.10.3 *Variance of Sample Mean*

The sample mean

$$\hat{\mu}_n(X)=\frac{1}{n}\sum_{0\le j<n}\ln\left\lfloor\frac{1}{T^jX}\right\rfloor$$

satisfies

$$\lim_{n\to\infty}\mathrm{E}\left(\hat{\mu}_n(X)\right)=\ln(K)=0.9878490568...=\mu,$$

$$\begin{aligned}
\lim_{n\to\infty}n\,\mathrm{Var}\left(\hat{\mu}_n(X)\right) &= \lim_{n\to\infty}\frac{1}{n}\sum_{\substack{0\le j<n,\\0\le k<n}}\mathrm{Cov}(\ln(T^jX),\ln(T^kX))=\sigma^2\\
&\approx \mathrm{Var}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\right)\left(1+\frac{2r_1}{1-r_1}\right)\approx 1.2
\end{aligned}$$

for a wide variety of initial distributions for X on $[0,1]$. (No negative sign is introduced this time in the definition of $\hat{\mu}_n(X)$, unlike before.)

4.10.4 *Continued Fraction Digits*

If $a_1, a_2, a_3, \ldots$ denote the partial denominators (digits) of X, then it is clear that

$$\ln\left((a_1 a_2 a_3 \cdots a_n)^{\frac{1}{n}}\right) = \frac{1}{n} \sum_{0 \le j < n} \ln \left\lfloor \frac{1}{T^j X} \right\rfloor$$

(no nonzero error ε_n is present here). Baladi & Vallée [4] proved that the following Central Limit Theorem is true:

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\frac{1}{n}(\ln a_1 + \ln a_2 + \cdots + \ln a_n) - \mu}{\frac{\sigma}{\sqrt{n}}} \le t \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du$$

and Lhote [5] computed that

$$\sigma^2 = 1.2297301427... = (1.1089319829...)^2.$$

What happens if we omit the logarithms on the left-hand side? Since a_k has infinite expectation, it is not surprising that asymptotic normality fails. Lévy [6], Philipp [7], Heinrich [8] and Hensley [9] proved that

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\ln(2)}{n} \sum_{k=1}^{n} a_k - (\ln(n) - \gamma - \ln(\ln(2))) \le t \right) = \int_{-\infty}^{t} f(u)\, du,$$

where the density f of the limiting stable distribution $S(1, 1, \pi/2, 0; 1)$ is given by

$$f(u) = \frac{1}{\pi} \int_{0}^{\infty} \sin(\pi v) \exp(-v \ln(v) - u v)\, dv.$$

See Figure 4.3. The median of f is 1.35578... and the mode of f is −0.22278.... Extreme asymmetry and a heavy right-tail are the most noticeable features here!

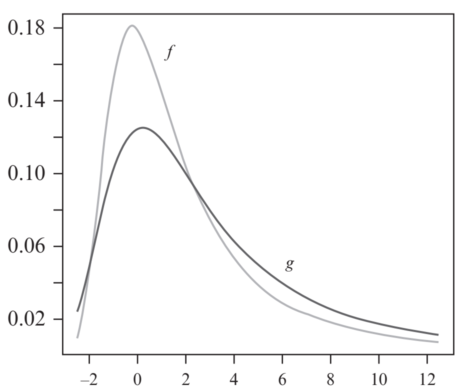

Figure 4.3 Two non-normal limiting stable distributions.

As a footnote, let us return to some very simple ideas. If $X_1, X_2, \ldots, X_n$ is an independent sample from the uniform distribution and $Y_1, Y_2, \ldots, Y_n$ is an independent sample from the Gauss–Kuzmin distribution, then

$$\mathrm{P}\left(\frac{\frac{1}{n}\sum\limits_{k=1}^{n} X_k - \frac{1}{2}}{\frac{1}{6}\sqrt{\frac{3}{n}}} \le t\right) \to \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du$$

$$\leftarrow \mathrm{P}\left(\frac{\frac{1}{n}\sum\limits_{k=1}^{n} Y_k - \left(\frac{1}{\ln(2)} - 1\right)}{\frac{1}{\ln(2)}\sqrt{\frac{(3/2)\ln(2)-1}{n}}} \le t\right)$$

as $n \to \infty$. Also, the distributions of reciprocals have densities

$$\frac{d}{dt}\mathrm{P}\left(\frac{1}{X} \le t\right) = \begin{cases} \dfrac{1}{t^2} & \text{if } t \ge 1, \\ 0 & \text{otherwise;} \end{cases}$$

$$\frac{d}{dt}\mathrm{P}\left(\frac{1}{Y} \le t\right) = \begin{cases} \dfrac{1}{\ln(2)}\dfrac{1}{t\,(t+1)} & \text{if } t \ge 1, \\ 0 & \text{otherwise.} \end{cases}$$

The expectations of $1/X$ and of $1/Y$ are infinite. Our ideas hence become vastly more complicated at this point [9]:

$$\mathrm{P}\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{X_k} - (\ln(n) + 1 - \gamma) \le t\right) \to \int\limits_{-\infty}^{t} f(u)\, du$$

where f is exactly as before, and

$$\mathrm{P}\left(\frac{1}{n}\sum_{k=1}^{n}\frac{1}{Y_k} - \frac{\ln(n) + 1 - \ln(2) - \gamma}{\ln(2)} \le t\right) \to \int\limits_{-\infty}^{t} g(u)\, du,$$

where

$$g(u) = \frac{1}{\pi}\int\limits_{0}^{\infty} \sin\left(\frac{\pi\, v}{\ln(2)}\right)\exp\left(-\frac{v}{\ln(2)}\ln(v) - u\, v\right) dv$$

is the density of the limiting stable distribution $S(1, 1, \pi/(2\ln(2)), 0; 1)$. The median of g is 2.48474... and the mode of g is 0.20735...; asymmetry and a heavy right-tail again dominate. A wealth of materials on calculating stable distributions is available [10–12].

Acknowledgments I thank Pascal Sebah for computing r_1, Loïck Lhote for computing σ^2, John Nolan for his STABLE software [12], and Doug Hensley for helpful discussions.

[1] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.
[2] S. R. Finch, Khintchine-Lévy constants, *first volume*, pp. 59–65.
[3] P. Sebah, Accelerating convergence of a double series, unpublished note (2007).
[4] V. Baladi and B. Vallée, Euclidean algorithms are Gaussian, *J. Number Theory* 110 (2005) 331–386; MR2122613 (2006e:11192).
[5] L. Lhote, The variance associated with Khintchine's constant, unpublished note (2007).
[6] P. Lévy, Fractions continues aléatoires, *Rend. Circ. Mat. Palermo* 1 (1952) 170–208; also in *Œuvres de Paul Lévy*, v. VI, *Théorie des jeux*, ed. D. Dugué, P. Deheuvels and M. Ibéro, Gauthier-Villars, 1980, pp. 361–399; MR0066583 (16,600e) and MR0586767 (83c:01070).
[7] W. Philipp, Limit theorems for sums of partial quotients of continued fractions, *Monatsh. Math.* 105 (1988) 195–206; MR0939942 (89e:60069).
[8] L. Heinrich, Rates of convergence in stable limit theorems for sums of exponentially ψ-mixing random variables with an application to metric theory of continued fractions, *Math. Nachr.* 131 (1987) 149–165; MR0908807 (89a:60061).
[9] D. Hensley, The statistics of the continued fraction digit sum, *Pacific J. Math.* 192 (2000) 103–120; MR1741027 (2000k:11092); Zbl. 1015.11038.
[10] J. P. Nolan, Numerical calculation of stable densities and distribution functions. Heavy tails and highly volatile phenomena, *Comm. Statist. Stochastic Models* 13 (1997) 759–774; MR1482292 (98f:60024).
[11] I. A. Belov, On the computation of the probability density function of α-stable distributions, *Proc. 10th Internat. Conf. Mathematical Modelling and Analysis – 2nd Internat. Conf. Computational Methods in Applied Mathematics*, Trakai, 2005, ed. R. Čiegis, pp. 333–341; MR2194689.
[12] J. P. Nolan, *Stable Distributions. Models for Heavy Tailed Data*, Birkhäuser, to appear.

4.11 Continued Fraction Transformation. III

We continue the discussion from our earlier essays [1, 2], turning attention first to two variations on regular continued fractions (RCFs). For reasons of space, only first-order results (means) will be presented. After this, we exhibit formulas connected with Lüroth representations and with ordinary decimal representations.

4.11.1 *Nearest Integer Continued Fractions*

Define $T: [-1/2, 1/2] \to [-1/2, 1/2]$ by

$$T(x) = \begin{cases} \dfrac{1}{x} - \left\lfloor \dfrac{1}{x} + \dfrac{1}{2} \right\rfloor & \text{if } -1/2 \le x \le 1/2 \text{ and } x \ne 0, \\ 0 & \text{if } x = 0. \end{cases}$$

For example,

$$\begin{aligned}
\pi - 3 &= 0.141592..., & \left\lfloor \tfrac{1}{\pi-3} + \tfrac{1}{2} \right\rfloor &= 7, \\
T(\pi - 3) &= 0.062513..., & \left\lfloor \tfrac{1}{T(\pi-3)} + \tfrac{1}{2} \right\rfloor &= 16, \\
T^2(\pi - 3) &= -0.003405..., & \left\lfloor \tfrac{1}{T^2(\pi-3)} + \tfrac{1}{2} \right\rfloor &= -294, \\
T^3(\pi - 3) &= 0.365409..., & \left\lfloor \tfrac{1}{T^3(\pi-3)} + \tfrac{1}{2} \right\rfloor &= 3, \\
T^4(\pi - 3) &= -0.263340..., & \left\lfloor \tfrac{1}{T^4(\pi-3)} + \tfrac{1}{2} \right\rfloor &= -4
\end{aligned}$$

and

$$\begin{aligned}
\pi &= 3 + \frac{1|}{|7} + \frac{1|}{|16} + \frac{1|}{|-294} + \frac{1|}{|3} + \frac{1|}{|-4} + \frac{1|}{|5} + \frac{1|}{|-15} + \frac{1|}{|-3} + \frac{1|}{|2} + \cdots \\
&= 3 + \frac{1|}{|7} + \frac{1|}{|16} - \frac{1|}{|294} - \frac{1|}{|3} - \frac{1|}{|4} - \frac{1|}{|5} - \frac{1|}{|15} + \frac{1|}{|3} - \frac{1|}{|2} + \cdots
\end{aligned}$$

is the **nearest integer continued fraction** (NICF) expansion for π. This is also called a **centered continued fraction**. Let X be a random variable in $[-1/2, 1/2]$ with density

$$\frac{d}{dx}\mathrm{P}(X \le x) = \begin{cases} \dfrac{1}{\ln(\varphi)} \dfrac{1}{\varphi + 1 + x} & \text{if } -1/2 \le x < 0, \\ \dfrac{1}{\ln(\varphi)} \dfrac{1}{\varphi + x} & \text{if } 0 \le x \le 1/2 \end{cases}$$

where $\varphi = (1+\sqrt{5})/2$ denotes the Golden mean [3]. What is the mean of $\ln(|X|)$? This is equal to the asymptotic mean of $(1/n)\ln q_n$, corresponding to denominators q_n in the partial convergents to x:

$$\frac{p_1}{q_1} = \frac{3}{1}, \; \frac{p_2}{q_2} = \frac{22}{7}, \; \frac{p_3}{q_3} = \frac{355}{113}, \; \frac{p_4}{q_4} = \frac{104348}{33215}, \; \frac{p_5}{q_5} = \frac{312689}{99532}, \; ...,$$

as $n \to \infty$. It follows that [4, 5]

$$\begin{aligned}
\mathrm{E}(\ln(|X|)) &= \frac{1}{\ln(\varphi)} \int_{-1/2}^{0} \frac{\ln(-x)}{\varphi + 1 + x} dx + \frac{1}{\ln(\varphi)} \int_{0}^{1/2} \frac{\ln(x)}{\varphi + x} dx \\
&= -\frac{\pi^2}{12 \ln(\varphi)} = -1.7091579853... = \mathrm{E}(\ln(|TX|)).
\end{aligned}$$

Also, what is the mean of $\ln(|a_1|)$, where a_1, a_2, a_3, ... denote the partial denominators (digits) of X? Using the substitution $y=\pm 1/x$, it follows that [6, 7]

$$\begin{aligned}
\mathrm{E}\left(\ln\left|\left\lfloor \frac{1}{X}+\frac{1}{2}\right\rfloor\right|\right) &= \frac{1}{\ln(\varphi)}\int_{-1/2}^{0}\frac{\ln\left\lfloor -\frac{1}{x}+\frac{1}{2}\right\rfloor}{\varphi+1+x}dx+\frac{1}{\ln(\varphi)}\int_{0}^{1/2}\frac{\ln\left\lfloor \frac{1}{x}+\frac{1}{2}\right\rfloor}{\varphi+x}dx \\
&= \frac{1}{\ln(\varphi)}\int_{2}^{\infty}\left(\frac{\ln\left\lfloor y+\frac{1}{2}\right\rfloor}{y((\varphi+1)y-1)}+\frac{\ln\left\lfloor y+\frac{1}{2}\right\rfloor}{y(\varphi y+1)}\right)dy \\
&= \frac{1}{\ln(\varphi)}\int_{2}^{5/2}\left(\frac{\ln(2)}{y((\varphi+1)y-1)}+\frac{\ln(2)}{y(\varphi y+1)}\right)dy \\
&\quad +\frac{1}{\ln(\varphi)}\sum_{n=3}^{\infty}\int_{n-1/2}^{n+1/2}\left(\frac{\ln(n)}{y((\varphi+1)y-1)}+\frac{\ln(n)}{y(\varphi y+1)}\right)dy \\
&= \frac{\ln(2)}{\ln(\varphi)}\ln\left(\frac{5\varphi+3}{5\varphi+2}\right) \\
&\quad +\frac{1}{\ln(\varphi)}\sum_{n=3}^{\infty}\ln(n)\ln\left(\frac{(\varphi+1)(n+\frac{1}{2})-1}{(\varphi+1)(n-\frac{1}{2})-1}\,\frac{\varphi(n-\frac{1}{2})+1}{\varphi(n+\frac{1}{2})+1}\right) \\
&= 1.6964441175...=\mathrm{E}\left(\ln\left|\left\lfloor \frac{1}{TX}+\frac{1}{2}\right\rfloor\right|\right).
\end{aligned}$$

These two constants are the NICF analogs of Lévy's constant and Khintchine's constant, respectively. A Central Limit Theorem exists in both cases [8], but the associated variances have not yet been numerically evaluated.

4.11.2 *Odd Digit Continued Fractions*

Define $T\colon [0,1]\to[0,1]$ by

$$T(x)=\begin{cases}\dfrac{1}{x}-\left\lfloor\dfrac{1}{x}\right\rfloor & \text{if } \left\lfloor\dfrac{1}{x}\right\rfloor\equiv 1 \bmod 2 \text{ and } x\neq 0,\\[2ex] \left\lceil\dfrac{1}{x}\right\rceil-\dfrac{1}{x} & \text{if } \left\lceil\dfrac{1}{x}\right\rceil\equiv 1 \bmod 2 \text{ and } x\neq 0,\\[2ex] 0 & \text{if } x=0.\end{cases}$$

For example,

$$\begin{aligned}
\pi-3&=0.141592..., & \left\lfloor \tfrac{1}{\pi-3} \right\rfloor &=7,\\
T(\pi-3)&=0.062513..., & \left\lfloor \tfrac{1}{T(\pi-3)} \right\rfloor &=15,\\
T^2(\pi-3)&=0.996594..., & \left\lfloor \tfrac{1}{T^2(\pi-3)} \right\rceil &=1,\\
T^3(\pi-3)&=0.003417..., & \left\lceil \tfrac{1}{T^3(\pi-3)} \right\rceil &=293,\\
T^4(\pi-3)&=0.365409..., & \left\lceil \tfrac{1}{T^4(\pi-3)} \right\rceil &=3,\\
T^5(\pi-3)&=0.263340..., & \left\lfloor \tfrac{1}{T^5(\pi-3)} \right\rfloor &=3,\\
T^6(\pi-3)&=0.797366..., & \left\lfloor \tfrac{1}{T^6(\pi-3)} \right\rfloor &=1
\end{aligned}$$

and

$$\begin{aligned}
\pi &= 3+\frac{1|}{|7}+\frac{1|}{|15}+\frac{1|}{|1}+\frac{1|}{|293}-\frac{1|}{|3}-\frac{1|}{|3}+\frac{1|}{|1}+\frac{1|}{|3}+\frac{1|}{|1}+\frac{1|}{|15}+\cdots\\
&= 3+\frac{1|}{|7}+\frac{1|}{|15}+\frac{1|}{|1}+\frac{1|}{|293}+\frac{1|}{|-3}+\frac{1|}{|3}+\frac{1|}{|1}+\frac{1|}{|3}+\frac{1|}{|1}+\frac{1|}{|15}+\cdots
\end{aligned}$$

is the **odd digit continued fraction** (ODCF) expansion for π. The phrase "partial denominator" or "partial quotient" often replaces the word "digit". Let X be a random variable in $[0,1]$ with density

$$\frac{d}{dx}\mathrm{P}(X\le x)=\frac{1}{3\ln(\varphi)}\left(\frac{1}{\varphi-1+x}+\frac{1}{\varphi+1-x}\right)$$

where φ is as before. What is the mean of $\ln(X)$? This is equal to the asymptotic mean of $(1/n)\ln q_n$, corresponding to denominators q_n in the partial convergents to x :

$$\frac{p_1}{q_1}=\frac{3}{1},\ \frac{p_2}{q_2}=\frac{22}{7},\ \frac{p_3}{q_3}=\frac{333}{106},\ \frac{p_4}{q_4}=\frac{355}{113},\ \frac{p_5}{q_5}=\frac{104348}{33215},\ ...,$$

as $n\to\infty$. It follows that [9, 10]

$$\begin{aligned}
\mathrm{E}(\ln(X)) &= \frac{1}{3\ln(\varphi)}\int_0^1\left(\frac{\ln(x)}{\varphi-1+x}+\frac{\ln(x)}{\varphi+1-x}\right)dx\\
&= -\frac{\pi^2}{18\ln(\varphi)}=-1.1394386568...=\mathrm{E}(\ln(TX)).
\end{aligned}$$

Also, what is the mean of $\ln(|a_1|)$, where $a_1, a_2, a_3, \ldots$ denote the digits of X? Let $\lfloor z\rceil=\lfloor z\rfloor$ if $\lfloor z\rfloor$ is odd and $\lfloor z\rceil=\lceil z\rceil$ otherwise. Using the substitution $y=1/x$, it follows that [11]

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rceil\right) &= \frac{1}{3\ln(\varphi)}\int_0^1 \ln\left\lfloor\frac{1}{x}\right\rceil\left(\frac{1}{\varphi-1+x}+\frac{1}{\varphi+1-x}\right)dx \\
&= \frac{1}{3\ln(\varphi)}\int_1^\infty \ln\lfloor y\rceil\left(\frac{1}{y((\varphi-1)y+1)}+\frac{1}{y((\varphi+1)y-1)}\right)dy \\
&= \frac{1}{3\ln(\varphi)}\sum_{n=1}^\infty\int_{2n}^{2n+2} \ln(2n+1)\left(\frac{1}{y((\varphi-1)y+1)}+\frac{1}{y((\varphi+1)y-1)}\right)dy \\
&= \frac{1}{3\ln(\varphi)}\sum_{n=1}^\infty \ln(2n+1)\ln\left(\frac{2(\varphi+1)(n+1)-1}{2(\varphi+1)n-1}\frac{2(\varphi-1)n+1}{2(\varphi-1)(n+1)+1}\right) \\
&= 1.0283554474... = \mathrm{E}\left(\ln\left\lfloor\frac{1}{TX}\right\rceil\right).
\end{aligned}$$

These two constants are the ODCF analogs of Lévy's constant and Khintchine's constant, respectively. A Central Limit Theorem exists in both cases [8], but again the associated variances have not yet been numerically evaluated.

4.11.3 *Lüroth Representations*

Define $A : [0, 1] \to [0, 1]$ by

$$A(x) = \begin{cases} \left\lfloor\frac{1}{x}\right\rfloor\left(x\left\lceil\frac{1}{x}\right\rceil - 1\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

and $B : [0, 1] \to [0, 1]$ by

$$B(x) = \begin{cases} \left\lceil\frac{1}{x}\right\rceil\left(1 - x\left\lfloor\frac{1}{x}\right\rfloor\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

For example,

$$\begin{aligned}
a_1 &= \left\lfloor\tfrac{1}{\pi-3}\right\rfloor = 7, & b_1 &= \left\lfloor\tfrac{1}{\pi-3}\right\rfloor = 7, \\
a_2 &= \left\lfloor\tfrac{1}{A(\pi-3)}\right\rfloor = 1, & b_2 &= \left\lfloor\tfrac{1}{B(\pi-3)}\right\rfloor = 14, \\
a_3 &= \left\lfloor\tfrac{1}{A^2(\pi-3)}\right\rfloor = 1, & b_3 &= \left\lfloor\tfrac{1}{B^2(\pi-3)}\right\rfloor = 7, \\
a_4 &= \left\lfloor\tfrac{1}{A^3(\pi-3)}\right\rfloor = 1, & b_4 &= \left\lfloor\tfrac{1}{B^3(\pi-3)}\right\rfloor = 1,
\end{aligned}$$

$$\begin{aligned}
a_5 &= \left\lfloor \tfrac{1}{A^4(\pi-3)} \right\rfloor = 2, & b_5 &= \left\lfloor \tfrac{1}{B^4(\pi-3)} \right\rfloor = 1,\\
a_6 &= \left\lfloor \tfrac{1}{A^5(\pi-3)} \right\rfloor = 1, & b_6 &= \left\lfloor \tfrac{1}{B^5(\pi-3)} \right\rfloor = 1,\\
a_7 &= \left\lfloor \tfrac{1}{A^6(\pi-3)} \right\rfloor = 4, & b_7 &= \left\lfloor \tfrac{1}{B^6(\pi-3)} \right\rfloor = 15,\\
a_8 &= \left\lfloor \tfrac{1}{A^7(\pi-3)} \right\rfloor = 23, & b_8 &= \left\lfloor \tfrac{1}{B^7(\pi-3)} \right\rfloor = 1
\end{aligned}$$

and

$$\begin{aligned}
\pi &= 3 + \frac{1}{a_1+1} + \sum_{n=2}^{\infty} \left(\prod_{k=1}^{n-1} \frac{1}{a_k(a_k+1)} \right) \frac{1}{a_n+1}\\
&= 3 + \frac{1}{b_1} + \sum_{n=2}^{\infty} \left(\prod_{k=1}^{n-1} \frac{1}{b_k(b_k+1)} \right) \frac{(-1)^{n-1}}{b_n}
\end{aligned}$$

are the **positive Lüroth** and **alternating Lüroth representations** for π, respectively. The limiting constants are the same whether we use as or bs. For uniformly distributed X, it follows that

$$\mathrm{E}\left(\ln\left\lfloor \frac{1}{X} \right\rfloor\right) = \sum_{n=1}^{\infty} \frac{\ln(n)}{n(n+1)} = -\sum_{k=2}^{\infty} (-1)^k \zeta'(k) = 0.7885305659...$$

(Lüroth analog of Khintchine's constant [4, 12, 13]),

$$\mathrm{E}\left(\ln\left\lceil \frac{1}{X} \right\rceil\right) = \sum_{n=1}^{\infty} \frac{\ln(n+1)}{n(n+1)} = -\sum_{k=2}^{\infty} \zeta'(k) = 1.2577468869...$$

(which appeared earlier [1]),

$$\begin{aligned}
\mathrm{E}\left(\ln\left\lfloor \frac{1}{X} \right\rfloor + \ln\left\lceil \frac{1}{X} \right\rceil\right) &= \sum_{n=1}^{\infty} \frac{\ln(n(n+1))}{n(n+1)}\\
&= -2\sum_{k=1}^{\infty} \zeta'(2k) = 2.0462774528...
\end{aligned}$$

(Lüroth analog of Lévy's constant [14]),

$$\mathrm{E}\left(\ln\left\lfloor \frac{1}{X} \right\rfloor^2\right) = \sum_{n=1}^{\infty} \frac{\ln(n)^2}{n(n+1)} = \sum_{k=2}^{\infty} (-1)^k \zeta''(k),$$

$$\mathrm{E}\left(\ln\left\lceil \frac{1}{X} \right\rceil^2\right) = \sum_{n=1}^{\infty} \frac{\ln(n+1)^2}{n(n+1)} = \sum_{k=2}^{\infty} \zeta''(k),$$

$$\mathrm{Var}\left(\ln\left\lfloor \frac{1}{X} \right\rfloor\right) = 1.1759638742... = (1.0844186803...)^2,$$

$$\operatorname{Var}\left(\ln\left\lceil\frac{1}{X}\right\rceil\right)=0.7543859444...=(0.8685539387...)^2,$$

$$\begin{aligned}\mathrm{E}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor\cdot\ln\left\lceil\frac{1}{X}\right\rceil\right) &= \sum_{n=1}^{\infty}\frac{\ln(n)\ln(n+1)}{n(n+1)}\\ &= \sum_{k=2}^{\infty}(-1)^k\zeta''(k)+\sum_{j=1}^{\infty}\frac{1}{j}\sum_{k=2}^{\infty}(-1)^{j+k}\zeta'(j+k),\end{aligned}$$

$$\operatorname{Var}\left(\ln\left\lfloor\frac{1}{X}\right\rfloor+\ln\left\lceil\frac{1}{X}\right\rceil\right)=3.8012096188...=(1.9496691049...)^2.$$

It can be proved whenever $i\neq j$ that digits a_i and a_j are independent random variables (unlike any of the continued fraction expansions we have examined), hence $\rho(\ln a_i,\ln a_j)=0$. As a consequence, two relevant Central Limit Theorems are easy to state: as $n\to\infty$, both of the distributions

$$\mathrm{P}\left(\frac{\left(\frac{1}{n}\sum_{i=1}^{n}\ln(a_i)\right)-0.7885305659...}{\frac{1.0844186803...}{\sqrt{n}}}\leq t\right),\quad \mathrm{P}\left(\frac{\left(\frac{1}{n}\sum_{i=1}^{n}\ln(a_i(a_i+1))\right)-2.0462774528...}{\frac{1.9496691049...}{\sqrt{n}}}\leq t\right)$$

tend to the standard normal. (For earlier expansions, the computation of σ was complicated by the existence of nonzero correlations.)

Here is an unexplained coincidence. Consider a random ordered (strongly) binary tree with N vertices, where N is odd. Janson [15, 16] proved that

$$\mathrm{E}\left(\frac{H}{\sqrt{N}}\cdot\frac{W}{\sqrt{N}}\right)\to 1+\sum_{n=1}^{\infty}\frac{\ln\left[n(n+1)\right]}{n(n+1)}=3.0462774528...$$

as $N\to\infty$ (which implies that the cross-correlation between height H and width W is asymptotically $-0.6428251027...$). The appearance of the same infinite series in two seemingly distant settings is fascinating! Why should the joint distribution of height and width of trees be at all related to the ergodic theory of numbers?

Since

$$\mathrm{P}(a_j=k)=\frac{1}{k(k+1)}=\int_{1/(k+1)}^{1/k}dx=\mathrm{P}\left(k<\frac{1}{X}<k+1\right)$$

where X is uniformly distributed, it follows that [2, 17, 18]

$$\mathrm{P}\left(\frac{1}{n}\sum_{j=1}^{n}a_j-(\ln(n)+1-\gamma)\leq t\right)\to\int_{-\infty}^{t}f(u)\,du$$

and f is the density function

$$f(u) = \frac{1}{\pi} \int_0^\infty \sin(\pi\, v) \exp(-v \ln(v) - u\, v)\, dv$$

of the limiting stable distribution $S(1, 1, \pi/2, 0; 1)$. Similarly precise characterizations of digit sums for NICF and ODCF remain open.

4.11.4 *Ordinary Decimal Representations*

At the risk of being anticlimatic, we define $T\colon [0,1] \to [0,1]$ by

$$T(x) = \{10\, x\} = 10\, x - \lfloor 10\, x \rfloor$$

and digits $a_1 = \lfloor 10\, x \rfloor$, $a_2 = \lfloor 10\, Tx \rfloor$, $a_3 = \lfloor 10\, T^2 x \rfloor$, ... For uniformly distributed X, it follows that

$$\mathrm{E}(\lfloor 10\, X \rfloor) = \frac{9}{2}, \quad \mathrm{Var}(\lfloor 10\, X \rfloor) = \frac{33}{4}$$

and, because a_i and a_j are independent random variables whenever $i \neq j$,

$$\mathrm{P}\left(\frac{\frac{1}{n} \sum_{j=1}^{n} a_j - \frac{9}{2}}{\frac{1}{2}\sqrt{\frac{33}{n}}} \le t \right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du.$$

We merely mention the Newcomb–Benford law [19–21], which is a different topic altogether (leading nonzero digit phenomenology) and yet seemingly related.

[1] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.
[2] S. R. Finch, Continued fraction transformation. II, *this volume*, §4.10.
[3] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[4] S. R. Finch, Khintchine-Lévy constants, *first volume*, pp. 59–65.
[5] G. J. Rieger, Über die mittlere Schrittanzahl bei Divisionsalgorithmen, *Math. Nachr.* 82 (1978) 157-180; MR0480366 (58 #533).
[6] A. M. Rockett, The metrical theory of continued fractions to the nearer integer, *Acta Arith.* 38 (1980/81) 97-103; MR0604225 (82d:10074).
[7] J. Bourdon, On the Khintchine constant for centred continued fraction expansions, *Appl. Math. E-Notes* 7 (2007) 167–174; MR2322809 (2008c:11101).
[8] V. Baladi and B. Vallée, Euclidean algorithms are Gaussian, *J. Number Theory* 110 (2005) 331–386; MR2122613 (2006e:11192).
[9] G. J. Rieger, Ein Heilbronn-Satz für Kettenbrüche mit ungeraden Teilnennern, *Math. Nachr.* 101 (1981) 295–307; MR0638347 (83c:10011).
[10] S. Kalpazidou, On a problem of Gauss-Kuzmin type for continued fraction with odd partial quotients, *Pacific J. Math*. 123 (1986) 103–114; MR0834141 (87k:11086).
[11] C. Ganatsiou, On the asymptotic behaviour of the digits of continued fractions with odd partial quotients, *Quaestiones Math.* 18 (1995) 517–526; MR1354127 (96k:11106).

[12] T. Salát, Zur metrischen Theorie der Lürothschen Entwicklungen der reellen Zahlen, *Czechoslovak Math. J.* 18 (1968) 489–522; MR0229605 (37 #5179).
[13] A. Dávid, Berechnung einer bestimmten Konstante aus der Theorie der Lürothschen Reihen, *Acta Fac. Rerum Natur. Univ. Comenian. Math.* 21 (1968) 15–20; MR0265298 (42 #208).
[14] H. Jager and C. de Vroedt, Lüroth series and their ergodic properties, *Proc. Konink. Nederl. Akad. Wetensch. Ser. A* 72 (1969) 31-42; *Indag. Math.* 31 (1969) 31-42; MR0238793 (39 #157).
[15] S. Janson, On the asymptotic joint distribution of height and width in random trees, *Studia Sci. Math. Hungar.* 45 (2008) 451–467; MR2641443 (2011d:60239).
[16] S. R. Finch, Shapes of binary trees, *this volume*, §4.6.
[17] D. Hensley, The statistics of the continued fraction digit sum, *Pacific J. Math.* 192 (2000) 103–120; MR1741027 (2000k:11092); Zbl. 1015.11038.
[18] J. P. Nolan, *Stable Distributions. Models for Heavy Tailed Data*, Birkhäuser, to appear.
[19] É. Janvresse and T. De la Rue, From uniform distributions to Benford's law, *J. Appl. Probab.* 41 (2004) 1203–1210; MR2122815 (2006b:60161).
[20] P. Schatte, On Benford's law for continued fractions, *Math. Nachr.* 148 (1990) 137–144; MR1127337 (92m:11077).
[21] S. R. Finch, Newcomb-Benford law, *this volume*, §4.22.

4.12 Continued Fraction Transformation. IV

Let $\lfloor x + iy \rfloor = \lfloor x \rfloor + i \lfloor y \rfloor$, where i is the imaginary unit. Extending the regular continued fraction algorithm [1] from the real interval $[0, 1]$ to the complex square $[0, 1] + i[0, 1]$ is problematic: the transformation

$$T(z) = \begin{cases} \dfrac{1}{z} - \left\lfloor \dfrac{1}{z} \right\rfloor & \text{if } z \neq 0, \\ 0 & \text{if } z = 0 \end{cases}$$

gives divergent continued fractions of the form

$$\frac{1|}{|-i} + \frac{1|}{|-i} + \frac{1|}{|-i} + \cdots$$

whenever

$$z = \frac{\sqrt{p}}{p-1} + i\frac{1}{2}$$

for any odd prime number p. This observation appears to be new. Nakada [2] noted divergence given any z satisfying both $|z| > 1$ and $|z - i| > 1$, for which $p = 3$ is a limiting case.

Extending the nearest integer continued fraction algorithm [3] to the complex square $[-1/2, 1/2] + i[-1/2, 1/2]$ at least makes sense! The transformation here is

$$T(z) = \begin{cases} \dfrac{1}{z} - \left\lfloor \dfrac{1}{z} + \dfrac{1}{2} \right\rfloor & \text{if } z \neq 0, \\ 0 & \text{if } z = 0 \end{cases}$$

and is called **Hurwitz's algorithm** [4, 5]. Consider the eight regions into which the four circular arcs $|z \pm 1| = 1$, $|z \pm i| = 1$ partition the square. The additional four circular arcs $|z \pm 1 \pm i| = 1$ subdivide four of the regions, making a total of twelve. Hensley [6–8] proved that the invariant density function for T is smooth on the interiors of the twelve regions and continuous everywhere except perhaps along the eight circular arcs. No closed-form expression for the density is known. For a complex random variable Z following this distribution, Monte Carlo simulation suggests that

$$\mathrm{E}(\ln(|Z|)) = 1.092766....$$

We shall not pursue this topic further, opting instead to discuss the most natural extension from $\mathbb{R}$ to $\mathbb{C}$ yet found of continued fraction theory.

4.12.1 *Schmidt's Complex Continued Fractions*

Define matrices

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Regular continued fractions can be thought of as infinite products of matrices; for example,

$$\pi = 3 + \frac{1|}{|7} + \frac{1|}{|15} + \frac{1|}{|1} + \frac{1|}{|292} + \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|2} + \frac{1|}{|1} + \frac{1|}{|3} + \cdots$$

is identified with

$$A^3BA^7BA^{15}BA^1BA^{292}BA^1BA^1BA^1BA^2BA^1BA^3B\cdots.$$

If the above product is multiplied on the right by

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

yielding

$$\begin{pmatrix} p^{(1)} & p^{(2)} & p^{(1)}+p^{(2)} \\ q^{(1)} & q^{(2)} & q^{(1)}+q^{(2)} \end{pmatrix},$$

then the ratios $p^{(1)}/q^{(1)}$, $p^{(2)}/q^{(2)}$ and $\left(p^{(1)}+p^{(2)}\right) / \left(q^{(1)}+q^{(2)}\right)$ each approach π as more terms are included in the product. For later convenience, let $p^{(3)} = p^{(1)} + p^{(2)}$ and $q^{(3)} = q^{(1)} + q^{(2)}$.

Define instead matrices [9–12]

$$V_1 = \begin{pmatrix} 1 & i \\ 0 & 1 \end{pmatrix}, \quad V_2 = \begin{pmatrix} 1 & 0 \\ -i & 1 \end{pmatrix}, \quad V_3 = \begin{pmatrix} 1-i & i \\ -i & 1+i \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & -1+i \\ 1-i & i \end{pmatrix},$$

$$E_1 = \begin{pmatrix} 1 & 0 \\ 1-i & i \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & -1+i \\ 0 & i \end{pmatrix}, \quad E_3 = \begin{pmatrix} i & 0 \\ 0 & 1 \end{pmatrix}.$$

With this enhanced "alphabet", the real number π can be represented by

$$E_2 V_1^2 V_3^7 V_1^{15} V_3^1 V_1^{292} V_3^1 V_1^1 V_3^1 V_1^2 V_3^1 V_1^3 \cdots$$

and the interpretation of convergence (ratios of first-row elements to second-row elements) is identical to before.

For the complex number e^i, the matrix representation can be proved to be [9, 11]

$$CCV_3^1 CCV_3^3 CCV_3^5 CCV_3^7 CCV_3^9 CCV_3^{11} CCV_3^{13} CCV_3^{15} CCV_3^{17} CCV_3^{19} \ldots$$

and for the number $\pi e^i/4$, it can be calculated to be

$$\begin{aligned}
&V_2^1 E_2 V_3^1 C V_1^2 E_3 V_2^1 C E_3 C E_1 C E_2 C E_1 V_3^6 C V_2^1 V_3^4 E_2 C E_2 V_1^2 C V_2^1 V_3^1 V_1^1 E_3 C \\
&E_1 V_2^2 C V_3^1 V_1^1 V_3^4 E_2 V_3^2 C E_1 V_2^3 C V_3^1 E_2 V_3^1 C V_2^1 E_1 V_2^2 C V_1^1 E_2 V_1^1 C V_2^6 E_3 V_2^{12} C \\
&V_3^1 V_1^1 V_2^1 V_1^1 E_3 C V_1^{80} E_3 V_1^{32} C V_2^1 V_1^1 E_3 C E_2 V_1^1 C E_1 V_2^3 C V_1^1 V_2^2 E_3 V_2^3 V_1^1 C V_1^1 V_2^1 \\
&E_3 V_2^1 V_1^2 V_2^2 C V_3^8 E_1 V_3^{19} V_2^5 V_3^1 C E_2 V_1^1 C V_3^6 E_1 V_3^6 C V_1^1 E_2 V_2^1 C V_3^2 E_2 V_3^3 C V_2^5 E_1 V_2^4 \\
&C V_2^1 E_1 V_2^2 C V_1^2 E_3 V_1^4 C V_2^1 V_1^3 E_3 C V_3^1 E_1 V_3^2 C E_2 V_2^1 V_1^1 V_2^1 V_3^1 C V_1^1 V_2^2 E_3 V_2^2 C V_3^1 \\
&E_2 C E_3 V_1^2 C V_2^2 E_3 C E_1 V_3^1 V_1^1 V_3^1 C V_2^8 E_3 C V_3^1 E_1 V_3^2 C V_2^3 V_1^3 E_3 V_1^3 V_3^1 C E_1 V_3^1 C \\
&V_1^1 E_2 C E_3 V_3^1 C E_1 V_3^2 V_2^3 V_3^1 C V_2^1 V_1^1 V_3^1 V_1^5 V_3^1 V_1^2 E_2 C E_1 V_2^2 C E_2 V_1^1 C V_3^2 V_1^1 E_3 V_1^1 \\
&C V_2^1 V_1^1 V_3^1 E_1 V_3^3 C E_3 C V_2^1 E_1 C E_3 C V_3^3 V_1^2 V_3^1 V_2^1 V_1^1 E_3 C V_2^1 E_1 V_1^1 V_2^1 V_3^1 C E_2 C E_1 \\
&V_3^1 C E_3 V_1^2 C E_2 C E_3 C E_1 C V_1^1 V_2^1 E_3 C E_1 V_2^1 C V_2^4 E_3 V_1^3 C V_3^2 E_2 V_1^4 C E_1 C V_1^2 V_3^2 \\
&V_1^1 V_3^1 E_2 V_3^5 C V_1^1 V_2^1 V_3^1 V_2^2 E_2 C V_3^1 E_1 V_3^1 C V_3^1 V_2^1 V_1^1 V_2^1 V_1^8 E_2 V_1^4 C V_2^2 E_3 C E_2 C V_1^1 \\
&E_2 V_1^2 C V_3^1 E_3 V_2^1 C V_1^1 V_3^1 V_1^2 V_3^4 E_2 V_3^3 C V_1^3 E_2 V_1^3 C V_1^1 E_2 V_3^1 V_2^1 V_1^1 C V_1^1 E_3 V_1^1 V_2^1 \\
&C V_1^2 V_2^1 V_1^1 E_3 V_1^3 C V_1^1 E_2 C V_3^1 E_2 V_1^2 C V_1^6 E_2 V_1^4 C C C V_1^2 E_3 V_1^2 C E_1 V_3^1 V_2^1 V_3^1 V_2^1 \ldots.
\end{aligned}$$

Note that powers of V_j are collected together, but not powers of C or E_j. The **terms** of the matrix representation are hence

$$T_1 = C, \quad T_2 = C, \quad T_3 = V_3^1, \quad T_4 = C, \quad T_5 = C, \quad T_6 = V_3^3, \; \ldots$$

for e^i and

$$T_1 = V_2^1, \quad T_2 = E_2, \quad T_3 = V_3^1, \quad T_4 = C, \quad T_5 = V_1^2, \quad T_6 = E_3, \; \ldots$$

for $\pi e^i/4$. This convention will be crucial later: the phrase "full terms" will sometimes be used for emphasis. We now give **Schmidt's algorithm** for generating such **chains** of matrices.

Let $\mathbb{C}$ and $\mathbb{C}^*$ denote two distinct complex planes. Define sets

$$F(I) = \{z \in \mathbb{C} : \operatorname{Im}(z) \geq 0\},$$

$$F^*(I) = \left\{ z \in \mathbb{C}^* : 0 \le \mathrm{Re}(z) \le 1,\ \mathrm{Im}(z) \ge 0,\ \left| z - \frac{1}{2} \right| \ge \frac{1}{2} \right\}$$

and subsets

$$F(V_1) = \{ z \in F(I) : \mathrm{Im}(z) \ge 1 \},$$

$$F(V_2) = \left\{ z \in F(I) : \left| z - \frac{i}{2} \right| \le \frac{1}{2} \right\},$$

$$F(V_3) = \left\{ z \in F(I) : \left| z - \left(1 + \frac{i}{2}\right) \right| \le \frac{1}{2} \right\},$$

$$F(C) = \left\{ \begin{array}{c} z \in F(I) : 0 < \mathrm{Re}(z) < 1,\ \frac{1}{2} < \mathrm{Im}(z) < 1, \\ \left| z - \frac{i}{2} \right| > \frac{1}{2},\ \left| z - \left(1 + \frac{i}{2}\right) \right| > \frac{1}{2} \end{array} \right\},$$

$$F(E_1) = \left\{ \begin{array}{c} z \in F(I) : 0 \le \mathrm{Re}(z) < 1,\ 0 \le \mathrm{Im}(z) < \frac{1}{2}, \\ \left| z - \frac{i}{2} \right| > \frac{1}{2},\ \left| z - \left(1 + \frac{i}{2}\right) \right| > \frac{1}{2} \end{array} \right\},$$

$$F(E_2) = \left\{ z \in F(I) : \mathrm{Re}(z) > 1,\ 0 \le \mathrm{Im}(z) < 1,\ \left| z - \left(1 + \frac{i}{2}\right) \right| > \frac{1}{2} \right\},$$

$$F(E_3) = \left\{ z \in F(I) : \mathrm{Re}(z) < 0,\ 0 \le \mathrm{Im}(z) < 1,\ \left| z - \frac{i}{2} \right| > \frac{1}{2} \right\},$$

$$F^*(V_1) = \left\{ z \in F^*(I) : 0 \le \mathrm{Re}(z) \le 1,\ \mathrm{Im}(z) > 1,\ \left| z - \left(\frac{1}{2} + i\right) \right| > \frac{1}{2} \right\},$$

$$F^*(V_2) = \left\{ \begin{array}{c} z \in F^*(I) : 0 \le \mathrm{Re}(z) < \frac{1}{2},\ 0 \le \mathrm{Im}(z) \le 1, \\ \left| z - \frac{1}{2} \right| \ge \frac{1}{2},\ \left| z - \left(\frac{1}{2} + i\right) \right| > \frac{1}{2} \end{array} \right\},$$

$$F^*(V_3) = \left\{ \begin{array}{c} z \in F^*(I) : \frac{1}{2} < \mathrm{Re}(z) \le 1,\ 0 \le \mathrm{Im}(z) \le 1, \\ \left| z - \frac{1}{2} \right| \ge \frac{1}{2},\ \left| z - \left(\frac{1}{2} + i\right) \right| > \frac{1}{2} \end{array} \right\},$$

$$F^*(C) = \left\{ z \in F^*(I) : \left| z - \left(\frac{1}{2} + i\right) \right| \le \frac{1}{2} \right\}.$$

The letter F suggests "Farey set" and $F(C)$, for instance, is the image of the interior of $F^*(I)$ under the action of C, where Cz is the value of the linear fractional function

$$Cz = \begin{pmatrix} 1 & -1+i \\ 1-i & i \end{pmatrix} z = \frac{z + (-1+i)}{(1-i)z + i}, \quad z \in F^*(I).$$

Note that each of the seven matrices is invertible and, for instance,

$$C^{-1}z = \begin{pmatrix} -1 & 1+i \\ -1-i & i \end{pmatrix} z = \frac{-z + (1+i)}{(-1-i)z + i}.$$

Schmidt's transformation T maps the disjoint union $F(I) \cup F^*(I)$ into $F(I) \cup F^*(I)$ via the following formula:

$$T(z,\varepsilon) = \begin{cases} \left(V_j^{-1}z, \varepsilon\right) & \text{if } (z \in F(V_j) \wedge \varepsilon = 1) \vee (z \in F^*(V_j) \wedge \varepsilon = 0), \\ \left(E_j^{-1}z, 1-\varepsilon\right) & \text{if } z \in F(E_j) \wedge \varepsilon = 1, \\ \left(C^{-1}z, 1-\varepsilon\right) & \text{if } (z \in F(C) \wedge \varepsilon = 1) \vee (z \in F^*(C) \wedge \varepsilon = 0), \end{cases}$$

$$= \begin{cases} (z-i, \varepsilon) & \text{if } (z \in F(V_1) \wedge \varepsilon = 1) \vee (z \in F^*(V_1) \wedge \varepsilon = 0), \\ \left(\dfrac{z}{iz+1}, \varepsilon\right) & \text{if } (z \in F(V_2) \wedge \varepsilon = 1) \vee (z \in F^*(V_2) \wedge \varepsilon = 0), \\ \left(\dfrac{(1+i)z-i}{iz+(1-i)}, \varepsilon\right) & \text{if } (z \in F(V_3) \wedge \varepsilon = 1) \vee (z \in F^*(V_3) \wedge \varepsilon = 0), \\ \left(\dfrac{z}{(1+i)z-i}, 1-\varepsilon\right) & \text{if } z \in F(E_1) \wedge \varepsilon = 1, \\ \left(\dfrac{z-(1+i)}{-i}, 1-\varepsilon\right) & \text{if } z \in F(E_2) \wedge \varepsilon = 1, \\ (-iz, 1-\varepsilon) & \text{if } z \in F(E_3) \wedge \varepsilon = 1, \\ \left(\dfrac{-z+(1+i)}{(-1-i)z+i}, 1-\varepsilon\right) & \text{if } (z \in F(C) \wedge \varepsilon = 1) \vee (z \in F^*(C) \wedge \varepsilon = 0) \end{cases}$$

where $j = 1, 2, 3$ and $\varepsilon = 0, 1$. The chains for π, e^i and $\pi e^i/4$ were obtained by iterating T with starting value $\varepsilon = 1$, meaning that π, e^i and $\pi e^i/4$ are thought of as residing in $F(I)$. Clearly $\pi \notin F^*(I)$ and $e^i \notin F^*(I)$, but $\pi e^i/4$ can thought of as residing in $F^*(I)$ as well. Starting with $\varepsilon = 0$ instead, the **dual chain** for $\pi e^i/4$ is

$$\begin{gathered} C V_1^2 E_2 V_3^1 V_1^1 C E_1 C V_3^1 E_1 C V_3^1 E_2 V_3^1 C V_2^5 E_1 V_1^1 C E_2 V_1^4 C V_3^2 E_2 V_2^2 C V_2^1 E_3 \\ V_2^1 C V_3^2 E_1 V_2^1 V_3^1 C V_1^1 E_2 V_1^3 C V_3^4 E_1 V_3^1 C E_2 C V_3^3 E_1 V_3^1 C E_3 V_1^1 C V_2^{12} E_1 V_2^7 \\ V_1^1 C E_1 V_3^1 V_2^1 V_3^1 C V_3^{32} E_1 V_3^{80} C E_1 V_2^1 C E_3 C V_3^1 V_1^3 E_2 C E_1 V_3^1 C V_1^2 E_3 V_1^1 C E_1 \ldots . \end{gathered}$$

Dual chains will not be mentioned again, since the ergodic results for chains we seek are the same as ergodic results for dual chains. The associated geometry of Schmidt's algorithm is well-illustrated in [6, 13].

4.12.2 *Invariant Density*

Let $h : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$h(x,y) = \frac{1}{x\,y} - \frac{1}{x^2} \arctan\left(\frac{x}{y}\right)$$

and $\tilde{f}: F(I) \cup F^*(I) \to F(I) \cup F^*(I)$ be given by

$$\tilde{f}(z) = \begin{cases} \frac{1}{2\pi^2}\left(h(x,y) + h(1-x,y) + h(x^2 - x + y^2, y)\right) & \text{if } z = x + iy \in F(I), \\ \frac{1}{2\pi}\frac{1}{y^2} & \text{if } z = x + iy \in F^*(I). \end{cases}$$

The probability density function $\tilde{f}$ is continuous everywhere except at the points $0, 1 \in F(I)$ and $0, 1 \in F^*(I)$.

Define a constant

$$\kappa = \frac{24}{\sqrt{15}} \arccos\left(\frac{1}{4}\right) - 2\pi$$

and the Jacobian determinant

$$\| V_j z \| = \left| \frac{d}{dz}(V_j z) \right|^2$$

for each $j = 1, 2, 3$. For example,

$$V_2 z = \frac{z}{-iz + 1} = \frac{x}{x^2 + (y+1)^2} + i\frac{x^2 + y(y+1)}{x^2 + (y+1)^2},$$

$$V_3 z = \frac{(1-i)z + i}{-iz + (1+i)} = \frac{x(x-1) + (y+1)^2}{(x-1)^2 + (y+1)^2} + i\frac{(x-1)^2 + y(y+1)}{(x-1)^2 + (y+1)^2}$$

and

$$\| V_2 z \| = \frac{1}{|-iz + 1|^4} = \frac{1}{(x^2 + (y+1)^2)^2},$$

$$\| V_3 z \| = \frac{1}{|-iz + (1+i)|^4} = \frac{1}{((x-1)^2 + (y+1)^2)^2}.$$

The invariant probability density function f is given by

$$f(z) = \begin{cases} \frac{\pi}{\kappa}\tilde{f}(z) & \text{if } z \in F(E_1) \cup F(E_2) \cup F(E_3) \cup F(C) \cup F^*(C), \\ \frac{\pi}{\kappa}\left(\tilde{f}(z) - \tilde{f}(V_j z)\, \| V_j z \|\right) & \text{if } z \in F(V_j) \cup F^*(V_j),\ 1 \le j \le 3 \end{cases}$$

where, as always, a union involving F and F^* is a disjoint one. Consequences of this remarkable explicit formula follow in the next two sections. Note, for example,

$$f(z) = \frac{\pi}{\kappa}\left(\frac{1}{y^2} - \frac{1}{(x^2 + y(y+1))^2}\right)$$

for $z \in F^*(V_2)$ and

$$f(z) = \frac{\pi}{\kappa}\left(\frac{1}{y^2} - \frac{1}{((x-1)^2 + y(y+1))^2}\right)$$

for $z \in F^*(V_3)$. Over and beyond the singularities at points $0, 1 \in F(I)$ and $0, 1 \in F^*(I)$, there are jump discontinuities at the boundaries of $F(V_j)$ and $F^*(C)$.

4.12.3 Analog of Khintchine's Constant

For each term T_n in the matrix representation of z, define the corresponding continued fraction "digit"

$$\alpha_n(z) = \begin{cases} m & \text{if } T_n = V_j^m \text{ for some } 1 \le j \le 3, \\ 1 & \text{otherwise.} \end{cases}$$

In the case $z = \pi e^i/4$, we have $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 1$, $\alpha_5 = 2$, $\alpha_6 = 1$ and $\alpha_{16} = 6$, $\alpha_{19} = 4$. Define also

$$\Psi(x) = \pi - \frac{2}{\sqrt{1-x^2}} \arccos(x).$$

It can be shown that

$$F(V_1^m) = \{z \in F(I) : \operatorname{Im}(z) \ge m\},$$

$$F(V_2^m) = \left\{ z \in F(I) : \left| z - \frac{i}{2m} \right| \le \frac{1}{2m} \right\},$$

$$F(V_3^m) = \left\{ z \in F(I) : \left| z - \left(1 + \frac{i}{2m}\right) \right| \le \frac{1}{2m} \right\},$$

$$F^*(V_1^m) = \left\{ z \in F^*(I) : 0 \le \operatorname{Re}(z) \le 1,\ \operatorname{Im}(z) > m,\ \left| z - \left(\frac{1}{2} + m\,i\right) \right| > \frac{1}{2} \right\},$$

$$F^*(V_2^m) = \left\{ \begin{array}{c} z \in F^*(I) : 0 \le \operatorname{Re}(z) < \dfrac{1}{m^2+1},\ 0 \le \operatorname{Im}(z) \le \dfrac{1}{m}, \\ \left| z - \dfrac{1}{2} \right| \ge \dfrac{1}{2},\ \left| z - \left(\dfrac{1}{2m^2} + \dfrac{i}{m} \right) \right| > \dfrac{1}{2m^2} \end{array} \right\},$$

$$F^*(V_3^m) = \left\{ \begin{array}{c} z \in F^*(I) : \dfrac{m^2}{m^2+1} < \operatorname{Re}(z) \le 1,\ 0 \le \operatorname{Im}(z) \le \dfrac{1}{m}, \\ \left| z - \dfrac{1}{2} \right| \ge \dfrac{1}{2},\ \left| z - \left(\dfrac{2m^2-1}{2m^2} + \dfrac{i}{m} \right) \right| > \dfrac{1}{2m^2} \end{array} \right\}$$

and hence

$$\int\limits_{F(V_j^m)} f(z)\,dz = \frac{1}{2\kappa} \left(\Psi\left(\frac{1}{2m}\right) - \Psi\left(\frac{1}{2(m+1)}\right) \right) = \int\limits_{F^*(V_j^m)} f(z)\,dz$$

for each $1 \leq j \leq 3$ and all $m \geq 1$. By ergodicity, the sum $(1/N)\sum_{n\leq N}\ln(\alpha_n(z))$ tends almost certainly as $N \to \infty$ to

$$\begin{aligned}
\int\limits_{F(I)\cup F^*(I)} \ln(\alpha_1(z))f(z)\,dz &= 2\sum_{j=1}^{3}\sum_{m=1}^{\infty}\int\limits_{F(V_j^m)-F(V_j^{m+1})} \ln(m)f(z)\,dz \\
&= \frac{3}{\kappa}\sum_{m=2}^{\infty}\ln(m)\left(\Psi\left(\frac{1}{2m}\right) - 2\Psi\left(\frac{1}{2(m+1)}\right) + \Psi\left(\frac{1}{2(m+2)}\right)\right) \\
&= \frac{3}{\kappa}\left(\ln(2)\Psi\left(\frac{1}{4}\right) + \sum_{m=3}^{\infty}\ln\left(1 - \frac{1}{(m-1)^2}\right)\Psi\left(\frac{1}{2m}\right)\right) \\
&= \ln(1.2617651749...) = 0.2325116730....
\end{aligned}$$

which is the Schmidt analog of Khintchine's constant [10].

In the real case, the almost-certain divergence of $(1/N)\sum_{n\leq N}\alpha_n(z)$ is well-known. It is interesting that in the complex case, the mean converges to

$$\begin{aligned}
&2\sum_{j=1}^{3}\left(\sum_{m=1}^{\infty}\int\limits_{F(V_j^m)-F(V_j^{m+1})} m f(z)\,dz\right) + \sum_{j=1}^{3}\int\limits_{F(E_j)} f(z)\,dz + \int\limits_{F(C)} f(z)\,dz + \int\limits_{F^*(C)} f(z)\,dz \\
&\quad = \frac{3}{\kappa}\Psi\left(\frac{1}{2}\right) + \left(\sum_{j=1}^{3}\int\limits_{F(E_j)} f(z)\,dz + \int\limits_{F(C)} f(z)\,dz\right) + \frac{\pi}{\kappa}\left(\frac{2}{\sqrt{3}} - 1\right) \\
&\quad = \frac{3\pi}{\kappa}\left(1 - \frac{4}{3\sqrt{3}}\right) + \frac{\pi}{\kappa}\left(\frac{2}{\sqrt{3}} - 1\right) + \frac{\pi}{\kappa}\left(\frac{2}{\sqrt{3}} - 1\right) = \frac{\pi}{\kappa} = 1.6667324083....
\end{aligned}$$

The variance, however, is divergent.

4.12.4 *Analog of Lévy's Constant*

We wish to compute the almost-certain limit of $(1/n)\ln\left|q_n^{(\ell)}\right|$ as $n \to \infty$, corresponding to denominators $q_n^{(\ell)}$ in the partial convergents to z. The limit turns out to be independent of $1 \leq \ell \leq 3$. There are two variations:

- the **powerless** scenario, in which $q_n^{(\ell)}$ is evaluated at each iteration of Schmidt's algorithm (powers of V_j are irrelevant)
- the **powerful** scenario, in which $q_n^{(\ell)}$ is evaluated only at iterations that "close" a term T_k (only those powers of V_j constituting full terms are relevant, as well as any terms E_j and C).

The first gives a simpler result, but the second is more consistent with the real case. As an example, look at

$$V_2^1 E_2 V_3^1 C V_1 = \begin{pmatrix} 1+3i & -6+i \\ 4+i & -3+7i \end{pmatrix}, \quad V_2^1 E_2 V_3^1 C V_1^2 = \begin{pmatrix} 1+3i & -9+2i \\ 4+i & -4+11i \end{pmatrix}$$

from the matrix representation of $\pi e^i/4$. In the powerless way of counting, the ratio $p_5^{(2)}/q_5^{(2)}$ is $(-6+i)/(-3+7i)$ and $p_6^{(2)}/q_6^{(2)}$ is $(-9+2i)/(-4+11i)$. In the powerful way of counting, the ratio $p_5^{(2)}/q_5^{(2)}$ is $(-9+2i)/(-4+11i)$. Both variations are interesting to us.

For the powerless scenario, let

$$\tilde{T}_1(z) = \begin{pmatrix} \tilde{a}_1(z) & \tilde{b}_1(z) \\ \tilde{c}_1(z) & \tilde{d}_1(z) \end{pmatrix}$$

be the initial output of Schmidt's algorithm, starting with input z, and let

$$\begin{aligned} \tilde{\varphi}(z) &= -\ln|\tilde{c}_1(z)z - \tilde{a}_1(z)| \\ &= \begin{cases} 0 & \text{if } z \in F(V_1) \cup F(E_2) \cup F(E_3) \cup F^*(V_1), \\ -\frac{1}{2}\ln\left(2\left((x-\frac{1}{2})^2 + (y-\frac{1}{2})^2\right)\right) & \text{if } z \in F(E_1) \cup F(C) \cup F^*(C), \\ -\frac{1}{2}\ln\left(x^2 + (y-1)^2\right) & \text{if } z \in F(V_2) \cup F^*(V_2), \\ -\frac{1}{2}\ln\left((x-1)^2 + (y-1)^2\right) & \text{if } z \in F(V_3) \cup F^*(V_3). \end{cases} \end{aligned}$$

Then $(1/n)\ln\left|q_n^{(\ell)}\right|$ converges to [10]

$$\int\limits_{F(I)\cup F^*(I)} \tilde{\varphi}(z)\tilde{f}(z)\,dz = 0.29156...$$

via numerical calculation of each component of the integral. Closed-form expressions for the components appear to be impossible. Nakada [14–16], however, proved by a different approach that the powerless Schmidt analog of Lévy's constant is

$$\frac{G}{\pi} = 0.2915609040... = \ln(1.3385151519...)$$

where G is Catalan's constant [17, 18].

For the powerful scenario, let

$$T_1(z) = \begin{pmatrix} a_1(z) & b_1(z) \\ c_1(z) & d_1(z) \end{pmatrix}$$

be the initial full term in the complex continued fraction expansion of z, and let

$$\varphi(z) = -\ln|c_1(z)z - a_1(z)|$$
$$= \begin{cases} 0 & \text{if } z \in F(V_1) \cup F(E_2) \cup F(E_3) \cup F^*(V_1), \\ -\frac{1}{2}\ln\left(2\left((x-\frac{1}{2})^2 + (y-\frac{1}{2})^2\right)\right) & \text{if } z \in F(E_1) \cup F(C) \cup F^*(C), \\ -\frac{1}{2}\ln\left(m^2x^2 + (my-1)^2\right) & \text{if } z \in \left(F(V_2^m) - F(V_2^{m+1})\right) \cup \\ & \left(F^*(V_2^m) - F^*(V_2^{m+1})\right), \\ -\frac{1}{2}\ln\left(m^2(x-1)^2 + (my-1)^2\right) & \text{if } z \in \left(F(V_3^m) - F(V_3^{m+1})\right) \cup \\ & \left(F^*(V_3^m) - F^*(V_3^{m+1})\right). \end{cases}$$

Then $(1/n)\ln\left|q_n^{(\ell)}\right|$ converges to [10]

$$\int\limits_{F(I)\cup F^*(I)} \varphi(z)f(z)\,dz = 0.4859...$$

via numerical calculation of each component of the integral and summation over $m \geq 1$. Closed-form expressions for the components again appear to be impossible. Nakada [15] proved, as a corollary of his aforementioned result, that the powerful Schmidt analog of Lévy's constant is

$$\frac{G}{\kappa} = 0.4859540077... = \ln(1.6257252237...).$$

These are magnificent formulas, needless to say!

Complex continued fractions built upon the Eisenstein–Jacobi integers (rather than the Gaussian integers) were introduced in [19], but no comparable ergodic theory has been published, as far as is known.

We merely mention the **Jacobi-Perron algorithm** [20–24]

$$T_{\text{JPA}}(x,y) = \left(\frac{y}{x} - \left\lfloor\frac{y}{x}\right\rfloor, \frac{1}{x} - \left\lfloor\frac{1}{x}\right\rfloor\right)$$

and the **Podsypanin algorithm** [25–28]

$$T_{\text{MJPA}}(x,y) = \begin{cases} \left(\frac{y}{x}, \frac{1}{x} - \left\lfloor\frac{1}{x}\right\rfloor\right) & \text{if } x \geq y \wedge x \neq 0, \\ \left(\frac{1}{y} - \left\lfloor\frac{1}{y}\right\rfloor, \frac{x}{y}\right) & \text{if } x < y \wedge y \neq 0, \\ 0 & \text{if } x = y = 0 \end{cases}$$

for $(x,y) \in [0,1] \times [0,1]$. Both possess unique invariant densities but only the latter has a closed-form expression:

$$f_{\text{MJPA}}(x,y) = \frac{1}{2c}\frac{2+x+y}{(1+x)(1+y)(1+x+y)}$$

where

$$c = \frac{\pi^2}{12} + \mathrm{Li}_2\left(-\frac{1}{2}\right) = 0.3740528265....$$

If q_n denotes the common denominator in the n^{th} partial convergent to (x, y), then

$$\lim_{n\to\infty} \frac{1}{n} \ln(q_n) = -\int_0^1 \int_0^1 \ln\left(\max\{x, y\}\right) f_{\mathrm{MJPA}}(x, y)\, dx\, dy = 0.6695004121...$$

almost certainly (we omit the complicated exact formula involving dilogarithms and $\zeta(3)$). A precise estimate of the entropy associated with T_{JPA} would be good to see someday.

Acknowledgment I thank Asmus Schmidt for sending me 1998 reprints of his papers and helpful handwritten notes.

[1] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.
[2] H. Nakada, On metric Diophantine approximation of complex numbers, complex continued fractions, *Séminaire de Théorie des Nombres, Talence, 1987–1988,* exp. 45, Univ. Bordeaux I, Talence, 1988; MR0993138 (90f:11067).
[3] S. R. Finch, Continued fraction transformation. III, *this volume*, §4.11.
[4] A. Hurwitz, Über die Entwicklung complexer Grössen in Kettenbrüche, *Acta Math.* 11 (1888) 187-200; also in *Mathematische Werke*, v. II, *Zahlentheorie, Algebra und Geometrie*, Birkhäuser Verlag, 1963, pp. 84–115; MR0154777 (27 #4723a).
[5] H. Nakada, On the Kuzmin's theorem for the complex continued fractions, *Keio Engineering. Rep.* 29 (1976) 93–108; MR0463401 (57 #3353).
[6] D. Hensley, *Continued Fractions*, World Scientific, 2006, pp. 67–98; MR2351741 (2009a:11019).
[7] D. Hensley, The Hurwitz continued fraction algorithm, unpublished note (2006).
[8] D. Hensley, The Hurwitz algorithm's 'Gauss-Kuzmin density', Mathematica program HurwitzCrossFinal.nb.
[9] A. L. Schmidt, Diophantine approximation of complex numbers, *Acta Math.* 134 (1975) 1–85; MR0422168 (54 #10160).
[10] A. L. Schmidt, Ergodic theory for complex continued fractions, *Monatsh. Math.* 93 (1982) 39–62; MR0648739 (83g:10036).
[11] A. L. Schmidt, Diophantine approximation of complex numbers, *Classical and Quantum Models and Arithmetic Problems*, ed. D. V. Chudnovsky and G. V. Chudnovsky, Dekker, 1984, pp. 353–377; MR0756249 (85i:11058).
[12] A. L. Schmidt, Ergodic theory of complex continued fractions, *Number Theory with an Emphasis on the Markoff Spectrum*, Proc. 1991 Provo conf., ed. A. D. Pollington and W. Moran, Dekker, 1993, pp. 215–226; MR1219336 (95f:11055).
[13] D. Hensley, Asmus Schmidt's complex continued fraction algorithm, Mathematica program SchmidtCircles.nb.
[14] H. Nakada, On ergodic theory of A. Schmidt's complex continued fractions over Gaussian field, *Monatsh. Math.* 105 (1988) 131–150; MR0930432 (89f:11113).
[15] H. Nakada, The metrical theory of complex continued fractions, *Acta Arith.* 56 (1990) 279–289; MR1096343 (92e:11081).

[16] H. Nakada, Dynamics of complex continued fractions and geodesics over H^3, *Dynamical Systems and Chaos*, v. 1, Proc. 1994 Tokyo conf., ed. N. Aoki, K. Shiraiwa, and Y. Takahashi, World Scientific, 1995, pp. 192–199; MR1479930 (99c:11103).
[17] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[18] S. R. Finch, Khintchine-Lévy constants, *first volume*, pp. 59–65.
[19] A. L. Schmidt, Diophantine approximation in the Eisensteinian field, *J. Number Theory* 16 (1983) 169–204; MR0698164 (84m:10024).
[20] W. A. Beyer and M. S. Waterman, Ergodic computations with continued fractions and Jacobi's algorithm, *Numer. Math.* 19 (1972) 195–205; MR0303696 (46 #2832).
[21] F. Schweiger, *The Metrical Theory of Jacobi-Perron Algorithm*, Lect. Notes in Math. 334, Springer-Verlag, 1973; MR0345925 (49 #10654).
[22] P. R. Baldwin, A convergence exponent for multidimensional continued-fraction algorithms, *J. Stat. Phys.* 66 (1992) 1507–1526; MR1156412 (93d:11083).
[23] V. Baladi and A. Nogueira, Lyapunov exponents for non-classical multidimensional continued fraction algorithms, *Nonlinearity* 9 (1996) 1529–1546; MR1419459 (98e:11085).
[24] R. Murray, Existence, mixing and approximation of invariant densities for expanding maps on $\mathbb{R}^r$, *Nonlinear Anal.* 45 (2001) *Ser. A: Theory Methods*, 37–72; MR1828066 (2003b:37038).
[25] F. Schweiger, A modified Jacobi-Perron algorithm with explicitly given invariant measure, *Ergodic Theory*, Proc. 1978 Oberwolfach conf., ed. M. Denker and K. Jacobs, Lect. Notes in Math. 729, Springer-Verlag, 1979, pp. 199–202; MR0550422 (80i:28035).
[26] S. Ito, M. Keane and M. Ohtsuki, Almost everywhere exponential convergence of the modified Jacobi-Perron algorithm, *Ergodic Theory Dynam. Systems* 13 (1993) 319–334; MR1235475 (94i:11058).
[27] T. Fujita, S. Ito, M. Keane and M. Ohtsuki, On almost everywhere exponential convergence of the modified Jacobi-Perron algorithm: a corrected proof, *Ergodic Theory Dynam. Systems* 16 (1996) 1345–1352; MR1424403 (98a:11105).
[28] R. Meester, A simple proof of the exponential convergence of the modified Jacobi-Perron algorithm, *Ergodic Theory Dynam. Systems* 19 (1999) 1077–1083; MR1709431 (2000f:37006).

4.13 Lyapunov Exponents

We are interested in iterates of the **logistic map** $T : [0, 1] \to [0, 1]$ defined by

$$T(x) = a\,x\,(1 - x)$$

where $0 \le a \le 4$ is constant. Actually, only the values $a = 4$ and

$$a = \frac{2}{3}\left(\left(19 + 3\sqrt{33}\right)^{1/3} + 4\left(19 + 3\sqrt{33}\right)^{-1/3} + 1\right) = 3.6785735104...$$

will be examined (the latter has minimal polynomial $a^3 - 2a^2 - 4a - 8$). Both correspond to chaotic maps for which invariant probability densities $f(x)$ provably exist. An important feature of chaos is sensitivity to initial conditions. The **Lyapunov exponent** for each map quantifies the exponential rate at which two initially close points x, y separate [1]:

$$|T(x) - T(y)| \approx |T'(x)| \cdot |x - y|$$

after the first iteration,

$$|T^n(x) - T^n(y)| \approx \prod_{0 \le j < n} \left|T'(T^j x)\right| \cdot |x - y|$$

after the n^{th} iteration, and hence

$$\frac{1}{n} \ln |T^n(x) - T^n(y)| \approx \frac{1}{n} \sum_{0 \le j < n} \ln \left|T'(T^j x)\right|.$$

For X distributed according to f, let us write

$$\hat{\mu}_n(X) = \frac{1}{n} \sum_{0 \le j < n} T^j X, \quad \hat{\lambda}_n(X) = \frac{1}{n} \sum_{0 \le j < n} \ln \left|T'(T^j X)\right|$$

which converge as $n \to \infty$ almost surely, by ergodicity, to

$$\mathrm{E}(X) = \int_0^1 x f(x)\, dx, \quad \mathrm{E}\, |\ln(T' X)| = \int_0^1 \ln |T'(x)| f(x)\, dx.$$

Our study will encompass not only means, but also variances and autocovariances of arbitrary time lag. A complete solution is possible for $a = 4$; only partial results exist for $a = 3.678\ldots$. The approach we take is similar to [2].

4.13.1 *Ulam–von Neumann Map*

When $a = 4$, the invariant density has a closed-form expression [3]:

$$f(x) = \frac{1}{\pi \sqrt{x(1-x)}}$$

and thus

$$\mathrm{E}(T^j X) = \frac{1}{2}, \quad \mathrm{Var}(T^j X) = \frac{1}{8}, \quad \mathrm{Cov}(T^j X, T^k X) = 0$$

for all $j < k$. Also [4],

$$\mathrm{E} \left|\ln(T'(T^j X))\right| = \ln(2), \quad \mathrm{Var} \left|\ln(T'(T^j X))\right| = \frac{\pi^2}{12},$$

$$\mathrm{Cov} \left(\left|\ln(T'(T^j X))\right|, \left|\ln(T'(T^k X))\right|\right) = -\frac{\pi^2}{24} \frac{1}{2^{k-j}}$$

for all $j < k$. Clearly

$$\lim_{n \to \infty} \mathrm{E}\left(\hat{\mu}_n(X)\right) = \frac{1}{2},$$

$$\lim_{n\to\infty} n\,\mathrm{Var}\left(\hat{\mu}_n(X)\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{\substack{0\le j<n,\\ 0\le k<n}} \mathrm{Cov}(T^j X, T^k X) = \frac{1}{8}$$

and the Central Limit Theorem holds:

$$\lim_{n\to\infty} \mathrm{P}\left(2\sqrt{2n}\left(\hat{\mu}_n(X) - \frac{1}{2}\right) \le t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du.$$

By contrast,

$$\lim_{n\to\infty} \mathrm{E}\left(\hat{\lambda}_n(X)\right) = \ln(2),$$

$$\begin{aligned}
\lim_{n\to\infty} n^2\,\mathrm{Var}\left(\hat{\lambda}_n(X)\right) &= \lim_{n\to\infty} \sum_{\substack{0\le j<n,\\ 0\le k<n}} \mathrm{Cov}\left(\left|\ln(T'(T^j X))\right|, \left|\ln(T'(T^k X))\right|\right) \\
&= \lim_{n\to\infty} \frac{\pi^2}{6}\left(1 - \frac{1}{2^n}\right) = \frac{\pi^2}{6}.
\end{aligned}$$

Estimates $\hat{\lambda}_n(X)$ of the Lyapunov exponent are anomalously precise [5]: they possess a standard deviation that scales as $1/n$ rather than $1/\sqrt{n}$. In this case, evidence points to a revised Central Limit Theorem of the form [4, 6]:

$$\lim_{n\to\infty} \mathrm{P}\left(n\left(\hat{\lambda}_n(X) - \ln(2)\right) \le t\right) = \frac{2}{\pi^2} \int_{-\infty}^{t} \ln\left(\coth\left(\frac{u}{2}\right)\right) du$$

but a rigorous proof seems to be open.

4.13.2 *Ruelle-Misiurewicz Map*

When $a = 3.678...$, no closed-form expression for the invariant density is known, even though its existence is certain [7–9]. A numerical approach is necessary. Let $y = \frac{1}{2}a - a\,x$, then under the change of variables, T becomes

$$S(y) = y^2 - c$$

where

$$c = \frac{1}{4}a^2 - \frac{1}{2}a = 1.5436890126...$$

(with minimal polynomial $c^3 - 2c^2 + 2c - 2$). Now let [10]

$$\theta = \frac{1}{\pi} \arccos\left(\frac{y}{c - c^2}\right);$$

under this second change of variables, S^2 becomes

$$\tau(\theta) = \frac{1}{\pi} \arccos\left(\cos(2\pi\theta) + \kappa \sin(2\pi\theta)^2\right)$$

where

$$\kappa = \frac{c(c^2 - c) - 1}{2} = 0.1477988712....$$

The invariant density φ associated with $\tau : [0, 1] \to [0, 1]$ satisfies the functional equation [10]

$$\frac{\varphi(\tau^{-1}(\theta))}{|\tau'(\tau^{-1}(\theta))|} + \frac{\varphi(1 - \tau^{-1}(\theta))}{|\tau'(1 - \tau^{-1}(\theta))|} = \varphi(\theta)$$

where

$$\tau^{-1}(\theta) = \frac{1}{2\pi} \arccos\left(\frac{1 - \sqrt{1 + 4\kappa^2 - 4\kappa\cos(\pi\theta)}}{2\kappa}\right),$$

$$\tau'(\theta) = \frac{2\sin(2\pi\theta)\,(1 - 2\kappa\cos(2\pi\theta))}{\sqrt{\sin(2\pi\theta)^2\,(1 - 2\kappa\cos(2\pi\theta) - \kappa^2\sin(2\pi\theta)^2)}}.$$

The left-hand side of this equation is a special case of the Frobenius–Perron operator $P_\tau\varphi(\theta)$. Starting with an initial guess $\varphi_0 \equiv 1$, the uniform density, iterates $\varphi_{n+1} = P_\tau\varphi_n$ converge to a limiting density φ. Backtracking through the two coordinate transformations, we obtain the desired invariant density f. It turns out to be supported on the intervals $[\frac{1}{a}, 1 - \frac{1}{a}]$ and $[1 - \frac{1}{a}, \frac{a}{4}]$, which are exchanged by T, with three vertical asymptotes.

Recall that $x = \frac{1}{2} - \frac{1}{a}y$ and $y = (c - c^2)\cos(\pi\theta)$. For X distributed according to f, we compute

$$\mathrm{E}(X) = \frac{1}{2}\int_0^1 (x + T(x))\,\varphi(\theta)d\theta = 0.6717404535...,$$

$$\mathrm{E}\,|\ln(T'X)| = \frac{1}{2}\int_0^1 (\ln|2y| + \ln|2S(y)|)\,\varphi(\theta)d\theta = 0.3421726886....$$

No one evidently has computed higher-order moments of X and $\ln(T'X)$, let alone $\hat{\mu}_n(X)$ and $\hat{\lambda}_n(X)$. Does the Central Limit Theorem need revision here too?

The value 3.678... is the simplest *Misiurewicz point*. For any $a \leq 3.678...$, the logistic map T admits no periodic point x of odd order > 1, i.e., it has no odd cycles. For any $a > 3.678...$, T has odd cycles [11, 12].

A graph of $\mathrm{E}(X)$, as a function of a, appears in [13]; the more familiar graph of $\mathrm{E}\,|\ln(T'X)|$ appears in [14]. In a sense, such plotting is meaningless, because there always exists finer detail than captured in whatever scale we choose [15].

Jakobson [16, 17] proved that the set $A = \{a \in [0, 4] : T$ has an absolutely continuous invariant density$\}$ has positive measure. Both $4 \in A$ and $3.678... \in A$, but the status of values like 3.6, 3.7, 3.8 or 3.9 is unknown. Note: the condition that a density be *absolutely continuous* is important, yet outside our scope of study.

What can be said about T for $a \notin A$? This question was satisfactorily answered only recently [18, 19].

The *metric entropy* of T can be proved to be equal to the Lyapunov exponent, but the *topological entropy* is altogether a different characterization [20–23]. For the regular continued fraction transformation $T_{\mathrm{RCF}}(x) = \{1/x\}$, the metric entropy is $\pi^2/(6\ln(2))$ while the topological entropy is infinite [24]. The limit of $\mathrm{E}\left(\hat{\lambda}_n(X)\right)$ as $n \to \infty$ is $\pi^2/(6\ln(2))$; the limit of $n\,\mathrm{Var}\left(\hat{\lambda}_n(X)\right)$ as $n \to \infty$ is equal to $4(0.8621470373...)$ and, in fact, the Central Limit Theorem holds [2].

One-dimensional maps of the interval have inspired much computation [25–35]. We mention, for example, the maps $T_\ell : [0,1] \to [0,1]$ defined by

$$T_\ell(x) = 1 - |2x - 1|^\ell$$

for real $\ell > 1$. Clearly the case $\ell = 2$ gives the Ulam–von Neumann map. Each T_ℓ has an absolutely continuous invariant density with metric entropies (Lyapunov exponents) equal to [32, 34]

$$\begin{cases} \ln(2) = 0.6931471805... & \text{if } \ell = 2, \\ 0.6908569334... & \text{if } \ell = 3, \\ 0.6844935750... & \text{if } \ell = 4, \\ 0.6756910613... & \text{if } \ell = 5. \end{cases}$$

As another example, consider the map $S_0 : [0,1] \to [0,1]$ defined by

$$S_0(x) = \left\{ 2x + \frac{1}{4\pi} \sin(2\pi x) \right\}.$$

The absolutely continuous invariant density of S_0 has entropy equal to $0.6837719602....$ It would be good someday to see such high-precision results for the logistic map, given values of a other than $3.678...$ and 4.

[1] G.-H. Choe, *Computational Ergodic Theory*, Springer-Verlag, 2005, pp. 47–84, 269–271; MR2130385 (2005m:37206).
[2] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.
[3] S. L. Ulam and J. von Neumann, On combination of stochastic and deterministic processes, *Bull. Amer. Math. Soc.* 53 (1947) 1120.
[4] C. Anteneodo, Statistics of finite-time Lyapunov exponents in the Ulam map, *Phys. Rev. E* 69 (2004) 016207; nlin/0309065.
[5] J. Theiler and L. A. Smith, Anomalous convergence of Lyapunov exponent estimates, *Phys. Rev. E* 51 (1995) 3738–3741.
[6] A. Prasad and R. Ramaswamy, Characteristic distributions of finite-time Lyapunov exponents, *Phys. Rev. E* 60 (1999) 2761–2766; chao-dyn/9906022.
[7] D. Ruelle, Applications conservant une mesure absolument continue par rapport à *dx* sur [0, 1], *Comm. Math. Phys.* 55 (1977) 47–51; MR0467840 (57 #7691).

[8] M. Misiurewicz, Absolutely continuous measures for certain maps of an interval, *Inst. Hautes Études Sci. Publ. Math.* 53 (1981) 17–51; MR0623533 (83j:58072).

[9] F. Schweiger, *Ergodic Theory of Fibred Systems and Metric Number Theory*, Oxford Univ. Press, 1995, pp. 121–122, 245–249; MR1419320 (97h:11083).

[10] B. R. Hunt, Estimating invariant measures and Lyapunov exponents, *Ergodic Theory Dynam. Systems* 16 (1996) 735–749; MR1406431 (98a:58103).

[11] S. R. Finch, Feigenbaum-Coullet-Tresser constants, *first volume*, pp. 65–76.

[12] J. C. Sprott, Misiurewicz point of the logistic map, unpublished note (2006).

[13] H. L. D. de S. Cavalcante and J. R. Rios Leite, Bifurcations and averages in the logistic map, *Dyn. Stab. Syst.* 15 (2000) 35–41; MR1760800 (2000m:37045).

[14] D. Ruelle, *Chaotic Evolution and Strange Attractors*, Cambridge Univ. Press, 1989, pp. 25–26, 40–43, 58–61; MR1017644 (91d:58169)

[15] P. Young, Logistic map (period doubling route to chaos), Mathematica program logistic.nb.

[16] M. V. Jakobson, Absolutely continuous invariant measures for one-parameter families of one-dimensional maps, *Comm. Math. Phys.* 81 (1981) 39–88; MR0630331 (83j:58070).

[17] M. Benedicks and L. Carleson, On iterations of $1 - ax^2$ on $(-1, 1)$, *Annals of Math.* 122 (1985) 1–25; MR0799250 (87c:58058).

[18] M. Lyubich, Almost every real quadratic map is either regular or stochastic, *Annals of Math.* 156 (2002) 1–78; MR1935840 (2003i:37032).

[19] A. Avila and C. G. Moreira, Statistical properties of unimodal maps: the quadratic family, *Annals of Math.* 161 (2005) 831–881; MR2153401 (2006d:37061).

[20] J. P. Crutchfield and N. H. Packard, Symbolic dynamics of one-dimensional maps: entropies, finite precision, and noise, *Internat. J. Theoret. Phys.* 21 (1981/82) 433–466; MR0658310 (84b:58064).

[21] L. Block, J. Keesling, S. H. Li and K. Peterson, An improved algorithm for computing topological entropy, *J. Stat. Phys.* 55 (1989) 929–939; MR1002478 (90k:58093).

[22] P. Góra and A. Boyarsky, Computing the topological entropy of general one-dimensional maps, *Trans. Amer. Math. Soc.* 323 (1991) 39–49; MR1062871 (92a:58083).

[23] G. Froyland, O. Junge and G. Ochs, Rigorous computation of topological entropy with respect to a finite partition, *Phys. D* 154 (2001) 68–84; MR1840806 (2002c:37021).

[24] R. M. Corless, G. W. Frank and J. G. Monroe, Chaos and continued fractions, *Phys. D* 46 (1990) 241–253; MR1083721 (92c:58074).

[25] T. Y. Li, Finite approximation for the Frobenius-Perron operator. A solution to Ulam's conjecture, *J. Approx. Theory* 17 (1976) 177–186; MR0412689 (54 #811).

[26] C. Chiu, Q. Du and T. Y. Li, Error estimates of the Markov finite approximation of the Frobenius-Perron operator, *Nonlinear Anal.* 19 (1992) 291–308; MR1178404 (93f:28014).

[27] L. R. Mead, Approximate solution of Fredholm integral equations by the maximum-entropy method, *J. Math. Phys.* 27 (1986) 2903–2907.

[28] J. Ding, A maximum entropy method for solving Frobenius-Perron operator equations, *Appl. Math. Comput.* 93 (1998) 155–168; MR1629441 (99g:65065).

[29] J. Ding and L. R. Mead, Maximum entropy approximation for Lyapunov exponents of chaotic maps, *J. Math. Phys.* 43 (2002) 2518–2522; MR1893684 (2003a:37044).

[30] F. Y. Hunt, A Monte Carlo approach to the approximation of invariant measures, *Random Comput. Dynam.* 2 (1994) 111–133; erratum, 5 (1997) 361–362; MR1265229 (95a:58074) and MR1483876 (98j:58069).
[31] P. Hall and R. C. L. Wolff, Properties of invariant distributions and Lyapunov exponents for chaotic logistic maps, *J. Royal Statist. Soc. Ser. B* 57 (1995) 439–452; MR1323349 (95k:62034).
[32] R. Murray, Discrete Approximation of Invariant Densities, Ph.D. thesis, University of Cambridge, 1997.
[33] M. Keane, R. Murray and L.-S. Young, Computing invariant measures for expanding circle maps, *Nonlinearity* 11 (1998) 27–46; MR1492949 (99a:58106).
[34] M. Pollicott and O. Jenkinson, Computing invariant densities and metric entropy, *Comm. Math. Phys.* 211 (2000) 687–703; MR1773814 (2001i:37015).
[35] O. Jenkinson and M. Pollicott, Orthonormal expansions of invariant densities for expanding maps, *Adv. Math.* 192 (2005) 1–34; MR2122279 (2005m:37053).

4.14 Lyapunov Exponents. II

Before discussing continuous-time systems, let us emphasize the definition of Lyapunov exponent λ for discrete-time systems in one-dimension [1]. If $f:\mathbb{R}\to\mathbb{R}$ is differentiable and

$$x_n = f(x_{n-1}), \quad x_0 = u,$$

then λ quantifies the exponential rate at which two initially close points u, u_0 separate under the iteration:

$$\lambda = \lim_{n\to\infty}\frac{1}{n}\ln|f^n(u) - f^n(u_0)| = \lim_{n\to\infty}\frac{1}{n}\ln|D_u f^n(u_0)\,(u-u_0)|$$

almost always. For example, $\lambda = \ln(2)$ is experimentally verified for the logistic case $f(x) = 4x\,(1-x)$ and $u_0 = 1/3$. This definition is meaningful as well for multi-dimensional maps $f:\mathbb{R}^m\to\mathbb{R}^m$. It is not true, however, that the norm of a product of Jacobian matrices is equal to the product of their norms; thus the calculational technique (based on the chain rule) used in [2] fails for $m>1$.

Consider the **classical Lorenz system** [3–7]

$$\begin{cases} dx/dt = -10(x-y), & x(0)=0, \\ dy/dt = 28x - y - xz, & y(0)=1, \\ dz/dt = xy - \frac{8}{3}z, & z(0)=0 \end{cases}$$

and define, for convenience,

$$X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad F(X) = \begin{pmatrix} -10(x-y) \\ 28x - y - xz \\ xy - \frac{8}{3}z \end{pmatrix}.$$

Let $U = (u, v, w)$ denote a point that is close to the initial state $U_0 = (0, 1, 0)$. The solution of the perturbed system

$$dX/dt = F(X), \quad X(0) = U$$

is written as $X(t; u, v, w)$. Differentiating both sides with respect to U, we obtain the **variational equation** [8–11]

$$d\Phi/dt = D_X F(X)\,\Phi(t), \quad \Phi(0) = I$$

where $\Phi(t; u, v, w) = D_U X(t; U)$ is a 3×3 matrix. The Lyapunov exponent λ for the Lorenz system satisfies [1]

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \ln |X(t; U) - X(t; U_0)| = \lim_{t\to\infty} \frac{1}{t} \ln |D_U X(t; U_0)\,(U - U_0)|$$

almost always. It follows that [12]

$$\lambda = \lim_{t\to\infty} \frac{1}{t} \ln |\Phi(t; U_0)|$$

where $|M|$ denotes the 2-norm (largest singular value) of a matrix M. Equivalently, $|M|$ is the square root of the largest eigenvalue of $M^T M$. To compute λ, therefore, we must possess not only $(x(t), y(t), z(t))$ but also

$$\Phi(t) = (\varphi_{i,j}(t))_{1\le i\le 3,\, 1\le j\le 3}$$

where

$$\begin{cases} d\varphi_{1,j}/dt = -10(\varphi_{1,j} - \varphi_{2,j}), & \varphi_{1,1}(0) = 1, \quad \varphi_{1,2}(0) = \varphi_{1,3}(0) = 0, \\ d\varphi_{2,j}/dt = (28 - z)\varphi_{1,j} - \varphi_{2,j} - x\,\varphi_{3,j}, & \varphi_{2,2}(0) = 1, \quad \varphi_{2,1}(0) = \varphi_{2,3}(0) = 0, \\ d\varphi_{3,j}/dt = y\,\varphi_{1,j} + x\,\varphi_{2,j} - \frac{8}{3}\varphi_{3,j}, & \varphi_{3,3}(0) = 1, \quad \varphi_{3,1}(0) = \varphi_{3,2}(0) = 0 \end{cases}$$

for $j = 1, 2, 3$. Difficulties arising from integrating this 12×12 ODE system include numerical overflow and numerical rank deficiency [12]. We obtain experimentally that $\lambda \approx 0.9$ via this approach; approximating $\Phi(t)$ as $t \to \infty$ to higher precision seems hopeless.

Using alternative approaches, Viswanath [12–14] and Sprott [15–17] independently computed that $\lambda = 0.90563\ldots$. It is known via rigorous numerics that the classical Lorenz system is chaotic [18–24] and that, indeed, almost all points in state space tend to a strange attractor (the famous *Lorenz butterfly*) [25–27]. No such behavior can possibly occur for continuous flows in one or two dimensions. The literature on calculating Lypanouv exponents is huge; we merely mention a few helpful surveys [28–34].

Although the Lorenz system was originally derived from a meteorological model of fluid convection, it can be more easily formulated in connection with the Malkus water wheel [4–6]. The wheel is free to rotate about a horizontal axis and its circumference is composed of small leaky cells. Water pours into the cells near the top of the wheel at a constant rate. Water leaks out of each cell at a rate proportional to the density of water inside. The mass of the wheel consists entirely of water confined to the circumference. As the wheel starts to rotate, new cells will move into position to receive the water. With the right balance between rates of water in-flow and out-flow, as well as frictional damping and gravitational acceleration, the Lorenz system emerges (governing, for example, angular velocity of the wheel). This is a fascinatingly simple illustration of chaos!

What is the algebraically simplest example of a dissipative chaotic flow? Sprott [35] conjectured that

$$\frac{d^3\xi}{dt^3}+\frac{2017}{1000}\frac{d^2\xi}{dt^2}-\left(\frac{d\xi}{dt}\right)^2+\xi=0,$$

with Lyapunov exponent 0.0551..., is one such case. For conservative flows,

$$\frac{d^3\xi}{dt^3}+\frac{d\xi}{dt}-\xi^2+\frac{1}{100}=0$$

may be algebraically simplest [36]. Upon setting $\eta=5\xi+1/2$, an equation resembling the logistic equation:

$$\frac{d^3\eta}{dt^3}+\frac{d\eta}{dt}+\frac{1}{5}\eta(1-\eta)=0$$

is the interesting outcome, with Lyapunov exponent 0.0964.... A survey of this line of thought is found in [15, 37, 38]. These examples deserve further analysis.

A single pendulum [39]

$$\begin{cases} d\theta/dt=\omega, \\ d\omega/dt=-(g/\ell)\sin(\theta) \end{cases}$$

cannot exhibit chaos. Only with the introduction of a nonautonomous driving term (and possibly a viscous damping term) can chaos arise: see [40–45]. By contrast, a double pendulum [39, 46–48]

$$\begin{cases} \dfrac{d\theta_1}{dt}=\omega_1, \\ \dfrac{d\theta_2}{dt}=\omega_2, \\ \dfrac{d\omega_1}{dt}=\dfrac{-m_2\sin(\theta_1-\theta_2)\left(\ell_1\cos(\theta_1-\theta_2)\omega_1^2+\ell_2\omega_2^2\right)-\frac{g}{2}\left((2m_1+m_2)\sin(\theta_1)+m_2\sin(\theta_1-2\theta_2)\right)}{\ell_1\left(m_1+m_2-m_2\cos(\theta_1-\theta_2)^2\right)}, \\ \dfrac{d\omega_2}{dt}=\sin(\theta_1-\theta_2)\dfrac{(m_1+m_2)\left(g\cos(\theta_1)+\ell_1\omega_1^2\right)+\ell_2 m_2\cos(\theta_1-\theta_2)\omega_2^2}{\ell_2\left(m_1+m_2-m_2\cos(\theta_1-\theta_2)^2\right)} \end{cases}$$

exhibits chaos if, for example, θ_1 is initially large ($\pi/2<\theta_1<\pi$) and $\theta_2=\omega_1=\omega_2=0$. (Point-mass m_1 determines angle θ_1 relative to a downward vertical axis at ceiling suspension; point-mass m_2 determines angle θ_2 relative to a downward vertical axis at m_1; the connecting rods of length ℓ_1, ℓ_2 are massless and no friction or forcing occurs; g is acceleration due to gravity.) The value of a Lyapunov exponent computed in [49] awaits confirmation.

We merely mention the interesting feedback control-theoretic problem of stabilizing an inverted pendulum on a moving cart [50–58]. Under the most ideal conditions, chaos cannot occur. If, however, we include realistic effects like time delay [59, 60], discrete sampling [61] or system friction [62], then chaos becomes possible again. More on a torque-driven pendulum (not cart-driven) and optimal control is found in [63].

[1] J.-P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, *Rev. Modern Phys.* 57 (1985) 617–656; addendum 57 (1985) 1115; MR0800052 (87d:58083a) and MR0809538 (87d:58083b).

[2] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.

[3] E. N. Lorenz, Deterministic nonperiodic flow, *J. Atmos. Sci.* 20 (1963) 130–141.

[4] E. N. Lorenz, On the prevalence of aperiodicity in simple systems, *Global Analysis*, Proc. 1978 Calgary conf., ed. M. Grmela and J. E. Marsden, Lect. Notes in Math. 755, Springer-Verlag, 1979, pp. 53–75; MR0564902 (81f:58026).

[5] C. Sparrow, *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*, Springer-Verlag, 1982, pp. 1–12, 194–195; MR0681294 (84b:58072).

[6] S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Westview Press, 1994, pp. 301–347, 462.

[7] H.-O. Peitgen, H. Jürgens and D. Saupe, *Chaos and Fractals. New Frontiers of Science*, 2nd ed., Springer-Verlag, 2004, pp. 647–670, 700–702; MR2031217 (2004m:37061).

[8] T. S. Parker and L. O. Chua, *Practical Numerical Algorithms for Chaotic Systems*, Springer-Verlag, 1989, pp. 57–82, 305–306; MR1009561 (91b:58161).

[9] M. Sandri, Numerical calculation of Lyapunov exponents, *Mathematica J.* 6 (1996) 78–84.

[10] K. T. Alligood, T. D. Sauer and J. A. Yorke, *Chaos. An Introduction to Dynamical Systems*, Springer-Verlag, 1997, pp. 105–123, 193–202, 256–265, 379–386; MR1418166 (98a:58113).

[11] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, 2nd ed., Springer-Verlag, 2003, pp. 726–735; MR2004534 (2004h:37002).

[12] D. Viswanath, Lyapunov Exponents from Random Fibonacci Sequences to the Lorenz Equations, Ph.D. thesis, Cornell Univ., 1998.

[13] D. Viswanath, Symbolic dynamics and periodic orbits of the Lorenz attractor. *Nonlinearity* 16 (2003) 1035–1056; MR1975795 (2004b:37063).

[14] D. Viswanath, The fractal property of the Lorenz attractor, *Physica D* 190 (2004) 115–128; MR2043795 (2005a:37057).

[15] J. C. Sprott, *Chaos and Time-Series Analysis*, Oxford Univ. Press, 2003, pp. 90–99, 104–126, 253–255; MR2014088 (2004k:37063).

[16] J. C. Sprott, Lyapunov exponent and dimension of Lorenz attractor, unpublished note (2005).

[17] J. C. Sprott, Lyapunov exponent spectrum software, unpublished note (2010).

[18] S. P. Hastings and W. C. Troy, A shooting approach to the Lorenz equations, *Bull. Amer. Math. Soc.* 27 (1992) 298–303; arXiv:math/9210217v1; MR1161275 (93f:58150).

[19] B. Hassard, J. Zhang, S. P. Hastings and W. C. Troy, A computer proof that the Lorenz equations have “chaotic” solutions, *Appl. Math. Lett.* 7 (1994) 79–83; MR1349899 (96d:58082).

[20] S. P. Hastings and W. C. Troy, A shooting approach to chaos in the Lorenz equations, *J. Differential Equations* 127 (1996) 41–53; MR1387256 (97e:34079).
[21] Z. Galias and P. Zgliczyński, Computer assisted proof of chaos in the Lorenz equations, *Physica D* 115 (1998) 165–188; MR1626596 (99h:58123).
[22] K. Mischaikow and M. Mrozek, Chaos in the Lorenz equations: a computer-assisted proof, *Bull. Amer. Math. Soc.* 32 (1995) 66–72; MR1276767 (95e:58121).
[23] K. Mischaikow and M. Mrozek, Chaos in the Lorenz equations: a computer assisted proof. II. Details, *Math. Comp.* 67 (1998) 1023–1046; MR1459392 (98m:58095).
[24] K. Mischaikow, M. Mrozek and A. Szymczak, Chaos in the Lorenz equations: a computer assisted proof. III. Classical parameter values, *J. Differential Equations* 169 (2001) 17–56; MR1808460 (2002f:37058).
[25] W. Tucker, The Lorenz attractor exists, *C. R. Acad. Sci. Paris Sér. I Math.* 328 (1999) 1197–1202; MR1701385 (2001b:37051).
[26] W. Tucker, A rigorous ODE solver and Smale's 14th problem, *Found. Comput. Math.* 2 (2002) 53–117; MR1870856 (2003b:37055).
[27] M. Viana, What's new on Lorenz strange attractors? *Math. Intelligencer* 22 (2000) 6–19; MR1773551 (2001h:37068).
[28] K. Geist, U. Parlitz and W. Lauterborn, Comparison of different methods for computing Lyapunov exponents, *Progr. Theoret. Phys.* 83 (1990) 875–893; MR1070085 (91j:65136).
[29] A. Karantonis and M. Pagitsas, Comparative study for the calculation of the Lyapunov spectrum from nonlinear experimental signals, *Phys. Rev. E* 53 (1996) 5428–5444.
[30] L. Dieci, R. D. Russell and E. S. Van Vleck, On the computation of Lyapunov exponents for continuous dynamical systems, *SIAM J. Numer. Anal.* 34 (1997) 402–423; MR1445743 (97m:65139).
[31] K. Ramasubramanian and M. S. Sriram, A comparative study of computation of Lyapunov spectra with different algorithms, *Physica D* 139 (2000) 72–86; arXiv:chao-dyn/9909029; MR1744799 (2000j:37120).
[32] G. Tancredi, A. Sanchez and F. Roig, A comparison between methods to compute Lyapunov exponents, *Astron. J.* 121 (2001) 1171–1179.
[33] J. B. Dingwell, Lyapunov exponents, *Wiley Encyclopedia of Biomedical Engineering.*, ed. M. Akay, Wiley, 2006.
[34] M. D. Hartl, Lyapunov exponents in constrained and unconstrained ordinary differential equations, arXiv:physics/0303077.
[35] J. C. Sprott, Simplest dissipative chaotic flow, *Phys. Lett. A* 228 (1997) 271–274; MR1442639 (97m:58141).
[36] J. C. Sprott, Some simple chaotic jerk functions, *Amer. J. Phys.* 65 (1997) 537–543.
[37] J. C. Sprott and S. J. Linz, Algebraically simple chaotic flows, *Internat. J. Chaos Theory Appl.* 5 (2000) 3–22.
[38] J. C. Sprott, *Elegant Chaos: Algebraically Simple Chaotic Flows*, World Scientific, 2010; MR2643315 (2011k:37001).
[39] S. W. McCuskey, *An Introduction to Advanced Dynamics*, Addison-Wesley, 1959, pp. 53–59; MR0099759 (20 #6197).
[40] G. L. Baker and J. P. Gollub, *Chaotic Dynamics. An Introduction*, 2nd ed., Cambridge Univ. Press, 1996, pp. 1–6, 120–125, 131; MR1415612 (97j:58096).
[41] G. L. Baker and J. A. Blackburn, *The Pendulum. A Case Study in Physics*, Oxford Univ. Press, 2005, pp. 121–152; MR2155144 (2006a:70001).
[42] R. Doerner, B. Hübinger, H. Heng and W. Martienssen, Approaching nonlinear dynamics by studying the motion of a pendulum. II, Analyzing

chaotic motion, *Internat. J. Bifur. Chaos Appl. Sci. Engrg.* 4 (1994) 761–771; MR1314077.

[43] J. C. Sprott, *Chaos and Time-Series Analysis*, Oxford Univ. Press, 2003, pp. 5–8, 87–89, 184–193, 428, 439; MR2014088 (2004k:37063).

[44] J. C. Sprott, Simple driven conservative chaotic flows, unpublished note (2001).

[45] H. P. W. Gottlieb and J. C. Sprott, Simplest driven conservative chaotic oscillator, *Phys. Lett. A* 291 (2001) 385–388; MR1917691 (2003e:70033).

[46] P. H. Richter and H.-J. Scholz, Chaos in classical mechanics: the double pendulum, *Stochastic Phenomena and Chaotic Behaviour in Complex Systems*, Proc. 1983 Flattnitz conf., ed. P. Schuster, Springer-Verlag, 1984, pp. 86–97; MR0771695.

[47] T. Shinbrot, C. Grebogi, J. Wisdom and J. A. Yorke, Chaos in a double pendulum, *Amer. J. Phys.* 60 (1992) 491–499.

[48] R. B. Levien and S. M. Tan, Double pendulum: An experiment in chaos, *Amer. J. Phys.* 61 (1993) 1038-1044.

[49] T. Stachowiak and T. Okada, A numerical analysis of chaos in the double pendulum, *Chaos Solitons Fractals* 29 (2006) 417–422.

[50] F. M. Phelps and J. H. Hunter, An analytical solution of the inverted pendulum, *Amer. J. Phys.* 33 (1965) 285–295; comments/reply 34 (1966) 533–535; MR0202363 (34 #2234).

[51] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, Wiley, 1972, pp. 4–7, 139–140, 195–196; MR0406607 (53 #10394).

[52] K. J. Åström and K. Furuta, Swinging up a pendulum by energy control, *Automatica* 36 (2000) 287–295; MR1827782 (2002f:93083).

[53] D. Chatterjee, A. Patra and H. K. Joglekar, Swing-up and stabilization of a cart-pendulum system under restricted cart track length, *Systems Control Lett.* 47 (2002) 355–364; MR2008713.

[54] C. Aguilar Ibañez, O. Gutiérrez Frias and M. Suárez Castañón, Lyapunov-based controller for the inverted pendulum cart system, *Nonlinear Dynam.* 40 (2005) 367–374; MR2144754.

[55] N. Muskinja and B. Tovornik, Swinging up and stabilization of a real inverted pendulum, *IEEE Trans. Industrial Electronics* 53 (2006) 631–639.

[56] P. Mason, M. Broucke and B. Piccoli, Time optimal swing-up of the planar pendulum, *IEEE Trans. Automat. Control* 53 (2008) 1876–1886; MR2454754 (2009j:49080).

[57] P. Paoletti and R. Genesio, Rate limited time optimal control of a planar pendulum, *Systems Control Lett.* 60 (2011) 264–270; MR2798163 (2012j:49053).

[58] O. Boubaker, The inverted pendulum: A fundamental benchmark in control theory and robotics, *Proc. IEEE Internat. Conf. on Education and e-Learning Innovations (ICEELI)*, Sousse, 2012, pp. 1–6; arXiv:1405.3094.

[59] J. Sieber and B. Krauskopf, Complex balancing motions of an inverted pendulum subject to delayed feedback control, *Physica D* 197 (2004) 332–345; MR2093577 (2005d:93075).

[60] J. Sieber and B. Krauskopf, Chaos induced by delay in a linearly controlled system, unpublished poster (2004).

[61] G. Haller and G. Stépán, Micro-chaos in digital control, *J. Nonlinear Sci.* 6 (1996) 415–448; MR1411342 (98b:93045).

[62] S. A. Campbell, S. Crawford and K. Morris, Time delay and feedback control of an inverted pendulum with stick slip friction, *Proc. ASME Internat. Design Engineering Technical Conf. (IDETC)*, Las Vegas, 2007.

[63] S. R. Finch, Swing-up control of a pendulum, *this volume*, §3.16.

4.15 Lyapunov Exponents. III

Our ongoing study encompasses both discrete iteration [1] and continuous flow [2]; the system dynamics can be either deterministic or stochastic. Let A denote a real $m \times m$ matrix and B, X denote real m-vectors. Consider the difference equation

$$X_n = A\,X_{n-1} + B\,\varepsilon_n, \quad X_0 \text{ arbitrary}$$

where ε_n is scalar $N(0,1)$ white noise. Order the complex eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ of A so that λ_1 has maximum modulus. When $|\lambda_1| > 1$, it follows that

$$\tfrac{1}{n} \ln|X_n| \to \ln|\lambda_1| > 0 \quad \text{almost surely as } n \to \infty$$

which indicates that no convergence to stationarity can occur. The quantity $\ln|\lambda_1|$ is the Lyapunov exponent of the system, since the derivative of the linear transformation $x \mapsto A\,x$ is itself.

Consider instead the differential equation

$$dX_t = A\,X_t\,dt + B\,dW_t, \quad X_0 \text{ arbitrary}$$

where W_t is scalar Brownian motion with unit variance. The corresponding flow is

$$X_t = e^{A\,t}\left(X_0 + \int_0^t e^{-A\,s} B\,dW_s\right)$$

and the complex eigenvalues of e^A are $e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_m}$. Here, however, we order $\lambda_1, \lambda_2, \ldots, \lambda_m$ so that λ_1 has maximum *real part* (which implies that e^{λ_1} has maximum modulus). When $\mathrm{Re}(\lambda_1) > 0$, the interpretation of

$$\tfrac{1}{t} \ln|X_t| \to \mathrm{Re}(\lambda_1) \quad \text{almost surely as } t \to \infty,$$

is exactly as before. An informal proof is to choose X_0 to be the dominant eigenvector of A or of e^A, respectively, and to choose $B = 0$; then

$$|X_n| = |A^n X_0| = |\lambda_1|^n |X_0| \quad \text{or} \quad |X_t| = |e^{A\,t} X_0| = |e^{\lambda_1}|^t |X_0|,$$

respectively. See [3] for special treatment of the case $m = 1$. The probability density for

$$\ln|X_n| - n\,\ln|\lambda_1|, \quad \ln|X_t| - t\,\mathrm{Re}(\lambda_1)$$

is also of interest, and turns out to be doubly-exponential [3, 4].

Additive noise does not enter the formula for Lyapunov exponents; *multiplicative* noise contrasts in this regard. Let A, B denote real $m \times m$ matrices. The equations

$$X_n = A\,X_{n-1} + B\,X_{n-1}\,\varepsilon_n, \quad X_0 \text{ arbitrary};$$

$$dX_t = A\,X_t\,dt + B\,X_t\,dW_t, \quad X_0 \text{ arbitrary}$$

require more intricate analysis. Let us focus only on the continuous-time case for now, leaving the discrete-time case for later.

In the event A and B commute, that is, $A\,B = B\,A$, it can be proved that [5]

$$X_t = \exp\left(\left(A - \tfrac{1}{2}B^2\right) t + B\,W_t\right) X_0.$$

There is, however, no consequential formula for the Lyapunov exponent that is valid for all $m \geq 1$ and all A, B.

Set $m = 1$ or $m = 2$. Let us adhere to the convention of replacing A by $A + \frac{1}{2}B^2$:

$$dX_t = \left(A + \tfrac{1}{2}B^2\right) X_t\,dt + B\,X_t\,dW_t, \quad X_0 \text{ arbitrary.}$$

If $m = 1$, $A = a$ and $B = \sigma > 0$, then the random variable $\ln|X_t/X_0|$ is normally distributed with mean $a\,t$ and variance $\sigma^2 t$ (the process X_t is often called geometric Brownian motion). Clearly

$$\tfrac{1}{t}\ln|X_t| \to a \quad \text{almost surely as } t \to \infty.$$

Stability is unchanged by noise in this example. The same can be said if $m = 2$,

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma \end{pmatrix}$$

where $a > b$ and $\sigma > 0$. If instead

$$B = \begin{pmatrix} 0 & -\sigma \\ \sigma & 0 \end{pmatrix}$$

then it can be proved that [6, 7]

$$\tfrac{1}{t}\ln|X_t| \to \tfrac{1}{2}(a+b) + \tfrac{1}{2}(a-b)\frac{I_1\left(\frac{a-b}{2\sigma^2}\right)}{I_0\left(\frac{a-b}{2\sigma^2}\right)} \quad \text{almost surely}$$

where I_0, I_1 are modified Bessel functions [8]. For example, when $a = 1$, $b = -2$ and $\sigma = 10$, the Lyapunov exponent has value $-0.4887503163...$ [9]. For the same a and b, the Lyapunov exponent has value $0.3941998582...$ when $\sigma = 1$, and is zero precisely when $\sigma = 1.4560286969...$ [6]. More noise implies enhanced stability in this example.

If instead [10, 11]

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ \sigma & 0 \end{pmatrix}$$

then

$$\tfrac{1}{t}\ln|X_t| \to \kappa\,\sigma^{2/3} = (0.2893082598...)\sigma^{2/3} \quad \text{almost surely}$$

and

$$\kappa = \frac{\pi}{12^{1/6}\Gamma(1/3)^2} = \frac{3^{1/3}\sqrt{\pi}}{2^{2/3}\Gamma(1/6)}.$$

What happens when the bottom row of A is nonzero? If

$$A = \begin{pmatrix} 0 & 1 \\ -\alpha & 2\beta \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ \sigma & 0 \end{pmatrix}$$

where $\beta^2 > \alpha$, then more complicated formulation emerges. We avoid the hypergeometric functions in [12, 13], preferring modified Bessel functions (of both integer and fractional types). Let

$$\gamma = \frac{\sigma^2}{2}, \quad \delta = \frac{4\left(\beta^2 - \alpha\right)}{9\gamma^2}$$

for convenience. Define

$$\begin{aligned}
f(\alpha,\beta,\gamma) &= \frac{3}{2\pi^{\frac{3}{2}}} \int_0^\infty \sqrt{z} \exp\left(-\frac{1}{12}\gamma^2 z^3 + \left(\beta^2-\alpha\right) z\right) dz \\
&= \frac{\delta^{\frac{1}{2}} I_{-\frac{2}{3}}\left(\sqrt{\delta}\right) I_{-\frac{1}{3}}\left(\sqrt{\delta}\right)}{\gamma} + \frac{2\left(\frac{2}{3}\right)^{\frac{1}{3}} \delta^{\frac{1}{3}} I_{\frac{1}{3}}\left(\sqrt{\delta}\right)^2}{\left(\beta^2-\alpha\right)\gamma^{\frac{1}{3}}} + \frac{\delta^{\frac{1}{2}} I_{\frac{1}{3}}\left(\sqrt{\delta}\right) I_{\frac{2}{3}}\left(\sqrt{\delta}\right)}{\gamma} \\
&\quad + \frac{6\left(\frac{2}{3}\right)^{\frac{2}{3}} \gamma^{\frac{1}{3}} \delta^{\frac{2}{3}} I_{\frac{2}{3}}\left(\sqrt{\delta}\right)^2}{\left(\beta^2-\alpha\right)^2} + \frac{2\left(\frac{2}{3}\right)^{\frac{2}{3}} \left(\beta^2-\alpha\right) \delta^{\frac{1}{6}} I_{\frac{1}{3}}\left(\sqrt{\delta}\right) I_{\frac{4}{3}}\left(\sqrt{\delta}\right)}{\gamma^{\frac{5}{3}}} \\
&\quad + \frac{2\left(\frac{2}{3}\right)^{\frac{1}{3}} \left(\beta^2-\alpha\right)^{\frac{1}{2}} \delta^{\frac{1}{3}} I_{\frac{2}{3}}\left(\sqrt{\delta}\right) I_{\frac{5}{3}}\left(\sqrt{\delta}\right)}{\gamma^{\frac{4}{3}}},
\end{aligned}$$

$$\begin{aligned}
g(\alpha,\beta,\gamma) &= \frac{3}{2\pi^{\frac{3}{2}}} \int_0^\infty \frac{1}{\sqrt{z}} \exp\left(-\frac{1}{12}\gamma^2 z^3 + \left(\beta^2-\alpha\right) z\right) dz \\
&= \frac{\left(\frac{2}{3}\right)^{\frac{1}{3}} \delta^{\frac{1}{3}} \left(I_{-\frac{1}{3}}\left(\sqrt{\delta}\right)^2 + I_{-\frac{1}{3}}\left(\sqrt{\delta}\right) I_{\frac{1}{3}}\left(\sqrt{\delta}\right) + I_{\frac{1}{3}}\left(\sqrt{\delta}\right)^2\right)}{\gamma^{\frac{1}{3}}} \\
&= \frac{3\left(\mathrm{Ai}\left(\left(\frac{3}{2}\right)^{\frac{2}{3}} \delta^{\frac{1}{3}}\right)^2 + \mathrm{Bi}\left(\left(\frac{3}{2}\right)^{\frac{2}{3}} \delta^{\frac{1}{3}}\right)^2\right)}{2\gamma^{\frac{1}{3}}}
\end{aligned}$$

then

$$\tfrac{1}{t}\ln|X_t| \to \beta + \frac{\gamma}{2}\frac{f(\alpha,\beta,\gamma)}{g(\alpha,\beta,\gamma)} \quad \text{almost surely.}$$

For example, when $\alpha = 1$ and $|\beta| > 1$ is fixed, the Lyapunov exponent is decreasing as a function of $\gamma \in (0, (\beta^2-1)^{3/2}\gamma_0)$ and increasing for

$\gamma \in ((\beta^2 - 1)^{3/2}\gamma_0, \infty)$, where $\gamma_0 = 1.6946141069....$ At criticality, we have

$$\begin{aligned}\frac{f(1, \beta, (\beta^2 - 1)^{3/2}\gamma_0)}{g(1, \beta, (\beta^2 - 1)^{3/2}\gamma_0)} &= \frac{1}{(\beta^2 - 1)\gamma_0}(1.4567743021...) \\ &= \frac{2}{(\beta^2 - 1)\gamma_0}(1 + 0.8848441574...)^{-1/2};\end{aligned}$$

further stabilization by noise beyond this point is impossible. As another example, when $\alpha = 1$ and $\gamma > 0$ is fixed, we have

$$\lim_{|\beta| \to 1^+} \frac{f(1, \beta, \gamma)}{g(1, \beta, \gamma)} = \left(\frac{4}{\gamma}\right)^{2/3} \kappa = \frac{2}{\gamma} \kappa\, \sigma^{2/3},$$

consistent with preceding zero-row results. The constant $\kappa = 0.2893082598...$ also appears in [14], but reasons for this connection are unclear.

Explicit expressions like the above are quite rare in this area. A promising approach is presented in [15] but unfortunately no examples are given.

Addendum As an illustration, Baxendale [6] determined the Lyapunov exponent $-0.48875...$ for

$$dX_t = \begin{pmatrix} a - \frac{1}{2}\sigma^2 & 0 \\ 0 & b - \frac{1}{2}\sigma^2 \end{pmatrix} X_t\, dt + \begin{pmatrix} 0 & -\sigma \\ \sigma & 0 \end{pmatrix} X_t\, dW_t$$

when $a = 1$, $b = -2$ and $\sigma = 10$. He suggested an approach for computing the corresponding central limit variance. No one has evaluated this variance until recently [16]; it turns out to be 0.011248... but has a more complicated expression than a simple ratio of modified Bessel functions.

[1] S. R. Finch, Lyapunov exponents, *this volume*, §4.13.

[2] S. R. Finch, Lyapunov exponents. II, *this volume*, §4.14.

[3] S. R. Finch, Another look at AR(1), arXiv:0710.5419.

[4] S. R. Finch, Extreme value constants, *first volume*, pp. 363–367.

[5] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Wiley, 1974, pp. 125–144, 176–201; MR0443083 (56 #1456)

[6] P. H. Baxendale, Moment stability and large deviations for linear stochastic differential equations, *Probabilistic Methods in Mathematical Physics*, Proc. 1985 Katata/Kyoto conf., ed. K. Itô and N. Ikeda, Academic Press, 1987, pp. 31–54; MR0933817 (89c:60068).

[7] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992, pp. 103–126, 148–152, 232–241, 396–399, 540–548; MR1214374 (94b:60069).

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, pp. 374–377; MR1225604 (94b:00012).

[9] D. Talay, Approximation of upper Lyapunov exponents of bilinear stochastic differential systems, *SIAM J. Numer. Anal.* 28 (1991) 1141–1164; MR1111458 (92i:60119).

[10] S. T. Ariaratnam and W. C. Xie, Lyapunov exponent and rotation number of a two-dimensional nilpotent stochastic system, *Dynam. Stability Systems* 5 (1990) 1–9; MR1057870 (91h:60086).
[11] W. C. Xie, Lyapunov exponents and moment Lyapunov exponents of a two-dimensional near-nilpotent system, *Trans. ASME J. Appl. Mech.* 68 (2001) 453–461; MR1836484 (2002b:74026).
[12] P. Imkeller and C. Lederer, An explicit description of the Lyapunov exponents of the noisy damped harmonic oscillator, *Dynam. Stability Systems* 14 (1999) 385–405; MR1746113 (2000m:34122).
[13] P. Imkeller and C. Lederer, Some formulas for Lyapunov exponents and rotation numbers in two dimensions and the stability of the harmonic oscillator and the inverted pendulum, *Dyn. Syst.* 16 (2001) 29–61; MR1835906 (2002j:60100).
[14] B. Derrida and E. Gardner, Lyapounov exponent of the one-dimensional Anderson model: weak disorder expansions, *J. Physique* 45 (1984) 1283–1295; MR0763431 (85m:82098).
[15] A. Leizarowitz, Exact results for the Lyapunov exponents of certain linear Ito systems, *SIAM J. Appl. Math.* 50 (1990) 1156–1165; MR1053924 (91j:93120).
[16] S. R. Finch, CLT variance associated with Baxendale's SDE, arXiv:0809.5274.

4.16 Lyapunov Exponents. IV

We are interested in the effects of multiplicative noise (continuing our study [1]). Let E_n denote matrix $N(0,1)$ white noise, that is, $E_1, E_2, E_3, \ldots$ is a sequence of independent $m \times m$ matrices and all m^2 entries of E_n, for each n, are independent standard normal variables. Cohen & Newman [2] proved that the recurrence

$$X_n = E_n X_{n-1}, \quad X_0 \neq 0 \text{ arbitrary}$$

gives rise to Lyapunov exponent

$$\tfrac{1}{n} \ln |X_n| \to \tfrac{1}{2}\left(\ln(2) + \psi(\tfrac{m}{2})\right) \quad \text{almost surely as } n \to \infty,$$

where $\psi(x)$ is the digamma function and $\gamma = -\psi(1)$ is the Euler–Mascheroni constant [3]. In particular, for $m = 1$,

$$x_n = \varepsilon_n x_{n-1}$$

has Lyapunov exponent $\lambda = -(\ln(2) + \gamma)/2$ and the following Central Limit Theorem holds:

$$\frac{\ln |x_n| - n\,\lambda}{\pi\sqrt{n/8}} \to N(0,1) \quad \text{as } n \to \infty;$$

for $m = 2$,

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \varepsilon_n & \varepsilon_n' \\ \varepsilon_n'' & \varepsilon_n''' \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}$$

has Lyapunov exponent $\lambda = (\ln(2) - \gamma)/2$ and

$$\frac{\ln\sqrt{x_n^2 + y_n^2} - n\,\lambda}{\pi\sqrt{n/24}} \to N(0,1) \quad \text{as } n \to \infty.$$

Upon constraining certain entries of E_n, relevant Lyapunov exponent calculations become more complicated. Wright & Trefethen [4] found that $\lambda = \ln(1.0574735537...)$ when

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \varepsilon_{n+1} & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix},$$

$\lambda = \ln(1.1149200917...)$ when

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & \varepsilon_{n+1} \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix},$$

and $\lambda = \ln(0.9949018837...)$ when

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \varepsilon'_{n+1} & \varepsilon_{n+1} \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix}.$$

Upon replacing standard normal variables ε_n by symmetric Bernoulli variables

$$\mathrm{P}\left(\varepsilon_n = 1\right) = \mathrm{P}\left(\varepsilon_n = -1\right) = 1/2,$$

the three preceding examples no longer possess distinct Lyapunov exponents. Viswanath [5, 6] proved that the three **random Fibonacci sequences** each have $\lambda = v$, where

$$v = \ln(1.1319882487...) = 0.1239755988...$$

was computed via a fractal invariance measure on the Stern-Brocot division of the real line. A high-precision estimate of v, due to Bai [7], was based on the cycle expansion method applied to a corresponding Ruelle dynamical zeta function [8–10]. It is interesting to compare the “almost-sure growth rate”

$$\tfrac{1}{n}\mathrm{E}\left(\ln|x_n|\right) \to v = \ln(1.1319882487...)$$

against the “average growth rate” [11, 12]

$$\tfrac{1}{n}\ln\left(\mathrm{E}\,|x_n|\right) \to \ln(\xi) = \ln(1.2055694304...)$$

where ξ has minimal polynomial $\xi^3 + \xi^2 - \xi - 2$. The latter value is larger due to outlying sequences that occur with very small probability. It is difficult to detect the difference experimentally since [13]

$$\tfrac{1}{n}\ln\left(\mathrm{Var}\,|x_n|\right) \to \ln(1 + \sqrt{5})$$

and hence $\sim (1 + \sqrt{5})^n$ datapoints are needed to estimate $\mathrm{E}\,|x_n|$ adequately.

Embree & Trefethen [14] examined the more general linear recurrence

$$x_{n+1} = x_n + \beta\,\varepsilon_{n+1}x_{n-1}$$

and determined that the critical threshold β^* (below which solutions decay exponentially almost surely; above which solutions grow exponentially almost surely) is $\beta^* = 0.70258....$ It also appears that the value $\tilde{\beta}$ corresponding to maximal decay is $\tilde{\beta} = 0.36747...$ with Lyapunov exponent $\ln(0.8951...)$.

Chassaing, Letac & Mora [15] examined a different kind of random Fibonacci sequence:

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{cases} \begin{pmatrix} x_{n-1} + y_{n-1} \\ y_{n-1} \end{pmatrix} & \text{with probability } 1/2, \\ \begin{pmatrix} x_{n-1} \\ x_{n-1} + y_{n-1} \end{pmatrix} & \text{with probability } 1/2, \end{cases}$$

which reduces to the study of random products of the two nonnegative matrices:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

Bai [16] computed that $\lambda = \ln(1.4861851938...) = 0.3962125642....$ Let $\varphi = (1 + \sqrt{5})/2$ denote the Golden mean [17]. Another variation is the random sequence:

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{cases} \begin{pmatrix} x_{n-1} + y_{n-1} \\ x_{n-1} \end{pmatrix} & \text{with probability } \varphi - 1 \approx 0.62, \\ \begin{pmatrix} y_{n-1} \\ x_{n-1} + y_{n-1} \end{pmatrix} & \text{with probability } 2 - \varphi \approx 0.38 \end{cases}$$

with associated nonnegative matrices:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

In this case, λ turns out to be $2v/(\varphi - 1)$, which constitutes another occurrence of Viswanath's constant [7].

Fix $\alpha > 0$. Chassaing, Letac & Mora [15, 18] proved that

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & \varepsilon_n \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix}$$

has Lyapunov exponent

$$\lambda = \frac{K_0(\alpha)}{\alpha K_1(\alpha)},$$

where ε_n is distributed according to $\mathrm{Exp}(\alpha/2)$ and K_0, K_1 are modified Bessel functions [19]. If $\alpha = 2$, then $2\lambda = K_0(2)/K_1(2) = 0.8143077587....$ A related ratio $I_1(2)/I_0(2)$ appears in [20]; see also [1].

Lyons [21, 22] studied

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & \varepsilon_n \\ 1 & 1 + \varepsilon_n \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix},$$

where $\varepsilon_n = 0$ with probability $1/2$ and $\varepsilon_n = \tau$ otherwise. It turns out that $\tau \mapsto \lambda(\tau)$ is a strictly increasing function of $\tau > 0$. An important threshold value

$\tau = 0.2688513727...$ is the solution of the equation [16]

$$2\lambda(\tau) = \ln(2)$$

and is connected with the distribution of certain random continued fractions.

Ishii [23, 24] proved that

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & c - \varepsilon_n \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix}$$

has Lyapunov exponent

$$\lambda(c) = \operatorname{arccosh}\left(\frac{\sqrt{(2+c)^2+\delta^2} + \sqrt{(2-c)^2+\delta^2}}{4}\right)$$

where ε_n is distributed according to Cauchy(δ). If instead ε_n follows a Unif$(-\sqrt{3}\sigma, \sqrt{3}\sigma)$ distribution or a $N(0, \sigma^2)$ distribution, then asymptotic results of Derrida & Gardner [25, 26] apply:

$$\lim_{\sigma \to 0^+} \frac{\lambda(c,\sigma)}{\sigma^{2/3}} = \frac{6^{1/3}\sqrt{\pi}}{2\Gamma(1/6)} = 0.2893082598... \quad \text{if } c = 2,$$

$$\lim_{\sigma \to 0^+} \frac{\lambda(c,\sigma)}{\sigma^2} = \begin{cases} 1/6 & \text{if } c = 1, \\ \dfrac{\Gamma(3/4)^2}{\Gamma(1/4)^2} = 0.1142366452... = \dfrac{12}{105.0451015308...} & \text{if } c = 0. \end{cases}$$

The constants 0.2893082598... and 0.1142366452... also appear in [27, 28], respectively, but reasons for these connections are unclear.

Fix an odd integer $k \geq 3$. Pincus [29, 30] and Lima & Rahibe [31] examined

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{cases} \begin{pmatrix} \cos(\frac{\pi}{k})x_{n-1} + \sin(\frac{\pi}{k})y_{n-1} \\ -\sin(\frac{\pi}{k})x_{n-1} + \cos(\frac{\pi}{k})y_{n-1} \end{pmatrix} & \text{with probability } 1-\eta, \\ \begin{pmatrix} x_{n-1} \\ 0 \end{pmatrix} & \text{with probability } \eta \end{cases}$$

and proved that

$$\lambda(k) = \frac{\eta^2}{1-(1-\eta)^{2k}} \sum_{j=1}^{2k-1} (1-\eta)^j \ln\left|\cos\left(\frac{j\pi}{k}\right)\right|.$$

The identical expression emerges if we replace the definition of the latter portion by

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \ell\, x_{n-1} \\ (1/\ell) y_{n-1} \end{pmatrix} \quad \text{with probability } \eta$$

for a fixed integer $\ell \geq 2$, and compute the asymptotic difference between $\lambda(k,\ell)$ and $\eta\ln(\ell)$ in the limit as $\ell \to \infty$. A precise numerical estimate of $\lambda(3,2) = 0.1794...$, however, is evidently open [16].

Ben-Naim & Krapivsky [32] studied two variations of random Fibonacci sequences:

$$x_n = \begin{cases} x_{n-1} + x_{n-2} & \text{with probability } 1-\eta \\ x_{n-1} + x_{n-3} & \text{with probability } \eta \end{cases}, \quad x_0 = 0, \quad x_1 = x_2 = 1;$$

$$x_n = \begin{cases} x_{n-1} + x_{n-2} & \text{with probability } 1-\eta \\ 2x_{n-1} & \text{with probability } \eta \end{cases}, \quad x_1 = x_2 = 1$$

and determined that

$$\lim_{\eta\to 0^+} \lambda(\eta) = \ln(\varphi)$$

for both cases. Second-order asymptotic terms differ, however:

$$\lim_{\eta\to 0^+} \frac{\lambda(\eta)-\ln(\varphi)}{\eta} = \begin{cases} \ln\left(\dfrac{2\varphi}{\varphi+2}\right) & \text{for case 1,} \\ \ln\left(\dfrac{2\varphi+1}{\varphi+2}\right) & \text{for case 2} \end{cases}$$

and a third-order term is possible for the latter.

Consider the random geometric sequence [33]

$$x_n = 2x_p, \quad x_0 = 1, \quad p \in \{0, 1, \ldots, n-1\}$$

where each of the n possible indices is given equal weight. The sequence is not necessarily increasing, but enjoys average growth $n+1$ and almost-sure growth

$$2^{\gamma} n^{\ln(2)} = (1.4919670404...)\exp(\ln(2)\ln(n)).$$

Consider instead two additional random Fibonacci models [34, 35]:

$$x_n = x_{n-1} + x_q, \quad x_0 = 1, \quad q \in \{0, 1, \ldots, n-1\};$$

$$x_n = x_p + x_q, \quad x_0 = 1, \quad p, q \in \{0, 1, \ldots, n-1\}.$$

Model 1 enjoys average growth

$$\frac{1}{2\sqrt{e\pi}} n^{-1/4}\exp(2\sqrt{n})$$

and almost-sure growth

$$C\exp\big((1.889...)\sqrt{n}\big)$$

where $C > 0$ is unknown. Model 2 is not necessarily increasing but enjoys average growth $n+1$; unlike the random geometric sequence, it seems not to display almost-sure behavior of any kind.

Kenyon & Peres [36] studied random products associated with two sets of matrices:

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

and

$$\begin{pmatrix}3 & 0\\ 2 & 0\end{pmatrix}, \quad \begin{pmatrix}2 & 0\\ 1 & 1\end{pmatrix}, \quad \begin{pmatrix}1 & 1\\ 0 & 2\end{pmatrix}, \quad \begin{pmatrix}0 & 2\\ 0 & 3\end{pmatrix}.$$

The three matrices in the first set are equiprobable, with Lyapunov exponent $\ln(2)/3 = 0.2310490601\ldots$. The four matrices in the second set are likewise equiprobable, with Lyapunov exponent [37]

$$\frac{1}{6}\ln\left(\frac{2}{3}\right) + \sum_{i=0}^{\infty} 4^{-i-1}\ln\left(\frac{(3\cdot 2^i)!}{(2^{i+1})!}\right) = 0.7974350484\ldots.$$

We wonder whether $\exp(0.7974350484\ldots)$ is transcendental. Moshe [38] studied random products associated with two equiprobable 3×3 matrices:

$$\begin{pmatrix}1 & 3 & 1\\ 1 & 2 & 0\\ -3 & -6 & 0\end{pmatrix}, \quad \begin{pmatrix}4 & 2 & 8\\ -2 & -1 & -4\\ 3 & 1 & 4\end{pmatrix}$$

and computed Lyapunov exponent

$$\frac{1}{16}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{1}{2^{j+k}}\ln\left|3\cdot 2^{3j} - 2(-1)^j - \tfrac{22}{9}2^{3j+k} + \tfrac{22}{9}(-1)^j 2^k\right| = 0.5897925607\ldots.$$

Many more similar examples are found in [39–42].

Up to now, the random mechanisms underlying sequences have been very simple. Here is a more complicated but well-known example [43, 44]:

$$x_{n+1} = a_n x_n + x_{n-1}, \quad x_0 = 0, \quad x_1 = 1$$

where the cofficients a_n are obtained by selecting a random $\theta \in [0,1]$ and computing its continued fraction digits:

$$\theta = \frac{1|}{|a_1} + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots.$$

For instance, if $\theta = \pi - 3$, then

$$\{a_1, a_2, a_3, a_4\} = \{7, 15, 1, 292\}, \quad \{x_2, x_3, x_4, x_5\} = \{7, 106, 113, 33102\};$$

note that x_n is simply the denominator of the n^{th} partial convergent to θ. Lévy [45] proved that this recurrence gives rise to Lyapunov exponent

$$\frac{\pi^2}{12\ln(2)} = 1.1865691104\ldots.$$

Another example involves the recurrence [46]

$$x_{n+1} = 2^{b_n}x_n + 2^{b_{n-1}}x_{n-1}, \quad x_0 = 0, \quad x_1 = 1$$

where the cofficients b_n are obtained via

$$\theta = \frac{2^{-b_1}|}{|1} + \frac{2^{-b_2}|}{|1} + \frac{2^{-b_3}|}{|1} + \cdots.$$

The corresponding Lyapunov exponent is

$$\frac{1}{\ln(4/3)}\left(\frac{\pi^2}{12} + \mathrm{Li}_2\left(-\frac{1}{2}\right)\right) = 1.3002298798...$$

where $\mathrm{Li}_2(y)$ is the dilogarithm function [47]. (This constant also appears in [48] without explanation.) Generalization to base $k \geq 2$ is possible, as well as formulation for Khintchine-type and Lochs-type constants in this broad setting.

Addendum The subject continues to expand [49–55]. Two earlier works deserve mention. Hope [56] examined

$$x_{n+1} = a_n x_n + x_{n-1}, \quad x_0 = 0, \quad x_1 = 1$$

like Lévy, but with a simple rule

$$\mathrm{P}\left(a_n = 1\right) = \mathrm{P}\left(a_n = 2\right) = 1/2$$

and independence assumed. The Lyapunov exponent is

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{a=1 \text{ or } 2} \ln\left(a_1 + \frac{1|}{|a_2} + \frac{1|}{|a_3} + \cdots + \frac{1|}{|a_{n-1}} + \frac{1|}{|a_n}\right) \approx 0.673 \approx \ln(1.96).$$

Davison [57] studied the same except with the rule

$$a_n = 1 + (\lfloor \theta\, n \rfloor \bmod 2)$$

for a random $\theta \in [0, 2]$, showing that

$$1.931 < \sqrt{2+\sqrt{3}} \leq \liminf_{n\to\infty} x_n^{1/n} \leq \limsup_{n\to\infty} x_n^{1/n} \leq \sqrt{\left(\frac{1+\sqrt{5}}{2}\right)\left(1+\sqrt{2}\right)} < 1.977.$$

We wonder how closely these examples might be connected.

The sequence of polynomials giving Pascal's rhombus [39] arises from a second-order recurrence

$$p_n(x) = (1 + x + x^2)p_{n-1}(x) + x^2 p_{n-2}(x), \quad p_1(x) = 1 + x + x^2, \quad p_0(x) = 1.$$

Let u_n to be the number of odd coefficients in $p_n(x)$. A numerical method gives "typical growth" $\lambda = 0.57331379313....$ While $\limsup_{n\to\infty} \ln(u_n)/\ln(n) = 1$ is trivial, the following was proved only recently [58]:

$$\begin{aligned}\liminf_{n\to\infty} \frac{\ln(u_n)}{\ln(n)} &= \rho\left(A^3B^3\right)^{1/6} \\ &= \frac{\ln(1.6376300574...)}{\ln(2)} = 0.7116094872...\end{aligned}$$

where A, B are known 5×5 integer matrices and ρ denotes spectral radius (the maximal modulus of eigenvalues). Consider instead the Fibonacci polynomials [39]

$$q_n(x) = x\, q_{n-1}(x) + q_{n-2}(x), \quad q_1(x) = x, \quad q_0(x) = 1.$$

The number v_n of odd coefficients in $q_n(x)$ is the n^{th} term of Stern's sequence [59]:

$$v_{2n+1} = v_n, \quad v_{2n} = v_n + v_{n-1}.$$

Again, $\lambda = 0.3962125642...$ via numerics; "typical dispersion" $\sigma^2 = 0.0221729451...$ can be found similarly [40]. The limit superior and limit inferior do not present any difficulties for $\{v_n\}$. An evaluation of σ^2 corresponding to $\{u_n\}$, however, remains open.

[1] S. R. Finch, Lyapunov exponents. III, *this volume*, §4.15.
[2] J. E. Cohen and C. M. Newman, The stability of large random matrices and their products, *Annals of Probab.* 12 (1984) 283–310; MR0735839 (86a:60013).
[3] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[4] T. G. Wright and L. N. Trefethen, Computing Lyapunov constants for random recurrences with smooth coefficients, *J. Comput. Appl. Math.* 132 (2001) 331–340; MR1840632 (2002d:65010).
[5] D. Viswanath, Random Fibonacci sequences and the number 1.13198824..., *Math. Comp.* 69 (2000) 1131–1155; MR1654010 (2000j:15040).
[6] B. Hayes, The Vibonacci numbers, *Amer. Scientist,* v. 87 (1999) n. 4, 296-301.
[7] Z.-Q. Bai, On the cycle expansion for the Lyapunov exponent of a product of random matrices, *J. Phys. A* 40 (2007) 8315–8328; MR2371235 (2009b:82033).
[8] R. Mainieri, Zeta function for the Lyapunov exponent of a product of random matrices, *Phys. Rev. Lett.* 68 (1992) 1965–1968; chao-dyn/9301001.
[9] R. Mainieri, Cycle expansion for the Lyapunov exponent of a product of random matrices, *Chaos* 2 (1992) 91–97; MR1158540 (93e:82029).
[10] J. L. Nielsen, Lyapunov exponent for products of random matrices, unpublished project (1997), *Chaos: Classical and Quantum*, Niels Bohr Institute.
[11] B. Rittaud, On the average growth of random Fibonacci sequences, *J. Integer Seq.* 10 (2007) 07.2.4; MR2276788 (2007j:11018).
[12] É. Janvresse, B. Rittaud and T. de la Rue, How do random Fibonacci sequences grow?, *Probab. Theory Related Fields* 142 (2008) 619–648; math.PR/0611860; MR2438703 (2009m:37148).
[13] E. Makover and J. McGowan, An elementary proof that random Fibonacci sequences grow exponentially, *J. Number Theory* 121 (2006) 40–44; math.NT/0510159; MR2268754 (2008a:11087).
[14] M. Embree and L. N. Trefethen, Growth and decay of random Fibonacci sequences, *Royal Soc. Lond. Proc. Ser. A* 455 (1999) 2471–2485; MR1807827 (2001i:11098).
[15] P. Chassaing, G. Letac and M. Mora, Brocot sequences and random walks in $\mathrm{SL}(2,\mathbb{R})$, *Probability Measures on Groups. VII*, Proc. 1983 Oberwolfach conf., ed. H. Heyer, Lect. Notes in Math. 1064, Springer-Verlag, 1984, pp. 36–48; MR0772400 (86g:60012).
[16] Z.-Q. Bai, Calculations of certain Lyapunov exponents, unpublished note (2007).
[17] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[18] D. Mannion, Products of 2×2 random matrices, *Annals Appl. Probab.* 3 (1993) 1189–1218; MR1241041 (94k:60021).

[19] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, pp. 374–377; MR1225604 (94b:00012).

[20] S. R. Finch, Euler-Gompertz constant, *first volume*, pp. 423–428.

[21] R. Lyons, Singularity of some random continued fractions, *J. Theoret. Probab.* 13 (2000) 535–545; MR1778585 (2002c:60138).

[22] K. Simon, B. Solomyak and M. Urbański, Invariant measures for parabolic IFS with overlaps and random continued fractions, *Trans. Amer. Math. Soc.* 353 (2001) 5145–5164; MR1852098 (2003c:37030a).

[23] K. Ishii, Localization of eigenstates and transport phenomena in the one-dimensional disordered system, *Prog. Theor. Phys. Suppl.* 53 (1973) 77–138.

[24] I. M. Lifshits, S. A. Gredeskul and L. A. Pastur, *Introduction to the Theory of Disordered Systems*, Wiley, 1988, pp. 124–163; MR1042095 (90k:82073).

[25] B. Derrida and E. Gardner, Lyapounov exponent of the one-dimensional Anderson model: weak disorder expansions, *J. Physique* 45 (1984) 1283–1295; MR0763431 (85m:82098).

[26] F. M. Izrailev, S. Ruffo and L. Tessieri, Classical representation of the one-dimensional Anderson model, *J. Phys. A* 31 (1998) 5263–5270; MR1634869 (99c:82037).

[27] S. T. Ariaratnam and W. C. Xie, Lyapunov exponent and rotation number of a two-dimensional nilpotent stochastic system, *Dynam. Stability Systems* 5 (1990) 1–9; MR1057870 (91h:60086).

[28] C. Sire and P. L. Krapivsky, Random Fibonacci sequences, *J. Phys. A* 34 (2001) 9065–9083; cond-mat/0106457; MR1876126 (2003b:60018).

[29] S. Pincus, Furstenberg-Kesten results: asymptotic analysis, *Random Matrices and their Applications*, Proc. 1984 Brunswick conf., ed. J. E. Cohen, H. Kesten and C. M. Newman, Amer. Math. Soc., 1986, pp. 79–86; MR0841083 (87m:60069).

[30] S. Pincus, Strong laws of large numbers for products of random matrices, *Trans. Amer. Math. Soc.* 287 (1985) 65–89; MR0766207 (86i:60087).

[31] R. Lima and M. Rahibe, Exact Lyapunov exponent for infinite products of random matrices, *J. Phys. A* 27 (1994) 3427–3437; MR1282183 (95d:82004).

[32] E. Ben-Naim and P. L. Krapivsky, Weak disorder in Fibonacci sequences, *J. Phys. A* 39 (2006) L301–L307; cond-mat/0603117; MR2238100 (2007c:82047).

[33] E. Ben-Naim and P. L. Krapivsky, Random geometric series, *J. Phys. A* 37 (2004) 5949–5957; cond-mat/0403157; MR2074617 (2005i:11098).

[34] E. Ben-Naim and P. L. Krapivsky, Growth and structure of stochastic sequences, *J. Phys. A* 35 (2002) L557–L563; cond-mat/0208072; MR1946873.

[35] I. Krasikov, G. J. Rodgers and C. E. Tripp, Growing random sequences. *J. Phys. A* 37 (2004) 2365–2370; MR2045930 (2005k:60167).

[36] R. Kenyon and Y. Peres, Intersecting random translates of invariant Cantor sets, *Invent. Math.* 104 (1991) 601–629; MR1106751 (92g:28018).

[37] P. Sebah, Series evaluation via Kummer acceleration, Stirling approximation and Euler-Maclaurin summation, unpublished note (2007).

[38] Y. Moshe, Random matrix products and applications to cellular automata, *J. d'Analyse Math.* 99 (2006) 267–294; MR2279553 (2008d:37090).

[39] S. Finch, P. Sebah and Z.-Q. Bai, Odd entries in Pascal's trinomial triangle, arXiv:0802.2654.

[40] S. Finch, Z.-Q. Bai and P. Sebah, Typical dispersion and generalized Lyapunov exponents, arXiv:0803.2611.

[41] M. Pollicott, Maximal Lyapunov exponents for random matrix products, *Invent. Math.* 181 (2010) 209–226; MR2651384 (2011g:37146).

[42] V. Yu. Protasov and R. M. Jungers, Lower and upper bounds for the largest Lyapunov exponent of matrices, *Linear Algebra Appl.* 438 (2013) 4448–4468; MR3034543.
[43] K. Dajani and C. Kraaikamp, *Ergodic Theory of Numbers*, Math. Assoc. Amer., 2002, pp. 20–31; MR1917322 (2003f:37014).
[44] S. R. Finch, Continued fraction transformation, *this volume*, §4.9.
[45] H.-C. Chan, The asymptotic growth rate of random Fibonacci type sequences, *Fibonacci Quart.* 43 (2005) 243–255; MR2171636 (2006f:11011).
[46] H.-C. Chan, The asymptotic growth rate of random Fibonacci type sequences. II, *Fibonacci Quart.* 44 (2006) 73–84; MR2209557 (2006m:11014).
[47] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[48] S. R. Finch, Continued fraction transformation. IV, *this volume*, §4.12.
[49] É. Janvresse, B. Rittaud and T. de la Rue, Growth rate for the expected value of a generalized random Fibonacci sequence, *J. Phys. A* 42 (2009) 085005; MR2525481 (2010e:11011).
[50] É. Janvresse, B. Rittaud and T. de la Rue, Almost-sure growth rate of generalized random Fibonacci sequences, *Annales de l'Institut Henri Poincaré Probab. Stat.* 46 (2010) 135–158; arXiv:0804.2400; MR2641774 (2012a:37106).
[51] E. B. Cureg and A. Mukherjea, Numerical results on some generalized random Fibonacci sequences, *Comput. Math. Appl.* 59 (2010) 233–246; MR2575510 (2010j:11024).
[52] Z.-Q. Bai, A transfer operator approach to random Fibonacci sequences, *J. Phys. A* 44 (2011) 115002; MR2773858 (2012a:37102).
[53] Y. Lan, Novel computation of the growth rate of generalized random Fibonacci sequences, *J. Stat. Phys.* 142 (2011) 847–861; MR2773789 (2012b:82056).
[54] C. Zhang and Y. Lan, Computation of growth rates of random sequences with multi-step memory, *J. Stat. Phys.* 150 (2013) 722–743; MR3024154.
[55] P. J. Forrester, Lyapunov exponents for products of complex Gaussian random matrices, *J. Stat. Phys.* 151 (2013) 796–808; arXiv:1206.2001; MR3055376.
[56] P. Hope, Exponential growth of random Fibonacci sequences, *Fibonacci Quart.* 33 (1995) 164–168; MR1329024 (96b:11115).
[57] J. L. Davison, A class of transcendental numbers with bounded partial quotients, *Number Theory and Applications*, Proc. 1988 Banff NATO Adv. Study Instit., ed. R. A. Mollin, Kluwer, 1989, pp. 365–371; MR1123082 (92h:11060).
[58] N. Guglielmi and V. Yu. Protasov, Exact computation of joint spectral characteristics of linear operators, *Found. Comput. Math.* 13 (2013) 37–97; MR3009529.
[59] S. R. Finch, Stolarsky-Harborth constant, *first volume*, pp. 145–151.

4.17 Stars and Watermelons

The p-**vicious walker model** of length $2n$ consists of p lattice paths W_1, W_2, ..., W_p in $\mathbb{Z}^2$ where

- W_k starts at the point $(0, a_k)$ and ends at the point $(2n, b_k)$ for $k = 1, \ldots, p$
- all steps are directed northeast or southeast (that is, from (i,j) to $(i+1, j+1)$ or to $(i+1, j-1)$)
- if $k \neq \ell$, then W_k and W_ℓ never intersect (hence $a_k \neq a_\ell$ and $b_k \neq b_\ell$, for instance).

In p-**star** configurations, $a_k = 2k - 2$ for each k (with no constraint on b_k); in p-**watermelon** configurations, $b_k = 2k - 2$ as well [1, 2]. We often think of the horizontal axis as time and the vertical axis as space, writing $W_k(0) = a_k$ and $W_k(2n) = b_k$. A p-watermelon with a **wall** has the additional property that

- $W_k(i) \geq 0$ for all $0 \leq i \leq 2n$, for all k.

Gillet [3] demonstrated that $\lim_{n\to\infty} W_k(\lfloor 2n\,t \rfloor)/\sqrt{2n}$ tends to a family of p non-intersecting Brownian excursions, $0 \leq t \leq 1$, as an extension of a principle given in [4].

The **height** of a path W_k in a p-watermelon with wall is the maximum value of $W_k(i)$ over all i. The **area** under a path W_k is the area of the polygonal region determined by the curve $j = W_k(i)$, the horizontal line $j = 0$, and the vertical lines $i = 0$, $i = 2n$. In the case $p = 2$, we will refer to the upper height and upper area (corresponding to W_2) and the lower height and lower area (corresponding to W_1).

Counting all 1-watermelons with wall (or Dyck paths) and 2-watermelons with wall give

$$\frac{(2n)!}{n!(n+1)!}, \quad \frac{6(2n)!(2n+2)!}{n!(n+1)!(n+2)!(n+3)!}$$

possible configurations of length $2n$, respectively. (The former is the n^{th} Catalan number.) The average height $H_1(n)$ for 1-watermelons with wall satisfies [5, 6]

$$H_1(n) \sim \sqrt{\pi n}$$

as $n \to \infty$ and the average area $A_1(n)$ satisfies [7, 8]

$$A_1(n) \sim \sqrt{\pi} n^{3/2}.$$

To go to the average L_∞-norm of Brownian excursion, divide the H_1 result by $\sqrt{2n}$ (space dimension only), yielding $\sqrt{\pi/2}$. To go to the average L_1-norm, divide the A_1 result by $(2n)^{3/2}$ (both time and space considered), yielding $\sqrt{\pi/8}$. Exact formulas for $H_1(n)$ and $A_1(n)$ are also available [9].

The average upper height $H_2(n)$ for 2-watermelons with wall satisfies

$$H_2(n) \sim (2.57758...)\sqrt{n} \sim (1.822625...)\sqrt{2n},$$

a new result due to Fulmek [6]. The coefficient can be expressed as a linear combination of several complicated integrals of theta functions; a certain double Dirichlet series also plays a role in the proof. Numerical results for $3 \leq p \leq 5$ and for higher moments were obtained by Feierl [10]. A different method was proposed in [11]. To go to the average upper L_∞-norm of Brownian excursion, divide the H_2 result by $\sqrt{2n}$. An exact formula for $H_2(n)$ is also available [12]. Similar information about the lower height is not known.

An exact formula for $A_2(n)$ seems to be an open problem. Interestingly, we have both average upper/lower L_1-norm results for Brownian excursion:

$$\tfrac{5}{8}\left(\sqrt{2}-1\right)\sqrt{\pi}, \quad \tfrac{5}{8}\sqrt{\pi}$$

due to Tracy & Widom [13]. Multiplying each constant by $(2n)^{3/2}$ therefore provides the main asymptotic terms for average upper/lower areas under 2-watermelons with wall. Numerical results in [13] also apply for $3 \le p \le 9$. In a study of average upper L_1-norms as $p \to \infty$, the constant 1.7710868074... arises [14, 15] and thus random matrix theory lurks nearby.

Counting all 1-watermelons without wall (or bilateral Dyck paths) and 2-watermelons without wall give [16]

$$\frac{(2n)!}{(n!)^2}, \quad \frac{(2n)!(2n+1)!}{(n!)^2((n+1)!)^2}$$

possible configurations of length $2n$, respectively. (The former is the n^{th} central binomial coefficient.) These tend to Brownian bridges as $n \to \infty$ [3, 17]. In the same way, p-stars with wall tend to Brownian meanders and p-stars without wall tend to Brownian motions. Corresponding questions about average heights and average areas (suitably generalized) for $p \ge 2$ seem to be unanswered.

[1] N. Bonichon and M. Mosbah, Watermelon uniform random generation with applications, *Theoret. Comput. Sci.* 307 (2003) 241–256; MR2022577 (2004m:05032).
[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A078920 and A103905.
[3] F. Gillet, Asymptotic behaviour of watermelons, math.PR/0307204.
[4] S. R. Finch, Variants of Brownian motion, *this volume*, §4.5.
[5] N. G. de Bruijn, D. E. Knuth and S. O. Rice, The average height of planted plane trees, *Graph Theory and Computing*, ed. R. C. Read, Academic Press, 1972, pp. 15–22; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 215–223; MR0505710 (58 #21737).
[6] M. Fulmek, Asymptotics of the average height of 2-watermelons with a wall, *Elec. J. Combin.* 14 (2007) R64; math.CO/0607163; MR2350454 (2008j:05031).
[7] D. Merlini, R. Sprugnoli and M. C. Verri, The area determined by underdiagonal lattice paths, *Proc. 1996 Colloq. on Trees in Algebra and Programming (CAAP)*, Linköping, ed. H. Kirchner, Lect. Notes in Comp. Sci. 1059, Springer-Verlag, 1996, pp. 59–71; MR1415900 (97f:68140).
[8] R. Chapman, Moments of Dyck paths, *Discrete Math.* 204 (1999) 113–117; MR1691864 (2000g:05011).
[9] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000108, A008549 and A136439.
[10] T. Feierl, The height of watermelons with wall, *J. Phys. A* 45 (2012) 095003; arXiv:0802.2691; MR2897031.
[11] M. Katori, M. Izumi and N. Kobayashi, Two Bessel bridges conditioned never to collide, double Dirichlet series, and Jacobi theta function, *J. Stat. Phys.* 131 (2008) 1067–1083; arXiv:0711.1710; MR2407380 (2009e:60184).
[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005700 and A136440.

[13] C. A. Tracy and H. Widom, Nonintersecting Brownian excursions, *Annals Appl. Probab.* 17 (2007) 953–979; math.PR/0607321; MR2326237 (2008j:60238).
[14] S. R. Finch, Longest subsequence constants, *first volume*, pp. 382–387.
[15] S. R. Finch, Hammersley's path process, *this volume*, §4.1.
[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000984 and A000891.
[17] T. Feierl, The height and range of watermelons without wall, *European J. Combin.* 34 (2013) 138–154; arXiv:0806.0037; MR2974277.

4.18 Prophet Inequalities

Suppose that you view a sequence X_1, X_2, ..., X_n of independent identically distributed nonnegative random variables and that you wish to stop at a value of X as large as possible. As in [1], revisiting earlier values is not permitted. If you are a prophet (meaning that you have complete foresight), then you know $\max\{X_1, \ldots, X_n\}$ beforehand; let M_n denote the average such "insider information" value. If you are a mortal (meaning that you have no choice but to select an X via stopping rules) and if you proceed optimally, then the value V_n obtained satisfies

$$\frac{M_n}{V_n} \leq 1 + \alpha_n$$

for best constants α_n with $0.1 < \alpha_n < 0.6$. Let us now be more precise [2–7].

Define

$$f_n(w, x) = \frac{n}{n-1} w^{(n-1)/n} + \frac{1}{n-1} x,$$

$$g_{k,n}(x) = \begin{cases} f_n(g_{k-1,n}(x), x) & \text{if } 1 \leq k \leq n, \\ f_n(0, x) & \text{if } k = 0 \end{cases}$$

then α_n is the unique solution of $g_{n-1,n}(x) = 1$, $0 < x < 1$. For example [5, 7],

$$g_{1,2}(x) = \frac{2}{1}\left(\frac{x}{1}\right)^{1/2} + \frac{x}{1},$$

$$g_{2,3}(x) = \frac{3}{2}\left(\frac{3}{2}\left(\frac{x}{2}\right)^{2/3} + \frac{x}{2}\right)^{2/3} + \frac{x}{2},$$

$$g_{3,4}(x) = \frac{4}{3}\left(\frac{4}{3}\left(\frac{4}{3}\left(\frac{x}{3}\right)^{3/4} + \frac{x}{3}\right)^{3/4} + \frac{x}{3}\right)^{3/4} + \frac{x}{3},$$

$$g_{4,5}(x) = \frac{5}{4}\left(\frac{5}{4}\left(\frac{5}{4}\left(\frac{5}{4}\left(\frac{x}{4}\right)^{4/5} + \frac{x}{4}\right)^{4/5} + \frac{x}{4}\right)^{4/5} + \frac{x}{4}\right)^{4/5} + \frac{x}{4}$$

give rise to $\alpha_2 = 0.17157...$, $\alpha_3 = 0.22138...$, $\alpha_4 = 0.24810...$, $\alpha_5 = 0.26495....$ Kertz [6] proved that α_n is strictly increasing and that

$$\alpha_\infty = \lim_{n\to\infty} \alpha_n = 0.3414889923...$$

is the unique solution of

$$\int_0^1 \frac{1}{u - u\ln(u) + x} du = 1.$$

Hence a prophet may never win more, on average, than 1.34... times the winnings of a mortal.

Suppose instead that you view a sequence $X_1, X_2, ..., X_n$ of independent identically distributed random variables taking values only in the interval $[0, 1]$. Everything else is the same. With this additional information, the optimal stopping value V_n now satisfies

$$M_n \le V_n + \beta_n$$

for best constants β_n with $0 < \beta_n < 1/4$. Again, let us be more precise [5, 7, 8]. Define β_n to be the unique solution of

$$(n-1)\left(g_{n,n}(x) - g_{n-1,n}(x)\right) = 1, \quad 0 < x < 1.$$

Sample $g_{n-1,n}(x)$ expressions were given earlier; sample $g_{n,n}(x)$ expressions are [5, 7]

$$g_{2,2}(x) = \frac{2}{1}\left(\frac{2}{1}\left(\frac{x}{1}\right)^{1/2} + \frac{x}{1}\right)^{1/2} + \frac{x}{1},$$

$$g_{3,3}(x) = \frac{3}{2}\left(\frac{3}{2}\left(\frac{3}{2}\left(\frac{x}{2}\right)^{2/3} + \frac{x}{2}\right)^{2/3} + \frac{x}{2}\right)^{2/3} + \frac{x}{2},$$

$$g_{4,4}(x) = \frac{4}{3}\left(\frac{4}{3}\left(\frac{4}{3}\left(\frac{4}{3}\left(\frac{x}{3}\right)^{3/4} + \frac{x}{3}\right)^{3/4} + \frac{x}{3}\right)^{3/4} + \frac{x}{3}\right)^{3/4} + \frac{x}{3}$$

and give rise to $\beta_2 = 1/16$, $\beta_3 = 0.07761...$, $\beta_4 = 0.08538....$ It seems likely that β_n is strictly increasing, but a proof that

$$\beta_\infty = \lim_{n\to\infty} \beta_n \approx 0.1113$$

exists is open. A high-precision estimate of β_∞ is also desired.

The nested radical expressions for $g_{k,n}(x)$ deserve more study. A helpful survey on general prophet inequalities [9] is recommended.

[1] S. R. Finch, Optimal stopping constants, *first volume*, pp. 361–363.

[2] U. Krengel and L. Sucheston, Semiamarts and finite values, *Bull. Amer. Math. Soc.* 83 (1977) 745–747; MR0436314 (55 #9261).

[3] U. Krengel and L. Sucheston, On semiamarts, amarts, and processes with finite value, *Probability on Banach Spaces*, ed. J. Kuelbs, Dekker, 1978, pp. 197–266; MR0515432 (80g:60053).

[4] T. P. Hill and R. P. Kertz, Ratio comparisons of supremum and stop rule expectations, *Z. Wahrsch. Verw. Gebiete* 56 (1981) 283–285; MR0618276 (82h:60078).
[5] T. P. Hill and R. P. Kertz, Comparisons of stop rule and supremum expectations of i.i.d. random variables, *Annals Probab.* 10 (1982) 336–345; MR0647508 (83g:60053).
[6] R. P. Kertz, Stop rule and supremum expectations of i.i.d. random variables: a complete comparison by conjugate duality, *J. Multivariate Anal.* 19 (1986) 88–112; MR0847575 (87m:60102).
[7] P. C. Allaart, Prophet inequalities for I.I.D. random variables with random arrival times, *Sequential Anal.* 26 (2007) 403–413; math.PR/0611664; MR2359862 (2008h:60150).
[8] T. P. Hill and R. P. Kertz, Additive comparisons of stop rule and supremum expectations of uniformly bounded independent random variables, *Proc. Amer. Math. Soc.* 83 (1981) 582–585; MR0627697 (82j:60071).
[9] T. P. Hill and R. P. Kertz, A survey of prophet inequalities in optimal stopping theory, *Strategies for Sequential Search and Selection in Real Time*, Proc. 1990 Amherst conf., ed. F. T. Bruss, T. S. Ferguson and S. M. Samuels, Amer. Math. Soc., 1992, pp. 191–207; MR1160620 (93g:60089).

4.19 Excursion Durations

This essay bears some resemblance to [1], but comes from a different viewpoint. Let $\{X_t : 0 \leq t \leq 1\}$ denote standard Brownian motion and fix a time $0 < \tau < 1$. The **excursion** straddling τ is $\{X_t : \alpha_\tau \leq t \leq \beta_\tau\}$, where

$$\alpha_\tau = \sup\{t < \tau : X_t = 0\}, \quad \beta_\tau = \inf\{t > \tau : X_t = 0\}.$$

We are interested in the **duration** $\beta_t - \alpha_\tau$ of this excursion, as well as all excursions straddling earlier times. More precisely, let

$$M_\tau - 1 = \#\{\text{excursions completed by time } \tau \text{ whose durations exceed } \tau - \alpha_\tau\},$$

$$N_\tau - 1 = \#\{\text{excursions completed by time } \tau \text{ whose durations exceed } \beta_\tau - \alpha_\tau\};$$

we wish to compute the probability that $M_\tau = 1$ (the current excursion, measured up to time τ, has a record duration) and the probability that $N_\tau = 1$ (the current excursion, measured to its completion, has a record duration). Since $\beta_\tau \geq \tau$, it is clear that $M_\tau \geq N_\tau$. Simple scaling arguments show that the distribution of M_τ and the distribution of N_τ are independent of τ.

Define functions

$$\varphi(x) = \frac{1}{2}\int_1^\infty e^{-xu}u^{-3/2}du = e^{-x} - \sqrt{\pi\, x}\,\operatorname{erfc}(\sqrt{x}),$$

$$\psi(x) = 1 + \frac{1}{2}\int_0^1 (1 - e^{-xu})u^{-3/2}du = e^{-x} + \sqrt{\pi\, x}\,\operatorname{erf}(\sqrt{x})$$

involving the error and complementary error functions [2]; then [3, 4]

$$\mathrm{P}(M_\tau = k) = \int_0^\infty e^{-x}\varphi(x)^{k-1}\psi(x)^{-k}dx,$$

$$\mathrm{P}(N_\tau = k) = \frac{1}{2}\int_0^\infty x^{-1}(1 - e^{-x})\varphi(x)^{k-1}\psi(x)^{-k}dx.$$

Numerical integration gives

$$\mathrm{P}(M_\tau = k) = \begin{cases} 0.6265075987... & \text{if } k = 1, \\ 0.1430092516... & \text{if } k = 2, \\ 0.0630157050... & \text{if } k = 3, \\ 0.0356483608... & \text{if } k = 4, \end{cases}$$

$$\mathrm{P}(N_\tau = k) = \begin{cases} 0.8003100322... & \text{if } k = 1, \\ 0.0812481569... & \text{if } k = 2, \\ 0.0334196946... & \text{if } k = 3, \\ 0.0184590943... & \text{if } k = 4 \end{cases}$$

and asymptotic analysis gives, as $k \to \infty$,

$$\mathrm{P}(M_\tau = k) \sim \frac{2}{\pi k^2}, \quad \mathrm{P}(N_\tau = k) \sim \frac{1}{\pi k^2}.$$

It is striking that the current excursion is, with fairly high probability, of duration greater than all preceding excursions!

Let $L_1 > L_2 > L_3 > \ldots > 0$ denote the ranked durations of excursions of X_t. Note that $\sum L_j = 1$ almost surely. The joint probability law of $(L_1, L_2, L_3, \ldots)$ follows what is called the Poisson–Dirichlet $(1/2, 0)$ distribution. If instead X_t is a Brownian bridge (meaning that $X_1 = 0$), then the Poisson–Dirichlet $(1/2, 1/2)$ distribution emerges. Can numerical results for $\mathrm{P}(M_\tau)$ and $\mathrm{P}(N_\tau)$ be found in this case? We also wonder what happens when X_t is an Ornstein–Uhlenbeck process [5].

The constant 0.6265... appears in [6], as well as the Golomb–Dickman constant 0.6243... [7].

[1] S. R. Finch, Zero crossings, *this volume*, §4.4.

[2] S. R. Finch, Notation: Error function, *this volume*, p. xi.

[3] C. L. Scheffer, The rank of the present excursion, *Stochastic Process. Appl.* 55 (1995) 101–118; MR1312151 (96m:60189).

[4] J. Pitman and M. Yor, The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator, *Annals of Probab.* 25 (1997) 855–900; MR1434129 (98f:60147).

[5] J. Pitman and M. Yor, On the lengths of excursions of some Markov processes, *Séminaire de Probabilités, XXXI*, ed. J. Azéma, M. Emery and M. Yor, Lect. Notes in Math. 1655, Springer, 1997, pp. 272–286; MR1478737 (98j:60108).

[6] S. N. Majumdar and R. M. Ziff, Universal record statistics of random walks and Lévy flights, *Phys. Rev. Lett.* 101 (2008) 050601; arXiv:0806.0057; MR2430298 (2009k:82048).
[7] S. R. Finch, Golomb-Dickman constant, *first volume*, pp. 284–292.

4.20 Gambler's Ruin

Consider two gamblers A, B with initial integer fortunes a, b. Let $m = a + b$ denote the initial sum of fortunes. In each round of a fair game, one player wins and is paid 1 by the other player:

$$(a, b) \mapsto \begin{cases} (a+1, b-1) & \text{with probability } 1/2, \\ (a-1, b+1) & \qquad '' \end{cases}$$

Assume that rounds are independent for the remainder of this essay. The **ruin probability** p_E for a gambler E is the probability that E's fortune reaches 0 before it reaches m. For the symmetric 2-player problem,

$$p_A = \frac{b}{a+b}, \quad p_B = \frac{a}{a+b}$$

and this can be proved using either discrete-time (1D random walk) methods or by continuous-time (1D Brownian motion) methods [1].

Before discussing the symmetric 3-player problem (which constitutes the most natural generalization of the preceding), let us examine the following **3-player C-centric game** [2, 3]:

$$(a, b, c) \mapsto \begin{cases} (a+1, b, c-1) & \text{with probability } 1/4, \\ (a-1, b, c+1) & \qquad '' \\ (a, b+1, c-1) & \qquad '' \\ (a, b-1, c+1) & \qquad '' \end{cases}$$

In each round, C plays against either A or B (with equal probability) and wins 1 or loses 1 (again with equal probability). Let $m = a + b + c$ denote the initial sum of fortunes. By discrete-time methods, it is known that [3]

$$p_A = f(b, a, m) - f(a, a+c, m)$$

where

$$f(a, b, m) = \frac{2}{m} \sum_{\substack{1 \le j < m \\ j \text{ odd}}} \sin\left(\frac{a j \pi}{m}\right) \cot\left(\frac{j\pi}{2m}\right) \frac{\sinh\left((m-b)\varphi_{j,m}\right)}{\sinh\left(m\,\varphi_{j,m}\right)},$$

$$\varphi_{j,m} = \operatorname{arccosh}\left(2 - \cos(j\pi/m)\right).$$

For example,

$$p_A = \begin{cases} \frac{295476041655}{716708481082} = 0.4122... & \text{if } a = 3, b = 3, c = 9; \\ \frac{2964404261421089}{8592617979692098} = 0.3449... & \text{if } a = 4, b = 4, c = 7; \\ \frac{93962873}{360352742} = 0.2607... & \text{if } a = 5, b = 5, c = 5 \end{cases}$$

and these numerical results are consistent with [2] (obtained by recurrences). From

$$p_A = \begin{cases} \frac{1}{4} = 0.25 & \text{if } a = b = c = 1; \\ \frac{17}{66} = 0.2575... & \text{if } a = b = c = 2; \\ \frac{365}{1406} = 0.2596... & \text{if } a = b = c = 3; \\ \frac{223655}{858958} = 0.2603... & \text{if } a = b = c = 4 \end{cases}$$

it is clear that 3-player problems differ from 2-player problems (because scaling is not invariant) and hence 2D Brownian motion methods will only approximate (but not exactly solve) 2D random walk probabilities. If we allow $m \to \infty$ in such a way that $a/m \to \alpha > 0$ and $b/m \to \beta > 0$, then [3]

$$p_A = g(\beta, \alpha) - g(\alpha, 1 - \beta)$$

where

$$g(\alpha, \beta) = 4 \sum_{\substack{1 \le j < \infty \\ j \text{ odd}}} \frac{\sin(\alpha j \pi)}{j \pi} \frac{\sinh((1 - \beta) j \pi)}{\sinh(j \pi)}.$$

For example,

$$p_A = \begin{cases} 0.2614366507... & \text{if } \alpha = 1/3, \beta = 1/3; \\ 0.4126822642... & \text{if } \alpha = 1/5, \beta = 1/5 \end{cases}$$

in this limiting case. If instead we allow $c \to \infty$ for fixed a, b, then [2]

$$p_A = \frac{1}{\pi} \int_0^{\pi} \frac{\sin(x) \sin(b\,x)}{1 - \cos(y)} e^{-a\,y} dx$$

where

$$\cos(x) + \cosh(y) = 2.$$

For example,

$$p_A = \begin{cases} 1/2 & \text{if } a = b; \\ 0.6976527263... & \text{if } a = 1, b = 2; \\ 0.6232861831... & \text{if } a = 2, b = 3; \\ 0.7906109052... & \text{if } a = 1, b = 3. \end{cases}$$

Let us turn attention to the **symmetric 3-player game**:

$$(a, b, c) \mapsto \begin{cases} (a + 2, b - 1, c - 1) & \text{with probability } 1/3, \\ (a - 1, b + 2, c - 1) & \quad '' \\ (a - 1, b - 1, c + 2) & \quad '' \end{cases}$$

One player wins and is paid 1 by each of the other players. A discrete-time solution was outlined in [4], but it is conceptually very different from C-centric game results. For small values of m, some results are known [5, 6]:

$$p_C = \begin{cases} \frac{2}{3} = 0.6666... & \text{if } a=b=c=1; \\ \frac{4}{9} = 0.4444... & \text{if } a=b=c=2; \\ \frac{8}{21} = 0.3809... & \text{if } a=b=c=3; \\ \frac{16}{45} = 0.3555... & \text{if } a=b=c=4; \\ \frac{848}{2457} = 0.3451... & \text{if } a=b=c=5; \\ \frac{49}{144} = 0.3402... & \text{if } a=b=c=6. \end{cases}$$

Asymptotic numerical evaluation is feasible when modeling the game as Brownian motion in the plane of the equilateral triangle given by

$$\left\{ x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}-1\\0\end{pmatrix} + z\begin{pmatrix}0\\\sqrt{3}\end{pmatrix} : x+y+z=m, x\geq 0, y\geq 0, z\geq 0 \right\}.$$

Computing p_C corresponds to finding the probability that Brownian motion first exits the triangle along the edge $z=0$, starting from $(x,y,z)=(a,b,c)$. In the event $a=b$, we determine $\eta>0$ so that

$$\frac{c}{m} = \frac{I\left(\frac{\eta^2}{1+\eta^2}, \frac{1}{2}, \frac{1}{6}\right)}{I\left(1, \frac{1}{2}, \frac{1}{6}\right)}$$

where

$$I(\xi,\alpha,\beta) = \int_0^{\xi} t^{\alpha-1}(1-t)^{\beta-1}dt$$

is the incomplete beta function; it follows that [7–9]

$$p_C = \frac{1}{\pi}\left(\frac{\pi}{2} - \arctan\left(\frac{\eta^2-1}{2\eta}\right)\right).$$

For example,

$$p_C = \begin{cases} 1/3 & \text{if } a=b=c, \text{ that is, } c/m=1/3; \\ 0.1421549761... & \text{if } 2a=2b=c, \text{ that is, } c/m=1/2; \\ 0.5617334934... & \text{if } a=b=2c, \text{ that is, } c/m=1/5. \end{cases}$$

In the event $a\neq b$, no such explicit formulas apply. A purely numerical approach [8–12] gives, for example,

$$p_A = 0.6542207068..., \quad p_B = 0.2923400189..., \quad p_C = 0.0534392741...$$

when $10a=5b=2c$.

The final game we mention, usually referred to as the 3**-tower problem** (or **Hanoi tower problem**), is [8]:

$$(a,b,c)\mapsto\begin{cases}(a-1,b+1,c) & \text{with probability } 1/6,\\ (a-1,b,c+1) & \qquad''\\ (a+1,b-1,c) & \qquad''\\ (a,b-1,c+1) & \qquad''\\ (a+1,b,c-1) & \qquad''\\ (a,b+1,c-1) & \qquad''\end{cases}$$

In each round, one player is randomly chosen as the loser and one player (distinct from the first) is randomly chosen as the winner. A study of corresponding ruin probabilities has evidently not been done.

Another quantity of interest is the **game duration** d, which is the expected number of rounds until one of the gamblers is ruined. For the symmetric 2-player and 3-player problems, we have [13–15]

$$d=a\,b,\quad d=\frac{a\,b\,c}{a+b+c-2}$$

respectively. For the 3-tower problem, we have [14–18]

$$d=\frac{3a\,b\,c}{a+b+c};$$

in fact, corresponding variance and probability distribution are also known. For the 3-player C-centric game, $d=ab+bc+ca$ [19]. No simple formulas for d can be anticipated when the number of players exceeds three [16, 20, 21].

Here is an interesting variation on the symmetric 2-player problem:

$$(a_1,a_2,b_1,b_2)\mapsto\begin{cases}(a_1+1,a_2,b_1-1,b_2) & \text{with probability } 1/4,\\ (a_1-1,a_2,b_1+1,b_2) & \qquad''\\ (a_1,a_2+1,b_1,b_2-1) & \qquad''\\ (a_1,a_2-1,b_1,b_2+1) & \qquad''\end{cases}$$

The gamblers use two different currencies, say dollars and euros. In each round, a currency and a winner are randomly chosen. When one of the players runs out of either currency, the game is over. Ruin probabilities p are not known; if $a_1=a_2=b_1=b_2=n$, then game durations d are $O(n^2)$ and, more precisely, [22]

$$\delta=\lim_{n\to\infty}\frac{d}{n^2}=\frac{256}{\pi^4}\sum_{k=0}^{\infty}\sum_{\ell=0}^{\infty}\frac{(-1)^{k+\ell}}{(2k+1)(2\ell+1)\left[(2k+1)^2+(2\ell+1)^2\right]}.$$

Another representation

$$\delta=2\left(1-\frac{32}{\pi^3}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)^3\cosh\left[\frac{\pi}{2}(2k+1)\right]}\right)=1.1787416525...$$

is rapidly convergent and possesses a straightforward generalization to an arbitrary number of different currencies.

The following question is similar to our asymptotic analysis of the symmetric 3-player game. Let $a \le b$. A particle at the center of an $a \times b$ rectangle undergoes Brownian motion until it hits the rectangular boundary. What is the probability that it hits an edge of length a (rather than an edge of length b)? The answer [23, 24]

$$P(b/a) = \frac{4}{\pi} \sum_{j=0}^{\infty} \frac{(-1)^j}{2j+1} \operatorname{sech}\left(\frac{(2j+1)\pi}{2} \frac{b}{a} \right)$$

is found via solution of a steady-state heat PDE problem. This has a closed-form expression in certain cases: [25–27]

$$P(r) = \begin{cases} \frac{1}{2} & \text{if } r = 1, \\ \frac{2}{\pi} \arcsin\left[(\sqrt{2}-1)^2 \right] & \text{if } r = 2, \\ \frac{2}{\pi} \arcsin\left[(\sqrt{2}-3^{1/4})(\sqrt{3}-1)/2 \right] & \text{if } r = 3, \\ \frac{2}{\pi} \arcsin\left[(\sqrt{2}+1)^2(2^{1/4}-1)^4 \right] & \text{if } r = 4, \\ \frac{2}{\pi} \arcsin\left[(\sqrt{5}-2)(3-2\cdot 5^{1/4})/\sqrt{2} \right] & \text{if } r = 5, \\ \frac{2}{\pi} \arcsin\left[(3-2\sqrt{2})^2(2+\sqrt{5})^2(\sqrt{10}-3)^2(5^{1/4}-\sqrt{2})^4 \right] & \text{if } r = 10, \end{cases}$$

which are based on singular moduli k_1, k_4, k_9, k_{16}, k_{25}, k_{100} appearing in the theory of elliptic functions. We wonder whether heat PDE-type analysis might assist in the asymptotic study of some 4-player games.

[1] H. M. Taylor and S. Karlin, *An Introduction to Stochastic Modeling*, 3rd ed., Academic Press, 1998, pp. 141–145, 509–514; MR1627763 (99c:60001).

[2] V. D. Barnett, A three-player extension of the gambler's ruin problem, *J. Appl. Probab.* 1 (1964) 321–334; MR0171330 (30 #1561).

[3] Y. Itoh and H. Maehara, A variation to the ruin problem, *Math. Japon.* 47 (1998) 97–102; MR1606328 (98m:60068).

[4] Y. C. Swan and F. T. Bruss, A matrix-analytic approach to the N-player ruin problem, *J. Appl. Probab.* 43 (2006) 755–766; MR2274798 (2007m:60221).

[5] A. L. Rocha and F. Stern, The gambler's ruin problem with n players and asymmetric play, *Statist. Probab. Lett.* 44 (1999) 87–95; MR1706327 (2000f:60063).

[6] A. L. Rocha and F. Stern, The asymmetric n-player gambler's ruin problem with equal initial fortunes, *Adv. Appl. Math.* 33 (2004) 512–530; MR2081041 (2005d:60067).

[7] T. Ferguson, Gambler's ruin in three dimensions, unpublished note (1995).

[8] Y. C. Swan, On Two Unsolved Problems in Probability, Ph.D. thesis, Université Libre de Bruxelles, 2007.

[9] Y. C. Swan and F. T. Bruss, The Schwarz-Christoffel transformation as a tool in applied probability, *Math. Sci.* 29 (2004) 21–32; MR2073566.

[10] L. N. Trefethen, Numerical computation of the Schwarz-Christoffel transformation, *SIAM J. Sci. Statist. Comput.* 1 (1980) 82–102; erratum 1 (1980) 302; MR0572542 (81g:30012a) and MR0594762 (81g:30012b).
[11] T. A. Driscoll, Algorithm 756: A MATLAB toolbox for Schwarz-Christoffel mapping, *ACM Trans. Math. Software* 22 (1996) 168–186.
[12] T. A. Driscoll, Algorithm 843: Improvements to the Schwarz-Christoffel toolbox for MATLAB, *ACM Trans. Math. Software* 31 (2005) 239–251; MR2266791 (2007f:30001).
[13] D. Sandell, A game with three players, *Statist. Probab. Lett.* 7 (1988) 61–63; MR0996854 (90g:60047).
[14] A. Engel, The computer solves the three tower problem, *Amer. Math. Monthly* 100 (1993) 62–64.
[15] D. Stirzaker, Tower problems and martingales, *Math. Sci.* 19 (1994) 52–59; MR1294785 (95h:60068).
[16] F. T. Bruss, G. Louchard and J. W. Turner, On the N-tower problem and related problems, *Adv. Appl. Probab.* 35 (2003) 278–294; MR1975514 (2004d:60181).
[17] A. Alabert, M. Farré and R. Roy, Exit times from equilateral triangles, *Appl. Math. Optim.* 49 (2004) 43–53; MR2023644 (2004j:60175).
[18] D. Stirzaker, Three-handed gambler's ruin, *Adv. Appl. Probab.* 38 (2006) 284–286; MR2213975 (2006k:60073).
[19] R. Cowan and J. Burridge, Gamblers' ruin with many players: the ratio law, unpublished note (2016).
[20] D. K. Chang, A game with four players, *Statist. Probab. Lett.* 23 (1995) 111–115; MR1341352 (96g:60055).
[21] D. Cho, A game with N players, *J. Korean Statist. Soc.* 25 (1996) 185–193; MR1423937 (98a:60051).
[22] A. Kmet and M. Petkovšek, Gambler's ruin problem in several dimensions, *Adv. Appl. Math.* 28 (2002) 107–118; MR1888839 (2003a:60012).
[23] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed., Oxford Univ. Press, 2001, pp. 554–563; MR2059709 (2004m:60002).
[24] F. Bornemann, Short remarks on the solution of the SIAM 100-digit challenge, unpublished note (2002).
[25] J. M. Borwein and P. B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, Wiley, 1987, pp. 69, 139, 162; MR1641658 (99h:11147).
[26] B. C. Berndt, H. H. Chan and L.-C. Zhang, Ramanujan's singular moduli, *Ramanujan J.* 1 (1997) 53–74; MR1607528 (2001b:11033).
[27] J. Boersma, Solution of Trefethen's problem 10, unpublished note (2002).

4.21 Self-Convolutions

Let f be a square-integrable probability density function supported on a subinterval of $\mathbb{R}$ of length $1/2$. Define the **self-convolution** of f to be

$$(f*f)(x) = \int_{-\infty}^{\infty} f(t)f(x-t)\,dt.$$

Thus $f*f$ is the probability density for a sum of two independent random variables, each distributed according to f, and is supported on an interval of length 1.

We are interested in the "size" of $f*f$, measured via both L_2 and L_∞ norms. Before doing this, however, let us examine f alone as a preliminary exercise.

For each integer $n \geq 1$, define

$$g_n(x) = \frac{n+1}{n}\left(\frac{1}{\sqrt{2x}}\right)^{\frac{n-1}{n}}, \quad 0 < x < 1/2$$

then clearly g_n is a probability density for all n,

$$\|g_n\|_2^2 = \int_0^{1/2} g_n(x)^2 dx = \frac{(n+1)^2}{2n} \to \infty$$

as $n \to \infty$, and $\|g_n\|_\infty = \infty$ always. Consequently

$$\sup_f \|f\|_2^2 = \infty = \sup_f \|f\|_\infty .$$

Also, suppose that there exists a probability density h on $[0, 1/2]$ with $\|h\|_2^2 < 2$. By the Cauchy–Schwarz inequality,

$$2 = \int_0^{1/2} h(x) \cdot 2\, dx \leq \|h\|_2 \cdot \|2\|_2 < \sqrt{2} \cdot \sqrt{2} = 2,$$

which is a contradiction. Consequently

$$\inf_f \|f\|_2^2 = 2 = \inf_f \|f\|_\infty .$$

The problem of assessing $f*f$ together is more difficult. Let us first discuss relevant infimums. Martin & O'Bryant [1, 2] conjectured that

$$\inf_f \|f*f\|_\infty = \pi/2 = 1.5707963267...$$

on the basis of their proof that the left-hand side must exceed $1.262 = (2)(0.638)$, plus their observation that $\|g*g\|_\infty = \pi/2$, where

$$g(x) = \lim_{n\to\infty} g_n(x) = 1/\sqrt{2x}.$$

Technically, g is not admissible (since it is not square-integrable). See [3–5] for discussion of a similar case.

Martin & O'Bryant [1] also proved that

$$\inf_f \|f*f\|_2^2 \geq 1.14915 = (2)(0.574575)$$

after elaborate computations. This may be nearly correct, since the probability density

$$k(x) = \frac{4}{\pi}\frac{1}{\sqrt{8x(1-2x)}}, \quad 0 < x < 1/2$$

satisfies

$$\|k * k\|_2^2 < 1.14939.$$

Again, k is not admissible for technical reasons. No exact formula is even conjectured in this case, which renders it especially interesting!

Here is a problem involving ratios of L_p norms. Hölder's inequality gives

$$\|f\|_2^2 \le \|f\|_\infty \cdot \|f\|_1$$

which is an equality if $f = 2$ on $[0, 1/2]$. Consequently

$$\inf_f \frac{\|f\|_\infty}{\|f\|_2^2} = 1.$$

Martin & O'Bryant [1, 2] conjectured that

$$\inf_f \frac{\|f * f\|_\infty}{\|f * f\|_2^2} = \frac{\pi}{4\ln(2)}$$

on the basis, in part, of their observation that $\|g * g\|_2^2 = 2\ln(2)$. This result gives a sense of how large $\|f * f\|_2^2$ can be, in terms of $\|f * f\|_\infty$. No other mention of relevant supremums in the literature has yet been found!

Addendum The first conjecture is false: in fact,

$$1.2748 \le \inf_f \|f * f\|_\infty \le 1.5098.$$

The second conjecture is also false: in fact,

$$\inf_f \frac{\|f * f\|_\infty}{\|f * f\|_2^2} \le \frac{1}{0.88922\ldots} < \frac{1}{0.88254\ldots} = \frac{\pi}{4\ln(2)}.$$

Such adjustments open up this subject considerably since no one knows what the extremal functions f now might be [6, 7]. A sequence of lower bounds defined in [8] and numerical optimization (on a simplex in $\mathbb{R}^{2n}$) suggest an improvement 1.28 over 1.2748; the upper bound 1.5098 is believed to be close to the true value.

[1] G. Martin and K. O'Bryant, The symmetric subset problem in continuous Ramsey theory, *Experim. Math.* 16 (2007) 145–165; arXiv:math/0410004; MR2339272 (2008h:05114).

[2] G. Martin and K. O'Bryant, The supremum of autoconvolutions, with applications to additive number theory, *Illinois J. Math.* 53 (2009) 219–235; arXiv:0807.5121; MR2584943 (2011c:42021).

[3] B. Green, The number of squares and $B_h[g]$ sets, *Acta Arith.* 100 (2001) 365–390; MR1862059 (2003d:11033).

[4] A. Schinzel and W. M. Schmidt, Comparison of L^1- and L^∞-norms of squares of polynomials, *Acta Arith.* 104 (2002) 283–296; MR1914723 (2003f:11035).

[5] J. Cilleruelo and C. Vinuesa, $B_2[g]$ sets and a conjecture of Schinzel and Schmidt, *Combin. Probab. Comput.* 17 (2008) 741–747; MR2463407 (2009h:11037).

[6] M. Matolcsi and C. Vinuesa, Improved bounds on the supremum of autoconvolutions, *J. Math. Anal. Appl.* 372 (2010) 439–447; arXiv:0907.1379; MR2678874 (2011j:11043).
[7] J. Cilleruelo, I. Ruzsa and C. Vinuesa, Generalized Sidon sets, *Adv. Math.* 225 (2010) 2786–2807; arXiv:0909.5024; MR2680183 (2011m:11032).
[8] A. Cloninger and S. Steinerberger, On suprema of autoconvolutions with an application to Sidon sets, *Proc. Amer. Math. Soc.* 145 (2017) 3191–3200; arXiv:1403.7988; MR3652775.

4.22 Newcomb–Benford Law

The literature for Benford's law is quite large and growing [1]; we avoid interesting foundational issues [2, 3] and turn attention instead to a specific scenario [4–7].

Let $\{a_n\}_{n=0}^{\infty}$ be an m^{th} order linear homogeneous recurrence. Consequently the sequence can be written as

$$a_n = p_1 q_1^n + P_2(n) q_2^n + P_3(n) q_3^n + \cdots + P_m(n) q_m^n,$$

where $q_1, q_2, \ldots, q_m$ are associated eigenvalues; q_1 is the largest eigenvalue (in absolute value); p_1 is constant and $P_2, P_3, \ldots, P_m$ are polynomials. The sequence $\{a_n\}$ is called **random-enough** if q_1 is real, positive, not a rational power of 10, of multiplicity 1 (as an eigenvalue) and p_1 is positive. Famous integer examples include

$$a_n = 2^n \quad \text{(powers of 2)},$$

$$a_n = \frac{\varphi^n - (1-\varphi)^n}{\sqrt{5}} \quad \text{(Fibonacci sequence)},$$

$$a_n = \varphi^n + (1-\varphi)^n \quad \text{(Lucas sequence)}$$

where φ is the Golden mean [8].

Consider the j^{th} leftmost decimal digit D_j of an integer a. If $j = 1$, then $1 \le D_j(a) \le 9$; if $j \ge 2$, then $0 \le D_j(a) \le 9$. Let $\{a_n\}_{n=0}^{\infty}$ be a random-enough sequence of positive integers. Benford's law states that [4]

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \# \{n \le N : D_1(a_n) = d\} &= \log_{10}\left(1 + \frac{1}{d}\right) \\ &= \sum_{k=0}^{0} \log_{10}\left(1 + \frac{1}{10k + d}\right)\end{aligned}$$

for $1 \le d \le 9$. In words, the first digit of an arbitrary term a_n is not uniformly distributed over $\{1, 2, \ldots, 9\}$, but instead favors small values:

$$\mathrm{P}\{D_1 = 1\} = 0.30103..., \quad \mathrm{P}\{D_1 = 2\} = 0.17609..., \quad \mathrm{P}\{D_1 = 3\} = 0.12493...$$

and, of course, $\mathrm{P}\{D_1 = 0\} = 0$.

Fix $j \geq 2$. A generalization of Benford's law states that [4]

$$\lim_{N\to\infty} \frac{1}{N} \# \{n \leq N : D_j(a_n) = d\} = \sum_{k=10^{j-2}}^{10^{j-1}-1} \log_{10}\left(1 + \frac{1}{10k+d}\right)$$

for $0 \leq d \leq 9$. The second digit of an arbitrary term a_n is not uniformly distributed over $\{0, 1, \ldots, 9\}$:

$$\mathrm{P}\{D_2 = 1\} = 0.11389..., \quad \mathrm{P}\{D_2 = 2\} = 0.10882..., \quad \mathrm{P}\{D_2 = 3\} = 0.10432...$$

and $\mathrm{P}\{D_2 = 0\} = 0.11967...$; each of the probabilities are, however, closer to $1/10$ than before. The same is true for the third digit of an arbitrary term a_n:

$$\mathrm{P}\{D_3 = 1\} = 0.10137..., \quad \mathrm{P}\{D_3 = 2\} = 0.10097..., \quad \mathrm{P}\{D_3 = 3\} = 0.10057...$$

and $\mathrm{P}\{D_3 = 0\} = 0.10178...$. Such numerical results were first tabulated in [9, 10]. For simplicity, we henceforth refer to Benford's law and its generalization together ($j \geq 1$) as NBL.

Another way to illustrate the approach to uniformity (as $j \to \infty$) makes use of moments. It is straightforward to show that [11]

$$\mathrm{E}(D_1) = 2\log_{10}(2) - 4\log_{10}(3) + 8\log_{10}(5) - \log_{10}(7) = 3.4402369671...,$$

$$\mathrm{E}(D_1^2) = 8\log_{10}(2) - 50\log_{10}(3) + 72\log_{10}(5) - 13\log_{10}(7),$$

$$\mathrm{Var}(D_1) = \mathrm{E}(D_1^2) - \mathrm{E}(D_1)^2 = 6.0565126313...,$$

$$\mathrm{E}(D_2) = 4.1873897069..., \quad \mathrm{Var}(D_2) = 8.2537786232...,$$

$$\mathrm{E}(D_3) = 4.4677656509..., \quad \mathrm{Var}(D_3) = 8.2500943647....$$

The means approach $9/2$ and the variances approach $33/4$, as anticipated. We also have

$$\mathrm{Cov}(D_1, D_2) = \mathrm{E}(D_1 D_2) - \mathrm{E}(D_1)\mathrm{E}(D_2) = 14.8019478993...,$$

for example. Correlation coefficients are small but positive; the largest is

$$\rho(D_1, D_2) = \frac{\mathrm{Cov}(D_1, D_2)}{\sqrt{\mathrm{Var}(D_1)}\sqrt{\mathrm{Var}(D_2)}} = 0.0560563403....$$

It is further known that the sequence $\{n!\}_{n=0}^{\infty}$ and triangular array $\{\binom{k}{\ell} : 0 \leq \ell \leq k,\ k \geq 1\}$ satisfy NBL [12]. The sequences $\{n^2\}_{n=0}^{\infty}$ and $\{n^3\}_{n=0}^{\infty}$ appear to offer special challenges, since the limiting digital probabilities evidently do not exist [3].

First-digit phenomena were mentioned in [13] without elaboration. In the language of [14], $\{a_n\}_{n=0}^{\infty}$ satisfies NBL if and only if the fractional parts of $\log_{10}(a_n)$ are uniformly distributed in $[0, 1]$, proved by Diaconis [12]. Our discussion can be extended to non-integer variables X, where we agree that $D_1(1/2) = 5 = D_1(1/20)$ (the first significant decimal digit). For example,

$$\mathrm{P}\{D_1(X) = 1\} = \frac{1}{9} < \log_{10}(2)$$

if X is Uniform$(0,1)$ and

$$\mathrm{P}\{D_1(X) = 1\} = \sum_{k=-\infty}^{\infty} \left(\exp(-10^k) - \exp(-2 \cdot 10^k)\right) = 0.32965... > \log_{10}(2)$$

if X is Exponential(1). Thus NBL does not apply in either case [2].

[1] A. Berger and T. P. Hill, Benford Online Bibliography.
[2] A. Berger and T. P. Hill, Benford's law strikes back: no simple explanation in sight for mathematical gem, *Math. Intelligencer* 33 (2011) 85–91; MR2774458 (2012j:62006).
[3] K. A. Ross, Benford's law, a growth industry, *Amer. Math. Monthly* 118 (2011) 571–583; MR2826449 (2012g:62014).
[4] S. Slijepčević, A note on initial digits of recurrence sequences, *Fibonacci Quart.* 36 (1998) 305–308; MR1640361.
[5] L. C. Washington, Benford's law for Fibonacci and Lucas numbers, *Fibonacci Quart.* 19 (1981) 175–177; MR0614056 (82f:10009).
[6] J. L. Brown and R. L. Duncan, Modulo one uniform distribution of the sequence of logarithms of certain recursive sequences, *Fibonacci Quart.* 8 (1970) 482–486; MR0360444 (50 #12894).
[7] L. Kuipers and J. S. Shiue, Remark on a paper by Duncan and Brown on the sequence of logarithms of certain recursive sequences, *Fibonacci Quart.* 11 (1973) 292–294; MR0332699 (48 #11025).
[8] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[9] S. Newcomb, Note on the frequency of use of the different digits in natural numbers, *Amer. J. Math.* 4 (1881) 39–40.
[10] F. Benford, The law of anomalous numbers, *Proc. Amer. Philos. Soc.* 78 (1938) 551–572.
[11] Z. Li, L. Cong and H. Wang, Discussion on Benford's law and its application, arXiv:math/0408057.
[12] P. Diaconis, The distribution of leading digits and uniform distribution mod 1, *Annals of Probab.* 5 (1977) 72–81; MR0422186 (54 #10178).
[13] S. R. Finch, Continued fraction transformation. III, *this volume*, §4.11.
[14] S. R. Finch, Pisot-Vijayaraghavan-Salem constants, *first volume*, pp. 192–199.

4.23 Electing a Leader

The following scenario was examined in [1]: we toss n ideal coins, then toss those which show tails after the first toss, then toss those which show tails after the second toss, etc. Observe that if, at a given toss, only heads appear, then the process immediately terminates. Suppose instead that we require the coins (all of which showed heads) to be tossed again? Under such a change of rules, it is clear that the final toss will always involve exactly one coin. This solitary coin is called the **leader** and the process of selecting such is called an **election**.

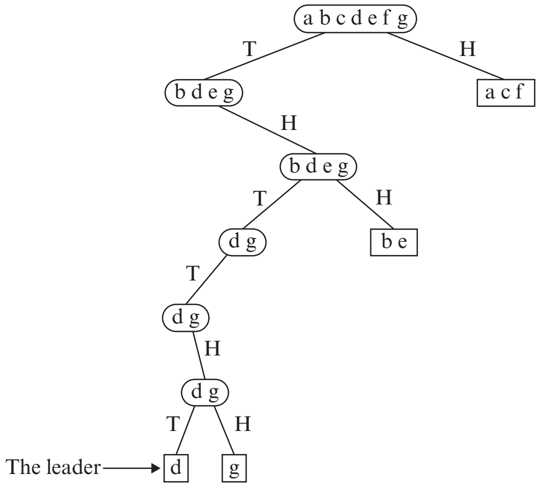

Figure 4.4 A typical election, starting with $n = 7$ candidates.

Certain parameters governing the election (a random incomplete trie) are of interest. In Figure 4.4, the **size** $v_7 = 10$ is the number of vertices in the tree. The **height** $h_7 = 6$ is the length of the longest root-to-leaf path, that is, the time duration to choose a leader. Finally, $c_7 = 21$ is the total number of coin tosses. Let C_n denote likewise, given arbitrary n and a random election. It is surprising that $\mathrm{E}(C_n) = 2n$ for $n \geq 2$; the random variables V_n and H_n are more complicated [2].

The following sums involving Bernoulli numbers [3] and binomial coefficients are relevant and interesting [2, 4]:

$$\frac{1}{n}\sum_{k=2}^{n-1}\binom{n}{k}\frac{B_k}{2^{k-1}-1} \sim \frac{\ln(n)}{2\ln(2)} - \left(\frac{\ln(\pi)}{2\ln(2)} - \frac{\gamma}{2\ln(2)} + \frac{3}{4}\right) + \delta_1\left(\frac{\ln(n)}{\ln(2)}\right),$$

$$\sum_{k=1}^{n-1}\binom{n}{k}\frac{B_k}{2^k-1} \sim -\frac{\ln(n)}{\ln(2)} + \frac{1}{2} + \delta_2\left(\frac{\ln(n)}{\ln(2)}\right),$$

$$n\sum_{k=0}^{n-1}\binom{n}{k}\frac{B_k}{2^{k+1}-1} \sim \frac{\pi^2}{6\ln(2)} + \delta_3\left(\frac{\ln(n)}{\ln(2)}\right)$$

where, for $m = 1, 2, 3$,

$$\delta_m(x) = \frac{1}{\ln(2)}\sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty}\zeta\left(m-1-\frac{2\pi i k}{\ln(2)}\right)\Gamma\left(m-1-\frac{2\pi i k}{\ln(2)}\right)\exp(2\pi i k x)$$

are periodic functions of period 1 and very small amplitude. For example, $|\delta_2(x)| < 1.927 \times 10^{-5}$ for all x. Each fluctuates symmetrically about 0. Define also [5]

$$\varepsilon(x) = \frac{2}{\ln(2)^2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \nu\left(\frac{2\pi ik}{\ln(2)}\right) \exp(2\pi ikx) - \delta_2^2(x)$$

where

$$\nu(s) = \zeta(1-s)\Gamma(-s) - s\zeta'(1-s)\Gamma(-s) - s\zeta(1-s)\Gamma'(-s).$$

This again has period 1 and small amplitude – we have $|\varepsilon(x)| < 1.398 \times 10^{-4}$ always – fluctuations are symmetrical not about 0, but instead about

$$\int_0^1 \varepsilon(x)dx = -\frac{1}{\ln(2)^2} \sum_{\substack{k=-\infty \\ k\neq 0}}^{\infty} \left|\zeta\left(1-\frac{2\pi ik}{\ln(2)}\right)\Gamma\left(1-\frac{2\pi ik}{\ln(2)}\right)\right|^2 \approx -1.856 \times 10^{-10}.$$

Let us return to coin tossing. Prodinger [2] showed that

$$\mathrm{E}(V_n) \sim \frac{2\ln(n)}{\ln(2)} + \left(2 - \frac{\ln(\pi)-\gamma}{\ln(2)}\right) + 2\delta_1\left(\frac{\ln(n)}{\ln(2)}\right) - \delta_2\left(\frac{\ln(n)}{\ln(2)}\right),$$

$$\mathrm{E}(H_n) = -\sum_{k=1}^{n-1} \binom{n}{k} \frac{B_k}{1-2^{-k}} \sim \frac{\ln(n)}{\ln(2)} + \frac{1}{2} - \delta_2\left(\frac{\ln(n)}{\ln(2)}\right)$$

asymptotically as $n \to \infty$, assuming that the election is conducted exactly as described earlier. If we alter the rules so that a draw between two coins is allowed (if precisely two coins are left, they *both* are declared leaders), then

$$\begin{aligned}\mathrm{E}(\tilde{V}_n) \sim{} & \frac{2\ln(n)}{\ln(2)} + \left(2 - \frac{\ln(\pi)-\gamma+\frac{\pi^2}{16}}{\ln(2)}\right) + 2\delta_1\left(\frac{\ln(n)}{\ln(2)}\right) - \delta_2\left(\frac{\ln(n)}{\ln(2)}\right) \\ & - \frac{3}{8}\delta_3\left(\frac{\ln(n)}{\ln(2)}\right),\end{aligned}$$

$$\mathrm{E}(\tilde{H}_n) \sim \frac{\ln(n)}{\ln(2)} + \frac{1}{2} - \frac{\pi^2}{12\ln(2)} - \delta_2\left(\frac{\ln(n)}{\ln(2)}\right) - \frac{1}{2}\delta_3\left(\frac{\ln(n)}{\ln(2)}\right),$$

$$\mathrm{E}(\tilde{C}_n) \sim 2n - \frac{\pi^2}{6\ln(2)} - \delta_3\left(\frac{\ln(n)}{\ln(2)}\right).$$

For example, the constant for $\mathrm{E}(V_n)$ is 1.1812500478...; the difference $\pi^2/(16\ln(2)) = 0.8899268328...$ quantifies how much is saved by stopping earlier to give $\mathrm{E}(\tilde{V}_n)$. For $\mathrm{E}(H_n)$ versus $\mathrm{E}(\tilde{H}_n)$, the difference $\pi^2/(12\ln(2)) = 1.1865691104...$ is slightly greater.

Fill, Mahmoud & Szpankowski [5] proved that

$$\mathrm{Var}(H_n) \sim \frac{1}{12} + \frac{\pi^2}{6\ln(2)^2} - \frac{\gamma^2+2\gamma_1}{\ln(2)^2} + \varepsilon\left(\frac{\ln(n)}{\ln(2)}\right)$$

asymptotically as $n \to \infty$, where γ_1 is the first Stieltjes constant [6]. As predicted in [2], this is a nontrivial result. The constant 3.1166951643... also appears in [7]; another treatment is given by [8]. As far as is known, evaluating $\mathrm{Var}(\tilde{H}_n)$ remains open. The parameters V_n, C_n, $\tilde{V}_n$, $\tilde{C}_n$ deserve more attention. Random elections yielding a predetermined number > 1 of leaders are examined in [9].

4.23.1 *Non-Ideal Coins*

Instead of assuming that coins are ideal (independent probability of tails $= 1/2$), let us suppose that coins "know" their count just before each toss. More precisely, if $n_1 = n$ is the count before the first toss and n_j is the count before the j^{th} toss, $j \geq 1$, then at time j, each coin enjoys independent probability of tails $= 1/n_j$. Since $n_{j+1} \leq n_j$, the odds that any active candidate becomes the leader improve with time. If $n_{j+1} = 1$, the election is over. If $n_{j+1} = 0$, then n_{j+1} is overwritten with n_j and the coins are tossed again.

Clearly $\mathrm{E}(H_1) = 0$. From the recursion

$$\left[1 - \left(1 - \frac{1}{n}\right)^n - \left(\frac{1}{n}\right)^n\right] \mathrm{E}(H_n) = 1 + \sum_{k=2}^{n-1} \binom{n}{k} \left(\frac{1}{n}\right)^k \left(1 - \frac{1}{n}\right)^{n-k} \mathrm{E}(H_k)$$

for $n \geq 2$, we obtain $\mathrm{E}(H_2) = 2$, $\mathrm{E}(H_3) = 13/6$, $\mathrm{E}(H_4) = 65/29$ and [10, 11]

$$\lim_{n\to\infty} \mathrm{E}(H_n) = 2.4417158788....$$

A more complicated recursion gives $\lim_{n\to\infty} \mathrm{Var}(H_n) = 2.832554383....$

Is $1/n$ the optimal probability? Replacing $1/n$ everywhere by t/n for $0 < t < 2$ in the preceding, we obtain

$$\mathrm{E}(H_2) = \frac{2}{(2-t)t}, \quad \mathrm{E}(H_3) = \frac{18 - 3t - 2t^2}{3(3-t)(2-t)t}.$$

Differentiating the recursion with respect to t allows us to find a minimum point $t^* = 1.0654388051...$ and thus [11]

$$\lim_{n\to\infty} \mathrm{E}(H_n^*) = 2.4348109638....$$

No one has evaluated $\lim_{n\to\infty} \mathrm{Var}(H_n^*)$, as far as is known. Related topics in random elections are found in [12].

4.23.2 *Number Games*

The following game, proposed by Gilbert [13], was revisited by Fokkink [14]. A player A chooses a secret integer from 1 to n. Another player B attempts to guess A's integer. After each guess, A tells B whether the guess is too high, too low or correct. If B has guessed A's integer, the game ends. If not, then A may change the secret integer, but the new integer must be consistent with all the information so far provided. Assuming both players adopt optimal, equilibrium, randomized strategies, the expected number ξ_n of guesses is conjectured to satisfy [14]

$$\xi_n \sim \frac{\ln(n)}{\ln(2)} - (0.487...)$$

asymptotically as $n \to \infty$. It is further acknowledged in [14] that this formula may require a small amplitude oscillation and [9] is cited. A verification of either claim would be good to see.

Here is a comparatively simple game, proposed by Häggström [15] as a model for the Swedish National Lottery. Every contestant chooses a positive integer. The person who submits the smallest integer not chosen by anybody else is the winner. (If no integer is chosen by exactly one person, then there is no winner.) Let us focus on the case where there are exactly three contestants. Assuming all three adopt optimal, equilibrium, randomized strategies, each of them independently draws an integer according to a shifted geometric distribution:

$$\mathrm{P}\,(\ell \text{ is selected}) = (1-r)r^{\ell-1},$$

where $\ell = 1, 2, 3, \ldots$ and $r = 0.5436890126...$ satisfies the cubic equation

$$\frac{1}{r^3} - \frac{1}{r^2} - \frac{1}{r} - 1 = 0.$$

This constant is the reciprocal growth rate for the so-called Tribonacci sequence [16]. What can be said if instead there are exactly four contestants? The only other reference found on this subject, [17], contains more elaborate analyses (assuming a Poisson random count of players or an upper bound on playable numbers, if not both).

[1] S. R. Finch, Feller's constant, *first volume*, pp. 339–342.

[2] H. Prodinger, How to select a loser, *Discrete Math.* 120 (1993) 149–159; MR1235902 (94g:05010).

[3] S. R. Finch, Fractional parts of Bernoulli numbers, *this volume*, §1.23.

[4] D. E. Knuth, *The Art of Computer Programming*, v. 3, *Sorting and Searching*, Addison-Wesley, 1973, pp. 500–507, 510, 727; MR0445948 (56 #4281).

[5] J. A. Fill, H. M. Mahmoud and W. Szpankowski, On the distribution for the duration of a randomized leader election algorithm, *Annals Appl. Probab.* 6 (1996) 1260–1283; MR1422986 (97k:05053).

[6] S. R. Finch, Stieltjes constants, *first volume*, pp. 166–171.

[7] S. Janson and W. Szpankowski, Analysis of an asymmetric leader election algorithm, *Elec. J. Combin.* 4 (1997) R17; MR1464071 (98d:90039).

[8] C. Knessl, Asymptotic and numerical studies of the leader election algorithm, *European J. Appl. Math.* 12 (2001) 645–664; MR1877345 (2002k:60033).

[9] P. J. Grabner, Searching for losers, *Random Structures Algorithms* 4 (1993) 99–110; MR1192529 (93i:05048).

[10] A. Itai and M. Rodeh, Symmetry breaking in distributed networks, *Inform. and Comput.* 88 (1990) 60–87; MR1068205 (91j:68011).

[11] C. Lavault and G. Louchard, Asymptotic analysis of a leader election algorithm, *Theoret. Comput. Sci.* 359 (2006) 239–254; MR2251613 (2008d:68148).
[12] S. Janson, C. Lavault and G. Louchard, Convergence of some leader election algorithms, *Discrete Math. Theor. Comput. Sci.* 10 (2008) 171–196; MR2464596 (2010c:60075).
[13] E. N. Gilbert, Games of identification or convergence, *SIAM Review* 4 (1962) 16–24.
[14] R. Fokkink, Tossing coins to guess a secret number, *Amer. Math. Monthly* 119 (2012) 337–339; MR2900981.
[15] P. Winkler, *Mathematical Puzzles: a Conoisseur's Collection*, A. K. Peters, Ltd., 2004, pp. 33, 96, 98; MR2034896 (2006c:00002).
[16] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[17] R. Östling, J. T. Wang, E. Y. Chou and C. F. Camerer, Testing game theory in the field: Swedish LUPI lottery games, *Amer. Econ. J.: Microecon.* 3 (2011) 1–33.

4.24 Substitution Dynamics

Starting with 0, the bit substitutions

$$\begin{cases} 0 \to 01 \\ 1 \to 10 \end{cases}, \quad \begin{cases} 0 \to 01 \\ 1 \to 0 \end{cases}$$

generate recursively the infinite Prouhet–Thue–Morse word 0110100110010110... and Fibonacci word 01001010010010100101..., respectively [1]. What can be said about the *entropy* (loosely, the amount of disorder) if we introduce some randomness into such definitions?

If [2, 3]

$$\begin{cases} 0 \to \begin{cases} 01 & \text{with probability } 1/2, \\ 10 & \text{with probability } 1/2 \end{cases} \\ 1 \to 0 \end{cases}$$

with independence assumed throughout, then the set of possible words at step $n-2$ is $\{001, 010, 100\}$ at $n=4$ and

$$\{00101, 00110, 01001, 01010, 01100, 10001, 10010, 10100\}$$

at $n=5$. Define

$$f_n = f_{n-1} + f_{n-2} \text{ for } n \geq 2, \quad f_0 = 0, \quad f_1 = 1$$

(Fibonacci's sequence) and [4]

$$a_n = (2a_{n-1} - a_{n-2}a_{n-3})\, a_{n-2} \text{ for } n \geq 3, \quad a_0 = 0, \quad a_1 = 1, \quad a_2 = 1.$$

At step 2, there are $a_4 = 3$ words, each of length $f_4 = 3$; at step 3, there are $a_5 = 8$ words, each of length $f_5 = 5$. The corresponding entropy is

$$\begin{aligned} \lim_{n\to\infty} \frac{\ln(a_n)}{f_n} &= \lim_{n\to\infty} \frac{1}{f_{n+1}} \left[\ln(n) + \sum_{k=2}^{n-1} f_{k-2} \ln(n-k+1) \right] \\ &= 0.4443987251... = \ln(1.5595521944...). \end{aligned}$$

Here is a somewhat artificial example on three symbols (with motivation to come later). If [5]

$$\begin{cases} 0 \to 01 \\ 1 \to \begin{cases} 10 & \text{with probability } 1/2, \\ 20 & \text{with probability } 1/2, \end{cases} \\ 2 \to 22 \end{cases}$$

with independence assumed throughout, then the set of possible words at step n is $\{0110, 0120\}$ at $n = 2$ and

$$\{01101001, 01102001, 01102201, 01201001, 01202001, 01202201\}$$

at $n = 3$. Define [4, 6]

$$\alpha_n = (\alpha_{n-1} + \alpha_{n-2})\,\alpha_{n-1} \text{ for } n \geq 3, \quad \alpha_1 = 1, \quad \alpha_2 = 2.$$

At step 2, there are $\alpha_2 = 2$ words, each of length $2^2 = 4$; at step 3, there are $\alpha_3 = 6$ words, each of length $2^3 = 8$. The corresponding entropy is

$$\begin{aligned} \lim_{n\to\infty} \frac{\ln(\alpha_n)}{2^n} &= \sum_{k=1}^{\infty} \frac{1}{2^{k+1}} \ln\left(1 + \frac{\alpha_{k-1}}{\alpha_k}\right) \\ &= (0.3547882102...)\ln(2). \end{aligned}$$

Imagine now replacing the symbol 2 in the preceding by the empty symbol. We obtain

$$\begin{cases} 0 \to 01 \\ 1 \to \begin{cases} 10 & \text{with probability } 1/2, \\ 0 & \text{with probability } 1/2, \end{cases} \end{cases}$$

which is recognized as an "intertwining" of the Prouhet–Thue–Morse and Fibonacci substitutions [5]. The set of possible words at step n is $\{0110, 010\}$ at $n = 2$ and

$$\{01101001, 0110001, 011001, 0101001, 010001, 01001\}$$

at $n = 3$. The sequence $\{\alpha_n\}$ remains relevant, but unfortunately the word lengths are no longer consistent. Because the word lengths are 2^n at most, we deduce that the entropy is $\geq (0.3547882102...)\ln(2)$. More precise bounds would be good to see someday.

More examples are found in [5, 7–9]. Let $\varphi = (1 + \sqrt{5})/2$ be the Golden mean [10]. Starting with 0, the substitution [11]

$$\begin{cases} 0 \to 02324 \\ 1 \to 32324 \\ 2 \to 323 \\ 3 \to 12324 \\ 4 \to 12323 \end{cases}$$

gives rise to 023243231232432312323.... Rewriting every positive digit via

$$1 \to ++, \quad 2 \to +-, \quad 3 \to -+, \quad 4 \to --$$

we obtain $0 + - + - - + - + + - + - - + - + + - + - + \ldots$ which turns out to be identical to the sequence

$$\varepsilon_n = \operatorname{sgn}\left(\sin\left(\frac{2\pi n}{\varphi^2}\right)\right) = \begin{cases} + & \text{if } \{n/\varphi^2\} < 1/2 \\ - & \text{if } \{n/\varphi^2\} > 1/2 \\ 0 & \text{if } n = 0 \end{cases}$$

where $\{x\}$ denotes the fractional part of $x > 0$. Letting

$$S(N) = \sum_{n=1}^{N} \varepsilon_n, \quad \Sigma(N) = \frac{1}{N}\sum_{n=1}^{N} S(n)^2$$

it appears that

$$\max_{1 \le n \le N} S(n) \sim -\min_{1 \le n \le N} S(n) \sim \frac{1}{6\ln(\varphi)} \ln(N)$$

as $N \to \infty$, but the existence and identity of $\lim_{N\to\infty} \Sigma(N)/\ln(N)$ remain open. This circle of ideas reminds us of the following question: is the series

$$\sum_{n=1}^{\infty} (-1)^n \frac{|\sin(n)|}{n}$$

convergent? The answer is yes; its delicate proof is connected with Diophantine approximation [12]. Another self-similar sequence appears in [13] (in a different context). See [14, 15] for related material.

4.24.1 *Penrose–Robinson Tilings*

Penrose [16–18] discovered a famous tiling of the plane that is nonperiodic and generated by two types of rhombi with equal edge length (one with acute angle $\pi/5$ and the other with acute angle $2\pi/5$). Bisecting the rhombi across the obtuse angles gives the Robinson triangles P and Q in Figure 4.5. More on this decomposition (P is also known as a Golden triangle) appears in [19–22]. Again, what can be said about the entropy if some randomness is introduced?

We proceed in close analogy with random Fibonacci words, omitting all details. Define [2, 4]

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = \begin{pmatrix} (2b_{n-1} - a_{n-1}b_{n-2})\, a_{n-1} \\ (2a_n - a_{n-1}a_{n-2}b_{n-2}^2)\, b_{n-1} \end{pmatrix} \quad \text{for } n \ge 2,$$

$$\begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.$$

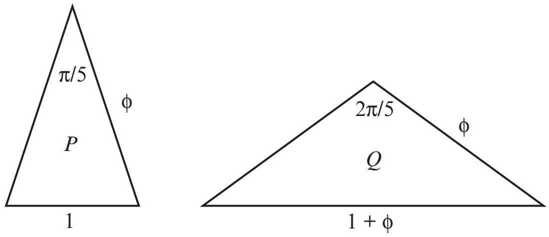

Figure 4.5 P and Q triangles.

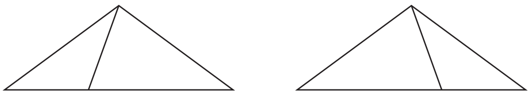

Figure 4.6 $a_1 = 2$.

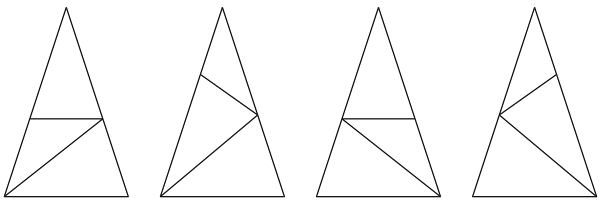

Figure 4.7 $b_1 = 4$.

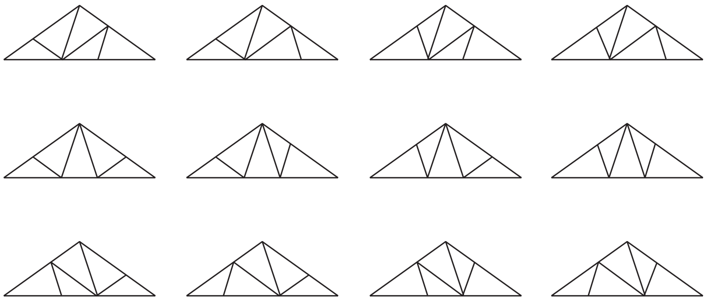

Figure 4.8 $a_2 = 12$ (four duplicates occurred among the original sixteen).

When $n = 1$, there are $a_1 = 2$ triangles of type Q, each partitioned into $f_3 = 2$ triangular subregions (Figure 4.6); next there are $b_1 = 4$ triangles of type P, each partitioned into $f_4 = 3$ subregions (Figure 4.7). When $n = 2$, there are $a_2 = 12$ triangles of type Q, each partitioned into $f_5 = 5$ subregions (Figure 4.8); next there are $b_2 = 88$ triangles of type P, each partitioned into $f_6 = 8$ subregions (not pictured). The corresponding entropy is

$$\lim_{n\to\infty} \frac{\ln(a_n)}{f_{2n+1}} = \lim_{n\to\infty} \frac{\ln(b_n)}{f_{2n+2}} = 0.606094....$$

A rapidly convergent expression for this constant would be welcome, as would a rigorous definition of *quasiperiodicity* in two dimensions.

Acknowledgments I thank Claude Godrèche and Johan Nilsson for their pictorial explanations of $a_2 = 12$ and David Wing for his helpful comments.

[1] S. R. Finch, Prouhet-Thue-Morse constant, *first volume*, pp. 436–441.
[2] C. Godrèche and J. M. Luck, Quasiperiodicity and randomness in tilings of the plane, *J. Statist. Phys.* 55 (1989) 1–28; MR1003500 (91c:52025).
[3] J. Nilsson, On the entropy of random Fibonacci words, arXiv:1001.3513.
[4] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A072042, A126023, A235857, and A235858.
[5] D. J. Wing, *Notions of Complexity in Substitution Dynamical Systems*, Ph.D. thesis, Oregon State Univ., 2011.
[6] S. R. Finch, Quadratic recurrence constants, *first volume*, pp. 443–448.
[7] J. Nilsson, On the entropy of a family of random substitutions, *Monatsh. Math.* 168 (2012) 563–577; arXiv:1103.4777; MR2993964.
[8] J. Nilsson, On the entropy of a two step random Fibonacci substitution, *Entropy* 15 (2013) 3312–3324; arXiv:1303.2526; MR3111435.
[9] M. Moll, On a family of random noble means substitutions, arXiv:1312.5136.
[10] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[11] C. Godrèche, J. M. Luck and F. Vallet, Quasiperiodicity and types of order: a study in one dimension, *J. Phys. A* 20 (1987) 4483–4499; MR0914286 (89c:82054).
[12] A. V. Kumchev, On the convergence of some alternating series, *Ramanujan J.* 30 (2013) 101–116; arXiv:1102.4644; MR3010465.
[13] S. R. Finch, Discrepancy and uniformity, *this volume*, §1.6.
[14] S. Aubry, C. Godrèche and F. Vallet, Incommensurate structure with no average lattice: an example of a one-dimensional quasicrystal, *J. Physique* 48 (1987) 327–334; MR0885441 (88c:82035).
[15] J. M. Luck, C. Godrèche, A. Janner and T. Janssen, The nature of the atomic surfaces of quasiperiodic self-similar structures, *J. Phys. A* 26 (1993) 1951–1999; MR1220802 (94a:82049).
[16] R. Penrose, The role of aesthetics in pure and applied mathematical research, *Bull. Instit. Math. Appl.* 10 (1974) 266–271; *The Physics of Quasicrystals*, ed. P. J. Steinhardt and S. Ostlund, World Scientific, 1987, pp. 655–661; MR0989615 (90b:82003).
[17] M. Gardner, Extraordinary nonperiodic tiling that enriches the theory of tiles, *Sci. Amer.*, v. 236 (1977) n. 1, 110–121; *Penrose Tiles to Trapdoor Ciphers*, W. H. Freeman, 1989, pp. 1–12; MR0968892 (89m:00006).
[18] R. Penrose, Pentaplexity: a class of nonperiodic tilings of the plane, *Eureka* 39 (1978) 16–22; *Math. Intelligencer* 2 (1979/80) 32–37; MR0558670 (81d:52012).
[19] J. Peyrière, Frequency of patterns in certain graphs and in Penrose tilings, *J. Physique* 47 (1986) C3-41–C3-62.
[20] B. Grünbaum and G. C. Shephard, *Tilings and Patterns*, W. H. Freeman, 1987, pp. 531–549; MR0857454 (88k:52018).
[21] M. Senechal, *Quasicrystals and Geometry*, Cambridge Univ. Press, 1995, pp. 170–206, 244–247; MR1340198 (96c:52038).
[22] N. P. Frank, A primer of substitution tilings of the Euclidean plane, *Expo. Math.* 26 (2008) 295–326; MR2462439 (2010d:52047).

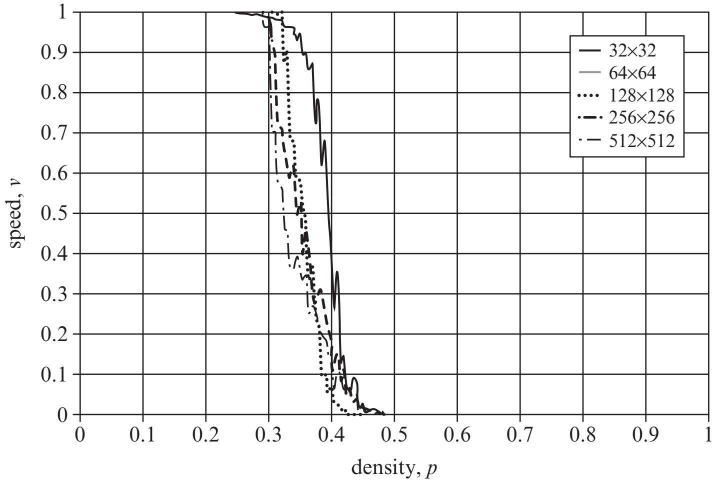

Figure 4.9 The phase transition from freely-flowing to fully-jammed is not sharp, even for large n. We desire more precise estimates of both upper and lower critical densities, as functions of p and n. Another plot in [2] gives not the mean of v over realizations, but the standard deviation (with well-defined peak).

4.25 Biham–Middleton–Levine Traffic

Consider two types of cars, red (east-bound) and blue (north-bound) which populate a two-dimensional $n \times n$ square lattice with periodic boundary conditions. Each lattice site is in one of three states: empty, occupied by a red car, or occupied by a blue car. The cars are initially distributed independently and uniformly at random over the lattice sites with spatial density p, implying that at each site,

$$\mathrm{P}\,(\text{red car}) = p/2, \quad \mathrm{P}\,(\text{blue car}) = p/2, \quad \mathrm{P}\,(\text{empty}) = 1 - p.$$

This is the only indeterminate step within the traffic model [1].

Time is integer-valued. At each time point, two steps occur, one immediately following the other. First, all red cars simultaneously attempt to move one lattice site to the east. If the site east of a red car is currently empty, it advances; otherwise it is blocked (even if the east site is becoming empty). Second, all blue cars simultaneously attempt to move one lattice site to the north. If the site north of a blue car is currently empty, it advances; otherwise it is blocked (even if the north site is becoming empty).

The velocity v of the system at each time t is the ratio between the number of cars that successfully moved and the total number of cars. If $v = 0$, then no car has moved at t; if $v = 1$, then all the cars have moved. The dependence of v for large t on both p and n is exceedingly interesting – see Figure 4.9 – depicted is an average of v over many realizations and over a large time interval [2].

Early in the study of this particular traffic model, it was thought that the phase transition exhibited by v would be comparable to other famous systems in statistical mechanics (for example, percolation). Such a belief seems, however, not to be supported by computer simulation. Intermediate stable phases, where regions of gridlock coexist with bands of unrestricted movement, seem to form effortlessly for $32 \leq n \leq 512$ [3, 4]. No one knows what truly happens as $n \to \infty$. Do such critical intervals slowly cascade to $p = 0$ in the limit? Or do they remain intact and disjoint from $p = 0$?

Additional references [5–12] cover both theoretical and experimental aspects of the subject.

[1] O. Biham, A. A. Middleton and D. Levine, Self-organization and a dynamical transition in traffic flow models, *Phys. Rev. A* 46 (1992) R6124–R6127.

[2] D. Cámpora, J. de la Torre, J. C. García Vázquezc and F. Sancho Caparrini, BML model on non-orientable surfaces, *Physica A* 389 (2010) 3290–3298.

[3] R. M. D'Souza, Coexisting phases and lattice dependence of a cellular automata model for traffic flow, *Phys. Rev. E* 71 (2005) 066112; arXiv:cond-mat/0402351.

[4] R. M. D'Souza, BML revisited: statistical physics, computer simulation, and probability, *Complexity* 12 (2006) 30–39.

[5] T. Nagatani, Anisotropic effect on jamming transition in traffic-flow model, *J. Phys. Soc. Japan* 62 (1993) 2656–2662.

[6] M. Fukui and Y. Ishibashi, Evolution of traffic jam in traffic flow model, *J. Phys. Soc. Japan* 62 (1993) 3841–3844.

[7] H. F. Chau, K. Y. Wan and K. K. Yan, An improved upper bound for critical car density of the two-dimensional Biham-Middleton-Levine traffic model, *Physica A* 254 (1998) 117–121.

[8] Y. Shi, Self-organization and phase transition in two-dimensional traffic-flow model, *Commun. Theor. Phys.* 31 (1999) 85–90; arXiv:cond-mat/9912386; MR1751791.

[9] O. Angel, A. E. Holroyd and J. B. Martin, The jammed phase of the Biham-Middleton-Levine traffic model, *Elec. Comm. Probab.* 10 (2005) 167–178; arXiv:math/0504001; MR2162816 (2006f:60090).

[10] T. D. Austin and I. Benjamini, For what number of cars must self organization occur in the Biham-Middleton-Levine traffic model from any possible starting configuration? arXiv:math/0607759.

[11] I. Benjamini, O. Gurel-Gurevich and R. Izkovsky, The Biham-Middleton-Levine traffic model for a single junction, *In and Out of Equilibrium. 2*, Proc. 10^{th} Brazilian School of Probability, Rio de Janeiro, 2006, ed. V. Sidoravicius and M. Eulália Vares, Birkhäuser, 2008, pp. 97–116; arXiv:math/0703201; MR2463481 (2009i:60001).

[12] S. Maerivoet and B. De Moor, Cellular automata models of road traffic, *Phys. Rep.* 419 (2005) 1–64; MR2174827.

4.26 Contact Processes

A one-dimensional **contact process** is a continuous-time Markov process on the lattice $\mathbb{Z}$ of integers. The state at time t is given by a set $\eta_t \subseteq \mathbb{Z}$ of the lattice sites which we visualize as being occupied by particles. The system evolves as follows:

- if $x \in \eta_t$, then x becomes vacant at rate 1
- if $x \notin \eta_t$, then x becomes occupied at rate $f(N_x)$

where $N_x = |\eta_t \cap \{x-1, x+1\}|$ is the number of nearest-neighbor sites that are occupied,

$$f(N) = \begin{cases} 0 & \text{if } N=0, \\ \lambda & \text{if } N=1, \\ 2\lambda & \text{if } N=2 \end{cases}$$

and $\lambda > 0$ is a fixed parameter. This process is a simple model of the spread of an infectious disease [1–5]. An individual at $x \in \mathbb{Z}$ is infected if $x \in \eta_t$ and healthy if $x \notin \eta_t$. Healthy individuals become infected at a rate which is proportional to the number of infected neighbors. Infected individuals recover at rate 1.

As λ increases from zero, the contact process undergoes an extinction–survival phase transition. There is a unique critical threshold λ_c such that $\lambda < \lambda_c$ implies $\eta_t = \emptyset$ for large t almost surely, whereas $\lambda > \lambda_c$ implies $\eta_t \neq \emptyset$ for all t almost surely. The best rigorous bounds for λ_c are $1.5517 < \lambda_c < 1.9412$ [6–11]; the best non-rigorous numerical estimate is

$$\lambda_c = 1.64892... = \frac{1}{2}(3.29784...) = \frac{1}{2}\frac{1}{0.30322...} = \frac{1}{0.60645...}$$

obtained via numerical means/simulation [12–17] and via lengthy series expansions [18–20].

One variation on the preceding is to replace $f(N_x)$ by $g(N_x)$, where

$$g(N) = \begin{cases} 0 & \text{if } N=0, \\ \lambda & \text{if } N=1, \\ \lambda & \text{if } N=2 \end{cases}$$

and λ_c here is $1.74173... = 1/0.57414...$. Another variation is to replace $f(N_x)$ by $h(N_x)$, where

$$h(N) = \begin{cases} 0 & \text{if } N=0, \\ \lambda/4 & \text{if } N=1, \\ \lambda & \text{if } N=2 \end{cases}$$

and λ_c here is $6.17066 = 1/0.16205...$. No closed-form expressions are known for any of these critical thresholds [21–25].

Such models are often referred to as *interacting particle systems* or asynchronously-updated *probabilistic cellular automata*. Our opening example (f) is often called the *basic* contact process and is clearly connected to epidemiology and ecology [26–28]. In statistical physics, it is closely related to Schlögl's first model of an autocatalytic chemical reaction, to directed percolation in two dimensions, and to Reggeon field theory. The other examples are associated with the poisoning of a catalytic surface (g) and the testing of an order-parameter exponent universality conjecture (h). To describe the latter idea – that a certain exponent $\beta = 0.277...$ is valid for a wide class of nonequilibrium systems with phase transition – would take us too far afield [19, 25, 29–31].

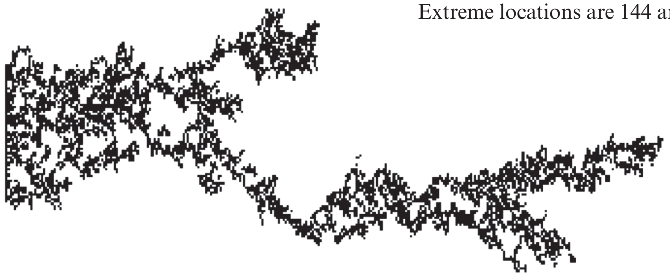

Figure 4.10 Subcritical example ($M = 180$, $N = 540$, $\lambda < \lambda_c$).

4.26.1 *Implementation*

The following discussion is based on what is called the *graphical representation* of the basic contact process [32–34]. Let M be a large positive integer. For every integer $1 \leq x \leq M$, let $\{t_n^x : n \geq 1\}$ be the arrival times of a Poisson process with rate 1. For every integer $1 \leq x \leq M - 1$, let $\{u_n^x : n \geq 1\}$ be the arrival times of a Poisson process with rate λ. Likewise, for every integer $2 \leq x \leq M$, let $\{v_n^x : n \geq 1\}$ be the arrival times of a Poisson process with rate λ. To generate times v_n^x up to a large value N, for example, simply generate a single random integer K via Poisson(λN), then generate K Uniform$[0, N]$ random values and sorted in increasing order [35]. Of course, K will usually be different for each x.

Let

$$W = \bigcup_{1 \leq x \leq M} \{t_n^x : n \geq 1\} \;\cup \bigcup_{1 \leq x \leq M-1} \{u_n^x : n \geq 1\} \;\cup \bigcup_{2 \leq x \leq M} \{v_n^x : n \geq 1\}$$

be sorted in increasing order, keeping track for each value the corresponding site x and whether it arose as a t, u or v. The event that two values coincide exactly has probability zero. The list W captures all changes occurring on the finite lattice $[1, M]$ over the finite time interval $[0, N]$.

Without loss of generality, assume M is divisible by 3. Figures 4.10 and 4.11 are constructed with initial state taken to be the binary M-vector

$$\xi_0 = \left\{ \underbrace{0, 0, \ldots, 0}_{M/3}, \overbrace{1, 1, \ldots, 1}^{M/3}, \underbrace{0, 0, \ldots, 0}_{M/3} \right\}$$

which serves as an indicator for the set η_0. Now select the first element w in the list W. If w arose as a t, then place a 0 at x (there is a death at x if x is occupied). If w arose as a u and if there is a 1 at x, then place a 1 at $x + 1$ (there is a birth at $x + 1$ if x is occupied and $x + 1$ is vacant). If w arose as a v and if there is a 1 at x, then place a 1 at $x - 1$ (there is a birth at $x - 1$ if x is occupied and $x - 1$ is vacant). This gives ξ_w and hence η_w. Now select the second element in W and

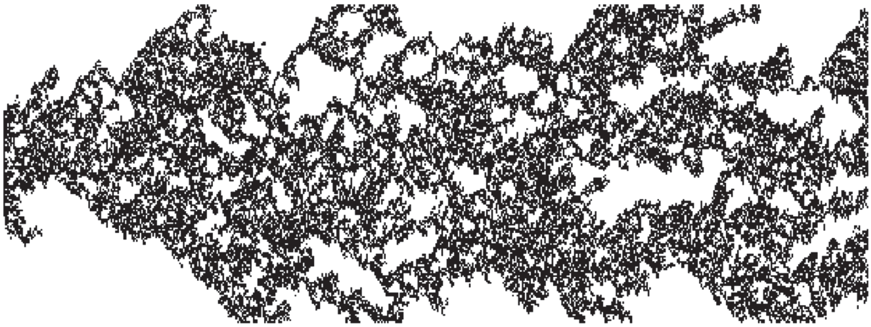

Figure 4.11 Supercritical example ($M = 180$, $N = 540$, $\lambda > \lambda_c$).

continue similarly until either the list is exhausted or all ξ_w are 0s (the vacuum state is absorbing) Figures 4.10 and 4.11 exhibit only a subsample of states, one per unit time. The vertical axis is space ($1 \le x \le M$) and the horizontal axis is time ($0 \le t \le N$).

In closing, we mention rigorous bounds $0.3597 < \lambda_c < 0.79$ for the contact process in two spatial dimensions [4, 36, 37], as well as a non-rigorous estimate $\lambda_c \approx 0.412$ [10, 12]. Every lattice site here has four nearest neighbors, complicating the analysis. Revisited calculations [37] of the upper bound 0.79 here would be good to see someday, as well as series expansions [38] giving precise results earlier in one dimension.

4.26.2 *Discrete Time Analog*

An exceedingly simple model, described in [26], deserves further study. The time interval $[0, N]$ from earlier is replaced by $\{0, 1, \ldots, N\}$; we need "collision rules" to decide the outcome when several events occur simultaneously in space and time.

For every integer $1 \le x \le M$, let $\{t_n^x : n \ge 1\}$ be the arrival times of a Bernoulli process with rate γ. Hence each t_n^x corresponds to a biased coin toss yielding heads. For every integer $1 \le x \le M - 1$, let $\{u_n^x : n \ge 1\}$ correspond to the coin tosses yielding tails. Likewise, for every integer $2 \le x \le M$, let $\{v_n^x : n \ge 1\}$ correspond to the coin tosses yielding tails. Note that only one Bernoulli process is involved here for each x, not three independent Poisson processes as before.

Form the multilist W as before – many coincident values appear here unlike before – keeping track for each value the corresponding site x and whether it arose as a t, u or v. Take the initial state ξ_0 as before. Select all the elements w in the multilist W equal to 1. First, for each $w = 1$ arising as a t, assign a 0 at x (there is a death at x if x is occupied). This gives a provisional state, called ξ_1, and we make a copy, called ξ_1', on which further changes are written. Second, for each $w = 1$ arising as a u, if there is a 1 at x, then assign a 1 at $x' + 1$ (there is a birth at $x' + 1$ if x is occupied and $x' + 1$ is vacant). Third, for each $w = 1$ arising as a v, if there is a 1 at x, then assign a 1 at $x' - 1$ (there is a birth at $x' - 1$ if x is occupied and $x' - 1$ is vacant). Finally, overwrite ξ_1 by ξ_1'. Now continue with all elements w in W equal to 2, assign deaths followed by births, and so forth.

Durrett & Levin [26] estimated the critical threshold γ_c to be approximately 0.47 for large M and N. A more accurate estimate is highly desirable!

4.26.3 *Oriented or Directed Percolation*

The graphs of one-dimensional discrete-time contact processes bear resemblance to two-dimensional percolation [39]. More precisely, they are similar to the *oriented* or *directed* case of percolation in which fluid must flow either north or east [2, 40–43]. For both bonds and sites, there exist critical probabilities p_{cb} and p_{cs} below which all clusters are finite and above which an infinite cluster must exist. No closed-form expressions are known in this case (unlike ordinary percolation). Without giving any details, we have rigorous bounds on bond critical probability [37, 44–53]

$$0.6383 \leq p_{cb} \leq 2/3;$$

rigorous bounds on site critical probability [37, 46, 47, 49, 50, 52, 54]

$$0.6977 \leq p_{cs} \leq 0.7491;$$

and numerical estimates [29, 55–65]

$$p_{cb} = 0.64470018... = 1 - 0.35529982..., \quad p_{cs} = 0.7054852... = 1 - 0.2945148...$$

for the square lattice. Different probabilities apply for the triangular and hexagonal (honeycomb) lattices in $\mathbb{R}^2$ as well as for the cubic lattice in $\mathbb{R}^3$. A percolation-theoretic analog of the connective constant for self-avoiding walks [66] is investigated in [67].

[1] T. E. Harris, Contact interactions on a lattice, *Annals Probab.* 2 (1974) 969–988; MR0356292 (50 #8763).

[2] R. Durrett, *Lecture Notes on Particle Systems and Percolation*, Wadsworth & Brooks/Cole, 1988, pp. 50–105; MR0940469 (89k:60157).

[3] T. M. Liggett, *Interacting Particle Systems*, Springer-Verlag, 1985, pp. 264–314; MR0776231 (86e:60089).

[4] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, Springer-Verlag, 1999, pp. 31–78, 125–134; MR1717346 (2001g:60247).

[5] T. M. Liggett, T. E. Harris' contributions to interacting particle systems and percolation, *Annals Probab.* 39 (2011) 407–416; arXiv:1103.1988; MR2789500 (2012a:60274).

[6] N. Konno, *Lecture Notes on Interacting Particle Systems*, Dept. of Math., Kobe Univ., 1997.

[7] R. Holley and T. M. Liggett, The survival of contact processes. *Annals Probab.* 6 (1978) 198–206; MR0488379 (58 #7927).
[8] T. M. Liggett, Improved upper bounds for the contact process critical value, *Annals Probab.* 23 (1995) 697–723; MR1334167 (96d:60157).
[9] H. Ziezold and C. Grillenberger, On the critical infection rate of the one-dimensional basic contact process: numerical results, *J. Appl. Probab*. 25 (1988) 1–8; MR0929499 (89d:60198).
[10] M. Katori and N. Konno, Correlation inequalities and lower bounds for the critical value λ_c of contact processes, *J. Phys. Soc. Japan* 59 (1990) 877–887; MR1055279 (91h:60114).
[11] A. Sudbury, Rigorous lower bounds for the critical infection rate in the diffusive contact process, *J. Appl. Probab.* 38 (2001) 1074–1078; MR1876561 (2002j:60187).
[12] R. C. Brower, M. A. Furman and M. Moshe, Critical exponents for the Reggeon quantum spin model, *Phys. Lett. B* 76 (1978) 213–219.
[13] P. Grassberger and A. de la Torre, Reggeon field theory (Schlögl's first model) on a lattice: Monte Carlo calculations of critical behaviour, *Annals of Phys*. 122 (1979) 373–396.
[14] N. Konno and M. Katori, Applications of the CAM based on a new decoupling procedure of correlation functions in the one-dimensional contact process, *J. Phys. Soc. Japan* 59 (1990) 1581–1592.
[15] R. Gouet, F. J. López and G. Sanz, Estimation of critical values in interacting particle systems, *J. Appl. Probab.* 37 (2000) 118–125; MR1761665 (2001e:60207).
[16] R. Dickman and M. M. de Oliveira, Quasi-stationary simulation of the contact process, *Physica A* 357 (2005) 134–141; arXiv:cond-mat/0411610.
[17] M. M. de Oliveira and R. Dickman, Quasi-stationary simulation: the subcritical contact process, *Brazilian J. Phys.* v. 36 (2006) n. 3A, 685–689; arXiv:cond-mat/0601163.
[18] R. Dickman, Nonequilibrium lattice models: series analysis of steady states, *J. Stat. Phys.* 55 (1989) 997–1026; MR1002482 (90h:82026).
[19] I. Jensen and R. Dickman, Time-dependent perturbation theory for nonequilibrium lattice models, *J. Stat. Phys.* 71 (1993) 89–127; MR1215851 (94c:82055).
[20] M. J. de Oliveira, Perturbative series expansion for the subcritical stationary properties of the contact process, *J. Phys. A* 39 (2006) 11131–11144; MR2275114 (2007h:82084).
[21] R. Dickman and M. A. Burschka, Nonequilibrium critical poisoning in a single-species model, *Phys. Lett. A* 127 (1988) 132–137.
[22] A. L. C. Ferreira and S. K. Mendiratta, Mean-field approximation with coherent anomaly method for a non-equilibrium model, *J. Phys. A* 26 (1993) L145–L150.
[23] M. Katori and N. Konno, Bounds for the critical line of the θ-contact processes with $1 \le \theta \le 2$, *J. Phys. A* 26 (1993) 6597–6614; MR1253878 (95a:60146).
[24] E. Ben-Naim and P. L. Krapivsky, Cluster approximation for the contact process, *J. Phys. A* 27 (1994) L481–L487; arXiv:cond-mat/9403079.
[25] I. Jensen and R. Dickman, Series analysis of the generalized contact process, *Physica A* 203 (1994) 175–188.
[26] R. Durrett and S. A. Levin, Stochastic spatial models: a user's guide to ecological applications, *Phil. Trans. Royal Soc. London B* 343 (1994) 329–350.
[27] R. Durrett, Stochastic spatial models, *SIAM Rev.* 41 (1999) 677–718; MR1722998 (2000h:60086).
[28] S. Krone and Y. Guan, Stochastic Spatial Simulator, Univ. of Idaho.

[29] J. W. Essam, A. J. Guttmann and K. De'Bell, On two-dimensional directed percolation, *J. Phys. A* 21 (1988) 3815–3832; MR0968312 (89i:82049).
[30] J. R. G. de Mendonça, Precise critical exponents for the basic contact process, *J. Phys. A* 32 (1999) L467–L473.
[31] A. Yu. Tretyakov, N. Inui and N. Konno, Phase transition for the one-sided contact process, *J. Phys. Soc. Japan* 66 (1997) 3764–3769.
[32] T. E. Harris, Additive set-valued Markov processes and graphical methods, *Annals Probab.* 6 (1978) 355–378; MR0488377 (58 #7925).
[33] D. Griffeath, The basic contact processes, *Stochastic Process. Appl.* 11 (1981) 151–185; MR0616064 (82i:60174).
[34] R. Durrett, The contact process, 1974–1989, *Mathematics of Random Media* , Proc. 1989 Blacksburg conf., ed. W. E. Kohler and B. S. White, Amer. Math. Soc., 1991, pp. 1–18; MR1117232 (92k:60223).
[35] K. Sigman, Notes on the Poisson process, lecture notes (2009).
[36] D. Griffeath, Ergodic theorems for graph interactions, *Adv. Appl. Probab.* 7 (1975) 179–194; MR0397926 (53 #1781).
[37] A. M. Stacey, *Bounds on the Critical Probabilities in Oriented Percolation Models*, Ph.D. thesis, Cambridge Univ., 1994.
[38] I. Jensen, *Computer Simulations and Analytical Studies of Nonequilibrium Phase Transitions in Interacting Particle Systems with Absorbing States*, Institute of Physics, Univ. of Århus, Denmark, 1993.
[39] S. R. Finch, Percolation cluster density constants, *first volume*, pp. 371–378.
[40] R. Durrett, Oriented percolation in two dimensions, *Annals Probab.* 12 (1984) 999–1040; MR0757768 (86g:60117).
[41] S. R. Broadbent and J. M. Hammersley, Percolation processes. I, Crystals and mazes, *Proc. Cambridge Philos. Soc.* 53 (1957) 629–641; MR0091567 (19,989e).
[42] J. M. Hammersley, Percolation processes. II, The connective constant, *Proc. Cambridge Philos. Soc.* 53 (1957) 642–645; MR0091568 (19,989f).
[43] H. Hinrichsen, On possible experimental realizations of directed percolation, *Brazilian J. Phys.* v. 30 (2000) n. 1, 69–82; arXiv:cond-mat/9910284.
[44] J. M. Hammersley, Percolation processes: Lower bounds for the critical probability, *Annals Math. Statist.* 28 (1957) 790–795; MR0101564 (21 #374).
[45] J. G. Mauldon, Asymmetric oriented percolation on a plane, *Fourth Berkeley Symp. Math. Stat. Probab.*, v. 2, ed. J. Neyman, Univ. of Calif. Press, 1961, pp. 337–345; MR0133892 (24 #A3716).
[46] J. Bishir, A lower bound for the critical probability in the one-quadrant oriented-atom percolation process., *J. Royal Statist. Soc. Ser. B* 25 (1963) 401–404; MR0168047 (29 #5312).
[47] L. Gray, J. C. Wierman and R. T. Smythe, Lower bounds for the critical probability in percolation models with oriented bonds, *J. Appl. Probab.* 17 (1980) 979–986; MR0587198 (81k:60104).
[48] D. Dhar, Diode-resistor percolation in two and three dimensions, I. Upper bounds on critical probability, *J. Phys. A* 15 (1982) 1849–1858.
[49] P. Balister, B. Bollobás and A. Stacey, Upper bounds for the critical probability of oriented percolation in two dimensions, *Proc. Royal Soc. London Ser. A* 440 (1993) 201–220; MR1232833 (94e:60085).
[50] P. Balister, B. Bollobás and A. Stacey, Improved upper bounds for the critical probability of oriented percolation in two dimensions, *Random Structures Algorithms* 5 (1994) 573–589; MR1293080 (95g:60141).

[51] B. Bollobás and A. Stacey, Approximate upper bounds for the critical probability of oriented percolation in two dimensions based on rapidly mixing Markov chains, *J. Appl. Probab.* 34 (1997) 859–867; MR1484020 (99j:60152).
[52] T. M. Liggett, Survival of discrete time growth models, with applications to oriented percolation, *Annals Appl. Probab.* 5 (1995) 613–636; MR1359822 (97c:60239).
[53] C. E. M. Pearce, Oriented bond percolation and phase transitions: an analytic approach, *Numerical Analysis and Applied Mathematics*, Proc. 2007 Corfu conf., ed. T. E. Simos, G. Psihoyios and Ch. Tsitouras, Amer. Instit. Phys., pp. 697-698.
[54] C. E. M. Pearce and F. K. Fletcher, Oriented site percolation, phase transitions and probability bounds, *J. Inequal. Pure Appl. Math.* 6 (2005) A 135; MR2191604 (2007d:60067).
[55] J. Blease, Pair-connectedness for directed bond percolation on some two-dimensional lattices by series methods, *J. Phys. C* 10 (1977) 3461–3476.
[56] J. Kertész and T. Vicsek, Oriented bond percolation, *J. Phys. C* 13 (1980) L343–L348.
[57] W. Kinzel and J. M. Yeomans, Directed percolation: a finite-size renormalisation group approach, *J. Phys. A* 14 (1981) L163–L168.
[58] D. Dhar and M. Barma, Monte Carlo simulation of directed percolation on a square lattice, *J. Phys. C* 14 (1981) Ll–L6.
[59] K. De'Bell and J. W. Essam, Directed percolation: mean field theory and series expansions for some two-dimensional lattices, *J. Phys. A* 16 (1983) 385–404; MR0701345 (85a:82051).
[60] K. De'Bell and J. W. Essam, Estimates of the site percolation probability exponents for some directed lattices, *J. Phys. A* 16 (1983) 3145–3147.
[61] R. N. Onody and U. P. C. Neves, Series expansion of the directed percolation probability, *J. Phys. A* 25 (1992) 6609–6615; MR1210881 (93m:82017).
[62] I. Jensen and A. J. Guttmann, Series expansions of the percolation probability for directed square and honeycomb lattices, *J. Phys. A* 28 (1995) 4813–4833; arXiv:cond-mat/9509121; MR1353145 (96g:82024).
[63] I. Jensen and A. J. Guttmann, Series expansions for two-dimensional directed percolation, *Nuclear Physics B* (Proc. Suppl.) 47 (1996) 835–837.
[64] I. Jensen, Low-density series expansions for directed percolation on square and triangular lattices, *J. Phys. A* 29 (1996) 7013–7040; MR1421922.
[65] I. Jensen, Low-density series expansions for directed percolation, I. A new efficient algorithm with applications to the square lattice, *J. Phys. A* 32 (1999) 5233–5249; arXiv:cond-mat/9906036; MR1720335 (2000g:82016).
[66] S. R. Finch, Self-avoiding walk constants, *first volume*, pp. 331–338.
[67] P. Balister, B. Bollobás and A. Stacey, Counting boundary paths for oriented percolation clusters, *Random Structures Algorithms* 14 (1999) 1–28; MR1662266 (99k:60243).

4.27 Interpolating between Max and Sum

Consider the stochastic process [1]

$$X_{n+1} = \max\left\{\alpha\,\beta\,X_n + Y_n, \beta\,X_n\right\}, \quad n = 1, 2, 3, \ldots$$

where $0 \leq \alpha \leq 1$, $0 < \beta < 1$, $X_1 \geq 0$ are constants and Y_1, Y_2, Y_3, … are nonnegative, independent, identically distributed random variables. What can be said about the long-range mean

$$\mu = \lim_{n\to\infty} \mathrm{E}(X_n)$$

if Y_n is Uniform[0, 1] or if Y_n is Exponential(1)?

Let $\beta = 1/2$ for concreteness. In the special case when $\alpha = 0$, we have [2]

$$\mu = \sum_{n=0}^{\infty} \frac{\beta^{n(n+3)/2}}{(n+1)(n+2)} = 0.5443705469... = \frac{1.0887410938...}{2}$$

for the uniform scenario and

$$\begin{aligned}\mu &= 1 + \sum_{n=1}^{\infty} \beta^n \int_0^{\infty} e^{-x}\left(1-e^{-\beta x}\right)\left(1-e^{-\beta^2 x}\right)\left(1-e^{-\beta^3 x}\right)\cdots\left(1-e^{-\beta^n x}\right) dx \\ &= 1.1962832643...\end{aligned}$$

for the exponential scenario. For the latter, when $\alpha > 0$, set $\gamma = 1/\beta$, $\delta = 1 - \alpha$ and define recursively

$$m_k = \begin{cases} \gamma\, m_{k/2} & \text{if } k \text{ is even,} \\ \gamma\, m_{(k-1)/2} + \delta & \text{if } k \text{ is odd,} \end{cases} \qquad m_1 = 1,$$

$$d_k = \begin{cases} d_{k/2} & \text{if } k \text{ is even,} \\ \dfrac{m_{(k-1)/2}\, d_{(k-1)/2}}{\alpha\beta - m_{(k-1)/2}} & \text{if } k \text{ is odd,} \end{cases} \qquad d_1 = 1.$$

Set also $p = \lfloor \ln(2)/\ln(\gamma) \rfloor + 1$. We have [3]

$$\mu = \lim_{n\to\infty} \frac{\sum_{k=1}^{n\,2^p-1} d_k/m_k}{\sum_{k=1}^{n\,2^p-1} d_k} = \begin{cases} 1.3749080780... & \text{if } \alpha = 2/5, \\ 1.6972298042... & \text{if } \alpha = 4/5 \end{cases}$$

but wonder whether the two upper summation limits can be simplified. (Analogous formulas for the uniform scenario are not known; numerical bounds are available [2]:

$$\begin{cases} 1.297 < \mu < 1.345 & \text{if } \alpha = 2/5, \\ 1.678 < \mu < 1.690 & \text{if } \alpha = 4/5 \end{cases}$$

although fairly loose.) As $\beta \to 1^-$, convergence becomes slower [3]; it would be good to understand the corresponding rate at which $\mu \to \infty$.

Consider now the stochastic process [4]

$$X_{n+1} = \max\left\{\alpha_n X_n + Y_n, X_n\right\}, \quad n = 1, 2, 3, \ldots$$

where $\alpha_n = 1 - 1/n$ and $X_1, Y_1, Y_2, Y_3, \ldots$ are as before. Hence $\alpha_n \to 1$ while what we called β before is fixed at 1. This process interpolates between finding a maximum ($\alpha_n \equiv 0$) and calculating a sum ($\alpha_n \equiv 1$). It is not surprising that $\mu \to \infty$ under the circumstances. More precisely, if $\mathrm{E}\,|Y_1| < \infty$, then

$$\frac{X_n}{n} \to \theta$$

almost surely, where θ is the unique solution of the remarkable equation

$$\theta = \mathrm{E}\,(\max\,(0,\, Y_1 - \theta)).$$

No examples are provided in [4], thus the following results are new. If Y_1 is Uniform$[0, 1]$, $Z = Y_1 - \theta$ and $0 < \theta < 1$, we obtain

$$\mathrm{P}\,(\max\,(0, Z) > 0) = 1 - \theta, \quad \mathrm{P}\,(\max\,(0, Z) = 0) = \theta$$

(a mixed distribution: partly discrete, partly continuous). It follows that

$$\theta = \mathrm{E}\,(\max\,(0, Z)) = 0 \cdot \theta + \int_0^{1-\theta} z\,dz = \frac{1}{2}(1 - \theta)^2$$

and thus $\theta = 2 - \sqrt{3} = 0.2679491924\ldots$. If instead Y_1 is Exponential(1) and $\theta > 0$, we obtain

$$\mathrm{P}\,(\max\,(0, Z) > 0) = e^{-\theta}, \quad \mathrm{P}\,(\max\,(0, Z) = 0) = 1 - e^{-\theta}.$$

It follows that

$$\theta = \mathrm{E}\,(\max\,(0, Z)) = 0 \cdot (1 - e^{-\theta}) + \int_0^{\infty} z\,e^{-(z+\theta)}dz = e^{-\theta}$$

and thus $\theta = W(1) = 0.5671432904\ldots$, where W is the Lambert or "product log" function [5].

Under additional conditions on Y_1, a Central Limit Theorem:

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{X_n - n\,\theta}{\sigma\sqrt{n/(2c+1)}} \le t\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \exp\left(-\frac{u^2}{2}\right) du$$

is valid as $n \to \infty$, where [4]

$$c = \mathrm{P}\,(Y_1 > \theta), \quad \sigma^2 = \mathrm{Var}\,(\max\,(0,\, Y_1 - \theta)).$$

Therefore more is known for this case $\alpha_n = 1 - 1/n$, $\beta = 1$ than for the original constant $0 \le \alpha \le 1$, $0 < \beta < 1$ case [1]. Also, the requirement that Y_1 is nonnegative can be lifted somewhat.

Different formulation applies if instead $\alpha_n = 1 - 1/n^{\ell}$, $\beta = 1$ for some $\ell > 1$. Let

$$\psi(\theta) = \mathrm{E}\,(\max\,(0,\, Y_1 - \theta))$$

and define recursively

$$a_{k+1} = a_k + \psi\left(\frac{a_k}{k^\ell}\right), \quad a_1 = 0.$$

If $\mathrm{E}\,|Y_1| < \infty$, then

$$\frac{X_n}{n} \to \psi(0) = \begin{cases} 1/2 & \text{if } Y_n \text{ is Uniform}[0, 1], \\ 1 & \text{if } Y_n \text{ is Exponential}(1) \end{cases}$$

almost surely. Under additional conditions on Y_1, a Central Limit Theorem

$$\frac{1}{\sigma\sqrt{n}}\left(X_n - \sum_{k=1}^{n} \psi\left(\frac{a_k}{k^\ell}\right)\right) \to \text{Normal}(0, 1)$$

is valid as $n \to \infty$, where $\sigma^2 = \operatorname{Var}(\max(0, Y_1))$; further,

$$\sum_{k=1}^{n} \psi\left(\frac{a_k}{k^\ell}\right) = \begin{cases} \psi(0)n + \dfrac{\psi'(0)\psi(0)}{2-\ell} n^{2-\ell} + o\left(n^{2-\ell}\right) & \text{if } 1 < \ell \le 3/2, \\ \psi(0)n + o\left(\sqrt{n}\right) & \text{if } \ell > 3/2. \end{cases}$$

We mention finally $\mathrm{MAR}(1)$ or ARMAX processes, for which addition in the classical $\mathrm{AR}(1)$ model

$$X_{n+1} = \rho X_n + Y_n,$$

$0 < \rho < 1$, is replaced by maximization [6–15]:

$$X_{n+1} = \max\{\rho X_n, Y_n\}.$$

Statistical time series procedures (for parameter estimation, prediction, and so forth) for $\mathrm{MAR}(1)$ still await careful development.

[1] J. Haslett, Problems in the stochastic storage of solar thermal energy, *Analysis and Optimisation of Stochastic Systems*, Proc. 1978 Oxford conf., ed. O. L. R. Jacobs, M. H. A. Davis, M. A. H. Dempster, C. J. Harris and P. C. Parks, Academic Press, 1980, pp. 553–562; MR0592970 (81j:93069).

[2] D. J. Daley and J. Haslett, A thermal energy storage process with controlled input, *Adv. Appl. Probab.* 14 (1982) 257–271; MR0650122 (83d:60107).

[3] G. Hooghiemstra and M. Keane, Calculation of the equilibrium distribution for a solar energy storage model, *J. Appl. Probab.* 22 (1985) 852–864; MR0808865 (87g:60086).

[4] F. den Hollander, G. Hooghiemstra, M. Keane and J. Resing, Strong law and central limit theorem for a process between maxima and sums, *Probab. Theory Related Fields* 90 (1991) 37–55; MR1124828 (93a:60030).

[5] S. R. Finch, Iterated exponential constants, *first volume*, pp. 448–452.

[6] M. T. Alpuim, An extremal Markovian sequence, *J. Appl. Probab.* 26 (1989) 219–232; MR1000283 (90j:60065).

[7] M. T. Alpuim and E. Athayde, On the stationary distribution of some extremal Markovian sequences, *J. Appl. Probab.* 27 (1990) 291–302; MR1052301 (92d:60025).

[8] M. T. Alpuim, N. A. Catkan and J. Hüsler, Extremes and clustering of nonstationary max-AR(1) sequences, *Stochastic Process. Appl.* 56 (1995) 171–184; MR1324329 (96b:60133).
[9] R. A. Davis and S. I. Resnick, Basic properties and prediction of max-ARMA processes, *Adv. Appl. Probab*. 21 (1989) 781–803; MR1039628 (91g:62069).
[10] P. A. W. Lewis and E. McKenzie, Minification processes and their transformations, *J. Appl. Probab.* 28 (1991) 45–57; MR1090446 (92c:60053).
[11] N. Inagaki, Asymptotic distribution of maximal autoregressive process with weight tending to 1, *Annals Inst. Statist. Math*. 46 (1994) 633–640; MR1325986 (97e:62120).
[12] M. A. Ancona-Navarrete and J. A. Tawn, A comparison of methods for estimating the extremal index, *Extremes* 3 (2000) 5–38; MR1815172.
[13] A. V. Lebedev, Statistical analysis of first-order MARMA processes (in Russian), *Mat. Zametki* 83 (2008) 552–558; Engl. transl. in *Math. Notes* 83 (2008) 506–511; MR2431621 (2009g:62119).
[14] M. Ferreira, On the dependence of tail behavior: a characterization of first-order max-autoregressive processes (in Russian), *Mat. Zametki* 90 (2011) 902–917; Engl. transl. in *Math. Notes* 90 (2011) 882–893; MR2962964.
[15] P. Naveau, Z. Zhang and B. Zhu, An extension of max autoregressive models, *Stat. Interface* 4 (2011) 253–266; MR2812820 (2012h:62342).

4.28 Mixing Time of Markov Chains

We will concentrate on two specific examples, leaving general theory aside. Consider the cycle Z_n (integers modulo n) as our state space. A **lazy random walk** is a particle that moves left or right, each with probability $1/4$, or remains motionless with probability $1/2$. Let us assume that the starting point is at 0. After how many time steps is the distribution of the particle close to uniform?

The transition matrix Q, whose ij^{th} element conveys the odds that the particle is at site j given it was at site i one step earlier, is

$$Q=\begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

when $n=6$. If we wish information on the odds over a separation of t (positive integer) steps, then the matrix product Q^t is required.

Let μ_t denote the first row of Q^t and ν denote the vector $(1/n, 1/n, \ldots, 1/n)$. Define

$$d(t)=\frac{1}{2}\left\|\mu_t-\nu\right\|_1,$$

one-half the L_1 norm of the vector difference (a sum of absolute values). This is called the **total variation distance**. Now define

$$t_{\text{mix}}(\varepsilon) = \min\left\{t \geq 1 : d(t) \leq \varepsilon\right\},$$

$$t_{\text{mix}} = t_{\text{mix}}(1/4)$$

the **mixing time**. For the case $n = 6$, we compute

$$\mu_3 = \begin{pmatrix} \frac{5}{16} & \frac{15}{64} & \frac{3}{32} & \frac{1}{32} & \frac{3}{32} & \frac{15}{64} \end{pmatrix},$$

$$\mu_4 = \begin{pmatrix} \frac{35}{128} & \frac{7}{32} & \frac{29}{256} & \frac{1}{16} & \frac{29}{256} & \frac{7}{32} \end{pmatrix}$$

and $d(3) = 9/32 > 0.28$, $d(4) = 27/128 < 0.22$, therefore $t_{\text{mix}} = 4$. Our interest is in the growth of t_{mix} as $n \to \infty$. It is known that [1]

$$c\,n^2 < t_{\text{mix}} \leq n^2$$

for some $c > 0$; simulation suggests that t_{mix}/n^2 approaches a constant ≈ 0.0949.

A **(non-lazy) random walk** is a particle that moves left or right, each with probability $1/2$. The transition matrix P is

$$P = \begin{pmatrix}
0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\
\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
\frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0
\end{pmatrix}$$

when $n = 7$. For technical reasons, we must restrict the cycle length n to be odd (to ensure aperiodicity). Let μ_t denote the first row of P^t and everything else be as before. For the case $n = 7$, we compute

$$\mu_8 = \begin{pmatrix} \frac{35}{128} & \frac{9}{256} & \frac{7}{32} & \frac{7}{64} & \frac{7}{64} & \frac{7}{32} & \frac{9}{256} \end{pmatrix},$$

$$\mu_9 = \begin{pmatrix} \frac{9}{256} & \frac{63}{256} & \frac{37}{512} & \frac{21}{128} & \frac{21}{128} & \frac{37}{512} & \frac{63}{256} \end{pmatrix}$$

and $d(8) = 253/896 > 0.28$, $d(9) = 223/896 < 0.24$, therefore $t_{\text{mix}} = 9$. Again, the growth rate of t_{mix} is quadratic in n; simulation suggests that t_{mix}/n^2 approaches a constant ≈ 0.1898. We also mention rigorous bounds [2–4]

$$\left(\frac{2n^2}{\pi^2} - 1\right)\ln(2) \leq t_{\text{mix}} \leq \frac{4n^2}{\pi^2}\ln(2)$$

which imply that the ratio falls between 0.14 and 0.28. Similar bounds could be determined for the lazy case. The non-lazy mixing time is at most twice the lazy mixing time, but may be less.

A remarkable equation for the lazy constant $C \approx 0.0949$ was announced in [5]:

$$\frac{1}{2}\int_0^1 \left| -1 + \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{C\pi}} \exp\left(-\frac{(x-k)^2}{C}\right) \right| dx = \frac{1}{4}$$

which gives a more accurate estimate $C = 0.0948705678....$ The justification involved passage from discrete (n-cycle) to continuous (circle), Fourier analysis, and reinterpretation of random walks as heat flow. Unfortunately the authors of [5] never completed their proof – their draft preprint is no longer available online – and we are left wondering if/how the challenging details can be brought together. It appears likely that $2C$ is the corresponding non-lazy constant, but verification remains open as well.

Setting $\varepsilon = 1/4$ is, of course, arbitrary. For many Markov chains (not our two examples), there is a more natural choice of threshold. In such scenarios, the variation distance $d(t)$ is fairly large and essentially flat for small t, then abruptly changes character and decays exponentially to zero as t increases beyond a certain point. It is believed that such *cut-off phenomena* are widespread, although they have been rigorously ascertained only sporadically (for example, riffle shuffles of 52 cards [6–9]). How are the group theoretic properties of the state space related to the existence or non-existence of a cut-off? This is a difficult question; we must often settle for the order of magnitude (as a function of n) of a possible threshold. Only rarely are these results so accurate as to yield tight bounds on the level of a constant.

On the one hand, given any $\varepsilon > 0$, the equation for $t_{\text{mix}}(\varepsilon)/n^2$ in the limit as $n \to \infty$ is the same as that for C except $1/4$ on the right-hand side is replaced by ε. For example, if $\varepsilon = 1/10$, then the limit is $0.1875465011....$

On the other hand, consider a random walk in which a particle moves left or right, each with probability $1/3$, or remains motionless with probability $1/3$. What does the heuristic in [5] predict for the value of $t_{\text{mix}}(\varepsilon)/n^2$? Intuition suggests that the variance of the walk generator is key. The walk with probabilities $\{1/4, 1/2, 1/4\}$ has variance $1/2$; the walk with probabilities $\{1/3, 1/3, 1/3\}$ has variance $2/3$; dividing $1/2$ by $2/3$ yields $3/4$. For example, if $\varepsilon = 1/4$, then the limit is ≈ 0.0712 via simulation; if $\varepsilon = 1/10$, then the limit is ≈ 0.1406. These compare well with multiplying 0.0948705678... and 0.1875465011... respectively by $3/4$.

Acknowledgments I am grateful to Peter Winkler, Aaron Smith, Stefan Steinerberger, Natesh Pillai and Ravi Montenegro for helpful discussions.

[1] D. Levin, Y. Peres and E. Wilmer, *Markov Chains and Mixing Times*, Amer. Math. Soc., 2009, pp. 3–5, 9, 47–48, 53–55, 65, 96; MR2466937 (2010c:60209).
[2] P. Diaconis, *Group Representations in Probability and Statistics*, Inst. Math. Stat., 1988, pp. 17–30; MR0964069 (90a:60001).
[3] R. Montenegro, The mixing time of simple random walk on a cycle, lecture notes (2004).
[4] R. Montenegro, Exponential decay and lower bounding mixing time, lecture notes (2004).
[5] A. Bloemendal and K. Yang, Mixing time of lazy simple random walk on the n-cycle, unpublished note (2010).
[6] D. Aldous and P. Diaconis, Strong uniform times and finite random walks, *Adv. Appl. Math.* 8 (1987) 69–97; MR0876954 (88d:60175).
[7] P. Diaconis, The cutoff phenomenon in finite Markov chains, *Proc. Nat. Acad. Sci. U.S.A.* 93 (1996) 1659–1664; MR1374011 (97b:60112).
[8] P. Diaconis, The mathematics of mixing things up, *J. Stat. Phys.* 144 (2011) 445–458; MR2826629 (2012j:60207).
[9] L. Saloff-Coste, Random walks on finite groups, *Probability on Discrete Structures*, ed. H. Kesten, Springer-Verlag, 2004, pp. 263–346; MR2023654 (2004k:60133).

4.29 Correlated Products

Fix $|\rho| < 1$ and ε_t to be $N(0,1)$ white noise. The stationary first-order autoregressive process

$$X_t = \rho X_{t-1} + \sqrt{1-\rho^2}\,\varepsilon_t$$

exhibits a surprising phase transition with respect to the correlation coefficient ρ. Define

$$A_n(\rho) = \mathrm{E}\left(X_1 X_2 \cdots X_n\right)$$

then

$$A_n(\rho) = \begin{cases} 0 & \text{if } n = 2k+1, \\ \dfrac{1}{k!}\dfrac{d^k}{dz^k} f(z,\rho)\Big|_{z=0} & \text{if } n = 2k \end{cases}$$

where $f(z,\rho)$ is the infinite continued fraction [1, 2]

$$f(z,\rho) = \frac{1|}{|1} - \frac{\rho z|}{|1} - \frac{2\rho^3 z|}{|1} - \frac{3\rho^5 z|}{|1} - \frac{4\rho^7 z|}{|1} - \cdots.$$

For example,

$$A_2 = \rho, \quad A_4 = \rho^2 + 2\rho^4, \quad A_6 = \rho^3 + 4\rho^5 + 4\rho^7 + 6\rho^9,$$

$$A_8 = \rho^4 + 6\rho^6 + 12\rho^8 + 20\rho^{10} + 24\rho^{12} + 18\rho^{14} + 24\rho^{16},$$

$$\begin{aligned} A_{10} = {} & \rho^5 + 8\rho^7 + 24\rho^9 + 50\rho^{11} + 88\rho^{13} + 108\rho^{15} + 156\rho^{17} + 150\rho^{19} + 144\rho^{21} \\ & + 96\rho^{23} + 120\rho^{25}. \end{aligned}$$

There exists a unique $0 < \rho_A < 1$ such that $A_{2k}(\rho) \to \infty$ if $\rho > \rho_A$ and $A_{2k}(\rho) \to 0$ if $\rho < \rho_A$ as $k \to \infty$. The critical threshold $\rho_A = 0.5630071693...$ is the smallest positive r for which $f(1, r) = \infty$. Also [1],

$$A_{2k}(\rho_A) \to \lim_{z \to 1}(1 - z)f(z, \rho_A) = 0.5090085224....$$

Finally, if [2]

$$\alpha(\rho) = \lim_{k \to \infty} A_{2k}(\rho)^{1/k}$$

then $\alpha(0^+) = 0$, $\alpha(1^-) = \infty$ and

$$\frac{\rho}{1 - \rho^2} \le \alpha(\rho) \le \frac{\rho + \rho^3}{1 - \rho^2}$$

for $0 < \rho < 1$. More accurate bounds on $\alpha(\rho)$ would be good to see someday.

Not as much is known about

$$\tilde{A}_n(\rho) = \mathrm{E}\,|X_1 X_2 \cdots X_n|\,.$$

It is possible to rewrite $\tilde{A}_n(\rho)$ as convolutions via a Hilbert–Schmidt kernel. Using the eigenanalysis method in [1], we deduce that the corresponding threshold $\tilde{\rho}_A < 0.5392$. In particular, it is strictly smaller than ρ_A.

It is important not to confuse $\ln\left(\mathrm{E}\,|X_1 X_2 \cdots X_n|\right)$ with

$$\mathrm{E}\left(\ln|X_1 X_2 \cdots X_n|\right) = n\mathrm{E}(\ln|X_t|) = \frac{n}{2}\left(-\ln(2) - \gamma\right)$$

where γ is Euler's constant [3]. The latter is independent of ρ; a Central Limit Theorem for $\ln|X_t|$ appears in [4].

Let ε_t now be an $m \times m$ symmetric matrix with independent $N(0, 1)$ white noise entries. The $m \times m$ symmetric matrix X_t satisfies the same recurrence as before – correlation $|\rho| < 1$ remains a scalar – our interest is in the (noncommutative) matrix product

$$Q_n = m^{-n/2} X_1 X_2 \cdots X_n$$

for large integer m. Define

$$B_n^m(\rho) = m^{-1}\mathrm{E}\left(\mathrm{tr}\,(Q_n)\right)$$

where $\mathrm{tr}\,(Q_n)$ is the trace of Q_n (sum of diagonal elements), then

$$B_n(\rho) = \lim_{m \to \infty} B_n^m(\rho)$$

satisfies

$$B_n(\rho) = \begin{cases} 0 & \text{if } n = 2k + 1, \\ \dfrac{\rho^k}{k!} \dfrac{d^k}{dz^k} g(z, \rho^2)\Big|_{z=0} & \text{if } n = 2k \end{cases}$$

where $g(z,q)$ is the (generalized Rogers–Ramanujan) continued fraction [2]

$$g(z,q) = \frac{1|}{|1} - \frac{z|}{|1} - \frac{q\,z|}{|1} - \frac{q^2\,z|}{|1} - \frac{q^3\,z|}{|1} - \cdots.$$

For example,

$$B_2 = \rho, \quad B_4 = \rho^2 + \rho^4, \quad B_6 = \rho^3 + 2\rho^5 + \rho^7 + \rho^9,$$

$$B_8 = \rho^4 + 3\rho^6 + 3\rho^8 + 3\rho^{10} + 2\rho^{12} + \rho^{14} + \rho^{16},$$

$$B_{10} = \rho^5 + 4\rho^7 + 6\rho^9 + 7\rho^{11} + 7\rho^{13} + 5\rho^{15} + 5\rho^{17} + 3\rho^{19} + 2\rho^{21} + \rho^{23} + \rho^{25}.$$

There exists a unique $0 < \rho_B < 1$ such that $B_{2k}(\rho) \to \infty$ if $\rho > \rho_B$ and $B_{2k}(\rho) \to 0$ if $\rho < \rho_B$ as $k \to \infty$. The critical threshold $\rho_B = 0.6629014851...$ is the smallest positive r for which $g\left(r, r^2\right) = \infty$. If [2]

$$\beta(\rho) = \lim_{k\to\infty} B_{2k}(\rho)^{1/k}$$

then $\beta(0^+) = 0$, $\beta(1^-) = 4$ and

$$\rho\left(1+\rho^2\right) \le \beta(\rho) \le \min\left\{\frac{1}{2}\rho\left(1+\sqrt{\frac{1+3\rho^2}{1-\rho^2}}\right), 2\rho\left(1+\rho^2\right)\right\}$$

for $0 < \rho < 1$. Again, more accurate bounds on $\beta(\rho)$ would be good to see.

We note that $g(z,q)$ can be expressed as a ratio of two q-hypergeometric functions:

$$\left[1 + \sum_{k=1}^{\infty} \frac{(-1)^k q^{k^2} z^k}{(1-q)\left(1-q^2\right)\cdots\left(1-q^k\right)}\right] \Big/ \left[1 + \sum_{\ell=1}^{\infty} \frac{(-1)^\ell q^{\ell^2-\ell} z^\ell}{(1-q)\left(1-q^2\right)\cdots\left(1-q^\ell\right)}\right]$$

which makes possible a high-precision calculation of the radius of convergence for $B_{2k}(\rho)$ generating series. There is no known analogous treatment for $A_{2k}(\rho)$. Logan, Mazo, Odlyzko & Shepp [1] studied the one-dimensional scenario as a toy model for correlated matrix products which arise in the analysis of learning curves for adaptive systems. It is ironic, as Mazza & Piau [2] wrote, that the infinite-dimensional scenario turns out to be easier to manage in this regard.

A formula for the expected product of components of a multivariate normally distributed vector (with arbitrary covariance matrix) appears in [5]. See additionally [6], which served as a starting point for [2], and [7] for possibly relevant study of free random variables.

[1] B. F. Logan, J. E. Mazo, A. M. Odlyzko and L. A. Shepp, On the average product of Gauss-Markov variables, *Bell System Tech. J.* 62 (1983) 2993–3006; MR0728852 (85m:60072).

[2] C. Mazza and D. Piau, Products of correlated symmetric matrices and q-Catalan numbers, *Probab. Theory Related Fields* 124 (2002) 574–594; MR1942324 (2003m:05014).

[3] S. R. Finch, Euler-Mascheroni constant, *first volume*, pp. 28–40.
[4] S. R. Finch, Another look at AR(1), arXiv:0710.5419.
[5] R. Kan, From moments of sum to moments of product, *J. Multivariate Anal.* 99 (2008) 542–554; MR2396978 (2008m:60028).
[6] A. Khorunzhy, Products of random matrices and q-Catalan numbers, arXiv:math/0104074.
[7] K.-P. Ho, Central limit for the product of free random variables, arXiv:1101.5220.

5

Geometry and Topology

5.1 Knots, Links and Tangles

We start with some terminology from differential topology [1]. Let C be a circle and $n \geq 2$ be an integer. An **immersion** $f: C \to \mathbb{R}^n$ is a smooth function whose derivative never vanishes. An **embedding** $g: C \to \mathbb{R}^n$ is an immersion that is one-to-one. It follows that $g(C)$ is a manifold but $f(C)$ need not be (f is only locally one-to-one, so consider the map that twists C into a figure of eight).

A **knot** is a smoothly embedded circle in $\mathbb{R}^3$; hence a knot is a closed spatial curve with no self-intersections. Two knots J and K are **equivalent** if there is a homeomorphism $\mathbb{R}^3 \to \mathbb{R}^3$ taking J onto K. This implies that the complements $\mathbb{R}^3 - J$ and $\mathbb{R}^3 - K$ are homeomorphic as well.

A **link** is a compact smooth 1-dimensional submanifold of $\mathbb{R}^3$. The connected components of a link are disjoint knots, often with intricate intertwinings. Two links L and M are **equivalent** if, likewise, there is a homeomorphism $\mathbb{R}^3 \to \mathbb{R}^3$ taking L onto M.

We can project a knot or a link into the plane in such a way that its only self-intersections are transversal double points. Ambiguity is removed by specifying at each double point which arc passes over and which arc passes under. Over all possible such projections of K or L, determine one with the minimum number of double points; this defines the **crossing number** of K or L.

There is precisely 1 knot with 0 crossings (the circle), 1 knot with 3 crossings (the trefoil), and 1 knot with 4 crossings. Note that, although the left-hand trefoil T_L is not ambiently isotopic (i.e., deformable) to the right-hand trefoil T_R, a simple reflection about a plane gives T_R as a homeomorphic image of T_L. Under our definition of equivalence, chiral pairs as such are counted only once.

There are precisely 2 knots with 5 crossings, and 5 knots with 6 crossings. In particular, there is no homeomorphism $\mathbb{R}^3 \to \mathbb{R}^3$ taking the granny knot $T_L \# T_L$ onto the square knot $T_L \# T_R$, where $\#$ denotes the connected sum of manifolds [2, 3]. (See Figure 5.1.) Also, there are precisely 8 knots with 7 crossings, and 25 knots with 8 crossings.

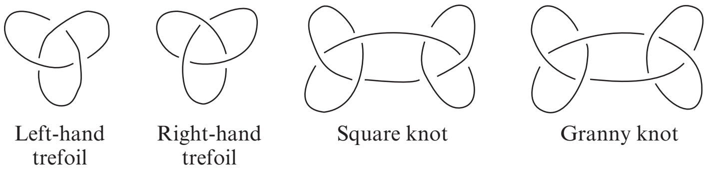

Figure 5.1 Four famous knots (T_L and T_R are prime and equivalent; $T_L\#T_R$ and $T_L\#T_L$ are composite and distinct).

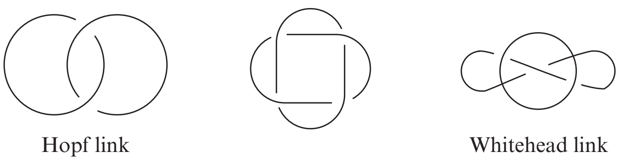

Figure 5.2 All two-component prime links with crossing number ≤ 5.

A link L is **splittable** if we can embed a plane in $\mathbb{R}^3$, disjoint from L, that separates one or more components of L from other components of L. There are precisely 1, 0, 1, 1, 3, 4, 15 nonsplittable links with 0, 1, 2, 3, 4, 5, 6 crossings, respectively.

A knot K or nonsplittable link L is **prime** if it is not a circle and if, for any plane P that intersects K or L transversely in exactly two points, P slices off merely an unknotted arc away from the rest. (See Figure 5.2.) Otherwise it is **composite**. For example, $T_L\#T_L$ and $T_L\#T_R$ are composite knots, each being nontrivial connected sums of knots. Every knot decomposes as a unique connected sum of prime knots [4].

People have known for a long time that there exist non-equivalent links with homeomorphic complements [5, 6]. This cannot happen for knots, as proved by Gordon & Luecke [7, 8].

Let B denote the compact unit ball in $\mathbb{R}^3$ and ∂B denote its boundary. A **tangle** U is a smooth 1-dimensional submanifold of B meeting ∂B transversely at the four points

$$NE=\left(\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}},0\right),\ NW=\left(\tfrac{-1}{\sqrt{2}},\tfrac{1}{\sqrt{2}},0\right),\ SW=\left(\tfrac{-1}{\sqrt{2}},\tfrac{-1}{\sqrt{2}},0\right),\ SE=\left(\tfrac{1}{\sqrt{2}},\tfrac{-1}{\sqrt{2}},0\right)$$

and meeting ∂B nowhere else. Thus U is a union of two smoothly embedded line segments in B with distinct endpoints on ∂B, together with an arbitrary number of smoothly embedded circles in the interior of B, all disjoint but often intertwined. Two tangles U and V are **(strongly) equivalent** if there is a homeomorphism $B \to B$ that takes U onto V, is orientation-preserving on B, and leaves ∂B fixed pointwise. The crossing number of a tangle is defined via projections as before. Tangles form the building blocks of knots and links [9–11]; the first precise

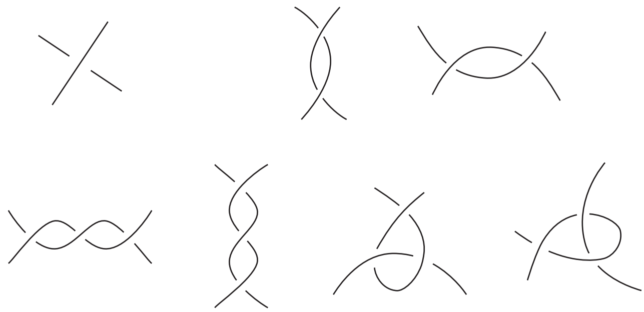

Figure 5.3 All prime alternating tangles with crossing number ≤ 3.

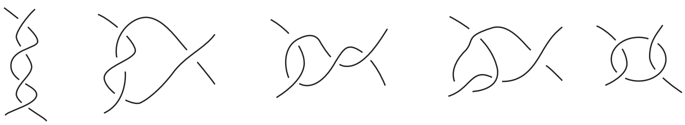

Figure 5.4 Five of the 4-crossing prime alternating tangles; the other five are obtained by rotating through 90° (and switching crossings to maintain the convention that the NW strand is an underpass).

asymptotic enumeration results discovered in this subject concerned tangles (as we shall soon see).

A tangle is **trivial** if it is only the union of the two line segments NW-NE and SW-SE, or the union of the two line segments SW-NW and SE-NE. A tangle U is **prime** if it is not trivial; if, for any sphere S in B that is disjoint from U, no portion of U is enclosed by S; and if, for any sphere S in B that intersects U transversely in exactly two points, S encloses merely an unknotted arc of U. (See Figures 5.3 and 5.4.)

Finally, a knot, link or tangle is **alternating** if, for some projection, as we proceed along any connected component in the projection plane from beginning to end, the sequence of underpasses and overpasses is strictly alternating. The first non-alternating knots appear with crossing number ≥ 8. General references on knot theory include [12–17].

5.1.1 *Prime Alternating Tangles*

Let a_n denote the number of prime alternating tangles with n crossings (up to strong equivalence) and let $A(x) = \sum_{n=1}^{\infty} a_n x^n$ be the corresponding generating function. Then [18]

$$\begin{aligned} A(x) &= x + 2x^2 + 4x^3 + 10x^4 + 29x^5 + 98x^6 + 372x^7 \\ &\quad + 1538x^8 + 6755x^9 + 30996x^{10} + \cdots \end{aligned}$$

satisfies the equation

$$A(x)(1+x)-A(x)^2-(A(x)+1)r(A(x))-x-2\frac{x^2}{1-x}=0$$

where the algebraic function $r(x)$ is defined by

$$r(x)=\frac{(1-4x)^{\frac{3}{2}}+(2x^2-10x-1)}{2(x+2)^3}-\frac{2}{1+x}-x+2.$$

Further, $A(x)$ satisfies the irreducible quintic equation

$$\begin{aligned}0 = {} & (x^4-2x^3+x^2)A(x)^5+(8x^4-14x^3+8x^2-2x)A(x)^4\\ & +(25x^4-16x^3-14x^2+8x+1)A(x)^3+(38x^4+15x^3-30x^2-x+2)A(x)^2\\ & +(28x^4+36x^3-5x^2-12x+1)A(x)+(8x^4+17x^3+8x^2-x).\end{aligned}$$

Sundberg & Thistlethwaite [19] proved the above remarkable formulas, as well as the following asymptotics:

$$a_n \sim \frac{3\alpha}{4\sqrt{\pi}}n^{-\frac{5}{2}}\lambda^{n-\frac{3}{2}} \sim \frac{3}{4}\sqrt{\frac{\beta}{\pi}}n^{-\frac{5}{2}}\lambda^n,$$

where

$$\alpha=\frac{5^{\frac{7}{2}}}{3^5\sqrt{2}}\sqrt{\frac{(21001+371\sqrt{21001})^3}{(17+3\sqrt{21001})^5}}=3.8333138762...$$

$$\beta=\alpha^2\lambda^{-3}=0.0632356411...$$

and

$$\lambda=\frac{101+\sqrt{21001}}{40}=6.1479304437...$$

A completely different approach to the solution of this problem appears in [20].

Let $\hat{a}_n$ denote the number of n-crossing prime alternating tangles with exactly two components. That is, no circles are allowed. A two-component tangle is also known as a **knot with four external legs**. The sequence [18, 21, 22]

$$\{\hat{a}_n\}_{n=1}^{\infty}=\{1,\ 2,\ 4,\ 8,\ 24,\ 72,\ 264,\ 1074,\ 4490,\ 20296,\ 92768,\ldots\},$$

is believed to possess a leading term of the form $\hat{\lambda}^n$ with $\hat{\lambda}<\lambda$, but more intensive analysis is needed to compute $\hat{\lambda}$.

5.1.2 *Prime Alternating Links*

Let b_n denote the number of prime alternating links with n crossings (up to equivalence), then the sequence [23, 24]

$$\{b_n\}_{n=1}^{\infty}=\{0,\ 1,\ 1,\ 2,\ 3,\ 8,\ 14,\ 39,\ 96,\ 297,\ 915,\ 3308,\ 12417,\ldots\}$$

satisfies the following asymptotics [25]:

$$b_n \sim \frac{3}{16\gamma}\sqrt{\frac{\beta}{\pi}}\,n^{-\frac{7}{2}}\lambda^n$$

where

$$\gamma = \frac{1}{2}\left(\frac{371}{\sqrt{21001}} - 1\right) = 0.7800411357...$$

and λ, β are as before. This is a somewhat more precise result than that proved in [19].

Let c_n denote the number of prime links with n crossings (including both alternating and non-alternating links), then we have [23, 26, 27]

$$\{c_n\}_{n=1}^{\infty} = \{0,\ 1,\ 1,\ 2,\ 3,\ 9,\ 16,\ 50,\ 132,\ 452,\ 1559,\ldots\}.$$

The value c_{12} is not known. Stoimenow [28], building on Ernst & Sumners [29] and Welsh [30], proved that

$$4 \le \liminf_{n\to\infty} c_n^{1/n} \le \limsup_{n\to\infty} c_n^{1/n} \le \frac{\sqrt{13681}+91}{20} = 10.3982903484...$$

but further improvements in the upper bound are likely. The two-component analogs [23]

$$\{\hat{b}_n\}_{n=1}^{\infty} = \{0,\ 1,\ 0,\ 1,\ 1,\ 3,\ 6,\ 14,\ 42,\ 121,\ 384,\ 1408,\ 5100,\ 21854,\ldots\},$$

$$\{\hat{c}_n\}_{n=1}^{\infty} = \{0,\ 1,\ 0,\ 1,\ 1,\ 3,\ 8,\ 16,\ 61,\ 185,\ 638\ldots\}$$

also await study.

5.1.3 *Prime Alternating Knots*

Let d_n denote the number of prime alternating knots with n crossings (up to equivalence), then the sequence [31]

$$\{d_n\}_{n=1}^{\infty} = \{0,\ 0,\ 1,\ 1,\ 2,\ 3,\ 7,\ 18,\ 41,\ 123,\ 367,\ 1288,\ 4878,\ 19536,\ldots\}$$

is more difficult and only *conjectured* to satisfy the following asymptotics [32]:

$$d_n \sim \eta\cdot n^{\xi}\cdot\kappa^n$$

where

$$\xi = -\frac{\sqrt{13}+1}{6} - 3 = -3.7675918792....$$

Thistlethwaite [33] proved that

$$\limsup_{n\to\infty} d_n^{1/n} < \lambda$$

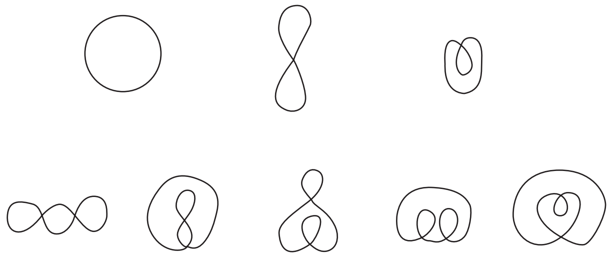

Figure 5.5 All closed planar curves with crossing number ≤ 2.

and further claimed that $\lim_{n\to\infty} d_n^{1/n}$ exists. If the conjectured asymptotic form for d_n is true, it would follow that $\kappa < \lambda$. Again, more intensive analysis is needed to compute κ. Might it be true that $\kappa = \hat{\lambda}$ [22]?

Let e_n denote the number of prime knots with n crossings (including both alternating and non-alternating knots), then we have [31]

$$\{e_n\}_{n=1}^{\infty} = \{0,\ 0,\ 1,\ 1,\ 2,\ 3,\ 7,\ 21,\ 49,\ 165,\ 552,\ 2176,\ 9988,\ 46972,\ldots\}.$$

The value e_{17} is not known. Welsh [30] proved that

$$2.68 \leq \liminf_{n\to\infty} e_n^{1/n}$$

and clearly Stoimenow's upper bound 10.40 applies to the limit superior. Sharper bounds for both $\{c_n\}$ and $\{e_n\}$ would be good to see.

5.1.4 *Planar Curves*

Here are enumeration problems that seem to be even more complicated than those in knot theory [34–38]. A **closed planar curve** is a smoothly immersed circle in $\mathbb{R}^2$ whose only self-intersections are transversal double points. Define an equivalence relation between closed planar curves in the same manner as between knots, with the additional condition that the homeomorphism $\mathbb{R}^2 \to \mathbb{R}^2$ is orientation-preserving. (See Figure 5.5.)

An **open planar curve** is a smoothly immersed line in $\mathbb{R}^2$, given by $h: \mathbb{R} \to \mathbb{R}^2$, whose only self-intersections are transversal double points and which satisfies $h(x) = (x, 0)$ for all sufficiently large $|x|$. Such a curve is also known as a **knot with two external legs**. Define an equivalence relation between open planar curves in the same manner as between closed planar curves. Note that, unlike closed curves, open curves are oriented from the initial point $(-\infty, 0)$ to the final point $(\infty, 0)$. (See Figure 5.6.)

Let p_n and q_n denote the number of n-crossing closed curves and open curves, respectively. The sequences [39, 40]

$$\{p_n\}_{n=0}^{\infty} = \{1,\ 2,\ 5,\ 20,\ 82,\ 435,\ 2645,\ 18489,\ 141326,\ 1153052,\ 9819315,\ldots\},$$

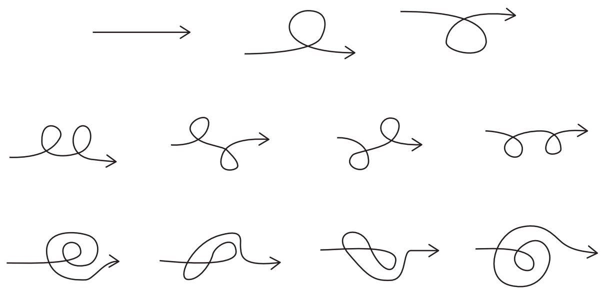

Figure 5.6 All open planar curves with crossing number ≤ 2.

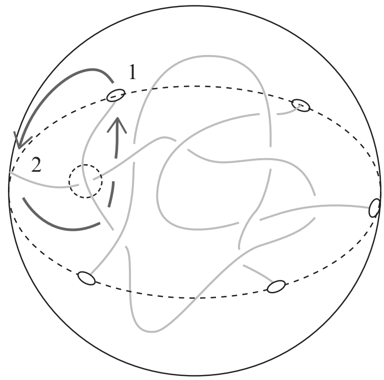

Figure 5.7 Positions of legs 1 and 2 can be reversed on the sphere (following the arrows), thus removing the crossing indicated by the dotted circle. Image courtesy of Vadim Meshkov.

$$\{q_n\}_{n=0}^{\infty} = \{1,\ 2,\ 8,\ 42,\ 260,\ 1796,\ 13396,\ 105706,\ 870772,\ 7420836,\ 65004584, \ldots\}$$

are *conjectured* to satisfy the following asymptotics [32]:

$$p_n \sim \frac{1}{4} q_n \sim \omega \cdot n^{\theta} \cdot \mu^n$$

where $\theta = \xi + 1 = -2.7675918792\ldots$. Numerically, we have $\mu = 11.4\ldots$ [22]. There is a great amount of work to be done in this area.

Addendum At the risk of potential confusion, let us generalize the word **tangle** to include smooth 1-dimensional submanifolds U of B meeting ∂B transversely at *any* four distinct points and meeting ∂B nowhere else. Two such tangles U and V are **weakly equivalent** if there is a homeomorphism $B \to B$ that takes U onto V, but need not be orientation-preserving on B nor need it leave endpoints fixed. Kanenobu, Saito & Satoh [41] gave the number of non-weakly equivalent prime tangles with 4, 5, 6, 7 crossings to be 0, 1, 4, 18 respectively. The four legs (small circles on the spherical surface depicted in Figure 5.7) of *classical* tangles are fixed on the equator, whereas the legs of weakly equivalent tangles can slide anywhere on the unit sphere, hence there are many more possible untangling strategies.

A different generalization of tangle was provided by Bogdanov, Meshkov, Omelchenko & Petrov [42], in which **2-tangles** correspond to classical tangles and k**-tangles**, $k > 2$, similarly possess $2k$ legs equally spaced on the equator. The number of non-equivalent prime alternating 2-tangles with 2, 3, 4, 5 crossings is given in [42] to be 1, 2, 5, 13 respectively, which at first glance appears to contradict the numbers 2, 4, 10, 29 from [19], until it is understood that 1, 2, 5, 13 do not distinguish projections that differ by only a sequence of flypes. The asymptotics of counts of prime alternating k-tangles, as the number n of crossings $\to \infty$, would be a challenging exercise.

[1] V. Guillemin and A. Pollack, *Differential Topology*, Prentice-Hall, 1974, pp. 13–18, 27–32; MR0348781 (50 #1276).
[2] H. Seifert, Verschlingungsinvarianten, *Sitzungsber. Preuss. Akad. Wiss.* 26 (1933) 811–823.
[3] R. H. Fox, On the complementary domains of a certain pair of inequivalent knots, *Proc. Konink. Nederl. Akad. Wetensch. Ser. A* 55 (1952) 37–40; *Indag. Math.* 14 (1952) 37–40; MR0048024 (13,966c).
[4] H. Schubert, Die eindeutige Zerlegbarkeit eines Knotens in Primknoten, *Sitzungsberichte der Heidelberger Akademie der Wissenschaften Mathematisch-Naturwissenschaftliche Klasse* (1949) n. 3, 57–104; MR0031733 (11,196f).
[5] J. H. C. Whitehead, On doubled knots, *J. London Math. Soc.* 12 (1937) 63–71.
[6] D. Rolfsen, *Knots and Links*, Publish or Perish, 1976, p. 49; MR0515288 (58 #24236).
[7] C. M. Gordon and J. Luecke, Knots are determined by their complements, *J. Amer. Math. Soc.* 2 (1989) 371–415; *Bull. Amer. Math. Soc.* 20 (1989) 83–87; MR0965210 (90a:57006a) and MR0972070 (90a:57006b).
[8] B. A. Cipra, To have and have knot: When are knots alike? *Science* 241 (1988) 1291–1292; MR0967958 (89k:57009).
[9] J. H. Conway, An enumeration of knots and links, and some of their algebraic properties, *Computational Problems in Abstract Algebra*, Proc. 1967 Oxford conf., ed. J. Leech, Pergamon Press, 1970, pp. 329–358; MR0258014 (41 #2661).
[10] W. B. R. Lickorish, Prime knots and tangles, *Trans. Amer. Math. Soc.* 267 (1981) 321–332; MR0621991 (83d:57004).
[11] M. Eudave Muñoz, Primeness and sums of tangles, *Trans. Amer. Math. Soc.* 306 (1988) 773–790; MR0933317 (89g:57005).
[12] R. H. Crowell and R. H. Fox, *Introduction to Knot Theory*, Springer-Verlag, 1977; MR0146828 (26 #4348).
[13] M. Thistlethwaite, Knot tabulations and related topics, *Aspects of Topology*, ed. I. M. James and E. H. Kronheimer, Cambridge Univ. Press, 1985, pp. 1–76; MR0787823 (86j:57004).
[14] G. Burde and H. Zieschang, *Knots*, de Gruyter, 1985; MR0808776 (87b:57004).
[15] C. C. Adams, *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots*, W. H. Freeman, 1994; MR1266837 (94m:57007).
[16] W. B. R. Lickorish, *An Introduction to Knot Theory*, Springer-Verlag, 1997; MR1472978 (98f:57015).
[17] J. Hoste, M. Thistlethwaite, and J. Weeks, The first 1,701,936 knots, *Math. Intelligencer* v. 20 (1998) n. 4, 33–48; MR1646740 (99i:57015).

[18] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A047051 and A067646.
[19] C. Sundberg and M. Thistlethwaite, The rate of growth of the number of prime alternating links and tangles, P*acific J. Math.* 182 (1998) 329–358; MR1609591 (98m:57013).
[20] P. Zinn-Justin and J.-B. Zuber, Matrix integrals and the counting of tangles and links, *Discrete Math.* 246 (2002) 343–360; math-ph/9904019; MR1887495 (2003i:57019).
[21] P. Zinn-Justin and J.-B. Zuber, On the counting of colored tangles, *J. Knot Theory Ramifications* 9 (2000) 1127–1141; math-ph/0002020; MR1807611 (2002b:57008).
[22] J. L. Jacobsen and P. Zinn-Justin, A transfer matrix approach to the enumeration of knots, *J. Knot Theory Ramifications* 11 (2002) 739–758; math-ph/0102015; MR1918811 (2003f:57014) .
[23] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A049344, A048952, A048953, A059739, A059741, and A086771.
[24] M. Thistlethwaite, Prime unoriented alternating links to 19 crossings, unpublished table (2001).
[25] S. Kunz-Jacques and G. Schaeffer, The asymptotic number of prime alternating links, presentation at *Formal Power Series and Algebraic Combinatorics (FPSAC)* conf., Tempe, Arizona, 2001.
[26] D. Bar-Natan, The Hoste-Thistlethwaite link table, The Knot Atlas, Univ. of Toronto.
[27] R. G. Scharein, Number of prime links, Knot Plot, Univ. of British Columbia.
[28] A. Stoimenow, On the number of links and link polynomials, *Quart. J. Math.* 55 (2004) 87–98; MR2043016 (2005a:57006).
[29] C. Ernst and D. W. Sumners, The growth of the number of prime knots, *Math. Proc. Cambridge Philos. Soc.* 102 (1987) 303–315; MR0898150 (88m:57006).
[30] D. J. A. Welsh, On the number of knots and links, *Sets, Graphs and Numbers*, Proc. 1991 Budapest conf., Colloq. Math. Soc. János Bolyai 60, North-Holland, 1992, pp. 713–718; MR1218230 (94f:57010).
[31] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002863 and A002864.
[32] G. Schaeffer and P. Zinn-Justin, On the asymptotic number of plane curves and alternating knots, *Experim. Math.* 13 (2004) 483–493; math-ph/0304034; MR2118273 (2006d:05018).
[33] M. Thistlethwaite, On the structure and scarcity of alternating links and tangles, *J. Knot Theory Ramifications* 7 (1998) 981–1004; MR1654669 (99k:57031).
[34] V. I. Arnold, *Topological Invariants of Plane Curves and Caustics*, Amer. Math. Soc., 1994; MR1286249 (95h:57003).
[35] V. I. Arnold, Plane curves, their invariants, perestroikas and classifications, *Singularities and Bifurcations*, ed. V. I. Arnold, Adv. Soviet Math. 21, Amer. Math. Soc., 1994, pp. 33–91; MR1310595 (95m:57009).
[36] S. M. Guseĭn-Zade, On the enumeration of curves from infinity to infinity, *Singularities and Bifurcations*, ed. V. I. Arnold, Adv. Soviet Math. 21, Amer. Math. Soc., 1994, pp. 189–198; MR1310602 (96d:57029).
[37] S. M. Guseĭn-Zade and F. S. Duzhin, On the number of topological types of plane curves (in Russian), *Uspekhi Mat. Nauk* v. 53 (1998) n. 3, 197–198; Engl. transl. in *Russian Math. Surveys* 53 (1998) 626–627; MR1657608 (99i:57001).
[38] S. K. Lando, On enumeration of unicursal curves, *Differential and Symplectic Topology of Knots and Curves*, ed. S. Tabachnikov, Amer. Math. Soc., 1999, pp. 77–81; MR1738391 (2000m:57012).
[39] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A008983 and A054993.

[40] F. S. Duzhin, Number of topological types of plane curves with simple double points, unpublished table (1998).
[41] T. Kanenobu, H. Saito and S. Satoh, Tangles with up to seven crossings, *Interdiscip. Inform. Sci.* 9 (2003) 127–140; MR2023113 (2004j:57006).
[42] A. Bogdanov, V. Meshkov, A. Omelchenko and M. Petrov, Enumerating the k-tangle projections, *J. Knot Theory Ramifications* 21 (2012) 1250069; MR2911085.

5.2 Convex Lattice Polygons

Let $n \geq 3$ be an integer. A convex lattice n-gon is a polygon whose n vertices are points on the integer lattice $\mathbb{Z}^2$ and whose interior angles are strictly less than π. Let a_n denote the least possible area enclosed by a convex lattice n-gon, then [1–3]

$$\{a_n\}_{n=3}^{\infty} = \left\{\tfrac{1}{2}, 1, \tfrac{5}{2}, 3, \tfrac{13}{2}, 7, \tfrac{21}{2}, 14, x, 24, \tfrac{65}{2}, 40, y, 59, z, 87, w, 121, ...\right\},$$

where the unknown values x, y, z, and w are known to satisfy

$$x \in \left\{\tfrac{39}{2}, \tfrac{41}{2}, \tfrac{43}{2}\right\}, \quad y \in \left\{\tfrac{99}{2}, \tfrac{101}{2}, \tfrac{103}{2}\right\},$$

$$z \in \left\{\tfrac{147}{2}, \tfrac{149}{2}, \tfrac{151}{2}\right\}, \quad w \in \left\{\tfrac{209}{2}, \tfrac{211}{2}, \tfrac{213}{2}\right\}.$$

On the one hand, Rabinowitz [4] and Colburn & Simpson [5] demonstrated that $a_n \leq Cn^3$ for some constant $C > 0$; Zunic [6] later proved that $C \leq 1/54$. On the other hand, Andrews [7] and Arnold [8] were the first to show that $a_n \geq cn^3$ for some $c > 0$; other proofs appear in [9–12]. Bárány & Tokushige [13] succeeded in proving that $\lim_{n\to\infty} a_n/n^3$ actually exists and computed that

$$\lim_{n\to\infty} \frac{a_n}{n^3} = 0.0185067... < \frac{1}{54}$$

via a heuristic solution of $\approx 10^{10}$ constrained minimization problems. Further, the shape of the minimizing n-gon is approximated by that of the ellipse

$$\frac{x^2}{A^2} + \frac{y^2}{B^2} = 1$$

where $A = (0.003573...)n^2$ and $B = (1.656...)n$.

Much less can be said about the higher dimensional analog. A d-dimensional convex lattice polytope with n vertices has volume v_n satisfying [7, 9, 14, 15]

$$v_n \geq c_d n^{\frac{d+1}{d-1}}$$

but little else is known.

5.2.1 *Integer Convex Hulls*

Before discussing integer convex hulls, let us mention ordinary convex hulls. Given n points chosen at random in the unit disk D, the convex hull C_n is the intersection of all convex sets containing all n points. The boundary of C_n is a polygon; let N_n denote the number of vertices of the polygon. It can be proved that [16–18]

$$\lim_{n\to\infty}\frac{\mathrm{E}(N_n)}{n^{1/3}}=2\pi\xi,\quad \lim_{n\to\infty}\frac{\mathrm{Var}(N_n)}{n^{1/3}}=2\pi\eta,$$

where

$$\xi=\left(\tfrac{3\pi}{2}\right)^{-\frac{1}{3}}\Gamma\left(\tfrac{5}{3}\right)=0.5384576135...,$$

$$\eta=\frac{16\pi^2\Gamma\left(\frac{2}{3}\right)^{-3}-57}{27}\xi=0.1316029298...=2(0.3350302716...)-\xi.$$

We point out that this is more complicated than the corresponding result when the unit disk is replaced by the unit square [16, 17, 19]:

$$\lim_{n\to\infty}\frac{\mathrm{E}(\tilde{N}_n)}{\ln(n)}=\frac{8}{3},\quad \lim_{n\to\infty}\frac{\mathrm{Var}(\tilde{N}_n)}{\ln(n)}=\frac{40}{27}.$$

In the integer case, we consider not n random points in D, but rather *all* lattice points in rD, the disk of radius r, where r is large. The convex hull C_r of all these lattice points is clearly a convex lattice polygon, together with its interior. Motivation for studying this polygon comes from integer programming: When maximizing a linear function φ on the lattice points in rD (or any given convex set in $\mathbb{R}^2$), one looks for the maximum point of φ on C_r. The size of the programming problem is hence proportional to N_r, the number of vertices of C_r, and thus we wish to have bounds on N_r.

Balog & Bárány [20, 21] proved that, for sufficiently large r,

$$0.33r^{2/3}\leq N_r\leq 5.54r^{2/3}$$

but confessed that it is not clear whether $\lim_{r\to\infty}N_r r^{-2/3}$ exists. It is possible, however, to obtain asymptotics for the average value of N_r, defined in a special way:

$$\mathrm{E}_\theta(N_r)=\frac{1}{r^\theta}\int\limits_r^{r+r^\theta}N_\rho\,d\rho$$

where the parameter θ satisfies $0<\theta<1$. (Actually, the only feature required of r^θ is that it increases with r, but less rapidly than r itself.) Balog & Deshouillers [22] proved that

$$\lim_{r\to\infty}\frac{\mathrm{E}_\theta(N_r)}{r^{2/3}}=\frac{6\cdot 2^{2/3}}{\pi}\chi=3.4536898915...$$

independently of θ, where χ is defined later. The growth rate $2/3$ is what we would expect on the basis of the probabilistic model (ordinary convex hull case), but the preceding constant 3.453... is slightly different from $2\pi\xi=3.383...$. In this sense, lattice points do not behave in the same way as random points.

Another occurrence of the constant χ is as follows. For real x, let $||x||$ denote the distance from x to the nearest integer. Then, for $0 \le a < b \le 1$, we have [22]

$$\lim_{\lambda\to 0^+} \frac{1}{(b-a)\lambda^{1/3}} \int\limits_a^b \min_{t\neq 0}\left(||\alpha t|| + \lambda t^2\right)\, d\alpha = \frac{6}{\pi^2}\chi.$$

If $\lambda = 0$, the integral clearly is zero since, for any α, the point $t = 1/\alpha$ gives the minimum. If $\lambda > 0$, this strategy no longer works because the penalty term $\lambda t^2 = \lambda/\alpha^2$ would be large.

Let Δ denote the triangular region bounded by the lines $y = x$, $y = 1 - x$ and $x = 1$. Partition Δ into four domains:

$$\Delta_1 = \{(x,y)\in\Delta : 1 \le xy(x+y)\},$$

$$\Delta_2 = \{(x,y)\in\Delta : xy(x+y) \le 1 \le x(x+y)(x+2y)\},$$
$$\Delta_3 = \{(x,y)\in\Delta : x(x+y)(x+2y) \le 1 \le x(x+y)(2x+y)\},$$
$$\Delta_4 = \{(x,y)\in\Delta : x(x+y)(2x+y) \le 1\}.$$

Define $F: \Delta \to \mathbb{R}$ by

$$F(x,y) = \begin{cases} 4 - x^3 - y^3 & \text{in } \Delta_1, \\ \dfrac{1}{xy(x+y)} + 2 - (x+y)(x-y)^2 & \text{in } \Delta_2, \\ \dfrac{1}{y(x+y)(x+2y)} + 6 - (x+y)(3x^2 + 2xy + y^2) & \text{in } \Delta_3, \\ \dfrac{1}{x(x+y)(2x+y)} + \dfrac{1}{y(x+y)(x+2y)} + 4 - (x+y)(x^2+xy+y^2) & \text{in } \Delta_4, \end{cases}$$

then χ is given by

$$\chi = \int\limits_{1/2}^{1} \int\limits_{1-x}^{x} F(x,y)\, dy\, dx.$$

Again, much less can be said about the higher dimensional analog. Let B_d denote the d-dimensional unit ball. The number of vertices, N_r, of the integer convex hull of rB_d satisfies [23]

$$c_d r^{\frac{d(d-1)}{d+1}} \le N_r \le C_d r^{\frac{d(d-1)}{d+1}}$$

but an asymptotic average value for N_r is not known for any $d \ge 3$.

5.2.2 *Cubes and Thresholds*

The d-dimensional unit cube has 2^d vertices. Randomly select $n = n(d)$ vertices with replacement and form the ordinary convex hull of these points. If V_d denotes its expected volume, then for any $\varepsilon > 0$, [24, 25]

$$\lim_{d\to\infty} V_d = \begin{cases} 0 & \text{if } n(d) \le \left(2/\sqrt{e} - \varepsilon\right)^d, \\ 1 & \text{if } n(d) \ge \left(2/\sqrt{e} + \varepsilon\right)^d. \end{cases}$$

This is an interesting occurrence of the constant $2/\sqrt{e} = 1.2130613194...$, which is surprisingly small (relative to 2)! If instead the n points are selected uniformly in the interior of the d-cube, then the same threshold phenomenon occurs, with constant $2/\sqrt{e}$ replaced by

$$\exp\left(\int_0^\infty \left(\frac{1}{x} - \frac{1}{e^x - 1}\right)^2 dx\right) = 2.1396909474....$$

In fact, a closed-form expression is possible since

$$\int_0^\infty \left(\frac{1}{x} - \frac{1}{e^x - 1}\right)^2 dx = \ln(2\pi) - \gamma - \frac{1}{2} = 0.7606614015...$$

and the details underlying this formula appear in [26]. See [25] for relevant discussion of the d-dimensional unit ball.

[1] R. J. Simpson, Convex lattice polygons of minimum area, *Bull. Austral. Math. Soc.* 42 (1990) 353–367; MR1083272 (91k:52023).

[2] C. Landauer, Computational search for minimum area n-gon, unpublished note (2008).

[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A063984, A070911, and A089187.

[4] S. Rabinowitz, On the number of lattice points inside a convex lattice n-gon, *Proc. 20th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1989, ed. F. Hoffman, R. C. Mullin, R. G. Stanton and K. B. Reid, Congr. Numer. 73, Utilitas Math., 1990, pp. 99–124; MR1041842 (91a:52019).

[5] C. J. Colbourn and R. J. Simpson, A note on bounds on the minimum area of convex lattice polygons, *Bull. Austral. Math. Soc.* 45 (1992) 237–240; MR1155481 (93d:52010).

[6] J. Zunic, Notes on optimal convex lattice polygons, *Bull. London Math. Soc*. 30 (1998) 377–385; MR1620817 (99k:11102).

[7] G. E. Andrews, A lower bound for the volume of strictly convex bodies with many boundary lattice points, *Trans. Amer. Math. Soc*. 106 (1963) 270–279; MR0143105 (26 #670).

[8] V. I. Arnold, Statistics of integral convex polygons (in Russian), *Funktsional. Anal. i Prilozhen*., v. 14 (1980) n. 2, 1–3; Engl. transl. in *Funct. Anal. Appl*., v. 14 (1980) n. 2, 79–81; MR0575199 (81g:52011).

[9] W. M. Schmidt, Integer points on curves and surfaces, *Monatsh. Math*. 99 (1985) 45–72; MR0778171 (86d:11081).

[10] I. Bárány and J. Pach, On the number of convex lattice polygons, *Combin. Probab. Comput.* 1 (1992) 295–302; MR1211319 (93m:52017).

[11] S. Rabinowitz, $O(n^3)$ bounds for the area of a convex lattice n-gon, *Geombinatorics,* v. 2 (1993) n. 4, 85–88; MR1214699 (94b:52028).

[12] T.-X. Cai, On the minimum area of convex lattice polygons, *Taiwanese J. Math*. 1 (1997) 351–354; MR1486557 (98m:52026).

[13] I. Bárány and N. Tokushige, The minimum area of convex lattice n-gons, *Combinatorica* 24 (2004) 171–185; MR2071331 (2005e:52024).

[14] S. V. Konyagin and K. A. Sevastyanov, Estimate of the number of vertices of a convex integral polyhedron in terms of its volume (in Russian), *Funktsional. Anal. i Prilozhen.* v. 18 (1984) n. 1, 13–15; Engl. transl. in *Funct. Anal. Appl.*, v. 18 (1984) n. 1, 11–13; MR0739085 (86g:52020).
[15] I. Bárány and A. M. Vershik, On the number of convex lattice polytopes, *Geom. Funct. Anal.* 2 (1992) 381–393; MR1191566 (93k:52013).
[16] A. Rényi and R. Sulanke, Über die konvexe Hülle von n zufällig gewählten Punkten. I, *Z. Wahrsch. Verw. Gebiete* 2 (1963) 75–84; II, 3 (1964) 138–147; also in *Selected Papers of Alfréd Rényi*, v. 3, Akadémiai Kiadó, 1976, pp. 143–152 and 242–251; MR0156262 (27 #6190) and MR0169139 (29 #6392).
[17] P. Groeneboom, Limit theorems for convex hulls, *Probab. Theory Relat. Fields* 79 (1988) 327–368; MR0959514 (89j:60024).
[18] S. Finch and I. Hueter, Random convex hulls: A variance revisited, *Adv. Appl. Probab.* 36 (2004) 981–986; MR2119851 (2005m:60015).
[19] S. R. Finch, Geometric probability constants, *first volume*, pp. 479–484.
[20] A. Balog and I. Bárány, On the convex hull of the integer points in a disc, *Discrete and Computational Geometry: Papers from the DIMACS Special Year*, Proc. 1989/1990 New Brunswick workshops, ed. J. E. Goodman, R. Pollack and W. Steiger, Amer. Math. Soc., 1991, pp. 39–44; *Proc. 7th ACM Symp. on Computational Geometry (SCG),* North Conway, ACM, 1991, pp. 162–165; MR1143287 (93b: 11083).
[21] I. Bárány, Random points, convex bodies, lattices, *International Congress of Mathematicians*, v. 3, Proc. 2002 Beijing conf., Higher Ed. Press, pp. 527–535; math.CO/0304462; MR1957558 (2004a:52003).
[22] A. Balog and J.-M. Deshouillers, On some convex lattice polytopes, *Number Theory In Progress*, v. 2, Proc. 1997 Zakopane-Kościelisko conf., ed. K. Győry, H. Iwaniec and J. Urbanowicz, de Gruyter, 1999, pp. 591–606; MR1689533 (2000f:11083).
[23] I. Bárány and D. G. Larman, The convex hull of the integer points in a large ball, *Math. Annalen* 312 (1998) 167–181; MR1645957 (99i:52014).
[24] M. E. Dyer, Z. Füredi and C. McDiarmid, Volumes spanned by random points in the hypercube, *Random Structures Algorithms* 3 (1992) 91–106; MR1139489 (92j:52010).
[25] M. E. Dyer, Z. Füredi and C. McDiarmid, Random volumes in the n-cube, *Polyhedral Combinatorics*, Proc. 1989 Morristown conf., ed. W. Cook and P. D. Seymour, Amer. Math. Soc., 1990, pp. 33–38; MR1105114 (92i:52006).
[26] S. Finch and P. Sebah, Comment on "Volumes spanned by random points in the hypercube", *Random Structures Algorithms* 35 (2009) 390–392; MR2548520 (2010i:52008).

5.3 Volumes of Hyperbolic 3-Manifolds

Hyperbolic n-space is the n-dimensional real upper half-space

$$\mathbb{H}^n = \{\xi \in \mathbb{R}^n : x_n > 0\}, \quad \xi = (x_1, x_2, x_3, \ldots, x_n),$$

endowed with the complete Riemannian metric $ds = |d\xi|/x_n$ of constant sectional curvature equal to -1. That is, the geodesics of $\mathbb{H}^n$ consist entirely of semicircles and vertical lines that are orthogonal to the $(n-1)$-dimensional boundary $\mathbb{R}^{n-1} \times \{0\}$.

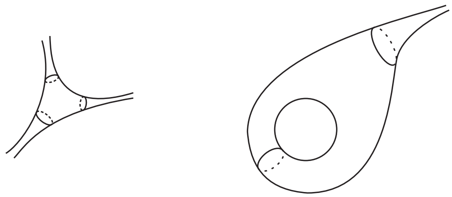

Figure 5.8 There exist two orientable surfaces with hyperbolic volume 2π: a sphere with 3 punctures and a torus with 1 puncture.

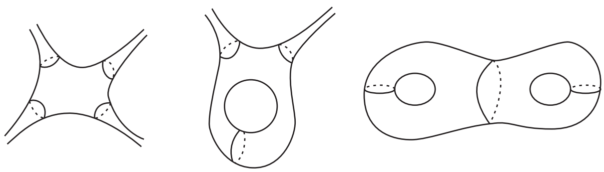

Figure 5.9 There exist three orientable surfaces with hyperbolic volume 4π: a sphere with 4 punctures, a torus with 2 punctures, and a (closed) connected sum of two tori.

A **hyperbolic *n*-manifold** M is an n-dimensional connected manifold with a complete Riemannian metric such that every point of M has a neighborhood isometric with an open subset of $\mathbb{H}^n$ [1]. Such a manifold may be either orientable or nonorientable. It is **open** if it has at least one cusp, for example, a puncture in $n=2$ (see Figures 5.8 and 5.9); otherwise it is **closed**.

From the notion of length along a geodesic proceeds the definition of volume $\mathrm{vol}(M)$ of a hyperbolic manifold. Unlike the Euclidean case, this is an important characteristic of M. If two finite-volume hyperbolic n-manifolds are homeomorphic, where $n \geq 3$, then they must be isometric. This surprising fact (false for $n=2$) is known as the Mostow–Prasad rigidity theorem [2, 3] and is believed to be crucial for the classification of 3-manifolds. We henceforth restrict attention only to manifolds with finite volume; the topological invariance of $\mathrm{vol}(M)$ follows from the Gauss–Bonnet theorem when $n=2$ and via Mostow–Prasad rigidity when $n \geq 3$.

Define the **volume spectrum** $\mathrm{spc}(n)$ to be the set of all volumes of finite-volume hyperbolic n-manifolds. It is known that [4, 5]

$$\mathrm{spc}(2) = \{2\pi k : k \geq 1\}, \quad \mathrm{spc}(4) = \left\{\frac{4\pi^2}{3}k : k \geq 1\right\}$$

but $\mathrm{spc}(3)$ is far more complicated. Let us restrict attention only to orientable 3-manifolds and call the consequential subset $\mathrm{spc}_o(3)$. Let ω denote the first infinite ordinal. Gromov, Jørgensen and Thurston [6–8] proved that $\mathrm{spc}_o(3)$ is a closed, non-discrete, well-ordered set of positive real numbers which looks like

$$\begin{aligned}v_1 &< v_2 < v_3 < \ldots < v_\omega < v_{\omega+1} < v_{\omega+2} < \ldots < v_{2\omega} < v_{2\omega+1} < \ldots \\ &< v_{3\omega} < v_{3\omega+1} < \ldots < v_{\omega^2} < v_{\omega^2+1} < \ldots < v_{\omega^3} < v_{\omega^3+1} < \ldots\end{aligned}$$

where

- v_1 is the least volume of a closed orientable 3-manifold,
- v_2 is the next smallest volume of a closed orientable 3-manifold,
- $v_\omega = \lim_{k\to\infty} v_k$ is the least volume of an (open) orientable 3-manifold with one cusp and is the first limit point in $\mathrm{spc}_o(3)$,
- $v_{2\omega} = \lim_{k\to\infty} v_{\omega+k}$ is the next smallest volume of an (open) orientable 3-manifold with one cusp and is the second limit point in $\mathrm{spc}_o(3)$,
- $v_{\omega^2} = \lim_{k\to\infty} v_{k\omega}$ is the least volume of an (open) orientable 3-manifold with two cusps and is the first limit point of limit points in $\mathrm{spc}_o(3)$.

The set $\mathrm{spc}_o(3)$ is said to have ordinal type ω^ω. For convenience, we will henceforth use the phrase "minimal manifold" to refer to a "least-volume manifold".

Weeks [9] and Matveev & Fomenko [10] independently discovered what is conjectured to be the unique minimal closed orientable 3-manifold. It has volume given by [11–13]

$$v_1 = \mathrm{Im}\left[\mathrm{Li}_2(z_0) + \ln(|z_0|)\ln(1-z_0)\right] = 0.9427073627...$$

where

$$\mathrm{Li}_2(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^2} = -\int_0^z \frac{\ln(1-u)}{u}du, \quad |z| \le 1$$

is the dilogarithm function [14] and z_0 is the zero of the cubic $z^3 - z^2 + 1$ with $\mathrm{Im}(z) > 0$. Evidence supporting this conjecture includes [15–30]; the previously best rigorous lower bound $v_1 \ge 0.324$ was strengthened to $v_1 \ge 0.547$ [31] upon confirmation of Perelman's proof of the Poincaré conjecture. The next smallest volume is conjectured to be $v_2 = 0.9813688288...$ [32]. Cao & Meyerhoff [33] proved that there exist two minimal 1-cusped orientable 3-manifolds; one of the manifolds is the complement of the figure-eight knot [34, 35] in $\mathbb{H}^3$ and has volume given by

$$\begin{aligned}v_\omega &= 2\,\mathrm{Im}\left[\mathrm{Li}_2(e^{i\pi/3})\right] = 2\,\mathrm{Cl}_2(\pi/3) = 3\,\mathrm{Cl}_2(2\pi/3) \\ &= \frac{9\sqrt{3}}{2}\sum_{n=0}^{\infty}\frac{2n+1}{(3n+1)^2(3n+2)^2} \\ &= 2(1.0149416064...) = 2.0298832128...,\end{aligned}$$

where Clausen's integral is defined by

$$\mathrm{Cl}_2(\theta) = \sum_{k=1}^{\infty} \frac{\sin(k\theta)}{k^2} = -\int_0^{\theta} \ln\left(2\sin\left(\frac{t}{2}\right)\right) dt = \mathrm{Im}\left[\mathrm{Li}_2(e^{i\theta})\right].$$

Broadhurst [36–38] found a series that can be used as a base-3 digit-extraction algorithm for v_ω:

$$v_\omega = \frac{2\sqrt{3}}{9} \sum_{n=0}^{\infty} \frac{(-1)^n}{27^n} \left(\frac{9}{(6n+1)^2} - \frac{9}{(6n+2)^2} - \frac{12}{(6n+3)^2} - \frac{3}{(6n+4)^2} + \frac{1}{(6n+5)^2} \right).$$

Define $L = v_\omega/2 = 1.0149416064...$ [39] to be **Lobachevsky's constant**, which we will need later. The next smallest volume of a 1-cusped orientable 3-manifold is conjectured to be $v_{2\omega} = 2.5689706009...$ [40, 41]. Finally, it is conjectured that the Whitehead link complement is a minimal 2-cusped orientable 3-manifold, which has volume given by [42]

$$v_{\omega^2} = 4\,\mathrm{Cl}_2(\pi/2) = 4G = 3.6638623767...$$

where G is Catalan's constant [43, 44]. Much more about $\mathrm{spc_o}(3)$ still awaits discovery.

The full set $\mathrm{spc}(n)$ is well-ordered but surprisingly different from $\mathrm{spc_o}(3)$. The minimal closed nonorientable 3-manifold appears to have volume $2L$ (the same as the figure-eight complement) [32], but the minimal 1-cusped nonorientable 3-manifold was proved by Adams [45, 46] to be what is called the Gieseking manifold, which has volume L (only half as large). The next smallest volume of a 1-cusped nonorientable 3-manifold is conjectured to be 1.8319311884.... It is known that $2L$ is also the volume of the minimal 2-cusped nonorientable 3-manifold [47].

The complement of a knot in $\mathbb{H}^3$ admits a hyperbolic structure unless it is a torus or satellite knot. Automated techniques [48] exist for computing volume and other hyperbolic invariants of 3-manifolds, which serve to distinguish knots up to homeomorphism [49–53]. The so-called "volume conjecture" relates, for any knot, the asymptotic behavior of its colored Jones polynomial evaluated at a root of unity to its volume [11, 54].

We now generalize. A **Kleinian group** is a discrete nonelementary subgroup of the group of all orientation-preserving isometries of $\mathbb{H}^3$. A **hyperbolic 3-orbifold** is a quotient of $\mathbb{H}^3$ by a Kleinian group, possibly with torsion. (An orientable 3-manifold is a special case of a 3-orbifold for which the Kleinian group is torsion-free.) The volume spectrum $\mathrm{spc}'_\mathrm{o}(3)$ of orientable 3-orbifolds is of ordinal type ω^ω [55] and is quite similar to before, where

- v'_1 is the least volume of a closed orientable 3-orbifold,
- $v'_{l\omega} = \lim_{k\to\infty} v'_{(l-1)\omega+k}$ is the l^{th} limit point in $\mathrm{spc}'_\mathrm{o}(3)$, where $l = 1, 2, 3, ...$

The unique minimal closed orientable 3-orbifold is conjectured to have volume [56–58]

$$v_1' = \frac{1}{60}\sum_{j=1}^{3} \mathrm{Im}\left[\mathrm{Li}_2(z_j) + \ln(|z_j|)\ln(1-z_j)\right] = 0.0390502856...$$

where z_1 is the zero of the quartic $z^4 - 2z^3 + z - 1$ with $\mathrm{Im}(z) > 0$, and z_2, z_3 are the two distinct zeroes of the octic $z^8 - 3z^7 + 5z^6 - 5z^5 + 3z^4 - z + 1$ satisfying both $\mathrm{Re}(z) < 1$ and $0 < \mathrm{Im}(z) < 1$. See [16, 59–62] for supporting evidence. Unlike what occurs for orientable manifolds, however, the volume u' of the minimal 1-cusped orientable 3-orbifold is not equal to the limit point v_ω'. Adams [63] and Meyerhoff [16, 64] proved that

$$u' = L/12 = 0.0845784672... < v_\omega' = G/3 = 0.3053218647....$$

In fact [65–67], the six open orientable orbifolds of volume less than $L/4$ have volumes $L/12$, $G/6$, $L/6$, $L/6$, $5L/24$, and $G/4$, whereas

$$v_{2\omega}' = \frac{7}{24}\left[\mathrm{Cl}_2\left(\frac{2\pi}{7}\right) + \mathrm{Cl}_2\left(\frac{4\pi}{7}\right) - \mathrm{Cl}_2\left(\frac{6\pi}{7}\right)\right] = 0.4444574639...,$$

$$v_{3\omega}' = \frac{G}{2} = 0.4579827970....$$

See [13, 57] for an interesting unsolved problem about linear relations involving Clausen function values. Finally [65], with regard to the full set $\mathrm{spc}'(3)$, the six open nonorientable orbifolds of volume less than $L/8$ have volumes $L/24$, $G/12$, $L/12$, $L/12$, $5L/48$, and $G/8$. The minimal closed nonorientable 3-orbifold appears not to be known. A remarkable connection between shortest geodesic lengths in closed arithmetic 3-orbifolds and Lehmer's conjecture from number theory [68] is described in [1, 69, 70].

[1] C. Maclachlan and A. W. Reid, *The Arithmetic of Hyperbolic 3-Manifolds*, Springer-Verlag, 2003; MR1937957 (2004i:57021).

[2] G. D. Mostow, *Strong Rigidity of Locally Symmetric Spaces*, Princeton Univ. Press and Univ. of Tokyo Press, 1973; MR0385004 (52 #5874).

[3] G. Prasad, Strong rigidity of *Q*-rank 1 lattices, *Invent. Math.* 21 (1973) 255–286; MR0385005 (52 #5875).

[4] J. G. Ratcliffe and S. T. Tschantz, The volume spectrum of hyperbolic 4-manifolds, *Experim. Math.* 9 (2000) 101–125; MR1758804 (2001b:57048).

[5] J. G. Ratcliffe, Hyperbolic manifolds, *Handbook of Geometric Topology*, ed. R. J. Daverman and R. B. Sher, North-Holland, 2002, pp. 899–920; MR1886683 (2003c:57017).

[6] W. P. Thurston, *The Geometry and Topology of Three-Manifolds*, unpublished manuscript, Princeton Univ., 1979, chs. 5–6.

[7] M. Gromov, Hyperbolic manifolds (according to Thurston and Jørgensen), *Séminaire Bourbaki: 1979/80*, Lect. Notes in Math. 842, Springer-Verlag, 1981, pp. 40–53; MR0636516 (84b:53046).

[8] W. P. Thurston, Three-dimensional manifolds, Kleinian groups and hyperbolic geometry, *Bull. Amer. Math. Soc.* 6 (1982) 357–381; MR0648524 (83h:57019).

[9] J. A. Weeks, *Hyperbolic Structures on 3-Manifolds*, Ph.D. thesis, Princeton Univ., 1985.

[10] S. V. Matveev and A. T. Fomenko, Isoenergetic surfaces of Hamiltonian systems, the enumeration of three-dimensional manifolds in order of growth of their complexity, and the calculation of the volumes of closed hyperbolic manifolds (in Russian), *Uspekhi Mat. Nauk* 43 (1988) 5–22, 247; Engl. transl. in *Russian Math. Surveys* 43 (1988) 3–24; MR0937017 (90a:58052).

[11] R. M. Kashaev, The hyperbolic volume of knots from the quantum dilogarithm, *Lett. Math. Phys.* 39 (1997) 269–275; q-alg/9601025; MR1434238 (98b:57012).

[12] D. J. Broadhurst, Solving differential equations for 3-loop diagrams: Relation to hyperbolic geometry and knot theory, hep-th/9806174.

[13] J. M. Borwein and D. J. Broadhurst, Determinations of rational Dedekind-zeta invariants of hyperbolic manifolds and Feynman knots and links, hep-th/9811173.

[14] S. R. Finch, Apéry's constant: Polylogarithms, *first volume*, pp. 40–53.

[15] R. Meyerhoff, A lower bound for the volume of hyperbolic 3-manifolds, *Canad. J. Math.* 39 (1987) 1038–1056; MR0918586 (88k:57049).

[16] R. Meyerhoff, Sphere-packing and volume in hyperbolic 3-space, *Comment. Math. Helv.* 61 (1986) 271–278; MR0856090 (88e:52023).

[17] F. W. Gehring and G. J. Martin, Inequalities for Möbius transformations and discrete groups, *J. Reine Angew. Math.* 418 (1991) 31–76; MR1111201 (92j:20047).

[18] F. W. Gehring and G. J. Martin, Precisely invariant collars and the volume of hyperbolic 3-folds, *J. Differential Geom.* 49 (1998) 411–435; MR1669657 (2000c:57030).

[19] A. Przeworski, Tubes in hyperbolic 3-manifolds, *Topology Appl.* 128 (2003) 103–122; MR1956607 (2004a:57023).

[20] A. Przeworski, Cones embedded in hyperbolic manifolds, *J. Differential Geom.* 58 (2001) 219–232; MR1913942 (2003e:57030).

[21] D. Gabai, R. Meyerhoff and P. Milley, Volumes of tubes in hyperbolic 3-manifolds, *J. Differential Geom.* 57 (2001) 23–46; MR1871490 (2002i:57017).

[22] T. H. Marshall and G. J. Martin, Volumes of hyperbolic 3-manifolds. Notes on a paper of D. Gabai, G. Meyerhoff, and P. Milley, *Conform. Geom. Dyn.* 7 (2003) 34–48; MR1992036 (2004e:57022).

[23] I. Agol, Volume change under drilling, *Geom. Topol.* 6 (2002) 905–916; MR1943385 (2004e:57021).

[24] A. Przeworski, Density of tube packings in hyperbolic space, *Pacific J. Math.* 214 (2004) 127–144; MR2039129 (2004k:51023).

[25] T. Chinburg, E. Friedman, K. N. Jones and A. W. Reid, The arithmetic hyperbolic 3-manifold of smallest volume, *Annali Scuola Norm. Sup. Pisa Cl. Sci.* 30 (2001) 1–40; MR1882023 (2003a:57027).

[26] S. Kojima and Y. Miyamoto, The smallest hyperbolic 3-manifolds with totally geodesic boundary, *J. Differential Geom.* 34 (1991) 175–192; MR1114459 (92f:57019).

[27] M. Culler and P. B. Shalen, Paradoxical decompositions, 2-generator Kleinian groups, and volumes of hyperbolic 3-manifolds, *J. Amer. Math. Soc.* 5 (1992) 231–288; MR1135928 (93a:57017).

[28] M. Culler and P. B. Shalen, The volume of a hyperbolic 3-manifold with Betti number 2, *Proc. Amer. Math. Soc.* 120 (1994) 1281–1288; MR1205485 (94f:57012).

[29] M. Culler, S. Hersonsky and P. B. Shalen, The first Betti number of the smallest closed hyperbolic 3-manifold, *Topology* 37 (1998) 805–849; MR1607748 (98m:57017).

[30] A. Przeworski, Volumes of hyperbolic 3-manifolds of Betti number at least 3, *Bull. London Math. Soc.* 34 (2002) 359–360; MR1887708 (2002k:57044).
[31] I. Agol, Testable predictions based on Perelman's proof, unpublished note (2004).
[32] C. D. Hodgson and J. R. Weeks, Symmetries, isometries and length spectra of closed hyperbolic three-manifolds, *Experim. Math.* 3 (1994) 261–274; MR1341719 (97a:57013).
[33] C. Cao and R. Meyerhoff, The orientable cusped hyperbolic 3-manifolds of minimum volume, *Invent. Math.* 146 (2001) 451–478; MR1869847 (2002i:57016).
[34] S. R. Finch, Knots, links and tangles, *this volume*, §5.1.
[35] J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*, Springer-Verlag, 1994; MR1299730 (95j:57011)
[36] D. J. Broadhurst, The master two-loop diagram with masses, *Z. Phys. C – Part. Fields* 47 (1990) 115–124.
[37] D. J. Broadhurst, Massive 3-loop Feynman diagrams reducible to SC* primitives of algebras of the sixth root of unity, *Europ. Phys. J. C Part. Fields* 8 (1999) 313–333; hep-th/9803091; MR1739510 (2002a:81180).
[38] J. M. Borwein and D. H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century*, A. K. Peters, 2004, pp. 88–91; MR2033012 (2005b:00012).
[39] J. Milnor, Hyperbolic geometry: The first 150 years, *Bull. Amer. Math. Soc.* 6 (1982) 9–24; MR0634431 (82m:57005).
[40] M. V. Hildebrand and J. A. Weeks, A computer generated census of cusped hyperbolic 3-manifolds, *Computers and Mathematics*, ed. E. Kaltofen and S. M. Watt, Springer-Verlag, 1989, pp. 53–59; MR1005959 (90f:57043).
[41] P. J. Callahan, M. V. Hildebrand and J. R. Weeks, A census of cusped hyperbolic 3-manifolds, *Math. Comp*. 68 (1999) 321–332; MR1620219 (99c:57035).
[42] H. Yoshida, Volumes of orientable 2-cusped hyperbolic 3-manifolds, *Kobe J. Math.* 18 (2001) 147–161; MR1907670 (2003f:57036).
[43] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[44] S. R. Finch, Kneser-Mahler polynomial constants, *first volume*, pp. 231–235.
[45] C. C. Adams, The noncompact hyperbolic 3-manifold of minimal volume, *Proc. Amer. Math. Soc.* 100 (1987) 601–606; MR0894423 (88m:57018).
[46] C. C. Adams, The newest inductee in the Number Hall of Fame, *Math. Mag.* 71 (1998) 341–349.
[47] C. C. Adams, Volumes of N-cusped hyperbolic 3-manifolds, *J. London Math. Soc.* 38 (1988) 555–565; MR0972138 (89k:22020).
[48] J. R. Weeks, SnapPea: A computer program for creating and studying hyperbolic 3-manifolds.
[49] C. C. Adams, *The Knot Book: An Elementary Introduction to the Mathematical Theory of Knots*, W. H. Freeman, 1994; MR1266837 (94m:57007).
[50] C. C. Adams, M. V. Hildebrand and J. R. Weeks, Hyperbolic invariants of knots and links, *Trans. Amer. Math. Soc.* 326 (1991) 1–56; MR0994161 (91j:57004).
[51] P. J. Callahan and A. W. Reid, Hyperbolic structures on knot complements, *Chaos Solitons Fractals* 9 (1998) 705–738; MR1628752 (99e:57022).
[52] P. J. Callahan, J. C. Dean and J. R. Weeks, The simplest hyperbolic knots, *J. Knot Theory Ramifications* 8 (1999) 279–297; MR1691433 (2000c:57005).
[53] A. Champanerkar, I. Kofman and E. Patterson, The next simplest hyperbolic knots, *J. Knot Theory Ramifications* 13 (2004) 965–987; math.GT/0311380; MR2101238 (2005k:57010).
[54] H. Murakami, J. Murakami, M. Okamoto, T. Takata and Y. Yokota, Kashaev's conjecture and the Chern-Simons invariants of knots and links, *Experim. Math.* 11 (2002) 427–435; MR1959752 (2004a:57016).

[55] W. D. Dunbar and R. Meyerhoff, Volumes of hyperbolic 3-orbifolds, *Indiana Univ. Math. J.* 43 (1994) 611–637; MR1291531 (95g:57024).
[56] A. Borel, Commensurability classes and volumes of hyperbolic 3-manifolds, *Annali Scuola Norm. Sup. Pisa Cl. Sci.* 8 (1981) 1–33; *Collected Works*, v. 3, Springer-Verlag, 1983, pp. 617–649; MR0616899 (82j:22008).
[57] D. Zagier, Hyperbolic manifolds and special values of Dedekind zeta-functions, *Invent. Math.* 83 (1986) 285–301; MR0818354 (87e:11069).
[58] D. J. Broadhurst, Volume of manifold m016$(-4,3)$ with quartic invariant trace field and discriminant -275, unpublished note (2004).
[59] R. Meyerhoff, A lower bound for the volume of hyperbolic 3-orbifolds, *Duke Math. J.* 57 (1988) 185–203; MR0952231 (89f:57059).
[60] F. W. Gehring and G. J. Martin, On the minimal volume hyperbolic 3-orbifold, *Math. Res. Lett.* 1 (1994) 107–114; MR1258496 (95b:30072).
[61] F. W. Gehring and G. J. Martin, The volume of hyperbolic 3-folds with p-torsion, $p \geq 6$, *Quart. J. Math. Oxford Ser.* 50 (1999) 1–12; MR1673252 (2000c:57031).
[62] T. Chinburg and E. Friedman, The smallest arithmetic hyperbolic three-orbifold, *Invent. Math.* 86 (1986) 507–527; MR0860679 (88a:22022).
[63] C. C. Adams, Limit volumes of hyperbolic three-orbifolds, *J. Differential Geom.* 34 (1991) 115–141; MR1114455 (92d:57029).
[64] R. Meyerhoff, The cusped hyperbolic 3-orbifold of minimum volume, *Bull. Amer. Math. Soc.* 13 (1985) 154–156; MR0799800 (87b:22022).
[65] C. C. Adams, Noncompact hyperbolic 3-orbifolds of small volume, *Topology '90,* ed. B. Apanasov, W. D. Neumann, A. W. Reid and L. Siebenmann, de Gruyter, 1992, pp. 1–15; MR1184398 (93h:57019).
[66] W. D. Neumann and A. W. Reid, Notes on Adams' small volume orbifolds, *Topology '90,* ed. B. Apanasov, W. D. Neumann, A. W. Reid and L. Siebenmann, de Gruyter, 1992, pp. 311–314; MR1184417 (94a:57028).
[67] C. C. Adams, Volumes of hyperbolic 3-orbifolds with multiple cusps, *Indiana Univ. Math. J.* 41 (1992) 149–172; MR1160907 (93c:57011).
[68] S. R. Finch, Pisot-Vijayaraghavan-Salem constants, *first volume*, pp. 192–199.
[69] W. D. Neumann and A. W. Reid, Arithmetic of hyperbolic manifolds, *Topology '90,* ed. B. Apanasov, W. D. Neumann, A. W. Reid and L. Siebenmann, de Gruyter, 1992, pp. 273–310; MR1184416 (94c:57024).
[70] E. Ghate and E. Hironaka, The arithmetic and geometry of Salem numbers, *Bull. Amer. Math. Soc.* 38 (2001) 293–314; MR1824892 (2002c:11137).

5.4 Poisson–Voronoi Tessellations

The **d-dimensional Poisson process** of **intensity** λ is a random scattering of points (called **particles**) in $\mathbb{R}^d$ that meets the following two requirements. Let $S \subseteq \mathbb{R}^d$ denote a measurable set of finite volume μ and $N(S)$ denote the number of particles falling in S. We have [1, 2]

- $\mathrm{P}\{N(S)=n\}=e^{-\lambda\mu}(\lambda\mu)^n/n!$ for any S, for any $n=0,1,2,\ldots$
- if $S_1,\ldots,S_k$ are disjoint measurable sets, then $N(S_1),\ldots,N(S_k)$ are independent random variables.

In particular, the location of S in $\mathbb{R}^d$ is immaterial (stationarity) and $\mathrm{E}(N(S))=\lambda\mu=\mathrm{Var}(N(S))$ (equality of mean and variance). An alternative characterization of the Poisson process involves the limit of the uniform distribution on expanding cubes $C \subseteq \mathbb{R}^d$. Let ν denote the volume of C. Given m independent uniformly distributed particles in C and a measurable set $S \subseteq C$ of volume μ, the probability that exactly n particles fall in S is

$$\frac{m!}{n!(m-n)!}\left(\frac{\mu}{\nu}\right)^n\left(1-\frac{\mu}{\nu}\right)^{m-n} \to e^{-\lambda\mu}\frac{(\lambda\mu)^n}{n!},$$

which occurs in the limit as $\nu \to \infty$ in such a way that $m/\nu \to \lambda$. The interpretation of λ as a rate or intensity is thus clear, as is the phrase *binomial process* to denote a Uniform (C) distribution.

Here is a sample problem involving the Poisson process; assume for simplicity henceforth that $\lambda = 1$. Let ξ be an arbitrary point in $\mathbb{R}^d$ and R denote the distance from ξ to its nearest neighboring particle. What can be said about R? If $\omega_d = \pi^{d/2}\Gamma(d/2+1)^{-1}$ is the volume of the unit d-ball, then [3–5]

$$\mathrm{P}\{R > r\} = \mathrm{P}\{d\text{-ball of radius } r \text{ contains no particles}\} = e^{-\omega_d r^d},$$

which implies that

$$\begin{aligned} \mathrm{E}(R) &= \omega_d^{-1/d}\Gamma\left(\frac{1}{d}+1\right) = \begin{cases} \dfrac{1}{2} & \text{if } d=1 \text{ or } 2, \\ \left(\dfrac{3}{4\pi}\right)^{1/3}\Gamma\left(\dfrac{4}{3}\right) & \text{if } d=3 \end{cases} \\ &= \begin{cases} 0.5 & \text{if } d=1 \text{ or } 2, \\ 0.5539602783... & \text{if } d=3. \end{cases} \end{aligned}$$

Likewise,

$$\mathrm{E}(R^2) = \omega_d^{-2/d}\Gamma\left(\frac{2}{d}+1\right) = \begin{cases} \dfrac{1}{2} & \text{if } d=1, \\ \dfrac{1}{\pi} & \text{if } d=2, \\ \left(\dfrac{3}{4\pi}\right)^{2/3}\Gamma\left(\dfrac{5}{3}\right) & \text{if } d=3 \end{cases}$$

and thus

$$\mathrm{Var}(R) = \mathrm{E}(R^2) - \mathrm{E}(R)^2 = \begin{cases} 0.25 & \text{if } d=1, \\ 0.0683098861... & \text{if } d=2, \\ 0.0405357524... & \text{if } d=3. \end{cases}$$

We will consider a vastly more difficult version of this problem shortly. Of all unit-intensity scattering methods, the Poisson process is the "most random"; hence the forthcoming constants deserve to be better understood!

5.4.1 *Cellular Parameters*

Given any set of distinct particles $\{p_i\}_{i=1}^{\infty}$ in $\mathbb{R}^d$, the corresponding **Voronoi tessellation** is the subdivision of $\mathbb{R}^d$ into convex polyhedral cells $\{\Pi_i\}_{i=1}^{\infty}$ with the property that Π_i contains all points in $\mathbb{R}^d$ closer to p_i than to any other $p_j, j \neq i$. If $d=1$, the cells are subintervals of the line characterized simply by length. If $d \geq 2$, the geometry is more elaborate. Our interest is in the scenario when the particles are realizations of a Poisson process of intensity 1; hence the cellular parameters are random variables. Applications of this material include any field involving pattern analysis: astronomy, geography, metallurgy, biology and socio-economic planning, to mention only a few [6, 7].

If $d=1$ and M denotes the length of a typical cell, then $\mathrm{E}(M)=1$ and $\mathrm{Var}(M)=1/2$ [8]. If $d=2$ or 3, the associated mean values are known exactly [8], but the derivation of second moment integrals is notoriously difficult. A closed-form expression has not been found for any of these integrals.

For the following, define expressions [9, 10]

$$\begin{aligned} f_V(x,y) \;=\; & 4\left[(\pi/2+x)(1+2\sin(x)^2)+3\sin(x)\cos(x)\right]\sec(x)^5 \cdot \\ & \left[(\pi/2+y)(1+2\sin(y)^2)+3\sin(y)\cos(y)\right]\sec(y)^5, \end{aligned}$$

$$f_L(x,y)=((\pi/2+x)\tan(x)+1)\sec(x)^2((\pi/2+y)\tan(y)+1)\sec(y)^2,$$

$$f_P(x,y)=(1+\sin(x))\sec(x)^4(1+\sin(y))\sec(y)^4,$$

$$f_M(x,y)=\sec(x)^3\sec(y)^3,$$

$$g(x,y)=(\pi/2+x+\sin(x)\cos(x))\sec(x)^2+(\pi/2+y+\sin(y)\cos(y))\sec(y)^2,$$

$$\begin{aligned} h(\rho,\theta) \;=\; & \rho^2(\pi-\theta+\sin(2\theta)/2)+(1+\rho^2-2\rho\cos(\theta))(\pi-\kappa(\rho,\theta) \\ & +\sin(2\kappa(\rho,\theta)/2) \end{aligned}$$

where

$$\kappa(\rho,\theta)=\arccos\left(\frac{1-\rho\cos(\theta)}{\sqrt{1+\rho^2-2\rho\cos(\theta)}}\right).$$

A geometric interpretation of $h(\rho,\theta)$ is as the area of the union of two overlapping planar disks with unit distance between their centers, one with radius ρ and the other with radius $\sqrt{1+\rho^2-2\rho\cos(\theta)}$. When $d=2$, we have [10–13]

$$\mathrm{E}(V)=6,$$

$$\mathrm{E}(V^2)=12\pi\int\limits_{-\pi/2}^{\pi/2}\int\limits_{-x}^{\pi/2} f_V(x,y)g(x,y)^{-4}\sin(x+y)\,dy\,dx+18,$$

$$\mathrm{Var}(V)=1.7808116990\ldots=37.7808116990\ldots-\mathrm{E}(V)^2,$$

where V is the number of vertices of the cell; [10, 11]

$$\mathrm{E}(L) = 5\pi^{3/2} \int_{-\pi/2}^{\pi/2} \int_{-x}^{\pi/2} f_L(x,y) g(x,y)^{-7/2} \left(\tan(x) + \tan(y)\right) dy\, dx = \frac{2}{3},$$

$$\mathrm{E}(L^2) = 16\pi \int_{-\pi/2}^{\pi/2} \int_{-x}^{\pi/2} f_L(x,y) g(x,y)^{-4} \left(\tan(x) + \tan(y)\right)^2 dy\, dx,$$

$$\mathrm{Var}(L) = 0.1856273347... = 0.6300717791... - \mathrm{E}(L)^2$$

where L is the length of an arbitrary edge;

$$\mathrm{E}(P) = 4,$$

$$\mathrm{E}(P^2) = 64\pi \int_{-\pi/2}^{\pi/2} \int_{-x}^{\pi/2} f_P(x,y) g(x,y)^{-3} \sin(x+y)\, dy\, dx + 6\mathrm{E}(L^2),$$

$$\mathrm{Var}(P) = 0.9454930107... = 16.9454930107... - \mathrm{E}(P)^2$$

where $P = \sum L$ is the total perimeter; and [9–11, 14–16]

$$\mathrm{E}(M) = 1,$$

$$\begin{aligned} \mathrm{E}(M^2) &= 2\pi \int_{-\pi/2}^{\pi/2} \int_{-x}^{\pi/2} f_M(x,y) g(x,y)^{-2} \sin(x+y)\, dy\, dx \\ &= 2\pi \int_0^{\infty} \int_0^{\pi} \rho\, h(\rho,\theta)^{-2}\, d\theta\, d\rho, \end{aligned}$$

$$\mathrm{Var}(M) = 0.2801760409... = 1.2801760409... - \mathrm{E}(M)^2$$

where M is the area of the cell. It is also known that $\mathrm{E}(M^3) = 1.999...$ [15].

For the following, define expressions [9, 17]

$$f_L(x,y) = \sec(x)^2 (\sec(x) + \tan(x))^2 \sec(y)^2 (\sec(y) + \tan(y))^2,$$

$$g(x,y) = \sec(x)^3 (2/3 + \sin(x) - \sin(x)^3/3) + \sec(y)^3 (2/3 + \sin(y) - \sin(y)^3/3),$$

$$\begin{aligned} h(\rho,\theta) &= \pi\rho^3 \left[2/3 + 3\cos(\theta)/4 - \cos(3\theta)/12\right] + \pi(1 + \rho^2 - 2\rho\cos(\theta))^{3/2} \cdot \\ &\quad \left[2/3 + 3\cos(\kappa(\rho,\theta))/4 - \cos(3\kappa(\rho,\theta))/12\right] \end{aligned}$$

and $\kappa(\rho,\theta)$ is as before. A geometric interpretation of $h(\rho,\theta)$ as the volume of the union of two spatial balls again holds. When $d=3$, we have [17]

$$\mathrm{E}(W)=\frac{144\pi^2}{24\pi^2+35}=5.2275734378...,$$

$$\mathrm{Var}(W)=2.4846406759...=29.8121647244-\mathrm{E}(W)^2$$

where W is the number of vertices of an arbitrary face of the cell; [12, 13, 17]

$$\mathrm{E}(V)=\frac{96\pi^2}{35}=27.0709149287...,$$

$$\mathrm{Var}(V)=44.4983886849...=777.3328237620-\mathrm{E}(V)^2$$

where $V=\sum W$ is the total number of vertices; [17]

$$\mathrm{E}(E)=\frac{144\pi^2}{35}=40.6063723930...,$$

$$\mathrm{Var}(E)=100.1213745412...=1748.9988534645...-\mathrm{E}(E)^2$$

where $E=3V/2$ is the number of edges;

$$\mathrm{E}(F)=\frac{48\pi^2}{35}+2=15.5354574643...,$$

$$\mathrm{Var}(F)=11.1245971712...=252.4750357979...-\mathrm{E}(F)^2$$

where $F=V/2+2$ is the number of faces;

$$\begin{aligned}\mathrm{E}(L) &= \frac{35}{36\pi^{1/3}}\Gamma\left(\frac{13}{3}\right)\int_{-\pi/2}^{\pi/2}\int_{-x}^{\pi/2} f_L(x,y)\,g(x,y)^{-13/3}\,(\tan(x)+\tan(y))\,dy\,dx\\ &= \frac{7}{9}\left(\frac{3}{4\pi}\right)^{1/3}\Gamma\left(\frac{4}{3}\right)=0.4308579942...,\end{aligned}$$

$$\mathrm{E}(L^2)=\frac{35}{36\pi^{2/3}}\Gamma\left(\frac{14}{3}\right)\int_{-\pi/2}^{\pi/2}\int_{-x}^{\pi/2} f_L(x,y)\,g(x,y)^{-14/3}\,(\tan(x)+\tan(y))^2\,dy\,dx,$$

$$\mathrm{Var}(L)=0.1052391356...=0.2908777468...-\mathrm{E}(L)^2$$

where L is the length of an arbitrary edge;

$$\mathrm{E}(Q)=\frac{21}{24\pi^2+35}\left(\frac{4\pi}{3}\right)^{5/3}\Gamma\left(\frac{1}{3}\right)=2.2523418064...,$$

$$\mathrm{Var}(Q)=1.4699757822...=6.5430193952...-\mathrm{E}(Q)^2$$

where Q is the perimeter of an arbitrary face;

$$\mathrm{E}(P)=\frac{3}{5}\left(\frac{4\pi}{3}\right)^{5/3}\Gamma\left(\frac{1}{3}\right)=17.4955801644...,$$

$$\mathrm{Var}(P)=13.6179400522...=319.7132653418...-\mathrm{E}(P)^2$$

where $P=\sum L=\sum Q/2$ is the total perimeter;

$$\mathrm{E}(B)=\frac{35}{24\pi^2+35}\left(\frac{256\pi}{81}\right)^{1/3}\Gamma\left(\frac{2}{3}\right)=0.3746830505...,$$

$$\mathrm{Var}(B)=0.1423896695...=0.2827770579-\mathrm{E}(B)^2$$

where B is the surface area of an arbitrary face;

$$\mathrm{E}(A)=\left(\frac{256\pi}{3}\right)^{1/3}\Gamma\left(\frac{5}{3}\right)=5.8208725950...,$$

$$\mathrm{Var}(A)=2.1914834552...=36.0740412231-\mathrm{E}(A)^2$$

where $A=\sum B$ is the total surface area; and [9, 14, 16, 17]

$$\mathrm{E}(M)=1,$$

$$\mathrm{E}(M^2)=\frac{8\pi^2}{3}\int_0^\infty\int_0^\pi \rho^2\sin(\theta)\,h(\rho,\theta)^{-2}\,d\theta\,d\rho,$$

$$\mathrm{Var}(M)=0.1790324378...=1.1790324378...-\mathrm{E}(M)^2$$

where M is the volume of the cell.

5.4.2 *Vertex Counts*

Thus far we have discussed only moments of distributions associated with Poisson–Voronoi cells. The computation of actual probabilities seems to be hard. If $d=2$, for example, what is the probability that an arbitrary cell is a triangle? The solution can be expressed as a complicated quadruple integral and turns out numerically to be [18–20]

$$\mathrm{P}(V=3)=0.01124001....$$

Simulation can be used to verify this result and the preceding moment estimates as well [21–29]; for example, it appears that $\mathrm{P}(V=4)=0.1608...$ and $\mathrm{P}(V=5)=0.2594...$ [30]. Integral formulas for these latter probabilities [30, 31] evidently require further simplification to be numerically feasible. The function

$\mathrm{P}(V=n)$ is apparently maximized when $n=6$ and falls off for $n\geq 7$; it is known that asymptotically [32, 33]

$$\mathrm{P}(V=n)=\frac{C}{4\pi^2}\frac{(8\pi^2)^n}{(2n)!}\left(1+O(n^{-1})\right)$$

as $n\to\infty$, where

$$\begin{aligned} C &= \prod_{j=1}^{\infty}\left(1-\frac{1}{j^2}+\frac{4}{j^4}\right)^{-1}=0.3443473089... \\ &= 4\cdot\left|\Gamma\left(\tfrac{\sqrt{5}}{2}+i\tfrac{\sqrt{3}}{2}\right)\right|^2\cdot\left|\Gamma\left(-\tfrac{\sqrt{5}}{2}-i\tfrac{\sqrt{3}}{2}\right)\right|^2=4\pi^2\left(\cosh(\pi\sqrt{3})-\cos(\pi\sqrt{5})\right)^{-1}. \end{aligned}$$

Other questions can be conditional in nature. If a cell is known to be a triangle, what is its expected area and its expected perimeter? Brakke [10] computed that these quantities are 0.343089... and 2.740297..., respectively, and subsequent study [15] confirmed these estimates to four decimal places. (The work in [10, 11, 17] has unfortunately remained quite obscure.) See also [34] for more about the distribution of edge lengths L in $\mathbb{R}^d$ and [35] for inradius/circumradius-type analysis of cells in the plane.

The Goudsmit–Miles tessellation of the plane, which is based on the Poisson line process (as opposed to a point process), is discussed in [36].

5.4.3 *Stienen Spheres*

Around each particle $p_i\in\mathbb{R}^d$, construct a sphere with diameter equal to the distance to the nearest neighbor p_j of p_i, $i\neq j$. The union of all such spheres and their interiors is called the **Stienen model**. Each sphere is a subset of a Voronoi cell; each cell is a superset of a Stienen sphere. For arbitrary d, if M' denotes the volume of a typical sphere, then $\mathrm{E}(M')=2^{-d}$ and $\mathrm{Var}(M')=2^{-2d}$. If $d=1$, the cross-correlation ρ between M and M' is simply $1/\sqrt{2}$. For $d=2$ and 3, Olsbo [37] computed $\rho=0.705143...$ and $\rho=0.677790...$ via complicated numerical integration. It is not obvious that these correlations are necessarily positive because two neighboring particles lying close together often yield small spheres and large cells.

Addendum Simplification of various double integrals in [38, 39] gives rise to closed-form expressions involving a new constant [40]:

$$\begin{aligned} c &= 2\int_1^{\sqrt{3}}\frac{(1+z^2)\arctan(z)}{1-\frac{14}{9}z^2+z^4}dz+\int_0^1\frac{(3z-1)\operatorname{arctanh}(z)}{1-\frac{2}{3}z+z^2}dz \\ &= 3.4954848920..., \end{aligned}$$

which can be rewritten as a sum of dilogarithms with complex algebraic arguments. For example, the second volume moment (for $d = 3$) becomes

$$\mathrm{E}(M^2) = -\frac{4}{243}\left(16c - 3\pi^2\right) + \frac{8\sqrt{3}}{27}\pi = 1.1790324378...$$

in agreement with before. The third edge-length moment, as another example, is 0.2451902663.... We will only mention the existence of other geometric characteristics: edges lengths in an s-dimensional section, $s < d$, and the linear contact distribution.

The recovery of Brakke's original integrals for $\mathrm{E}(W^2)$, $\mathrm{E}(V^2)$, $\mathrm{E}(Q^2)$, $\mathrm{E}(P^2)$, $\mathrm{E}(B^2)$, $\mathrm{E}(A^2)$ and $\mathrm{E}(M^2)$ when $d = 3$ is a longstanding challenge!

[1] M. H. DeGroot, *Probability and Statistics*, Addison-Wesley, 1975, pp. 209–211.
[2] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer-Verlag, 1988, pp. 18–24; MR0950166 (90e:60060).
[3] P. Hertz, Über den gegenseitigen durchschnittlichen Abstand von Punkten, die mit bekannter mittlerer Dichte im Raume angeordnet sind, *Math. Annalen* 67 (1909) 387–398.
[4] S. Chandrasekhar, Stochastic problems in physics and astronomy, *Rev. Mod. Phys.* 15 (1943) 1–89; also in *Selected Papers on Noise and Stochastic Processes*, ed. N. Wax, Dover, 1954, pp. 3–91.
[5] E. Parzen, *Stochastic Processes*, Holden-Day, 1962, pp. 32–34; MR0139192 (25 #2628).
[6] A. Okabe, B. Boots, K. Sugihara and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed., Wiley, 2000, pp. 31–33, 44–50, 291–294, 311–318; MR1210959 (94a:52033) and MR1770006 (2001c:52013).
[7] D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, Wiley, 1987, pp. 328–333; MR0895588 (88j:60034a).
[8] J. L. Meijering, Interface area, edge length, and number of vertices in crystal aggregates with random nucleation, *Philips Research Reports* 8 (1953) 270–290.
[9] E. N. Gilbert, Random subdivisions of space into crystals, *Annals Math. Statist.* 33 (1962) 958–972; MR0144253 (26 #1800).
[10] K. A. Brakke, Statistics of random plane Voronoi tessellations, unpublished manuscript (1985).
[11] K. A. Brakke, Plane Voronoi tessellation second order statistics, *Abstracts Amer. Math. Soc.* 7 (1986) 235–236.
[12] L. Heinrich, R. Körner, N. Mehlhorn and L. Muche, Numerical and analytical computation of some second-order characteristics of spatial Poisson-Voronoi tessellations, *Statistics* 31 (1998) 235–259; MR1718316 (2000g:60016).
[13] L. Heinrich and L. Muche, Second-order properties of the point process of nodes in a stationary Voronoi tessellation, *Math. Nachr.* 281 (2008) 350–375; erratum 283 (2010) 1674–1676; MR2392118 (2008k:60025).
[14] J. Møller, *Lectures on Random Voronoĭ Tessellations.*, Lect. Notes in Statistics 87, Springer-Verlag, 1994, pp. 83–103; MR1295245 (96b:60028).
[15] A. Hayen and M. P. Quine, Areas of components of a Voronoi polygon in a homogeneous Poisson process in the plane, *Adv. Appl. Probab.* 34 (2002) 281–291; MR1909915 (2003d:60025).

[16] E. Pineda, P. Bruna and D. Crespo, Cell size distribution in random tessellations of space, *Phys. Rev. E* 70 (2004) 066119.
[17] K. A. Brakke, Statistics of three dimensional random Voronoi tessellations, unpublished manuscript (1985).
[18] R. E. Miles and R. J. Maillardet, The basic structures of Voronoi and generalized Voronoi polygons, *J. Appl. Probab.* special volume 19A, (1982) 97–111; MR0633183 (83d:60021).
[19] A. Hayen and M. P. Quine, The proportion of triangles in a Poisson-Voronoi tessellation of the plane, *Adv. Appl. Probab.* 32 (2000) 67–74; MR1765171 (2001h:60019).
[20] A. Hayen and M. P. Quine, Calculating the proportion of triangles in a Poisson-Voronoi tessellation of the plane, *J. Statist. Comput. Simul.* 67 (2000) 351–358; MR1815169.
[21] A. L. Hinde and R. E. Miles, Monte-Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process, *J. Statist. Comput. Simul.* 10 (1980) 205–223.
[22] H. G. Hanson, Voronoi cell properties from simulated and real random spheres and points, *J. Stat. Phys.* 30 (1983) 591–605; MR0710751 (85c:52012).
[23] K. A. Brakke, 2,000,000,000 random Voronoi polygons, unpublished manuscript (1985).
[24] U. Lorz, Distribution of cell characteristics of the spatial Poisson-Voronoi tessellation and plane sections, *Geometrical Problems of Image Processing*, Proc. 1991 Georgenthal workshop, ed. U. Eckhardt, A. Hübler, W. Nagel and G. Werner, Akademie-Verlag, 1991, pp. 171–178; MR1111686 (91m:68007).
[25] U. Lorz and U. Hahn, Geometric characteristics of spatial Voronoi tessellations and planar sections, Fakultät für Mathematik und Informatik, Technische Universität Bergakademie Freiberg preprint 93-05 (1993).
[26] S. Kumar, S. K. Kurtz, J. R. Banavar and M. G. Sharma, Properties of a three-dimensional Poisson-Voronoi tessellation: a Monte Carlo study, *J. Stat. Phys.* 67 (1992) 523–551.
[27] S. Kumar and S. K. Kurtz, Properties of a two-dimensional Poisson–Voronoi tessellation: a Monte-Carlo study, *Materials Characterization* 31 (1993) 55–68.
[28] S. Kumar and S. K. Kurtz, Monte-Carlo study of angular and edge length distributions in a three-dimensional Poisson-Voronoi tesselation, *Materials Characterization* 34 (1995) 15–27.
[29] M. Tanemura, Statistical distributions of Poisson Voronoi cells in two and three dimensions, *Forma* 18 (2003) 221–247; MR2040084.
[30] P. Calka, An explicit expression for the distribution of the number of sides of the typical Poisson-Voronoi cell, *Adv. Appl. Probab.* 35 (2003) 863–870; MR2014258 (2004g:60019).
[31] P. Calka, Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional Poisson-Voronoi tessellation and a Poisson line process, *Adv. Appl. Probab.* 35 (2003) 551–562; MR1990603 (2004k:60021).
[32] H. J. Hilhorst, The perimeter of large planar Voronoi cells: a double-stranded random walk, *J. Stat. Mech.* (2005) L02003; cond-mat/0412569.
[33] H. J. Hilhorst, New Monte Carlo method for planar Poisson-Voronoi cells, *J. Phys. A* 40 (2007) 2615–2638; cond-mat/0612422; MR2325498 (2008f:60016).
[34] M. Schlather, A formula for the edge length distribution function of the Poisson Voronoi tessellation, *Math. Nachr.* 214 (2000) 113–119; MR1762055 (2001b: 60021).

[35] P. Calka, The distributions of the smallest disks containing the Poisson-Voronoi typical cell and the Crofton cell in the plane, *Adv. Appl. Probab.* 34 (2002) 702–717; MR1938938 (2003j:60013).
[36] S. R. Finch, Random triangles. V, *this volume*, §5.14.
[37] V. Olsbo, On the correlation between the volumes of the typical Poisson-Voronoi cell and the typical Stienen sphere, *Adv. Appl. Probab.* 39 (2007) 883–892; MR2381579 (2009a:60009).
[38] L. Muche, The Poisson-Voronoi tessellation: relationships for edges, *Adv. Appl. Probab.* 37 (2005) 279–296; MR2144554 (2005m:60017).
[39] L. Muche, Contact and chord length distribution functions of the Poisson-Voronoi tessellation in high dimensions, *Adv. Appl. Probab.* 42 (2010) 48–68; MR2666918 (2011e:60021).
[40] L. Muche and F. Ballani, The second volume moment of the typical cell and higher moments of edge lengths of the spatial Poisson-Voronoi tessellation, *Monatsh. Math.* 163 (2011) 71–80; MR2787583 (2012h:60041).

5.5 Optimal Escape Paths

A summary of Bellman's "Lost in a Forest" problem appears in [1]. Certain allied constants are described in [2, 3] and research is ongoing [4–6]. We will focus on just one facet of the problem for now, namely the following:

A hiker is lost in a forest whose shape is known to be a half-plane. What is the best path for him to follow to escape from the forest?

This is equivalent to:

A swimmer is lost in a dense fog at sea, and she knows that the shore is a line. What is the best path for her to follow to search for the shore?

Since no information is available concerning the initial distance or orientation of the boundary, a candidate path must be unbounded. Baeza-Yates, Culberson & Rawlins [7–9] claimed that the best path (which minimizes the maximum escape time) is a logarithmic spiral. Their argument was based on symmetry; a proof via the calculus of variations is still sought after [5, 6].

Speed is constant, thus escape time is proportional to arclength. If we assume that a logarithmic spiral $r = e^{\kappa\theta}$ is indeed optimal, then straightforward analysis leads to the best value of the parameter κ. Let the initial (unknown) distance from the boundary be R. Then the min-max logarithmic spiral can be shown to have parameter

$$\kappa = \tan\alpha = 0.2124695594... = \ln(1.2367284662...)$$

with arclength

$$R\csc\alpha\sec\beta = (13.8111351795...)R,$$

where α, β satisfy the simultaneous equations

$$\frac{1}{\tan\alpha} + \frac{1}{\tan\beta} = \frac{2\pi - \alpha - \beta}{\cos^2\alpha}, \quad \frac{\cos\alpha}{\cos\beta} = e^{(2\pi-\alpha-\beta)\tan\alpha}.$$

It is surprising that such interesting constants emerge here, yet frustrating that a gap in the proof (for such a simple forest/sea) should persist.

5.5.1 *Growth of Squares*

While on the subject of logarithmic spirals, it seems natural to continue a discussion begun in [10]. Let $f_1 = 1, f_2 = 1, f_3 = 2, \ldots$ denote the Fibonacci sequence and $\varphi = (1 + \sqrt{5})/2$ denote the Golden mean. In the xy plane, draw the 1×1 square with center $(1/2, 1/2)$, then the adjacent 1×1 square with center $(-1/2, 1/2)$, then the adjacent 2×2 square with center $(0, -1)$, then the adjacent 3×3 square with center $(5/2, -1/2)$, then the adjacent 5×5 square with center $(3/2, 7/2)$, and so forth (in a counterclockwise manner). The n^{th} square is $f_n \times f_n$ and shares an edge between the two squares preceding it. Supposing we now translate the origin to the point $(2/5, 1/5)$, the logarithmic spiral $r = e^{\kappa\theta + \lambda}$ then asymptotically approaches the $f_n \times f_n$ square centers as $n \to \infty$, where [11]

$$\kappa = \frac{2}{\pi}\ln(\varphi) = 0.3063489625...,$$

$$\lambda_{\text{center}} = \frac{1}{2}\ln\left(\frac{\varphi+1}{10}\right) - \arctan(3)\kappa = -1.0527245979....$$

In the squares just constructed, consider instead the **leading vertices**

$$(0,1),\quad (-1,0),\quad (1,-2),\quad (4,1),\quad (-1,6),\quad \ldots$$

and the **trailing vertices**

$$(1,1),\quad (-1,1),\quad (-1,-2),\quad (4,-2),\quad (4,6),\quad \ldots$$

in the original coordinate system [11]. After translation (as before), the two associated asymptotic spirals possess the same κ but different λs:

$$\lambda_{\text{lead}} = \frac{1}{2}\ln\left(\frac{2(\varphi+2)}{25}\right) - \arctan(2\varphi-3)\kappa = -0.6909179135...,$$

$$\lambda_{\text{trail}} = \frac{1}{2}\ln\left(\frac{11\varphi+7}{25}\right) - \arctan(\varphi)\kappa = -0.3156737662....$$

There exists a nice duality between this material (starting with a square and concatenating) and earlier material (starting with a Golden rectangle and partitioning). In Figure 1.2 of [10], supposing we translate the origin to the point $((1+3\varphi)/5, (3-\varphi)/5)$, the spiral pictured there possesses the same κ but yet another λ:

$$\lambda'_{\text{lead}} = \frac{1}{2}\ln\left(\frac{2(\varphi+2)}{5}\right) - (\pi + \arctan(2\varphi-3))\,\kappa = -0.8486226074....$$

Other variations suggest themselves.

[1] S. R. Finch and J. E. Wetzel, Lost in a forest, *Amer. Math. Monthly* 11 (2004) 645–654; MR2091541 (2006d:51016).
[2] S. R. Finch, Moser's worm constant, *first volume*, pp. 491–497.
[3] S. R. Finch, Beam detection constant, *first volume*, pp. 515–519.
[4] S. R. Finch and J. A. Shonder, Lost at sea, math.OC/0411518.
[5] S. R. Finch and L.-Y. Zhu, Searching for a shoreline, math.OC/0501123.
[6] S. R. Finch, The logarithmic spiral conjecture, math.OC/0501133.
[7] R. A. Baeza-Yates, J. C. Culberson and G. J. E. Rawlins, Searching in the plane, *Inform. and Comput.* 106 (1993) 234–252; MR1241311 (94h:90019).
[8] R. A. Baeza-Yates, J. C. Culberson and G. J. E. Rawlins, Searching with uncertainty (extended abstract), *First Scandinavian Workshop on Algorithm Theory*, Proc. 1988 Halmstad conf., ed. R. Karlsson and A. Lingas, Lecture Notes in Comput. Sci. 318, Springer-Verlag, 1988, pp. 176–189; MR1019368.
[9] R. A. Baeza-Yates, J. C. Culberson and G. J. E. Rawlins, Searching with uncertainty (full report), Research Report CS–87–68, Dept. of Computer Sci., Univ. of Waterloo, 1987.
[10] S. R. Finch, The Golden mean, *first volume*, pp. 5–11.
[11] R. Fletcher, Investigation of the growth of squares, unpublished note (2004).

5.6 Minkowski–Siegel Mass Constants

Let X denote either a vector space over the real numbers $\mathbb{R}$ or a module over the integers $\mathbb{Z}$. A symmetric positive definite bilinear form f on X is an **inner product** if, for any linear form g on X, there exists a unique $x \in X$ such that $g(y) = f(x, y)$ for all $y \in X$. This nondegeneracy condition is superfluous when X is a finite-dimensional vector space [1, 2]. The pair (X, f) is called an **inner product space** or an **inner product module**, respectively. Two pairs (X, f) and (X', f') are **isomorphic** if there is a bijective linear transformation $h: X \to X'$ satisfying

$$f'(h(x), h(y)) = f(x, y)$$

for all $x, y \in X$. In the special case $(X, f) = (X', f')$, the map h is called an **automorphism**. The set of all such maps forms a group $\mathrm{Aut}(X, f)$ under composition, known as the **automorphism group**. We will need the cardinality $|\mathrm{Aut}(X, f)|$ later when defining the Minkowski–Siegel mass constants.

If X is an n-dimensional $\mathbb{R}$-vector space, then (X, f) is isomorphic to $(\mathbb{R}^n, \cdot)$, that is, Euclidean n-space equipped with the standard dot product [1, 2]. If X is a free $\mathbb{Z}$-module of rank n, then for $n \le 7$, (X, f) is isomorphic to $(\mathbb{Z}^n, \cdot)$. What happens for $n \ge 8$? A partial answer to this question will occupy us for the remainder of this essay [3–5].

An inner product module (X, f) over $\mathbb{Z}$ is said to be **even** if $f(x, x) \equiv 0 \bmod 2$ for all $x \in X$. Otherwise it is said to be **odd**. The phrases **Type II** and **Type I** (for even and odd, respectively) are also often used.

There is a more geometric approach to this subject. A **lattice** in $\mathbb{R}^n$ is a subset $\Lambda \subseteq \mathbb{R}^n$ such that, for some basis $\{e_1, e_2, \ldots, e_n\}$ of $\mathbb{R}^n$, we have

$$\Lambda = \left\{ \sum_{j=1}^{n} i_j e_j : i_j \in \mathbb{Z},\ 1 \le j \le n \right\}.$$

The **volume** of Λ is the Lebesgue measure of the fundamental parallelepiped

$$\left\{ \sum_{j=1}^{n} r_j e_j : r_j \in \mathbb{R},\ 0 \le r_j \le 1,\ 1 \le j \le n \right\}$$

or, equivalently, the absolute value of the determinant of the matrix whose rows are the vectors $e_1, e_2, \ldots, e_n$. The lattice Λ is **unimodular** or **self-dual** if the dot product $e_k \cdot e_l \in \mathbb{Z}$ for all $1 \le k, l \le n$ and if the volume of Λ is 1. It can be proved that the unimodular lattices in $\mathbb{R}^n$ are "representations" of the free inner product $\mathbb{Z}$-modules of rank n. All properties of one language carry over to the other. For example, a unimodular lattice Λ is even if $v \cdot v \equiv 0 \bmod 2$ for all $v \in \Lambda$; otherwise it is odd [3–5].

We merely mention that this subject is closely connected with the construction of dense sphere packings in $\mathbb{R}^n$ [6].

5.6.1 *Classification of Inner Product Modules*

Classifying pairs (X, f) up to isomorphism, where X is a free $\mathbb{Z}$-module of rank n and f is an inner product, becomes interesting starting at $n = 8$. There is a unique odd module when $n = 8$, namely $(\mathbb{Z}^8, \cdot)$. There is also a unique even module $\mathbb{E}_8$ when $n = 8$; it is easiest to describe $\mathbb{E}_8$ as a certain unimodular lattice in $\mathbb{R}^8$. Let $\{e_1, e_2, \ldots, e_n\}$ denote the standard orthonormal basis of $\mathbb{R}^8$ and define the following to be the basis for $\mathbb{E}_8$:

$$\begin{array}{cccc} 2e_1, & e_2 - e_1, & e_3 - e_2, & e_4 - e_3, \\ e_5 - e_4, & e_6 - e_5, & e_7 - e_6, & \frac{1}{2}(e_1 + e_2 + \cdots + e_8). \end{array}$$

In words, $\mathbb{E}_8$ consists of all points in $\mathbb{R}^8$ whose coordinates are either all integers or all halves of odd integers, and sum to an even integer. We emphasize that $\mathbb{E}_8 \approx \mathbb{Z}^8$ as modules, but $\mathbb{E}_8 \not\approx \mathbb{Z}^8$ as inner product modules [5, 7, 8].

Table 5.1 gives the number a_n of odd unimodular lattices and the number b_n of even unimodular lattices, where $8 \le n \le 25$ [7, 9]. For $9 \le n \le 11$, the only odd unimodular lattices are $\mathbb{Z}^n$ and $\mathbb{E}_8 \oplus \mathbb{Z}^{n-8}$. When $n = 12$, a third odd lattice $\mathbb{D}_{12}^{+}$ appears. Even unimodular lattices exist if and only if $n \equiv 0 \bmod 8$. When $n = 16$, the only even lattices are $\mathbb{E}_8 \oplus \mathbb{E}_8$ and another new case $\mathbb{D}_{16}^{+}$. The famous Leech lattice $\mathbb{L}$ corresponds to $n = 24$ and is the unique even case with the property that $v \cdot v \ge 4$ for every nonzero $v \in \mathbb{L}$. It is known [10] that $a_{26} \ge 2307$, $a_{27} \ge 14179$, $a_{28} \ge 327972$ and $b_{32} \ge 1162109024 > 10^9$; no one expects a complete classification of even lattices for $n = 32$ to be achieved in the near future.

Table 5.1 *Number of free inner product $\mathbb{Z}$-modules of rank n (Type I and Type II)*

n	a_n	b_n	n	a_n	b_n	n	a_n	b_n
8	1	1	14	4		20	28	
9	2		15	5		21	40	
10	2		16	6	2	22	68	
11	2		17	9		23	117	
12	3		18	13		24	273	24
13	3		19	16		25	665	

Against such difficult enumerations, it is surprising that exact formulas, valid for all n, involving the reciprocal sum of automorphism group orders should exist. Let $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, ... denote the Bernoulli numbers and $E_0 = 1$, $E_1 = 0$, $E_2 = -1$, $E_3 = 0$, $E_4 = 5$, ... denote the Euler numbers. The following sum is taken over all nonisomorphic odd unimodular lattices in $\mathbb{R}^n$ [11, 12]:

$$M_n = \sum_{\Lambda} \frac{1}{|\operatorname{Aut}(\Lambda)|}$$

$$= \begin{cases} \dfrac{1}{2} & \text{if } n = 1, \\ \dfrac{(1-2^{-k})(1+2^{1-k})}{k! \cdot 2} \left| B_k \cdot B_2 B_4 \cdots B_{2k-2} \right| & \text{if } n = 2k \equiv 0 \bmod 8, \\ \dfrac{2^k + 1}{k! \cdot 2^{2k+1}} \left| B_2 B_4 \cdots B_{2k} \right| & \text{if } 1 < n = 2k+1 \equiv \pm 1 \bmod 8, \\ \dfrac{1}{(k-1)! \cdot 2^{2k+1}} \left| E_{k-1} \cdot B_2 B_4 \cdots B_{2k-2} \right| & \text{if } n = 2k \equiv \pm 2 \bmod 8, \\ \dfrac{2^k - 1}{k! \cdot 2^{2k+1}} \left| B_2 B_4 \cdots B_{2k} \right| & \text{if } n = 2k+1 \equiv \pm 3 \bmod 8, \\ \dfrac{(1-2^{-k})(1-2^{1-k})}{k! \cdot 2} \left| B_k \cdot B_2 B_4 \cdots B_{2k-2} \right| & \text{if } n = 2k \equiv 4 \bmod 8. \end{cases}$$

In particular, $M_n = 1/(n!\,2^n)$ for $1 \le n \le 8$. Milnor & Husemoller [3] provided a corresponding asymptotic formula:

$$M_n \sim C \cdot \left(\frac{n}{2\pi e \sqrt{e}} \right)^{n^2/4} \left(\frac{8\pi e}{n} \right)^{n/4} \left(\frac{1}{n} \right)^{1/24} = C \cdot F(n)$$

as $n \to \infty$, where $C \approx 0.705$, but no precise expression for C was given. We will return to this issue momentarily. For nonisomorphic even unimodular lattices in $\mathbb{R}^n$, the analogous sum is [4, 11, 12]

$$N_n = \sum_{\Lambda} \frac{1}{|\operatorname{Aut}(\Lambda)|} = \frac{|B_k|}{2k} \prod_{l=1}^{k-1} \frac{|B_{2l}|}{4l}$$

Table 5.2 *Type I Minkowski–Siegel mass constants M_n*

n	Exact	Decimal
8	$\frac{1}{10321920}$	$9.688... \times 10^{-8}$
9	$\frac{17}{2786918400}$	$6.099... \times 10^{-9}$
10	$\frac{1}{2229534720}$	$4.485... \times 10^{-10}$
$\vdots$		
16	$\frac{505121}{12340763622899712000}$	$4.093... \times 10^{-14}$
17	$\frac{642332179}{18881368343036559360000}$	$3.401... \times 10^{-14}$
18	$\frac{692319119}{15105094674429247488000}$	$4.583... \times 10^{-14}$
$\vdots$		
24	$\frac{70187670795628001881586 2361}{2107902862678499821906978406 4000000}$	$3.329... \times 10^{-8}$
25	$\frac{84715059480304651623612272842147}{30465396080006318014267329085440000000}$	$2.780... \times 10^{-6}$
26	$\frac{14616335635894388876188472684851927}{31871491283698917307233513504768000000}$	$4.586... \times 10^{-4}$
27	$\frac{18943527517721468674304869954629232 65007}{12429881600642577749821070266859520000000}$	$1.524... \times 10^{-1}$
28	$\frac{103450603774276940430378894822230239502 03227}{99439052805140621998568562134876160000000}$	$1.040... \times 10^{2}$
29	$\frac{42850098239595906821156287393561695866872 20752159}{2883732531349078037958488301911408640000 0000}$	$1.485... \times 10^{5}$

if $n = 2k \equiv 0 \bmod 8$, with asymptotics

$$N_n \sim D \cdot \left(\frac{n}{2\pi e\sqrt{e}}\right)^{n^2/4} \left(\frac{\pi e}{2n}\right)^{n/4} \left(\frac{1}{n}\right)^{1/24}.$$

Such **mass formulas** are useful in verifying that a candidate listing of isomorphism classes of unimodular lattices, for a prescribed genus, is correct. See Tables 5.2 and 5.3.

Although M_n and N_n are initially very small and are decreasing, they eventually reverse direction and increase dramatically. The asymptotics for M_n are similar to the asymptotics for the product of even-subscripted Bernoulli numbers:

$$\prod_{j=1}^{n} |B_{2j}| \sim C \cdot n! \cdot 2^{n+1} \cdot F(2n+1).$$

Table 5.3 *Type II Minkowski–Siegel mass constants N_n*

n	Exact	Decimal
8	$\dfrac{1}{696729600}$	$1.435... \times 10^{-9}$
16	$\dfrac{691}{277667181515243520000}$	$2.488... \times 10^{-18}$
24	$\dfrac{1027637932586061520960267}{129477933340026851560636148613120000000}$	$7.936... \times 10^{-15}$
32	$\dfrac{4890529010450384254108570593011950899382291953107314413193123}{121325280941552041649762780685623131486814208000000000}$	$4.030... \times 10^{7}$

It turns out that the constants C and D can be written as [13]

$$C = 2^{-5/4} e^{1/24} A^{-1/2} Z = 0.7048648734...,$$
$$D = 4C = 2.8194594938... = 2^{1/24} \cdot 2.7391949550...$$

where $A = \exp(\frac{1}{12} - \zeta'(-1)) = 1.2824271291...$ is the Glaisher–Kinkelin constant [14] and

$$Z = \prod_{i=1}^{\infty} \zeta(2i) = 1.8210174514...$$

bears resemblance to certain constants arising when enumerating abelian groups [15].

5.6.2 *Products and Sums of Factorials*

While determining C and D, Kellner [13] examined the product of factorials

$$\prod_{\nu=1}^{n} (k\,\nu)! \sim W_k \left(\frac{k\,n}{e\sqrt{e}}\right)^{\frac{k\,n^2}{2}} \left(\frac{k\,n}{e}\right)^{\frac{k\,n}{2}} \left(\frac{2\pi k\,n}{e}\right)^{\frac{n}{2}} n^{\frac{1}{4}+\frac{k}{12}+\frac{1}{12k}}$$

and computed the constants $F_k = (2\pi)^{-1/4} A^{-k} W_k$ to be

$$F_k = k^{\frac{5}{12k}} (2\pi)^{\frac{k}{4}-\frac{1}{2}+\frac{1}{2k}} e^{\frac{1}{12k}} A^{-k-\frac{1}{k}} \prod_{m=2}^{k-1} \Gamma(\tfrac{m}{k})^{-\frac{m-1}{k}}$$

for each positive integer k. In particular, we have

$$F_1 = (2\pi)^{1/4} e^{1/12} A^{-2} = 1.0463350667...,$$

$$F_2 = 2^{5/24} (2\pi)^{1/4} e^{1/24} A^{-5/2} = 1.0239374116...,$$

$$F_3 = 3^{5/36} (2\pi)^{5/12} e^{1/36} A^{-10/3} \Gamma(2/3)^{-1/3} = 1.0160405370...,$$

$$F_4 = 2^{1/3} (2\pi)^{1/2} e^{1/48} A^{-17/4} \Gamma(3/4)^{-1/2} = 1.0120458980....$$

The case $k=1$ corresponds to the asymptotics of the well-known Barnes G-function [14]. As k grows without bound, we also have

$$\lim_{k\to\infty} F_k = 1, \quad \lim_{k\to\infty} F_k^k = e^{\gamma/12},$$

where γ is the Euler–Mascheroni constant, and

$$\lim_{l\to\infty} l^{-\gamma/12} \prod_{k=1}^{l} F_k = 1.0246068826....$$

An exact evaluation of the final limit remains open. By way of contrast, the sum of factorials

$$\sum_{\nu=1}^{n} (k\nu)! \sim (kn)! \sim (2\pi k n)^{1/2} \left(\frac{kn}{e}\right)^{kn}$$

does not involve any new constants.

[1] E. Hewitt and K. Stromberg, *Real and Abstract Analysis*, Springer-Verlag, 1965, pp. 242, 248; MR0188387 (32 #5826).
[2] A. E. Taylor, *General Theory of Functions and Integration*, Blaisdell, 1965, pp. 154, 161–162; MR0178100 (31 #2358).
[3] J. Milnor and D. Husemoller, *Symmetric Bilinear Forms*, Springer-Verlag, 1973, pp. 1–3, 15–18, 26–29, 49–51; MR0506372 (58 #22129).
[4] J.-P. Serre, *A Course in Arithmetic*, Springer-Verlag, 1973, pp. 48–55; MR0344216 (49 #8956).
[5] W. Ebeling, *Lattices and Codes*, 2nd ed., Friedr. Vieweg & Sohn, 2002, pp. 1–5; MR1280458 (95c:11084) and MR1938666 (2003i:11093).
[6] S. R. Finch, Hermite's constants, *first volume*, pp. 506–508.
[7] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, 1988, pp. 41–55, 408–413; MR0920369 (89a:11067), MR1194619 (93h:11069) and MR1662447 (2000b:11077).
[8] G. Nebe and N. J. A. Sloane, Unimodular lattices, together with a table of the best such lattices, Catalogue of Lattices.
[9] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A054909 and A054911.
[10] O. D. King, A mass formula for unimodular lattices with no roots, *Math. Comp.* 72 (2003) 839–863; math.NT/0012231; MR1954971 (2003m:11101).
[11] J. H. Conway and N. J. A. Sloane, The unimodular lattices of dimension up to 23 and the Minkowski-Siegel mass constants, *European J. Combin.* 3 (1982) 219–231; MR0679207 (84b:10045).
[12] V. E. Voskresenskii, Calculation of local volumes in the Siegel-Tamagawa formula (in Russian), *Mat. Sbornik* 180 (1989) 443–455, 559; Engl. transl. in *Math. USSR Sbornik* 66 (1990) 447–460; MR0997894 (90h:11055).
[13] B. Kellner, On asymptotic constants related to products of Bernoulli numbers and factorials, *Integers* 9 (2009) 83–106; math.NT/0604505; MR2506140 (2010d:11027).
[14] S. R. Finch, Glaisher-Kinkelin constant, *first volume*, pp. 135–145.
[15] S. R. Finch, Abelian group enumeration constants, *first volume*, pp. 273–276.

5.7 Slicing Problem

Before stating the slicing problem, let us examine a related problem with known solution. Let K be a compact convex set in $\mathbb{R}^n$ with nonempty interior. Assume that the n-dimensional volume of K is unity, that is, $\mathrm{vol}_n(K) = 1$. The **centroid** of K is $\mu = \mathrm{E}(X)$, where X is a uniformly distributed random point in K. Let H be any $(n-1)$-dimensional plane passing through μ with corresponding half-spaces H^+ and H^-. Grünbaum [1], Hammer [2] and Mityagin [3] independently proved that

$$\min\left\{\mathrm{vol}_n(K \cap H^+), \mathrm{vol}_n(K \cap H^-)\right\} \geq \left(\frac{n}{n+1}\right)^n \to \frac{1}{e} = 0.3678794411...$$

and, further, the bound $(n/(n+1))^n$ is best possible. In words, at least a proportion $1/e$ of the convex set volume lies on each side of any planar cut through the centroid. Applications of this result appear in [4–9]. Grünbaum wrote that it would be interesting to find the analog of this result when substituting $(n-1)$-dimensional surface area for n-dimensional volume, and added that this problem is unsolved even for $n = 2$.

We now give the slicing problem (which is perhaps related to Grünbaum's foreshadowing but likewise unsolved). Let K be as before, with the additional condition that K is **isotropic**:

$$\Sigma = \mathrm{Cov}(X) = \mathrm{E}\left((X-\mu)(X-\mu)^T\right) = \sigma^2 I,$$

where X is a uniformly distributed random point in K and I is the $n \times n$ identity matrix. This latter condition is equivalent to saying that, for every vector $v \in \mathbb{R}^n$,

$$\mathrm{E}\left([v^T(X-\mu)]^2\right) = \sigma^2 |v|^2.$$

The vector μ is often called the **barycenter** of K, the matrix Σ the **inertia matrix** and the scalar σ the **isotropic constant**. Let H be as before. It is conjectured that such an H exists so that

$$\mathrm{vol}_{n-1}(K \cap H) > c$$

for some constant $c > 0$ independent of n and K. (Note that Grünbaum's theorem was true for all H and involved vol_n, not vol_{n-1}.) Bourgain [10, 11] and Paouris [12] proved that

$$\mathrm{vol}_{n-1}(K \cap H) > \frac{b}{n^{1/4}\ln(n)}$$

for some constant $b > 0$. The slicing problem is also known as the hyperplane conjecture; an equivalent formulation is that the isotropic constant $\sigma < a$ for some constant $a < \infty$ independent of n and K.

We mention an obvious converse of Grünbaum's theorem: There exists H for which

$$\mathrm{vol}_n(K \cap H^+) = \mathrm{vol}_n(K \cap H^-) = \frac{1}{2}.$$

A converse of the slicing problem can be expressed as [13, 14]

$$\sigma \geq \frac{1}{\sqrt{n+2}}\omega_n^{-1/n} \to \frac{1}{\sqrt{2\pi e}} = 0.2419707245... = (4.1327313541...)^{-1}$$

where $\omega_n = \pi^{n/2}\Gamma(n/2+1)^{-1}$ is the volume of the unit n-ball. The requirement that K be isotropic is not too restrictive, since every convex set has a linear image which is isotropic. See [15, 16] for applications and [17–20] for recent progress.

[1] B. Grünbaum, Partitions of mass-distributions and of convex bodies by hyperplanes, *Pacific J. Math.* 10 (1960) 1257–1261; MR0124818 (23 #A2128).

[2] P. Hammer, Volumes cut from convex bodies by hyperplanes, unpublished manuscript (1960).

[3] B. S. Mityagin, Two inequalities for volumes of convex bodies (in Russian), *Mat. Zametki* 5 (1969) 99–106; Engl. transl. in *Math. Notes* 5 (1968) 61–65; MR0243429 (39 #4751).

[4] J. Kahn and N. Linial, Balancing extensions via Brunn-Minkowski, *Combinatorica* 11 (1991) 363–368; MR1137768 (93e:52017).

[5] G. Brightwell, Balanced pairs in partial orders, *Discrete Math.* 201 (1999) 25–52; MR1687862 (2000d:06004).

[6] P. Gritzmann and V. Klee, Mathematical programming and convex geometry, in *Handbook of Convex Geometry*, v. A, ed. P. M. Gruber and J. M. Wills, North-Holland, 1993, pp. 627–674; MR1242992 (95a:52013).

[7] W. D. Smith and N. C. Wormald, Geometric separator theorems and applications, abbrev. version in *Proc. 39th Symp. on Foundations of Computer Science (FOCS),* Palo Alto, IEEE, 1998, pp. 232–243; full version at Monash Univ.

[8] S. G. Bobkov, On concentration of distributions of random weighted sums, *Annals of Probab*. 31 (2003) 195–215; MR1959791 (2004a:60049).

[9] D. Bertsimas and S. Vempala, Solving convex programs by random walks, *J. ACM* 51 (2004) 540–556; *Proc. 34th ACM Symp. on the Theory of Computing (STOC)*, Montreal, ACM, 2002, pp. 109–115.

[10] J. Bourgain, On the distribution of polynomials on high-dimensional convex sets, *Geometric Aspects of Functional Analysis (1989–90)*, ed. J. Lindenstrauss and V. D. Milman, Lect. Notes in Math. 1469, Springer-Verlag, 1991, pp. 127–137; MR1122617 (92j:52007).

[11] S. Dar, Remarks on Bourgain's problem on slicing of convex bodies, *Geometric Aspects of Functional Analysis (1992–1994)*, Oper. Theory Adv. Appl., v. 77, Birkhäuser, 1995, pp. 61–66; MR1353449 (96j:46010).

[12] G. Paouris, On the isotropic constant of non-symmetric convex bodies, *Geometric Aspects of Functional Analysis (1996–2000)*, ed. by V. D. Milman and G. Schechtman, Lect. Notes in Math. 1745, Springer-Verlag, 2000, pp. 239–243; MR1796722 (2002b:52013).

[13] V. D. Milman and A. Pajor, Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space, *Geometric Aspects of Functional Analysis (1987–88)*, ed. J. Lindenstrauss and V. D. Milman, Lect. Notes in Math. 1376, Springer-Verlag, 1989, pp. 64–104; MR1008717 (90g:52003).

[14] S. Dar, On the isotropic constant of non-symmetric convex bodies, *Israel J. Math.* 97 (1997) 151–156; MR1441244 (98a:52012).

[15] R. J. Gardner, *Geometric Tomography*, Cambridge Univ. Press, 1995, pp. 298–299, 302, 308, 342–344; MR1356221 (96j:52006).
[16] R. Kannan, L. Lovász, and M. Simonovits, Random walks and an $O(n^5)$ volume algorithm for convex bodies, *Random Structures Algorithms* 11 (1997) 1–50; MR1608200 (99h:68078).
[17] K. Ball, A remark on the slicing problem, *Geometric Aspects of Functional Analysis (1996–2000)*, ed. by V. D. Milman and G. Schechtman, Lect. Notes in Math. 1745, Springer-Verlag, 2000, pp. 21–26; MR1796710 (2002b:52009).
[18] J. Bourgain, B. Klartag and V. Milman, A reduction of the slicing problem to finite volume ratio bodies, *C. R. Math. Acad. Sci. Paris* 336 (2003) 331–334; MR1976313 (2004d:52005).
[19] J. Bourgain, B. Klartag and V. Milman, Symmetrization and isotropic constants of convex bodies, *Geometric Aspects of Functional Analysis*, ed. V. D. Milman and G. Schechtman, Lect. Notes in Math. 1850, Springer-Verlag, 2004, pp. 101–115; MR2087154 (2005i:52005).
[20] B. Klartag, An isomorphic version of the slicing problem, *J. Funct. Anal.* 218 (2005) 372–394; MR2108116 (2005i:52007).

5.8 Constant of Theodorus

In the complex plane, consider the recursive sequence

$$z_n = \left(1 + \frac{i}{\sqrt{n}}\right) z_{n-1}, \quad n \geq 1,$$

with starting point $z_0 = 1$. The points z_{n-1} and z_n determine a right triangle relative to the origin 0, with legs 1 and $\sqrt{n}$. Clearly the polar coordinates (r_n, θ_n) of z_n are given by

$$r_n = \sqrt{n+1}, \quad \theta_n = \begin{cases} \sum\limits_{j=0}^{n-1} \arctan\left(\frac{1}{\sqrt{j+1}}\right) & \text{if } n \geq 1, \\ 0 & \text{if } n = 0. \end{cases}$$

A closed-form expression for z_n is

$$z_n = \prod_{k=1}^{n} \left(1 + \frac{i}{\sqrt{k}}\right) \quad n \geq 1,$$

and determines what is called the **discrete spiral of Theodorus**.

Davis [1, 2] and Heuvers, Moak & Boursaw [3] independently constructed the continuous analog of this spiral. A parametric representation is [1, 2]

$$\begin{aligned} f(t) &= \prod_{k=1}^{\infty} \frac{1 + \frac{i}{\sqrt{k}}}{1 + \frac{i}{\sqrt{k+t}}}, \quad -1 < t < \infty, \\ &= \sqrt{1+t}\, \exp\left(i \sum_{k=1}^{\infty} \left(\arctan\left(\sqrt{k+t}\right) - \arctan\left(\sqrt{k}\right)\right)\right) \end{aligned}$$

and a polar representation is [3]

$$\theta(r) = \sum_{j=0}^{\infty}\left(\arctan\left(\frac{1}{\sqrt{j+1}}\right) - \arctan\left(\frac{1}{\sqrt{j+r^2}}\right)\right), \quad r>0.$$

Gronau [2] proved that $f(t)$ is the unique solution of the functional equation

$$f(t) = \left(1+\frac{i}{\sqrt{t}}\right)f(t-1), \quad f(0)=1, \quad 0<t<\infty$$

such that $|f(t)|$ is increasing and $\arg(f(t))$ is both increasing and continuous.

Among many possible questions, Davis [1] asked: What is the slope of the spiral at the point 1? Clearly

$$\left.\frac{dy}{dx}\right|_{(x,y)=(1,0)} = \left.\frac{d\theta}{dr}\right|_{(r,\theta)=(1,0)} = \sum_{k=1}^{\infty}\frac{1}{k^{3/2}+k^{1/2}},$$

which Gautschi [4] evaluated to be 1.8600250792.... This is called the **constant of Theodorus**.

Also, what can be said about the growth of θ_n as $n\to\infty$? For convenience, given a real number ξ, let $\{\xi\}=\xi \bmod 1$ denote the fractional part of ξ. Hlawka [5] proved that

$$\theta_n = 2\sqrt{n+1}+K+\frac{1}{6\sqrt{n+1}}+O\left(n^{-3/2}\right),$$

where the **square root spiral constant** $K=K_0-1-3\pi/8=-2.1577829966...$ and

$$K_0=\frac{1}{8}\int_2^{\infty}\{x\}(1-\{x\})(3x-2)\frac{1}{x^2(x-1)^{3/2}}dx = 0.0203142484....$$

The numerical estimate of K was obtained by Grünberg [6], correcting an apparent error in [5].

In more detail, the series

$$K=\frac{\pi}{4}+\sum_{m=0}^{\infty}(-1)^m\frac{\zeta\left(m+\frac{1}{2}\right)-1}{2m+1}$$

converges quickly [7], as does

$$K'=\sum_{m=0}^{\infty}\frac{\zeta\left(m+\frac{1}{2}\right)-1}{2m+1} = -1.8265078108...$$

(associated with the growth of θ_n', obtained by replacing arctan by arctanh in the definition of θ_n). Similarly, the series

$$\sum_{k=1}^{\infty}\frac{1}{(k+1)\sqrt{k}} = \frac{1}{2}+\sum_{m=1}^{\infty}(-1)^{m+1}\left\{\zeta\left(m+\frac{1}{2}\right)-1\right\}$$

converges quickly (to Theodorus' constant), as does

$$\sum_{k=2}^{\infty} \frac{1}{(k-1)\sqrt{k}} = \sum_{m=1}^{\infty} \left\{ \zeta\left(m+\frac{1}{2}\right) - 1 \right\} = 2.1840094702...$$

(obtained by simply replacing $+$ by $-$ and removing the term for $k = 1$).

[1] P. J. Davis, *Spirals: From Theodorus to Chaos*, A K Peters, 1993, pp. 7–11, 37–43, 220; MR1224447 (94e:00001).
[2] D. Gronau, The spiral of Theodorus, *Amer. Math. Monthly* 111 (2004) 230–237; MR2042127 (2005c:51022).
[3] K. J. Heuvers, D. S. Moak and B. Boursaw, The functional equation of the square root spiral, *Functional Equations and Inequalities*, ed. T. M. Rassias, Kluwer, 2000, pp. 111–117; MR1792078 (2001k:39033).
[4] W. Gautschi, The spiral of Theodorus, special functions, and numerical analysis, in Davis, *op cit.*, pp. 67–87; MR1224447 (94e:00001).
[5] E. Hlawka, Gleichverteilung und Quadratwurzelschnecke, *Monatsh. Math.* 89 (1980) 19–44; Engl. transl. in Davis, *op cit.*, pp. 157–167; MR0566292 (81h:10069).
[6] D. Grünberg, Euler–Maclaurin summation examples, unpublished note (2005).
[7] D. Brink, The spiral of Theodorus and sums of zeta-values at the half-integers, *Amer. Math. Monthly* 119 (2012) 779–786; MR2990936.

5.9 Nearest-Neighbor Graphs

Consider a set P of n points that are independently and uniformly distributed in the d-dimensional unit cube. Let $p \in P$. There exists almost-surely $q \in P$ such that $q \neq p$ and $|p-q| < |p-r|$ for all $r \in P$, $r \neq p$, $r \neq q$. The point q is called the **nearest neighbor** of p and we write $p \prec q$. Note that $p \prec q$ does not imply $q \prec p$. Draw an edge connecting p and q if and only if $p \prec q$; the resulting graph of n vertices and $\leq n$ edges is called the **nearest-neighbor graph** G on P.

What is the probability, $\alpha(d)$, given $p \in P$, that $p \prec q$ implies $q \prec p$? Such a pair is **isolated** from the rest of G, in the sense that the only edge touching p or q is the edge that connects p and q. We have [1–15]

$$\alpha(1) = \frac{2}{3}, \quad \alpha(2) = \frac{6\pi}{8\pi + 3\sqrt{3}} = 0.6215048968..., \quad \alpha(3) = \frac{16}{27}$$

and, more generally [9],

$$\alpha(d) = \begin{cases} \left[\dfrac{3}{2} + \dfrac{1}{2} \displaystyle\sum_{k=1}^{\ell} \frac{1 \cdot 3 \cdots (2k-1)}{2 \cdot 4 \cdots (2k)} \left(\frac{3}{4}\right)^k \right]^{-1} & \text{if } d = 2\ell + 1, \\ \left[\dfrac{4}{3} + \dfrac{\sqrt{3}}{2\pi} \left(1 + \displaystyle\sum_{k=1}^{\ell-1} \frac{2 \cdot 4 \cdots (2k)}{3 \cdot 5 \cdots (2k+1)} \left(\frac{3}{4}\right)^k \right) \right]^{-1} & \text{if } d = 2\ell. \end{cases}$$

Here is a variation of the preceding. Draw an edge connecting p and q if and only if $q \prec p$; the resulting graph of n vertices and $\leq n$ edges is called the **nearest-neighbor anti-graph** H on P. What is the probability, $\beta(d)$, that $p \in P$ is isolated from the rest of H? That is, what proportion of points in P are not nearest neighbors of any other points? We have [16–21]

$$\beta(1)=\frac{1}{4},\quad \beta(2)\approx 0.28,\quad \beta(3)\approx 0.30$$

but the latter two estimates are only simulation-based. To further understand $\beta(2)$ will occupy us for the remainder of this essay.

Define constants $C(0,d)=1$ and

$$C(k,d)=\int\limits_{\Omega(k,d)}\exp\left[-\operatorname{Vol}\left(\bigcup_{j=1}^{k}S(x_j)\right)\right]dx_1dx_2\ldots dx_k$$

for $k\geq 1$, where $S(x_j)$ is the ball in $\mathbb{R}^d$ of radius $|x_j|$, centered at x_j, and

$$\Omega(k,d)=\left\{(x_1,x_2,\ldots,x_k)\in\mathbb{R}^{dk}:|x_i|\leq|x_i-x_j|\text{ for all }1\leq i\neq j\leq k\right\}.$$

It is known that [19, 22–25]

$$\beta(2)=\sum_{k=0}^{6}\frac{(-1)^k}{k!}C(k,2),\quad \beta(3)=\sum_{k=0}^{12}\frac{(-1)^k}{k!}C(k,3)$$

and clearly $C(1,d)=1$, $C(2,1)=1/2$. The upper limits of summation are the *kissing numbers* in $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. A proof that 24 is the kissing number in $\mathbb{R}^4$ was given only recently [26, 27]. Also, $C(6,2)=0$ since $\Omega(6,2)$ is of measure zero.

Henze [24, 25] showed that

$$C(2,d)=\frac{2^{d+1}\pi^{d-1}}{\Gamma(d-1)}\int\limits_0^{\infty}\int\limits_0^{\xi}\int\limits_{\theta_0}^{\pi}\xi^{d-1}\eta^{d-1}\sin(\theta)^{d-2}F_d(\xi,\eta)\,d\theta\,d\eta\,d\xi$$

where

$$\theta_0=\arccos\left(\frac{\eta}{2\xi}\right),$$

$$F_d(\xi,\eta)=\exp\left[-f_d(\xi,\gamma)-f_d(\eta,\delta)\right],$$

$$\gamma=\frac{\xi(\xi-\eta\cos(\theta))}{\sqrt{\xi^2+\eta^2-2\xi\,\eta\cos(\theta)}},\quad \delta=\frac{\eta(\eta-\xi\cos(\theta))}{\sqrt{\xi^2+\eta^2-2\xi\,\eta\cos(\theta)}},$$

$$f_d(x,y)=\frac{\pi^{d/2}x^d}{2\Gamma(d/2+1)}\left[1+I\left(\frac{y^2}{x^2},\frac{1}{2},\frac{d+1}{2}\right)\right]$$

and I is the regularized beta function

$$I(z,a,b)=\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int\limits_0^{z}w^{a-1}(1-w)^{b-1}dw.$$

Table 5.4 *Old and new calculations of constants*

k	Tao & Wu estimate of $C(k,2)/k!$	Current estimate of $C(k,2)/k!$
2	0.3163335...	0.316585...
3	0.0329390...	0.033056...
4	0.0006575...	still open
5	0.0000010...	still open

(In [24], the definitions of γ and δ were mistakenly reversed; also, the expression within square brackets for $f_d(x,y)$ was unclear.) We obtain

$$C(2,2) = 0.63317... = 2(0.316585...), \quad C(2,3) = 0.70888....$$

Tao & Wu [19] independently showed that

$$\begin{aligned} C(2,2) \;&=\; \pi \int\limits_{\pi/2}^{\pi} \int\limits_{0}^{\infty} \frac{\tau}{\left(g(\tau,\theta) + \tau^2 h(\tau,\theta)\right)^2} d\tau \, d\theta \\ &\quad + \pi \int\limits_{\pi/3}^{\pi/2} \int\limits_{2\cos(\theta)}^{1/(2\cos(\theta))} \frac{\tau}{\left(g(\tau,\theta) + \tau^2 h(\tau,\theta)\right)^2} d\tau \, d\theta \end{aligned}$$

where

$$g(\tau,\theta) = \pi - \varphi + \tfrac{1}{2}\sin(2\varphi), \quad h(\tau,\theta) = \pi - \psi + \tfrac{1}{2}\sin(2\psi),$$

$$\varphi = \arcsin\left(\frac{\tau \sin(\theta)}{\sqrt{1+\tau^2 - 2\tau\cos(\theta)}}\right), \quad \psi = \arcsin\left(\frac{\sin(\theta)}{\sqrt{1+\tau^2 - 2\tau\cos(\theta)}}\right).$$

(Several underlying details in [19] are clarified in [28].) Even more elaborate integral formulas apply for $C(3,2)$, $C(4,2)$, $C(5,2)$. Given the discrepancy between our estimate of $C(2,2)$ and their estimate (see Table 5.4), it seems doubtful that their approximation $\beta(2) = 0.284051...$ is entirely correct.

A discrete version of the latter problem appears in [29–32]. Let all the vertices of the lattice $\mathbb{Z}^d$ be initially occupied by particles which can annihilate one-by-one their $2d$ nearest neighbors. More precisely, for each unit-length edge $\{u, v\}$ of the lattice, there is a Uniform $[0,1]$ random variable $T_{\{u,v\}}$ representing the time of an attack along the edge. If vertices u, v are both occupied immediately prior to time $T_{\{u,v\}}$, then at time $T_{\{u,v\}}$ either vertex u or vertex v (each with probability $1/2$) becomes vacant (that is, one particle annihilates the other). If u, v are not both occupied at time $T_{\{u,v\}}$, then there is no change. Once a vertex becomes vacant, it remains vacant permanently. The variables $T_{\{u,v\}}$, considered over all unit-length edges $\{u, v\}$, are independent. By time 1, no two surviving particles can be adjacent. When $d = 1$, the probability that a given vertex remains occupied is $1/e = 0.3678794411\ldots$. When $d = 2$, this probability is known to be in the interval $(0.227, 0.306)$ and is approximately 0.25 via simulation. Greater accuracy is desired.

[1] P. J. Clark and F. C. Evans, On some aspects of spatial pattern in biological populations, *Science* 121 (1955) 397–398.
[2] P. J. Clark, Grouping in spatial distributions, *Science* 123 (1956) 373–374.
[3] M. F. Dacey, Proportion of reflexive n^{th} order neighbors in spatial distribution, *Geographical Analysis* 1 (1969) 385–388.
[4] G. F. Schwarz and A. Tversky, On the reciprocity of proximity relations, *J. Math. Psych.* 22 (1980) 157–175; MR0609119 (82f:92050).
[5] T. F. Cox, Reflexive nearest neighbours, *Biometrics* 37 (1981) 367–369; MR0673043 (83k:62122).
[6] D. P. Shine and J. Herbert, Birds on a wire, *J. Recreational Math.* 11 (1978-79) 227–228; 15 (1982-83) 232.
[7] S. Morris, Competition winners, *Omni* v. 2 (1980) n. 9, p. 108.
[8] C. Kluepfel, Birds on a wire, cows in the field, and stars in the heavens, *J. Recreational Math.* 13 (1980-81) 241–245.
[9] D. K. Pickard, Isolated nearest neighbors, *J. Appl. Probab.* 19 (1982) 444–449; MR0649985 (83g:60063).
[10] M. F. Schilling, Mutual and shared neighbor probabilities: finite- and infinite-dimensional results, *Adv. Appl. Probab.* 18 (1986) 388–405; MR0840100 (87k:60041).
[11] N. Henze, On the probability that a random point is the j^{th} nearest neighbour to its own k^{th} nearest neighbour, *J. Appl. Probab.* 23 (1986) 221–226; MR0826925 (87k:60133).
[12] D. Eppstein, M. S. Paterson and F. F. Yao, On nearest-neighbor graphs, *Discrete Comput. Geom.* 17 (1997) 263–282; MR1432064 (98d:05121).
[13] D. P. Shine and M. P. Cohen, Spread the news, *J. Recreational Math.* 36 (2007) 277–278.
[14] C. A. S. Tercariol, F. de Moura Kiipper and A. Souto Martinez, An analytical calculation of neighbourhood order probabilities for high dimensional Poissonian processes and mean field models, *J. Phys. A* 40 (2007) 1981–1989; cond-mat/0609210; MR2316309 (2008a:82035).
[15] P. J. Campbell and B. Atwood, The farmer problem, *UMAP Journal* 33 (2012) 313–331.
[16] F. D. K. Roberts, Nearest neighbours in a Poisson ensemble, *Biometrika* 56 (1969) 401–406.
[17] R. Abilock and M. Goldberg, *N* riflemen, *Amer. Math. Monthly* 75 (1968) 1009; 89 (1982) 274–275.
[18] S. Morris, Rifle puzzle, *Omni* v. 8 (1986) n. 4, p. 113; v. 9 (1987) n. 7, p. 141.
[19] R. Tao and F. Y. Wu, The vicious neighbour problem, *J. Phys. A* 20 (1987) L299–L306; MR0888078 (88d:82021).
[20] E. G. Enns, P. F. Ehlers and T. Misi, A cluster problem as defined by nearest neighbours, *Canad. J. Statist.* 27 (1999) 843–851; MR1767151 (2001b:60017).
[21] S. Portnoy, A squirtgun battle, *J. Recreational Math.* 37 (2008) 39–45.
[22] C. M. Newman, Y. Rinott and A. Tversky, Nearest neighbors and Voronoi regions in certain point processes, *Adv. Appl. Probab.* 15 (1983) 726–751; MR0721703 (85m:60023).

[23] C. M. Newman and Y. Rinott, Nearest neighbors and Voronoi volumes in high-dimensional point processes with various distance functions, *Adv. Appl. Probab.* 17 (1985) 794–809; MR0809431 (87d:60048).
[24] N. Henze, Über die Anzahl von Zufallspunkten mit typ-gleichem nächsten Nachbarn und einen multivariaten Zwei-Stichproben-Test, *Metrika* 31 (1984) 259–273; MR0773815 (86i:62075).
[25] N. Henze, On the fraction of random points with specified nearest-neighbour interrelations and degree of attraction, *Adv. Appl. Probab.* 19 (1987) 873–895; MR0914597 (89c:60063).
[26] O. R. Musin, The kissing number in four dimensions, *Annals of Math.* 168 (2008) 1–32; math.MG/0309430; MR2415397.
[27] F. Pfender and G. M. Ziegler, Kissing numbers, sphere packings, and some unexpected proofs, *Notices Amer. Math. Soc.* 51 (2004) 873–883; MR2145821 (2006a:52015).
[28] S. R. Finch, Union of n disks: remote centers, common origin, arXiv:1511.04968.
[29] M. O'Hely and A. Sudbury, The annihilating process, *J. Appl. Probab.* 38 (2001) 223–231; MR1816125 (2001m:60226).
[30] A. Sudbury, Inclusion-exclusion methods for treating annihilating and deposition processes, *J. Appl. Probab.* 39 (2002) 466–478; MR1928883 (2003k:60266).
[31] A. Sudbury, The annihilating process on random trees and the square lattice, *J. Appl. Probab.* 41 (2004) 816–831; MR2074826 (2005g:82090).
[32] M. D. Penrose and A. Sudbury, Exact and approximate results for deposition and annihilation processes on graphs, *Annals Appl. Probab.* 15 (2005) 853–889; MR2114992 (2005k:60307).

5.10 Random Triangles

Let X_1, X_2, X_3, Y_1, Y_2, Y_3, Z_1, Z_2, Z_3 be independent normally distributed random variables with mean 0 and variance 1. The points (X_1, Y_1), (X_2, Y_2), (X_3, Y_3) constitute the vertices of a triangle in Euclidean 2-space (the plane); the points (X_1, Y_1, Z_1), (X_2, Y_2, Z_2), (X_3, Y_3, Z_3) constitute the vertices of a triangle in Euclidean 3-space. A number of parameters (for example, sides, angles, perimeter and area) describe the triangle, but the corresponding probability density functions are not well-known. We attempt to remedy this situation in this essay. Perhaps the most famous results for random Gaussian triangles are the following [1, 2]:

$$\mathrm{P}(\text{a Gaussian triangle in 2-space is obtuse}) = 3/4 = 0.75,$$

$$\mathrm{P}(\text{a Gaussian triangle in 3-space is obtuse}) = 1 - 3\sqrt{3}/(4\pi) = 0.5865033284...$$

which translate into statements about the maximum angle exceeding $\pi/2$. Consider, however, an arbitrary angle α in a triangle. What is its first moment $\mathrm{E}(\alpha)$? This turns out to be trivial. What is its second moment $\mathrm{E}(\alpha^2)$? This is more difficult, even in 2 dimensions, and the answer is apparently new. Our essay, the first in a series, arises in an effort to expand upon [3].

5.10.1 *Sides*

Let a, b, c denote the sides of a random Gaussian triangle. The trivariate density $f(x, y, z)$ for a, b, c in 2 dimensions is [4]

$$\begin{cases} \dfrac{2}{3\pi}\dfrac{x\,y\,z}{\sqrt{(x+y+z)(-x+y+z)(x-y+z)(x+y-z)}}\exp\left(-\dfrac{1}{6}\left(x^2+y^2+z^2\right)\right) \\ \qquad \text{if } |x-y|<z<x+y, \\ 0 \qquad \text{otherwise} \end{cases}$$

and we shall give an elementary proof of this later. The condition $|x-y|<z<x+y$ is equivalent to $|x-z|<y<x+z$ and to $|y-z|<x<y+z$ via the Law of Cosines. As a consequence, the univariate density for a corresponds to Rayleigh's distribution:

$$\frac{x}{2}\exp\left(-\frac{x^2}{4}\right), \quad x>0$$

and [5, 6]

$$\mathrm{E}(a)=\sqrt{\pi}=1.7724538509..., \quad \mathrm{E}(a^2)=4,$$

$$\mathrm{E}(a\,b)=4E\left(\frac{1}{2}\right)-\frac{3}{2}K\left(\frac{1}{2}\right)=3.3412233051...$$

where

$$K(\xi)=\int_0^{\pi/2}\frac{1}{\sqrt{1-\xi^2\sin(\theta)^2}}\,d\theta=\int_0^1\frac{1}{\sqrt{(1-t^2)(1-\xi^2t^2)}}\,dt,$$

$$E(\xi)=\int_0^{\pi/2}\sqrt{1-\xi^2\sin(\theta)^2}\,d\theta=\int_0^1\sqrt{\frac{1-\xi^2t^2}{1-t^2}}\,dt$$

are complete elliptic integrals of the first and second kind [7]. The cross-correlation coefficient

$$\rho(a,b)=\frac{\mathrm{Cov}(a,b)}{\sqrt{\mathrm{Var}(a)\,\mathrm{Var}(b)}}=\frac{\mathrm{E}(a\,b)-\pi}{4-\pi}=0.2325593465...$$

is quite small, indicating weak positive dependency. Interestingly, $\rho(a^2,b^2)=1/4=0.25$ since a^2, b^2 are quadratic forms in normal variables and classical theory applies [8, 9].

The trivariate density for a, b, c in 3 dimensions is [4]

$$\begin{cases} \dfrac{\sqrt{3}}{9\pi}\,x\,y\,z\exp\left(-\dfrac{1}{6}\left(x^2+y^2+z^2\right)\right) & \text{if } |x-y|<z<x+y, \\ 0 & \text{otherwise} \end{cases}$$

which is surprisingly simpler than the corresponding result in 2 dimensions. As a consequence, the univariate density for a corresponds to the

Maxwell–Boltzmann distribution:

$$\frac{x^2}{2\sqrt{\pi}} \exp\left(-\frac{x^2}{4}\right), \quad x > 0$$

and

$$\mathrm{E}(a) = \frac{4}{\sqrt{\pi}} = 2.2567583341..., \quad \mathrm{E}(a^2) = 6,$$

$$\mathrm{E}(a\,b) = 2 + \frac{6\sqrt{3}}{\pi} = 5.3079733725...,$$

$$\rho(a,b) = \frac{-8 + 3\sqrt{3} + \pi}{-8 + 3\pi} = 0.2370510252..., \quad \rho(a^2,b^2) = \frac{1}{4} = 0.25.$$

5.10.2 *Perimeter and Area*

For perimeter $a + b + c$, the density is a double integral:

$$\int_0^x \int_0^{x-v} f(x-u-v,u,v)\,du\,dv, \quad x > 0$$

which we have not attempted to evaluate. Thus only moments are given. In 2 dimensions,

$$\mathrm{E}(\text{perimeter}) = 3\sqrt{\pi} = 5.3173615527...,$$

$$\begin{aligned}\mathrm{E}(\text{perimeter}^2) &= \mathrm{E}((a+b+c)^2)\\ &= 3\mathrm{E}(a^2) + 6\mathrm{E}(a\,b)\\ &= 12 + 24E\left(\frac{1}{2}\right) - 9K\left(\frac{1}{2}\right) = 32.0473398308...\end{aligned}$$

and in 3 dimensions,

$$\mathrm{E}(\text{perimeter}) = \frac{12}{\sqrt{\pi}} = 6.7702750025...,$$

$$\mathrm{E}(\text{perimeter}^2) = 30 + \frac{36\sqrt{3}}{\pi} = 49.8478402351....$$

More can be said about area $(1/4)\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}$. In 2 dimensions, area can be proved to be exponentially distributed, with density [10]

$$\frac{2}{\sqrt{3}} \exp\left(-\frac{2}{\sqrt{3}}x\right), \quad x > 0.$$

The formula given in [11] is unfortunately incorrect. In particular,

$$\mathrm{E}(\text{area}) = \frac{\sqrt{3}}{2} = 0.8660254037..., \quad \mathrm{E}(\text{area}^2) = \frac{3}{2} = 1.5.$$

A proposed density in [12] for 3 dimensional area also seems to be wrong. We find instead

$$\mathrm{E}(\text{area}) = \sqrt{3} = 1.7320508075..., \quad \mathrm{E}(\text{area}^2) = \frac{9}{2} = 4.5.$$

5.10.3 *Angles*

Let α, β, γ denote the angles of a random Gaussian triangle. Of course, $\alpha+\beta+\gamma=\pi$, thus γ can be eliminated from consideration. The bivariate density $\varphi(x,y)$ for α, β in 2 dimensions is [13]

$$\begin{cases} \dfrac{6}{\pi}\dfrac{\sin(x)\sin(y)\sin(x+y)}{\left(\sin(x)^2+\sin(y)^2+\sin(x+y)^2\right)^2} & \text{if } 0<x<\pi,\ 0<y<\pi \text{ and } x+y<\pi, \\ 0 & \text{otherwise} \end{cases}$$

and we shall confirm this later. The univariate density for α was first discovered by W. S. Kendall [14], via a fairly geometric argument, but has never appeared explicitly in the open literature (the closest was [15]; see also [16]). Starting from the bivariate density, we obtain the univariate density via

$$\begin{aligned} &\frac{6}{\pi}\int_0^{\pi-x} \frac{\sin(x)\sin(y)\sin(x+y)}{\left(\sin(x)^2+\sin(y)^2+\sin(x+y)^2\right)^2}dy \\ &= \tfrac{6}{\pi}\int_0^{\pi-x} \tfrac{\cos(x)\sin(x)}{2(4-\cos(x)^2)(\sin(x)^2+\sin(y)^2+\sin(x+y)^2)}dy \\ &\quad+\tfrac{6}{\pi}\int_0^{\pi-x}\left(\tfrac{\sin(x)\sin(y)\sin(x+y)}{(\sin(x)^2+\sin(y)^2+\sin(x+y)^2)^2}-\tfrac{\cos(x)\sin(x)}{2(4-\cos(x)^2)(\sin(x)^2+\sin(y)^2+\sin(x+y)^2)}\right)dy \\ &= \frac{3}{\pi}\frac{\cos(x)}{\left(4-\cos(x)^2\right)^{3/2}}\left(\frac{\pi}{2}+\arcsin\left(\frac{\cos(x)}{2}\right)\right)+\frac{3}{\pi}\frac{1}{4-\cos(x)^2}. \end{aligned}$$

Call this latter expression $g(x)$. Now, since $3\mathrm{E}(\alpha)=\mathrm{E}(\alpha+\beta+\gamma)=\pi$, we have $\mathrm{E}(\alpha)=\pi/3$. It is harder to show that

$$\mathrm{E}(\alpha^2) = \frac{7}{36}\pi^2 - \frac{1}{2}\,\mathrm{Li}_2\left(\frac{1}{4}\right) = 1.7852634251...$$

where

$$\mathrm{Li}_2(\xi) = \sum_{k=1}^{\infty}\frac{\xi^k}{k^2} = -\int_0^{\xi}\frac{\ln(1-t)}{t}dt$$

is the dilogarithm function [18]. Also, since $3\,\mathrm{Var}(\alpha)+6\,\mathrm{Cov}(\alpha,\beta)=\mathrm{Var}(\alpha+\beta+\gamma)=0$, we have $\rho(\alpha,\beta)=-1/2$; therefore

$$\mathrm{E}(\alpha\,\beta) = \frac{5}{72}\pi^2 + \frac{1}{4}\,\mathrm{Li}_2\left(\frac{1}{4}\right) = 0.7523023542....$$

Finally,

$$G(x) = \int_0^x g(\xi)\, d\xi = \frac{1}{\pi} \frac{\sin(x)}{(4 - \cos(x)^2)^{1/2}} \left(\frac{\pi}{2} + \arcsin\left(\frac{\cos(x)}{2}\right)\right) + \frac{1}{\pi} x,$$

which implies that $\mathrm{P}(\alpha > \pi/2) = 1 - G(\pi/2) = 1/4 = 0.25$, where α is arbitrary. This is equal to $(1/3)\mathrm{P}(\max(\alpha, \beta, \gamma) > \pi/2)$ because a triangle can have at most one obtuse angle.

The bivariate density for α, β in 3 dimensions is new, as far as we know:

$$\begin{cases} \dfrac{24\sqrt{3}}{\pi} \dfrac{\sin(x)^2 \sin(y)^2 \sin(x+y)^2}{(\sin(x)^2 + \sin(y)^2 + \sin(x+y)^2)^3} & \text{if } 0 < x < \pi, 0 < y < \pi \\ & \text{and } 0 < x + y < \pi, \\ 0 & \text{otherwise.} \end{cases}$$

The univariate density for α is obtained similarly:

$$\begin{aligned} &\frac{24\sqrt{3}}{\pi} \int_0^{\pi - x} \frac{\sin(x)^2 \sin(y)^2 \sin(x+y)^2}{(\sin(x)^2 + \sin(y)^2 + \sin(x+y)^2)^3} dy \\ &= \frac{24\sqrt{3}}{\pi} \int_0^{\pi - x} \frac{\left(2 + \cos(x)^2\right) \sin(x)^2}{4\left(4 - \cos(x)^2\right)^2 (\sin(x)^2 + \sin(y)^2 + \sin(x+y)^2)} dy \\ &\quad + \frac{24\sqrt{3}}{\pi} \int_0^{\pi - x} \left(\begin{array}{l} \dfrac{\sin(x)^2 \sin(y)^2 \sin(x+y)^2}{(\sin(x)^2 + \sin(y)^2 + \sin(x+y)^2)^3} \\ \quad - \dfrac{\left(2 + \cos(x)^2\right) \sin(x)^2}{4\left(4 - \cos(x)^2\right)^2 (\sin(x)^2 + \sin(y)^2 + \sin(x+y)^2)} \end{array} \right) dy \\ &= \frac{6\sqrt{3}}{\pi} \frac{\left(2 + \cos(x)^2\right) \sin(x)}{\left(4 - \cos(x)^2\right)^{5/2}} \left(\frac{\pi}{2} + \arcsin\left(\frac{\cos(x)}{2}\right)\right) + \frac{9\sqrt{3}}{\pi} \frac{\cos(x)\sin(x)}{\left(4 - \cos(x)^2\right)^2}. \end{aligned}$$

Call this latter expression $h(x)$. We observe that $h(x) = -\sqrt{3} g'(x)$ and wonder about the meaning of such a connection. As before, $\mathrm{E}(\alpha) = \pi/3$. It follows that

$$\mathrm{E}(\alpha^2) = \frac{\pi}{3}\left(\pi - \sqrt{3}\right) = 1.4760687694...,$$

$$\mathrm{E}(\alpha\,\beta) = \frac{\pi}{6}\sqrt{3} = 0.9068996821....$$

Finally,

$$\mathrm{P}(\alpha > \pi/2) = 1 + \sqrt{3}\,(g(\pi/2) - g(0)) = \frac{1}{3} - \frac{\sqrt{3}}{4\pi} = 0.1955011094...$$

where α is arbitrary. This again is equal to $(1/3)\mathrm{P}(\max(\alpha, \beta, \gamma) > \pi/2)$.

5.10.4 *Order Statistics*

We will, for brevity's sake, study only maximum/minimum angles in two dimensions and only maximum/minimum sides in three dimensions. Define $\tilde{g}(x)$ to be

$$\begin{aligned}&\frac{3}{\pi}\frac{\cos(x)}{(4-\cos(x)^2)^{3/2}}\left(\frac{\pi}{2}-\arcsin\left(\frac{\cos(x)}{2}\right)-2\arctan\left(\frac{3\cos(x)}{\sqrt{4-\cos(x)^2}}\right)\right)\\&\quad+\frac{3}{\pi}\frac{1-4\cos(x)^2}{(4-\cos(x)^2)(1+2\cos(x)^2)}\end{aligned}$$

which is positive for $\pi/3 < x < \pi/2$. Given $\alpha > 0$, $\beta > 0$, $\alpha + \beta < \pi$, the angle α is maximum if $\alpha > \beta$ and $\alpha > \pi - \alpha - \beta$. Hence the density for the maximum angle is

$$\begin{cases} 3\displaystyle\int_{\pi-2x}^{x} \varphi(x,y)dy & \text{if } \pi/3 < x < \pi/2, \\ 3\displaystyle\int_{0}^{\pi-x} \varphi(x,y)dy & \text{if } \pi/2 < x < \pi \end{cases} = \begin{cases} 3\tilde{g}(x) & \text{if } \pi/3 < x < \pi/2, \\ 3g(x) & \text{if } \pi/2 < x < \pi \end{cases}$$

after breaking up the integral of $\varphi(x,y)$ precisely as outlined earlier. This density again was first discovered by Kendall [14] using a different approach. Incidentally, the identity

$$\arcsin\left(\frac{\cos(x)}{2}\right) = \arctan\left(\frac{\cos(x)}{\sqrt{4-\cos(x)^2}}\right)$$

might lead to a more natural expression for $\tilde{g}(x)$. The value $3g(\pi) = 3/\pi - 1/\sqrt{3} = 0.3775793893...$ is called the shape constant (or first collinearity constant) for planar Gaussian triangles [15, 16].

The function $\tilde{g}(x)$ is negative for $0 < x < \pi/3$ and the angle α is minimum if $\alpha < \beta$ and $\alpha < \pi - \alpha - \beta$. By a similar breakup, the density for the minimum angle is

$$3\int_{x}^{\pi-2x} \varphi(x,y)dy = -3\tilde{g}(x).$$

Moments for these distributions remain open.

Advancing up to three dimensions, the density for the maximum side is [4]

$$\frac{3x}{2\sqrt{\pi}}\left[2\sqrt{\frac{3}{\pi}}\left(e^{-x^2/2}-e^{-x^2/3}\right)+x\,e^{-x^2/4}\operatorname{erf}\left(\frac{\sqrt{3}x}{6}\right)\right]$$

for $x>0$, and the density for the minimum side is

$$\frac{3x}{2\sqrt{\pi}}\left[2\sqrt{\frac{3}{\pi}}\left(e^{-x^2/2}-e^{-x^2}\right)+x\,e^{-x^2/4}\operatorname{erfc}\left(\frac{\sqrt{3}x}{2}\right)\right]$$

where erf, erfc are the error and complementary error functions [17].

5.10.5 *Trivariate Details*

Our proof closely follows [19]. Consider sides a, b of a random Gaussian triangle in the plane. Using

$$a^2=(X_2-X_1)^2+(Y_2-Y_1)^2,\quad b^2=(X_3-X_1)^2+(Y_3-Y_1)^2$$

we picture vectors $\vec{a},\vec{b}$ emanating from (X_1,Y_1) to (X_2,Y_2), (X_3,Y_3), respectively. Define $0<\theta_a<2\pi$ to be the angle between vector $\vec{a}$ and the x-axis; define $0<\theta_b<2\pi$ likewise. Observe that

$$(u_a,u_b)=\left(\frac{X_2-X_1}{\sqrt{2}},\frac{X_3-X_1}{\sqrt{2}}\right),\quad (v_a,v_b)=\left(\frac{Y_2-Y_1}{\sqrt{2}},\frac{Y_3-Y_1}{\sqrt{2}}\right)$$

are independent random vectors satisfying

$$(u_a,u_b),(v_a,v_b)\sim N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1&\frac{1}{2}\\\frac{1}{2}&1\end{pmatrix}\right).$$

Define $s_a=a^2/4$ and $s_b=b^2/4$. Then

$$u_a=\sqrt{2s_a}\cos(\theta_a),\quad v_a=\sqrt{2s_a}\sin(\theta_a),\quad u_b=\sqrt{2s_b}\cos(\theta_b),\quad v_b=\sqrt{2s_b}\sin(\theta_b)$$

and conversely

$$s_a=\frac{u_a^2+v_a^2}{2},\quad s_b=\frac{u_b^2+v_b^2}{2},\quad \tan(\theta_a)=\frac{v_a}{u_a},\quad \tan(\theta_b)=\frac{v_b}{u_b}.$$

The Jacobian matrix of the transformation $(u_a,v_a,u_b,v_b)\mapsto(s_a,s_b,\theta_a,\theta_b)$ is

$$J=\begin{pmatrix}u_a&v_a&0&0\\0&0&u_b&v_b\\-\dfrac{v_a}{u_a^2+v_a^2}&\dfrac{u_a}{u_a^2+v_a^2}&0&0\\0&0&-\dfrac{v_b}{u_b^2+v_b^2}&\dfrac{u_b}{u_b^2+v_b^2}\end{pmatrix}.$$

For example,

$$\sec(\theta_a)^2\frac{\partial\theta_a}{\partial u_a}=\frac{\partial}{\partial u_a}\tan(\theta_a)=\frac{\partial}{\partial u_a}\frac{v_a}{u_a}=-\frac{v_a}{u_a^2}$$

implies that

$$\frac{\partial\theta_a}{\partial u_a}=-\cos(\theta_a)^2\frac{v_a}{u_a^2}=-\frac{u_a^2}{2s_a}\frac{v_a}{u_a^2}=-\frac{v_a}{u_a^2+v_a^2}.$$

As another example,

$$\sec(\theta_a)^2 \frac{\partial\theta_a}{\partial v_a} = \frac{\partial}{\partial v_a}\tan(\theta_a) = \frac{\partial}{\partial v_a}\frac{v_a}{u_a} = \frac{1}{u_a}$$

implies that

$$\frac{\partial\theta_a}{\partial v_a} = \cos(\theta_a)^2\frac{1}{u_a} = \frac{u_a^2}{2s_a}\frac{1}{u_a} = \frac{u_a}{u_a^2+v_a^2}.$$

Since the absolute determinant $|J|=1$, changing variables from (u_a, v_a, u_b, v_b) to $(s_a, s_b, \theta_a, \theta_b)$ is easily performed. The density for (u_a, u_b) gives rise to

$$\begin{aligned}
&\frac{1}{2\pi\sqrt{1-(\frac{1}{2})^2}}\exp\left[-\frac{1}{2\left(1-(\frac{1}{2})^2\right)}\left(u_a^2 - 2(\tfrac{1}{2})u_a u_b + u_b^2\right)\right] \\
&\quad = \frac{1}{\sqrt{3}\pi}\exp\left[-\frac{2}{3}\left(u_a^2 - u_a u_b + u_b^2\right)\right] \\
&\quad = \frac{1}{\sqrt{3}\pi}\exp\left[-\frac{2}{3}\left(2s_a\cos(\theta_a)^2 - \sqrt{2s_a}\sqrt{2s_b}\cos(\theta_a)\cos(\theta_b) + 2s_b\cos(\theta_b)^2\right)\right] \\
&\quad = \frac{1}{\sqrt{3}\pi}\exp\left[-\frac{4}{3}\left(s_a\cos(\theta_a)^2 - \sqrt{s_a s_b}\cos(\theta_a)\cos(\theta_b) + s_b\cos(\theta_b)^2\right)\right]
\end{aligned}$$

and the density for (v_a, v_b) likewise gives rise to

$$\begin{aligned}
&\frac{1}{\sqrt{3}\pi}\exp\left[-\frac{2}{3}\left(v_a^2 - v_a v_b + v_b^2\right)\right] \\
&\quad = \frac{1}{\sqrt{3}\pi}\exp\left[-\frac{4}{3}\left(s_a\sin(\theta_a)^2 - \sqrt{s_a s_b}\sin(\theta_a)\sin(\theta_b) + s_b\sin(\theta_b)^2\right)\right].
\end{aligned}$$

By independence, the density for (u_a, u_b, v_a, v_b) is

$$\frac{1}{3\pi^2}\exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\theta_a-\theta_b) + s_b\right)\right]$$

where $0<\theta_a<2\pi$, $0<\theta_b<2\pi$.

We move toward integrating out θ_a. Let $\omega = \theta_a - \theta_b$. The Jacobian matrix of the transformation $(s_a, s_b, \theta_a, \theta_b) \mapsto (s_a, s_b, \omega, \theta_a)$ is

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

and $|K|=1$, hence the density for $(s_a, s_b, \omega, \theta_a)$ is

$$\frac{1}{3\pi^2}\exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\omega) + s_b\right)\right]$$

where $-2\pi < \omega < 2\pi$ plus an additional condition. If $\omega < 0$, then $\theta_b < 2\pi$ forces $\theta_a < 2\pi + \theta_a - \theta_b = 2\pi + \omega$, thus

$$\frac{1}{3\pi^2} \int\limits_0^{2\pi+\omega} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\omega) + s_b\right)\right] d\theta_a$$
$$= \frac{2\pi+\omega}{3\pi^2} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\omega) + s_b\right)\right];$$

if $\omega > 0$, then $\theta_b > 0$ forces $\theta_a > \theta_a - \theta_b = \omega$, thus

$$\frac{1}{3\pi^2} \int\limits_\omega^{2\pi} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\omega) + s_b\right)\right] d\theta_a$$
$$= \frac{2\pi-\omega}{3\pi^2} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\omega) + s_b\right)\right].$$

In either case, the coefficient numerator is $2\pi - |\omega|$ and the density is symmetric in ω about 0. Let $\gamma = |\omega|$, then we multiply by 2 to obtain the density for (s_a, s_b, γ):

$$\frac{2(2\pi-\gamma)}{3\pi^2} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\gamma) + s_b\right)\right]$$

where $0 < \gamma < 2\pi$. Adding contributions at γ and $2\pi - \gamma$ yields

$$\frac{4}{3\pi} \exp\left[-\frac{4}{3}\left(s_a - \sqrt{s_a s_b}\cos(\gamma) + s_b\right)\right]$$

for $0 < \gamma < \pi$, which works since $2(2\pi - \gamma) + 2\gamma = 4\pi$ and $\cos(\gamma) = \cos(2\pi - \gamma)$. Replacing s_a, s_b by $a^2/4$, $b^2/4$ yields

$$\frac{4}{3\pi} \exp\left[-\frac{4}{3}\left(\frac{a^2}{4} - \frac{a\,b}{4}\cos(\gamma) + \frac{b^2}{4}\right)\right] \frac{a}{2}\frac{b}{2}$$
$$= \frac{1}{3\pi} a\,b \exp\left[-\frac{1}{3}\left(a^2 - a\,b\cos(\gamma) + b^2\right)\right].$$

This is already useful for computing moments of area:

$$\mathrm{E}\left(\left(\frac{1}{2} a\,b\sin(\gamma)\right)^m\right) = m!\left(\frac{\sqrt{3}}{2}\right)^m$$

for all positive integers m. Also, an initial step in calculating $\mathrm{E}(a\,b)$ is to evaluate

$$\frac{1}{3\pi} \int\limits_0^{\pi} a^2 b^2 \exp\left[-\frac{1}{3}\left(a^2 - a\,b\cos(\gamma) + b^2\right)\right] d\gamma$$

$$= \frac{a^2 b^2}{3} \exp\left[-\frac{1}{3}\left(a^2 + b^2\right)\right] I_0\left(\frac{a\,b}{3}\right)$$

where $I_0(z)$ is the modified Bessel function of the first kind [20]. Note that the angle γ is adjacent to sides a, b and opposite to side c, as is traditional. The analogous density for (α, β, c) appears in the next section.

We now bring c into the trivariate density, removing γ. Differentiating the Law of Cosines

$$c^2 = a^2 - 2\,a\,b\cos(\gamma) + b^2$$

with respect to γ, it is clear that

$$\begin{aligned} 2\,c\,dc &= 2\,a\,b\sin(\gamma)\,d\gamma \\ &= \sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}\,d\gamma \end{aligned}$$

by a formula for area, and hence the density becomes

$$\begin{aligned} &\frac{1}{3\pi}\,a\,b\exp\left[-\frac{1}{3}\left(a^2 - a\,b\cos(\gamma) + b^2\right)\right] da\,db\,d\gamma \\ &= \frac{1}{3\pi}\,a\,b\exp\left[-\frac{1}{6}\left(a^2 + b^2 + (a^2 - 2\,a\,b\cos(\gamma) + b^2)\right)\right] da\,db\,d\gamma \\ &= \frac{2}{3\pi}\,\frac{a\,b\,c}{\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}} \\ &\quad \times \exp\left[-\frac{1}{6}\left(a^2 + b^2 + c^2\right)\right] da\,db\,dc \end{aligned}$$

assuming $0 < \gamma < \pi$, that is, $a^2 - 2\,a\,b + b^2 < c^2 < a^2 + 2\,a\,b + b^2$. The required condition $|a-b| < c < a+b$ does not change upon permutation of sides a, b, c.

Note that the variables s_a, s_b are each exponentially distributed with mean 1, with cross-correlation $1/4$. A closed-form expression for the density for (s_a, s_b) is not possible [19], but an infinite series representation [21]

$$\sum_{n=0}^{\infty} \frac{1}{4^n}\Phi(-n,1,s_a)\Phi(-n,1,s_b)\exp(-(s_a+s_b))$$

is valid, where $\Phi(u,v,w)$ is the confluent hypergeometric function of the first kind [22]. In this special case,

$$\Phi(-n,1,t) = \sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k}{k!}t^k.$$

Proving the series representation makes use of

$$s_a = \left(\frac{u_a}{\sqrt{2}}\right)^2 + \left(\frac{v_a}{\sqrt{2}}\right)^2, \quad s_b = \left(\frac{u_b}{\sqrt{2}}\right)^2 + \left(\frac{v_b}{\sqrt{2}}\right)^2$$

and the fact that $u_a/\sqrt{2}$, $u_b/\sqrt{2}$ are jointly normal with mean 0, variance $1/2$ and cross-correlation $1/2$. Other multivariate generalizations of the exponential distribution are found in [23].

For the (a,b,c)-density associated with random Gaussian triangles in 3-space, we refer to [4].

5.10.6 *Bivariate Details*

Let $\Delta=(a+b+c)(-a+b+c)(a-b+c)(a+b-c)$ for convenience. The transformation $(a,b,c)\mapsto(\alpha,\beta,c)$ is prescribed via

$$\cos(\alpha)=\frac{-a^2+b^2+c^2}{2\,b\,c},\quad \cos(\beta)=\frac{-b^2+a^2+c^2}{2\,a\,c}.$$

We have, for example,

$$-\sin(\alpha)\frac{\partial\alpha}{\partial a}=-\frac{a}{b\,c},\quad -\sin(\alpha)\frac{\partial\alpha}{\partial b}=\frac{a^2+b^2-c^2}{2\,b^2c},\quad -\sin(\alpha)\frac{\partial\alpha}{\partial c}=\frac{a^2-b^2+c^2}{2\,b\,c^2}$$

hence

$$\frac{\partial\alpha}{\partial a}=\frac{a}{b\,c}\frac{1}{\sin(\alpha)}=\frac{a}{b\,c}\frac{1}{\sqrt{1-\cos(\alpha)^2}}=\frac{a}{b\,c}\frac{2b\,c}{\sqrt{\Delta}}=\frac{2a}{\sqrt{\Delta}},$$

$$\frac{\partial\alpha}{\partial b}=-\frac{a^2+b^2-c^2}{2b^2c}\frac{1}{\sin(\alpha)}=-\frac{a^2+b^2-c^2}{2b^2c}\frac{2\,b\,c}{\sqrt{\Delta}}=-\frac{a^2+b^2-c^2}{b\sqrt{\Delta}},$$

$$\frac{\partial\alpha}{\partial c}=-\frac{a^2-b^2+c^2}{2b\,c^2}\frac{1}{\sin(\alpha)}=-\frac{a^2-b^2+c^2}{2b\,c^2}\frac{2\,b\,c}{\sqrt{\Delta}}=-\frac{a^2-b^2+c^2}{c\sqrt{\Delta}}.$$

The corresponding Jacobian matrix is

$$L=\begin{pmatrix} \dfrac{2a}{\sqrt{\Delta}} & \dfrac{-a^2-b^2+c^2}{b\sqrt{\Delta}} & \dfrac{-a^2+b^2-c^2}{c\sqrt{\Delta}} \\ \dfrac{-a^2-b^2+c^2}{a\sqrt{\Delta}} & \dfrac{2b}{\sqrt{\Delta}} & \dfrac{a^2-b^2-c^2}{c\sqrt{\Delta}} \\ 0 & 0 & 1 \end{pmatrix}$$

and $|L|=1/(a\,b)$. By the Law of Sines,

$$a=c\frac{\sin(\alpha)}{\sin(\gamma)}=c\frac{\sin(\alpha)}{\sin(\alpha+\beta)},\quad b=c\frac{\sin(\beta)}{\sin(\gamma)}=c\frac{\sin(\beta)}{\sin(\alpha+\beta)}$$

and, under the change of variables,

$$\sqrt{\Delta}=2c^2\frac{\sin(\alpha)\sin(\beta)}{\sin(\alpha+\beta)}.$$

The density for (α,β,c) in two dimensions is

$$\begin{aligned}
&\frac{2}{3\pi}\frac{a^2b^2c}{\sqrt{\Delta}}\exp\left[-\frac{1}{6}\left(a^2+b^2+c^2\right)\right]\\
&=\frac{2c^5}{3\pi}\frac{\sin(\alpha)^2\sin(\beta)^2}{\sin(\alpha+\beta)^4\sqrt{\Delta}}\exp\left[-\frac{c^2}{6\sin(\alpha+\beta)^2}\left(\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2\right)\right]\\
&=\frac{c^3}{3\pi}\frac{\sin(\alpha)\sin(\beta)}{\sin(\alpha+\beta)^3}\exp\left[-\frac{c^2}{6}\frac{\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2}{\sin(\alpha+\beta)^2}\right].
\end{aligned}$$

Integrating out c is facilitated by observing that

$$\int_0^\infty c^3 \exp\left(-\frac{c^2}{6} r\right) dc = \frac{18}{r^2}$$

for $r > 0$, therefore the density for (α, β) in two dimensions is

$$\begin{aligned}&\frac{18}{3\pi}\frac{\sin(\alpha)\sin(\beta)}{\sin(\alpha+\beta)^3}\left(\frac{\sin(\alpha+\beta)^2}{\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2}\right)^2\\ &\quad=\frac{6}{\pi}\frac{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}{(\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2)^2}.\end{aligned}$$

Similarly, the density for (α, β, c) in three dimensions is

$$\begin{aligned}&\frac{\sqrt{3}}{9\pi}a^2b^2c\exp\left(-\frac{1}{6}\left(a^2+b^2+c^2\right)\right)\\ &\quad=\frac{\sqrt{3}c^5}{9\pi}\frac{\sin(\alpha)^2\sin(\beta)^2}{\sin(\alpha+\beta)^4}\exp\left[-\frac{c^2}{6}\frac{\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2}{\sin(\alpha+\beta)^2}\right].\end{aligned}$$

Here we observe that

$$\int_0^\infty c^5 \exp\left(-\frac{c^2}{6} r\right) dc = \frac{216}{r^3}$$

for $r > 0$, therefore the density for (α, β) in three dimensions is

$$\begin{aligned}&\frac{216\sqrt{3}}{9\pi}\frac{\sin(\alpha)^2\sin(\beta)^2}{\sin(\alpha+\beta)^4}\left(\frac{\sin(\alpha+\beta)^2}{\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2}\right)^3\\ &\quad=\frac{24\sqrt{3}}{\pi}\frac{\sin(\alpha)^2\sin(\beta)^2\sin(\alpha+\beta)^2}{(\sin(\alpha)^2+\sin(\beta)^2+\sin(\alpha+\beta)^2)^3}.\end{aligned}$$

We turn attention to the most interesting of our moment evaluations, that concerning $\mathrm{E}(\alpha^2)$. First,

$$\int_0^\pi \arcsin\left(\frac{\cos(x)}{2}\right) dx = 0$$

because $\arcsin(\cos(\pi - x)/2) = \arcsin(-\cos(x)/2) = -\arcsin(\cos(x)/2)$. Consequently

$$\begin{aligned}\int_0^\pi \frac{x\sin(x)}{\sqrt{4-\cos(x)^2}}dx &= -x\arcsin\left(\frac{\cos(x)}{2}\right)\Bigg|_0^\pi + \int_0^\pi \arcsin\left(\frac{\cos(x)}{2}\right)dx\\ &= \frac{\pi^2}{6}\end{aligned}$$

using integration by parts. Second,

$$\begin{aligned}
&\int_0^\pi \left(\arcsin\left(\frac{\cos(x)}{2}\right)\right)^2 dx \\
&= \frac{1}{4}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{1}{16^{m+n}}\binom{2m}{m}\binom{2n}{n}\frac{1}{2m+1}\frac{1}{2n+1}\int_0^\pi \cos(x)^{2m+2n+2}dx \\
&= \frac{\pi}{16}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{1}{64^{m+n}}\binom{2m}{m}\binom{2n}{n}\binom{2m+2n+2}{m+n+1}\frac{1}{2m+1}\frac{1}{2n+1} \\
&= \frac{\pi}{2}\operatorname{Li}_2\left(\frac{1}{4}\right)
\end{aligned}$$

which is a curious generalization of sums found in [24]. Consequently

$$\begin{aligned}
&\int_0^\pi \frac{x\sin(x)}{\sqrt{4-\cos(x)^2}}\arcsin\left(\frac{\cos(x)}{2}\right)dx \\
&= -\frac{x}{2}\left(\arcsin\left(\frac{\cos(x)}{2}\right)\right)^2\Bigg|_0^\pi + \frac{1}{2}\int_0^\pi \left(\arcsin\left(\frac{\cos(x)}{2}\right)\right)^2 dx \\
&= -\frac{\pi^3}{72} + \frac{\pi}{4}\operatorname{Li}_2\left(\frac{1}{4}\right)
\end{aligned}$$

using integration by parts again. Third, $G(\pi) = 1$ and $G(0) = 0$, where $G'(x) = g(x)$. Finally,

$$\begin{aligned}
\int_0^\pi x^2 G'(x)dx &= x^2 G(x)\big|_0^\pi - 2\int_0^\pi x\,G(x)\,dx \\
&= \pi^2 - \frac{2}{\pi}\int_0^\pi x\frac{\sin(x)}{\sqrt{4-\cos(x)^2}}\left(\frac{\pi}{2}+\arcsin\left(\frac{\cos(x)}{2}\right)\right)dx - \frac{2}{\pi}\int_0^\pi x^2 dx \\
&= \pi^2 - \frac{\pi^2}{6} - 2\left(-\frac{\pi^2}{72}+\frac{1}{4}\operatorname{Li}_2\left(\frac{1}{4}\right)\right) - \frac{2}{3}\pi^2 \\
&= \frac{7}{36}\pi^2 - \frac{1}{2}\operatorname{Li}_2\left(\frac{1}{4}\right)
\end{aligned}$$

as was to be shown.

A random Gaussian triangle *captures* a location (ξ,η) with probability

$$\frac{3}{(2\pi)^{5/2}}\left[\varphi(\delta)+\psi(\delta)\right]=\begin{cases}0.250000... & \text{if } \delta=0,\\ 0.197171... & \text{if } \delta=1/2,\\ 0.098289... & \text{if } \delta=1,\\ 0.032455... & \text{if } \delta=3/2,\\ 0.007626... & \text{if } \delta=2,\end{cases}$$

where $\delta=\sqrt{\xi^2+\eta^2}$ and

$$\varphi=\int_0^\infty\int_0^\infty\int_{-\infty}^0 \exp\left(-\tfrac{(a_1+\delta)^2+(b_1+\delta)^2+(c_1+\delta)^2}{2}\right)\left[\pi+2\arctan\left(\tfrac{a_1 b_1}{c_1\sqrt{a_1^2+b_1^2+c_1^2}}\right)\right]dc_1\,db_1\,da_1,$$

$$\psi=\int_{-\infty}^0\int_{-\infty}^0\int_0^\infty \exp\left(-\tfrac{(a_1+\delta)^2+(b_1+\delta)^2+(c_1+\delta)^2}{2}\right)\left[\pi-2\arctan\left(\tfrac{a_1 b_1}{c_1\sqrt{a_1^2+b_1^2+c_1^2}}\right)\right]dc_1\,db_1\,da_1.$$

The specific result $1/4$ for capturing $(0,0)$ is well-known [25]; the general result is less so [26]. See also [27–29].

We conclude with an unsolved problem: what is an exact expression for

$$\mathrm{E}(a\,\gamma)=\frac{1}{3\pi}\int_0^\infty\int_0^\infty\int_0^\pi x^2\,y\,\theta\exp\left[-\frac{1}{3}\left(x^2-x\,y\cos(\theta)+y^2\right)\right]d\theta\,dy\,dx=1.6377...$$

(in two dimensions)? An answer for $\mathrm{E}(a\,\alpha)$ is believed to be even more difficult.

Acknowledgments I am grateful to Wilfrid Kendall, Stephen Herschkorn, James Ferry and William Gosper for their kind help, and to the creators of both Mathematica and Maple.

[1] S. Portnoy, A Lewis Carroll pillow problem: probability of an obtuse triangle, *Statist. Sci.* 9 (1994) 279–284; MR1293297 (95h:60003).
[2] B. Eisenberg and R. Sullivan, Random triangles in *n* dimensions, *Amer. Math. Monthly* 103 (1996) 308–318; MR1383668 (96m:60025).
[3] S. R. Finch, Geometric probability constants, *first volume*, pp. 479–484.
[4] P. Clifford and N. J. B. Green, Distances in Gaussian point sets, *Math. Proc. Cambridge Philos. Soc.* 97 (1985) 515–524; MR0778687 (86i:62091).
[5] K. S. Miller, Complex Gaussian processes, *SIAM Rev.* 11 (1969) 544–567; MR0258109 (41 #2756).
[6] K. S. Miller, *Complex Stochastic Processes. An Introduction to Theory and Application*, Addison-Wesley, 1974, pp. 86–100; MR0368118 (51 #4360).
[7] S. R. Finch, Archimedes' constant, *first volume*, pp. 17–28.
[8] H. O. Lancaster, Traces and cumulants of quadratic forms in normal variables, *J. Royal Statist. Soc. Ser. B* 16 (1954) 247–254; MR0070088 (16,1128h).
[9] S. R. Searle, *Linear Models*, Wiley, 1971, pp. 54–64; MR0293792 (45 #2868).
[10] S. J. Herschkorn and J. Ferry, Random triangle problem, USENET sci.math newsgroup posting, 2004.

[11] M. I. Stoka, Alcuni problemi di probabilità geometriche non uniformi, *Annali della Facoltà di Economia e Commercio Univ. Bari* 28 (1989) 525–541.
[12] E. Bosetto, Systems of stochastically independent and normally distributed random points in the Euclidean space E_3, *Beiträge Algebra Geom.* 40 (1999) 291–301; MR1720105 (2000g:60015).
[13] K. V. Mardia, Recent directional distributions with applications, *Statistical Distributions in Scientific Work*, v. 6, *Applications in Physical, Social, and Life Sciences*, Proc. 1980 NATO Advanced Study Institute, Univ. Trieste, ed. C. Taillie, G. P. Patil and B. A. Baldessari, Reidel, 1981, pp. 1–19; MR0656354 (84b:62071).
[14] W. S. Kendall, *Some Problems in Probability Theory*, Ph.D. thesis, Oxford Univ., 1979.
[15] D. G. Kendall and W. S. Kendall, Alignments in two-dimensional random sets of points, *Adv. Appl. Probab.* 12 (1980) 380–424; MR0569434 (81d:60014).
[16] D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, Wiley, 1987, pp. 265–273; MR0895588 (88j:60034a).
[17] S. R. Finch, Notation: Error function, *this volume*, p. xi.
[18] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[19] J. Lee, J. Kim and S.-H. Jung, Bayesian analysis of paired survival data using a bivariate exponential distribution, *Lifetime Data Anal.* 13 (2007) 119–137; MR2355299.
[20] S. R. Finch, Sobolev isoperimetric constants, *first volume*, pp. 219–225.
[21] P. A. P. Moran, Testing for correlation between non-negative variates, *Biometrika* 54 (1967) 385–394; MR0221711 (36 #4763).
[22] S. R. Finch, Ornstein-Uhlenbeck process, *this volume*, §4.3.
[23] T. P. Hutchinson and C. D. Lai, *Continuous Bivariate Distributions, Emphasising Applications*, Rumsby Scientific Publishing, 1990, pp. 138–158, 167, 223; MR1070715 (92e:62097).
[24] S. R. Finch, Central binomial coefficients, *this volume*, §1.22.
[25] R. Howard and P. Sisson, Capturing the origin with random points: generalizations of a Putnam problem, *College Math. J.* 27 (1996) 186–192; MR1390366.
[26] S. R. Finch, Capturing, ordering and Gaussianity in 2D, arXiv:1601.04937.
[27] S. R. Finch, Expected n-step product for Gaussian tours, arXiv:1512.05592.
[28] S. R. Finch, Pins, stakes, anchors and Gaussian triangles, arXiv:1410.6742.
[29] S. R. Finch, Random Gaussian tetrahedra, arXiv:1005.1033.

5.11 Random Triangles. II

Let S denote the unit sphere in Euclidean 3-space. A spherical triangle T is a region enclosed by three great circles on S; a great circle is a circle whose center is at the origin. The sides of T are arcs of great circles and have length a, b, c. Each of these is $\leq \pi$. The angle α opposite side a is the dihedral angle between the two planes passing through the origin and determined by arcs b, c. The angles β, γ opposite sides b, c are similarly defined. Each of these is $\leq \pi$ too [1].

The sum of the angles is $\leq 3\pi$ yet $\geq \pi$. In particular, the sum need not be the constant π. Define the spherical excess $E = \alpha + \beta + \gamma - \pi$. The sum of the sides is ≥ 0 yet $\leq 2\pi$. Define the spherical defect $D = 2\pi - (a + b + c)$. It can be shown that the area of T is E and a calculus-based proof appears in [2]; see also [3]. Clearly the perimeter of T is $2\pi - D$.

The probability density functions for sides, angles, excess and defect on S will occupy us in this essay. Random triangles are defined here by selecting three independent uniformly distributed points on the sphere to be vertices. One way to do this is to let X_1, X_2, X_3, Y_1, Y_2, Y_3, Z_1, Z_2, Z_3 be independent normally distributed random variables with mean 0 and variance 1; then the points

$$\frac{(X_1, Y_1, Z_1)}{\sqrt{X_1^2 + Y_1^2 + Z_1^2}}, \quad \frac{(X_2, Y_2, Z_2)}{\sqrt{X_2^2 + Y_2^2 + Z_2^2}}, \quad \frac{(X_3, Y_3, Z_3)}{\sqrt{X_3^2 + Y_3^2 + Z_3^2}}$$

satisfy our requirements. Any spherically-symmetric underlying distribution will do, in fact, but we shall refer to the normal variables X_i, Y_j, Z_k again at a later time.

5.11.1 *Sides*

The trivariate density $f(x, y, z)$ for sides a, b, c is [4]

$$\begin{cases} \dfrac{1}{4\pi} \dfrac{\sin(x)\sin(y)\sin(z)}{\sqrt{1 - \cos(x)^2 - \cos(y)^2 - \cos(z)^2 + 2\cos(x)\cos(y)\cos(z)}} \\ \qquad \text{if } x + y + z < 2\pi,\ x + y > z,\ y + z > x \text{ and } z + x > y, \\ 0 \qquad \text{otherwise.} \end{cases}$$

As a consequence, the univariate density for a is

$$\frac{1}{2}\sin(x), \quad 0 < x < \pi$$

and

$$\mathrm{E}(a) = \frac{\pi}{2} = 1.5707963267\ldots, \quad \mathrm{E}(a^2) = \frac{\pi^2}{2} - 2 = 2.9348022005\ldots.$$

Sides a, b, c are uncorrelated and, moreover, pairwise independent. They are, however, mutually dependent, since [5–7]

$$\mathrm{P}\left(a < \frac{\pi}{2}, b < \frac{\pi}{2}, c < \frac{\pi}{2}\right) = \frac{1}{4}\left(1 - \frac{1}{\pi}\right) > \frac{1}{8},$$

$$\mathrm{P}\left(a > \frac{\pi}{2}, b > \frac{\pi}{2}, c > \frac{\pi}{2}\right) = \frac{1}{4\pi} < \frac{1}{8}$$

and since $\mathrm{E}(a\,b\,c) = 3.694\ldots < \pi^3/8$.

5.11.2 *Angles*

The trivariate density $g(x,y,z)$ for angles α, β, γ is [4]

$$\begin{cases} \dfrac{1}{4\pi}\dfrac{1-\cos(x)^2-\cos(y)^2-\cos(z)^2-2\cos(x)\cos(y)\cos(z)}{\sin(x)^2\sin(y)^2\sin(z)^2} \\ \qquad \text{if } x+y+z>\pi,\ x+y<\pi+z,\ y+z<\pi+x \text{ and } z+x<\pi+y, \\ 0 \qquad \text{otherwise.} \end{cases}$$

As a consequence, α is uniformly distributed on $[0,\pi]$ and

$$\mathrm{E}(\alpha)=\frac{\pi}{2}=1.5707963267..., \quad \mathrm{E}(\alpha^2)=\frac{\pi^2}{3}=3.8757845850....$$

Angles α, β, γ are uncorrelated but, unlike before, pairwise *dependent*. Integrating out z, the bivariate density for α, β is

$$\frac{1}{2\pi}\frac{1}{\sin(x)^2\sin(y)^2}\cdot\begin{cases} -\cos(y)\sin(y)+y & \text{if } x-y>0 \text{ and } x+y<\pi, \\ \pi+\cos(y)\sin(y)-y & \text{if } x-y<0 \text{ and } x+y>\pi, \\ -\cos(x)\sin(x)+x & \text{if } x-y<0 \text{ and } x+y<\pi, \\ \pi+\cos(x)\sin(x)-x & \text{if } x-y>0 \text{ and } x+y>\pi \end{cases}$$

which is not uniform on $[0,\pi]\times[0,\pi]$. The mutual dependence can also be seen from [5–7]

$$\mathrm{P}\left(\alpha<\frac{\pi}{2},\beta<\frac{\pi}{2},\gamma<\frac{\pi}{2}\right)=\frac{1}{2}\left(\frac{1}{\pi}-\frac{1}{4}\right)<\frac{1}{8},$$
$$\mathrm{P}\left(\alpha>\frac{\pi}{2},\beta>\frac{\pi}{2},\gamma>\frac{\pi}{2}\right)=\frac{1}{2}\left(\frac{3}{4}-\frac{1}{\pi}\right)>\frac{1}{8}$$

and from $\mathrm{E}(\alpha\,\beta\,\gamma)=4.688...>\pi^3/8$.

5.11.3 *Excess and Defect*

In this section, we gather several results which seem to defy easy analysis. A proof that angle α is uncorrelated with either adjacent side b or c is known, hence $\mathrm{E}(\alpha\,b)=\pi^2/4=\mathrm{E}(\alpha\,c)$ immediately. The joint moment of α with its opposite side a is obviously a triple integral:

$$\begin{aligned}\mathrm{E}(\alpha\,a) &= \frac{1}{4\pi}\int_0^{\pi}\int_0^{\pi}\int_0^{\pi}\sin(x)\sin(y)z \\ &\quad\times\arccos\left[\cos(x)\cos(y)+\sin(x)\sin(y)\cos(z)\right]dx\,dy\,dz\end{aligned}$$

whose exact evaluation seems difficult. Miles [4] proved, via stochastic geometry, that $\mathrm{E}(\alpha\,a)=\pi^2/2-2$ as a special case of a more general theorem. As a consequence, the correlation coefficient between α and a is

$$\rho(\alpha,a)=\frac{\sqrt{3(\pi^2-8)}}{\pi}=0.7538511740....$$

Recall from [8] that analogous results for Gaussian triangles in the plane remain open.

Clearly

$$\mathrm{E}(\alpha+\beta+\gamma-\pi)=\frac{\pi}{2},\quad \mathrm{E}((\alpha+\beta+\gamma-\pi)^2)=\frac{\pi^2}{2},$$

$$\mathrm{E}(2\pi-a-b-c)=\frac{\pi}{2},\quad \mathrm{E}((2\pi-a-b-c)^2)=\pi^2-6$$

however the verification of

$$\mathrm{E}((\alpha+\beta+\gamma-\pi)(2\pi-a-b-c))=6-\frac{\pi^2}{2},$$

$$\rho(E,D)=-\frac{\sqrt{3(\pi^2-8)}}{\pi}=-0.7538511740...$$

rests on the aforementioned nontrivial result.

A proposed density $h(x)$ for excess E was published in 1867 [9]:

$$-\frac{(x^2-4\pi x+3\pi^2-6)\cos(x)-6(x-2\pi)\sin(x)-2(x^2-4\pi x+3\pi^2+3)}{16\pi\cos(x/2)^4}$$

for $0<x<2\pi$ and remained obscure until it was cited in a recent paper [10]. The supporting proof is geometric. No analytic proof using our trivariate density for α, β, γ has yet been found. In some relevant 1928 calculations, Burnside [11] remarked that, “in a similar way”, the probability that the area of T should lie between x and $x+dx$ “may be determined”. Miles [4] confessed in 1971 that the functional form of $h(x)$ has “so far eluded the author”, but then mentioned (in a footnote) pertinent work of J. N. Boots.

With regard to defect D, Jones & Benyon-Tinker [12] expressed the perimeter density in terms of elliptic integrals [8]:

$$k(x)=\frac{1}{4\pi}\int_0^{x/2}\frac{E\left(\sin\left(\frac{t}{2}\right)\right)-\cos\left(\frac{x-t}{2}\right)^2 K\left(\sin\left(\frac{t}{2}\right)\right)}{\sqrt{\cos\left(\frac{t}{2}\right)^2-\cos\left(\frac{x-t}{2}\right)^2}}\sin(t)\,dt.$$

No closed-form evaluation of this integral is known. Finch & Jones [13] recognized the value $k(\pi)=3\sqrt{2}/32$ and revisited the proof of the area density $h(x)$. See also [14].

Miles’ [4] proof that $\mathrm{E}(\alpha\, a)=\pi^2/2-2$ is clarified in [15]. Let Li_3 denote the trilogarithm function [16] and G denote Catalan’s constant [17]. It is interesting that the conditional moment

$$\mathrm{E}(\alpha\, a\,|\, b=\pi/2)=3.0538319164...=\frac{\pi^2}{8}-\frac{\ln(2)^2}{2}-\frac{4G}{\pi}+\frac{8}{\pi}\,\mathrm{Im}\left(\mathrm{Li}_3(1+i)\right)$$

remains complicated whereas

$$\mathrm{E}(\alpha\, a\,|\,\beta=\pi/2)=2.8708787614...=\frac{\pi}{4}\left[2+(1+\ln(2))\,\pi-4G\right]$$

is simple.

5.11.4 *Proof for (a,b,γ)*

We will demonstrate that a, b, γ are independent random variables; the sides a, b each have the sine density on $[0,\pi]$ and the angle γ is uniformly distributed on $[0,\pi]$. Our starting point is the fact that a is an angle between two vectors (X_1, Y_1, Z_1) and (X_3, Y_3, Z_3), where X_i, Y_j, Z_k were defined earlier, and b likewise for the vectors (X_2, Y_2, Z_2) and (X_3, Y_3, Z_3). The formulas [18–20]

$$\cos(a) = \frac{X_1X_3 + Y_1Y_3 + Z_1Z_3}{\sqrt{X_1^2 + Y_1^2 + Z_1^2}\sqrt{X_3^2 + Y_3^2 + Z_3^2}},$$

$$\cos(b) = \frac{X_2X_3 + Y_2Y_3 + Z_2Z_3}{\sqrt{X_2^2 + Y_2^2 + Z_2^2}\sqrt{X_3^2 + Y_3^2 + Z_3^2}}$$

are familiar: $\cos(a)$ is the sample correlation coefficient r_{13} between two samples of size three (each sample coming from a population of known mean $=0$) and $\cos(b)$ is likewise the sample correlation coefficient r_{23}. Also, by the Law of Cosines for Sides:

$$\cos(c) = \cos(a)\cos(b) + \sin(a)\sin(b)\cos(\gamma)$$

we obtain

$$\cos(\gamma) = \frac{\cos(c) - \cos(a)\cos(b)}{\sin(a)\sin(b)} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$

and recognize this as the sample partial correlation coefficient $r_{12\cdot3}$ between samples 1 and 2, holding variable 3 fixed. An exercise in [21] states that r_{13}, r_{23}, $r_{12\cdot3}$ are independent because X_i, Y_j, Z_k are independent and normally distributed. Hence a, b, γ are independent as well.

The sample correlation coefficient r_{13} is uniformly distributed on $[-1,1]$, as a special case of results given in [22–25], hence

$$\begin{aligned} \mathrm{P}(a < \xi) &= \mathrm{P}(\cos(a) > \cos(\xi)) = \mathrm{P}(r_{13} > \cos(\xi)) = \frac{1}{2}\int_{\cos(\xi)}^{1} d\eta \\ &= \frac{1 - \cos(\xi)}{2} \end{aligned}$$

and $d\mathrm{P}(a < \xi)/d\xi = \sin(\xi)/2$. The sample partial correlation coefficient $r_{12\cdot3}$ has the arcsine distribution on $[-1,1]$, hence

$$\begin{aligned} \mathrm{P}(\gamma < \xi) &= \mathrm{P}(\cos(\gamma) > \cos(\xi)) = \mathrm{P}(r_{12\cdot3} > \cos(\xi)) = \frac{1}{\pi}\int_{\cos(\xi)}^{1} \frac{d\eta}{\sqrt{1 - \eta^2}} \\ &= \frac{1}{2} - \frac{1}{\pi}\arcsin(\cos(\xi)) = \frac{1}{\pi}\xi \end{aligned}$$

and $d\mathrm{P}(\gamma < \xi)/d\xi = 1/\pi$, as was to be shown.

Geisser & Mantel [26] were the first to notice that the correlations r_{13}, r_{23}, r_{12} are pairwise but not mutually independent (for samples of arbitrary size). This "natural" example has been justly celebrated and is of "valuable pedagogical use" [27]. Recasting the example in terms of spherical triangle sides a, b, c makes it even more remarkable, in our opinion. No one seems to have linked Miles' paper [4] in geometric probability to ongoing research in theoretical statistics.

5.11.5 *Proof for* (a, β, γ)

We bring β into the trivariate density $\sin(a)\sin(b)/(4\pi)$, removing b. From the Law of Cosines for Angles:

$$-\cos(\alpha) = \cos(\beta)\cos(\gamma) - \sin(\beta)\sin(\gamma)\cos(a)$$

we have

$$\begin{aligned}\sin(\alpha)^3 &= \left(1-\cos(\alpha)^2\right)^{3/2} \\ &= \left(1-(\cos(\beta)\cos(\gamma)-\sin(\beta)\sin(\gamma)\cos(a))^2\right)^{3/2}\end{aligned}$$

since $0 < \alpha < \pi$. Differentiating the identity [1, 4]

$$\sin(a)\cot(b) = \cot(\beta)\sin(\gamma) + \cos(\gamma)\cos(a)$$

with respect to b, we obtain

$$-\sin(a)\csc(b)^2 db = -\csc(\beta)^2\sin(\gamma)d\beta$$

hence

$$db = \frac{\sin(b)^2\sin(\gamma)}{\sin(a)\sin(\beta)^2}d\beta.$$

Via the Law of Sines:

$$\frac{\sin(a)}{\sin(\alpha)} = \frac{\sin(b)}{\sin(\beta)} = \frac{\sin(c)}{\sin(\gamma)}$$

the density $\sin(a)\sin(b)/(4\pi)$ becomes

$$\begin{aligned}\frac{1}{4\pi}\sin(a)\frac{\sin(b)^3\sin(\gamma)}{\sin(a)\sin(\beta)^2} &= \frac{1}{4\pi}\frac{\sin(b)^3}{\sin(\beta)^3}\sin(\beta)\sin(\gamma) \\ &= \frac{1}{4\pi}\frac{\sin(a)^3}{\sin(\alpha)^3}\sin(\beta)\sin(\gamma) \\ &= \frac{1}{4\pi}\frac{\sin(\beta)\sin(\gamma)\sin(a)^3}{(1-(\cos(\beta)\cos(\gamma)-\sin(\beta)\sin(\gamma)\cos(a))^2)^{3/2}}.\end{aligned}$$

More elaborate arguments lead to the trivariate densities of (a, b, c) and (α, β, γ).

This preceding expression is helpful in computing the bivariate density for (β, γ). Integrating out a gives

$$\frac{|\sin(\beta-\gamma)|\cos(\beta+\gamma)-|\sin(\beta+\gamma)|\cos(\beta-\gamma)+\arcsin(\cos(\beta-\gamma))-\arcsin(\cos(\beta+\gamma))}{4\pi\sin(\beta)^2\sin(\gamma)^2}$$

which seems complicated at first glance. Everything simplifies if we partition the square $[0,\pi]\times[0,\pi]$ into four isosceles right triangles according to the diagonal lines $\beta-\gamma=0$, $\beta+\gamma=\pi$. For example, if $\beta-\gamma>0$ and $\beta+\gamma<\pi$, then the numerator becomes $-2\cos(\gamma)\sin(\gamma)+2\gamma$. As another example, if $\beta-\gamma<0$ and $\beta+\gamma>\pi$, then the numerator becomes $2\pi+2\cos(\gamma)\sin(\gamma)-2\gamma$. For the remaining two triangles, γ is merely replaced by β, by symmetry. Such formulas can be used to confirm directly that β, γ are each uniformly distributed on $[0,\pi]$ and $\mathrm{E}(\beta\,\gamma)=\pi^2/4$.

A joint density for (a,α) might assist in evaluating the triple integral mentioned earlier, but finding this (and the joint density for $(r_{13}, r_{13.2})$) seems to be hard.

Any spherical triangle T determines a unique chordal triangle T' (with sides as straight lines through the interior of S) and vice versa. Let r' denote the radius of the unique circle passing through the three vertices of T'. The density for two T' angles is given in [28], as well as the trivariate density for two T' sides coupled with r'. Such results lead to progress in answering an open question: What is the exact probability that four random circular caps of angular radius 88° completely cover S? The progress is, however, insignificant if 88° is replaced by, say, 71°. We hope to see resolution of this issue someday.

[1] I. Todhunter, *Spherical Trigonometry for the Use of Colleges and Schools*, 5th ed., Macmillan, 1886, pp. 6–9, 16–32, 67–74.

[2] J. Makesh Pravin, Symmetry on a sphere, *Resonance* 7 (2002) 84–88.

[3] J. C. Polking, The area of a spherical triangle. Girard's theorem, *The Geometry of the Sphere*, lecture notes (1996), Rice Univ.

[4] R. E. Miles, Random points, sets and tessellations on the surface of a sphere, *Sankhya Ser. A* 33 (1971) 145–174; MR0321150 (47 #9683).

[5] Y. Maeda, Seven types of random spherical triangle in S^n and their probabilities, *Computational Geometry and Graph Theory*, Proc. 2007 Kyoto conf., Lect. Notes in Comp. Sci. 4535, Springer-Verlag, 2008, pp. 119–126.

[6] E. J. Watson and E. H. Lockwood, Probability of an acute-angled spherical triangle, *Math. Gazette* 58 (1974) 217–218.

[7] M. W. Crofton and Exhumatus, Problem 2326, *Mathematical Questions and Solutions from the "Educational Times"*, v. 7, ed. W. J. Miller, Hodgson & Son, Jan.-Jul. 1867, pp. 105–106.

[8] S. R. Finch, Random triangles, *this volume*, §5.10.

[9] M. W. Crofton and Exhumatus, Problem 2370, *Mathematical Questions and Solutions from the "Educational Times"*, v. 8, ed. W. J. Miller, Hodgson & Son, Jul.-Dec. 1867, pp. 21–23.

[10] M. Hindmarsh, A.-C. Davis and R. Brandenberger, Formation of topological defects in first order phase transitions, *Phys. Rev. D* 49 (1994) 1944–1950; arXiv:hep-ph/9307203.
[11] W. Burnside, *Theory of Probability*, Cambridge Univ. Press, 1928, pp. 74–79.
[12] A. J. Jones and G. Benyon-Tinker, The jet-set travelling salesman: (1) Triangular tours, unpublished manuscript (1994).
[13] S. R. Finch and A. J. Jones, Random spherical triangles, arXiv:1009.5329.
[14] J. Philip, The area and perimeter of a random spherical triangle, unpublished manuscript (2014).
[15] S. R. Finch, Correlation between angle and side, arXiv:1012.0781.
[16] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.
[17] S. R. Finch, Catalan's constant, *first volume*, pp. 53–59.
[18] D. Jackson, The trigonometry of correlation, *Amer. Math. Monthly* 31 (1924) 275–280.
[19] J. O. Irwin, Note on the addition formula for the Jacobian elliptic functions and their connection with the theory of correlation in statistics, *Math. Gazette* 49 (1965) 425–427.
[20] G. Thomas, G., and J. O'Quigley, A geometric interpretation of partial correlation using spherical triangles, *Amer. Statist.* 47 (1993) 30–32.
[21] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 2nd ed., Wiley, 1984, pp. 102–109, 130–133, 251, 266–268, 282–283 (esp. exercises 21 & 22 in sect. 7.6); MR0771294 (86b:62079).
[22] N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions*, v. II, 2nd ed., Wiley, 1995, pp. 545–638; MR1326603 (96j:62029).
[23] W. A. Chance, A geometric derivation of the distribution of the correlation coefficient $|r|$ when $\rho = 0$, *Amer. Math. Monthly* 93 (1986) 94–98; MR0827582 (87d:62102).
[24] M. S. Srivastava and C. G. Khatri, *An Introduction to Multivariate Statistics*, North-Holland, 1979, pp. 87–92, 100–101; MR0544670 (81j:62001).
[25] T. W. Anderson, R. A. Fisher and multivariate analysis, *Statist. Sci.* 11 (1996) 20–34; MR1437125.
[26] S. Geisser and N. Mantel, Pairwise independence of jointly dependent variables, *Annals Math. Statist.* 33 (1962) 290–291; MR0137188 (25 #644).
[27] R. A. Schaufele, An application of pairwise independence of random variables to regression analysis, *Canad. Math. Bull.* 18 (1975) 397–404; MR0405733 (53 #9525).
[28] S. R. Finch, Covering a sphere with four random circular caps, arXiv:1502.06931.

5.12 Random Triangles. III

Let Ω be a compact convex set in Euclidean n-space with nonempty interior. Random triangles are defined here by selecting three independent uniformly distributed points in Ω to be vertices. Generating such points for $(n, \Omega) = (2,\text{unit square})$ or $(n, \Omega) = (3,\text{unit cube})$ is straightforward. For $(n, \Omega) = (2,\text{unit disk})$ or $(n, \Omega) = (3,\text{unit ball})$, we use the following result [1]. Let X_1, X_2, X_3, Y_1, Y_2, Y_3, Z_1, Z_2, Z_3 be independent normally distributed random variables with mean 0 and variance $1/2$. Let W_1, W_2, W_3 be exponential random variables, independent of the others, with mean 1. Then the points

$$\frac{(X_1, Y_1)}{\sqrt{X_1^2+Y_1^2+W_1}}, \quad \frac{(X_2, Y_2)}{\sqrt{X_2^2+Y_2^2+W_2}}, \quad \frac{(X_3, Y_3)}{\sqrt{X_3^2+Y_3^2+W_3}}$$

are uniform in the disk, and the points

$$\frac{(X_1, Y_1, Z_1)}{\sqrt{X_1^2+Y_1^2+Z_1^2+W_1}}, \quad \frac{(X_2, Y_2, Z_2)}{\sqrt{X_2^2+Y_2^2+Z_2^2+W_2}}, \quad \frac{(X_3, Y_3, Z_3)}{\sqrt{X_3^2+Y_3^2+Z_3^2+W_3}}$$

are uniform in the ball. Compared with the intricate joint distributions of sides and angles for Gaussian triangles [2] and for spherical triangles [3], little is known for uniform triangles in Ω.

5.12.1 *Disk*

The density $f(x)$ for an arbitrary side a of a random uniform triangle in the unit disk is [4–11]

$$\frac{4x}{\pi}\arccos\left(\frac{x}{2}\right)-\frac{x^2}{\pi}\sqrt{4-x^2}, \quad 0<x<2$$

and

$$\mathrm{E}(a)=\frac{128}{45\pi}=0.9054147873..., \quad \mathrm{E}(a^2)=1.$$

No one has attempted to extend this univariate result to a bivariate or trivariate density, as far as is known.

The density $g(x)$ for an arbitrary angle α is [12, 13]

$$\frac{\cos(5x)-(1-12\pi x+12x^2)\cos(x)}{12\pi^2\sin(x)^3} - \frac{(\pi-x)\cos(4x)+10(\pi-x)\cos(2x)+(\pi-13x)}{12\pi^2\sin(x)^2}$$

when $0<x<\pi$ and

$$\mathrm{E}(\alpha)=\frac{\pi}{3}=1.0471975511..., \quad \mathrm{E}(\alpha^2)=\frac{\pi^2}{6}+\frac{1}{12}=1.7282674001....$$

We can also give partial results for the maximum angle (analogous to the Gaussian case [2]). Corresponding to the density for $\max\{\alpha,\beta,\gamma\}$, the expression $3g(x)$ holds when $\pi/2<x<\pi$; an expression when $\pi/3<x<\pi/2$ remains open, although a numerical approach is employed in [14]. It can also be shown that [15–17]

$$\mathrm{P}(\text{a uniform triangle in the disk is acute})=\frac{4}{\pi^2}-\frac{1}{8}=0.2802847345....$$

Moments of area are known [18–25]:

$$\mathrm{E}(\text{area}) = \frac{35}{48\pi} = 0.2321009586... = (0.0738800297...)\pi,$$

$$\mathrm{E}(\text{area}^2) = \frac{3}{32} = 0.09375.$$

Via the generalised hypergeometric function ${}_pF_q$ [26], define

$$\begin{aligned}\Phi_n(y) = {} & \frac{9}{16\pi}\frac{2^{n/2}}{(n-2)!}\Gamma\left(1+\frac{n}{2}\right)^3 y^{(n-3)/2}\left\{-\frac{4\pi^{3/2}}{\sqrt{3}}\frac{1}{\Gamma(\frac{1}{2}+\frac{n}{4})\Gamma(1+\frac{n}{4})}\left(\frac{4y}{27}\right)^{1/2}\right. \\ & -\frac{3\Gamma(-\frac{5}{6})\Gamma(\frac{1}{6})}{2^{2/3}\sqrt{\pi}}\frac{1}{\Gamma(\frac{1}{6}+\frac{n}{4})\Gamma(\frac{2}{3}+\frac{n}{4})}\left(\frac{4y}{27}\right)^{5/6}{}_4F_3\left(\frac{1}{3},\frac{1}{3},\frac{5}{6}-\frac{n}{4},\frac{1}{3}-\frac{n}{4};\frac{2}{3},\frac{11}{6},\frac{4}{3};\frac{4y}{27}\right) \\ & +\frac{9\sqrt{\pi}\Gamma(-\frac{7}{6})}{2^{1/3}\Gamma(\frac{1}{6})}\frac{1}{\Gamma(-\frac{1}{6}+\frac{n}{4})\Gamma(\frac{1}{3}+\frac{n}{4})}\left(\frac{4y}{27}\right)^{7/6}{}_4F_3\left(\frac{2}{3},\frac{2}{3},\frac{7}{6}-\frac{n}{4},\frac{2}{3}-\frac{n}{4};\frac{4}{3},\frac{13}{6},\frac{5}{3};\frac{4y}{27}\right) \\ & \left.+\frac{4\sqrt{\pi}}{3}\frac{1}{\Gamma(1+\frac{n}{4})\Gamma(\frac{3}{2}+\frac{n}{4})}{}_4F_3\left(-\frac{1}{2},-\frac{1}{2},-\frac{n}{4},-\frac{1}{2}-\frac{n}{4};\frac{1}{6},-\frac{1}{6},\frac{1}{2};\frac{4y}{27}\right)\right\}\end{aligned}$$

for $n \geq 2$ and $0 < y < 27/4$. Then the density for area is given by $8x\Phi_2\left(4x^2\right)$, a result due to Mathai [27]. We shall see Φ_3 shortly.

The bivariate density for (a, b) in the unit disk can be found, imitating Parry's [28] analysis. Finch [29] concluded that

$$\mathrm{E}(a\,b) = 0.8378520652..., \quad \rho(a, b) = 0.1002980835...,$$

$$\mathrm{E}(\text{perimeter}^2) = 8.0271123917...$$

but exact evaluation of these constants remains open.

5.12.2 *Ball*

The density $f(x)$ for an arbitrary side a of a random uniform triangle in the unit ball is [6, 10, 11]

$$\frac{3}{16}x^5 - \frac{9}{4}x^3 + 3x^2, \quad 0 < x < 2$$

and

$$\mathrm{E}(a) = \frac{36}{35} = 1.0285714285..., \quad \mathrm{E}(a^2) = \frac{6}{5} = 1.2.$$

A recent extraordinary calculation [28, 30] gives a trivariate density for the sides (a, b, c). For reasons of space, we report only the bivariate density $f(x, y)$ for (a, b):

$$f(x,y) = \begin{cases} \varphi(x,y) & \text{if } x + y \leq 2, \\ \psi(x,y) & \text{if } x + y > 2 \text{ and } x \leq 2 \end{cases}$$

when $0 \le y \le x$ (use symmetry otherwise) where

$$\begin{aligned}\varphi(x,y) &= \frac{9}{16}x^5y^2 - \frac{27}{4}x^3y^2 + \frac{27}{16}x^3y^3 + \frac{9}{16}x^3y^4 + 9x^2y^2 \\ &\quad -\frac{27}{8}x^2y^3 + \frac{9}{32}x^2y^5 - \frac{9}{8}xy^4 + \frac{9}{160}xy^6,\end{aligned}$$

$$\begin{aligned}\psi(x,y) &= -\frac{9}{160}x^6y + \frac{9}{32}x^5y^2 + \frac{9}{8}x^4y - \frac{9}{16}x^4y^3 - \frac{9}{4}x^3y - \frac{27}{8}x^3y^2 \\ &\quad + \frac{27}{16}x^3y^3 + \frac{9}{2}x^2y^2 + \frac{9}{5}xy - \frac{9}{4}xy^3.\end{aligned}$$

It follows that

$$\mathrm{E}(a\,b) = \frac{884}{825} = 1.0715..., \quad \rho(a,b) = \frac{884/825 - (36/35)^2}{6/5 - (36/35)^2} = \frac{274}{2871} = 0.0954...$$

and the cross-correlation coefficient is somewhat smaller than that found in the Gaussian case [2].

With regard to angles, apart from $\mathrm{E}(\alpha) = \pi/3$, all we know is that [16, 17]

$$\mathrm{P}(\text{a uniform triangle in the ball is acute}) = \frac{33}{70} = 0.4714285714....$$

Mathai [27] showed that the density for area is given by $8x\Phi_3\left(4x^2\right)$, hence [31]

$$\mathrm{E}(\text{area}) = \frac{9\pi}{77} = 0.3671991413..., \quad \mathrm{E}(\text{area}^2) = \frac{9}{50} = 0.18,$$

$$\mathrm{E}(\text{perimeter}) = \frac{108}{35}, \quad \mathrm{E}(\text{perimeter}^2) = \frac{2758}{275}.$$

It is surprising that E(perimeter^2) is known in exactly in three dimensions but not in two dimensions.

5.12.3 *Square*

The density $f(x)$ for an arbitrary side a of a random uniform triangle in the unit square is [10, 32]

$$\begin{cases} 2x^3 - 8x^2 + 2\pi x & \text{if } 0 \le x \le 1, \\ 8x\sqrt{x^2-1} - 2x^3 + 2(\pi-2)x - 8x\arctan\left(\sqrt{x^2-1}\right) & \text{if } 1 < x \le \sqrt{2} \end{cases}$$

and

$$\mathrm{E}(a) = \frac{1}{15}\left(2 + \sqrt{2} + 5\ln(1+\sqrt{2})\right) = 0.5214054331..., \quad \mathrm{E}(a^2) = \frac{1}{3}.$$

Nothing comparable is known for an arbitrary angle α; we have only [33, 34]

$$\begin{aligned}\mathrm{P}(\text{a uniform triangle in the square is acute}) &= 1-\left(\frac{97}{150}+\frac{\pi}{40}\right)\\ &= 1-0.7252064830...\\ &= 0.2747935169....\end{aligned}$$

A remarkable formula holds for the density $h(x)$ for area [22, 35–38]:

$$\begin{aligned}&(-16\pi^2x^2-16\pi^2x-24x+12)+(240x^2-96x-12)\ln(1-2x)\\ &\quad -240x^2\ln(2x)+48x^2\ln(2x)^2+(96x^2+96x)\,\mathrm{Li}_2(2x),\end{aligned}$$

where $\mathrm{Li}_2(\xi)$ is the dilogarithm function [2]. As a consequence,

$$\mathrm{E}(\text{area})=\frac{11}{144},\quad \mathrm{E}(\text{area}^2)=\frac{1}{96}.$$

Evaluating $\mathrm{E}(\text{perimeter}^2)$ remains open.

5.12.4 *Cube*

The density $f(x)$ for an arbitrary side a of a random uniform triangle in the unit cube is [39–41]

$$\begin{cases} -x^5+8x^4-6\pi x^3+4\pi x^2 & \text{if } 0\le x\le 1,\\[2ex] -8x(2x^2+1)\sqrt{x^2-1}+2x^5+6x^3-8\pi x^2 & \\ \quad +(6\pi-1)x+24x^3\arctan\left(\sqrt{x^2-1}\right) & \text{if } 1< x\le\sqrt{2},\\[2ex] 8x(x^2+1)\sqrt{x^2-2}-x^5+6(\pi-1)x^3-8\pi x^2+(6\pi-5)x & \\ \quad -24x(x^2+1)\arctan\left(\sqrt{x^2-2}\right)+24x^2\arctan\left(x\sqrt{x^2-2}\right) & \text{if } \sqrt{2}< x\le\sqrt{3}\end{cases}$$

and

$$\begin{aligned}\mathrm{E}(a) &= \frac{1}{105}\left(4+17\sqrt{2}-6\sqrt{3}+21\ln(1+\sqrt{2})+42\ln(2+\sqrt{3})-7\pi\right)\\ &= 0.6617071822...,\end{aligned}$$

$$\mathrm{E}(a^2)=\frac{1}{2}.$$

Essentially nothing else is known: it would be good someday to learn more about the associated acuteness probability and E(area). See [42] for related discussion.

[1] F. Barthe, O. Guédon, S. Mendelson and A. Naor, A probabilistic approach to the geometry of the ℓ_p^n-ball, *Annals Probab.* 33 (2005) 480–513; MR2123199 (2006g:46014).

[2] S. R. Finch, Random triangles, *this volume*, § 5.10.
[3] S. R. Finch, Random triangles. II, *this volume*, §5.11.
[4] E. Borel, *Principes et formules classiques du calcul des probabilités*, t. 1, *Traité du calcul des Probabilités et de ses Applications*, f. 1, Gauthier-Villars, 1925, pp. 74–90.
[5] R. Deltheil, *Probabilités Géométriques*, t. 2, *Traité du calcul des Probabilités et de ses Applications*, f. 2, ed E. Borel, Gauthier-Villars, 1926, pp. 40–42, 114–120.
[6] J. M. Hammersley, The distribution of distance in a hypersphere, *Annals Math. Statist.* 21 (1950) 447–452; MR0037481 (12,268e).
[7] R. D. Lord, The distribution of distance in a hypersphere, *Annals Math. Statist.* 25 (1954) 794–798; MR0065048 (16,377d).
[8] V. S. Alagar, The distribution of the distance between random points, *J. Appl. Probab.* 13 (1976) 558–566; MR0418183 (54 2#6225).
[9] H. Solomon, *Geometric Probability*, SIAM, 1978, pp. 35–36, 128–129; MR0488215 (58 #7777).
[10] S. R. Dunbar, The average distance between points in geometric figures, *College Math. J.* 28 (1997) 187–197; MR1444006 (98a:52007).
[11] S.-J. Tu and E. Fischbach, Random distance distribution for spherical objects: general theory and applications to physics, *J. Phys. A* 35 (2002) 6557–6570; MR1928848.
[12] R. Sullivan, *Crofton's Theorem for Parametrized Families of Convex Polygons*, Ph.D. thesis, Lehigh Univ., 1996.
[13] B. Eisenberg and R. Sullivan, Random triangles in n dimensions, *Amer. Math. Monthly* 103 (1996) 308–318; MR1383668 (96m:60025).
[14] C. Small, Random uniform triangles and the alignment problem, *Math. Proc. Cambridge Philos. Soc.* 91 (1982) 315–322; MR0641532 (83b:62033).
[15] W. S. B. Woolhouse, Problem 1350, *Mathematical Questions and Solutions from the "Educational Times*", v. 1, ed. W. J. Miller, Hodgson & Son, Jul. 1863-Jun. 1864, pp. 49–51.
[16] G. R. Hall, Acute triangles in the n-ball, *J. Appl. Probab.* 19 (1982) 712–715; MR0664859 (83h:60016).
[17] C. Buchta, A note on the volume of a random polytope in a tetrahedron, *Illinois J. Math.* 30 (1986) 653–659; MR0857217 (87m:60038).
[18] C. M. Ingleby, Correction on the four-point problem, *Mathematical Questions and Solutions from the "Educational Times*", v. 5, ed. W. J. Miller, Hodgson & Son, Jan.-Jul. 1866, pp. 108–109.
[19] W. S. B. Woolhouse, Some additional observations, *Mathematical Questions and Solutions from the "Educational Times*", v. 7, ed. W. J. Miller, Hodgson & Son, Jan.-Jul. 1867, pp. 81–83.
[20] W. S. B. Woolhouse, Problem 2471, *Mathematical Questions and Solutions from the "Educational Times*", v. 8, ed. W. J. Miller, Hodgson & Son, Jul.-Dec. 1867, pp. 100–105.
[21] M. G. Kendall and P. A. P. Moran, *Geometrical Probability*, Hafner, 1963, pp. 42–46; MR0174068 (30 #4275).
[22] N. Henze, Random triangles in convex regions, *J. Appl. Probab.* 20 (1983) 111–125; MR0688085 (84g:60019).
[23] A. M. Mathai, *An Introduction to Geometrical Probability*, Gordon and Breach, 1999, pp. 159–171; MR1737197 (2001a:60014).
[24] V. Klee, What is the expected volume of a simplex whose vertices are chosen at random from a given convex body? *Amer. Math. Monthly* 76 (1969) 286–288; MR1535340.

[25] R. E. Pfiefer, The historical development of J. J. Sylvester's four point problem, *Math. Mag.* 62 (1989) 309–317; MR1031429 (90m:52001).
[26] S. R. Finch, Notation: Hypergeometric function, *this volume*, p. xi.
[27] A. M. Mathai, On a conjecture in geometric probability regarding asymptotic normality of a random simplex, *Annals Probab.* 10 (1982) 247–251; MR0637392 (82m:60019).
[28] M. Parry, *Application of Geometric Probability Techniques to Elementary Particle and Nuclear Physics*, Ph.D. thesis, Purdue Univ., 1998.
[29] S. R. Finch, Perimeter variance of uniform random triangles, arXiv:1007.0261.
[30] M. Parry and E. Fischbach, Probability distribution of distance in a uniform ellipsoid: theory and applications to physics, *J. Math. Phys.* 41 (2000) 2417–2433; MR1751899 (2001j:81267).
[31] C. Buchta and J. Müller, Random polytopes in a ball, *J. Appl. Probab.* 21 (1984) 753–762; MR0766813 (86e:60013).
[32] B. Ghosh, Random distances within a rectangle and between two rectangles, *Bull. Calcutta Math. Soc.* 43 (1951) 17–24; MR0044765 (13,475a).
[33] E. Langford, The probability that a random triangle is obtuse, *Biometrika* 56 (1969) 689–690.
[34] E. Langford, A problem in geometrical probability, *Math. Mag.* 43 (1970) 237–244; MR0298725 (45 #7774).
[35] P. Valtr, Probability that n random points are in convex position, *Discrete Comput. Geom.* 13 (1995) 637–643; MR1318803 (96c:60017).
[36] M. Trott, The area of a random triangle, *Mathematica J.* 7 (1998) 189–198.
[37] Z. F. Seidov, Random triangle problem: geometrical approach, arXiv:math/0002134.
[38] J. Philip, The area of a random convex polygon, unpublished manuscript (2004).
[39] D. P. Robbins and T. S. Bolis, Average distance between two points in a box, *Amer. Math. Monthly* 85 (1978) 277–278.
[40] A. M. Mathai, P. Moschopoulos and G. Pederzoli, Distance between random points in a cube, *Statistica (Bologna)* 59 (1999) 61–81; MR1793157 (2001h:60031).
[41] J. Philip, The probability distribution of the distance between two random points in a box, unpublished manuscript (2007).
[42] S. R. Finch, Rectilinear crossing constant, *first volume*, pp. 532–534.

5.13 Random Triangles. IV

We step back momentarily to gain perspective. By **parabolic geometry** is meant the study of distances, angles, etc., in a Riemannian manifold having zero scalar curvature; for example, geometry in two-dimensional Euclidean space $\mathbb{R}^2$ (the planar model).

By **elliptic geometry** is meant the study of such properties in a Riemannian manifold having positive scalar curvature. Given a line (geodesic) L and a point P not on L, there is no line parallel to L passing through P. The sum of the three angles of a triangle is greater than π; the quantity $(\alpha + \beta + \gamma) - \pi$ is called angular excess. The simplest example of this geometry is the spherical model S embedded in three-dimensional Euclidean space $\mathbb{R}^3$. Geodesics are great circles, that is, intersections of S with two-dimensional subspaces of $\mathbb{R}^3$.

By **hyperbolic geometry** is meant the study of such properties in a Riemannian manifold having negative scalar curvature. Given a line (geodesic) L and a point P not on L, there are at least two distinct lines parallel to L passing through P. The sum of the three angles of a triangle is less than π; the quantity $\pi-(\alpha+\beta+\gamma)$ is called angular defect. The simplest example of this geometry is the hyperboloidal model H embedded in three-dimensional Minkowski space $\mathbb{M}^3$. Geodesics are great hyperbolas, that is, *nonempty* intersections of H with two-dimensional subspaces of $\mathbb{M}^3$.

With regard to the latter, $\mathbb{M}^3$ is the vector space of ordered real triples (just like $\mathbb{R}^3$) equipped with the symmetric bilinear form [1–3]

$$q\left[(x,y,z),(u,v,w)\right]=-zw+xu+yv$$

instead of the usual (positive definite) inner product

$$p\left[(x,y,z),(u,v,w)\right]=xu+yv+zw.$$

Define the unit hyperboloid H to be the positive sheet ($z>0$) of points satisfying $q[(x,y,z),(x,y,z)]=-1$; equivalently,

$$H=\left\{(x,y,z)\in\mathbb{M}^3:z=\sqrt{1+x^2+y^2}\right\}.$$

This is analogous to the unit sphere S of points satisfying $p[(x,y,z),(x,y,z)]=1$; equivalently,

$$S=\left\{(x,y,z)\in\mathbb{R}^3:z=\pm\sqrt{1-x^2-y^2}\right\}.$$

Distance between two points in H:

$$\operatorname{arccosh}\left(-q\left[(x,y,z),(u,v,w)\right]\right)$$

is analogous to distance between two points in S:

$$\arccos\left(p\left[(x,y,z),(u,v,w)\right]\right)$$

(the latter is the angle at the origin determined by the two vectors).

A hyperbolic triangle T is a region enclosed by three geodesics on H. The sides of T are arcs of great hyperbolas and have length a, b, c. Since H is non-compact, there is no upper bound on these. To define a uniform distribution, we will need to introduce some restrictions. The angle α opposite side a is the dihedral angle between the two planes passing through the origin and determined by arcs b, c. The angles β, γ opposite sides b, c are similarly defined. Each of these is $\leq\pi$. By the Law of Cosines for Sides:

$$\cosh(c)=\cosh(a)\cosh(b)-\sinh(a)\sinh(b)\cos(\gamma)$$

we obtain

$$\cos(\gamma)=-\frac{\cosh(c)-\cosh(a)\cosh(b)}{\sinh(a)\sinh(b)}$$

analogous to an expression for $\cos(\gamma)$ in spherical trigonometry [4].

The disk of radius $R > 0$ on H is

$$\begin{aligned}\Delta_R &= \{(x, y, z) \in H : \operatorname{arccosh}(-q[(x, y, z), (0, 0, 1)]) \le R\} \\ &= \{(x, y, z) \in H : z \le \cosh(R)\}.\end{aligned}$$

This is analogous to the disk of radius $0 < R < \pi$ on S:

$$\{(x, y, z) \in S : \arccos(p[(x, y, z), (0, 0, 1)]) \le R\} = \{(x, y, z) \in S : z \ge \cos(R)\};$$

the special case when $R = \pi/2$ is a hemisphere on S.

The orthogonal projection of Δ_R ($\subset H$) into the xy-plane gives simply the disk $x^2 + y^2 \le \sinh(R)^2$ because

$$\sqrt{1 + x^2 + y^2} = z \le \cosh(R) \quad \text{implies} \quad x^2 + y^2 \le \cosh(R)^2 - 1 = \sinh(R)^2.$$

It is hence apparent [5] that circular circumference is proportional to $\sinh(R)$.

An alternative mapping from Δ_R into the xy-plane is nonlinear:

$$\begin{pmatrix} \sqrt{z^2 - 1}\cos(\theta) \\ \sqrt{z^2 - 1}\sin(\theta) \\ z \end{pmatrix} \mapsto \begin{pmatrix} \operatorname{arccosh}(z)\cos(\theta) \\ \operatorname{arccosh}(z)\sin(\theta) \end{pmatrix}$$

but has the advantage that Δ_R is mapped onto the (even simpler) disk $x^2 + y^2 \le R^2$. The inverse mapping

$$\begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix} \mapsto \begin{pmatrix} \sinh(r)\cos(\theta) \\ \sinh(r)\sin(\theta) \\ \cosh(r) \end{pmatrix}$$

will be helpful soon; call this Φ for convenience.

We now discuss the random generation of uniform points in Δ_R. Here it is useful to first review the generation of points in the Euclidean planar disk of radius R. We want distance ξ between a random point and the center $(0, 0)$ to possess density function

$$f(\xi) = \frac{2}{R^2}\xi, \quad 0 < \xi < R$$

(proportional to circular circumference, radius ξ). The cumulative distribution is

$$\eta = F(\xi) = \int_0^\xi \frac{2}{R^2} t\, dt = \frac{1}{R^2}\xi^2, \quad 0 < \eta < 1$$

hence $\xi = R\sqrt{\eta}$. By the inverse CDF method, the point

$$\begin{pmatrix} R\sqrt{\eta}\cos(\theta) \\ R\sqrt{\eta}\sin(\theta) \end{pmatrix} \quad \text{where} \quad \eta \sim \text{Unif}[0, 1], \ \theta \sim \text{Unif}[0, 2\pi]$$

satisfies the desired uniformity condition.

Returning now to Δ_R, we want distance ξ between a random point and the center $(0, 0, 1)$ to possess density function [6, 7]

$$f(\xi) = \frac{\sinh(\xi)}{\cosh(R) - 1}, \quad 0 < \xi < R$$

(again by proportionality). The cumulative distribution is

$$\eta = F(\xi) = \int_0^\xi \frac{\sinh(t)}{\cosh(R) - 1}\, dt = \frac{\cosh(\xi) - 1}{\cosh(R) - 1}, \quad 0 < \eta < 1$$

hence $\xi = \operatorname{arccosh}(1 + (\cosh(R) - 1)\eta)$. In the planar disk of radius R, the point

$$\begin{pmatrix} \operatorname{arccosh}(1 + (\cosh(R) - 1)\eta)\cos(\theta) \\ \operatorname{arccosh}(1 + (\cosh(R) - 1)\eta)\sin(\theta) \end{pmatrix} \quad \text{where} \quad \eta \sim \mathrm{Unif}[0, 1],\ \theta \sim \mathrm{Unif}[0, 2\pi]$$

is more likely to appear near the circular boundary than near the center. Applying the transformation Φ, we obtain that

$$\begin{pmatrix} \sqrt{(1 + (\cosh(R) - 1)\eta)^2 - 1}\cos(\theta) \\ \sqrt{(1 + (\cosh(R) - 1)\eta)^2 - 1}\sin(\theta) \\ 1 + (\cosh(R) - 1)\eta \end{pmatrix} \quad \text{where} \quad \eta \sim \mathrm{Unif}[0, 1],\ \theta \sim \mathrm{Unif}[0, 2\pi]$$

satisfies the desired uniformity condition in Δ_R.

5.13.1 *Sides*

We do not know the trivariate density $f(x, y, z)$ for sides a, b, c of a uniform random triangle in Δ_R. Let

$$X = \frac{\cosh(a)}{L^2}, \quad Y = \frac{\cosh(b)}{L^2}, \quad Z = \frac{\cosh(c)}{L^2}$$

denote normalized sides, where $L = \cosh(R) - 1$. The trivariate characteristic function

$$\mathrm{E}\left(\exp\left(irX + isY + itZ\right)\right)$$

has a complicated quintuple integral expression [6, 7] that we choose not to reproduce here. Setting $r = s = 0$, the following expression for the univariate characteristic function for Z emerges:

$$\begin{aligned} &\frac{1}{2\pi}\int_0^{2\pi}\int_{1/L}^{1+1/L}\int_{1/L}^{1+1/L} \exp\left[it\left(uv - \cos(\varphi)\sqrt{u^2 - \frac{1}{L^2}}\sqrt{v^2 - \frac{1}{L^2}}\right)\right] du\, dv\, d\varphi \\ &= \int_{1/L}^{1+1/L}\int_{1/L}^{1+1/L} J_0\left(t\sqrt{u^2 - \frac{1}{L^2}}\sqrt{v^2 - \frac{1}{L^2}}\right)\exp\left(ituv\right) du\, dv, \end{aligned}$$

where $J_0(\theta)$ is the zeroth Bessel function of the first kind. It follows that

$$\mathrm{E}(Z) = \left(\frac{L+2}{2L}\right)^2, \quad \mathrm{E}(Z^2) = \frac{L^4 + 6L^3 + 13L^2 + 12L + 6}{6L^4}$$

and, in the limit as $R \to \infty$, the univariate density for Z tends to

$$-1 + \frac{2}{\pi}\sqrt{\frac{2}{\zeta} - 1} + \frac{1}{\pi}\arccos(1-\zeta), \quad 0 < \zeta < 2.$$

It also follows that

$$\mathrm{E}(YZ) = \frac{(L+2)^2(L^2+3L+3)}{12L^4}$$

from the biivariate characteristic function for Y, Z:

$$\int\limits_{1/L}^{1+1/L}\int\limits_{1/L}^{1+1/L}\int\limits_{1/L}^{1+1/L} J_0\left(s\sqrt{u^2 - \frac{1}{L^2}}\sqrt{w^2 - \frac{1}{L^2}}\right)$$
$$\times J_0\left(t\sqrt{u^2 - \frac{1}{L^2}}\sqrt{v^2 - \frac{1}{L^2}}\right)\exp\left(isuw + ituv\right)du\,dv\,dw.$$

A complicated expression for the limiting trivariate density for X, Y, Z exists [6] in terms of a certain elliptic integral, but again we omit this.

5.13.2 *Angles*

We know even less about the density for angles α, β, γ of a uniform random triangle in Δ_R. This is unfortunate since the angular defect $\pi - (\alpha + \beta + \gamma)$ is equal to the area of the triangle and this is an important quantity to understand.

By the Law of Cosines for Sides, a triangle is acute if and only if the three inequalities

$$\cosh(a)\cosh(b) > \cosh(c),$$

$$\cosh(a)\cosh(c) > \cosh(b),$$

$$\cosh(b)\cosh(c) > \cosh(a)$$

hold, which permits a proof of [7]

$$\lim_{R\to\infty} \mathrm{P}(\text{a uniform triangle in } \Delta_R \text{ is acute}) = 1.$$

We close with an interesting variation. The **circumscribed circle** of a triangle is a circle that goes through the three vertices of the triangle. If such a circle exists, its center is called the **circumcenter** (which coincides with the intersection of the three perpendicular bisectors of the sides). We say, under such a condition, that the triangle possesses a circumcenter. This is true if and only if the three inequalities

$$\sinh\left(\frac{a}{2}\right) < \sinh\left(\frac{b}{2}\right) + \sinh\left(\frac{c}{2}\right),$$

$$\sinh\left(\frac{b}{2}\right) < \sinh\left(\frac{a}{2}\right) + \sinh\left(\frac{c}{2}\right),$$

$$\sinh\left(\frac{c}{2}\right) < \sinh\left(\frac{a}{2}\right) + \sinh\left(\frac{b}{2}\right)$$

hold, which inspires a numerical computation [7]

$$\lim_{R\to\infty} \mathrm{P}(\text{a uniform triangle in } \Delta_R \text{ possesses a circumcenter}) = 0.4596203....$$

No exact expression for this constant is known.

[1] W. F. Reynolds, Hyperbolic geometry on a hyperboloid, *Amer. Math. Monthly* 100 (1993) 442–455; MR1215530 (94f:51037).

[2] M. Reid and B. Szendroi, *Geometry and Topology*, Cambridge Univ. Press, 2005, pp. 34–61; MR2194744 (2006g:51002).

[3] J. G. Ratcliffe, *Foundations of Hyperbolic Manifolds*, 2nd ed., Springer-Verlag, 2006, pp. 12, 54–99; MR2249478 (2007d:57029).

[4] S. R. Finch, Random triangles. II, *this volume*, §5.11.

[5] J. W. Anderson, *Hyperbolic Geometry*, 2nd ed., Springer-Verlag, 2005, pp. 189–208; MR2161463 (2006b:51001).

[6] Y. Isokawa, Limit distributions of random triangles in hyperbolic planes, *Bull. Faculty Educ. Kagoshima Univ. Natur. Sci.* 49 (1997) 1–16; MR1653095 (99k:60020).

[7] Y. Isokawa, Geometric probabilities concerning large random triangles in the hyperbolic plane, *Kodai Math. J.* 23 (2000) 171–186; MR1768179 (2001f:60014).

5.14 Random Triangles. V

We defined the d-dimensional Poisson(λ) point process in an earlier essay [1] and exhibited moment formulas for various cellular parameters of the corresponding Voronoi tessellation. Many of these formulas are analytically intractible; numerical integration is sometimes necessary. For example, when $d = 2$, the probability that a typical cell (a convex polygon) is a triangle is 0.01124001.... Monte Carlo simulation often provides the only window for study. For example, when $d = 2$ or $d = 3$, the value of the density function f_V for the cellular volume V tends to 0 for small arguments [2–4], although no workable expression for either density is known.

Any $d + 1$ particles from the point process define almost surely an open ball which contains the $d + 1$ particles on its boundary. If no other particles from the process are contained in the ball, then let C denote the convex polyhedron with vertices at the $d + 1$ particles. The collection of all such cells C constitute almost surely a subdivision of $\mathbb{R}^d$, called the **Poisson–Delaunay tessellation**. This can be regarded as dual to what we discussed earlier. Formulas here are more accessible than in [1]. When $d=2$ and $d=3$, the cells are almost surely triangles and tetrahedra, respectively [5, 6]. The value of the density f_V tends to 0 for small arguments in both cases [7, 9]. Since our interest is in random triangles, we will focus on the scenario $d=2$.

The shortest distance between a line L and the origin is the length $|r|$ of the perpendicular segment from $(0,0)$ to L. If the intersection point is $(r\cos(\theta), r\sin(\theta))$, then clearly the equation for L is

$$x\cos(\theta)+y\sin(\theta)=r.$$

There is a one-to-one correspondence between the set of points

$$Q=\{(r,\theta):-\infty<r<\infty \text{ and } 0\leq\theta<\pi\}$$

and the set of all lines L. For arbitrary $\lambda>0$, the Poisson point process of intensity λ in Q induces the **Poisson line process** of intensity λ. The resulting subdivision of $\mathbb{R}^2$ is called the **Goudsmit–Miles tessellation** of the plane (Figure 5.10). Formulas here are again more accessible than before. The probability that a typical cell (again a convex polygon) is a triangle is $2-\pi^2/6=0.3550659331...$ [10–12]. The value of the density f_V tends to ∞ (not 0) for small arguments; more precisely [13–15],

$$\begin{aligned}\lim_{x\to0^+}\sqrt{x}f_V(x) &= \lambda\frac{2\sqrt{2}}{12-\pi^2}\int_0^{\pi}\int_0^{\pi-\varphi}\sqrt{\sin(\varphi)\sin(\psi)\sin(\varphi+\psi)}d\psi\,d\varphi\\ &= \frac{6\pi}{12-\pi^2}(0.3231100260...)\lambda.\end{aligned}$$

One of our goals is to explain why this interesting constant arises! It turns out that "small" polygons of the tessellation are almost all triangles, therefore knowledge about triangular areas carries over to limiting polygonal areas. We shall discuss both.

There are effectively no stationary line processes except the one with Poisson structure. This is somewhat exaggerated – mixtures of the λ parameter in Poisson line processes lead to Cox line processes and there are pathological examples with many parallel lines – but essentially it is not worthwhile to consider any other tessellation based on a line process [16, 17].

Characteristics of **Delaunay triangles** and **Miles triangles** will dominate this essay. There is a third type of random triangle – called **Miles intriangles** – that we shall touch upon as well. Let us review: The largest circle inscribed in a given convex polygon is called the incircle. This circle will almost surely be tangent to three sides of the polygon. Let T denote the triangle determined by these three sides, extended as far as required. Given a typical cell C in a planar Goudsmit–Miles tessellation, the intriangle T might be considerably larger than C (the prefix "in"

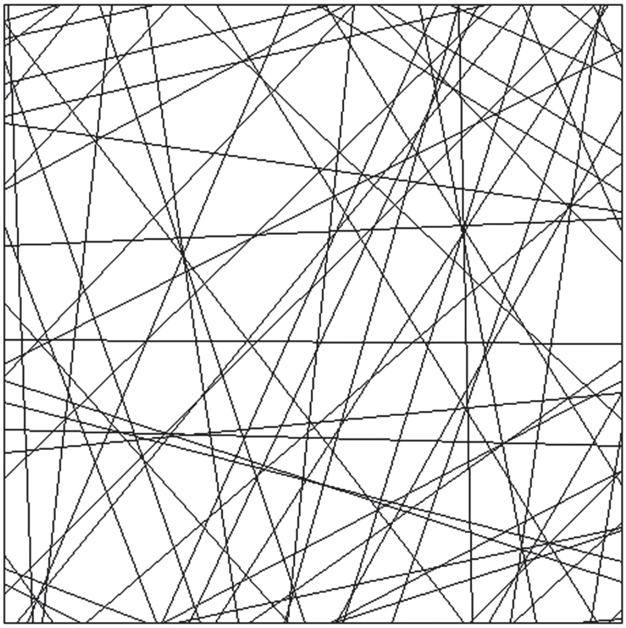

Figure 5.10 A Poisson line process of intensity 17 in the unit square.

refers to the fact that C and T share an incircle, *not* that T is inscribed in anything). Clearly no other line is allowed to intersect the incircle. However, it is possible that one or more lines might hit the intriangle elsewhere. It can be shown that both the area V of T and the perimeter S of T have infinite expectation [18].

Buried in an appendix to [18], we find these three types of random triangles listed in a table. Clarifying Miles' table is the second goal of this essay. For simplicity, we shall assume $\lambda = 1$ henceforth.

5.14.1 *Delaunay Triangles*

Unlike the examples in [19], it is easier to start with angles than with sides. The bivariate density for arbitrary angles α, β in a typical (triangular) cell of a Poisson–Delaunay tessellation is [5, 18, 20]

$$\frac{8}{3\pi}\sin(x)\sin(y)\sin(x+y),$$

where $x > 0$, $y > 0$, $x + y < \pi$. Integrating out y, we obtain the density $g(x)$ for α:

$$\frac{4}{3\pi}\left[(\pi - x)\cos(x) + \sin(x)\right]\sin(x)$$

and

$$\mathrm{E}(\alpha) = \frac{\pi}{3} = 1.0471975511..., \quad \mathrm{E}(\alpha^2) = \frac{2\pi^2}{9} - \frac{5}{6} = 1.3599120891....$$

Corresponding to the density for $\max\{\alpha,\beta,\gamma\}$, the expression $3g(x)$ holds when $\pi/2<x<\pi$; the expression when $\pi/3<x<\pi/2$ is [21, 22]

$$\frac{4}{\pi}\left[(3x-\pi)\cos(x)-\sin(3x)\right]\sin(x).$$

It thus follows that

$$\mathrm{P}(\text{a typical Delaunay triangle is acute})=\frac{1}{2}=0.5.$$

As in [23], the cross-correlation coefficient $\rho(\alpha,\beta)=-1/2$, hence

$$\mathrm{E}(\alpha\,\beta)=\frac{\pi^2}{18}+\frac{5}{12}=0.9649780222....$$

The density for an arbitrary side a is [6, 24, 25]

$$\frac{\pi\,x}{3}\left[x\exp\left(-\frac{\pi\,x^2}{4}\right)+\operatorname{erfc}\left(\frac{\sqrt{\pi}x}{2}\right)\right],$$

where $x>0$ and erfc is the complementary error function [8]; also

$$\mathrm{E}(a)=\frac{32}{9\pi}=1.1317684842\,...,\quad \mathrm{E}(a^2)=\frac{5}{\pi}=1.5915494309....$$

No such simple density formula exists for perimeter $S=a+b+c$. Muche [7] gave

$$\frac{\pi\,x^3}{12}\int\limits_0^{2\pi}\int\limits_0^{2\pi-\varphi}\frac{\sin\left(\frac{\varphi}{2}\right)\sin\left(\frac{\psi}{2}\right)\sin\left(\frac{\varphi+\psi}{2}\right)}{\left[\sin\left(\frac{\varphi}{2}\right)+\sin\left(\frac{\psi}{2}\right)+\sin\left(\frac{\varphi+\psi}{2}\right)\right]^4}$$
$$\times\exp\left\{\frac{-\pi\,x^2}{4\left[\sin\left(\frac{\varphi}{2}\right)+\sin\left(\frac{\psi}{2}\right)+\sin\left(\frac{\varphi+\psi}{2}\right)\right]^2}\right\}d\psi\,d\varphi,$$

where $x>0$ and

$$\mathrm{E}(S)=\frac{32}{3\pi}=3.3953054526\,...,\quad \mathrm{E}(S^2)=\frac{125}{3\pi}=13.2629119243....$$

By contrast, for area $V=(1/4)\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}$, we have a formula due to Rathie [9]:

$$\frac{8\,\pi\,x}{9}K_{1/6}\left(\frac{2\,\pi\,x}{3\sqrt{3}}\right)^2,$$

where $x>0$ and

$$K_{1/6}(w)=\sum_{i=0}^{\infty}\frac{\pi}{i!\Gamma\left(\frac{5}{6}+i\right)}\left(\frac{w}{2}\right)^{-\frac{1}{6}+2i}-\sum_{j=0}^{\infty}\frac{\pi}{j!\Gamma(\frac{7}{6}+j)}\left(\frac{w}{2}\right)^{\frac{1}{6}+2j}$$

is the modified Bessel function of the second kind; also

$$\mathrm{E}(V) = \frac{1}{2} = 0.5, \quad \mathrm{E}(V^2) = \frac{35}{8\pi^2} = 0.4432801784....$$

5.14.2 *Miles Triangles*

Cells of a Goudsmit–Miles tessellation are sampled until we obtain a triangular one. The bivariate density for arbitrary angles α, β in such a typical triangle is [18, 26]

$$\frac{4}{12-\pi^2}\frac{\sin(x)\sin(y)\sin(x+y)}{\sin(x)+\sin(y)+\sin(x+y)} = \frac{8}{12-\pi^2}\sin\left(\frac{x}{2}\right)\sin\left(\frac{y}{2}\right)\cos\left(\frac{x+y}{2}\right)$$

where $x > 0$, $y > 0$, $x + y < \pi$. Integrating out y, we obtain the density $h(x)$ for α:

$$\frac{2}{12-\pi^2}\left[2\sin(x) - (\pi - x)(1-\cos(x))\right]$$

and

$$\mathrm{E}(\alpha) = \frac{\pi}{3} = 1.0471975511..., \quad \mathrm{E}(\alpha^2) = \frac{8}{12-\pi^2} + \frac{\pi^2}{6} - 4 = 1.4001051740....$$

Corresponding to the density for $\max\{\alpha, \beta, \gamma\}$, the expression $3h(x)$ holds when $\pi/2 < x < \pi$; the expression when $\pi/3 < x < \pi/2$ is

$$\frac{6}{12-\pi^2}\left[2(1-2\cos(x))\sin(x) - (3x-\pi)(1-\cos(x))\right].$$

It thus follows that

$$\mathrm{P}(\text{a typical Miles triangle is acute}) = \frac{1}{4} + \frac{3(\pi-3)}{12-\pi^2} = 0.4493892406....$$

As in [23], $\rho(\alpha, \beta) = -1/2$, hence

$$\mathrm{E}(\alpha\,\beta) = -\frac{4}{12-\pi^2} + \frac{\pi^2}{12} + 2 = 0.9448814798....$$

Miles [18] gave the trivariate density for sides a, b, c:

$$\frac{1}{12-\pi^2}\frac{(x+y+z)(-x+y+z)(x-y+z)(x+y-z)}{x^2y^2z^2}\exp\left(-(x+y+z)\right)$$

if $|x-y| < z < x+y$, and we shall verify this later. The condition $|x-y| < z < x+y$ is equivalent to $|x-z| < y < x+z$ and to $|y-z| < x < y+z$ via the Law of Cosines. As a consequence,

$$\mathrm{E}(a) = \frac{1}{3} = 0.3333333333..., \quad \mathrm{E}(a^2) = -\frac{96}{\pi^2(12-\pi^2)} + 4 + \frac{8}{\pi^2} = 0.2448288927...,$$

$$\mathrm{E}(a\,b) = \frac{48}{\pi^2(12-\pi^2)} - \frac{5}{3} - \frac{4}{\pi^2} = 0.2109188869....$$

The cross-correlation coefficient

$$\rho(a,b)=\frac{48/(12-\pi^2)-16\pi^2/9-4}{-96/(12-\pi^2)+35\pi^2/9+8}=0.7464061592...$$

is quite large, indicating strong positive dependency. Integrating out z gives the bivariate density for a, b:

$$\frac{\exp\left(-2(x+y)\right)}{(12-\pi^2)x^2y^2}$$
$$\times\left\{\left(x^2-y^2\right)^2\exp\left(x+y\right)\left[\mathrm{Ei}\left(-(x+y)\right)-\mathrm{Ei}\left(-(y-x)\right)\right]+\pi\sum_{k=0}^{3}x^k h_k(x,y)\right\}$$

for $0<x<y$, where

$$h_0(x,y)=(-1+\exp(2x))(-y^3+y^2-2y-2),\quad h_2(x,y)=(-1+\exp(2x))(y+1),$$
$$h_1(x,y)=(1+\exp(2x))(-y^2+2y+2),\qquad h_3(x,y)=1+\exp(2x)$$

and Ei is the exponential integral

$$\mathrm{Ei}(w)=\int_{-\infty}^{w}\frac{\exp(t)}{t}dt,\quad w<0.$$

For $0<y<x$, simply use symmetry. Unlike Delaunay triangles, a closed-form expression for the univariate density for a seems out of reach.

Perimeter S is exponentially distributed, with density

$$\exp\left(-s\right),\quad s>0$$

and moments $\mathrm{E}(S)=1$, $\mathrm{E}(S^2)=2$. A starting point for area V was provided by Miles [26]: if $U=\sqrt{V}$ and

$$\chi=2\sqrt{\cot\left(\frac{\alpha}{2}\right)\cot\left(\frac{\beta}{2}\right)\tan\left(\frac{\alpha+\beta}{2}\right)}=\sqrt{2}\,\frac{\sin(\alpha)+\sin(\beta)+\sin(\alpha+\beta)}{\sqrt{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}}$$

then the conditional density for U, given α and β, is

$$\chi\exp\left(-\chi u\right),\quad u>0.$$

Therefore the unconditional density for U is

$$\int_0^{\pi}\int_0^{\pi-\alpha}\chi\exp\left(-\chi u\right)\frac{4}{12-\pi^2}\frac{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}{\sin(\alpha)+\sin(\beta)+\sin(\alpha+\beta)}d\beta\,d\alpha$$
$$=\frac{4\sqrt{2}}{12-\pi^2}\int_0^{\pi}\int_0^{\pi-\alpha}\sqrt{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}\exp\left(-\chi u\right)d\beta\,d\alpha$$

and, transforming to V, we obtain

$$\frac{2\sqrt{2}}{12-\pi^2}\frac{1}{\sqrt{v}}\int\limits_0^{\pi}\int\limits_0^{\pi-\alpha}\sqrt{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}\exp\left(-\chi\sqrt{v}\right)d\beta\,d\alpha$$

as the area density. Integrating first with respect to v over $(0,\infty)$, it follows that

$$\begin{aligned}\mathrm{E}(V) &= \frac{\sqrt{2}}{12-\pi^2}\int\limits_0^{\pi}\int\limits_0^{\pi-\alpha}\left(\frac{2}{\chi}\right)^3\sqrt{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}d\beta\,d\alpha \\ &= \frac{\pi\,(25-36\ln(2))}{12-\pi^2}=0.0688684716...,\end{aligned}$$

$$\begin{aligned}\mathrm{E}(V^2) &= \frac{3\sqrt{2}}{12-\pi^2}\int\limits_0^{\pi}\int\limits_0^{\pi-\alpha}\left(\frac{2}{\chi}\right)^5\sqrt{\sin(\alpha)\sin(\beta)\sin(\alpha+\beta)}d\beta\,d\alpha \\ &= \frac{3\left(15\pi^2-148\right)}{2(12-\pi^2)}=0.0310266433....\end{aligned}$$

The density formula for V also serves to motivate the constant 0.3231100260... at the beginning of this essay (asymptotics for polygonal cells as $v\to 0^+$).

5.14.3 *A Verification*

Well-known formulas give angles α, β in terms of sides a, b, c:

$$\sin\left(\frac{\alpha}{2}\right)^2=\frac{\left(\frac{S}{2}-b\right)\left(\frac{S}{2}-c\right)}{b\,c}=\frac{(S-2b)(2a+2b-S)}{4b(S-a-b)},$$

$$\sin\left(\frac{\beta}{2}\right)^2=\frac{\left(\frac{S}{2}-a\right)\left(\frac{S}{2}-c\right)}{a\,c}=\frac{(S-2a)(2a+2b-S)}{4a(S-a-b)}$$

where $S=a+b+c$. To compute the Jacobian determinant J of $(a,b,S)\to(\alpha,\beta,S)$, we differentiate $\sin(\alpha/2)^2$:

$$\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)\frac{\partial\alpha}{\partial a}=\frac{\partial}{\partial a}\frac{(S-2b)(2a+2b-S)}{4b(S-a-b)},$$

$$\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)\frac{\partial\alpha}{\partial b}=\frac{\partial}{\partial b}\frac{(S-2b)(2a+2b-S)}{4b(S-a-b)},$$

$$\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)\frac{\partial\alpha}{\partial S}=\frac{\partial}{\partial S}\frac{(S-2b)(2a+2b-S)}{4b(S-a-b)}$$

and likewise differentiate $\sin(\beta/2)^2$. Additional formulas

$$\sin\left(\frac{\alpha}{2}\right)\cos\left(\frac{\alpha}{2}\right)=\sin\left(\frac{\alpha}{2}\right)\sqrt{\frac{\frac{S}{2}\left(\frac{S}{2}-a\right)}{b\,c}}=\frac{\sqrt{S(S-2a)(S-2b)(2a+2b-S)}}{4b(S-a-b)},$$

$$\sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\beta}{2}\right)=\sin\left(\frac{\beta}{2}\right)\sqrt{\frac{\frac{S}{2}\left(\frac{S}{2}-b\right)}{a\,c}}=\frac{\sqrt{S(S-2a)(S-2b)(2a+2b-S)}}{4a(S-a-b)}$$

permit the expression of $\partial\alpha/\partial a$, $\partial\beta/\partial a$, $\ldots$ entirely in terms of a, b, S. We find that

$$J=\frac{S}{a\,b(S-a-b)}=\frac{S}{a\,b\,c}.$$

Let $\gamma=\pi-\alpha-\beta$, then $\sin(\gamma/2)=\cos((\alpha+\beta)/2)$. The conditional density for a and b, given S, is thus

$$\begin{aligned}\frac{8}{12-\pi^2}\sin\left(\frac{\alpha}{2}\right)\sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\gamma}{2}\right)J &= \frac{8}{12-\pi^2}\frac{\left(\frac{S}{2}-a\right)\left(\frac{S}{2}-b\right)\left(\frac{S}{2}-c\right)}{a\,b\,c}\frac{S}{a\,b\,c}\\ &= \frac{S}{12-\pi^2}\frac{S-2a}{a^2}\frac{S-2b}{b^2}\frac{S-2c}{c^2}\end{aligned}$$

where $\max\{2a,2b,2c\}<S$. Our formula corrects an error in [18], which inexplicably gives $4/S^2$ as the first factor. The inequality $2c<S$ implies $S<2a+2b$. It follows that the unconditional density for a, b, S is

$$\frac{S}{12-\pi^2}\frac{S-2a}{a^2}\frac{S-2b}{b^2}\frac{S-2c}{c^2}\exp\left(-S\right)$$

as was to be shown.

5.14.4 *Miles Intriangles*

Three sides (tangential to the incircle) of a typical cell in a Goudsmit–Miles tessellation are extended until they intersect. Not much is known about the triangle so formed; we shall be brief. The bivariate density for arbitrary angles α, β in such a typical triangle is [13, 18]

$$\frac{1}{3\pi}\left[\sin(x)+\sin(y)+\sin(x+y)\right]=\frac{4}{3\pi}\cos\left(\frac{x}{2}\right)\cos\left(\frac{y}{2}\right)\sin\left(\frac{x+y}{2}\right)$$

where $x>0$, $y>0$, $x+y<\pi$. Integrating out y, we obtain the density for α:

$$\frac{1}{3\pi}\left[(\pi-x)\sin(x)+2\left(1+\cos(x)\right)\right]$$

and $\mathrm{E}(\alpha)=\pi/3$. A density for side a is not known, but $\mathrm{E}(a)=\infty$ as implied earlier. We mention two results:

$$\begin{aligned}\mathrm{P}(\text{a typical intriangle is hit by just one line}) &= 2-\pi^2\left(\tfrac{17}{2}-12\ln(2)\right)\\ &= 0.2014241570...,\end{aligned}$$

$$\begin{aligned}\mathrm{P}(\text{a typical intriangle is hit by at least two lines}) &= -3+\pi^2\left(\tfrac{26}{3}-12\ln(2)\right)\\ &= 0.4435099098....\end{aligned}$$

5.14.5 *Goudsmit–Miles Cells*

We move away from triangles and talk about convex polygons with N vertices. Goudsmit [27] proved that a typical such cell, determined via a Poisson line process of intensity $\lambda = 1$, has the following mean values:

$$\mathrm{E}(N) = 4, \quad \mathrm{E}(S) = 2, \quad \mathrm{E}(V) = \frac{1}{\pi} = 0.3183098861....$$

Miles [13, 14] proved that

$$\mathrm{P}(\text{a typical cell has } N = 3) = 2 - \frac{\pi^2}{6} = 0.3550659331...,$$

$$\mathrm{E}(N^2) = \frac{\pi^2}{2} + 12 = 16.9348022005...,$$

$$\mathrm{E}(S^2) = \frac{\pi^2}{2} + 2 = \frac{68.4437543191...}{\pi^2} = 6.9348022005...,$$

$$\mathrm{E}(V^2) = \frac{1}{2} = \frac{48.7045455170...}{\pi^4} = 0.5$$

and announced that D. G. Kendall (unpublished) had obtained

$$\mathrm{E}(V^3) = \frac{4\pi}{7} = \frac{1725.8818444438...}{\pi^6} = 1.7951958020....$$

Cross-moments between N, S, V were also given.

Let us focus on N-results first. Tanner [28, 29] computed that

$$\mathrm{P}(\text{a typical cell has } N = 4) = -\frac{1}{3} - \frac{7\pi^2}{36} + \pi^2 \ln(2) - \frac{7}{2}\zeta(3) = 0.3814662248...,$$

where $\zeta(3)$ is Apéry's constant [30],

$$\begin{aligned} \mathrm{E}(N^3) &= \frac{232}{7} + \frac{39\pi^2}{14} + \frac{\pi^4}{21} + \frac{12\pi^2}{7}\ln(2) - 6\zeta(3) \\ &\quad - \frac{192}{7}\int_0^{\pi/2} x^2 \tan(x)\ln(\sin(x))dx \\ &= 76.0364049460..., \end{aligned}$$

and $\mathrm{E}(N^4) = 362.08446....$ The fourth moment can be expressed as an elaborate quadruple integral and deserves more attention. Simulation [12, 15, 31, 32] suggests that $\mathrm{P}(N = 5) \approx 0.196$ and $\mathrm{P}(N = 6) \approx 0.062$. The function $\mathrm{P}(N = k)$ is apparently maximized when $k = 4$ and falls off for $k \geq 5$; it is known that asymptotically [33]

$$\mathrm{P}(N = k) \sim \frac{8}{3k} \frac{2\left(4\pi^2\right)^{k-1}}{(2k)!}$$

as $k \to \infty$.

Results for perimeter S include [29]

$$\mathrm{E}(S^3) = \mathrm{E}(N^3) - \frac{3\pi^2}{2} - 28 = \frac{1030.4005353057...}{\pi^3} = 33.2319983444...,$$

$$\mathrm{E}(S^4) = \mathrm{E}(N^4) - 2\mathrm{E}(N^3) - \frac{\pi^2}{2} - 4 = \frac{19586.7132...}{\pi^4} = 201.07685...$$

but nothing comparable is known for the fourth moment of area V. Simulation [12, 15, 32] suggests that $\mathrm{E}(V^4) \approx 11.4$. No formulas for the density for either S or V are known.

Our study has been devoted to "typical" cells C; an alternative is the Crofton cell C_0, which is the unique polygon containing the origin. The Crofton cell is *not* typical for Goudsmit–Miles (unlike Poisson–Voronoi, for which typicality *does* hold [34, 35]). Even less is known here. Matheron [36, 37] proved that

$$\mathrm{E}(N_0) = \mathrm{E}(S_0) = \frac{\pi^2}{2} = 4.9348022005...$$

and Miles [18] proved that

$$\mathrm{P}(\text{the Crofton cell has } N_0 = 3) = \frac{\pi^2\,(25 - 36\ln(2))}{6} = 0.0768208880....$$

The following moments

$$\mathrm{E}(V_0) = \frac{\pi}{2} = 1.5707963267..., \quad \mathrm{E}(V_0^2) = \frac{4\pi^2}{7} = 5.6397739434...$$

are listed in [32], but a reference cannot be found. Simulation [31, 32] suggests that $\mathrm{P}(N_0 = 4) \approx 0.297$, $\mathrm{P}(N_0 = 5) \approx 0.341$ and $\mathrm{P}(N_0 = 6) \approx 0.196$. The function $\mathrm{P}(N_0 = k)$ is apparently maximized when $k = 5$ and falls off for $k \geq 6$; it is known that asymptotically [33]

$$\mathrm{P}(N_0 = k) \sim \frac{2k}{3} \frac{2\left(4\pi^2\right)^{k-1}}{(2k)!}$$

as $k \to \infty$. The distribution of N_0 has a thicker tail (greater weight for large k) than N does. Simulation [32] further suggests that $\mathrm{E}(N_0^2) \approx 25.72$, $\mathrm{E}(S_0^2) \approx 30.51$, $\mathrm{E}(V_0^3) \approx 36.03$ and $\mathrm{E}(V_0^4) \approx 357.8$. Again, no density formulas are known.

Hilhorst & Calka [33] wrote about cloud chamber experiments in physics motivating the work of Goudsmit [27]: The problem was "to calculate the probability for three independent lines to nearly pass through the same point, or, put differently, for a typical triangular cell to have an area less than ε in the limit of very small ε". It would seem that the solution came almost twenty years later [13, 14], with asymptotics

$$\int_0^{\varepsilon} f_V(x)dx \sim \frac{12\pi}{12 - \pi^2}(0.3231100260...)\lambda\sqrt{\varepsilon}$$

valid as $\varepsilon \to 0^+$.

Let Ω be a planar convex region with area V and perimeter S. A Poisson point process yields K points in Ω satisfying $\mathrm{E}(K) = \mathrm{Var}(K) = \lambda V$ whereas a Poisson line process yields L lines hitting Ω satisfying $\mathrm{E}(L) = \mathrm{Var}(L) = \lambda S$. The total length M of the line segments crossing Ω is a sum of L independent identically distributed chord lengths and hence is approximately normally distributed with $\mathrm{E}(M) = \pi\lambda V$ and

$$\mathrm{Var}(M) = \begin{cases} \frac{2}{3}\pi\lambda & \text{if } \Omega \text{ is a disk of unit diameter,} \\ \frac{4}{3}\left(1 - \sqrt{2} + 3\,\mathrm{arcsinh}(1)\right)\lambda & \text{if } \Omega \text{ is a square of unit side,} \\ \frac{3}{4}\ln(3)\lambda & \text{if } \Omega \text{ is an equilateral triangle of} \\ & \text{unit side} \end{cases}$$

for suitably large λ. Studies on such chord lengths for regular polygons and ellipses include [38–49].

The number I of intersection points (between the $\binom{L}{2}$ pairs of lines) in Ω satisfies $\mathrm{E}(I) = \pi\,\lambda^2 V$ in general and

$$\mathrm{E}(I) = \frac{1}{4}\pi^2\lambda^2, \quad \mathrm{Var}(I) = \frac{1}{4}\pi^2\lambda^2 + \frac{8}{3}\pi\,\lambda^3$$

in the special case when Ω is a disk of unit diameter. How is the general formula proved? Under the condition that $L = \ell$ is fixed, we have [10, 50]

$$\mathrm{E}(I \,|\, L = \ell) = \ell(\ell - 1)\pi\frac{V}{S^2},$$

thus, allowing L to vary,

$$\begin{aligned} \mathrm{E}(I) &= \left[\mathrm{E}(L^2) - \mathrm{E}(L)\right]\pi\frac{V}{S^2} \\ &= \left[\mathrm{Var}(L) + \mathrm{E}(L)^2 - \mathrm{E}(L)\right]\pi\frac{V}{S^2} \\ &= \pi\mathrm{E}(L)^2\frac{V}{S^2} = \pi\,\lambda^2 S^2\frac{V}{S^2} = \pi\,\lambda^2 V. \end{aligned}$$

Finding the variance expression for the disk is more complicated; it is possible to do likewise for the square and equilateral triangle.

An unrelated new method for generating random triangles, taking one vertex pinned at the origin O and the other two vertices as Poisson particles closest to O, is discussed in [51, 52]. Different results emerge if we focus instead on sides [53, 54], drawn from a linear family of varying slopes and intercepts. Superposition of particles and lines is the subject of [55, 56].

Acknowledgments I am grateful to Richard Cowan for a helpful discussion about the Poisson line process and Miles' fundamental work on the resulting tessellations. Thanks are also due to Adrian Baddeley, Ege Rubak and Rolf Turner for writing an R package *spatstat* [57], which enables testing of numerical predictions in this essay via simulation.

[1] S. R. Finch, Poisson-Voronoi tessellations, *this volume*, §5.4.
[2] A. L. Hinde and R. E. Miles, Monte-Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process, *J. Statist. Comput. Simul.* 10 (1980) 205–223.
[3] M. Tanemura, Random Voronoi cells of higher dimensions, *Symmetry: Intersections of Art and Science*, Proc. 2001 Sydney conf., *Visual Math.*, v. 3 (2001) n. 3.
[4] E. Pineda , P. Bruna and D. Crespo, Cell size distribution in random tessellations of space, *Phys. Rev. E* 70 (2004) 066119.
[5] A. Okabe, B. Boots, K. Sugihara and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed., Wiley, 2000, pp. 389–400; MR1210959 (94a:52033) and MR1770006 (2001c:52013).
[6] D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, 2nd ed., Wiley, 1995, pp. 248–255, 307–315, 324–333; MR0895588 (88j:60034a).
[7] L. Muche, Distributional properties of the three-dimensional Poisson Delaunay cell, *J. Stat. Phys.* 84 (1996) 147–167; MR1401254 (97j:60027).
[8] S. R. Finch, Notation: Error function, *this volume*, p. xi.
[9] P. N. Rathie, On the volume distribution of the typical Poisson-Delaunay cell, *J. Appl. Probab.* 29 (1992) 740–744; MR1174447 (93f:60016).
[10] L. A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley, 1976, pp. 27–30, 45–58; MR0433364 (55 #6340).
[11] H. Solomon, *Geometric Probability*, SIAM, 1978, pp. 39–60; MR0488215 (58 #7777).
[12] E. I. George, Sampling random polygons, *J. Appl. Probab.* 24 (1987) 557–573; MR0900299 (88g:60030).
[13] R. E. Miles, Random polygons determined by random lines in a plane, *Proc. Nat. Acad. Sci. U.S.A.* 52 (1964) 901–907; MR0168000 (29 #5265).
[14] R. E. Miles, Random polygons determined by random lines in a plane. II, *Proc. Nat. Acad. Sci. U.S.A.* 52 (1964) 1157–1160; MR0169258 (29 #6510).
[15] I. K. Crain and R. E. Miles, Monte Carlo estimates of the distributions of the random polygons determined by random lines in the plane, *J. Statist. Comput. Simul.* 4 (1976) 293–325.
[16] L. Breiman, The Poisson tendency in traffic distribution, *Annals Math. Statist.* 34 (1963) 308–311; MR0144785 (26 #2326).
[17] T. Thedéen, A note on the Poisson tendency in traffic distribution, *Annals Math. Statist.* 35 (1964) 1823–1824; MR0169688 (29 #6933).
[18] R. E. Miles, The various aggregates of random polygons determined by random lines in a plane, *Adv. Math.* 10 (1973) 256–290; MR0319232 (47 #7777).
[19] S. R. Finch, Random triangles. III, *this volume*, §5.12.
[20] R. E. Miles, On the homogeneous planar Poisson point process, *Math. Biosci.* 6 (1970) 85–127; MR0279853 (43 #5574).
[21] K. V. Mardia, R. Edwards and M. L. Puri, Analysis of central place theory, Proc. 41st Session of the International Statistical Institute, New Delhi, 1977, v. 2, *Bull. Inst. Internat. Statist.* 47 (1977) 93–110, 138–147; MR0617573 (82i:52014).
[22] B. N. Boots, Using angular properties of Delaunay triangles to evaluate point patterns, *Geographical Analysis* 18 (1986) 252–259.
[23] S. R. Finch, Random triangles, *this volume*, §5.10.

[24] R. Collins, A geometrical sum rule for two-dimensional fluid correlation functions, *J. Phys. C* 1 (1968) 1461–1471.

[25] R. Sibson, The Dirichlet tessellation as an aid in data analysis, *Scand. J. Statist.* 7 (1980) 14–20; MR0565960 (81b:62007).

[26] R. E. Miles, A heuristic proof of a long-standing conjecture of D. G. Kendall concerning the shapes of certain large random polygons, *Adv. Appl. Probab.* 27 (1995) 397–417; MR1334821 (96h:60017).

[27] S. Goudsmit, Random distribution of lines in a plane, *Rev. Mod. Phys.* 17 (1945) 321–322.

[28] J. C. Tanner, The proportion of quadrilaterals formed by random lines in a plane, *J. Appl. Probab.* 20 (1983) 400–404; MR0698543 (84e:60023).

[29] J. C. Tanner, Polygons formed by random lines in a plane: some further results, *J. Appl. Probab.* 20 (1983) 778–787; MR0720469 (84k:60020).

[30] S. R. Finch, Apéry's constant, *first volume*, pp. 40–53.

[31] P. Calka, Precise formulae for the distributions of the principal geometric characteristics of the typical cells of a two-dimensional Poisson-Voronoi tessellation and a Poisson line process, *Adv. Appl. Probab.* 35 (2003) 551–562; MR1990603 (2004k:60021).

[32] J. Michel and K. Paroux, Empirical polygon simulation and central limit theorems for the homogeneous Poisson line process, *Methodol. Comput. Appl. Probab.* 9 (2007) 541–556; MR2404742 (2009c:60026).

[33] H. J. Hilhorst and P. Calka, Random line tessellations of the plane: statistical properties of many-sided cells, *J. Stat. Phys.* 132 (2008) 627–647; MR2429698 (2009k:60029).

[34] J. Møller, *Lectures on Random Voronoi Tessellations*, Lect. Notes in Statist. 87, Springer-Verlag, 1994, p. 112; MR1295245 (96b:60028).

[35] J. Mecke, On the relationship between the 0-cell and the typical cell of a stationary random tessellation, *Pattern Recognition* 32 (1999) 1645–1648.

[36] G. Matheron, Ensembles fermés aléatoires, ensembles semi-markoviens et polyèdres poissoniens, *Adv. Appl. Probab.* 4 (1972) 508–541; MR0348819 (50 #1314).

[37] G. Matheron, *Random Sets and Integral Geometry*, Wiley, 1975 pp. 177–185; MR0385969 (52 #6828).

[38] R. Sulanke, Die Verteilung der Sehnenlangen an ebenen und raumlichen Figuren, *Math. Nachr.* 23 (1961) 51–74; MR0150720 (27 #707).

[39] R. Coleman, Random paths through convex bodies, *J. Appl. Probab.* 6 (1969) 430–441; MR0254890 (40 #8097).

[40] A. M. Kellerer, Considerations on the random traversal of convex bodies and solutions for general cylinders, *Radiation Research* 47 (1971) 359–376.

[41] F. Piefke, The chord length distribution of the ellipse, *Litovsk. Mat. Sb.* v. 19 (1979) n. 3, 45–54; Engl. transl. in *Lithuanian Math. J.* 19 (1979) 325–332; MR0551926 (80m:60019).

[42] P. W. Kuchel and R. J. Vaughan, Average lengths of chords in a square, *Math. Mag.* 54 (1981) 261–269; MR0644078 (83f:60026).

[43] K. Voss, Powers of chords for convex sets, *Biometrical J.* 24 (1982) 513–516; MR690489 (84d:52006).

[44] W. Gille, Chord length distribution density of an infinitely long circular hollow cylinder, *Math. Comput. Modelling* 34 (2001) 423–431; MR1835837.

[45] N. G. Agaronyan and V. K. Oganyan, Chord length distribution functions for polygons, *Izv. Nats. Akad. Nauk Armenii Mat.* v. 40 (2005) n. 4, 42–57; Engl. transl. in *J. Contemp. Math. Anal.* v. 40 (2005) n. 4, 43–56; MR2310110 (2007k:52015).

[46] E. Geciauskas, Recurrent approach to Blaschke's problem, *Geom. Dedicata* 121 (2006) 9–18; MR2276232 (2007k:52017).
[47] L. Heinrich, On lower bounds of second-order chord power integrals of convex discs, *Rend. Circ. Mat. Palermo, Suppl.* 81 (2009) 213–222; MR2809426 (2012d:60024).
[48] H. S. Harutyunyan and V. K. Ohanyan, The chord length distribution function for regular polygons, *Adv. Appl. Probab.* 41 (2009) 358–366; MR2541181 (2011b:60039).
[49] U. Bäsel, Random chords and point distances in regular polygons, *Acta Math. Univ. Comenian.* 83 (2014) 1–18; MR3178155.
[50] L. A. Santaló, Random lines and tessellations in a plane, *Stochastica* 4 (1980) 3–13; MR573721 (81e:60012).
[51] D. J. Daley, S. Ebert and R. J. Swift, Size distributions in random triangles, *J. Appl. Probab.* 51A (2014) 283–295; MR3317364.
[52] S. R. Finch, Triangles formed via Poisson nearest neighbors, arXiv:1712.08481.
[53] J. Gates, Shape distributions for planar triangles by dual construction, *Adv. Appl. Probab.* 26 (1994) 324–333; MR1272714 (95b:60012).
[54] S. R. Finch, Three random intercepts of a segment, arXiv:1802.02466.
[55] R. V. Ambartzumian, Random shapes by factorization, *Statistics in Theory and Practise, Essays in Honour of Bertil Matérn*, ed. B. Ranneby, Swedish University of Agricultural Sciences, Section of Forest Biometry, 1982, pp. 35-41; MR0688997 (84c:62004).
[56] S. R. Finch, 0-pierced triangles within a Poisson overlay, arXiv:1804.01353.
[57] A. Baddeley, E. Rubak and R. Turner, *Spatial Point Patterns: Methodology and Applications with R*, Chapman and Hall/CRC Press, 2015.

5.15 Random Triangles. VI

As a conclusion of our survey, we gather various results for random triangles in the plane subject to constraints. If we break a line segment L in two places at random, the three pieces can be configured as a triangle with probability $1/4$ [1–4]. If we instead select three points on a circle Γ at random, a triangle can almost surely be formed by connecting each pair of points with a line. Assuming L has length 1 and Γ has radius 1, what can be said about sides and angles of such triangles?

5.15.1 *Unit Perimeter*

Consider the broken L model, with the condition that triangle inequalities are satisfied. The bivariate density for two arbitrary sides a, b is [5, 6]

$$\begin{cases} 8 & \text{if } 0<x<1/2,\ 0<y<1/2 \text{ and } x+y>1/2, \\ 0 & \text{otherwise.} \end{cases}$$

Integrating on y from $1/2-x$ to $1/2$, the univariate density for a is

$$\begin{cases} 8\,x & \text{if } 0<x<1/2, \\ 0 & \text{otherwise} \end{cases}$$

and corresponding moments are

$$\mathrm{E}(a) = 1/3 = 0.3333333333\ldots, \quad \mathrm{E}(a^2) = 1/8 = 0.125.$$

As in [7], the cross-correlation coefficient $\rho(a, b) = -1/2$, hence

$$\mathrm{E}(a\,b) = 5/48 = 0.1041666666\ldots.$$

The Law of Cosines (with third side $c = 1 - a - b$) and a Jacobian determinant calculation imply that the bivariate density for two angles α, β is

$$\begin{cases} 8\dfrac{\sin(x)\sin(y)\sin(x+y)}{(\sin(x)+\sin(y)+\sin(x+y))^3} & \text{if } 0 < x < \pi,\ 0 < y < \pi \text{ and } x + y < \pi, \\ 0 & \text{otherwise.} \end{cases}$$

This is a new result, as far as is known, although it bears resemblance to formulas in [7]. Integrating on y from 0 to $\pi - x$, the univariate density for α is

$$\begin{cases} -8\dfrac{(3-\cos(x))\sin(x)}{(1+\cos(x))^3}\ln\left(\sin\left(\dfrac{x}{2}\right)\right) - 8\dfrac{\sin(x)}{(1+\cos(x))^2} & \text{if } 0 < x < \pi, \\ 0 & \text{otherwise} \end{cases}$$

and corresponding moments are

$$\mathrm{E}(\alpha) = \pi/3 = 1.0471975511\ldots, \quad \mathrm{E}(\alpha^2) = 8/3 - \pi^2/9 = 1.5700439554\ldots.$$

Because $\rho(\alpha, \beta) = -1/2$, we have

$$\mathrm{E}(\alpha\,\beta) = -4/3 + 2\pi^2/9 = 0.8599120891\ldots.$$

It is feasible to calculate the density for the maximum angle (omitted). The probability that a broken L triangle is obtuse can be shown to be [8–10]

$$9 - 12\ln(2) = 0.6822338332\ldots = 1 - 0.3177661667\ldots.$$

Let $h(z) < 0 < f(z) < w < g(z) < 1$ be the three zeroes of the cubic polynomial $(1-w)w^2 - 64z^2$. For area $z = \sqrt{(1/2)(1/2-a)(1/2-b)(a+b-1/2)}$, the density is [11]

$$\frac{256z}{\sqrt{(1-f)(g-h)}}K\left(\sqrt{\frac{(g-f)(1-h)}{(1-f)(g-h)}}\right), \quad 0 < z < \frac{1}{12\sqrt{3}}$$

where K is the complete elliptic integral of the first kind [7]. This again is a new result, but the moments [6]

$$\mathrm{E}(\text{area}) = \frac{\pi}{105} = 0.0299199300\ldots, \quad \mathrm{E}(\text{area}^2) = \frac{1}{960} = 0.0010416666\ldots$$

are well-known. A similar set of side/angle computations for unit area triangles (à la “throwing paint”) is attempted in [12].

5.15.2 *Unit Circumradius*

Consider the selection Γ model, equivalently, all triangles inscribing the unit circle. The bivariate density for two arbitrary angles α, β is [13–15]

$$\begin{cases} 2/\pi^2 & \text{if } 0 < x < \pi,\, 0 < y < \pi \text{ and } x + y < \pi, \\ 0 & \text{otherwise.} \end{cases}$$

To prove this, use the fact that an inscribed angle is one-half the length of its intercepted circular arc [16, 17]. Integrating on y from 0 to $\pi - x$, the univariate density for α is

$$\begin{cases} 2(\pi - x)/\pi^2 & \text{if } 0 < x < \pi, \\ 0 & \text{otherwise} \end{cases}$$

and corresponding moments are

$$\mathrm{E}(\alpha) = \pi/3 = 1.0471975511\ldots, \quad \mathrm{E}(\alpha^2) = \pi^2/6 = 1.6449340668\ldots.$$

As before, the cross-correlation coefficient $\rho(\alpha, \beta) = -1/2$, hence

$$\mathrm{E}(\alpha\,\beta) = \pi^2/12 = 0.8224670334\ldots.$$

The angle α is maximum if $\alpha > \beta$ and $\alpha > \pi - \alpha - \beta$ [7]. Hence the density for the maximum angle is

$$\begin{cases} 3\displaystyle\int_{\pi-2x}^{x} 2/\pi^2\, dy & \text{if } \pi/3 < x < \pi/2, \\ 3\displaystyle\int_{0}^{\pi-x} 2/\pi^2\, dy & \text{if } \pi/2 < x < \pi \end{cases} = \begin{cases} 6(3x - \pi)/\pi^2 & \text{if } \pi/3 < x < \pi/2, \\ 6(\pi - x)/\pi^2 & \text{if } \pi/2 < x < \pi \end{cases}$$

and the probability that a selection Γ triangle is obtuse [8, 9, 15] is $3/4 = 0.75$.

The univariate density for a is [18, 19]

$$\begin{cases} \dfrac{2}{\pi}\dfrac{1}{\sqrt{4 - x^2}} & \text{if } 0 < x < 2, \\ 0 & \text{otherwise} \end{cases}$$

and corresponding moments are

$$\mathrm{E}(a) = 4/\pi = 1.2732395447\ldots, \quad \mathrm{E}(a^2) = 2.$$

It can be shown that sides a, b are independent, which is delightfully paradoxical since angles α, β are *dependent* and

$$a = 2\sin(\alpha), \quad b = 2\sin(\beta).$$

The remaining side c satisfies

$$c = \begin{cases} \frac{1}{2}\left(a\sqrt{4-b^2} + b\sqrt{4-a^2}\right) & \text{with probability } 1/2, \\ \frac{1}{2}\left|a\sqrt{4-b^2} - b\sqrt{4-a^2}\right| & \text{with probability } 1/2 \end{cases}$$

but a simple expression for the trivariate density for all three sides a, b, c seems unlikely.

For area $z = (1/4)\sqrt{(a+b+c)(-a+b+c)(a-b+c)(a+b-c)}$, the density is $8z\Psi\left(4z^2\right)$, where

$$\Psi(y) = \frac{1}{4\pi^3}\frac{1}{\sqrt{y}}\left\{\Gamma\left(\frac{1}{3}\right)^3\left(\frac{4y}{27}\right)^{-1/6} {}_2F_1\left(\frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{4y}{27}\right)\right.$$
$$\left. - 3\Gamma\left(\frac{2}{3}\right)^3\left(\frac{4y}{27}\right)^{1/6} {}_2F_1\left(\frac{2}{3}, \frac{2}{3}, \frac{4}{3}, \frac{4y}{27}\right)\right\},$$

${}_2F_1$ is the Gauss hypergeometric function [20] and $0 < y < 27/4$. This formula corrects that which appears in Case III of [21]. The moments [22–24]:

$$\mathrm{E}(\text{area}) = \frac{3}{2\pi} = 0.4774648292..., \quad \mathrm{E}(\text{area}^2) = \frac{3}{8} = 0.375$$

are well-known. We mention that analogous results for random tetrahedra inscribing the unit sphere [23, 25, 26] are $\mathrm{E}(\text{volume}) = 4\pi/105 \approx 0.11968$ and $\mathrm{E}(\text{volume}^2) = 2/81 \approx 0.02469$. See [27] for a related coverage probability issue.

A study of triangles circumscribing the unit circle Γ was undertaken in [28]. On the one hand, the bivariate density for angles in the unit inradius scenario is the same as that in the unit circumradius scenario. On the other hand, a side has infinite mean and a more complicated density.

5.15.3 *Side-Angle-Side Example*

Thus far we have examined cases when three sides are given or three angles are given. Portnoy [29] studied an example in which two sides $a = \cos(\theta)$, $b = \sin(\theta)$ are given, where θ is Uniform $[0, \pi/2]$, as well as the included angle γ, which is independent and Uniform $[0, \pi]$. Let us focus solely on the obtuseness probability. By the Law of Cosines,

$$b^2 = a^2 + c^2 - 2a\,c\cos(\beta),$$

$$c^2 = a^2 + b^2 - 2a\,b\cos(\gamma).$$

If $\beta \geq \pi/2$, then $\cos(\beta) \leq 0$ and $b^2 \geq a^2 + c^2$, hence

$$b^2 - a^2 \geq c^2 = a^2 + b^2 - 2a\,b\cos(\gamma)$$

hence

$$2a\,b\cos(\gamma) \geq 2\,a^2$$

hence

$$\cos(\gamma) \geq a/b = \cot(\theta)$$

and conversely. The probability that $\beta \geq \pi/2$ is thus

$$\mathrm{P}\{\cos(\gamma) - \cot(\theta) \geq 0\} = 1 - \mathrm{P}\{\cos(\gamma) + \cot(\theta) \geq 0\}$$

by symmetry, and the latter probability (of a sum) is a convolution integral:

$$\frac{2}{\pi^2} \int_0^{\infty} \int_{\xi(x)}^{x+1} \frac{1}{\sqrt{1-(x-y)^2}} \frac{1}{1+y^2} dy\, dx$$

where $\xi(x) = \max\{x-1, 0\}$. Reversing the order of integration, we obtain

$$\frac{3}{4} + \frac{1}{\pi^2} \ln\left(1+\sqrt{2}\right)^2 = 1 - 0.1712917389...$$

as the value of the integral. Finally, the obtuseness probability for the triangle is

$$\mathrm{P}\{\theta \geq \pi/2\} + \mathrm{P}\{\alpha \geq \pi/2\} + \mathrm{P}\{\beta \geq \pi/2\}$$

which becomes

$$1 - \frac{2}{\pi^2} \ln\left(1+\sqrt{2}\right)^2 = 0.8425834778....$$

This exact evaluation is new, as far as is known, improving on [29]. See also [30].

Random convex quadrilaterals inscribing the unit circle Γ behave differently than random triangles. Any two sides are negatively correlated (rather than independent). Any two adjacent angles are uncorrelated yet dependent (rather than negatively correlated). Formulas like $\mathrm{E}(\text{area}) = 3/\pi$ and $\mathrm{E}(\text{area}^2) = 1/2 + 105/(16\pi^2)$ are merely conjectured, not yet proved [31]. By contrast, when breaking a unit line segment L in three places at random, a quadrilateral is formed with probability 1/2; formulas like $\mathrm{E}(\text{area}) = 17\pi/525 - \pi^2/160$ and $\mathrm{E}(\text{area}^2) = 1/560$ are demonstrably true if, further, the four vertices lie on a common circle [11].

[1] E. Seneta and F. Jongmans, The problem of the broken rod and Ernesto Cesàro's early work in probability, *Math. Sci.* 30 (2005) 67–76; MR2185363 (2006h:60006).

[2] G. S. Goodman, The problem of the broken stick reconsidered, *Math. Intelligencer* 30 (2008) 43–49; MR2437199 (2010a:60026).

[3] J. J. Hunter, Random triangles, order statistics and singular distributions, *Internat. J. Math. Ed. Sci. Tech.* 17 (1986) 253–257; MR0827358.

[4] J. Whittaker, Random triangles, *Amer. Math. Monthly* 97 (1990) 228–230; MR1048437 (91i:60038).

[5] R. Pyke, Spacings, *J. Royal Statist. Soc. Ser. B* 27 (1965) 395–449; MR0216622 (35 #7451).

[6] D. E. Dobbs, The average area of a triangle, *Mathematics and Computer Education* 21 (1987) 178–181.

[7] S. R. Finch, Random triangles, *this volume*, §5.10.

[8] F. Budden, Some connected problems on probability, *Math. Gazette*, v. 66 (1982) n. 435, 4–10.

[9] R. K. Guy, There are three times as many obtuse-angled triangles as there are acute-angled ones, *Math. Mag.* 66 (1993) 175–179; MR1572957.
[10] D. E. Dobbs and J. L Zarestky, Acute random triangles, *Nieuw Arch. Wisk.* (4) 15 (1997) 141–162; MR1609502.
[11] S. R. Finch, Median area for broken sticks, arXiv:1804.09602.
[12] S. R. Finch, Lognormality and triangles of unit area, arXiv:1412.5860.
[13] R. E. Miles, The various aggregates of random polygons determined by random lines in a plane, *Adv. Math.* 10 (1973) 256–290; MR0319232 (47 #7777).
[14] J. S. Rao, Some tests based on arc-lengths for the circle, *Sankhya Ser. B* 38 (1976) 329–338; MR0652731 (58 #31571).
[15] T. Moore, Random triangle problem (long summary), USENET sci.math newsgroup posting, 1997.
[16] E. E. Moise, *Elementary Geometry from an Advanced Standpoint*, Addison-Wesley, 1963, pp. 192–196; MR0149339 (26 #6829).
[17] S. Lang and G. Murrow, *Geometry: A High School Course*, Springer-Verlag, 1983, pp. 192–197.
[18] S. S. Chiu and R. C. Larson, Bertrand's paradox revisited: more lessons about that ambiguous word, random, *J. Industrial and Systems Engineering* 3 (2009) 1–26.
[19] E. W. Weisstein, Circle Line Picking, Wolfram MathWorld posting.
[20] S. R. Finch, Notation: Hypergeometric function, *this volume*, p. xi.
[21] A. M. Mathai and D. S. Tracy, On a random convex hull in an n-ball, *Comm. Statist. A – Theory Methods* 12 (1983) 1727–1736; MR0704849 (85c:60013).
[22] J. Edwards, *A Treatise on the Integral Calculus with Applications, Examples and Problems*, v. 2, MacMillan and Co., 1922, p. 772.
[23] R. E. Miles, Isotropic random simplices, *Adv. Appl. Probab.* 3 (1971) 353–382; MR0309164 (46 #8274).
[24] E. W. Weisstein, Circle Triangle Picking, Wolfram MathWorld posting.
[25] L. Heinrich, R. Körner, N. Mehlhorn and L. Muche, Numerical and analytical computation of some second-order characteristics of spatial Poisson-Voronoi tessellations, *Statistics* 31 (1998) 235–259; MR1718316 (2000g:60016).
[26] E. W. Weisstein, Sphere Tetrahedron Picking, Wolfram MathWorld posting.
[27] S. R. Finch, Covering a sphere with four random circular caps, arXiv:1502.06931.
[28] S. R. Finch, Three random tangents to a circle, arXiv:1101.3931.
[29] S. Portnoy, A Lewis Carroll pillow problem: probability of an obtuse triangle, *Statist. Sci.* 9 (1994) 279–284; MR1293297 (95h:60003).
[30] S. R. Finch, Uniform triangles with equality constraints, arXiv:1411.5216.
[31] S. R. Finch, Random cyclic quadrilaterals, arXiv:1610.00510.

5.16 Colliding Dice Probabilities

Let K, L be congruent regular polyhedra in $\mathbb{R}^3$. Let g denote a rigid motion of $\mathbb{R}^3$, that is, $g(x) = \Phi x + \tau$ where Φ is a 3×3 rotation matrix and τ is a translation 3-vector. The polyhedra K, $g(L)$ are said to **touch** if $K \cap g(L) \neq \varnothing$ but $\mathrm{int}(K) \cap \mathrm{int}(g(L)) = \varnothing$. Alternatively, we may think of ΦL moving toward K in the direction τ, stopping precisely when the two polyhedra collide.

Let us sample the space SO_3 of matrices Φ according to the uniform distribution (Haar measure, normalized to 1). The space of vectors τ is slightly harder to describe. Let

$$K-\Phi L=\{y-\Phi x: y\in K \text{ and } x\in L\}$$

be the Minkowski sum of K and the reflected image $-\Phi L$ of ΦL. Another way to characterize $K-\Phi L$ is as the convex hull of all pairwise sums of vertices of K and $-\Phi L$. Clearly

$$\tau\in\mathrm{bd}(K-\Phi L) \quad \text{if and only if} \quad \text{the polyhedra } K, g(L) \text{ touch.}$$

Thus we sample the space $\mathrm{bd}(K-\Phi L)$ uniformly (area measure), which is complicated only by the intricate variety of possible faces of $K-\Phi L$.

With independent Φ and τ as described, it is clear that

$$\mathrm{P}\{\text{collision is edge-to-edge}\}>0,$$

$$\mathrm{P}\{\text{collision is vertex-to-face or face-to-vertex}\}>0$$

and that no other types of collisions occur with positive likelihood. What is unclear is the relative magnitude of these two probabilities.

Answering a question asked by Firey, McMullen [1, 2] proved that the edge-to-edge collisions are strictly more likely than vertex-to-face collisions. In the case of two cubes (cubical dice), the exact values of the probabilities are

$$\frac{3\pi}{3\pi+8}=0.5408836762...>0.4591163237...=\frac{8}{3\pi+8}.$$

More generally, we have [3]

$$\frac{\pi V_1^2}{8V_0V_2+\pi V_1^2}>\frac{8V_0V_2}{8V_0V_2+\pi V_1^2}$$

where $V_0=1$ is the Euler characteristic of K, $\frac{1}{2}V_1$ is the mean width b (to be defined shortly), $2V_2$ is the surface area a and V_3 is the volume. For the unit cube, it follows that $b=3/2$ and $a=6$.

In the case of two regular tetrahedra (tetrahedral dice), we have

$$b=\tfrac{3}{2\pi}\arccos\left(-\tfrac{1}{3}\right), \quad a=\sqrt{3}$$

and hence

$$\begin{aligned}\frac{9\arccos\left(-\frac{1}{3}\right)^2}{4\sqrt{3}\pi+9\arccos\left(-\frac{1}{3}\right)^2} &= 0.6015106899...>0.3984893100...\\ &= \frac{4\sqrt{3}\pi}{4\sqrt{3}\pi+9\arccos\left(-\frac{1}{3}\right)^2}.\end{aligned}$$

In the case of two regular octahedra (octahedral dice), we have

$$b=\tfrac{3}{\pi}\arccos\left(\tfrac{1}{3}\right), \quad a=2\sqrt{3}$$

and hence

$$\frac{9\arccos\left(\frac{1}{3}\right)^2}{2\sqrt{3}\pi + 9\arccos\left(\frac{1}{3}\right)^2} = 0.5561691925\ldots > 0.4438308074\ldots$$
$$= \frac{2\sqrt{3}\pi}{2\sqrt{3}\pi + 9\arccos\left(\frac{1}{3}\right)^2}.$$

These specific numerical results are apparently new. For tetrahedra, verification by simulation is done using [4, 5]. The touching is vertex-to-face or face-to-vertex if and only if τ lies in a triangular face of $K - \Phi L$. (All other faces of $K - \Phi L$ are parallelograms.) Hence it suffices to assess the ratio of surface area of triangles only to surface area of the whole. The cases of two cubes or of two octahedra are more difficult.

5.16.1 *Mean Width*

Let C be a convex body in $\mathbb{R}^3$. In earlier essays [6–8], the words "width" or "breadth" were used to denote the *minimum* distance between all pairs of parallel C-supporting planes. Here, we instead take the *mean* of all such distances, calling this b. The phrase **mean width** [9, 10] is used, as well as **mean breadth** [11] and **mean caliper diameter** [12, 13].

Closed-form expressions for b exist when C is a convex polyhedron. Numerical confirmation of such formulas is possible via quadratic programming (since the optimization constraints are linear).

5.16.2 *Intrinsic Volumes*

Let P be a rectangular parallelepiped in $\mathbb{R}^3$ of dimensions z_1, z_2, z_3. It is well-known that

$$V_3(P) = z_1 z_2 z_3,$$
$$V_2(P) = z_1 z_2 + z_1 z_3 + z_2 z_3 = \tfrac{1}{2}a,$$
$$V_1(P) = z_1 + z_2 + z_3 = 2b$$

are the elementary symmetric polynomials in three variables. In $\mathbb{R}^n$, there are n such intrinsic volumes, corresponding to the n elementary symmetric polynomials [10]. Little is known about higher-dimensional intrinsic volumes and the isoperimetric inequalities among them. Limiting approximation arguments enable us to compute $V_j(C)$ for arbitrary convex C. Additionally, let $V_0(C) = 1$. Hadwiger's famous theorem [3] gives that V_0, V_1, …, V_n are a basis of the space of all additive continuous measures that are invariant under rigid motions.

Acknowledgment Rolf Schneider generously proposed the method underlying the tetrahedral simulation. More about mean width computations for convex polyhedra is found in [14–19], for certain other convex bodies in [20–23], and a specific non-convex body in [24].

[1] P. McMullen, A dice probability problem, *Mathematika* 21 (1974) 193–198; MR0362049 (50 #14491).
[2] P. M. Gruber, Seven small pearls from convexity, *Math. Intelligencer* 5 (1983) 16–19; MR0737676 (85h:52001).
[3] R. Schneider and W. Weil, *Stochastic and Integral Geometry*, Springer-Verlag, 2008, pp. 349–359, 599–602, 628–630; MR2455326 (2010g:60002).
[4] J. Arvo, Fast random rotation matrices, *Graphics Gems III*, ed. D. Kirk, Academic Press, 1992, pp. 117–120.
[5] C. B. Barber, D. P. Dobkin, and H. T. Huhdanpaa, The Quickhull algorithm for convex hulls, *ACM Trans. Math. Software* 22 (1996) 469–483; MR1428265 (97g:65292).
[6] S. R. Finch, Moser's worm constant: Broadest curve of unit length, *first volume*, p. 493.
[7] S. R. Finch, Reuleaux triangle constants, *first volume*, pp. 513–515.
[8] S. R. Finch, Circumradius-inradius constants, *first volume*, pp. 535–537.
[9] M. Månsson and M. Rudemo, Random patterns of nonoverlapping convex grains, *Adv. Appl. Probab.* 34 (2002) 718–738; MR1938939 (2003j:60014).
[10] G.-C. Rota, Mathematical snapshots, unpublished note (1997).
[11] D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, Wiley, 1987, pp. 11–19; MR0895588 (88j:60034a).
[12] J. E. Hilliard, The calculation of the mean caliper diameter of a body for use in the analysis of the number of particles per unit volume, *Stereology*, Proc. 1967 Chicago conference, ed. H. Elias, Springer-Verlag, 1967, pp. 211–215.
[13] L. A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley, 1976, pp. 226–230; MR0433364 (55 #6340).
[14] S. R. Finch, Width distributions for convex regular polyhedra, arXiv:1110.0671.
[15] S. R. Finch, Mean width of a regular simplex, arXiv:1111.4976.
[16] S. R. Finch, Mean width of a regular cross-polytope, arXiv:1112.0499.
[17] S. R. Finch, Planar projections and second intrinsic volume, arXiv:1203.2464.
[18] S. R. Finch, Rank-3 projections of a 4-cube, arXiv:1204.3468.
[19] Z. Kabluchko, A. E. Litvak and D. Zaporozhets, Mean width of regular polytopes and expected maxima of correlated Gaussian variables (in Russian), *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 442 (2015) 75–96; Engl. transl. in *J. Math. Sci.* 225 (2017) 770–787; arXiv:1511.08479; MR3506845.
[20] S. R. Finch, Convex hull of two orthogonal disks, arXiv:1211.4514.
[21] S. R. Finch, Oblique circular cones and cylinders, arXiv:1212.5946.
[22] S. R. Finch, Mutually equidistant spheres that intersect, arXiv:1301.5515.
[23] U. Bäsel, The mean width of the oloid and integral geometric applications of it, arXiv:1604.07245.
[24] S. R. Finch, Problem 11687 (Slicing a doughnut at random), *Amer. Math. Monthly* 120 (2013) 77; 122 (2015) 177–178.

5.17 Gergonne–Schwarz Surface

We mentioned Plateau's problem in [1] but did not give a nontrivial example. Let

$$F[\phi, m] = \int_0^{\sin(\phi)} \frac{d\tau}{\sqrt{1-\tau^2}\sqrt{1-m\tau^2}}$$

denote the incomplete elliptic integral of the first kind and $K[m]=F[\pi/2,m]$; the latter is admittedly incompatible with [2] but we purposefully choose formulas here to be consistent with the computer algebra package Mathematica. The three basic Jacobi elliptic functions are defined via

$$u = \int_0^{\mathrm{sn}(u,m)} \frac{d\tau}{\sqrt{1-\tau^2}\sqrt{1-m\tau^2}} = \int_{\mathrm{cn}(u,m)}^{1} \frac{d\tau}{\sqrt{1-\tau^2}\sqrt{m\tau^2+(1-m)}}$$
$$= \int_{\mathrm{dn}(u,m)}^{1} \frac{d\tau}{\sqrt{1-\tau^2}\sqrt{\tau^2-(1-m)}}$$

and two (of nine) others we require are

$$\mathrm{sc}(u,m)=\frac{\mathrm{sn}(u,m)}{\mathrm{cn}(u,m)}, \quad \mathrm{sd}(u,m)=\frac{\mathrm{sn}(u,m)}{\mathrm{dn}(u,m)}.$$

Our work supplements [3] very closely, even down to the level of notation. The setting is three-dimensional xyz-space.

5.17.1 *Six Edges of a Cube*

Consider a polygonal wire loop with six line segments:

$$(0,0,0)\to(1,0,0)\to(1,0,1)\to(1,1,1)\to(0,1,1)\to(0,1,0)\to(0,0,0).$$

What is the minimal area for any surface spanning this fixed boundary? Equivalently, what is the outcome of dipping the wire loop in a soap solution?

Define

$$\rho_0 = K[1/4] = 1.6857503548...$$

and let $t=\mathcal{E}(\xi)$ denote the functional inverse of the elliptic integral

$$\xi = \int_0^t \frac{d\tau}{\sqrt{1+\tau^2+\tau^4}}.$$

The desired minimal surface is given implicitly by the equation [3]

$$\mathcal{E}(x)\mathcal{E}(y)=\mathcal{E}(z)$$

where $0\le x,y,z\le\rho_0$.

This is as far as Nitsche [3] went in describing his calculations. Solving for z and rescaling (so that the surface spans the $1\times1\times1$ cube), we find that

$$z=\frac{1}{2\rho_0}F\left[\arccos\left(\frac{\mathrm{cn}\left(2\rho_0 x,\frac{1}{4}\right)+\mathrm{cn}\left(2\rho_0 y,\frac{1}{4}\right)}{1+\mathrm{cn}\left(2\rho_0 x,\frac{1}{4}\right)\mathrm{cn}\left(2\rho_0 y,\frac{1}{4}\right)}\right),\frac{1}{4}\right], \quad 0\le x,y\le 1$$

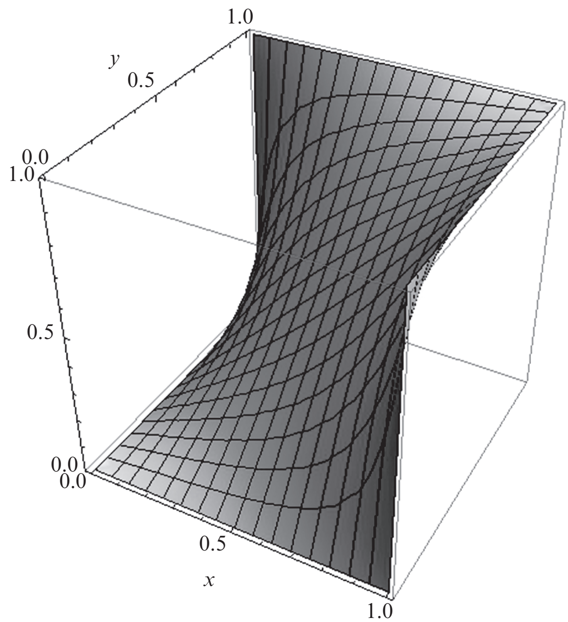

Figure 5.11 "Six edges" minimal surface

and the surface area is

$$2\int_0^1\int_0^{1-x}\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\,dy\,dx=\frac{3}{2}\frac{K[3/4]}{K[1/4]}=1.9188923567...,$$

as predicted in [4]. See Figure 5.11.

5.17.2 *Four Edges of a Regular Tetrahedron*

Consider a polygonal wire loop with four line segments:

$$(0,0,0)\to(1,0,1)\to(1,1,0)\to(0,1,1)\to(0,0,0).$$

Again, what is the minimal area for any surface spanning this fixed boundary?

With ρ_0 as before, let $s=\mathcal{F}(\eta)$ denote the functional inverse of the elliptic integral

$$\eta=\int_0^s\frac{d\sigma}{\sqrt{\frac{3}{4}+\frac{5}{2}\sigma^2+\frac{3}{4}\sigma^4}}.$$

The desired minimal surface is given implicitly by the equation [3]

$$\mathcal{F}(y)\mathcal{F}(z)+\mathcal{F}(z)\mathcal{F}(x)+\mathcal{F}(x)\mathcal{F}(y)+1=0$$

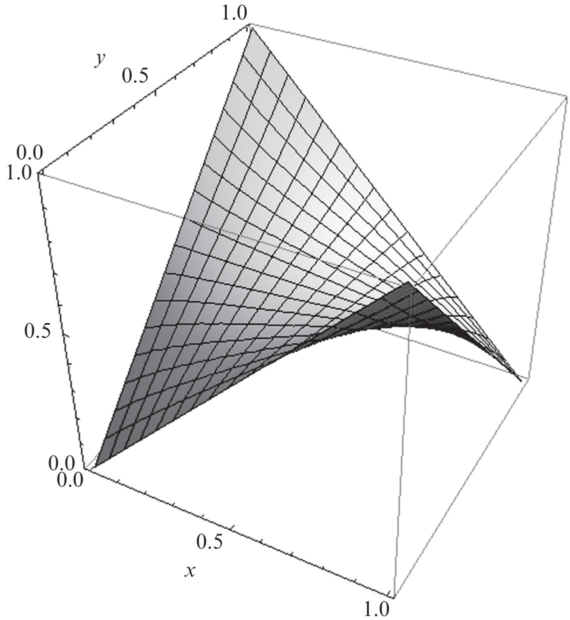

Figure 5.12 Tetrahedral "four edges" minimal surface

where $0 \le x, y \le \rho_0$ and $-\rho_0 \le z \le 0$. Dalpe [5] introduced one correction in the preceding: the cube has side ρ_0, not $2\rho_0$.

This is as far as described in [3]. Solving for z and rescaling (so that the surface spans the $1 \times 1 \times 1$ cube), we find that

$$z = \frac{1}{\sqrt{3}\rho_0} F\left[\arccos\left(\frac{\operatorname{cn}\left(\sqrt{3}\rho_0 x, -\frac{1}{3}\right)\operatorname{cn}\left(\sqrt{3}\rho_0 y, -\frac{1}{3}\right)}{1 + \operatorname{sn}\left(\sqrt{3}\rho_0 x, -\frac{1}{3}\right)\operatorname{sn}\left(\sqrt{3}\rho_0 y, -\frac{1}{3}\right)}\right), -\frac{1}{3}\right], \; 0 \le x, y \le 1$$

(note multiplication in the numerator and sn in the denominator, unlike before) and the surface area is

$$2\int_0^1 \int_0^{1-x} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dy\, dx = \frac{K[3/4]}{K[1/4]} = 1.2792615711...,$$

as predicted in [4]. See Figure 5.12. This example and the first one feature portions of what is known as the *Schwarz D surface* (D stands for "Diamond").

5.17.3 *Two Diagonals and Free Boundaries*

Consider the soap film (resembling a twisted curtain) formed between two skew line segments:

$$(2, 0, 0) \to (0, 2, 0) \quad \text{and} \quad (0, 0, 2) \to (2, 2, 2).$$

Understanding that two remaining boundaries are unspecified, what is the minimal area for any surface spanning the diagonals? [6] This is a famous question due to Gergonne (1816) and answered by Schwarz (1872).

For fixed $\kappa > 0$, let $t = Q(\varphi, \kappa)$ and $t = R(\psi, \kappa)$ denote functional inverses of the elliptic integrals

$$\varphi = \int_0^t \frac{d\tau}{\sqrt{\kappa - \tau^2 - \tau^4}}, \quad \psi = \int_0^t \frac{d\tau}{\sqrt{\kappa + (1+2\kappa)\tau^2 + \kappa\tau^4}}.$$

Define also

$$\lambda(\kappa) = \frac{\sqrt{1+4\kappa} - 1}{2\sqrt{1+4\kappa}}, \quad \mu(\kappa) = \sqrt{\frac{\sqrt{1+4\kappa} - 1}{2}}.$$

We have, in particular,

$$\int_0^{\mu(\kappa)} \frac{d\tau}{\sqrt{\kappa - \tau^2 - \tau^4}} = \frac{K[\lambda(\kappa)]}{(1+4\kappa)^{1/4}},$$

$$\int_0^1 \frac{d\tau}{\sqrt{\kappa + (1+2\kappa)\tau^2 + \kappa\tau^4}} = \frac{K\left[-\frac{1}{4\kappa}\right]}{2\sqrt{\kappa}}$$

and these two expressions, when set equal, force $\kappa = \kappa_0 = 0.2092861374....$ Denote the former integral by φ_0 and latter by ψ_0; consequently $\varphi_0 = \psi_0 = 1.3970394887....$ The desired minimal surface is given implicitly by the equation [3]

$$Q(x - \varphi_0)R(z - \psi_0) + Q(y - \varphi_0) = 0$$

where $0 \le x, y \le 2\varphi_0$ and $0 \le z \le 2\psi_0$. We have introduced two corrections in the preceding: the upper integration limit of ψ_0 is 1 (not $\mu(\kappa)$, which was a typographical error in [3]) and the denominator underlying $K\left[-\frac{1}{4\kappa}\right]$ is $2\sqrt{\kappa}$ (not merely 2, which was a computational error in [3]). More on the second correction will be mentioned shortly.

This, again, is as far as described in [3]. Let

$$\theta_0 = (1+4\kappa_0)^{1/4}\varphi_0, \quad \lambda_0 = \lambda(\kappa_0), \quad \varepsilon(x,y) = \begin{cases} 1 & \text{if } (x-1)(y-1) > 0, \\ -1 & \text{otherwise.} \end{cases}$$

Solving for z and rescaling (so that the surface spans the $2 \times 2 \times 2$ cube), we find that

$$z = 1 + \frac{\varepsilon(x,y)}{2\sqrt{\kappa}\,\psi_0} F\left[\arccos\left(\frac{\mathrm{sd}\,(\theta_0(x-1),\lambda_0)^2 - \mathrm{sd}\,(\theta_0(y-1),\lambda_0)^2}{\mathrm{sd}\,(\theta_0(x-1),\lambda_0)^2 + \mathrm{sd}\,(\theta_0(y-1),\lambda_0)^2}\right), -\frac{1}{4\kappa_0}\right]$$

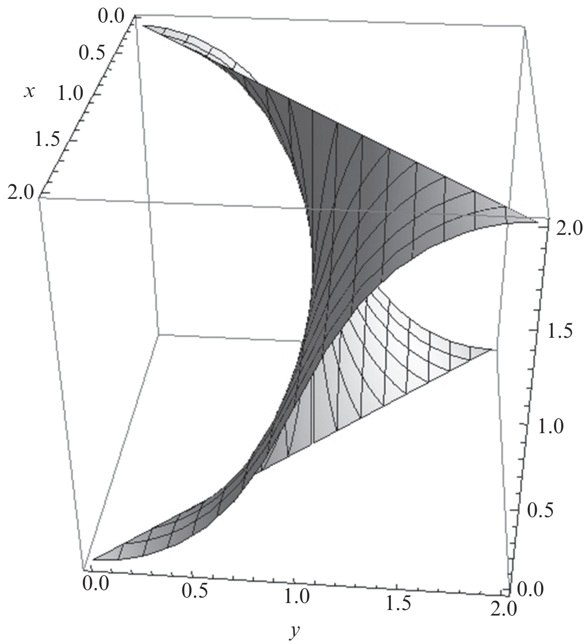

Figure 5.13 “Two diagonals” minimal surface

assuming $(y > x$ and $x < 2 - y)$ or $(y < x$ and $x > 2 - y)$; elsewhere on $0 \leq x, y \leq 2$, no definition for z is given. The surface area is

$$4 \int_0^1 \int_0^{1-x} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dy\, dx = 4.9348196582... = 4\,(1.2337049145...)$$

and a closed-form expression remains open. See Figure 5.13. We have not attempted to establish consistency with [7].

5.17.4 *Details of Elliptic Functions*

We can compute $\mathcal{E}(\xi)$ and $\mathcal{F}(\eta)$ using results in [8]:

$$\xi = \int_0^t \frac{d\tau}{\sqrt{1 + \tau^2 + \tau^4}} = \frac{1}{2} F\left[\arccos\left(\frac{1 - t^2}{1 + t^2}\right), \frac{1}{4}\right],$$

$$\eta = \int_0^s \frac{d\sigma}{\sqrt{\frac{3}{4} + \frac{5}{2}\sigma^2 + \frac{3}{4}\sigma^4}} = \frac{1}{\sqrt{3}} F\left[\arccos\left(\frac{1 - s^2}{1 + s^2}\right), -\frac{1}{3}\right]$$

since each quartic has four imaginary zeroes; hence

$$t = \sqrt{\frac{1 - \mathrm{cn}\,(2\xi, 1/4)}{1 + \mathrm{cn}\,(2\xi, 1/4)}},$$

$$s = \sqrt{\frac{1 - \mathrm{cn}\left(\sqrt{3}\eta, -1/3\right)}{1 + \mathrm{cn}\left(\sqrt{3}\eta, -1/3\right)}}$$

and thus

$$z = \frac{1}{2}F\left[\arccos\left(\frac{1 - \mathcal{E}(x)^2\mathcal{E}(y)^2}{1 + \mathcal{E}(x)^2\mathcal{E}(y)^2}\right), \frac{1}{4}\right]$$

gives the "six edges" result. From

$$\mathcal{F}(z) = -\frac{1 + \mathcal{F}(x)\mathcal{F}(y)}{\mathcal{F}(x) + \mathcal{F}(y)}$$

we obtain

$$z = \frac{1}{\sqrt{3}}F\left[\arccos\left(\frac{1 - \left(\frac{1+\mathcal{F}(x)\mathcal{F}(y)}{\mathcal{F}(x)+\mathcal{F}(y)}\right)^2}{1 + \left(\frac{1+\mathcal{F}(x)\mathcal{F}(y)}{\mathcal{F}(x)+\mathcal{F}(y)}\right)^2}\right), -\frac{1}{3}\right]$$

and, because $\mathrm{sn}(u,m)^2 + \mathrm{cn}(u,m)^2 = 1$, the "four edges" result follows.

Computing $Q(\varphi, \kappa)$ is somewhat different [9]:

$$\begin{aligned}\varphi &= \int_0^t \frac{d\tau}{\sqrt{\kappa - \tau^2 - \tau^4}} \\ &= \frac{1}{(1+4\kappa)^{1/4}}\left\{K[\lambda(\kappa)] - F\left[\arcsin\left(\sqrt{\frac{\sqrt{1+4\kappa} - 2t^2 - 1}{\sqrt{1+4\kappa} - 1}}\right), \lambda(\kappa)\right]\right\}\end{aligned}$$

since the quartic has two real zeroes and two imaginary zeroes. Observe that, when $t = \mu(\kappa)$, the second term vanishes. Inverting, we obtain

$$t = \frac{\kappa}{(1+4\kappa)^{1/4}}\,\mathrm{sd}\left((1+4\kappa)^{1/4}\varphi, \lambda(\kappa)\right)$$

and therefore

$$-\frac{Q(y-\varphi_0, \kappa)}{Q(x-\varphi_0, \kappa)} = -\frac{\mathrm{sd}\left((1+4\kappa)^{1/4}(y-\varphi_0), \lambda(\kappa)\right)}{\mathrm{sd}\left((1+4\kappa)^{1/4}(x-\varphi_0), \lambda(\kappa)\right)}.$$

Only the inverse of $R(\psi, \kappa)$ is required:

$$\psi = \int_0^t \frac{d\tau}{\sqrt{\kappa + (1+2\kappa)\tau^2 + \kappa\tau^4}} = \frac{\mathrm{sign}(t)}{2\sqrt{\kappa}}F\left[\arccos\left(\frac{1-t^2}{1+t^2}\right), -\frac{1}{4\kappa}\right]$$

which generalizes the earlier cases $\kappa = -1$ and $\kappa = 3/4$. Note the specialization $t = 1$, as well as the need here to track whether $t = -Q(y - \varphi_0, \kappa)/Q(x - \varphi_0, \kappa)$ is positive or negative.

5.17.5 *Approximations of Minimal Surfaces*

A surprisingly good fit to the "four edges" surface is provided by the hyperbolic paraboloid

$$z = x + y - 2xy$$

and the corresponding surface area is

$$2\int_0^1 \int_0^{1-x} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dy\, dx = 1.2807... > 1.2792....$$

See [10] for more on approximating the Schwarz D surface, which (upon suitable transformation) should enable a reasonable fit to the "six edges" surface.

Fairly coarse fits to the "two diagonals" surface are provided by

$$z = 1 + \frac{y-1}{x-1}, \quad z = 1 + \frac{4}{\pi}\arctan\left(\frac{y-1}{x-1}\right)$$

if ($y > x$ and $x < 2 - y$) or ($y < x$ and $x > 2 - y$), and the corresponding surface areas are 5.1231... and 5.0307..., respectively. We mentioned earlier that Nitsche [3] mistakenly solved the equation

$$\frac{K[\lambda(\kappa)]}{(1 + 4\kappa)^{1/4}} = \frac{K\left[-\frac{1}{4\kappa}\right]}{2};$$

the denominator underlying $K\left[-\frac{1}{4\kappa}\right]$ is missing a factor $\sqrt{\kappa}$. It is nevertheless instructive to follow through to the end. We find $\kappa = \widetilde{\kappa}_0 = 6.6061877190...$ and consequently $\widetilde{\varphi}_0 = \widetilde{\psi}_0 = 0.7781217795....$ The surface obtained *is* a minimal surface (with mean curvature everywhere equal to zero) and correctly spans the diagonals. The two free contours, however, are not best possible: the surface area for $\widetilde{\kappa}_0$ is 4.9480..., which is larger than the surface area 4.9348... for κ_0.

The constant 1.9188... appears in [11, 12], 1.2792... in [13, 14] and a rough estimate for $\frac{1}{4}(4.9348...)$ in [15]. See [16, 17] for introductory materials, as well as Schwarz's complete works [18]. Other polygonal wire loops, with more solutions of Plateau's problem, are surveyed in [19].

Addendum Another portion of the Schwarz D surface arises as a soap film spanning two parallel equilateral triangles with vertices

$$\{(1, -1, -1), (-1, 1, -1), (-1, -1, 1)\} \quad \text{and} \quad \{(-1, 1, 1), (1, -1, 1), (1, 1, -1)\}.$$

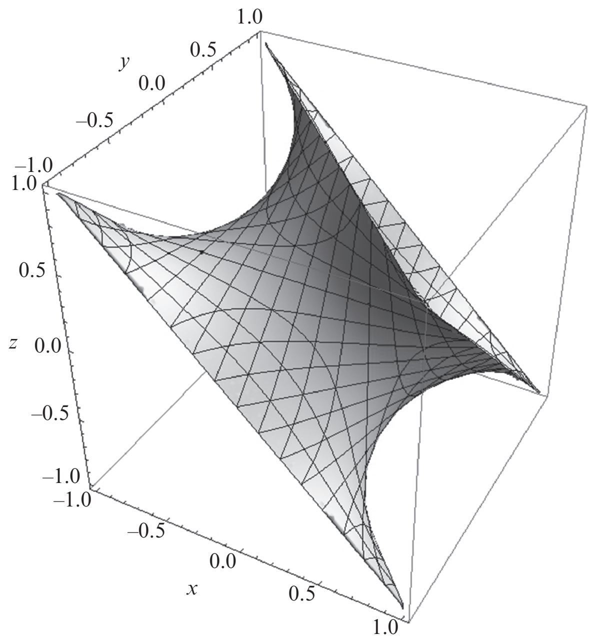

Figure 5.14 "Two twisted triangles" minimal surface

One triangle is a copy of the other, rotated 60° about its center. Each of the six edges has length $2\sqrt{2}$ and the perpendicular distance between triangular centers is $2/\sqrt{3}$; the ratio of these is $\sqrt{6}$. Define $\zeta_0 = K[8/9]$. The desired minimal annulus is given implicitly by [18, 20]

$$\mathrm{sc}(\zeta_0 y, \tfrac{8}{9})\,\mathrm{sc}(\zeta_0 z, \tfrac{8}{9}) + \mathrm{sc}(\zeta_0 z, \tfrac{8}{9})\,\mathrm{sc}(\zeta_0 x, \tfrac{8}{9}) + \mathrm{sc}(\zeta_0 x, \tfrac{8}{9})\,\mathrm{sc}(\zeta_0 y, \tfrac{8}{9}) + 3 = 0$$

where $-1 \leq x, y, z \leq 1$ and its surface area is $6K[3/4]/K[1/4]$. See Figure 5.14. (This result contradicts a statement in [21] that, for Schwarz D to appear, the ratio of edge length to distance should be $2\sqrt{3}$.)

A more difficult task is to represent the minimal annulus corresponding to parallel triangles that are aligned [22–26], that is, with no rotation. This is a member of the family of *Schwarz H surfaces* (H stands for "Hexagonal"). Determination of such representations, for a range of perpendicular distances between triangular centers, and associated numerical calculation of surface areas, is a worthy challenge.

Acknowledgments Ulrike Bücking was so kind as to point out two errors in [3]; I also appreciate correspondence with Djurdje Cvijović and Stefan Hildebrandt.

[1] S. R. Finch, Planar harmonic mappings, *this volume*, §3.4.
[2] S. R. Finch, Random triangles, *this volume*, §5.10.

[3] J. C. C. Nitsche, *Vorlesungen über Minimalflächen*, Springer-Verlag, 1975; revised Engl. transl., *Lectures on Minimal Surfaces*, v. I, Cambridge Univ. Press, 1989, pp. 68–88, 233–247, 349–358; MR0448224 (56 #6533) and MR1015936 (90m:49031).

[4] P. J. F. Gandy, D. Cvijović, A. L. Mackay and J. Klinowski, Exact computation of the triply periodic D ('diamond') minimal surface, *Chem. Phys. Lett.* 314 (1999) 543–551.

[5] D. Dalpe, *Schwarz's Surface and the Theory of Minimal Surfaces*, MS thesis, Concordia Univ., Montreal, 1998.

[6] U. Dierkes, S. Hildebrandt and A. J. Tromba, *Regularity of Minimal Surfaces*, 2nd ed., Springer-Verlag, 2010, pp. 45–46; MR2760441 (2012b:49003).

[7] D. Cvijović and J. Klinowski, The T and CLP families of triply periodic minimal surfaces. II, The properties and computation of T surfaces, *J. Physique I* 2 (1992) 2191–2205; MR1203044 (94a:53013).

[8] H. Hancock, *Elliptic Integrals*, John Wiley & Sons, Chapman & Hall, 1917, pp. 55–56.

[9] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists*, 2nd ed., Springer-Verlag, 1971, p. 50; MR0277773 (43 #3506).

[10] I. S. Barnes, A useful trigonometric approximation to periodic minimal surfaces?, *Aust. Math. Soc. Gaz.*, v. 17 (1990) n. 4, 99-105; Zbl 0714.53012.

[11] A. L. Mackay, Periodic minimal surfaces, *Nature* 314 (1985) 604–606.

[12] D. M. Anderson, H. T. Davis, J. C. C. Nitsche and L. E. Scriven, Periodic surfaces of prescribed mean curvature, *Physics of Amphiphilic Layers*, Proc. Les Houches workshop, 1987, ed. J. Meunier, D. Langevin and N. Boccara, Springer-Verlag, 1987, p. 130 (abstract only); *Advances in Chemical Physics*, v. 77, ed. I. Prigogine and S. A. Rice, Wiley, 1990, pp. 337–396.

[13] S. Furui and B. Masud, On the colour confinement and the minimal surface, arXiv:hep-lat/0006003.

[14] S. Furui and B. Masud, Numerical calculation of a minimal surface using bilinear interpolations and Chebyshev polynomials, arXiv:math-ph/0608043.

[15] J. Monterde, The Plateau-Bézier problem, *Mathematics of Surfaces*, Proc. 10th IMA conf., Leeds, UK, ed. M. J. Wilson and R. R. Martin, Lect. Notes in Comput. Sci. 2768, Springer-Verlag, 2003, pp. 262–273.

[16] S. Hildebrandt and A. Tromba, *Mathematics and Optimal Form*, Scientific American Library, 1985; MR0854269 (87k:00009).

[17] S. Hyde, S. Andersson, K. Larsson, Z. Blum, T. Landh, S. Lidin and B. W. Ninham, *The Language of Shape: The Role of Curvature in Condensed Matter: Physics, Chemistry and Biology*, Elsevier, 1997.

[18] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, v. 1, Springer-Verlag, 1890.

[19] T. Kawasaki, Classification of spatial polygons that could possibly generate embedded triply periodic minimal surfaces, *Tokyo J. Math.* 26 (2003) 23–53; MR1981998 (2004d:53013).

[20] A. G. Greenhill, *The Applications of Elliptic Functions*, MacMillan and Co., 1892, pp. 26–28.

[21] U. Dierkes, S. Hildebrandt and F. Sauvigny, *Minimal Surfaces*, 2nd ed., Springer-Verlag, 2010, pp. 223–224; MR2566897 (2012b:49002).

[22] S. Lidin, Ring-like minimal surfaces, *J. Phys. France* 49 (1988) 421–427; MR0937640 (89j:53007).

[23] S. Lidin and S. Larsson, Bonnet transformation of infinite periodic minimal surfaces with hexagonal symmetry, *J. Chem. Soc. Faraday Trans.* 86 (1990) 769–775.

[24] A. Fogden and S. T. Hyde, Parametrization of triply periodic minimal surfaces. II, Regular class solutions, *Acta Crystallographica* A48 (1992) 575–591.
[25] G. E. Schröder-Turk, T. Varslot, L. de Campo, S. C. Kapfer and W. Mickel, A bicontinuous mesophase geometry with hexagonal symmetry, *Langmuir* 27 (2011) 10475–10483.
[26] H. Karcher, The triply periodic minimal surfaces of Alan Schoen and their constant mean curvature companions, *Manuscripta Math.* 64 (1989) 291–357; MR1003093 (90g:53010).

5.18 Partitioning Problem

Let us begin with a two-dimensional problem. Consider an equilateral triangular region T with edges of unit length. What is the minimum length of a smooth curve that partitions T into two subregions of equal area? Assuming the vertices of T are $(-1/2, 0)$, $(0, \sqrt{3}/2)$, $(1/2, 0)$, a solution is given by one-sixth of the circumference of the circle

$$x^2 + \left(y - \frac{\sqrt{3}}{2}\right)^2 = r^2 = \frac{3\sqrt{3}}{4\pi}$$

and hence the desired length is [1–4]

$$\frac{1}{6}(2\pi r) = \frac{\pi}{3}\sqrt{\frac{3\sqrt{3}}{4\pi}} = 0.6733868435....$$

See Figure 5.15. The solution is a curve of constant curvature and meets the boundary ∂T of T orthogonally.

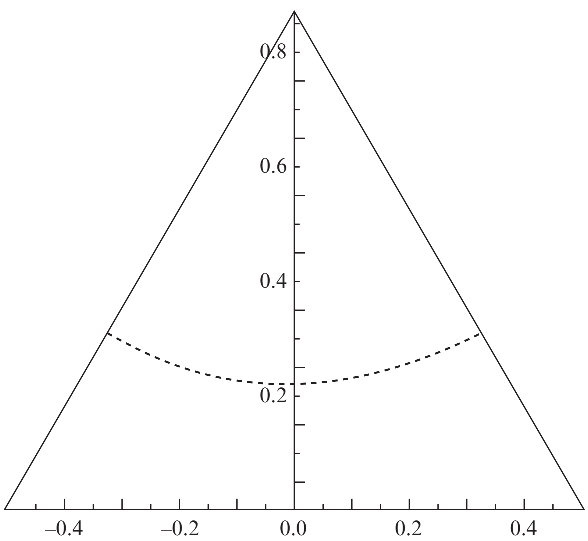

Figure 5.15 Optimally partitioning an equilateral triangle in half

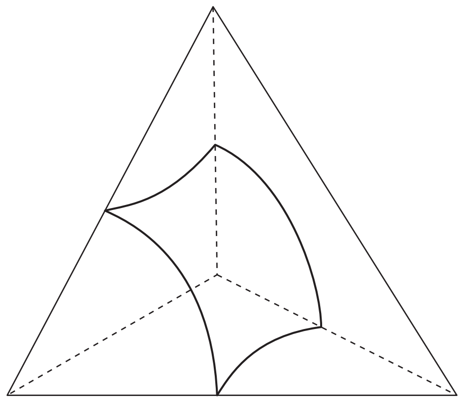

Figure 5.16 Optimally partitioning a regular tetrahedron in half (Smyth [5])

Let us now move up one dimension. Consider a regular tetrahedral region T with edges of unit length. What is the minimum area of a smooth surface that partitions T into two subregions of equal volume? An easy upper bound for the surface area is $1/4 = 0.25$, given by a planar square with vertices coinciding with edge midpoints. A graph of the minimal surface appears in [5] without elaboration – see Figure 5.16 – and a purely numerical approach [6] yields that its area is 0.2172341554.... It is a surface of constant *mean* curvature (in fact, zero) and meets the boundary ∂T of T orthogonally *everywhere*, but its Weierstrass–Enneper representation is unknown. We will not discuss this particular tetrahedron further; additional words are found in [7–9].

Consider instead the irregular tetrahedral region T with vertices $(0,0,0)$, $(1,0,0)$, $(0,0,1)$, $(0,1,1)$. We pose the same problem as before. This is a classical example [10], solved in 1872, and features a portion of what is known as the *Schwarz P surface* (P stands for "Primitive"). The surface has zero mean curvature and thus is a minimal surface in the same sense as the Schwarz D surface. In the following, the functions $F[\phi, m]$ and $K[m]$ are defined exactly as in [11].

5.18.1 *Tetrahedral Dissection*

Unlike our treatment of the Schwarz D surface [11], an expression for the Schwarz P surface in x, y, z solely does not seem possible. We thus turn to a parametric approach using the Weierstrass–Enneper representation [12]:

$$x(u,v) = \kappa \operatorname{Re} \int_0^{u+iv} \frac{1-\omega^2}{\sqrt{1+14\omega^4+\omega^8}}\, d\omega,$$

$$y(u,v) = \kappa \operatorname{Re} \int_0^{u+iv} \frac{i(1+\omega^2)}{\sqrt{1+14\omega^4+\omega^8}}\, d\omega,$$

$$z(u,v) = \frac{1}{2} + \kappa \operatorname{Re} \int_0^{u+iv} \frac{2\omega}{\sqrt{1+14\omega^4+\omega^8}}\, d\omega,$$

where the complex line integrals have endpoint $u+iv$ satisfying

$$u \geq 0, \quad v \leq 0, \quad (u+1)^2 + v^2 \leq 2, \quad u^2 + (v-1)^2 \leq 2$$

– call this planar domain Ω – and the normalization constant is

$$\kappa = \frac{3}{2K[1/9]} = 0.9274219745....$$

This is as far as Nitsche [7, 12] went in characterizing the surface; calculations based on [13] further yield that

$$x = \frac{\kappa}{4} \operatorname{Re}\left(-iF\left[\theta(u,v), \frac{1}{4}\right] + F\left[\theta(u,v), \frac{3}{4}\right]\right),$$

$$y = \frac{\kappa}{4} \operatorname{Re}\left(iF\left[\theta(u,v), \frac{1}{4}\right] + F\left[\theta(u,v), \frac{3}{4}\right]\right),$$

$$z = \frac{1}{2} + \left(2-\sqrt{3}\right) \kappa \operatorname{Im}\left(F\left[\arcsin\left(i\left(2+\sqrt{3}\right)(u+iv)^2\right), \left(2-\sqrt{3}\right)^4\right]\right)$$

where

$$\theta(u,v) = \arcsin\left(\frac{2(1+i)(u+iv)}{\sqrt{1+4i(u+iv)^2-(u+iv)^4}}\right).$$

See Figures 5.17, 5.18, 5.19. The four corners of Ω are mapped to the surface as follows:

$$\begin{array}{ll}
(u,v) = (0,0) \mapsto (x,y,z) = \left(0,0,\tfrac{1}{2}\right) & \text{[front left]} \\
(u,v) = \left(\frac{\sqrt{3}-1}{2}, -\frac{\sqrt{3}-1}{2}\right) \mapsto (x,y,z) = \left(\tfrac{1}{2},\tfrac{1}{2},\tfrac{1}{2}\right) & \text{[back right]} \\
(u,v) = (\sqrt{2}-1,0) \mapsto (x,y,z) = (\xi,0,1-\xi) & \text{[front right]} \\
(u,v) = (0,-(\sqrt{2}-1)) \mapsto (x,y,z) = (0,\xi,\xi) & \text{[back left]}
\end{array}$$

where $\xi \approx 0.350$. Letting $x_u, y_u, z_u, x_v, y_v, z_v$ denote partial derivatives and

$$e = (x_u,y_u,z_u)\cdot(x_u,y_u,z_u), \quad g = (x_v,y_v,z_v)\cdot(x_v,y_v,z_v),$$

$$f = (x_u,y_u,z_u)\cdot(x_v,y_v,z_v)$$

we have surface area

$$\iint_\Omega \sqrt{eg-f^2}\, dv\, du = \frac{1}{4}\frac{K[1/4]}{K[3/4]} = \frac{1}{12}(2.3451028840...) = 0.1954...$$

as predicted in [13].

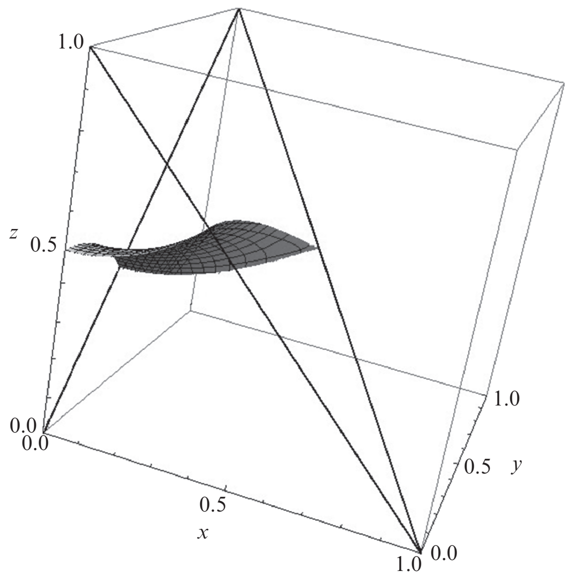

Figure 5.17 Optimally partitioning an irregular tetrahedron in half

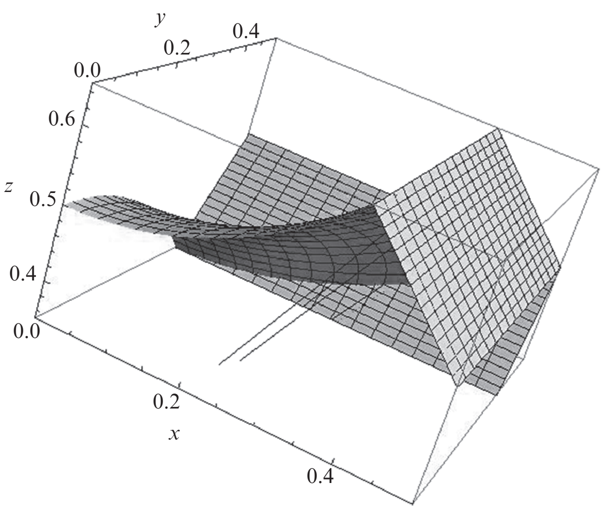

Figure 5.18 First closeup of tetrahedral partition

5.18.2 *Four Edges of a Regular Octahedron*

We return to a variation of Plateau's problem in [11]. Consider a polygonal wire loop with four line segments:

$$(0,0,1/2) \to (1/2,-1/2,1/2) \to (1/2,0,1) \to (1/2,1/2,1/2) \to (0,0,1/2).$$

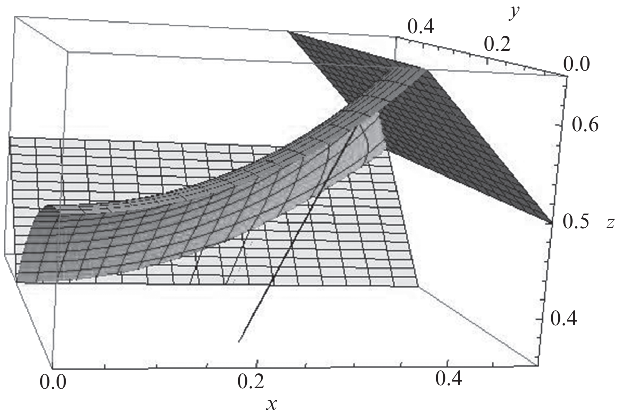

Figure 5.19 Second closeup of tetrahedral partition

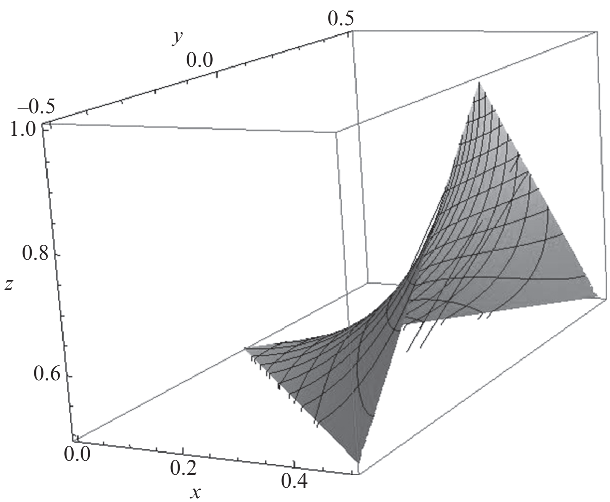

Figure 5.20 Octahedral "four edges" surface

What is the minimal area for any surface spanning this fixed boundary? Equivalently, what is the outcome of dipping the wire loop in a soap solution? [14, 15]

The same formulas for x, y, z apply here, but a new domain $\widetilde{\Omega}$ is needed:

$$u \geq |v|, \quad u^2 + (v+1)^2 \leq 2, \quad u^2 + (v-1)^2 \leq 2.$$

See Figure 5.20. The two corners of $\widetilde{\Omega}$ not in Ω are mapped to the surface as follows:

$$(u, v) = \left(\frac{\sqrt{3}-1}{2}, \frac{\sqrt{3}-1}{2}\right) \mapsto (x, y, z) = \left(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right),$$
$$(u, v) = (1, 0) \mapsto (x, y, z) = \left(\frac{1}{2}, 0, 1\right)$$

and the corresponding area is

$$\iint\limits_{\widetilde{\Omega}} \sqrt{eg-f^2}\,dv\,du = \frac{1}{2}\frac{K[1/4]}{K[3/4]} = \frac{1}{6}(2.3451028840...) = 0.3908...$$

again as predicted in [13].

5.18.3 *Integration Details*

To prove our formulas for x, y, z, we must evaluate the hyperelliptic integral

$$I_p(\eta) = \int\limits_0^{\eta} \frac{\omega^p}{\sqrt{1-14\omega^4+\omega^8}}\,d\omega$$

for $p = 0, 1, 2$. Note that the coefficient of ω^4 in $I_p(\eta)$ is -14 whereas it is $+14$ in the definitions of x, y, z. This is chosen so that we may follow [13] closely and then, at the end, perform a transformation to align with [12].

Let $t = \omega^2$, then $dt = 2\omega\,d\omega$ and

$$I_p(\eta) = \frac{1}{2}\int\limits_0^{\eta^2} \frac{t^{p/2}}{\sqrt{1-14t^2+t^4}}\frac{dt}{\sqrt{t}}.$$

Let $s = t + 1/t$, then assuming $0 < \mathrm{Re}(t) < 1$, we have

$$t = \frac{1}{2}\left(s - \sqrt{s^2-4}\right),$$

$$t^4 - 14t^2 + 1 = \left(s^2-16\right)t^2,$$

$$dt = \frac{1}{2}\frac{\sqrt{s^2-4}-s}{\sqrt{s^2-4}}ds = -\frac{1}{\sqrt{s^2-4}}\,t\,ds$$

hence

$$\begin{aligned}\frac{dt}{t^{3/2}} &= -\frac{1}{\sqrt{s^2-4}}\sqrt{\frac{2}{s-\sqrt{s^2-4}}}\,ds \\ &= -\frac{1}{\sqrt{s-2}\sqrt{s+2}}\frac{\sqrt{s-2}+\sqrt{s+2}}{2}\,ds \\ &= -\frac{1}{2}\left(\frac{1}{\sqrt{s-2}}+\frac{1}{\sqrt{s+2}}\right)ds\end{aligned}$$

hence

$$\begin{aligned}\frac{1}{\sqrt{t^4-14t^2+1}}\frac{dt}{\sqrt{t}} &= \frac{1}{\sqrt{s-4}\sqrt{s+4}}\frac{dt}{t^{3/2}} \\ &= -\frac{1}{2}\frac{1}{\sqrt{s-4}\sqrt{s+4}}\left(\frac{1}{\sqrt{s-2}}+\frac{1}{\sqrt{s+2}}\right)ds\end{aligned}$$

hence

$$I_p(\eta) = -\frac{1}{4}\int\limits_{s=\infty}^{\eta^2+1/\eta^2} \frac{1}{\sqrt{s-4}\sqrt{s+4}}\left(\frac{1}{\sqrt{s-2}}+\frac{1}{\sqrt{s+2}}\right) t^{p/2}ds$$

$$= \frac{1}{2^{2+p/2}}\int\limits_{\eta^2+1/\eta^2}^{\infty}\left(\frac{\left(s-\sqrt{s-2}\sqrt{s+2}\right)^{p/2}}{\sqrt{s-4}\sqrt{s-2}\sqrt{s+4}}+\frac{\left(s-\sqrt{s-2}\sqrt{s+2}\right)^{p/2}}{\sqrt{s-4}\sqrt{s+2}\sqrt{s+4}}\right)ds.$$

Define $\zeta=\eta^2+1/\eta^2$. For the case $p=0$, we need [16]

$$\int\limits_{\zeta}^{\infty}\frac{1}{\sqrt{s-4}\sqrt{s-2}\sqrt{s+4}}\,ds=\frac{1}{\sqrt{2}}F\left[\arcsin\left(\frac{2\sqrt{2}}{\sqrt{\zeta+4}}\right),\frac{3}{4}\right],$$

$$\int\limits_{\zeta}^{\infty}\frac{1}{\sqrt{s-4}\sqrt{s+2}\sqrt{s+4}}\,ds=\frac{1}{\sqrt{2}}F\left[\arcsin\left(\frac{2\sqrt{2}}{\sqrt{\zeta+4}}\right),\frac{1}{4}\right]$$

which together imply that $I_0(\eta)$ is equal to

$$\frac{1}{4\sqrt{2}}\left(F\left[\arcsin\left(\frac{2\sqrt{2}\eta}{\sqrt{\eta^4+4\eta^2+1}}\right),\frac{1}{4}\right]+F\left[\arcsin\left(\frac{2\sqrt{2}\eta}{\sqrt{\eta^4+4\eta^2+1}}\right),\frac{3}{4}\right]\right).$$

Similar work implies that $I_2(\eta)$ is equal to

$$\frac{1}{4\sqrt{2}}\left(-F\left[\arcsin\left(\frac{2\sqrt{2}\eta}{\sqrt{\eta^4+4\eta^2+1}}\right),\frac{1}{4}\right]+F\left[\arcsin\left(\frac{2\sqrt{2}\eta}{\sqrt{\eta^4+4\eta^2+1}}\right),\frac{3}{4}\right]\right).$$

For the case $p=1$, it is best to factor an earlier representation of $2I_1(\eta)$:

$$\int\limits_{0}^{\eta^2}\frac{dt}{\sqrt{t-\left(2+\sqrt{3}\right)}\sqrt{t-\left(2-\sqrt{3}\right)}\sqrt{t+\left(2-\sqrt{3}\right)}\sqrt{t+\left(2+\sqrt{3}\right)}}$$

and employ [16] to simplify this integral to

$$\left(2-\sqrt{3}\right)F\left[\arcsin\left(\left(2+\sqrt{3}\right)\eta^2\right),\left(2-\sqrt{3}\right)^4\right].$$

Our expression for $2I_1(\eta)$ corrects an error that appears in [13].

From

$$\int\limits_{0}^{\eta}\frac{\omega^p}{\sqrt{1+14\omega^4+\omega^8}}\,d\omega=\left(\frac{1-i}{\sqrt{2}}\right)^{p+1}I_p\left(\frac{1+i}{\sqrt{2}}\eta\right)$$

(that is, a rotation of the domain by 45°) and

$$\arcsin\left(\frac{2\sqrt{2}\omega}{\sqrt{\omega^4+4\omega^2+1}}\right)\Bigg|_{\omega=\frac{1+i}{\sqrt{2}}(u+iv)}=\theta(u,v),$$

we deduce that

$$
\begin{aligned}
\frac{x}{\kappa} &= \mathrm{Re}\left\{\left(\frac{1-i}{\sqrt{2}}\right) I_0\left(\frac{1+i}{\sqrt{2}}(u+iv)\right) - \left(\frac{1-i}{\sqrt{2}}\right)^3 I_2\left(\frac{1+i}{\sqrt{2}}(u+iv)\right)\right\} \\
&= \mathrm{Re}\left\{\left(\frac{1-i}{\sqrt{2}}\right)\frac{F\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4\sqrt{2}}\right. \\
&\quad \left. + \left(\frac{1+i}{\sqrt{2}}\right)\frac{-F\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4\sqrt{2}}\right\} \\
&= \mathrm{Re}\left\{\frac{-iF\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{y}{\kappa} &= \mathrm{Re}\left\{i\left(\frac{1-i}{\sqrt{2}}\right) I_0\left(\frac{1+i}{\sqrt{2}}(u+iv)\right) + i\left(\frac{1-i}{\sqrt{2}}\right)^3 I_2\left(\frac{1+i}{\sqrt{2}}(u+iv)\right)\right\} \\
&= \mathrm{Re}\left\{\left(\frac{1+i}{\sqrt{2}}\right)\frac{F\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4\sqrt{2}}\right. \\
&\quad \left. + \left(\frac{1-i}{\sqrt{2}}\right)\frac{-F\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4\sqrt{2}}\right\} \\
&= \mathrm{Re}\left\{\frac{iF\left[\theta(u,v),\frac{1}{4}\right]+F\left[\theta(u,v),\frac{3}{4}\right]}{4}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\frac{z-\frac{1}{2}}{\kappa} &= \mathrm{Re}\left\{\left(\frac{1-i}{\sqrt{2}}\right)^2 2I_1\left(\frac{1+i}{\sqrt{2}}(u+iv)\right)\right\} \\
&= \mathrm{Re}\left\{-i\left(2-\sqrt{3}\right)F\left[\arcsin\left(\left(2+\sqrt{3}\right)\left(\frac{1+i}{\sqrt{2}}(u+iv)\right)^2\right),\left(2-\sqrt{3}\right)^4\right]\right\} \\
&= \mathrm{Im}\left\{\left(2-\sqrt{3}\right)F\left[\arcsin\left(i\left(2+\sqrt{3}\right)(u+iv)^2\right),\left(2-\sqrt{3}\right)^4\right]\right\}
\end{aligned}
$$

as was to be shown.

5.18.4 *Approximations*

With regard to tetrahedral dissection, a reasonable approximation is provided by the plane containing $V_1=\left(0,0,\frac{1}{2}\right)$, $V_2=(\xi,0,1-\xi)$, $V_3=(0,\xi,\xi)$, which also contains

$$V_4=\left(\frac{1}{2},\frac{1}{2},\frac{1}{2}\right)=\left(1-\frac{1}{\xi}\right)V_1+\frac{1}{2\xi}V_2+\frac{1}{2\xi}V_3.$$

The plane cuts the tetrahedron into two polyhedra of equal area, and the area of the quadrilateral slice is written in terms of the cross-product of its diagonals:

$$\frac{1}{2}\left|(V_3 - V_1) \times (V_4 - V_2)\right| = \frac{1}{2\sqrt{2}}\sqrt{(1-2\xi)^2 + 2\xi^2} = 0.2046\ldots > 0.1954\ldots.$$

With regard to the octahedral "four edges" surface, an excellent approximation is given in [17]:

$$z = \frac{1}{\pi}\arccos\left(\cos(\pi x) - \cos(\pi y)\right), \quad x \geq |y|, \quad -\frac{1}{2} \leq y \leq \frac{1}{2}$$

and the corresponding surface area is

$$\int_0^{1/2}\int_{-x}^{x}\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\,dy\,dx = 0.3920\ldots > 0.3908\ldots.$$

The constant 2.3451... appears in [18–23].

[1] Y. Tomonaga, *Geometry of Length and Area*, Dept. of Math., Utsunomiya Univ., 1974, pp 29–33; MR0430973 (55 #3977).
[2] M. N. Bleicher, Isoperimetric divisions into several cells with natural boundary, *Intuitive Geometry*, Proc. 1985 Siófok conf., ed. K. Böröczky and G. Fejes Tóth, North-Holland, 1987, pp. 63–84; MR0910701 (88j:52023).
[3] S. J. Cox and E. Flikkema, The minimal perimeter for *N* confined deformable bubbles of equal area, *Elec. J. Combin.* 17 (2010) R45; MR2607331 (2011e:05054).
[4] B. Eduard, Dissection of regular polygons in equal area pieces with minimal cut length, unpublished tables (2009).
[5] B. Smyth, Stationary minimal surfaces with boundary on a simplex, *Invent. Math.* 76 (1984) 411–420; MR0746536 (86j:53015).
[6] K. A. Brakke, The Surface Evolver, downloadable software, Susquehanna Univ.
[7] J. C. C. Nitsche, Stationary partitioning of convex bodies, *Arch. Rational Mech. Anal.* 89 (1985) 1–19; corrigendum 95 (1986) 389; MR0784101 (86j:53013) and MR0853784 (87h:53008).
[8] J. Jost, The geometric calculus of variations: A short survey and a list of open problems, *Exposition. Math.* 6 (1988) 111–143; MR0938179 (89h:58036).
[9] U. Dierkes, S. Hildebrandt and A. J. Tromba, *Regularity of Minimal Surfaces*, 2nd ed., Springer-Verlag, 2010, pp. 39–41; MR2760441 (2012b:49003).
[10] H. A. Schwarz, *Gesammelte Mathematische Abhandlungen*, v. 1, Springer-Verlag, 1890.
[11] S. R. Finch, Gergonne-Schwarz surface, *this volume*, §5.17.
[12] J. C. C. Nitsche, Periodic surfaces that are extremal for energy functionals containing curvature functions, *Statistical Thermodynamics and Differential Geometry of Microstructured Materials*, Proc. 1991 Minneapolis conf., ed. H. T. Davis and J. C. C. Nitsche, Springer-Verlag, 1993, pp. 69–98; MR1226921 (94m:58052).
[13] P. J. F. Gandy and J. Klinowski, Exact computation of the triply periodic Schwarz P minimal surface, *Chem. Phys. Lett.* 322 (2000) 579–586.
[14] T. H. Colding and W. P. Minicozzi, *A Course in Minimal Surfaces*, Amer. Math. Soc., 2011, pp. 17–19; MR2780140.

[15] W. H. Meeks and J. Pérez, The classical theory of minimal surfaces, *Bull. Amer. Math. Soc.* 48 (2011) 325–407; MR2801776 (2012m:53008).

[16] P. F. Byrd and M. D. Friedman, *Handbook of Elliptic Integrals for Engineers and Scientists*, 2nd ed., Springer-Verlag, 1971, pp. 84, 112; MR0277773 (43 #3506).

[17] I. S. Barnes, A useful trigonometric approximation to periodic minimal surfaces?, *Aust. Math. Soc. Gaz.*, v. 17 (1990) n. 4, 99–105; Zbl 0714.53012.

[18] A. H. Schoen, *Infinite Periodic Minimal Surfaces Without Self-Intersections*, NASA technical note TN D-5541, 1970.

[19] A. L. Mackay, Periodic minimal surfaces, *Nature* 314 (1985) 604–606.

[20] D. M. Anderson, H. T. Davis, J. C. C. Nitsche and L. E. Scriven, Periodic surfaces of prescribed mean curvature, *Physics of Amphiphilic Layers*, Proc. Les Houches workshop, 1987, ed. J. Meunier, D. Langevin and N. Boccara, Springer-Verlag, 1987, p. 130 (abstract only); *Advances in Chemical Physics*, v. 77, ed. I. Prigogine and S. A. Rice, Wiley, 1990, pp. 337–396.

[21] A. Fogden and S. Lidin, Description of cubic liquid-crystalline structures using simple surface foliations, *J. Chem. Soc. Faraday Trans.* 90 (1994) 3423–3431.

[22] L. Hauswirth, J. Pérez, P. Romon and A. Ros, The periodic isoperimetric problem, *Trans. Amer. Math. Soc.* 356 (2004) 2025–2047; MR2031051 (2004j:53014).

[23] S. Rajagopalan and R. A. Robb, Schwarz meets Schwann: Design and fabrication of biomorphic and durataxic tissue engineering scaffolds, *Medical Image Analysis* 10 (2006) 693–712.

5.19 Soap Film Experiments

We conclude our brief survey of minimal surfaces, started in [1, 2], with more solutions of Plateau's problem. The functions $F[\phi, m]$ and $K[m]$ are defined exactly as before.

5.19.1 *Ramp Inside a Cube*

Consider a polygonal wire loop with six line segments:

$$(0,0,0) \to (1,0,0) \to (1,1,0) \to (1,1,1) \to (0,1,1) \to (0,1,0) \to (0,0,0).$$

What is the minimal area for any surface spanning this fixed boundary? Equivalently, what is the outcome of dipping the wire loop in a soap solution? Following [3–5], we numerically solve the equation

$$\frac{1}{\sqrt{2}} = \frac{1}{\sqrt{2+\sqrt{2-\lambda}}} \frac{K\left[\frac{8\sqrt{2-\lambda}}{\left(2+\sqrt{2-\lambda}\right)^2}\right]}{K\left[\frac{2-\sqrt{2-\lambda}}{2+\sqrt{2-\lambda}}\right]}$$

and obtain $\lambda = 1.5733414653\ldots$. Define

$$\begin{aligned} x(u,v) &= \kappa \,\mathrm{Re} \int_0^{u+iv} \frac{1-\tau^2}{\sqrt{1+\lambda\tau^4+\tau^8}}\, d\tau \\ &= \kappa \,\mathrm{Re} \left\{ \frac{F\left[\arcsin\left(\sqrt{2+\sqrt{2-\lambda}}\,\frac{\omega}{1+\omega^2}\right), \frac{2-\sqrt{2-\lambda}}{2+\sqrt{2-\lambda}}\right]}{\sqrt{2+\sqrt{2-\lambda}}} \right\}, \end{aligned}$$

$$\begin{aligned} y(u,v) &= \kappa \,\mathrm{Re} \int_0^{u+iv} \frac{i(1+\tau^2)}{\sqrt{1+\lambda\tau^4+\tau^8}}\, d\tau \\ &= \kappa \,\mathrm{Re} \left\{ \frac{F\left[\arcsin\left(\sqrt{2+\sqrt{2-\lambda}}\,\frac{i\omega}{1-\omega^2}\right), \frac{2-\sqrt{2-\lambda}}{2+\sqrt{2-\lambda}}\right]}{\sqrt{2+\sqrt{2-\lambda}}} \right\}, \end{aligned}$$

$$\begin{aligned} z(u,v) &= \kappa \,\mathrm{Re} \int_0^{u+iv} \frac{2\tau}{\sqrt{1+\lambda\tau^4+\tau^8}}\, d\tau \\ &= \sqrt{2}\kappa \,\mathrm{Re} \left\{ \frac{F\left[\arcsin\left(\sqrt{\frac{-\lambda+\sqrt{-4+\lambda^2}}{2}}\,\omega^2\right), \frac{\left(\lambda+\sqrt{-4+\lambda^2}\right)^2}{4}\right]}{\sqrt{-\lambda+\sqrt{-4+\lambda^2}}} \right\} \end{aligned}$$

where the complex line integrals have endpoint $\omega = u + iv$ satisfying

$$u^2 + v^2 \le 1, \quad |v| \ge u$$

– call this planar domain Ω – and the normalization constant κ satisfies

$$\frac{1}{\kappa} = 2\sqrt{2}\,\mathrm{Re}\left\{ \frac{1}{\sqrt{-\lambda+\sqrt{-4+\lambda^2}}} K\left[\frac{\left(\lambda+\sqrt{-4+\lambda^2}\right)^2}{4} \right] \right\}.$$

These expressions give the top portion ($z > 0$) of the surface in Figures 5.21 and 5.22. A reflection provides the bottom portion; a rotation would further align the surface with our six prescribed vertices. This is a representative of the *Schwarz*

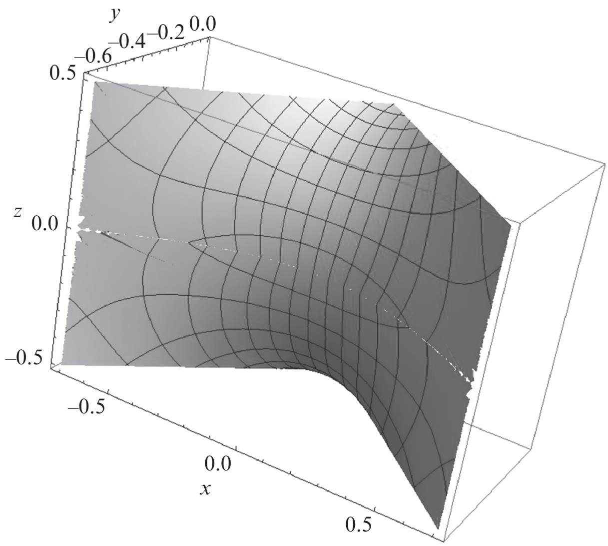

Figure 5.21 First view of CLP surface

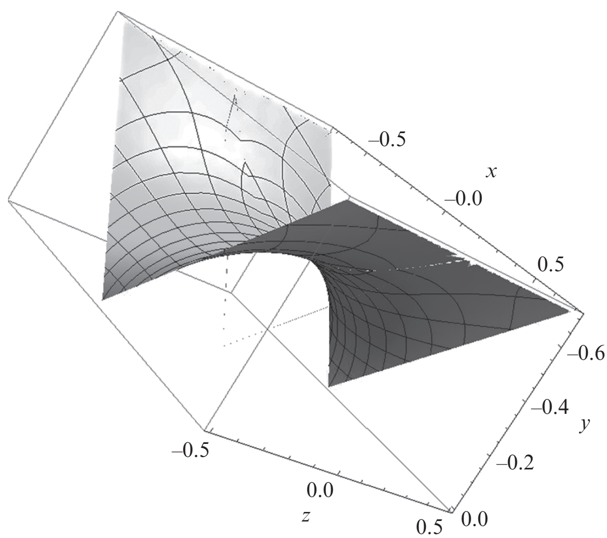

Figure 5.22 Second view of CLP surface

CLP family of minimal surfaces; a nice contrast exists with the Schwarz D surface [6]. We also have surface area

$$2\iint\limits_{\Omega} \sqrt{eg-f^2}\,dv\,du = 1.7816507345\ldots$$

where e, f, g are as in [2]. Brakke and Weber duplicated this calculation, using Surface Evolver software [7] and conformal mapping techniques [8] respectively.

5.19.2 *Saddle Inside a Cube*

Consider a polygonal wire loop with eight line segments:

$$(0,0,1) \to (1,0,1) \to (1,0,0) \to (1,1,0) \to (1,1,1) \to (0,1,1) \to (0,1,0) \to (0,0,0) \to (0,0,1).$$

Again, what is the minimal area for any surface spanning this fixed boundary? Following [9], we numerically solve the equation

$$\frac{1}{2} = \sqrt{\frac{\sqrt{2-\lambda}}{-\lambda + \sqrt{-4+\lambda^2}}}\,\frac{K\left[\dfrac{\left(\lambda + \sqrt{-4+\lambda^2}\right)^2}{4}\right]}{K\left[\dfrac{1}{2} - \dfrac{1}{\sqrt{2-\lambda}}\right]}$$

and obtain $\lambda = -5.3485781991\ldots$. Define

$$\begin{aligned} x(u,v) &= \kappa\,\mathrm{Re}\int_0^{u+iv} \frac{1-\tau^2}{\sqrt{1+\lambda\tau^4+\tau^8}}\,d\tau \\ &= \kappa\,\mathrm{Re}\left\{\frac{F\left[\arcsin\left(\sqrt{\dfrac{2\sqrt{2-\lambda}}{1+\sqrt{2-\lambda}\,\omega^2+\omega^4}}\,\omega\right), \dfrac{1}{2} - \dfrac{1}{\sqrt{2-\lambda}}\right]}{\sqrt{2\sqrt{2-\lambda}}}\right\}, \end{aligned}$$

$$\begin{aligned} y(u,v) &= \kappa\,\mathrm{Re}\int_0^{u+iv} \frac{i(1+\tau^2)}{\sqrt{1+\lambda\tau^4+\tau^8}}\,d\tau \\ &= \kappa\,\mathrm{Re}\left\{\frac{i\,F\left[\arcsin\left(\sqrt{\dfrac{2\sqrt{2-\lambda}}{1+\sqrt{2-\lambda}\,\omega^2+\omega^4}}\,\omega\right), \dfrac{1}{2} + \dfrac{1}{\sqrt{2-\lambda}}\right]}{\sqrt{2\sqrt{2-\lambda}}}\right\}, \end{aligned}$$

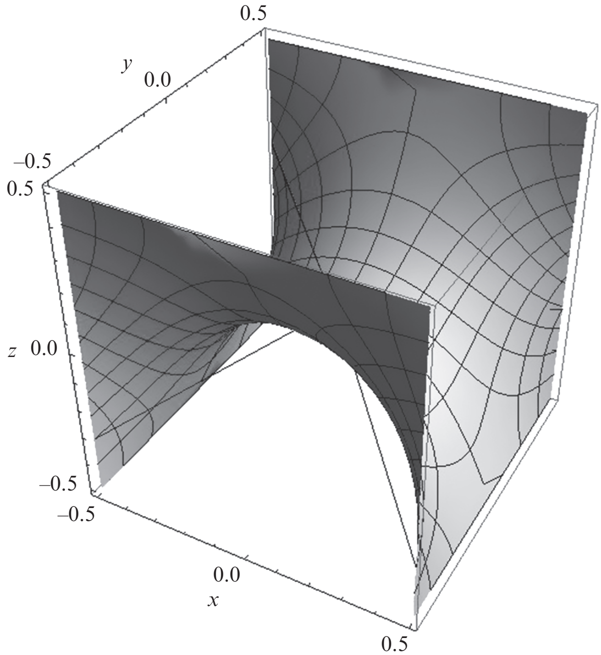

Figure 5.23 First view of T surface

$$z(u,v) = \kappa\,\mathrm{Re}\int_0^{u+iv} \frac{2\tau}{\sqrt{1+\lambda\,\tau^4+\tau^8}}\,d\tau$$

$$= \sqrt{2}\kappa\,\mathrm{Re}\left\{\frac{F\left[\arcsin\left(\sqrt{\dfrac{-\lambda+\sqrt{-4+\lambda^2}}{2}}\,\omega^2\right),\ \dfrac{\left(\lambda+\sqrt{-4+\lambda^2}\right)^2}{4}\right]}{\sqrt{-\lambda+\sqrt{-4+\lambda^2}}}\right\}$$

where the complex line integrals have endpoint $\omega = u + iv$ satisfying

$$u^2 + v^2 \leq 1, \quad |v| \geq u$$

– call this planar domain Ω – and the normalization constant κ satisfies

$$\frac{1}{\kappa} = \frac{2\sqrt{2}}{\sqrt{-\lambda+\sqrt{-4+\lambda^2}}}K\left[\frac{\left(\lambda+\sqrt{-4+\lambda^2}\right)^2}{4}\right].$$

(No call to the Re function is needed here, unlike before.) These expressions give a quarter-wedge of the surface in Figures 5.23 and 5.24. Reflections provide the

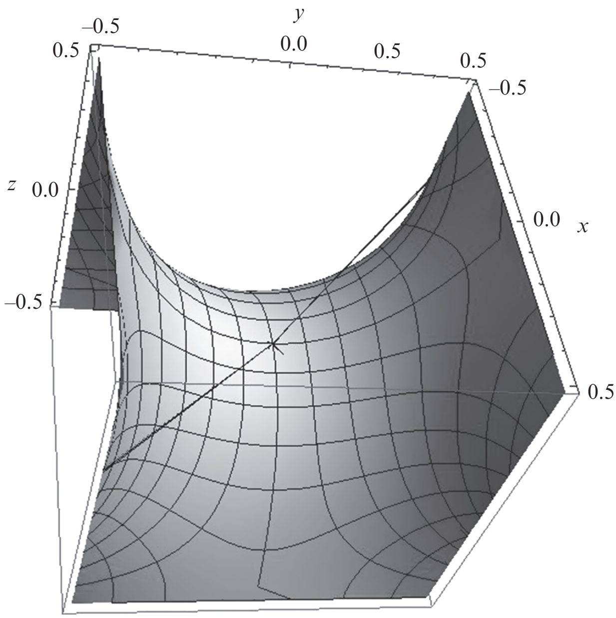

Figure 5.24 Second view of T surface

other three quarter-wedges; a rotation would further align the surface with our eight prescribed vertices. This is a representative of the *Schwarz T* family of minimal surfaces, also known as *tD surfaces* (generalizing the D surface). We finally have surface area

$$4\iint_{\Omega} \sqrt{eg-f^2}\, dv\, du = 2.4674098291... = 2(1.2337049145...),$$

duplicating a calculation by Brakke [7]. The CLP expression for z is identical to the T expression for z; this is true for x and y too (although less apparently so). The latter expressions for $\{x, y\}$ give elliptic parameters $\{1/4, 3/4\}$ when $\lambda = -14$, consistent with our earlier work [2]. The former expressions, which come from [3], give $\{-1/3, -1/3\}$ instead. Yet another set of expressions appear in [9], which we have not attempted to use.

The presence of the constant 1.2337049145..., which also appeared in [1], indicates that the T surface is related to Gergonne's surface [9, 10]. This is surprising because the T surface is the solution of a fixed boundary problem whereas Gergonne's surface solves a problem involving a partially free boundary.

5.19.3 *Other Problems*

Consider a smooth wire loop C given parametrically by

$$x = \cos(\theta), \quad y = \sin(\theta), \quad z = \cos(\theta)^2, \quad 0 \le \theta < 2\pi.$$

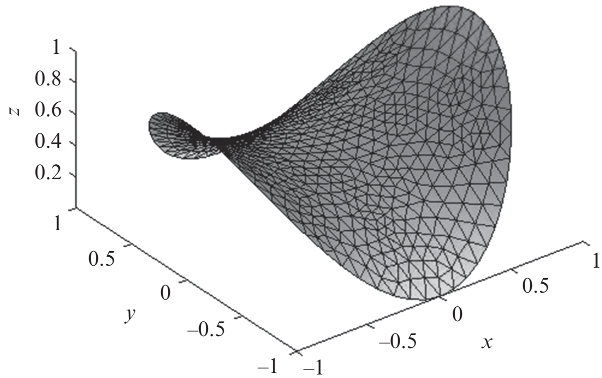

Figure 5.25 Surface from Matlab help pages with boundary C

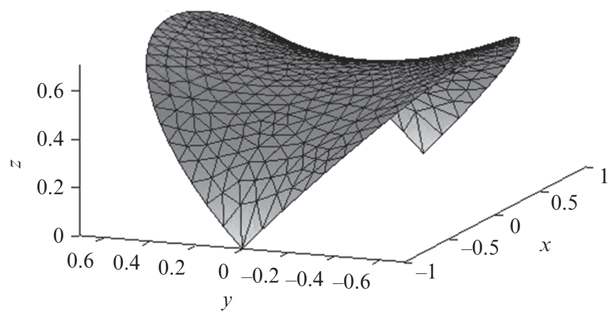

Figure 5.26 Surface spanning folded circular loop

The projection of C into the xy-plane is the unit circle; its projection into the xz-plane is the parabola $z = x^2$; its projection into the yz-plane is the parabola $z = 1 - y^2$. The arclength of C is

$$4\sqrt{2}E\left[\frac{1}{2}\right] = 7.6403955780... = 4\,(1.9100988945...) > 2\pi,$$

which incidentally is the arclength of the planar sine curve (one period). A closed-form expression for the area $3.8269736664... > \pi$ of the minimal surface spanning C is unknown [11, 12]. See Figure 5.25.

Consider instead the folded circular loop, that is, the outcome of orthogonally mounting two unit semicircles along common diameters. For the boundary configuration shown in Figure 5.26, we deduce that its projection in the xy-plane is the ellipse $x^2 + 2y^2 = 1$ and its height z is simply $|y|$. The arclength is obviously 2π; the surface area $2.4822844847... < (2 + \pi)/2$ is again unknown [13, 14].

We wonder finally what can be said about minimal surfaces that span three disjoint perpendicular cubic edges. This topic is believed to be more difficult than the "two diagonals" analog (Gergonne's surface) and progress would be good to see someday..

Acknowledgments Kenneth Brakke generously computed all surface areas in this essay, verifying my results for the CLP and T cases, and providing a reliable standard (against which to compare various approaches) for other cases. Matthias Weber demonstrated an impressive new technique to evaluate surface area for the CLP case.

[1] S. R. Finch, Gergonne-Schwarz surface, *this volume*, §5.17.
[2] S. R. Finch, Partitioning problem, *this volume*, §5.18.
[3] D. Cvijović and J. Klinowski, The T and CLP families of triply periodic minimal surfaces. III, The properties and computation of CLP surfaces, *J. Physique I* 2 (1992) 2207–2220; MR1203045 (94a:53014).
[4] H. G. von Schnering and R. Nesper, How Nature adapts chemical structures to curved surfaces, *Angewandte Chemie* 26 (1987) 1059-1200.
[5] A. H. Schoen, *Infinite Periodic Minimal Surfaces Without Self-Intersections*, NASA technical note TN D-5541, 1970.
[6] H. Karcher, Introduction to conjugate Plateau constructions, *Global Theory of Minimal Surfaces*, Proc. 2001 Clay Math. Institute Summer School, Berkeley, ed. D. Hoffman, Amer. Math. Soc., 2005, pp. 137–161; MR2167258 (2006d:53005).
[7] K. A. Brakke, The Surface Evolver, downloadable software, Susquehanna Univ.
[8] M. Weber, Classical minimal surfaces in Euclidean space by examples: Geometric and computational aspects of the Weierstrass representation, *Global Theory of Minimal Surfaces*, Proc. 2005 Clay Math. Institute Summer School, Berkeley, ed. D. Hoffman, Amer. Math. Soc., 2005, pp. 19–63; MR2167255 (2006e:53025).
[9] D. Cvijović and J. Klinowski, The T and CLP families of triply periodic minimal surfaces. II, The properties and computation of T surfaces, *J. Physique I* 2 (1992) 2191–2205; MR1203044 (94a:53013).
[10] T. Kawasaki, Classification of spatial polygons that could possibly generate embedded triply periodic minimal surfaces, *Tokyo J. Math.* 26 (2003) 23–53; MR1981998 (2004d:53013).
[11] MathWorks Documentation Team, Minimal surface problem on the unit disk, *Partial Differential Equation Toolbox Examples*.
[12] S. Finch, A. Weiss and B. Greene, Surface area of *pdedemo3*, Matlab Central thread 330079.
[13] S. Finch and B. Greene, *pdenonlin* "Stepsize too small", Matlab Central thread 330468.
[14] S. Finch, Convex hull of two orthogonal disks, arXiv:1211.4514.

5.20 Inflating an Inelastic Membrane

Starting with two circular unit disks (made of Mylar, a thin material that does not stretch nor shrink), we sew these together along their boundaries and then fill the interior with a fluid (air or helium) to capacity. What is the shape of the resulting three-dimensional solid of revolution (Mylar balloon)? [1–3]

Without loss of generality, assume that the solid is centered at the origin and its axis of revolution is the z axis. In the plane $y=0$, the boundary curve $z=z(x)$ solves the following calculus of variations problem: Maximize volume

$$4\pi \int_0^\rho x\, z(x) dx$$

subject to the constraint

$$\int_0^\rho \sqrt{1+z'(x)^2}\, dx = 1$$

where $0<\rho<1$ is fixed. It turns out that the optimal value of ρ is

$$\rho = \frac{4\sqrt{2\pi}}{\Gamma(1/4)^2} = \frac{\sqrt{2}}{K[1/2]} = 0.7627597635... = (1.3110287771...)^{-1}$$

and the parametric representation for the associated boundary surface is

$$x=\rho\,\mathrm{cn}\left(u,\tfrac{1}{2}\right)\cos(v),\quad y=\rho\,\mathrm{cn}\left(u,\tfrac{1}{2}\right)\sin(v),$$
$$z=\sqrt{2}\rho\left(E\left[\arcsin\left(\mathrm{sn}\left(u,\tfrac{1}{2}\right)\right),\tfrac{1}{2}\right]-\tfrac{1}{2}u\right)$$

for $-K[1/2]<u<K[1/2]$, $0<v<2\pi$. In the preceding, $K[m]$, $\mathrm{sn}(u,m)$, $\mathrm{cn}(u,m)$ are defined exactly as in [4] and

$$E[\phi,m] = \int_0^{\sin(\phi)} \sqrt{\frac{1-m\,t^2}{1-t^2}}\, dt$$

denotes the incomplete elliptic integral of the second kind. These are admittedly incompatible with [5] but we purposefully choose formulas here to be consistent with the computer algebra package Mathematica. See Figure 5.27. Let $E[m]=E[\pi/2,m]$. Clearly ρ is the equatorial radius and

$$\begin{aligned}\tau &= \frac{(2\pi)^{3/2}}{\Gamma(1/4)^2}\rho = 2\sqrt{2}\left(E\left[\tfrac{1}{2}\right]-\tfrac{1}{2}K\left[\tfrac{1}{2}\right]\right)\rho = (1.1981402347...)\rho\\ &= \frac{16\pi^2}{\Gamma(1/4)^4} = \frac{\pi}{K[1/2]^2} = 0.9138931620...\end{aligned}$$

is the polar diameter (thickness). Note that $\tau/(2\rho)=0.5990701173...$, the ratio of extreme distances through the origin.

The volume

$$\begin{aligned}V &= \sqrt{\frac{\pi}{2}}\frac{\Gamma(1/4)^2}{6}\rho^3 = \frac{\sqrt{2}\pi}{3}K\left[\frac{1}{2}\right]\rho^3 = (2.7458122499...)\rho^3\\ &= \frac{64\pi^2}{3\Gamma(1/4)^4} = \frac{4\pi}{3K[1/2]^2} = 1.2185242161...\end{aligned}$$

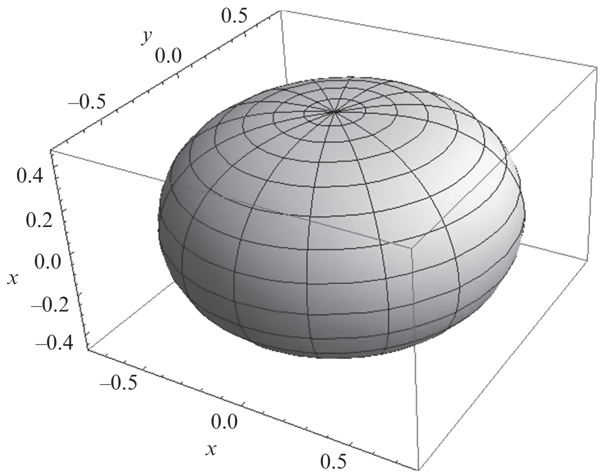

Figure 5.27 Mylar balloon, starting from two unit disks

is considerably less than $(\sqrt{2}/3)\pi = 1.48...$, the volume of the sphere with surface area equal to that of the two original disks. ($4\pi r^2 = 2\pi$, hence $r = 1/\sqrt{2}$, hence $(4/3)\pi r^3 = (\sqrt{2}/3)\pi$.) It seems reasonable to call V the **Mylar balloon constant**. The surface area A possesses an elementary expression: $\pi^2\rho^2$. Comparing the original area 2π with A:

$$\frac{2\pi}{A} = \frac{2}{\pi\rho^2} = \frac{1}{\tau} = 1.0942198076...$$

reveals a remarkable fact. We seem to have lost some of the 2D area, despite the 1D restriction on Mylar stretching/shrinking. There must be crimping or wrinkling of the inflated balloon in order to accommodate $\approx 9.42\%$ area of the deflated balloon. Most of the crimping occurs at the equator; none occurs at the poles. More precisely, the crimping is governed by a local distribution function [6]

$$\delta(x) = \frac{\rho^2}{x}\int_0^x \frac{dt}{\sqrt{\rho^4 - t^4}} = \frac{\rho}{\sqrt{2}x}\left(K\left[\frac{1}{2}\right] - F\left[\arccos\left(\frac{x}{\rho}\right), \frac{1}{2}\right]\right)$$

over $0 < x < \rho$, where $F[\phi, m]$ is defined exactly as in [4]. See Figure 5.28. We have $\delta(\rho) = 1/r = 1.311...$ whereas $\delta(0) = 1$. Implicit in all our analysis is an assumption that the wrinkles do not affect the volume of the balloon. We wonder about the realism of such, given that the wrinkles *do* affect the surface area significantly.

The unit square analog of the Mylar balloon gives rise to a teabag or paper bag [7, 8], whose optimal volume appears to be approximately 0.208 [9, 10]. More work will be needed to confirm that the actual **teabag constant** is no larger than this value.

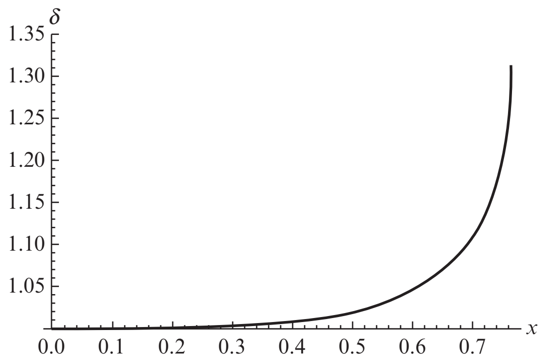

Figure 5.28 Local distribution of the 9.42% excess area

[1] W. H. Paulsen, What is the shape of a Mylar balloon? *Amer. Math. Monthly* 101 (1994) 953–958; MR1304319 (95i:49067).
[2] I. M. Mladenov and J. Oprea, The Mylar balloon revisited, *Amer. Math. Monthly* 110 (2003) 761–784; MR2024747 (2004k:33042).
[3] J. Oprea, *The Mathematics of Soap Films: Explorations with Maple®*, Amer. Math. Soc., 2000, pp. 154–157, 258–260; MR1779498 (2002c:53013).
[4] S. R. Finch, Gergonne-Schwarz surface, *this volume*, §5.17.
[5] S. R. Finch, Random triangles, *this volume*, §5.10.
[6] M. Hadzhilazova and I. M. Mladenov, Once more the Mylar balloon, *C. R. Acad. Bulgare Sci.* 61 (2008) 847–856; MR2450943 (2009h:53007).
[7] M. A. B. Deakin, The filled inelastic membrane: a set of challenging problems in the calculus of variations, *Bull. Aust. Math. Soc.* 80 (2009) 506–509; MR2569925 (2010m:49060).
[8] Wikipedia contributors, Paper bag problem, Wikipedia entry (2012).
[9] A. Gammel, The teabag constant, unpublished note (2003).
[10] A. Gammel, Simulation of inflating a teabag, YouTube video file (2010).

5.21 Enumerative Geometry

Given a complex projective variety $\mathbb{V}$ (as defined in [1]), we wish to count the curves in $\mathbb{V}$ that satisfy certain prescribed conditions. Let $\widetilde{\mathbb{C}}^n$ denote complex projective n-dimensional space. In our first example, $\mathbb{V} = \widetilde{\mathbb{C}}^2$, the complex projective plane; in the second and third, $\mathbb{V}$ is a general hypersurface in $\widetilde{\mathbb{C}}^n$ of degree $2n - 3$. Call such $\mathbb{V}$ a **cubic twofold** when $n = 3$ and a **quintic threefold** when $n = 4$.

Our interest is in **rational curves**, which include all lines (degree 1), conics (degree 2) and singular cubics (degree 3). No elliptic curves are rational. The word "rational" here refers to the affine parametrization of the curve – a ratio of polynomials – and the curve is of degree d if the polynomials are of degree at most d. For instance, the circle $x^2 + y^2 = 1$ is represented as

$$x = \frac{1 - t^2}{1 + t^2}, \quad y = \frac{2t}{1 + t^2}, \quad -\infty < t < \infty.$$

The lemniscate of Bernoulli has degree 4 and is represented as

$$x = \frac{1 - t^4}{1 + 6t^2 + t^4}, \quad y = \frac{2t\left(1 - t^2\right)}{1 + 6t^2 + t^4}, \quad -\infty < t < \infty.$$

It is also defined implicitly:

$$\left(x^2 + y^2\right)^2 = x^2 - y^2$$

and clearly possesses a singularity (vanishing gradient) at the origin. The semi-cubical parabola $y^2 = x^3$ and four-petal rose

$$\left(x^2 + y^2\right)^3 = 4x^2y^2$$

possess likewise. All rational curves, smooth or not, have genus 0.

5.21.1 *Rational Plane Curves Passing Through Points*

In the following, we use homogeneous coordinates. Given two distinct points (X_1, Y_1, Z_1), (X_2, Y_2, Z_2) in $\widetilde{\mathbb{C}^2}$, there is exactly one line passing through both because the simultaneous system of equations

$$aX_j + bY_j + cZ_j = 0, \quad j \in \{1, 2\}$$

has a unique solution (a, b, c) in $\widetilde{\mathbb{C}^2}$ (up to a common scalar). It is a little harder to prove the corresponding result for conics. Given five points (X_j, Y_j, Z_j) in *general position*, there is exactly one conic passing through all five via study of

$$aX_j^2 + bX_jY_j + cY_j^2 + dX_jZ_j + eY_jZ_j + fZ_j^2 = 0, \quad j \in \{1, 2, 3, 4, 5\}$$

in $\widetilde{\mathbb{C}^5}$. Hence we have $K_1 = K_2 = 1$, where K_d is defined as the number of rational curves in $\widetilde{\mathbb{C}^2}$ of degree d passing through $3d - 1$ general points. The quantity $3d - 1$ turns out to be the critical threshold for our question: less would give an answer of infinity, more would give an answer of zero [2].

Proving that $K_3 = 12$ involves a heavy dose of algebraic geometry [3, 4]. Credit for this accomplishment (in the mid-1800s) is assigned variously to Chasles [5] and Steiner [6].

Kontsevich's famous recursion [7–9]:

$$K_d = \sum_{\substack{d_1 + d_2 = d, \\ d_1 \geq 1, d_2 \geq 1}} K_{d_1} K_{d_2} \left[d_1^2 d_2^2 \binom{3d - 4}{3d_1 - 2} - d_1^3 d_2 \binom{3d - 4}{3d_1 - 1} \right], \quad d > 1$$

was not found until recently (in 1994). Its astonishing proof drew upon ideas not from geometry but from mathematical physics, specifically, quantum field theory and string theory. Other relevant recursions for curve counting are known [7, 10–12] but these are too complicated for us to discuss here.

The asymptotics for K_d are [11, 13]

$$\frac{K_d}{(3d-1)!} \sim \frac{(0.1380093466...)^d}{d^{7/2}}$$
$$\times \left(\frac{6.0358078488...}{1} - \frac{2.2352424409...}{d} + \frac{0.0543137879...}{d^2} + \cdots \right)$$

as $d \to \infty$, obtained using a device due to Zagier called the "asymp$_k$ trick". No closed-form expression for these constants is known.

5.21.2 *Lines On a Hypersurface*

The fact that exactly 27 lines lie on a cubic twofold in $\widetilde{\mathbb{C}}^3$ is a well-known theorem [14, 15] due to Cayley & Salmon (in 1849). Somewhat later, Schubert proved (in 1886) that exactly 2875 lines lie on a quintic threefold in $\widetilde{\mathbb{C}}^4$. Thus we have $M_3 = 27$ and $M_4 = 2875$, where M_n is defined as the number of lines on a general hypersurface in $\widetilde{\mathbb{C}}^n$ of degree $2n-3$. Expanding on these results, van der Waerden proved (in 1933) that

$$M_n = \frac{1}{(n-1)!} \frac{d^{n-1}}{dx^{n-1}} \left((1-x) \prod_{k=0}^{2n-3} (2n-3-k+kx) \right) \Bigg|_{x=0}$$

and Zagier [9, 13] obtained asymptotics

$$M_n \sim \sqrt{\frac{27}{\pi}} (2n-3)^{2n-7/2} \left(1 - \frac{9}{8n} - \frac{111}{640n^2} - \frac{9999}{25600n^3} + \cdots \right)$$

as $n \to \infty$. In this case, closed-form expressions are available.

5.21.3 *Rational Curves On a Quintic Threefold*

The number of conics on a cubic twofold is infinity. By contrast, the number of conics on a quintic threefold is 609250. Our discussion at this point becomes highly speculative – it is merely conjectured (by Clemens [8]) that the number n_d of degree d rational curves on a quintic threefold is finite – but the following calculations are known to be valid at least for $d \le 9$. Define $f_0(q), f_1(q), f_2(q)$ via power series expansion of a certain hypergeometric function [4]:

$$\sum_{d=0}^{\infty} q^d \frac{\prod_{j=1}^{5d} (5w+j)}{\prod_{k=1}^{d} (5w+k)^5} = f_0(q) + f_1(q)w + f_2(q)w^2 + \cdots.$$

It follows that

$$f_0(q) = \sum_{d=0}^{\infty} q^d \frac{(5d)!}{(d!)^5}, \quad f_1(q) = \sum_{d=0}^{\infty} q^d \left(\frac{(5d)!}{(d!)^5} \sum_{i=d+1}^{5d} \frac{1}{i} \right)$$

(a similar expression for $f_2(q)$ would be good to see). We then define rational numbers N_d recursively from

$$f_2(q) = \frac{1}{2}\frac{f_1(q)^2}{f_0(q)} + \frac{1}{5}\sum_{d=0}^{\infty} d\, N_d q^d f_0(q) \exp\left(d\frac{f_1(q)}{f_0(q)}\right),$$

yielding

$$\{N_d\}_{d=1}^{\infty} = \left\{2875, \frac{4876875}{8}, \frac{8564575000}{27}, \frac{15517926796875}{64}, 229305888887648, \ldots\right\}.$$

Such numbers are examples of **Gromov–Witten invariants**, which count not only the rational curves we desire, but also capture (unwanted) additional structure [8]. The final step is another recursion [4, 16]:

$$N_d = \sum_{h\,|\,d} \frac{n_{d/h}}{h^3}$$

yielding

$$\{n_d\}_{d=1}^{\infty} = \{2875, 609250, 317206375, 242467530000, 229305888887625, \ldots\}.$$

It is, again, merely conjectured (by Gopakumar & Vafa [8, 9]) that all numbers n_d obtained in this manner are indeed integers. Much work lies ahead to rigorously confirm everything written here. The asymptotics for n_d remain open.

Let S be a cubic twofold and let H_d be the number of rational curves on S of degree d passing through $d-1$ general points on S. Traves [9, 17] gave the values

$$\{H_d\}_{d=1}^{\infty} = \{27, 27, 72, 216, 459, 936, \ldots\}$$

and conjectured that H_d is always finite. A recursive formula for H_d (à la Kontsevich for K_d?) also remains open.

[1] S. R. Finch, Elliptic curves over $\mathbb{Q}$, *this volume*, §1.14.

[2] I. Strachan, How to count curves: from nineteenth-century problems to twenty-first-century solutions, *Royal Soc. Lond. Philos. Trans. Ser. A* 361 (2003) 2633–2647; MR2050669 (2005c:53111).

[3] L. Caporaso, Counting curves on surfaces: a guide to new techniques and results, *European Congress of Mathematics*, v. I, Proc. 1996 Budapest conf., ed. A. Balog, G. O. H. Katona, A. Recski and D. Szász, Birkhäuser, 1998, pp. 136–149; arXiv:alg-geom/9611029; MR1645804 (99j:14050).

[4] A. Gathmann, Enumerative geometry, unpublished lecture notes (2003).

[5] E. Arbarello, M. Cornalba and P. A. Griffiths, *Geometry of Algebraic Curves*, v. II, Springer-Verlag, 2011, pp. 768–769; MR2807457 (2012e:14059).

[6] S. Lanzat and M. Polyak, Counting real curves with passage/tangency conditions, *J. London Math. Soc.* 85 (2012) 838–854; arXiv:1011.1686; MR2927810.

[7] M. Kontsevich and Yu. Manin, Gromov-Witten classes, quantum cohomology, and enumerative geometry, *Comm. Math. Phys.* 164 (1994) 525–562; arXiv:hep-th/9402147; MR1291244 (95i:14049).

[8] S. Katz, *Enumerative Geometry and String Theory*, Amer. Math. Soc., 2006, pp. 1–10, 27–31, 125–143, 192–196; MR2218550 (2007i:14054).
[9] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A013587, A027363, A060041, A076912, and A238370.
[10] C. Itzykson, Counting rational curves on rational surfaces, *Internat. J. Modern Phys. B* 8 (1994) 3703–3724; MR1305168 (96e:14061).
[11] P. Di Francesco and C. Itzykson, Quantum intersection rings, *The Moduli Space of Curves*, Proc. 1994 Texel Island conf., ed. R. Dijkgraaf, C. Faber and G. van der Geer, Birkhäuser Boston, 1995, pp. 81–148; arXiv:hep-th/9412175v1; MR1363054 (96k:14041a).
[12] P. Di Francesco, Enumerative geometry from string theory, *Math. Comput. Modelling* 26 (1997) 75–96; MR1492503 (98j:14072).
[13] D. B. Grünberg and P. Moree, Sequences of enumerative geometry: congruences and asymptotics, *Experim. Math.* 17 (2008) 409–426; arXiv:math/0610286; MR2484425 (2010k:11149).
[14] J. J. O'Connor and E. F. Robertson, Cubic surfaces, MacTutor History of Mathematics Archive.
[15] B. Hunt, *The Geometry of Some Special Arithmetic Quotients*, Springer-Verlag, 1996, pp. 108–167, 302–310; MR1438547 (98c:14033).
[16] M. Kontsevich, Enumeration of rational curves via torus actions. *The Moduli Space of Curves*, Proc. 1994 Texel Island conf., ed. R. Dijkgraaf, C. Faber and G. van der Geer, Birkhäuser Boston, 1995, pp. 335–368; arXiv:hep-th/9405035; MR1363062 (97d:14077).
[17] W. Traves, Conics on the cubic surface, unpublished slide presentation (2004).

5.22 Distance-Avoiding Sets in the Plane

Fix a real number $d>0$. Let $D=\{1,d\}$ if $d\neq 1$; otherwise $D=\{1\}$ may simply be written as 1. A subset $S\subseteq\mathbb{R}^n$ is said to **avoid** D if $\|x-y\|\notin D$ for all $x,y\in S$. For example, the union of open balls of radius $1/2$ with centers in $(2\mathbb{Z})^n$ avoids the distance 1. If instead the balls have centers in $(3\mathbb{Z})^n$, then their union avoids $\{1,2\}$.

It is natural to ask about the "largest possible" S that avoids D. Let B_R denote the ball of radius R with center 0. Assuming S is Lebesgue measurable, its **density**

$$\delta(S)=\limsup_{R\to\infty}\frac{\mu(B_R\cap S)}{\mu(B_R)}$$

quantifies the asymptotic proportion of $\mathbb{R}^n$ occupied by S. We wish to know

$$m_D(\mathbb{R}^n)=\sup\{\delta(S):S\text{ is measurable and avoids }D\}.$$

The shortage of information regarding $m_D(\mathbb{R}^n)$ is surprising. Until further notice, let $n=2$ and $d=1$ for simplicity [1–3].

On the one hand, the number of $\mathbb{Z}^2$ points within B_R is $\sim\pi R^2$ [4], hence the number of $(2\mathbb{Z})^2$ points within B_R is $\sim(\pi/4)R^2$. Each open disk in our example has area $\pi/4$ and B_R has area πR^2, thus $m_1(\mathbb{R}^2)\geq\pi/16\approx 0.196$. It turns out we can do better by arranging the disks with centers according to an equilateral triangle lattice, giving $m_1(\mathbb{R}^2) \geq \pi/\left(8\sqrt{3}\right) \approx 0.227$. An additional improvement (replacing six portions of each circular circumference by linear segments) gives $m_1(\mathbb{R}^2) \geq 0.229365$. This is the best lower bound currently known [5, 6].

On the other hand, a configuration called the Moser spindle implies that $m_1(\mathbb{R}^2) \leq 2/7 \approx 0.286$ [7, 8]. Székely [9, 10] improved the upper bound to $12/43 \approx 0.279$. The best result currently known is $m_1(\mathbb{R}^2) \leq 0.258795$ via linear programming techniques [11, 12]. Erdős' conjecture that $m_1(\mathbb{R}^2) < 1/4$ seems out of reach.

Sets avoiding 1 have been studied by combinatorialists because of their association with the *measurable chromatic number* of the plane. What is the minimum number of colors $\chi_m(\mathbb{R}^2)$ required to color all points of $\mathbb{R}^2$ so that any two points at distance 1 receive distinct colors and so that points receiving the same color form Lebesgue measurable sets? It is known only that $5 \leq \chi_m(\mathbb{R}^2) \leq 7$ [13].

Let us now consider the case $n=2$ and $d=2$. The number of $(3\mathbb{Z})^2$ points within B_R is $\sim (\pi/9)\,R^2$. Each open disk in our example has area $\pi/4$ and B_R has area $\pi\,R^2$, thus $m_{1,2}(\mathbb{R}^2) \geq \pi/36 \approx 0.087$. Better lower bounds can surely be found, akin to before. We also know that $m_{1,2}(\mathbb{R}^2) \leq 2/9 \approx 0.222$ [9]. No one appears to have pursued this case further.

A more interesting problem is to allow d to vary, in an effort to determine

$$\inf_{d>0} m_{1,d}(\mathbb{R}^2).$$

One line of research gave $m_{1,\sqrt{3}}(\mathbb{R}^2) \leq 2/11 \approx 0.182$ [9], now improved to $m_{1,\sqrt{3}}(\mathbb{R}^2) \leq 0.170213$ [11]. Another direction gives $m_{1,c}(\mathbb{R}^2) \leq 0.141577$, where

$$c = \frac{j_{1,2}}{j_{1,1}} = 1.8309303282...$$

is a ratio of the first two positive zeroes of the Bessel function J_1 [14, 15]. There is no indication [11] that c is necessarily an optimal choice for d.

For $n=3$ and $d=1$, a configuration called the Moser–Raiskii spindle implies that $m_1(\mathbb{R}^3) \leq 3/14 \approx 0.214$ [8]. Székely [16] improved the upper bound to $7/37 \approx 0.189$; this was further diminished to $3/16 = 0.1875$ in [13]. The best result currently known is $m_1(\mathbb{R}^3) \leq 0.165609$ [11].

For $n=4$ and $d=1$, an early result $m_1(\mathbb{R}^4) \leq 16/125 = 0.128$ [13] was superseded later by 0.112937 [11] and more recently improved to 0.100062 [17]. Upper bounds on $m_1(\mathbb{R}^n)$ are now known up to $n=24$; lower bounds seem to be relatively neglected.

Let us return finally to a lower bound, mentioned in [13]:

$$\inf_{d>0} m_{1,d}(\mathbb{R}^2) \geq \left(\frac{1}{\chi_m(\mathbb{R}^2)}\right)^2 \geq \left(\frac{1}{7}\right)^2 = \frac{1}{49}$$

and proved in [9]. The gap between $1/49 \approx 0.02$ and ≈ 0.14 deserves to be bridged! We are hopeful that someone will accept this challenge.

An unrelated problem is as follows. Let I be a Lebesgue surface measurable subset of the unit sphere in $\mathbb{R}^3$ with the property that no two vectors in I are orthogonal. Let α denote the largest possible area of such sets I, normalized by 4π. It is known [18] that $0.2928 < \alpha < 0.313$ and the upper bound is (again) the outcome of linear programming techniques.

[1] L. A. Székely, Erdős on unit distances and the Szemerédi-Trotter theorems, *Paul Erdős and His Mathematics. II*, Proc. 1999 Budapest conf., ed. G. Halász, L. Lovász, M. Simonovits and V. T. Sós, János Bolyai Math. Soc., 2002, pp. 649–666; MR1954746 (2004a:52028).

[2] H. T. Croft, K. J. Falconer and R. K. Guy, Maximal sets avoiding certain distance configurations/The chromatic number of the plane and of space, *Unsolved Problems in Geometry*, Springer-Verlag, 1991, sect. G4/G10; MR1107516 (92c:52001).

[3] P. Brass, W. Moser and J. Pach, *Research Problems in Discrete Geometry*, Springer-Verlag, 2005, pp. 234–237; MR2163782 (2006i:52001).

[4] S. R. Finch, Sierpinski's constant: Circle and divisor problems, *first volume*, pp. 122–125.

[5] H. T. Croft, Incidence incidents, *Eureka (Cambridge)* 30 (1967) 22–26.

[6] E. R. Scheinerman and D. H. Ullman, *Fractional Graph Theory. A Rational Approach to the Theory of Graphs*, Dover Publ., 2011, pp. 45–46, 54–55; MR2963519.

[7] L. Moser and W. Moser, 7 planar points which cannot be 3-colored, *Canad. Math. Bull.* 4 (1961) 187–189.

[8] D. G. Larmon and C. A. Rogers, The realization of distances within sets in Euclidean space, *Mathematika* 19 (1972) 1–24; MR0319055 (47 #7601).

[9] L. A. Székely, Measurable chromatic number of geometric graphs and sets without some distances in Euclidean space, *Combinatorica* 4 (1984) 213–218; MR0771730 (85m:52008).

[10] L. A. Székely, Inclusion-exclusion formulae without higher terms, *Ars Combin.* 23 (1987) B, 7–20; MR0890334 (88e:05006).

[11] F. M. de Oliveira Filho and F. Vallentin, Fourier analysis, linear programming, and densities of distance avoiding sets in $\mathbb{R}^n$, *J. European Math. Soc.* 12 (2010) 1417–1428; arXiv:0808.1822; MR2734347 (2011m:42012).

[12] T. Keleti, M. Matolcsi, F. M. de Oliveira Filho and I. Z. Ruzsa, Better bounds for planar sets avoiding unit distances; *Discrete Comput. Geom.* 55 (2016) 642–661; arXiv:1501.00168; MR3473672.

[13] L. A. Székely and N. C. Wormald, Bounds on the measurable chromatic number of $\mathbb{R}^n$, *Discrete Math.* 75 (1989) 343–372; MR1001407 (90i:05040).

[14] S. R. Finch, Bessel function zeroes, *this volume*, §2.2.

[15] B. D. Kotlyar, On the packing density of a fuzzy set, *Cybernetics and Systems Analysis* 29 (1993) 648–655.

[16] L. A. Székely, Remarks on the chromatic number of geometric graphs, *Graphs and Other Combinatorial Topics. Proc. 3rd Czechoslovak Sympos. on Graph Theory*, Prague, 1982, ed. M. Fiedler, Teubner, 1983, pp. 312–315; MR0737057 (85e:05076).

[17] C. Bachoc, A. Passuello and A. Thiery, The density of sets avoiding distance 1 in Euclidean space, *Discrete Comput. Geom.* 53 (2015) 783–808; arXiv:1401.6140; MR3341578.

[18] E. DeCorte and O. Pikhurko, Spherical sets avoiding a prescribed set of angles, arXiv:1502.05030.

5.23 Fraenkel Asymmetry

For simplicity, we restrict attention to subregions of the plane. Let $\Omega \subseteq \mathbb{R}^2$ be the closure of a bounded, open, connected set of area $|\Omega|$ with piecewise continuously differentiable boundary and perimeter p. The classical isoperimetric inequality:

$$p(\Omega) \geq (4\pi\,|\Omega|)^{1/2} \quad \text{with equality iff } \Omega \text{ is a disk}$$

can be expressed as

$$\delta(\Omega) \geq 0 \quad \text{with equality iff } \Omega \text{ is a disk}$$

where the **isoperimetric deficit** is

$$\delta(\Omega) = \frac{p(\Omega)}{(4\pi\,|\Omega|)^{1/2}} - 1.$$

We wish to refine $\delta(\Omega) \geq 0$ so that the right-hand side vanishes only on disks and measures to what degree Ω deviates from a disk. Out of many possible choices, we examine **Fraenkel asymmetry** [1–3]

$$\alpha(\Omega) = \inf\left\{ \frac{|(\Omega \setminus D) \cup (D \setminus \Omega)|}{|\Omega|} : D \text{ a disk with } |D| = |\Omega| \right\}.$$

Note the symmetric difference of sets in the numerator (some authors employ $|\Omega \setminus D|$ instead, hence their results are off by a factor of 2). Before understanding best constants for the inequality $\delta(\Omega) \geq c\,\alpha(\Omega)^2$, that is, extreme values of the ratio $\delta(\Omega)/\alpha(\Omega)^2$, let us first examine $\alpha(\Omega)$ for several polygonal regions.

The Fraenkel asymmetry of a regular hexagon (side length 1) is

$$\frac{1}{3\sqrt{3}/2} \cdot 12 \int_{\sqrt{3}/2}^{\sqrt{3\sqrt{3}/(2\pi)}} \sqrt{\frac{3\sqrt{3}}{2\pi} - x^2}\, dx$$
$$= \frac{-9\sqrt{\left(2\sqrt{3}-\pi\right)\pi} + 18\sqrt{3}\arccos\left(\sqrt{\pi/\left(2\sqrt{3}\right)}\right)}{\left(3\sqrt{3}/2\right)\pi} = 0.0744657545...$$

which is quite close to zero (Figure 5.29). The square has greater asymmetry

$$16 \int_{1/2}^{1/\sqrt{\pi}} \sqrt{\frac{1}{\pi} - x^2}\, dx$$
$$= 4 - \frac{2\sqrt{(4-\pi)\pi} + 8\arcsin\left(\sqrt{\pi}/2\right)}{\pi} = 0.1810919376...$$

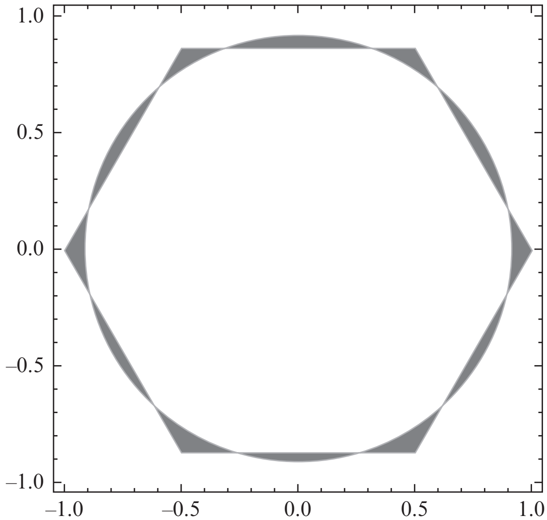

Figure 5.29 Symmetric difference between regular hexagon and Fraenkel disk.

and the equilateral triangle has still greater asymmetry

$$\frac{1}{\sqrt{3}/4}\cdot 12 \int_0^{1/4-\sqrt{3\pi(3\sqrt{3}-\pi)}/(12\pi)} \left(\left(\frac{1}{\sqrt{3}}-\sqrt{3}x\right)-\sqrt{\frac{\sqrt{3}}{4\pi}-x^2}\right)dx$$

$$=0.3649426110...$$

(omitting the exact expression, which is complicated).

Let $\ell \geq 2/\sqrt{\pi}$. If Ω is the rectangle with vertices $(\pm\ell/2, \pm 1/(2\ell))$, clearly $|\Omega| = 1$ and

$$\alpha(\Omega) = -\frac{1}{\ell^2}\sqrt{\frac{4\ell^2-\pi}{\pi}}+\frac{4}{\pi}\arcsin\left(\sqrt{\frac{4\ell^2-\pi}{4\ell^2}}\right)\to 2$$

as $\ell \to \infty$. Fraenkel asymmetry can never exceed 2; from

$$p(\Omega) = 2\left(\ell+\frac{1}{\ell}\right)\sim 2\ell$$

we deduce

$$\alpha(\Omega)\sim 2-\frac{8}{\sqrt{\pi}}\frac{1}{p}+\frac{4\sqrt{\pi}}{3}\frac{1}{p^3}.$$

This example is inefficient (in terms of perimeter) by comparison with the following.

Let $0 < \theta \leq \arctan(\pi/4)$ and

$$f(\theta) = \frac{\sqrt{\pi}}{4}\frac{\cos(\theta)^2}{\sin(\theta)}, \quad g(\theta) = \frac{1}{\sqrt{\pi}}\sin(\theta).$$

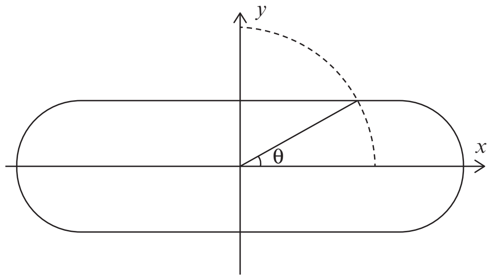

Figure 5.30 For a biscuit (or stadium or racetrack) of unit area, θ is the angle determined by the intersection between its boundary and the circle with common center, radius $1/\sqrt{\pi}$.

Consider the rectangle with vertices $(\pm f(\theta), \pm g(\theta))$, capped on the right and left by semicircles. The equation of the boundary in the first quadrant only is

$$y = \begin{cases} g(\theta) & \text{if } 0 \le x \le f(\theta), \\ \sqrt{g(\theta)^2 - (x - f(\theta))^2} & \text{if } f(\theta) < x \le f(\theta) + g(\theta). \end{cases}$$

The region Ω' in Figure 5.30, called a **biscuit**, satisfies $|\Omega'| = 1$ and [4, 5]

$$\alpha(\Omega') = \frac{2}{\pi}\left(\pi - 2\theta - 2\sin(\theta)\cos(\theta)\right) \to 2$$

as $\theta \to 0^+$. From

$$p(\Omega') = \sqrt{\pi}\,\frac{1 + \sin(\theta)^2}{\sin(\theta)} \sim \frac{\sqrt{\pi}}{\theta}$$

we deduce

$$\alpha(\Omega') \sim 2 - \frac{8}{\sqrt{\pi}}\frac{1}{p} + \frac{8\sqrt{\pi}}{3}\frac{1}{p^3}.$$

The third term when expanding $\alpha(\Omega')$ is greater than that for $\alpha(\Omega)$. These asymptotics are consistent with a theorem that, among all *convex* sets Ω of unit area and fixed perimeter

$$p \ge p_0 = \frac{2}{\sqrt{\pi}}\frac{\pi^2 + 8}{\sqrt{\pi^2 + 16}} = 3.9643784229...,$$

the biscuit maximizes α. Write $E_p = \Omega'$ for convenience. Since $\delta(\Omega) = p(4\pi)^{-1/2} - 1$ is fixed, E_p coincides with the solution of a restricted version of the earlier optimization problem.

If $2\sqrt{\pi} < p < p_0$, then the maximizing convex set E_p is called an **oval** whose boundary consists of four symmetrically placed circular arcs. We omit all details except to remark that $\arctan(\pi/4) < \theta < \pi/4$ for these. Also of interest is [5–7]

$$\min_{p > 2\sqrt{\pi}} \frac{\delta(E_p)}{\alpha(E_p)^2} = 0.4055851970... = \frac{1}{4}(1.6223407880...)$$

which is achieved for a specific biscuit. Allowing non-convex sets to enter the discussion,

$$\frac{\delta(E_{\mathrm{nc}})}{\alpha(E_{\mathrm{nc}})^2} \approx 0.39314$$

is achieved by a certain set, called a **mask**, whose boundary involves eight circular arcs. Proof of this latter new assertion has not yet appeared.

Finally, we turn to an older topic: the calculation of maximal coefficients c_k in the asymptotic estimate

$$\delta(\Omega) \geq \sum_{k=1}^{m} c_k \alpha(\Omega)^k + o\left(\alpha(\Omega)^m\right)$$

for arbitrary Ω. The fact that $c_k = 0$ for odd k and [8–10]

$$c_2 = \frac{\pi}{8(4-\pi)} = 0.4574740457... = \frac{1}{4}(1.8298961831...)$$

has been known since the 1990s; the fact that [6]

$$c_4 = -\frac{\pi^3(3\pi-14)(5\pi-16)}{96(4-\pi)^4(\pi-2)} = -0.6962146734...,$$

$$\begin{aligned} c_6 &= \frac{\pi^5(-759808+1619648\pi-1386576\pi^2+612992\pi^3-148024\pi^4+18552\pi^5-945\pi^5)}{2880(4-\pi)^7(\pi-2)^4} \\ &= -1.7607874382... \end{aligned}$$

was found only in 2013. Verification makes use of a sequence of ovals converging to the disk $(\theta \to (\pi/4)^-)$.

We witnessed two measures of asymmetry (in a different context) in [11]; Reuleaux polygons are mentioned in [12]. Yet another measure – *Hausdorff asymmetry* – is found in [13].

5.23.1 *Geometric Uncertainty Principle*

For the following, an assumption of finite perimeter is not needed, thus hypotheses may be weakened. Let $\Omega \subseteq \mathbb{R}^2$ be an open bounded region with a given decomposition

$$\Omega = \bigcup_{j=1}^{N} \Omega_j$$

into disjoint Lebesgue measurable sets Ω_j. Define the j^{th} **area deviation**

$$\sigma(\Omega_j) = \frac{|\Omega_j| - \min_{1\leq i\leq N}|\Omega_i|}{|\Omega_j|}$$

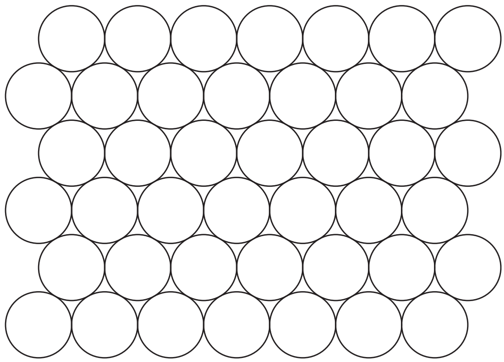

Figure 5.31 Tiling of the plane using disks and hourglasses in equal proportion.

which satisfies $0 \le \sigma(\Omega_j) \le 1$ and, like $\alpha(\Omega_j)$, is scale-invariant. Steinerberger [14] proved the remarkable existence of a universal constant $\kappa > 0$ such that, for sufficiently large N depending only on Ω, the sum

$$\left(\sum_{j=1}^{N} \frac{|\Omega_j|}{|\Omega|}\alpha(\Omega_j)\right) + \left(\sum_{j=1}^{N} \frac{|\Omega_j|}{|\Omega|}\sigma(\Omega_j)\right) \ge \kappa.$$

It is known that κ is at least $1/60000$ and conjectured that $\kappa = 0.0744657545...$, which corresponds to the regular hexagonal tiling of the plane. Another candidate tiling of the plane – Kepler's circle packing with exactly one adjacent **hourglass** per disk (Figure 5.31) – gives a considerably larger sum.

5.23.2 *Bisecting Chords*

As an aside, given a planar measurable convex set Ω, a **bisecting chord** is a line segment whose endpoints lie on the boundary of Ω and which partitions Ω into two subsets of equal area. For example, a disk D of radius $1/2$ possesses infinitely many bisecting chords, all of length 1. The area of such a disk is $\pi/4 = 0.7853981633....$ For most sets Ω, we expect bisecting chord lengths to vary. Suppose Ω has the property that its maximum bisecting chord length is 1. How small can the area of such a set Ω be? Is D the area-minimizing set Ω?

The answer to the second question is no. Define the **Auerbach triangle** Δ (or rounded triangle) to consist of six parts, three linear and three nonlinear, with the topmost part (the dashed curve in Figure 5.32) given parametrically by [15–17]

$$x(t) = \frac{e^{4t}-1}{e^{4t}+1} - t, \quad y(t) = 2\frac{e^{2t}}{e^{4t}+1}, \quad -\frac{\ln(3)}{4} \le t \le \frac{\ln(3)}{4}.$$

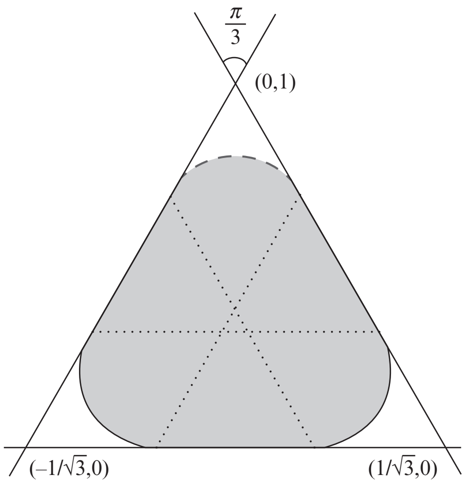

Figure 5.32 Auerbach triangle with unit bisecting (halving) chords.

Then Δ satisfies the required property, but its area is

$$\frac{\sqrt{3}}{8}\left(8\ln(3)-\ln(3)^2-4\right)=0.7755147827...=\frac{1}{4}(3.1020591308...)<\frac{\pi}{4}.$$

This numerical value is the answer to the first question. A third question is: How large can the perimeter of such a set Ω be? Note that the perimeter of Δ is $3\ln(3)=3.2958368660...>\pi$ and Δ evidently is the perimeter-maximizing set Ω as well. Related materials include [18–23].

Addendum Let Ω be the ellipse $x^2/\ell^2+\ell^2y^2\leq 1/\pi$ and Ω' be the rhombus with vertices $(\pm\ell,0)$, $(0,\pm 1/(2\ell))$. Clearly $|\Omega|=|\Omega'|=1$ and

$$\alpha(\Omega)=\frac{4}{\pi}\left[\arcsin\left(\frac{\ell}{\sqrt{1+\ell^2}}\right)-\arcsin\left(\frac{1}{\sqrt{1+\ell^2}}\right)\right],$$

$$\alpha(\Omega')=8\int\limits_0^{\xi}\left[\sqrt{\frac{1}{\pi}-x^2}-\frac{1}{2\ell^2}(\ell-x)\right]dx$$

where

$$\xi=\frac{\ell}{1+4\ell^4}+\frac{2\ell^2\sqrt{1+(4\ell^2-\pi)\ell^2}}{(1+4\ell^4)\sqrt{\pi}}$$

(the exact expression for $\alpha(\Omega')$ is complicated). From

$$p(\Omega)=\frac{4\ell}{\sqrt{\pi}}\int\limits_0^{\pi/2}\sqrt{1-\left(1-\frac{1}{\ell^4}\right)\cos(\theta)^2}\,d\theta\sim\frac{4\ell}{\sqrt{\pi}}$$

(an elliptic integral of the second kind) and

$$p(\Omega') = 4\sqrt{\ell^2 + \frac{1}{4\ell^2}} \sim 4\ell$$

we deduce that, as $\ell \to \infty$,

$$\alpha(\Omega) \sim 2 - \frac{32}{\pi^{3/2}}\frac{1}{p}, \quad \alpha(\Omega') \sim 2 - \frac{16}{\sqrt{\pi}}\frac{1}{p}$$

which again are inefficient by comparison with a biscuit. More computations of Fraenkel asymmetry are found in [24], related to the study of various *triangle centers* [25, 26].

Acknowledgment I am thankful to Stefan Steinerberger for his generous correspondence and kind help.

[1] N. Fusco, F. Maggi and A. Pratelli, The sharp quantitative isoperimetric inequality, *Annals of Math.* 168 (2008) 941–980; MR2456887 (2009k:52021).
[2] F. Maggi, Some methods for studying stability in isoperimetric type problems, *Bull. Amer. Math. Soc.* 45 (2008) 367–408; MR2402947 (2009b:49105).
[3] M. Cicalese and G. P. Leonardi, A selection principle for the sharp quantitative isoperimetric inequality, *Arch. Ration. Mech. Anal.* 206 (2012) 617–643; arXiv:1007.3899; MR2980529.
[4] S. Campi, Isoperimetric deficit and convex plane sets of maximum translative discrepancy, *Geom. Dedicata* 43 (1992) 71–81; MR1169365 (93d:52012).
[5] A. Alvino, V. Ferone and C. Nitsch, A sharp isoperimetric inequality in the plane, *J. European Math. Soc.* 13 (2011) 185–206; MR2735080 (2011k:52007).
[6] M. Cicalese and G. P. Leonardi, Best constants for the isoperimetric inequality in quantitative form, *J. European Math. Soc.* 15 (2013) 1101–1129; arXiv:1101.0169; MR3085102.
[7] C. Bianchini, G. Croce and A. Henrot, On the quantitative isoperimetric inequality in the plane, arXiv:1507.08189.
[8] R. R. Hall, W. K. Hayman and A. W. Weitsman, On asymmetry and capacity, *J. d'Analyse Math.* 56 (1991) 87–123; MR1243100 (95h:31004).
[9] R. R. Hall, A quantitative isoperimetric inequality in n-dimensional space, *J. Reine Angew. Math.* 428 (1992) 161–176; MR1166511 (93d:51041).
[10] R. R. Hall and W. K. Hayman, A problem in the theory of subordination, *J. d'Analyse Math.* 60 (1993) 99–111; MR1253231 (94m:42010).
[11] S. R. Finch, Reuleaux triangle constants, *first volume*, pp. 513–515.
[12] F. Maggi, M. Ponsiglione and A. Pratelli, Quantitative stability in the isodiametric inequality via the isoperimetric inequality, *Trans. Amer. Math. Soc.* 366 (2014) 1141–1160; arXiv:1104.4074; MR3145725.
[13] A. Alvino, V. Ferone and C. Nitsch, A sharp isoperimetric inequality in the plane involving Hausdorff distance, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* 20 (2009) 397–412; MR2550854 (2010i:52012).
[14] S. Steinerberger, A geometric uncertainty principle with an application to Pleijel's estimate, *Annales Henri Poincaré* 15 (2014) 2299–2319; arXiv:1306.3103; MR3272823.

[15] N. Fusco and A. Pratelli, On a conjecture by Auerbach, *J. European Math. Soc.* 13 (2011) 1633–1676; MR2835326 (2012h:52021).

[16] L. Esposito, V. Ferone, B. Kawohl, C. Nitsch and C. Trombetti, The longest shortest fence and sharp Poincaré-Sobolev inequalities, *Arch. Ration. Mech. Anal.* 206 (2012) 821–851; arXiv:1011.6248; MR2989444.

[17] A. Dumitrescu, A. Ebbers-Baumann, A. Grüne, R. Klein and G. Rote, On the geometric dilation of closed curves, graphs, and point sets, *Comput. Geom.* 36 (2007) 16–38; MR2264047 (2008g:52024).

[18] H. T. Croft, K. J. Falconer and R. K. Guy, Dividing up a piece of land by a short fence, *Unsolved Problems in Geometry*, Springer-Verlag, 1991, sect. A26; MR1107516 (92c:52001).

[19] P. Goodey, Area and perimeter bisectors of planar convex sets, *Integral Geometry and Convexity*, Proc. 2004 Wohan conf., ed. E. L. Grinberg, S. Li, G. Zhang and J. Zhou, World Sci. Publ., 2006, 29–35; MR2240971 (2007e:52010).

[20] A. Ebbers-Baumann, A. Grüne and R. Klein, Geometric dilation of closed planar curves: new lower bounds, *Comput. Geom.* 37 (2007) 188–208; MR2331064 (2008b:52023).

[21] A. Ebbers-Baumann, A. Grüne and R. Klein, On the geometric dilation of finite point sets, *Algorithms and Computation*, Proc. 14^{th} Internat. Sympos. (ISAAC 2003), Kyoto, 2003, ed. T. Ibaraki, N. Katoh and H. Ono, Lect. Notes in Comp. Sci. 2906, Springer-Verlag, 2003, pp. 250–259; MR2087565 (2005d:68136).

[22] A. Ebbers-Baumann, A. Grüne and R. Klein, The geometric dilation of finite point sets, *Algorithmica* 44 (2006) 137–149; MR2194472 (2006m:68159).

[23] S. Steinerberger, A remark on disk packings and numerical integration of harmonic functions, arXiv:1403.8002.

[24] S. R. Finch, In limbo: Three triangle centers, arXiv:1406.0836.

[25] S. R. Finch, Least capacity point of triangles, arXiv:1407.4105.

[26] S. R. Finch, Appell F1 and conformal mapping, arXiv:1408.1074.

Index