] N. S. Landkoff, *Foundations of Modern Potential Theory*, Springer-Verlag, 1972; MR 50 #2520.
] T. J. Ransford, *Potential Theory in the Complex Plane*, Cambridge Univ. Press, 1995; MR 96e:31001.
] J. Rostand, Computing logarithmic capacity with linear programming, *Experim. Math.* 6 (1997) 221–238; MR 98f:65032.
] L. V. Ahlfors, *Conformal Invariants: Topics in Geometric Function*, McGraw-Hill, 1973; MR 50 #10211.
] J. B. Conway, *Functions of One Complex Variable II*, Springer-Verlag, 1995; MR 96i:30001.
] P. K. Kythe, *Computational Conformal Mapping*, Birkhäuser, 1998; MR 99k:65027.

5

Constants Associated with Enumerating Discrete Structures

5.1 Abelian Group Enumeration Constants

Every finite abelian group is a direct sum of cyclic subgroups. A corollary of this fundamental theorem is the following. Given a positive integer n, the number $a(n)$ of non-isomorphic abelian groups of order n is given by [1,2]

$$a(n) = P(\alpha_1)P(\alpha_2)P(\alpha_3)\cdots P(\alpha_r),$$

where $n = p_1^{\alpha_1}p_2^{\alpha_2}p_3^{\alpha_3}\cdots p_r^{\alpha_r}$ is the prime factorization of n, $p_1, p_2, p_3, \ldots, p_r$ are distinct primes, each α_k is positive, and $P(\alpha_k)$ denotes the number of unrestricted partitions of α_k. For example, $a(p^4) = 5$ for any prime p since there are five partitions of 4:

$$4 = 1 + 3 = 2 + 2 = 1 + 1 + 2 = 1 + 1 + 1 + 1.$$

As another example, $a(p^4q^4) = 25$ for any distinct primes p and q, but $a(p^8) = 22$.

It is clear that

$$\liminf_{n\to\infty} a(n) = 1,$$

but it is more difficult to see that [3–6]

$$\limsup_{n\to\infty} \ln(a(n))\frac{\ln(\ln(n))}{\ln(n)} = \frac{\ln(5)}{4}.$$

A number of authors have examined the average behavior of $a(n)$ over all positive integers. The most precise known results are [7–10]

$$\sum_{n=1}^{N} a(n) = A_1N + A_2N^{\frac{1}{2}} + A_3N^{\frac{1}{3}} + O\left(N^{\frac{50}{199}+\varepsilon}\right),$$

where $\varepsilon > 0$ is arbitrarily small,

$$A_k = \prod_{\substack{j=1 \\ j \neq k}}^{\infty} \zeta\left(\frac{j}{k}\right) = \begin{cases} 2.2948565916\ldots & \text{if } k=1, \\ -14.6475663016\ldots & \text{if } k=2, \\ 118.6924619727\ldots & \text{if } k=3, \end{cases}$$

and $\zeta(x)$ is Riemann's zeta function [1.6]. We cannot help but speculate about the following estimate:

$$\sum_{n=1}^{N} a(n) \sim \sum_{k=1}^{\infty} A_k N^{\frac{1}{k}} + \Delta(N),$$

but an understanding of the error $\Delta(N)$ has apparently not yet been achieved [11, 12]. Similar enumeration results for finite semisimple associative rings appear in [5.1.1].

If, instead, focus is shifted to the sum of the reciprocals of $a(n)$, then [13, 14]

$$\sum_{n=1}^{N} \frac{1}{a(n)} = A_0 N + O\left(N^{\frac{1}{2}} \ln(N)^{-\frac{1}{2}}\right),$$

where A_0 is an infinite product over all primes p:

$$A_0 = \prod_p \left[1 - \sum_{k=2}^{\infty} \left(\frac{1}{P(k-1)} - \frac{1}{P(k)}\right) \frac{1}{p^k}\right] = 0.7520107423\ldots.$$

In summary, the average number of non-isomorphic abelian groups of any given order is $A_1 = 2.2948$ if "average" is understood in the sense of arithmetic mean, and $A_0^{-1} = 1.3297$ if "average" is understood in the sense of harmonic mean. We cannot even hope to obtain analogous statistics for the general (not necessarily abelian) case at present. Some interesting bounds are known [15–19] and are based on the classification theorem of finite simple groups.

The constant A_1 also appears in [20] in connection with the arithmetical properties of class numbers of quadratic fields.

Erdös & Szekeres [21, 22] examined $a(n)$ and the following generalization: $a(n, i)$ is the number of representations of n as a product (of an arbitrary number of terms, with order ignored) of factors of the form p^j, where $j \geq i$. They proved that

$$\sum_{n=1}^{N} a(n, i) = C_i N^{\frac{1}{i}} + O\left(N^{\frac{1}{i+1}}\right), \quad \text{where } C_i = \prod_{k=1}^{\infty} \zeta\left(1 + \frac{k}{i}\right),$$

and surely someone has tightened this estimate by now. See also the discussion of square-full and cube-full integers in [2.6.1].

5.1.1 Semisimple Associative Rings

A finite associative ring R with identity element $1 \neq 0$ is said to be **simple** if R has no proper (two-sided) ideals and is **semisimple** if R is a direct sum of simple ideals.

Simple rings generalize fields. Semisimple rings, in turn, generalize simple rings. Every (finite) semisimple ring is, in fact, a direct sum of full matrix rings over finite fields. Consequently, given a positive integer n, the number $s(n)$ of non-isomorphic semisimple rings of order n is given by

$$s(n) = Q(\alpha_1) Q(\alpha_2) Q(\alpha_3) \cdots Q(\alpha_r),$$

where $n = p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \cdots p_r^{\alpha_r}$ is the prime factorization of n, $p_1, p_2, p_3, \ldots, p_r$ are distinct primes, each α_k is positive, and $Q(\alpha_k)$ denotes the number of (unordered) sets of integer pairs (r_j, m_j) for which

$$\alpha_k = \sum_j r_j m_j^2 \text{ and } r_j m_j^2 > 0 \text{ for all } j.$$

As an example, $s(p^5) = 8$ for any prime p since

$$\begin{aligned} 5 &= 1 \cdot 1^2 + 1 \cdot 2^2 = 5 \cdot 1^2 = 2 \cdot 1^2 + 3 \cdot 1^2 = 1 \cdot 1^2 + 4 \cdot 1^2 \\ &= 1 \cdot 1^2 + 1 \cdot 1^2 + 3 \cdot 1^2 = 1 \cdot 1^2 + 2 \cdot 1^2 + 2 \cdot 1^2 \\ &= 1 \cdot 1^2 + 1 \cdot 1^2 + 1 \cdot 1^2 + 2 \cdot 1^2 = 1 \cdot 1^2 + 1 \cdot 1^2 + 1 \cdot 1^2 + 1 \cdot 1^2 + 1 \cdot 1^2. \end{aligned}$$

Asymptotically, there are extreme results [23, 24]:

$$\liminf_{n \to \infty} s(n) = 1,$$

$$\limsup_{n \to \infty} \ln(s(n)) \frac{\ln(\ln(n))}{\ln(n)} = \frac{\ln(6)}{4}$$

and average results [25–30]:

$$\sum_{n=1}^{N} s(n) = A_1 B_1 N + A_2 B_2 N^{\frac{1}{2}} + A_3 B_3 N^{\frac{1}{3}} + O\left(N^{\frac{50}{199}+\varepsilon}\right),$$

where $\varepsilon > 0$ is arbitrarily small, A_k is as defined in the preceding, and

$$B_k = \prod_{r=1}^{\infty} \prod_{m=2}^{\infty} \zeta\left(\frac{rm^2}{k}\right).$$

In particular, there are, on average,

$$A_1 B_1 = \prod_{rm^2 > 1} \zeta\left(rm^2\right) = 2.4996161129\ldots$$

non-isomorphic semisimple rings of any given order ("average" in the sense of arithmetic mean).

[1] D. G. Kendall and R. A. Rankin, On the number of abelian groups of a given order, *Quart. J. Math.* 18 (1947) 197–208; MR 9,226c.

[2] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000688 and A004101.

[3] E. Krätzel, Die maximale Ordnung der Anzahl der wesentlich verschiedenen abelschen Gruppen n^{ter} Ordnung, *Quart. J. Math.* 21 (1970) 273–275; MR 42 #3171.

[4] W. Schwarz and E. Wirsing, The maximal number of non-isomorphic abelian groups of order n, *Arch. Math. (Basel)* 24 (1973) 59–62; MR 47 #4953.
[5] A. Ivić, The distribution of values of the enumerating function of non-isomorphic abelian groups of finite order, *Arch. Math. (Basel)* 30 (1978) 374–379; MR 58 #16562.
[6] E. Krätzel, The distribution of values of $a(n)$, *Arch. Math. (Basel)* 57 (1991) 47–52; MR 92e:11097.
[7] H.-E. Richert, Über die Anzahl Abelscher Gruppen gegebener Ordnung. I, *Math. Z.* 56 (1952) 21–32; MR 14,349e.
[8] B. R. Srinivasan, On the number of abelian groups of a given order, *Acta Arith.* 23 (1973) 195–205; MR 49 #2610.
[9] G. Kolesnik, On the number of abelian groups of a given order, *J. Reine Angew. Math.* 329 (1981) 164–175; MR 83b:10055.
[10] H.-Q. Liu, On the number of abelian groups of a given order, *Acta Arith.* 59 (1991) 261–277; supplement 64 (1993) 285–296; MR 92i:11103 and MR 94g:11073.
[11] E. Krätzel, *Lattice Points*, Kluwer, 1988, pp. 293–303; MR 90e:11144.
[12] A. Ivić, *The Riemann Zeta-Function*, Wiley, 1985, pp. 37–38, 413–420, 438–439; MR 87d:11062.
[13] 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; MR 82a:10047.
[14] P. Zimmermann, Estimating the infinite product A_0, unpublished note (1996).
[15] P. M. Neumann, An enumeration theorem for finite groups, *Quart. J. Math.* 20 (1969) 395–401; MR 40 #7344.
[16] L. Pyber, Enumerating finite groups of given order, *Annals of Math.* 137 (1993) 203–220; MR 93m:11097.
[17] M. Ram Murty, Counting finite groups, *Analysis, Geometry and Probability*, ed. R. Bhatia, Hindustan Book Agency, 1996, pp. 161–172; MR 98j:11076.
[18] L. Pyber, Group enumeration and where it leads us, *European Congress of Mathematics*, v. 2, Budapest, 1996, ed. A. Balog, G. O. H. Katona, A. Recski, and D. Szász, Birkhäuser, 1998, pp. 187–199; MR 99i:20037.
[19] B. Eick and E. A. O'Brien, Enumerating p-groups, *J. Austral. Math. Soc. Ser. A* 67 (1999) 191–205; MR 2000h:20033.
[20] H. Cohen, *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993, pp. 290–293; MR 94i:11105.
[21] P. Erdös and G. Szekeres, Über die Anzahl der Abelschen Gruppen gegebener Ordnung und über ein verwandtes zahlentheoretisches Problem, *Acta Sci. Math. (Szeged)* 7 (1934-35) 95–102.
[22] I. Z. Ruzsa, Erdös and the integers, *J. Number Theory* 79 (1999) 115–163; MR 2002e:11002.
[23] J. Knopfmacher, Arithmetical properties of finite rings and algebras, and analytic number theory. IV, *J. Reine Angew. Math.* 270 (1974) 97–114; MR 51 #389.
[24] J. Knopfmacher, Arithmetical properties of finite rings and algebras, and analytic number theory. VI, *J. Reine Angew. Math.* 277 (1975) 45–62; MR 52 #3088.
[25] J. Knopfmacher, Arithmetical properties of finite rings and algebras, and analytic number theory. I, *J. Reine Angew. Math.* 252 (1972) 16–43; MR 47 #1769.
[26] J. Knopfmacher, *Abstract Analytic Number Theory*, North-Holland, 1975; MR 91d:11110.
[27] J. Duttlinger, Eine Bemerkung zu einer asymptotischen Formel von Herrn Knopfmacher, *J. Reine Angew. Math.* 266 (1974) 104–106; MR 49 #2605.
[28] C. Calderón and M. J. Zárate, The number of semisimple rings of order at most x, *Extracta Math.* 7 (1992) 144–147; MR 94m:11111.
[29] M. Kühleitner, Comparing the number of abelian groups and of semisimple rings of a given order, *Math. Slovaca* 45 (1995) 509–518; MR 97d:11143.
[30] W. G. Nowak, On the value distribution of a class of arithmetic functions, *Comment. Math. Univ. Carolinae* 37 (1996) 117–134; MR 97h:11104.

5.2 Pythagorean Triple Constants

The positive integers a, b, c are said to form a **primitive Pythagorean triple** if $a \le b$, $\gcd(a, b, c) = 1$, and $a^2 + b^2 = c^2$. Clearly any such triple can be interpreted geometrically as the side lengths of a right triangle with commensurable sides. Define $P_h(n)$, $P_p(n)$, and $P_a(n)$ respectively as the number of primitive Pythagorean triples whose hypotenuses, perimeters, and areas do not exceed n. D. N. Lehmer [1] showed that

$$\lim_{n\to\infty} \frac{P_h(n)}{n} = \frac{1}{2\pi}, \quad \lim_{n\to\infty} \frac{P_p(n)}{n} = \frac{\ln(2)}{\pi^2}$$

and Lambek & Moser [2] showed that

$$\lim_{n\to\infty} \frac{P_a(n)}{\sqrt{n}} = C = \frac{1}{\sqrt{2\pi^5}}\Gamma\left(\frac{1}{4}\right)^2 = 0.5313399499\ldots,$$

where $\Gamma(x)$ is the Euler gamma function [1.5.4].

What can be said about the error terms? D. H. Lehmer [3] demonstrated that

$$P_p(n) = \frac{\ln(2)}{\pi^2}n + O\left(n^{\frac{1}{2}}\ln(n)\right),$$

and Lambek & Moser [2] and Wild [4] further demonstrated that

$$P_h(n) = \frac{1}{2\pi}n + O\left(n^{\frac{1}{2}}\ln(n)\right), \quad P_a(n) = Cn^{\frac{1}{2}} - Dn^{\frac{1}{3}} + O\left(n^{\frac{1}{4}}\ln(n)\right),$$

where

$$D = -\frac{1+2^{-\frac{1}{3}}}{1+4^{-\frac{1}{3}}}\frac{\zeta(\frac{1}{3})}{\zeta(\frac{4}{3})} = 0.2974615529\ldots$$

and $\zeta(x)$ is the Riemann zeta function [1.6]. Sharper estimates for $P_a(n)$ were obtained in [5–8].

It is obvious that the hypotenuse c and the perimeter $a + b + c$ of a primitive Pythagorean triple a, b, c must both be integers. If ab was odd, then both a and b would be odd and hence $c^2 \equiv 2 \bmod 4$, which is impossible. Thus the area $ab/2$ must also be an integer. If $P_a'(n)$ is the number of primitive Pythagorean triples whose areas $\le n$ are integers, then $P_a'(n) = P_a(n)$. Such an identity does not hold for non-right triangles, of course.

A somewhat related matter is the ancient **congruent number problem** [9], the solution of which Tunnell [10] has reduced to a weak form of the Birch–Swinnerton–Dyer conjecture from elliptic curve theory. In the congruent number problem, the right triangles are permitted to have rational sides (rather than just integer sides). For a prescribed integer n, does there exist a rational right triangle with area n?

There is also the problem of enumerating **primitive Heronian triples**, equivalently, coprime integers $a \le b \le c$ that are side lengths of an *arbitrary* triangle with commensurable sides. What can be said asymptotically about the numbers $H_h(n)$, $H_p(n)$, $H_a(n)$, and $H_a'(n)$ (analogously defined)? A starting point for answering this question might be [11, 12].

[1] D. N. Lehmer, Asymptotic evaluation of certain totient sums, *Amer. J. Math.* 22 (1900) 293–335.
[2] J. Lambek and L. Moser, On the distribution of Pythagorean triples, *Pacific J. Math.* 5 (1955) 73–83; MR 16,796h.
[3] D. H. Lehmer, A conjecture of Krishnaswami, *Bull. Amer. Math. Soc.* 54 (1948) 1185–1190; MR 10,431c.
[4] R. E. Wild, On the number of primitive Pythagorean triangles with area less than n, *Pacific J. Math.* 5 (1955) 85–91; MR 16,797a.
[5] J. Duttlinger and W. Schwarz, Über die Verteilung der pythagoräischen Dreiecke, *Colloq. Math.* 43 (1980) 365–372; MR 83e:10018.
[6] H. Menzer, On the number of primitive Pythagorean triangles, *Math. Nachr.* 128 (1986) 129–133; MR 87m:11022.
[7] W. Müller, W. G. Nowak, and H. Menzer, On the number of primitive Pythagorean triangles, *Annales Sci. Math. Québec* 12 (1988) 263–273; MR 90b:11020.
[8] W. Müller and W. G. Nowak, Lattice points in planar domains: Applications of Huxley's 'Discrete Hardy-Littlewood method,' *Number-Theoretic Analysis: Vienna 1988-89*, ed. E. Hlawka and R. F. Tichy, Lect. Notes in Math. 1452, Springer-Verlag, 1990, pp. 139–164; MR 92d:11113.
[9] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Springer-Verlag, 1984; MR 94a:11078.
[10] J. B. Tunnell, A classical Diophantine problem and modular forms of weight 3/2, *Invent. Math.* 72 (1983) 323–334; MR 85d:11046.
[11] C. P. Popovici, Heronian triangles (in Russian), *Rev. Math. Pures Appl.* 7 (1962) 439–457; MR 31 #121.
[12] D. Singmaster, Some corrections to Carlson's "Determination of Heronian triangles," *Fibonacci Quart.* 11 (1973) 157–158; MR 45 #156 and MR 47 #4922.

5.3 Rényi's Parking Constant

Consider the one-dimensional interval $[0, x]$ with $x > 1$. Imagine it to be a street for which parking is permitted on one side. Cars of unit length are one-by-one parked *completely at random* on the street and obviously no overlap is allowed with cars already in place. What is the mean number, $M(x)$, of cars that can fit?

Rényi [1–3] determined that $M(x)$ satisfies the following integrofunctional equation:

$$M(x) = \begin{cases} 0 & \text{if } 0 \le x < 1, \\ 1 + \dfrac{2}{x-1} \displaystyle\int_0^{x-1} M(t)\,dt & \text{if } x \ge 1. \end{cases}$$

By a Laplace transform technique, Rényi proved that the limiting mean density, m, of cars in the interval $[0, x]$ is

$$m = \lim_{x\to\infty} \frac{M(x)}{x} = \int_0^\infty \beta(x)\,dx = 0.7475979202\ldots,$$

where

$$\beta(x) = \exp\left(-2\int_0^x \frac{1-e^{-t}}{t}\,dt\right) = e^{-2(\ln(x)-\mathrm{Ei}(-x)+\gamma)}, \quad \alpha(x) = m - \int_0^x \beta(t)\,dt,$$

γ is the Euler–Mascheroni constant [1.5], and Ei is the exponential integral [6.2.1]. Several alternative proofs appear in [4, 5].

What can be said about the variance, $V(x)$, of the number of cars that can fit on the street? Mackenzie [6], Dvoretzky & Robbins [7], and Mannion [8, 9] independently addressed this question and deduced that

$$v = \lim_{x\to\infty} \frac{V(x)}{x} = 4\int_0^\infty \left[e^{-x}(1-e^{-x})\frac{\alpha(x)}{x} - e^{-2x}(x+e^{-x}-1)\frac{\alpha(x)^2}{\beta(x)x^2}\right] dx - m$$
$$= 0.0381563991\ldots.$$

A central limit theorem holds [7], that is, the total number of cars is approximately normally distributed with mean mx and variance vx for large enough x.

It is natural to consider the parking problem in a higher dimensional setting. Consider the two-dimensional rectangle of length $x > 1$ and width $y > 1$ and imagine cars to be unit squares with sides parallel to the sides of the parking rectangle. What is the mean number, $M(x, y)$, of cars that can fit? Palasti [10–12] conjectured that

$$\lim_{x\to\infty}\lim_{y\to\infty} \frac{M(x,y)}{xy} = m^2 = (0.7475979202\ldots)^2 = 0.558902\ldots.$$

Despite some determined yet controversial attempts at analysis [13, 14], the conjecture remains unproven. The mere existence of the limiting parking density was shown only recently [15]. Intensive computer simulation [16–18] suggests, however, that the conjecture is false and the true limiting value is $0.562009\ldots$.

Here is a variation in the one-dimensional setting. In Rényi's problem, a car that lands in a parking position overlapping with an earlier car is discarded. Solomon [14, 19–21] studied a revised rule in which the car "rolls off" the earlier car immediately to the left or to the right, whichever is closer. It is then parked if there exists space for it; otherwise it is discarded. The mean car density is larger:

$$m = \int_0^\infty (2x+1)\exp\left[-2(x+e^{-x}-1)\right]\beta(x)\,dx = 0.8086525183\ldots$$

since cars are permitted greater flexibility to park bumper to bumper. If Rényi's problem is thought of as a model for sphere packing in a three-dimensional volume, then Solomon's variation corresponds to packing with "shaking" allowed for the spheres to settle, hence creating more space for additional spheres.

Another variation involves random car lengths [22, 23]. If the left and right endpoints of the k^{th} arriving car are taken as the smaller and larger of two independent uniform draws from $[0, x]$, then the asymptotic expected number of cars successfully parked is $C \cdot k^{(\sqrt{17}-3)/4}$, where [24, 25]

$$C = \left(1 - \frac{1}{2^{(\sqrt{17}-1)/4}}\right)\sqrt{\pi}\,\frac{\Gamma\left(\frac{\sqrt{17}}{2}\right)}{\Gamma\left(\frac{\sqrt{17}+1}{2}\right)\Gamma\left(\frac{\sqrt{17}+3}{2}\right)^2} = 0.9848712825\ldots$$

and Γ is the gamma function [1.5.4]. Note that x is only a scale factor in this variation and does not figure in the result.

Applications of the parking problem (or, more generally, the sequential packing or space-filling problem) include such widely separated disciplines as:

- Physics: models of liquid structure [26–29];
- Chemistry: adsorption of a fluid film on a crystal surface [5.3.1];
- Monte Carlo methods: evaluation of definite integrals [30];
- Linguistics: frequency of one-syllable, length-n English words [31];
- Sociology: models of elections in Japan and lengths of gaps generated in parking problems [32–35];
- Materials science: intercrack distance after multiple fracture of reinforced concrete [36];
- Computer science: optimal data placement on a CD [37] and linear probing hashing [38].

See also [39–41]. Note the similarities in formulation between the Golomb–Dickman constant [5.4] and the Rényi constant.

5.3.1 Random Sequential Adsorption

Consider the case in which the interval $[0, x]$ is replaced by the discrete finite linear lattice $1, 2, 3, \ldots, n$. Each car is a line segment of unit length and covers two lattice points when it parks. No car is permitted to touch points that have already been covered. The process stops when no adjacent pairs of lattice points are left uncovered. It can be proved that, as $n \to \infty$ [19, 42–45],

$$m = \frac{1 - e^{-2}}{2} = 0.4323323583\ldots, \quad v = e^{-4} = 0.0183156388\ldots,$$

both of which are smaller than their continuous-case counterparts. The two-dimensional discrete analog involves unit square cars covering four lattice points, and is analytically intractable just like the continuous case. Palasti's conjecture appears to be false here too: The limiting mean density in the plane is not $m^2 = 0.186911\ldots$ but rather $0.186985\ldots$ [46–48].

For simplicity's sake, we refer to the infinite linear lattice $1, 2, 3, \ldots$ as the $1 \times \infty$ strip. The $2 \times \infty$ strip is the infinite ladder lattice with two parallel lines and crossbeams, the $3 \times \infty$ strip is likewise with three parallel lines, and naturally the $\infty \times \infty$ strip is the infinite square lattice. Thus we have closed-form expressions for m and v on $1 \times \infty$, but only numerical corrections to Palasti's estimate on $\infty \times \infty$.

If a car is a unit line segment (**dimer**) on the $2 \times \infty$ strip, then the mean car density is $\frac{1}{2}(0.91556671\ldots)$. If instead the car is on the $\infty \times \infty$ strip, then the corresponding mean density is $\frac{1}{2}(0.90682\ldots)$ [49–55]. Can exact formulas be found for these two quantities?

If the car is a line segment of length two (linear **trimer**) on the $1 \times \infty$ strip, then the mean density of vacancies is $\mu(3) = 0.1763470368\ldots$, where [6, 56–58]

$$\mu(r) = 1 - r \int_0^1 \exp\left(-2 \sum_{k=1}^{r-1} \frac{1 - x^k}{k}\right) dx.$$

More generally, $\mu(r)$ is the mean density of vacancies for linear r-mers on the $1 \times \infty$ strip, for any integer $r \geq 2$. A corresponding formula for the variance is not known.

Now suppose that the car is a single particle and that no other cars are allowed to park in any adjacent lattice points (**monomer** with **nearest neighbor exclusion**). The mean car density for the $1 \times \infty$ strip is $m_1 = \frac{1}{2}(1 - e^{-2})$ as before, of course. The mean densities for the $2 \times \infty$ and $3 \times \infty$ strips are [59–61]

$$m_2 = \frac{2 - e^{-1}}{4} = 0.4080301397\ldots, \quad m_3 = \frac{1}{3} = 0.3333333333\ldots,$$

and the corresponding density for the $\infty \times \infty$ strip is $m_\infty = 0.364132\ldots$ [47, 48, 50, 53, 55, 62]. Again, can exact formulas for m_4 or m_∞ be found?

The continuous case can be captured from the discrete case by appropriate limiting arguments [6, 58, 63]. Exhaustive surveys of random sequential adsorption models are provided in [64–66].

[1] A. Rényi, On a one-dimensional problem concerning random space-filling (in Hungarian), *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 3 (1958) 109–127; Engl. transl. in *Selected Transl. Math. Statist. and Probab.*, v. 4, Inst. Math. Stat. and Amer. Math. Soc., 1963, pp. 203–224; also in *Selected Papers of Alfréd Rényi*, v. 2, Akadémiai Kiadó, 1976, pp. 173–188; MR 21 #3039.

[2] M. Lal and P. Gillard, Evaluation of a constant associated with a parking problem, *Math. Comp.* 28 (1974) 561–564; MR 49 #6560.

[3] G. Marsaglia, A. Zaman, and J. C. W. Marsaglia, Numerical solution of some classical differential-difference equations, *Math. Comp.* 53 (1989) 191–201; MR 90h:65124.

[4] P. C. Hemmer, The random parking problem, *J. Stat. Phys.* 57 (1989) 865–869; MR 92e:82054.

[5] P. L. Krapivsky, Kinetics of random sequential parking on a line, *J. Stat. Phys.* 69 (1992) 135–150; MR 93h:82063.

[6] J. K. Mackenzie, Sequential filling of a line by intervals placed at random and its application to linear adsorption, *J. Chem. Phys.* 37 (1962) 723–728.

[7] A. Dvoretzky and H. Robbins, On the parking problem, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 9 (1964) 209–224; MR 30 #3488.

[8] D. Mannion, Random space-filling in one dimension, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 9 (1964) 143–154; MR 31 #1698.

[9] D. Mannion, Random packing of an interval, *Adv. Appl. Probab.* 8 (1976) 477–501; 11 (1979) 591–602; MR 55 #1401 and MR 80i:60014.

[10] I. Palasti, On some random space filling problems, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5 (1960) 353–359; MR 26 #4466.

[11] I. Palasti, On a two-dimensional random space filling problem, *Studia Sci. Math. Hungar.* 11 (1976) 247–252; MR 81d:52011.

[12] I. Palasti, A two-dimensional case of random packing and covering, *Random Fields*, v. 2, ed. J. Fritz, J. L. Lebowitz, and D. Szász, Colloq. Math. Soc. János Bolyai 27, North-Holland

[13] H. J. Weiner, Sequential random packing in the plane, *J. Appl. Probab.* 15 (1978) 803–814; letters to the editor and replies, 16 (1979) 697–707 and 17 (1980) 878–892; MR 80a:60015, MR 81b:60013, and MR 82m:60021.

[14] H. Solomon and H. J. Weiner, A review of the packing problem, *Commun. Statist. Theory Methods* 15 (1986) 2571–2607; MR 88a:60028.

[15] M. D. Penrose, Random parking, sequential adsorption, and the jamming limit, *Commun. Math. Phys.* 218 (2001) 153–176; MR 2002a:60015.

[16] B. E. Blaisdell and H. Solomon, On random sequential packing in the plane and a conjecture of Palasti, *J. Appl. Probab.* 7 (1970) 667–698; MR 43 #8101.

[17] B. J. Brosilow, R. M. Ziff, and R. D. Vigil, Random sequential adsorption of parallel squares, *Phys. Rev. A* 43 (1991) 631–638.

[18] B. Bonnier, M. Hontebeyrie, and C. Meyers, On the random filling of $\mathbb{R}^d$ by nonoverlapping d-dimensional cubes, *Physica A* 198 (1993) 1–10; MR 94g:82045.

[19] H. Solomon, Random packing density, *Proc. Fifth Berkeley Symp. Math. Stat. Probab.*, v. 3, ed. L. M. Le Cam and J. Neyman, Univ. of Calif. Press, 1967, pp. 119–134; MR 41 #1187.

[20] J. Talbot and S. M. Ricci, Analytic model for a ballistic deposition process, *Phys. Rev. Lett.* 68 (1992) 958–961.

[21] P. Viot, G. Tarjus, and J. Talbot, Exact solution of a generalized ballistic-deposition model, *Phys. Rev. E* 48 (1993) 480–488.

[22] P. E. Ney, A random interval filling problem, *Annals of Math. Statist.* 33 (1962) 702–718; MR 25 #1561.

[23] J. P. Mullooly, A one dimensional random space-filling problem, *J. Appl. Probab.* 5 (1968) 427–435; MR 38 #5261.

[24] E. G. Coffman, C. L. Mallows, and B. Poonen, Parking arcs on the circle with applications to one-dimensional communication networks, *Annals of Appl. Probab.* 4 (1994) 1098–1111; MR 95k:60245.

[25] E. G. Coffman, L. Flatto, P. Jelenkovic, and B. Poonen, Packing random intervals on-line, *Algorithmica* 22 (1998) 448–476; MR 2000h:60012.

[26] J. D. Bernal, A geometrical approach to the structure of liquids, *Nature* 183 (1959) 141–147.

[27] J. D. Bernal, Geometry of the structure of monatomic liquids, *Nature* 185 (1960) 68–70.

[28] J. D. Bernal and J. Mason, Coordination of randomly packed spheres, *Nature* 188 (1959) 910–911.

[29] J. D. Bernal, J. Mason, and K. R. Knight, Radial distribution of the random close packing of equal spheres, *Nature* 194 (1962) 956–958.

[30] G. Bánkövi, Evaluation of integrals by Monte Carlo methods based on the one-dimensional random space filling, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 5 (1960) 339–352; MR 26 #4465.

[31] J. Dolby and H. Solomon, Information density phenomena and random packing, *J. Appl. Probab.* 12 (1975) 364–370.

[32] Y. Itoh and S. Ueda, A random packing model for elections, *Annals Inst. Statist. Math.* 31 (1979) 157–167.

[33] G. Bánkövi, On gaps generated by a random space filling procedure, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* 7 (1962) 395–407; MR 27 #6314.

[34] Y. Itoh, On the minimum of gaps generated by one-dimensional random packing, *J. Appl. Probab.* 17 (1980) 134–144; MR 81a:60019.

[35] J. A. Morrison, The minimum of gaps generated by random packing of unit intervals into a large interval, *SIAM J. Appl. Math.* 47 (1987) 398–410; MR 88g:60033.

[36] A. C. Kimber, An application of random packing to the multiple fracture of composite materials, *J. Appl. Probab.* 31 (1994) 564–569; MR 94m:73093.

[37] E. G. Coffman and G. S. Lueker, *Probabilistic Analysis of Packing and Partitioning Algorithms*, Wiley, 1991; MR 92h:68038.

[38] P. Flajolet, P. Poblete, and A. Viola, On the analysis of linear probing hashing, *Algorithmica* 22 (1998) 490–515; MR 2000h:68056.

[39] J. M. F. Chamayou, On two cascade models, *Stochastic Process. Appl.* 7 (1978) 153– MR 58 #24621.

[40] M. Sibuya and Y. Itoh, Random sequential bisection and its associated binary tree, *An Inst. Statist. Math.* 39 (1987) 69–84; MR 88e:60120.

[41] F. Komaki and Y. Itoh, A unified model for Kakutani's interval splitting and Rényi's ran packing, *Adv. Appl. Probab.* 24 (1992) 502–505; MR 93c:60009.

[42] P. J. Flory, Intramolecular reaction between neighboring substituents of vinyl polyme *Amer. Chem. Soc.* 61 (1939) 1518–1521.

[43] E. S. Page, The distribution of vacancies on a line, *J. Royal Statist. Soc. Ser. B* 21 (1 364–374; MR 22 #9984.

[44] F. Downton, A note on vacancies on a line, *J. Royal Statist. Soc. Ser. B* 23 (1961) 207– MR 24 #A3672.

[45] J. Texter, Alternative solution to a discrete car parking problem, *J. Chem. Phys.* 91 (1 6295–6301; MR 92c:82118.

[46] V. Privman, J.-S. Wang, and P. Nielaba, Continuum limit in random sequential adsorp *Phys. Rev. B* 43 (1991) 3366–3371.

[47] R. Dickman, J.-S. Wang, and I. Jensen, Random sequential adsorption: Series and v expansions, *J. Chem. Phys.* 94 (1991) 8252–8257.

[48] A. Baram and M. Fixman, Random sequential adsorption: Long time dynamics, *J. C Phys.* 103 (1995) 1929–1933.

[49] R. S. Nord and J. W. Evans, Irreversible immobile random adsorption of dimers, trin . . . on 2D lattices, *J. Chem. Phys.* 82 (1985) 2795–2810.

[50] J. W. Evans and R. S. Nord, Random and cooperative sequential adsorption on inf ladders and strips, *J. Stat. Phys.* 69 (1992) 151–162.

[51] M. J. de Oliveira, T. Tomé, and R. Dickman, Anisotropic random sequential adsorptic dimers on a square lattice, *Phys. Rev. A* 46 (1992) 6294–6299.

[52] J.-S. Wang and R. B. Pandey, Kinetics and jamming coverage in a random seque adsorption of polymer chains, *Phys. Rev. Lett.* 77 (1996) 1773–1776.

[53] C. K. Gan and J.-S. Wang, Extended series expansions for random sequential adsorp *J. Chem. Phys.* 108 (1998) 3010–3012.

[54] C. Fusco, P. Gallo, A. Petri, and M. Rovere, Random sequential adsorption and diffu of dimers and k-mers on a square lattice, *J. Chem. Phys.* 114 (2001) 7563–7569.

[55] J.-S. Wang, Series expansion and computer simulation studies of random sequentia sorption, *Colloids and Surfaces A* 165 (2000) 325–343.

[56] M. Gordon and I. H. Hillier, Statistics of random placement, subject to restrictions, linear lattice, *J. Chem. Phys.* 38 (1963) 1376–1380.

[57] E. A. Boucher, Kinetics and statistics of occupation of linear arrays: A model for pol reactions, *J. Chem. Phys.* 59 (1973) 3848–3852.

[58] J. J. González, P. C. Hemmer, and J. S. Høye, Cooperative effects in random seque polymer reactions, *Chem. Phys.* 3 (1974) 228–238.

[59] A. Baram and D. Kutasov, Random sequential adsorption on a quasi-one-dimensi lattice: An exact solution, *J. Phys. A* 25 (1992) L493–L498.

[60] Y. Fan and J. K. Percus, Random sequential adsorption on a ladder, *J. Stat. Phys.* 66 (1 263–271; MR 92j:82075.

[61] A. Baram and D. Kutasov, Random sequential adsorption on a $3 \times \infty$ lattice: An e solution, *J. Phys. A* 27 (1994) 3683–3687.

[62] P. Meakin, J. L. Cardy, E. Loh, and D. J. Scalapino, Maximal coverage in random seque adsorption, *J. Chem. Phys.* 86 (1987) 2380–2382.

[63] M. C. Bartelt, J. W. Evans, and M. L. Glasser, The car-parking limit of random seque adsorption: Expansions in one dimension, *J. Chem. Phys.* 99 (1993) 1438–1439.

[64] J. W. Evans, Random and cooperative sequential adsorption, *Rev. Mod. Phys.* 65 (1 1281–1329.

[65] J. Talbot, G. Tarjus, P. R. van Tassel, and P. Viot, From car parking to protein adsorptior overview of sequential adsorption processes, *Colloids and Surfaces A* 165 (2000) 287–

[66] V. Privman, Recent theoretical results for nonequilibrium deposition of submicron particles, *J. Adhesion* 74 (2000) 421–440; cond-mat/0003062.

5.4 Golomb–Dickman Constant

Every permutation on n symbols can be uniquely expressed as a product of disjoint cycles. For example, the permutation π on $\{0, 1, 2, \ldots, 9\}$ defined by $\pi(x) = 3x$ mod 10 has cycle structure

$$\pi = (0)\,(1\ 3\ 9\ 7)\,(2\ 6\ 8\ 4)\,(5).$$

In this case, the permutation π has $\alpha_1(\pi) = 2$ cycles of length 1, $\alpha_2(\pi) = 0$ cycles of length 2, $\alpha_3(\pi) = 0$ cycles of length 3, and $\alpha_4(\pi) = 2$ cycles of length 4. The total number $\sum_{j=1}^{\infty} \alpha_j$ of cycles in π is equal to 4 in the example.

Assume that n is fixed and that the $n!$ permutations on $\{0, 1, 2, \ldots, n-1\}$ are assigned equal probability. Picking π at random, we have the classical results [1–4]:

$$\mathrm{E}\left(\sum_{j=1}^{\infty} \alpha_j\right) = \sum_{i=1}^{n} \frac{1}{i} = \ln(n) + \gamma + O\left(\frac{1}{n}\right),$$

$$\mathrm{Var}\left(\sum_{j=1}^{\infty} \alpha_j\right) = \sum_{i=1}^{n} \frac{i-1}{i^2} = \ln(n) + \gamma - \frac{\pi^2}{6} + O\left(\frac{1}{n}\right),$$

$$\lim_{n\to\infty} \mathrm{P}(\alpha_j = k) = \frac{1}{k!} \exp\left(-\frac{1}{j}\right)\left(\frac{1}{j}\right)^k \quad \text{(asymptotic Poisson distribution)},$$

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{\sum_{j=1}^{\infty} \alpha_j - \ln(n)}{\sqrt{\ln(n)}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt \quad \text{(asymptotic normal distribution)},$$

where γ is the Euler–Mascheroni constant [1.5].

What can be said about the limiting distribution of the **longest cycle** and the **shortest cycle**,

$$M(\pi) = \max\{j \ge 1 : \alpha_j > 0\}, \quad m(\pi) = \min\{j \ge 1 : \alpha_j > 0\},$$

given a random permutation π? Goncharov [1,2] and Golomb [5–7] both considered the average value of $M(\pi)$. Golomb examined the constant [8–10]

$$\lambda = \lim_{n\to\infty} \frac{\mathrm{E}(M(\pi))}{n} = 1 - \int_1^{\infty} \frac{\rho(x)}{x^2} dx = 0.6243299885\ldots,$$

where $\rho(x)$ is the unique continuous solution of the following delay-differential equation:

$$\rho(x) = 1 \text{ for } 0 \le x \le 1, \; x\rho'(x) + \rho(x-1) = 0 \quad \text{for } x > 1.$$

(Actually, he worked with the function $\rho(x-1)$.) Shepp & Lloyd [11] and others [6] discovered additional expressions:

$$\lambda = \int_0^{\infty} e^{-x+\mathrm{Ei}(-x)}\, dx = \int_0^1 e^{\mathrm{Li}(x)}\, dx = G(1, 1),$$

where

$$G(a, r) = \frac{1}{a} \int_0^{\infty} \left(1 - \exp(a\,\mathrm{Ei}(-x)) \sum_{k=0}^{r-1} \frac{(-a)^k}{k!} \mathrm{Ei}(-x)^k\right) dx,$$

Ei is the exponential integral [6.2.1], and Li is the logarithmic integral [6.2.2]. Gourdon [12] determined the complete asymptotic expansion for $\mathrm{E}(M(\pi))$:

$$\mathrm{E}(M(\pi)) = \lambda n + \frac{\lambda}{2} - \frac{e^\gamma}{24}\frac{1}{n} + \left[\frac{e^\gamma}{48} - \frac{(-1)^n}{8}\right]\frac{1}{n^2} + \left[\frac{17e^\gamma}{3840} + \frac{(-1)^n}{8} + \frac{e^{\frac{2(2n+1)\pi}{3}i}}{6} + \frac{e^{\frac{2(n+2)\pi}{3}i}}{6}\right]\frac{1}{n^3} + O\left(\frac{1}{n^4}\right).$$

Note the periodic fluctuations involving roots of unity.

A similar integral formula for $\lim_{n\to\infty} \mathrm{Var}(M(\pi))/n^2 = 0.0369078300\ldots = H(1, 1)$ holds, where [12]

$$H(a, r) = \frac{2}{a(a+1)} \int_0^{\infty} \left(1 - \exp(a\,\mathrm{Ei}(-x)) \sum_{k=0}^{r-1} \frac{(-a)^k}{k!} \mathrm{Ei}(-x)^k\right) x\, dx - G(a, r)^2.$$

We will need values of $G(a, r)$ and $H(a, r)$, $a \ne 1 \ne r$, later in this essay. An analog of λ appears in [13, 14] in connection with polynomial factorization.

The arguments leading to asymptotic average values of $m(\pi)$ are more complicated. Shepp & Lloyd [11] proved that

$$\lim_{n\to\infty} \frac{\mathrm{E}(m(\pi))}{\ln(n)} = e^{-\gamma} = 0.5614594835\ldots$$

as well as formulas for higher moments. A complete asymptotic expansion for $\mathrm{E}(m(\pi))$, however, remains open.

The mean and variance of the r^{th} longest cycle (normalized by n and n^2, as $n \to \infty$) are given by $G(1, r)$ and $H(1, r)$. For example, $G(1, 2) = 0.2095808742\ldots$, $H(1, 2) = 0.0125537906\ldots$ and $G(1, 3) = 0.0883160988\ldots$, $H(1, 3) = 0.0044939231\ldots$ [11, 12].

There is a fascinating connection between λ and prime factorization algorithms

n between 1 and N, Dickman [17–20] determined that

$$\lim_{N\to\infty} \mathrm{P}(f(n) \le n^x) = \rho(\tfrac{1}{x})$$

for $0 < x \le 1$. With this in mind, what is the average value of x such that $f(n) = n^x$? Dickman obtained numerically that

$$\mu = \lim_{N\to\infty} \mathrm{E}(x) = \lim_{N\to\infty} \mathrm{E}\left(\frac{\ln(f(n))}{\ln(n)}\right) = \int_0^1 x\, d\rho\left(\frac{1}{x}\right) = 1 - \int_1^\infty \frac{\rho(y)}{y^2} dy = \lambda,$$

which is indeed surprising! Dickman's constant μ and Golomb's constant λ are identical! Knuth & Trabb Pardo [15] described this result as follows: λn is the *asymptotic average number of digits* in the largest prime factor of an n-digit number. More generally, if we are factoring a random n-digit number, the distribution of digits in its prime factors is approximately the same as the distribution of the cycle lengths in a random permutation on n elements. This remarkable and unexpected fact is explored in greater depth in [21, 22].

Other asymptotic formulas involving the largest prime factor function $f(n)$ include [15, 23, 24]

$$\mathrm{E}(f(n)^k) \sim \frac{\zeta(k+1)}{k+1} \frac{N^k}{\ln(N)}, \quad \mathrm{E}(\ln(f(n))) \sim \lambda \ln(N) - \lambda(1-\gamma),$$

where $\zeta(x)$ is the zeta function [1.6]. See also [25–29]. Note the curious coincidence [15] involving integral and sum:

$$\int_0^\infty \rho(x)\, dx = e^\gamma = \sum_{n=1}^\infty n\rho(n).$$

Dickman's function is important in the study of y-smooth numbers [24, 30–32], that is, integers whose prime divisors never exceed y. It appears in probability theory as the density function (normalized by e^γ) of [33, 34]

$$X_1 + X_1X_2 + X_1X_2X_3 + \cdots, \quad X_j \text{ independent uniform random variables on } [0,1].$$

See [35–40] for other applications of $\rho(x)$. A closely-allied function, due to Buchstab, satisfies [24, 34, 41–45]

$$\omega(x) = \frac{1}{x} \text{ for } 1 \le x \le 2, \ x\omega'(x) + \omega(x) - \omega(x-1) = 0 \ \text{ for } x > 2,$$

which arises when estimating the frequency of integers n whose *smallest* prime factor $\ge n^x$. Both functions are positive everywhere, and special values include [46]

$$\rho(\tfrac{3+\sqrt{5}}{2}) = 1 - \ln(\tfrac{3+\sqrt{5}}{2}) + \ln(\tfrac{1+\sqrt{5}}{2})^2 - \tfrac{\pi^2}{60}, \qquad \lim_{x\to\infty} \rho(x) = 0,$$
$$\tfrac{5+\sqrt{5}}{2}\omega(\tfrac{5+\sqrt{5}}{2}) = 1 + \ln(\tfrac{3+\sqrt{5}}{2}) + \ln(\tfrac{1+\sqrt{5}}{2})^2 - \tfrac{\pi^2}{60}, \quad \lim_{x\to\infty} \omega(x) = e^{-\gamma}.$$

Whereas $\rho(x)$ is nonincreasing, the difference $\omega(x) - e^{-\gamma}$ changes sign (at most twice) in every interval of length 1. Its oscillatory behavior plays a role in understanding

Note the similarity in formulation between the Golomb–Dickman constant and Rényi's parking constant [5.3].

5.4.1 Symmetric Group

Here are several related questions. Given π, a permutation on n symbols, define its **order** $\theta(\pi)$ to be the least positive integer m such that π^m = identity. Clearly $1 \le \theta(\pi) \le n!$. What is its mean value, $\mathrm{E}(\theta(\pi))$? Goh & Schmutz [47], building upon the work of Erdös & Turán [48], proved that

$$\ln(\mathrm{E}(\theta(\pi))) = B\sqrt{\frac{n}{\ln(n)}} + o(1),$$

where $B = 2\sqrt{2b} = 2.9904703993\ldots$, and

$$b = \int_0^\infty \ln\left(1 - \ln(1 - e^{-x})\right) dx = 1.1178641511\ldots.$$

Stong [49] improved the $o(1)$ estimate and gave alternative representations for b:

$$b = \int_0^\infty \frac{xe^{-x}}{(1-e^{-x})(1-\ln(1-e^{-x}))} dx = \int_0^\infty \frac{\ln(x+1)}{e^x - 1} dx = -\sum_{k=1}^\infty \frac{e^k}{k} \mathrm{Ei}(-k).$$

A typical permutation π can be shown to satisfy $\ln(\theta(\pi)) \sim \frac{1}{2}\ln(n)^2$; hence a few exceptional permutations contribute significantly to the mean. What can be said about the variance of $\theta(\pi)$?

Also, define $g(n)$ to be the maximum order $\theta(\pi)$ of all n-permutations π. Landau [50, 51] proved that $\ln(g(n)) \sim \sqrt{n\ln(n)}$, and greatly refined estimates of $g(n)$ appeared in [52].

A natural equivalence relation can be defined on the symmetric group S_n via conjugacy. In the limit as $n \to \infty$, for almost all conjugacy classes C, the elements of C have order equal to $\exp(\sqrt{n}(A + o(1)))$, where [48, 53, 54]

$$A = \frac{2\sqrt{6}}{\pi} \sum_{j\ne 0} \frac{(-1)^{j+1}}{3j^2 + j} = 4\sqrt{2} - \frac{6\sqrt{6}}{\pi}.$$

Note that the summation involves reciprocals of nonzero pentagonal numbers.

Let s_n denote the probability that two elements of the symmetric group, S_n, chosen at random (with replacement) actually generate S_n. The first several values are $s_1 = 1$, $s_2 = 3/4$, $s_3 = 1/2$, $s_4 = 3/8, \ldots$ [55]. What can be said about the asymptotics of s_n? Dixon [56] proved an 1892 conjecture by Netto [57] that $s_n \to 3/4$ as $n \to \infty$. Babai [58] gave a more refined estimate.

5.4.2 Random Mapping Statistics

We now generalize the discussion from permutations (bijective functions) on n symbols

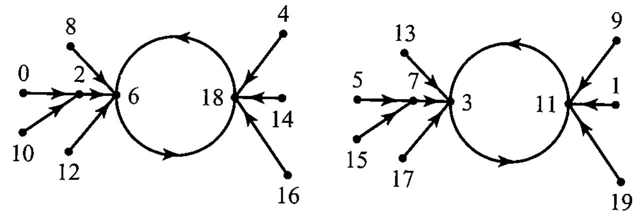


Figure 5.1. The functional graph for $\psi(x) = x^2 + 2 \bmod 20$ has two components, each containing a cycle of length 2.

defined by $\varphi(x) = 2x \bmod 10$ has cycles (0) and (2 4 8 6). The remaining symbols 1, 3, 5, 7, and 9 are transient in the sense that if one starts with 3, one is absorbed into the cycle (2 4 8 6) and never returns to 3. We can nevertheless define cycle lengths α_j as before; in this simple example, $\alpha_1(\varphi) = 1$, $\alpha_2(\varphi) = \alpha_3(\varphi) = 0$, and $\alpha_4(\varphi) = 1$.

The lengths of the longest and shortest cycles, $M(\varphi)$ and $m(\varphi)$, are clearly of interest in pseudo-random number generation. Purdom & Williams [59–61] found that

$$\lim_{n\to\infty} \frac{\mathrm{E}(M(\varphi))}{\sqrt{n}} = \lambda\sqrt{\frac{\pi}{2}} = 0.7824816009\ldots, \quad \lim_{n\to\infty} \frac{\mathrm{E}(m(\varphi))}{\ln(n)} = \frac{1}{2}e^{-\gamma}.$$

Observe that $\mathrm{E}(M(\varphi))$ grows on the order of only $\sqrt{n}$ rather than n as earlier.

As another example, consider the function ψ on $\{0, 1, 2, \ldots, 19\}$ defined by $\psi(x) = x^2 + 2 \bmod 20$. From Figure 5.1, clearly $\alpha_2(\psi) = 2$. Here are other interesting quantities [62]. Note that the transient symbols 0, 5, 10, and 15 each require 2 steps to reach a cycle, and this is the maximum such distance. Thus define the **longest tail** $L(\psi) = 2$. Note also that 4 is the number of vertices in the nonrepeating trajectory for each of 0, 5, 10, and 15, and this is the maximum such length. Thus define the **longest rho-path** $R(\psi) = 4$. Clearly, for the earlier example, $L(\varphi) = 1$ and $R(\varphi) = 5$. It can be proved that, for arbitrary n-mappings φ [61],

$$\lim_{n\to\infty} \frac{\mathrm{E}(L(\varphi))}{\sqrt{n}} = \sqrt{2\pi}\ln(2) = 1.7374623212\ldots,$$

$$\lim_{n\to\infty} \frac{\mathrm{E}(R(\varphi))}{\sqrt{n}} = \sqrt{\frac{\pi}{2}} \int_0^\infty \left(1 - e^{\mathrm{Ei}(-x) - I(x)}\right) dx = 2.4149010237\ldots,$$

where

$$I(x) = \int_0^x \frac{e^{-y}}{y}\left(1 - \exp\left(\frac{-2y}{e^{x-y} - 1}\right)\right) dy.$$

Another quantity associated with a mapping φ is the **largest tree** $P(\varphi)$. Each vertex in each cycle of φ is the root of a unique maximal tree [5.6]. Select the tree with the greatest number of vertices, and call this number $P(\varphi)$. For the two examples, clearly

$P(\varphi) = 2$ and $P(\psi) = 6$. It is known that, for arbitrary n-mappings φ [12,61],

$$\nu = \lim_{n\to\infty} \frac{\mathrm{E}(P(\varphi))}{n} = 2\int_0^\infty \left[1 - (1 - F(x))^{-1}\right] dx = 0.4834983471\ldots,$$

$$\lim_{n\to\infty} \frac{\mathrm{Var}(P(\varphi))}{n^2} = \frac{8}{3}\int_0^\infty \left[1 - (1 - F(x))^{-1}\right] x\, dx - \nu^2 = 0.0494698522\ldots,$$

where

$$F(x) = \frac{-1}{2\sqrt{\pi}} \int_x^\infty e^{-t} t^{-\frac{3}{2}}\, dt = 1 - \frac{1}{\sqrt{\pi x}}\exp(-x) - \mathrm{erf}(\sqrt{x})$$

and erf is the error function [4.6]. Gourdon [63] mentioned a coin-tossing game, the analysis of which yields the preceding two constants.

Finally, let us examine the connected component structure of a mapping. We have come full circle, in a sense, because components relate to mappings as cycles relate to permutations. For the two examples, the counting function is $\beta_2(\varphi) = 1$, $\beta_8(\varphi) = 1$ while $\beta_{10}(\psi) = 2$. In the interest of analogy, here are more details. The total number $\sum_{j=1}^\infty \beta_j$ of components is equal to 2 in both cases. Picking φ at random, we have [64–67]

$$\mathrm{E}\left(\sum_{j=1}^\infty \beta_j\right) = \sum_{i=1}^n c_{n,0,i} = \frac{1}{2}\ln(n) + \frac{1}{2}(\ln(2) + \gamma) + o(1),$$

$$\begin{aligned}\mathrm{Var}\left(\sum_{j=1}^\infty \beta_j\right) &= \sum_{i=1}^n c_{n,0,i} - \left(\sum_{i=1}^n c_{n,0,i}\right)^2 + \sum_{i=1}^n c_{n,0,i} \sum_{j=1}^{n-i} c_{n,i,j} \\ &= \frac{1}{2}\ln(n) + o(\ln(n)),\end{aligned}$$

$$\lim_{n\to\infty} \mathrm{P}(\beta_j = k) = \frac{1}{k!}\exp(-d_j)\, d_j^k, \quad \text{(asymptotic Poisson distribution)},$$

where

$$c_{n,p,q} = \binom{n-p}{q}\frac{(q-1)!}{n^q}, \quad d_j = \frac{e^{-j}}{j}\sum_{i=0}^{j-1}\frac{j^i}{i!},$$

and a corresponding Gaussian limit also holds. Define the **largest component** $Q(\varphi) = \max\{j \ge 1 : \beta_j > 0\}$; then [12,61,68]

$$\lim_{n\to\infty} \frac{\mathrm{E}(Q(\varphi))}{n} = G(\tfrac{1}{2}, 1) = 0.7578230112\ldots,$$

$$\lim \frac{\mathrm{Var}(Q(\varphi))}{} = H(\tfrac{1}{2}, 1) = 0.0370072165$$

Such results answer questions raised in [69–71]. It seems fitting to call 0.75782... the **Flajolet–Odlyzko constant**, owing to its importance. The mean and variance of the r^{th} largest component (again normalized by n and n^2, as $n \to \infty$) are given by $G(\frac{1}{2}, r)$ and $H(\frac{1}{2}, r)$. For example, $G(\frac{1}{2}, 2) = 0.1709096198\ldots$ and $H(\frac{1}{2}, 2) = 0.0186202233\ldots$. A discussion of smallest components appears in [72].

[1] W. Goncharov, Sur la distribution des cycles dans les permutations, *C. R. (Dokl.) Acad. Sci. URSS* 35 (1942) 267–269; MR 4,102g.

[2] W. Goncharov, On the field of combinatory analysis (in Russian), *Izv. Akad. Nauk SSSR* 8 (1944) 3–48; Engl. transl. in *Amer. Math. Soc. Transl.* 19 (1962) 1–46; MR 6,88b and MR 24 A1221.

[3] R. E. Greenwood, The number of cycles associated with the elements of a permutation group, *Amer. Math. Monthly* 60 (1953) 407–409; MR 14,939b.

[4] H. S. Wilf, *generatingfunctionology*, Academic Press, 1990; MR 95a:05002.

[5] S. W. Golomb, Random permutations, *Bull. Amer. Math. Soc.* 70 (1964) 747.

[6] S. W. Golomb, *Shift Register Sequences*, Holden-Day, 1967; MR 39 #3906.

[7] S. W. Golomb and P. Gaal, On the number of permutations on n objects with greatest cycle length k, *Adv. Appl. Math.* 20 (1998) 98–107; *Probabilistic Methods in Discrete Mathematics*, Proc. 1996 Petrozavodsk conf., ed. V. F. Kolchin, V. Ya. Kozlov, Yu. L. Pavlov, and Yu. V. Prokhorov, VSP, 1997, pp. 211–218; MR 98k:05003 and MR 99j:05001.

[8] W. C. Mitchell, An evaluation of Golomb's constant, *Math. Comp.* 22 (1968) 411–415.

[9] D. E. Knuth, *The Art of Computer Programming*, v. 1, *Fundamental Algorithms*, Addison-Wesley, 1969, pp. 180–181, 519–520; MR 51 #14624.

[10] A. MacLeod, Golomb's constant to 250 decimal places, unpublished note (1997).

[11] L. A. Shepp and S. P. Lloyd, Ordered cycle lengths in a random permutation, *Trans. Amer. Math. Soc.* 121 (1966) 350–557; MR 33 #3320.

[12] X. Gourdon, *Combinatoire, Algorithmique et Géométrie des Polynômes*, Ph.D. thesis, École Polytechnique, 1996.

[13] P. Flajolet, X. Gourdon, and D. Panario, Random polynomials and polynomial factorization, *Proc. 1996 Int. Colloq. on Automata, Languages and Programming (ICALP)*, Paderborn, ed. F. Meyer auf der Heide and B. Monien, Lect. Notes in Comp. Sci. 1099, Springer-Verlag, pp. 232–243; MR 98e:68123.

[14] P. Flajolet, X. Gourdon, and D. Panario, The complete analysis of a polynomial factorization algorithm over finite fields, *J. Algorithms* 40 (2001) 37–81; INRIA report RR3370; MR 2002f:68193.

[15] D. E. Knuth and L. Trabb Pardo, Analysis of a simple factorization algorithm, *Theoret. Comput. Sci.* 3 (1976) 321–348; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 303–339; MR 58 #16485.

[16] D. E. Knuth, *The Art of Computer Programming*, v. 2, *Seminumerical Algorithms*, 2nd ed., Addison-Wesley, 1981, pp. 367–368, 395, 611; MR 83i:68003.

[17] K. Dickman, On the frequency of numbers containing prime factors of a certain relative magnitude, *Ark. Mat. Astron. Fysik*, v. 22A (1930) n. 10, 1–14.

[18] S. D. Chowla and T. Vijayaraghavan, On the largest prime divisors of numbers, *J. Indian Math. Soc.* 11 (1947) 31–37; MR 9,332d.

[19] V. Ramaswami, On the number of positive integers less than x and free of prime divisors greater than x^c, *Bull. Amer. Math. Soc.* 55 (1949) 1122–1127; MR 11,233f.

[20] N. G. de Bruijn, On the number of positive integers $\leq x$ and free of prime factors $> y$, *Proc. Konink. Nederl. Acad. Wetensch. Ser. A* 54 (1951) 50–60; *Indag. Math.* 13 (1951) 50–60; MR 13,724e.

[21] R. Arratia, A. D. Barbour, and S. Tavaré, Random combinatorial structures and prime factorizations, *Notices Amer. Math. Soc.* 44 (1997) 903–910; MR 98i:60007.

[22] R. Arratia, A. D. Barbour, and S. Tavaré, *Logarithmic Combinatorial Structures: A Probabilistic Approach*, unpublished manuscript (2001).

[23] K. Alladi and P. Erdös, On an additive arithmetic function, *Pacific J. Math.* 71 (1977) 275–294; MR 56 #5401.

[24] G. Tenenbaum, *Introduction to Analytic and Probabilistic Number Theory*, Cambridge Univ. Press, 1995, pp. 365–366, 393, 399–400; MR 97e:11005b.

[25] P. Erdös 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; MR 84a:10046.

[26] T. Z. Xuan, On a result of Erdös and Ivić, *Arch. Math. (Basel)* 62 (1994) 143–154; MR 94m:11109.

[27] H. Z. Cao, Sums involving the smallest prime factor of an integer, *Utilitas Math.* 45 (1994) 245–251; MR 95d:11126.

[28] S. R. Finch, Moments of the Smarandache function, *Smarandache Notions J.* 10 (1999) 95–96; 11 (2000) 140–141; MR 2000a:11129.

[29] K. Ford, The normal behavior of the Smarandache function, *Smarandache Notions J.* 10 (1999) 81–86; MR 2000a:11130.

[30] K. K. Norton, Numbers with small prime factors, and the least k^{th} power non-residue, *Memoirs Amer. Math. Soc.* 106 (1971) 1–106; MR 44 #3948.

[31] A. Hildebrand and G. Tenenbaum, Integers without large prime factors, *J. Théorie Nombres Bordeaux* 5 (1993) 411–484; MR 95d:11116.

[32] P. Moree, *On the Psixyology of Diophantine Equations*, Ph.D. thesis, Univ. of Leiden, 1993.

[33] J.-M.-F. Chamayou, A probabilistic approach to a differential-difference equation arising in analytic number theory, *Math. Comp.* 27 (1973) 197–203; MR 49 #1725.

[34] G. Marsaglia, A. Zaman, and J. C. W. Marsaglia, Numerical solution of some classical differential-difference equations, *Math. Comp.* 53 (1989) 191–201; MR 90h:65124.

[35] H. Davenport and P. Erdös, The distribution of quadratic and higher residues, *Publ. Math. (Debrecen)* 2 (1952) 252–265; MR 14,1063h.

[36] L. I. Pál and G. Németh, A statistical theory of lattice damage in solids irradiated by high-energy particles, *Nuovo Cimento* 12 (1959) 293–309; MR 21 #7630.

[37] G. A. Watterson, The stationary distribution of the infinitely-many neutral alleles diffusion model, *J. Appl. Probab.* 13 (1976) 639–651; correction 14 (1977) 897; MR 58 #20594a and b.

[38] D. Hensley, The convolution powers of the Dickman function, *J. London Math. Soc.* 33 (1986) 395–406; MR 87k:11097.

[39] D. Hensley, Distribution of sums of exponentials of random variables, *Amer. Math. Monthly* 94 (1987) 304–306.

[40] C. J. Lloyd and E. J. Williams, Recursive splitting of an interval when the proportions are identical and independent random variables, *Stochastic Process. Appl.* 28 (1988) 111–122; MR 89e:60025.

[41] A. A. Buchstab, Asymptotic estimates of a general number theoretic function (in Russian), *Mat. Sbornik* 2 (1937) 1239–1246.

[42] S. Selberg, The number of cancelled elements in the sieve of Eratosthenes (in Norwegian), *Norsk Mat. Tidsskr.* 26 (1944) 79–84; MR 8,317a.

[43] N. G. de Bruijn, On the number of uncancelled elements in the sieve of Eratosthenes, *Proc. Konink. Nederl. Acad. Wetensch. Sci. Sect.* 53 (1950) 803–812; *Indag. Math.* 12 (1950) 247–256; MR 12,11d.

[44] A. Y. Cheer and D. A. Goldston, A differential delay equation arising from the sieve of Eratosthenes, *Math. Comp.* 55 (1990) 129–141; MR 90j:11091.

[45] A. Hildebrand and G. Tenenbaum, On a class of differential-difference equations arising in analytic number theory, *J. d'Analyse Math.* 61 (1993) 145–179; MR 94i:11069.

[46] P. Moree, A special value of Dickman's function, *Math. Student* 64 (1995) 47–50.

[47] W. M. Y. Goh and E. Schmutz, The expected order of a random permutation, *Bull. London Math. Soc.* 23 (1991) 34–42; MR 93a:11080.

[48] P. Turán, Combinatorics, partitions, group theory, *Colloquio Internazionale sulle Teorie Combinatorie*, t. 2, Proc. 1973 Rome conf., Accad. Naz. Lincei, 1976, pp. 181–200; also in *Collected Works*, v. 3, ed. P. Erdös, Akadémiai Kiadó, pp. 2302–2321; MR 58 #21978.

[49] R. Stong, The average order of a permutation, *Elec. J. Combin.* 5 (1998) R41; MR 99f:11122.

[50] W. Miller, The maximum order of an element of a finite symmetric group, *Amer. Math. Monthly* 94 (1987) 497–506; MR 89k:20005.

[51] J.-L. Nicolas, On Landau's function $g(n)$, *The Mathematics of Paul Erdős*, v. 1, ed. R. L. Graham and J. Nesetril, Springer-Verlag, 1997, pp. 228–240; MR 98b:11096.

[52] J.-P. Massias, J.-L. Nicolas, and G. Robin, Effective bounds for the maximal order of an element in the symmetric group, *Math. Comp.* 53 (1989) 665–678; MR 90e:11139.

[53] D. H. Lehmer, On a constant of Turán and Erdös, *Acta Arith.* 37 (1980) 359–361; MR 82b:05014.

[54] W. Schwarz, Another evaluation of an Erdös-Turán constant, *Analysis* 5 (1985) 343–345; MR 86m:40003.

[55] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A040173 and A040174.

[56] J. D. Dixon, The probability of generating the symmetric group, *Math. Z.* 110 (1969) 199–205; MR 40 #4985.

[57] E. Netto, *The Theory of Substitutions and Its Applications to Algebra*, 2nd ed., Chelsea, 1964; MR 31 #184.

[58] L. Babai, The probability of generating the symmetric group, *J. Combin. Theory Ser. A* 52 (1989) 148–153; MR 91a:20007.

[59] P. W. Purdom and J. H. Williams, Cycle length in a random function, *Trans. Amer. Math. Soc.* 133 (1968) 547–551; MR 37 #3616.

[60] D. E. Knuth, *The Art of Computer Programming*, v. 2, *Seminumerical Algorithms*, 2nd ed., Addison-Wesley, 1981; pp. 7–8, 517–520; MR 83i:68003.

[61] P. Flajolet and A. M. Odlyzko, Random mapping statistics, *Advances in Cryptology – EUROCRYPT '89*, ed. J.-J. Quisquater and J. Vandewalle, Lect. Notes in Comp. Sci. 434, Springer-Verlag, 1990, pp. 329–354; MR 91h:94003.

[62] R. Sedgewick and P. Flajolet, *Introduction to the Analysis of Algorithms*, Addison-Wesley, 1996, pp. 357–358, 454–465.

[63] X. Gourdon, Largest components in random combinatorial structures, *Discrete Math.* 180 (1998) 185–209; MR 99c:60013.

[64] M. D. Kruskal, The expected number of components under a random mapping function, *Amer. Math. Monthly* 61 (1954) 392–397; MR 16,52b.

[65] S. M. Ross, A random graph, *J. Appl. Probab.* 18 (1981) 309–315; MR 81m:05124.

[66] P. M. Higgins and E. J. Williams, Random functions on a finite set, *Ars Combin.* 26 (1988) A, 93–102; MR 90g:60008.

[67] V. F. Kolchin, *Random Mappings*, Optimization Software Inc., 1986, pp. 46, 79, 164; MR 88a:60022.

[68] D. Panario and B. Richmond, Exact largest and smallest size of components, *Algorithmica* 31 (2001) 413–432; MR 2002j:68065.

[69] G. A. Watterson and H. A. Guess, Is the most frequent allele the oldest?, *Theoret. Populat. Biol.* 11 (1977) 141–160.

[70] P. J. Donnelly, W. J. Ewens, and S. Padmadisastra, Functionals of random mappings: Exact and asymptotic results, *Adv. Appl. Probab.* 23 (1991) 437–455; MR 92k:60017.

[71] P. M. Higgins, *Techniques of Semigroup Theory*, Oxford Univ. Press, 1992; MR 93d:20101.

[72] D. Panario and B. Richmond, Smallest components in decomposable structures: exp-log class, *Algorithmica* 29 (2001) 205–226.

5.5 Kalmár's Composition Constant

An **additive composition** of an integer n is a sequence $x_1, x_2, \ldots, x_k$ of integers (for some $k \ge 1$) such that

$$n = x_1 + x_2 + \cdots + x_k, \quad x_j \ge 1 \text{ for all } 1 \le j \le k.$$

A **multiplicative composition** of n is the same except

$$n = x_1 x_2 \cdots x_k, \quad x_j \ge 2 \text{ for all } 1 \le j \le k.$$

The number $a(n)$ of additive compositions of n is trivially 2^{n-1}. The number $m(n)$ of multiplicative compositions does not possess a closed-form expression, but asymptotically satisfies

$$\sum_{n=1}^{N} m(n) \sim \frac{-1}{\rho \zeta'(\rho)} N^{\rho} = (0.3181736521\ldots) \cdot N^{\rho},$$

where $\rho = 1.7286472389\ldots$ is the unique solution of $\zeta(x) = 2$ with $x > 1$ and $\zeta(x)$ is Riemann's zeta function [1.6]. This result was first deduced by Kalmár [1,2] and refined in [3–8].

An **additive partition** of an integer n is a sequence $x_1, x_2, \ldots, x_k$ of integers (for some $k \ge 1$) such that

$$n = x_1 + x_2 + \cdots + x_k, \quad 1 \le x_1 \le x_2 \le \cdots \le x_k.$$

Partitions naturally represent equivalence classes of compositions under sorting. The number $A(n)$ of additive partitions of n is mentioned in [1.4.2], while the number $M(n)$ of **multiplicative partitions** asymptotically satisfies [9, 10]

$$\sum_{n=1}^{N} M(n) \sim \frac{1}{2\sqrt{\pi}} N \exp\left(2\sqrt{\ln(N)}\right) \ln(N)^{-\frac{3}{4}}.$$

Thus far we have dealt with *unrestricted* compositions and partitions. Of many possible variations, let us focus on the case in which each x_j is restricted to be a prime number. For example, the number $M_{\mathrm{p}}(n)$ of **prime multiplicative partitions** is trivially 1 for $n \ge 2$. The number $a_{\mathrm{p}}(n)$ of **prime additive compositions** is [11]

$$a_{\mathrm{p}}(n) \sim \frac{1}{\xi f'(\xi)} \left(\frac{1}{\xi}\right)^n = (0.3036552633\ldots) \cdot (1.4762287836\ldots)^n,$$

where $\xi = 0.6774017761\ldots$ is the unique solution of the equation

$$f(x) = \sum_{p} x^p = 1, \quad x > 0,$$

and the sum is over all primes p. The number $m_{\mathrm{p}}(n)$ of **prime multiplicative compositions** satisfies [12]

$$\sum_{n=1}^{N} m_{\mathrm{p}}(n) \sim \frac{-1}{\eta g'(\eta)} N^{-\eta} = (0.4127732370\ldots) \cdot N^{-\eta},$$

where $\eta = -1.3994333287\ldots$ is the unique solution of the equation

$$g(y) = \sum_{p} p^y = 1, \quad y < 0.$$

Not much is known about the number $A_{\mathrm{p}}(n)$ of **prime additive partitions** [13–16] except that $A_{\mathrm{p}}(n+1) > A_{\mathrm{p}}(n)$ for $n \ge 8$.

Here is a related, somewhat artificial topic. Let p_n be the n^{th} prime, with $p_1 = 2$, and define formal series

$$P(z) = 1 + \sum_{n=1}^{\infty} p_n z^n, \quad Q(z) = \frac{1}{P(z)} = \sum_{n=0}^{\infty} q_n z^n.$$

Some people may be surprised to learn that the coefficients q_n obey the following asymptotics [17]:

$$q_n \sim \frac{1}{\theta P'(\theta)} \left(\frac{1}{\theta}\right)^n = (-0.6223065745\ldots) \cdot (-1.4560749485\ldots)^n.$$

where $\theta = -0.6867778344\ldots$ is the unique zero of $P(z)$ inside the disk $|z| < 3/4$. By way of contrast, $p_n \sim n \ln(n)$ by the Prime Number Theorem. In a similar spirit, consider the coefficients c_k of the $(n-1)^{\text{st}}$ degree polynomial fit

$$c_0 + c_1(x-1) + c_2(x-1)(x-2) + \cdots + c_{n-1}(x-1)(x-2)(x-3)\cdots(x-n+1)$$

to the dataset [18]

$$(1,2), (2,3), (3,5), (4,7), (5,11), (6,13), \ldots, (n, p_n).$$

In the limit as $n \to \infty$, the sum $\sum_{k=0}^{n-1} c_k$ converges to $3.4070691656\ldots$.

Let us return to the counting of compositions and partitions, and merely mention variations in which each x_j is restricted to be square-free [12] or where the xs must be distinct [8]. Also, compositions/partitions $x_1, x_2, \ldots, x_k$ and $y_1, y_2, \ldots, y_l$ of n are said to be **independent** if proper subsequence sums/products of xs and ys never coincide. How many such pairs are there (as a function of n)? See [19] for an asymptotic answer.

Cameron & Erdős [20] pointed out that the number of sequences $1 \le z_1 < z_2 < \cdots < z_k = n$ for which $z_i | z_j$ whenever $i < j$ is $2m(n)$. The factor 2 arises because we can choose whether or not to include 1 in the sequence. What can be said about the number $c(n)$ of sequences $1 \le w_1 < w_2 < \cdots < w_k \le n$ for which $w_i \nmid w_j$ whenever $i \ne j$? It is conjectured that $\lim_{n\to\infty} c(n)^{1/n}$ exists, and it is known that $1.55967^n \le c(n) \le 1.59^n$ for sufficiently large n. For more about such sequences, known as **primitive sequences**, see [2.27].

Finally, define $h(n)$ to be the number of ways to express 1 as a sum of $n+1$ elements of the set $\{2^{-i} : i \ge 0\}$, where repetitions are allowed and order is immaterial. Flajolet & Prodinger [21] demonstrated that

$$h(n) \sim (0.2545055235\ldots)\kappa^n,$$

where $\kappa = 1.7941471875\ldots$ is the reciprocal of the smallest positive root x of the equation

$$\sum_{j=1}^{\infty} (-1)^{j+1} \frac{x^{2^{j+1}-2-j}}{(1-x)(1-x^3)(1-x^7)\cdots(1-x^{2^j-1})} - 1 = 0.$$

This is connected to enumerating level number sequences associated with binary trees [5.6].

[1] L. Kalmár, A "factorisatio numerorum" problémájáról, *Mat. Fiz. Lapok* 38 (1931) 1–15.
[2] L. Kalmár, Über die mittlere Anzahl Produktdarstellungen der Zahlen, *Acta Sci. Math. (Szeged)* 5 (1930-32) 95–107.
[3] E. Hille, A problem in "Factorisatio Numerorum," *Acta Arith.* 2 (1936) 136–144.
[4] P. Erdös, On some asymptotic formulas in the theory of the "factorisatio numerorum," *Annals of Math.* 42 (1941) 989–993; MR 3,165b.
[5] P. Erdös, Corrections to two of my papers, *Annals of Math.* 44 (1943) 647–651; MR 5,172c.
[6] S. Ikehara, On Kalmár's problem in "Factorisatio Numerorum." II, *Proc. Phys.-Math. Soc. Japan* 23 (1941) 767–774; MR 7,365h.
[7] R. Warlimont, Factorisatio numerorum with constraints, *J. Number Theory* 45 (1993) 186–199; MR 94f:11098.
[8] H.-K. Hwang, Distribution of the number of factors in random ordered factorizations of integers, *J. Number Theory* 81 (2000) 61–92; MR 2001k:11183.
[9] A. Oppenheim, On an arithmetic function. II, *J. London Math. Soc.* 2 (1927) 123–130.
[10] G. Szekeres and P. Turán, Über das zweite Hauptproblem der "Factorisatio Numerorum," *Acta Sci. Math. (Szeged)* 6 (1932-34) 143–154; also in *Collected Works of Paul Turán*, v. 1, ed. P. Erdös, Akadémiai Kiadó, pp. 1–12.
[11] P. Flajolet, Remarks on coefficient asymptotics, unpublished note (1995).
[12] A. Knopfmacher, J. Knopfmacher, and R. Warlimont, Ordered factorizations for integers and arithmetical semigroups, *Advances in Number Theory*, Proc. 1991 Kingston conf., ed. F. Q. Gouvêa and N. Yui, Oxford Univ. Press, 1993, pp. 151–165; MR 97e:11118.
[13] P. T. Bateman and P. Erdös, Monotonicity of partition functions, *Mathematika* 3 (1956) 1–14; MR 18,195a.
[14] P. T. Bateman and P. Erdös, Partitions into primes, *Publ. Math. (Debrecen)* 4 (1956) 198–200; MR 18,15c.
[15] J. Browkin, Sur les décompositions des nombres naturels en sommes de nombres premiers, *Colloq. Math.* 5 (1958) 205–207; MR 21 #1956.
[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000041, A000607, A001055, A002033, A002572, A008480, and A023360.
[17] N. Backhouse, Formal reciprocal of a prime power series, unpublished note (1995).
[18] F. Magata, Newtonian interpolation and primes, unpublished note (1998).
[19] P. Erdös, J.-L. Nicolas, and A. Sárközy, On the number of pairs of partitions of n without common subsums, *Colloq. Math.* 63 (1992) 61–83; MR 93c:11087.
[20] P. J. Cameron and P. Erdös, On the number of sets of integers with various properties, *Number Theory*, Proc. 1990 Canad. Number Theory Assoc. Banff conf., ed. R. A. Mollin, Gruyter, pp. 61–79; MR 92g:11010.
[21] P. Flajolet and H. Prodinger, Level number sequences for trees, *Discrete Math.* 65 (1987) 149–156; MR 88e:05030.

5.6 Otter's Tree Enumeration 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 **adjacent**.

A **forest** is a graph that is **acyclic**, meaning that there is no sequence of adjacent vertices $v_0, v_1, \ldots, v_m$ such that $v_i \ne v_j$ for all $i < j < m$ and $v_0 = v_m$.

A **tree** (or **free tree**) is a forest that is **connected**, meaning that 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 $v_m = w$.

Two trees σ and τ are **isomorphic** if there is a one-to-one map from the vertices of σ to the vertices of τ that preserves adjacency (see Figure 5.2). Diagrams for all

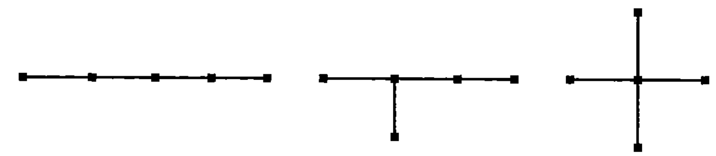


Figure 5.2. There exist three non-isomorphic trees of order 5.

What can be said about the asymptotics of t_n, the number of non-isomorphic trees of order n? Building upon the work of Cayley and Pólya, Otter [3–6] determined that

$$\lim_{n\to\infty} \frac{t_n n^{\frac{5}{2}}}{\alpha^n} = \beta,$$

where $\alpha = 2.9557652856\ldots = (0.3383218568\ldots)^{-1}$ is the unique positive solution of the equation $T(x^{-1}) = 1$ involving a certain function T to be defined shortly, and

$$\beta = \frac{1}{\sqrt{2\pi}}\left(1 + \sum_{k=2}^{\infty} \frac{1}{\alpha^k} T'\left(\frac{1}{\alpha^k}\right)\right)^{\frac{3}{2}} = 0.5349496061\ldots$$

where T' denotes the derivative of T. Although α and β can be calculated efficiently to great accuracy, it is not known whether they are algebraic or transcendental [6, 7].

A **rooted tree** is a tree in which precisely one vertex, called the **root**, is distinguished from the others (see Figure 5.3). We agree to draw the root as a tree's topmost vertex and that an isomorphism of rooted trees maps a root to a root. What can be said about the asymptotics of T_n, the number of non-isomorphic rooted trees of order n? Otter's corresponding result is

$$\lim_{n\to\infty} \frac{T_n n^{\frac{3}{2}}}{\alpha^n} = \left(\frac{\beta}{2\pi}\right)^{\frac{1}{3}} = 0.4399240125\ldots = \left(\frac{1}{4\pi\alpha}\right)^{\frac{1}{2}}(2.6811281472\ldots).$$

In fact, the generating functions

$$\begin{aligned} t(x) &= \sum_{n=1}^{\infty} t_n x^n \\ &= x + x^2 + x^3 + 2x^4 + 3x^5 + 6x^6 + 11x^7 + 23x^8 + 47x^9 + 106x^{10} + \cdots, \\ T(x) &= \sum_{n=1}^{\infty} T_n x^n \\ &= x + x^2 + 2x^3 + 4x^4 + 9x^5 + 20x^6 + 48x^7 + 115x^8 + 286x^9 + \cdots \end{aligned}$$

are related by the formula $t(x) = T(x) - \frac{1}{2}(T(x)^2 - T(x^2))$, the constant α^{-1} is the radius of convergence for both, and the coefficients T_n can be computed using

$$T(x) = x\exp\left(\sum_{k=1}^{\infty} \frac{T(x^k)}{k}\right), \quad T_{n+1} = \frac{1}{n}\sum_{k=1}^{n}\left(\sum_{d\mid k} d T_d\right) T_{n-k+1}.$$

There are many varieties of trees and the elaborate details of enumerating them are best left to [4, 5]. Here is the first of many examples. A **weakly binary tree** is a rooted

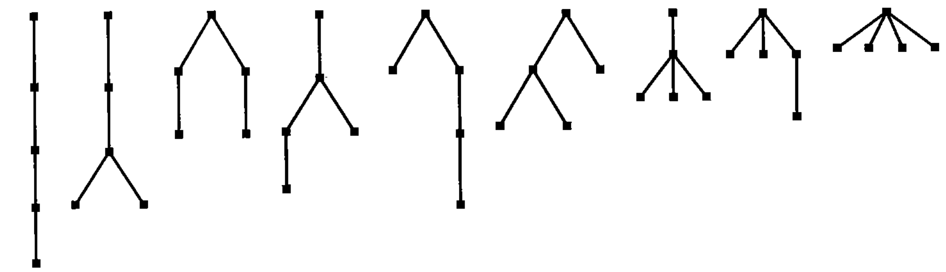


Figure 5.3. There exist nine non-isomorphic rooted trees of order 5.

tree for which the root is adjacent to at most two vertices and all non-root vertices are adjacent to at most three vertices. For instance, there exist six non-isomorphic weakly binary trees of order 5. The asymptotics of B_n, the number of non-isomorphic weakly binary trees of order n, were obtained by Otter [3, 8–10]:

$$\lim_{n\to\infty} \frac{B_n n^{\frac{3}{2}}}{\xi^n} = \eta,$$

where $\xi^{-1} = 0.4026975036\ldots = (2.4832535361\ldots)^{-1}$ is the radius of convergence for

$$\begin{aligned} B(x) &= \sum_{n=0}^{\infty} B_n x^n \\ &= 1 + x + x^2 + 2x^3 + 3x^4 + 6x^5 + 11x^6 + 23x^7 + 46x^8 + 98x^9 + \cdots \end{aligned}$$

and

$$\begin{aligned} \eta &= \sqrt{\frac{\xi}{2\pi}}\left(1 + \frac{1}{\xi}B\left(\frac{1}{\xi^2}\right) + \frac{1}{\xi^3}B'\left(\frac{1}{\xi^2}\right)\right)^{\frac{1}{2}} \\ &= 0.7916031835\ldots = (0.3187766258\ldots)\xi. \end{aligned}$$

The series coefficients arise from

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

$$B_k = \begin{cases} \dfrac{B_i(B_i+1)}{2} + \displaystyle\sum_{j=0}^{i-1} B_{k-j-1}B_j & \text{if } k = 2i+1, \\ \displaystyle\sum_{j=0}^{i-1} B_{k-j-1}B_j & \text{if } k = 2i. \end{cases}$$

Otter showed, in this special case, that $\xi = \lim_{n\to\infty} c_n^{2^{-n}}$, where the sequence $\{c_n\}$ obeys the quadratic recurrence

$$c_0 = 2, \qquad c_n = c_{n-1}^2 + 2 \quad \text{for } n \ge 1,$$

and consequently

$$\eta = \frac{1}{2}\sqrt{\frac{\xi}{\pi}}\sqrt{3 + \frac{1}{c_1} + \frac{1}{c_1c_2} + \frac{1}{c_1c_2c_3} + \frac{1}{c_1c_2c_3c_4} + \cdots}.$$

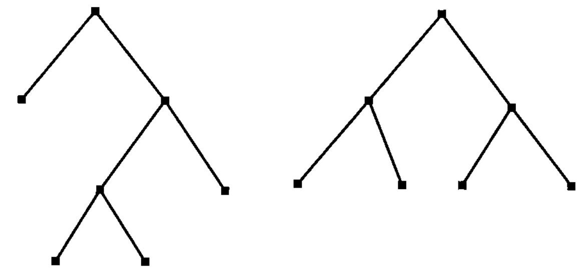


Figure 5.4. There exist two non-isomorphic strongly binary trees of order 7.

Here is a slight specialization of the preceding. Define a **strongly binary tree** to be a rooted tree for which the root is adjacent to either zero or two vertices, and all non-root vertices are adjacent to either one or three vertices (see Figure 5.4). These trees, also called **binary trees**, are discussed further in [5.6.9] and [5.13]. The number of non-isomorphic strongly binary trees of order $2n+1$ turns out to be exactly B_n. The one-to-one correspondence is obtained, in the forward direction, by deleting all the **leaves** (terminal nodes) of a strongly binary tree. To go in reverse, starting with a weakly binary tree, add two leaves to any vertex of degree 1 (or to the root if it has degree 0), and add one leaf to any vertex of degree 2 (or to the root if it has degree 1). Hence the same asymptotics apply in both weak and strong cases.

Also, in a commutative non-associative algebra, the expression x^4 is ambiguous and could be interpreted as xx^3 or x^2x^2. The expression x^5 likewise could mean xxx^3, xx^2x^2, or x^2x^3. Clearly B_{n-1} is the number of possible interpretations of x^n; thus $\{B_n\}$ is sometimes called the Wedderburn-Etherington sequence [11–15].

5.6.1 Chemical Isomers

A **weakly ternary tree** is a rooted tree for which the root is adjacent to at most three vertices and all non-root vertices are adjacent to at most four vertices. For instance, there exist eight non-isomorphic weakly ternary trees of order 5. The asymptotics of R_n, the number of non-isomorphic weakly ternary trees of order n, were again obtained by Otter [3, 15–17]:

$$\lim_{n\to\infty}\frac{R_n n^{\frac{3}{2}}}{\xi_R^n}=\eta_R,$$

where $\xi_R^{-1}=0.3551817423\ldots=(2.8154600332\ldots)^{-1}$ is the radius of convergence for

$$\begin{aligned}R(x)&=\sum_{n=0}^{\infty}R_nx^n\\&=1+x+x^2+2x^3+4x^4+8x^5+17x^6+39x^7+89x^8+211x^9+\cdots,\\ \eta_R&=\sqrt{\frac{\xi_R}{2\pi}}\left(-1+\rho+\frac{1}{\xi_R^3}R'\left(\frac{1}{\xi_R^2}\right)\rho+\frac{1}{\xi_R^4}R'\left(\frac{1}{\xi_R^3}\right)\right)^{\frac{1}{2}}\rho^{-\frac{1}{2}}\\&=0.5178759064\ldots,\end{aligned}$$

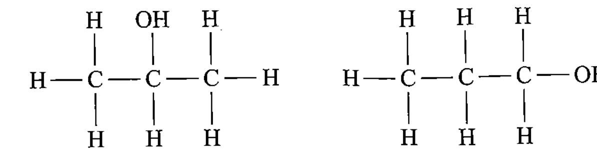


Figure 5.5. The formula C_3H_7OH (propanol) has two isomers.

and $\rho=R(\xi_R^{-1})$. The series coefficients arise from

$$R(x)=1+\frac{1}{6}x\left(R(x)^3+3R(x)R(x^2)+2R(x^3)\right).$$

An application of this material involves organic chemistry [18–21]: R_n is the number of **constitutional isomers** of the molecular formula $C_nH_{2n+1}OH$ (alcohols – see Figure 5.5). Constitutional isomeric pairs differ in their atomic connectivity, but the relative positioning of the OH group is immaterial.

Further, if we define [18, 19, 22, 23]

$$\begin{aligned}r(x)=&\frac{1}{24}x\left(R(x)^4+6R(x)^2R(x^2)+8R(x)R(x^3)+3R(x^2)^2+6R(x^4)\right)\\&-\frac{1}{2}\left(R(x)^2-R(x^2)\right)+R(x)\end{aligned}$$

then

$$\begin{aligned}r(x)&=\sum_{n=0}^{\infty}r_nx^n\\&=1+x+x^2+x^3+2x^4+3x^5+5x^6+9x^7+18x^8+35x^9+75x^{10}+\cdots\end{aligned}$$

and r_n is the number of constitutional isomers of the molecular formula C_nH_{2n+2} (alkanes – see Figure 5.6). The series $r(x)$ is related to $R(x)$ as $t(x)$ is related to $T(x)$ (in the sense that r, t are free and R, T are rooted); its radius of convergence is likewise ξ_R^{-1} and

$$\lim_{n\to\infty}\frac{r_nn^{\frac{5}{2}}}{\xi_R^n}=2\pi\frac{\eta_R^3}{\xi_R}\rho=0.6563186958\ldots.$$

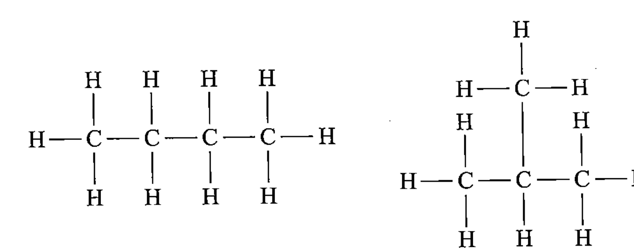


Figure 5.6. The formula C_4H_{10} (butane) has two isomers.

generating function is

$$H(x) = \sum_{n=1}^{\infty} H_n x^n$$
$$= x^2 + x^4 + x^5 + 2x^6 + 3x^7 + 6x^8 + 10x^9 + 19x^{10} + 35x^{11} + \cdots = x\tilde{H}(x).$$

It can be proved that [5,29]

$$\lim_{n\to\infty} \frac{H_n n^{\frac{3}{2}}}{\xi_H^n} = \eta_H = \frac{1}{\xi_H\sqrt{2\pi}}\left(\frac{\xi_H}{\xi_H+1} + \sum_{k=2}^{\infty}\frac{1}{\xi_H^k}\tilde{H}'\left(\frac{1}{\xi_H^k}\right)\right)^{\frac{1}{2}} = 0.1924225474\ldots,$$
$$\lim_{n\to\infty}\frac{h_n n^{\frac{5}{2}}}{\xi_H^n} = 2\pi\xi_H^2(\xi_H+1)\eta_H^3 = 0.6844472720\ldots,$$

where $\xi_H^{-1} = 0.4567332095\ldots = (2.1894619856\ldots)^{-1}$ is the radius of convergence for both $H(x)$ and $h(x)$, and further

$$\tilde{H}(x) = \frac{x}{x+1}\exp\left(\sum_{k=1}^{\infty}\frac{\tilde{H}(x^k)}{k}\right),$$
$$h(x) = (x+1)\tilde{H}(x) - \frac{x+1}{2}\tilde{H}(x)^2 - \frac{x-1}{2}\tilde{H}(x^2).$$

If we take into account the ordering (from left to right) of the subtrees of any vertex, then **ordered trees** arise and different enumeration problems occur. For example, define two ordered rooted trees σ and τ to be **cyclically isomorphic** if σ and τ are isomorphic as rooted trees, and if τ can be obtained from σ by circularly rearranging all the subtrees of any vertex, or likewise for each of several vertices. The equivalence classes under this relation are called **mobiles**. There exist fifty-one mobiles of order 7 but only forty-eight rooted trees of order 7 (see Figure 5.8).

The generating function for mobiles is [22,26,30]

$$M(x) = \sum_{n=1}^{\infty} M_n x^n$$
$$= x + x^2 + 2x^3 + 4x^4 + 9x^5 + 20x^6 + 51x^7 + 128x^8 + 345x^9 + \cdots,$$

$$M(x) = x\left(1 - \sum_{k=1}^{\infty}\frac{\varphi(k)}{k}\ln(1 - M(x^k))\right), \quad M_n \sim \eta_M n^{-\frac{3}{2}}\xi_M^n,$$

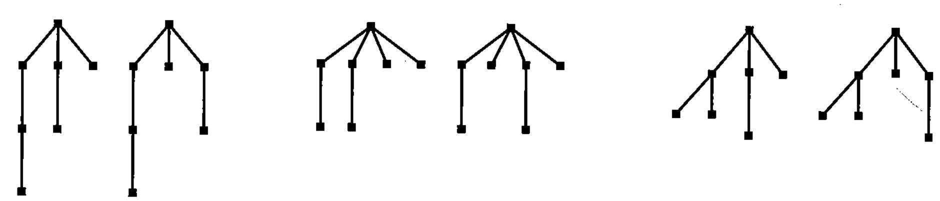


Figure 5.8. There exist three pairs of distinct mobiles (of order 7) that are identical as rooted trees.

where φ is the Euler totient function [2.7] and $\xi_M^{-1} = 0.3061875165\ldots = (3.2659724710\ldots)^{-1}$.

If we **label** the vertices of a graph distinctly with the integers $1, 2, \ldots, n$, the corresponding enumeration problems often simplify; for example, there are exactly n^{n-2} labeled free trees and n^{n-1} labeled rooted trees. For labeled mobiles, the problem becomes quite interesting, with exponential generating function [31]

$$\hat{M}(x) = \sum_{n=1}^{\infty}\frac{\hat{M}_n}{n!}x^n$$
$$= x + \frac{2}{2!}x^2 + \frac{9}{3!}x^3 + \frac{68}{4!}x^4 + \frac{730}{5!}x^5 + \frac{10164}{6!}x^6 + \frac{173838}{7!}x^7 + \cdots,$$
$$\hat{M}(x) = x\,(1 - \ln(1 - \hat{M}(x))), \quad \hat{M}_n \sim \hat{\eta}\,\hat{\xi}^n n^{n-1},$$

where $\hat{\xi} = e^{-1}(1-\mu)^{-1} = 1.1574198038\ldots$, $\hat{\eta} = \sqrt{\mu(1-\mu)} = 0.4656386467\ldots$, and $\mu = 0.6821555671\ldots$ is the unique solution of the equation $\mu(1-\mu)^{-1} = 1 - \ln(1-\mu)$.

An **increasing tree** is a labeled rooted tree for which the labels along any branch starting at the root are increasing. The root must be labeled 1. Again, for increasing mobiles, enumeration provides interesting constants [32]:

$$\tilde{M}(x) = \sum_{n=1}^{\infty}\frac{\tilde{M}_n}{n!}x^n$$
$$= x + \frac{1}{2!}x^2 + \frac{2}{3!}x^3 + \frac{7}{4!}x^4 + \frac{36}{5!}x^5 + \frac{245}{6!}x^6 + \frac{2076}{7!}x^7 + \cdots,$$
$$\tilde{M}'(x) = 1 - \ln(1 - \tilde{M}(x)), \quad \tilde{M}_n \sim \tilde{\xi}^{n-1} n!\left(\frac{1}{n^2} - \frac{1}{n^2\ln(n)} + O\left(\frac{1}{n^2\ln(n)^2}\right)\right)$$

where $\tilde{\xi}^{-1} = -e\,\mathrm{Ei}(-1) = 0.5963473623\ldots = e^{-1}(0.6168878482\ldots)^{-1}$ is the Euler–Gompertz constant [6.2]. See a strengthening of these asymptotics in [31,33].

5.6.3 Attributes

Thus far, we have discussed only enumeration issues. Otter's original constants α and β, however, appear in several asymptotic formulas governing other attributes of trees. By the **degree** (or **valency**) of a vertex, we mean the number of vertices that are adjacent to it. Given a random rooted tree with n vertices, the expected degree of the root is [34]

$$\theta = 1 + \sum_{i=1}^{\infty} T\left(\frac{1}{\alpha^i}\right) = 2 + \sum_{j=1}^{\infty} T_j\frac{1}{\alpha^j(\alpha^j - 1)} = 2.1918374031\ldots$$

as $n \to \infty$, and the variance of the degree of the root is

$$\sum_{i=1}^{\infty} iT\left(\frac{1}{\alpha^i}\right) = 1 + \sum_{j=1}^{\infty} T_j\frac{2\alpha^j - 1}{\alpha^j(\alpha^j - 1)^2} = 1.4741726868\ldots.$$

By the **distance** between two vertices, we mean the number of edges in the shortest path connecting them. The average distance between a vertex and the root is

$$\frac{1}{2}\left(\frac{2\pi}{\beta}\right)^{\frac{1}{3}} n^{\frac{1}{2}} = (1.1365599187\ldots)n^{\frac{1}{2}}$$

as $n \to \infty$, and the variance of the distance is

$$\frac{4-\pi}{4\pi}\left(\frac{2\pi}{\beta}\right)^{\frac{2}{3}} n = (0.3529622229\ldots)n.$$

Let v be an arbitrary vertex in a random free tree with n vertices and let p_m denote the probability, in the limit as $n \to \infty$, that v is of degree m. Then [35]

$$p_1 = \frac{\alpha^{-1} + \sum\limits_{k=1}^{\infty} D_k \frac{\alpha^{-2k}}{1-\alpha^{-k}}}{1 + \sum\limits_{k=1}^{\infty} kT_k \frac{\alpha^{-2k}}{1-\alpha^{-k}}} = 0.4381562356\ldots,$$

where $D_1 = 1$ and $D_{k+1} = \sum_{j=1}^{n}\left(\sum_{d|j} D_d\right) T_{k-j+1}$. Clearly $p_m \to 0$ as $m \to \infty$. More precisely, if

$$\omega = \prod_{i=1}^{\infty}\left(1-\frac{1}{\alpha^i}\right)^{-T_{i+1}} = \exp\left(\sum_{j=1}^{\infty}\frac{1}{j}\left[\alpha^j T\left(\frac{1}{\alpha^j}\right) - 1\right]\right) = 7.7581602911\ldots$$

then $\lim_{m\to\infty} \alpha^m p_m$ is given by [36,37]

$$(2\pi\beta^2)^{-\frac{1}{3}}\omega = (1.2160045618\ldots)^{-1}\omega = 6.3800420942\ldots.$$

We will need both θ and ω later. See also [38,39].

Let G be a graph and let $A(G)$ be the automorphism group of G. A vertex v of G is a **fixed point** if $\varphi(v) = v$ for every $\varphi \in A(G)$. Let q denote the probability, in the limit as $n \to \infty$, that an arbitrary vertex in a random tree of order n is a fixed point. Harary & Palmer [7,40] proved that

$$q = (2\pi\beta^2)^{-\frac{1}{3}}\left(1 - E\left(\frac{1}{\alpha^2}\right)\right) = 0.6995388700\ldots,$$

where $E(x) = T(x)(1 + F(x) - F(x^2))$. Interestingly, the same value q applies for rooted trees as well.

For reasons of space, we omit discussion of constants associated with covering and packing [41–43], as well as counting maximally independent sets of vertices [44–47], games [48], and equicolorable trees [49].

5.6.4 Forests

Let f_n denote the number of non-isomorphic forests of order n; then the generating function [26]

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} f_n x^n \\ &= x + 2x^2 + 3x^3 + 6x^4 + 10x^5 + 20x^6 + 37x^7 + 76x^8 + 153x^9 + 329x^{10} + \cdots \end{aligned}$$

satisfies

$$1 + f(x) = \exp\left(\sum_{k=1}^{\infty}\frac{t(x^k)}{k}\right), \quad f_n = \frac{1}{n}\sum_{k=1}^{n}\left(\sum_{d|k} d t_d\right) f_{n-k}$$

and $f_0 = 1$ for the sake only of the latter formula. Palmer & Schwenk [50] showed that

$$f_n \sim c\, t_n = \left(1 + f\left(\frac{1}{\alpha}\right)\right) t_n = (1.9126258077\ldots)t_n.$$

If a forest is chosen at random, then as $n \to \infty$, the expected number of trees in the forest is

$$1 + \sum_{i=1}^{\infty} t\left(\frac{1}{\alpha^i}\right) = \frac{3}{2} + \frac{1}{2}T\left(\frac{1}{\alpha^2}\right) + \sum_{j=1}^{\infty} t_j \frac{1}{\alpha^j(\alpha^j - 1)} = 1.7555101394\ldots.$$

The corresponding number for rooted trees is $\theta = 2.1918374031\ldots$, a constant that unsurprisingly we encountered earlier [5.6.3]. The probability of exactly k rooted trees in a random forest is asymptotically $\omega\alpha^{-k} = (7.7581602911\ldots)\alpha^{-k}$. For free trees, the analogous probability likewise drops off geometrically as α^{-k} with coefficient

$$\frac{\alpha}{c}\prod_{i=1}^{\infty}\left(1-\frac{1}{\alpha^i}\right)^{-t_{i+1}} = \frac{\alpha}{c}\exp\left(\sum_{j=1}^{\infty}\frac{1}{j}\left[\alpha^j t\left(\frac{1}{\alpha^j}\right) - 1\right]\right) = 3.2907434386\ldots.$$

Also, the asymptotic probability that two rooted forests of order n have no tree in common is [51]

$$\prod_{i=1}^{\infty}\left(1-\frac{1}{\alpha^{2i}}\right)^{T_i} = \exp\left(-\sum_{j=1}^{\infty}\frac{1}{j}T\left(\frac{1}{\alpha^{2j}}\right)\right) = 0.8705112052\ldots.$$

5.6.5 Cacti and 2-Trees

We now examine graphs that are not trees but are nevertheless tree-like. A **cactus** is a connected graph in which no edge lies on more than one (minimal) cycle [52–54]. See Figure 5.9. If we further assume that every edge lies on exactly one cycle and that all cycles are polygons with m sides for a fixed integer m, the cactus is called an m-**cactus**. By convention, a 2-cactus is simply a tree. Discussions of 3-cacti appear in [4], 4-cacti in [55], and m-cacti with vertex coloring in [56]; we will not talk about such special

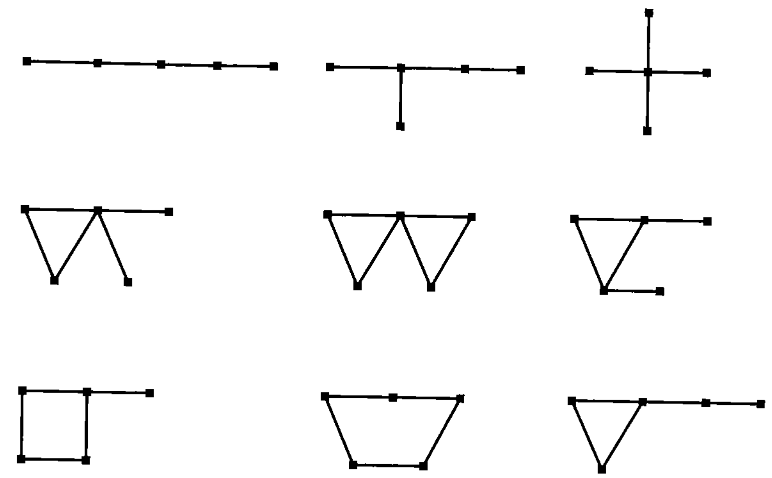


Figure 5.9. There exist nine non-isomorphic cacti of order 5.

cases. The generating functions for cacti and rooted cacti are [57]

$$c(x) = \sum_{n=1}^{\infty} c_n x^n$$
$$= x + x^2 + 2x^3 + 4x^4 + 9x^5 + 23x^6 + 63x^7 + 188x^8 + 596x^9 + 1979x^{10} + \cdots,$$
$$C(x) = \sum_{n=1}^{\infty} C_n x^n$$
$$= x + x^2 + 3x^3 + 8x^4 + 26x^5 + 84x^6 + 297x^7 + 1066x^8 + 3976x^9 + \cdots,$$

and these satisfy [58–60]

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

$$c(x) = C(x) - \frac{1}{2}\sum_{k=1}^{\infty}\frac{\varphi(k)}{k}\ln(1 - C(x^k)) + \frac{(C(x)+1)(C(x)^2 - 2C(x) + C(x^2))}{4(C(x)-1)(C(x^2)-1)},$$

with radius of convergence 0.2221510651 .... For the labeled case, we have

$$\hat{c}(x) = \sum_{n=1}^{\infty}\frac{\hat{c}_n}{n!}x^n$$
$$= x + \frac{1}{2!}x^2 + \frac{4}{3!}x^3 + \frac{31}{4!}x^4 + \frac{362}{5!}x^5 + \frac{5676}{6!}x^6 + \frac{111982}{7!}x^7 + \cdots,$$
$$\hat{C}(x) = \sum_{n=1}^{\infty}\frac{\hat{C}_n}{n!}x^n$$
$$= x + \frac{2}{2!}x^2 + \frac{12}{3!}x^3 + \frac{124}{4!}x^4 + \frac{1810}{5!}x^5 + \frac{34056}{6!}x^6 + \frac{783874}{7!}x^7 + \cdots,$$

and these satisfy

$$\hat{C}(x) = x\exp\left(\frac{\hat{C}(x)}{2}\frac{2-\hat{C}(x)}{1-\hat{C}(x)}\right), \quad x\hat{c}'(x) = \hat{C}(x),$$

with radius of convergence 0.2387401436 ....

A **2-tree** is defined recursively as follows [4]. A 2-tree of rank 1 is a triangle (a graph with three vertices and three edges), and a 2-tree of rank $n \geq 2$ is built from a 2-tree of rank $n-1$ by creating a new vertex of degree 2 adjacent to each of two existing adjacent vertices. Hence a 2-tree of rank n has $n+2$ vertices and $2n+1$ edges. The generating function for 2-trees is [61]

$$w(x) = \sum_{n=0}^{\infty} w_n x^n$$
$$= 1 + x + x^2 + 2x^3 + 5x^4 + 12x^5 + 39x^6 + 136x^7 + 529x^8 + 2171x^9 + \cdots$$
$$w(x) = \frac{1}{2}\left[W(x) + \exp\left(\sum_{k=1}^{\infty}\frac{1}{2k}(2x^k W(x^{2k}) + x^{2k}W(x^{2k})^2 - x^{2k}W(x^{4k}))\right)\right]$$
$$+ \frac{1}{3}x\left(W(x^3) - W(x)^3\right),$$

where $W(x)$ is the generating function for 2-trees with a distinguished and oriented edge:

$$W(x) = \sum_{n=0}^{\infty} W_n x^n$$
$$= 1 + x + 3x^2 + 10x^3 + 39x^4 + 160x^5 + 702x^6 + 3177x^7 + 14830x^8 + \cdots$$

$$W(x) = \exp\left(\sum_{k=1}^{\infty}\frac{x^k W(x^k)^2}{k}\right), \quad w_n \sim \eta_w n^{-\frac{5}{2}}\xi_w^n.$$

Further, $w(x)$ has radius of convergence $\xi_w^{-1} = 0.1770995223\ldots = (5.6465426162\ldots)^{-1}$ and

$$\eta_w = \frac{1}{16\xi\sqrt{\pi}}\left(\xi + 2\tilde{W}'\left(\frac{1}{\xi}\right)\tilde{W}\left(\frac{1}{\xi}\right)^{-1}\right)^{\frac{3}{2}} = 0.0948154165\ldots,$$

$$\tilde{W}(x) = e^{-xW(x)^2}W(x).$$

5.6.6 *Mapping Patterns*

We studied labeled functional graphs on n vertices in [5.4]. Let us remove the labels and consider only graph isomorphism classes, called **mapping patterns**. Observe that the original Otter constants α and β play a crucial role here. The generating function labeled acyclic digraphs is [65, 71–74]

$$A(x) = \sum_{n=0}^{\infty} \frac{A_n}{n!2^{\binom{n}{2}}} x^n = 1 + x + \frac{3}{2! \cdot 2} x^2 + \frac{25}{3! \cdot 2^3} x^3 + \frac{543}{4! \cdot 2^6} x^4 + \frac{29281}{5! \cdot 2^{10}} x^5 + \cdots,$$

$$A'(x) = A(x)^2 A(\tfrac{1}{2}x)^{-1}, \quad A_n \sim \frac{n! 2^{\binom{n}{2}}}{\eta_A \, \xi_A^n},$$

where $\xi_A = 1.4880785456\ldots$ is the smallest positive zero of the function

$$\lambda(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n! 2^{\binom{n}{2}}} x^n = A(x)^{-1}, \quad \lambda'(x) = -\lambda(\tfrac{1}{2}x),$$

and $\eta_A = \xi_A \lambda(\xi_A/2) = 0.5743623733\ldots = (1.7410611252\ldots)^{-1}$. It is curious that the function $\lambda(-x)$ was earlier studied by Mahler [75] with regard to enumerating partitions of integers into powers of 2. See [76, 77] for discussion of the unlabeled acyclic digraph analog.

5.6.8 Data Structures

To a combinatorialist, the phrase "(strongly) binary tree with $2n + 1$ vertices" means an isomorphism class of trees. To a computer scientist, however, the same phrase virtually always includes the word "ordered," whether stated explicitly or not. Hence the phrase "random binary tree" is sometimes ambiguous in the literature: The sample space has B_n elements for the former person but $\binom{2n}{n}/(n + 1)$ elements for the latter! We cannot hope here to survey the role of trees in computer algorithms, only to provide a few constants.

A **leftist tree** of size n is an ordered binary tree with n leaves such that, in any subtree σ, the leaf closest to the root of σ is in the right subtree of σ. The generating function of leftist trees is [6, 65, 78, 79]

$$\begin{aligned} L(x) &= \sum_{n=0}^{\infty} L_n x^n \\ &= x + x^2 + x^3 + 2x^4 + 4x^5 + 8x^6 + 17x^7 + 38x^8 + 87x^9 + 203x^{10} + \cdots \end{aligned}$$

$$L(x) = x + \frac{1}{2} L(x)^2 + \frac{1}{2} \sum_{m=1}^{\infty} l_m(x)^2 = \sum_{m=1}^{\infty} l_m(x),$$

where the auxiliary generating functions $l_m(x)$ satisfy

$$l_1(x) = x, \quad l_2(x) = x L(x), \quad l_{m+1}(x) = l_m(x) \left(L(x) - \sum_{k=1}^{m-1} l_k(x) \right), \quad m \ge 2.$$

It can be proved (with difficulty) that

$$L_n \sim (0.2503634293\ldots) \cdot (2.7494879027\ldots)^n n^{-\frac{3}{2}}.$$

Leftist trees are useful in certain sorting and merging algorithms.

A **2,3-tree** of size n is a rooted ordered tree with n leaves satisfying the following:

- Each non-leaf vertex has either 2 or 3 successors.
- All of the root-to-leaf paths have the same length.

The generating function of 2,3-trees (no relation to 2-trees!) is [65, 80, 81]

$$\begin{aligned} Z(x) &= \sum_{n=0}^{\infty} Z_n x^n \\ &= x + x^2 + x^3 + x^4 + 2x^5 + 2x^6 + 3x^7 + 4x^8 + 5x^9 + 8x^{10} + 14x^{11} + \cdots \end{aligned}$$

$$Z(x) = x + Z(x^2 + x^3), \quad Z_n = \sum_{k=\lceil n/3 \rceil}^{\lfloor n/2 \rfloor} \binom{k}{3k-n} Z_k \sim \varphi^n n^{-1} f(\ln(n)),$$

where φ is the Golden mean [1.2] and $f(x)$ is a nonconstant, positive, continuous function that is periodic with period $\ln(4 - \varphi) = 0.867\ldots$, has mean $(\varphi \ln(4 - \varphi))^{-1} = 0.712\ldots$, and oscillates between $0.682\ldots$ and $0.806\ldots$. These are also a particular type of **B-trees**. A similar analysis [82] uncovers the asymptotics of what are known as AVL-trees (or height-balanced trees). Such trees support efficient database searches, deletions, and insertions; other varieties are too numerous to mention.

If τ is an ordered binary tree, then its **height** and **register functions** are recursively defined by [83]

$$\mathrm{ht}(\tau) = \begin{cases} 0 & \text{if } \tau \text{ is a point}, \\ 1 + \max(\mathrm{ht}(\tau_L), \mathrm{ht}(\tau_R)) & \text{otherwise}, \end{cases}$$

$$\mathrm{rg}(\tau) = \begin{cases} 0 & \text{if } \tau \text{ is a point}, \\ 1 + \mathrm{rg}(\tau_L) & \text{if } \mathrm{rg}(\tau_L) = \mathrm{rg}(\tau_R), \\ \max(\mathrm{rg}(\tau_L), \mathrm{rg}(\tau_R)) & \text{otherwise}, \end{cases}$$

where τ_L and τ_R are the left and right subtrees of the root. That is, $\mathrm{ht}(\tau)$ is the number of edges along the longest branch from the root, whereas $\mathrm{rg}(\tau)$ is the minimum number of registers needed to evaluate the tree (thought of as an arithmetic expression). If we randomly select a binary tree τ with $2n + 1$ vertices, then the asymptotics of $\mathrm{E}(\mathrm{ht}(\tau))$ involve $2\sqrt{\pi n}$ as mentioned in [1.4], and those of $\mathrm{E}(\mathrm{rg}(\tau))$ involve $\ln(n)/\ln(4)$ plus a zero mean oscillating function [2.16]. Also, define $\mathrm{ym}(\tau)$ to be the number of maximal subtrees of τ having register function exactly 1 less than $\mathrm{rg}(\tau)$. Prodinger [84], building upon the work of Yekutiele & Mandelbrot [85], proved that $\mathrm{E}(\mathrm{ym}(\tau))$ is asymptotically

$$\frac{2G}{\pi \ln(2)} + \frac{5}{2} = 3.3412669407\ldots$$

plus a zero mean oscillating function, where G is Catalan's constant [1.7]. This is also known as the **bifurcation ratio** at the root, which quantifies the hierarchical complexity of more general branching structures.

5.6.9 Galton–Watson Branching Process

Thus far, by "random binary trees," it is meant that we select binary trees with n vertices from a population endowed with the uniform probability distribution. The integer n is fixed.

It is also possible, however, to *grow* binary trees (rather than to merely select them). Fix a probability $0 < p < 1$ and define recursively a (strongly) binary tree τ in terms of left and right subtrees of the root as follows: Take $\tau_L = \emptyset$ with probability $1 - p$, and independently take $\tau_R = \emptyset$ with probability $1 - p$. It can be shown [86–88] that this process terminates, that is, τ is a finite tree, with **extinction probability** 1 if $p \leq 1/2$ and $1/p - 1$ if $p > 1/2$. Of course, the number of vertices N is here a random variable, called the **total progeny**.

Much can be said about the Bienaymé–Galton–Watson process (which is actually more general than described here). We focus on just one detail. Let N_k denote the number of vertices at distance k from the root, that is, the size of the k^{th} generation. Consider the subcritical case $p < 1/2$. Let a_k denote the probability that $N_k = 0$; then the sequence $a_0, a_1, a_2, \ldots$ obeys the quadratic recurrence [6.10]

$$a_0 = 0, \qquad a_k = (1-p) + p a_{k-1}^2 \quad \text{for } k \geq 1, \qquad \lim_{k\to\infty} a_k = 1.$$

What can be said about the convergence rate of $\{a_k\}$? It can be proved that

$$C(p) = \lim_{k\to\infty} \frac{1-a_k}{(2p)^k} = \prod_{l=0}^{\infty} \frac{1+a_l}{2},$$

which has no closed-form expression in terms of p, as far as is known. This is over and beyond the fact, of greatest interest to us here, that $\mathrm{P}(N_k > 0) \sim C(p)(2p)^k$ for $0 < p < 1/2$. Other interesting parameters are the **moment of extinction** $\min\{k : N_k = 0\}$ or tree height, and the **maximal generation size** $\max\{N_k : k \geq 0\}$ or tree width.

5.6.10 Erdös–Rényi Evolutionary Process

Starting with n initially disconnected vertices, define a random graph by successively adding edges between pairs of distinct points, chosen uniformly from $\binom{n}{2}$ candidates without replacement. Continue with this process until no candidate edges are left [89–92].

At some stage of the evolution, a **complex component** emerges, that is, the first component possessing more than one cycle. It is remarkable that this complex component will usually remain unique throughout the entire process, and the probability that this is true is $5\pi/18 = 0.8726\ldots$ as $n \to \infty$. In other words, the first component that acquires more edges than vertices is quite likely to become the **giant component** of the random graph. The probability that exactly two complex components emerge is $50\pi/1296 = 0.1212\ldots$, but the probability ($> 0.9938\ldots$) that the evolving graph never has more than two complex components at any time is not precisely known [93].

There are many related results, but we mention only one. Start with an $m \times n$ rectangular grid of rooms, each with four walls. Successively remove interior walls in a random manner such that, at some step in the procedure, the associated graph (with all

mn rooms as vertices and all neighboring pairs of rooms with open passage as edges) becomes a tree. Stop when this condition is met; the result is a **random maze** [94]. The difficulty lies in detecting whether the addition of a new edge creates an unwanted cycle. An efficient way of doing this (maintaining equivalence classes that change over time) is found in QF and QFW, two of a class of **union-find algorithms** in computer science. Exact performance analyses of QF and QFW appear in [95–97], using random graph theory and a variant of the Erdös-Rényi process.

[1] F. Harary, *Graph Theory*, Addison-Wesley, 1969; MR 41 #1566.

[2] R. Sedgewick and P. Flajolet, *Introduction to the Analysis of Algorithms*, Addison-Wesley, 1996.

[3] R. Otter, The number of trees, *Annals of Math.* 49 (1948) 583–599; MR 10,53c.

[4] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, 1973; MR 50 #9682.

[5] F. Harary, R. W. Robinson, and A. J. Schwenk, Twenty-step algorithm for determining the asymptotic number of trees of various species, *J. Austral. Math. Soc. Ser. A* 20 (1975) 483–503; corrigenda 41 (1986) 325; MR 53 #10644 and MR 87j:05091.

[6] A. M. Odlyzko, Asymptotic enumeration methods, *Handbook of Combinatorics*, ed. R. L. Graham, M. Groetschel, and L. Lovasz, MIT Press, 1995, pp. 1063–1229; MR 97b:05012.

[7] D. J. Broadhurst and D. Kreimer, Renormalization automated by Hopf algebra, *J. Symbolic Comput.* 27 (1999) 581–600; hep-th/9810087; MR 2000h:81167.

[8] E. A. Bender, Asymptotic methods in enumeration, *SIAM Rev.* 16 (1974) 485–515; errata 18 (1976) 292; MR 51 #12545 and MR 55 #10276.

[9] D. E. Knuth, *The Art of Computer Programming*, v. 1, *Fundamental Algorithms*, Addison-Wesley, 1997, pp. 396–397, 588–589; MR 51 #14624.

[10] J. N. Franklin and S. W. Golomb, A function-theoretic approach to the study of nonlinear recurring sequences, *Pacific J. Math.* 56 (1975) 455–468; MR 51 #10212.

[11] J. H. M. Wedderburn, The functional equation $g(x^2) = 2ax + [g(x)]^2$, *Annals of Math.* 24 (1922-23) 121–140.

[12] I. M. H. Etherington, Non-associate powers and a functional equation, *Math. Gazette* 21 (1937) 36–39, 153.

[13] H. W. Becker, Genetic algebra, *Amer. Math. Monthly* 56 (1949) 697–699.

[14] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Reidel, 1974, pp. 52–55, 273–279; MR 57 #124.

[15] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000055, A000081, A000598, A000602, A000621, A000625, and A001190.

[16] C. Bailey, E. Palmer, and J. Kennedy, Points by degree and orbit size in chemical trees. II, *Discrete Appl. Math.* 5 (1983) 157–164; MR 85b:05098.

[17] E. A. Bender and S. Gill Williamson, *Foundations of Applied Combinatorics*, Addison-Wesley, 1991, pp. 372–373, 395–396.

[18] G. Pólya and R. C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, 1987, pp. 4–8, 78–86, 130–134; MR 89f:05013.

[19] A. T. Balaban, *Chemical Applications of Graph Theory*, Academic Press, 1976; MR 58 #21870.

[20] N. Trinajstic, *Chemical Graph Theory*, v. 2, CRC Press, 1992; MR 93g:92034.

[21] D. H. Rouvray, Combinatorics in chemistry, *Handbook of Combinatorics*, ed. R. L. Graham, M. Groetschel, and L. Lovasz, MIT Press, 1995, pp. 1955–1981; MR 97e:92017.

[22] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-Like Structures*, Cambridge Univ. Press, 1998; MR 2000a:05008.

[23] E. M. Rains and N. J. A. Sloane, On Cayley's enumeration of alkanes (or 4-valent trees), *J. Integer Seq.* 2 (1999) 99.1.1; math.CO/0207176; MR 99j:05010.

[24] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, unpublished manuscript, 2001.

[25] R. W. Robinson, F. Harary, and A. T. Balaban, The numbers of chiral and achiral alkanes and monosubstituted alkanes, *Tetrahedron* 32 (1976) 355–361.

[26] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000014, A000220, A001678, A004111, A005200, A005201, A029768, A032200, A038037, A005195, and A059123.
[27] A. Meir, J. W. Moon, and J. Mycielski, Hereditarily finite sets and identity trees, *J. Combin. Theory Ser. B* 35 (1983) 142–155; MR 85h:05009.
[28] F. Harary and G. Prins, The number of homeomorphically irreducible trees, and other species, *Acta Math.* 101 (1959) 141–162; MR 21 #653.
[29] S. R. Finch, Counting topological trees, unpublished note (2001).
[30] C. G. Bower, Mobiles (cyclic trees), unpublished note (2001).
[31] I. M. Gessel, B. E. Sagan, and Y. N. Yeh, Enumeration of trees by inversions, *J. Graph Theory* 19 (1995) 435–459; MR 97b:05070.
[32] F. Bergeron, P. Flajolet, and B. Salvy, Varieties of increasing trees, *17th Colloq. on Trees in Algebra and Programming (CAAP)*, Proc. 1992 Rennes conf., ed. J.-C. Raoult, Lect. Notes in Comp. Sci. 581, Springer-Verlag, 1992, pp. 24–48; MR 94j:68233.
[33] W. Y. C. Chen, The pessimistic search and the straightening involution for trees, *Europ. J. Combin.* 19 (1998) 553–558; MR 99g:05103.
[34] A. Meir and J. W. Moon, On the altitude of nodes in random trees, *Canad. J. Math.* 30 (1978) 997–1015; MR 80k:05043.
[35] R. W. Robinson and A. J. Schwenk, The distribution of degrees in a large random tree, *Discrete Math.* 12 (1975) 359–372; MR 53 #10629.
[36] A. J. Schwenk, An asymptotic evaluation of the cycle index of a symmetric group, *Discrete Math.* 18 (1977) 71–78; MR 57 #9598.
[37] A. Meir and J. W. Moon, On an asymptotic evaluation of the cycle index of the symmetric group, *Discrete Math.* 46 (1983) 103–105; MR 84h:05071.
[38] C. K. Bailey, Distribution of points by degree and orbit size in a large random tree, *J. Graph Theory* 6 (1982) 283–293; MR 83i:05037.
[39] W. M. Y. Goh and E. Schmutz, Unlabeled trees: Distribution of the maximum degree, *Random Structures Algorithms* 5 (1994) 411–440; MR 95c:05111.
[40] F. Harary and E. M. Palmer, The probability that a point of a tree is fixed, *Math. Proc. Cambridge Philos. Soc.* 85 (1979) 407–415; MR 80f:05020.
[41] A. Meir and J. W. Moon, Packing and covering constants for certain families of trees. I, *J. Graph Theory* 1 (1977) 157–174; MR 57 #2965.
[42] A. Meir and J. W. Moon, Packing and covering constants for certain families of trees. II, *Trans. Amer. Math. Soc.* 233 (1977) 167–178; MR 57 #157.
[43] A. Meir and J. W. Moon, Path edge-covering constants for certain families of trees, *Utilitas Math.* 14 (1978) 313–333; MR 80a:05073.
[44] A. Meir and J. W. Moon, On maximal independent sets of nodes in trees, *J. Graph Theory* 12 (1988) 265–283; MR 89f:05064.
[45] P. Kirschenhofer, H. Prodinger, and R. F. Tichy, Fibonacci numbers of graphs. II, *Fibonacci Quart.* 21 (1983) 219–229; MR 85e:05053.
[46] P. Kirschenhofer, H. Prodinger, and R. F. Tichy, Fibonacci numbers of graphs. III: Planted plane trees, *Fibonacci Numbers and Their Applications*, Proc. 1984 Patras conf., ed. A. N. Philippou, G. E. Bergum, and A. F. Horadam, Reidel, 1986, pp. 105–120, MR 88k:05065.
[47] M. Drmota, On generalized Fibonacci numbers of graphs, *Applications of Fibonacci Numbers*, v. 3, Proc. 1988 Pisa conf., ed. G. E. Bergum, A. N. Philippou, and A. F. Horadam, Kluwer, 1990, pp. 63–76; MR 92e:05032.
[48] A. Meir and J. W. Moon, Games on random trees, *Proc. 15th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, Baton Rouge, 1984, ed. F. Hoffman, K. B. Reid, R. C. Mullin, and R. G. Stanton, Congr. Numer. 44, Utilitas Math., 1984, pp. 293–303; MR 86h:05045.
[49] N. Pippenger, Enumeration of equicolorable trees, *SIAM J. Discrete Math.* 14 (2001) 93–115; MR 2001j:05073.
[50] E. M. Palmer and A. J. Schwenk, On the number of trees in a random forest, *J. Combin. Theory Ser. B* 27 (1979) 109–121; MR 81j:05071.

[51] S. Corteel, C. D. Savage, H. S. Wilf, and D. Zeilberger, A pentagonal number sieve, *J. Combin. Theory Ser. A* 82 (1998) 186–192; MR 99d:11111.
[52] K. Husimi, Note on Mayers' theory of cluster integrals, *J. Chem. Phys.* 18 (1950) 682–684; MR 12,467i.
[53] G. E. Uhlenbeck and G. W. Ford, *Lectures in Statistical Mechanics*, Lect. in Appl. Math., v. 1, Amer. Math. Soc., 1963; MR 27 #1241.
[54] G. E. Uhlenbeck and G. W. Ford, The theory of linear graphs with applications to the theory of the virial development of the properties of gases, *Studies in Statistical Mechanics*, v. 1, ed. J. de Boer and G. Uhlenbeck, North-Holland, 1962, pp. 119–211; MR 24 #B2420.
[55] F. Harary and G. E. Uhlenbeck, On the number of Husimi trees. I, *Proc. Nat. Acad. Sci. USA* 39 (1953) 315–322; MR 14,836a.
[56] M. Bóna, M. Bousquet, G. Labelle, and P. Leroux, Enumeration of m-ary cacti, *Adv. Appl. Math.* 24 (2000) 22–56; MR 2001c:05072.
[57] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000083, A000237, A000314, A001205, A001372, A002861, A005750, A035351, and A054581.
[58] C. G. Bower, Cacti (mixed Husimi trees), unpublished note (2001).
[59] G. W. Ford and G. E. Uhlenbeck, Combinatorial problems in the theory of graphs. III, *Proc. Nat. Acad. Sci. USA* 42 (1956) 529–535; MR 18,326c.
[60] C. Domb, Graph theory and embeddings, *Phase Transitions and Critical Phenomena*, v. 3, ed. C. Domb and M. S. Green, Academic Press, 1974, pp. 33–37; MR 50 #6393.
[61] T. Fowler, I. Gessel, G. Labelle, and P. Leroux, Specifying 2-trees, *Formal Power Series and Algebraic Combinatorics (FPSAC)*, Proc. 2000 Moscow conf., ed. D. Krob, A. A. Mikhalev, and A. V. Mikhalev, Springer-Verlag, 2000, pp. 202–213; MR 2001m:05132.
[62] A. Meir and J. W. Moon, On random mapping patterns, *Combinatorica* 4 (1984) 61–70; MR 85g:05125.
[63] L. R. Mutafchiev, Large trees in a random mapping pattern, *Europ. J. Combin.* 14 (1993) 341–349; MR 94f:05072.
[64] L. R. Mutafchiev, Limit theorem concerning random mapping patterns, *Combinatorica* 8 (1988) 345–356; MR 90a:05170.
[65] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001205, A003024, A003087, A005217, A006196, A008483, A015701, A014535, and A037026.
[66] P. Hanlon, Counting interval graphs, *Trans. Amer. Math. Soc.* 272 (1982) 383–426; MR 83h:05050.
[67] F. S. Roberts, *Discrete Mathematical Models*, Prentice-Hall, 1976, pp. 111–140.
[68] P. C. Fishburn, *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets*, Wiley, 1985; MR 86m:06001.
[69] H. S. Wilf, *generatingfunctionology*, Academic Press, 1990, pp. 76–77, 151; MR 95a:05002.
[70] R. M. Robinson, A new absolute geometric constant, *Amer. Math. Monthly* 58 (1951) 462–469; addendum 59 (1952) 296–297; MR 13,200b.
[71] R. W. Robinson, Counting labeled acyclic digraphs, *New Directions in the Theory of Graphs*, Proc. 1971 Ann Arbor conf., ed. F. Harary, Academic Press, 1973, pp. 239–273; MR 51 #249.
[72] R. P. Stanley, Acyclic orientations of graphs, *Discrete Math.* 5 (1973) 171–178; MR 47 #6537.
[73] V. A. Liskovec, The number of maximal vertices of a random acyclic digraph (in Russian), *Teor. Veroyatnost. i Primenen.* 20 (1975) 412–421; Engl. transl. in *Theory Probab. Appl.* 20 (1975) 401–421; MR 52 #1822.
[74] E. A. Bender, L. B. Richmond, R. W. Robinson, and N. C. Wormald, The asymptotic number of acyclic digraphs. I, *Combinatorica* 6 (1986) 15–22; MR 87m:05102.
[75] K. Mahler, On a special functional equation, *J. London Math. Soc.* 15 (1940) 115–123; MR 2,133e.
[76] R. W. Robinson, Counting unlabeled acyclic digraphs, *Combinatorial Mathematics V*, Proc. 1976 Melbourne conf., ed. C. H. C. Little, Lect. Notes in Math. 622, Springer-Verlag, 1977, pp. 28–43; MR 57 #16129.

[77] E. A. Bender and R. W. Robinson, The asymptotic number of acyclic digraphs. II, *J. Combin. Theory Ser. B* 44 (1988) 363–369; MR 90a:05098.
[78] R. Kemp, A note on the number of leftist trees, *Inform. Process. Lett.* 25 (1987) 227–232.
[79] R. Kemp, Further results on leftist trees, *Random Graphs '87*, Proc. 1987 Poznan conf., ed. M. Karonski, J. Jaworski, and A. Rucinski, Wiley, 1990, pp. 103–130; MR 92e:05034.
[80] R. E. Miller, N. Pippenger, A. L. Rosenberg, and L. Snyder, Optimal 2,3-trees, *SIAM J. Comput.* 8 (1979) 42–59; MR 80c:68050.
[81] A. M. Odlyzko, Periodic oscillations of coefficients of power series that satisfy functional equations, *Adv. Math.* 44 (1982) 180–205; MR 84a:30042.
[82] A. M. Odlyzko, Some new methods and results in tree enumeration, *Proc. 13th Manitoba Conf. on Numerical Mathematics and Computing*, Winnipeg, 1983, ed. D. S. Meek and G. H. J. van Rees, Congr. Numer. 42, Utilitas Math., 1984, pp. 27–52; MR 85g:05061.
[83] H. Prodinger, Some recent results on the register function of a binary tree, *Random Graphs '85*, Proc. 1985 Poznan conf, ed. M. Karonski and Z. Palka, Annals of Discrete Math. 33, North-Holland, 1987, pp. 241–260; MR 89g:68058.
[84] H. Prodinger, On a problem of Yekutieli and Mandelbrot about the bifurcation ratio of binary trees, *Theoret. Comput. Sci.* 181 (1997) 181–194; also in *Proc. 1995 Latin American Theoretical Informatics Conf. (LATIN)*, Valparáiso, ed. R. A. Baeza-Yates, E. Goles Ch., and P. V. Poblete, Lect. Notes in Comp. Sci. 911, Springer-Verlag, 1995, pp. 461–468; MR 98i:68212.
[85] I. Yekutieli and B. B. Mandelbrot, Horton-Strahler ordering of random binary trees, *J. Phys. A* 27 (1994) 285–293; MR 94m:82022.
[86] T. E. Harris, *The Theory of Branching Processes*, Springer-Verlag, 1963; MR 29 #664.
[87] K. B. Athreya and P. Ney, *Branching Processes*, Springer-Verlag, 1972; MR 51 #9242.
[88] G. Sankaranarayanan, *Branching Processes and Its Estimation Theory*, Wiley, 1989; MR 91m:60156
[89] P. Erdös and A. Rényi, On random graphs. I, *Publ. Math. (Debrecen)* 6 (1959) 290–297; also in *Selected Papers of Alfréd Rényi*, v. 2, Akadémiai Kiadó, 1976, pp. 308–315; MR 22 #10924.
[90] B. Bollobás, The evolution of random graphs, *Trans. Amer. Math. Soc.* 286 (1984) 257–274; MR 85k:05090.
[91] B. Bollobás, *Random Graphs*, Academic Press, 1985; MR 87f:05152.
[92] S. Janson, T. Luczak, and A. Rucinski, *Random Graphs*, Wiley, 2000; MR 2001k:05180.
[93] S. Janson, D. E. Knuth, T. Luczak, and B. Pittel, The birth of the giant component, *Random Structures Algorithms* 4 (1993) 231–358; MR 94h:05070.
[94] M. A. Weiss, *Data Structures and Algorithm Analysis in C++*, 2nd ed., Addison-Wesley, 1999, pp. 320–322.
[95] A. C. C. Yao, On the average behavior of set merging algorithms, *8th ACM Symp. on Theory of Computing (STOC)*, Hershey, ACM, 1976, pp. 192–195; MR 55 #1819.
[96] D. E. Knuth and A. Schönhage, The expected linearity of a simple equivalence algorithm, *Theoret. Comput. Sci.* 6 (1978) 281–315; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 341–389; MR 81a:68049.
[97] B. Bollobás and I. Simon, Probabilistic analysis of disjoint set union algorithms, *SIAM J. Comput.* 22 (1993) 1053–1074; also in *17th ACM Symp. on Theory of Computing (STOC)*, Providence, ACM, 1985, pp. 124–231; MR 94j:05110.

5.7 Lengyel's Constant

5.7.1 Stirling Partition Numbers

Let S be a set with n elements. The set of all subsets of S has 2^n elements. By a **partition** of S we mean a disjoint set of nonempty subsets (called **blocks**) whose union is S. The set of partitions of S that possess exactly k blocks has $S_{n,k}$ elements, where $S_{n,k}$ is a

Stirling number of the second kind. The set of *all* partitions of S has B_n elements, where B_n is a **Bell number**:

$$B_n = \sum_{k=1}^{n} S_{n,k} = \frac{1}{e}\sum_{j=0}^{\infty}\frac{j^n}{j!} = \frac{d^n}{dx^n}\exp\left(e^x - 1\right)\Big|_{x=0}.$$

For example, $S_{4,1} = 1$, $S_{4,2} = 7$, $S_{4,3} = 6$, $S_{4,4} = 1$, and $B_4 = 15$. More generally, $S_{n,1} = 1$, $S_{n,2} = 2^{n-1} - 1$, and $S_{n,3} = \frac{1}{2}(3^{n-1}+1) - 2^{n-1}$. The following recurrences are helpful [1–4]:

$$S_{n,0} = \begin{cases} 1 & \text{if } n = 0, \\ 0 & \text{if } n \geq 1, \end{cases} \quad S_{n,k} = kS_{n-1,k} + S_{n-1,k-1} \quad \text{if } n \geq k \geq 1,$$

$$B_0 = 1, \quad B_n = \sum_{k=0}^{n-1}\binom{n-1}{k}B_k,$$

and corresponding asymptotics are discussed in [5–9].

5.7.2 Chains in the Subset Lattice of S

If U and V are subsets of S, write $U \subset V$ if U is a proper subset of V. This endows the set of all subsets of S with a **partial ordering**; in fact, it is a **lattice** with maximum element S and minimum element $\emptyset$. The number of **chains** $\emptyset = U_0 \subset U_1 \subset \cdots \subset U_{k-1} \subset U_k = S$ of length k is $k!S_{n,k}$. Hence the number of all chains from $\emptyset$ to S is [1, 6, 10]

$$\sum_{k=0}^{n} k!S_{n,k} = \sum_{j=0}^{\infty}\frac{j^n}{2^{j+1}} = \frac{1}{2}\operatorname{Li}_{-n}\left(\frac{1}{2}\right) = \frac{d^n}{dx^n}\frac{1}{2-e^x}\Big|_{x=0} \sim \frac{n!}{2}\left(\frac{1}{\ln(2)}\right)^{n+1},$$

where $\operatorname{Li}_m(x)$ is the polylogarithm function. Wilf [10] marveled at how accurate this asymptotic approximation is.

If we further insist that the chains are **maximal**, equivalently, that additional proper insertions are impossible, then the number of such chains is $n!$ A general technique due to Doubilet, Rota & Stanley [11], involving what are called *incidence algebras*, can be used to obtain the two aforementioned results, as well as to enumerate chains within more complicated posets [12].

As an aside, we give a deeper application of incidence algebras: to enumerating chains of linear subspaces within finite vector spaces [6]. Define the q-**binomial coefficient** and q-**factorial** by

$$\binom{n}{k}_q = \frac{\prod_{j=1}^{n}(q^j - 1)}{\prod_{j=1}^{k}(q^j - 1)\cdot\prod_{j=1}^{n-k}(q^j - 1)},$$

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

where $q > 1$. Note the special case in the limit as $q \to 1^+$. Consider the n-dimensional vector space $\mathbb{F}_q^n$ over the finite field $\mathbb{F}_q$, where q is a prime power [12–16]. The number of k-dimensional linear subspaces of $\mathbb{F}_q^n$ is $\binom{n}{k}_q$ and the total number of linear subspaces of $\mathbb{F}_q^n$ is asymptotically $c_e q^{n^2/4}$ if n is even and $c_o q^{n^2/4}$ if n is odd, where [17, 18]

$$c_e = \frac{\sum\limits_{k=-\infty}^{\infty} q^{-k^2}}{\prod\limits_{j=1}^{\infty}(1-q^{-j})}, \quad c_o = \frac{\sum\limits_{k=-\infty}^{\infty} q^{-\left(k+\frac{1}{2}\right)^2}}{\prod\limits_{j=1}^{\infty}(1-q^{-j})}.$$

We give a recurrence for the number χ_n of chains of proper subspaces (again, ordered by inclusion):

$$\chi_1 = 1, \quad \chi_n = 1 + \sum_{k=1}^{n-1} \binom{n}{k}_q \chi_k \quad \text{for } n \geq 2.$$

For the asymptotics, it follows that [6, 17]

$$\chi_n \sim \frac{1}{\zeta_q'(r)r}\left(\frac{1}{r}\right)^n \prod_{j=1}^{n}(q^j - 1) = \frac{A}{r^n}(q-1)(q^2-1)(q^3-1)\cdots(q^n-1),$$

where $\zeta_q(x)$ is the zeta function for the poset of subspaces:

$$\zeta_q(x) = \sum_{k=1}^{\infty} \frac{x^k}{(q-1)(q^2-1)(q^3-1)\cdots(q^k-1)}$$

and $r > 0$ is the unique solution of the equation $\zeta_q(r) = 1$. In particular, when $q = 2$, we have $c_e = 7.3719688014\ldots$, $c_o = 7.3719494907\ldots$, and

$$\chi_n \sim \frac{A}{r^n} \cdot Q \cdot 2^{\frac{n(n+1)}{2}},$$

where $r = 0.7759021363\ldots$, $A = 0.8008134543\ldots$, and

$$Q = \prod_{k=1}^{\infty}\left(1 - \frac{1}{2^k}\right) = 0.2887880950\ldots$$

is one of the digital search tree constants [5.14]. If we further insist that the chains are maximal, then the number of such chains is $[n!]_q$.

5.7.3 Chains in the Partition Lattice of S

We have discussed chains in the poset of subsets of the set S. There is, however, another poset associated naturally with S that is less familiar and more difficult to study: the **poset of partitions** of S. Here is the partial ordering: Assuming P and Q are two partitions of S, then $P < Q$ if $P \neq Q$ and if $p \in P$ implies that p is a subset of q for some $q \in Q$. In other words, P is a *refinement* of Q in the sense that each of its blocks fits within a block of Q. For arbitrary n, the poset is, in fact, a lattice with minimum element $m = \{\{1\}, \{2\}, \ldots, \{n\}\}$ and maximum element $M = \{\{1, 2, \ldots, n\}\}$.

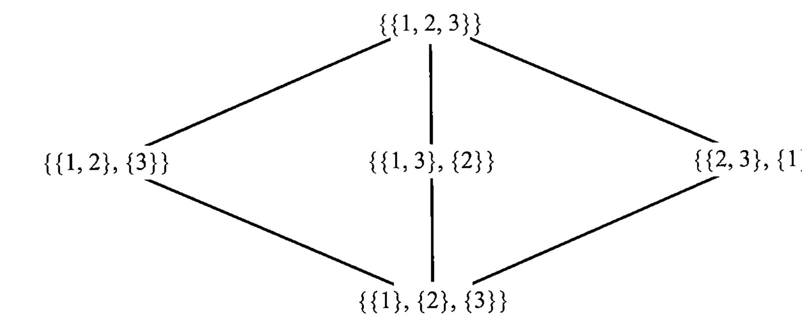


Figure 5.10. The number of chains $m < P_1 < M$ in the partition lattice of the set $\{1, 2, 3\}$ is three.

What is the number of chains $m = P_0 < P_1 < P_2 < \cdots < P_{k-1} < P_k = M$ of length k in the partition lattice of S? In the case $n = 3$, there is only one chain for $k = 1$, specifically, $m < M$. For $k = 2$, there are three such chains as pictured in Figure 5.10.

Let Z_n denote the number of all chains from m to M of any length; clearly $Z_1 = Z_2 = 1$ and, by the foregoing, $Z_3 = 4$. We have the recurrence

$$Z_n = \sum_{k=1}^{n-1} S_{n,k} Z_k$$

and exponential generating function

$$Z(x) = \sum_{n=1}^{\infty} \frac{Z_n}{n!} x^n, \quad 2Z(x) = x + Z(e^x - 1),$$

but techniques of Doubilet, Rota & Stanley and Bender do not apply here to give asymptotic estimates of Z_n. The partition lattice is the first natural lattice without the structure of a *binomial lattice*, which implies that well-known generating function techniques are no longer helpful.

Lengyel [19] formulated a different approach to prove that the quotient

$$r_n = \frac{Z_n}{(n!)^2(2\ln(2))^{-n} n^{-1-\ln(2)/3}}$$

must be bounded between two positive constants as $n \to \infty$. He presented numerical evidence suggesting that r_n tends to a unique value. Babai & Lengyel [20] then proved a fairly general convergence criterion that enabled them to conclude that $\Lambda = \lim_{n\to\infty} r_n$ exists and $\Lambda = 1.09\ldots$. The analysis in [19] involves intricate estimates of the Stirling numbers; in [20], the focus is on nearly convex linear recurrences with finite retardation and active predecessors.

In an ambitious undertaking, Flajolet & Salvy [21] computed $\Lambda = 1.0986858055\ldots$. Their approach is based on (complex fractional) analytic iterates of $\exp(x) - 1$ and much more, but unfortunately their paper is presently incomplete. See [5.8] for related discussion of the Takeuchi-Prellberg constant.

By way of contrast, the number of *maximal* chains is given exactly by $n!(n-1)!/2^{n-1}$ and Lengyel [19] observed that Z_n exceeds this by an exponentially large factor.

5.7.4 Random Chains

Van Cutsem & Ycart [22] examined random chains in both the subset and partition lattices. It is remarkable that a common framework exists for studying these and that, in a certain sense, the limiting distributions of both types of chains are *identical*. We mention only one consequence: If $\kappa_n = k/n$ is the normalized length of the random chain, then

$$\lim_{n\to\infty} \mathrm{E}(\kappa_n) = \frac{1}{2\ln(2)} = 0.7213475204\ldots$$

and a corresponding Central Limit Theorem also holds.

[1] L. Lovász, *Combinatorial Problems and Exercises*, 2nd ed., North-Holland, 1993, pp. 16–18, 162–173; MR 94m:05001.

[2] L. Comtet, *Advanced Combinatorics: The Art of Finite and Infinite Expansions*, Reidel, 1974, pp. 59–60, 204–211; MR 57 #124.

[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000110, A000670, A005121, and A006116.

[4] G.-C. Rota, The number of partitions of a set, *Amer. Math. Monthly* 71 (1964) 498–504; MR 28 #5009.

[5] N. G. de Bruijn, *Asymptotic Methods in Analysis*, Dover, 1958, pp. 102–109; MR 83m:41028.

[6] E. A. Bender, Asymptotic methods in enumeration, *SIAM Rev.* 16 (1974) 485–515; errata 18 (1976) 292; MR 51 #12545 and MR 55 #10276.

[7] A. M. Odlyzko, Asymptotic enumeration methods, *Handbook of Combinatorics*, v. I, ed. R. Graham, M. Grötschel, and L. Lovász, MIT Press, 1995, pp. 1063–1229; MR 97b:05012.

[8] B. Salvy and J. Shackell, Symbolic asymptotics: multiseries of inverse functions, *J. Symbolic Comput.* 27 (1999) 543–563; MR 2000h:41039.

[9] B. Salvy and J. Shackell, Asymptotics of the Stirling numbers of the second kind, *Studies in Automatic Combinatorics*, v. 2, Algorithms Project, INRIA, 1997.

[10] H. S. Wilf, *generatingfunctionology*, Academic Press, 1990, pp. 21–24, 146–147; MR 95a:05002.

[11] P. Doubilet, G.-C. Rota, and R. Stanley, On the foundations of combinatorial theory. VI: The idea of generating function, *Proc. Sixth Berkeley Symp. Math. Stat. Probab.*, v. 2, ed. L. M. Le Cam, J. Neyman, and E. L. Scott, Univ. of Calif. Press, 1972, pp. 267–318; MR 53 #7796.

[12] M. Aigner, *Combinatorial Theory*, Springer-Verlag, 1979, pp. 78–79, 142–143; MR 80h:05002.

[13] J. Goldman and G.-C. Rota, The number of subspaces of a vector space, *Recent Progress in Combinatorics*, Proc. 1968 Waterloo conf., ed. W. T. Tutte, Academic Press, 1969, pp. 75–83; MR 40 #5453.

[14] G. E. Andrews, *The Theory of Partitions*, Addison-Wesley, 1976; MR 99c:11126.

[15] H. Exton, *q-Hypergeometric Functions and Applications*, Ellis Horwood, 1983; MR 85g:33001.

[16] M. Sved, Gaussians and binomials, *Ars Combin.* 17 *A* (1984) 325–351; MR 85j:05002.

[17] T. Slivnik, Subspaces of $\mathbb{Z}_2^n$, unpublished note (1996).

[18] M. Wild, The asymptotic number of inequivalent binary codes and nonisomorphic binary matroids, *Finite Fields Appl.* 6 (2000) 192–202; MR 2001i:94077.

[19] T. Lengyel, On a recurrence involving Stirling numbers, *Europ. J. Combin.* 5 (1984) 313–321; MR 86c:11010.

[20] L. Babai and T. Lengyel, A convergence criterion for recurrent sequences with application to the partition lattice, *Analysis* 12 (1992) 109–119; MR 93f:05005.

[21] P. Flajolet and B. Salvy, Hierarchical set partitions and analytic iterates of the exponential function, unpublished note (1990).

[22] B. Van Cutsem and B. Ycart, Renewal-type behavior of absorption times in Markov chains, *Adv. Appl. Probab.* 26 (1994) 988–1005; MR 96f:60118.

5.8 Takeuchi–Prellberg Constant

In 1978, Takeuchi defined a triply recursive function [1,2]

$$t(x, y, z) = \begin{cases} y & \text{if } x \le y, \\ t(t(x-1, y, z), t(y-1, z, x), t(z-1, x, y)) & \text{otherwise} \end{cases}$$

that is useful for benchmark testing of programming languages. The value of $t(x, y, z)$ is of no practical significance; in fact, McCarthy [1,2] observed that the function can be described more simply as

$$t(x, y, z) = \begin{cases} y & \text{if } x \le y, \\ \begin{cases} z & \text{if } y \le z, \\ x & \text{otherwise,} \end{cases} & \text{otherwise.} \end{cases}$$

The interesting quantity is not $t(x, y, z)$, but rather $T(x, y, z)$, defined to be the number of times the *otherwise* clause is invoked in the recursion. We assume that the program is memoryless in the sense that previously computed results are not available at any time in the future. Knuth [1,3] studied the **Takeuchi numbers** $T_n = T(n, 0, n+1)$:

$$T_0 = 0,\ T_1 = 1,\ T_2 = 4,\ T_3 = 14,\ T_4 = 53,\ T_5 = 223, \ldots$$

and deduced that

$$e^{n\ln(n) - n\ln(\ln(n)) - n} < T_n < e^{n\ln(n) - n + \ln(n)}$$

for all sufficiently large n. He asked for more precise asymptotic information about the growth of T_n.

Starting with Knuth's recursive formula for the Takeuchi numbers

$$T_{n+1} = \sum_{k=0}^{n} \left[\binom{n+k}{n} - \binom{n+k}{n+1} \right] T_{n-k} + \sum_{k=1}^{n-1} \binom{2k}{k} \frac{1}{k+1}$$

and the somewhat related Bell numbers [5.7]

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k},\quad B_0 = 1,\ B_1 = 1,\ B_2 = 2,\ B_3 = 5,\ B_4 = 15,\ B_5 = 52, \ldots,$$

Prellberg [4] observed that the following limit exists:

$$c = \lim_{n\to\infty} \frac{T_n}{B_n \exp\left(\frac{1}{2} W_n^2\right)} = 2.2394331040\ldots,$$

where $W_n \exp(W_n) = n$ are special values of the Lambert W function [6.11].

Since both the Bell numbers and the W function are well understood, this provides an answer to Knuth's question. The underlying theory is still under development, but

Prellberg's numerical evidence is persuasive. Recent theoretical work [5] relates the constant c to an associated functional equation,

$$T(z) = \sum_{n=0}^{\infty} T_n z^n, \quad T(z) = \frac{T(z - z^2)}{z} - \frac{1}{(1-z)(1-z+z^2)},$$

in a manner parallel to how Lengyel's constant [5.7] is obtained.

[1] D. E. Knuth, Textbook examples of recursion, *Artificial Intelligence and Mathematical Theory of Computation*, ed. V. Lifschitz, Academic Press, 1991, pp. 207–229; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 391–414; MR 93a:68093.

[2] I. Vardi, *Computational Recreations in Mathematica*, Addison-Wesley, 1991, pp. 179–199; MR 93e:00002.

[3] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000651.

[4] T. Prellberg, On the asymptotics of Takeuchi numbers, *Symbolic Computation, Number Theory, Special Functions, Physics and Combinatorics*, Proc. 1999 Gainesville conf., ed. F. G. Garvan and M. Ismail, Kluwer, 2001, pp. 231–242; math.CO/0005008.

[5] T. Prellberg, On the asymptotic analysis of a class of linear recurrences, presentation at *Formal Power Series and Algebraic Combinatorics (FPSAC)* conf., Univ. of Melbourne, 2002.

5.9 Pólya's Random Walk Constants

Let L denote the d-dimensional cubic lattice whose vertices are precisely all integer points in d-dimensional space. A **walk** ω on L, beginning at the origin, is an infinite sequence of vertices $\omega_0, \omega_1, \omega_2, \omega_3, \ldots$ with $\omega_0 = 0$ and $|\omega_{j+1} - \omega_j| = 1$ for all j. Assume that the walk is random and symmetric in the sense that, at each time step, all $2d$ directions of possible travel have equal probability. What is the likelihood that $\omega_n = 0$ for some $n > 0$? That is, what is the **return probability** p_d?

Pólya [1–4] proved the remarkable fact that $p_1 = p_2 = 1$ but $p_d < 1$ for $d > 2$. McCrea & Whipple [5], Watson [6], Domb [7] and Glasser & Zucker [8] each contributed facets of the following evaluations of $p_3 = 1 - 1/m_3 = 0.3405373295\ldots$, where the expected number m_3 of returns to the origin, plus one, is

$$\begin{aligned} m_3 &= \frac{3}{(2\pi)^3} \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \frac{1}{3 - \cos(\theta) - \cos(\varphi) - \cos(\psi)}\, d\theta\, d\varphi\, d\psi \\ &= \frac{12}{\pi^2}\left(18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6}\right) K\left[(2-\sqrt{3})(\sqrt{3}-\sqrt{2})\right]^2 \\ &= 3\left(18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6}\right)\left[1 + 2\sum_{k=1}^{\infty} \exp(-\sqrt{6}\pi k^2)\right]^4 \\ &= \frac{\sqrt{6}}{32\pi^3}\Gamma\left(\frac{1}{24}\right)\Gamma\left(\frac{5}{24}\right)\Gamma\left(\frac{7}{24}\right)\Gamma\left(\frac{11}{24}\right) = 1.5163860591\ldots. \end{aligned}$$

Hence the **escape probability** for a random walk on the three-dimensional cubic lattice is $1 - p_3 = 0.6594626704\ldots$. In these expressions, K denotes the complete elliptic integral of the first kind [1.4.6] and Γ denotes the gamma function [1.5.4]. Return and escape probabilities can also be computed for the body-centered or face-centered cubic

Table 5.1. *Expected Number of Returns and Return Probabilities*

d	m_d	p_d
4	1.2394671218...	0.1932016732...
5	1.1563081248...	0.1351786098...
6	1.1169633732...	0.1047154956...
7	1.0939063155...	0.0858449341...
8	1.0786470120...	0.0729126499...

lattices (as opposed to the simple cubic lattice), but we will not discuss these or other generalizations [9].

What can be said about p_d for $d > 3$? Closed-form expressions do not appear to exist here. Montroll [10–12] determined that $p_d = 1 - 1/m_d$, where

$$m_d = \frac{d}{(2\pi)^d}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\left(d - \sum_{k=1}^{d}\cos(\theta_k)\right)^{-1} d\theta_1\, d\theta_2 \cdots d\theta_d = \int_0^{\infty} e^{-t}\left(I_0\left(\frac{t}{d}\right)\right)^d dt$$

and $I_0(x)$ denotes the zeroth modified Bessel function [3.6]. The corresponding numerical approximations, as functions of d, are listed in Table 5.1 [10, 13–17].

What is the length of travel required for a return? Let $U_{d,l,n}$ be the number of d-dimensional n-step walks that start from the origin and end at a lattice point l. Let $V_{d,l,n}$ be the number of d-dimensional n-step walks that start from the origin and reach the lattice point $l \neq 0$ for the *first time* at the end (second time if $l = 0$). Then the generating functions

$$U_{d,l}(x) = \sum_{n=0}^{\infty}\frac{U_{d,l,n}}{(2d)^n}x^n, \quad V_{d,l}(x) = \sum_{n=0}^{\infty}\frac{V_{d,l,n}}{(2d)^n}x^n$$

satisfy $V_{d,l}(x) = U_{d,l}(x)/U_{d,0}(x)$ if $l \neq 0$, $V_{d,l}(x) = 1 - 1/U_{d,0}(x)$ if $l = 0$, and $U_{d,0}(1) = m_d$, $V_{d,0}(1) = p_d$. For example,

$$U_{1,l}(x) = \sum_{n=0}^{\infty}\frac{1}{2^n}\binom{n}{\frac{l+n}{2}}x^n, \quad U_{2,l}(x) = \sum_{n=0}^{\infty}\frac{1}{4^n}\binom{n}{\frac{l_1+l_2+n}{2}}\binom{n}{\frac{l_1-l_2+n}{2}}x^n,$$

where we agree to set the binomial coefficients equal to 0 if $l + n$ is odd for $d = 1$ or $l_1 + l_2 + n$ is odd for $d = 2$. If $d = 3$, then $a_n = U_{3,0,2n}$ satisfies [18]

$$a_n = \binom{2n}{n}\sum_{k=0}^{n}\binom{n}{k}^2\binom{2k}{k} = \sum_{k=0}^{n}\frac{(2n)!(2k)!}{(n-k)!^2 k!^4}, \quad \sum_{n=0}^{\infty}\frac{a_n}{(2n)!}y^{2n} = I_0(2y)^3,$$

$$(n+2)^3 a_{n+2} - 2(2n+3)(10n^2 + 30n + 23)a_{n+1} + 36(n+1)(2n+1)(2n+3)a_n = 0,$$

and if $d = 4$, then $b_n = U_{4,0,2n}$ satisfies [19]

$$\begin{aligned}(n+2)^4 b_{n+2} &- 4(2n+3)^2(5n^2 + 15n + 12)b_{n+1} \\ &+ 256(n+1)^2(2n+1)(2n+3)b_n = 0.\end{aligned}$$

For any d, the mean **first-passage time** to arrive at any lattice point l is infinite (in spite of the fact that the associated probability $V_{d,l}(1) = 1$ for $d = 1$ or 2). There are several alternative ways of quantifying the length of required travel. Using our formulas for $V_{d,l}(x)$, the median first-passage times are 2-4, 1-3, 6-8, and 17-19 steps for $l = 0$, 1, 2, and 3 when $d = 1$, and 2-4, 25-27, and 520-522 steps for $l = (0, 0), (1, 0)$, and $(1, 1)$ when $d = 2$. Hughes [3,20] examined the conditional mean time to return to the origin (conditional upon return eventually occurring). Also, for $d = 1$, the mean time for the earliest of three independent random walkers to return to the origin is finite and has value [6,21–23]

$$2\sum_{n=0}^{\infty}\frac{1}{2^{6n}}\binom{2n}{n}^3 = \frac{2}{\pi^3}\int_0^\pi\int_0^\pi\int_0^\pi \frac{1}{1-\cos(\theta)\cos(\varphi)\cos(\psi)}\,d\theta\,d\varphi\,d\psi$$
$$= \frac{8}{\pi^2}K\left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2\pi^3}\Gamma\left(\frac{1}{4}\right)^4 = 2(1.3932039296\ldots),$$

whereas for $d = 2$, the mean time for the earliest of an *arbitrary* number of independent random walkers is infinite. More on multiple random walkers, of both the friendly and vicious kinds, is found in [24].

It is known that

$$U_{d,l}(x) = \frac{d}{(2\pi)^d}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\cdots\int_{-\pi}^{\pi}\left(d - x\sum_{k=1}^{d}\cos(\theta_k)\right)^{-1}$$
$$\times \exp\left(i\sum_{k=1}^{d}\theta_k l_k\right)d\theta_1\,d\theta_2\cdots d\theta_d,$$

which can be numerically evaluated for small d. Here are some sample probabilities [11,16] that a three-dimensional walk reaches a point l:

$$V_{3,l}(1) = \frac{U_{3,l}(1)}{m_3} = \begin{cases} 0.3405373295\ldots & \text{if } l = (1,0,0),\\ 0.2183801414\ldots & \text{if } l = (1,1,0),\\ 0.1724297877\ldots & \text{if } l = (1,1,1).\end{cases}$$

An asymptotic expansion for these probabilities is [11,12]

$$V_{3,l}(1) = \frac{3}{2\pi m_3|l|}\left[1 + \frac{1}{8|l|^2}\left(-3 + \frac{5(l_1^4 + l_2^4 + l_3^4)}{|l|^2} + \cdots\right)\right] \sim \frac{0.3148702313\ldots}{|l|}$$

and is valid as $|l|^2 = l_1^2 + l_2^2 + l_3^2 \to \infty$.

Let $W_{d,n}$ be the average number of distinct vertices visited during a d-dimensional n-step walk. It can be shown that [25–28]

$$W_d(x) = \sum_{n=0}^{\infty}W_{d,n}x^n = \frac{1}{(1-x)^2U_{d,0}(x)}, \quad W_{d,n} \sim \begin{cases}\sqrt{\dfrac{8n}{\pi}} & \text{if } d = 1,\\ \dfrac{\pi n}{\ln(n)} & \text{if } d = 2,\\ (1-p_3)n & \text{if } d = 3\end{cases}$$

as $n \to \infty$. Higher-order asymptotics for $W_{3,n}$ are possible using the expansion [11,12,29–31]

$$U_{3,0}(x) = m_3 - \frac{3\sqrt{3}}{2\pi}(1-x^2)^{\frac{1}{2}} + c(1-x^2) - \frac{3\sqrt{3}}{4\pi}(1-x^2)^{\frac{3}{2}} + \cdots,$$

where $x \to 1^-$ and

$$c = \frac{9}{32}\left(m_3 + \frac{6}{\pi^2 m_3}\right) = 0.5392381750\ldots.$$

Other parameters, for example, the average growth of distance from the origin [32],

$$\lim_{n\to\infty}\frac{1}{\ln(n)}\sum_{j=1}^{n}\frac{j^{-1/2}}{1+|\omega_j|} = \lambda_1 \text{ with probability 1, if } d = 1,$$

$$\lim_{n\to\infty}\frac{1}{\ln(n)^2}\sum_{j=1}^{n}\frac{1}{1+|\omega_j|^2} = \lambda_2 \text{ with probability 1, if } d = 2,$$

$$\lim_{n\to\infty}\frac{1}{\ln(n)}\sum_{j=1}^{n}\frac{1}{1+|\omega_j|^2} = \lambda_d \text{ with probability 1, if } d \geq 3,$$

are more difficult to analyze. The constants λ_d are known only to be finite and positive.

For a one-dimensional n-step walk ω, define M_n^+ to be the maximum value of ω_j and M_n^- to be the maximum value of $-\omega_j$. Then M_n^+ and M_n^- each follow the half-normal distribution [6.2] in the limit as $n \to \infty$, and [33,34]

$$\lim_{n\to\infty}\mathrm{E}\left(n^{-\frac{1}{2}}M_n^+\right) = \sqrt{\frac{2}{\pi}} = \lim_{n\to\infty}\mathrm{E}\left(n^{-\frac{1}{2}}M_n^-\right).$$

Further, if T_n^+ is the smallest value of j for which $\omega_j = M_n^+$ and T_n^- is the smallest value of k for which $-\omega_k = M_n^-$, then the **arcsine law** applies:

$$\lim_{n\to\infty}\mathrm{P}\left(n^{-1}T_n^+ < x\right) = \frac{2}{\pi}\arcsin\sqrt{x} = \lim_{n\to\infty}\mathrm{P}\left(n^{-1}T_n^- < x\right),$$

which implies that a one-dimensional random walk tends to be either highly negative or highly positive (not both). Such detailed information about d-dimensional walks is not yet available. Define also $\tau_{d,r}$ to be the smallest value of j for which $|\omega_j| \geq r$, for any positive integer r. Then [35]

$$\tau_{1,r} = r^2, \quad \tau_{2,2} = \tfrac{9}{2}, \quad \tau_{2,3} = \tfrac{135}{13}, \quad \tau_{2,4} = \tfrac{11791}{668},$$

but a pattern is not evident. What precisely can be said about $\tau_{d,r}$ as $r \to \infty$?

As a computational aside, we mention a result of Odlyzko's [36–38]: Any algorithm that determines M_n^+ (or M_n^-) exactly must examine at least $(A+o(1))\sqrt{n}$ of the ω_j values on average, where $A=\sqrt{8/\pi}\ln(2)=1.1061028674\ldots$.

On the one hand, the waiting time N_n for a one-dimensional random walk to hit a new vertex, not visited in the first n steps, satisfies [39]

$$\limsup_{n\to\infty}\frac{N_n}{n\ln(\ln(n))^2}=\frac{1}{\pi^2}\quad\text{with probability 1.}$$

On the other hand, if F_n denotes the set of vertices that are maximally visited by the random walk up to step n, called **favorite sites**, then $|F_n|\geq 4$ only finitely often, with probability 1 [40].

For two-dimensional random walks, we may define F_n analogously. The number of visits to a selected point in F_n within the first n steps is $\sim\ln(n)^2/\pi$ with probability 1, as $n\to\infty$. This can be rephrased as the asymptotic number of times a drunkard drops by his favorite watering hole [41,42]. Dually, the length of time C_r required to totally cover all vertices of the $r\times r$ torus (square with opposite sides identified) satisfies [43]

$$\lim_{r\to\infty}\mathrm{P}\left(\left|\frac{C_r}{r^2\ln(r)^2}-\frac{4}{\pi}\right|<\varepsilon\right)=1$$

for every $\varepsilon>0$ (convergence in probability). This solves what is known as the "white screen problem" [44].

If a three-dimensional random walk ω is restricted to the region $x\geq y\geq z$, then the analogous series coefficients are

$$\tilde{a}_n=\sum_{k=0}^{n}\frac{(2n)!(2k)!}{(n-k)!(n+1-k)!k!^2(k+1)!^2},$$

and from this we have [45]

$$\tilde{m}_3=\sum_{n=0}^{\infty}\frac{\tilde{a}_n}{6^{2n}}=1.0693411205\ldots,\quad \tilde{p}_3=1-\frac{1}{\tilde{m}_3}=0.0648447153\ldots$$

characterizing the return. What can be said concerning other regions, for example, a half-space, quarter-space, or octant?

Here is one variation. Let $X_1, X_2, X_3, \ldots$ be independent normally distributed random variables with mean μ and variance 1. Consider the partial sums $S_j=\sum_{k=1}^{j}X_k$, which constitute a random walk on the real line (rather than the one-dimensional lattice) with Gaussian increments (rather than Bernoulli increments). There is an enormous literature on $\{S_j\}$, but we shall mention only one result. Let H be the first positive value of S_j, called the **first ladder height** of the process; then the moments of H when $\mu=0$ are [46]

$$\mathrm{E}_0(H)=\frac{1}{\sqrt{2}},\quad \mathrm{E}_0(H^2)=-\frac{\zeta(\frac{1}{2})}{\sqrt{\pi}}=\sqrt{2}\rho=\sqrt{2}(0.5825971579\ldots)$$

and, for arbitrary μ in a neighborhood of 0,

$$\mathrm{E}_\mu(H)=\frac{1}{\sqrt{2}}\exp\left[-\frac{\mu}{\sqrt{2\pi}}\sum_{k=0}^{\infty}\frac{\zeta(\frac{1}{2}-k)}{k!(2k+1)}\left(-\frac{\mu^2}{2}\right)^k\right],$$

where $\zeta(x)$ is the Riemann zeta function [1.6]. Other occurrences of the interesting constant ρ in the statistical literature are in [47–50].

Here is another variation. Let $Y_1, Y_2, Y_3, \ldots$ be independent Uniform $[-1,1]$ random variables, $S_0=0$, and $S_j=\sum_{k=1}^{j}Y_k$. Then the expected maximum value of $\{S_0, S_1, \ldots, S_n\}$ is [51]

$$\mathrm{E}\left(\max_{0\leq j\leq n}S_j\right)=\sqrt{\frac{2}{3\pi}}n^{\frac{1}{2}}+\sigma+\frac{1}{5}\sqrt{\frac{2}{3\pi}}n^{-\frac{1}{2}}+O\left(n^{-\frac{3}{2}}\right)$$

as $n\to\infty$, where $\sigma=-0.2979521902\ldots$ is given by

$$\sigma=\frac{\zeta(\frac{1}{2})}{\sqrt{6\pi}}+\frac{\zeta(\frac{3}{2})}{20\sqrt{6\pi}}+\sum_{k=1}^{\infty}\left(\frac{t_k}{k}-\frac{k^{-\frac{1}{2}}}{\sqrt{6\pi}}-\frac{k^{-\frac{3}{2}}}{20\sqrt{6\pi}}\right)$$

and

$$t_k=\frac{2(-1)^k}{(k+1)!}\sum_{k/2\leq j\leq k}(-1)^j\binom{k}{j}\left(j-\frac{k}{2}\right)^{k+1}.$$

A deeper connection between $\zeta(x)$ and random walks is discussed in [52].

5.9.1 Intersections and Trappings

A walk ω on the lattice L is **self-intersecting** if $\omega_i=\omega_j$ for some $i<j$, and the **self-intersection time** is the smallest value of j for which this happens. Computing self-intersection times is more difficult than first-passage times since the entire history of the walk requires memorization. If $d=1$, then clearly the mean self-intersection time is 3. If $d=2$, the mean self-intersection time is [53]

$$\begin{aligned}\frac{2\cdot 4}{4^2}+\frac{3\cdot 12}{4^3}+\frac{4\cdot 44}{4^4}+\frac{5\cdot 116}{4^5}+\cdots&=\sum_{n=2}^{\infty}\frac{n(4c_{n-1}-c_n)}{4^n}\\&=\frac{c_1}{2}+\sum_{n=2}^{\infty}\frac{c_n}{4^n}=4.5860790989\ldots,\end{aligned}$$

where the sequence $\{c_n\}$ is defined in [5.10]. When n is large, no exact formula for evaluating c_n is known, unlike the sequences $\{a_n\}$, $\{\tilde{a}_n\}$, and $\{b_n\}$ discussed earlier. We are, in this example, providing foreshadowing of difficulties to come later. See the generalization in [54, 55].

A walk ω is **self-trapping** if, for some k, $\omega_i\neq\omega_j$ for all $i<j\leq k$ and ω_k is completely surrounded by previously visited vertices. If $d=2$, there are eight self-trapping walks when $k=7$ and sixteen such walks when $k=8$. A Monte Carlo simulation in [56, 57] gave a mean self-trapping time of approximately $70.7\ldots$.

Two walks ω and ω' **intersect** if $\omega_i = \omega'_j$ for some nonzero i and j. The probability q_n that two n-step independent random walks never intersect satisfies [58–61]

$$\ln(q_n) \sim \begin{cases} -\frac{5}{8}\ln(n) & \text{if } d = 2, \\ -\xi \ln(n) & \text{if } d = 3, \\ -\frac{1}{2}\ln(\ln(n)) & \text{if } d = 4 \end{cases}$$

as $n \to \infty$, where the exponent ξ is approximately $0.29\ldots$ (again obtained by simulation). For each $d \geq 5$, it can be shown [62] that $\lim_{n\to\infty} q_n$ lies strictly between 0 and 1. Further simulation [63] yields $q_5 = 0.708\ldots$ and $q_6 = 0.822\ldots$, and we shall refer to these in [5.10].

5.9.2 Holonomicity

A **holonomic function** (in the sense of Zeilberger [45,64,65]) is a solution $f(z)$ of a linear homogeneous differential equation

$$f^{(n)}(z) + r_1(z)f^{(n-1)}(z) + \cdots + r_{n-1}(z)f'(z) + r_n(z)f(z) = 0,$$

where each $r_k(z)$ is a rational function with rational coefficients. **Regular holonomic constants** are values of f at algebraic points z_0 where each r_k is analytic; f can be proved to be analytic at z_0 as well. **Singular holonomic constants** are values of f at algebraic points z_0 where each r_k has, at worst, a pole of order k at z_0 (called Fuchsian or "regular" singularities [66–68]). The former include π, $\ln(2)$, and the tetralogarithm $\mathrm{Li}_4(1/2)$; the latter include Apéry's constant $\zeta(3)$, Catalan's constant G, and Pólya's constants p_d, $d > 2$. Holonomic constants of either type fall into the class of polynomial-time computable constants [69]. We merely mention a somewhat related theory of EL numbers due to Chow [70].

[1] G. Pólya, Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Stassennetz, *Math. Annalen* 84 (1921) 149–160; also in *Collected Papers*, v. 4, ed. G.-C. Rota, MIT Press, 1984, pp. 69–80, 609.

[2] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks*, Math. Assoc. Amer., 1984; math.PR/0001057; MR 89a:94023.

[3] B. D. Hughes, *Random Walks and Random Environments*, v. 1, Oxford Univ. Press, 1995; MR 96i:60070.

[4] F. Spitzer, *Principles of Random Walks*, 2nd ed, Springer-Verlag, 1976, p. 103; MR 52 #9383.

[5] W. H. McCrea and F. J. W. Whipple, Random paths in two and three dimensions, *Proc. Royal Soc. Edinburgh* 60 (1940) 281–298; MR 2,107f.

[6] G. N. Watson, Three triple integrals, *Quart. J. Math.* 10 (1939) 266–276; MR 1,205b.

[7] C. Domb, On multiple returns in the random-walk problem, *Proc. Cambridge Philos. Soc.* 50 (1954) 586–591; MR 16,148f.

[8] M. L. Glasser and I. J. Zucker, Extended Watson integrals for the cubic lattices, *Proc. Nat. Acad. Sci. USA* 74 (1977) 1800–1801; MR 56 #686.

[9] D. J. Daley, Return probabilities for certain three-dimensional random walks, *J. Appl. Probab.* 16 (1979) 45–53; MR 80e:60083.

[10] E. W. Montroll, Random walks in multidimensional spaces, especially on periodic lattices, *J. SIAM* 4 (1956) 241–260; MR 19,470d.

[11] E. W. Montroll, Random walks on lattices, *Stochastic Processes in Mathematical Physics and Engineering*, ed. R. Bellman, Proc. Symp. Appl. Math. 16, Amer. Math. Soc., 1964, pp. 193–220; MR 28 #4585.

[12] E. W. Montroll and G. H. Weiss, Random walks on lattices. II, *J. Math. Phys.* 6 (1965) 167–181; MR 30 #2563.

[13] K. Kondo and T. Hara, Critical exponent of susceptibility for a class of general ferromagnets in $d > 4$ dimensions, *J. Math. Phys.* 28 (1987) 1206–1208; MR 88k:82052.

[14] D. I. Cartwright, Some examples of random walks on free products of discrete groups, *Annali Mat. Pura Appl.* 151 (1988) 1–15; MR 90f:60018.

[15] P. Griffin, Accelerating beyond the third dimension: Returning to the origin in simple random walk, *Math. Sci.* 15 (1990) 24–35; MR 91g:60083.

[16] T. Hara, G. Slade, and A. D. Sokal, New lower bounds on the self-avoiding-walk connective constant, *J. Stat. Phys.* 72 (1993) 479–517; erratum 78 (1995) 1187–1188; MR 94e:82053.

[17] J. Keane, Pólya's constants p_d to 80 digits for $3 \leq d \leq 64$, unpublished note (1998).

[18] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002896, A039699, A049037, and A063888.

[19] M. L. Glasser and A. J. Guttmann, Lattice Green function (at 0) for the 4D hypercubic lattice, *J. Phys. A* 27 (1994) 7011–7014; MR 95i:82047.

[20] B. D. Hughes, On returns to the starting site in lattice random walks, *Physica* 134A (1986) 443–457; MR 87c:60057.

[21] K. Lindenberg, V. Seshadri, K. E. Shuler, and G. H. Weiss, Lattice random walks for sets of random walkers: First passage times, *J. Stat. Phys.* 23 (1980) 11–25; MR 83c:60100.

[22] W. D. Fryer and M. S. Klamkin, Comment on problem 612, *Math. Mag.* 40 (1967) 52–53.

[23] B. C. Berndt, *Ramanujan's Notebooks: Part II*, Springer-Verlag, 1989, p. 24; MR 90b:01039.

[24] M. E. Fisher, Walks, walls, wetting, and melting, *J. Stat. Phys.* 34 (1984) 667–729; MR 85j:82022.

[25] A. Dvoretzky and P. Erdös, Some problems on random walk in space, *Proc. Second Berkeley Symp. Math. Stat. Probab.*, ed. J. Neyman, Univ. of Calif. Press, 1951, pp. 353–367; MR 13,852b.

[26] G. H. Vineyard, The number of distinct sites visited in a random walk on a lattice, *J. Math. Phys.* 4 (1963) 1191–1193; MR 27 #4610.

[27] M. N. Barber and B. W. Ninham, *Random and Restricted Walks: Theory and Applications*, Gordon and Breach, 1970.

[28] E. W. Montroll and B. J. West, On an enriched collection of stochastic processes, *Studies in Statistical Mechanics*, v. 7, *Fluctuation Phenomena*, ed. E. W. Montroll and J. L. Lebowitz, North-Holland, 1979, pp. 61–175; MR 83g:82001.

[29] A. A. Maradudin, E. W. Montroll, G. H. Weiss, R. Herman, and H. W. Milnes, Green's functions for monatomic simple cubic lattices, *Acad. Royale Belgique Classe Sci. Mém. Coll. 4°, 2e Sér.*, v. 14 (1960) n. 7, 1–176; MR 22 #7440.

[30] G. S. Joyce, Lattice Green function for the simple cubic lattice, *J. Phys. A* (1972) L65–L68.

[31] G. S. Joyce, On the simple cubic lattice Green function, *Philos. Trans. Royal Soc. London Ser. A* 273 (1973) 583–610; MR 58 #28708.

[32] P. Erdös and S. J. Taylor, Some problems concerning the structure of random walk paths, *Acta Math. Acad. Sci. Hungar.* 11 (1960) 137–162; MR 22 #12599.

[33] P. Révész, *Random Walk in Random and Nonrandom Environments*, World Scientific, 1990; MR 92c:60096.

[34] W. Feller, *An Introduction to Probability Theory and Its Applications*, v. 1, 3rd ed., Wiley, 1968; MR 37 #3604.

[35] I. Kastanas, Simple random walk barriers, unpublished note (1997).

[36] A. M. Odlyzko, Search for the maximum of a random walk, *Random Structures Algorithms* 6 (1995) 275–295; MR 97b:60117.

[37] H.-K. Hwang, A constant arising from the analysis of algorithms for determining the maximum of a random walk, *Random Structures Algorithms* 10 (1997) 333–335; MR 98i:05007.

[38] P. Chassaing, How many probes are needed to compute the maximum of a random walk?, *Stochastic Process. Appl.* 81 (1999) 129–153; MR 2000k:60087.
[39] E. Csáki, A note on: "Three problems on the random walk in Z^d" by P. Erdös and P. Révész, *Studia Sci. Math. Hungar.* 26 (1991) 201–205; MR 93k:60172.
[40] B. Tóth, No more than three favorite sites for simple random walk, *Annals of Probab.* 29 (2001) 484–503; MR 2002c:60076.
[41] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni, Thick points for planar Brownian motion and the Erdős-Taylor conjecture on random walk, *Acta Math.* 186 (2001) 239–270; math.PR/0105107.
[42] I. Stewart, Where drunkards hang out, *Nature* 413 (2001) 686–687.
[43] A. Dembo, Y. Peres, J. Rosen, and O. Zeitouni, Cover times for Brownian motion and random walks in two dimensions, math.PR/0107191.
[44] H. S. Wilf, The white screen problem, *Amer. Math. Monthly* 96 (1989) 704–707.
[45] J. Wimp and D. Zeilberger, How likely is Pólya's drunkard to stay in $x > y > z$?, *J. Stat. Phys.* 57 (1989) 1129–1135; MR 90m:82083.
[46] J. T. Chang and Y. Peres, Ladder heights, Gaussian random walks and the Riemann zeta function, *Annals of Probab.* 25 (1997) 787–802; MR 98c:60086.
[47] H. Chernoff, Sequential tests for the mean of a normal distribution. IV. (Discrete case), *Annals of Math. Statist.* 36 (1965) 55–68; MR 30 #681.
[48] T. L. Lai, Asymptotic moments of random walks with applications to ladder variables and renewal theory, *Annals of Probab.* 4 (1976) 51–66; MR 52 #12086.
[49] D. Siegmund, Corrected diffusion approximations in certain random walk problems, *Adv. Appl. Probab.* 11 (1979) 701–719; MR 80i:60096.
[50] S. Asmussen, P. Glynn, and J. Pitman, Discretization error in simulation of one-dimensional reflecting Brownian motion, *Annals of Appl. Probab.* 5 (1995) 875–896; MR 97e:65156.
[51] E. G. Coffman, P. Flajolet, L. Flatto, and M. Hofri, The maximum of a random walk and its application to rectangle packing, *Probab. Engin. Inform. Sci.* 12 (1998) 373–386; MR 99f:60127.
[52] 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.
[53] A. J. Guttmann, Padé approximants and SAW generating functions, unpublished note (2001).
[54] D. J. Aldous, Self-intersections of random walks on discrete groups, *Math. Proc. Cambridge Philos. Soc.* 98 (1985) 155–177; MR 86j:60157.
[55] D. J. Aldous, Self-intersections of 1-dimensional random walks, *Probab. Theory Relat. Fields* 72 (1986) 559–587; MR 88a:60125.
[56] R. S. Lehman and G. H. Weiss, A study of the restricted random walk, *J. SIAM* 6 (1958) 257–278; MR 20 #4891.
[57] S. Hemmer and P. C. Hemmer, An average self-avoiding random walk on the square lattice lasts 71 steps, *J. Chem. Phys.* 81 (1984) 584–585; MR 85g:82099.
[58] K. Burdzy, G. F. Lawler, and T. Polaski, On the critical exponent for random walk intersections, *J. Stat. Phys.* 56 (1989) 1–12; MR 91h:60073.
[59] G. F. Lawler, *Intersections of Random Walks*, Birkhäuser, 1991; MR 92f:60122.
[60] G. Slade, Random walks, *Amer. Scientist*, v. 84 (1996) n. 2, 146–153.
[61] G. F. Lawler, O. Schramm, and W. Werner, Values of Brownian intersection exponents. II: Plane exponents, *Acta Math.* 187 (2001) 275–308; math.PR/0003156.
[62] G. F. Lawler, A self-avoiding random walk, *Duke Math. J.* 47 (1980) 655–693; MR 81j:60081.
[63] T. Prellberg, Intersection probabilities for high dimensional walks, unpublished note (2002).
[64] D. Zeilberger, A holonomic systems approach to special functions identities, *J. Comput. Appl. Math.* 32 (1990) 321–348; MR 92b:33014.
[65] H. S. Wilf and D. Zeilberger, Towards computerized proofs of identities, *Bull. Amer. Math. Soc.* 23 (1990) 77–83; MR 91a:33003.

[66] W. Wasow, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley, 1965, pp. 1–29; MR 34 #3401.
[67] E. L. Ince, *Ordinary Differential Equations*, Dover, 1956, pp. 356–365; MR 6,65f.
[68] G. F. Simmons, *Differential Equations with Applications and Historical Notes*, McGraw-Hill, 1972, pp. 153–174; MR 58 #17258.
[69] 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; MR 2001h:11161.
[70] T. Y. Chow, What is a closed-form number?, *Amer. Math. Monthly* 106 (1999) 440–448; math.NT/9805045; MR 2000e:11156.

5.10 Self-Avoiding Walk Constants

Let L denote the d-dimensional cubic lattice whose vertices are precisely all integer points in d-dimensional space. An n-step **self-avoiding walk** ω on L, beginning at the origin, is a sequence of vertices $\omega_0, \omega_1, \omega_2, \ldots, \omega_n$ with $\omega_0 = 0$, $|\omega_{j+1} - \omega_j| = 1$ for all j and $\omega_i \neq \omega_j$ for all $i \neq j$. The number of such walks is denoted by c_n. For example, $c_0 = 1, c_1 = 2d, c_2 = 2d(2d-1), c_3 = 2d(2d-1)^2$, and $c_4 = 2d(2d-1)^3 - 2d(2d-2)$. Self-avoiding walks are vastly more difficult to study than ordinary walks [1–6], and historically arose as a model for linear polymers in chemistry [7, 8]. No exact combinatorial enumerations are possible for large n. The methods for analysis hence include finite series expansions and Monte Carlo simulations.

For simplicity's sake, we have suppressed the dependence of c_n on d; we will do this for associated constants too whenever possible.

What can be said about the asymptotics of c_n? Since $c_{n+m} \leq c_n c_m$, on the basis of Fekete's submultiplicativity theorem [9–12], it is known that the **connective constant**

$$\mu_d = \lim_{n\to\infty} c_n^{\frac{1}{n}} = \inf_n c_n^{\frac{1}{n}}$$

exists and is nonzero. Early attempts to estimate $\mu = \mu_d$ included [13–15]; see [2] for a detailed survey. The current best rigorous lower and upper bounds for μ, plus the best-known estimate, are given in Table 5.2 [16–24]. The extent of our ignorance is fairly surprising: Although we know that $\mu^2 = \lim_{n\to\infty} c_{n+2}/c_n$ and $c_{n+1} \geq c_n$ for all n and all d, proving that $\mu = \lim_{n\to\infty} c_{n+1}/c_n$ for $2 \leq d \leq 4$ remains an open problem [25, 26].

Table 5.2. *Estimates for Connective Constant* μ

d	Lower Bound	Best Estimate for μ	Upper Bound
2	2.6200	2.6381585303	2.6792
3	4.5721	4.68404	4.7114
4	6.7429	6.77404	6.8040
5	8.8285	8.83854	8.8602
6	10.8740	10.87809	10.8886

It is believed that there exists a positive constant $\gamma = \gamma_d$ such that the following limit exists and is nonzero:

$$A = \begin{cases} \lim\limits_{n\to\infty} \dfrac{c_n}{\mu^n n^{\gamma-1}} & \text{if } d \neq 4, \\ \lim\limits_{n\to\infty} \dfrac{c_n}{\mu^n n^{\gamma-1}\ln(n)^{1/4}} & \text{if } d = 4. \end{cases}$$

The **critical exponent** γ is conjectured to be [27–29]

$$\gamma_2 = \tfrac{43}{32} = 1.34375, \quad \gamma_3 = 1.1575\ldots, \quad \gamma_4 = 1$$

and has been proved [1,30] to equal 1 for $d > 4$. For small d, we have bounds [1,25,31]

$$c_n \le \begin{cases} \mu^n \exp\left(C n^{1/2}\right) & \text{if } d = 2, \\ \mu^n \exp\left(C n^{2/(d+2)} \ln(n)\right) & \text{if } 3 \le d \le 4, \end{cases}$$

which do not come close to proving the existence of A. It is known [32] that, for $d = 5$, $1 \le A \le 1.493$ and, for sufficiently large d, $A = 1 + (2d)^{-1} + d^{-2} + O(d^{-3})$.

Another interesting object of study is the **mean square end-to-end distance**

$$r_n = \mathrm{E}\left(|\omega_n|^2\right) = \frac{1}{c_n}\sum_{\omega} |\omega_n|^2,$$

where the summation is over all n-step self-avoiding walks ω on L. Like c_n, it is believed that there is a positive constant $\nu = \nu_d$ such that the following limit exists and is nonzero:

$$B = \begin{cases} \lim\limits_{n\to\infty} \dfrac{r_n}{n^{2\nu}} & \text{if } d \neq 4, \\ \lim\limits_{n\to\infty} \dfrac{r_n}{n^{2\nu}\ln(n)^{1/4}} & \text{if } d = 4. \end{cases}$$

As before, it is conjectured that [27,33,34]

$$\nu_2 = \tfrac{3}{4} = 0.75, \quad \nu_3 = 0.5877\ldots, \quad \nu_4 = \tfrac{1}{2} = 0.5$$

and has been proved [1,30] that $\nu = 1/2$ for $d > 4$. This latter value is the same for Pólya walks, that is, the self-avoidance constraint has little effect in high dimensions. It is known [32] that, for $d = 5$, $1.098 \le B \le 1.803$ and, for sufficiently large d, $B = 1 + d^{-1} + 2d^{-2} + O(d^{-3})$. Hence a self-avoiding walk moves away from the origin faster than a Pólya walk, but only at the level of the amplitude and not at the level of the exponent.

If we accept the conjectured asymptotics $c_n \sim A\mu^n n^{\gamma-1}$ and $r_n \sim Bn^{2\nu}$ as truth (for $d \neq 4$), then the calculations shown in Table 5.3 become possible [23,24,33,35–37].

Table 5.3. *Estimates for Amplitudes A and B*

d	Estimate for A	Estimate for B	d	Estimate for A	Estimate for B
2	1.177043	0.77100	5	1.275	1.4767
3	1.205	1.21667	6	1.159	1.2940

(The logarithmic correction for $d = 4$ renders any reliable estimation of A or B very difficult.) Here is an application. Two walks ω and ω' **intersect** if $\omega_i = \omega'_j$ for some nonzero i and j. The probability that two n-step independent random self-avoiding walks never intersect is [1,38]

$$\frac{c_{2n}}{c_n^2} \sim \begin{cases} A^{-1}2^{\gamma-1}n^{1-\gamma} \to 0 & \text{if } 2 \le d \le 3, \\ A^{-1}\ln(n)^{-1/4} \to 0 & \text{if } d = 4, \\ A^{-1} > 0 & \text{if } d \ge 5 \end{cases}$$

as $n \to \infty$. This conjectured behavior is consistent with intuition: c_{2n}/c_n^2 is (slightly) larger than the corresponding probability q_n for ordinary walks [5.9.1] since self-avoiding walks tend to be more thinly dispersed in space.

Other interesting measures of the size of a walk include the **mean square radius of gyration,**

$$s_n = \mathrm{E}\left(\frac{1}{n+1}\sum_{i=0}^{n}\left|\omega_i - \frac{1}{n+1}\sum_{j=0}^{n}\omega_j\right|^2\right) = \mathrm{E}\left(\frac{1}{2(n+1)^2}\sum_{i=0}^{n}\sum_{j=0}^{n}|\omega_i - \omega_j|^2\right),$$

and the **mean square distance of a monomer from the endpoints,**

$$t_n = \mathrm{E}\left(\frac{1}{n+1}\sum_{i=0}^{n}\frac{|\omega_i|^2 + |\omega_n - \omega_i|^2}{2}\right).$$

The radius of gyration, for example, can be experimentally measured for polymers in a dilute solution via light scattering, but the end-to-end distance is preferred for theoretical simplicity [33,39–41]. It is conjectured that $s_n \sim En^{2\nu}$ and $t_n \sim Fn^{2\nu}$, where ν is the same exponent as for r_n, and $E/B = 0.14026\ldots$, $F/B = 0.43961\ldots$ for $d = 2$ and $E/B = 0.1599\ldots$ for $d = 3$.

One can generalize this discussion to arbitrary lattices L in d-dimensional space. For example, in the case $d = 2$, there is a rigorous upper bound $\mu < 4.278$ and an estimate $\mu = 4.1507951\ldots$ for the equilateral triangular lattice [17,35,42–45], and it is conjectured that $\mu = \sqrt{2+\sqrt{2}} = 1.8477590650\ldots$ for the hexagonal (honeycomb) lattice [46–48]. The critical exponents γ, ν and amplitude ratios E/B, F/B, however, are thought to be *universal* in the sense that they are lattice-independent (although dimension-dependent). An important challenge, therefore, is to better understand the nature of such exponents and ratios, and certainly to prove their existence in low dimensions.

5.10.1 Polygons and Trails

The connective constant μ values given previously apply not only to the asymptotic growth of the number of self-avoiding walks, but also to the asymptotic growth of numbers of **self-avoiding polygons** and of self-avoiding walks with prescribed endpoints [2,49]. See [5.19] for discussion of lattice animals or polyominoes, which are related to self-avoiding polygons.

No site or bond may be visited more than once in a self-avoiding walk. By way of contrast, a **self-avoiding trail** may revisit sites, but not bonds. Thus walks are a proper subset of trails [50–55]. The number h_n of trails is conjectured to satisfy $h_n \sim G\lambda^n n^{\gamma-1}$, where γ is the same exponent as for c_n. The connective constant λ provably exists as before and, in fact, satisfies $\lambda \geq \mu$. For the square lattice, there are rigorous bounds $2.634 < \lambda < 2.851$ and an estimate $\lambda = 2.72062\ldots$; the amplitude is approximately $G = 1.272\ldots$. For the cubic lattice, there is an upper bound $\lambda < 4.929$ and an estimate $\lambda = 4.8426\ldots$. Many related questions can be asked.

5.10.2 Rook Paths on a Chessboard

How many self-avoiding walks can a rook take from a fixed corner of an $m \times n$ chessboard to the opposite corner without ever leaving the chessboard? Denote the number of such **paths** by $p_{m-1,n-1}$; clearly $p_{k,1} = 2^k$, $p_{2,2} = 12$, and [56–58]

$$p_{k,2} \sim \frac{4+\sqrt{13}}{2\sqrt{13}}\left(\sqrt{\frac{3+\sqrt{13}}{2}}\right)^{2k} = 1.0547001962\ldots\cdot(1.8173540210\ldots)^{2k}$$

as $k \to \infty$. More broadly, the generating function for the sequence $\{p_{k,l}\}_{k=1}^{\infty}$ is rational for any integer $l \geq 1$ and thus relevant asymptotic coefficients are all algebraic numbers. What can be said about the asymptotics of $p_{k,k}$ as $k \to \infty$? Whittington & Guttmann [59] proved that

$$p_{k,k} \sim (1.756\ldots)^{k^2}$$

and conjectured the following [60,61]. If $\pi_{j,k}$ is the number of j-step paths with generating function

$$P_k(x) = \sum_{j=1}^{\infty} \pi_{j,k} x^j, \quad P_k(1) = p_{k,k}$$

then there is a *phase transition* in the sense that

$$\begin{aligned} &0 < \lim_{k\to\infty} P_k(x)^{\frac{1}{k}} < 1 \quad \text{exists for } 0 < x < \mu^{-1} = 0.3790522777\ldots, \\ &\lim_{k\to\infty} P_k(\mu^{-1})^{\frac{1}{k}} = 1, \\ &1 < \lim_{k\to\infty} P_k(x)^{\frac{1}{k^2}} < \infty \text{ exists for } x > \mu^{-1}. \end{aligned}$$

A proof was given by Madras [62]. This is an interesting occurrence of the connective constant $\mu = \mu_2$; an analogous theorem involving a d-dimensional chessboard also holds and naturally makes use of μ_d.

5.10.3 Meanders and Stamp Foldings

A **meander** of order n is a planar self-avoiding loop (road) crossing an infinite line (river) $2n$ times ($2n$ bridges). Define two meanders as equivalent if one may be deformed continuously into the other, keeping the bridges fixed. The number of inequivalent meanders M_n of order n satisfy $M_1 = 1$, $M_2 = 2$, $M_3 = 8$, $M_4 = 42$, $M_5 = 262, \ldots$.

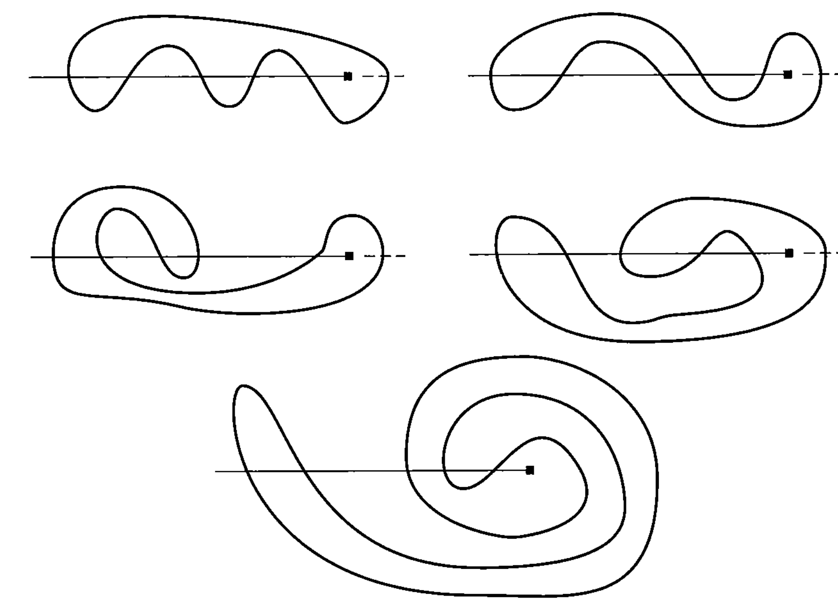


Figure 5.11. There are eight meanders of order 3 and ten semi-meanders of order 5; reflections across the river are omitted.

A **semi-meander** of order n is a planar self-avoiding loop (road) crossing a semi-infinite line (river with a source) n times (n bridges). Equivalence of semi-meanders is defined similarly. The number of inequivalent semi-meanders $\tilde{M}_n$ of order n satisfy $\tilde{M}_1 = 1$, $\tilde{M}_2 = 1$, $\tilde{M}_3 = 2$, $\tilde{M}_4 = 4$, $\tilde{M}_5 = 10, \ldots$.

Counting meanders and semi-meanders has attracted much attention [63–73]. See Figure 5.11. As before, we expect asymptotic behavior

$$M_n \sim C\frac{R^{2n}}{n^{\alpha}}, \quad \tilde{M}_n \sim \tilde{C}\frac{R^n}{n^{\tilde{\alpha}}},$$

where $R = 3.501838\ldots$, that is, $R^2 = 12.262874\ldots$. No exact formula for the connective constant R is known. In contrast, there is a conjecture [74–76] that the critical exponents are given by

$$\alpha = \sqrt{29}\frac{\sqrt{29}+\sqrt{5}}{12} = 3.4201328816\ldots,$$

$$\tilde{\alpha} = 1 + \sqrt{11}\frac{\sqrt{29}+\sqrt{5}}{24} = 2.0531987328\ldots,$$

but doubt has been raised [77–79] about the semi-meander critical exponent value. The sequences $\tilde{M}_n$ and M_n are also related to enumerating the ways of folding a linear or circular row of stamps onto one stamp [80–87].

[1] N. Madras and G. Slade, *The Self-Avoiding Walk*, Birkhäuser, 1993; MR 94f:82002.
[2] B. D. Hughes, *Random Walks and Random Environments*, v. 1, Oxford Univ. Press, 1995; MR 96i:60070.
[3] G. Slade, Self-avoiding walks, *Math. Intellig.* 16 (1994) 29–35.

[4] G. Slade, Random walks, *Amer. Scientist*, v. 84 (1996) n. 2, 146–153.
[5] B. Hayes, How to avoid yourself, *Amer. Scientist*, v. 86 (1998) n. 4, 314–319.
[6] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001334, A001411, A001412, and A001668.
[7] P. J. Flory, The configuration of real polymer chains, *J. Chem. Phys.* 17 (1949) 303–310.
[8] K. F. Freed, Polymers as self-avoiding walks, *Annals of Probab.* 9 (1981) 537–556; MR 82j:60123.
[9] M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Math. Z.* 17 (1923) 228–249.
[10] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, v. 1, Springer-Verlag, 1972, ex. 98, pp. 23, 198; MR 81e:00002.
[11] J. M. Hammersley and K. W. Morton, Poor man's Monte Carlo, *J. Royal Statist. Soc. Ser. B.* 16 (1954) 23–38; discussion 61–75; MR 16,287i.
[12] J. M. Hammersley, Percolation processes. II: The connective constant, *Proc. Cambridge Philos. Soc.* 53 (1957) 642–645; MR 19,989f.
[13] M. E. Fisher and M. F. Sykes, Excluded-volume problem and the Ising model of ferromagnetism, *Phys. Rev.* 114 (1959) 45–58; MR 22 #10274.
[14] W. A. Beyer and M. B. Wells, Lower bound for the connective constant of a self-avoiding walk on a square lattice, *J. Combin. Theory Ser. A* 13 (1972) 176–182; MR 46 #1605.
[15] F. T. Wall and R. A. White, Macromolecular configurations simulated by random walks with limited orders of non-self-intersections, *J. Chem. Phys.* 65 (1976) 808–812.
[16] T. Hara, G. Slade, and A. D. Sokal, New lower bounds on the self-avoiding-walk connective constant, *J. Stat. Phys.* 72 (1993) 479–517; erratum 78 (1995) 1187–1188; MR 94c:82053.
[17] S. E. Alm, Upper bounds for the connective constant of self-avoiding walks, *Combin. Probab. Comput.* 2 (1993) 115–136; MR 94g:60126.
[18] A. R. Conway and A. J. Guttmann, Lower bound on the connective constant for square lattice self-avoiding walks, *J. Phys. A* 26 (1993) 3719–3724.
[19] J. Noonan, New upper bounds for the connective constants of self-avoiding walks, *J. Stat. Phys.* 91 (1998) 871–888; MR 99g:82030.
[20] J. Noonan and D. Zeilberger, The Goulden-Jackson cluster method: Extensions, applications and implementations, *J. Differ. Eq. Appl.* 5 (1999) 355–377; MR 2000f:05005.
[21] I. Jensen and A. J. Guttmann, Self-avoiding polygons on the square lattice, *J. Phys. A* 32 (1999) 4867–4876.
[22] A. Pönitz and P. Tittmann, Improved upper bounds for self-avoiding walks in $\mathbb{Z}^d$, *Elec. J. Combin.* 7 (2000) R21; MR 2000m:05015.
[23] D. MacDonald, S. Joseph, D. L. Hunter, L. L. Moseley, N. Jan, and A. J. Guttmann, Self-avoiding walks on the simple cubic lattice, *J. Phys. A* 33 (2000) 5973–5983; MR 2001g:82051.
[24] A. L. Owczarek and T. Prellberg, Scaling of self-avoiding walks in high dimensions, *J. Phys. A* 34 (2001) 5773–5780; cond-mat/0104135.
[25] H. Kesten, On the number of self-avoiding walks, *J. Math. Phys.* 4 (1963) 960–969; 5 (1964) 1128–1137; MR 27 #2006 and 29 #4118.
[26] G. L. O'Brien, Monotonicity of the number of self-avoiding walks, *J. Stat. Phys.* 59 (1990) 969–979; MR 91k:8007.
[27] P. Butera and M. Comi, N-vector spin models on the simple-cubic and the body-centered-cubic lattices: A study of the critical behavior of the susceptibility and of the correlation length by high-temperature series extended to order 21, *Phys. Rev. B* 56 (1997) 8212–8240; hep-lat/9703018.
[28] S. Caracciolo, M. S. Causo, and A. Pelissetto, Monte Carlo results for three-dimensional self-avoiding walks, *Nucl. Phys. Proc. Suppl.* 63 (1998) 652–654; hep-lat/9711051.
[29] S. Caracciolo, M. S. Causo, and A. Pelissetto, High-precision determination of the critical exponent gamma for self-avoiding walks, *Phys. Rev. E* 57 (1998) R1215–R1218; cond-mat/9703250.
[30] T. Hara and G. Slade, Self-avoiding walk in five or more dimensions. I: The critical behaviour, *Commun. Math. Phys.* 147 (1992) 101–136; MR 93j:82032.

[31] J. M. Hammersley and D. J. A. Welsh, Further results on the rate of convergence to the connective constant of the hypercubical lattice, *Quart. J. Math.* 13 (1962) 108–110; MR 25 #2967.
[32] T. Hara and G. Slade, The self-avoiding-walk and percolation critical points in high dimensions, *Combin. Probab. Comput.* 4 (1995) 197–215; MR 96i:82081.
[33] B. Li, N. Madras, and A. D. Sokal, Critical exponents, hyperscaling and universal amplitude ratios for two- and three-dimensional self-avoiding walks, *J. Stat. Phys.* 80 (1995) 661–754; hep-lat/9409003; MR 96e:82046.
[34] T. Prellberg, Scaling of self-avoiding walks and self-avoiding trails in three dimensions, *J. Phys. A* 34 (2001) L599–L602.
[35] I. G. Enting and A. J. Guttmann, Self-avoiding rings on the triangular lattice, *J. Phys. A* 25 (1992) 2791–2807; MR 93b:82030.
[36] A. R. Conway, I. G. Enting, and A. J. Guttmann, Algebraic techniques for enumerating self-avoiding walks on the square lattice, *J. Phys. A* 26 (1993) 1519–1534; MR 94f:82040.
[37] A. R. Conway and A. J. Guttmann, Square lattice self-avoiding walks and corrections-to-scaling, *Phys. Rev. Lett.* 77 (1996) 5284–5287.
[38] G. F. Lawler, *Intersections of Random Walks*, Birkhäuser, 1991; MR 92f:60122.
[39] S. Caracciolo, A. Pelissetto, and A. D. Sokal, Universal distance ratios for two-dimensional self-avoiding walks: Corrected conformal-invariance predictions, *J. Phys. A* 23 (1990) L969–L974; MR 91g:81130.
[40] J. L. Cardy and A. J. Guttmann, Universal amplitude combinations for self-avoiding walks, polygons and trails, *J. Phys. A* 26 (1993) 2485–2494; MR 94f:82039.
[41] K. Y. Lin and J. X. Huang, Universal amplitude ratios for self-avoiding walks on the Kagomé lattice, *J. Phys. A* 28 (1995) 3641–3643.
[42] A. J. Guttmann, On two-dimensional self-avoiding random walks, *J. Phys. A* 17 (1984) 455–468; MR 85g:82096.
[43] T. Ishinabe, Critical exponents and corrections to scaling for the two-dimensional self-avoiding walk, *Phys. Rev. B* 37 (1988) 2376–2379.
[44] T. Ishinabe, Reassessment of critical exponents and corrections to scaling for self-avoiding walks, *Phys. Rev. B* 39 (1989) 9486–9495.
[45] A. J. Guttmann and J. Wang, The extension of self-avoiding random walk series in two dimensions, *J. Phys. A* 24 (1991) 3107–3109; MR 92b:82073.
[46] B. Nienhuis, Exact critical point and critical exponents of $O(n)$ models in two dimensions, *Phys. Rev. Lett.* 49 (1982) 1062–1065.
[47] I. G. Enting and A. J. Guttmann, Polygons on the honeycomb lattice, *J. Phys. A* 22 (1989) 1371–1384; MR 90c:82050.
[48] D. MacDonald, D. L. Hunter, K. Kelly, and N. Jan, Self-avoiding walks in two to five dimensions: Exact enumerations and series study, *J. Phys. A* 25 (1992) 1429–1440.
[49] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006817, A006818, A006819, and A006851.
[50] A. Malakis, The trail problem on the square lattice, *J. Phys. A* 9 (1976) 1283–1291.
[51] A. J. Guttmann and S. S. Manna, The geometric properties of planar lattice trails, *J. Phys. A* 22 (1989) 3613–3619.
[52] A. J. Guttmann, Lattice trails. I: Exact results, *J. Phys. A* 18 (1985) 567–573; MR 87f:82069.
[53] A. J. Guttmann, Lattice trails. II: Numerical results, *J. Phys. A* 18 (1985) 575–588; MR 87f:82070.
[54] A. R. Conway and A. J. Guttmann, Enumeration of self-avoiding trails on a square lattice using a transfer matrix technique, *J. Phys. A* 26 (1993) 1535–1552; MR 94f:82041.
[55] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001335, A001413, A002931, and A036418.
[56] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A006192, A007764, A007786, and A007787.
[57] H. L. Abbott and D. Hanson, A lattice path problem, *Ars Combin.* 6 (1978) 163–178; MR 80m:05009.

[58] K. Edwards, Counting self-avoiding walks in a bounded region, *Ars Combin.* 20 *B* (1985) 271–281; MR 87h:05015.
[59] S. G. Whittington and A. J. Guttmann, Self-avoiding walks which cross a square, *J. Phys. A* 23 (1990) 5601–5609; MR 92e:82038.
[60] T. W. Burkhardt and I. Guim, Self-avoiding walks that cross a square, *J. Phys. A* 24 (1991) L1221–L1228.
[61] J. J. Prentis, Renormalization theory of self-avoiding walks which cross a square, *J. Phys. A* 24 (1991) 5097–5103.
[62] N. Madras, Critical behaviour of self-avoiding walks that cross a square, *J. Phys. A* 28 (1995) 1535–1547; MR 96d:82032.
[63] H. Poincaré, Sur un théorème de géométrie, *Rend. Circ. Mat. Palmero* 33 (1912) 375–407; also in *Oeuvres*, t. 6, Gauthier-Villars, 1953, pp. 499–538.
[64] J. Touchard, Contribution à l'étude du problème des timbres poste, *Canad. J. Math.* 2 (1950) 385–398; MR 12,312i.
[65] V. I. Arnold, The branched covering $CP^2 \to S^4$, hyperbolicity and projective topology (in Russian), *Sibirskii Mat. Zh.*, v. 29 (1988) n. 5, 36–47, 237; Engl. transl. in *Siberian Math. J.* 29 (1988) 717–726; MR 90a:57037.
[66] S. K. Lando and A. K. Zvonkin, Meanders, *Selecta Math. Soviet.* 11 (1992) 117–144; MR 93k:05013.
[67] S. K. Lando and A. K. Zvonkin, Plane and projective meanders, *Theoret. Comput. Sci.* 117 (1993) 227–241; MR 94i:05004.
[68] P. Di Francesco, O. Golinelli, and E. Guitter, Meanders: A direct enumeration approach, *Nucl. Phys. B* 482 (1996) 497–535; hep-th/9607039; MR 97j:82074.
[69] P. Di Francesco, O. Golinelli, and E. Guitter, Meander, folding and arch statistics, *Math. Comput. Modelling* 26 (1997) 97–147; hep-th/9506030; MR 99f:82029.
[70] P. Di Francesco, Meander determinants, *Commun. Math. Phys.* 191 (1998) 543–583; hep-th/9612026; MR 99e:05007.
[71] M. G. Harris, A diagrammatic approach to the meander problem, hep-th/9807193.
[72] J. Reeds and L. Shepp, An upper bound on the meander constant, unpublished note (1999).
[73] O. Golinelli, A Monte-Carlo study of meanders, *Europ. Phys. J. B* 14 (2000) 145–155; cond-mat/9906329.
[74] P. Di Francesco, O. Golinelli, and E. Guitter, Meanders: Exact asymptotics, *Nucl. Phys. B* 570 (2000) 699–712; cond-mat/9910453; MR 2001f:82032.
[75] P. Di Francesco, Folding and coloring problems in mathematics and physics, *Bull. Amer. Math. Soc.* 37 (2000) 251–307; MR 2001g:82004.
[76] P. Di Francesco, E. Guitter, and J. L. Jacobsen, Exact meander asymptotics: A numerical check, *Nucl. Phys. B* 580 (2000) 757–795; MR 2002b:82024.
[77] I. Jensen, Enumerations of plane meanders, cond-mat/9910313.
[78] I. Jensen, A transfer matrix approach to the enumeration of plane meanders, *J. Phys. A* 33 (2000) 5953–5963; cond-mat/0008178; MR 2001e:05008.
[79] I. Jensen and A. J. Guttmann, Critical exponents of plane meanders, *J. Phys. A* 33 (2000) L187–L192; MR 2001b:82024.
[80] W. F. Lunnon, A map-folding problem, *Math. Comp.* 22 (1968) 193–199; MR 36 #5009.
[81] J. E. Koehler, Folding a strip of stamps, *J. Combin. Theory* 5 (1968) 135–152; errata *J. Combin. Theory Ser. A* 19 (1975) 367; MR 37 #3945 and MR 52 #2910.
[82] W. F. Lunnon, Multi-dimensional map-folding, *Comput. J.* 14 (1971) 75–80; MR 44 #2627.
[83] W. F. Lunnon, Bounds for the map-folding function, unpublished note (1981).
[84] M. Gardner, The combinatorics of paper-folding, *Wheels, Life, and Other Mathematical Amusements*, W. H. Freeman, 1983, pp. 60–73.
[85] M. Gardner, *The Sixth Book of Mathematical Games from Scientific American*, Univ. of Chicago Press, 1984, pp. 21, 26–27.
[86] N. J. A. Sloane, My favorite integer sequences, *Sequences and Their Applications (SETA)*, Proc. 1998 Singapore conf., ed. C. Ding, T. Helleseth, and H. Niederreiter, Springer-Verlag, 1999, pp. 103–130; math.CO/0207175.
[87] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000682, A001011, A005315, and A005316.

5.11 Feller's Coin Tossing Constants

Let w_n denote the probability that, in n independent tosses of an ideal coin, no run of three consecutive heads appears. Clearly $w_0 = w_1 = w_2 = 1$, $w_n = \frac{1}{2}w_{n-1} + \frac{1}{4}w_{n-2} + \frac{1}{8}w_{n-3}$ for $n \geq 3$, and $\lim_{n\to\infty} w_n = 0$. Feller [1] proved the following more precise asymptotic result:

$$\lim_{n\to\infty} w_n \alpha^{n+1} = \beta,$$

where

$$\alpha = \frac{\left(136 + 24\sqrt{33}\right)^{\frac{1}{3}} - 8\left(136 + 24\sqrt{33}\right)^{-\frac{1}{3}} - 2}{3} = 1.0873780254\ldots$$

and

$$\beta = \frac{2-\alpha}{4-3\alpha} = 1.2368398446\ldots.$$

We first examine generalizations of these formulas. If runs of k consecutive heads, $k > 1$, are disallowed, then the analogous constants are [1,2]

$$\alpha \text{ is the smallest positive root of } 1 - x + \left(\frac{x}{2}\right)^{k+1} = 0$$

and

$$\beta = \frac{2-\alpha}{k+1-k\alpha}.$$

Equivalently, the generating function that enumerates coin toss sequences with no runs of k consecutive heads is [3]

$$S_k(z) = \frac{1-z^k}{1-2z+z^{k+1}}, \quad \frac{1}{n!}\frac{d^n}{dz^n}S_k(z)\bigg|_{z=0} \sim \frac{\beta}{\alpha}\left(\frac{2}{\alpha}\right)^n.$$

See [4–8] for more material of a combinatorial nature.

If the coin is non-ideal, that is, if $\mathrm{P}(H) = p$, $\mathrm{P}(T) = q$, $p + q = 1$, but p and q are not equal, then the asymptotic behavior of w_n is governed by

$$\alpha \text{ is the smallest positive root of } 1 - x + qp^k x^{k+1} = 0$$

and

$$\beta = \frac{1-p\alpha}{(k+1-k\alpha)q}.$$

A further generalization involves time-homogeneous two-state Markov chains. It makes little sense here to talk of coin tosses, so we turn attention to a different application. Imagine that a ground-based sensor determines once per hour whether a fixed line-of-sight through the atmosphere is cloud-obscured (0) or clear (1). Since meteorological events often display persistence through time, the sensor observations are not independent. A simple model for the time series $X_1, X_2, X_3, \ldots$ of observations might be a Markov chain with transition probability matrix

$$\begin{pmatrix} P(X_{j+1}=0|X_j=0) & P(X_{j+1}=1|X_j=0) \\ P(X_{j+1}=0|X_j=1) & P(X_{j+1}=1|X_j=1) \end{pmatrix} = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix},$$

where conditional probability parameters satisfy $\pi_{00}+\pi_{01}=1=\pi_{10}+\pi_{11}$. The special case when $\pi_{00}=\pi_{10}$ and $\pi_{01}=\pi_{11}$ is equivalent to the Bernoulli trials scenario discussed in connection with coin tossing. Let $w_{n,k}$ denote the probability that no cloudy intervals of length $k>1$ occur, and assume that initially $P(X_0=1)=\theta_1$. The asymptotic behavior is similar to before, where α is the smallest positive root of [9, 10]

$$1-(\pi_{11}+\pi_{00})x+(\pi_{11}-\pi_{01})x^2+\pi_{10}\pi_{01}\pi_{11}^{k-1}x^{k+1}=0$$

and

$$\beta=\frac{[-1+(2\pi_{11}-\pi_{01})\alpha-(\pi_{11}-\pi_{01})\pi_{11}\alpha^2][\theta_1+(\pi_{01}-\theta_1)\alpha]}{\pi_{10}\pi_{01}[-1-k+(\pi_{11}+\pi_{00})k\alpha+(\pi_{11}-\pi_{01})(1-k)\alpha^2]}.$$

See [11] for a general technique for analysis of pattern statistics, with applications in molecular biology.

Of many possible variations on this problem, we discuss one. How many patterns of n children in a row are there if every girl is next to at least one other girl? If we denote the answer by Y_n, then $Y_1=1$, $Y_2=2$, $Y_3=4$, and $Y_n=2Y_{n-1}-Y_{n-2}+Y_{n-3}$ for $n\geq 4$; hence

$$\lim_{n\to\infty}\frac{Y_{n+1}}{Y_n}=\frac{\left(100+12\sqrt{69}\right)^{\frac{1}{3}}+4\left(100+12\sqrt{69}\right)^{-\frac{1}{3}}+4}{6}=1.7548776662\ldots.$$

A generalization of this, in which the girls must appear in groups of at least k, is given in [12, 13]. Similar cubic irrational numbers occur in [1.2.2].

Let us return to coin tossing. What is the expected length of the longest run of consecutive heads in a sequence of n ideal coin tosses? The answer is surprisingly complicated [14–21]:

$$\sum_{k=1}^{n}(1-w_{n,k})=\frac{\ln(n)}{\ln(2)}-\left(\frac{3}{2}-\frac{\gamma}{\ln(2)}\right)+\delta(n)+o(1)$$

as $n\to\infty$, where γ is the Euler-Mascheroni constant and

$$\delta(n)=\frac{1}{\ln(2)}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\Gamma\left(\frac{2\pi ik}{\ln(2)}\right)\exp\left(-2\pi ik\frac{\ln(n)}{\ln(2)}\right).$$

That is, the expected length is $\ln(n)/\ln(2)-0.6672538227\ldots$ plus an oscillatory, small-amplitude correction term. The function $\delta(n)$ is periodic ($\delta(n)=\delta(2n)$), has zero mean, and is "negligible" ($|\delta(n)|<1.574\times 10^{-6}$ for all n). The corresponding variance is $C+c+\varepsilon(n)+o(1)$, where $\varepsilon(n)$ is another small-amplitude function and

$$C=\frac{1}{12}+\frac{\pi^2}{6\ln(2)^2}=3.5070480758\ldots,$$

$$c=\frac{2}{\ln(2)}\sum_{k=0}^{\infty}\ln\left[1-\exp\left(-\frac{2\pi^2}{\ln(2)}(2k+1)\right)\right]=(-1.237412\ldots)\times 10^{-12}.$$

Functions similar to $\delta(n)$ and $\varepsilon(n)$ appear in [2.3], [2.16], [5.6], and [5.14].

Also, if 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., what is the probability that the final toss involves exactly one coin? Again, the answer is complicated [22–25]:

$$\frac{n}{2}\sum_{j=0}^{\infty}2^{-j}(1-2^{-j})^{n-1}\sim\frac{1}{2\ln(2)}+\rho(n)+o(1)$$

as $n\to\infty$, where

$$\rho(n)=\frac{1}{2\ln(2)}\sum_{\substack{k=-\infty\\k\neq 0}}^{\infty}\Gamma\left(1-\frac{2\pi ik}{\ln(2)}\right)\exp\left(2\pi ik\frac{\ln(n)}{\ln(2)}\right).$$

That is, the probability of a unique survivor (no ties) at the end is $1/(2\ln(2))=0.7213475204\ldots$ plus an oscillatory function satisfying $|\rho(n)|<7.131\times 10^{-6}$ for all n. The expected length of the longest of the n coin toss sequences is $\sum_{j=0}^{\infty}[1-(1-2^{-j})^n]$ and can be analyzed similarly [26]. Related discussion is found in [27–31].

[1] W. Feller, *An Introduction to Probability Theory and Its Applications*, v. 1, 3rd ed., Wiley, 1968, pp. 278–279, 322–325; MR 37 #3604.

[2] W. Feller, Fluctuation theory of recurrent events, *Trans. Amer. Math. Soc.* 67 (1949) 98–119; MR 11,255c.

[3] R. Sedgewick and P. Flajolet, *Introduction to the Analysis of Algorithms*, Addison-Wesley, 1996, pp. 366–373.

[4] R. A. Howard, *Dynamic Probabilistic Systems: Markov Models*, v. 1, Wiley, 1971.

[5] A. N. Philippou and F. S. Makri, Successes, runs and longest runs, *Statist. Probab. Lett.* 4 (1986) 211–215; MR 88c:60035b.

[6] K. Hirano, Some properties of the distributions of order k, *Fibonacci Numbers and Their Applications*, Proc. 1984 Patras conf., ed. A. N. Philippou, G. E. Bergum, and A. F. Horadam, Reidel, 1986; MR 88b:62031.

[7] A. P. Godbole, Specific formulae for some success run distributions, *Statist. Probab. Lett.* 10 (1990) 119–124; MR 92e:60021.

[8] K. A. Suman, The longest run of any letter in a randomly generated word, *Runs and Patterns in Probability: Selected Papers*, ed. A. P. Godbole and S. G. Papastavnidis, Kluwer, 1994, pp. 119–130; MR 95k:60028.

[9] M. B. Rajarshi, Success runs in a two-state Markov chain, *J. Appl. Probab.* 11 (1974) 190–192; correction 14 (1977) 661; MR 51 #4402 and MR 57 #7777.

[10] C. Banderier and M. Vandenbogaert, A Markovian generalization of Feller's coin tossing constants, unpublished note (2000).

[11] P. Nicodème, B. Salvy, and P. Flajolet, Motif statistics, *Proc. 1999 European Symp. on Algorithms (ESA)*, Prague, ed. J. Nesetril, Lect. Notes in Comp. Sci. 1643, Springer-Verlag, 1999, pp. 194–211; INRIA research report RR-3606.

[12] J. H. Conway and R. K. Guy, *The Book of Numbers*, Springer-Verlag, 1996, pp. 205–206; MR 98g:00004.
[13] R. Austin and R. Guy, Binary sequences without isolated ones, *Fibonacci Quart.* 16 (1978) 84–86; MR 57 #5778.
[14] D. W. Boyd, Losing runs in Bernoulli trials, unpublished note (1972).
[15] P. Erdös and P. Révész, On the length of the longest head-run, *Topics in Information Theory*, Proc. 1975 Keszthely conf., ed. I. Csiszár and P. Elias, Colloq. Math. Soc. János Bolyai 16, North-Holland, 1977, pp. 219–228; MR 57 #17788.
[16] D. E. Knuth, The average time for carry propagation, *Proc. Konink. Nederl. Akad. Wetensch. Ser. A* 81 (1978) 238–242; *Indag. Math.* 40 (1978) 238–242; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 467–471; MR 81b:68030.
[17] L. J. Guibas and A. M. Odlyzko, Long repetitive patterns in random sequences, *Z. Wahrsch. Verw. Gebiete* 53 (1980) 241–262; MR 81m:60047.
[18] L. Gordon, M. F. Schilling, and M. S. Waterman, An extreme value theory for long head runs, *Probab. Theory Relat. Fields* 72 (1986) 279–287; MR 87i:60023.
[19] H. Prodinger, Über längste 1-Teilfolgen in 0-1-Folgen, *Zahlentheoretische Analysis II*, Wien Seminar 1984-86, ed. E. Hlawka, Lect. Notes in Math. 1262, Springer-Verlag, 1987, pp. 124–133; MR 90g:11106.
[20] P. Hitczenko and G. Stengle, Expected number of distinct part sizes in a random integer composition, *Combin. Probab. Comput.* 9 (2000) 519–527; MR 2002d:05008.
[21] M. F. Schilling, The longest run of heads, *College Math J.* 21 (1990) 196–206.
[22] F. T. Bruss and C. A. O'Cinneide, On the maximum and its uniqueness for geometric random samples, *J. Appl. Probab.* 27 (1990) 598–610; MR 92a:60096.
[23] L. Räde, P. Griffin, and O. P. Lossers, Tossing coins until all show heads, *Amer. Math. Monthly* 101 (1994) 78–80.
[24] P. Kirschenhofer and H. Prodinger, The number of winners in a discrete geometrically distributed sample, *Annals of Appl. Probab.* 6 (1996) 687–694; MR 97g:60016.
[25] N. J. Calkin, E. R. Canfield, and H. S. Wilf, Averaging sequences, deranged mappings, and a problem of Lampert and Slater, *J. Combin. Theory Ser. A* 91 (2000) 171–190; MR 2002c:05019.
[26] C. F. Woodcock, On the asymptotic behaviour of a function arising from tossing coins, *Bull. London Math. Soc.* 28 (1996) 19–23; MR 96i:26001.
[27] B. Eisenberg, G. Stengle, and G. Strang, The asymptotic probability of a tie for first place, *Annals of Appl. Probab.* 3 (1993) 731–745; MR 95d:60044.
[28] J. J. A. M. Brands, F. W. Steutel, and R. J. G. Wilms, On the number of maxima in a discrete sample, *Statist. Probab. Lett.* 20 (1994) 209–217; MR 95e:60010.
[29] Y. Baryshnikov, B. Eisenberg, and G. Stengle, A necessary and sufficient condition for the existence of the limiting probability of a tie for first place, *Statist. Probab. Lett.* 23 (1995) 203–209; MR 96d:60015.
[30] Y. Qi, A note on the number of maxima in a discrete sample, *Statist. Probab. Lett.* 33 (1997) 373–377; MR 98f:60061.
[31] D. E. Lampert and P. J. Slater, Parallel knockouts in the complete graph, *Amer. Math. Monthly* 105 (1998) 556–558.

5.12 Hard Square Entropy Constant

Consider the set of all $n \times n$ binary matrices. What is the number $F(n)$ of such matrices with no pairs of adjacent 1s? Two 1s are said to be adjacent if they lie in positions (i, j) and $(i+1, j)$, or if they lie in positions (i, j) and $(i, j+1)$, for some i, j. Equivalently, $F(n)$ is the number of configurations of non-attacking Princes on an $n \times n$ chessboard, where a "Prince" attacks the four adjacent, non-diagonal places. Let $N = n^2$; then [1–3]

$$\kappa = \lim_{n\to\infty} F(n)^{\frac{1}{N}} = 1.5030480824\ldots = \exp(0.4074951009\ldots)$$

is the **hard square entropy constant**. Earlier estimates were obtained by both physicists [4–9] and mathematicians [10–13]. Some related combinatorial enumeration problems appear in [14–16].

Instead of an $n \times n$ binary matrix, consider an $n \times n$ binary array that looks like

$$\begin{pmatrix} a_{11} & & a_{23} & \\ & a_{22} & & a_{34} \\ a_{21} & & a_{33} & \\ & a_{32} & & a_{44} \\ a_{31} & & a_{43} & \\ & a_{42} & & a_{54} \\ a_{41} & & a_{53} & \\ & a_{52} & & a_{64} \end{pmatrix}$$

(here $n = 4$). What is the number $G(n)$ of such arrays with no pairs of adjacent 1s? Two 1s here are said to be adjacent if they lie in positions (i, j) and $(i+1, j)$, or in (i, j) and $(i, j+1)$, or in (i, j) and $(i+1, j+1)$, for some i, j. Equivalently, $G(n)$ is the number of configurations of non-attacking Kings on an $n \times n$ chessboard with regular hexagonal cells. It is surprising that the **hard hexagon entropy constant**

$$\kappa = \lim_{n\to\infty} G(n)^{\frac{1}{N}} = 1.3954859724\ldots = \exp(0.3332427219\ldots)$$

is *algebraic* (in fact, is solvable in radicals [17–22]) with minimal integer polynomial [23]

$$\begin{aligned} &25937424601x^{24} + 2013290651222784x^{22} + 250506231172067379 2x^{20} \\ &+ 79772669886665837977 6x^{18} + 744948831013108310016 0x^{16} \\ &+ 29580150383769582305 28x^{14} - 72405670285649161617408x^{12} \\ &+ 10715544815044338804 3264x^{10} - 71220809441400405884928x^{8} \\ &- 73347491183630103871488x^{6} + 97143135277377575190528x^{4} \\ &- 32751691810479015985152. \end{aligned}$$

This is a consequence of Baxter's exact solution of the hard hexagon model [24–27] via theta elliptic functions and the Rogers–Ramanujan identities from number theory [28–31]! The expression for κ, in fact, comes out of a more general expression for

$$\kappa(z) = \lim_{n\to\infty} Z_n(z)^{\frac{1}{N}},$$

where $Z_n(z)$ is known as the **partition function** for the model and $G(n) = Z_n(1)$, $\kappa = \kappa(1)$. More on the physics of phase transitions in lattice gas models is found in [5.12.1].

McKay and Calkin independently calculated that, if we replace Princes by Kings on the chessboard with square cells, then the corresponding constant κ is $1.3426439511\ldots$; see also [32–34]. Note that the distinction between Princes and Kings on a chessboard with regular hexagonal cells is immaterial. (Clarification: If a Prince occupies cell c, then any cell sharing an edge with c is vulnerable to attack. If a King occupies cell c, by contrast, then any cell sharing *either* an edge or corner with c is vulnerable.)

If we examine instead a chessboard with equilateral triangular cells, then $\kappa = 1.5464407087\ldots$ for Princes [3]. This may be called the **hard triangle entropy constant**. The value of κ when replacing Princes by Kings here is not known.

What are the constants κ for non-attacking Knights or Queens on chessboards with square cells? The analysis for Knights should be similar to that for Princes and Kings, but for Queens everything is different since interactions are no longer local [35].

The hard square entropy constant also appears in the form $\ln(\kappa)/\ln(2) = 0.5878911617\ldots$ in several coding-theoretic papers [36–41], with applications including holographic data storage and retrieval.

5.12.1 Phase Transitions in Lattice Gas Models

Statistical mechanics is concerned with the average properties of a large system of particles. We consider here, for example, the phase transition from a disordered fluid state to an ordered solid state, as temperature falls or density increases.

A simple model for this phenomenon is a **lattice gas**, in which particles are placed on the sites of a regular lattice and only adjacent particles interact. This may appear to be hopelessly idealized, as rigid molecules could not possibly satisfy such strict symmetry requirements. The model is nevertheless useful in understanding the link between microscopic and macroscopic descriptions of matter.

Two types of lattice gas models that have been studied extensively are the **hard square** model and the **hard hexagon** model. Once a particle is placed on a lattice site, no other particle is allowed to occupy the same site or any next to it, as pictured in Figure 5.12. Equivalently, the indicated squares and hexagons cannot overlap, hence giving rise to the adjective "hard."

Given a (square or triangular) lattice of N sites, assign a variable $\sigma_i = 1$ if site i is occupied and $\sigma_i = 0$ if it is vacant, for each $1 \le i \le N$. We study the **partition function**

$$Z_n(z) = \sum_{\sigma} \left(z^{\sigma_1+\sigma_2+\sigma_3+\cdots+\sigma_N} \cdot \prod_{(i,j)} (1 - \sigma_i \sigma_j) \right),$$

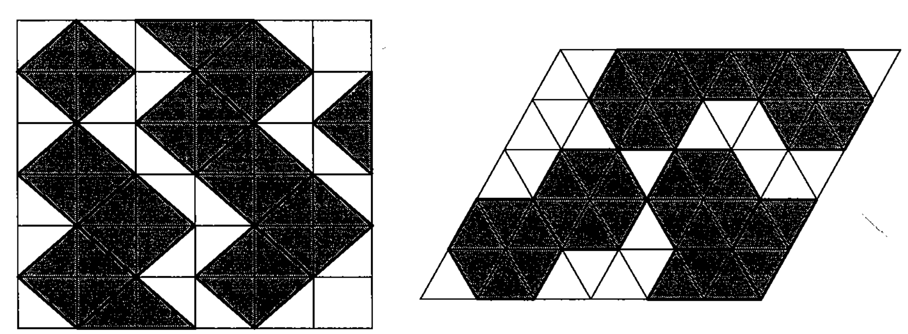


Figure 5.12. Hard squares and hard hexagons sit, respectively, on the square lattice and triangular lattice.

where the sum is over all 2^N possible values of the vector $\sigma = (\sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_N)$ and the product is over all edges of the lattice (sites i and j are distinct and adjacent). Observe that the product enforces the nearest neighbor exclusion: If a configuration has two particles next to each other, then zero contribution is made to the partition function.

It is customary to deal with boundary effects by wrapping the lattice around to form a torus. More precisely, for the square lattice, $2n$ new edges are created to connect the n rightmost and n topmost points to corresponding n leftmost and n bottommost points. Hence there are a total of $2N$ edges in the square lattice, each site "looking like" every other. For the triangular lattice, $4n - 1$ new edges are created, implying a total of $3N$ edges. In both cases, the number of boundary sites, relative to N, is vanishingly small as $n \to \infty$, so this convention does not lead to any error.

Clearly the following combinatorial expressions are true [4, 42, 43]: For the square lattice,

$$Z_n = \sum_{k=0}^{\lfloor N/2 \rfloor} f_{k,n} z^k, \quad f_{0,n} = 1, \quad f_{1,n} = N, \quad f_{2,n} = \begin{cases} 2 & \text{if } n = 2, \\ \frac{1}{2}N(N-5) & \text{if } n \ge 3, \end{cases}$$

$$f_{3,n} = \begin{cases} 6 & \text{if } n = 3, \\ \frac{1}{6}(N(N-10)(N-13) + 4N(N-9) + 4N(N-8)) & \text{if } n \ge 4, \end{cases}$$

where $f_{k,n}$ denotes the number of allowable tilings of the N-site lattice with k squares, and for the triangular lattice,

$$Z_n = \sum_{k=0}^{\lfloor N/3 \rfloor} g_{k,n} z^k, \quad g_{0,n} = 1, \quad g_{1,n} = N, \quad g_{2,n} = \tfrac{1}{2}N(N-7),$$

$$g_{3,n} = \begin{cases} 0 & \text{if } n = 3, \\ \frac{1}{6}(N(N-14)(N-19) + 6N(N-13) + 6N(N-12)) & \text{if } n \ge 4, \end{cases}$$

where $g_{k,n}$ denotes the corresponding number of hexagonal tilings.

Returning to physics, we remark that the partition function is important since it acts as the "denominator" in probability calculations. For example, consider the two sublattices A and B of the square lattice with sites as shown in Figure 5.13. The probability that an arbitrary site α in the sublattice A is occupied is

$$\rho_A(z) = \lim_{n\to\infty} \frac{1}{Z_n} \sum_{\sigma} \left(\sigma_\alpha \cdot z^{\sigma_1+\sigma_2+\sigma_3+\cdots+\sigma_N} \cdot \prod_{(i,j)} (1 - \sigma_i \sigma_j) \right),$$

which is also called the local density at α. We can define analogous probabilities for the three sublattices A, B, and C of the triangular lattice.

We are interested in the behavior of these models as a function of the positive variable z, known as the **activity**. Figure 5.14, for example, exhibits a graph of the mean density for the hard hexagon case:

$$\rho(z) = z\frac{d}{dz}(\ln(\kappa(z))) = \frac{\rho_A(z) + \rho_B(z) + \rho_C(z)}{3}$$

using the exact formulation given in [18].

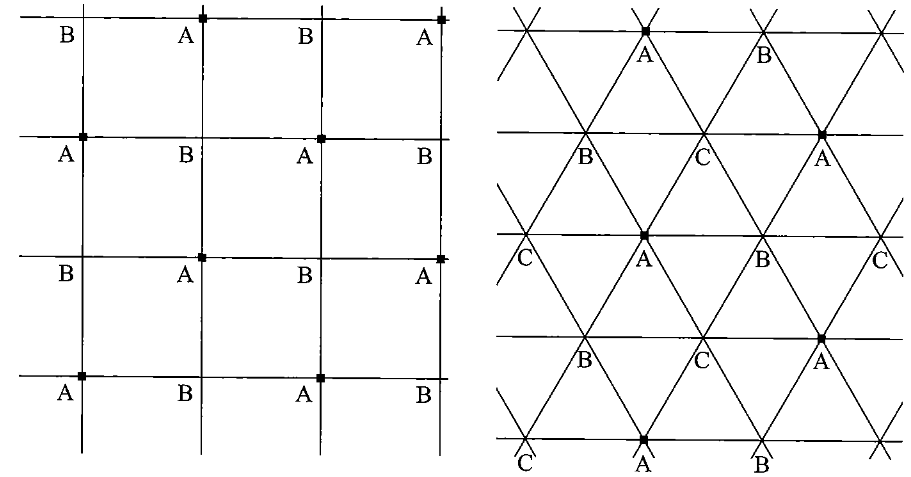


Figure 5.13. Two sublattices of the square lattice and three sublattices of the triangular lattice.

The existence of a phase transition is visually obvious. Let us look at the extreme cases: closely-packed configurations (large z) and sparsely-distributed configurations (small z). For infinite z, one of the possible sublattices is completely occupied, assumed to be the A sublattice, and the others are completely vacant; that is,

$$\rho_A = 1, \ \rho_B = 0 \ \text{ (for the square model)}$$

and

$$\rho_A = 1, \ \rho_B = \rho_C = 0 \ \text{ (for the hexagon model)}.$$

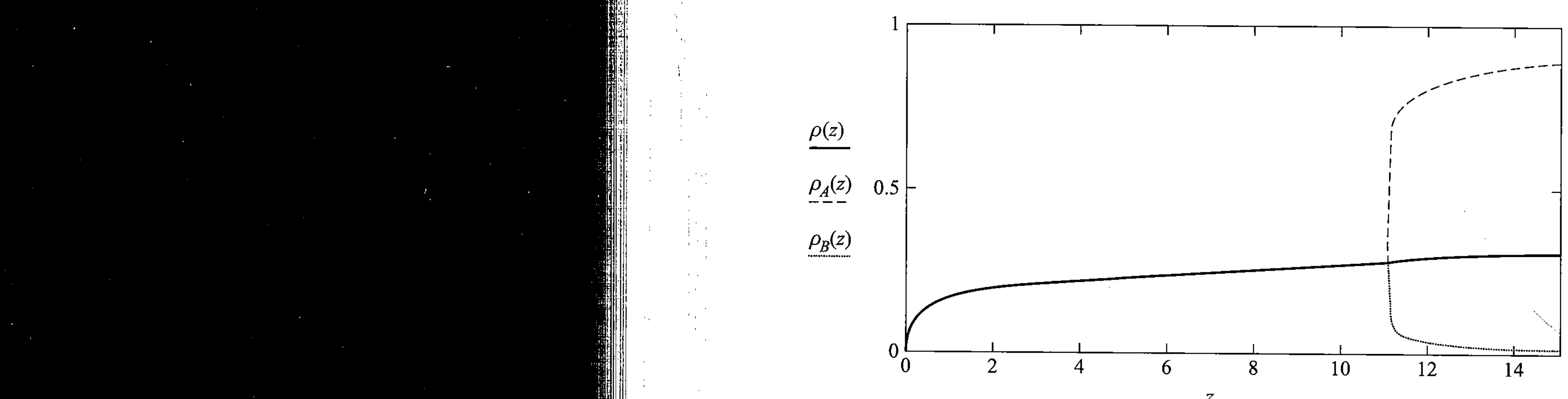


Figure 5.14. Graph of the mean density and sublattice densities, as functions of z, for the hexagon model.

For z close to zero, there is no preferential ordering on the sublattices; that is,

$$\rho_A = \rho_B \text{ (for the square model) and } \rho_A = \rho_B = \rho_C \text{ (for the hexagon model).}$$

Low activity corresponds to homogeneity and high activity corresponds to heterogeneity; thus there is a critical value, z_c, at which a phase transition occurs. Define the **order parameter**

$$R = \rho_A - \rho_B \text{ (for squares) and } R = \rho_A - \rho_B = \rho_A - \rho_C \text{ (for hexagons);}$$

then $R = 0$ for $z < z_c$ and $R > 0$ for $z > z_c$.

Elaborate numerical computations [7,44,45] have shown that, in the limit as $n \to \infty$,

$$z_c = 3.7962\ldots \text{ (for squares) and } z_c = 11.09\ldots \text{ (for hexagons),}$$

assuming site α to be infinitely deep within the lattice. The computations involved highly-accurate series expansions for R and what are known as corner transfer matrices, which we cannot discuss here for reasons of space.

In a beautiful development, Baxter [24,25] provided an exact solution of the hexagon model. The full breadth of this accomplishment cannot be conveyed here, but one of many corollaries is the exact formula

$$z_c = \frac{11+5\sqrt{5}}{2} = \left(\frac{1+\sqrt{5}}{2}\right)^5 = 11.0901699437\ldots$$

for the hexagon model. No similar theoretical breakthrough has occurred for the square model and thus the identity of $3.7962\ldots$ remains masked from sight. The critical value $z_c = 7.92\ldots$ for the triangle model (on the hexagonal or honeycomb lattice) likewise is not exactly known [46].

For hard hexagons, the behavior of $\rho(z)$ and $R(z)$ at criticality is important [24,26,27]:

$$\rho \sim \rho_c - 5^{-3/2}\left(1 - \frac{z}{z_c}\right)^{2/3} \text{ as } z \to z_c^-, \quad \rho_c = \frac{5-\sqrt{5}}{10} = 0.2763932022\ldots,$$

$$R \sim \frac{3}{\sqrt{5}}\left[\frac{1}{5\sqrt{5}}\left(\frac{z}{z_c} - 1\right)\right]^{1/9} \text{ as } z \to z_c^+,$$

and it is conjectured that the exponents $1/3$ and $1/9$ are universal. For hard squares and hard triangles, we have only numerical estimates $\rho_c = 0.368\ldots$ and $0.422\ldots$, respectively. Far away from criticality, computations at $z = 1$ are less difficult [3,47]:

$$\rho(1) = \begin{cases} 0.1624329213\ldots & \text{for hard hexagons,} \\ 0.2265708154\ldots & \text{for hard squares,} \\ 0.2424079763\ldots & \text{for hard triangles,} \end{cases}$$

and the first of these is algebraic of degree 12 [18,22]. A generalization of $\rho(1)$ is the probability that an arbitrary point α and a specified configuration of neighboring points α' are all occupied; sample computations can be found in [3].

Needless to say, three-dimensional analogs of the models discussed here defy any attempt at exact solution [44].

[1] N. J. Calkin and H. S. Wilf, The number of independent sets in a grid graph, *SIAM J. Discrete Math.* 11 (1998) 54–60; MR 99e:05010.
[2] B. D. McKay, On Calkin and Wilf's limit theorem for grid graphs, unpublished note (1996).
[3] R. J. Baxter, Planar lattice gases with nearest-neighbour exclusion, *Annals of Combin.* 3 (1999) 191–203; cond-mat/9811264; MR 2001h:82010.
[4] D. S. Gaunt and M. E. Fisher, Hard-sphere lattice gases. I: Plane-square lattice, *J. Chem. Phys.* 43 (1965) 2840–2863; MR 32 #3716.
[5] L. K. Runnels and L. L. Combs, Exact finite method of lattice statistics. I: Square and triangular lattice gases of hard molecules, *J. Chem. Phys.* 45 (1966) 2482–2492.
[6] B. D. Metcalf and C. P. Yang, Degeneracy of antiferromagnetic Ising lattices at critical magnetic field and zero temperature, *Phys. Rev. B* 18 (1978) 2304–2307.
[7] R. J. Baxter, I. G. Enting, and S. K. Tsang, Hard-square lattice gas, *J. Stat. Phys.* 22 (1980) 465–489; MR 81f:82037.
[8] P. A. Pearce and K. A. Seaton, A classical theory of hard squares, *J. Stat. Phys.* 53 (1988) 1061–1072; MR 89j:82076.
[9] S. Milosevic, B. Stosic, and T. Stosic, Towards finding exact residual entropies of the Ising antiferromagnets, *Physica A* 157 (1989) 899–906.
[10] N. G. Markley and M. E. Paul, Maximal measures and entropy for $\mathbb{Z}^\nu$ subshifts of finite type, *Classical Mechanics and Dynamical Systems*, ed. R. L. Devaney and Z. H. Nitecki, Dekker, 1981, pp. 135–157; MR 83c:54059.
[11] N. J. Calkin, *Sum-Free Sets and Measure Spaces*, Ph.D. thesis, Univ. of Waterloo, 1988.
[12] K. Weber, On the number of stable sets in an $m \times n$ lattice, *Rostock. Math. Kolloq.* 34 (1988) 28–36; MR 89i:05172.
[13] K. Engel, On the Fibonacci number of an $m \times n$ lattice, *Fibonacci Quart.* 28 (1990) 72–78; MR 90m:11033.
[14] H. Prodinger and R. F. Tichy, Fibonacci numbers of graphs, *Fibonacci Quart.* 20 (1982) 16–21; MR 83m:05125.
[15] Y. Kong, General recurrence theory of ligand binding on three-dimensional lattice, *J. Chem. Phys.* 111 (1999) 4790–4799.
[16] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A000045, A001333, A006506, A051736, A051737, A050974, A066863, and A066864.
[17] G. S. Joyce, Exact results for the activity and isothermal compressibility of the hard-hexagon model, *J. Phys. A* 21 (1988) L983–L988; MR 90a:82070.
[18] G. S. Joyce, On the hard hexagon model and the theory of modular functions, *Philos. Trans. Royal Soc. London A* 325 (1988) 643–702; MR 90c:82055.
[19] G. S. Joyce, On the icosahedral equation and the locus of zeros for the grand partition function of the hard hexagon model, *J. Phys. A* 22 (1989) L237–L242; MR 90e:82101.
[20] M. P. Richey and C. A. Tracy, Equation of state and isothermal compressibility for the hard hexagon model in the disordered regime, *J. Phys. A* 20 (1987) L1121–L1126; MR 89b:82121.
[21] C. A. Tracy, L. Grove, and M. F. Newman, Modular properties of the hard hexagon model, *J. Stat. Phys.* 48 (1987) 477; MR 89b:82125.
[22] M. P. Richey and C. A. Tracy, Algorithms for the computation of polynomial relationships for the hard hexagon model, *Nucl. Phys. B* 330 (1990) 681–704; MR 91c:82029.
[23] P. Zimmermann, The hard hexagon entropy constant, unpublished note (1996).
[24] R. J. Baxter, Hard hexagons: Exact solution, *J. Phys. A* 13 (1980) L61–L70; MR 80m:82052.
[25] R. J. Baxter, Rogers-Ramanujan identities in the hard hexagon model, *J. Stat. Phys.* 26 (1981) 427–452; MR 84m:82104.
[26] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982; MR 90b:82001.

[27] R. J. Baxter, Ramanujan's identities in statistical mechanics, *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. 69–84; MR 89h:82043.
[28] G. E. Andrews, The hard-hexagon model and Rogers-Ramanujan type identities, *Proc. Nat. Acad. Sci. USA* 78 (1981) 5290–5292; MR 82m:82005.
[29] G. E. Andrews, *q-Series: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics and Computer Algebra*, Amer. Math. Soc., 1986; MR 88b:11063.
[30] G. E. Andrews, The reasonable and unreasonable effectiveness of number theory in statistical mechanics, *The Unreasonable Effectiveness of Number Theory*, Proc. 1991 Orono conf., ed. S. A. Burr, Amer. Math. Soc., 1992, pp. 21–34; MR 94c:82021.
[31] G. E. Andrews and R. J. Baxter, A motivated proof of the Rogers-Ramanujan identities, *Amer. Math. Monthly* 96 (1989) 401–409; 97 (1990) 214–215; MR 90e:11147 and MR 91j:11086.
[32] D. E. Knuth, Nonattacking kings on a chessboard, unpublished note (1994).
[33] H. S. Wilf, The problem of the kings, *Elec. J. Combin.* 2 (1995) R3; MR 96b:05012.
[34] M. Larsen, The problem of kings, *Elec. J. Combin.* 2 (1995) R18; MR 96m:05011.
[35] I. Rivin, I. Vardi, and P. Zimmermann, The n-queens problem, *Amer. Math. Monthly* 101 (1994) 629–638; MR 95d:05009.
[36] W. Weeks and R. E. Blahut, The capacity and coding gain of certain checkerboard codes, *IEEE Trans. Inform. Theory* 44 (1998) 1193–1203; MR 98m:94061.
[37] S. Forchhammer and J. Justesen, Entropy bounds for constrained two-dimensional random fields, *IEEE Trans. Inform. Theory* 45 (1999) 118–127; MR 99k:94023.
[38] A. Kato and K. Zeger, On the capacity of two-dimensional run-length constrained channels, *IEEE Trans. Inform. Theory* 45 (1999) 1527–1540; MR 2000c:94012.
[39] Z. Nagy and K. Zeger, Capacity bounds for the 3-dimensional (0, 1) runlength limited channel, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (AAECC-13)*, Proc. 1999 Honolulu conf., ed. M. Fossorier, H. Imai, S. Lin, and A. Poli, Lect. Notes in Comp. Sci. 1719, Springer-Verlag, 1999, pp. 245–251.
[40] H. Ito, A. Kato, Z. Nagy, and K. Zeger, Zero capacity region of multidimensional run length constraints, *Elec. J. Combin.* 6 (1999) R33; MR 2000i:94036.
[41] R. M. Roth, P. H. Siegel, and J. K. Wolf, Efficient coding schemes for the hard-square model, *IEEE Trans. Inform. Theory* 47 (2001) 1166–1176; MR 2002b:94017.
[42] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A007197, A027683, A066866, and A067967.
[43] R. Hardin, Computation of $f_{k,n}$ and $g_{k,n}$, $1 \leq n \leq 9$, unpublished note (2002).
[44] D. S. Gaunt, Hard-sphere lattice gases. II: Plane-triangular and three-dimensional lattices, *J. Chem. Phys.* 46 (1967) 3237.
[45] R. J. Baxter and S. K. Tsang, Entropy of hard hexagons, *J. Phys. A* 13 (1980) 1023–1030.
[46] L. K. Runnels, L. L. Combs, and J. P. Salvant, Exact finite method of lattice statistics. II: Honeycomb lattice gas of hard molecules, *J. Chem. Phys.* 47 (1967) 4015–4020.
[47] C. Richard, M. Höffe, J. Hermisson, and M. Baake, Random tilings: Concepts and examples, *J. Phys. A* 31 (1998) 6385–6408; MR 99g:82033.

5.13 Binary Search Tree Constants

We first define a certain function f. The formulation may seem a little abstruse, but f has a natural interpretation as a path length along a type of weakly binary tree (an application of which we will discuss subsequently) [5.6].

Given a vector $V = (v_1, v_2, \ldots, v_k)$ of k distinct integers, define two subvectors V_L and V_R by

$$V_L = (v_j : v_j < v_1,\ 2 \leq j \leq k),\quad V_R = (v_j : v_j > v_1,\ 2 \leq j \leq k).$$

The subscripts L and R mean "left" and "right"; we emphasize that the sublists V_L and V_R preserve the ordering of the elements as listed in V.

Now, over all integers x, define the recursive function

$$f(x,V)=\begin{cases} 0 & \text{if } V=\emptyset \qquad (\emptyset \text{ is the empty vector}),\\ \begin{cases} 1 & \text{if } x=v_1,\\ 1+f(x,V_L) & \text{if } x<v_1,\\ 1+f(x,V_R) & \text{if } x>v_1. \end{cases} & \text{otherwise} \quad (v_1 \text{ is the first vector component}), \end{cases}$$

Clearly $0 \le f(x,V) \le k$ always and the ordering of $v_1, v_2, \ldots, v_k$ is crucial in determining the value of $f(x,V)$. For example, $f(7,(3,9,5,1,7)) = 4$ and $f(4,(3,9,5,1,7)) = 3$.

Let V be a random permutation of $(1, 3, 5, \ldots, 2n-1)$. We are interested in the probability distribution of $f(x,V)$ in two regimes:

- random odd x satisfying $1 \le x \le 2n-1$ (successful search),
- random even x satisfying $0 \le x \le 2n$ (unsuccessful search).

Note that both V and x are random; it is assumed that they are drawn independently with uniform sampling. The expected value of $f(x,V)$ is, in the language of computer science [1–3],

- the average number of comparisons required to *find* an existing random record x in a data structure with n records,
- the average number of comparisons required to *insert* a new random record x into a data structure with n records,

where it is presumed the data structure follows that of a **binary search tree**. Figure 5.15 shows how such a tree is built starting with V as prescribed. Define also $g(l,V) = |\{x : f(x,V) = l,\ 1 \le x \le 2n-1,\ x \text{ odd}\}|$, the number of vertices occupying the l^{th} level of the tree ($l = 1$ is the root level). For example, $g(2,(3,9,5,1,7)) = 2$ and $g(3,(3,9,5,1,7)) = 1$.

$$V = \begin{pmatrix} 3\\ 9\\ 5\\ 1\\ 7 \end{pmatrix}$$

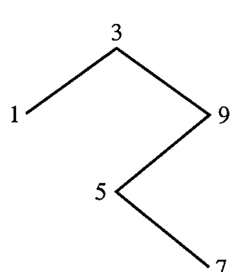


Figure 5.15. Binary search tree constructed using V.

In addition to the two average-case parameters, we want the probability distribution of

$$h(V) = \max\{f(x,V) : 1 \le x \le 2n-1,\ x \text{ odd}\} - 1,$$

the **height** of the tree (which captures the worst-case scenario for finding the record x, given V), and

$$s(V) = \max\{l : g(l,V) = 2^{l-1}\} - 1,$$

the **saturation level** of the tree (which provides the number of full levels of vertices in the tree, minus one). Thus $h(V)$ is the longest path length from the root of the tree to a leaf whereas $s(V)$ is the shortest such path. For example, $h(3,9,5,1,7) = 3$ and $s(3,9,5,1,7) = 1$.

Define, as is customary, the harmonic numbers

$$H_n = \sum_{k=1}^{n} \frac{1}{k} = \ln(n) + \gamma + \frac{1}{n} + O\left(\frac{1}{n^2}\right), \quad H_n^{(2)} = \sum_{k=1}^{n} \frac{1}{k^2} = \frac{\pi^2}{6} - \frac{1}{n} + O\left(\frac{1}{n^2}\right),$$

where γ is the Euler–Mascheroni constant [1.5]. Then the expected number of comparisons in a successful search (random, odd $1 \le x \le 2n-1$) of a random tree is [2–4]

$$\mathrm{E}(f(x,V)) = 2\left(1+\frac{1}{n}\right)H_n - 3 = 2\ln(n) + 2\gamma - 3 + O\left(\frac{\ln(n)}{n}\right),$$

and in an unsuccessful search (random, even $0 \le x \le 2n$) the expected number is

$$\mathrm{E}(f(x,V)) = 2(H_{n+1} - 1) = 2\ln(n) + 2\gamma - 2 + O\left(\frac{1}{n}\right).$$

The corresponding variances are, for odd x,

$$\begin{aligned}\mathrm{Var}(f(x,V)) &= \left(2+\frac{10}{n}\right)H_n - 4\left(1+\frac{1}{n}\right)\left(H_n^{(2)} + \frac{H_n^2}{4}\right) + 4\\ &\sim 2\left(\ln(n) + \gamma - \frac{\pi^3}{3} + 2\right)\end{aligned}$$

and, for even x,

$$\mathrm{Var}(f(x,V)) = 2(H_{n+1} - 2H_{n+1}^{(2)} + 1) \sim 2\left(\ln(n) + \gamma - \frac{\pi^3}{3} + 1\right).$$

A complete analysis of $h(V)$ and $s(V)$ remained unresolved until 1985 when Devroye [3, 5–7], building upon work of Robson [8] and Pittel [9], proved that

$$\frac{h(V)}{\ln(n)} \to c, \quad \frac{s(V)}{\ln(n)} \to d,$$

almost surely as $n \to \infty$, where $c = 4.3110704070\ldots$ and $d = 0.3733646177\ldots$ are the only two real solutions of the equation

$$\frac{2}{x}\exp\left(1 - \frac{1}{x}\right) = 0.$$

Observe that the rate of convergence for $h(V)/\ln(n)$ and $s(V)/\ln(n)$ is slow; hence a numerical verification requires efficient simulation [10]. Considerable effort has been devoted to making these asymptotics more precise [11–14]. Reed [15, 16] and Drmota [17–19] recently proved that

$$\mathrm{E}(h(V)) = c\ln(n) - \frac{3c}{2(c-1)}\ln(\ln(n)) + O(1),$$
$$\mathrm{E}(s(V)) = d\ln(n) + O(\sqrt{\ln(n)}\ln(\ln(n)))$$

and $\mathrm{Var}(h(V)) = O(1)$ as $n \to \infty$. No numerical estimates of the latter are yet available. See also [20].

It is curious that for digital search trees [5.14], which are somewhat more complicated than binary search trees, the analogous limits

$$\frac{h(V)}{\ln(n)} \to \frac{1}{\ln(2)}, \quad \frac{s(V)}{\ln(n)} \to \frac{1}{\ln(2)}$$

do not involve new constants. The fact that limiting values for $h(V)/\ln(n)$ and $s(V)/\ln(n)$ are equal means that the trees are almost perfect (with only a small "fringe" around $\log_2(n)$). This is a hint that search/insertion algorithms on digital search trees are, on average, more efficient than on binary search trees.

Here is one related subject [21–23]. Break a stick of length r into two parts at random. Independently, break each of the two substicks into two parts at random as well. Continue inductively, so that at the end of the n^{th} step, we have 2^n pieces. Let $P_n(r)$ denote the probability that all of the pieces have length < 1. For fixed r, clearly $P_n(r) \to 1$ as $n \to \infty$. More interestingly,

$$\lim_{n\to\infty} P_n(r^n) = \begin{cases} 0 & \text{if } r > e^{1/c}, \\ 1 & \text{if } 0 < r < e^{1/c}, \end{cases}$$

where $e^{1/c} = 1.2610704868\ldots$ and c is as defined earlier. The techniques for proving this are similar to those utilized in [5.3].

We merely mention a generalization of binary search trees called **quadtrees** [24–30], which also possess intriguing asymptotic constants. Quadtrees are useful for storing and retrieving multidimensional real data, for example, in cartography, computer graphics, and image processing [31–33].

[1] R. Sedgewick and P. Flajolet, *Introduction to the Analysis of Algorithms*, Addison-Wesley, 1996, pp. 236–240, 246–250, 261.
[2] D. E. Knuth, *The Art of Computer Programming*, v. 3, *Sorting and Searching*, Addison-Wesley, 1973, pp. 427, 448, 671–672; MR 56 #4281.
[3] H. M. Mahmoud, *Evolution of Random Search Trees*, Wiley 1992, pp. 68–85, 92–95, 177–206; MR 93f:68045.

[4] G. G. Brown and B. O. Shubert, On random binary trees, *Math. Oper. Res.* 9 (1984) 43–65; MR 86c:05099.
[5] L. Devroye, A note on the height of binary search trees, *J. ACM* 33 (1986) 489–498; MR 87i:68009.
[6] L. Devroye, Branching processes in the analysis of the heights of trees, *Acta Inform.* 24 (1987) 277–298; MR 88j:68022.
[7] J. D. Biggins, How fast does a general branching random walk spread?, *Classical and Modern Branching Processes,* Proc. 1994 Minneapolis conf., ed. K. B. Athreya and P. Jagers, Springer-Verlag, 1997, pp. 19–39; MR 99c:60186.
[8] J. M. Robson, The height of binary search trees, *Austral. Comput. J.* 11 (1979) 151–153; MR 81a:68071.
[9] B. Pittel, On growing random binary trees, *J. Math. Anal. Appl.* 103 (1984) 461–480; MR 86c:05101.
[10] L. Devroye and J. M. Robson, On the generation of random binary search trees, *SIAM J. Comput.* 24 (1995) 1141–1156; MR 96j:68040.
[11] L. Devroye and B. Reed, On the variance of the height of random binary search trees, *SIAM J. Comput.* 24 (1995) 1157–1162; MR 96k:68033.
[12] J. M. Robson, On the concentration of the height of binary search trees, *Proc. 1997 Int. Colloq. on Automata, Languages and Programming (ICALP)*, Bologna, ed. P. Degano, R. Gorrieri, and A. Marchettio-Spaccamela, Lect. Notes in Comp. Sci. 1256, Springer-Verlag, pp. 441–448; MR 98m:68047.
[13] J. M. Robson, Constant bounds on the moments of the height of binary search trees, *Theoret. Comput. Sci.* 276 (2002) 435–444.
[14] M. Drmota, An analytic approach to the height of binary search trees, *Algorithmica* 29 (2001) 89–119.
[15] B. Reed, How tall is a tree?, *Proc. 32nd ACM Symp. on Theory of Computing (STOC)*, Portland, ACM, 2000, pp. 479–483.
[16] B. Reed, The height of a random binary search tree, *J. ACM*, to appear.
[17] M. Drmota, The variance of the height of binary search trees, *Theoret. Comput. Sci.* 270 (2002) 913–919.
[18] M. Drmota, An analytic approach to the height of binary search trees. II, *J. ACM*, to appear.
[19] M. Drmota, The saturation level in binary search trees, *Mathematics and Computer Science. Algorithms, Trees, Combinatorics and Probabilities*, Proc. 2000 Versailles conf., ed. D. Gardy and A. Mokkadem, Birkhäuser, 2000, pp. 41–51; MR 2001j:68026.
[20] C. Knessl and W. Szpankowski, The height of a binary search tree: The limiting distribution perspective, *Theoret. Comput. Sci.*, to appear.
[21] M. Sibuya and Y. Itoh, Random sequential bisection and its associated binary tree, *Annals Inst. Statist. Math.* 39 (1987) 69–84; MR 88e:60120.
[22] Y. Itoh, Binary search trees and 1-dimensional random packing, INRIA Project Algorithms Seminar 1997-1998, lecture summary.
[23] T. Hattori and H. Ochiai, A note on the height of binary search trees, random sequential bisection, and successive approximation for differential equation with moving singularity, unpublished note (1998).
[24] L. Laforest, *Étude des arbes hyperquaternaires*, LaCIM Tech. Report 3, Université du Québec à Montréal, 1990.
[25] L. Devroye and L. Laforest, An analysis of random d-dimensional quad trees, *SIAM J. Comput.* 19 (1990) 821–832; MR 91f:68018.
[26] P. Flajolet, G. Gonnet, C. Puech, and J. M. Robson, Analytic variations on quadtrees, *Algorithmica* 10 (1993) 473–500; MR 94i:68052.
[27] P. Flajolet, G. Labelle, L. Laforest, and B. Salvy, Hypergeometrics and the cost structure of quadtrees, *Random Structures Algorithms*, 7 (1995) 117–144; MR 96m:68034.
[28] G. Labelle and L. Laforest, Combinatorial variations on multidimensional quadtrees, *J. Combin. Theory Ser. A* 69 (1995) 1–16; MR 95m:05018.
[29] G. Labelle and L. Laforest, Sur la distribution de l'artié de la racine d'une arborescence hyperquaternaire à d dimensions, *Discrete Math.* 139 (1995) 287–302; MR 96f:05093.

[30] G. Labelle and L. Laforest, Etude de constantes universelles pour les arborescences hyperquaternaires de recherche, *Discrete Math.* 153 (1996) 199–211; MR 97c:05003.
[31] H. Samet, The quadtree and related hierarchical structures, *ACM Comput. Surveys* 16 (1984) 187–206.
[32] H. Samet, *The Design and Analysis of Spatial Data Structures*, Addison-Wesley, 1990.
[33] H. Samet, *Applications of Spatial Data Structures*, Addison-Wesley, 1990.

5.14 Digital Search Tree Constants

Prior acquaintance with binary search trees [5.13] is recommended before reading this essay. Given a binary $k \times n$ matrix $M = (m_{i,j}) = (m_1, m_2, \ldots, m_k)$ of k distinct rows, define two submatrices $M_{L,p}$ and $M_{R,p}$ by

$$M_{L,p} = (m_i : m_{i,p} = 0,\ 2 \le i \le k), \quad M_{R,p} = (m_i : m_{i,p} = 1,\ 2 \le i \le k)$$

for any integer $1 \le p \le n$. That is, the p^{th} column of $M_{L,p}$ is all zeros and the p^{th} column of $M_{R,p}$ is all ones. The subscripts L and R mean "left" and "right"; we emphasize that the sublists $M_{L,p}$ and $M_{R,p}$ preserve the ordering of the rows as listed in M.

Now, over all binary n-vectors x, define the recursive function

$$f(x, M, p) = \begin{cases} 0 & \text{if } M = \emptyset, \\ \begin{cases} 1 & \text{if } x = m_1, \\ 1 + f(x, M_{L,p}, p+1) & \text{if } x \ne m_1 \text{ and } x_p = 0, \\ 1 + f(x, M_{R,p}, p+1) & \text{if } x \ne m_1 \text{ and } x_p = 1. \end{cases} & \text{otherwise,} \end{cases}$$

Clearly $0 \le f(x, M, p) \le k$ always and the ordering of $m_1, m_2, \ldots, m_k$, as well as the value of p, is crucial in determining the value of $f(x, M, p)$.

Let $M = (m_1, m_2, \ldots, m_k)$ be a random binary $n \times n$ matrix with n distinct rows, and let x denote a binary n-vector. We are interested in the probability distribution of $f(x, M, 1)$ in two regimes:

- random x satisfying $x = m_i$ for some i, $1 \le i \le n$ (successful search),
- random x satisfying $x \ne m_i$ for all i, $1 \le i \le n$ (unsuccessful search).

There is double randomness here as with binary search trees [5.13], but note that x depends on M more intricately than before. The expected value of $f(x, M, 1)$ is, in the language of computer science, [1–6]

- the average number of comparisons required to *find* an existing random record x in a data structure with n records,
- the average number of comparisons required to *insert* a new random record x into a data structure with n records,

where it is presumed the data structure follows that of a **digital search tree**. Figure 5.16 shows how such a tree is built starting with M as prescribed.

$$M = \begin{pmatrix} 0\,1\,0\,0\,0\,0 \\ 1\,0\,1\,0\,1\,1 \\ 0\,1\,0\,1\,1\,0 \\ 0\,1\,1\,0\,1\,1 \\ 1\,1\,0\,1\,0\,0 \\ 1\,0\,0\,0\,1\,1 \end{pmatrix}$$

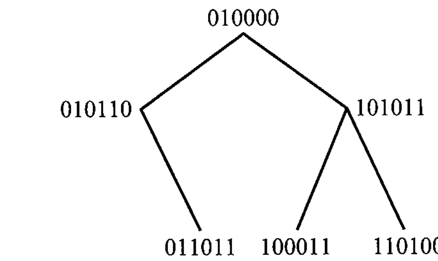


Figure 5.16. Digital search tree constructed using M.

Another parameter of some interest is the number A_n of non-root vertices of degree 1, that is, nodes without children. For binary search trees [3, 7], it is known that $\mathrm{E}(A_n) = (n+1)/3$. For digital search trees, the corresponding result is more complicated, as we shall soon see. Because digital search trees are usually better "balanced" than binary search trees, one anticipates a linear coefficient closer to $1/2$ than $1/3$.

Let γ denote the Euler–Mascheroni constant [1.5] and define a new constant

$$\alpha = \sum_{k=1}^{\infty} \frac{1}{2^k - 1} = 1.6066951524\ldots.$$

Then the expected number of comparisons in a successful search (random, $x = m_i$ for some i) of a random tree is [3–6, 8, 9]

$$\begin{aligned} \mathrm{E}(f(x, M, 1)) &= \frac{1}{\ln(2)}\ln(n) + \frac{3}{2} + \frac{\gamma - 1}{\ln(2)} - \alpha + \delta(n) + O\left(\frac{\ln(n)}{n}\right) \\ &\sim \log_2(n) - 0.716644\ldots + \delta(n), \end{aligned}$$

and in an unsuccessful search (random $x \ne m_i$ for all i) the expected number is

$$\begin{aligned} \mathrm{E}(f(x, M, 1)) &= \frac{1}{\ln(2)}\ln(n) + \frac{1}{2} + \frac{\gamma}{\ln(2)} - \alpha + \delta(n) + O\left(\frac{\ln(n)}{n}\right) \\ &\sim \log_2(n) - 0.273948\ldots + \delta(n), \end{aligned}$$

where

$$\delta(n) = \frac{1}{\ln(2)} \sum_{\substack{k=-\infty \\ k \ne 0}}^{\infty} \Gamma\left(-1 - \frac{2\pi i k}{\ln(2)}\right) \exp\left(2\pi i k \frac{\ln(n)}{\ln(2)}\right).$$

The function $\delta(n)$ is oscillatory ($\delta(n) = \delta(2n)$), has zero mean, and is "negligible" ($|\delta(n)| < 1.726 \times 10^{-7}$ for all n). Similar functions $\varepsilon(n)$, $\rho(n)$, $\sigma(n)$ and $\tau(n)$ will be needed later. These arise in the analysis of many algorithms [3, 4, 6], as well as in problems discussed in [2.3], [2.16], [5.6], and [5.11]. Although such functions can be safely ignored for practical purposes, they need to be included in certain treatments for the sake of theoretical rigor.

The corresponding variances are, for searching,

$$\mathrm{Var}(f(x,M,1)) \sim \frac{1}{12} + \frac{\pi^2+6}{6\ln(2)^2} - \alpha - \beta + \varepsilon(n) \sim 2.844383\ldots + \varepsilon(n)$$

and, for inserting,

$$\mathrm{Var}(f(x,M,1)) \sim \frac{1}{12} + \frac{\pi^2}{6\ln(2)^2} - \alpha - \beta + \varepsilon(n) \sim 0.763014\ldots + \varepsilon(n),$$

where the new constant β is given by

$$\beta = \sum_{k=1}^{\infty} \frac{1}{(2^k-1)^2} = 1.1373387363\ldots.$$

Flajolet & Sedgewick [3, 8, 10] answered an open question of Knuth's regarding the parameter A_n:

$$\mathrm{E}(A_n) = \left[\theta + 1 - \frac{1}{Q}\left(\frac{1}{\ln(2)} + \alpha^2 - \alpha\right) + \rho(n)\right] n + O(n^{1/2}),$$

where the new constants Q and θ are given by

$$Q = \prod_{k=1}^{\infty}\left(1 - \frac{1}{2^k}\right) = 0.2887880950\ldots = (3.4627466194\ldots)^{-1},$$

$$\theta = \sum_{k=1}^{\infty} \frac{k2^{k+1}}{1\cdot 3\cdot 7\cdots(2^k-1)} \sum_{j=1}^{k} \frac{1}{2^j-1} = 7.7431319855\ldots.$$

The linear coefficient of $\mathrm{E}(A_n)$ fluctuates around

$$c = \theta + 1 - \frac{1}{Q}\left(\frac{1}{\ln(2)} + \alpha^2 - \alpha\right) = 0.3720486812\ldots,$$

which is not as close to $1/2$ as one might have anticipated! Here also [11] is an integral representation for c:

$$c = \frac{1}{\ln(2)} \int_0^{\infty} \frac{x}{1+x}\left(1+\frac{x}{1}\right)^{-1}\left(1+\frac{x}{2}\right)^{-1}\left(1+\frac{x}{4}\right)^{-1}\left(1+\frac{x}{8}\right)^{-1}\cdots dx.$$

There are three main types of m-ary search trees: digital search trees, radix search tries (tries), and Patricia tries. We have assumed that $m = 2$ throughout. What, for example, is the variance for searching corresponding to Patricia tries? If we omit the fluctuation term, the remaining coefficient

$$\nu = \frac{1}{12} + \frac{\pi^2}{6\ln(2)^2} + \frac{2}{\ln(2)} \sum_{k=1}^{\infty} \frac{(-1)^k}{k(2^k-1)}$$

is interesting because, at first glance, it seems to be exactly 1! In fact, $\nu > 1 + 10^{-12}$ and this can be more carefully explained via the Dedekind eta function [12, 13].

5.14.1 Other Connections

In number theory, the divisor function $d(n)$ is the number of integers d, $1 \le d \le n$, that divide n. A special value of its generating function [4, 14, 15]

$$\sum_{n=1}^{\infty} d(n)q^n = \sum_{k=1}^{\infty} \frac{q^k}{1-q^k} = \sum_{k=1}^{\infty} \frac{q^{k^2}(1+q^k)}{1-q^k}$$

is α when $q = 1/2$. Erdös [16, 17] proved that α is irrational; forty years passed while people wondered about constants such as

$$\sum_{n=1}^{\infty} \frac{1}{2^n-3} \text{ and } \sum_{n=1}^{\infty} \frac{1}{2^n+1}$$

(the former appears in [18] whereas the latter is connected to tries [6] and mergesort asymptotics [19, 20]). Borwein [21, 22] proved that, if $|a| \ge 2$ is an integer, $b \ne 0$ is a rational number, and $b \ne -a^n$ for all n, then the series

$$\sum_{n=1}^{\infty} \frac{1}{a^n+b} \text{ and } \sum_{n=1}^{\infty} \frac{(-1)^n}{a^n+b}$$

are both irrational. Under the same conditions, the product

$$\prod_{n=1}^{\infty}\left(1 + \frac{b}{a^n}\right)$$

is irrational [23, 24], and hence so is Q. See [25] for recent computer-aided irrationality proofs.

On the one hand, from the combinatorics of integer partitions, we have Euler's pentagonal number theorem [14, 26–28]

$$\prod_{n=1}^{\infty}(1-q^n) = \sum_{n=-\infty}^{\infty}(-1)^n q^{\frac{1}{2}(3n+1)n} = 1 + \sum_{n=1}^{\infty}(-1)^n\left(q^{\frac{1}{2}(3n-1)n} + q^{\frac{1}{2}(3n+1)n}\right)$$

and

$$\begin{aligned}\prod_{n=1}^{\infty}(1-q^n)^{-1} &= 1 + \sum_{n=1}^{\infty} \frac{q^n}{(1-q)(1-q^2)(1-q^3)\cdots(1-q^n)} \\ &= 1 + \sum_{n=1}^{\infty} \frac{q^{n^2}}{(1-q)^2(1-q^2)^2(1-q^3)^2\cdots(1-q^n)^2} = 1 + \sum_{n=1}^{\infty} p(n)q^n,\end{aligned}$$

where $p(n)$ denotes the number of unrestricted partitions of n. If $q = 1/2$, these specialize to Q and $1/Q$. On the other hand, in the theory of finite vector spaces, Q appears in the asymptotic formula [5.7] for the number of linear subspaces of $\mathbb{F}_{q,n}$ when $q = 2$.

A substantial theory has emerged involving q-analogs of various classical mathematical objects. For example, the constant α is regarded as a $1/2$-analog of the Euler–Mascheroni constant [11]. Other constants (e.g., Apéry's constant $\zeta(3)$ or Catalan's constant G) can be similarly generalized.

Out of many more possible formulas, we mention three [4, 14, 26, 29]:

$$Q = \frac{1}{3} - \frac{1}{3\cdot 7} + \frac{1}{3\cdot 7\cdot 15} - \frac{1}{3\cdot 7\cdot 15\cdot 31} + - \cdots$$
$$= \exp\left(-\sum_{n=1}^{\infty}\frac{1}{n(2^n-1)}\right)$$
$$= \sqrt{\frac{2\pi}{\ln(2)}}\exp\left(\frac{\ln(2)}{24} - \frac{\pi^2}{6\ln(2)}\right)\prod_{n=1}^{\infty}\left[1-\exp\left(\frac{-4\pi^2 n}{\ln(2)}\right)\right].$$

The second makes one wonder if a simple relationship between Q and α exists. It can be shown that Q is the asymptotic probability that the determinant of a random $n \times n$ binary matrix is odd. A constant P similar to Q appears in [2.8]; exponents in P are constrained to be odd integers.

The reciprocal sum of repunits [30]

$$9\sum_{n=1}^{\infty}\frac{1}{10^n-1} = \frac{1}{1} + \frac{1}{11} + \frac{1}{111} + \frac{1}{1111} + \cdots = 1.1009181908\ldots$$

is irrational by Borwein's theorem. The reciprocal series of Fibonacci numbers can be expressed as [31–33]

$$\sum_{k=1}^{\infty}\frac{1}{f_k} = \sqrt{5}\sum_{n=0}^{\infty}\frac{(-1)^n}{\varphi^{2n+1}-(-1)^n} = 3.3598856662\ldots,$$

where φ is the Golden mean, and this sum is known to be irrational [34–37]. Note that the subseries of terms with even subscripts can similarly be evaluated [26, 31]:

$$\sum_{k=1}^{\infty}\frac{1}{f_{2k}} = \sqrt{5}\left(\sum_{n=1}^{\infty}\frac{1}{\lambda^n-1} - \sum_{n=1}^{\infty}\frac{1}{\mu^n-1}\right) = 1.5353705088\ldots,$$

where $2\lambda = \sqrt{3}+5$ and $2\mu = 7+3\sqrt{5}$. A completely different connection to the Fibonacci numbers (this time resembling the constant Q) is found in [1.2].

A certain normalizing constant [38–40]

$$K = \sqrt{\prod_{n=0}^{\infty}\left(1+\frac{1}{2^{2n}}\right)} = 1.6467602581\ldots$$

occurs in efficient *binary cordic* implementations of two-dimensional vector rotation. Products such as Q and K, however, have no known closed-form expression except when $q = \exp(-\pi\xi)$, where $\xi > 0$ is an algebraic number [26, 41].

Observe that $2^{n+1}-1$ is the smallest positive integer not representable as a sum of n integers of the form 2^i, $i \ge 0$. Define h_n to be the smallest positive integer not representable as a sum of n integers of the form $2^i 3^j$, $i \ge 0$, $j \ge 0$, that is, $h_0 = 1$, $h_1 = 5$, $h_2 = 23$, $h_3 = 431, \ldots$ [42, 43]. What is the precise growth rate of h_n as $n \to \infty$? What is the numerical value of the reciprocal sum of h_n (what might be called the *2-3 analog* of the constant α)? This is vaguely related to our discussion in [2.26] and [2.30.1].

5.14.2 Approximate Counting

Returning to computer science, we discuss **approximate counting**, an algorithm due to Morris [44]. Approximate counting involves keeping track of a large number, N, of events in only $\log_2(\log_2(N))$ bit storage, where accuracy is not paramount. Consider the integer time series $X_0, X_1, \ldots, X_N$ defined recursively by

$$X_n = \begin{cases} 1 & \text{if } n = 0,\\ \begin{cases} 1 + X_{n-1} & \text{with probability } 2^{-X_{n-1}},\\ X_{n-1} & \text{with probability } 1 - 2^{-X_{n-1}}. \end{cases} & \text{otherwise,} \end{cases}$$

It is not hard to prove that

$$\mathrm{E}(2^{X_N} - 2) = N \text{ and } \mathrm{Var}(2^{X_N}) = \tfrac{1}{2}N(N+1);$$

hence probabilistic updates via this scheme give an unbiased estimator of N. Flajolet [45–50] studied the distribution of X_N in much greater detail:

$$\mathrm{E}(X_N) = \frac{1}{\ln(2)}\ln(N) + \frac{1}{2} + \frac{\gamma}{\ln(2)} - \alpha + \sigma(n) + O\left(\frac{\ln(N)}{N}\right)$$
$$\sim \log_2(N) - 0.273948\ldots + \sigma(N),$$

$$\mathrm{Var}(X_N) \sim \frac{1}{12} + \frac{\pi^2}{6\ln(2)^2} - \alpha - \beta - \chi + \tau(n) \sim 0.763014\ldots + \tau(N),$$

where α and β are as before, the new constant χ is given by

$$\chi = \frac{1}{\ln(2)}\sum_{n=1}^{\infty}\frac{1}{n}\operatorname{csch}\left(\frac{2\pi^2 n}{\ln(2)}\right) = (1.237412\ldots)\times 10^{-12},$$

and $\sigma(n)$ and $\tau(n)$ are oscillatory "negligible" functions. In particular, since $\chi > 0$, the constant coefficient for $\mathrm{Var}(X_N)$ is (slightly) smaller than that for $\mathrm{Var}(f(x, M, 1))$ given earlier. Similar ideas in probabilistic counting algorithms are found in [6.8].

[1] E. G. Coffman and J. Eve, File structures using hashing functions, *Commun. ACM* 13 (1970) 427–432.

[2] A. G. Konheim and D. J. Newman, A note on growing binary trees, *Discrete Math.* 4 (1973) 57–63; MR 47 #1650.

[3] P. Flajolet and R. Sedgewick, Digital search trees revisited, *SIAM Rev.* 15 (1986) 748–767; MR 87m:68014.

[4] D. E. Knuth, *The Art of Computer Programming*, v. 3, *Sorting and Searching*, Addison-Wesley, 1973, pp. 21, 134, 156, 493–502, 580, 685–686; MR 56 #4281.

[5] G. Louchard, Exact and asymptotic distributions in digital and binary search trees, *RAIRO Inform. Théor. Appl.* 21 (1987) 479–495; MR 89h:68031.

[6] H. M. Mahmoud, *Evolution of Random Search Trees*, Wiley, 1992, pp. 226–227, 260–291; MR 93f:68045.

[7] G. G. Brown and B. O. Shubert, On random binary trees, *Math. Oper. Res.* 9 (1984) 43–65; MR 86c:05099.

[8] P. Kirschenhofer and H. Prodinger, Further results on digital search trees, *Theoret. Comput. Sci.* 58 (1988) 143–154; MR 89j:68022.

[9] P. Kirschenhofer, H. Prodinger, and W. Szpankowski, Digital search trees again revisited: The internal path length perspective, *SIAM J. Comput.* 23 (1994) 598–616; MR 95i:68034.

[10] H. Prodinger, External internal nodes in digital search trees via Mellin transforms, *SIAM J. Comput.* 21 (1992) 1180–1183; MR 93i:68047.

[11] P. Flajolet and B. Richmond, Generalized digital trees and their difference-differential equations, *Random Structures Algorithms* 3 (1992) 305–320; MR 93f:05086.

[12] P. Kirschenhofer and H. Prodinger, Asymptotische Untersuchungen über charakteristische Parameter von Suchbäumen, *Zahlentheoretische Analysis II*, ed. E. Hlawka, Lect. Notes in Math. 1262, Springer-Verlag, 1987, pp. 93–107; MR 90h:68028.

[13] P. Kirschenhofer, H. Prodinger, and J. Schoissengeier, Zur Auswertung gewisser numerischer Reihen mit Hilfe modularer Funktionen, *Zahlentheoretische Analysis II*, ed. E. Hlawka, Lect. Notes in Math. 1262, Springer-Verlag, 1987, pp. 108–110; MR 90g:11057.

[14] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford Univ. Press, 1985; Thms. 310, 350, 351, 353; MR 81i:10002.

[15] B. C. Berndt, *Ramanujan's Notebooks: Part I*, Springer-Verlag, 1985, p. 147; MR 86c:01062.

[16] P. Erdös, On arithmetical properties of Lambert series, *J. Indian Math. Soc.* 12 (1948) 63–66; MR 10,594c.

[17] D. H. Bailey and R. E. Crandall, Random generators and normal numbers, Perfectly Scientific Inc. preprint (2001).

[18] P. Erdös and R. L. Graham, *Old and New Problems and Results in Combinatorial Number Theory*, Enseignement Math. Monogr. 28, 1980, p. 62; MR 82j:10001.

[19] P. Flajolet and M. Golin, Exact asymptotics of divide-and-conquer recurrences, *Proc. 1993 Int. Colloq. on Automata, Languages and Programming (ICALP)*, Lund, ed. A. Lingas, R. Karlsson, and S. Carlsson, Lect. Notes in Comp. Sci. 700, Springer-Verlag, 1993, pp. 137–149.

[20] P. Flajolet and M. Golin, Mellin transforms and asymptotics: The mergesort recurrence, *Acta Inform.* 31 (1994) 673–696; MR 95h:68035.

[21] P. B. Borwein, On the irrationality of $\sum(1/(q^n + r))$, *J. Number Theory* 37 (1991) 253–259; MR 92b:11046.

[22] P. B. Borwein, On the irrationality of certain series, *Math. Proc. Cambridge Philos. Soc.* 112 (1992) 141–146; MR 93g:11074.

[23] J. Lynch, J. Mycielski, P. A. Vojta, and P. Bundschuh, Irrationality of an infinite product, *Amer. Math. Monthly* 87 (1980) 408–409.

[24] R. Wallisser, Rational approximation of the q-analogue of the exponential function and irrationality statements for this function, *Arch. Math. (Basel)* 44 (1985) 59–64; MR 86i:11036.

[25] T. Amdeberhan and D. Zeilberger, q-Apéry irrationality proofs by q-WZ pairs, *Adv. Appl. Math.* 20 (1998) 275–283; MR 99d:11074.

[26] J. M. Borwein and P. B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, Wiley, 1987, pp. 62–75, 91–101; MR 99h:11147.

[27] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Cambridge Univ. Press, 1990; MR 91d:33034.

[28] G. E. Andrews, *q-Series: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics and Computer Algebra*, CBMS, v. 66, Amer. Math. Soc., 1986; MR 88b:11063.

[29] R. J. McIntosh, Some asymptotic formulae for q-hypergeometric series, *J. London Math. Soc.* 51 (1995) 120–136; MR 95m:11112.

[30] A. H. Beiler, *Recreations in the Theory of Numbers*, Dover, 1966.

[31] A. F. Horadam, Elliptic functions and Lambert series in the summation of reciprocals in certain recurrence-generated sequences, *Fibonacci Quart.* 26 (1988) 98–114; MR 89e:11013.

[32] P. Griffin, Acceleration of the sum of Fibonacci reciprocals, *Fibonacci Quart.* 30 (1992) 179–181; MR 93d:11024.

[33] F.-Z. Zhao, Notes on reciprocal series related to Fibonacci and Lucas numbers, *Fibonacci Quart.* 37 (1999) 254–257; MR 2000d:11020.

[34] R. André-Jeannin, Irrationalité de la somme des inverses de certaines suites récurrentes, *C. R. Acad. Sci. Paris Sér. I Math.* 308 (1989) 539–541; MR 90b:11012.

[35] P. Bundschuh and K. Väänänen, Arithmetical investigations of a certain infinite product, *Compositio Math.* 91 (1994) 175–199; MR 95e:11081.

[36] D. Duverney, Irrationalité de la somme des inverses de la suite de Fibonacci, *Elem. Math.* 52 (1997) 31–36; MR 98c:11069.

[37] M. Prevost, On the irrationality of $\sum t^n/(A\alpha^n + B\beta^n)$, *J. Number Theory* 73 (1998) 139–161; MR 2000b:11085.

[38] J. E. Volder, The CORDIC trigonometric computing technique, *IRE Trans. Elec. Comput.* EC-8 (1959) 330–334; also in *Computer Arithmetic*, v. 1, ed. E. E. Swartzlander, IEEE Computer Soc. Press, 1990, pp. 226–230.

[39] J. S. Walther, A unified algorithm for elementary functions, *Spring Joint Computer Conf. Proc.* (1971) 379–385; also in *Computer Arithmetic*, v. 1, ed. E. E. Swartzlander, IEEE Computer Soc. Press, 1990, pp. 272–278.

[40] H. G. Baker, Complex Gaussian integers for 'Gaussian graphics,' *ACM Sigplan Notices*, v. 28 (1993) n. 11, 22–27.

[41] R. W. Gosper, Closed forms for $\prod_{n=1}^{\infty}(1 + \exp(-\pi\xi n))$, unpublished note (1998).

[42] V. S. Dimitrov, G. A. Jullien, and W. C. Miller, Theory and applications for a double-base number system, *IEEE Trans. Comput.* 48 (1999) 1098–1106; also in *Proc. 13th IEEE Symp. on Computer Arithmetic (ARITH)*, Asilomar, 1997, ed. T. Lang, J.-M. Muller, and N. Takagi, IEEE, 1997, pp. 44–51.

[43] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A018899.

[44] R. Morris, Counting large numbers of events in small registers, *Commun. ACM* 21 (1978) 840–842.

[45] P. Flajolet, Approximate counting: A detailed analysis, *BIT* 25 (1985) 113–134; MR 86j:68053.

[46] P. Kirschenhofer and H. Prodinger, Approximate counting: An alternative approach, *RAIRO Inform. Théor. Appl.* 25 (1991) 43–48; MR 92e:68070.

[47] H. Prodinger, Hypothetical analyses: Approximate counting in the style of Knuth, path length in the style of Flajolet, *Theoret. Comput. Sci.* 100 (1992) 243–251; MR 93j:68076.

[48] P. Kirschenhofer and H. Prodinger, A coin tossing algorithm for counting large numbers of events, *Math. Slovaca* 42 (1992) 531–545; MR 93j:68073.

[49] H. Prodinger, Approximate counting via Euler transform, *Math. Slovaca* 44 (1994) 569–574; MR 96h:11013.

[50] P. Kirschenhofer, A note on alternating sums, *Elec. J. Combin.* 3 (1996) R7; MR 97e:05025.

5.15 Optimal Stopping Constants

Consider the well-known **secretary problem**. An unordered sequence of **applicants** (distinct real numbers) $s_1, s_2, \ldots, s_n$ are interviewed by you one at a time. You have no prior information about the ss. You know the value of n, and as s_k is being interviewed, you must either accept s_k and end the process, or reject s_k and interview s_{k+1}. The decision to accept or reject s_k must be based solely on whether $s_k > s_j$ for all $1 \le j < k$ (that is, on whether s_k is a **candidate**). An applicant once rejected cannot later be recalled.

If your objective is to select the most highly qualified applicant (the largest s_k), then the optimal strategy is to reject the first $m - 1$ applicants and accept the next candidate, where [1–4]

$$m = \min\left\{k \ge 1 : \sum_{j=k+1}^{n} \frac{1}{j-1} \le 1\right\} \sim \frac{n}{e}$$

as $n \to \infty$. The asymptotic probability of obtaining the best applicant via this strategy is hence $1/e = 0.3678794411\ldots$, where e is the natural logarithmic base [1.3]. See a generalization of this in [5–7].

If your objective is instead to minimize the expected rank R_n of the chosen applicant (the largest s_k has rank 1, the second-largest has rank 2, etc.), then different formulation applies. Lindley [8] and Chow et al. [9] derived the optimal strategy in this case and proved that [10]

$$\lim_{n\to\infty} R_n = \prod_{k=1}^{\infty}\left(1+\frac{2}{k}\right)^{\frac{1}{k+1}} = 3.8695192413\ldots = C.$$

A variation might include you knowing in advance that $s_1, s_2, \ldots, s_n$ are independent, uniformly distributed variables on the interval $[0, 1]$. This is known as a **full-information problem** (as opposed to the no-information problems just discussed). How does knowledge of the distribution improve your chances of success? For the "nothing but the best" objective, Gilbert & Mosteller [11] calculated the asymptotic probability of success to be [12, 13]

$$e^{-a} - (e^a - a - 1)\,\mathrm{Ei}(-a) = 0.5801642239\ldots,$$

where $a = 0.8043522628\ldots$ is the unique real solution of the equation $\mathrm{Ei}(a) - \gamma - \ln(a) = 1$, Ei is the exponential integral [6.2], and γ is the Euler-Mascheroni constant [1.5].

The full-information analog for $\lim_{n\to\infty} R_n$ appears to be an open problem [14–16]. Yet another objective, however, might be to maximize the hiree's expected quality Q_n itself (the k^{th} applicant has quality s_k). Clearly

$$Q_0 = 0, \quad Q_n = \tfrac{1}{2}(1 + Q_{n-1}^2) \quad \text{if } n \geq 1,$$

and $Q_n \to 1$ as $n \to \infty$. Moser [11, 17–19] deduced that

$$Q_n \sim 1 - \frac{2}{n + \ln(n) + b},$$

where the constant b is estimated [10] to be $1.76799378\ldots$.

Here is a closely related problem. Assume $s_1, s_2, \ldots, s_n$ are independent, uniformly distributed variables on the interval $[0, N]$. Your objective is to minimize the number T_N of interviews necessary to select an applicant of expected quality $\geq N - 1$. Gum [20] sketched a proof that $T_N = 2N - O(\ln(N))$ as $N \to \infty$. Alternatively, assume everything as before except that $s_1', s_2', \ldots, s_n'$ are drawn with replacement from the set $\{1, 2, \ldots, N\}$. It can be proved here that $T_N' = cN + O(\sqrt{N})$, where [10]

$$c = 2\sum_{k=3}^{\infty}\frac{\ln(k)}{k^2-1} - \frac{\ln(2)}{3} = 1.3531302722\ldots = \ln(C).$$

The secretary problem and its offshoots fall within the theory of **optimal stopping** [19]. Here is a sample exercise: We observe a fair coin being tossed repeatedly and can stop observing at any time. When we stop, the payoff is the average number of heads observed. What is the best strategy to maximize the expected payoff? Chow & Robbins [21, 22] described a strategy that achieves an expected payoff $> 0.79 = (0.59 + 1)/2$.

[1] T. S. Ferguson, Who solved the secretary problem?, *Statist. Sci.* 4 (1989) 282–296; MR 91k:01011.
[2] P. R. Freeman, The secretary problem and its extensions: A review, *Int. Statist. Rev.* 51 (1983) 189–206; MR 84k:62115.
[3] F. Mosteller, *Fifty Challenging Problems in Probability with Solutions*, Addison-Wesley, 1965, pp. 12, 74–79; MR 53 #1666.
[4] M. Gardner, *New Mathematical Diversions from Scientific American*, Simon and Schuster, 1966, pp. 35–36, 41–43.
[5] A. Q. Frank and S. M. Samuels, On an optimal stopping problem of Gusein-Zade, *Stochastic Process. Appl.* 10 (1980) 299–311; MR 83f:60067.
[6] M. P. Quine and J. S. Law, Exact results for a secretary problem, *J. Appl. Probab.* 33 (1996) 630–639; MR 97k:60127.
[7] G. F. Yeo, Duration of a secretary problem, *J. Appl. Probab.* 34 (1997) 556–558; MR 98a:60056.
[8] D. V. Lindley, Dynamic programming and decision theory, *Appl. Statist.* 10 (1961) 39–51; MR 23 #A740.
[9] Y. S. Chow, S. Moriguti, H. Robbins, and S. M. Samuels, Optimal selection based on relative rank (the "secretary problem"), *Israel J. Math.* 2 (1964) 81–90; MR 31 #855.
[10] P. Sebah, Computations of several optimal stopping constants, unpublished note (2001).
[11] J. P. Gilbert and F. Mosteller, Recognizing the maximum of a sequence, *J. Amer. Statist. Assoc.* 61 (1966) 35–73; MR 33 #6792.
[12] S. M. Samuels, Exact solutions for the full information best choice problem, Purdue Univ. Stat. Dept. report 82–17 (1982).
[13] S. M. Samuels, Secretary problems, in *Handbook of Sequential Analysis*, ed. B. K. Ghosh and P. K. Sen, Dekker, 1991, pp. 381–405; MR 93g:62102.
[14] F. T. Bruss and T. S. Ferguson, Minimizing the expected rank with full information, *J. Appl. Probab.* 30 (1993) 616–626; MR 94m:60090.
[15] F. T. Bruss and T. S. Ferguson, Half-prophets and Robbins' problem of minimizing the expected rank, *Proc. 1995 Athens Conf. on Applied Probability and Time Series Analysis*, v. 1, ed. C. C. Heyde, Yu. V. Prohorov, R. Pyke, and S. T. Rachev, Lect. Notes in Statist. 114, Springer-Verlag, 1996, pp. 1–17; MR 98k:60066.
[16] D. Assaf and E. Samuel-Cahn, The secretary problem: Minimizing the expected rank with i.i.d. random variables, *Adv. Appl. Probab.* 28 (1996) 828–852; MR 97f:60089.
[17] L. Moser, On a problem of Cayley, *Scripta Math.* 22 (1956) 289–292.
[18] I. Guttman, On a problem of L. Moser, *Canad. Math. Bull.* 3 (1960) 35–39; MR 23 #B2064.
[19] T. S. Ferguson, *Optimal Stopping and Applications*, unpublished manuscript (2000).
[20] B. Gum, The secretary problem with a uniform distribution, presentation at *Tenth SIAM Conf. on Discrete Math.*, Minneapolis, 2000.
[21] Y. S. Chow and H. Robbins, On optimal stopping rules for s_n/n, *Illinois J. Math.* 9 (1965) 444–454; MR 31 #4134.
[22] T. S. Ferguson, *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, 1967, p. 314; MR 35 #6231.

5.16 Extreme Value Constants

Let $X_1, X_2, \ldots, X_n$ denote a random sample from a population with continuous probability density function $f(x)$. Many interesting results exist concerning the distribution of the **order statistics**

$$X^{(1)} < X^{(2)} < \cdots < X^{(n)},$$

where $X^{(1)} = \min\{X_1, X_2, \dots, X_n\} = m_n$ and $X^{(n)} = \max\{X_1, X_2, \dots, X_n\} = M_n$. We will focus only on the extreme values M_n for brevity's sake.

If $X_1, X_2, \dots, X_n$ are taken from a Uniform [0, 1] distribution (i.e., $f(x)$ is 1 for $0 \le x \le 1$ and is 0 otherwise), then the probability distribution of M_n is prescribed by

$$\mathrm{P}(M_n < x) = \begin{cases} 0 & \text{if } x < 0, \\ x^n & \text{if } 0 \le x \le 1, \\ 1 & \text{if } x > 1 \end{cases}$$

and its moments are given by

$$\mu_n = \mathrm{E}(M_n) = \frac{n}{n+1}, \quad \sigma_n^2 = \mathrm{Var}(M_n) = \frac{n}{(n+1)^2(n+2)}.$$

These are all exact results [1–3]. Note that clearly

$$\lim_{n\to\infty} \mathrm{P}\left(n\left(M_n - 1\right) < y\right) = \lim_{n\to\infty} \mathrm{P}\left(M_n < 1 + \frac{1}{n}y\right) = \begin{cases} e^y & \text{if } y < 0, \\ 1 & \text{if } y \ge 0. \end{cases}$$

This asymptotic result is a special case of a far more general theorem due to Fisher & Tippett [4] and Gnedenko [5]. Under broad circumstances, the asymptotic distribution of M_n (suitably normalized) must belong to one of just three possible families. We see another, less trivial, example in the following.

If $X_1, X_2, \dots, X_n$ are from a Normal (0, 1) distribution, that is,

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \quad F(x) = \int_{-\infty}^{x} f(\xi)\, d\xi = \frac{1}{2}\operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2},$$

then the probability distribution of M_n is prescribed by

$$\mathrm{P}(M_n < x) = F(x)^n = n \int_{-\infty}^{x} F(\xi)^{n-1} f(\xi)\, d\xi$$

and its moments are given by

$$\mu_n = n \int_{-\infty}^{\infty} x F(x)^{n-1} f(x)\, dx, \quad \sigma_n^2 = n \int_{-\infty}^{\infty} x^2 F(x)^{n-1} f(x)\, dx - \mu_n^2.$$

For small n, exact expressions are possible [2, 3, 6–11]:

$$\begin{aligned}
\mu_2 &= \tfrac{1}{\sqrt{\pi}} = 0.564\dots, & \sigma_2^2 &= 1 - \mu_2^2 = 0.681\dots, \\
\mu_3 &= \tfrac{3}{2\sqrt{\pi}} = 0.846\dots, & \sigma_3^2 &= 1 + \tfrac{\sqrt{3}}{2\pi} - \mu_3^2 = 0.559\dots, \\
\mu_4 &= \tfrac{3}{\sqrt{\pi}}(1 - 2S_2) = 1.029\dots, & \sigma_4^2 &= 1 + \tfrac{\sqrt{3}}{\pi} - \mu_4^2 = 0.491\dots, \\
\mu_5 &= \tfrac{5}{\sqrt{\pi}}(1 - 3S_2) = 1.162\dots, & \sigma_5^2 &= 1 + \tfrac{5\sqrt{3}}{2\pi}(1 - 2S_3) - \mu_5^2 = 0.447\dots, \\
\mu_6 &= \tfrac{15}{2\sqrt{\pi}}(1 - 4S_2 + 2T_2) = 1.267\dots, & \sigma_6^2 &= 1 + \tfrac{5\sqrt{3}}{\pi}(1 - 3S_3) - \mu_6^2 = 0.415\dots, \\
\mu_7 &= \tfrac{21}{2\sqrt{\pi}}(1 - 5S_2 + 5T_2) = 1.352\dots, & \sigma_7^2 &= 1 + \tfrac{35\sqrt{3}}{4\pi}(1 - 4S_3 + 2T_3) - \mu_7^2 \\
& & &= 0.391\dots,
\end{aligned}$$

where

$$S_k = \frac{\sqrt{k}}{\pi} \int_0^{\frac{\pi}{4}} \frac{dx}{\sqrt{k + \sec(x)^2}} = \frac{1}{\pi} \arcsin\sqrt{\frac{k}{2(1+k)}},$$

$$T_k = \frac{\sqrt{k}}{\pi^2} \int_0^{\frac{\pi}{4}} \int_0^{\frac{\pi}{4}} \frac{dx\, dy}{\sqrt{k + \sec(x)^2 + \sec(y)^2}} = \frac{1}{\pi^2} \int_0^{\pi S(k)} \arcsin\sqrt{\frac{1}{2}\frac{k(k+1)}{k(k+2) - \tan(z)^2}}\, dz.$$

Similar expressions for $\mu_8 = 1.423\dots$ and $\sigma_8^2 = 0.372\dots$ remain to be found. Ruben [12] demonstrated a connection between moments of order statistics and volumes of certain hyperspherical simplices (generalized spherical triangles). Calkin [13] discovered a binomial identity that, in a limiting case, yields the exact expression for μ_3.

We turn now to the asymptotic distribution of M_n. Let

$$a_n = \sqrt{2\ln(n)} - \frac{1}{2}\frac{\ln(\ln(n)) + \ln(4\pi)}{\sqrt{2\ln(n)}}.$$

It can be proved [14–18] that

$$\lim_{n\to\infty} \mathrm{P}\left(\sqrt{2\ln(n)}(M_n - a_n) < y\right) = \exp(-e^{-y}),$$

and the resulting doubly exponential density function $g(y) = \exp(-y - e^{-y})$ is skewed to the right (called the Gumbel density or Fisher–Tippett Type I extreme values density). A random variable Y, distributed according to Gumbel's expression, satisfies [4]

$$\mathrm{E}(Y) = \gamma = 0.577215\dots, \quad \mathrm{Skew}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^3\right]}{\mathrm{Var}(Y)^{3/2}} = \frac{12\sqrt{6}}{\pi^3}\zeta(3) = 1.139547\dots,$$

$$\mathrm{Var}(Y) = \frac{\pi^2}{6} = 1.644934\dots, \quad \mathrm{Kurt}(Y) = \frac{\mathrm{E}\left[(Y - \mathrm{E}(Y))^4\right]}{\mathrm{Var}(Y)^2} - 3 = \frac{12}{5} = 2.4,$$

where γ is the Euler–Mascheroni constant [1.5] and $\zeta(3)$ is Apéry's constant [1.6]. (Some authors report the *square* of skewness; this explains the estimate 1.2986 in [2] and 1.3 in [19].) The constant $\zeta(3)$ also appears in [20]. Doubly exponential functions like $g(y)$ occur elsewhere (see [2.13], [5.7], and [6.10]).

The well-known Central Limit Theorem implies an asymptotic normal distribution for the *sum* of many independent, identically distributed random variables, whatever their common original distribution. A similar situation holds in extreme value theory. The asymptotic distribution of M_n (normalized) must belong to one of the following families [2, 14–17]:

$$G_{1,\alpha}(y) = \begin{cases} 0 & \text{if } y \le 0, \\ \exp(-y^{-\alpha}) & \text{if } y > 0, \end{cases} \qquad \text{“Fréchet” or Type II,}$$

$$G_{2,\alpha}(y) = \begin{cases} \exp(-(-y)^{\alpha}) & \text{if } y \le 0, \\ 1 & \text{if } y > 0, \end{cases} \qquad \text{“Weibull” or Type III,}$$

$$G_3(y) = \exp(-e^{-y}), \qquad \text{“Gumbel” or Type I,}$$

where $\alpha > 0$ is an arbitrary shape parameter. Note that $G_{2,1}(y)$ arose in our discussion of uniformly distributed X and $G_3(y)$ with regard to normally distributed X. It turns out to be unnecessary to know much about the distribution F of X to ascertain to which "domain of attraction" it belongs; the behavior of the tail of F is the crucial element. These three families can be further combined into a single one:

$$H_\beta(y) = \exp\left(-(1+\beta y)^{-1/\beta}\right) \text{ if } 1+\beta y > 0, \quad H_0(y) = \lim_{\beta \to 0} H_\beta(y),$$

which reduces to the three cases accordingly as $\beta > 0$, $\beta < 0$, or $\beta = 0$.

There is a fascinating connection between the preceding and random matrix theory (RMT). Consider first an $n \times n$ diagonal matrix with random diagonal elements $X_1, X_2, \ldots, X_n$; of course, its largest eigenvalue is equal to M_n. Consider now a random $n \times n$ complex Hermitian matrix. This means $X_{ij} = \bar{X}_{ij}$, so diagonal elements are real and off-diagonal elements satisfy a symmetry condition; further, all eigenvalues are real. A "natural" way of generating such matrices follows what is called the Gaussian Unitary Ensemble (GUE) probability distribution [21]. Exact moment formulas for the largest eigenvalue exist here for small n just as for the diagonal normally-distributed case discussed earlier [22]. The eigenvalues are independent in the diagonal case, but they are strongly dependent in the full Hermitian case. RMT is important in several ways: First, the spacing distribution between nontrivial zeros of the Riemann zeta function appears to be close to the eigenvalue distribution coming from GUE [2.15.3]. Second, RMT is pivotal in solving the longest increasing subsequence problem discussed in [5.20], and its tools are useful in understanding the two-dimensional Ising model [5.22]. Finally, RMT is associated with the physics of atomic energy levels, but elaboration on this is not possible here.

[1] J. D. Gibbons, *Nonparametric Statistical Inference*, McGraw-Hill, 1971; MR 86m:62067.
[2] H. A. David, *Order Statistics*, 2nd ed., Wiley, 1981; MR 82i:62073.
[3] N. Balakrishnan and A. Clifford Cohen, *Order Statistics and Inference*, Academic Press, 1991; MR 92k:62098.
[4] R. A. Fisher and L. H. C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Philos. Soc.* 24 (1928) 180–190.
[5] B. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, *Annals of Math.* 44 (1943) 423–453; MR 5,41b.

[6] H. L. Jones, Exact lower moments of order statistics in small samples from a normal distribution, *Annals of Math. Statist.* 19 (1948) 270–273; MR 9,601d.
[7] H. J. Godwin, Some low moments of order statistics, *Annals of Math. Statist.* 20 (1949) 279–285; MR 10,722f.
[8] H. Ruben, On the moments of the range and product moments of extreme order statistics in normal samples, *Biometrika* 43 (1956) 458–460; MR 18,607d.
[9] J. K. Patel and C. B. Read, *Handbook of the Normal Distribution*, Dekker, 1982, pp. 238–241; MR 83j:62002.
[10] Y. Watanabe, M. Isida, S. Taga, Y. Ichijo, T. Kawase, G. Niside, Y. Takeda, A. Horisuzi, and I. Kuriyama, Some contributions to order statistics, *J. Gakugei, Tokushima Univ.* 8 (1957) 41–90; MR 20 #5545.
[11] Y. Watanabe, T. Yamamoto, T. Sato, T. Fujimoto, M. Inoue, T. Suzuki, and T. Uno, Some contributions to order statistics (continued), *J. Gakugei, Tokushima Univ.* 9 (1958) 31–86; MR 21 #2330.
[12] H. Ruben, On the moments of order statistics in samples from normal populations, *Biometrika* 41 (1954) 200–227; also in *Contributions to Order Statistics*, ed. A. E. Sarhan and B. G. Greenberg, Wiley, 1962, pp. 165–190; MR 16,153c.
[13] N. J. Calkin, A curious binomial identity, *Discrete Math.* 131 (1994) 335–337; MR 95i:05002.
[14] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, 2nd ed., Krieger, 1987; MR 89a:60059.
[15] R.-D. Reiss, *Approximate Distributions of Order Statistics*, Springer-Verlag, 1989; MR 90e:62001.
[16] J. Galambos, Order statistics, *Handbook of Statistics*, v. 4, *Nonparametric Methods*, ed. P. R. Krishnaiah and P. K. Sens, Elsevier Science, 1984, pp. 359–382; MR 87g:62001.
[17] M. R. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, 1983; MR 84h:60050.
[18] P. Hall, On the rate of convergence of normal extremes, *J. Appl. Probab.* 16 (1979) 433–439; MR 80d:60025.
[19] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, 1972, p. 930; MR 94b:00012.
[20] P. C. Joshi and S. Chakraborty, Moments of Cauchy order statistics via Riemann zeta functions, *Statistical Theory and Applications*, ed. H. N. Nagaraja, P. K. Sen, and D. F. Morrison, Springer-Verlag, 1996, pp. 117–127; MR 98f:62149.
[21] C. A. Tracy and H. Widom, Universality of the distribution functions of random matrix theory, *Integrable Systems: From Classical to Quantum*, Proc. 1999 Montréal conf., ed. J. Harnad, G. Sabidussi, and P. Winternitz, Amer. Math. Soc, 2000, pp. 251–264; math-ph/9909001; MR 2002f:15036.
[22] J. Gravner, C. A. Tracy, and H. Widom, Limit theorems for height fluctuations in a class of discrete space and time growth models, *J. Stat. Phys.* 102 (2001) 1085–1132; math.PR/0005133; MR 2002d:82065.

5.17 Pattern-Free Word Constants

Let $a, b, c, \ldots$ denote the letters of a finite alphabet. A **word** is a finite sequence of letters; two examples are $abcacbacbc$ and $abcacbabcb$. A **square** is a word of the form xx, with x a nonempty word. A word is **square-free** if it contains no squares as factors. The first example contains the square $acbacb$ whereas the second is square-free. We ask the following question: How many square-free words of length n are there?

Over a two-letter alphabet, the only square-free words are a, b, ab, ba, aba, and bab; thus **binary** square-free words are not interesting. There do, however, exist arbitrarily long **ternary** square-free words, that is, over a three-letter alphabet. This fact was first proved by Thue [1, 2] using what is now called the Prouhet–Thue–Morse sequence [6.8]. Precise asymptotic enumeration of such words is complicated [3–7]. Brandenburg [8] proved that the number $s(n)$ of ternary square-free words of length $n > 24$ satisfies

$$6 \cdot 1.032^n < 6 \cdot 2^{\frac{n}{22}} \leq s(n) \leq 6 \cdot 1172^{\frac{n-2}{22}} < 3.157 \cdot 1.379^n,$$

and Brinkhuis [9] showed that $s(n) \leq A \cdot 1.316^n$ for some constant $A > 0$. Noonan & Zeilberger [10] improved the upper bound to $A' \cdot 1.302128^n$ for some constant $A' > 0$, and obtained a non-rigorous estimate of the limit

$$S = \lim_{n\to\infty} s(n)^{\frac{1}{n}} = 1.302\ldots.$$

An independent computation [11] gave $S = \exp(0.263719\ldots) = 1.301762\ldots$, as well as estimates of S for k-letter alphabets, $k > 3$. Ekhad & Zeilberger [12] recently demonstrated that $1.041^n < 2^{n/17} \leq s(n)$, the first improvement in the lower bound in fifteen years. Note that S is a connective constant in the same manner as certain constants μ associated with self-avoiding walks [5.10]. In fact, Noonan & Zeilberger's computation of S is based on the same Goulden-Jackson technology used in bounding μ.

A **cube-free word** is a word that contains no factors of the form xxx, where x is a nonempty word. The Prouhet–Thue–Morse sequence gives examples of arbitrarily long binary cube-free words. Brandenburg [8] proved that the number $c(n)$ of binary cube-free words of length $n > 18$ satisfies

$$2 \cdot 1.080^n < 2 \cdot 2^{\frac{n}{9}} \leq c(n) \leq 2 \cdot 1251^{\frac{n-1}{17}} < 1.315 \cdot 1.522^n,$$

and Edlin [13] improved the upper bound to $B \cdot 1.45757921^n$ for some constant $B > 0$. Edlin also obtained a non-rigorous estimate of the limit:

$$C = \lim_{n\to\infty} c(n)^{\frac{1}{n}} = 1.457\ldots.$$

A word is **overlap-free** if it contains no factor of the form $xyxyx$, with x nonempty. The Prouhet–Thue–Morse sequence, again, gives examples of arbitrarily long binary overlap-free words. Observe that a square-free word must be overlap-free, and that an overlap-free word must be cube-free. In fact, overlapping is the lowest pattern avoidable in arbitrarily long binary words. The number $t(n)$ of binary overlap-free words of length n satisfies [14, 15]

$$p \cdot n^{1.155} \leq t(n) \leq q \cdot n^{1.587}$$

for certain constants p and q. Therefore, $t(n)$ experiences only polynomial growth, unlike $s(n)$ and $c(n)$. Cassaigne [16] proved the interesting fact that $\lim_{n\to\infty} \ln(t(n))/\ln(n)$ does not exist, but

$$1.155 < T_L = \liminf_{n\to\infty} \frac{\ln(t(n))}{\ln(n)} < 1.276 < 1.332 < T_U = \limsup_{n\to\infty} \frac{\ln(t(n))}{\ln(n)} < 1.587$$

(actually, he proved much more). We observed similar asymptotic misbehavior in [2.16].

An **abelian square** is a word xx', with x a nonempty word and x' a permutation of x. A word is **abelian square-free** if it contains no abelian squares as factors. The word $abcacbabcb$ contains the abelian square $abcacb$. In fact, any ternary word of length at least 8 must contain an abelian square. Pleasants [17] proved that arbitrarily long abelian square-free words, based on five letters, exist. The four-letter case remained an open question until recently. Keränen [18] proved that arbitrarily long quaternary abelian square-free words also exist. Carpi [19] went farther to show that their number $h(n)$ must satisfy

$$\liminf_{n\to\infty} h(n)^{\frac{1}{n}} > 1.000021,$$

and he wrote, "...the closeness of this value to 1 leads us to think that, probably, it is far from optimal."

A ternary word w is a **partially abelian square** if $w = xx'$, with x a nonempty word and x' a permutation of x that leaves the letter b fixed, and that allows only adjacent letters a and c to commute. For example, the word $bacbca$ is a partially abelian square. A word is **partially abelian square-free** if it contains no partially abelian squares as factors. Cori & Formisano [20] used Kobayashi's inequalities for $t(n)$ to derive bounds for the number of partially abelian square-free words.

Kolpakov & Kucherov [21, 22] asked: What is the minimal proportion of one letter in infinite square-free ternary words? Follow-on work by Tarannikov suggests [23] that the answer is $0.2746\ldots$.

A word over a k-letter alphabet is **primitive** if it is not a power of any subword [24]. The number of primitive words of length n is $\sum_{d|n} \mu(d)k^{n/d}$, where $\mu(d)$ is the Möbius mu function [2.2]. Hence, on the one hand, the proportion of words that are primitive is easily shown to approach 1 as $n \to \infty$. On the other hand, the problem of all counting words not *containing* a power is probably about as difficult as enumerating square-free words, cube-free words, etc.

A binary word $w_1 w_2 w_3 \ldots w_n$ of length n is said to be **unforgeable** if it never matches a left or right shift of itself, that is, it is never the same as any of $u_1 u_2 \ldots u_m w_1 w_2 \ldots w_{n-m}$ or $w_{m+1} w_{m+2} \ldots w_n v_1 v_2 \ldots v_m$ for any possible choice of u_is or v_js and any $1 \leq m \leq n-1$. For example, we cannot have $w_1 = w_n$ because trouble would arise when $m = n-1$. Let $f(n)$ denote the number of unforgeable words of length n. The example shows immediately that

$$0 \leq \rho = \lim_{n\to\infty} \frac{f(n)}{2^n} \leq \frac{1}{2}.$$

Further, via generating functions [7, 25–27],

$$\begin{aligned}\rho &= \sum_{n=1}^{\infty} (-1)^{n-1} \frac{2}{2^{(2^{n+1}-1)} - 1} \prod_{m=2}^{n} \frac{2^{(2^m-1)}}{2^{(2^m-1)} - 1} = 0.2677868402\ldots \\ &= 1 - 0.7322131597\ldots,\end{aligned}$$

and this series is extremely rapidly convergent.

[1] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, *Videnskapsselskapets Skrifter I, Matematisk-Naturvidenskapelig Klasse, Kristiania*, n. 1, Dybwad, 1912,

pp. 1–67; also in *Selected Mathematical Papers*, ed. T. Nagell, A. Selberg, S. Selberg, and K. Thalberg, Universitetsforlaget, 1977, pp. 413–478; MR 57 #46.
[2] J.-P. Allouche and J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, *Sequences and Their Applications (SETA)*, Proc. 1998 Singapore conf., ed. C. Ding, T. Helleseth, and H. Niederreiter, Springer-Verlag, 1999, pp. 1–16; MR 2002e:11025.
[3] M. Lothaire, *Combinatorics on Words*, Addison-Wesley, 1983; MR 98g:68134.
[4] J. Berstel, Some recent results on squarefree words, *Proc. 1984 Symp. on Theoretical Aspects of Computer Science (STACS)*, Paris, ed. M. Fontet and K. Mehlhorn, Lect. Notes in Comp. Sci. 166, Springer-Verlag, 1984, pp. 14–25; MR 86e:68056.
[5] J. Currie, Open problems in pattern avoidance, *Amer. Math. Monthly* 100 (1993) 790–793.
[6] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge Univ. Press, 2002.
[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A003000, A006156, A028445, and A007777.
[8] F.-J. Brandenburg, Uniformly growing k^{th} power-free homomorphisms, *Theoret. Comput. Sci.* 23 (1983) 69–82; MR 84i:68148.
[9] J. Brinkhuis, Non-repetitive sequences on three symbols, *Quart. J. Math.* 34 (1983) 145–149; MR 84e:05008.
[10] J. Noonan and D. Zeilberger, The Goulden-Jackson cluster method: Extensions, applications and implementations, *J. Differ. Eq. Appl.* 5 (1999) 355–377; MR 2000f:05005.
[11] M. Baake, V. Elser, and U. Grimm, The entropy of square-free words, *Math. Comput. Modelling* 26 (1997) 13–26; math-ph/9809010; MR 99d:68202.
[12] S. B. Ekhad and D. Zeilberger, There are more than $2^{n/17}$ n-letter ternary square-free words, *J. Integer Seq.* 1 (1998) 98.1.9.
[13] A. E. Edlin, The number of binary cube-free words of length up to 47 and their numerical analysis, *J. Differ. Eq. Appl.* 5 (1999) 353–354.
[14] A. Restivo and S. Salemi, Overlap-free words on two symbols, *Automata on Infinite Words*, ed. M. Nivat and D. Perrin, Lect. Notes in Comp. Sci. 192, Springer-Verlag, 1985, pp. 198–206; MR 87c:20101.
[15] Y. Kobayashi, Enumeration of irreducible binary words, *Discrete Appl. Math.* 20 (1988) 221–232; MR 89f:68036.
[16] J. Cassaigne, Counting overlap-free binary words, *Proc. 1993 Symp. on Theoretical Aspects of Computer Science (STACS)*, Würzburg, ed. P. Enjalbert, A. Finkel, and K. W. Wagner, Lect. Notes in Comp. Sci. 665, Springer-Verlag, 1993, pp. 216–225; MR 94j:68152.
[17] P. A. B. Pleasants, Non-repetitive sequences, *Proc. Cambridge Philos. Soc.* 68 (1970) 267–274; MR 42 #85.
[18] V. Keränen, Abelian squares are avoidable on 4 letters, *Proc. 1992 Int. Colloq. on Automata, Languages and Programming (ICALP)*, Vienna, ed. W. Kuich, Lect. Notes in Comp. Sci. 623, Springer-Verlag, 1992, pp. 41–52; MR 94j:68244.
[19] A. Carpi, On the number of abelian square-free words on four letters, *Discrete Appl. Math.* 81 (1998) 155–167; MR 98j:68139.
[20] R. Cori and M. R. Formisano, On the number of partially abelian square-free words on a three-letter alphabet, *Theoret. Comput. Sci.* 81 (1991) 147–153; MR 92i:68129.
[21] R. Kolpakov and G. Kucherov, Minimal letter frequency in n^{th} power-free binary words, *Mathematical Foundations of Computer Science (MFCS)*, Proc. 1997 Bratislava conf., ed. I. Prívara and P. Ruzicka, Lect. Notes in Comp. Sci. 1295, Springer-Verlag, pp. 347–357; MR 99d:68206.
[22] R. Kolpakov, G. Kucherov, and Y. Tarannikov, On repetition-free binary words of minimal density, *Theoret. Comput. Sci.* 218 (1999) 161–175; MR 2000b:68180.
[23] Y. Tarannikov, Minimal letter density in infinite ternary square-free word is 0.2746 . . . , *Proc. Two Joint French-Russian Seminars on Combinatorial and Algorithmical Properties of Discrete Structures*, Moscow State Univ., 2001, pp. 51–56.
[24] H. Petersen, On the language of primitive words, *Theoret. Comput. Sci.* 161 (1996) 141–156; MR 97e:68065.
[25] P. T. Nielsen, A note on bifix-free sequences, *IEEE Trans. Inform. Theory* 19 (1973) 704–706; MR 52 #2724.

[26] G. Blom and O. P. Lossers, Overlapping binary sequences, *SIAM Rev.* 37 (1995) 619–620.
[27] D. J. Greaves and S. J. Montgomery-Smith, Unforgeable marker sequences, unpublished note (2000).

5.18 Percolation Cluster Density Constants

Percolation theory is concerned with fluid flow in random media, for example, molecules penetrating a porus solid or wildfires consuming a forest. Broadbent & Hammersley [1–3] wondered about the probable number and structure of open channels in media for fluid passage. Answering their question has created an entirely new field of research [4–10]. Since the field is vast, we will attempt only to present a few constants.

Let $M = (m_{ij})$ be a random $n \times n$ binary matrix satisfying the following:

- $m_{ij} = 1$ with probability p, 0 with probability $1 - p$ for each i, j,
- m_{ij} and m_{kl} are independent for all $(i, j) \neq (k, l)$.

An s-**cluster** is an isolated grouping of s adjacent 1s in M, where adjacency means horizontal or vertical neighbors (not diagonal). For example, the 4×4 matrix

$$M = \begin{pmatrix} 1\,0\,1\,1 \\ 1\,1\,0\,0 \\ 0\,1\,0\,1 \\ 1\,0\,0\,1 \end{pmatrix}$$

has one 1-cluster, two 2-clusters, and one 4-cluster. The total number of clusters K_4 is 4 in this case. For arbitrary n, the total cluster count K_n is a random variable. The limit $\kappa_S(p)$ of the normalized expected value $\mathrm{E}(K_n)/n^2$ exists as $n \to \infty$, and $\kappa_S(p)$ is called the **mean cluster density** for the **site percolation model**. It is known that $\kappa_S(p)$ is twice continuously differentiable on $[0, 1]$; further, $\kappa_S(p)$ is analytic on $[0, 1]$ except possibly at one point $p = p_c$. Monte Carlo simulation and numerical Padé approximants can be used to compute $\kappa_S(p)$. For example [11], it is known that $\kappa_S(1/2) = 0.065770\ldots$.

Instead of an $n \times n$ binary matrix M, consider a binary array A of $2n(n-1)$ entries that looks like

$$A = \begin{pmatrix} & a_{12} & & a_{14} & & a_{16} & \\ a_{11} & & a_{13} & & a_{15} & & a_{17} \\ & a_{22} & & a_{24} & & a_{26} & \\ a_{21} & & a_{23} & & a_{25} & & a_{27} \\ & a_{32} & & a_{34} & & a_{36} & \\ a_{31} & & a_{33} & & a_{35} & & a_{37} \\ & a_{42} & & a_{44} & & a_{46} & \end{pmatrix}$$

(here $n = 4$). We associate a_{ij} not with a site of the $n \times n$ square lattice (as we do for m_{ij}) but with a bond. An s-cluster here is an isolated, connected subgraph of the graph of all bonds associated with 1s. For example, the array

$$A = \left(\begin{array}{ccccccc} & 1 & & 0 & & 0 & \\ 1 & & 0 & & 0 & & 0 \\ & 0 & & 1 & & 0 & \\ 0 & & 1 & & 1 & & 0 \\ & 0 & & 1 & & 0 & \\ 1 & & 0 & & 0 & & 0 \\ & 0 & & 0 & & 0 & \end{array}\right)$$

has one 1-cluster, one 2-cluster, and one 4-cluster. For **bond percolation models** such as this, we include 0-clusters in the total count as well, that is, isolated sites with no attached 1s bonds. In this case there are seven 0-clusters; hence the total number of clusters K_4 is 10. The mean cluster density $\kappa_B(p) = \lim_{n\to\infty} \mathrm{E}(K_n)/n^2$ exists and similar smoothness properties hold. Remarkably, however, an exact integral expression can be found at $p = 1/2$ for the mean cluster density [13, 14]:

$$\kappa_B\left(\frac{1}{2}\right) = -\frac{1}{8}\cot(y)\cdot\frac{d}{dy}\left\{\frac{1}{y}\int_{-\infty}^{\infty}\operatorname{sech}\left(\frac{\pi x}{2y}\right)\ln\left(\frac{\cosh(x)-\cos(2y)}{\cosh(x)-1}\right)dx\right\}\Bigg|_{y=\frac{\pi}{3}},$$

which Adamchik [11, 12] recently simplified to

$$\kappa_B\left(\frac{1}{2}\right) = \frac{3\sqrt{3}-5}{2} = 0.0980762113\ldots.$$

This constant is sometimes reported as $0.0355762113\ldots$, which is $\kappa_B(1/2) - 1/16$, if 0-clusters are not included in the total count. It may alternatively be reported as $0.0177881056\ldots$, which occurs if one normalizes not by the number of sites, n^2, but by the number of bonds, $2n(n-1)$. Caution is needed when reviewing the literature. Other occurrences of this integral are in [15–18].

An expression for the limiting variance of bond cluster density is not known, but a Monte Carlo estimate $0.164\ldots$ and relevant discussion appear in [11]. The bond percolation model on the *triangular* lattice gives a limiting mean cluster density $0.111\ldots$ at a specific value $p = 0.347\ldots$ (see the next section for greater precision). The associated variance $0.183\ldots$, again, is not known.

5.18.1 Critical Probability

Let us turn attention away from mean cluster density $\kappa(p)$ and instead toward **mean cluster size** $\sigma(p)$. In the examples given earlier, $S_4 = (1+2+2+4)/4 = 9/4$ for the site case, $S_4 = (1+2+4)/3 = 7/3$ for the bond case, and $\sigma(p)$ is the limiting value of $\mathrm{E}(S_n)$ as $n \to \infty$. The **critical probability** or **percolation threshold** p_c is defined to be [5, 6, 10]

$$p_c = \inf_{\substack{0<p<1\\ \sigma(p)=\infty}} p,$$

that is, the concentration p at which an ∞-cluster appears in the infinite lattice. There are other possible definitions that turn out to be equivalent under most conditions. For example, if $\theta(p)$ denotes the **percolation probability**, that is, the probability that an ∞-cluster contains a prescribed site or bond, then p_c is the unique point for which $p < p_c$ implies $\theta(p) = 0$, and $p > p_c$ implies $\theta(p) > 0$. The critical probability indicates a phase transition in the system, analogous to that observed in [5.12] and [5.22].

For site percolation on the square lattice, there are rigorous bounds [19–24]

$$0.556 < p_c < 0.679492$$

and an estimate [25, 26] $p_c = 0.5927460\ldots$. based on extensive simulation. Ziff [11] additionally calculated that $\kappa_S(p_c) = 0.0275981\ldots$ via simulation. Parameter bounds for the cubic lattice and higher dimensions appear in [27–30].

In contrast, for bond percolation on the square and triangular lattices, there are exact results due to Sykes & Essam [31, 32]. Kesten [33] proved that $p_c = 1/2$ on the square lattice, corresponding to the expression $\kappa_B(1/2)$ in the previous section. On the triangular lattice, Wierman [34] proved that

$$p_c = 2\sin\left(\frac{\pi}{18}\right) = 0.3472963553\ldots,$$

and this corresponds to another exact expression [11, 35–37],

$$\begin{aligned}\kappa_B(p_c) &= -\frac{3}{8}\csc(2y)\cdot\frac{d}{dy}\left\{\int_{-\infty}^{\infty}\frac{\sinh((\pi-y)x)\sinh(\frac{2}{3}yx)}{x\sinh(\pi x)\cosh(yx)}dx\right\}\Bigg|_{y=\frac{\pi}{3}} + \frac{3}{2} - \frac{2}{1+p_c}\\ &= \frac{35}{4} - \frac{3}{p_c} = \frac{23}{4} - \frac{3}{2}\cdot\left\{\sqrt[3]{4\left(1+i\sqrt{3}\right)} + \sqrt[3]{4\left(1-i\sqrt{3}\right)}\right\}\\ &= 0.1118442752\ldots.\end{aligned}$$

Similar results apply for the hexagonal (honeycomb) lattice by duality.

It is also known that $p_c = 1/2$ for site percolation on the triangular lattice [10] and, in this case, $\kappa_S(1/2) = 0.0176255\ldots$ via simulation [11, 38]. For site percolation on the hexagonal lattice, we have bounds [39]

$$0.6527 < 1 - 2\sin\left(\frac{\pi}{18}\right) \le p_c \le 0.8079$$

and an estimate $p_c = 0.6962\ldots$ [40, 41].

5.18.2 Series Expansions

Here are details on how the functions $\kappa_S(p)$ and $\kappa_B(p)$ may be computed [6, 42, 43]. We will work on the square lattice, focusing mostly on site percolation. Let g_{st} denote the number of lattice animals [5.19] with area s and perimeter t, and let $q = 1 - p$. The probability that a fixed site is a 1-cluster is clearly pq^4. Because a 2-cluster can be oriented either horizontally or vertically, the average 2-cluster count per site is $2p^2q^6$. A 3-cluster can be linear (two orientations) or L-shaped (four orientations); hence the

Table 5.4. *Mean s-Cluster Densities*

s	Mean s-Cluster Density for Site Model	Mean s-Cluster Density for Bond Model
0	0	q^4
1	pq^4	$2pq^6$
2	$2p^2q^6$	$6p^2q^8$
3	$p^3(2q^8+4q^7)$	$p^3(18q^{10}+4q^9)$
4	$p^4(2q^{10}+8q^9+9q^8)$	$p^4(55q^{12}+32q^{11}+q^8)$

average 3-cluster count per site is $p^3(2q^8+4q^7)$. More generally, the mean s-cluster density is $\sum_t g_{st}p^sq^t$. Summing the left column entries in Table 5.4 [44,45] gives $\kappa_S(p)$ as the number of entries $\to\infty$:

$$\kappa_S(p) = p - 2p^2 + p^4 + p^8 - p^9 + 2p^{10} - 4p^{11} + 11p^{12} + - \cdots$$
$$\sim \kappa_S(p_c) + a_S(p-p_c) + b_S(p-p_c)^2 + c_S\,|p-p_c|^{2-\alpha}.$$

Likewise, summing the right column entries in the table gives $\kappa_B(p)$:

$$\kappa_B(p) = q^4 + 2p - 6p^2 + 4p^3 + 2p^6 - 2p^7 + 7p^8 - 12p^9 + 28p^{10} + - \cdots$$
$$\sim \kappa_B(\tfrac{1}{2}) + a_B(p-\tfrac{1}{2}) + b_B(p-\tfrac{1}{2})^2 + c_B\,|p-\tfrac{1}{2}|^{2-\alpha},$$

where $a_B = -0.50\ldots$, $b_B = 2.8\ldots$, and $c_B = -8.48\ldots$ [46]. The exponent α is conjectured to be $-2/3$, that is, $2-\alpha = 8/3$.

If instead of $\sum_{s,t} g_{st}p^sq^t$, we examine $\sum_{s,t} s^2 g_{st}p^{s-1}q^t$, then for the site model,

$$\sigma_S(p) = 1 + 4p + 12p^2 + 24p^3 + 52p^4 + 108p^5 + 224p^6 + 412p^7 + - \cdots$$
$$\sim C\,|p-p_c|^{-\gamma}$$

is the mean cluster size series (for low concentration $p < p_c$). The exponent γ is conjectured to be $43/18$.

The expression $1 - \sum_{s,t} s g_{st}p^{s-1}q^t$, when expanded in terms of q, gives

$$\theta_S(p) = 1 - q^4 - 4q^6 - 8q^7 - 23q^8 - 28q^9 - 186q^{10} + 48q^{11} - + \cdots$$
$$\sim D\,|p-p_c|^{\beta},$$

which is the site percolation probability series (for high concentration $p > p_c$). The exponent β is conjectured to be $5/36$.

Smirnov & Werner [47] recently proved that α, β, and γ indeed exist and are equal to their conjectured values, for site percolation on the triangular lattice. A proof of universality would encompass both site and bond cases on the square lattice, but this has not yet been achieved.

5.18.3 Variations

Let the sites of an infinite lattice be independently labeled A with probability p and B with probability $1-p$. Ordinary site percolation theory involves clusters of As. Let us instead connect adjacent sites that possess *opposite* labels and leave adjacent sites with

the same labels disconnected. This is known as AB **percolation** or **antipercolation**. We wish to know what can be said of the probability $\theta(p)$ that an infinite AB cluster contains a prescribed site. It turns out that $\theta(p) = 0$ for all p for the infinite square lattice [48], but $\theta(p) > 0$ for all p lying in some nonempty subinterval containing 1/2, for the infinite triangular lattice [49]. The exact extent of this interval is not known: Mai & Halley [50] gave [0.2145, 0.7855] via Monte Carlo simulation whereas Wierman [51] gave [0.4031, 0.5969]. The function $\theta(p)$, for the triangular lattice, is nondecreasing on [0, 1/2] and therefore was deemed unimodal on [0, 1] by Appel [52].

Ordinary bond percolation theory is concerned with models in which any selected bond is either open (1) or closed (0). First-passage percolation [53] assigns not a binary random variable to each bond, but rather a nonnegative *real* random variable, thought of as length. Consider the square lattice in which each bond is independently assigned a length from the Uniform [0, 1] probability distribution. Let T_n denote the shortest length of all lattice path lengths starting at the origin (0, 0) and ending at $(n, 0)$; then it can be proved that the limit

$$\tau = \lim_{n\to\infty}\frac{\mathrm{E}(T_n)}{n} = \inf_n \frac{\mathrm{E}(T_n)}{n}$$

exists. Building upon earlier work [54–58], Alm & Parviainen [59] obtained rigorous bounds $0.243666 \le \tau \le 0.403141$ and an estimate $\tau = 0.312\ldots$ via simulation. If, instead, lengths are taken from the exponential distribution with unit mean, then we have bounds $0.300282 \le \tau \le 0.503425$ and an estimate $\tau = 0.402$. Godsil, Grötschel & Welsh [9] suggested the exact evaluation of τ to be a "hopelessly intractable problem."

We mention finally a constant $\lambda_c = 0.359072\ldots$ that arises in **continuum percolation** [5, 60]. Consider a homogeneous Poisson process of intensity λ on the plane, that is, points are uniformly distributed in the plane such that

- the probability of having exactly n points in a subset S of measure μ is $e^{-\lambda\mu}(\lambda\mu)^n/n!$ and
- the counts n_i of points in any collection of disjoint measurable subsets S_i are independent random variables.

Around each point, draw a disk of unit radius. The disks are allowed to overlap; that is, they are fully penetrable. There exists a unique critical intensity λ_c such that an unbounded connected cluster of disks develops with probability 1 if $\lambda > \lambda_c$ and with probability 0 if $\lambda < \lambda_c$. Hall [61] proved the best-known rigorous bounds $0.174 < \lambda_c < 0.843$, and the numerical estimate $0.359072\ldots$ is found in [62–64]. Among several alternative representations, we mention $\varphi_c = 1 - \exp(-\pi\lambda_c) = 0.676339\ldots$ [65] and $\pi\lambda_c = 1.128057\ldots$ [66]. The latter is simply the normalized total area of all the disks, disregarding whether they overlap or not, whereas φ_c takes overlapping portions into account. Continuum percolation shares many mathematical properties with lattice percolation, yet in many ways it is a more accurate model of physical disorder. Interestingly, it has also recently been applied in pure mathematics itself, to the study of gaps in the set of Gaussian primes [67].

[1] S. R. Broadbent and J. M. Hammersley, Percolation processes. I: Crystals and mazes, *Proc. Cambridge Philos. Soc.* 53 (1957) 629–641; MR 19,989e.
[2] J. M. Hammersley, Percolation processes. II: The connective constant, *Proc. Cambridge Philos. Soc.* 53 (1957) 642–645; MR 19,989f.
[3] J. M. Hammersley, Percolation processes: Lower bounds for the critical probability, *Annals of Math. Statist.* 28 (1957) 790–795; MR 21 #374.
[4] D. Stauffer, Scaling theory of percolation clusters, *Phys. Rep.* 54 (1979) 1–74.
[5] G. Grimmett, *Percolation*, Springer-Verlag, 1989; MR 90j:60109.
[6] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2^{nd} ed., Taylor and Francis, 1992; MR 87k:82093.
[7] M. Sahimi, *Applications of Percolation Theory*, Taylor and Francis, 1994.
[8] S. Havlin and A. Bunde, Percolation, *Contemporary Problems in Statistical Physics*, ed. G. H. Weiss, SIAM, 1994, pp. 103–146.
[9] C. Godsil, M. Grötschel, and D. J. A. Welsh, Combinatorics in statistical physics, *Handbook of Combinatorics*, v. II, ed. R. Graham, M. Grötschel, and L. Lovász, MIT Press, 1995, pp. 1925–1954; MR 96h:05001.
[10] B. D. Hughes, *Random Walks and Random Environments*, v. 2, Oxford Univ. Press, 1996; MR 98d:60139.
[11] R. M. Ziff, S. R. Finch, and V. S. Adamchik, Universality of finite-size corrections to the number of critical percolation clusters, *Phys. Rev. Lett.* 79 (1997) 3447–3450.
[12] V. S. Adamchik, S. R. Finch, and R. M. Ziff, The Potts model on the square lattice, unpublished note (1996).
[13] H. N. V. Temperley and E. H. Lieb, Relations between the 'percolation' and 'colouring' problem and other graph-theoretical problems associated with regular planar lattices: Some exact results for the 'percolation' problem, *Proc. Royal Soc. London* A 322 (1971) 251–280; MR 58 #16425.
[14] J. W. Essam, Percolation and cluster size, *Phase Transitions and Critical Phenomena*, v. II, ed. C. Domb and M. S. Green, Academic Press, 1972, pp. 197–270; MR 50 #6392.
[15] M. L. Glasser, D. B. Abraham, and E. H. Lieb, Analytic properties of the free energy for the "ice" models, *J. Math. Phys.* 13 (1972) 887–900; MR 55 #7227.
[16] R. J. Baxter, Potts model at the critical temperature, *J. Phys. C* 6 (1973) L445–L448.
[17] H. N. V. Temperley, The shapes of domains occurring in the droplet model of phase transitions, *J. Phys. A* 9 (1976) L113–L117.
[18] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982, pp. 322–341; MR 90b:82001.
[19] B. Tóth, A lower bound for the critical probability of the square lattice site percolation, *Z. Wahr. Verw. Gebiete* 69 (1985) 19–22; MR 86f:60118.
[20] S. A. Zuev, Bounds for the percolation threshold for a square lattice (in Russian), *Teor. Veroyatnost. i Primenen.* 32 (1987) 606–609; Engl. transl. in *Theory Probab. Appl.* 32 (1987) 551–553; MR 89h:60171.
[21] S. A. Zuev, A lower bound for a percolation threshold for a square lattice (in Russian), *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* (1988) n. 5, 59–61; Engl. transl. in *Moscow Univ. Math. Bull.*, v. 43 (1988) n. 5, 66–69; MR 91k:82035.
[22] M. V. Menshikov and K. D. Pelikh, Percolation with several defect types: An estimate of critical probability for a square lattice, *Math. Notes Acad. Sci. USSR* 46 (1989) 778–785; MR 91h:60116.
[23] J. C. Wierman, Substitution method critical probability bounds for the square lattice site percolation model, *Combin. Probab. Comput.* 4 (1995) 181–188; MR 97g:60136.
[24] J. van den Berg and A. Ermakov, A new lower bound for the critical probability of site percolation on the square lattice, *Random Structures Algorithms* 8 (1996) 199–212; MR 99b:60165.
[25] R. M. Ziff and B. Sapoval, The efficient determination of the percolation threshold by a frontier-generating walk in a gradient, *J. Phys. A* 19 (1986) L1169–L1172.

[26] R. M. Ziff, Spanning probability in 2D percolation, *Phys. Rev. Lett.* 69 (1992) 2670–2673.
[27] M. Campanino and L. Russo, An upper bound on the critical percolation probability for the three-dimensional cubic lattice, *Annals of Probab.* 13 (1985) 478–491; MR 86j:60222.
[28] J. Adler, Y. Meir, A. Aharony, and A. B. Harris, Series study of percolation moments in general dimension, *Phys. Rev. B* 41 (1990) 9183–9206.
[29] B. Bollobás and Y. Kohayakawa, Percolation in high dimensions, *Europ. J. Combin.* 15 (1994) 113–125; MR 95c:60092.
[30] T. Hara and G. Slade, The self-avoiding-walk and percolation critical points in high dimensions, *Combin. Probab. Comput.* 4 (1995) 197–215; MR 96i:82081.
[31] M. F. Sykes and J. W. Essam, Exact critical percolation probabilities for site and bond problems in two dimensions, *J. Math. Phys.* 5 (1964) 1117–1121; MR 29 #1977.
[32] J. W. Essam and M. F. Sykes, Percolation processes. I: Low-density expansion for the mean number of clusters in a random mixture, *J. Math. Phys.* 7 (1966) 1573–1581; MR 34 #3952.
[33] H. Kesten, The critical probability of bond percolation on the square lattice equals 1/2, *Commun. Math. Phys.* 74 (1980) 41–59; MR 82c:60179.
[34] J. C. Wierman, Bond percolation on honeycomb and triangular lattices, *Adv. Appl. Probab.* 13 (1981) 298–313; MR 82k:60216.
[35] R. J. Baxter, H. N. V. Temperley, and S. E. Ashley, Triangular Potts model at its transition temperature, and related models, *Proc. Royal Soc. London A* 358 (1978) 535–559; MR 58 #20108.
[36] V. S. Adamchik, S. R. Finch, and R. M. Ziff, The Potts model on the triangular lattice, unpublished note (1996).
[37] V. S. Adamchik, A class of logarithmic integrals, *Proc. 1997 Int. Symp. on Symbolic and Algebraic Computation (ISSAC)*, ed. W. W. Küchlin, Maui, ACM, 1997, pp. 1–8; MR 2001k:33043.
[38] D. W. Erbach, Advanced Problem 6229, *Amer. Math. Monthly* 85 (1978) 686.
[39] T. Luczak and J. C. Wierman, Critical probability bounds for two-dimensional site percolation models, *J. Phys. A* 21 (1988) 3131–3138; MR 89i:82053.
[40] Z. V. Djordjevic, H. E. Stanley, and A. Margolina, Site percolation threshold for honeycomb and square lattices, *J. Phys. A* 15 (1982) L405–L412.
[41] R. P. Langlands, C. Pichet, P. Pouliot, and Y. Saint-Aubin, On the universality of crossing probabilities in two-dimensional percolation, *J. Stat. Phys.* 67 (1992) 553–574; MR 93e:82028.
[42] J. Adler, Series expansions, *Comput. in Phys.* 8 (1994) 287–295.
[43] A. R. Conway and A. J. Guttmann, On two-dimensional percolation, *J. Phys. A* 28 (1995) 891–904.
[44] M. F. Sykes and M. Glen, Percolation processes in two dimensions. I: Low-density series expansions, *J. Phys. A* 9 (1976) 87–95.
[45] M. F. Sykes, D. S. Gaunt, and M. Glen, Perimeter polynomials for bond percolation processes, *J. Phys. A* 14 (1981) 287–293; MR 81m:82045.
[46] C. Domb and C. J. Pearce, Mean number of clusters for percolation processes in two dimensions, *J. Phys. A* 9 (1976) L137–L140.
[47] S. Smirnov and W. Werner, Critical exponents for two-dimensional percolation, *Math. Res. Lett.* 8 (2001) 729–744; math.PR/0109120.
[48] M. J. Appel and J. C. Wierman, On the absence of infinite AB percolation clusters in bipartite graphs, *J. Phys. A* 20 (1987) 2527–2531; MR 89b:82061.
[49] J. C. Wierman and M. J. Appel, Infinite AB percolation clusters exist on the triangular lattice, *J. Phys. A* 20 (1987) 2533–2537; MR 89b:82062
[50] T. Mai and J. W. Halley, AB percolation on a triangular lattice, *Ordering in Two Dimensions*, ed. S. K. Sinha, North-Holland, 1980, pp. 369–371.
[51] J. C. Wierman, AB percolation on close-packed graphs, *J. Phys. A* 21 (1988) 1939–1944; MR 89m:60258.

[52] M. J. B. Appel, *AB* percolation on plane triangulations is unimodal, *J. Appl. Probab.* 31 (1994) 193–204; MR 95e:60101.
[53] R. T. Smythe and J. C. Wierman, *First-Passage Percolation on the Square Lattice*, Lect. Notes in Math. 671, Springer-Verlag, 1978; MR 80a:60135.
[54] J. M. Hammersley and D. J. A. Welsh, First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory, *Bernoulli, 1713; Bayes, 1763; Laplace, 1813. Anniversary Volume*, Proc. 1963 Berkeley seminar, ed. J. Neyman and L. M. Le Cam, Springer-Verlag, 1965, pp. 61–110; MR 33 #6731.
[55] S. Janson, An upper bound for the velocity of first-passage percolation, *J. Appl. Probab.* 18 (1981) 256–262; MR 82c:60178.
[56] D. J. A. Welsh, An upper bound for a percolation constant, *Z. Angew Math. Phys.* 16 (1965) 520–522.
[57] R. T. Smythe, Percolation models in two and three dimensions, *Biological Growth and Spread*, Proc. 1979 Heidelberg conf., ed. W. Jäger, H. Rost, and P. Tautu, Lect. Notes in Biomath. 38, Springer-Verlag, 1980, pp. 504–511; MR 82i:60183.
[58] R. Ahlberg and S. Janson, Upper bounds for the connectivity constant, unpublished note (1981).
[59] S. E. Alm and R. Parviainen, Lower and upper bounds for the time constant of first-passage percolation, *Combin. Probab. Comput*, to appear.
[60] R. Meester and R. Roy, *Continuum Percolation*, Cambridge Univ. Press, 1996; MR 98d:60193.
[61] P. Hall, On continuum percolation, *Annals of Probab.* 13 (1985) 1250–1266; MR 87f:60018.
[62] S. W. Haan and R. Zwanzig, Series expansions in a continuum percolation problem, *J. Phys. A* 10 (1977) 1547–1555.
[63] E. T. Gawlinski and H. E. Stanley, Continuum percolation in two dimensions: Monte Carlo tests of scaling and universality for noninteracting discs, *J. Phys. A* 14 (1981) L291–L299; MR 82m:82026.
[64] J. Quintanilla, S. Torquato, and R. M. Ziff, Efficient measurement of the percolation threshold for fully penetrable discs, *J. Phys. A* 33 (2000) L399–L407.
[65] C. D. Lorenz and R. M. Ziff, Precise determination of the critical percolation threshold for the three-dimensional "Swiss cheese" model using a growth algorithm, *J. Chem. Phys.* 114 (2001) 3659–3661.
[66] P. Hall, *Introduction to the Theory of Coverage Processes*, Wiley, 1988; MR 90f:60024.
[67] I. Vardi, Prime percolation, *Experim. Math.* 7 (1998) 275–289; MR 2000h:11081.

5.19 Klarner's Polyomino Constant

A **domino** is a pair of adjacent squares. Generalizing, we say that a **polyomino** or **lattice animal** of order n is a connected set of n adjacent squares [1–7]. See Figures 5.17 and 5.18.

Define $A(n)$ to be the number of polyominoes of order n, where it is agreed that two polyominoes are distinct if and only if they have different shapes *or* different

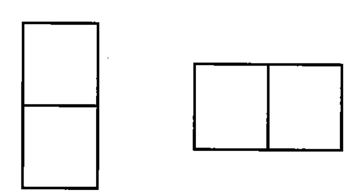


Figure 5.17. All dominoes (polyominoes of order 2); $A(2) = 2$.

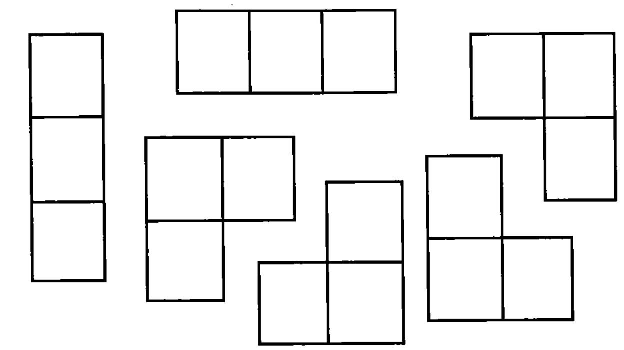


Figure 5.18. All polyominoes of order 3; $A(3) = 6$.

orientations:

$$A(1) = 1, \ A(2)=2, \ A(3)=6, \ A(4)=19, \ A(5)=63, \ A(6) = 216, \ A(7)=760, \ldots.$$

There are different senses in which polyominoes are defined, for example, free versus fixed, bond versus site, simply-connected versus not necessarily so, and others. For brevity, we focus only on the fixed, site, possibly multiply-connected case.

Redelmeier [8] computed $A(n)$ up to $n = 24$, and Conway & Guttmann [9] found $A(25)$. In a recent flurry of activity, Oliveira e Silva [10] computed $A(n)$ up to $n = 28$, Jensen & Guttmann [11, 12] extended this to $A(46)$, and Knuth [13] found $A(47)$. Klarner [14, 15] proved that the limit

$$\alpha = \lim_{n\to\infty} A(n)^{\frac{1}{n}} = \sup_n A(n)^{\frac{1}{n}}$$

exists and is nonzero, although Eden [16] numerically investigated α several years earlier. The best-known bounds on α are $3.903184 \le \alpha \le 4.649551$, as discussed in [17–20]. Improvements are possible using the new value $A(47)$. The best-known estimate, obtained via series expansion analysis by differential approximants [11], is $\alpha = 4.062570\ldots$. A more precise asymptotic expression for $A(n)$ is

$$A(n) \sim \left(\frac{0.316915\ldots}{n} - \frac{0.276\ldots}{n^{3/2}} + \frac{0.335\ldots}{n^2} - \frac{0.25\ldots}{n^{5/2}} + O\left(\frac{1}{n^3}\right)\right)\alpha^n,$$

but such an empirical result is far from being rigorously proved.

Satterfield [5, 21] reported a lower bound of 3.91336 for α, using one of several algorithms he developed with Klarner and Shende. Details of their work unfortunately remain unpublished.

We mention that parallel analysis can be performed on the triangular and hexagonal lattices [7, 22].

Any self-avoiding polygon [5.10] determines a polyomino, but the converse is false since a polyomino can possess holes. A polyomino is **row-convex** if every (horizontal) row consists of a single strip of squares, and it is **convex** if this requirement is met for every column as well. Note that a convex polyomino does not generally determine a convex polygon in the usual sense. Counts of row-convex polyominoes obey a third-order linear recurrence [23–28], but counts $\tilde{A}(n)$ of convex polyominoes are somewhat more difficult to analyze [29, 30]:

$$\tilde{A}(1) = 1,\ \tilde{A}(2) = 2,\ \tilde{A}(3) = 6,\ \tilde{A}(4) = 19,\ \tilde{A}(5) = 59,$$
$$\tilde{A}(6) = 176,\ \tilde{A}(7) = 502, \ldots,$$
$$\tilde{A}(n) \sim (2.67564\ldots)\tilde{\alpha}^n,$$

where $\tilde{\alpha} = 2.3091385933\ldots = (0.4330619231\ldots)^{-1}$. Exact generating function formulation for $\tilde{A}(n)$ was discovered only recently [31–33] but is too complicated to include here. Bender [30] further analyzed the expected shape of convex polyominoes, finding that, when viewed from a distance, most convex polyominoes resemble rods tilted 45° from the vertical with horizontal (and vertical) thickness roughly equal to $2.37597\ldots$. More results like this are found in [34–36].

It turns out that the growth constant $\tilde{\alpha}$ for convex polyominoes is the same as the growth constant α' for **parallelogram polyominoes**, that is, polyominoes whose left and right boundaries both climb in a northeasterly direction:

$$A'(1) = 1,\ A'(2) = 2,\ A'(3) = 4,\ A'(4) = 9,\ A'(5) = 20,$$
$$A'(6) = 46,\ A'(7) = 105, \ldots.$$

These have the virtue of a simpler generating function $f(q)$. Let $(q)_0 = 1$ and $(q)_n = \prod_{j=1}^{n}(1 - q^j)$; then $f(q)$ is a ratio $J_1(q)/J_0(q)$ of q-analogs of Bessel functions:

$$J_0(q) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n q^{\binom{n+1}{2}}}{(q)_n (q)_n}, \quad J_1(q) = -\sum_{n=1}^{\infty} \frac{(-1)^n q^{\binom{n+1}{2}}}{(q)_{n-1}(q)_n},$$

which gives $\alpha' = \tilde{\alpha}$, but a different multiplicative constant $0.29745\ldots$.

There are many more counting problems of this sort than we can possibly summarize! Here is one more example, studied independently by Glasser, Privman & Svrakic [37] and Odlyzko & Wilf [38–40]. An ***n*-fountain** (Figure 5.19) is best pictured as a connected, self-supporting stacking of n coins in a triangular lattice array against a vertical wall.

Note that the bottom row cannot have gaps but the higher rows can; each coin in a higher row must touch two adjacent coins in the row below. Let $B(n)$ be the number of n-fountains. The generating function for $B(n)$ satisfies a beautiful identity involving

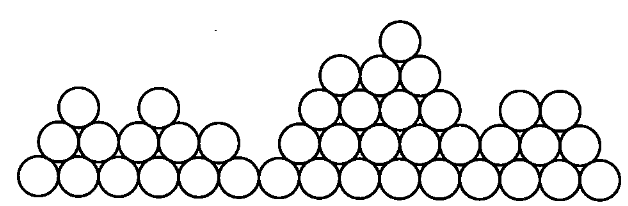


Figure 5.19. An example of an n-fountain.

Ramanujan's continued fraction:

$$1 + \sum_{n=1}^{\infty} B(n)x^n = 1 + x + x^2 + 2x^3 + 3x^4 + 5x^5 + 9x^6 + 15x^7 + 26x^8 + 45x^9 + \cdots$$
$$= \frac{1|}{|1} - \frac{x|}{|1} - \frac{x^2|}{|1} - \frac{x^3|}{|1} - \frac{x^4|}{|1} - \frac{x^5|}{|1} - \cdots,$$

and the following growth estimates arise:

$$\lim_{n\to\infty} B(n)^{\frac{1}{n}} = \beta = 1.7356628245\ldots = (0.5761487691\ldots)^{-1},$$
$$B(n) = (0.3123633245\ldots)\beta^n + O\left(\left(\frac{5}{3}\right)^n\right).$$

See [41] for other related counting problems.

[1] S. W. Golomb, *Polyominoes*, 2nd ed., Princeton Univ. Press, 1994; MR 95k:00006.

[2] A. J. Guttmann, Planar polygons: Regular, convex, almost convex, staircase and row convex, *Computer-Aided Statistical Physics*, Proc. 1991 Taipei symp., ed. C.-K. Hu, Amer. Inst. Phys., 1992, pp. 12–43.

[3] M. P. Delest, Polyominoes and animals: Some recent results, *J. Math. Chem.* 8 (1991) 3–18; MR 93e:05024.

[4] C. Godsil, M. Grötschel, and D. J. A. Welsh, Combinatorics in statistical physics, *Handbook of Combinatorics*, v. II, ed. R. Graham, M. Grötschel, and L. Lovász, MIT Press, 1995, pp. 1925–1954; MR 96h:05001.

[5] D. A. Klarner, Polyominoes, *Handbook of Discrete and Computational Geometry*, ed. J. E. Goodman and J. O'Rourke, CRC Press, 1997, pp. 225–240; MR 2000j:52001.

[6] D. Eppstein, Polyominoes and Other Animals (Geometry Junkyard).

[7] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001168, A001169, A001207, A001420, A005169, A006958, and A067675.

[8] D. H. Redelmeier, Counting polyominoes: Yet another attack, *Discrete Math.* 36 (1981) 191–203; MR 84g:05049.

[9] A. R. Conway and A. J. Guttmann, On two-dimensional percolation, *J. Phys. A* 28 (1995) 891–904.

[10] T. Oliveira e Silva, Enumeration of animals on the (4^4) regular tiling of the Euclidean plane, unpublished note (1999).

[11] I. Jensen and A. J. Guttmann, Statistics of lattice animals (polyominoes) and polygons, *J. Phys. A* 33 (2000) L257–L263; MR 2001e:82025.

[12] I. Jensen, Enumerations of lattice animals and trees, *J. Stat. Phys.* 102 (2001) 865–881; MR 2002b:82026.

[13] D. E. Knuth, The number of 47-ominoes, unpublished note (2001).

[14] D. A. Klarner, Cell growth problems, *Canad. J. Math.* 19 (1967) 851–863; MR 35 #5339.

[15] D. A. Klarner, My life among the polyominoes, *Nieuw Arch. Wisk.* 29 (1981) 156–177; also in *The Mathematical Gardner*, ed. D. A. Klarner, Wadsworth, 1981, pp. 243–262; MR 83e:05038.

[16] M. Eden, A two-dimensional growth process, *Proc. Fourth Berkeley Symp. Math. Stat. Probab.*, v. 4, ed. J. Neyman, Univ. of Calif. Press, 1961, pp. 223–239; MR 24 #B2493.

[17] D. A. Klarner and R. L. Rivest, A procedure for improving the upper bound for the number of n-ominoes, *Canad. J. Math.* 25 (1973) 585–602; MR 48 #1943.

[18] B. M. I. Rands and D. J. A. Welsh, Animals, trees and renewal sequences, *IMA J. Appl. Math.* 27 (1981) 1–17; corrigendum 28 (1982) 107; MR 82j:05049 and MR 83c:05041.

[19] S. G. Whittington and C. E. Soteros, Lattice animals: Rigorous results and wild guesses, *Disorder in Physical Systems: A Volume in Honour of J. M. Hammersley*, ed. G. R. Grimmett and D. J. A. Welsh, Oxford Univ. Press, 1990, pp. 323–335; MR 91m:82061.

[20] A. J. Guttmann, On the number of lattice animals embeddable in the square lattice, *J. Phys. A* 15 (1982) 1987–1990.
[21] D. A. Klarner, S. Shende, and W. Satterfield, Lower-bounding techniques for lattice animals, unpublished work (2000).
[22] W. F. Lunnon, Counting hexagonal and triangular polyominoes, *Graph Theory and Computing*, ed. R. C. Read, Academic Press, 1972, pp. 87–100; MR 49 #2439.
[23] G. Pólya, On the number of certain lattice polygons, *J. Combin. Theory* 6 (1969) 102–105; also in *Collected Papers*, v. 4, ed. G.-C. Rota, MIT Press, 1984, pp. 441–444, 630; MR 38 #4329.
[24] H. N. V. Temperley, Combinatorial problems suggested by the statistical mechanics of domains and of rubber-like molecules, *Phys. Rev.* 103 (1956) 1–16; MR 17,1168f.
[25] D. A. Klarner, Some results concerning polyominoes, *Fibonacci Quart.* 3 (1965) 9–20; MR 32 #4028.
[26] M. P. Delest, Generating functions for column-convex polyominoes, *J. Combin. Theory Ser. A* 48 (1988) 12–31; MR 89e:05013.
[27] V. Domocos, A combinatorial method for the enumeration of column-convex polyominoes, *Discrete Math.* 152 (1996) 115–123; MR 97c:05005.
[28] D. Hickerson, Counting horizontally convex polyominoes, *J. Integer Seq.* 2 (1999) 99.1.8; MR 2000k:05023.
[29] D. A. Klarner and R. L. Rivest, Asymptotic bounds for the number of convex n-ominoes, *Discrete Math.* 8 (1974) 31–40; MR 49 #91.
[30] E. A. Bender, Convex n-ominoes, *Discrete Math.* 8 (1974) 219–226; MR 49 #67.
[31] M. Bousquet-Mélou and J.-M. Fédou, The generating function of convex polyominoes: The resolution of a q-differential system, *Discrete Math.* 137 (1995) 53–75; MR 95m:05009.
[32] M. Bousquet-Mélou and A. Rechnitzer, Lattice animals and heaps of dimers, *Discrete Math.*, to appear.
[33] K. Y. Lin, Exact solution of the convex polygon perimeter and area generating function, *J. Phys. A* 24 (1991) 2411–2417; MR 92f:82023.
[34] G. Louchard, Probabilistic analysis of some (un)directed animals, *Theoret. Comput. Sci.* 159 (1996) 65–79; MR 97b:68164.
[35] G. Louchard, Probabilistic analysis of column-convex and directed diagonally-convex animals, *Random Structures Algorithms* 11 (1997) 151–178; MR 99c:05047.
[36] G. Louchard, Probabilistic analysis of column-convex and directed diagonally-convex animals. II: Trajectories and shapes, *Random Structures Algorithms* 15 (1999) 1–23; MR 2000e:05042.
[37] M. L. Glasser, V. Privman, and N. M. Svrakic, Temperley's triangular lattice compact cluster model: Exact solution in terms of the q series, *J. Phys. A* 20 (1987) L1275–L1280; MR 89b:82107.
[38] R. K. Guy, The strong law of small numbers, *Amer. Math. Monthly* 95 (1988) 697–712; MR 90c:11002.
[39] A. M. Odlyzko and H. S. Wilf, The editor's corner: n coins in a fountain, *Amer. Math. Monthly* 95 (1988) 840–843.
[40] A. M. Odlyzko, Asymptotic enumeration methods, *Handbook of Combinatorics*, v. I, ed. R. Graham, M. Grötschel, and L. Lovász, MIT Press, 1995, pp. 1063–1229; MR 97b:05012.
[41] V. Privman and N. M. Svrakic, Difference equations in statistical mechanics. I: Cluster statistics models, *J. Stat. Phys.* 51 (1988) 1091–1110; MR 89i:82079.

5.20 Longest Subsequence Constants

5.20.1 Increasing Subsequences

Let π denote a random permutation on the symbols $1, 2, \ldots, N$. An **increasing subsequence** of π is a sequence $(\pi(j_1), \pi(j_2), \ldots, \pi(j_k))$ satisfying both $1 \le j_1 < j_2 < \ldots < j_k \le N$ and $\pi(j_1) < \pi(j_2) < \ldots < \pi(j_k)$. Define L_N to be the length of the

longest increasing subsequence of π. For example, the permutation $\pi = (2, 7, 4, 1, 6, 3, 9, 5, 8)$ has longest increasing subsequences $(2, 4, 6, 9)$ and $(1, 3, 5, 8)$; hence $L_9 = 4$. What can be said about the probability distribution of L_N (e.g., its mean and variance) as $N \to \infty$?

This question has inspired an avalanche of research [1–4]. Vershik & Kerov [5] and Logan & Shepp [6] proved that

$$\lim_{N\to\infty} N^{-\frac{1}{2}}\, \mathrm{E}(L_N) = 2,$$

building upon earlier work in [7–10]. Odlyzko & Rains [11] conjectured in 1993 that both limits

$$\lim_{N\to\infty} N^{-\frac{1}{3}}\, \mathrm{Var}(L_N) = c_0, \quad \lim_{N\to\infty} N^{-\frac{1}{6}}(\mathrm{E}(L_N) - 2\sqrt{N}) = c_1$$

exist and are finite and nonzero; numerical approximations were computed via Monte Carlo simulation. In a showcase of analysis (using methods from mathematical physics), Baik, Deift & Johansson [12] obtained

$$c_0 = 0.81318\ldots \text{ (i.e., } \sqrt{c_0} = 0.90177\ldots), \quad c_1 = -1.77109\ldots,$$

confirming the predictions in [11]. These constants are defined exactly in terms of the solution to a Painlevé II equation. (Incidently, Painlevé III arises in [5.22] and Painlevé V arises in [2.15.3].) The derivation involves a relationship between random matrices and random permutations [13, 14]. More precisely, Tracy & Widom [15–17] derived a certain probability distribution function $F(x)$ characterizing the largest eigenvalue of a random Hermitian matrix, generated according to the Gaussian Unitary Ensemble (GUE) probability law. Baik, Deift & Johansson proved that the limiting distribution of L_N is Tracy & Widom's $F(x)$, and then obtained estimates of the constants c_0 and c_1 via moments quoted in [16].

Before presenting more details, we provide a generalization. A **2-increasing subsequence** of π is a union of two disjoint increasing subsequences of π. Define $\tilde{L}_N$ to be the length of the longest 2-increasing subsequence of π, minus L_N. For example, the permutation $\pi = (2, 4, 7, 9, 5, 1, 3, 6, 8)$ has longest increasing subsequence $(2, 4, 5, 6, 8)$ and longest 2-increasing subsequence $(2, 4, 7, 9) \cup (1, 3, 6, 8)$; hence $\tilde{L}_9 = 8 - 5 = 3$. As before, both

$$\lim_{N\to\infty} N^{-\frac{1}{3}}\, \mathrm{Var}(\tilde{L}_N) = \tilde{c}_0, \quad \lim_{N\to\infty} N^{-\frac{1}{6}}(\mathrm{E}(\tilde{L}_N) - 2\sqrt{N}) = \tilde{c}_1$$

exist and can be proved [18] to possess values

$$\tilde{c}_0 = 0.5405\ldots, \quad \tilde{c}_1 = -3.6754\ldots.$$

The corresponding distribution function $\tilde{F}(x)$ characterizes the second-largest eigenvalue of a random Hermitian matrix under GUE. Such proofs were extended to m-increasing subsequences, for arbitrary $m > 2$, and to the joint distribution of row lengths from random Young tableaux in [19–21].

Here are the promised details [12, 18]. Fix $0 < t \le 1$. Let $q_t(x)$ be the solution of the Painlevé II differential equation

$$q_t''(x) = 2q_t(x)^3 + xq_t(x), \quad q_t(x) \sim \frac{1}{2}\left(\frac{t}{\pi}\right)^{\frac{1}{2}} x^{-\frac{1}{4}} \exp\left(-\frac{2}{3}x^{\frac{3}{2}}\right) \text{ as } x \to \infty,$$

and define

$$\Phi(x,t) = \exp\left[-\int_x^{\infty}(y-x)q_t(y)^2 dy\right].$$

The Tracy-Widom functions are

$$F(x) = \Phi(x,1), \quad \tilde{F}(x) = \Phi(x,1) - \frac{\partial\Phi}{\partial t}(x,t)\bigg|_{t=1}$$

and hence

$$c_0 = \int_{-\infty}^{\infty} x^2 F'(x)dx - \left(\int_{-\infty}^{\infty} xF'(x)dx\right)^2, \quad c_1 = \int_{-\infty}^{\infty} xF'(x)dx,$$

$$\tilde{c}_0 = \int_{-\infty}^{\infty} x^2 \tilde{F}'(x)dx - \left(\int_{-\infty}^{\infty} x\tilde{F}'(x)dx\right)^2, \quad \tilde{c}_1 = \int_{-\infty}^{\infty} x\tilde{F}'(x)dx$$

are the required formulas. Note that the values of c_0, c_1, $\tilde{c}_0$, and $\tilde{c}_1$ appear in the caption of Figure 2 of [16]. Hence these arguably should be called Odlyzko–Rains–Tracy–Widom constants.

What makes this work especially exciting [1, 22] is its connection with the common cardgame of solitaire (for which no successful analysis has yet been performed) and possibly with the unsolved Riemann hypothesis [1.6] from prime number theory. See [23, 24] for other applications.

5.20.2 Common Subsequences

Let a and b be random sequences of length n, with terms a_i and b_j taking values from the alphabet $\{0, 1, \ldots, k-1\}$. A sequence c is a **common subsequence** of a and b if c is a subsequence of both a and b, meaning that c is obtained from a by deleting zero or more terms a_i and from b by deleting zero or more terms b_j. Define $\lambda_{n,k}$ to be the length of the longest common subsequence of a and b. For example, the sequences $a = (1, 0, 0, 2, 3, 2, 1, 1, 0, 2)$, $b = (0, 1, 1, 1, 3, 3, 3, 0, 2, 1)$ have longest common subsequence $c = (0, 1, 1, 0, 2)$ and $\lambda_{10,3} = 5$. What can be said about the mean of $\lambda_{n,k}$ as $n \to \infty$, as a function of k?

It can be proved that $\mathrm{E}(\lambda_{n,k})$ is superadditive with respect to n, that is, $\mathrm{E}(\lambda_{m,k}) + \mathrm{E}(\lambda_{n,k}) \le \mathrm{E}(\lambda_{m+n,k})$. Hence, by Fekete's theorem [25, 26], the limit

$$\gamma_k = \lim_{n\to\infty}\frac{\mathrm{E}(\lambda_{n,k})}{n} = \sup_n \frac{\mathrm{E}(\lambda_{n,k})}{n}$$

Table 5.5. *Estimates for Ratios γ_k*

k	Lower Bound	Numerical Estimate	Upper Bound
2	0.77391	0.8118	0.83763
3	0.63376	0.7172	0.76581
4	0.55282	0.6537	0.70824
5	0.50952	0.6069	0.66443

exists. Beginning with Chvátal & Sankoff [27–30], a number of researchers [31–37] have investigated γ_k. Table 5.5 contains rigorous lower and upper bounds for γ_k, as well as the best numerical estimates of γ_k presently available [37].

It is known [27, 31] that $1 \le \gamma_k\sqrt{k} \le e$ for all k and conjectured [38] that $\lim_{k\to\infty}\gamma_k\sqrt{k} = 2$. There is interest in the rate of convergence of the limiting ratio [39–41]

$$\gamma_k n - O(\sqrt{n\ln(n)}) \le \mathrm{E}(\lambda_{n,k}) \le \gamma_k n$$

as well as in $\mathrm{Var}(\lambda_{n,k})$, which is conjectured [39, 41, 42] to grow linearly with n.

A sequence c is a **common supersequence** of a and b if c is a supersequence of both a and b, meaning that both a and b are subsequences of c. The shortest common subsequence length $\Lambda_{n,k}$ of a and b can be shown [34, 43, 44] to satisfy

$$\lim_{n\to\infty}\frac{\mathrm{E}(\Lambda_{n,k})}{n} = 2 - \gamma_k.$$

Such nice duality as this fails, however, if we seek longest subsequences/shortest supersequences from a set of > 2 random sequences.

[1] D. Aldous and P. Diaconis, Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem, *Bull. Amer. Math. Soc.* 36 (1999) 413–432; MR 2000g:60013.

[2] P. Deift, Integrable systems and combinatorial theory, *Notices Amer. Math. Soc.* 47 (2000) 631–640; MR 2001g:05012.

[3] C. A. Tracy and H. Widom, Universality of the distribution functions of random matrix theory, *Integrable Systems: From Classical to Quantum*, Proc. 1999 Montréal conf., ed. J. Harnad, G. Sabidussi, and P. Winternitz, Amer. Math. Soc, 2000, pp. 251–264; math-ph/9909001; MR 2002f:15036.

[4] R. P. Stanley, Recent progress in algebraic combinatorics, presentation at *Mathematical Challenges for the 21st Century* conf., Los Angeles, 2000; *Bull. Amer. Math. Soc.*, to appear; math.CO/0010218.

[5] A. M. Vershik and S. V. Kerov, Asymptotic behavior of the Plancherel measure of the symmetric group and the limit form of Young tableaux (in Russian), *Dokl. Akad. Nauk SSSR* 233 (1977) 1024–1027; Engl. transl. in *Soviet Math. Dokl.* 233 (1977) 527–531; MR 58 #562.

[6] B. F. Logan and L. A. Shepp, A variational problem for random Young tableaux, *Adv. Math.* 26 (1977) 206–222; MR 98e:05108.

[7] S. M. Ulam, Monte Carlo calculations in problems of mathematical physics, *Modern Mathematics for the Engineer: Second Series*, ed. E. F. Beckenbach, McGraw-Hill, 1961, pp. 261–281; MR 23 #B2202.

[8] R. M. Baer and P. Brock, Natural sorting over permutation spaces, *Math. Comp.* 22 (1968) 385–410; MR 37 #3800.

[9] 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; MR 53 #9457.
[10] J. F. C. Kingman, Subadditive ergodic theory, *Annals of Probab.* 1 (1973) 883–909; MR 50 #8663.
[11] A. M. Odlyzko and E. M. Rains, On longest increasing subsequences in random permutations, *Analysis, Geometry, Number Theory: The Mathematics of Leon Ehrenpreis*, Proc. 1998 Temple Univ. conf., ed. E. L. Grinberg, S. Berhanu, M. Knopp, G. Mendoza, and E. T. Quinto, Amer. Math. Soc., 2000, pp. 439–451; MR 2001d:05003.
[12] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of random permutations, *J. Amer. Math. Soc.* 12 (1999) 1119–1178; math.CO/9810105; MR 2000e:05006.
[13] C. A. Tracy and H. Widom, Random unitary matrices, permutations and Painlevé, *Commun. Math. Phys.* 207 (1999) 665–685; math.CO/9811154; MR 2001h:15019.
[14] E. M. Rains, Increasing subsequences and the classical groups, *Elec. J. Combin.* 5 (1998) R12; MR 98k:05146.
[15] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Phys. Lett. B* 305 (1993) 115–118; hep-th/9210074; MR 94f:82046.
[16] C. A. Tracy and H. Widom, Level-spacing distributions and the Airy kernel, *Commun. Math. Phys.* 159 (1994) 151–174; hep-th/9211141; MR 95e:82003.
[17] C. A. Tracy and H. Widom, The distribution of the largest eigenvalue in the Gaussian ensembles: $\beta = 1, 2, 4$, *Calogero-Moser-Sutherland Models*, Proc. 1997 Montréal conf., ed. J. F. van Diejen and L. Vinet, Springer-Verlag, 2000, pp. 461–472; solv-int/9707001; MR 2002g:82021.
[18] J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the second row of a Young diagram under Plancherel measure, *Geom. Func. Anal.* 10 (2000) 702–731; addendum 10 (2000) 1606–1607; math.CO/9901118; MR 2001m:05258.
[19] A. Okounkov, Random matrices and random permutations, *Int. Math. Res. Notices* (2000) 1043–1095; math.CO/9903176; MR 2002c:15045.
[20] A. Borodin, A. Okounkov, and G. Olshanski, Asymptotics of Plancherel measures for symmetric groups. *J. Amer. Math. Soc.* 13 (2000) 481–515; math.CO/9905032; MR 2001g:05103
[21] K. Johansson, Discrete orthogonal polynomial ensembles and the Plancherel measure, *Annals of Math.* 153 (2001) 259–296; math.CO/9906120; MR 2002g:05188.
[22] D. Mackenzie, From solitaire, a clue to the world of prime numbers, *Science* 282 (1998) 1631–1632; 283 (1999) 794–795.
[23] I. M. Johnstone, On the distribution of the largest principal component, *Annals of Statist.* 29 (2001) 295–327; Stanford Univ. Dept. of Statistics TR 2000-27.
[24] C. A. Tracy and H. Widom, A distribution function arising in computational biology, *MathPhys Odyssey 2001: Integrable Models and Beyond*, Proc. 2001 Okayama/Kyoto conf., ed. M. Kashiwara and T. Miwa, Birkhäuser, 2002, pp. 467–474; math.CO/0011146.
[25] M. Fekete, Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten, *Math. Z.* 17 (1923) 228–249.
[26] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, v. 1, Springer-Verlag, 1972, ex. 98, pp. 23, 198; MR 81e:00002.
[27] V. Chvátal and D. Sankoff, Longest common subsequences of two random sequences, *J. Appl. Probab.* 12 (1975) 306–315; MR 53 #9324.
[28] V. Chvátal and D. Sankoff, An upper-bound technique for lengths of common subsequences, *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison*, ed. D. Sankoff and J. B. Kruskal, Addison-Wesley, 1983, pp. 353–357; MR 85h:68073.
[29] R. Durrett, *Probability: Theory and Examples*, Wadsworth, 1991, pp. 361–373; MR 91m:60002.
[30] I. Simon, Sequence comparison: Some theory and some practice, *Electronic Dictionaries and Automata in Computational Linguistics*, Proc. 1987 St. Pierre d'Oléron conf., ed. M. Gross and D. Perrin, Lect. Notes in Comp. Sci. 377, Springer-Verlag, 1987, pp. 79–92.

[31] J. G. Deken, Some limit results for longest common subsequences, *Discrete Math.* 26 (1979) 17–31; MR 80e:68100.
[32] J. G. Deken, Probabilistic behavior of longest-common-subsequence length, *Time Warps, String Edits, and Macromolecules*, ed. D. Sankoff and J. B. Kruskal, Addison-Wesley, 1983, pp. 359–362; MR 85h:68073.
[33] V. Dancik and M. S. Paterson, Upper bounds for the expected length of a longest common subsequence of two binary sequences, *Random Structures Algorithms* 6 (1995) 449–458; also in *Proc. 1994 Symp. on Theoretical Aspects of Computer Science (STACS)*, Caen, ed. P. Enjalbert, E. W. Mayr, and K. W. Wagner, Lect. Notes in Comp. Sci. 775, Springer-Verlag, 1994, pp. 669–678; MR 95g:68040 and MR 96h:05016.
[34] V. Dancik, *Expected Length of Longest Common Subsequences*, Ph.D. thesis, Univ. of Warwick, 1994.
[35] M. Paterson and V. Dancík, Longest common subsequences, *Mathematical Foundations of Computer Science (MFCS)*, Proc. 1994 Kosice conf., ed. I. Privara, B. Rovan, and P. Ruzicka, Lect. Notes in Comp. Sci. 841, Springer–Verlag, Berlin, 1994, pp. 127–142; MR 95k:68180.
[36] V. Dancík, Upper bounds for the expected length of longest common subsequences, unpublished note (1996); abstract in *Bull. EATCS* 54 (1994) 248.
[37] R. A. Baeza-Yates, R. Gavaldà, G. Navarro, and R. Scheihing, Bounding the expected length of longest common subsequences and forests, *Theory Comput. Sys.* 32 (1999) 435–452; also in *Proc. Third South American Workshop on String Processing (WSP)*, Recife, ed. R. Baeza-Yates, N. Ziviani, and K. Guimarães, Carleton Univ. Press, 1996, pp. 1–15; MR 2000k:68076.
[38] D. Sankoff and S. Mainville, Common subsequences and monotone subsequences, *Time Warps, String Edits, and Macromolecules*, ed. D. Sankoff and J. B. Kruskal, Addison-Wesley, 1983, pp. 363–365; MR 85h:68073.
[39] K. S. Alexander, The rate of convergence of the mean length of the longest common subsequence, *Annals of Appl. Probab.* 4 (1994) 1074–1082; MR 95k:60020.
[40] W. T. Rhee, On rates of convergence for common subsequences and first passage time, *Annals of Appl. Probab.* 5 (1995) 44–48; MR 96e:60023.
[41] J. Boutet de Monvel, Extensive simulations for longest common subsequences: Finite size scaling, a cavity solution, and configuration space properties, *Europ. Phys. J. B* 7 (1999) 293–308; cond-mat/9809280.
[42] M. S. Waterman, Estimating statistical significance of sequence alignments, *Philos. Trans. Royal Soc. London Ser. B* 344 (1994) 383–390.
[43] V. Dancik, Common subsequences and supersequences and their expected length, *Combin. Probab. Comput.* 7 (1998) 365–373; also in *Sixth Combinatorial Pattern Matching Symp. (CPM)*, Proc. 1995 Espoo conf., ed. Z. Galil and E. Ukkonen, Lect. Notes in Comp. Sci. 937, Springer-Verlag, 1995, pp. 55–63; MR 98e:68100 and MR 99m:05007.
[44] T. Jiang and M. Li, On the approximation of shortest common supersequences and longest common subsequences, *SIAM J. Comput.* 24 (1995) 1122–1139; also in *Proc. 1994 Int. Colloq. on Automata, Languages and Programming (ICALP)*, Jerusalem, ed. S. Abiteboul and E. Shamir, Lect. Notes in Comp. Sci. 820, Springer-Verlag, 1994, pp. 191–202; MR 97a:68081.

5.21 k-Satisfiability Constants

Let $x_1, x_2, \ldots, x_n$ be Boolean variables. Choose k elements randomly from the set $\{x_1, \neg x_1, x_2, \neg x_2, \ldots, x_n, \neg x_n\}$ under the restriction that x_j and $\neg x_j$ cannot both be selected. These k **literals** determine a **clause**, which is the disjunction ($\bigvee$, that is, "inclusive or") of the literals.

Perform this selection process m times. The m independent clauses determine a **formula**, which is the conjunction ($\bigwedge$, that is, "and") of the clauses. A sample formula, in the special case $n = 5$, $k = 3$, and $m = 4$, is

$$[x_1 \vee (\neg x_5) \vee (\neg x_2)] \wedge [(\neg x_3) \vee x_2 \vee (\neg x_1)] \wedge [x_5 \vee x_2 \vee x_4] \wedge [x_4 \vee (\neg x_3) \vee x_1].$$

A formula is **satisfiable** if there exists an assignment of 0s and 1s to the xs so that the formula is true (that is, has value 1). The design of efficient algorithms for discovering such an assignment, given a large formula, or for proving that the formula is **unsatisfiable**, is an important topic in theoretical computer science [1–3].

The k-satisfiability problem, or k-SAT, behaves differently for $k = 2$ and $k \geq 3$. For $k = 2$, the problem can be solved by a linear time algorithm, whereas for $k \geq 3$, the problem is NP-complete.

There is another distinction involving ideas from percolation theory [5.18]. As $m \to \infty$ and $n \to \infty$ with limiting ratio $m/n \to r$, empirical evidence suggests that the random k-SAT problem undergoes a phase transition at a critical value $r_c(k)$ of the parameter r. For $r < r_c$, a random formula is satisfiable with probability $\to 1$ as $m, n \to \infty$. For $r > r_c$, a random formula is likewise unsatisfiable almost surely. Away from the boundary, k-SAT is relatively easy to solve; computational difficulties appear to be maximized at the threshold $r = r_c$ itself. This observation may ultimately help in improving algorithms for solving the traveling salesman problem [8.5] and other combinatorial nightmares.

In the case of 2-SAT, it has been proved [4–6] that $r_c(2) = 1$. A rigorous understanding of 2-SAT from a statistical mechanical point-of-view was achieved in [7].

In the case of k-SAT, $k \geq 3$, comparatively little has been proved. Here is an inequality [4] valid for all $k \geq 3$:

$$\frac{3}{8}\frac{2^k}{k} \leq r_c(k) \leq \ln(2) \cdot \ln\left(\frac{2^k}{2^k - 1}\right)^{-1} \sim \ln(2) \cdot 2^k.$$

Many researchers have contributed to placing tight upper bounds [8–16] and lower bounds [17–20] on the 3-SAT threshold:

$$3.26 \leq r_c(3) \leq 4.506.$$

Large-scale computations [21–23] give an estimate $r_c(3) = 4.25\ldots$. Estimates for larger k [1] include $r_c(4) = 9.7\ldots$, $r_c(5) = 20.9\ldots$, and $r_c(6) = 43.2\ldots$, but these can be improved. Unlike 2-SAT, we do not yet possess a proof that $r_c(k)$ exists, for $k \geq 3$, but Friedgut [24] took an important step in this direction. Sharp phase transitions, corresponding to certain properties of random graphs, play an essential role in his paper. The possibility that $r_c(k)$ oscillates between the bounds $O(2^k/k)$ and $O(2^k)$ has not been completely ruled out, but this would be unexpected.

We mention a similar instance of threshold phenomena for random graphs. When $m \to \infty$ and $n \to \infty$ with limiting ratio $m/n \to s$, then in a random graph G on n vertices and with m edges, it appears that G is k-colorable with probability $\to 1$ for $s < s_c(k)$ and G is not k-colorable with probability $\to 1$ for $s > s_c(k)$. As before, the

existence of $s_c(k)$ is only conjectured if $k \geq 3$, but we have bounds [25–33]

$$1.923 \leq s_c(3) \leq 2.495, \quad 2.879 \leq s_c(4) \leq 4.587,$$
$$3.974 \leq s_c(5) \leq 6.948, \quad 5.190 \leq s_c(6) \leq 9.539$$

and an estimate [34] $s_c(3) = 2.3$.

Consider also the discrete n-cube Q of vectors of the form $(\pm1, \pm1, \pm1, \ldots, \pm1)$. The **half cube** H_v generated by any $v \in Q$ is the set of all vectors $w \in Q$ having negative inner product with v. If a vector $u \in H_v$, it is natural to say that H_v **covers** u. Let $v_1, v_2, \ldots, v_m$ be drawn randomly from Q. When $m \to \infty$ and $n \to \infty$ with limiting ratio $m/n \to t$, it appears that $\bigcup_{k=1}^m H_{v_k}$ covers all of Q with probability $\to 1$ for $t > t_c$ but fails to do so with probability $\to 1$ for $t < t_c$. The existence of t_c was conjectured in [35] but a proof is not known. We have bounds [36,37]

$$0.005 \leq t_c \leq 0.9963 = 1 - 0.0037$$

and an estimate [38,39] $t_c = 0.82$. The motivation for studying this problem arises in binary neural networks.

Here is an interesting variation that encompasses both 2-SAT and 3-SAT. Fix a number $0 \leq p \leq 1$. When selecting m clauses at random, choose a 3-clause with probability p and a 2-clause with probability $1 - p$. This is known as $(2 + p)$-SAT and is useful in understanding the onset of complexity when moving from 2-SAT to 3-SAT [3,40–42]. Clearly the critical value for this model satisfies

$$r_c(2 + p) \leq \min\left\{\frac{1}{1-p}, \frac{1}{p} r_c(3)\right\}$$

for all p. Further [43], if $p \leq 2/5$, then with probability $\to 1$, a random $(2 + p)$-SAT formula is satisfiable if and only if its 2-SAT subformula is satisfiable. This is a remarkable result: A random mixture containing 60% 2-clauses and 40% 3-clauses behaves like 2-SAT! Evidence for a conjecture that the critical threshold $p_c = 2/5$ appears in [44]. See also [45].

Another variation involves replacing "inclusive or" when forming clauses by "exclusive or." By way of contrast with k-SAT, $k \geq 3$, the XOR-SAT problem can be solved in polynomial time, and its transition from satisfiability to unsatisfiability is completely understood [46].

[1] S. Kirkpatrick and B. Selman, Critical behavior in the satisfiability of random Boolean expressions, *Science* 264 (1994) 1297–1301; MR 96e:68063.
[2] B. Hayes, Can't get no satisfaction, *Amer. Scientist*, v. 85 (1997) n. 2, 108–112.
[3] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, Determining computational complexity from characteristic phase transitions, *Nature* 400 (1999) 133–137; MR 2000f:68055.
[4] V. Chvátal and B. Reed, Mick gets some (the odds are on his side), *Proc. 33rd Symp. on Foundations of Computer Science (FOCS)*, Pittsburgh, IEEE, 1992, pp. 620–627.
[5] A. Goerdt, A threshold for unsatisfiability, *J. Comput. Sys. Sci.* 53 (1996) 469–486; also in *Mathematical Foundations of Computer Science (MFCS)*, Proc. 1992 Prague conf., ed. I. M. Havel and V. Koubek, Lect. Notes in Comp. Sci. 629, Springer-Verlag, 1992, pp. 264–274; MR 94j:03011 and MR 98i:03012.

[6] W. Fernandez de la Vega, On random 2-SAT, unpublished note (1992).
[7] B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson, The scaling window of the 2-SAT transition, *Random Structures Algorithms* 18 (2001) 201–256; MR 2002a:68052.
[8] J. Franco and M. Paull, Probabilistic analysis of the Davis-Putnam procedure for solving the satisfiability problem, *Discrete Appl. Math.* 5 (1983) 77–87; correction 17 (1987) 295–299; MR 84e:68038 and MR 88e:68050.
[9] A. Z. Broder, A. M. Frieze, and E. Upfal, On the satisfiability and maximum satisfiability of random 3-CNF formulas, *Proc. 4th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, Austin, ACM, 1993, pp. 322–330; MR 94b:03023.
[10] A. El Maftouhi and W. Fernandez de la Vega, On random 3-SAT, *Combin. Probab. Comput.* 4 (1995) 189–195; MR 96f:03007.
[11] A. Kamath, R. Motwani, K. Palem, and P. Spirakis, Tail bounds for occupancy and the satisfiability threshold conjecture, *Random Structures Algorithms* 7 (1995) 59–80; MR 97b:68091.
[12] L. M. Kirousis, E. Kranakis, D. Krizanc, and Y. C. Stamatiou, Approximating the unsatisfiability threshold of random formulas, *Random Structures Algorithms* 12 (1998) 253–269; MR 2000c:68069.
[13] O. Dubois and Y. Boufkhad, A general upper bound for the satisfiability threshold of random r-SAT formulae, *J. Algorithms* 24 (1997) 395–420; MR 98e:68103.
[14] S. Janson, Y. C. Stamatiou, and M. Vamvakari, Bounding the unsatisfiability threshold of random 3-SAT, *Random Structures Algorithms* 17 (2000) 103–116; erratum 18 (2001) 99–100; MR 2001c:68065 and MR 2001m:68064.
[15] A. C. Kaporis, L. M. Kirousis, Y. C. Stamatiou, M. Vamvakari, and M. Zito, Coupon collectors, q-binomial coefficients and the unsatisfiability threshold, *Seventh Italian Conf. on Theoretical Computer Science (ICTCS)*, Proc. 2001 Torino conf., ed. A. Restivo, S. Ronchi Della Rocca, and L. Roversi, Lect. Notes in Comp. Sci. 2202, Springer-Verlag, 2001, pp. 328–338.
[16] O. Dubois, Y. Boufkhad, and J. Mandler, Typical random 3-SAT formulae and the satisfiability threshold, *Proc. 11th ACM-SIAM Symp. on Discrete Algorithms (SODA)*, San Francisco, ACM, 2000, pp. 126–127.
[17] M.-T. Chao and J. Franco, Probabilistic analysis of two heuristics for the 3-satisfiability problem, *SIAM J. Comput.* 15 (1986) 1106–1118; MR 88b:68079.
[18] A. Frieze and S. Suen, Analysis of two simple heuristics on a random instance of k-SAT, *J. Algorithms* 20 (1996) 312–355; MR 97c:68062.
[19] D. Achlioptas, Setting 2 variables at a time yields a new lower bound for random 3-SAT, *Proc. 32nd ACM Symp. on Theory of Computing (STOC)*, Portland, ACM, 2000, pp. 28–37.
[20] D. Achlioptas and G. B Sorkin, Optimal myopic algorithms for random 3-SAT, *Proc. 41st Symp. on Foundations of Computer Science (FOCS)*, Redondo Beach, IEEE, 2000, pp. 590–600.
[21] T. Larrabee and Y. Tsuji, Evidence for a satisfiability threshold for random 3CNF formulas, presentation at *AAAI Spring Symp. on AI and NP-Hard Problems*, Palo Alto, 1993; Univ. of Calif. at Santa Cruz Tech. Report UCSC-CRL-92-42.
[22] B. Selman, D. G. Mitchell, and H. J. Levesque, Generating hard satisfiability problems, *Artificial Intellig.* 81 (1996) 17–29; Hard and easy distributions of SAT problems, *Proc. Tenth Nat. Conf. on Artificial Intelligence*, San Jose, AAAI Press, 1992, pp. 459–465; MR 98b:03013.
[23] J. M. Crawford and L. D. Auton, Experimental results on the crossover point in random 3-SAT, *Artificial Intellig.* 81 (1996) 31–57; also in *Proc. Eleventh Nat. Conf. on Artificial Intelligence,* Washington DC, AAAI Press, 1993, pp. 21–27; MR 97d:03055.
[24] E. Friedgut, Sharp thresholds of graph properties, and the k-SAT problem (appendix by J. Bourgain), *J. Amer. Math. Soc.* 12 (1999) 1017–1054; MR 2000a:05183.
[25] T. Luczak, Size and connectivity of the k-core of a random graph, *Discrete Math.* 91 (1991) 61–68; MR 92m:05171.
[26] V. Chvátal, Almost all graphs with $1.44n$ edges are 3-colorable, *Random Structures Algorithms* 2 (1991) 11–28; MR 92c:05056.

[27] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, *Random Structures Algorithms* 6 (1995) 161–179; MR 97a:05191.
[28] B. Pittel, J. Spencer, and N. Wormald, Sudden emergence of a giant k-core in a random graph, *J. Combin. Theory Ser. B* 67 (1996) 111–151; MR 97e:05176.
[29] M. Molloy, A gap between the appearances of a k-core and a $(k+1)$-chromatic graph, *Random Structures Algorithms* 8 (1996) 159–160.
[30] D. Achlioptas and M. Molloy, The analysis of a list-coloring algorithm on a random graph, *Proc. 38th Symp. on Foundations of Computer Science (FOCS)*, Miami Beach, IEEE, 1997, pp. 204–212.
[31] P. E. Dunne and M. Zito, An improved upper bound on the non-3-colourability threshold, *Inform. Process. Lett.* 65 (1998) 17–23; MR 98i:05073.
[32] D. Achlioptas and M. Molloy, Almost all graphs with $2.522n$ edges are not 3-colorable, *Elec. J. Combin.* 6 (1999) R29; MR 2000e:05140.
[33] A. C. Kaporis, L. M. Kirousis, and Y. C. Stamatiou, A note on the non-colorability threshold of a random graph, *Elec. J. Combin.* 7 (2000) R29; MR 2001e:05116.
[34] T. Hogg and C. P. Williams, The hardest constraint problems: A double phase transition, *Artificial Intellig.* 69 (1994) 359–377.
[35] E. Gardner, Maximum storage capacity in neural networks, *Europhys. Lett.* 4 (1987) 481–485.
[36] J. H. Kim and J. R. Roche, Covering cubes by random half cubes, with applications to binary neural networks, *J. Comput. Sys. Sci.* 56 (1998) 223–252; MR 2000g:68129.
[37] M. Talagrand, Intersecting random half cubes, *Random Structures Algorithms* 15 (1999) 436–449; MR 2000i:60011.
[38] W. Krauth and M. Opper, Critical storage capacity of the $J = \pm 1$ neural network, *J. Phys. A* 22 (1989) L519–L523.
[39] W. Krauth and M. Mézard, Storage capacity of memory networks with binary couplings, *J. Physique* 50 (1989) 3057–3066.
[40] R. Monasson and R. Zecchina, The entropy of the k-satisfiability problem, *Phys. Rev. Lett.* 76 (1996) 3881–3885; cond-mat/9603014; MR 97a:82053.
[41] R. Monasson and R. Zecchina, Statistical mechanics of the random k-satisfiability model, *Phys. Rev. E* 56 (1997) 1357–1370; MR 98g:82022.
[42] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyansky, $(2+p)$-SAT: Relation of typical-case complexity to the nature of the phase transition, *Random Structures Algorithms* 15 (1999) 414–435; cond-mat/9910080; MR 2000i:68076.
[43] D. Achlioptas, L. M. Kirousis, E. Kranakis, and D. Krizanc, Rigorous results for $(2+p)$-SAT, *Theoret. Comput. Sci.* 265 (2001) 109–129; MR 2002g:68047.
[44] G. Biroli, R. Monasson, and M. Weigt, A variational description of the ground state structure in random satisfiability problems, *Europ. Phys. J. B* 14 (2000) 551–568; cond-mat/9907343.
[45] I. Gent and T. Walsh, The SAT phase transition, *Proc. 11th European Conf. on Artificial Intelligence (ECAI)*, Amsterdam, Wiley, 1994, pp. 105–109.
[46] C. Creignou and H. Daudé, Satisfiability threshold for random XOR-CNF formulas, *Discrete Appl. Math.* 96–97 (1999) 41–53; MR 2000i:68072.

5.22 Lenz–Ising Constants

The Ising model is concerned with the physics of phase transitions, for example, the tendency of a magnet to lose strength as it is heated, with total loss occurring above a certain finite critical temperature. This essay can barely introduce the subject. Unlike hard squares [5.12] and percolation clusters [5.18], a concise complete problem statement here is not possible. We are concerned with large arrays of 1s and -1s whose joint distribution passes through a singularity as a parameter T increases. The definition and characterization of the joint distribution is elaborate; our treatment is combinatorial and focuses on series expansions. See [1–10] for background.

Let L denote the regular d-dimensional cubic lattice with $N = n^d$ sites. For example, in two dimensions, L is the $n \times n$ square lattice with $N = n^2$. To eliminate boundary effects, L is wrapped around to form a d-dimensional torus so that, without exception, every site has $2d$ nearest neighbors. This convention leads to negligible error for large N.

5.22.1 Low-Temperature Series Expansions

Suppose that the N sites of L are colored black or white at random. The dN edges of L fall into three categories: black-black, black-white, and white-white. What can be said jointly about the relative numbers of these? Over all possible such colorings, let $A(p, q)$ be the number of colorings for which there are exactly p black sites and exactly q black-white edges. (See Figure 5.20.)

Then, for large enough N [11–14],

$$\begin{aligned}
&A(0,0) = 1 && \text{(all white)},\\
&A(1,2d) = N && \text{(one black)},\\
&A(2,4d-2) = dN && \text{(two black, adjacent)},\\
&A(2,4d) = \tfrac{1}{2}(N-2d-1)N && \text{(two black, not adjacent)},\\
&A(3,6d-4) = (2d-1)dN && \text{(three black, adjacent)}.
\end{aligned}$$

Properties of this sequence can be studied via the bivariate generating function

$$a(x, y) = \sum_{p,q} A(p, q)x^p y^q$$

and the formal power series

$$\begin{aligned}
\alpha(x, y) &= \lim_{n\to\infty} \frac{1}{N} \ln(a(x, y))\\
&= xy^{2d} + dx^2y^{4d-2} - \frac{2d+1}{2}x^2y^{4d} + (2d-1)dx^3y^{6d-4} + \cdots
\end{aligned}$$

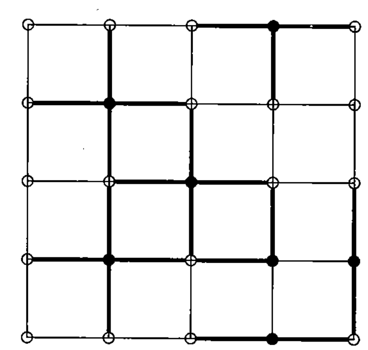


Figure 5.20. Sample coloring with $d = 2$, $N = 25$, $p = 7$, and $q = 21$ (ignoring wraparound).

obtained by merely collecting the coefficients that are linear in N. The latter is sometimes written as [15]

$$\exp(\alpha(x, y)) = 1 + xy^{2d} + dx^2y^{4d-2} - dx^2y^{4d} + (2d-1)dx^3y^{6d-4} + \cdots,$$

a series whose coefficients are integers only. This is what physicists call the **low-temperature series** for the **Ising free energy per site**. The letters x and y are not dummy variables but are related to temperature and magnetic field; the series $\alpha(x, y)$ is not merely a mathematical construct but is a thermodynamic function with properties that can be measured against physical experiment [16]. In the special case when $x = 1$, known as the **zero magnetic field case**, we write $\alpha(y) = \alpha(1, y)$ for convenience.

When $d = 2$, we have [11, 17]

$$\exp(\alpha(y)) = 1 + y^4 + 2y^6 + 5y^8 + 14y^{10} + 44y^{12} + 152y^{14} + 566y^{16} + \cdots.$$

Onsager [18–23] discovered an astonishing closed-form expression:

$$\alpha(y) = \frac{1}{2}\int_0^1\int_0^1 \ln\left[(1+y^2)^2 - 2y(1-y^2)(\cos(2\pi u) + \cos(2\pi v))\right] du\, dv$$

that permits computation of series coefficients to arbitrary order [24] and much more.

When $d = 3$, we have [11, 25–30]

$$\exp(\alpha(y)) = 1 + y^6 + 3y^{10} - 3y^{12} + 15y^{14} - 30y^{16} + 101y^{18} - 261y^{16} + - \cdots$$

No closed-form expression for this series has been found, and the required computations are much more involved than those for $d = 2$.

5.22.2 High-Temperature Series Expansions

The associated high-temperature series arises via a seemingly unrelated combinatorial problem. Let us assume that a nonempty *subgraph* of L is connected and contains at least one edge. Suppose that several subgraphs are drawn on L with the property that

- each edge of L is used at most once, and
- each site of L is used an *even* number of times (possibly zero).

Call such a configuration on L an **even polygonal drawing**. (See Figure 5.21.) An even polygonal drawing is the union of simple, closed, edge-disjoint polygons that need not be connected.

Let $B(r)$ be the number of even polygonal drawings for which there are exactly r edges. Then, for large enough N [4, 11, 31],

$$\begin{aligned}
&B(4) = \tfrac{1}{2}d(d-1)N && \text{(square)},\\
&B(6) = \tfrac{1}{3}d(d-1)(8d-13)N && \text{(two squares, adjacent)},\\
&B(8) = \tfrac{1}{8}d(d-1)\left(d(d-1)N + 216d^2 - 848d + 850\right)N && \text{(many possibilities)}.
\end{aligned}$$

On the one hand, for $d \ge 3$, the drawings can intertwine and be knotted [32], so computing $B(r)$ for larger r is quite complicated! On the other hand, for $d = 2$, clearly

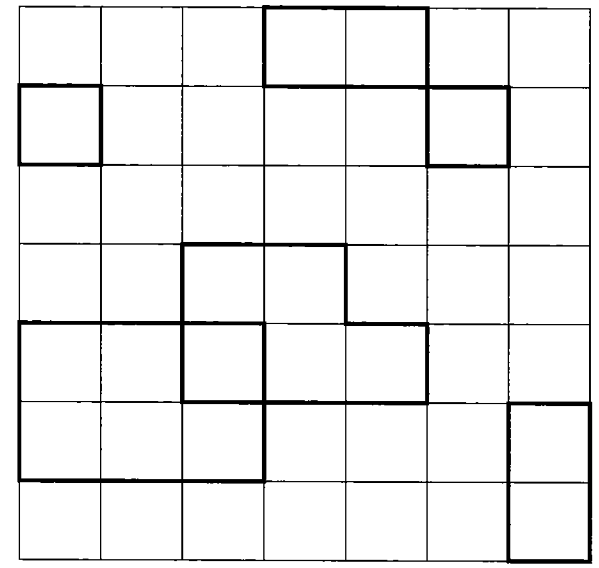


Figure 5.21. An even polygonal drawing for $d = 2$; other names include closed or Eulerian subgraph.

$B(q) = \sum_p A(p, q)$ always. As before, we define a (univariate) generating function

$$b(z) = 1 + \sum_r B(r) z^r$$

and a formal power series

$$\begin{aligned}\beta(z) &= \lim_{n\to\infty} \frac{1}{N} \ln(b(z)) \\ &= \frac{1}{2}d(d-1)z^4 + \frac{1}{3}d(d-1)(8d-13)z^6 + \frac{1}{4}d(d-1)(108d^2 - 424d + 425)z^8 \\ &\quad + \frac{2}{15}d(d-1)(2976d^3 - 19814d^2 + 44956d - 34419)z^{10} + \cdots\end{aligned}$$

called the **high-temperature zero-field series** for the **Ising free energy**. When $d = 3$ [11,25,29,33–36],

$$\exp(\beta(z)) = 1 + 3z^4 + 22z^6 + 192z^8 + 2046z^{10} + 24853z^{12} + 329334z^{14} + \cdots,$$

but again our knowledge of the series coefficients is limited.

5.22.3 Phase Transitions in Ferromagnetic Models

The two major unsolved problems connected to the Ising model are [4,31,37]:

- Find a closed-form expression for $\alpha(x, y)$ when $d = 2$.
- Find a closed-form expression for $\beta(z)$ when $d = 3$.

Why are these so important? We discuss now the underlying physics, as well its relationship to the aforementioned combinatorial problems.

Place a bar of iron in an external magnetic field at constant absolute temperature T. The field will induce a certain amount of magnetization into the bar. If the external field is then slowly turned off, we empirically observe that, for small T, the bar retains some of its internal magnetization, but for large T, the bar's internal magnetization disappears completely.

There is a unique **critical temperature**, T_c, also called the **Curie point**, where this qualitative change in behavior occurs. The Ising model is a simple means for explaining the physical phenomena from a microscopic point of view.

At each site of the lattice L, define a "spin variable" $\sigma_i = 1$ if site i is "up" and $\sigma_i = -1$ if site i is "down." This is known as the **spin-1/2 model**. We study the **partition function**

$$Z(T) = \sum_\sigma \exp\left[\frac{1}{\kappa T}\left(\sum_{(i,j)} \xi \sigma_i \sigma_j + \sum_k \eta \sigma_k\right)\right],$$

where ξ is the coupling (or interaction) constant between nearest neighbor spin variables, $\eta \geq 0$ is the intensity constant of the external magnetic field, and $\kappa > 0$ is Boltzmann's constant.

The function $Z(T)$ captures all of the thermodynamic features of the physical system and acts as a kind of "denominator" when calculating state probabilities. Observe that the first summation is over all 2^N possible values of the vector $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_N)$ and the second summation is over all edges of the lattice (sites i and j are distinct and adjacent). Henceforth we will assume $\xi > 0$, which corresponds to the **ferromagnetic case**. A somewhat different theory emerges in the antiferromagnetic case ($\xi < 0$), which we will not discuss.

How is Z connected to the combinatorial problems discussed earlier? If we assign a spin 1 to the color white and a spin -1 to the color black, then

$$\sum_{(i,j)} \sigma_i \sigma_j = (dN - q)\cdot 1 + q \cdot (-1) = dN - 2q,$$

$$\sum_k \sigma_k = (N - p)\cdot 1 + p \cdot (-1) = N - 2p,$$

and therefore

$$Z = x^{-\frac{1}{2}N} y^{-\frac{d}{2}N} \alpha(x, y),$$

where

$$x = \exp\left(-\frac{2\eta}{\kappa T}\right), \quad y = \exp\left(-\frac{2\xi}{\kappa T}\right).$$

Since small T gives small values of x and y, the phrase low-temperature series for $\alpha(x, y)$ is justified. (Observe that $T = \infty$ corresponds to the case when lattice site colorings are assigned equal probability, which is precisely the combinatorial problem described earlier. The range $0 < T < \infty$ corresponds to unequal weighting, accentuating the states with small p and q. The point $T = 0$ corresponds to an ideal case when all spins are aligned; heat introduces disorder into the system.)

For the high-temperature case, rewrite Z as

$$Z = \left(\frac{4}{(1-z^2)^d(1-w^2)}\right)^{\frac{N}{2}} \frac{1}{2^N} \sum_{\sigma} \left(\prod_{(i,j)} (1 + \sigma_i \sigma_j z) \cdot \prod_k (1 + \sigma_k w) \right),$$

where

$$z = \tanh\left(\frac{\xi}{\kappa T}\right), \quad w = \tanh\left(\frac{\eta}{\kappa T}\right).$$

In the zero-field scenario ($\eta = 0$), this expression simplifies to

$$Z = \left(\frac{4}{(1-z^2)^d}\right)^{\frac{N}{2}} b(z),$$

and since large T gives small z, the phraseology again makes sense.

5.22.4 Critical Temperature

We turn attention to some interesting constants. The radius of convergence y_c in the complex plane of the low-temperature series $\alpha(y) = \sum_{k=0}^{\infty} \alpha_k y^k$ is given by [29]

$$y_c = \lim_{k\to\infty} |\alpha_{2k}|^{-\frac{1}{2k}} = \begin{cases} \sqrt{2} - 1 = 0.4142135623\ldots & \text{if } d = 2, \\ \sqrt{0.2853\ldots} = 0.5341\ldots & \text{if } d = 3; \end{cases}$$

hence, if $d = 2$, the ferromagnetic critical temperature T_c satisfies

$$K_c = \frac{\xi}{\kappa T_c} = \frac{1}{2}\ln\left(\frac{1}{y_c}\right) = \frac{1}{2}\ln(\sqrt{2}+1) = 0.4406867935\ldots.$$

The two-dimensional result is a famous outcome of work by Kramers & Wannier [38] and Onsager [18]. For $d = 3$, the singularity at $y^2 = -0.2853\ldots$ is nonphysical and thus is not relevant to ferromagnetism; a second singularity at $y^2 = 0.412048\ldots$ is what we want but it is difficult to compute directly [29, 39]. To accurately obtain the critical temperature here, we examine instead the high-temperature series $\beta(z) = \sum_{k=0}^{\infty} \beta_k z^k$ and compute

$$z_c = \lim_{k\to\infty} \beta_{2k}^{-\frac{1}{2k}} = 0.218094\ldots, \quad K_c = \tfrac{1}{2}\ln\left(\frac{1+z_c}{1-z_c}\right) = 0.221654\ldots.$$

There is a huge literature of series and Monte Carlo analyses leading to this estimate [40–53]. (A conjectured exact expression for z_c in [54] appears to be false [55].) For $d > 3$, the following estimates are known [56–65]:

$$z_c = \begin{cases} 0.14855\ldots & \text{if } d = 4, \\ 0.1134\ldots & \text{if } d = 5, \\ 0.0920\ldots & \text{if } d = 6, \\ 0.0775\ldots & \text{if } d = 7, \end{cases} \qquad K_c = \begin{cases} 0.14966\ldots & \text{if } d = 4, \\ 0.1139\ldots & \text{if } d = 5, \\ 0.0923\ldots & \text{if } d = 6, \\ 0.0777\ldots & \text{if } d = 7. \end{cases}$$

An associated critical exponent γ will be discussed shortly.

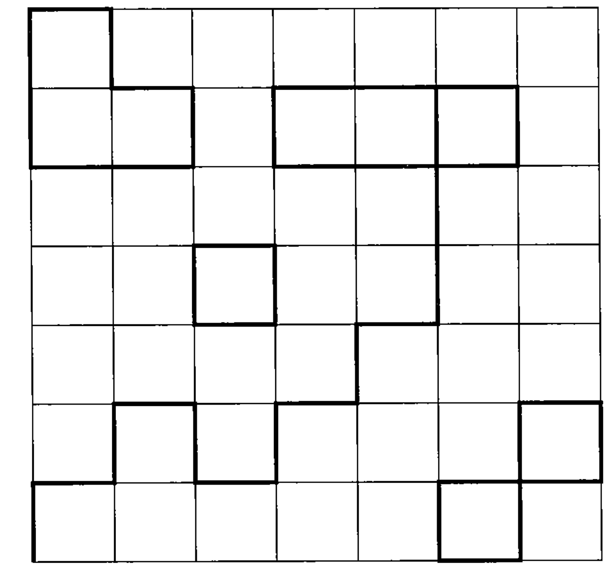


Figure 5.22. An odd polygonal drawing for $d = 2$.

5.22.5 Magnetic Susceptibility

Here is another combinatorial problem. Suppose that several subgraphs are drawn on L with the property that

- each edge of L is used at most once,
- all sites of L, except two, are even, and
- the two remaining sites are odd and must lie in the same (connected) subgraph.

Call this configuration an **odd polygonal drawing**. (See Figure 5.22.) Note that an odd polygonal drawing is the edge-disjoint union of an even polygonal drawing and an (undirected) self-avoiding walk [5.10] linking the two odd sites.

Let $C(r)$ be twice the number of odd polygonal drawings for which there are exactly r edges. Then, for large enough N [12, 66],

$$\begin{aligned} C(1) &= 2dN && \text{(SAW)}, \\ C(2) &= 2d(2d-1)N && \text{(SAW)}, \\ C(3) &= 2d(2d-1)^2 N && \text{(SAW)}, \\ C(4) &= 2d\left(2d(2d-1)^3 - 2d(2d-2)\right)N && \text{(SAW)}, \\ C(5) &= d^2(d-1)N^2 + 2d\left(16d^4 - 32d^3 + 16d^2 + 4d - 3\right)N && \text{(square and/or SAW)}. \end{aligned}$$

As before, we may define a generating function and a formal power series

$$c(z) = N + \sum_r C(r) z^r, \quad \chi(z) = \lim_{n\to\infty} \frac{1}{N}\ln(c(z)) = \sum_{k=0}^{\infty} \chi_k z^k,$$

which is what physicists call the **high-temperature zero-field series** for the **Ising magnetic susceptibility per site**. The radius of convergence z_c of $\chi(z)$ is the same as that for $\beta(z)$ for $d > 1$. For example, when $d = 3$, analyzing the series [67–73]

$$\chi(z) = 1 + 6z + 30z^2 + 150z^3 + 726z^4 + 3510z^5 + 16710z^6 + \cdots$$

is the preferred way to obtain critical parameter estimates (being the best behaved of several available series). Further, the limit

$$\lim_{k\to\infty} \frac{\chi_k}{z_c^{-k} k^{\gamma-1}}$$

appears to exist and is nonzero for a certain positive constant γ depending on dimensionality. As an example, if $d = 2$, numerical evidence surrounding the series [67,74,75]

$$\chi(z) = 1 + 4z + 12z^2 + 36z^3 + 100z^4 + 276z^5 + 740z^6 + 1972z^7 + 5172z^8 + \cdots$$

suggests that the **critical susceptibility exponent** γ is 7/4 and that γ is *universal* (in the sense that it is independent of the choice of lattice). No analogous exact expressions appear to be valid for γ when $d \geq 3$; for $d = 3$, the consensus is that $\gamma = 1.238\ldots$ [40,44,46,49–52,71,73].

We finally make explicit the association of $\chi(z)$ with the Ising model [76]:

$$\lim_{n\to\infty} \frac{1}{N} \ln(Z(z, w)) = \ln(2) - \frac{d}{2}\ln(1 - z^2) - \frac{1}{2}\ln(1 - w^2) + \beta(z) + \frac{1}{2}(\chi(z) - 1)\, w^2 + O(w^4),$$

where the big O depends on z. Therefore $\chi(z)$ occurs when evaluating a second derivative with respect to w, specifically, when computing the variance of P (defined momentarily).

5.22.6 *Q and P Moments*

Let us return to the random coloring problem, suitably generalized to incorporate temperature. Let

$$Q = d - \frac{2}{N} q = \frac{1}{N} \sum_{(i,j)} \sigma_i \sigma_j, \quad P = 1 - \frac{2}{N} p = \frac{1}{N} \sum_{k} \sigma_k$$

for convenience and assume henceforth that $d = 2$. To study the asymptotic distribution of Q, define

$$F(z) = \lim_{n\to\infty} \frac{1}{N} \ln(Z(z)).$$

Then clearly

$$\lim_{n\to\infty} \mathrm{E}(Q) = (\kappa T)\frac{dF}{d\xi}, \quad \lim_{n\to\infty} N\,\mathrm{Var}(Q) = (\kappa T)^2 \frac{d^2 F}{d\xi^2}$$

via term-by-term differentiation of $\ln(Z)$. Exact expressions for both moments are

possible using Onsager's formula:

$$F(z) = \ln\left(\frac{2}{1 - z^2}\right) + \frac{1}{2}\int_0^1\!\!\int_0^1 \ln\left[(1 + z^2)^2 - 2z(1 - z^2)(\cos(2\pi u) + \cos(2\pi v))\right] du\, dv,$$

but we give results at only two special temperatures. In the case $T = \infty$, for which states are assigned equal weighting, $\mathrm{E}(Q) \to 0$ and $N\,\mathrm{Var}(Q) \to 2$, confirming reasoning in [77]. In the case $T = T_c$, note that the singularity is fairly subtle since F and its first derivative are both well defined [11]:

$$F(z_c) = \frac{\ln(2)}{2} + \frac{2G}{\pi} = 0.9296953983\ldots = \frac{1}{2}(\ln(2) + 1.1662436161\ldots),$$

$$\lim_{n\to\infty} \mathrm{E}(Q) = \sqrt{2},$$

where G is Catalan's constant [1.7]. The second derivative of F, however, is unbounded in the vicinity of $z = z_c$ and, in fact [5],

$$\lim_{n\to\infty} N\,\mathrm{Var}(Q) \approx -\frac{8}{\pi}\left(\ln\left|\frac{T}{T_c} - 1\right| + g\right),$$

where g is the constant

$$g = 1 + \frac{\pi}{4} + \ln\left(\frac{\sqrt{2}}{4}\ln(\sqrt{2} + 1)\right) = 0.6194036984\ldots.$$

This is related to what physicists call the **logarithmic divergence** of the **Ising specific heat**. (See Figure 5.23.)

As an aside, we mention that corresponding values of $F(z_c)$ on the triangular and hexagonal planar lattices are, respectively [11],

$$\ln(2) + \frac{\ln(3)}{4} + \frac{H}{2} = 0.8795853862\ldots,$$

$$\frac{3\ln(2)}{4} + \frac{\ln(3)}{2} + \frac{H}{4} = 1.0250590965\ldots.$$

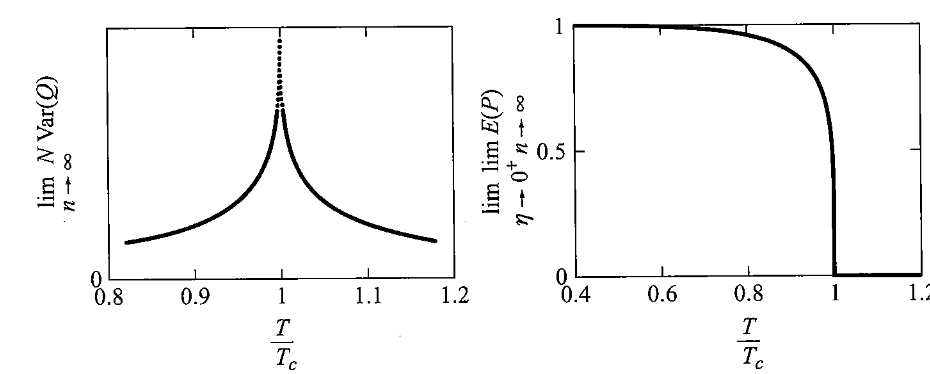


Figure 5.23. Graphs of Ising specific heat and spontaneous magnetization.

Both results feature a new constant [78, 79]:

$$H = \frac{5\sqrt{3}}{6\pi}\psi'\left(\frac{1}{3}\right) - \frac{5\sqrt{3}}{9}\pi - \ln(6) = \frac{\sqrt{3}}{6\pi}\psi'\left(\frac{1}{6}\right) - \frac{\sqrt{3}}{3}\pi - \ln(6)$$
$$= -0.1764297331\ldots,$$

where $\psi'(x)$ is the trigamma function (derivative of the digamma function $\psi(x)$ [1.5.4]). See [80–82] for other occurrences of H; note that the formula

$$\ln(2) + \ln(3) + H = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \ln\left[6 - 2\cos(\theta) - 2\cos(\varphi) - 2\cos(\theta+\varphi)\right]\, d\theta\, d\varphi$$
$$= \frac{3\sqrt{3}}{\pi}\left(1 - \frac{1}{5^2} + \frac{1}{7^2} - \frac{1}{11^2} + \frac{1}{13^2} - \frac{1}{17^2} + - \cdots\right)$$
$$= 1.6153297360\ldots$$

parallels nicely similar results in [3.10] and [5.23].

A more difficult analysis allows us to compute the corresponding two moments of P and also to see more vividly the significance of magnetic susceptibility and critical exponents. Let

$$F(z, w) = \lim_{n\to\infty}\frac{1}{N}\ln(Z(z, w));$$

then clearly

$$\lim_{\eta\to 0^+}\lim_{n\to\infty} \mathrm{E}(P) = (\kappa T)\left.\frac{\partial F}{\partial \eta}\right|_{\eta=0}, \quad \lim_{\eta\to 0^+}\lim_{n\to\infty} N\,\mathrm{Var}(P) = (\kappa T)^2\left.\frac{\partial^2 F}{\partial \eta^2}\right|_{\eta=0}$$

as before. Of course, we do not know $F(z, w)$ exactly when $w \neq 0$. Its derivative at $w = 0$, however, has a simple expression valid for all z:

$$\lim_{\eta\to 0^+}\lim_{n\to\infty} \mathrm{E}(P) = \begin{cases} \left[1 - \sinh\left(\dfrac{2\xi}{\kappa T}\right)^{-4}\right]^{\frac{1}{8}} & \text{if } T < T_c, \\ 0 & \text{if } T > T_c, \end{cases}$$
$$= \begin{cases} (1+y^2)^{\frac{1}{4}}(1-6y^2+y^4)^{\frac{1}{8}}(1-y^2)^{-\frac{1}{2}} & \text{if } T < T_c, \\ 0 & \text{if } T > T_c \end{cases}$$

due to Onsager and Yang [83–85]. A rigorous justification is found in [86–88]. For the special temperature $T = \infty$, we have $\mathrm{E}(P) \to 0$ and $N\,\mathrm{Var}(P) \to 1$ since p is Binomial $(N, 1/2)$ distributed. At criticality, $\mathrm{E}(P) \to 0$ as well, but the second derivative exhibits fascinatingly complicated behavior:

$$\lim_{\eta\to 0^+}\lim_{n\to\infty} N\,\mathrm{Var}(P) = \chi(z) \approx c_0^+ t^{-\frac{7}{4}} + c_1^+ t^{-\frac{3}{4}} + d_0 + c_2^+ t^{\frac{1}{4}} + e_0\, t\ln(t) + d_1 t + c_3^+ t^{\frac{5}{4}},$$

where $0 < t = 1 - T_c/T$, $c_0^+ = 0.9625817323\ldots$, $d_0 = -0.1041332451\ldots$, $e_0 = 0.0403255003\ldots$, $d_1 = -0.14869\ldots$, and

$$c_1^+ = \tfrac{\sqrt{2}}{8}K_c c_0^+, \quad c_2^+ = \tfrac{151}{192}K_c^2 c_0^+, \quad c_3^+ = \tfrac{615\sqrt{2}}{512}K_c^3 c_0^+.$$

Wu, McCoy, Tracy & Barouch [89–99] determined exact expressions for these series coefficients in terms of the solution to a Painlevé III differential equation (described in the next section). Different numerical values of the coefficients apply for $T < T_c$, as well as for the antiferromagnetic case [100, 101]. For example, when $t < 0$, the corresponding leading coefficient is $c_0^- = 0.0255369745\ldots$. The study of magnetic susceptibility $\chi(z)$ is far more involved than the other thermodynamic functions mentioned in this essay, and there are still gaps in the rigorous line of thought [102]. Also, in a recent breakthrough [103, 104], the entire asymptotic structure of $\chi(z)$ has now largely been determined.

5.22.7 *Painlevé III Equation*

Let $f(x)$ be the solution of the Painlevé III differential equation [105]

$$\frac{f''(x)}{f(x)} = \left(\frac{f'(x)}{f(x)}\right)^2 - \frac{1}{x}\frac{f'(x)}{f(x)} + f(x)^2 - \frac{1}{f(x)^2}$$

satisfying the boundary conditions

$$f(x) \sim 1 - \frac{e^{-2x}}{\sqrt{\pi x}} \text{ as } x \to \infty, \quad f(x) \sim x\,(2\ln(2) - \gamma - \ln(x)) \text{ as } x \to 0^+,$$

where γ is Euler's constant [1.5]. Define

$$g(x) = \left[\frac{xf'(x)}{2f(x)} + \frac{x^2}{4f(x)^2}\left(\left(1 - f(x)^2\right)^2 - f'(x)^2\right)\right]\ln(x).$$

Then exact expressions for c_0^+ and c_0^- are

$$c_0^+ = 2^{\frac{5}{8}}\pi\ln(\sqrt{2}+1)^{-\frac{7}{4}}\int_0^\infty y(1 - f(y))$$
$$\times \exp\left[\int_y^\infty x\ln(x)\left(1 - f(x)^2\right)dx - g(y)\right]dy,$$
$$c_0^- = 2^{\frac{5}{8}}\pi\ln(\sqrt{2}+1)^{-\frac{7}{4}}\int_0^\infty y$$
$$\times \left\{(1 + f(y))\exp\left[\int_y^\infty x\ln(x)\left(1 - f(x)^2\right)dx - g(y)\right] - 2\right\}dy.$$

Painlevé II arises in our discussion of the longest increasing subsequence problem [5.20], and Painlevé V arises in connection with the GUE hypothesis [2.15.3].

Here is a slight variation of these results. Define

$$h(x) = -\ln\left(f\left(\frac{x}{c}\right)\right)$$

for any constant $c > 0$; then the function $h(x)$ satisfies what is known as the sinh-Gordon differential equation:

$$h''(x) + \frac{1}{x}h'(x) = \frac{2}{c^2}\sinh(2h(x)),$$

$$h(x) \sim \sqrt{\frac{c}{\pi x}}\exp\left(-\frac{2x}{c}\right) \quad \text{as } x \to \infty.$$

Finally, we mention a beautiful formula:

$$\int_0^\infty x\ln(x)\left(1 - f(x)^2\right)dx = \frac{1}{4} + \frac{7}{12}\ln(2) - 3\ln(A),$$

where A is Glaisher's constant [2.15]. Conceivably, c_0^+ and c_0^- may someday be related to A as well.

[1] G. F. Newell and E. W. Montroll, On the theory of the Ising model of ferromagnetism, *Rev. Mod. Phys.* 25 (1953) 353–389; MR 15,88b.

[2] M. E. Fisher, The theory of cooperative phenomena, *Rep. Prog. Phys.* 30 (1967) 615–730.

[3] S. G. Brush, History of the Lenz-Ising model, *Rev. Mod. Phys.* 39 (1967) 883–893.

[4] C. Thompson, *Mathematical Statistical Mechanics*, Princeton Univ. Press, 1972; MR 80h:82001.

[5] B. M. McCoy and T. T. Wu, *The Two-Dimensional Ising Model*, Harvard Univ. Press, 1973.

[6] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982; MR 90b:82001.

[7] G. A. Baker, *Quantitative Theory of Critical Phenomena*, Academic Press, 1990; MR 92d:82039.

[8] C. Domb, *The Critical Point: A Historical Introduction to the Modern Theory of Critical Phenomena*, Taylor and Francis, 1996.

[9] G. M. Bell and D. A. Lavis, *Statistical Mechanics of Lattice Systems*, v. 1-2, Springer-Verlag, 1999.

[10] D. S. Gaunt and A. J. Guttmann, Asymptotic analysis of coefficients, *Phase Transitions and Critical Phenomena*, v. 3, ed. C. Domb and M. S. Green, Academic Press, 1974, pp. 181–243; A. J. Guttmann, Asymptotic analysis of power-series expansions, *Phase Transitions and Critical Phenomena*, v. 13, ed. C. Domb and J. L. Lebowitz, Academic Press, 1989, pp. 1–234; MR 50 #6393 and MR 94i:82003.

[11] C. Domb, On the theory of cooperative phenomena in crystals, *Adv. Phys.* 9 (1960) 149–361.

[12] T. Oguchi, Statistics of the three-dimensional ferromagnet. II and III, *J. Phys. Soc. Japan* 6 (1951) 27–35; MR 15,188h.

[13] A. J. Wakefield, Statistics of the simple cubic lattice, *Proc. Cambridge Philos. Soc.* 47 (1951) 419–429, 719–810; MR 13,308a and MR 13,417c.

[14] A. Rosengren and B. Lindström, A combinatorial series expansion for the Ising model, *Europ. J. Combin.* 8 (1987) 317–323; MR 89c:82062.

[15] R. J. Baxter and I. G. Enting, Series expansions for corner transfer matrices: The square lattice Ising model, *J. Stat. Phys.* 21 (1979) 103–123.

[16] C. Domb, Some statistical problems connected with crystal lattices, *J. Royal Statist. Soc. B* 26 (1964) 367–397.

[17] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A001393, A002891, A001408, and A002916.

[18] L. Onsager, Crystal physics. I: A two-dimensional model with an order-disorder transition, *Phys. Rev.* 65 (1944) 117–149.

[19] M. Kac and J. C. Ward, A combinatorial solution of the two-dimensional Ising model, *Phys. Rev.* 88 (1952) 1332–1337.

[20] S. Sherman, Combinatorial aspects of the Ising model for ferromagnetism. I: A conjecture of Feynman on paths and graphs, *J. Math. Phys.* 1 (1960) 202–217, addendum 4 (1963) 1213–1214; MR 22 #10273 and MR 27 #6560.

[21] N. V. Vdovichenko, A calculation of the partition function for a plane dipole lattice (in Russian), *Z. Èksper. Teoret. Fiz.* 47 (1965) 715–719; Engl. transl. in *Soviet Physics JETP* 20 (1965) 477–488; MR 31 #5622.

[22] M. L. Glasser, Exact partition function for the two-dimensional Ising model, *Amer. J. Phys.* 38 (1970) 1033–1036.

[23] F. Harary, A graphical exposition of the Ising problem, *J. Austral. Math. Soc.* 12 (1971) 365–377; MR 45 #5032.

[24] P. D. Beale, Exact distribution of energies in the two-dimensional Ising model, *Phys. Rev. Lett.* 76 (1996) 78–81.

[25] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002890, A029872, A029873, and A029874.

[26] M. Creutz, State counting and low-temperature series, *Phys. Rev. B* 43 (1991) 10659–10662.

[27] G. Bhanot, M. Creutz, and J. Lacki, Low temperature expansion for the 3-d Ising model, *Phys. Rev. Lett.* 69 (1992) 1841–1844; hep-lat/9206020.

[28] G. Bhanot, M. Creutz, I. Horvath, J. Lacki, and J. Weckel, Series expansions without diagrams, *Phys. Rev. E* 49 (1994) 2445–2453; hep-lat/9303002.

[29] A. J. Guttmann and I. G. Enting, Series studies of the Potts model. I: The simple cubic Ising model, *J. Phys. A* 26 (1993) 807–821; MR 94a:82009.

[30] C. Vohwinkel, Yet another way to obtain low temperature expansions for discrete spin systems, *Phys. Lett. B* 301 (1993) 208–212.

[31] B. A. Cipra, An introduction to the Ising model, *Amer. Math. Monthly* 94 (1987) 937–959; MR 89g:82001.

[32] M. E. Fisher, The nature of critical points, *Statistical Physics, Weak Interactions, Field Theory – Lectures in Theoretical Physics*, v. 7C, ed. W. E. Brittin, Univ. of Colorado Press, 1964, pp. 1–159.

[33] G. S. Rushbrooke and J. Eve, High-temperature Ising partition function and related non-crossing polygons for the simple cubic lattice, *J. Math. Phys.* 3 (1962) 185–189; MR 25 #1907.

[34] S. McKenzie, Derivation of high temperature series expansions: Ising model, *Phase Transitions – Cargèse 1980*, ed. M. Lévy, J.-C. Le Guillou, and J. Zinn-Justin, Plenum Press, 1982, pp. 247–270.

[35] G. Bhanot, M. Creutz, U. Glässner, and K. Schilling, Specific heat exponent for the 3-d Ising model from a 24-th order high temperature series, *Phys. Rev. B* 49 (1994) 12909–12914; *Nucl. Phys. B (Proc. Suppl.)* 42 (1995) 758–760; hep-lat/9312048.

[36] A. J. Guttmann and I. G. Enting, The high-temperature specific heat exponent of the 3D Ising model, *J. Phys. A* 27 (1994) 8007–8010.

[37] F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, 1973, pp. 11–16, 113–117, 236–237; MR 50 #9682.

[38] H. A. Kramers and G. H. Wannier, Statistics of the two-dimensional ferromagnet. I and II, *Phys. Rev.* 60 (1941) 252–276; MR 3,63i and MR 3,64a.

[39] C. Domb and A. J. Guttmann, Low-temperature series for the Ising model, *J. Phys. C* 3 (1970) 1652–1660.

[40] J. Zinn-Justin, Analysis of Ising model critical exponents from high temperature series expansions, *J. Physique* 40 (1979) 969–975.

[41] J. Adler, Critical temperatures of the $d = 3, s = 1/2$ Ising model: The effect of confluent corrections to scaling, *J. Phys. A* 16 (1983) 3585–3599.

[42] A. J. Liu and M. E. Fisher, The three-dimensional Ising model revisited numerically, *Physica A* 156 (1989) 35–76; MR 90c:82059.

[43] A. M. Ferrenberg and D. P. Landau, Critical behavior of the three-dimensional Ising model: A high-resolution Monte Carlo study, *Phys. Rev. B* 44 (1991) 5081–5091.
[44] B. G. Nickel, Confluent singularities in 3D continuum ϕ^4 theory: Resolving critical point discrepancies, *Physica A* 177 (1991) 189–196.
[45] N. Ito and M. Suzuki, Monte Carlo study of the spontaneous magnetization of the three-dimensional Ising model, *J. Phys. Soc. Japan* 60 (1991) 1978–1987.
[46] Y. Kinosita, N. Kawahima, and M. Suzuki, Coherent-anomaly analysis of series expansions and its application to the Ising model, *J. Phys. Soc. Japan* 61 (1992) 3887–3901; MR 93j:82037.
[47] C. F. Baillie, R. Gupta, K. A. Hawick, and G. S. Pawley, Monte Carlo renomalization-group study of the three-dimensional Ising model, *Phys. Rev. B* 45 (1992) 10438–10453.
[48] H. W. J. Blöte and G. Kamieniarz, Finite-size calculations on the three-dimensional Ising model, *Physica A* 196 (1993) 455–460.
[49] D. P. Landau, Computer simulation studies of critical phenomena, *Physica A* 205 (1994) 41–64.
[50] M. Kolesik and M. Suzuki, Accurate estimates of 3D Ising critical exponents using the coherent-anomaly method, *Physica A* 215 (1995) 138–151.
[51] H. W. J. Blöte, E. Luijten, and J. R. Heringa, Ising universality in three dimensions: A Monte Carlo study, *J. Phys. Math. A* 28 (1995) 6289–6313.
[52] R. Gupta and P. Tamayo, Critical exponents of the 3D Ising model, *Int. J. Mod. Phys. C* 7 (1996) 305–319.
[53] A. L. Talapov and H. W. J. Blöte, The magnetization of the 3D Ising model, *J. Phys. A* 29 (1996) 5727–5733.
[54] A. Rosengren, On the combinatorial solution of the Ising model, *J. Phys. A* 19 (1986) 1709–1714; MR 87k:82149.
[55] M. E. Fisher, On the critical polynomial of the simple cubic Ising model, *J. Phys. A.* 28 (1995) 6323–6333; corrigenda 29 (1996) 1145; MR 97b:82018a and MR 97b:82018b.
[56] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A010556, A010579, A010580, A030008, A030044, A030045, A030046, A030047, A030048, and A030049.
[57] M. E. Fisher and D. S. Gaunt, Ising model and self-avoiding walks on hypercubical lattices and high density expansions, *Phys. Rev.* 133 (1964) A224–A239.
[58] M. A. Moore, Critical behavior of the four-dimensional Ising ferromagnet and the breakdown of scaling, *Phys. Rev. B* 1 (1970) 2238–2240.
[59] D. S. Gaunt, M. F. Sykes, and S. McKenzie, Susceptibility and fourth-field derivative of the spin-1/2 Ising model for $T > T_c$ and $d = 4$, *J. Phys. A* 12 (1979) 871–877.
[60] M. F. Sykes, Derivation of low-temperature expansions for Ising model. X: The four-dimensional simple hypercubic lattice, *J. Phys. A* 12 (1979) 879–892.
[61] C. Vohwinkel and P. Weisz, Low-temperature expansion in the $d = 4$ Ising model, *Nucl. Phys. B* 374 (1992) 647–666; MR 93a:82015.
[62] A. J. Guttmann, Correction to scaling exponents and critical properties of the n-vector model with dimensionality > 4, *J. Phys. A* 14 (1981) 233–239.
[63] A. B. Harris and Y. Meir, Recursive enumeration of clusters in general dimension on hypercubic lattices, *Phys. Rev. A* 36 (1987) 1840–1848; MR 88m:82037.
[64] C. Münkel, D. W. Heermann, J. Adler, M. Gofman, and D. Stauffer, The dynamical critical exponent of the two-, three- and five-dimensional kinetic Ising model, *Physica A* 193 (1993) 540–552.
[65] M. Gofman, J. Adler, A. Aharony, A. B. Harris, and D. Stauffer, Series and Monte Carlo study of high-dimensional Ising models, *J. Stat. Phys.* 71 (1993) 1221–1230.
[66] M. F. Sykes, Some counting theorems in the theory of the Ising model and the excluded volume problem, *J. Math. Phys.* 2 (1961) 52–59; MR 22 #8749.
[67] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A002906, A002913, A002926, and A002927.
[68] D. S. Gaunt and M. F. Sykes, The critical exponent γ for the three-dimensional Ising model, *J. Phys. A* 12 (1979) L25–L28.

[69] D. S. Gaunt, High temperature series analysis for the three-dimensional Ising model: A review of some recent work, *Phase Transitions – Cargèse 1980*, ed. M. Lévy, J.-C. Le Guillou, and J. Zinn-Justin, Plenum Press, 1982, pp. 217–246.
[70] A. J. Guttmann, The high temperature susceptibility and spin-spin correlation function of the three-dimensional Ising model, *J. Phys. A* 20 (1987) 1855–1863; MR 88h:82075.
[71] P. Butera and M. Comi, N-vector spin models on the simple-cubic and the body-centered-cubic lattices: A study of the critical behavior of the susceptibility and of the correlation length by high-temperature series extended to order 21, *Phys. Rev. B* 56 (1997) 8212–8240; hep-lat/9703018.
[72] M. Campostrini, Linked-cluster expansion of the Ising model, *J. Stat. Phys.* 103 (2001) 369–394; cond-mat/0005130.
[73] P. Butera and M. Comi, Extension to order β^{23} of the high-temperature expansions for the spin-1/2 Ising model on the simple-cubic and the body-centered-cubic lattices, *Phys. Rev. B* 62 (2000) 14837–14843; hep-lat/0006009.
[74] M. F. Sykes, D. S. Gaunt, P. D. Roberts, and J. A. Wyles, High temperature series for the susceptibility of the Ising model. I: Two dimensional lattices, *J. Phys. A* 5 (1972) 624–639.
[75] B. G. Nickel, On the singularity structure of the 2D Ising model, *J. Phys. A* 32 (1999) 3889–3906; addendum 33 (2000) 1693–1711; MR 2000d:82013 and MR 2001a:82022.
[76] C. Domb, Ising model, *Phase Transitions and Critical Phenomena*, v. 3, ed. C. Domb and M. S. Green, Academic Press, 1974, pp. 357–484; MR 50 #6393.
[77] P. A. P. Moran, Random associations on a lattice, *Proc. Cambridge Philos. Soc.* 34 (1947) 321–328; 45 (1949) 488; MR 8,592b.
[78] R. M. F. Houtappel, Order-disorder in hexagonal lattices, *Physica* 16 (1950) 425–455; MR 12,576j.
[79] V. S. Adamchik, Exact formulas for some Ising-related constants, unpublished note (1997).
[80] R. Burton and R. Pemantle, Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances, *Annals of Probab.* 21 (1993) 1329–1371; MR 94m:60019.
[81] F. Y. Wu, Number of spanning trees on a lattice, *J. Phys. A* 10 (1977) L113–L115; MR 58 #8974.
[82] R. Shrock and F. Y. Wu, Spanning trees on graphs and lattices in d dimensions, *J. Phys. A* 33 (2000) 3881–3902; MR 2001b:05111.
[83] L. Onsager, Statistical hydrodynamics, *Nuovo Cimento Suppl.* 6 (1949) 279–287.
[84] C. N. Yang, The spontaneous magnetization of a two-dimensional Ising model, *Phys. Rev.* 85 (1952) 808–817; MR 14,522e.
[85] E. W. Montroll and R. B. Potts, Correlations and spontaneous magnetization of the two-dimensional Ising model, *J. Math. Phys.* 4 (1963) 308–319; MR 26 #5913.
[86] G. Benettin, G. Gallavotti, G. Jona-Lasinio, and A. L. Stella, On the Onsager-Yang value of the spontaneous magnetization, *Commun. Math. Phys.* 30 (1973) 45–54.
[87] D. B. Abraham and A. Martin-Löf, The transfer matrix for a pure phase in the two-dimensional Ising model, *Commun. Math. Phys.* 32 (1973) 245–268; MR 49 #6851.
[88] G. Gallavotti, *Statistical Mechanics: A Short Treatise*, Springer-Verlag, 1999.
[89] M. E. Fisher, The susceptibility of the plane Ising model, *Physica* 25 (1959) 521–524.
[90] E. Barouch, B. M. McCoy, and T. T. Wu, Zero-field susceptibility of the two-dimensional Ising model near T_c, *Phys. Rev. Lett.* 31 (1973) 1409–1411.
[91] C. A. Tracy and B. M. McCoy, Neutron scattering and the correlation functions of the Ising model near T_c, *Phys. Rev. Lett.* 31 (1973) 1500–1504.
[92] T. T. Wu, B. M. McCoy, C. A. Tracy, and E. Barouch, Spin-spin correlation functions for the two-dimensional Ising model: Exact theory in the scaling region, *Phys. Rev. B* 13 (1976) 316–374.
[93] B. M. McCoy, C. A. Tracy, and T. T. Wu, Painlevé functions of the third kind, *J. Math. Phys.* 18 (1977) 1058–1092; MR 57 #12993.

[94] C. A. Tracy, Painlevé transcendents and scaling functions of the two-dimensional Ising model, *Nonlinear Equations in Physics and Mathematics*, ed. A. O. Barut, Reidel, 1978, pp. 221–237; MR 84k:35001.
[95] X.-P. Kong, H. Au-Yang, and J. H. H. Perk, New results for the susceptibility of the two-dimensional Ising model at criticality, *Phys. Lett. A* 116 (1986) 54–56.
[96] X.-P. Kong, *Wave-Vector Dependent Susceptibility of the Two-Dimensional Ising Model*, Ph.D. thesis, State Univ. of New York at Stony Brook, 1987.
[97] S. Gartenhaus and W. S. McCullough, Higher order corrections for the quadratic Ising lattice susceptibility at criticality, *Phys. Rev. B* 38 (1988) 11688–11703; *Phys. Lett. A* 127 (1988) 315–318; MR 89b:82105.
[98] C. A. Tracy, Asymptotics of a tau-function arising in the two-dimensional Ising model, *Commun. Math. Phys.* 142 (1991) 298–311; MR 93c:82014.
[99] O. Babelon and D. Bernard, From form factors to correlation functions; the Ising model, *Phys. Lett. B* 288 (1992) 113–120; MR 94e:82015.
[100] X.-P. Kong, H. Au-Yang, and J. H. H. Perk, Comment on a paper by Yamadi and Suzuki, *Prog. Theor. Phys.* 77 (1987) 514–516.
[101] S. S. C. Burnett and S. Gartenhaus, Zero-field susceptibility of an antiferromagnetic square Ising lattice, *Phys. Rev. B* 47 (1993) 7944–7956.
[102] J. Palmer and C. A. Tracy, Two-dimensional Ising correlations: Convergence of the scaling limit, *Adv. Appl. Math* 2 (1981) 329–388; MR 84m:82024.
[103] W. P. Orrick, B. G. Nickel, A. J. Guttmann, and J. H. H. Perk, Critical behaviour of the two-dimensional Ising susceptibility, *Phys. Rev. Lett.* 86 (2001) 4120–4123.
[104] W. P. Orrick, B. G. Nickel, A. J. Guttmann, and J. H. H. Perk, The susceptibility of the square lattice Ising model: New developments, *J. Stat. Phys.* 102 (2001) 795–841; MR 2002e:82013.
[105] C. A. Tracy and H. Widom, Universality of the distribution functions of random matrix theory, *Integrable Systems: From Classical to Quantum*, Proc. 1999 Montréal conf., ed. J. Harnad, G. Sabidussi, and P. Winternitz, Amer. Math. Soc, 2000, pp. 251–264; math-ph/9909001; MR 2002f:15036.

5.23 Monomer-Dimer Constants

Let L be a graph [5.6]. A **dimer** consists of two adjacent vertices of L and the (non-oriented) bond connecting them. A **dimer arrangement** is a collection of disjoint dimers on L. Uncovered vertices are called **monomers**, so dimer arrangements are also known as **monomer-dimer coverings**. We will discuss such coverings only briefly at the beginning of the next section.

A **dimer covering** is a dimer arrangement whose union contains all the vertices of L. Dimer coverings and the closely-related topic of tilings will occupy the remainder of this essay.

5.23.1 2D Domino Tilings

Let a_n denote the number of distinct monomer-dimer coverings of an $n \times n$ square lattice L and $N = n^2$; then $a_1 = 1$, $a_2 = 7$, $a_3 = 131$, $a_4 = 10012$ [1,2], and asymptotically [3–6]

$$A = \lim_{n\to\infty} a_n^{\frac{1}{N}} = 1.940215351\ldots = (3.764435608\ldots)^{\frac{1}{2}}.$$

No exact expression for the constant A is known. Baxter's approach for estimating A was based on the corner transfer matrix variational approach, which also played a

role in [5.12]. A natural way for physicists to discuss the monomer-dimer problem is to associate an activity z with each dimer; A thus corresponds to the case $z = 1$. The mean number ρ of dimers per vertex is 0 if $z = 0$ and $1/2$ if $z = \infty$; when $z = 1$, ρ is $0.3190615546\ldots$, for which again there is no closed-form expression [3]. Unlike other lattice models (see [5.12], [5.18], and [5.22]), monomer-dimer systems do not have a phase transition [7].

Computing a_n is equivalent to counting (not necessarily perfect) **matchings** in L, that is, to counting independent sets of edges in L. This is related to the difficult problem of computing permanents of certain binary incidence matrices [8–14]. Kenyon, Randall & Sinclair [15] gave a randomized polynomial-time approximation algorithm for computing the number of monomer-dimer coverings of L, assuming ρ to be given.

Let us turn our attention henceforth to the zero monomer density case, that is, $z = \infty$. If b_n is the number of distinct dimer coverings of L, then $b_n = 0$ if n is odd and

$$b_n = 2^{N/2} \prod_{j=1}^{n/2} \prod_{k=1}^{n/2} \left(\cos^2 \frac{j\pi}{n+1} + \cos^2 \frac{k\pi}{n+1} \right)$$

if n is even. This exact expression is due to Kastelyn [16] and Fisher & Temperley [17, 18]. Further,

$$\lim_{\substack{n\to\infty \\ n \text{ even}}} \frac{1}{N} \ln(b_n) = \frac{1}{16\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \ln\left[4 + 2\cos(\theta) + 2\cos(\varphi)\right] d\theta\, d\varphi$$
$$= \frac{G}{\pi} = 0.2915609040\ldots;$$

that is,

$$B = \lim_{\substack{n\to\infty \\ n \text{ even}}} b_n^{\frac{1}{N}} = \exp\left(\frac{G}{\pi}\right) = 1.3385151519\ldots = (1.7916228120\ldots)^{\frac{1}{2}},$$

where G is Catalan's constant [1.7]. This is a remarkable solution, in graph theoretic terms, of the problem of counting **perfect matchings** on the square lattice. It is also an answer to the following question: What is the number of ways of tiling an $n \times n$ chessboard with 2×1 or 1×2 **dominoes**? See [19–26] for more details. The constant B^2 is called δ in [3.10] and appears in [1.8] too; the expression $4G/\pi$ arises in [5.22], $G/(\pi \ln(2))$ in [5.6], and $8G/\pi^2$ in [7.7].

If we wrap the square lattice around to form a torus, the counts b_n differ somewhat, but the limiting constant B remains the same [16, 27]. If, instead, we assume the chessboard to be shaped like an Aztec diamond [28], then the associated constant $B = 2^{1/4} = 1.189\ldots < 1.338\ldots = e^{G/\pi}$. Hence, even though the square chessboard has slightly less area than the diamond chessboard, the former possesses many more domino tilings [29]. Lattice boundary effects are thus seen to be nontrivial.

5.23.2 Lozenges and Bibones

The analog of $\exp(2G/\pi)$ for dimers on a hexagonal (honeycomb) lattice with wraparound is [30–32]

$$C^2 = \lim_{n\to\infty} c_n^{\frac{2}{N}} = \exp\left(\frac{1}{8\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \ln\left[3+2\cos(\theta)+2\cos(\varphi)+2\cos(\theta+\varphi)\right]\,d\theta\,d\varphi\right)$$
$$= 1.3813564445\ldots.$$

This constant is called β in [3.10] and can be expressed by other formulas too. It characterizes lozenge tilings on a chessboard with triangular cells satisfying periodic boundary conditions. See [33–38] as well.

If there is no wraparound, then the sequence [39]

$$c_n = \prod_{j=1}^{n}\prod_{k=1}^{n}\frac{n+j+k-1}{j+k-1}$$

emerges, and a different growth constant $3\sqrt{3}/4$ applies. We have assumed that the hexagonal grid is center-symmetric with sides n, n, and n (i.e., the simplest possible boundary conditions). The sequence further enumerates plane partitions contained within an $n \times n \times n$ box [40,41].

The corresponding analog for dimers on a triangular lattice with wraparound is [30,42,43]

$$D^2 = \lim_{n\to\infty} d_n^{\frac{2}{N}} = \exp\left(\frac{1}{8\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} \ln\left[6+2\cos(\theta)+2\cos(\varphi)+2\cos(\theta+\varphi)\right]\,d\theta\,d\varphi\right)$$
$$= 2.3565273533\ldots.$$

The expression $4\ln(D)$ bears close similarity to a constant $\ln(6) + H$ described in [5.22]. It also characterizes bibone tilings on a chessboard with hexagonal cells satisfying periodic boundary conditions. The case of no wraparound [1,44,45] apparently remains open.

5.23.3 3D Domino Tilings

Let h_n denote the number of distinct dimer coverings of an $n \times n \times n$ cubic lattice L and $N = n^3$. Then $h_n = 0$ if n is odd, $h_2 = 9$, and $h_4 = 5051532105$ [46,47]. An important unsolved problem in solid-state chemistry is the estimation of

$$\lim_{\substack{n\to\infty\\ n\text{ even}}} h_n^{\frac{1}{N}} = \exp(\lambda)$$

or, equivalently,

$$\lambda = \lim_{\substack{n\to\infty\\ n\text{ even}}} \frac{1}{N}\ln(h_n).$$

Hammersley [48] proved that λ exists and $\lambda \geq 0.29156$. Lower bounds were improved by Fisher [49] to 0.30187, Hammersley [50,51] to 0.418347, and Priezzhev [52,53] to 0.419989. In a review of [54], Minc pointed out that a conjecture due to Schrijver & Valiant on lower bounds for permanents of certain binary matrices would imply that $\lambda \geq 0.44007584$. Schrijver [55] proved this conjecture, and this is the best-known result.

Fowler & Rushbrooke [56] gave an upper bound of 0.54931 for λ over sixty years ago (assuming λ exists). Upper bounds have been improved by Minc [8,57,58] to 0.5482709, Ciucu [59] to 0.463107, and Lundow [60] to 0.457547.

A sequence of nonrigorous numerical estimates by Nagle [30], Gaunt [31], and Beichl & Sullivan [61] has culminated with $\lambda = 0.4466\ldots$. As with a_n, computing h_n for even small values of n is hard and matrix permanent approximation schemes offer the only hope. The field is treacherously difficult: Conjectured exact asymptotic formulas for h_n in [62,63] are incorrect.

A related topic is the number, k_n, of dimer coverings of the n-dimensional unit cube, whose 2^n vertices consist of all n-tuples drawn from $\{0, 1\}$ [47,64]. The term $k_6 = 16332454526976$ was computed independently by Lundow [46] and Weidemann [65]. In this case, we know the asymptotic behavior of k_n rather precisely [44,65,66]:

$$\lim_{n\to\infty}\frac{1}{n}k_n^{2^{1-n}} = \frac{1}{e} = 0.3678794411\ldots,$$

where e is the natural logarithmic base [1.3].

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A004003, A006125, A008793, A028420, and A039907.
[2] J. J. Henry, Monomer-dimer counts, $1 \leq n \leq 21$, unpublished note (1997).
[3] R. J. Baxter, Dimers on a rectangular lattice, *J. Math. Phys.* 9 (1968) 650–654.
[4] L. K. Runnels, Exact finite method of lattice statistics. III: Dimers on the square lattice, *J. Math. Phys.* 11 (1970) 842–850.
[5] M. Heise, Upper and lower bounds for the partition function of lattice models, *Physica A* 157 (1989) 983–999; MR 91c:82014.
[6] F. Cazals, Monomer-dimer tilings, *Studies in Automatic Combinatorics*, v. 2, Algorithms Project, INRIA, 1997.
[7] O. J. Heilmann and E. H. Lieb, Theory of monomer-dimer systems, *Commun. Math. Phys.* 25 (1972) 190–232; MR 45 #6337.
[8] H. Minc, *Permanents*, Addision-Wesley, 1978; MR 80d:15009.
[9] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, Cambridge Univ. Press, 1991; MR 93a:05087.
[10] J. M. Hammersley, A. Feuerverger, A. Izenman, and K. Makani, Negative finding for the three-dimensional dimer problem, *J. Math. Phys.* 10 (1969) 443–446.
[11] M. Jerrum, Two-dimensional monomer-dimer systems are computationally intractable, *J. Stat. Phys.* 48 (1987) 121–134; erratum 59 (1990) 1087–1088; MR 89d:82008 and MR 91h:82002.
[12] M. Luby, A survey of approximation algorithms for the permanent, *Sequences: Combinatorics, Compression, Security and Transmission*, Proc. 1988 Naples workshop, ed. R. M. Capocelli, Springer-Verlag, 1990, pp. 75–91; MR 90m:68059.
[13] M. Jerrum and A. Sinclair, Approximating the permanent, *SIAM J. Comput.* 18 (1989) 1149–1178; MR 91a:05075.
[14] A. Frieze and M. Jerrum, An analysis of a Monte-Carlo algorithm for estimating the permanent, *Combinatorica* 15 (1995) 67–83; MR 96g:68052.

[15] C. Kenyon, D. Randall, and A. Sinclair, Approximating the number of monomer-dimer coverings of a lattice, *J. Stat. Phys.* 83 (1996) 637–659; MR 97g:82003.
[16] P. W. Kasteleyn, The statistics of dimers on a lattice. I: The number of dimer arrangements on a quadratic lattice, *Physica* 27 (1961) 1209–1225.
[17] M. E. Fisher, Statistical mechanics of dimers on a plane lattice, *Phys. Rev.* 124 (1961) 1664–1672; MR 24 #B2437.
[18] H. N. V. Temperley and M. E. Fisher, Dimer problem in statistical mechanics – An exact result, *Philos. Mag.* 6 (1961) 1061–1063; MR 24 #B2436.
[19] E. W. Montroll, Lattice statistics, *Applied Combinatorial Mathematics*, ed. E. F. Beckenbach, Wiley, 1964, pp. 96–143; MR 30 #4687.
[20] P. W. Kasteleyn, Graph theory and crystal physics, *Graph Theory and Theoretical Physics*, ed. F. Harary, Academic Press, 1967, pp. 43–110; MR 40 #6903.
[21] J. K. Percus, *Combinatorial Methods*, Springer-Verlag, 1971; MR 49 #10555.
[22] L. Lovász and M. D. Plummer, *Matching Theory*, North-Holland, 1986; MR 88b:90087.
[23] P. John, H. Sachs, and H. Zernitz, Counting perfect matchings in polyminoes with an application to the dimer problem, *Zastos. Mat.* 19 (1987) 465–477; MR 89e:05158.
[24] H. Sachs and H. Zernitz, Remark on the dimer problem, *Discrete Appl. Math.* 51 (1994) 171–179; MR 97e:05067.
[25] D. M. Cvetkovic, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Applications*, 3rd ed., Johann Ambrosius Barth Verlag, 1995, pp. 245–251; MR 96b:05108.
[26] D. A. Lavis and G. M. Bell, *Statistical Mechanics of Lattice Systems*, v. 2, Springer-Verlag, 1999; MR 2001g:82002.
[27] J. Propp, Dimers and dominoes, unpublished note (1992).
[28] N. Elkies, G. Kuperberg, M. Larsen, and J. Propp, Alternating-sign matrices and domino tilings. I, *J. Algebraic Combin.* 1 (1992) 111–132; MR 94f:52035.
[29] H. Cohn, R. Kenyon, and J. Propp, A variational principle for domino tilings, *J. Amer. Math. Soc.* 14 (2001) 297–346; math.CO/0008220.
[30] J. F. Nagle, New series-expansion method for the dimer problem, *Phys. Rev.* 152 (1966) 190–197.
[31] D. S. Gaunt, Exact series-expansion study of the monomer-dimer problem, *Phys. Rev.* 179 (1969) 174–186.
[32] A. J. Phares and F. J. Wunderlich, Thermodynamics and molecular freedom of dimers on plane honeycomb and Kagomé lattices, *Nuovo Cimento B* 101 (1988) 653–686; MR 89m:82067.
[33] G. H. Wannier, Antiferromagnetism: The triangular Ising net, *Phys. Rev.* 79 (1950) 357–364; MR 12,576e.
[34] P. W. Kasteleyn, Dimer statistics and phase transitions, *J. Math. Phys.* 4 (1963) 287–293; MR 27 #3394.
[35] F. Y. Wu, Remarks on the modified potassium dihydrogen phosphate model of a ferroelectric, *Phys. Rev.* 168 (1968) 539–543.
[36] J. F. Nagle, Critical points for dimer models with 3/2-order transitions, *Phys. Rev. Lett.* 34 (1975) 1150–1153.
[37] H. W. J. Blöte and H. J. Hilhorst, Roughening transitions and the zero-temperature triangular Ising antiferromagnet, *J. Phys. A* 15 (1982) L631–L637; MR 83k:82046.
[38] C. Richard, M. Höffe, J. Hermisson, and M. Baake, Random tilings: Concepts and examples, *J. Phys. A* 31 (1998) 6385–6408; MR 99g:82033.
[39] V. Elser, Solution of the dimer problem on a hexagonal lattice with boundary, *J. Phys. A* 17 (1984) 1509–1513.
[40] G. David and C. Tomei, The problem of the calissons, *Amer. Math. Monthly* 96 (1989) 429–431; MR 90c:51024.
[41] P. A. MacMahon, *Combinatory Analysis*, Chelsea, 1960, pp. 182–183, 239–242; MR 25 #5003.
[42] A. J. Phares and F. J. Wunderlich, Thermodynamics and molecular freedom of dimers on plane triangular lattices, *J. Math. Phys.* 27 (1986) 1099–1109; MR 87c:82071.

[43] R. Kenyon, The planar dimer model with boundary: A survey, *Directions in Mathematical Quasicrystals*, ed. M. Baake and R. V. Moody, CRM Monogr. 13, Amer. Math. Soc., 2000, pp. 307–328; MR 2002e:82011.
[44] J. Propp, Enumeration of matchings: Problems and progress, *New Perspectives in Algebraic Combinatorics*, Proc. 1996-1997 MSRI Berkeley Program on Combinatorics, ed. L. J. Billera, A. Björner, C. Greene, R. E. Simion, and R. P. Stanley, Cambridge Univ. Press, 1999, pp. 255–291; math.CO/9904150; MR 2001c:05008.
[45] H. Sachs, A contribution to problem 18 in "Twenty open problems in enumeration of matchings," unpublished note (1997).
[46] P. H. Lundow, Computation of matching polynomials and the number of 1-factors in polygraphs, research report 12-1996, Umeå Universitet.
[47] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005271, A028446, A033535, and A045310.
[48] J. M. Hammersley, Existence theorems and Monte Carlo methods for the monomer-dimer problem, *Research Papers in Statistics: Festschrift for J. Neyman*, ed. F. N. David, Wiley, 1966, pp. 125–146; MR 35 #2595.
[49] J. A. Bondy and D. J. A. Welsh, A note on the monomer-dimer problem, *Proc. Cambridge Philos. Soc.* 62 (1966) 503–505; MR 34 #3958.
[50] J. M. Hammersley, An improved lower bound for the multidimensional dimer problem, *Proc. Cambridge Philos. Soc.* 64 (1968) 455–463; MR 38 #5639.
[51] J. M. Hammersley and V. V. Menon, A lower bound for the monomer-dimer problem, *J. Inst. Math. Appl.* 6 (1970) 341–364; MR 44 #3886.
[52] V. B. Priezzhev, The statistics of dimers on a three-dimensional lattice. I: An exactly solved model, *J. Stat. Phys.* 26 (1981) 817–828; MR 84h:82059a.
[53] V. B. Priezzhev, The statistics of dimers on a three-dimensional lattice. II: An improved lower bound, *J. Stat. Phys.* 26 (1981) 829–837; MR 84h:82059b.
[54] A. Schrijver and W. G. Valiant, On lower bounds for permanents, *Proc. Konink. Nederl. Akad. Wetensch. Ser. A* 83 (1980) 425–427; *Indag. Math.* 42 (1980) 425–427; MR 82a:15004.
[55] A. Schrijver, Counting 1-factors in regular bipartite graphs, *J. Combin. Theory B* 72 (1998) 122–135; MR 99b:05117.
[56] R. H. Fowler and G. S. Rushbrooke, An attempt to extend the statistical theory of perfect solutions, *Trans. Faraday Soc.* 33 (1937) 1272–1294.
[57] H. Minc, An upper bound for the multidimensional dimer problem, *Math. Proc. Cambridge Philos. Soc.* 83 (1978) 461–462; MR 58 #289.
[58] H. Minc, An asymptotic solution of the multidimensional dimer problem, *Linear Multilinear Algebra* 8 (1980) 235–239; MR 81e:82063.
[59] M. Ciucu, An improved upper bound for the 3-dimensional dimer problem, *Duke Math. J.* 94 (1998) 1–11; MR 99f:05026.
[60] P. H. Lundow, Compression of transfer matrices, *Discrete Math.* 231 (2001) 321–329; MR 2002a:05200.
[61] I. Beichl and F. Sullivan, Approximating the permanent via importance sampling with application to the dimer covering problem, *J. Comput. Phys.* 149 (1999) 128–147; MR 99m:82021.
[62] A. J. Phares and F. J. Wunderlich, Thermodynamics of dimers on a rectangular $L \times M \times N$ lattice, *J. Math. Phys.* 26 (1985) 2491–2499; comment by J. C. Wheeler, 28 (1987) 2739–2740; MR 86m:82050 and MR 88m:82038.
[63] H. Narumi, H. Kita, and H. Hosoya, Expressions for the perfect matching numbers of cubic $l \times m \times n$ lattices and their asymptotic values, *J. Math. Chem.* 20 (1996) 67–77; MR 98c:82030.
[64] N. Graham and F. Harary, The number of perfect matchings in a hypercube, *Appl. Math. Lett.* 1 (1988) 45–48.
[65] L. H. Clark, J. C. George, and T. D. Porter, On the number of 1-factors in the n-cube, *Proc. 28th Southeastern Conf. on Combinatorics, Graph Theory and Computing*, Boca Raton, 1997, Congr. Numer. 127, Utilitas Math., 1997, pp. 67–69; MR 98i:05127.

[66] H. Sachs, Problem 298, How many perfect matchings does the graph of the n-cube have?, *Discrete Math.* 191 (1998) 251.

5.24 Lieb's Square Ice Constant

Let L denote the $n \times n$ planar square lattice with wraparound and let $N = n^2$. An **orientation** of L is an assignment of a direction (or arrow) to each edge of L. What is the number, f_n, of orientations of L such that at each vertex there are exactly two inward and two outward pointing edges? Such orientations are said to obey the **ice rule** and are also called **Eulerian orientations**. The sequence $\{f_n\}$ starts with the terms $f_1 = 4$, $f_2 = 18$, $f_3 = 148$, and $f_4 = 2970$ [1,2]. After intricate analysis, Lieb [3–5] proved that

$$\lim_{n\to\infty} f_n^{\frac{1}{N}} = \left(\frac{4}{3}\right)^{\frac{3}{2}} = \sqrt{\frac{64}{27}} = 1.5396007178\ldots.$$

This constant is known as the **residual entropy for square ice**. A brief discussion of the underlying physics appears in [5.24.3]. The model is also called a **six-vertex model** since, at each vertex, there are six possible configurations of arrows [6–9]. See Figure 5.24.

We turn to several related results. Let $\tilde{f}_n$ denote the number of orientations of L such that at each vertex there are an even number of edges pointing in and an even number pointing out. Clearly $\tilde{f}_n \geq f_n$ and the model is called an **eight-vertex model**. In this case, however, the analysis is not quite so intricate and we have $\tilde{f}_n = 2^{N+1}$ via elementary linear algebra. The corresponding expression for the **sixteen-vertex model** (with no restrictions on the arrows) is obviously 2^{2N}.

Let us focus instead on the planar triangular lattice L with N vertices. What is the number, g_n, of orientations of L such that at each vertex there are exactly three inward and three outward pointing edges? (The phrase *Eulerian orientation* applies here, but not *ice rule*.) Baxter [10] proved that this **twenty-vertex model** satisfies

$$\lim_{n\to\infty} g_n^{\frac{1}{N}} = \sqrt{\frac{27}{4}} = 2.5980762113\ldots.$$

The problem of computing f_n and g_n is the same as counting nowhere-zero flows modulo

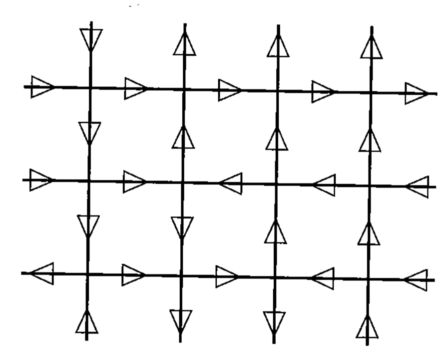


Figure 5.24. A sample planar configuration of arrows satisfying the ice rule.

3 on L [9,11,12]. Mihail & Winkler [13] studied related computational complexity issues.

One of several solutions of the famous alternating sign matrix conjecture [1, 14–16] is closely related to the square ice model. This achievement serves to underscore (once again) the commonality of combinatorial theory and statistical physics.

5.24.1 Coloring

Here is a fascinating topic that anticipates the next essay [5.25]. Let u_n denote the number of ways of coloring the vertices of the square lattice with three colors so that no two adjacent vertices are colored alike. Lenard [5] pointed out that $u_n = 3f_n$. In words, the number of 3-colorings of a square map is thrice the number of square ice configurations. We will return to u_n momentarily, with generalization in mind.

Replace the square lattice by the triangular lattice L and fix an integer $q \geq 4$. Let v_n denote the number of ways of coloring the vertices of L with q colors so that no two adjacent vertices are colored alike. Baxter [17, 18] proved that, if a parameter $-1 < x < 0$ is defined for $q > 4$ by $q = 2 - x - x^{-1}$, then

$$\lim_{n\to\infty} v_n^{\frac{1}{N}} = -\frac{1}{x}\prod_{j=1}^{\infty}\frac{(1-x^{6j-3})(1-x^{6j-2})^2(1-x^{6j-1})}{(1-x^{6j-5})(1-x^{6j-4})(1-x^{6j})(1-x^{6j+1})}.$$

In particular, letting $q \to 4^+$ (note that the formula makes sense for real q), we obtain

$$\begin{aligned} C^2 = \lim_{n\to\infty} v_n^{\frac{1}{N}} &= \prod_{j=1}^{\infty}\frac{(3j-1)^2}{(3j-2)(3j)} = \frac{3}{4\pi^2}\Gamma\left(\tfrac{1}{3}\right)^3 \\ &= 1.4609984862\ldots = (1.2087177032\ldots)^2, \end{aligned}$$

which we call **Baxter's 4-coloring constant** for a triangular lattice.

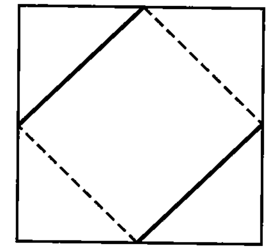
Define likewise u_n and w_n for the number of q-colorings of the square lattice and the hexagonal (honeycomb) lattice with N vertices, respectively. Analytical expressions for the corresponding limiting values are not available, but numerical assessment of certain series expansions provide the list in Table 5.6 [19–21]. The only known quantity in this table is Lieb's constant in the upper left corner. See [5.25] for related discussion on chromatic polynomials.

Table 5.6. *Limiting Values of Roots $u_n^{1/N}$ and $w_n^{1/N}$*

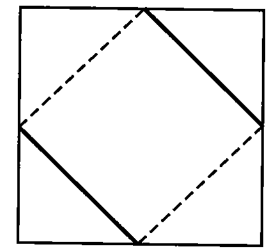
q	$\lim_{n\to\infty} u_n^{1/N}$	$\lim_{n\to\infty} w_n^{1/N}$
3	1.5396...	1.6600...
4	2.3360...	2.6034...
5	3.2504...	3.5795...
6	4.2001...	4.5651...
7	5.1667...	5.5553...

5.24.2 Folding

The *square-diagonal folding* problem may be translated into the following coloring problem. Cover the faces of the square lattice with either of the two following square tiles.



Tile 1



Tile 2

Tile 1: Alternating black and white segments join the centers of the consecutive edges around the square; west-to-north segment is black, north-to-east is white, east-to-south is black, and south-to-west is white. Tile 2: The opposite convention is adopted; west-to-north segment is white, north-to-east is black, east-to-south is white, and south-to-west is black.

There are $2N$ such coverings for a lattice made of N squares. Now, surrounding each vertex of the original lattice, there is a square *loop* formed from the four neighboring tiles. Count the number K_w of purely white loops and the number K_b of purely black loops, assuming wraparound. In the sample covering of Figure 5.25, both K_w and K_b are zero. Define

$$s = \lim_{n\to\infty} \frac{1}{4N} \ln\left(\sum_{\text{coverings}} 2^{K_w+K_b}\right)$$

to be the **entropy of folding** of the **square-diagonal lattice**, where the sum is over all 2^N tiling configurations. (This entropy is per triangle rather than per tile, which explains the additional factor of $1/4$.)

An obvious lower bound for s is

$$s \geq \lim_{n\to\infty} \frac{1}{4N} \ln(2^N + 2^N) = \lim_{n\to\infty} \frac{N+1}{4N} \ln(2) = \frac{1}{4}\ln(2) = 0.1732\ldots,$$

which is obtained by allowing the tiling configurations to alternate like a chessboard. There are two such possibilities (by simple exchanging of all tile 1s by tile 2s and all tile 2s by tile 1s). A more elaborate argument [22, 23] gives $s = 0.2299\ldots$.

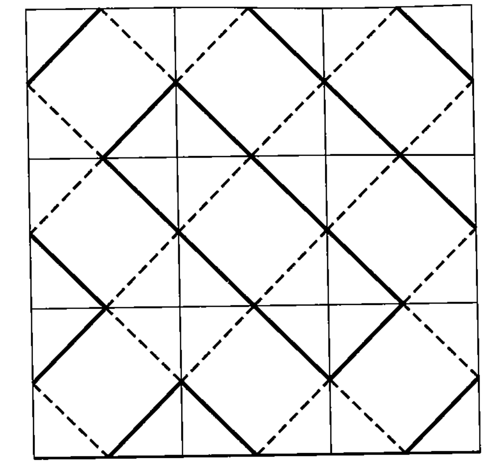


Figure 5.25. Sample covering of a lattice by tiles of both types.

The corresponding **entropy of folding** of the **triangular lattice** is $\ln(C) = 0.1895600483\ldots$ due to Baxter [17, 18] and possesses a simpler coloring interpretation, as already mentioned.

5.24.3 Atomic Arrangement in an Ice Crystal

Square ice is a two-dimensional idealization of water (H_2O) in its solid phase. The oxygen (O) atoms are pictured as the vertices of the square lattice, with outward pointing edges interpreted as the hydrogen (H) atoms. In actuality, however, there are several kinds of three-dimensional ice, depending on temperature and pressure [24, 25]. The residual entropies W for *ordinary hexagonal ice* Ice-Ih and for *cubic ice* Ice-Ic satisfy [3, 26–30]

$$1.5067 < W < 1.5070$$

and are equal within the limits of Nagle's estimation error. These complicated three-dimensional lattices are not the same as the simple models mathematicians tend to focus on.

It would be interesting to see the value of W for the customary $n \times n \times n$ cubic lattice, either with the ice rule in effect (two arrows point out, two arrows point in, and two null arrows) or with Eulerian orientation (three arrows point out and three arrows point in). No one appears to have done this.

[1] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A005130, A050204, and A054759.
[2] J. J. Henry, Ice configuration counts, $1 \leq n \leq 13$, unpublished note (1998).
[3] E. H. Lieb and F. Y. Wu, Two-dimensional ferroelectric models, *Phase Transitions and Critical Phenomena*, v. 1, ed. C. Domb and M. S. Green, Academic Press, 1972, pp. 331–490; MR 50 #6391.

[4] E. H. Lieb, Exact solution of the problem of the entropy of two-dimensional ice, *Phys. Rev. Lett.* 18 (1967) 692–694.
[5] E. H. Lieb, Residual entropy of square ice, *Phys. Rev.* 162 (1967) 162–172.
[6] J. K. Percus, *Combinatorial Methods*, Springer-Verlag, 1971; MR 49 #10555.
[7] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982; MR 90b:82001.
[8] G. M. Bell and D. A. Lavis, *Statistical Mechanics of Lattice Systems*, v. 1-2, Springer-Verlag, 1999.
[9] C. Godsil, M. Grötschel, and D. J. A. Welsh, Combinatorics in Statistical Physics, *Handbook of Combinatorics*, v. II, ed. R. Graham, M. Grötschel, and L. Lovász, MIT Press, 1995, pp. 1925–1954; MR 96h:05001.
[10] R. J. Baxter, *F* model on a triangular lattice, *J. Math. Phys.* 10 (1969) 1211–1216.
[11] D. H. Younger, Integer flows, *J. Graph Theory* 7 (1983) 349–357; MR 85d:05223.
[12] C.-Q. Zhang, *Integer Flows and Cycle Covers of Graphs*, Dekker, 1997; MR 98a:05002.
[13] M. Mihail and P. Winkler, On the number of Eulerian orientations of a graph, *Algorithmica* 16 (1996) 402–414; also in *Proc. 3rd ACM-SIAM Symp. on Discrete Algorithms (SODA)*, Orlando, ACM, 1992, pp. 138–145; MR 93f:05059 and MR 97i:68096.
[14] D. Bressoud and J. Propp, How the Alternating Sign Matrix conjecture was solved, *Notices Amer. Math. Soc.* 46 (1999) 637–646; corrections 1353–1354, 1419; MR 2000k:15002.
[15] D. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, Math. Assoc. Amer. and Cambridge Univ. Press, 1999; MR 2000i:15002.
[16] A. Lascoux, Square-ice enumeration, *Sémin. Lothar. Combin.* 42 (1999) B42p; MR 2000f:05086.
[17] R. J. Baxter, Colorings of a hexagonal lattice, *J. Math. Phys.* 11 (1970) 784–789; MR 42 #1457.
[18] R. J. Baxter, q colourings of the triangular lattice, *J. Phys. A* 19 (1986) 2821–2839; MR 87k:82123.
[19] J. F. Nagle, A new subgraph expansion for obtaining coloring polynomials for graphs, *J. Combin. Theory* 10 (1971) 42–59; MR 44 #2663.
[20] A. V. Bakaev and V. I. Kabanovich, Series expansions for the q-colour problem on the square and cubic lattices, *J. Phys. A.* 27 (1994) 6731–6739; MR 95i:82014.
[21] R. Shrock and S.-H. Tsai, Asymptotic limits and zeros of chromatic polynomials and ground state entropy of Potts antiferromagnets, *Phys. Rev. E* 55 (1997) 5165; cond-mat/9612249.
[22] P. Di Francesco, Folding the square-diagonal lattice, *Nucl. Phys. B* 525 (1998) 507–548; cond-mat/9803051; MR 99k:82033.
[23] P. Di Francesco, Folding and coloring problems in mathematics and physics, *Bull. Amer. Math. Soc.* 37 (2000) 251–307; MR 2001g:82004.
[24] L. Pauling, *The Nature of the Chemical Bond*, 3rd ed., Cornell Univ. Press, 1960.
[25] *Advances in Physics* 7 (1958) 171–297 (entire issue devoted to the physics of water and ice).
[26] H. Takahasi, Zur Slaterschen Theorie der Umwandlung von KH2PO4 (Teil 1), *Proc. Physico-Math. Soc. Japan* 23 (1941) 1069–1079.
[27] E. A. DiMarzio and F. H. Stillinger, Residual entropy of ice, *J. Chem. Phys.* 40 (1964) 1577–1581.
[28] J. F. Nagle, Lattice statistics of hydrogen bonded crystals. I: The residual entropy of ice, *J. Math. Phys.* 7 (1966) 1484–1491.
[29] J. F. Nagle, Ferroelectric models, *Phase Transitions and Critical Phenomena*, v. 3, ed. C. Domb and M. S. Green, Academic Press, 1974, pp. 653–666; MR 50 #6393.
[30] L. K. Runnels, Ice, *Sci. Amer.*, v. 215 (1966) n. 6, 118–126.

5.25 Tutte–Beraha Constants

Let G be a graph with n vertices v_j [5.6] and let λ be a positive integer. A **λ-coloring** of G is a function $\{v_1, v_2, \ldots, v_n\} \to \{1, 2, \ldots, \lambda\}$ with the property that adjacent

vertices must be colored differently. Define $P(\lambda)$ to be the number of λ-colorings of G. Then $P(\lambda)$ is a polynomial of degree n, called the **chromatic polynomial** (or **chromial**) of G. For example, if G is a triangle (three vertices with each pair connected), then $P(\lambda) = \lambda(\lambda - 1)(\lambda - 2)$. Chromatic polynomials were first extensively studied by Birkhoff & Lewis [1]; see [2–6] for introductory material.

A graph G is **planar** if it can be drawn in the plane in such a way that no two edges cross except at a common vertex. The famous Four Color Theorem for geographic maps can be restated as follows: If G is a planar graph, then $P(4) > 0$. Among several restatements of the theorem, we mention Kauffman's combinatorial three-dimensional vector cross product result [7–9].

We can ask about the behavior of $P(\lambda)$ at other real values. Clearly $P(0) = 0$ and, if G is connected, then $P(1) = 0$ and $P(\lambda) \neq 0$ for $\lambda < 0$ or $0 < \lambda < 1$. Further, $P(\varphi + 1) \neq 0$, where φ is the Golden mean [1.2]; more concerning φ will be said shortly.

A connected planar graph G determines a subdivision of the 2-sphere (under stereographic projection) into simply connected regions (faces). If each such region is bounded by a simple closed curve made up of exactly three edges of G, then G is called a **spherical triangulation**. We henceforth assume that this condition is always met.

Clearly $P(2) = 0$ for any spherical triangulation G. Empirical studies of typical G suggest that $P(\lambda) \neq 0$ for $1 < \lambda < 2$, but a single zero is expected in the interval $2 < \lambda < 3$. Tutte [10, 11] proved that

$$0 < |P(\varphi + 1)| \leq \varphi^{5-n};$$

hence $\varphi + 1$, although not itself a zero of $P(\lambda)$, is arbitrarily close to being a zero for large enough n. For this reason, the constant $\varphi + 1$ is called the **golden root**.

It is known that $P(3) > 0$ if and only if G is Eulerian; that is, the number of edges incident with each vertex is even [5]. Hence for non-Eulerian triangulations, we have $P(3) = 0$.

Tutte [12–14] subsequently proved a remarkable identity:

$$P(\varphi + 2) = (\varphi + 2)\varphi^{3n-10} \left(P(\varphi + 1)\right)^2,$$

which implies that $P(\varphi + 2) > 0$. Note that $\varphi + 2 = \sqrt{5}\varphi = 3.6180339887\ldots$. As stated earlier, $P(4) > 0$, and $P(\lambda) > 0$ for $\lambda \geq 5$ [1]. It is natural to ask about the possible whereabouts of the next accumulation point for zeros (after $2.618\ldots$).

Rigorous theory fails us here, so numerical evidence must suffice [15–18]. In the following, fix a family $\{G_k\}$ of spherical triangulations, where n_k is the order of G_k and $n_k \to \infty$ as $k \to \infty$. Typically, the graph G_k is recursively constructed from G_{k-1} for each k, but this is not essential. Experimental results indicate that the next batch of chromatic zeros might cluster around the point

$$\psi = 2 + 2\cos\left(\frac{2\pi}{7}\right) = 4\cos\left(\frac{\pi}{7}\right)^2 = 3.2469796037\ldots,$$

that is, a solution of the cubic equation $\psi^3 - 5\psi^2 + 6\psi - 1 = 0$. The constant ψ is called the **silver root** by analogy with the golden root $\varphi + 1$.

Or the zeros might cluster around some other point $> \psi$, but ≤ 4. Beraha [19] observed a pattern in the potential accumulation points, independent of the choice of $\{G_k\}$. He conjectured that, for arbitrary $\{G_k\}$, if chromatic zeros z_k cluster around a real number x, then $x = B_N$ for some $N \geq 1$, where

$$B_N = 2 + 2\cos\left(\frac{2\pi}{N}\right) = 4\cos\left(\frac{\pi}{N}\right)^2.$$

In words, the limiting values x cannot fall outside of a certain countably infinite set. Note that the **Tutte–Beraha constants** B_N include all the roots already discussed:

$$\begin{aligned} &B_2 = 0, \quad B_3 = 1, \qquad B_4 = 2, \qquad\quad B_5 = \varphi + 1, \\ &B_6 = 3, \quad B_7 = \psi, \quad B_{10} = \varphi + 2, \quad \lim_{N\to\infty} B_N = 4. \end{aligned}$$

Specific families $\{G_k\}$ have been constructed that can be proved to possess B_5, B_7, or B_{10} as accumulation points [20–23]. The marvel of Beraha's conjecture rests in its generality: It applies regardless of the configuration of G_k.

Beraha & Kahane also built a family $\{G_k\}$ possessing $B_1 = 4$ as an accumulation point. This is surprising since we know $P(4) > 0$ always, but $P_k(z_k) = 0$ for all k and $\lim_{k\to\infty} z_k = 4$. Hence the Four Color Theorem, although true, is nearly false [24].

The Tutte–Beraha constants also arise in mathematical physics [25–28] since evaluating $P(\lambda)$ over a lattice is equivalent to solving the λ-state zero-temperature antiferromagnetic Potts model. A heuristic explanation of the Beraha conjecture in [27] is insightful but is not a rigorous proof [8]. See [5.24] for related discussion on coloring and ice models. Other expressions containing $\cos(\pi/7)$ are mentioned in [2.23] and [8.2].

[1] G. D. Birkhoff and D. C. Lewis, Chromatic polynomials, *Trans. Amer. Math. Soc.* 60 (1946) 355–451; MR 8,284f.
[2] T. L. Saaty and P. C. Kainen, *The Four-Color Problem: Assaults and Conquest*, McGraw-Hill, 1977; MR 87k:05084.
[3] R. C. Read and W. T. Tutte, Chromatic polynomials, *Selected Topics in Graph Theory. 3*, ed. L. W. Beineke and R. J. Wilson, Academic Press, 1988, pp. 15–42; MR 93h:05003.
[4] W. T. Tutte, Chromials, *Hypergraph Seminar*, Proc. 1972 Ohio State Univ. conf., ed. C. Berge and D. Ray-Chaudhuri, Lect. Notes in Math. 411, Springer-Verlag, 1974, pp. 243–266; MR 51 #5370.
[5] W. T. Tutte, The graph of the chromial of a graph, *Combinatorial Mathematics III*, Proc. 1974 Univ. of Queensland conf., ed. A. P. Street and W. D. Wallis, Lect. Notes in Math. 452, Springer-Verlag, 1975, pp. 55–61; MR 51 #10152.
[6] W. T. Tutte, Chromials, *Studies in Graph Theory. Part II*, ed. D. R. Fulkerson, Math. Assoc. Amer., 1975, pp. 361–377; MR 53 #10637.
[7] L. H. Kauffman, Map coloring and the vector cross product, *J. Combin. Theory Ser. B* 48 (1990) 145–154; MR 91b:05078.
[8] L. H. Kauffman and H. Saluer, An algebraic approach to the planar coloring problem, *Commun. Math. Phys.* 152 (1993) 565–590; MR 94f:05056.
[9] R. Thomas, An update on the four-color theorem, *Notices Amer. Math. Soc.* 45 (1998) 848–859; MR 99g:05082.
[10] G. Berman and W. T. Tutte, The golden root of a chromatic polynomial, *J. Combin. Theory* 6 (1969) 301–302; MR 39 #98.

[11] W. T. Tutte, On chromatic polynomials and the golden ratio, *J. Combin. Theory* 9 (1970) 289–296; MR 42 #7557.
[12] W. T. Tutte, More about chromatic polynomials and the golden ratio, *Combinatorial Structures and Their Applications*, Proc. 1969 Univ. of Calgary conf., ed. R. K. Guy, H. Hanani, N. Sauer, and J. Schönheim, Gordon and Breach, 1970, pp. 439–453; MR 41 #8299.
[13] W. T. Tutte, The golden ratio in the theory of chromatic polynomials, *Annals New York Acad. Sci.* 175 (1970) 391–402; MR 42 #130.
[14] D. W. Hall, On golden identities for constrained chromials, *J. Combin. Theory Ser. B* 11 (1971) 287–298; MR 44 #6510.
[15] D. W. Hall and D. C. Lewis, Coloring six-rings, *Trans. Amer. Math. Soc.* 64 (1948) 184–191; MR 10,136g.
[16] D. W. Hall, J. W. Siry, and B. R. Vanderslice, The chromatic polynomial of the truncated icosahedron, *Proc. Amer. Math. Soc.* 16 (1965) 620–628; MR 31 #3361.
[17] D. W. Hall, Coloring seven-circuits, *Graphs and Combinatorics*, Proc. 1973 George Washington Univ. conf., ed. R. A. Bari and F. Harary, Lect. Notes in Math. 406, Springer-Verlag, 1974, pp. 273–290; MR 51 #5366.
[18] W. T. Tutte, Chromatic sums for rooted planar triangulations. V: Special equations, *Canad. J. Math.* 26 (1974) 893–907; MR 50 #167.
[19] S. Beraha, *Infinite Non-trivial Families of Maps and Chromials*, Ph.D. thesis, Johns Hopkins Univ., 1974.
[20] S. Beraha, J. Kahane, and R. Reid, B_7 and B_{10} are limit points of chromatic zeroes, *Notices Amer. Math. Soc.* 20 (1973) A-5.
[21] S. Beraha, J. Kahane, and N. J. Weiss, Limits of zeroes of recursively defined polynomials, *Proc. Nat. Acad. Sci. USA* 72 (1975) 4209; MR 52 #5946.
[22] S. Beraha, J. Kahane, and N. J. Weiss, Limits of zeros of recursively defined families of polynomials, *Studies in Foundations and Combinatorics*, ed. G.-C. Rota, Academic Press, 1978, pp. 213–232; MR 80c:30005.
[23] S. Beraha, J. Kahane, and N. J. Weiss, Limits of chromatic zeros of some families of maps, *J. Combin. Theory Ser. B* 28 (1980) 52–65; MR 81f:05076.
[24] S. Beraha and J. Kahane, Is the four-color conjecture almost false?, *J. Combin. Theory Ser. B* 27 (1979) 1–12; MR 80j:05061.
[25] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982; MR 90b:82001.
[26] R. J. Baxter, Chromatic polynomials of large triangular lattices, *J. Phys. A.* 20 (1987) 5241–5261; MR 89a:82033.
[27] H. Saleur, Zeroes of chromatic polynomials: A new approach to Beraha conjecture using quantum groups, *Commun. Math. Phys.* 132 (1990) 657–679; MR 91k:17014.
[28] H. Saleur, The antiferromagnetic Potts model in two dimensions: Berker Kadanoff phases, antiferromagnetic transition and the role of Beraha numbers, *Nucl. Phys. B* 360 (1991) 219–263; MR 92j:82016.