Averaging Sequences, Deranged Mappings, and a Problem of Lambert and Slater

Neil J. Calkin
Department of Mathematical Sciences
Clemson University
Clemson, SC 29634-1907
calkin@math.clemson.edu

E. Rodney Canfield
Department of Computer Science
The University of Georgia
Athens, GA 30602, USA
erc@cs.uga.edu

Herbert S. Wilf
Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104-6395
wilf@math.upenn.edu

Dedicated to the memory of Gian-Carlo Rota

Abstract

We answer a question posed by Lambert and Slater [7]. Consider a sequence of real numbers q_n in the interval $[0,1]$ defined by $q_0 = 0$, $q_1 = 1$, and, for $n \geq 1$, q_{n+1} equals an average of preceding terms in the sequence. The weights used in the average are provided by a triangular array $p_{n,k}$ of probabilities whose row sums are 1. What is the limiting behavior of a sequence q_n so defined? For the Lambert-Slater sequence the weight $p_{n,k}$ is the probability that a randomly chosen fixed-point free mapping of $[n+1]$ omits exactly k elements from its image. To gain some insight into this averaging process, we first analyze what happens with a simpler array of weights $p_{n,k}$ defined in terms of binomial coefficients. One of our theorems states that if the weights $p_{n,k}$ are closely concentrated and the sequence q_n exhibits oscillatory behavior up to a certain computable point, then it will exhibit oscillatory behavior from then on. We carry out the computations necessary to verify that the Lambert-Slater sequence satisfies the hypotheses of the latter theorem. A result on martingales [1] is used to prove the close concentration of the weights $p_{n,k}$.

AMS-MOS Subject Classification (1990): 05A16, 60C05, 60G46

1 Introduction

In [7], the following question is raised. Begin with n players, and repeat the following "knockout" procedure while there remain two or more players. Each remaining player chooses another player uniformly at random; the set of players so chosen drop out of the game; that is, we knockout the chosen players. The game terminates when there is a single player, or no player. Lambert and Slater in their paper consider more general knockout processes on a graph G during a single round of which each vertex chooses at random a neighbor for knocking out. The process described above corresponds to G being the complete graph on n vertices.

The question is, as a function of n = the number of players who begin the game, what is the expected number q_n of players remaining at the end of the knockout process? Clearly, the numbers q_n all lie between 0 and 1. Our initial computations through $n = 170$ revealed a gentle oscillation between maxima in the neighborhood of 0.53 and minima in the neighborhood of 0.47, with an apparent slight tendency towards convergence. One might conjecture at that point that the limit is 1/2, and that the convergence is slow.

To the contrary, however, we have proven that for appropriate constants a, b, ϕ and all large n,

$$| \; a + b\cos(2\pi \log n + \phi) \; - \; q_n \; | \quad < \quad b. \tag{1}$$

This implies that the sequence q_n has no limit. (All logarithms in this paper are to the base e.)

What is the connection between the Lambert-Slater sequence q_n and deranged mappings? A *deranged mapping* of the set $[n] = \{1, 2, \ldots n\}$ is a function $f : [n] \to [n]$ such that $f(i) \neq i$ for all i. The number of such functions is $(n-1)^n$. Let $p_{n,k}$ be the probability that a randomly chosen deranged mapping of $[n+1]$ omits exactly k elements from its image. Then the Lambert-Slater sequence q_n is given by

$$q_0 = 0, \quad q_1 = 1, \quad q_{n+1} = \sum_{k=0}^{n} p_{n,k}\, q_k,\; n \geq 1. \tag{2}$$

The reader may note that the probabilities $p_{n,k}$ so defined are nonzero only for k in the range $0 \leq k \leq n-1$, because any mapping of $[n+1]$ which omits exactly n elements from its image is constant, and hence not deranged. Nevertheless, we write the basic recursion (2) as above, with the summation over the range $0 \leq k \leq n$, because we will wish to consider other underlying arrays $p_{n,k}$. In Section 2, in particular, we analyze a sequence q_n defined by (2) with a different underlying triangular array $p_{n,k}$. In Section 3 we gather the necessary results on deranged mappings. In Section 4, we prove (Theorem 3) that if the given probabilities $p_{n,k}$ satisfy certain properties (foremost of which is close concentration), and if a sufficiently long segment of the sequence q_n exhibits a type of oscillatory behavior, then, in fact, the sequence q_n oscillates indefinitely. Finally, in Section 5, we describe the computations used to verify that the probabilities $p_{n,k}$ and sequence q_n posed by Lambert and Slater do indeed satisfy the hypotheses of Theorem 3. We conclude by proving (1).

The roles of q_n and $p_{n,k}$ change from section to section. The definition of $p_{n,k}$ is given at the start of each section, and always q_n is determined by (2).

2 The Coin Flipping Game

In this section we consider the following problem: initially n coins are all heads up on a table top. Repeat the following process until only one or none of the coins is heads up: flip exactly once all of the coins that are still showing heads. What is the probability q_n that we terminate with exactly one head?

This question was raised as a MONTHLY Problem in 1991 and a solution was given in 1994 [9]. The answer was given in just sufficient detail to answer the question precisely as asked, which was to decide if the limit exists or not (it doesn't). Here we give a more detailed solution to give some additional insight into the sort of behavior that such problems exhibit, in a context that is simpler than the knockouts problem that will be discussed below, and is the main object of this paper.

It is clear that the sequence q_n is given by the recursion (2) with the underlying probabilities:

$$p_{n,k} \quad = \quad (2^{n+1}-1)^{-1}\binom{n+1}{k}, \quad 0 \leq k \leq n.$$

We shall compute the ordinary generating function

$$f(x) \quad = \quad \sum_{n=0}^{\infty} q_n x^n$$

for this sequence. Summing

$$q_n x^n \quad = \quad \sum_{k=0}^{n} \binom{n}{k} 2^{-n} q_k x^n$$

over $n \geq 2$, we obtain

$$\begin{aligned} f(x) - x \quad &= \quad \sum_{n\geq 0}\sum_{k=0}^{n} \binom{n}{k} 2^{-n} q_k x^n - q_0 - x\left(\frac{q_0+q_1}{2}\right) \\ &= \quad -\frac{x}{2} \;+\; \frac{2}{x-2}\sum_{n\geq 0} q_n \left(\frac{x}{2-x}\right)^n, \end{aligned}$$

so

$$f(x) \quad = \quad \frac{x}{2} + \frac{2}{x-2} f\left(\frac{x}{2-x}\right)$$

and now, by iterating this, we obtain

$$\begin{aligned} f(x) \quad &= \quad \frac{x}{2} \;+\; \frac{x}{(2-x)^2} \;+\; \frac{2x}{(4-3x)^2} \;+\; \frac{4x}{8-7x} \;+\; \frac{8x}{(16-15x)^2} \;+\; \cdots \\ &= \quad \frac{x}{2}\sum_{n=0}^{\infty} \frac{2^n}{(2^n-(2^n-1)x)^2}. \end{aligned}$$

With this generating function we are now ready to prove the nonexistence of the limit. In fact we will give three different proofs of this. Let's state it as a theorem.

Theorem 1 *The probability q_n of terminating with exactly one head in an n-head game doesn't tend to a limit.*

2.1 First proof of Theorem 1

Now we can get an explicit, finite formula for the coefficients q_n. Indeed we have

$$\begin{aligned}
f(x) &= \frac{x}{2}\sum_{m\geq 0} 2^{-m}\frac{1}{(1-(1-2^{-m})x)^2} \\
&= \frac{x}{2}\sum_{m\geq 0} 2^{-m}\sum_{\ell\geq 0}(\ell+1)(1-2^{-m})^{\ell}x^{\ell} \\
&= \sum_{\ell\geq 0}\frac{(\ell+1)}{2}x^{\ell+1}\sum_{m\geq 0}2^{-m}(1-2^{-m})^{\ell} \\
&= \sum_{\ell\geq 0}\frac{(\ell+1)}{2}x^{\ell+1}\sum_{m\geq 0}2^{-m}\sum_{j}\binom{\ell}{j}(-1)^j 2^{-mj} \\
&= \sum_{\ell\geq 0}\frac{(\ell+1)}{2}x^{\ell+1}\sum_{j}\binom{\ell}{j}(-1)^j\frac{1}{1-2^{-j-1}}.
\end{aligned}$$

The coefficient of x^n gives us

$$q_n = \frac{n}{2}\sum_j\binom{n-1}{j}\frac{(-1)^j}{1-2^{-j-1}} = \frac{n}{2}\left\{\delta_{n,1}+\sum_j\binom{n-1}{j}\frac{(-1)^j}{2^{j+1}-1}\right\}$$

a finite, explicit formula for our coefficients.

Sums of this type have appeared elsewhere. Precisely this sequence q_n is treated in the analysis of a probabilistic model in number theory [8]. In [4] we find a discussion of the quantities

$$\Sigma_2(j) = \sum_k\binom{n}{k}(-1)^k\frac{1}{2^{k+j}-1}.$$

In terms of them, our q_n's are given by

$$q_{n+1} = \frac{n}{2}(\delta_{n,1}+\Sigma_2(1)).$$

Now a formula for our sequence can be read off from that of the Σ's, as given in the latter paper. We have

$$q_n = \frac{1}{2\log 2}+\frac{1}{2}\delta_{n,1}+\frac{ne^{-H_{n-1}}}{2\log 2}\sum_{m\neq 0}\frac{e^{\frac{2m\pi i H_{n-1}}{\log 2}}}{(1-\frac{2m\pi i}{\log 2})\prod_{k=1}^{n-1}\left\{\left(1+\frac{1-\frac{2m\pi i}{\log 2}}{k}\right)e^{(\frac{2m\pi i}{\log 2}-1)/k}\right\}},$$

where the H's are the harmonic numbers, and γ is Euler's constant. As $n \to \infty$, the above formula is

$$q_n = \frac{1}{2\log 2}\left(1 + \sum_{m\neq 0} \Gamma\left(1 - \frac{2m\pi i}{\log 2}\right) e^{2m\pi i \log n}\right) + o(1).$$

Note that q_n is a periodic function of $\log n$, to accuracy $o(1)$. A similar formula appears in [5, Section 5.2.2, equation (47)], where it arises in the asymptotic analysis of a sorting algorithm.

2.2 Second proof of Theorem 1

This time we use the fact that if

$$\lim_{n\to\infty} q_n = L$$

exists, then

$$\lim_{\epsilon\to 0^+} \epsilon f(1-\epsilon) = L.$$

Hence

$$\begin{aligned} \epsilon f(1-\epsilon) &= \frac{\epsilon(1-\epsilon)}{2}\sum_{k=0}^{\infty} \frac{2^k}{(2^k-(2^k-1)(1-\epsilon))^2} \\ &= \frac{\epsilon(1-\epsilon)}{2}\sum_{k=0}^{\infty} \frac{2^k}{(1+(2^k-1)\epsilon)^2}. \end{aligned}$$

It is easy to see that the behaviour of this when ϵ is small is close to the behaviour of

$$F(\epsilon) = \frac{\epsilon}{2}\sum_k \frac{2^k}{(1+2^k\epsilon)^2}.$$

Write $u(t) = 2^t/(1+\epsilon 2^t)^2$, and $\psi(t) = t - \lfloor t \rfloor - 1/2$. By the Euler-Maclaurin sum formula,

$$\begin{aligned} F(\epsilon) &= \frac{\epsilon}{2}\left\{\int_{-\infty}^{\infty} u(t)dt + \int_{-\infty}^{\infty} \psi(t)u'(t)dt\right\} \\ &= \frac{1}{2\log 2} + \frac{\epsilon}{2}\int_{-\infty}^{\infty} \psi(t)u'(t)dt \\ &= \frac{1}{2\log 2} + \frac{1}{2}\int_0^{\infty} \psi(\log\frac{t}{\epsilon})\frac{1-t}{(1+t)^3}dt \end{aligned}$$

In the integral, if we cut ϵ in half, we add 1 to the argument of ψ, which doesn't change its value. So the integral is invariant under halving of ϵ. To learn more, we replace ψ by its Fourier series

$$\psi(t) = -\sum_{n\geq 1} \frac{\sin 2n\pi t}{n\pi},$$

and integrate termwise (which Hardy is fond of noting cannot be justified by absolute convergence but it can by bounded convergence) to obtain,

$$\begin{aligned}
F(\epsilon) &= \frac{1}{2\log 2} + \frac{1}{2}\int_0^\infty \psi(\log\frac{t}{\epsilon})\frac{1-t}{(1+t)^3}dt \\
&= \frac{1}{2\log 2} - \frac{1}{2}\sum_{n\geq 1}\frac{1}{n\pi}\int_0^\infty \sin\left(2n\pi\log\frac{t}{\epsilon}\right)\frac{1-t}{(1+t)^3}dt \\
&= \frac{1}{2\log 2} - \frac{1}{2}\sum_{n\geq 1}\frac{1}{n\pi}\int_0^\infty \sin(2n\pi\log t)\cos(2n\pi\log\epsilon)\frac{1-t}{(1+t)^3}dt + \cdots \\
&\qquad \cdots + \frac{1}{2}\sum_{n\geq 1}\frac{1}{n\pi}\int_0^\infty \cos(2n\pi\log t)\sin(2n\pi\log\epsilon)\frac{1-t}{(1+t)^3}dt \\
&= \frac{1}{2\log 2} - \frac{1}{2}\sum_{n\geq 1}\frac{1}{n\pi}\cos(2n\pi\log\epsilon)\left(\frac{-4n^2\pi^3}{(\log 2)^2\sinh(2n\pi^2/\log 2)}\right) \\
&= \frac{1}{2\log 2} + \frac{1}{(\log 2)^2}\sum_{n\geq 1}\frac{2n\pi^2}{\sinh(2n\pi^2/\log 2)}\cos(2n\pi\log\epsilon),
\end{aligned}$$

showing that $F(\epsilon)$ is periodic in $\log\epsilon$ (and not constant), and hence that

$$\lim_{\epsilon\to 0^+}\epsilon f(1-\epsilon)$$

doesn't exist.

2.3 Third proof of Theorem 1

$$q_n \sim \frac{1}{2}\sum_{k=-\infty}^{\infty} 2^{\theta+k}e^{-2^{\theta+k}},$$

where θ is the fractional part of the base-2 logarithm of n:

$$n = 2^{t+\theta}.$$

This implies that q_n doesn't tend to a limit. To prove the asserted asymptotic formula, we extract the coefficient of x^n in $f(x)$, to find

$$q_n = \frac{n}{2}\sum_{k=0}^{\infty} 2^{-k}\left(1-2^{-k}\right)^{n-1}.$$

Hence,

$$\begin{aligned} q_n &= \frac{1}{2}\sum_{k=0}^{\infty} 2^{t+\theta-k}\left(1-\frac{2^{t+\theta-k}}{n}\right)^{n-1} \\ &= \frac{1}{2}\sum_{k=-t}^{\infty} 2^{\theta-k}\left(1-\frac{2^{\theta-k}}{n}\right)^{n-1}. \end{aligned}$$

Now, if $-t \leq k < -t/3$, then

$$\begin{aligned} \left(1-\frac{2^{\theta-k}}{n}\right)^{n-1} &< \left(1-\frac{2^{\frac{\theta+t}{3}}}{n}\right)^{n-1} \\ &= \left(1-\frac{n^{1/3}}{n}\right)^{n-1} \\ &\sim e^{-n^{1/3}}, \end{aligned}$$

and so

$$q_n = \sum_{k\geq -t/3} 2^{\theta-k}\left(1-\frac{2^{\theta-k}}{n}\right)^{n-1} + o\left(n^{-1}\right).$$

Now, if $k \geq -t/3$, then

$$\left(1-\frac{2^{\theta-k}}{n}\right)^{n-1} = e^{-2^{\theta-k}}\left(1+o\left(\frac{2^{\theta-k}+2^{2\theta-2k}}{n}\right)\right),$$

and so

$$\begin{aligned} q_n &= \frac{1}{2}\sum_{k\geq t/3} 2^{\theta-k}e^{-2^{\theta-k}}\left(1+o\left(\frac{2^{\theta-k}+2^{2\theta-2k}}{n}\right)\right) \quad + \quad o\left(\frac{1}{n}\right) \\ &= \frac{1}{2}\sum_{k=-\infty}^{\infty} 2^{\theta-k}e^{-2^{\theta-k}} \; + \; O\left(\frac{1}{n}\right) \\ &= \frac{1}{2}\sum_{k=-\infty}^{\infty} 2^{\theta+k}e^{-2^{\theta+k}} \; + \; O\left(\frac{1}{n}\right), \end{aligned}$$

as claimed.

By taking more care with the expansion of $(1-\frac{x}{n})^{n-1}$, we obtain

$$q_n = s_1(\theta)+\frac{1}{n}\left(s_2(\theta)-\frac{1}{2}s_3(\theta)\right)+\frac{1}{n^2}\left(s_3(\theta)-\frac{5}{6}s_4(\theta)+\frac{1}{8}s_5(\theta)\right)+O\left(\frac{1}{n^3}\right),$$

in which

$$s_j(\theta) \;=\; \frac{1}{2}\sum_{k=-\infty}^{\infty} 2^{j(\theta+k)}e^{-2^{\theta+k}}.$$

This can, of course, be expanded to any fixed number of terms. By way of example, actual computation gives $q_{512} = 0.721352430446\ldots$, and

$$s_1(0) + \frac{1}{n}\left(s_2(0) - \frac{1}{2}s_3(0)\right) + \frac{1}{n^2}\left(s_3(0) - \frac{5}{6}s_4(0) + \frac{1}{8}s_5(0)\right) = 0.7213524304938\ldots.$$

3 Properties of Deranged Mappings

Throughout this section $p_{n,k}$ equals the probability that a randomly chosen deranged mapping of $[n+1]$ omits exactly k points from its image, $0 \le k \le n$. In Section 5 we give an algorithm for computing these probabilities. The literature on random mappings is vast, see for example [6], although the fixed-point-free property does not appear often. In [2] the asymptotic distribution for many statistics on mappings is derived, including the normality of the image size. Of course, the number of elements omitted from the range differs from the latter only by a constant. It turns out we do not need information on the distribution of this statistic so much as we need a bound on deviation from average behavior; that is, a close concentration result. The necessary theorem appears in Chapter 7 (Martingales) of the book [1] by Alon, Spencer, and Erdős.

The total number of deranged mappings of $[n+1]$ is n^{n+1}, since in constructing a function f, for each integer i in the domain there are n choices for $f(i)$. Assigning each of these mappings the probability n^{-n-1} is an instance of the following general situation: let there be given a domain A, a range B, and an $|A| \times |B|$ matrix of probabilities whose row sums are all 1. The entry appearing at row a and column b equals the probability that a function $f : A \to B$ satisfies $f(a) = b$, the latter events being independent over different $a \in A$. The given matrix determines a probability distribution on the set of all functions $f : A \to B$. The uniform probability space which we are considering for deranged mappings is the case in which both A and B have size $n+1$, all diagonal entries of the matrix are 0, and all off diagonal entries are $1/n$.

Continuing to follow [1, p. 89], let $L : A^B \to \Re$ be a functional (we shall be interested in $L(f)$ being the number of points omitted from the image of f), and let

$$\emptyset = B_0 \subseteq B_1 \subseteq \cdots \subseteq B_m = B \tag{3}$$

be an m-gradation of the range B. The sequence $X_0, X_1, \ldots X_m$ defined on functions $f : A \to B$ by

$$X_i(f) \;\stackrel{\text{def}}{=}\; E\left(L(g) \mid g(b) = f(b) \text{ for all } b \in B_i\right)$$

is a martingale. (See [1] for undefined terms.) Note that $X_0(f)$ is the constant $E(L)$, independently of f, and $X_m(f)$ is $L(f)$. We need the following concentration result:

Theorem 2 *[1, p. 90]. Let a probability measure on a finite function space A^B be given as above by an $|A| \times |B|$ matrix of non negative numbers whose row sums are 1, and let $L(f)$ be a functional satisfying the Lipschitz condition*

$$f, f' \text{ differ only on } B_{i+1} - B_i \quad \Rightarrow \quad |L(f) - L(f')| \leq 1$$

with respect to a given gradation (3). Then, for $\mu = E(L)$ and all $\lambda > 0$ we have

$$\begin{aligned} Pr\left(L(f) > \mu + \lambda m^{1/2}\right) &< e^{-\lambda^2/2} \\ Pr\left(L(f) < \mu - \lambda m^{1/2}\right) &< e^{-\lambda^2/2}. \end{aligned}$$

Let us remark that with $B_i = [i]$, $0 \leq i \leq n+1$, and $L(f) =$ the number of points omitted from the range of f, the Lipschitz condition is satisfied. Thus, we have a close concentration result for the probabilities $p_{n,k}$.

We complete this section by computing the mean μ and standard deviation σ of the distribution:

$$\begin{aligned} \mu &= \sum_{k=0}^{n} k\, p_{n,k} \\ \sigma^2 &= \sum_{k=0}^{n} (k-\mu)^2\, p_{n,k}. \end{aligned}$$

Of course the mean is simply $n+1$ times the probability that a particular element, say 1, is omitted from the range:

$$\mu = (n+1)(1-1/n)^n.$$

Turning to σ, we start with

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

which is obtained by counting triples (i, j, f), with f a fixed-point-free mapping, i, j two distinct elements not in the range of f, and then dividing by the total number of fixed-point-free mappings. Adding $\sum_{k=0}^{n} kp_{n,k}$, and subtracting μ^2, we find

$$\sigma^2 = (n+1)n(1-1/n)^2(1-2/n)^{n-1} \; + \; (n+1)(1-1/n)^n \; - \; (n+1)^2(1-1/n)^{2n}.$$

Some careful calculation shows that σ^2 differs from $(n+1)e^{-1}(1-2e^{-1})$ by less than 1. We record for future use:

$$\begin{aligned} e^{-1}(1-n^{-1}) &< (1-1/n)^n < e^{-1} && \text{for } n \geq 2 \\ |\sigma^2 - (n+1)e^{-1}(1-2e^{-1})| &< 1 && \text{for } n \geq 1. \end{aligned} \tag{4}$$

4 A Theorem About Oscillatory Behavior

Throughout this section $p_{n,k}$ denotes a generic array of probabilities about which we shall assume certain hypotheses and prove certain results. We shall use $F(X)$ to denote the function

$$F(X) = a + b\cos(\lambda X + \phi),$$

the constants a, b, λ, ϕ being given. With no loss one may take the constants b, λ, ϕ positive, and we do so.

An array $p_{n,k}$, determining a sequence q_n, and a function $F(X)$ understood to have been given, for integers I and J we define

$$\Delta(I, J] \quad \stackrel{\text{def}}{=} \quad \max\{|q_k - F(\log k)| \ : \ I < k \leq J\},$$

the maximum absolute difference between q_k and the approximation $F(\log k)$ over the half open interval $(I, J]$.

We introduce one additional convenient notation involving the symbol θ. Whenever θ appears in an equation, it stands for a real number whose value is in the interval $[-1, +1]$. It is, of course, not necessarily the same value at each appearance; moreover, in a given equation, its value may depend on the free variables found in the equation. If we want to say, for example, that two real valued functions $f(x)$ and $g(x)$ differ in absolute value by no more than $10/x^2$ for all $x \in S$, we would write

$$f(x) \;=\; g(x) \;+\; 10\theta x^{-2}, \quad x \in S.$$

Precisely, this says that θ, defined on S by $\theta(x) = (f(x) - g(x))x^2/10$, never exceeds 1 in absolute value. Two further examples of this new notation, both used in the sequel, are

$$\log(1 + y) \;=\; y \;+\; 0.65\theta y^2, \quad y \geq -1/3 \tag{5}$$

$$F(X + y) \;=\; F(X) \;+\; F'(X)y \;+\; \theta b\lambda^2 y^2/2. \tag{6}$$

The first is proven by noting that $(\log(1 + y) - y)/y^2$ is an increasing function of $y > -1$. The second follows from Taylor's formula with remainder.

The next theorem gives conditions on $p_{n,k}$ to quantify and prove the notion that if $F(\log k)$ approximates q_k well on a sufficiently large interval $(I, J\,]$, then there is a substantially larger interval $(I, K]$ where the approximation is only slightly less good.

Theorem 3 *Given: a triangular array $p_{n,k}$, $0 \leq k \leq n$, $n \geq 1$, of probabilities whose row sums equal* 1*, and six positive constants α, β, a, b, λ, ϕ. Let q_n be defined by the recursion (2), $F(X) = a + b\cos(\lambda X + \phi)$, $\Delta(I, J]$ be defined by (4), μ $(= \mu(n))$ be the mean $\sum_k kp_{n,k}$ of the n-th row, σ^2 $(= \sigma^2(n))$ be the variance $\sum_k (k - \mu)^2 p_{n,k}$ of the n-th row, and $\Re_n$ be the set of k such that*

$|k-\mu| \leq (n+1)^{2/3}$. *Assume the following conditions:*

$$\begin{aligned}
\mu &= (n+1)\,\alpha\,(1+\theta/n) \\
\lambda\,\log\alpha &= -2\pi \\
\sigma^2 &< \beta(n+1) \;+\; 1, \qquad (n \geq 2) \\
\sum_{k \notin \Re_n} p_{n,k} &\leq 2e^{-n^{1/3}/2}.
\end{aligned}$$

Then there exist constants C_1, C_2, $\delta > 0$, *and* N *such that for every pair of integers* $I < J$ *satisfying*

$$J \;\geq\; N \quad \text{and} \quad \alpha\,J\,(1 - J^{-1} - \alpha^{-1}J^{-1/3}) \;\geq\; I+1, \tag{7}$$

there exists an integer $K \geq (1+\delta)J$ *such that*

$$\Delta(J,K] \;\leq\; \Delta(I,J] \;+\; C_1/J \;+\; C_2 e^{-J^{1/3}/2}. \tag{8}$$

Proof. We begin by telling how to choose C_1, C_2, δ, and N. Let

$$c \;=\; b\lambda\,(0.75\lambda \;+\; 0.65).$$

Choose N so large and δ positive but so small that

$$N \;\geq\; \max\Big(\,(3/\alpha + 0.01)^3,\; 90/\beta,\; 1000\,\Big)$$

and

$$\alpha\,(1+\delta+N^{-1})\,(1+N^{-1})\,(1+N^{-1}+\alpha^{-1}N^{-1/3}) \;=\; 1.$$

Then, choose C_1 and C_2 by the formulas

$$\begin{aligned}
C_1 &= 1.02\,c\beta\alpha^{-2} \;+\; b\lambda \\
C_2 &= 2\,(1 \;+\; |a| \;+\; b \;+\; m(b\lambda + cm))\,, \quad m = \max(\,1,\,1.0001\alpha^{-1} - 1\,).
\end{aligned}$$

Let $I < J$ be two integers satisfying condition (7). We claim that K may be taken as $\lceil(1+\delta)J\,\rceil$. To see this, let $n+1$ be an integer in the half open interval $(J,K]$. (In what follows, we sometimes use $n \geq J \geq N$ without explicit mention.) We have

$$q_{n+1} \;=\; \sum_{k \in \Re_n} p_{n,k}\,q_k \;+\; \sum_{k \notin \Re_n} p_{n,k}\,q_k.$$

Letting $E = 2e^{-n^{1/3}/2}$, we see that the second sum on the right of the previous equation equals θE. To bound the first sum, we first check that $\Re_n \subseteq (I,J]$. By assumption, $\mu - (n+1)^{2/3} \geq$

$(n-n^{-1})\alpha-(n+1)^{2/3}$; since the latter is increasing for $n \geq N$,

$$\begin{aligned}
\mu-(n+1)^{2/3} &\geq (J-J^{-1})\alpha-(J+1)^{2/3} \\
&= \alpha(J+1)(1-J^{-1}-\alpha^{-1}(J+1)^{-1/3}) \\
&\geq \alpha J(1-J^{-1}-\alpha^{-1}J^{-1/3}) \\
&\geq I+1,
\end{aligned}$$

and $\Re_n \subseteq (I,+\infty)$. In the other direction, by assumption, $\mu \leq (n+1)\alpha(1+n^{-1})$; since the latter is an increasing function of n,

$$\begin{aligned}
\mu \;+\; (n+1)^{2/3} &\leq (K+1)\alpha(1+K^{-1}) \;+\; (K+1)^{2/3} \\
&= \alpha K(1+K^{-1})(1+K^{-1}+\alpha^{-1}(K+1)^{-1/3}) \\
&\leq \alpha K(1+K^{-1})(1+K^{-1}+\alpha^{-1}K^{-1/3}) \\
&\leq \alpha K(1+N^{-1})(1+N^{-1}+\alpha^{-1}N^{-1/3}) \\
&\leq J\alpha(1+\delta+J^{-1})(1+N^{-1})(1+N^{-1}+\alpha^{-1}N^{-1/3}) \\
&\leq J,
\end{aligned}$$

and $\Re_n \subseteq (-\infty, J]$. Hence, as asserted, $\Re_n \subseteq (I,J]$, and

$$\sum_{k\in\Re_n} p_{n,k}\, q_k \;=\; \sum_{k\in\Re_n} p_{n,k}\, F(\log k) \;+\; \theta\Delta(I,J].$$

For the duration of the proof, let x $(= x(n,k))$ be implicitly defined by

$$k \;=\; \mu \;+\; x(n+1)^{1/2}.$$

Using (for $k \in \Re_n$)

$$\begin{aligned}
\left|\frac{k}{\mu} \;-\; 1\right| &\leq \frac{(n+1)^{2/3}}{\mu} \\
&\leq \frac{(n+1)^{2/3}}{(n+1)\alpha(1-1/n)} \\
&\leq \frac{n^{2/3}}{\alpha(n-1)},
\end{aligned}$$

and the assumption that $n^{1/3} \geq 3/\alpha + n^{-2/3}$, we have (again, for $k \in \Re_n$),

$$\begin{aligned}
k &= \mu\left(1 \;+\; \frac{x(n+1)^{1/2}}{\mu}\right) \\
&= \mu\,(1 \;+\; \theta/3)\,.
\end{aligned}$$

Thus, (5) is applicable, and

$$\log(k) \;=\; \log(\mu) \;+\; \frac{x(n+1)^{1/2}}{\mu} \;+\; 0.65\theta\frac{x^2(n+1)}{\mu^2}, \quad \text{for } k \in \Re_n.$$

Using (6) we calculate, for $k \in \Re_n$,

$$F(\log k) \;=\; F(\log\mu) \;+\; F'(\log\mu)\frac{x(n+1)^{1/2}}{\mu} \;+\; \theta\frac{cx^2(n+1)}{\mu^2}.$$

To continue, we compute

$$\begin{aligned}\sum_{k\in\Re_n} x\,p_{n,k} \;&=\; \sum_k x\,p_{n,k} \;-\; \sum_{k\notin\Re_n} x\,p_{n,k} \\ &=\; \theta\,E\,\frac{\max(\mu, n-\mu)}{(n+1)^{1/2}},\end{aligned}$$

as well as

$$\begin{aligned}\sum_{k\in\Re_n} x^2\,p_{n,k} \;&=\; \sum_k x^2\,p_{n,k} \;-\; \sum_{k\notin\Re_n} x^2\,p_{n,k} \\ &=\; \frac{\sigma^2}{n+1} \;+\; \theta\,E\,\frac{\max(\mu^2, (n-\mu)^2)}{n+1}.\end{aligned}$$

With $m = \max(1, n/\mu - 1)$, we have altogether

$$q_{n+1} \;=\; F(\log\mu) \;+\; \theta\frac{c\sigma^2}{\mu^2} \;+\; \theta\Delta(I,J] \;+\; \theta E\,(1 + |\,a| + b + m(b\lambda + cm))\,. \tag{9}$$

Using the assumptions about μ, α, and λ, we have

$$F(\log\mu) \;=\; F\,(\log(n+1)) \;+\; \theta b\lambda n^{-1},$$

which in conjunction with (9) implies (8). This completes the proof of Theorem 3.

Corollary Let q_n and $F(X)$ satisfy the hypotheses of Theorem 3, and let C_1, C_2, δ, N be the constants given by that theorem. Let I be the largest integer possible, subject to

$$\alpha\,N\,(1 - N^{-1} - \alpha^{-1}N^{-1/3}) \;\geq\; I + 1.$$

Then, for every $n > I$,

$$|\,q_n \;-\; F(\log n)\,| \;\leq\; \Delta(I,N] \;+\; C_1 N^{-1}(1+\delta^{-1}) \;+\; C_2\,e^{-N^{1/3}/2}\,\frac{1}{1 - e^{-\delta N^{1/3}/6}}.$$

Proof. By the definition of Δ the inequality holds for $I < n \le N$. Let $J_0 = N$ and $J_1 = \lceil (1+\delta) J_0 \rceil$. For $I < n \le J_1$, we have by Theorem 3,

$$|\, q_n \;-\; F(\log n)\,| \;\;\le\;\; \Delta(I,N] \;+\; C_1 J_0^{-1} \;+\; C_2\, e^{-J_0^{1/3}/2}.$$

Let $J_2 = \lceil (1+\delta) J_1 \rceil$; again by Theorem 3, for $I < n \le J_2$, we have

$$|\, q_n \;-\; F(\log n)\,| \;\;\le\;\; \Delta(I,N] \;+\; C_1 (J_0^{-1} + J_1^{-1}) \;+\; C_2\, (e^{-J_0^{1/3}/2} \;+\; e^{-J_1^{1/3}/2}).$$

By induction, with $J_{i+1} = \lceil (1+\delta) J_i \rceil$, we find that for all $n > I$,

$$|\, q_n \;-\; F(\log n)\,| \;\;\le\;\; \Delta(I,N] \;+\; C_1 \sum_{i=0}^{\infty} J_i \;+\; C_2 \sum_{i=0}^{\infty} e^{-J_i^{1/3}/2}.$$

Since $J_i \ge (1+\delta)^i N$, we have

$$\sum_{i=0}^{\infty} J_i \;\;\le\;\; J_0^{-1}(1+\delta^{-1});$$

and, since $(1+\delta)^{i/3} \ge (1 + i\delta/3)$,

$$\begin{aligned} \sum_{i=0}^{\infty} e^{-J_i^{1/3}/2} \;&\le\; \sum_{i=0}^{\infty} e^{-J_0^{1/3}(1+\delta)^{i/3}/2} \\ &\le\; \sum_{i=0}^{\infty} e^{-J_0^{1/3}(1+i\delta/3)/2} \\ &=\; e^{-N^{1/3}/2}\, \frac{1}{1 - e^{-N^{1/3}\delta/6}}. \end{aligned}$$

The Corollary follows.

5 Computations

Throughout this section $p_{n,k}$ again equals the probability that a randomly chosen deranged mapping on the set $[n+1]$ omits exactly k points from its image. Define $t(n,k)$ to be the number of ways to partition the set $[n]$ into an *ordered* collection of $n-k$ blocks, such that for $1 \le i \le n-k$ element i does not belong to the i-th block. Such an ordered partition corresponds in a natural way to a fixed-point-free mapping of $[n]$ whose image is exactly the set $\{1, 2, \ldots n-k\}$. Hence,

$$n^{n+1} \;\;=\;\; \sum_{k=0}^{n} \binom{n+1}{k} t(n+1,k),$$

and

$$p_{n,k} \quad = \quad \binom{n+1}{k} \; t(n+1,k) \; n^{-n-1}.$$

A mapping which omits no element from its image is a fixed-point-free permutation, also known as a *derangement*. The recursion for counting derangements is well known [3], and we have

$$t(n,0) = d_n, \quad d_0 = 1, \quad d_1 = 0, \quad d_{n+1} = n(d_n + d_{n-1}), \; n \geq 1.$$

Now let us consider $t(n+1,k)$ when k is at least 1. The ordered partitions counted by $t(n+1,k)$ are of two varieties. In the first variety, we have those partitions in which element $n+1$ is a singleton block; in the second variety we have those partitions in which element $n+1$ belongs to a block of size two or greater. To create an ordered partition of the first type, we proceed in three steps: (1) choose an integer j, in the range $1 \leq j \leq n-k+1$; (2) choose an ordered partition of $[n]$ into $n-k$ blocks such that i is not in the i-th block for $1 \leq i < j$ and such that $i+1$ is not in the i-th block for $j \leq i \leq n-k$; (3) insert $\{n+1\}$ as the j-th block. We defined $t(n,k)$ as counting ordered partitions of $[n]$ into $n-k$ blocks such that element i is forbidden from the i-the block. However, a moment's reflection will reveal that $t(n,k)$ will also count correctly *any* scheme of forbidding in which a certain element is denied membership in the i-th block, and the $n-k$ elements so singled out are all distinct. Hence, the number of ordered partitions in step (2) above equals $t(n,k)$, and the total number of partitions of the first variety is $(n-k+1)\,t(n,k)$.

To create an ordered partition of the second type: (1) partition $[n]$ into $n-k+1$ blocks, keeping i out of the i-th block; (2) choose a block into which $n+1$ is then inserted. Note that step (1) is feasible, and that step (2) does not create any forbidden memberships, due to the assumption $k \geq 1$. Summarizing, we have the nice recursion:

$$t(n+1,k) \quad = \quad (n-k+1) \quad (\; t(n,k) \; + \; t(n,k-1) \;), \qquad n \geq 0, k \geq 1. \tag{10}$$

Let us remark that if we fill the $k=0$ column of the array with the factorials $1, 1, 2, 6, 24, \ldots$ instead of the derangement numbers $1, 0, 1, 2, 9, \ldots$, and then fill the rest of the table (where $k \geq 1$) by exactly the same recursion (10), the resulting table contains $(n-k)!\,S(n, n-k)$, where $S(n,b)$ is the Stirling number of the second kind. The probabilities $p_{n,k}$ associated with $(n-k)!\,S(n,n-k)$ correspond to choosing a mapping of $[n+1]$ at random, with no requirement the mapping be fixed-point-free. This amounts to a Lambert-Slater knockout game in which self-elimination is permitted.

Altogether, then, we can compute initial rows of the $t(n,k)$ array in a number of arithmetic operations which is proportional to the number of values computed. This suggests that the first n values of the sequence q_n can be computed in quadratic time, but such a conclusion ignores a further multiplicative factor of $n \log n$ in the complexity due to the size of the integer operands involved. It is, however, feasible to compute the first $1,776$ values of q_n using the above scheme, provided the calculation is done in floating point, and not exactly. There arises the question of rounding error. To confirm reliability, the computations have been carried out in two different precisions, first with Digits = 22, then with Digits = 32. (For those unfamiliar with the symbolic computation system Maple, "Digits" is a global variable set by the user which controls the number of digits kept in floating point computations.) The two results agree, out to $n = 1776$, in the first sixteen places always. For anyone wishing to repeat the calculations, we report the following sample values, obtained with Digits = 32,

$$\begin{aligned} q_{1000} &= 0.47675343531232572822205635018666 \\ q_{1776} &= 0.52829933875860791739826500429501. \end{aligned}$$

Note that the four assumptions needed to apply Theorem 3 are fulfilled: the inequalities needed for μ and σ are implied by (4), and the concentration result for $p_{n,k}$ is given in Theorem 2. Because $\alpha = e^{-1}$, we take $\lambda = 2\pi$. The values chosen for a, b, and ϕ were determined by a least squares fit, although the source of these numbers is irrelevant to the use of the Theorem. Here is the summary of all computations:

$$\begin{aligned} a &= 0.5029602 \\ b &= 0.0268190 \\ \lambda &= 2\pi \\ \phi &= 3.5514971 \\ \alpha &= e^{-1} \\ \beta &= e^{-1}(1-2e^{-1}) \\ N &= 1776 \\ c &= 0.9036094 \\ C_1 &= 0.8305359 \\ m &= \max(\,1, 1.0001\alpha^{-1} - 1\,) \quad = \quad 1.7185537 \\ C_2 &= 8.9762295 \\ \delta &= 1.2171447 \\ I &= 505 \\ \Delta(I,N] &< 0.00036 \end{aligned}$$

One final act of arithmetics reveals

$$\Delta(I,N] \;+\; C_1 N^{-1}(1+\delta^{-1}) \;+\; C_2\, e^{-N^{1/3}/2}\,\frac{1}{1-e^{-N^{1/3}\delta/6}} \quad < \quad 0.0243;$$

since 0.0243 is smaller than $b = 0.026\cdots$, we have proven (1).

References

[1] Noga Alon, Joel H. Spencer, and Paul Erdős, The Probabilistic Method, John Wiley & Sons, Inc., New York, 1992.

[2] Jim Arney and Edward A. Bender, Random mappings with constraints on coalescence and number of origin, *Pacific. J. Math.* **103** (1982), 269–294.

[3] Louis P. Comtet, Advanced Combinatorics, D. Reidel, 1974.

[4] Peter Kirschenhofer, A Note on Alternating Sums, *The Electronic Journal of Combinatorics*, **3**(2) (1995), #R7.

[5] Donald Ervin Knuth, The Art of Computer Programming, Volume III: Sorting and Searching, Addison-Wesley, Reading, Massachusetts, (1973).

[6] Valentin Kolchin, Random Mappings, Optimization Software, New York (1986).

[7] Douglas E. Lambert and Peter J. Slater, Parallel knockouts in the complete graph, *American Math. Monthly* **105** (1998), 556–558.

[8] Shuguang Li, On Artin's Conjecture for Composite Moduli, University of Georgia, 1998

[9] Lennart Räde, proposer, Problem E3436, The Amer. Math. Monthly **98** (1991), p. 366; solution by O. P. Lossers, *ibid.* **101** (1994), p. 78.