AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **78(3)** (2020), Pages 376–398

Probabilistic aspects of r-Stirling numbers

Gregory J. Morrow

Department of Mathematics
University of Colorado Colorado Springs
Colorado Springs, CO 80918
U.S.A.
gmorrow@uccs.edu

Abstract

Let $k \geq 2$ and $1 \leq \ell \leq k$. Consider an infinite sequence of independent trials such that at each trial one of the cells of $[k]$ is selected at random and a new ball is placed in the cell. Define the random variable T_ℓ as the first trial on which each cell of $[\ell]$ contains at least one ball. Some identities for r-Stirling numbers of the second kind, and some relations with p-Stirling numbers of the first kind, are obtained by developing the distribution of T_ℓ in terms of r-Stirling numbers of the second kind. We introduce a subcollection (indexed by q) of partitions of $[\beta+\sigma]$ into β blocks, with r distinguished blocks. The size of this subcollection is shown to be $S(\beta+\sigma,\beta;r,q) = \sum_{\alpha_0+\cdots+\alpha_{\beta-r}=\sigma}^{\prime q} r^{\alpha_0}(r+1)^{\alpha_1}\cdots\beta^{\alpha_{\beta-r}}$, where $\alpha_i \geq 0$ and $'q$ denotes that exactly q of the α_i satisfy $\alpha_i \geq 1$. We express this sum in terms of r-Stirling numbers of both the first and second kinds.

1 Introduction

We are inspired by the following problem, posed by Engbers and Hammett [8]; cf. [14, (9.26)].

Problem: "Let k and n be positive integers, and let $m = \min\{k,n\}$. Prove that for $1 \leq \ell \leq m$, we have

$$\sum_{x=\ell}^{m} \left\{ {n \atop x} \right\} (x)_\ell \; (k-\ell)_{x-\ell} \; = \; \sum_{i=0}^{\ell} \binom{\ell}{i} (-1)^i (k-i)^n. \tag{1.1}$$

Here, for $a \in \mathbb{R}$ and $b \in \mathbb{Z}^+$, we have $(a)_b := a(a-1)\cdots(a-b+1)$, and $(a)_0 := 1$, and $\left\{ {n \atop x} \right\}$ is the Stirling number of the second kind, i.e. the number of ways to partition the set $[n] := \{1,2,\ldots,n\}$ into x nonempty blocks."

ISSN: 2202-3518

The number of partitions of a finite set S of n elements has many combinatorial interpretations, including the number of equivalence relations on S; these numbers (ϵ_n) are often called the *exponential numbers* or Bell numbers [1, p. 417], [15]. Here a partition of S is a collection of disjoint nonempty subsets of S, or *blocks*, whose union is S. For any $1 \le k \le n$, the number of partitions of S into exactly k blocks is denoted $\left\{ {n \atop k} \right\}$, and is called a Stirling number of the second kind. By definition $\left\{ {n \atop k} \right\}$ $= 0$, if $k \le 0$ or $k > n$, except $\left\{ {0 \atop 0} \right\} = 1$ (for the empty partition consisting of no blocks). For example $\left\{ {4 \atop 2} \right\} = 7$, as can be checked by taking 3 different partitions of $S = \{a, b, c, d\}$ consisting of two blocks each of size 2, and 4 partitions of two blocks such that one block is a singleton set and the other block is the remaining 3 elements of S. In general we have

$$\left\{ {n+1 \atop k} \right\} = k \left\{ {n \atop k} \right\} + \left\{ {n \atop k-1} \right\}.$$

Indeed, for the purpose of establishing this recurrence, we may write $S = [n] = \{1, 2, \ldots, n\}$. On the right-hand side, either $\{n+1\}$ is a singleton block and we have $k - 1$ blocks in a partition of $[n]$, or we have that the element $n + 1$ is thrown into one of the blocks of a partition of $[n]$ into k blocks, and for each such partition there are k ways to do this.

The original definition of Stirling, [12, p. 67], [3, p. 255], determines these numbers as the coefficients of the falling factorial powers $(x)_k = x(x-1)\cdots(x-k+1)$ that recover the ordinary powers: $x^n = \sum_{k=1}^{n} i\left\{ {n \atop k} \right\} (x)_k$. Via the 2-term recurrence, the partition definition and Stirling's definition of $\left\{ {n \atop k} \right\}$ are equivalent. Define the exponential or Touchard polynomials, [2, 16], by $\varphi_n(x) = \sum_{k=0}^{n} \left\{ {n \atop k} \right\} x^k$. Then, as in [2, (3.4)], we have $\left(x\frac{d}{dx}\right)^n e^{ax} = \varphi_n(ax)e^{ax}$, and in this same reference [2] further properties of these polynomials are discussed with applications. See [13] for applications of Stirling numbers of both first and second kinds, where Definition 5.2 covers Stirling numbers of the first kind.

Our main interest will be in a certain generalization of $\left\{ {n \atop k} \right\}$, as follows.

Definition 1.1 Define the r-Stirling number of the second kind $\left\{ {n \atop m} \right\}_r$, $r,\ m,\ n \ge 0$, as the number of partitions of the set $\{1, 2, \ldots, n\}$ into m nonempty disjoint subsets (or blocks) such that each of $1, 2, \ldots, r$ appears in a different block. The ordinary Stirling number of the second kind, namely the case $r = 1$, is denoted without subscript. By convention, $\left\{ {n \atop m} \right\}_0 = \left\{ {n \atop m} \right\}$. Further $\left\{ {n \atop m} \right\}_r = 0$, if $n < r$ or $m < r$, while $\left\{ {r \atop m} \right\}_r = \delta_{m,r}$, for $r \ge 0$, and $\left\{ {n \atop 0} \right\}_r = 0$, for $n > r$.

Suppose now that we have k boxes or *cells*, labeled by $[k] = \{1, 2, \ldots, k\}$. We consider a sequence of independent *trials* wherein at each trial a new ball is placed in one of the cells uniformly at random. Let $1 \le \ell \le k$ be given, and define $A_\ell^{(n)}$ as the event that each box labeled with $1, 2, \ldots, \ell$ contains at least one ball after n trials. In Section 2 we show how to solve the *Problem* by writing out two different approaches to the calculation of $P(A_\ell^{(n)})$. One approach is by the inclusion-exclusion principle to find the right-hand side of (1.1). For another approach, define the random variable L_n as the number of cells from $[k]$ that contain at least one ball after n trials. One calculates the probability of the event $\{L_n = x\} \cap A_\ell^{(n)}$ to find the left side of (1.1) and thus the solution, since by disjoint events $P(A_\ell^{(n)}) = \sum_{x=\ell}^{k} P(\{L_n = x\} \cap A_\ell^{(n)})$. Still our goal is to find another representation of $P(A_\ell^{(n)})$ that leads to r-Stirling numbers, for $r := k - \ell$. To arrive naturally at r-Stirling numbers in the present context, define the random waiting time T_ℓ as the number of trials at which, for the first time, each cell of $[\ell]$ contains at least one ball. We find the probability function of T_ℓ in terms of r-Stirling numbers by Lemma 2.3. But we also have

$$\sum_{\tau=\ell}^{n} P(T_\ell = \tau) = P(A_\ell^{(n)}), \tag{1.2}$$

and therefore we have an r-Stirling numbers approach to the representation of $P(A_\ell^{(n)})$. We introduce a sequential coding (many to one mapping of members of $[k] \times \cdots \times [k]$ to words) in Lemma 2.4 to write yet another representation $P(A_\ell^{(n)})$, again in terms of r-Stirling numbers, and so obtain Theorem 2.5. The paper then uses the waiting time construct as a jumping off point. For instance we discuss in Section 3 the distribution of the random variable $T_{\ell,t}$ defined as the number of trials at which, for the first time, exactly $\ell - t$ cells of $[\ell]$ contain at least one ball. This leads to the identity of Theorem 3.2.

The main applications of this paper are some new identities that arise from counting events in different ways. These identities are mainly for the r-Stirling numbers of the second kind, but we also find in Propositions 5.4 and 5.6 identities involving r-Stirling numbers of the second kind and p-Stirling numbers of the first kind, the latter introduced by Definition 5.2. Our aim is first to find relations between the r-Stirling numbers whose proofs are ultimately based on bijective arguments. In Section 5 we will also apply generating function arguments. While our goal is not to develop probability distributions involving the Stirling numbers, the reader may wish to consult [10].

In Section 4 we find a combinatorial proof of the probability function of T_ℓ. In Sections 5 and 5.1 we elaborate on the sequential coding device for decomposing events, and show its applications via the generating function method. In Section 5.2 we introduce by Definition 5.9 a subcollection, indexed by nonnegative integers q, of the set of partitions of $[\beta + \sigma]$ into β blocks with r distinguished blocks. We give a combinatorial proof of Corollary 5.10 for the size of this subcollection in terms of the sums $S(\beta + \sigma, \beta; r, q)$ of Definition 5.7. These sums are calculated by a formula involving r-Stirling numbers of both first and second kinds in Corolllary 5.8.

2 Probabilistic Approach to the Problem

We take a probabilistic point of view for (1.1) that leads naturally to the context of r-Stirling numbers of the second kind. Let there be $k \geq 2$ cells that we fill on successive independent trials with n balls, placing each ball in a randomly chosen cell. It is convenient to name the cells by the *digits* $1, 2, \ldots, k$. We say a cell is *filled* if it contains at least one ball. For each n, denote the collection of sequences of n digits $\Omega_n = \{(\omega_1, \omega_2, \ldots, \omega_n) : \ \omega_1, \ldots, \omega_n \in [k]\}$. The set Ω_n is an elementary probability space in which each point $\underline{\omega} := (\omega_1, \omega_2, \ldots, \omega_n)$ has probability k^{-n}. Let $1 \leq \ell \leq k$ determine a fixed subset $[\ell] \subset [k]$ of cells identified by their digits. Assume $n \geq \ell$. Define the event $A_\ell = A_\ell^{(n)} \subset \Omega_n$ by "each cell of $[\ell]$ is filled after n trials". In terms of sample points, we have that $(\omega_1, \ldots, \omega_n) \in A_\ell$ if and only if for each digit $\omega \in [\ell]$, there exists a trial $i \leq n$ such that $\omega_i = \omega$. In words, A_ℓ consists of n-long sequences of digits such that each digit of $[\ell]$ appears somewhere in the sequence. The identity (1.1) counts the number of sample points of A_ℓ in two different ways. To begin, the right side of the identity represents a counting via the inclusion-exclusion law, applied as $P(A_\ell) = 1 - P(A'_\ell)$, for $A_\ell = \bigcap_{j=1}^{\ell} F_j$, with $F_j := \{\text{ cell } j \text{ contains at least one ball }\}$. Here A' denotes the complement of A, and it is easy to calculate $k^n P(F'_{j_1} \cap F'_{j_2} \cap \cdots \cap F'_{j_i}) = (k-i)^n$, by independence.

To understand the left side of the identity, one may count according to how many different digits appear in the sample points of A_ℓ. That is, we decompose the event according to the values x of the random variable $L = L_n(\underline{\omega})$, defined as the number of cells of $[k]$ that have any balls in them after n placements. If $x \geq \ell$, the number of sample points in the event $\{\underline{\omega} : L(\underline{\omega}) = x\} \cap A_\ell$ is $\left\{ {n \atop x} \right\} (x)_\ell \ (k-\ell)_{x-\ell}$; this may be proved directly by a combinatorial argument. A sequential coding for this enumeration is mentioned at the start of Section 5, just before Definition 5.1. In the sequel we suppress the dependence on $\underline{\omega}$ and instead write for example $\{L_n = x\}$ for the event that exactly x cells are filled after n trials. Thus, because by disjoint events, $P(A_\ell^{(n)}) = \sum_{x=\ell}^{k} P(\{L = x\} \cap A_\ell)$, (1.1) has in principle been solved. Yet our interest lies in further developing this probabilistic approach due to yet other representations of $P(A_\ell^{(n)})$ that we introduce next.

2.1 The random waiting time T_ℓ

Recall the Definition 1.1 of r-Stirling numbers of the second kind. By [4, Theorem 2] these numbers satisfy: $\left\{ {n+1 \atop m} \right\}_r = m \left\{ {n \atop m} \right\}_r + \left\{ {n \atop m-1} \right\}_r$. Extension to higher r follows by Broder [4, Theorem 4]: $\left\{ {n \atop m} \right\}_{r+1} = \left\{ {n \atop m} \right\}_r - r \left\{ {n-1 \atop m} \right\}_r$. Triangular tables of r-Stirling numbers are displayed in [4, Table 1] for $r = 1, 2, 3$. We shall often apply the following arithmetical identity, [4, Theorem 8]:

Lemma 2.1 *Let $0 \leq r \leq k$ and $m \geq 0$. Then, with the following sum equal to* 1 *if*

$m = 0$,

$$\left\{ \begin{matrix} k+m \\ k \end{matrix} \right\}_r = \sum_{r \le i_1 \le i_2 \le \cdots \le i_m \le k} i_1 i_2 \cdots i_m. \tag{2.1}$$

For example, if $r = 2$, $k = 4$, and $m = 2$, $\left\{ \begin{smallmatrix} 6 \\ 4 \end{smallmatrix} \right\}_2 = 2^2 + 2\cdot 3 + 2\cdot 4 + 3^2 + 3\cdot 4 + 4^2 = 55$.

Definition 2.2 Define T_ℓ as "the first trial n such that each cell of $[\ell]$ is filled".

We call T_ℓ a stopping time since each event $\{T_\ell = n\}$ may be regarded as an event in Ω_n. This event is the collection of points $\underline{\omega} = (\omega_1, \ldots, \omega_n)$ such that $\omega_n \in [\ell]$, yet $\omega_n \ne \omega_i$ for all $1 \le i < n$, and $\underline{\omega} \in A_\ell^{(n)}$; that is every element of $[\ell]$ appears at least once in the sequence of digits $\underline{\omega}$. We note that, by disjoint events, (1.2) holds. This provides us motivation to find a simple expression for $P(T_\ell = \tau)$. Technically T_ℓ is defined on the infinite product probability space $[k] \times [k] \times \cdots$, though decomposition arguments come down to counting subsets of Ω_τ for some finite τ.

To calculate the distribution of T_ℓ, we introduce independent geometric random variables $G_1, \ldots, G_\ell$ with respective *success* parameters $\ell/k, (\ell-1)/k, \ldots, 1/k$. Here a geometric random variable G with success parameter p is the number of independent tosses of a coin with probability p for heads that are required to first obtain a heads; thus $P(G = t) = (1-p)^{t-1}p$, $t = 1, 2, \ldots$. It is easy to see that, in distribution, $T_\ell = G_1 + \cdots + G_\ell$. Indeed, in ascertaining T_ℓ, first one of the levels of $[\ell]$ must be filled, and this event has probability ℓ/k. Next, if $\ell \ge 2$, one of the remaining levels, of which there are now $\ell - 1$, must be filled, and the probability of success is now $(\ell-1)/k$, etc. By this representation, the r-Stirling numbers of the second kind come into play as follows.

Lemma 2.3 *Put $r = k - \ell$. Then for any $\tau \ge \ell$,*

$$P(T_\ell = \tau) = \frac{\ell!}{k^\tau} \left\{ \begin{matrix} r+\tau-1 \\ r+\ell-1 \end{matrix} \right\}_r .$$

Proof. Write $T_\ell = G_1 + \cdots + G_\ell$, where the $G_1, \ldots, G_\ell$ are independent geometric random variables such that G_j has success parameter $p_j = \frac{\ell-j+1}{k}$. Therefore, by direct computation, we have that $P(T_\ell = \tau)$ is given by:

$$\prod_{j=1}^{\ell} p_j \sum_{\alpha_1+\cdots+\alpha_\ell=\tau-\ell} \prod_{j=1}^{\ell} (1-p_j)^{\alpha_j} = \frac{\ell!}{k^\tau} \sum_{\alpha_1+\cdots+\alpha_\ell=\tau-\ell} (k-\ell)^{\alpha_1}(k-(\ell-1))^{\alpha_2} \cdots (k-1)^{\alpha_\ell}, \tag{2.2}$$

where the α_j are nonnegative integers in the sum. Thus, taking into account the fact that $\alpha_j = 0$ is possible for some j, the last sum is simply rewritten

$$\sum_{k-\ell \le i_1 \le i_2 \le \cdots \le i_{\tau-\ell} \le k-1} i_1 i_2 \cdots i_{\tau-\ell} = \left\{ \begin{matrix} k-1+\tau-\ell \\ k-1 \end{matrix} \right\}_{k-\ell} = \left\{ \begin{matrix} r+\tau-1 \\ r+\ell-1 \end{matrix} \right\}_r ,$$

by Lemma 2.1 and the definition of r. The proof is complete upon substituting this r-Stirling number on the right side of (2.2). □

For the remainder of this section, we first develop a sequential coding scheme for calculating $P(A_\ell^{(n)})$ in the proof of Lemma 2.4 that is different from the decomposition (1.2). This leads to the identity for r-Stirling numbers of Theorem 2.5. Afterwards, by Remark 2.7 we mention an alternative proof of this lemma, which is stated as follows.

Lemma 2.4 *Suppose $n \geq \ell$, and denote $r = k - \ell$. Then*

$$k^n P\left(A_\ell^{(n)}\right) = \ell! \left\{ {r+n \atop r+\ell} \right\}_r .$$

Proof. In order to fill the cells of $[\ell]$ we must place a ball in a *new* cell from $[\ell]$, that is a cell whose digit hasn't appeared before in sequence, each of precisely ℓ times. Denote by the letter N the occurrence of a new digit for the particular sequence of cells, whenever the new cell is from $[\ell]$. Here J denotes either an element of $[k]\setminus[\ell]$, or an old N. For example, if $\ell = 2$, $k \geq 4$, and $n = 6$, the sequence $(3, 1, 1, 4, 2, 3)$ falls under the event $JNJJNJ$, while $(4, 2, 3, 2, 1, 1)$ also falls under this event. Hence $A_\ell^{(n)}$ is written as a disjoint union of events consisting of words of length n from the alphabet $\{N, J\}$ with precisely ℓ many N's in the word. The total number of J's is therefore $n - \ell$. Call a maximum length sequence of consecutive J's of at least one J in a finite word on the alphabet $\{N, J\}$ a *string* of J's. The *length* of the string is simply the number of consecutive J's. By convention we allow an *empty string* of length 0 to mean there is not at least one J at a particular position in the word. For each $1 \leq i \leq \ell$, and for the i^{th} N, we account for a string of J's of length $\alpha_i \geq 0$ after the N. We also allow a string of length $\alpha_0 \geq 0$ J's before the initial N. We therefore account for (possibly empty) strings of J's in the $\ell + 1$ spaces around the skeletal word $N\ N\ \cdots\ N$, that consists of ℓ many N's. The total number of sample points in the union of all the elementary events on the alphabet $\{N, J\}$ is thus

$$\ell! \sum_{\alpha_0+\alpha_1+\alpha_2+\cdots+\alpha_\ell=n-\ell} r^{\alpha_0}(r+1)^{\alpha_1}\cdots k^{\alpha_\ell},$$

where $\ell!$ is the number of ways of choosing the digits for the N's in any given word, and again $r := k-\ell$. By Lemma 2.1, this counting yields the statement of the lemma by the definition of r. Indeed, for each $0 \leq i \leq \ell$, $\alpha_i \geq 0$ is the number of times the integer $r + i$ is repeated in the product of $m = n - \ell$ integers of the arithmetic formula (2.1) for $\left\{ {r+\ell+m \atop r+\ell} \right\}_r = \left\{ {r+n \atop r+\ell} \right\}_r$. □

By combining the different ways of calculating $P(A_\ell)$ according to r-Stirling numbers, we obtain the following extension to $r \geq 2$ of a well known identity, [9, (6.20)] or [4, (35)], for the ordinary Stirling numbers (the case $r = 1$) of the second kind.

Theorem 2.5 *Let $n \geq \ell$ and denote $r = k - \ell$. Then*

$$\sum_{\tau=\ell}^{n} (r+\ell)^{n-\tau} \left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r = \left\{ {r+n \atop r+\ell} \right\}_r .$$

Proof. By (1.2) and Lemma 2.3 we have that $k^n P(A_\ell^{(n)}) = \sum_{\tau=\ell}^{n} k^{n-\tau} \ell! \left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r$. On the other hand we have the statement of Lemma 2.4. Hence the theorem follows by canceling the factors of $\ell!$ on the two sides and by rewriting $k^{n-\tau} = (r+\ell)^{n-\tau}$. □

There is yet another way to calculate $P(A_\ell)$, but in terms of ordinary Stirling numbers that we mention now for completeness. Define the random variable X_ℓ as the total number of balls that fall into the cells of $[\ell]$ after n trials. By decomposing the event A_ℓ according to the values of X_ℓ, we obtain the following.

Lemma 2.6 *For all* $1 \le \ell \le k$, *and* $\ell \le n$, *we have (with* $0^0 = 1$*),*

$$k^n P\left(A_\ell^{(n)}\right) = \ell! \sum_{x=\ell}^{n} \binom{n}{x} \left\{ {x \atop \ell} \right\} (k-\ell)^{n-x}.$$

Proof. The event $\{X_\ell = x\} \cap A_\ell$ is written

"after n trials the cells of $[\ell]$ contain exactly x balls that also fill each cell of $[\ell]$".

We count this event by choosing x trials, and partitioning these trials into ℓ nonempty blocks. Fix for the moment a particular choice of x trials in time. Given any particular ordering of the blocks from the $\ell!$ orderings of partitions into ℓ blocks of these trials, place the balls from the i^{th} block into the i^{th} cell, $i = 1, \dots, \ell$. Also for each choice of x trials or balls in time, the remaining $n-x$ balls must each fall into the cells of $[k] \setminus [\ell]$. Hence for each choice of x trials, there are $\ell! \left\{ {x \atop \ell} \right\} (k-\ell)^{n-x}$ points $\underline{\omega}$ of the event $\{X_\ell = x\} \cap A_\ell$ in which the coordinate digits of $\underline{\omega}$ fall into and fill $[\ell]$ precisely on the times of these x trials. Since there are $\binom{n}{x}$ many choices of the x trials, we have $k^n P(\{X_\ell = x\} \cap A_\ell) = \ell! \left\{ {x \atop \ell} \right\} (k-\ell)^{n-x} \binom{n}{x}$. Since $P(A_\ell) = \sum_{x=\ell}^{n} P(\{X_\ell = x\} \cap A_\ell)$, the proof is complete. □

By [4, (32)], we have the following relationship between ordinary Stirling numbers and r-Stirling numbers of the second kind.

$$\sum_{x=\ell}^{n} \binom{n}{x} \left\{ {x \atop \ell} \right\} r^{n-x} = \left\{ {n+r \atop r+\ell} \right\}_r. \tag{2.3}$$

Remark 2.7 By (2.3) and Lemma 2.6, we obtain an alternative proof of Lemma 2.4.

The r-Stirling number of the second kind was introduced by Carlitz [6, (3.1)–(3.6)] in the form $R(n,k,\lambda) = \left\{ {\lambda+n \atop \lambda+k} \right\}_\lambda$; compare [4, Corollary 10] and [6, (3.10)]. An alternative proof of Carlitz's [6, (3.8)] follows. This result was rediscovered in disguised form as [11, Theorem 4.5].

Proposition 2.8 *For $\ell \le n$ and $r \ge 0$,*

$$\left\{ {r+n \atop r+\ell} \right\}_r = \frac{1}{\ell!} \sum_{i=0}^{\ell} (-1)^i \binom{\ell}{i} (r+\ell-i)^n.$$

Proof. By the inclusion-exclusion calculation, $k^n P(A_\ell)$ is a representation of the right side of (1.1) with $k = r+\ell$. So both sides of the identity are equal to $\frac{k^n}{\ell!} P(A_\ell)$ by Lemma 2.6 together with (2.3) for the left side. The identity also follows from (1.1) and Lemma 2.4. □

3 The random time $T_{\ell,t}$

We generalize our approach to the event A_ℓ as follows. For each $0 \le t \le \ell-1$, define the event $A_{\ell,t} = A_{\ell,t}^{(n)}$ by the condition that "exactly $\ell - t$ cells of $[\ell]$ are filled by trial n". We have $A_\ell = A_{\ell,0}$. Also define $B_{\ell,t} = B_{\ell,t}^{(n)}$ as the event that "at least $\ell - t$ cells of $[\ell]$ are filled by trial n". By the same proof as given by Lemma 2.4, only with $(\ell)_{\ell-t}$ in place of $\ell!$ for the number of choices for assigning distinct digits from $[\ell]$ to a skeleton word consisting of $(\ell - t)$ N's, we have

$$k^n P\left(A_{\ell,t}^{(n)}\right) = (\ell)_{\ell-t} \left\{ {r+n \atop r+\ell-t} \right\}_r. \tag{3.1}$$

Definition 3.1 Let $0 \le t \le \ell - 1$. Define $T_{\ell,t}$ as "the first trial on which $\ell - t$ cells of $[\ell]$ are filled".

By the proof of Lemma 2.3 with just the sum of the first $\ell - t$ geometric random variables G_j representing $T_{\ell,t}$, or by the proof of Lemma 2.4, in which the words on N's and J's are constructed from exactly $(\ell-t)$ many N's and terminate in an N, so there are only $\ell - t$ possible spaces preceding N's where a string of J's may appear, we have:

$$k^x P(T_{\ell,t} = x) = (\ell)_{\ell-t} \left\{ {r+x-1 \atop r+\ell-t-1} \right\}_r. \tag{3.2}$$

Theorem 3.2 *Let $0 \le t \le \ell - 1$, assume $n \ge \ell$, and put $r = k - \ell$. Then*

$$(\ell)_{\ell-t} \sum_{x=\ell-t}^{n} k^{n-x} \left\{ {r+x-1 \atop r+\ell-t-1} \right\}_r = \sum_{m=\ell-t}^{\ell} (\ell)_m \left\{ {r+n \atop r+m} \right\}_r.$$

Proof. Since the event $B_{\ell,t}^{(n)}$ may be written:

$$B_{\ell,t}^{(n)} = \text{“}\ell - t \text{ cells of } [\ell] \text{ are filled for the first time at a time } \tau \le n\text{”},$$

we have $B_{\ell,t}^{(n)} = \{T_{\ell,t} \le n\}$. Since the events $\{T_{\ell,t} = x\}$ are disjoint in x, we therefore have

$$P(B_{\ell,t}) = \sum_{x=\ell-t}^{n} P(T_{\ell,t} = x). \tag{3.3}$$

Furthermore since the events $A^{(n)}_{\ell,t}$ are disjoint in t, we have

$$P(B_{\ell,t}) = \sum_{j=0}^{t} P(A_{\ell,j}). \tag{3.4}$$

Therefore by equating the two expressions for $P(B_{\ell,t})$ coming from (3.3)–(3.4), and substituting (3.1)–(3.2) to write r-Stirling expressions for enumerations of the events $A_{\ell,t}$ and $\{T_{\ell,t} = x\}$, we obtain the desired statement. Here the summation index j of (3.4) is replaced by $m = \ell - j$ for the right side of the statement of the theorem. □

For illustration of Theorem 3.2, let $\ell = 4$, $k = 6$, $t = 2$, and $n = 5$, so that $r = 2$ and $\ell - t = 2$. We have

$$4 \cdot 3 \left(6^3 + 6^2 \left\{ {4 \atop 3} \right\}_2 + 6 \left\{ {5 \atop 3} \right\}_2 + \left\{ {6 \atop 3} \right\}_2\right) = 4 \cdot 3\,(216 + 36 \cdot 5 + 6 \cdot 19 + 65),$$

while on the other hand

$$4 \cdot 3 \left\{ {7 \atop 4} \right\}_2 + 4 \cdot 3 \cdot 2 \left\{ {7 \atop 5} \right\}_2 + 4! \left\{ {7 \atop 6} \right\}_2 = 12 \cdot 285 + 24 \cdot 125 + 24 \cdot 20;$$

both expressions evaluate to $12 \cdot 575$.

4 Combinatorial Proof of Lemma 2.3

In the sequel, we continue to write $r := k - \ell \geq 0$. Recall by Lemma 2.3 that $k^n P(T_\ell = \tau) = \ell! \left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r$. We give a direct combinatorial proof of this enumeration as follows. Similar to the proof of Lemma 2.4, we have a many to one mapping (sequential coding) of any $\underline{\omega} = (\omega_1, \ldots, \omega_\tau) \in \{T_\ell = \tau\}$ to a τ-letter word on the alphabet $\{N, J\}$ by the rule that ω_i is coded by N if the digit $\omega_i \in [\ell]$ has the property that this digit has not appeared as ω_j for all $j < i$. Else, the digit ω_i is coded by J. Note however that in this construction for the event $\{T_\ell = \tau\}$, automatically ω_τ is coded by N.

Definition 4.1 Let λ be a permutation of of $[\ell]$. Define the subevent $E = E_{\lambda,\tau}$ of $\{T_\ell = \tau\}$ by the property that for any $\underline{\omega} \in E$, we have that the sequence of new digits, $\omega_{i(1)}, \omega_{i(2)}, \ldots, \omega_{i(\ell)}$, each coded by N, is equal to $\lambda(1), \lambda(2) \ldots \lambda(\ell)$; in particular $\omega_{i(\ell)} = \omega_\tau = \lambda(\ell)$.

For illustration in the following argument let $r = 2$ and let $\underline{\omega} \in E$. Define an *extended sequence* of digits as $\underline{\omega}^+ := (a, b; \omega_1, \omega_2, \ldots, \omega_\tau)$, where a and b are distinct abstract digits and not equal to any digit from $[k]$. For illustration again, say $\tau = 6$ and $\ell = 3$. Define the *coordinate positions* of each of the digits of $\underline{\omega}^+$ respectively as *1, 2; 3, . . . , 8*. Here we have $r + \tau = 8$ coordinate positions. Note that digits are denoted $1, 2, 3, \ldots, 5$ (since $k = \ell + r = 5$), while coordinate positions are denoted

1, *2*, *3*, ... For illustration we take λ to be the identity permutation, $\lambda = \text{id}$, on [3]. Let three example points $\underline{\omega}$, or 6-tuples, for our event E be:

$$(1,\ 2,\ 4,\ 4,\ 2,\ 3)$$
$$(1,\ 2,\ 2,\ 2,\ 2,\ 3)$$
$$(5,\ 1,\ 1,\ 1,\ 2,\ 3)$$

In each example case, the digit 1 appears before digit 2, and digit 3 only appears at the last place. Now extend these 6-tuples to 8-tuples as follows:

$$(a,\ b;\ 1,\ 2,\ 4,\ 4,\ 2,\ 3)$$
$$(a,\ b;\ 1,\ 2,\ 2,\ 2,\ 2,\ 3)$$
$$(a,\ b;\ 5,\ 1,\ 1,\ 1,\ 2,\ 3)$$

We now partition the coordinate positions $\{\mathit{1}, \mathit{2}, \ldots, \mathit{7}\}$, where we have omitted the last coordinate position *8*, into $r + \ell - 1 = 4$ nonempty blocks according to the following rule: the block B_a contains *1* and in addition any coordinate positions (if they so exist) of the digit $\ell + 1 = 4$; the block B_b contains *2* and in addition any coordinate positions of the digit $\ell + 2 = 5$. The remaining $\ell - 1 = 2$ blocks B_1, B_2 are defined by: B_1 consists of the coordinate positions of the digit 1, and B_2 consists of the coordinate positions of the digit 2. Since digits 1 and 2 must appear, all blocks are nonempty and they exhaust all coordinate positions $\{\mathit{1}, \mathit{2}, \ldots, \mathit{7}\}$, since digit 3 only appears in coordinate position *8* of the extended sequence. Therefore we have the following correspondences in our example:

$$\{\ \{\mathit{1},\ \mathit{5},\ \mathit{6}\ \},\ \{\mathit{2}\},\ \{\mathit{3}\},\ \{\mathit{4},\ \mathit{7}\}\ \}\ \leftrightarrow\ (a,\ b;\ 1,\ 2,\ 4,\ 4,\ 2,\ 3);$$

$$\{\ \{\mathit{1}\},\ \{\mathit{2}\},\ \{\mathit{3}\},\ \{\mathit{4},\ \mathit{5},\ \mathit{6},\ \mathit{7}\}\ \}\ \leftrightarrow\ (a,\ b;\ 1,\ 2,\ 2,\ 2,\ 2,\ 3);$$

$$\{\ \{\mathit{1}\},\ \{\mathit{2},\ \mathit{3}\},\ \{\mathit{4},\ \mathit{5},\ \mathit{6}\},\ \{\mathit{7}\}\ \}\ \leftrightarrow\ (a,\ b;\ 5,\ 1,\ 1,\ 1,\ 2,\ 3).$$

For illustration let $\mathbf{R}$ denote all partitions of the coordinate positions $\{\mathit{1}, \mathit{2}, \ldots, \mathit{7}\}$, into $r + \ell - 1 = 4$ blocks such that, for $r = 2$, coordinate positions *1* and *2* belong to different blocks. We mean that there are $r + \tau - 1$ coordinate positions to be partitioned, and the first r coordinate positions each belong to a different block, and we have $\beta := r + \ell - 1$ blocks in general. We claim that for the given rule of constructing blocks, we have a bijection between $\mathbf{R}$ and all extended $(r + \tau)$-tuples constructed from E, which is trivially bijective to E. Indeed, obviously we have a map φ from E to $\mathbf{R}$ defined by the construction of blocks. Suppose we would have two image partitions the same from two members $\underline{\omega}$, $\underline{\omega}' \in E$ under φ. Then the coordinate positions of all digits appearing in $\underline{\omega}$ except the digit $\text{id}(\ell) = \ell$ would have been determined precisely by the common partition. For example, if there is a digit $\ell + 1 = 4$ in $\underline{\omega}$, then the coordinate positions of all occurrences of this digit are determined by the partition block B_a. Since $\varphi(\underline{\omega}) = \varphi(\underline{\omega}')$, B_a also determines the same coordinate positions of digit $\ell + 1 = 4$ for $\underline{\omega}'$, by definition of B_a. Clearly by going through all blocks of the common partition, and since there is only one place for digit $\text{id}(\ell)$ to go, we have $\underline{\omega} = \underline{\omega}'$. Therefore the mapping φ is one to one.

To see that it is onto, simply let $\rho \in \mathbf{R}$ be a partition of the type described. Let $\mu(B) = \min B$, that is the minimum of coordinate positions in B, for any block B of ρ. By definition of $\mathbf{R}$, for $r = 2$ we must have a block B_a whose minimum coordinate position is *1* and another block B_b whose minimum coordinate position is *2*. If block B_a has any coordinate positions other than *1* in it, then place digit $\ell + 1 = 4$ in each of these other coordinate positions of an extended $(r + \tau)$-tuple. Similarly, if B_b has any coordinate positions other than *2* in it, then place digit $\ell + 2 = 5$ in each of these other coordinate positions of the same extended $(r + \tau)$-tuple; there can be no overlap of coordinate positions by disjoint blocks. Now we have $\ell - 1$ blocks remaining. Order these blocks B by the order of their minima, $\mu(B)$, from least to largest. Accordingly we define blocks $B_1, \dots, B_{\ell-1}$ as the remaining blocks of ρ. Place digit 1 in the coordinate positions of B_1, and in general digit i in the coordinate positions of B_i, $i = 1, \dots, \ell - 1$. Finally place digit ℓ in the last place of the extended $(r+\tau)$-tuple. By construction, digit 1 appears for the first time at coordinate position $\mu(B_1)$, and in general digit i appears for the first time at coordinate position $\mu(B_i)$, $i = 1, \dots, \ell - 1$, and $\mu(B_1) < \mu(B_2) < \cdots < \mu(B_{\ell-1}) < r + \tau$. Therefore the extended $(r + \tau)$-tuple we have constructed comes from extending an element of E. Thus the claim is established: we have a bijection between E and $\mathbf{R}$.

In conclusion, by symmetry we have that for any permutation λ of $[\ell]$, the cardinality of $E_{\lambda,\tau}$ is the same as $E_{\mathrm{id},\tau}$. We showed that the cardinality of E equals the cardinality of $\mathbf{R}$, which by Definition 1.1 is precisely $\left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r$. Therefore by disjoint and exhaustive events $E_{\lambda,\tau}$, as λ varies over all $\ell!$ permutations of $[\ell]$, we have that indeed the cardinality of $\{T_\ell = \tau\}$ is $\ell! \left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r$.

Remark 4.2 By combining the above direct combinatorial proof of Lemma 2.3 with our original proof of this lemma, we obtain a proof of [4, Theorem 8], that is Lemma 2.1.

4.1 Application of the Combinatorial Proof

Here and in the sequel, assume $r \geq 1$ and let π denote an integer such that $0 \leq \pi \leq r$.

Definition 4.3 Define the statistic $W_{\pi,\ell}$ as "the number of cells that have been filled among all cells of $[k - \pi]$ by time T_ℓ".

The random variable $W_{0,\ell}$ is the number among all cells $[k]$ that are filled at time T_ℓ. The random variable $W_{r,\ell} \equiv \ell$ is trivial; we consider the case $\pi = r$ only in Section 5.1. So let now $0 \leq \pi < r$. By definition $\ell \leq W_{\pi,\ell} \leq k - \pi$. Our goal in this section is to establish the joint distribution of $W_{\pi,\ell}$ and T_ℓ in Proposition 4.5. This has an immediate application in Corollary 4.6. We derive this joint distribution in the present section by taking advantage of the bijection of Section 4. In Section 5 we will observe another proof based on an extension of the sequential coding device.

Let $D \subset [k - \pi] \setminus [\ell]$ be a specific subset (allowing the empty set) of size d, so that $0 \leq d \leq r - \pi$. We define a certain event $B^{(\tau)}_{\ell,D} \subset \{T_\ell = \tau\}$ as follows. For each subset $C \subset [\ell]$ of size $|C| = \ell - 1$, denote by A_C the subevent of $\{T_\ell = \tau\}$ defined by $A_C :=$ "the cells of $[k-\pi]$ filled at time $\tau - 1$ equals $D \cup C$, the cell $[\ell] \setminus C$ is filled at time τ". The events A_C, as C varies, are obviously disjoint. Define

$$B_{\ell,D}^{(\tau)} := \bigcup_{C \subset [\ell],\ |C| = \ell - 1} A_C. \tag{4.1}$$

This union is clearly the subset of $\{T_\ell = \tau\}$ such that at time $\tau - 1$, D has been filled, but no cells of $[k-\pi] \setminus [\ell]$ that lie outside of D have been filled, while cells of $[k] \setminus [k-\pi]$ may have been filled.

Lemma 4.4 *Let $\tau \geq \ell$, and let $D \subset [k-\pi] \setminus [\ell]$ be a fixed set of cardinality d. Then*

$$k^\tau P\left(B_{\ell,D}^{(\tau)}\right) = \ell(d + \ell - 1)! \left\{ \begin{matrix} \pi + \tau - 1 \\ \pi + d + \ell - 1 \end{matrix} \right\}_\pi .$$

Proof. The proof follows from the combinatorial proof of Lemma 2.3 given in Section 4. Indeed, fix a subset $C \subset [\ell]$ of size $|C| = \ell - 1$. We calculate the cardinality of A_C by replacing $[\ell]$ by $[\ell] \cup D$ in that proof. So take $\widetilde{\ell} := \ell + d$ to play the role of ℓ. By definition we do not allow any digits of $[k-\pi] \setminus ([\ell] \cup D)$ to enter into the coordinates of a point $\underline{\omega} \in A_C$, so we really just ignore this fixed set of digits. Also we take the role of r in that proof to be played by π, since the size of the set of digits allowed to be included in τ-tuples of A_C, besides the designated set of digits $[\ell] \cup D$ that must be represented, and besides the digits removed from consideration, is simply $k - (k - \pi) = \pi$. Since C is given, there is only 1 way to fill the last digit of a τ-tuple in A_C, while there are $(d + \ell - 1)!$ ways to permute the elements of $D \cup C$. Therefore by the bijection of Section 4, we have that $|A_C|/(d + \ell - 1)! = \left\{ \begin{matrix} \pi + \tau - 1 \\ \pi + \widetilde{\ell} - 1 \end{matrix} \right\}_\pi$, where $|A|$ denotes the cardinality of a set A. Here we have used that $|A_C|/(d + \ell - 1)! = |E_{\lambda,\tau}|$ for a specific permutation λ on $[\ell] \cup D$, since the order of first appearance of the $(\ell + d - 1)$ digits in $C \cup D$ is unspecified for A_C. Now there are exactly ℓ subsets C of the type that define $B_{\ell,D}^{(\tau)}$ in (4.1). Therefore by disjoint events and equal cardinalities of the various A_C, the lemma is proved upon substitution of $\widetilde{\ell} = \ell + d$ in the formula for $|A_C|$. □

Proposition 4.5 *For any $0 \leq \pi \leq r$, we have the following joint distribution of $W_{\pi,\ell}$ and T_ℓ:*

$$k^\tau P(W_{\pi,\ell} = w,\ T_\ell = \tau) = \ell(w-1)! \binom{r - \pi}{w - \ell} \left\{ \begin{matrix} \pi + \tau - 1 \\ \pi + w - 1 \end{matrix} \right\}_\pi , \quad \ell \leq w \leq k - \pi,\ \tau \geq w.$$

Proof. Let $D \subseteq [k-\pi] \setminus [\ell]$ denote a set of $d := w - \ell$ elements. Since the joint event in question means precisely that exactly w cells of $[k-\pi]$ are filled at time τ, and $[\ell]$ is not yet filled by time $\tau - 1$, but is completely filled at time τ, then, by disjoint events,

$$P(W_{\pi,\ell} = w,\ T_\ell = \tau) = \sum_D P\left(B_{\ell,D}^{(\tau)}\right),$$

where the event $B_{\ell,D}^{(\tau)}$ is defined by (4.1) and the sum runs over all d-element subsets D of $[k-\pi]\setminus[\ell]$. Since there are $\binom{r-\pi}{d}$ such subsets D, the result follows by Lemma 4.4 and the definition of d. □

Corollary 4.6 *For any $\tau \geq \ell$, we have*

$$\sum_{w=\ell}^{k-\pi}(w-1)!\binom{r-\pi}{w-\ell}\left\{{\pi+\tau-1 \atop \pi+w-1}\right\}_{\pi} = (\ell-1)!\left\{{r+\tau-1 \atop r+\ell-1}\right\}_{r}.$$

Proof. The proof follows by writing

$$\sum_{w=\ell}^{k} P(W_{\pi,\ell}=w,\ T_\ell=\tau) = P(T_\ell=\tau).$$

Now apply Proposition 4.5 to the left side and Lemma 2.3 to the right side, and cancel a factor of each of ℓ and k^τ on both sides. □

For illustration of Corollary 4.6, take $\ell=3$, $k=7$, $\pi=2$, and $\tau=5$. Thus $r=4$ and we find:

$$2!\binom{2}{0}\left\{{6 \atop 4}\right\}_2 + 3!\binom{2}{1}\left\{{6 \atop 5}\right\} + 4!\binom{2}{2}\left\{{6 \atop 6}\right\}_2 = 2\cdot 55 + 12\cdot 14 + 24\cdot 1 = 2\cdot 151 = 2!\left\{{8 \atop 6}\right\}_4.$$

As a reminder that $\pi=0$ indicates an ordinary Stirling number on the left side of Corollary 4.6, let $\ell=3$, $k=5$, $\pi=0$, and $\tau=7$:

$$2!\binom{2}{0}\left\{{6 \atop 2}\right\} + 3!\binom{2}{1}\left\{{6 \atop 3}\right\} + 4!\binom{2}{2}\left\{{6 \atop 4}\right\} = 2\cdot 31 + 12\cdot 90 + 24\cdot 65 = 2\cdot 1351.$$

It is easily checked that this is the same as: $2!\left\{{8 \atop 4}\right\}_2$.

5 Joint Probability Generating Function

We vary the sequential approach to counting A_ℓ in the proof of Lemma 2.4 to uncover a certain representation of the joint probability generating function of T_ℓ and $W_{\pi,\ell}$ defined by Definitions 2.2 and 4.3. So assume again that $0 \leq \pi \leq r$ and consider again the event $\{W_{\pi,\ell}=w, T_\ell=\tau\}$, where $r \geq 1$ and $0 \leq \pi < r$. We consider the case $\pi = r$ in Section 5.1. For convenience, since we may regard π as fixed, as are r and ℓ, we now drop the subscripts on the W and the T in what follows.

We first generalize the sequential coding scheme of the proof of Lemma 2.4 to handle the joint event $\{W=w, T=\tau\}$. Our subsequent plan is to introduce some strings in the new coding, define a corresponding statistic $Q = Q_{\pi,\ell}$ by Definition 5.1, and in turn determine the joint probability generating function of T, Q and W.

As before, to fill the cells of $[\ell]$ for the first time by trial τ we must place a ball in a *new* cell from $[\ell]$ each of precisely ℓ times. For the event $\{W=w, T=\tau\}$

we may also introduce *new* digits in sequence that *aren't* from $[\ell]$, but *are* from $[k-\pi]\setminus[\ell]$, as long as $w>\ell$. As in the proof of Lemma 2.4, denote by the letter N the occurrence of a new digit along a particular sequence, whenever the new digit is from $[\ell]$. Denote now by K a new digit that comes from $[k-\pi]\setminus[\ell]$. Finally, if a digit in sequence is either from $[k]\setminus[k-\pi]$, or is *old*, because its digit matches a previous digit in sequence (from either an N or a K), denote this occurrence by O; when $\pi=0$, the set $[k]\setminus[k-\pi]=\emptyset$ and O represents only an old cell. Introduce elementary events that partition $\{W=w,T=\tau\}$ by spelling τ-long words from the alphabet $\{N,K,O\}$; words may start in any letter if $\pi\geq 1$, but must end at the τ^{th} coordinate in the letter N. For example, let $\ell=2$, $k=7$ (so $r=5$), $\pi=3$, $w=4$, and $\tau=8$. Both points $(5,4,4,7,3,3,1,2)$ and $(6,3,3,5,4,4,2,1)$ belong to the event $OKOOKONN$. In this word there are $\ell=2$ N's, $w-\ell=2$ K's, and $\tau-w=4$ O's. There are $\pi=3$ ways to choose the first O, $r-\pi=5-3=2$ ways to choose the first K, $\pi+1=4$ ways to choose the second and third O's, 1 way to choose the second K, etc. Hence the cardinality of $OKOOKONN$ is $\ell!(r-\pi)_{w-\ell}\cdot\pi(\pi+1)^2(\pi+2)$.

Before each occurrence of an N or a K, we may in general have zero or more O's. Call any maximal length sequence of consecutive O's of at least one O from a word of finite length on the alphabet $\{N,K,O\}$ a *string* of O's. In the example above there are 3 strings of O's. Before the final N of the example there is an *empty string* of O's, whose length is 0. To enumerate the event $\{W=w,T=\tau\}$, first we count the number of ways of laying down a skeleton pattern of ℓ many N's and $(w-\ell)$ many K's with the last letter N. There are precisely $\binom{w-1}{\ell-1}$ such words. For each such pattern the number of ways to fill in digits for the K's is $(r-\pi)_{w-\ell}$, while there are $\ell!$ ways to fill in digits for the N's. For any fixed skeleton word, we count the number of ways to put in patterns of O's and digits for the O's as follows.

$$\sum_{\alpha_0+\alpha_1+\cdots+\alpha_{w-1}=\tau-w}\pi^{\alpha_0}(\pi+1)^{\alpha_1}\cdots(\pi+w-1)^{\alpha_\ell}. \tag{5.1}$$

Here $\alpha_i\geq 0$ is the length of the (possibly empty) string of O's before the the $(i+1)^{\text{st}}$ letter of type either N or K, $i=0,\ldots,w-1$. We may think of the α_i, $0\leq i\leq w-1$ as a function of the word that represents a subevent of $\{W=w,T=\tau\}$, and so in the example we have $\alpha_0=1$, $\alpha_1=2$, $\alpha_2=1$, $\alpha_3=0$. In case $\pi=0$ we take $0^0=1$ in the expression (5.1), corresponding to the condition that O can not appear as a first letter. If $\tau-w=0$, then the sum (5.1) is taken to equal 1 since there are no positions for O's to occupy. Indeed $\tau-w$ denotes the total length of the strings of O's on $\{W=w,T=\tau\}$.

By Lemma 2.1 the sum (5.1) is the π-Stirling number

$$\left\{{\pi+w-1+(\tau-w) \atop \pi+w-1}\right\}_\pi=\left\{{\pi+\tau-1 \atop \pi+w-1}\right\}_\pi.$$

Noting that $\ell!(r-\pi)_{w-\ell}\binom{w-1}{w-\ell}=\ell(w-1)!\binom{r-\pi}{w-\ell}$, by the above discussion we have a sequential coding proof of Proposition 4.5. One can count the event $\{L_n=x\}\cap A_\ell^{(n)}$ also by this device, corresponding to the case $\pi=0$, by removing the condition that the last digit is an N at trial n. This leads to the summands on the left side of (1.1).

Definition 5.1 Define the statistic $Q_{\pi,\ell}$ as "the number of strings of O's until time T_ℓ."

We now calculate the joint probability generating function $E\left\{s^T t^Q u^W\right\}$. For each w with $\ell \le w \le k-\pi$, there are $\ell!(r-\pi)_{w-\ell}\binom{w-1}{w-\ell}$ ways to choose digits for a skeletal sequence of $(w-\ell)$ many K's and ℓ many N's. We must have $\tau \ge w$ and $q \le w$. If $\tau - w = 0$ we must have $q = 0$, since there are no positions for O's to occupy, and therefore the sum on q in the following (5.2) is simply $1 \cdot t^0$. We break up the sum (5.1) for the contribution of O's according to the value of q with the $'q$ notation. If $\tau - w \ge 1$, then $q \ge 1$ and thus at $q = 0$ the $'q$ sum in (5.2) must be zero, as an empty sum. Thus we have that $E\left\{s^T t^Q u^W\right\}$ is written

$$\sum_{w=\ell}^{k-\pi}(su)^w\frac{(r-\pi)_{w-\ell}}{k^w}\binom{w-1}{w-\ell}\sum_{\tau=w}^{\infty}\frac{\ell!}{k^{\tau-w}}s^{\tau-w}\sum_{q=0}^{w}t^q\sum_{\alpha_0+\cdots+\alpha_{w-1}=\tau-w}^{'q}\pi^{\alpha_0}\cdots(\pi+w-1)^{\alpha_{w-1}}, \tag{5.2}$$

where the abbreviated notation $'q$ in the upper index of the inner sum of products indicates that exactly q of the nonnegative integer exponents α_i satisfy $\alpha_i \ge 1$, and where $k^{-\tau}$ accounts for probability. For example if $\pi = 2$, $w = 3$, $\tau - w = 3$, and $q = 2$, then the $'q$ sum is $\sum_{\alpha_0+\alpha_1+\alpha_2=3}^{'2} 2^{\alpha_0}\cdot 3^{\alpha_1}\cdot 4^{\alpha_2} = 2^1\cdot 3^2+2^2\cdot 3^1+2^1\cdot 4^2+2^2\cdot 4^1+3^1\cdot 4^2+3^2\cdot 4^1$. Now fix w in the outer sum of (5.2), and make the change of variables $\sigma = \tau - w$ and write the double sum over τ and q in (5.2) by

$$\sum_{\sigma=0}^{\infty}\frac{\ell!}{k^\sigma}s^\sigma\sum_{q=0}^{w}t^q\sum_{\alpha_0+\cdots+\alpha_{w-1}=\sigma}^{'q}\pi^{\alpha_0}\cdots(\pi+w-1)^{\alpha_{w-1}} = \ell!\prod_{i=0}^{w-1}\left(1+t\sum_{\alpha=1}^{\infty}\left(\frac{\pi(i)s}{k}\right)^{\alpha}\right), \tag{5.3}$$

where $\pi(i) := \pi + i$, $i = 0, 1, \dots, w-1$. Here again, if $\sigma = 0$ then we must have $q = 0$, which corresponds to the constant term 1 in the product. This product form identity is a key motivation for introducing the statistic Q. Rewrite the factor under the product on the right side of (5.3) as

$$1+t\sum_{\alpha=1}^{\infty}\left(\frac{\pi(i)s}{k}\right)^{\alpha} = \frac{k+(t-1)\pi(i)s}{k-\pi(i)s}. \tag{5.4}$$

Therefore, by (5.2)–(5.4), we have

$$E\left\{s^T t^Q u^W\right\} = \ell!\sum_{w=\ell}^{k}(su)^w\frac{(r-\pi)_{w-\ell}}{k^w}\binom{w-1}{w-\ell}\prod_{i=0}^{w-1}\left(\frac{k+(t-1)\pi(i)s}{k-\pi(i)s}\right). \tag{5.5}$$

Now write

$$\prod_{i=0}^{w-1}\left(\frac{k+(t-1)\pi(i)s}{k-\pi(i)s}\right) = \prod_{i=0}^{k-\pi-1}\left(\frac{1}{k-\pi(i)s}\right)\prod_{i=0}^{w-1}(k+(t-1)\pi(i)s)\prod_{i=w}^{k-\pi-1}(k-\pi(i)s), \tag{5.6}$$

and abbreviate these last three products respectively as Π_1, Π_2, Π_3. To represent Π_1, it is straightforward to see by Lemma 2.1, because the maximal base factor in the product of powers is $\pi + k - \pi - 1 = k - 1$, that

$$\Pi_1 = \prod_{i=0}^{k-\pi-1} \left(\frac{1}{k-\pi(i)s} \right) = \frac{1}{k^{k-\pi}} \prod_{i=0}^{k-\pi-1} \left(\sum_{\alpha=0}^{\infty} \left(\frac{\pi(i)s}{k} \right)^{\alpha} \right) = \frac{1}{k^{k-\pi}} \sum_{h=0}^{\infty} \frac{s^h}{k^h} \left\{ {k-1+h \atop k-1} \right\}_{\pi}. \tag{5.7}$$

To expand Π_3, we introduce the definition of a p-Stirling number of the first kind.

Definition 5.2 Define $\left[{n \atop m} \right]_p$ as the number of permutations of n into m cycles such that each of $1, 2, \ldots, p$ belong to distinct cycles. The ordinary Stirling number of the first kind is the case $p = 1$ and is denoted $\left[{n \atop m} \right]$. Also by convention $\left[{n \atop m} \right]_0 = \left[{n \atop m} \right]$ and $\left[{0 \atop 0} \right] = 1$.

By [4, Theorem 7], we have the following arithmetic representation of the p-Stirling number of the first kind.

Lemma 5.3 *Let $0 \le p \le n - m \le n$. Then, taking the following sum equal to 1 if $m = 0$,*

$$\left[{n \atop n-m} \right]_p = \sum_{p \le i_1 < i_2 < \cdots < i_m < n} i_1 i_2 \cdots i_m.$$

By Lemma 5.3 we have, by $\pi(w) = \pi + w$ and $\pi(k - \pi - 1) + 1 = k$, that

$$\Pi_3 = \prod_{i=w}^{k-\pi-1} (k - \pi(i)s) = \sum_{\nu=0}^{k-\pi-w} \left[{k \atop k-\nu} \right]_{\pi+w} (-1)^{\nu} s^{\nu} k^{k-\pi-w-\nu}. \tag{5.8}$$

Therefore, by (5.5)–(5.8), we have that $E\left\{s^T t^Q u^W\right\}$ is given by

$$\ell! \sum_{w=\ell}^{k} (su)^w \frac{(r-\pi)_{w-\ell}}{k^w} \binom{w-1}{w-\ell} \Pi_2 \sum_{h=0}^{\infty} \frac{s^h}{k^h} \left\{ {k-1+h \atop k-1} \right\}_{\pi} \sum_{\nu=0}^{k-\pi-w} \left[{k \atop k-\nu} \right]_{\pi+w} (-1)^{\nu} s^{\nu} k^{-w-\nu} \tag{5.9}$$

where we have canceled powers of $k^{k-\pi}$ in denominator from Π_1 of (5.7) and numerator from Π_3 of (5.8). We now set $t = 1$, so that $\Pi_2 = k^w$ by (5.6). As a consequence, we have by (5.9) that, for fixed $w \le \tau$, the $s^{\tau} u^w$ term of $E\left\{s^T u^W\right\}$ is

$$\frac{(r-\pi)_{w-\ell}}{k^{k-\pi}} \binom{w-1}{w-\ell} \ell!(su)^w \left(\sum_{h+\nu=\tau-w} \frac{s^{h+\nu}}{k^h} \left\{ {k-1+h \atop k-1} \right\}_{\pi} \left[{k \atop k-\nu} \right]_{\pi+w} (-1)^{\nu} k^{k-\pi-w-\nu} \right). \tag{5.10}$$

Hence, by equating coefficients in the bivariate probability generating function of T and W, we obtain, after observing that, for fixed w and τ, the power of k in (5.10) is $k^{-(w+h+\nu)} = k^{-\tau}$,

$$k^\tau P(W\!=\!w, T\!=\!\tau) = \ell!(r-\pi)_{w-\ell}\binom{w-1}{w-\ell}\sum_{h+\nu=\tau-w}\left\{ {k-1+h \atop k-1} \right\}_\pi \left[{k \atop k-\nu} \right]_{\pi+w}(-1)^\nu. \tag{5.11}$$

Now plug in Proposition 4.5 to the left side of (5.11). Notice that the combinatorial coefficients in front of the Stirling numbers on the resulting two sides are equal as noted in the alternative proof of Proposition 4.4 just before Definition 5.1. Therefore, by (5.11) and Proposition 4.5, we obtain the following identity.

Proposition 5.4 *Let* $0 \le \pi < r$, $\ell \le w \le k-\pi$, *and* $\tau \ge w$. *Then*

$$\sum_\nu \left\{ {k-1+\tau-w-\nu \atop k-1} \right\}_\pi \left[{k \atop k-\nu} \right]_{\pi+w}(-1)^\nu = \left\{ {\pi+\tau-1 \atop \pi+w-1} \right\}_\pi .$$

For illustration of this result, take $\ell = 1$, $k = 6$, $w = 2$, $\pi = 2$, and $\tau = 5$; thus $r = 5 > \pi$. We obtain $\left\{ {8 \atop 5} \right\}_2 \left[{6 \atop 6} \right]_4 - \left\{ {7 \atop 5} \right\}_2 \left[{6 \atop 5} \right]_4 + \left\{ {6 \atop 3} \right\}_2 \left[{6 \atop 4} \right]_4 = \left\{ {6 \atop 3} \right\}_2$; that is, $910 \cdot 1 - 125 \cdot 9 + 14 \cdot 20 = 65 = \left\{ {6 \atop 3} \right\}_2$.

5.1 Strings

In this section we eliminate the variable $W_{\pi,\ell}$ from the analysis and go back to the construction in the proof of Lemma 2.4 involving words on the alphabet $\{N, J\}$. This may be regarded as the case $\pi = r$ for the three letter alphabet, so K no longer exists (the 3 letter alphabet collapses), and the role of O is played now by J. As in the proof of Lemma 2.4, call a consecutive sequence of J's of maximal length, consisting of at least one J in a finite word from the alphabet $\{N, J\}$ a *string* of J's in this word.

Definition 5.5 Define the statistic $\widetilde{Q}_\ell$ as "the number of strings of J's until time T_ℓ."

By (5.2)–(5.3), with $u = 1$ and $\pi = r$ or $k - \pi = \ell$, and $w \equiv \ell$, we have by the left side of (5.3) that

$$E\left\{s^T t^{\widetilde{Q}}\right\} = \sum_{\tau=\ell}^{\infty}\frac{\ell!}{k^\tau}s^\tau\sum_{q=0}^{\ell}t^q\left(\sum_{\alpha_0+\alpha_1+\cdots+\alpha_{\ell-1}=\tau-\ell}^{'q} r^{\alpha_0}(r+1)^{\alpha_1}\cdots(r+\ell-1)^{\alpha_{\ell-1}}\right), \tag{5.12}$$

where again the notation $'q$ in the upper index of the inner sum of products indicates that exactly q of the nonnegative integer exponents α_i satisfy $\alpha_i \ge 1$. The case $q = 0$ is only possible if $\tau = \ell$, in which case the $'q$ sum is taken as 1. Similar to (5.3), we find the product form for (5.12) after introducing $\sigma = \tau - \ell$, as follows. The generating function $E\left\{s^T t^{\widetilde{Q}}\right\}$ is written

$$\frac{\ell!}{k^\ell}s^\ell\sum_{\sigma=0}^{\infty}\frac{s^\sigma}{k^\sigma}\sum_{q=0}^{\ell}t^q\sum_{\alpha_0+\cdots+\alpha_{\ell-1}=\tau-\ell}^{\prime q} r^{\alpha_0}\cdots(r+\ell-1)^{\alpha_{\ell-1}} = \frac{\ell!}{k^\ell}s^\ell\prod_{i=0}^{\ell-1}\left(1+t\sum_{\alpha=1}^{\infty}\left(\frac{r(i)s}{k}\right)^\alpha\right), \tag{5.13}$$

where $r(i) = r+i$, $i = 0,\dots\ell-1$; so $r(i)$ ranges from r to $r+\ell-1 = k-1$. By applying the identity of (5.4), we have

$$\prod_{i=0}^{\ell-1}\left(1+t\sum_{\alpha=1}^{\infty}\left(\frac{r(i)s}{k}\right)^\alpha\right) = \prod_{i=0}^{\ell-1}\frac{k+(t-1)r(i)s}{k-r(i)s}. \tag{5.14}$$

Here, for the product of denominators on the right side of (5.14), by identity (5.7) with $\pi = r$, we have

$$\prod_{i=0}^{\ell-1}\frac{1}{k-r(i)s} = k^{-\ell}\sum_{h=0}^{\infty}\frac{s^h}{k^h}\left\{\begin{matrix}k-1+h\\k-1\end{matrix}\right\}_r \tag{5.15}$$

For the numerator on the right side of (5.14), that is the equivalent of Π_2 of (5.6) with $w\equiv\ell$, we write by Lemma 5.3 that

$$\prod_{i=0}^{\ell-1}(k+(t-1)r(i)s) = \sum_{\nu=0}^{\ell}\begin{bmatrix}k\\k-\nu\end{bmatrix}_r s^\nu(t-1)^\nu k^{\ell-\nu}. \tag{5.16}$$

Now plug (5.14)–(5.16) into (5.13) to obtain

$$E\left\{s^T t^{\widetilde{Q}}\right\} = \frac{\ell!}{k^\ell}s^\ell\left(k^{-\ell}\sum_{h=0}^{\infty}\frac{s^h}{k^h}\left\{\begin{matrix}k-1+h\\k-1\end{matrix}\right\}_r\right)\left(\sum_{\nu=0}^{\ell}\begin{bmatrix}k\\k-\nu\end{bmatrix}_r s^\nu(t-1)^\nu k^{\ell-\nu}\right). \tag{5.17}$$

Finally, rewrite the double sum by $h+\nu=\sigma$, and expand $(t-1)^\nu$ by the binomial theorem. So, writing in general $\ell\wedge\sigma = \min\{\ell,\sigma\}$,

$$E\left\{s^T t^{\widetilde{Q}}\right\} = \frac{\ell!}{k^\ell}s^\ell\sum_{\sigma=0}^{\infty}\frac{s^\sigma}{k^\sigma}\sum_{\nu=0}^{\ell\wedge\sigma}\left\{\begin{matrix}k-1+\sigma-\nu\\k-1\end{matrix}\right\}_r\begin{bmatrix}k\\k-\nu\end{bmatrix}_r\sum_{q=0}^{\nu}t^q(-1)^{\nu-q}\binom{\nu}{q}. \tag{5.18}$$

Here σ represents $\tau-\ell$. Now read off the $s^\tau t^q$ coefficient in (5.18) to obtain

$$P(\widetilde{Q}_\ell = q, T_\ell = \tau) = \frac{\ell!}{k^\tau}\sum_{\nu=0}^{\ell\wedge(\tau-\ell)}\left\{\begin{matrix}k-1+\tau-\ell-\nu\\k-1\end{matrix}\right\}_r\begin{bmatrix}k\\k-\nu\end{bmatrix}_r(-1)^{\nu-q}\binom{\nu}{q}. \tag{5.19}$$

If $\tau=\ell$ we must have $\nu=0$ in this sum and $q=0$, and we simply recover the trivial fact that $P(T_\ell=\ell) = \frac{\ell!}{k^\ell} = \frac{\ell!}{k^\ell}\left\{\begin{smallmatrix}k-1\\k-1\end{smallmatrix}\right\}_r\left[\begin{smallmatrix}k\\k\end{smallmatrix}\right]_r$. If on the other hand $\tau>\ell$, then we must have $q \geq 1$. Since $\sum_{q\geq 1}(-1)^{\nu-q}\binom{\nu}{q} = (1-1)^\nu - (-1)^\nu = (-1)^{\nu-1}$, if $\nu \geq 1$, while this sum is zero if $\nu = 0$, we obtain by (5.19) that

$$P(T_\ell = \tau) = \frac{\ell!}{k^\tau} \sum_{\nu=1}^{\ell\wedge(\tau-\ell)} \left\{ {k-1+\tau-\ell-\nu \atop k-1} \right\}_r \left[{k \atop k-\nu} \right]_r (-1)^{\nu-1}, \ \tau > \ell. \tag{5.20}$$

Hence by (5.20) and Lemma 2.3 we obtain the following.

Proposition 5.6 *If $\tau > \ell$, then*

$$\sum_{\nu=1}^{\ell\wedge(\tau-\ell)} \left\{ {r+\tau-1-\nu \atop r+\ell-1} \right\}_r \left[{r+\ell \atop r+\ell-\nu} \right]_r (-1)^{\nu-1} = \left\{ {r+\tau-1 \atop r+\ell-1} \right\}_r.$$

Since the right side of this result can be incorporated as a $\nu = 0$ term in the sum, Proposition 5.6 is called an orthogonality relation between r-Stirling numbers of the first and second kinds. This result extends [4, Theorem 25 (54) case $r = p$] to non-matching rows and columns in the triangular tables [4, Table 1] of r-Stirling numbers of the first and second kinds. For example, if $\ell = 4$, $k = 6$, and $\tau = 6$, we obtain $\left\{ {6 \atop 5} \right\}_2 \left[{6 \atop 5} \right]_2 - \left\{ {5 \atop 5} \right\}_2 \left[{6 \atop 4} \right]_2 = 14 \cdot 14 - 1 \cdot 71 = 125 = \left\{ {7 \atop 5} \right\}_2$.

Definition 5.7 Let $\beta \geq r \geq 1$, $\sigma \geq 0$, and $q \geq 0$. Define

$$S(\beta+\sigma, \beta; r, q) = \sideset{}{'^{q}}\sum_{\alpha_0+\alpha_1+\cdots+\alpha_{\beta-r}=\sigma} r^{\alpha_0}(r+1)^{\alpha_1}\cdots\beta^{\alpha_{\beta-r}},$$

where $\alpha_i \geq 0$ and $'q$ denotes that exactly q of the α_i satisfy $\alpha_i \geq 1$. If both $\sigma = 0$ and $q = 0$ define $S(\beta, \beta; r, 0) = 1$. Otherwise, if the sum is empty for $\sigma \geq 1$, then $S(\beta+\sigma, \beta; r, 0) = 0$.

We observe, by Lemma 2.1, that $\sum_{q=0}^{\beta-r} S(\beta+\sigma, \beta; r, q) = \left\{ {\beta+\sigma \atop \beta} \right\}_r$. By comparing (5.13) and (5.19), we have the following.

Corollary 5.8 *Let $r \geq 0$, $1 \leq \ell \leq \tau$, and $0 \leq q \leq \ell \wedge \sigma$, where we denote $\sigma = \tau - \ell$. Then*

$$S(r+\ell+\sigma-1, r+\ell-1; r, q) = \sum_{\nu=q}^{\ell\wedge\sigma} \left\{ {r+\ell-1+\sigma-\nu \atop r+\ell-1} \right\}_r \left[{r+\ell \atop r+\ell-\nu} \right]_r (-1)^{\nu-q}\binom{\nu}{q}.$$

For example, take $\ell = 3$, $r = 2$, $q = 2$ and $\sigma = 4$ in Corollary 5.8. Then $S(8, 4; 2, 2) = \sideset{}{'^{2}}\sum_{\alpha_0+\alpha_1+\alpha_2=4} 2^{\alpha_0}3^{\alpha_1}4^{\alpha_2} = 2^13^3+2^23^2+2^33^1+2^14^3+2^24^2+2^34^1+3^14^3+3^24^2+3^34^1 = 782$, while $\left\{ {6 \atop 4} \right\}_2 \left[{5 \atop 3} \right]_2 \binom{2}{2} - \left\{ {5 \atop 4} \right\}_2 \left[{5 \atop 2} \right]_2 \binom{3}{2} = 55 \cdot 26 - 9 \cdot 24 \cdot 3 = 782$.

5.2 Strings of a Partition of $[\beta+\sigma]$ into β Blocks

We reconsider the combinatorial proof of Lemma 2.3 at the beginning of Section 4. Our purpose is to extend the bijection of that proof to a bijection of the subevent $\{\widetilde{Q}=q\}\cap E_{\lambda,\tau}$, where $E_{\lambda,\tau}$ is given by Definition 4.1 and $\widetilde{Q}$ is given by Definition 5.5. In our first illustration of that proof, with $r=2$, $\ell=3$, and $\tau=6$, we took $\underline{\omega}=(1,\ 2,\ 4,\ 4,\ 2,\ 3)$, of the form $NNJJJN$, where the sequence of new digits (N) is the identity permutation id on $[\ell]=[3]$. The bijection associates $\underline{\omega}\in E_{\mathrm{id},\tau}$ in this case with a partition into $r+\ell-1=4$ blocks of the $(r+\tau-1)$ coordinate positions *1, 2,…, 7*, as follows: $\{\ \{1,\ 5,\ 6\},\ \{2\},\ \{3\},\ \{4,\ 7\}\ \}$. We aim to relate this partition to the number q of strings of J's in the word on N's and J's that codes $\underline{\omega}$; so we want to determine this value of q from the partition itself. It turns out that we should remove the minimum coordinate position from each block, and take the union S of remaining positions and find the number of strings of consecutive integers in the set S. For the present example we have $S=\{5,\ 6\ \}\cup\{7\}$, consistent with $q=1$. Here, for any finite subset S of positive integers, define a *string* in S as a sequence of consecutive integers in S that is of maximal length.

Definition 5.9 Let $\beta\ge r\ge 1$, and $\sigma\ge 0$, and let ρ be a partition of $[\beta+\sigma]$ into β nonempty disjoint blocks with r distinguished blocks, meaning each element of $[r]$ belongs to a different block. For each block B denote $\mu(B)=\min B$, the minimal element of B. Define the number of strings of the partition ρ as the number of (maximal) strings of consecutive integers in $S=\bigcup_{B\text{ a block of }\rho}(B\setminus\{\mu(B)\})$. If S is empty, that is $\sigma=0$ and we have only singletons in the partition of $[\beta]$, then the number of strings is defined to be zero.

To establish the equivalence of the two ways of counting strings, consider an illustration with the same parameters $r=2$ and $\ell=3$ as before, but now with $\tau=9$, and the example sequence $\underline{\omega}=(4,1,4,5,1,2,1,1,3)$, whose coded sequence is $JNJJJNJJN$, and whose extended sequence is $\underline{\omega}^{+}=(a,b;4,1,4,5,1,2,1,1,3)$. The associated partition on coordinate positions $[10]$ (by $r+\tau-1=10$) is determined as: $\{\{1,\ 3,\ 5\ \},\ \{2,\ 6\},\ \{4,\ 7,\ 9,\ 10\},\ \{8\}\}$. The integers *1* and *2* are coordinate positions of a and b, that will fall in different or distinguished blocks of the partition by definition. Recall that in our correspondence in Section 4, the r blocks B such that $\mu(B)\in\{1,2\ldots,r\}$ are defined to include coordinate positions, respectively, of digits $\ell+i$, $1\le i\le r$, that exist in $\underline{\omega}$. Therefore the set of coordinate positions of the extended sequence giving rise to J's in the coded sequence due to digits from $[r+\ell]\setminus[\ell]$ appearing in $\underline{\omega}$ equals: $\bigcup_{\mu(B)\in\{1,2\ldots,r\}}(B\setminus\{\mu(B)\})$, that is the union of the distinguished blocks after the respective extra coordinate positions $1,2,\ldots,r$ have been removed. In our example these coordinate positions are $\{3,\ 5\}\cup\{6\}$. There are also J's entering the coded sequence due to digits $i\in[\ell]$ being repeated in $\underline{\omega}$. By definition, the remaining blocks $B_1,B_2,\ldots,B_{\ell-1}$ after the distinguished blocks are given by: B_i equals the set of coordinate positions of digit i in the extended sequence. Therefore by removing the minimum from each of these blocks B_i as well, and taking the union, we obtain all other extended coordinate positions corresponding to J's, so at positions $\{7,\ 9,\ 10\}$ of $\underline{\omega}^+$. Hence, by combining the two sources of J's, we count the number of strings of J's in the coded sequence by the number of strings of consecutive integers in $S = \{3,\ 5,\ 6,\ 7,\ 9,\ 10\}$; we have $q = 3$ strings of consecutive integers in S.

Given $r \geq 1$ and $1 \leq \ell \leq \tau$, we write $\beta = r + \ell - 1$ and $\sigma = \tau - \ell$, so that $\beta + \sigma = r + \tau - 1$. It remains to establish a bijection between $\{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$ and the set of all partitions $\mathbf{R}_q$ such that each $\rho \in \mathbf{R}_q$ is a partition of the coordinate positions $\{1, 2, \ldots, \beta + \sigma\}$ into β blocks, where there are r distinguished blocks, and where also there are q strings of consecutive integers in the set $S = \bigcup_{\text{blocks } B \text{ of } \rho} (B \setminus \{\mu(B)\})$. For this we simply recall by Definition 5.5 that $\widetilde{Q}$ is the number of strings of J's in the coded sequence for a given element $\underline{\omega} \in \{T = \tau\}$. Since we already established a bijection $\varphi : E_{\mathrm{id},\tau} \to \mathbf{R}$ in Section 4, where $\mathbf{R} = \bigcup_{q=0}^{\sigma} \mathbf{R}_q$, it is only a matter of showing that the restriction of φ to $\{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$ maps onto $\mathbf{R}_q$. But this is now obvious by the construction at the end of the combinatorial proof of Section 4, together with the above correspondence between strings of consecutive coordinate positions in S and strings of J's in a coded sequence. That is, we let $\rho \in \mathbf{R}_q$ and find an element $\underline{\omega}^+$ such that $\underline{\omega} \in E_{\mathrm{id},\tau}$ and $\varphi(\underline{\omega}) = \rho$. Then by virtue of $\rho \in \mathbf{R}_q$, the coded sequence of $\underline{\omega}$ has q strings of J's, and hence $\underline{\omega} \in \{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$, as required.

Finally, we want to compute the size of $\mathbf{R}_q$ by an arithmetic sum, in analogy with Lemma 2.1. Since the sums $S(\beta + \sigma; r, q)$ of Definition 5.7 themselves have a sum over q equal to the r-Stirling number $\left\{ {\beta+\sigma \atop \beta} \right\}_r$, as noted just after this definition, it is natural to believe that

$$|\mathbf{R}_q| = S(\beta + \sigma, \beta; r, q), \tag{5.21}$$

where $\beta + \sigma = r + \tau - 1$, $\beta = r + \ell - 1$, and the left side is the cardinality of $\mathbf{R}_q$. To prove that indeed (5.21) is correct, we may use the generating function approach. By the bijection above between $\{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$ and $\mathbf{R}_q$, the number $|\mathbf{R}_q|$ is the same as the coefficient of $s^\tau t^q$ in $\frac{k^\tau}{\ell!} E(s^T t^{\widetilde{Q}})$. Therefore (5.21) holds by (5.12) and Definition 5.7. We also prove (5.21) by a direct enumeration of $\{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$. Indeed, fix the permutation $\lambda = \mathrm{id}$ for the sequence of new digits, and as before consider a possibly empty string of J's before each N of the skeletal sequence of ℓ many N's in a coded sequence for $\underline{\omega}$; the lengths of these strings are $\alpha_0, \ldots, \alpha_{\ell-1}$, with $\alpha_i \geq 0$. Then the number of sample points $\underline{\omega} \in \{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}$ such that there are exactly q strings of J's in the coded sequence can be computed by accounting for all patterns with $\sum_{i=0}^{\ell-1} \alpha_i = q$ and using the fact that there are $r + i$ possible values for each J in the i^{th} string of J's (empty or not). So we obtain $|\{\widetilde{Q} = q\} \cap E_{\mathrm{id},\tau}|$ as the sum of the products $r^{\alpha_0}(r+1)^{\alpha_1} \cdots \beta^{\alpha_{\beta-r}}$ over the eligible patterns, consistent with Definition 5.7.

Corollary 5.10 *Let $\beta \geq r \geq 1$ and $\sigma \geq 0$. Then $S(\beta + \sigma, \beta; r, q)$ is the number of* *partitions of* $[\beta+\sigma]$ *into* β *nonempty blocks, such that the distinguished elements* $[r]$ *fall in distinct blocks, and such that the number of strings of this partition is* q*.*

For example, by Definition 5.7, we compute $S(5,3;2,2) = 2^1 3^1 = 6$. This can also be seen directly by finding all partitions of $\{\mathit{1},\ \mathit{2},\ \mathit{3},\ \mathit{4},\ \mathit{5}\}$ into 3 blocks, with digits *1* and *2* in distinct blocks, and with the total number of strings equal to $q = 2$, as follows.

$$\{\mathit{1}\},\{\mathit{2},\ \mathit{3}\},\{\mathit{4},\ \mathit{5}\};\quad \{\mathit{1},\ \mathit{3}\},\{\mathit{2}\},\{\mathit{4},\ \mathit{5}\};\quad \{\mathit{1},\ \mathit{3}\},\{\mathit{2},\ \mathit{5}\},\{\mathit{4}\};$$

$$\{\mathit{1},\ \mathit{5}\},\{\mathit{2},\ \mathit{3}\},\{\mathit{4}\};\quad \{\mathit{1},\ \mathit{3},\ \mathit{5}\},\{\mathit{2}\},\{\mathit{4}\};\quad \{\mathit{1}\},\{\mathit{2},\ \mathit{3},\ \mathit{5}\},\{\mathit{4}\}.$$

6 Discussion

The r-Stirling number of the second kind $\left\{ {\beta+\sigma \atop \beta} \right\}_r$ is defined for $1 \le r \le \beta$ and $\sigma \ge 0$ as the number of partitions of $[\beta+\sigma]$ into β blocks with r distinguished blocks. We enumerate a subcollection of these partitions determined by the number of strings q in terms of r-Stirling numbers of both first and second kinds via Corollaries 5.8 and 5.10.

An open-ended problem to handle further classes of partitions is motivated by the paper [5, p. 136], which discusses the number $e(n,k)$ of partitions of $[n]$ into k blocks, but with no singleton blocks. We have $e(n+k,k) = H(n,n-k)$, where $H(n,k)$, $0 \le k \le n$, are called the Ward numbers [7, p. 41]. In particular it is expected that there is a formula for $e(n,k)$ in terms of Stirling numbers akin to the formula of Corollary 5.8.

Acknowledgement

The author wishes to thank the referees for their insightful comments and encouragement to make improvements in the presentation.

References

[1] E. T. Bell, Exponential numbers, *Amer. Math. Monthly* 41 (1934), 411–419.

[2] K. N. Boyadzhiev, Exponential polynomials, Stirling numbers, and evaluation of some Gamma integrals, *Abstr. Appl. Anal.* 2009.

[3] K. N. Boyadzhiev, Close Encounters with the Stirling Numbers of the Second Kind, *Math. Mag.* 85 (4) (2012), 252–266.

[4] A. Z. Broder, The r-Stirling numbers, *Discrete Math.* 49 (3) (1984), 241–259.

[5] D. Callan, T. Mansour and M. Shattuck, Some identities for derangement and Ward number sequences and related bijections, *Pure Math. Appl. (PU.M.A.)* 25 (2) (2015), 132–143.

[6] L. Carlitz, Weighted Stirling numbers of the first and second kind—I, *Fibonacci Quart.* 18 (2) (1980), 147–162.

[7] L. Clark, Asymptotic normality of the Ward numbers, *Discrete Math.* 203 (1-3) (1999), 41–48.

[8] J. Engbers and A. Hammett, Problem 1146, *College Math. J.* 50 (2) (2019), p. 143.

[9] R. L. Graham, D. R. Knuth and O. Patashnik, *Concrete Mathematics*, 2nd Edition, New York, Addison-Wesley (1994).

[10] N. L. Johnson, A. W. Kemp and S. Kotz, *Univariate Discrete Distributions*, 3rd ed., Hoboken, N.J., Wiley (2005).

[11] Z. Kereskényi-Balogh and G. Nyul, Stirling numbers of the second kind and Bell numbers for graphs, *Australas. J. Combin.* 58 (2014), 264–274.

[12] D. Knuth, *The Art of Computer Programming, Vol. 1*, Third Ed., Reading, Massachusetts, Addison Wesley Longman (1997).

[13] N. Pippenger, The Hypercube of Resistors, Asymptotic Expansions, and Preferential Arrangements, *Math. Mag.* 83 (5) (2010), 331–346.

[14] J. Quaintance and H. W. Gould, *Combinatorial Identities for Stirling Numbers*, World Scientific, Singapore (2016).

[15] G.-C. Rota, The number of partitions of a set, *Amer. Math. Monthly* 71 (5) (1964), 498–504.

[16] J. Touchard, Nombres exponentiels et nombres de Bernoulli, *Canadian J. Math.* 8 (1956), 305–320.

(Received 1 Aug 2019; revised 3 July 2020, 7 Aug 2020)