Restricted Stirling and Lah numbers and their inverses

John Engbers* David Galvin† Cliff Smyth‡

October 20, 2016

Abstract

Let ${n \brace k}_{[r]}$ denote the number of ways of partitioning a set of size n into k non-empty blocks, each of which has size at most r. For all r we find a combinatorial interpretation for the entries of the inverse of the matrix $\left[{n \brace k}_{[r]}\right]_{n,k\geq 1}$. For even r we exhibit sets of forests counted by the entries of the inverse. For odd r our interpretation is as the difference in size between two sets of forests. This answers a question raised by Choi, Long, Ng and Smith in 2006.

More generally we consider restricted Stirling numbers of the second and first kinds ${n \brace k}_R$, ${n \brack k}_R$, and Lah numbers $L(n,k)_R$, for $R \subseteq \mathbb{N}$. These are defined to be the number of ways of partitioning a set of size n into k non-empty blocks (for Stirling numbers of the second kind), cycles (for Stirling numbers of the first kind) or lists (for Lah numbers) with the size of each block, cycle or list in R. For any R satisfying $1 \in R$ (a necessary condition for the inverses to exist) we find combinatorial interpretations for the entries of the inverses of the matrices $\left[{n \brace k}_R\right]_{n,k\geq 1}$, $\left[{n \brack k}_R\right]_{n,k\geq 1}$ and $[L(n,k)_R]_{n,k\geq 1}$, as the difference in size between two sets of forests.

In the case of Stirling numbers of the second kind and Lah numbers, for certain R we can do better, interpreting the inverse entries directly as counts of single sets of forests. Among these R's are those which include 1 and 2 and which have the property that for all odd $n \in R$, $n \geq 3$, we have $n \pm 1 \in R$.

Our proofs depend in part on two combinatorial interpretations of the coefficients of the reversion (compositional inverse) of a power series.

1 Introduction

Let ${n \brace k}$ denote the number of partitions of a set of size n into k non-empty blocks; this is a *Stirling number of the second kind.* The doubly-infinite matrix $\left[{n \brace k}\right]_{n,k\geq 1}$ is lower-triangular

*Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee WI.

†Department of Mathematics, University of Notre Dame, Notre Dame IN. Research supported by the Simons Foundation grant 360240 and by the National Security Agency grant NSA H98230-13-1-0248.

‡Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro NC. Research supported by the Simons Foundation grant 360468.

with integer entries and 1's on the main diagonal, so its inverse also has these properties. The inverse is well understood combinatorially:

$$\left[\begin{Bmatrix} n \\ k \end{Bmatrix}\right]^{-1}_{n,k\geq 1} = \left[(-1)^{n-k}\begin{bmatrix} n \\ k \end{bmatrix}\right]_{n,k\geq 1},$$

where $\begin{bmatrix} n \\ k \end{bmatrix}$ denotes the number of partitions of a set of size n into k non-empty cyclicly ordered sets; this is a *Stirling number of the first kind.*

In this note we consider restricted Stirling numbers of the second kind. For $R \subseteq \mathbb{N}$ the *R-restricted Stirling number of the second kind* $\begin{Bmatrix} n \\ k \end{Bmatrix}_R$ is the number of partitions of a set of size n into k non-empty blocks with all block sizes belonging to R (so $\begin{Bmatrix} n \\ k \end{Bmatrix}_{\mathbb{N}} = \begin{Bmatrix} n \\ k \end{Bmatrix}$; here and throughout $\mathbb{N} = \{1, 2, 3, \ldots\}$). Various instances of restricted Stirling numbers have appeared in the literature. Comtet [7, page 222] introduced *r-associated Stirling numbers*, corresponding to $R = \{r, r+1, r+2, \ldots\}$, and obtained recurrence relations and generating functions for them, and Choi and Smith [5] considered the complementary case $R = [r] := \{1, \ldots, r\}$. We note that if $1 \in R$ then the matrix $\left[\begin{Bmatrix} n \\ k \end{Bmatrix}_R\right]_{n,k\geq 1}$ is lower-triangular with integer entries and 1's on the main diagonal, so its inverse also has these properties.

Notation. *Let* $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_R$ *denote the* (n, k) *entry of the inverse matrix of* $\left[\begin{Bmatrix} n \\ k \end{Bmatrix}_R\right]_{n,k\geq 1}$.

Question 1.1. *Is there (up to sign) a combinatorial interpretation for* $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_R$*?*

In [4] Choi, Long, Ng and Smith consider Question 1.1 in the case $R = [r]$. For $r = 2$ they observe that it is quite classical, with $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[2]}$ being a Bessel number [9, A100861] and having many combinatorial interpretations. For example, $(-1)^{n-k}\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[2]}$ counts the number of size $n-k$ matchings of the complete graph K_{2n-1-k} [6].

For $r > 2$ Choi et al. observe that the sign-pattern of the matrix $\left[\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[r]}\right]_{n,k\geq 1}$ is not periodic or predictable, precluding a combinatorial interpretation of $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[r]}$, and they illustrate this with a portion of the $r = 3$ inverse matrix ([4, (1.2)], see also Figure 1).

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 7 & 6 & 1 & 0 & 0 \\ 0 & 10 & 25 & 10 & 1 & 0 \\ 0 & 10 & 75 & 65 & 15 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 & 0 \\ -5 & 11 & -6 & 1 & 0 & 0 \\ 10 & -45 & 35 & -10 & 1 & 0 \\ 35 & 175 & -210 & 85 & -15 & 1 \end{bmatrix}$$

Figure 1: The matrix $\left[\begin{Bmatrix} n \\ k \end{Bmatrix}_{[3]}\right]^6_{n,k=1}$ (left) and its inverse $\left[\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[3]}\right]^6_{n,k=1}$ (right); notice that the right-hand matrix does not display a periodic sign pattern.

If we relax the goal a little, however, then we can combinatorially understand not just $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[3]}$, but $\begin{Bmatrix} n \\ k \end{Bmatrix}^{-1}_{[r]}$ for any r, as the difference in sizes of two sets. To state this precisely we introduce labelled Schröder trees, a central object in much of what follows.

A *Schröder tree* is a tree with a distinguished root vertex in which no vertex has exactly one child (a *child* of a vertex v is any neighbor of v that is further from the root than v is). Notice that in this definition no order is placed on the children of a vertex. A *labelled Schröder tree* is a Schröder tree in which the leaves are labelled (for our purposes, a *leaf* of a tree is a vertex either of degree 1 or of degree 0; so in particular the tree on a single vertex is considered to have a leaf). It should also be said that whenever we say one of our combinatorial objects is labeled, we mean that all the labels used are distinct. These trees are sometimes referred to as *evolutionary trees* or *phylogenetic trees* (see for example [8]). The number of labelled Schröder trees with n leaves for $n = 1, 2, \ldots$ is the sequence $(1, 1, 4, 26, 236, 2752, 39208, \ldots)$ [9, A000311].

As an aside, note that if we extend the labelling of a labelled Schröder tree recursively by labelling each vertex with the union of the labels of its children, the result is an encoding of a total partition of the set of leaf labels [10, Example 5.2.5]. A *total partition* of a finite non-empty set X begins with a partition of X into at least two non-empty blocks, and then proceeds by furnishing each non-singleton block with a total partition; see Figure 2.

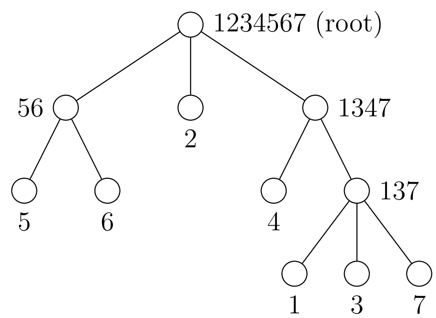

Figure 2: A total partition tree representing a total partition of $\{1, \ldots, 7\}$; on removing the labels from the non-leaf vertices we obtain a labelled Schröder tree.

A *labelled Schröder forest* is a forest whose leaves are labelled and all of whose components are labelled Schröder trees. The *down-degree* $d(v)$ of a non-leaf vertex v in a labelled Schröder forest is the number of children that v has. For each $n, k, r \geq 1$ denote by $\mathcal{F}_{[r]}(n, k)$ the set of labelled Schröder forests with label set $[n]$ and with n leaves and k components in which for every non-leaf vertex v, $d(v) \leq r$. Denote by $\mathcal{F}_{[r]}^{\text{even}}(n, k)$ those forests in $\mathcal{F}_{[r]}(n, k)$ in which there are an even number of components that have an odd number of non-leaf vertices, and denote by $\mathcal{F}_{[r]}^{\text{odd}}(n, k)$ those in which there are an odd number of components that have an odd number of non-leaf vertices.

Theorem 1.2. *For all $r \geq 1$ and all $n, k \geq 1$ we have*

$$\left\{ {n \atop k} \right\}_{[r]}^{-1} = \left| \mathcal{F}_{[r]}^{\text{even}}(n, k) \right| - \left| \mathcal{F}_{[r]}^{\text{odd}}(n, k) \right| .$$

Consider, for example, $r = 3$, $n = 4$ and $k = 2$. There are 19 labelled Schröder forests with four leaves and two components in which no non-leaf vertex has more than three children. Of these, 12 have no components with an odd number of non-leaf vertices, 4 have one such component and 3 have two such, yielding $|\mathcal{F}^{\text{even}}_{[3]}(4,2)| - |\mathcal{F}^{\text{odd}}_{[3]}(4,2)| = (12+3)-4 = 11$. This explains the $(4,2)$ entry in the right-hand matrix of Figure 1.

It would be more satisfying to have a single set of forests that is counted by $\left\{ {n \atop k} \right\}^{-1}_{[r]}$, rather than interpreting it as the difference in sizes between two sets. For $r = 3$, and more generally for odd r we cannot at the moment achieve this aim, in large part because of the aperiodic sign pattern of the associated matrix observed in [4]. Things are rather different for even r, however; it turns out that unlike the case $r = 3$ (Figure 1), when r is even the sign-pattern of $\left[\left\{ {n \atop k} \right\}^{-1}_{[r]}\right]_{n,k\geq 1}$ is very predictable; specifically, as in the case $r = 2$ we have that $(-1)^{n-k}\left\{ {n \atop k} \right\}^{-1}_{[r]}$ is positive for all n, k and even r. See Figure 3 for the case $r = 4$.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 & 0 & 0 \\ 1 & 7 & 6 & 1 & 0 & 0 \\ 0 & 15 & 25 & 10 & 1 & 0 \\ 0 & 25 & 90 & 65 & 15 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 2 & -3 & 1 & 0 & 0 & 0 \\ -6 & 11 & -6 & 1 & 0 & 0 \\ 25 & -50 & 35 & -10 & 1 & 0 \\ -140 & 280 & -225 & 85 & -15 & 1 \end{bmatrix}$$

Figure 3: The matrix $\left[\left\{ {n \atop k} \right\}_{[4]}\right]^{6}_{n,k=1}$ (left) and its inverse $\left[\left\{ {n \atop k} \right\}^{-1}_{[4]}\right]^{6}_{n,k=1}$ (right); notice that the right-hand matrix does display a periodic sign pattern.

This opens up the possibility of achieving the aim identified in [4], of a clean combinatorial interpretation of $\left\{ {n \atop k} \right\}^{-1}_{[r]}$, a possibility that we now realize.

We need the notion of an *r-good* labelled Schröder tree, which we define inductively as follows. A labelled Schröder tree consisting of a single vertex labelled with a natural number is r-good. A labelled Schröder tree T with more than one leaf (again with all leaves labelled from $\mathbb{N}$) is r-good if

- all non-leaf vertices of T have at most r children;
- along the path $v_1, v_2, \ldots, v_k$ from the root (v_1) to the leaf with largest label (v_k) all vertices (except v_k) have two children, and if v'_j, v_{j+1} are the two children of v_j ($1 \leq j \leq k-1$), then v'_j is either a leaf or has r children; and
- each subtree of T rooted at a v'_j (if v'_j is a leaf) or at a child of v'_j (if v'_j has r children) is an r-good labelled Schröder tree.

So, for example, a 2-good labelled Schröder tree is exactly a rooted tree with leaves labelled from $\mathbb{N}$ in which all non-leaf vertices have down-degree 2. An *r-good labelled Schröder forest* is a forest whose leaves are labelled in which each component is an r-good labelled Schröder tree. For $n, k, r \geq 1$, let $\mathcal{F}^{\text{good}}_{[r]}(n,k)$ be the set of those forests in $\mathcal{F}_{[r]}(n,k)$ that are r-good.

Theorem 1.3. *For all even $r \geq 2$ and all $n, k \geq 1$ we have*

$$(-1)^{n-k}\left\{ {n \atop k} \right\}^{-1}_{[r]} = \left|\mathcal{F}^{\text{good}}_{[r]}(n,k)\right|.$$

Note that the conclusion of Theorem 1.3 also (vacuously) holds when $r = 1$. In this case $\left\{ {n \atop k} \right\}_{[1]} = \left\{ {n \atop k} \right\}_{[1]}^{-1} = \left| \mathcal{F}_{[1]}^{\text{good}}(n,k) \right| = \mathbf{1}_{\{n=k\}}$ as the only 1-good tree is an isolated vertex.

As we will see in the sequel, for even r the exponential generating function

$$g_{r,1}(x) = \sum_{n=1}^{\infty} \left| \mathcal{F}_{[r]}^{\text{good}}(n,1) \right| \frac{x^n}{n!}$$

(of the sequence of counts of r-good trees by number of vertices) satisfies the functional equation

$$\sum_{j=1}^{r} (-1)^{j-1} \frac{(g_{r,1}(x))^j}{j!} = x,$$

making terms of these sequences very easy to calculate. For $r = 2$, the sequence begins $(1, 1, 3, 15, 105, 945, 10395, \ldots)$, and is the well-known sequence of double factorials of odd numbers [9, A001147]. For $r = 4$ it begins $(1, 1, 2, 6, 25, 140, 1015, \ldots)$, and for $r = 6$ it begins $(1, 1, 2, 6, 24, 120, 721, 5075, \ldots)$; neither of these sequences appear in [9].

We can generalize Theorems 1.2 and 1.3 considerably. We begin with Theorem 1.2, which we generalize to an interpretation of $\left\{ {n \atop k} \right\}_R^{-1}$ as the difference in sizes between two combinatorially defined sets, for any $R \subseteq \mathbb{N}$ with $1 \in R$ (a condition necessary and sufficient to ensure that $\left[\left\{ {n \atop k} \right\}_R \right]_{n,k \geq 1}$ is invertible). For such an R, and for $n, k \geq 1$, denote by $\mathcal{F}_R(n,k)$ the set of labelled Schröder forests with label set $[n]$ and with n leaves, k components, and all non-leaf down-degrees in R. Denote by $\mathcal{F}_R^{\text{even}}(n,k)$ those forests in $\mathcal{F}_R(n,k)$ in which there are an even number of components that have an odd number of non-leaf vertices, and denote by $\mathcal{F}_R^{\text{odd}}(n,k)$ those in which there are an odd number of components that have an odd number of non-leaf vertices.

Theorem 1.4. *For all $R \subseteq \mathbb{N}$ with $1 \in R$ and all $n, k \geq 1$ we have*

$$\left\{ {n \atop k} \right\}_R^{-1} = \left| \mathcal{F}_R^{\text{even}}(n,k) \right| - \left| \mathcal{F}_R^{\text{odd}}(n,k) \right|.$$

We turn now to the generalization of Theorem 1.3. For $d \geq 1$ let $h_d : \mathbb{N} \to \mathbb{N}$ be the map that sends s to $(s-1)d + 1$, and for $R \subseteq \mathbb{N}$ let $R(d) = \{h_d(s) : s \in R\}$. We think of $R(d)$ as a copy of R that has been "stretched" along the arithmetic progression $\{1, d+1, 2d+1, \ldots\}$. We can extend the notion of an r-good labelled Schröder forest (which corresponds to the case $R = \{1, \ldots, r\}$, $d = 1$) to arbitrary R and d, to create a notion of an $R(d)$-good labelled Schröder forest. The details are easy but lengthy, so we defer them to Section 4. We just mention the detail here that such structures with n leaves and k components will only exist when $d | (n-k)$. We define $\mathcal{F}_{R(d)}^{\text{good}}(n,k)$ to be the set of those forests in $\mathcal{F}_{R(d)}(n,k)$ that are $R(d)$-good.

Say that $R \subseteq \mathbb{N}$ has *no exposed odds* if whenever $n \in R$ is odd, we have $\{n-1, n+1\} \cap \mathbb{N} \subseteq R$. Note that $\{1, \ldots, r\}$ has no exposed odds whenever r is even.

Theorem 1.5. *For all $R \subseteq \mathbb{N}$ with $1 \in R$ and with no exposed odds and all $n, k \geq 1$ we have*

$$(-1)^{(n-k)/d} \left\{ {n \atop k} \right\}_{R(d)}^{-1} = \left| \mathcal{F}_{R(d)}^{\text{good}}(n,k) \right|.$$

For $d \geq 1$ the *hyperbolic function of order d of the first kind* (see for example [11]) is the function $H_{d,1}(z)$ defined by the power series

$$H_{d,1}(x) = \sum_{\ell \geq 0} \frac{x^{d\ell+1}}{(d\ell+1)!};$$

so for example $H_{1,1}(x) = e^x - 1$ and $H_{2,1}(x) = \sinh x$. The study of these functions goes back to the mid-1700's. As an immediate by-product of our proof of Theorem 1.5 we obtain combinatorial interpretations for the coefficients of the compositional inverses of these functions.

Corollary 1.6. *For $d \geq 1$, let $h_{d,1}(x)$ be the compositional inverse of $H_{d,1}(x)$ (satisfying $h_{d,1}(H_{d,1}(x))) = H_{d,1}(h_{d,1}(x))) = x$ for all x). Then writing $h_{d,1}(x)$ in the form*

$$h_{d,1}(x) = \sum_{\ell \geq 0} h_\ell \frac{x^{d\ell+1}}{(d\ell+1)!}$$

we have that $(-1)^\ell h_\ell$ is the number of $\mathbb{N}(d)$-good labelled Schröder trees with $d\ell+1$ leaves.

The set of $\mathbb{N}(d)$-good labelled Schröder trees with $d\ell + 1$ leaves is fairly straightforward to describe (inductively): the unique tree with one leaf ($\ell = 0$) is $\mathbb{N}(d)$-good, and for $\ell > 0$ a tree is $\mathbb{N}(d)$-good if

- along the path from the root to the largest leaf, all non-leaf vertices have $d+1$ children;
- at each such vertex v, there is a child v' that is a leaf with label smaller than that of at least one other leaf below each of the other descendants of v; and
- at each such vertex v, the subtrees rooted at each of its children are $\mathbb{N}(d)$-good.

An $\mathbb{N}(1)$-good labelled Schröder tree with $\ell + 1$ leaves thus consists of a path from a root to a leaf with label $\ell + 1$, with each vertex along the path (other than the leaf labelled $\ell + 1$) having one other child that is also a leaf. There are clearly $\ell!$ such structures, and $\sum_{\ell \geq 0} (-1)^\ell \ell! x^{\ell+1}/(\ell+1)! = \log(1+x)$, which is indeed the compositional inverse of $e^x - 1$.

For $d = 2$ the sequence (starting from $\ell = 0$) whose ℓth term is the number of $\mathbb{N}(2)$-good labelled Schröder trees with $2\ell + 1$ leaves begins $(1, 1, 9, 225, 11025, 893025, 108056025, \ldots)$, and is the sequence of squares of double factorials of odd numbers [9, A001818] (and it is well-known that this sequence arises in the power series of the inverse of the hyperbolic sine function). For $d = 3$ it begins $(1, 1, 34, 5446, 2405116, 2261938588, 3887833883752, \ldots)$; this does not appear in [9].

Now we turn to restricted Lah numbers. The *Lah number* $L(n, k)$ is the number of ways of partitioning a set of size n into k non-empty lists, or ordered sets, and for $R \subseteq \mathbb{N}$ the *R-restricted Lah number* $L(n, k)_R$ is the number of partitions of a set of size n into k non-empty lists with all list sizes belonging to R (so $L(n, k)_{\mathbb{N}} = L(n, k)$). Belbachir and Bousbaa [2] have studied r-associated Lah numbers, corresponding to $R = \{r, r+1, r+2, \ldots\}$. Note that if $1 \in R$ then $[L(n, k)_R]_{n,k \geq 1}$ is invertible, as it has integer entries and 1's on the diagonal.

Notation. *Let* $L(n,k)_R^{-1}$ *denote the* (n,k) *entry of the inverse matrix of* $[L(n,k)_R]_{n,k\geq 1}$.

Question 1.7. *Is there (up to sign) a combinatorial interpretation for* $L(n,k)_R^{-1}$*?*

In the case $R=\mathbb{N}$ an answer to Question 1.7 is quite classical — it is well-known that $(-1)^{n-k}L(n,k)_{\mathbb{N}}^{-1} = L(n,k)$. As with Stirling numbers of the second kind we can give an interpretation of $L(n,k)_R^{-1}$ for all $R\subseteq\mathbb{N}$ with $1\in R$ as a difference between the sizes of two sets, in this case sets of labelled Schröder forests in which linear orders are given to the children of each non-leaf vertex.

An *l.o.c. labelled Schröder forest* is a labelled Schröder forest in which each non-leaf vertex is endowed with a linear order on its set of children; here "l.o.c." stands for "linearlly ordered children". Note that this makes each component a plane tree; we eschew this more standard language here and later because when we move from trees to forests in our structures, we do not wish to impose an order on the components. For each $R\subseteq\mathbb{N}$ with $1\in R$ and each $n,k\geq 1$, denote by $\mathcal{F}_R^{\text{l.o.c.}}(n,k)$ the set of l.o.c. labelled Schröder forests with label set $[n]$ and with n leaves, k components, and all non-leaf down-degrees in R. Denote by $\mathcal{F}_R^{\text{l.o.c., even}}(n,k)$ those forests in $\mathcal{F}_R^{\text{l.o.c.}}(n,k)$ in which there are an even number of components that have an odd number of non-leaf vertices, and denote by $\mathcal{F}_R^{\text{l.o.c., odd}}(n,k)$ those in which there are an odd number of components that have an odd number of non-leaf vertices.

Theorem 1.8. *For* $R\subseteq\mathbb{N}$ *with* $1\in R$ *and all* $n,k\geq 1$ *we have*

$$L(n,k)_R^{-1} = \left|\mathcal{F}_R^{\text{l.o.c., even}}(n,k)\right| - \left|\mathcal{F}_R^{\text{l.o.c., odd}}(n,k)\right|.$$

For R with $1\in R$ and no exposed odds, and sets of the form $R(d)$ for those R, we again obtain a more satisfying interpretation of $L(n,k)_R^{-1}$ as a count of a single set of forests. To state our results, we introduce the structure of trees with linearly ordered children.

An *l.o.c. tree* is a rooted tree in which an ordering is specified for the children of each vertex, and an *l.o.c. forest* is a forest in which each component is an l.o.c. tree. Given a sequence $\mathcal{D}=(D_1,D_2,\ldots)$ of sets, a $\mathcal{D}$-decorated l.o.c. forest is one in which each non-leaf vertex v has associated with it an element from $D_{d(v)}$ (where recall $d(v)$ indicates the number of children, or down-degree, of v). A *labelled* l.o.c. tree or forest ($\mathcal{D}$-decorated or otherwise) is one in which every vertex receives a label; note that for Schröder structures we only labelled the leaves.

Let $\mathcal{G}_{\mathcal{D}}^{\text{l.o.c.}}(n,k)$ be the set of all labelled $\mathcal{D}$-decorated l.o.c forests with label set $[n]$ and with n vertices and k components. (We chose to use $\mathcal{G}$ for this set of forests, reserving $\mathcal{F}$ exclusively for notation associated with Schröder forests.) In many cases $\mathcal{G}_{\mathcal{D}}^{\text{l.o.c.}}(n,k)$ takes a particularly appealing form. For example if $(d_1,d_2,\ldots)$ happens to be a 0-1 sequence then $\mathcal{G}_{\mathcal{D}}^{\text{l.o.c.}}(n,k)$ is simply the set of l.o.c. forests in which the down-degree of each non-leaf vertex is some index in the support of $(d_1,d_2,\ldots)$ (viewed as a function from $\{1,2,\ldots\}$ to $\{0,1\}$).

Definition 1.9. *Given* $R\subseteq\mathbb{N}$ *define the sequence* $d(R)=(d_1,d_2,\ldots)$ *by setting*

$$\frac{x}{\sum_{r\in R}x^n} = 1+\sum_{n\geq 1}(-1)^n d_n x^n$$

and denote by $\mathcal{D}(R)$ *a sequence* $(D_1,D_2,D_3,\ldots)$ *of sets with* $|D_n|=d_n$ *for each* n.

Theorem 1.10. *Fix $R \subseteq \mathbb{N}$ with $1 \in R$ and with no exposed odds. Then $d(R)$ is a sequence of non-negative integers and for each $n, k \geq 1$,*

$$(-1)^{n-k} L(n,k)^{-1}_R = \left|\mathcal{G}^{\text{l.o.c.}}_{\mathcal{D}(R)}(n,k)\right| .$$

After the proof of Theorem 1.10 some special cases are highlighted. We mention just one here.

Corollary 1.11. *For $r \geq 2$ even, $d([r]) = (\mathbf{1}_{\{n \equiv 0 \text{ or } 1 \pmod r\}})_{n \geq 1}$ and*

$$(-1)^{n-k} L(n,k)^{-1}_{[r]} = \left|\mathcal{G}^{\text{l.o.c.}}_{\mathcal{D}([r])}(n,k)\right| .$$

Notice that (informally) in the limit as r goes to infinity the restriction placed on l.o.c. trees here approaches the restriction that each non-leaf vertex has exactly one child; thus we recover the classical result that the (n,k) entry of the matrix $[L(n,k)]_{n,k \geq 1}$ is $(-1)^{n-k} L(n,k)$.

For $r = 2$ the sequence whose nth term (starting from $n = 1$) is the number of (unlabelled) l.o.c. trees with n vertices in which the number of children of each non-leaf vertex is congruent to either 0 or 1 modulo r begins $(1, 1, 2, 5, 14, 42, 132, \ldots)$, and is the sequence of Catalan numbers [9, A000108]; when we move from unlabelled to labelled trees (multiplying the nth term by $n!$) we arrive at the sequence which begins $(1, 2, 12, 120, 1680, 30240, 665280, \ldots)$ [9, A001813]. For $r = 4$ these two sequences begin $(1, 1, 1, 1, 2, 7, 22, \ldots)$ [9, A063019] and $(1, 2, 6, 24, 240, 5040, 110880, \ldots)$ (not in [9]).

We now extend to sets of the form $R(d)$ described earlier.

Theorem 1.12. *Let $R \subseteq \mathbb{N}$ with $1 \in R$ have no exposed odds. For each $d \geq 1$, $d(R(d))$ is supported on the indices that are the multiples of d and*

$$(-1)^{(n-k)/d} L(n,k)^{-1}_{R(d)} = \left|\mathcal{G}^{\text{l.o.c.}}_{\mathcal{D}(R(d))}(n,k)\right| .$$

We will derive the following corollary.

Corollary 1.13. *For $d \geq 1$, if $R = \{1, d+1, 2d+1, \ldots\}$ then $d(R) = (\mathbf{1}_{\{n=d\}})_{n \geq 1}$ and so $(-1)^{(n-k)/d} L(n,k)^{-1}_R$ counts the number of n vertex, k component l.o.c forests labelled with $[n]$ in which all non-leaf vertices have exactly d children.*

Writing $a_{n,d}$ for the number of unlabelled l.o.c. trees with $dn+1$ vertices in which all non-leaf vertices have exactly d children, we have the recurrence $a_{0,d} = 1$ and for $n \geq 1$

$$a_{n,d} = \sum \left\{ a_{n_1,d} a_{n_2,d} \cdots a_{n_d,d} : \text{weak compositions } (n_1, \ldots, n_d) \text{ of } n-1 \right\} .$$

It follows that the ordinary generating function $A_d(x) = \sum_{n \geq 0} a_{n,d} x^n$ of the sequence $(a_{n,d})_{n \geq 0}$ satisfies the functional equation $A_d(x) = 1 + x A_d(x)^d$, so that $a_{n,d}$ is the Fuss-Catalan number $\binom{dn}{n}/(dn+1)$.

Given their common structure, it should be no surprise that Theorems 1.4 and 1.8 are special cases of the same more general result. That result, Corollary 3.3, appears in Section 3. Here we mention one more special case, that of restricted Stirling numbers of the first kind.

Let $\begin{bmatrix} n \\ k \end{bmatrix}$ denote the number of partitions of a set of size n into k non-empty cycles; this is a *Stirling number of the first kind*. For $R \subseteq \mathbb{N}$ define the *R-restricted Stirling number of the first kind* $\begin{bmatrix} n \\ k \end{bmatrix}_R$ to be the number of partitions of a set of size n into k non-empty cycles with all cycle sizes belonging to R. Note that if $1 \in R$ then as before the matrix $\left[\begin{bmatrix} n \\ k \end{bmatrix}_R\right]_{n,k\geq 1}$ is invertible.

Notation. *Let* $\begin{bmatrix} n \\ k \end{bmatrix}_R^{-1}$ *denote the* (n, k) *entry of the inverse matrix of* $\left[\begin{bmatrix} n \\ k \end{bmatrix}_R\right]_{n,k\geq 1}$.

A *c.o.c. labelled Schröder forest* is a labelled Schröder forest in which a cyclic order is given to the children of each non-leaf vertex; "c.o.c." here stands for "cyclically ordered children". For each $R \subseteq \mathbb{N}$ with $1 \in R$ and each $n, k \geq 1$, define $\mathcal{F}_R^{\text{c.o.c.}}(n, k)$, $\mathcal{F}_R^{\text{c.o.c., even}}(n, k)$ and $\mathcal{F}_R^{\text{c.o.c., odd}}(n, k)$ exactly in analogy with $\mathcal{F}_R^{\text{l.o.c.}}(n, k)$, $\mathcal{F}_R^{\text{l.o.c., even}}(n, k)$ and $\mathcal{F}_R^{\text{l.o.c., odd}}(n, k)$.

Theorem 1.14. *For all* $R \subseteq \mathbb{N}$ *with* $1 \in R$ *and all* $n, k \geq 1$ *we have*

$$\begin{bmatrix} n \\ k \end{bmatrix}_R^{-1} = \left|\mathcal{F}_R^{\text{c.o.c., even}}(n, k)\right| - \left|\mathcal{F}_R^{\text{c.o.c., odd}}(n, k)\right|.$$

Here is how the rest of this note is laid out. The results of Section 2 reduce the problem of understanding the inverse of a matrix A of the type under consideration to that of understanding the compositional inverse of the exponential generating function of the first column of A. Much of this has appeared in various forms in the literature, but we aim to give a self-contained treatment. In Section 3 we present combinatorial interpretations for the compositional inverses of the power series $\sum_{n\geq 1} a_n x^n/n!$ (in terms of labelled Schröder trees) and $\sum_{n\geq 1} a_n x^n$ (in terms of l.o.c. trees and labelled l.o.c. trees). From this we derive Theorems 1.4, 1.8 and 1.14. In Section 4 we apply the results of Section 3 to restricted Stirling numbers of the second kind, and prove Theorem 1.5. In Section 5 we apply them to restricted Lah numbers and prove Theorems 1.10 and 1.12. Finally in Section 6 we discuss some directions in which this work may be taken, and mention some problems.

2 Interpretation of inverse matrix entries

Let $a = (a_1, a_2, \ldots)$ be a sequence with entries in $\mathbb{C}$. For $n, k \geq 1$ set

$$a_{n,k} = \sum \left\{ \prod_{i=1}^{k} a_{|P_i|} : \{P_1, \ldots, P_k\} \text{ a partition of } \{1, \ldots, n\} \text{ into non-empty blocks} \right\}. \quad (1)$$

If $a = (1, 1, \ldots)$ then $a_{n,k} = \left\{ n \atop k \right\}$, a Stirling number of the second kind; if $a = ((n-1)!)_{n\geq 1}$ then $a_{n,k} = \begin{bmatrix} n \\ k \end{bmatrix}$, a Stirling number of the first kind; and if $a = (n!)_{n\geq 1}$ then $a_{n,k} = L(n, k)$, a Lah number.

As long as $a_1 \neq 0$ (as will always be the case in this note), the doubly infinite matrix

$$A_a := [a_{n,k}]_{n,k\geq 1}$$

is lower-triangular with non-zero entries down the diagonal, so it has an inverse A_a^{-1} which is also lower-triangular with non-zero entries down the diagonal. Moreover if $a_1 = 1$ and if

$a_i \in \mathbb{Z}$ for all i (as will mostly be the case in this note) then A_a is an integer matrix with 1's down the diagonal, and so by Cramer's rule A_a^{-1} is also an integer matrix with 1's down the diagonal.

The main aim of this section is to lay the groundwork for combinatorial interpretations of the entries A_a^{-1}. We begin by showing that A_a is determined by the exponential generating function of a in a straightforward way. This is essentially the exponential formula, but for completeness we provide a proof. Here and throughout, for a sequence $a = (a_1, a_2, \ldots)$ with $a_1 \neq 0$ and with all entries in $\mathbb{C}$ we denote by $E_a(x)$ the exponential generating function of a; that is,

$$E_a(x) = \sum_{j=1}^{\infty} a_j \frac{x^j}{j!}.$$

Lemma 2.1. *Let $a_{n,k}$ be defined as in (1). Then for $k \geq 1$, the exponential generating function of $(a_{n,k})_{n\geq 1}$ is $E_a^k(x)/k!$.*

Proof. We have

$$\begin{aligned} \frac{E_a^k(x)}{k!} &= \frac{1}{k!}\sum_{n\geq k}\left[\sum\left\{\prod_{i=1}^{k}\frac{a_{\ell_i}}{\ell_i!} : \text{compositions } (\ell_1,\ldots,\ell_k) \text{ of } n\right\}\right] x^n \\ &= \sum_{n\geq k}\left[\frac{1}{k!}\sum\left\{\binom{n}{\ell_1,\ldots,\ell_k}\prod_{i=1}^{k} a_{\ell_i} : \text{compositions } (\ell_1,\ldots,\ell_k) \text{ of } n\right\}\right]\frac{x^n}{n!}. \end{aligned}$$

Now from (1) we have

$$k!a_{n,k} = \sum\left\{\binom{n}{\ell_1,\ldots,\ell_k}\prod_{i=1}^{k} a_{\ell_i} : \text{compositions } (\ell_1,\ldots,\ell_k) \text{ of } n\right\},$$

so that

$$\frac{E_a^k(x)}{k!} = \sum_{n\geq k} a_{n,k}\frac{x^n}{n!}$$

as claimed. □

A formal power series $F(x)$ over $\mathbb{C}$ with zero constant term and non-zero linear term has a unique *reversion*, that is, there is a unique formal power series over $\mathbb{C}$, which we denote by $F^{<-1>}(x)$, satisfying $F(F^{<-1>}(x)) = F^{<-1>}(F(x)) = x$. For a sequence a, here and throughout we denote by $b = (b_1, b_2, \ldots)$ the sequence whose exponential generating function is $E_a^{<-1>}(x)$; that is, we define b via

$$E_a^{<-1>}(x) = \sum_{j=1}^{\infty} b_j \frac{x^j}{j!}.$$

We now show that A_a^{-1} is determined by b in a straightforward way; indeed, A_a^{-1} is exactly the matrix $A_b = [b_{n,k}]_{n,k\geq 1}$ that is produced from b via (1) (with $a_{n,k}$ and a_i replaced by $b_{n,k}$ and b_i in that formula). This is an easy application of the theory of exponential Riordan arrays (see for example [1]), but again for completeness we provide a proof.

Lemma 2.2. *With the notation as above,*

$$A_a^{-1} = A_b.$$

Proof. For $k \geq 1$ we have

$$\frac{E_a^k(E_a^{<-1>}(x))}{k!} = \frac{x^k}{k!}. \qquad (2)$$

But also, since $E_a^k(x)/k!$ is the exponential generating function of $(a_{n,k})_{n\geq 1}$, we have

$$\frac{E_a^k(E_a^{<-1>}(x))}{k!} = \sum_{n\geq 1} a_{n,k}\frac{(E_a^{<-1>}(x))^n}{n!}.$$

For $\ell \geq 1$,

$$[x^\ell]\left(\frac{(E_a^{<-1>}(x))^n}{n!}\right) = \frac{b_{\ell,n}}{\ell!},$$

so

$$[x^\ell]\left(\frac{E_a^k(E_a^{<-1>}(x))}{k!}\right) = \frac{1}{\ell!}\sum_{n\geq 1} b_{\ell,n}a_{n,k}.$$

Since by (2) this coefficient is either 0 (if $\ell \neq k$) or $1/\ell!$ (if $\ell = k$), we get that

$$\sum_{n\geq 1} b_{\ell,n}a_{n,k} = \begin{cases} 1 & \text{if } \ell = k, \\ 0 & \text{otherwise;} \end{cases}$$

in other words A_bA_a is the identity, as required. □

From here on, for a sequence a we denote by $b_{n,k}$ the number produced from b via (1) (with $a_{n,k}$ and a_i replaced by $b_{n,k}$ and b_i in that formula).

As an illustration of Lemma 2.2, consider the sequence $a = (1, 1, \ldots)$. Here the (n, k) entry of A_a is the Stirling number of the second kind $\left\{ {n \atop k} \right\}$. We have $E_a(x) = \sum_{\ell\geq 1} x^\ell/\ell! = e^x - 1$, a power series whose reversion is $E_a^{<-1>}(x) = \ln(1 + x) = \sum_{\ell\geq 1}(-1)^{\ell-1}x^\ell/\ell$, so that $b = ((-1)^{n-1}(n-1)!)_{n\geq 1}$. With $\{P_1, \ldots, P_k\}$ running over partitions of $\{1, \ldots, n\}$ into k non-empty blocks in both sums below, it follows that the (n, k) entry of A_a^{-1} is

$$\begin{aligned} b_{n,k} &= \sum \prod_{i=1}^{k} (-1)^{|P_i|-1}(|P_i| - 1)! \\ &= \sum (-1)^{|P_1|+\ldots+|P_k|-k} \prod_{i=1}^{k} (|P_i| - 1)! \qquad (3) \\ &= (-1)^{n-k}\left[{n \atop k} \right]. \end{aligned}$$

Similarly we can recover from Lemma 2.2 that if A is the matrix whose (n, k) entry is $\left[{n \atop k} \right]$ then the (n, k) entry of A^{-1} is $(-1)^{n-k}\left\{ {n \atop k} \right\}$, and if A is the matrix whose (n, k) entry is $L(n, k)$ then the (n, k) entry of A^{-1} is $(-1)^{n-k}L(n, k)$.

Notice that in the chain of equalities (3) we have that the sign of

$$\prod_{i=1}^{k} b_{|P_i|} = \prod_{i=1}^{k} (-1)^{|P_i|-1}(|P_i|-1)!,$$

namely $(-1)^{|P_1|+\ldots+|P_k|-k} = (-1)^{n-k}$, depends only only n and k. This useful observation allows us to conclude that

$$b_{n,k} = (-1)^{n-k} \sum \prod_{i=1}^{k} |b_{|P_i|}|,$$

and motivates the following definition.

Definition 2.3. *A sequence $(u_1, u_2, \ldots)$ of reals is* sign coherent *if there is a sign function $c : \mathbb{N} \times \mathbb{N} \to \{-1, 1\}$ with the following property: for each n and k, and for each composition $(i_1, \ldots, i_k)$ of n for which $u_{i_j} \neq 0$ for each j, the sign of $\prod_{j=1}^{k} u_{i_j}$ is $c(n, k)$ (independent of the choice of composition).*

As an example, if the u_i's are alternating (with u_1 positive) then the chain of equalities (3) can easily be modified to show that $(u_1, u_2, \ldots)$ is sign coherent with sign function $c(n, k) = (-1)^{n-k}$.

If a is a sequence such that the sequence b is sign coherent, say with sign function c, then a combination of Lemmas 2.1 and 2.2 yields

$$b_{n,k} = c(n, k) \sum \prod_{i=1}^{k} |b_{|P_i|}|,$$

with $\{P_1, \ldots, P_k\}$ running over partitions of $\{1, \ldots, n\}$ into k non-empty blocks in the sums. This leads to the following, which drives much of the remainder of this note.

Corollary 2.4. *Let a be a sequence of integers with $a_1 = 1$. Suppose that b is sign coherent, with sign function c, and suppose further that there is a collection $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \ldots\}$ of sets, with $|\mathcal{B}_j| = |b_j|$ for each j. Then $c(n, k)b_{n,k}$ may be combinatorially interpreted as follows: it counts the number of ways of partitioning $\{1, \ldots, n\}$ into k non-empty blocks $\{P_1, \ldots, P_k\}$ and associating with each block P_i an element of $\mathcal{B}_{|P_i|}$.*

In the absence of sign coherence, we have the following quasi-combinatorial interpretation of $b_{n,k}$ as the difference between the sizes of two sets. Let $\{P_1, \ldots, P_k\}$ be a partition of $\{1, \ldots, n\}$ into k non-empty blocks such that $b_{|P_i|} \neq 0$ for all i. Say that the partition is *positive* if there are an even number of blocks P_i with $b_{|P_i|}$ negative, and say that it is *negative* otherwise. Then $b_{n,k}$ counts the number of ways of partitioning $\{P_1, \ldots, P_k\}$ into a positive partition and associating with each block P_i an element of $\mathcal{B}_{|P_i|}$, minus the same count for negative partitions. We explore this idea in more detail at the end of Section 3.

We end this section by establishing a condition that ensures sign coherence of a sequence. Specializing to $d = 1$ we recover our observation that alternating sequences are sign coherent.

Lemma 2.5. *Let $(u_1, u_2, \ldots)$ be a sequence of reals. Suppose that for some $d \geq 1$, for all m*

$$\mathrm{sign}(u_m) = \begin{cases} 1 & \text{if } m = ds+1,\ s \text{ even,} \\ -1 & \text{if } m = ds+1,\ s \text{ odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Then $(u_1, u_2, \ldots)$ is sign coherent, and any function c satisfying $c(n,k) = (-1)^{(n-k)/d}$ whenever d divides $n-k$ is a valid sign function.

Proof. We need only consider pairs (n,k) for which $n = ds + k$ for some integer $s \geq 0$ (for all other pairs there is no composition $(i_1, \ldots, i_k)$ of n with $\prod_{j=1}^k u_{i_j} \neq 0$). For such a pair (n,k), consider the composition $(ds_1+1, \ldots, ds_k+1)$ of n (again, for all other compositions $(i_1, \ldots, i_k)$ of n we have $\prod_{j=1}^k u_{i_j} = 0$). Noting that $n - k = d\sum_{i=1}^k s_i$ we have

$$\prod_{j=1}^k \mathrm{sign}(u_{ds_j+1}) = (-1)^{\sum_{j=1}^k s_j} = (-1)^{(n-k)/d}.$$

□

We will only ever consider sign coherence of sequences $(u_1, u_2, \ldots)$ for which $u_1 > 0$ and there is some $d \geq 1$ such that the support of $(u_1, u_2, \ldots)$ (viewed as a function from $\mathbb{N}$) is $\{1, d+1, 2d+1, \ldots\}$. For these sequences it is fairly easy to establish that in fact $(u_1, u_2, \ldots)$ is sign coherent if and only if it satisfies the hypothesis of Lemma 2.5.

3 Remarks on the reversion

Section 2 has established the important role that power series reversion plays in understanding A_a^{-1}. In this section we give two results that furnish combinatorial interpretations of the coefficients of the series reversion. The first will be relevant for understanding inverse restricted Lah numbers; the second for inverse restricted Stirling numbers of the second kind.

Let $F(x)$ be a formal power series in a single variable x over $\mathbb{C}$. Recall that if $F(x)$ has zero constant term but non-zero linear term then we say that $F(x)$ is *reversible*, and denote by $F^{<-1>}(x)$ the unique formal power series in x over $\mathbb{C}$ satisfying $F(F^{<-1>}(x)) = F^{<-1>}(F(x)) = x$. Also denote by $G(x)$ the unique formal power series in x over $\mathbb{C}$ satisfying $G(x) = x/F(x)$.

Recall that a *l.o.c. tree* is an unlabelled tree with a distinguished root vertex in which an ordering is given to the children of each vertex v. Denote by $d(v)$ the number of children (down-degree) of vertex v. Given a sequence $w = (w_0, w_1, \ldots)$ of complex numbers, for a l.o.c. tree T the w-weight of T is defined by

$$w(T) = \prod_{v \in V(T)} w_{d(v)}$$

(here considering a leaf to have 0 children). Denote by $\mathcal{T}^{\text{l.o.c.}}(n)$ the set of l.o.c. trees on n vertices. Proposition 3.1 below is contained in the second proof of the Lagrange inversion formula given in [10, Chapter 5]; for completeness we provide a self-contained proof.

Proposition 3.1. *If $F(x)$ has zero constant term and non-zero linear term, and if the sequence $w = (w_0, w_1, \ldots)$ is defined via $G(x) = x/F(x) = \sum_{n\geq 0} w_n x^n$ then setting*

$$f_n^{<-1>} = \sum_{T \in \mathcal{T}^{\text{l.o.c.}}(n)} w(T)$$

for $n \geq 1$ we have $F^{<-1>}(x) = \sum_{n\geq 1} f_n^{<-1>} x^n$.

Proof. Set $B(x) = \sum_{n\geq 1} f_n^{<-1>} x^n$; we aim to show $F(B(x)) = x$. For $n \geq 2$, an l.o.c. tree T can be uniquely encoded by a tuple $(e, T_1, T_2, \ldots, T_e)$, where e represents the number of children of the root, and takes a value between 1 and $n-1$, and T_ℓ (for $1 \leq \ell \leq e$) is the l.o.c. tree whose root is the ℓth child of the root of T. Note that the vector $(|T_1|, \ldots, |T_e|)$ forms a composition of $n-1$. From the definition of $w(T)$ we have

$$w(T) = w_e \prod_{i=1}^{e} w(T_i),$$

and so for $n \geq 2$ we have

$$f_n^{<-1>} = \sum_{1\leq e\leq n-1} w_e \sum \left\{ \prod_{i=1}^{e} f_{m_i}^{<-1>} : \text{compositions } (m_1, \ldots, m_e) \text{ of } n-1 \right\}$$

while $f_1^{<-1>} = w_0$. This says that for all $n \geq 1$ the coefficient of x^n in $xG(B(x))$ is $f_n^{<-1>}$, and so $B(x) = xG(B(x))$. But now from the definition of $G(x)$ we have

$$F(B(x)) = \frac{B(x)}{G(B(x))} = x.$$

It follows that $B(x) = F^{<-1>}(x)$, as required. ☐

An immediate corollary is that

$$n! f_n^{<-1>} = \sum_{T \in \mathcal{T}^{\text{labelled l.o.c.}}(n)} w(T) \tag{4}$$

where $\mathcal{T}^{\text{labelled l.o.c.}}(n)$ is the set of labelled l.o.c. trees on vertex set $\{1, \ldots, n\}$; this is because each unlabelled tree gives rise to $n!$ labelled ones. The quantity $n! f_n^{<-1>}$ will be of more immediate relevance than $f_n^{<-1>}$ in Section 5.

Recall that a *Schröder tree* is a rooted tree (a tree with a distinguished root vertex) in which no vertex has exactly one child. Given a sequence $a = (a_1, a_2, \ldots)$ of complex numbers with $a_1 \neq 0$, for a Schröder tree T with n leaves and $m+n$ vertices the a-weight of T is defined by

$$a(T) = (-1)^m a_1^{-(m+n)} \prod \left\{ a_{d(v)} : v \in V(T),\ v \text{ not a leaf} \right\}$$

(recall here that an isolated vertex is considered a leaf). Denote by $\mathcal{T}^{\text{labelled Schröder}}(n)$ the set of Schröder trees with n leaves labelled by some set of size n. A result similar to Proposition 3.2 below appears in [3, Theorem 5.2], although in that reference Chen is working with a different family of trees that he also refers to as Schröder trees.

Proposition 3.2. *If $F(x) = \sum_{n\geq 1} a_n x^n/n!$ (with $a_1 \neq 0$) then setting*

$$f_n^{<-1>} = \sum_{T\in\mathcal{T}^{\text{labelled Schröder}}(n)} a(T)$$

for $n \geq 1$ we have $F^{<-1>}(x) = \sum_{n\geq 1} f_n^{<-1>} x^n/n!$.

Proof. Since $F(F^{<-1>}(x)) = x$ we have

$$\sum_{n=1}^{\infty} a_n \frac{(F^{<-1>}(x))^n}{n!} = x.$$

Comparing coefficients of x we get that

$$[x]F^{<-1>}(x) = a_1^{-1} = \sum \left\{a(T) : \text{labelled Schröder trees } T \text{ with one leaf}\right\}.$$

We now proceed by induction on n. For $n \geq 2$, comparing coefficients of x^n on both sides of $F(F^{<-1>}(x)) = x$ we get

$$[x^n]F^{<-1>}(x) = -a_1^{-1} \sum_{k=2}^{n} a_k \left(\sum_{(i_1,\ldots,i_k)} \frac{1}{k!} \binom{n}{i_1,\ldots,i_k} \prod_{j=1}^{k} f_{i_j}^{<-1>} \right), \tag{5}$$

where $\sum_{(i_1,\ldots,i_k)}$ is a sum over compositions $(i_1,\ldots,i_k)$ of n. We obtain a labelled Schröder tree with n labelled leaves by first choosing k, the number of children of the root, with k varying between 2 and n. We then choose an unordered partition of $\{1,\ldots,n\}$ (the set of leaf labels), with each block being the set of leaf labels used on the leaves that are descendants of the same child of the root. For each composition $(i_1,\ldots,i_k)$ of n there are $\frac{1}{k!}\binom{n}{i_1,\ldots,i_k}$ such partitions in which the block sizes are $i_1,\ldots,i_k$. Finally, we choose labelled Schröder trees of appropriate sizes and with appropriate labels to be rooted at each of the children of the root.

The a-weight of the labelled Schröder tree thus constructed is the product of the a-weights of the chosen labelled Schröder trees, times $-a_k/a_1$ (to account for the root, which has k children, is an additional vertex, and moreover is an additional non-leaf vertex). Summing over all choices, we get that

$$\sum_{T\in\mathcal{T}^{\text{labelled Schröder}}(n)} a(T) = -a_1^{-1} \sum_{k=2}^{n} a_k \left(\sum_{(i_1,\ldots,i_k)} \frac{1}{k!} \binom{n}{i_1,\ldots,i_k} \prod_{j=1}^{k} \left(\sum_{T_{i_j}} a(T_{i_j}) \right) \right) \tag{6}$$

where $\sum_{T_{i_j}}$ is a sum over $\mathcal{T}^{\text{labelled Schröder}}(i_j)$ and $\sum_{(i_1,\ldots,i_k)}$ is, as before, a sum over compositions $(i_1,\ldots,i_k)$ of n. By induction the right-hand side of (6) above is

$$-a_1^{-1} \sum_{k=2}^{n} a_k \left(\sum_{(i_1,\ldots,i_k)} \frac{1}{k!} \binom{n}{i_1,\ldots,i_k} \prod_{j=1}^{k} f_{i_j}^{<-1>} \right),$$

which by (5) is $f_n^{<-1>}$, so

$$f_n^{<-1>} = \sum_{T\in\mathcal{T}^{\text{labelled Schröder}}(n)} a(T)$$

as claimed. ∎

We end this section with a discussion of Theorems 1.4, 1.8 and 1.14, all of which follow quickly from the following corollary of Proposition 3.2.

Let $a = (a_1, a_2, a_3, \ldots)$ be a sequence of non-negative numbers with $a_1 \neq 0$ and with support R. Let $a_{n,k}$ be as in (1), and as before let $b_{n,k}$ be the (n,k) entry of the inverse of the matrix $[a_{n,k}]_{n,k\geq 1}$. Let $\mathcal{F}_R(n,k)$, $\mathcal{F}_R^{\text{even}}(n,k)$ and $\mathcal{F}_R^{\text{odd}}(n,k)$ be as defined just before the statement of Theorem 1.4.

Corollary 3.3. *For all $n, k \geq 1$ we have*

$$b_{n,k} = \sum_{F \in \mathcal{F}_R^{\text{even}}(n,k)} |a(F)| - \sum_{F \in \mathcal{F}_R^{\text{odd}}(n,k)} |a(F)|$$

where for a forest F with components $T_1, \ldots, T_k$, $a(F) = \prod_{i=1}^k a(T_i)$.

Proof. Let $\mathcal{T}_R(n)$ be the collection of labelled Schröder trees with n leaves in which every non-leaf vertex v has $d(v) \in R$. For $T \in \mathcal{T}_R(n)$ denote by $\text{int}(T)$ the number of non-leaf vertices in T. Recalling that here $a_i \geq 0$, an immediate application of Proposition 3.2 yields that if $(b_1, b_2, \ldots)$ is the sequence whose exponential generating function is the series reversion of $\sum_{n \in R} a_n x^n/n!$ then

$$b_n = \sum_{T \in \mathcal{T}_R(n)} (-1)^{\text{int}(T)} |a(T)|.$$

But now the results of Section 2 tell us that $b_{n,k}$ is a weighted sum of elements of $\mathcal{F}_R(n,k)$, where the weight of an $F \in \mathcal{F}_R(n,k)$ whose components have $m_1, \ldots, m_k$ non-leaf vertices is evidently $(-1)^{m_1+\ldots+m_k}|a(F)|$. This is $+1$ if an even number of the m_i's are odd (so $F \in \mathcal{F}_R^{\text{even}}(n,k)$) and -1 if an odd number of them are odd (so $F \in \mathcal{F}_R^{\text{odd}}(n,k)$). The result follows. □

Theorem 1.4 (all $a_n = 1$) is immediate since $|a(F)| = 1$ always in this case. For Theorem 1.8 ($a_n = n!$), note that each $F \in \mathcal{F}_R(n,k)$ corresponds in a natural way to a collection of $|a(F)|$ forests in $\mathcal{F}_R^{\text{l.o.c.}}(n,k)$ (independently assign a linear order the the children of each non-leaf vertex in F), all with the same collection of down-degrees, and that the resulting collections of forests partition $\mathcal{F}_R^{\text{l.o.c.}}(n,k)$ as F runs over $\mathcal{F}_R(n,k)$; Theorem 1.8 follows. The derivation of Theorem 1.14 is similar.

4 Restricted Stirling numbers — proof of Theorem 1.5

Here we present the proof of Theorem 1.5, which includes our first main result, Theorem 1.3, as a special case. Fix a set R of positive integers that includes 1 and has no exposed odds and fix $d \geq 1$ (the reader may find it helpful to initially consider the special case $d = 1$, which contains most of the novelty of the proof). Throughout this section R and d will remain fixed, and we mostly suppress dependence on these parameters in our notation. It will be useful to have the following alternate characterization of sets R with $1 \in R$ and with no exposed odds: when such an R is written as a union of maximal intervals of consecutive integers, each such interval is of the form $[1, \infty)$, or $[1, b]$ for even b, or $[a, b]$ for even a and b

(with possibly $a = b$), or $[a, \infty)$ for even a. Note that R may be the union of finitely many such intervals, or infinitely many.

We will be considering Schröder trees T with the property that for each non-leaf vertex v of T there is $s_v \in R$ such that $d(v) = d(s_v - 1) + 1$ (in the case $d = 1$, we are just saying that for each non-leaf vertex v, $d(v) \in R$). It is immediate that the number of leaves of T is one more than a multiple of d; specifically, it is $d(\sum_v (s_v - 1)) + 1$ where the sum is over the non-leaf vertices v.

So, for each $\ell = 0, 1, 2, \ldots$, let $\mathcal{T}(\ell)$ be the set of labelled Schröder trees with $d\ell + 1$ leaves in which all non-leaf vertices v have $d(v) = d(s_v - 1) + 1$ for some $s_v \in R$ (so for all $T \in \mathcal{T}(\ell)$ we have $\ell = \sum_v (s_v - 1)$). Refining this, let $\mathcal{T}^E(\ell)$ be the set of $T \in \mathcal{T}(\ell)$ for which the quantity $\sum_v s_v$ is even, and let $\mathcal{T}^O(\ell)$ be the set of those T for which it is odd. When $d = 1$ an easy computation shows that this reduces to $T \in \mathcal{T}^E(\ell)$ if and only if T has an even number of edges in total.

For each non-leaf vertex v of T, its *leaf-label* is the label of the largest leaf among all the leaves that are descendants of v.

We now describe the notion of $R(d)$-goodness alluded to in the introduction. As before, the definition is inductive. The tree consisting of a single vertex labelled with a natural number is $R(d)$-good. For a labelled Schröder tree T with more than one leaf, again leaves labelled from the natural numbers, we say that it is $R(d)$-good if

- it is in $T(\ell)$ for some ℓ;
- along the path $v_1, v_2, \ldots, v_k$ from the root (v_1) to the leaf with largest label (v_k) all vertices v_j (except v_k) have either $d + 1$ or $d(s_{v_j} - 1) + 1$ children where $s_{v_j} = a$ for some $a > 2$ that is at the lower end of one of the intervals that comprise R;
- for v_j ($1 \leq j \leq k - 1$) with $d + 1$ children, if v'_j is the child of v_j with the least leaf-label (note $v'_j \neq v_{j+1}$) then v'_j either is a leaf or has $d(s_{v'_j} - 1) + 1$ children where $s_{v'_j} = b$ for some b that is at the upper end of one of the intervals that comprise R; and
- each subtree of T rooted at a child of a v_j (if v_j has more than $d + 1$ children) or at a child of v_j other than v'_j (if v_j has $d + 1$ children), is an $R(d)$-good labelled Schröder tree, and each subtree of T rooted at a child of a v'_j is an $R(d)$-good labelled Schröder tree.

An *$R(d)$-good labelled Schröder forest* is a forest whose leaves are labelled in which each component is an $R(d)$-good labelled Schröder tree.

As an illustration of this definition, observe that if $R = \{1, 2, \ldots, r\}$ and $d = 1$ then the notions of $[r](1)$-goodness and r-goodness (defined in the introduction) coincide.

As another illustration, consider the case $R = \mathbb{N}$ and $d = 1$. Here an $\mathbb{N}(1)$-good labelled Schröder tree with n leaves consists of a path from the root to a leaf labelled n, with each vertex along this path (other than the terminal leaf) having exactly one other child, which is also a leaf. There are thus $(n - 1)!$ $\mathbb{N}(1)$-good Schröder trees with n leaves. Once Theorem 1.5 has been proven, this example will illustrate the classical inverse relationship between Stirling numbers of the first and second kinds.

Lemma 4.1. *If $R \subseteq \mathbb{N}$ has $1 \in R$ and has no exposed odds then $R(d)$-good labelled Schröder trees in $\mathcal{T}(\ell)$ all lie in $\mathcal{T}^E(\ell)$.*

Proof. We proceed by induction on ℓ. For the base case, there is a unique $T \in \mathcal{T}(0)$, and it has one leaf. It has no non-leaf vertices, so $\sum_v s_v = 0$ (being an empty sum), putting $T \in \mathcal{T}^E(\ell)$; and by definition T is $R(d)$-good. For the induction step, consider an $R(d)$-good labelled Schröder tree $T \in \mathcal{T}(\ell)$ with $\ell > 0$. It has a path $v_1, v_2, \ldots, v_k$ from the root (v_1) to the leaf with largest label (v_k). The contribution from each non-leaf vertex along this path to $\sum_v s_v$ is even (either $s_{v_j} = 2$ if v_j has $d+1$ children, or $s_{v_j} = a$ for some even a otherwise). If v_j has $d+1$ children, v_j' is its child with least leaf-label, and v_j' is not a leaf, then v_j' has $d(b-1)+1$ children for some even b, so the contribution from v_j' to $\sum_v s_v$ is even. By definition of $R(d)$-goodness all other non-leaf vertices of T are non-leaf vertices of some $R(d)$-good labelled Schröder tree with fewer leaves than T, and so by induction the contribution to $\sum_v s_v$ from all of these vertices is a sum of even numbers and so even, putting $T \in \mathcal{T}^E(\ell)$. □

Here is the main result of this section, from which Theorem 1.5 will easily follow using the results of Sections 2 and 3.

Claim 4.2. *If $R \subseteq \mathbb{N}$ has $1 \in R$ and has no exposed odds then for all $\ell \geq 0$ there is an involution $f_\ell : \mathcal{T}(\ell) \to \mathcal{T}(\ell)$ whose set of fixed points is exactly the set of $R(d)$-good labelled Schröder trees with $d\ell + 1$ leaves.*

Proof. We will define, by induction on ℓ, an involution $f_\ell : \mathcal{T}(\ell) \to \mathcal{T}(\ell)$ with the properties that

- all of its fixed points lie in $\mathcal{T}^E(\ell)$;
- the set of fixed points coincides with the set of $R(d)$-good labelled Schröder trees; and
- all orbits of size 2 involve one element each from $\mathcal{T}^E(\ell), \mathcal{T}^O(\ell)$.

The base case of the induction, $\ell = 0$, is trivial, as $\mathcal{T}^O(0)$ is empty and $\mathcal{T}^E(0)$ consists of a single element that is $R(d)$-good.

For $\ell > 0$, consider $T \in \mathcal{T}(\ell)$. Let $v_1, v_2, \ldots, v_k$ be the path from the root (v_1) to the leaf with largest label (v_k).

Case 1: There is some j ($1 \leq j \leq k-1$) such that either

- (subcase 1) v_j has $d+1$ children, and v_j', the child of v_j with the least leaf-label is not a leaf and has $d(s_{v_j'} - 1) + 1$ children where $s_{v_j'} \neq b$ for any b (i.e., $s_{v_j'}$ is not at the upper end of one of the intervals that comprise R)

or

- (subcase 2) v_j has $d(s_{v_j} - 1) + 1$ children where $s_{v_j} > 2$ and $s_{v_j} \neq a$ for any a (i.e., s_{v_j} is not at the lower end of one of the intervals that comprise R).

Choose j to be the least index where one of subcase 1, subcase 2 occurs (note that they are mutually exclusive at a particular vertex). If v_j satisfies the condition of subcase 1, we modify T by the operation of identifying v_j and v_j' (and giving the resulting vertex label v_j).

We refer to this operation as *contraction* at v_j, because it corresponds to the usual graph-theoretic operation of contracting the edge $v_j v_j'$. Observe that the result of this operation is still a labelled Schröder tree, with the same set of leaves as T. See Figure 4.

Moreover, the resulting tree is in $\mathcal{T}^E(\ell)$ if T was in $\mathcal{T}^O(\ell)$, and is in $\mathcal{T}^O(\ell)$ if T was in $\mathcal{T}^E(\ell)$. To see this, note that we have removed two vertices, one with $d+1$ children (v_j) and one with $d(s_{v_j'}-1)+1$ (v_j'), and replaced them with one vertex with $(d(s_{v_j'}-1)+1)+(d+1)-1 = d((s_{v_j'}+1)-1)+1$ children. Since $s_{v_j'}$ is not at the upper end of one of the intervals that comprise R we have $s_{v_j'}+1 \in R$, so that the resulting tree is in $\mathcal{T}(\ell)$. Since the parity of $\sum_v s_v$ determines membership of a tree in $\mathcal{T}^E(\ell)$ and $\mathcal{T}^O(\ell)$ and the effect of this operation is to increase this sum by one, we get the statement about toggling between these two sets.

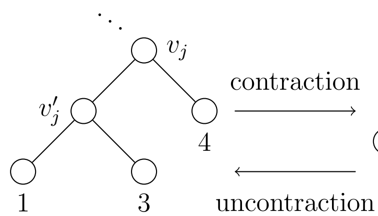

Figure 4: The tree on the left contracts to the one on the right, while the tree on the right uncontracts to the one on the left. Here $d = 1$.

If v_j satisfies the condition of subcase 2, we modify T by the following operation. Locate the d children of v_j that have the largest leaf-labels (v_{j+1} will be among those d). Let L be the set of children that are left, and note that $|L| = d(s_{v_j}-1)+1-d > 1$ as $s_{v_j} > 2$. Delete all edges from v_j to L. Add a new vertex v_j', and make v_j' a child of v_j and the parent of every vertex in L. We refer to this operation as *uncontraction* at v_j. Observe that the result of this operation is still a labelled Schröder tree, with the same set of leaves as T. See Figure 4.

Moreover, the resulting tree is in $\mathcal{T}^E(\ell)$ if T was in $\mathcal{T}^O(\ell)$, and is in $\mathcal{T}^O(\ell)$ if T was in $\mathcal{T}^E(\ell)$. To see this, note that where in T we had a vertex (v_j) with $d(s_{v_j}-1)+1$ children, in the new tree we have two vertices, one with $d+1$ children (v_j) and one with $d(s_{v_j}-1)+1-d = d((s_{v_j}-1)-1)+1$ children (v_j'). Since s_{v_j} is not at the lower end of one of the intervals that comprise R we have $s_{v_j}-1 \in R$, so that the resulting tree is in $\mathcal{T}(\ell)$. Note that, like contraction, uncontraction also changes the parity of $\sum_v s_v$ (by adding one) and we get the statement about toggling between $\mathcal{T}^E(\ell)$, and $\mathcal{T}^O(\ell)$.

Let $f_\ell(T)$ be the tree obtained from T by the process described above. In $f_\ell(T)$ the path from root to vertex of largest degree is $v_1, v_2, \ldots, v_k$, exactly as it was in T. Further, for $i < j$ (where v_j was the vertex of T at which contraction/uncontraction was performed to obtain $f_\ell(T)$) the number of children of v_i remains unchanged from T to $f_\ell(T)$, as does the location of the child of v_i with least leaf-label. Since this is the data that determines whether a contraction/uncontraction is to be performed at v_i, it follows that if the algorithm that is defining f_ℓ is applied to $f_\ell(T)$, it does not call for contraction/uncontraction at v_i for any $i < j$. However, at v_j, if in T we performed a contraction (subcase 1), then the algorithm calls for an uncontraction at v_j in $f_\ell(T)$ (subcase 2), while if in T we performed an uncontraction (subcase 2), then the algorithm calls for a contraction at v_j in $f_\ell(T)$ (subcase 1). Moreover, performing a contraction at v_j in T followed by an uncontraction at v_j in $f_\ell(T)$ returns T; while performing an uncontraction at v_j in T followed by a contraction at v_j in $f_\ell(T)$ also returns T. It follows that for those T currently under consideration, $f_\ell(f_\ell(T)) = T$.

Case 2: If the conditions of Case 1 do not occur, then

- along the path $v_1, v_2, \ldots, v_k$ all vertices v_j (except v_k) have either $d+1$ or $d(s_{v_j}-1)+1$ children where $s_{v_j} = a$ for some $a > 2$ that is a lower endpoint of one of the intervals comprising R; and
- for v_j $(1 \leq j \leq k-1)$ with $d+1$ children, if v'_j is the child of v_j with the least leaf-label then v'_j either is a leaf or has $d(s_{v'_j}-1)+1$ children where $s_{v'_j} = b$ for some b that is an upper endpoint of one of the intervals comprising R.

In this case, consider the collection of subtrees of T rooted at the various children of each v_j (if v_j has more than $d+1$ children), at the various children of v_j other than v'_j (if v_j has $d+1$ children), at each such v'_j (if it is a leaf), and at the various children of each such v'_j (otherwise). Order these subtrees by decreasing leaf-label at the root.

If all of these subtrees are $R(d)$-good, then set $f_\ell(T) = T$.

If they are not all $R(d)$-good then locate the first, T', that is not. Since T' has fewer leaves than T it is in $\mathcal{T}(\ell')$ for some $\ell' < \ell$, and so by induction there is a tree $f_{\ell'}(T')$ that is in $\mathcal{T}^E(\ell')$ if $T' \in \mathcal{T}^O(\ell')$, and in $\mathcal{T}^O(\ell')$ if $T' \in \mathcal{T}^E(\ell')$. In T replace T' with $f_{\ell'}(T')$ and declare the result to be $f_\ell(T)$.

Because f_ℓ in this case only changes T', it is immediate that $f_\ell(T) \in \mathcal{T}(\ell)$ and moreover that it is in $\mathcal{T}^E(\ell)$ if $T \in \mathcal{T}^O(\ell)$, and in $\mathcal{T}^O(\ell)$ if $T \in \mathcal{T}^E(\ell)$.

Consider what happens when we apply f_ℓ to $f_\ell(T)$ in this case. The path $v_1, \ldots, v_k$ stays the same from T to $f_\ell(T)$, as does the number of children of each v_j, the location of v'_j (the child of v_j with the least leaf-label), and crucially, the number of children of such a v'_j. Indeed, if v'_j is a leaf in T it evidently stays a leaf in $f_\ell(T)$, and if not, then, since the subtrees are rooted not at the v'_j but at the children of the v'_j, the number of children of each v'_j remains unchanged. Since we were not in Case 1 for T, it follows that we are not in Case 1 for $f_\ell(T)$, so we are in Case 2. The collection of subtrees examined when applying f_ℓ to $f_\ell(T)$ is evidently the same at those examined for T, except that T' has become $f_{\ell'}(T')$. The ordering remains unchanged, so now $f_{\ell'}(T')$ is the first component that is not $R(d)$-good. By induction $f_{\ell'}(f_{\ell'}(T')) = T'$, so $f_\ell(f_\ell(T)) = T$.

We have established the existence of an $f_\ell : \mathcal{T}(\ell) \to \mathcal{T}(\ell)$ whose orbits of size two include one element each of $\mathcal{T}^E(\ell), \mathcal{T}^O(\ell)$. That fixed points of the map just defined are $R(d)$-good labelled Schröder trees is immediate from the construction; that all $R(d)$-good labelled Schröder trees are fixed is a simple induction from the definition of $R(d)$-goodness. $\square$

In light of the results of Section 2, to deduce Theorem 1.5 from Claim 4.2 we need to establish the following.

Claim 4.3. *Set*

$$F(x) = \sum_{s \in R} \frac{x^{d(s-1)+1}}{(d(s-1)+1)!}$$

and let $F^{<-1>}(x) := \sum_{n\geq 1} f_n^{<-1>} x^n/n!$ *be the series reversion of* $F(x)$*. Then for* n *which is not one more than a multiple of* d*, we have* $f_n^{<-1>} = 0$*, while for* $n = d\ell + 1$*,* $\ell = 0, 1, 2, \ldots$ *we have*

$$f_n^{<-1>} = (-1)^\ell \left|\mathcal{T}^{\text{good}}(\ell)\right|,$$

where $\mathcal{T}^{\text{good}}(\ell)$ *is the set of* $R(d)$*-good labelled Schröder trees with* $d\ell + 1$ *leaves.*

Proof. An immediate application of Proposition 3.2, coupled with the observation that a Schröder tree each of whose vertices has a number of children that is one greater than a multiple of d must have a number of leaves that is one greater than a multiple of d, yields

$$f_n^{<-1>} = \begin{cases} 0 & \text{if } n \not\equiv 1 \pmod{d}, \\ \sum_{T\in\mathcal{T}(\ell)} a(T) & \text{if } n = d\ell + 1, \end{cases}$$

where

$$a(T) = \begin{cases} +1 & \text{if } T \text{ has an even number of non-leaf vertices,} \\ -1 & \text{if } T \text{ has an odd number of non-leaf vertices.} \end{cases}$$

Recall that for any $T \in \mathcal{T}(\ell)$ we have $\ell = \sum_v (s_v - 1) = \sum_v s_v - \text{int}(T)$ where $\text{int}(T)$ is the number of non-leaf vertices of T and $\sum_v s_v$ is the sum whose parity determines whether $T \in \mathcal{T}^E(\ell)$ or $\mathcal{T}^O(\ell)$. From this relation we see that if ℓ is even then trees in $\mathcal{T}^E(\ell)$ have $\text{int}(T)$ even (and those in $\mathcal{T}^O(\ell)$ have $\text{int}(T)$ odd), so that for even ℓ we have

$$f_{d\ell+1}^{<-1>} \left(= \sum_{T\in\mathcal{T}(\ell)} a(T) \right) = \left|\mathcal{T}^E(\ell)\right| - \left|\mathcal{T}^O(\ell)\right| = \left|\mathcal{T}^{\text{good}}(\ell)\right|,$$

the last equality from Claim 4.2. Similarly for odd ℓ we have

$$f_{d\ell+1}^{<-1>} = -|\mathcal{T}^{\text{good}}(\ell)|.$$

The claim follows. □

5 Restricted Lah numbers

Here we present the proofs of Theorem 1.10 and its generalization Theorem 1.12. We begin by applying Proposition 3.1 in the case where a is a 0-1 sequence with $a_1 = 1$. Evidently $w_0 = 1$ in this case and it is an easy induction that w_i is an integer for all i. Say that the sequence $(w_1, w_2, \ldots)$ is *tree sign coherent* if there is a sign function $e : \mathbb{N} \to \{+1, -1\}$ with the following property: for each n for which there is at least one $T \in \mathcal{T}^{\text{l.o.c.}}(n)$ with $w(T) = \prod_{i=1}^k w_{d(i)}$ non-zero (here $d(1), \ldots, d(k)$ are the down-degrees of the non-leaf vertices), the sign of $w(T)$ is $e(n)$ for every such T.

In the presence of tree sign coherence we can combinatorially interpret $e(n)b_n$ as the count of $\mathcal{D}$-decorated labelled l.o.c. trees on n vertices, where $\mathcal{D} = (D_1, D_2, \ldots)$ is any sequence of sets with $|D_i| = |w_i|$ (Note that this means that if $w_\ell = 0$ for some ℓ, we get no contribution to $e(n)b_n$ from any labelled l.o.c. trees that has a non-leaf vertex with down-degree ℓ. Using Defintion 1.9 and the results of Section 2 we immediately have the following corollary of Proposition 3.1 and (4).

Corollary 5.1. *Let $R \subseteq \mathbb{N}$ with $1 \in R$ satisfy the following properties.*

1. *The sequence $(w_1, w_2, \ldots)$, defined via $x/\sum_{n\in R} x^n = 1 + \sum_{n\geq 1} w_n x^n$, is tree sign coherent.*

2. *The sequence $(b_1, b_2, \ldots)$ is sign coherent with sign function $c(n,k)$.*

Then $d(R) = ((-1)^n w_n)_{n\geq 1}$ and $c(n,k)L(n,k)_R^{-1} = \left|\mathcal{G}^{\text{l.o.c.}}_{\mathcal{D}(R)}(n,k)\right|$.

5.1 Proof of Theorem 1.10

Recall that now $1 \in R$ and that R has no exposed odds, meaning that R is a union of maximal intervals of consecutive integers of the form $[1,\infty)$, or $[1,b]$, or $[a,b]$, or $[b,\infty)$, with a, b even (and, in the case of the interval $[a,b]$, not necessarily distinct). Writing $O_R(x)$ for the ordinary generating function of the sequence $(\mathbf{1}_{\{n\in R\}})_{n\geq 1}$, that is, $O_R(x) = \sum_{n\in R} x^n$, we now argue that the coefficients of $x/O_R(x)$ are alternating, with coefficients of odd powers being negative. This is equivalent to saying that the coefficients of $-x/O_R(-x)$ are all positive.

If $R = \mathbb{N}$ this is immediate, since $-x/\sum_{n\geq 1}(-x)^n = 1 + x$. Otherwise R is of the form $[1,b_1] \cup [a_2,b_2] \cup \ldots$, with all a_i, b_i even, with $1 < b_1 < a_2 \leq b_2 < a_3 \leq b_3 \ldots$, with either infinitely many a_i's and b_i's (R the union of infinitely many finite intervals), or finitely many with both sequences terminating at the same index (R the union of finitely many finite intervals), or finitely many with the sequence of a_i's terminating at an index one higher than that at which the sequence of b_i's terminates (R the union of finitely many finite intervals, together with one infinite interval). In all three of these situations, using the evenness of all the a_i's and b_i's we have

$$
\begin{aligned}
\frac{-x}{O_R(-x)} &= \frac{1}{(1+\cdots+(-x)^{b_1-1}) + (-x)^{a_2-1}(1+\cdots+(-x)^{b_2-a_2}) + \cdots} \\
&= \frac{1-(-x)}{1-(-x)^{b_1} + (-x)^{a_2-1} - (-x)^{b_2} + \cdots} \\
&= \frac{1+x}{1-(x^{b_1} + x^{a_2-1} + x^{b_2} + \cdots)} \\
&= (1+x)\sum_{k=1}^{\infty}(x^{b_1} + x^{a_2-1} + x^{b_2} + \cdots)^k
\end{aligned}
$$

which evidently has all positive coefficients.

We now argue that a sequence $(w_1, w_2, \ldots)$ that alternates in sign with w_1 negative is tree sign coherent. Indeed, consider an l.o.c. tree T on n vertices. If n is even, then, since the number of edges of T is odd, T has an odd number of non-leaf vertices that have an odd number of children and $w(T)$ is negative. Similarly, if n is odd T has an even number of non-leaf vertices that have an odd number of children and $w(T)$ is positive.

Sign coherence of the sequence $(b_1, b_2, \ldots)$ now follows immediately; since the b_i are alternating in sign with b_1 positive (this follows from the form of $e(n)$), we have sign coherence with sign function $c(n,k) = (-1)^{n-k}$. Observing finally that (comparing Definition 1.9 and the definition of w_n from the statement of Proposition 3.1) we have $w_n = (-1)^n d_n$, we have proven Theorem 1.10.

We give three illustrative examples.

- $R = \mathbb{N}$ (ordinary Lah numbers): Here $x/O_{\mathbb{N}}(x) = 1/(1+x+x^2+\ldots) = 1-x$, so that $w_i = -1$ if $i = 1$, and $w_i = 0$ otherwise. In this case $\mathcal{D}$-decorated l.o.c. trees are simply rooted paths. There are $n!$ labelled rooted paths on n vertices, and so we recover the classical result that $(-1)^{n-k}L(n,k)^{-1} = L(n,k)$.

- $R = \{1,\ldots,r\}$, $r \geq 2$ even ($[r]$-restricted Lah numbers): Here
$$x/O_{[r]}(x) = x/(x+\ldots+x^r) = \frac{1}{1-x^r} - \frac{x}{1-x^r}$$
so that
$$w_i = \begin{cases} 1 & \text{if } i = 0 \pmod r \\ -1 & \text{if } i = 1 \pmod r \\ 0 & \text{otherwise.} \end{cases}$$
In this case $\mathcal{D}$-decorated l.o.c. trees are simply l.o.c. trees all of whose non-leaf vertices have a number of children that is congruent to either 0 or 1 modulo r. This yields Corollary 1.11. (Note in this case that if r is odd we do not have alternation of signs of the w_i).

- $R = \{1,2,4,6,8,\ldots\}$: Here
$$x/O_R(x) = \frac{1-x^2}{1+x-x^2}$$
(an easy calculation) so that $w_i = (-1)^i F_i$, where F_i is a Fibonacci number defined by $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 3$, and $(-1)^{n-k}L(n,k)_R^{-1}$ counts the number of labelled l.o.c. forests on n vertices with k components in which each non-leaf vertex with (say) k children has associated with it one element from a set of F_k objects.

5.2 Proof of Theorem 1.12

Fix $d \geq 1$. Given a set R with $1 \in R$ that has no exposed odds, as before denote by $R(d)$ the set in which each element k of R is replaced with $d(k-1)+1$. Here we argue that the coefficients of $x/O_{R(d)}(x)$ are supported on the arithmetic progression $1, d+1, 2d+1, \ldots$, are alternating along that progression, and start with a positive term. In the case $R = \mathbb{N}$ this is again immediate. For all other R's, using the notation from the proof of Theorem 1.10 we have

$$\begin{aligned} \frac{x}{O_{R(d)}(x)} &= \frac{1}{(1+\cdots+x^{(b_1-1)d}) + x^{(a_2-1)d}(1+\cdots+x^{(b_2-a_2)d}) + \cdots} \\ &= \frac{1-x^d}{1-x^{b_1 d} + x^{(a_2-1)d} - x^{b_2 d} + \cdots} \\ &= \frac{1-y}{1-y^{b_1} + y^{a_2-1} - y^{b_2} + \cdots} \end{aligned}$$

where in the last equality we have set $y = x^d$. As in the $d = 1$ case the evenness of all the a_i's and b_i's implies that the above series (in y) is alternating with first coefficient positive, so we indeed have that in all cases

$$\frac{x}{O_{R(d)}(x)} = 1 + \sum_{n\geq 1} w_{dn} x^{dn}$$

with the w_{dn} positive for all even n and negative for all odd n.

We claim that the sequence $(0, 0, \ldots, w_d, 0, 0, \ldots, w_{2d}, 0, 0, \ldots)$ is tree sign coherent. Indeed, there is some $T \in \mathcal{T}^{\text{l.o.c.}}(n)$ with $w(T)$ non-zero only when n is of the form $dm + 1$ (m a positive integer). Fix such an n, and consider such a T; all non-leaf vertices must have a number of children that is a multiple of d. If m is even then there must be an even number of non-leaf vertices v for which $d(v)$ is an odd multiple of d, and so $w(T)$ must be positive. On the other hand, if m is odd then there must be an odd number for which $d(v)$ is an odd multiple of d, and so $w(T)$ must be negative. Tree sign coherence follows with

$$e(n) = (-1)^{\frac{n-1}{d}}$$

whenever $d|(n-1)$ (with $e(n)$ unconstrained everywhere else). From Lemma 2.5 it follows that the corresponding sequence $(b_1, b_2, \ldots)$ is sign coherent with $c(n, k) = (-1)^{(n-k)/d}$ whenever $d|(n - k)$ (with $c = 0$ everywhere else). Theorem 1.12 follows.

We mention just one special case of this result, when $R = \{1, d + 1, 2d + 1, \ldots\}$. Exactly as in the example $R = \mathbb{N}$ from Section 5.1 we easily get that $w_{dn} = (-1)^n$ (with all other w_i's zero), so that, as asserted in Corollary 1.13, $(-1)^{(n-k)/d} L(n, k)_R^{-1}$ counts the number of n vertex, k-component labelled l.o.c. forests in which all non-leaf vertices have down-degree d.

6 Concluding remarks

We have given combinatorial interpretations for each each of $\left\{ {n \atop k} \right\}_R^{-1}$, $\left[{n \atop k} \right]_R^{-1}$ and $L(n, k)_R^{-1}$ for all R with $1 \in R$, but for many R these interpretations are as the difference in sizes of two sets of forests. Here we present some questions related to the more satisfactory situation where the interpretations are as counts of single sets of forests.

6.1 Stirling numbers of the second kind

We have observed that if a set R of natural numbers that includes 1 has the property that the series reversion of $\sum_{s\in R} x^s/s!$ is alternating (or is supported on and alternating along an arithmetic progression), then there is the potential to obtain a clean combinatorial interpretation of the numbers $\left\{ {n \atop k} \right\}_R^{-1}$ along the lines of those presented in Theorems 1.3 and 1.5. The central issue is understanding, combinatorially, the coefficients of the series reversion when it is viewed as an exponential generating function.

We have resolved this central issue for many sets R, but there is most likely more to be done. For example, a Mathematica computation reveals that the series reversion of $x + x^2/2 + x^4/24 + x^5/120$ (corresponding to the set $R = \{1, 2, 4, 5\}$, which is not covered by our results) is alternating at least up to the x^{1200} term, so it is highly likely to be an alternating series.

Question 6.1. *Can we characterize those $R \subseteq \mathbb{N}$ and $d \geq 1$ for which the series reversion of $\sum_{n \in R(d)} x^n/n!$ is sign coherent, and so for which it is potentially possible to find a combinatorial interpretation of ${n \brace k}^{-1}_{R(d)}$ as a count of a set of forests? And can we find such a interpretation in each case?*

6.2 Lah numbers

The same question may be asked of restricted Lah numbers, where the goal is to find R's and d's such that $x/\sum_{s \in R(d)} x^s$ is alternating (or tree sign coherent). Here we can say definitively that there is more to be done. Consider, for example, the set $R = \{1, 2, r+1, r+2\}$, which is not covered by our results for any $r \geq 3$. We have

$$\begin{aligned} \frac{x}{x + x^2 + x^{r+1} + x^{r+2}} &= \frac{1}{(1+x)(1+x^r)} \\ &= \begin{cases} \sum_{k=1}^{\infty} (-1)^{k-1} k \sum_{j=0}^{r-1} (-1)^j x^{(k-1)r+j} & \text{if } r \text{ odd} \\ \sum_{k=1}^{\infty} \sum_{j=0}^{r-1} (-1)^j x^{2(k-1)r+j} & \text{if } r \text{ even,} \end{cases} \end{aligned}$$

which is alternating.

Question 6.2. *Can we characterize those $R \subseteq \mathbb{N}$ and $d \geq 1$ for which the series $x/\sum_{s \in R(d)} x^s$ is alternating (or tree sign coherent), and so for which it is potentially possible to find a combinatorial interpretation of $L(n,k)^{-1}_{R(d)}$ as a count of a set of forests? And can we find such a interpretation in each case?*

6.3 Stirling numbers of the first kind

Another avenue of exploration is restricted Stirling numbers of the first kind, ${n \brack k}_R$, defined to be the number of partitions of $[n]$ into k cyclically ordered non-empty blocks with all block sizes in R. Evidently the central issue is understanding the series reversion of the power series $\sum_{s \in R} x^s/s$. Mathematica computations suggest that if $1 \in R$ and R has no exposed odds, then the series reversion is indeed alternating.

Question 6.3. *Can we find combinatorial interpretations for the numbers ${n \brack k}^{-1}_R$ for various R, in particular those R for which we understand ${n \brace k}^{-1}_R$ and $L(n,k)^{-1}_R$?*

7 Acknowledgements

The second author thanks Hannah Porter and David Radnell for helpful discussions.

References

[1] P. Barry, On a Family of Generalized Pascal Triangles Defined by Exponential Riordan Arrays, *J. Integer Seq.* **10** (2007), Article 07.3.5.

[2] H. Belbachir and I. E. Bousbaa, Associated Lah numbers and r-Stirling numbers, arXiv:1404.5573 (2014).

[3] W. Chen, A general bijective algorithm for trees, *Proc. Nati. Acad. Sci. USA* **87** (1990), 9635–9639.

[4] J. Y. Choi, L. Long, S.-H. Ng and J. Smith, Reciprocity for multirestricted Stirling numbers, *J. Combin. Theory Ser. A* **113** (2006), 1050–1060.

[5] J. Y. Choi and J. D. H. Smith, On the combinatorics of multi-restricted numbers, *Ars Combin.* **75** (2005), 45–63.

[6] J. Y. Choi and J. D. H. Smith, On the unimodality and combinatorics of Bessel numbers, *Discrete Math.* **264** (2003) 45–53.

[7] L. Comtet, *Advanced Combinatorics*, Reidel, Boston, 1974.

[8] J. Felsenstein, The Number of Evolutionary Trees, *Systematic Zoology* **27** (1978), 27–33.

[9] N. Sloane, The on-line encyclopedia of integer sequences, `oeis.org`.

[10] R. Stanley, *Enumerative Combinatorics, vol. 2*, Cambridge University Press, Cambridge, 1999.

[11] A. Ungar, Generalized Hyperbolic Functions, *The American Mathematical Monthly* **89** (1982), 688–691.