Combinatorially interpreting generalized Stirling numbers

John Engbers* David Galvin† Justin Hilyard‡

August 17, 2016

Abstract

The Stirling numbers of the second kind $\left\{ {n \atop k} \right\}$ (counting the number of partitions of a set of size n into k non-empty classes) satisfy the relation

$$(xD)^n f(x) = \sum_{k \geq 0} \left\{ {n \atop k} \right\} x^k D^k f(x)$$

where f is an arbitrary function and D is differentiation with respect to x. More generally, for every word w in alphabet $\{x, D\}$ the identity

$$wf(x) = x^{(\#(x\text{'s in } w) - \#(D\text{'s in } w))} \sum_{k \geq 0} S_w(k) x^k D^k f(x)$$

defines a sequence $(S_w(k))_k$ of *Stirling numbers (of the second kind)* of w. Explicit expressions for, and identities satisfied by, the $S_w(k)$ have been obtained by numerous authors, and combinatorial interpretations have been presented.

Here we provide a new combinatorial interpretation that, unlike previous ones, retains the spirit of the familiar interpretation of $\left\{ {n \atop k} \right\}$ as a count of partitions. Specifically, we associate to each w a quasi-threshold graph G_w, and we show that $S_w(k)$ enumerates partitions of the vertex set of G_w into classes that do not span an edge of G_w. We use our interpretation to re-derive a known explicit expression for $S_w(k)$, and in the case $w = (x^s D^s)^n$ to find a new summation formula linking $S_w(k)$ to ordinary Stirling numbers. We also explore a natural q-analog of our interpretation.

In the case $w = (x^r D)^n$ it is known that $S_w(k)$ counts increasing, n-vertex, k-component r-ary forests. Motivated by our combinatorial interpretation we exhibit bijections between increasing r-ary forests and certain classes of restricted partitions.

*john.engbers@marquette.edu; Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee WI 53201. Research supported by the Simons Foundation.

†dgalvin1@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556, + 1 574 6814181 (corresponding author). Research supported by NSA grant H98230-13-1-0248, and by the Simons Foundation.

‡jhilyard@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556. Research supported by NSA grant H98230-13-1-0248.

1 Introduction

The Stirling number of the second kind, $\left\{ n \atop k \right\}$, counts the number of ways of partitioning a set of n elements into k non-empty classes. It satisfies the recurrence

$$\left\{ n \atop k \right\} = \begin{cases} \left\{ n-1 \atop k-1 \right\} + k\left\{ n-1 \atop k \right\} & \text{if both } n > 0 \text{ and } k > 0 \text{, and} \\ \mathbf{1}_{\{n=k\}} & \text{if } nk = 0. \end{cases} \tag{1}$$

The numbers $\left\{ n \atop k \right\}$ satisfy numerous algebraic identities; indeed, it was through the identity

$$x^n = \sum_{k \geq 0} \left\{ n \atop k \right\} x^{\underline{k}}$$

for $n \geq 0$, where $x^{\underline{k}}$ is the k^{th} falling power $x(x-1)\dots(x-(k-1))$, that James Stirling originally introduced the numbers, in his 1730 *Methodus differentialis* [25]. Central to the present paper is the following identity, probably first observed by Scherk in his 1823 thesis [23] (see [6] for details), which arises when one repeatedly applies the operator xD to an infinitely differentiable function $f(x)$ (where D is differentiation with respect to x). For all $n \geq 0$ we have

$$(xD)^n f(x) = \sum_{k \geq 0} \left\{ n \atop k \right\} x^k D^k f(x) \tag{2}$$

(where here and throughout we interpret D^0 as the identity). One way to verify (2) is to prove by induction on n that $(xD)^n f(x)$ takes the form $\sum_{k \geq 0} S(n,k) x^k D^k f(x)$ for some numbers $S(n,k)$, and then show that these numbers satisfy (1) (with $\left\{ \cdot \atop \cdot\cdot \right\}$ replaced everywhere by $S(\cdot, \cdot\cdot)$).

More generally, for each word w in alphabet $\{x, D\}$, with m x's and n D's, we have a unique expansion of the form

$$w = x^{m-n} \sum_{k \geq 0} S_w(k) x^k D^k, \tag{3}$$

with both sides being viewed as operators on a space of infinitely differentiable functions. One may easily verify (3) by induction on the length of the word w. Uniqueness comes from considering two different expressions of the form of the right-hand side of (3), and applying their difference to $f(x) = x^{k_0}$, where k_0 is the smallest index k for which the coefficients of $x^k D^k$ differ between the two expressions. The result is a power series in x with at least one non-zero coefficient, and so cannot be identically zero.

The integer sequence $(S_w(k))_{k \geq 0}$ that arises from (3) is what we call the *Stirling sequence (of the second kind)* of w; Scherk's identity (2) states that if $w = (xD)^n$ then the Stirling sequence of w coincides with the ordinary Stirling sequence of the second kind.

The set of words on alphabet $\{x, D\}$ forms a representation of the *Weyl algebra*, as the operators x and D satisfy the Weyl algebra's defining relation $Dx = xD + 1$. In this context the right-hand side of (3) is referred to as the *normal order* of the word w. The normal order problem arises in quantum mechanics, where x is viewed as a "creation operator" and D as an "annihilation operator" in a space of polynomials. Because these operators do not commute, it is desirable from a computational point of view to find expansions of words that are presented as a sum of words, all of which have the annihilation part completely to the right and the creation part completely to the left. See [7] for an introduction to this perspective.

The study of $S_w(k)$ has a long history. Some instances (beyond $w = (xD)^n$) were studied by Scherk in his 1823 thesis [23]. Carlitz [10, 11] derived summation formulae and identities for some instances in the 1930's, and Comtet [13] considered the case $w = (x^r D)^n$ in the 1970's, as did Lang [18] in 2000.

All of these references deal with the problem through generating functions and recurrences. In 1973 Navon [22] provided a lovely combinatorial interpretation of $S_w(k)$ for all w, associating a Ferrers board to w and realizing $S_w(k)$ as the number of ways of placing non-attacking rooks on the board (see Section 4 for more details); Varvak thoroughly explored this interpretation, and obtained q-analogs for it, in [28]. Very recently Codara et al. [14] gave a combinatorial interpretation in terms of graph coloring in the case $w = (x^s D^s)^n$; our Theorem 2.3, which was developed independently, generalizes this interpretation to arbitrary w.

The arrival of the quantum mechanics community to the problem in the early 2000's has led to a flurry of activity. Blasiak et al. [8, 9] thoroughly studied identities and recurrences for $S_w(k)$ for certain words w, and Schork [24] looked at q-analogs. Mendez et al. [21] looked at $S_w(k)$ for general w, and gave a combinatorial interpretation in terms of certain generalized tree structures. This interpretation was explored more by Lang [19] and recently quite thoroughly by Blasiak and Flajolet [6]. The connection to rook polynomials was revisited by Blasiak et al. in [5], and in [27] the same authors explored connections to Feynman diagrams. In [3], Asakly et al. showed how to read off the normal order of a word from a certain labeled tree, and in [20] Ma et al. explored connections between normal ordering and context-free grammars.

One contribution of the present paper is a simple new graph theoretic interpretation of $S_w(k)$ for any w and k, that very naturally generalizes the standard interpretation of $\left\{ n \atop k \right\}$ as a count of partitions. Specifically, we show how to associate to any w a ground-set and a set of forbidden pairs (which we encode as a graph on the ground-set), in such a way that $S_w(k)$ counts the number of partitions of the ground set that avoid putting both members of a forbidden pair into the same block. The statements of all our results can be found in Section 2, and the remaining sections are devoted to the proofs.

2 Statement of results

Here we outline the main results of the present paper.

2.1 A new combinatorial interpretation of $S_w(k)$

We begin by giving a new combinatorial interpretation of $S_w(k)$. We associate to each w a graph G_w with the property that $S_w(k)$ enumerates the partitions of the vertex set of G_w into a specified number of non-empty classes, with the property that no class includes both endvertices of an edge of G_w. In the case $w = (xD)^n$, G_w turns out to be the empty graph on n vertices and we recover the usual combinatorial interpretation of $\left\{ {n \atop k} \right\}$. In the case $w = (x^s D^s)^n$, G_w is the disjoint union of n copies of the complete graph on s vertices, and we recover a recent result of Codara et al. [14, Proposition 2.2]. To the best of our knowledge, ours is the first combinatorial interpretation of $S_w(k)$ for arbitrary w as a count of (restricted) partitions.

To define G_w, we first introduce the notion of a *Dyck word*.

Definition 2.1. *A word w in alphabet $\{x, D\}$ is a* Dyck word *if it satisfies the following:*

1. *it has the same number of x's as D's, and*
2. *reading the word from left to right, every initial segment has at least as many x's as D's.*

We say that a Dyck word is irreducible *if either it is the word xD or it is of the form $xw'D$ with w' a non-empty Dyck word, and we say that it is* reducible *otherwise.*

Observe that a reducible word may be written (in unique way) as $w_1 \dots w_\ell$, where each w_i is irreducible.

A *Dyck path* in $\mathbb{R}^2$ is a staircase path (a path that proceeds by taking unit steps, either in the positive x direction or the positive y direction) that starts at $(0, 0)$, ends on the line $x = y$, any never goes below this line. There is a natural correspondence between Dyck paths and Dyck words, given by mapping steps in the positive y direction to x, and steps in the positive x direction to D (see Figure 1 for an example). Irreducible Dyck words correspond to Dyck paths that meet the line $x = y$ only at their initial and terminal points, and reducible Dyck words correspond to Dyck paths that meet the line at some intermediate points as well.

We now associate to each Dyck word w an unlabeled graph G_w inductively, as follows:

1. If $w = xD$, then $G_w = K_1$ (the isolated vertex).
2. If w is irreducible with $w = xw'D$ for some non-empty Dyck word w', then $G_w = G_{w'} + K_1$ (the graph obtained from $G_{w'}$ by adding a dominating vertex).
3. If w is reducible, say $w = w_1 \dots w_\ell$ with each w_i irreducible, then $G_w = G_{w_1} \cup \dots \cup G_{w_\ell}$ (the disjoint union of the G_{w_i}'s).

In particular clause 3 above tells us that if $w = w'w''$ with w' and w'' both Dyck words (not necessarily irreducible) then $G_w = G_{w'} \cup G_{w''}$.

We note in passing that G_w belongs to the well-known family of *quasi-threshold*, or *trivially perfect* graphs (see e.g. [29] for a survey); recall that the family of quasi-threshold graphs is the smallest family that contains K_1 and is closed under adding dominating vertices and taking disjoint unions (if unions that do not amount to the addition of an isolated vertex are forbidden, we obtain the smaller class of *threshold* graphs).

As an example, consider the word $w = xxDxxDxDDD$. Writing this as

$$x\left([xD]\left[x\left\{\langle xD\rangle\langle xD\rangle\right\}D\right]\right)D,$$

we quickly see that G_w is constructed by taking two isolated vertices (say a and b), adding a dominating vertex (say c), taking the union of the resulting graph with an isolated vertex (say d), and then adding a final dominating vertex (say e). (See Figure 1.)

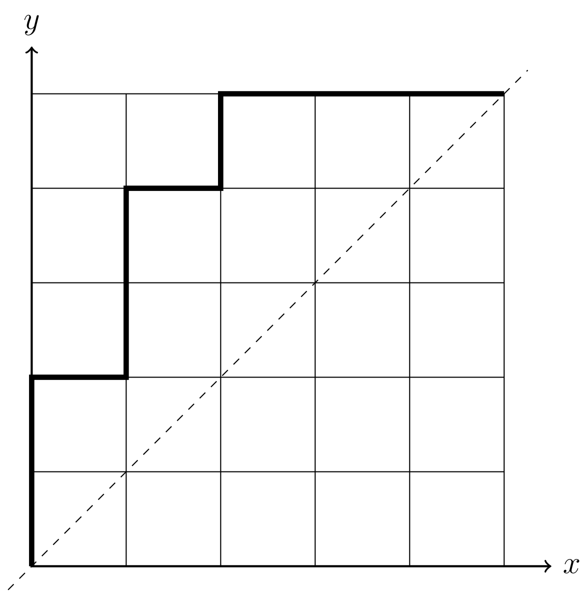


Figure 1: The Dyck path corresponding to the Dyck word $w = xxDxxDxDDD$ and the associated graph G_w.

To state our first theorem, it is convenient to generalize the symbol $\left\{ n \atop k \right\}$ to graphs.

Definition 2.2. *Let G be a graph and $k \geq 0$ an integer. The k^{th}* graph Stirling number *of G, denoted $\left\{ G \atop k \right\}$, is the number of ways of partitioning the vertex set of G into k non-empty classes, none of which contains both endvertices of an edge of G.*

Equivalently $\left\{ G \atop k \right\}$ counts partitions of G into k non-empty *independent sets*, that is, sets of mutually non-adjacent vertices; it also counts the number of different proper k-colorings of G using all k colors, with two colorings considered to be the same if they differ only on the names of the colors. These numbers were first studied in their own right (to the best of our knowledge) by Tomescu [26], although they were probably first introduced (in the setting of planar graphs) by Birkhoff [4]. Note that $\left\{ G \atop k \right\}$ does indeed generalize the Stirling numbers $\left\{ {n \atop k} \right\}$, since when $G = E_n$, the graph on n vertices with no edges, we have $\left\{ {E_n \atop k} \right\} = \left\{ {n \atop k} \right\}$.

We can now state of first theorem, which gives a simple new combinatorial interpretation of $S_w(k)$ in terms of restricted partitions.

Theorem 2.3. *Let w be a Dyck word in the alphabet $\{x, D\}$. For all $k \geq 0$ we have*

$$S_w(k) = \left\{ {G_w \atop k} \right\}.$$

Our proof of Theorem 2.3 will be direct, in the sense that we do not rely on previously known formulae or combinatorial interpretations for $S_w(k)$.

As an illustration of Theorem 2.3, let us return to $xxDxxDxDDD$. A little computation with the relation $Dx = xD + 1$ yields

$$xxDxxDxDDD = 2x^3D^3 + 4x^4D^4 + x^5D^5.$$

On the other hand, we have $\left\{ {G_w \atop 3} \right\} = 2$ (the two partitions of $V(G_w)$ into three non-empty independent sets are $ab|cd|e$ and $abd|c|e$), $\left\{ {G_w \atop 4} \right\} = 4$ (the partitions being $ab|c|d|e$, $ad|b|c|e$, $a|bd|c|e$ and $a|b|cd|e$), $\left\{ {G_w \atop 5} \right\} = 1$ (the unique partition being $a|b|c|d|e$), and $\left\{ {G_w \atop k} \right\} = 0$ for all other k, exactly as predicted by Theorem 2.3.

In the case $w = (x^{t_i}D^{t_i})^n$ we get a particularly appealing combinatorial interpretation. Here $G_w = K_{t_1} \cup \ldots \cup K_{t_n}$, the disjoint union of cliques of various sizes, and Theorem 2.3 gives that $S_w(k)$ counts the number of partitions of this union of cliques into k non-empty independent sets. The case when all $t_i = 2$ was observed in [6], in slightly different language, and the case of all $t_i = t$ for general t has appeared recently in [14]. Note that in this last case we may whimsically interpret $\left\{ {G_w \atop k} \right\}$ as the number of ways of breaking up a gathering of n sets of t-tuplets into k non-empty blocks, in such a way that no block contains more than one member from each set of t-tuplets. Variants of this problem for twins, or $t = 2$, were considered by Griffiths in [17]; our work on these notes began after reading that paper.

Theorem 2.3 can easily be extended to give a combinatorial interpretation of $S_w(k)$ for arbitrary w.

Definition 2.4. *Let w be a word in the alphabet $\{x, D\}$. Let $a = a(w)$ be the least non-negative integer such that all initial segments of $x^a w$ have at least as many x's as D's, and let $b = b(w)$ be the unique non-negative integer such that $x^a w D^b$ is a Dyck word. We refer to $x^a w D^b$ as the Dyck word* associated with *w.*

The normal order of w is easily obtained from that of $x^a w D^b$. Indeed, suppose that

w has m x's and n D's. From (3) we have $w = x^{m-n}\sum_{k\geq 0} S_w(k)x^kD^k$. Then

$$\begin{aligned} x^awD^b &= x^a\left(x^{m-n}\sum_{k\geq 0} S_w(k)x^kD^k\right)D^b \\ &= x^{a+m-n}\sum_{k\geq 0} S_w(k)x^kD^{k+b} \\ &= \sum_{k\geq 0} S_w(k)x^{k+b}D^{k+b}, \end{aligned}$$

the last equality using $a+m=b+n$ (valid since x^awD^b is a Dyck word). Since also $x^awD^b = \sum_{k\geq 0} S_{x^awD^b}(k)x^kD^k$, we get the identity $S_w(k) = S_{x^awD^b}(k+b)$. Therefore, the following is an immediate corollary of Theorem 2.3.

Corollary 2.5. *Let w be an arbitrary word in the alphabet $\{x, D\}$, and let x^awD^b be its associated Dyck word, as in Definition 2.4. For all $k \geq 0$ we have*

$$S_w(k) = \left\{ {G_{x^awD^b} \atop k+b} \right\}.$$

Theorem 2.3 and Corollary 2.5 yield an explicit expression for $S_w(k)$. Let w be any word with, say, m x's, and let x^awD^b be its associated Dyck word.

Definition 2.6. *The* height a_i *of the i^{th} x in x^awD^b is the excess of x's over D's in the initial segment of x^awD^b that ends immediately prior to the i^{th} x.*

Equivalently, the height of each x can be calculated from the Dyck path associated with x^awD^b: if the step in the positive y direction corresponding to a particular x of the word goes from (a,b) to $(a,b+1)$, then the height of that x is $b-a$.

Theorem 2.7. *Let w be an arbitrary word in the alphabet $\{x, D\}$, and let x^awD^b be its associated Dyck word. With the a_i's as in Definition 2.6, we have*

$$S_w(k) = \frac{1}{(k+b)!}\sum_{\ell=0}^{k+b}(-1)^\ell\binom{k+b}{\ell}\prod_{i=1}^{m+a}(k+b-\ell-a_i).$$

For example, if $w = xxDxxDxDDD$ then $m=5$, $a=b=0$, $a_1=0$, $a_2=1$, $a_3=1$, $a_4=2$, $a_5=2$ and

$$S_w(k) = \frac{1}{k!}\sum_{\ell=0}^{k}(-1)^\ell\binom{k}{\ell}(k-\ell)(k-\ell-1)^2(k-\ell-2)^2.$$

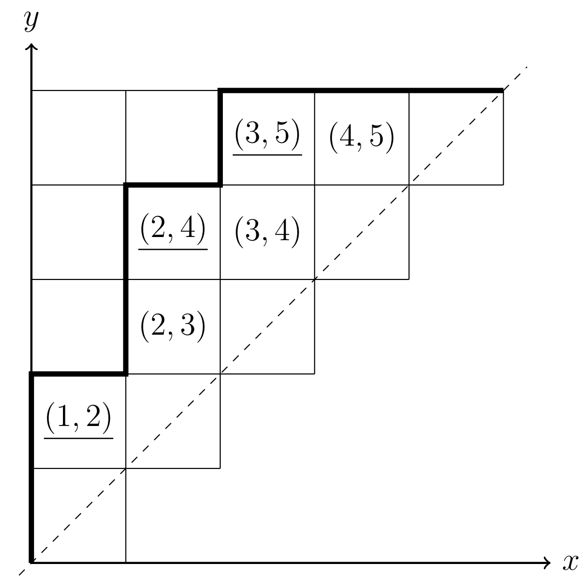
Similar explicit expressions have appeared in [28] and [21]. Note that when $w=(xD)^n$, Theorem 2.7 immediately reduces to the familiar

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

The proofs of Theorems 2.3 and 2.7 appear in Section 3.

2.2 A closely related combinatorial interpretation

By combining a result of Navon [22] with one of Goldman, Joichi and White [16], we find another graph H_w that can naturally be associated to a Dyck word w, such that $S_w(k) = \left\{ {H_w \atop k} \right\}$ for all k. To define this graph, label each unit square in $\mathbb{Z}^2$ with the coordinates of its top-right corner (so, for example, the square with corners at $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$ gets label $(1,1)$). Given a Dyck word w, let $\mathcal{W}_w$ be the set of (labels of) unit squares that lie below the staircase path of w, and completely above the line $x = y$. For example, if $w = (xD)^n$ then $\mathcal{W}_w = \emptyset$, and if $w = xxDxxDxDDD$ then $\mathcal{W}_w = \{(1,2),(2,3),(2,4),(3,4),(3,5),(4,5)\}$. Define a graph H_w on vertex set $\{1,\ldots,n\}$ (where n is the number of x's in w) by putting an edge from i to j ($i < j$) if and only if $(i,j) \in \mathcal{W}_w$. (See Figure 2.)



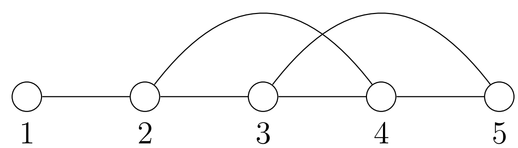


Figure 2: The staircase path of w and the graph H_w for $w = xxDxxDxDDD$.

It is worth noting that H_w is determined by the locations of the peaks of the Dyck path of w, that is, by the places where the path takes a step up followed by a step to the right. To make this precise, say that the Dyck path of w *has a peak* at (x,y) if it takes a step from $(x-1,y-1)$ to $(x-1,y)$ and then steps to (x,y). Let $\mathcal{T}_w = \{(x_1,y_1),\ldots,(x_k,y_k)\}$ be the set of peaks of the path of w. Then it is easy to see that the edge set of H_w can be covered by putting a clique on each of the consecutive segments $\{x_i,\ldots,y_i\}$, $1 \leq i \leq k$. For example, if $w = (xD)^n$ then $\mathcal{T}_w = \{(1,1),\ldots,(n,n)\}$ and the edge set of H_w is empty; while if $w = xxDxxDxDDD$ then $\mathcal{T}_w = \{(1,2),(2,4),(3,5)\}$, and the edge set of H_w is $\{\{1,2\}\} \cup \{\{2,3\},\{2,4\},\{3,4\}\} \cup \{\{3,4\},\{3,5\},\{4,5\}\}$, i.e., it is composed of cliques on the vertex sets $\{1,2\}$, $\{2,3,4\}$, and $\{3,4,5\}$ (see Figure 2, where the coordinates of the peaks of the path are underlined).

We note in passing that H_w belongs to the family of *indifference graphs*. Indeed, a characterization of indifference graphs mentioned in [15] is that they are exactly those graphs H on vertex set $\{v_1, \ldots, v_d\}$ for which there is some collection $\mathcal{C}$ of intervals from $\{1, \ldots, d\}$ such that the edge set of H is $\{v_i v_j : i, j$ are in some element of $\mathcal{C}\}$. Notice that the graphs H_w and G_w are sometimes isomorphic (for example, when $w = \prod_{i=1}^n x^{t_i} D^{t_i}$ for arbitrary t_i's), but not always (for example, when $w = xxDxxDxDDD$).

Theorem 2.8. *Let w be Dyck word in the alphabet $\{x, D\}$. For all $k \geq 0$ we have*

$$S_w(k) = \left\{ {H_w \atop k} \right\}.$$

As in Section 2.1, the following is an immediate corollary of Theorem 2.8.

Corollary 2.9. *Let w be an arbitrary word in the alphabet $\{x, D\}$, and let $x^a w D^b$ be its associated Dyck word. For all $k \geq 0$ we have*

$$S_w(k) = \left\{ {H_{x^a w D^b} \atop k + b} \right\}.$$

We give the proof of Theorem 2.8 in Section 4, where we also give a direct proof (not using the results of [22] and [16]) of the identity $\left\{ {H_w \atop k} \right\} = \left\{ {G_w \atop k} \right\}$ for all Dyck words w and $k \geq 0$.

2.3 A new summation formula for $S_w(k)$ when $w = (x^s D^s)^n$

All explicit expressions for $S_w(k)$ that have appeared in the literature have taken the form of alternating sums. Using Theorem 2.3, we can obtain a new expression for $S_w(k)$, in the special case $w = (x^s D^s)^n$, as a positive linear combination of ordinary Stirling numbers. In what follows we use $[x^\ell]p(x)$ for the coefficient of x^ℓ in the polynomial $p(x)$. The proof of the following theorem is given in Section 5.

Theorem 2.10. *Let $w = (x^s D^s)^n$. For each $k \geq 0$ we have*

$$S_w(k) = \sum_{\ell=0}^{(s-1)(n-1)} f(n, s, \ell) \left\{ {s(n-1) + 1 - \ell \atop k - (s-1)} \right\}$$

where

$$f(n, s, \ell) = \sum_{i_1 + \ldots + i_{s-1} = \ell} \binom{n-1}{i_1} \binom{2(n-1) - i_1}{i_2} \cdots \binom{(s-1)(n-1) - i_1 - \ldots - i_{s-2}}{i_{s-1}}. \tag{4}$$

In Section 5 we also establish the following alternate expression for $f(n, s, \ell)$:

$$f(n, s, \ell) = [x^\ell] \left((1 + x)(1 + 2x) \ldots (1 + (s-1)x) \right)^{n-1}. \tag{5}$$

The (unsigned) Stirling number of the first kind $\left[{a \atop b}\right]$ counts the number of permutations of a symbols that decompose into exactly b cycles. Using a well known identity satisfied by the Stirling numbers of the first kind, (5) immediately gives the following nice connection between generalized Stirling numbers of the second kind, and ordinary Stirling numbers of the first kind:

$$f(n,s,\ell) = [x^\ell]\left(\sum_{j=0}^{s-1}\left[{s \atop s-j}\right]x^j\right)^{n-1}.$$

The *Bell number* $B(w)$ of a word w is defined by $B(w)=\sum_k S_w(k)$ (so the Bell number of the word $(xD)^n$ is B_n, the n^{th} ordinary Bell number, counting the number of partitions of a set of size n into non-empty classes). From Theorem 2.10 we easily obtain

$$B((x^sD^s)^n) = \sum_{\ell=0}^{(s-1)(n-1)} f(n,s,\ell)B_{s(n-1)+1-\ell}. \tag{6}$$

In [9] the comment is made that the Bell number $B((x^sD^s)^n)$ can always be expressed in terms of conventional Bell numbers and r-nomial (binomial, trinomial, ...) coefficients, and the illustrative example $B((x^2D^2)^n)=\sum_{\ell=0}^{n-1}\binom{n-1}{\ell}B_{2n-1-\ell}$ is given; (6) illustrates this comment explicitly for arbitrary s.

2.4 Increasing forests

An *r-ary tree* is a tree in which every vertex, including a designated root, has some number i ($0 \le i \le r$) of children, with the set of children equipped with a bijection to some subset of a fixed set of size r (when $r=2$ this set is often taken to be $\{\text{left}, \text{right}\}$, for example, and for $r=3$ it might be $\{\text{left}, \text{middle}, \text{right}\}$; for general r we take it to be $\{1,\ldots r\}$). An *r-ary forest* is a forest in which each component is an r-ary tree. An *increasing r-ary forest* is an r-ary forest on, say, n vertices, together with a bijection from the vertices to $\{1,\ldots,n\}$, with the property that the labels go in increasing order when read along any path starting from a root vertex of a component.

Let $F(r,n,k)$ denote the set of increasing r-ary forests with n vertices and k components. It is easy to see that $|F(1,n,k)| = \left\{{n \atop k}\right\}$; in other words, writing $w(r,n)$ for the word $(x^rD)^n$, we have $|F(1,n,k)| = S_{w(1,n)}(k)$. More generally, Mendez et al [21, Section 5] have shown that the Stirling sequence of $(x^rD)^n$ enumerates increasing r-ary forests with n vertices by number of components; specifically, for all n, r and k,

$$|F(r,n,k)| = S_{w(r,n)}(k). \tag{7}$$

Corollaries 2.5 and 2.9 provide alternate combinatorial interpretations of $S_w(k)$, in the case $w=w(r,n)$, that are quite appealing. Indeed, in this case G_w is simply the threshold graph obtained from the graph on no vertices by iterating n times the operation of adding an isolated vertex and then adding $r-1$ dominating vertices, and H_w is the graph on vertex set $\{1,\ldots,rn\}$, with edges covered by cliques on vertices i through ir, for $1 \le i \le n$ (see Figures 3, 4 and 5 in Section 6). We refer to these graphs as $G(n,r)$ and $H(n,r)$ respectively. The following identities follow immediately from Theorems 2.3 and 2.8, via (7):

$$\begin{aligned} \left\{ \begin{matrix} G(n,r) \\ k+(r-1)n \end{matrix} \right\} &= |F(r,n,k)| \\ \left\{ \begin{matrix} H(n,r) \\ k+(r-1)n \end{matrix} \right\} &= |F(r,n,k)|. \end{aligned}$$

In Section 6 we give combinatorial proofs of both of these identities, by exhibiting bijections from the set of increasing r-ary forests with n vertices and k components to the set of partitions of both $G(n,r)$ and $H(n,r)$ into $k+(r-1)n$ non-empty independent sets.

2.5 q-analogs

The ordinary Weyl algebra on alphabet $\{x,D\}$ is generated by the relation $Dx = xD+1$. The *q-deformed Weyl algebra* is instead generated by the relation $Dx = qxD+1$ (where q is an indeterminate). This relation has been studied, for example, in the context of quantum harmonic oscillators [2]. If w is a word in the q-deformed Weyl algebra, with m x's and n D's, then we have the following analog of the normal order equation (3) (again easily verified by induction, and again unique)

$$w = x^{m-n} \sum_{k \geq 0} S_w^q(k) x^k D^k, \tag{8}$$

where now $S_w^q(k)$ is a polynomial in q with non-negative integer coefficients, that evaluates to $S_w(k)$ when $q=1$.

In the case $w=(xD)^n$, (8) leads to a q-analog of the Stirling numbers of the second kind. From a combinatorial perspective, a more natural way to define a q-analog of the Stirling numbers is through the recurrence

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\}_q = \begin{cases} q^{k-1}\left\{ \begin{smallmatrix} n-1 \\ k-1 \end{smallmatrix} \right\}_q + [k]_q \left\{ \begin{smallmatrix} n-1 \\ k \end{smallmatrix} \right\}_q & \text{if both } n>0 \text{ and } k>0, \text{ and} \\ \mathbf{1}_{\{n=k\}} & \text{if } nk=0 \end{cases}$$

where $[k]_q$ is, as usual, the polynomial $1+q+\ldots q^{k-1}$. This formulation goes back to Carlitz [12]. Happily, these two refinements of the Stirling numbers coincide: for all n, k and q we have $S^q_{(xD)^n}(k) = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}_q$ (with the proof of this following the lines of the proof of (2)).

Varvak [28] extended Navon's combinatorial interpretation of $S_w(k)$ for arbitrary w to a combinatorial interpretation of (the coefficients of) $S_w^q(k)$. We are also able to extend our interpretation (from Section 2.1) to the realm of the q-deformed Weyl algebra. Specifically, given a Dyck word w, by Theorem 2.3 there is a quasi-threshold graph G_w with the property that $\left\{ \begin{smallmatrix} G_w \\ k \end{smallmatrix} \right\} = S_w(k)$. Let $\mathcal{P}(w,k)$ be the set of partitions of

G_w into k non-empty independent sets. We show now how to associate a weight $\mathrm{wt}(P)$ to each $P \in \mathcal{P}(w,k)$ so that $\sum_{P\in\mathcal{P}(w,k)} q^{\mathrm{wt}(P)} = S^q_w(k)$.

Not unexpectedly, given the inductive construction of G_w, the definition of $\mathrm{wt}(P)$ will be inductive. It will also depend on a predetermined total order on the set of independent sets of G_w (which may be chosen arbitrarily); note that such a total order naturally induces a total order on the set of independent sets of any subgraph of G_w. Before giving the definition, we make some preliminary observations.

For $w = xD$, the only value of k for which $\mathcal{P}(w,k) \neq \emptyset$ is $k = 1$, and there is a unique partition in $\mathcal{P}(w,1)$. If $w = xw'D$ is irreducible, then G_w can be written as $G_{w'} + K_1$. Let v be the dominating vertex in this decomposition. For each k for which $\mathcal{P}(w,k) \neq \emptyset$, each $P \in \mathcal{P}(w,k)$ must include v as a singleton part, and the map $\varphi : \mathcal{P}(w,k) \to \mathcal{P}(w',k-1)$ that removes that singleton part is a bijection.

If $w = w_1w_2\dots w_\ell$ is reducible (with each w_i irreducible), then G_w can be written as the disjoint union of G_{w_1} and $G_{w_2\dots w_\ell}$. Fix a k for which $\mathcal{P}(w,k) \neq \emptyset$, and consider $P \in \mathcal{P}(w,k)$. There are numbers r,s such that exactly r of the parts of P have non-empty intersection with G_{w_1}, exactly s of them have non-empty intersection with $G_{w_2\dots w_\ell}$, and exactly $(r+s)-k$ of them have non-empty intersection with both G_{w_1} and $G_{w_2\dots w_\ell}$. By projection P induces partitions $P_1 \in \mathcal{P}(w_1,r)$ with parts $x_1,\dots,x_r$ (written in increasing order with respect to the total order on independent sets) and $P_2 \in \mathcal{P}(w_2\dots w_\ell, s)$ with parts $y_1,\dots,y_s$ (also in increasing order). We may encode how P is constructed from P_1 and P_2 using an r by s matrix whose ij entry is 1 if $x_i \cup y_j$ is one of the parts of P, and 0 otherwise. Call this matrix M; note that it contains a dimension $(r+s)-k$ permutation matrix as a minor, and all other entries are 0. When we come to define $\mathrm{wt}(P)$ in the reducible case, we will need to associate a weight to this matrix M.

Definition 2.11. *Let M be an r by s matrix that contains a dimension $(r+s)-k$ permutation matrix as a minor, and has all other entries 0. Mark all of the 0's in M that are either below a 1 (in the same column) or to the right of a 1 (in the same row). The* weight *of M, denoted $f(M)$, is the number of unmarked 0's in M.*

See Figure 6 in Section 7 for an example M.

We are now in a position to define the weight of a partition. For w a Dyck word, and P a partition of G_w into non-empty independent sets, associate a weight $\mathrm{wt}(P)$ to P inductively as follows.

1. If $w = xD$, then $\mathrm{wt}(P) = 0$.

2. If w is irreducible with $w = xw'D$ for some non-empty Dyck word w', then $\mathrm{wt}(P) = \mathrm{wt}(\varphi(P))$.

3. If w is reducible, say $w = w_1 \dots w_\ell$ with each w_i irreducible, then

$$\mathrm{wt}(P) = \mathrm{wt}(P_1) + \mathrm{wt}(P_2) + f(M).$$

We prove the following result in Section 7.

Theorem 2.12. *For each Dyck word w and each $k \geq 0$, we have*

$$\sum_{P \in \mathcal{P}(w,k)} q^{\mathrm{wt}(P)} = S_w^q(k).$$

While the construction above may seem involved, the following example illustrates a subtlety of q-analogs that must be captured by any interpretation. For $w = xDxDxD$, we have $S_w^q(2) = 2q + q^2$, despite the fact that all three partitions of the vertices of $G_w = E_3$ into two nonempty independent sets are isomorphic.

3 Interpreting $S_w(k)$ in terms of partitions of G_w

We begin with the proof of Theorem 2.3, which depends on the following two claims.

Claim 3.1. *Let w' be a word in the alphabet $\{x, D\}$, and let G' be a graph with the property that for all $k \geq 0$,*

$$S_{w'}(k) = \left\{ {G' \atop k} \right\}.$$

Let $w = xw'D$ and let G be obtained from G' by adding a dominating vertex. For all $k \geq 0$,

$$S_w(k) = \left\{ {G \atop k} \right\}.$$

Proof. Using

$$w' f(x) = \sum_{k \geq 0} \left\{ {G' \atop k} \right\} x^k D^k f(x)$$

we easily get (applying the above with $f(x)$ replaced by $f'(x)$, and setting $\left\{ {G' \atop -1} \right\} = 0$)

$$xw'Df(x) = \sum_{k \geq 0} \left\{ {G' \atop k-1} \right\} x^k D^k f(x).$$

The proof is completed by noting that $\left\{ {G \atop k} \right\} = \left\{ {G' \atop k-1} \right\}$ for all $k \geq 0$, since in any partition of G into non-empty independent sets, the dominating vertex added in going from G' to G must form a singleton block. $\square$

Claim 3.2. *Let w_1, w_2 be words in the alphabet $\{x, D\}$, and let G_1, G_2 be graphs with the property that for each $i \in \{1, 2\}$ and all $k \geq 0$,*

$$S_{w_i}(k) = \left\{ {G_i \atop k} \right\}. \tag{9}$$

For all $k \geq 0$,

$$S_{w_1 w_2}(k) = \left\{ \begin{matrix} G_1 \cup G_2 \\ k \end{matrix} \right\}$$

where $w_1 w_2$ is the concatenation of w_1 and w_2, and $G_1 \cup G_2$ is the disjoint union of G_1 and G_2.

Proof. Using (9) we have, for arbitrary f,

$$\begin{aligned} w_1 w_2 f(x) &= w_1 \sum_{k_2 \geq 0} \left\{ \begin{matrix} G_2 \\ k_2 \end{matrix} \right\} x^{k_2} D^{k_2} f(x) \\ &= \sum_{k_1 \geq 0} \left\{ \begin{matrix} G_1 \\ k_1 \end{matrix} \right\} x^{k_1} D^{k_1} \sum_{k_2 \geq 0} \left\{ \begin{matrix} G_2 \\ k_2 \end{matrix} \right\} x^{k_2} D^{k_2} f(x) \\ &= \sum_{k_1, k_2 \geq 0} \left\{ \begin{matrix} G_1 \\ k_1 \end{matrix} \right\} \left\{ \begin{matrix} G_2 \\ k_2 \end{matrix} \right\} x^{k_1} D^{k_1} x^{k_2} D^{k_2} f(x). \end{aligned} \tag{10}$$

Now by Leibniz' rule ($(fg)^{(n)}(x) = \sum_{k \geq 0} \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x)$) for the iterated derivative of a product we have

$$x^{k_1} D^{k_1} x^{k_2} D^{k_2} f(x) = \sum_{j \geq 0} \frac{k_1^{\underline{j}} k_2^{\underline{j}}}{j!} x^{k_1 + k_2 - j} D^{k_1 + k_2 - j} f(x). \tag{11}$$

(Recall that $x^{\underline{j}}$ is the j^{th} falling power of x, that is, $x(x-1)\dots(x-j+1)$.) Inserting this into (10) and extracting the coefficient of $x^k D^k$ from each side we get

$$\begin{aligned} S_{w_1 w_2}(k) &= \sum_{k_1, k_2 \geq 0} \left\{ \begin{matrix} G_1 \\ k_1 \end{matrix} \right\} \left\{ \begin{matrix} G_2 \\ k_2 \end{matrix} \right\} \frac{k_1^{\underline{k_1 + k_2 - k}} k_2^{\underline{k_1 + k_2 - k}}}{(k_1 + k_2 - k)!} \\ &= \sum_{k_1, k_2 \geq 0} \left\{ \begin{matrix} G_1 \\ k_1 \end{matrix} \right\} \left\{ \begin{matrix} G_2 \\ k_2 \end{matrix} \right\} \binom{k_1}{k_1 + k_2 - k} \binom{k_2}{k_1 + k_2 - k} (k_1 + k_2 - k)!. \end{aligned} \tag{12}$$

We claim that the right-hand side of (12) is exactly $\left\{ \begin{smallmatrix} G_1 \cup G_2 \\ k \end{smallmatrix} \right\}$. Indeed, one way to generate all partitions of $G_1 \cup G_2$ into k nonempty independent sets is as follows. First, fix a pair (k_1, k_2). Then, for each i, partition G_i into k_i non empty independent sets (there are $\left\{ \begin{smallmatrix} G_1 \\ k_1 \end{smallmatrix} \right\} \left\{ \begin{smallmatrix} G_2 \\ k_2 \end{smallmatrix} \right\}$ ways to do this). Next, choose $k_1 + k_2 - k$ of the classes from G_1 and $k_1 + k_2 - k$ of the classes from G_2 (there are $\binom{k_1}{k_1+k_2-k}\binom{k_2}{k_1+k_2-k}$ ways to do this). Finally, merge the chosen classes in pairs, one from G_1 and one from G_2 (there are $(k_1 + k_2 - k)!$ ways to do this), thereby creating $k_1 + k_2 - (k_1 + k_2 - k) = k$ non-empty independent sets. □

We are now ready to prove Theorem 2.3.

Proof. (Theorem 2.3) We proceed by induction on the length of w. If w has length 2 then $w = xD$ and $G_w = K_1$ and the result is trivial.

If w is irreducible, and of length greater than 2, then $w = xw'D$ for some Dyck word w which (by induction) has an associated graph $G_{w'}$, constructed as described in Section 2.1, with $S_{w'}(k) = \left\{ {G_{w'} \atop k} \right\}$ for all $k \geq 0$. That $S_w(k) = \left\{ {G_w \atop k} \right\}$ for all $k \geq 0$, where G_w is obtained from $G_{w'}$ by adding a dominating vertex follows from Claim 3.1.

If w is reducible, and of length greater than 2, then $w = w_1 w_2 \dots w_k$ for some (irreducible) Dyck words w_i which (by induction) have associated graphs G_{w_i}, constructed as described in Section 2.1, with $S_{w_i}(k) = \left\{ {G_{w_i} \atop k} \right\}$ for all $k \geq 0$. That $S_w(k) = \left\{ {G_w \atop k} \right\}$ for all $k \geq 0$ where G_w is the disjoint union of the G_{w_i}'s follows from repeated applications of Claim 3.2. $\square$

For Theorem 2.7 we utilize the chromatic polynomial. Recall that associated to a graph G there is a polynomial $\chi_G(q)$, the *chromatic polynomial*, whose value at each positive integer q is the number of proper q-colorings of G, that is, the number of functions $f : V \to \{1, \dots, q\}$ satisfying $f(u) \neq f(v)$ whenever $uv \in E$. A key observation is that for all G, $\chi_G(q)$ determines $(\left\{ {G \atop k} \right\})_{k \geq 0}$, and vice-versa. Indeed, on the one hand inclusion-exclusion gives

$$\left\{ {G \atop k} \right\} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} \chi_G(k - i), \tag{13}$$

while on the other hand

$$\chi_G(q) = \sum_{k \geq 0} \left\{ {G \atop k} \right\} q^{\underline{k}}.$$

To see this second relation, note that given a palette of q colors, for each k there are $\left\{ {G \atop k} \right\}$ ways to partition the vertex set into k non-empty color classes, and $q^{\underline{k}}$ ways to assign colors the classes. A particular consequence of (13) that we will use later is that

$$\text{if } \chi_G(q) = \chi_{G'}(q) \text{ for all } q \text{ then } \left\{ {G \atop k} \right\} = \left\{ {G' \atop k} \right\} \text{ for all } k \geq 0. \tag{14}$$

Theorem 2.7 follows immediately from (13) and the following claim that expresses the chromatic polynomial of G_w in terms of the heights of the x's in w (recall Definition 2.6 for the definition of height).

Claim 3.3. *For any Dyck word w, with m x's having heights $a_1, \dots, a_m$,*

$$\chi_{G_w}(q) = \prod_{i=1}^{m} (q - a_i).$$

Proof. We proceed by induction on the length of w, with length 2 trivial. Consider now a word w of length at least 4. If w is reducible, say $w = w_1 \dots w_k$ with each w_i an (irreducible) Dyck word, then by Claim 3.2 we have $G_w = G_{w_1} \cup \dots \cup G_{w_k}$ and so $\chi_{G_w}(q) = \prod_{i=1}^{k} \chi_{G_{w_i}}(q)$. The height of an x in w_i is the same as the height of the corresponding x in w, and so the claim follows by induction. If instead $w = xw'D$ is irreducible then by Claim 3.1, G_w is obtained from $G_{w'}$ by adding a dominating vertex, so $\chi_{G_w}(q) = q\chi_{G_{w'}}(q-1)$. The height of an x in w' is now one less than the height of the corresponding x in w, and so again the claim follows by induction. $\square$

4 Interpreting $S_w(k)$ in terms of partitions of H_w

To prove Theorem 2.8, we need only combine two old results. The first is due to Goldman et al., and forms part of their series of results on rook polynomials. An *n-board* is a subset of $\{1,\ldots,n\} \times \{1,\ldots,n\}$, and it is said to be *proper* if 1) it includes only pairs (i,j) with $i > j$, and 2) it satisfies the transitivity property that if (i,j) and (j,k) are both elements of the board, then so too is (i,k). To a proper n-board B associate a graph $\Gamma_n(B)$ on vertex set $\{1,\ldots,n\}$ by putting an edge from i to j (for $i > j$) if and only if $(i,j) \notin B$. Denote by $r_k(B)$ the number of ways of placing k non-attacking rooks on B; that is, the number of ways of selecting a subset of B of size k, with no two elements of the subset sharing a first coordinate, and no two sharing a second coordinate. The relevant result of Goldman et al. [16, Theorem 2] is that for all $k \geq 0$, $r_k(B) = \left\{ {\Gamma_n(B) \atop n-k} \right\}$. (Goldman et al. use the notation q_{n-k} for $\left\{ {\Gamma_n(B) \atop n-k} \right\}$.)

To interpret this result in the language of the present paper, let w be a Dyck word with n x's, and let $\mathcal{F}_w$ be the set of (labels of) unit squares that lie above the staircase path of w, and inside the $[0,n] \times [0,n]$ square (note that $\mathcal{F}_w$ forms what is often called a *Ferrers board*). While $\mathcal{F}_w$ does not form a proper n-board, it is easy to check that if we let $\tilde{\mathcal{F}}_w$ be the reflection of $\mathcal{F}_w$ across the line $x = y$, then $\tilde{\mathcal{F}}_w$ does, and that the graph $\Gamma_n(\tilde{\mathcal{F}}_w)$ is isomorphic to H_w (via the identity map on the labels). It is also clear that $r_k(\tilde{\mathcal{F}}_w) = r_k(\mathcal{F}_w)$. Thus Goldman et al.'s result is that for all $k \geq 0$,

$$r_k(\mathcal{F}_w) = \left\{ {H_w \atop n-k} \right\}. \tag{15}$$

The second result we need is Navon's combinatorial interpretation of $S_w(k)$ from [22]. Just as we associated a Dyck path with a Dyck word in Section 2.1, we may associate a staircase path with an arbitrary word w by starting at $(0,0)$ and, reading w from left to right, taking a step in the positive y direction each time an x is encountered in w, and a step in the positive x direction each time a D is encountered. Let B_w be the set of labels of the unit squares that lie above this staircase path and inside the box $[0,n] \times [0,m]$, where w has m x's and n D's. As before, let $r_k(B_w)$ be the number of ways of placing k non-attacking rooks on B_w. Navon's combinatorial interpretation of the numbers $S_w(k)$, as stated (and reproved) by Varvak in [28, Theorem 3.1], is that $S_w(n-k) = r_k(B_w)$. (Note that Varvak uses "U" in place of x).

It is clear that if w is a Dyck word with n x's then B_w from Navon's interpretation is exactly our $\mathcal{F}_w$, and so Navon's interpretation becomes in this case

$$S_w(n-k) = r_k(\mathcal{F}_w) \tag{16}$$

for all $k \geq 0$. Combining (15) and (16) we get Theorem 2.8.

It is also possible to give a direct proof (not using Navon's interpretation) of the identity $\left\{ {H_w \atop k} \right\} = \left\{ {G_w \atop k} \right\}$ for all Dyck words w and integers $k \geq 0$ (and so also an alternate proof of Theorem 2.8 via Theorem 2.3). We have already calculated (in Section 3) the chromatic polynomial of G_w to be

$$\chi_{G_w}(q) = \prod_{i=1}^{m} (q - a_i)$$

where w has m x's (and so also m D's, since we are assuming it to be a Dyck word), and a_i is the height of the i^{th} x. If we can show that H_w has the same chromatic polynomial, then we get $\left\{ {H_w \atop k} \right\} = \left\{ {G_w \atop k} \right\}$ using (14).

To compute $\chi_{H_w}(q)$, consider the set $\mathcal{T}_w = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ (defined in Section 2.2) of peaks of the Dyck path of w. As discussed in Section 2.2, the edge set of H_w is obtained by putting cliques on each of the consecutive integer segments $\{x_i, \ldots, y_i\}$, $1 \leq i \leq k$ (see Figure 2 in Section 2.2). We properly q-color H_w sequentially, starting with the clique on segment $\{x_1, \ldots, y_1\}$, which can be colored in $q(q-1)\ldots(q-(y_1-x_1))$ ways. Notice, by our alternate characterization from Section 3 of the heights of the x_i's in a word (if the step in the positive y direction corresponding to an x goes from (a, b) to $(a, b+1)$, then the height of that x is $b - a$), that this is the same as $\prod_{i=1}^{y_1}(q - a_i)$ (where a_i is the height of the i^{th} x).

Next we move on to the clique on segment $\{x_2, \ldots, y_2\}$. The first $y_1 - x_2 + 1$ vertices of this clique have already been colored (since they are part of the clique on segment $\{x_1, \ldots, y_1\}$), so it remains to color the last $y_2 - y_1$ vertices. The palette of colors available has size $q - (y_1 - x_2 + 1)$, so the number of ways in which these last $y_2 - y_1$ vertices of the second clique can be colored is $(q - (y_1 - x_2 + 1))(q - (y_1 - x_2 + 1) - 1)\ldots(q - (y_1 - x_2 + 1) - (y_2 - y_1 - 1))$; this is the same as $\prod_{i=y_1+1}^{y_2}(q - a_i)$. Continuing along the integer segment cliques in this manner, and noting that all proper q-colorings of H_w can be achieved by this sequential scheme, we get that indeed the number of proper q-colorings of H_w is $\prod_{i=1}^{m}(q - a_i)$.

5 A new summation formula when $w = (x^s D^s)^n$

We now turn to Theorem 2.10, which deals with the case $w = (x^s D^s)^n$. From Theorem 2.3 we know that

$$S_w(k) = \left\{ {nK_s \atop k} \right\}$$

in this case, where nK_s is the disjoint union of n copies of K_s.

The key observation that allows us to say something sensible about $\left\{ {nK_s \atop k} \right\}$ is (14), which we use in the following way. If G consists of a disjoint copies of K_b together with r isolated vertices, then its chromatic polynomial is

$$\chi_G(q) = q^{r+a}(q-1)^a(q-2)^a \ldots (q-(b-1))^a.$$

On the other hand, if G consists of a fan of a copies of K_b (a copies of K_b with a single vertex that is in common to all the copies) together with $r+a-1$ isolated vertices, then the chromatic polynomial is also

$$\chi_G(q) = q^{r+a}(q-1)^a(q-2)^a \dots (q-(b-1))^a.$$

So whenever we see a graph of the first kind described above, we may replace it with a graph of the second kind, without changing the values of the graph Stirling numbers.

As a warm-up we use this observation first to obtain an expression for $\left\{ {nK_2 \atop k} \right\}$. The chromatic polynomial of nK_2 is $q^n(q-1)^n$, which is the same as the chromatic polynomial of the graph G consisting of a star on $n+1$ vertices together with $n-1$ isolated vertices. So $\left\{ {nK_2 \atop k} \right\} = \left\{ {G \atop k} \right\}$, and we can find $\left\{ {G \atop k} \right\}$ easily: first, decide on a subset of size ℓ of the isolated vertices (perhaps empty) to be in the same block as the center of the star (which cannot be in the same block as any of the leaves of the star). The remaining vertices now form an independent set of size $2n-1-\ell$, so we use an ordinary Stirling number of the second kind to count the number of partitions of these vertices into $k-1$ classes. This leads to

$$\left\{ {nK_2 \atop k} \right\} = \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \left\{ {2n-1-\ell \atop k-1} \right\}.$$

Now we deal with the general case (with K_s replacing K_2). Here it will be convenient to define $G_{a,b,c}$ as the graph consisting of a fan of a copies of K_b together with c isolated vertices (for example, $G_{n,2,n-1}$ is the graph G that replaced nK_2 in the last paragraph).

The chromatic polynomial of nK_s is $(q(q-1)(q-2)\dots(q-(s-1)))^n$, and this is the same as the chromatic polynomial of $G_{n,s,n-1}$. To partition this graph into k non-empty independent sets, we first decide on a subset of size i_1, $0 \le i_1 \le n-1$, of the isolated vertices to be in the same block as the center of the fan (which cannot be in a block with any of the other vertices of the fan). The remaining vertices now form the following structure: n copies of K_{s-1}, together with $n-1-i_1$ isolated vertices. We may replace this with $G_{n,s-1,2(n-1)-i_1}$ without changing independent set partition counts, and so

$$\left\{ {nK_s \atop k} \right\} = \sum_{i_1=0}^{n-1} \binom{n-1}{i_1} \left\{ {G_{n,s-1,2(n-1)-i_1} \atop k-1} \right\}.$$

What is $\left\{ {G_{n,s-1,2(n-1)-i_1} \atop k-1} \right\}$? By first partnering the center vertex of the fan with some subset of size i_2 $(0 \le i_2 \le 2(n-1)-i_1)$ of the isolated vertices, then replacing the remaining structure (n disjoint copies of K_{s-2}, together with $2(n-1)-i_1-i_2$ isolated vertices) with $G_{n,s-2,3(n-1)-i_1-i_2}$, we get

$$\left\{ {G_{n,s-1,2(n-1)-i_1} \atop k-1} \right\} = \sum_{i_2=0}^{2(n-1)-i_1} \binom{2(n-1)-i_1}{i_2} \left\{ {G_{n,s-2,3(n-1)-i_1-i_2} \atop k-2} \right\}$$

and so

$$\left\{ {nK_s \atop k} \right\} = \sum_{i_1=0}^{n-1} \binom{n-1}{i_1} \sum_{i_2=0}^{2(n-1)-i_1} \binom{2(n-1)-i_1}{i_2} \left\{ {G_{n,s-2,3(n-1)-i_1-i_2} \atop k-2} \right\}.$$

Continuing this process we eventually reach

$$\begin{gathered} \left\{ {G_{n,2,(s-1)(n-1)-i_1-i_2-\ldots-i_{t-2}} \atop k-(s-2)} \right\} \\ = \\ \sum_{i_{s-1}=0}^{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}} \binom{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}{i_{s-1}} \left\{ {s(n-1)+1-i_1-i_2-\ldots-i_{s-1}} \atop {k-(s-1)} \right\}. \end{gathered}$$

(the final graph we consider on the right-hand side above is $G_{n,1,s(n-1)-i_1-i_2-\ldots-i_{s-1}}$, which is a collection of $s(n-1)+1-i_1-i_2-\ldots-i_{s-1}$ isolated vertices, so we are able to use an ordinary Stirling number to count partitions of it into $k-(s-1)$ blocks). We conclude that $\left\{ {nK_s \atop k} \right\}$ equals

$$\begin{gathered} \sum_{i_1=0}^{n-1} \binom{n-1}{i_1} \sum_{i_2=0}^{2(n-1)-i_1} \binom{2(n-1)-i_1}{i_2} \ldots \\ \sum_{i_{s-2}=0}^{(s-2)(n-1)-i_1-i_2-\ldots-i_{s-3}} \binom{(s-2)(n-1)-i_1-i_2-\ldots-i_{s-3}}{i_{s-2}} \\ \sum_{i_{s-1}=0}^{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}} \binom{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}{i_{s-1}} \left\{ {s(n-1)+1-i_1-i_2-\ldots-i_{s-1}} \atop {k-(s-1)} \right\}. \end{gathered} \tag{17}$$

Set $\ell = i_1 + i_2 + \ldots + i_{s-1}$ (so ℓ ranges from 0 to $(s-1)(n-1)$). For each ℓ in this range the coefficient of $\left\{ {s(n-1)+1-\ell} \atop {k-(s-1)} \right\}$ in (17) is

$$\sum_{i_1+i_2+\ldots+i_{s-1}=\ell} \binom{n-1}{i_1}\binom{2(n-1)-i_1}{i_2} \ldots \binom{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}{i_{s-1}}$$

(note that if ever we consider a $(s-1)$-tuple in the sum above that fails to satisfy one of the conditions $0 \leq i_j \leq j(n-1) - i_1 - i_2 - \ldots - i_{j-1}$, then the corresponding binomial coefficient will be 0). This establishes (4), and completes the proof of Theorem 2.10.

To establish (5), we must show that the right-hand side of (4) equals the right-hand side of (5), for which we employ a combinatorial argument. Let $A_1, \ldots, A_{s-1}$ be $s-1$ disjoint sets, each of size $n-1$. An *ℓ-selection* from A_1 through A_{s-1} is a specification of sets A_{11}, A_{21}, A_{22}, A_{31}, A_{32}, A_{33}, $\ldots$ $A_{(s-1)1}$, $\ldots$, $A_{(s-1)(s-1)}$, pairwise disjoint, with $A_{11} \subseteq A_1$, $A_{21} \cup A_{22} \subseteq A_2$, $\ldots$, $A_{(s-1)1} \cup \ldots \cup A_{(s-1)(s-1)} \subseteq A_{s-1}$, and $|A_{11}| + |A_{21}| + \ldots + |A_{(s-1)(s-1)}| = \ell$.

To count the number of ℓ-selections, we first specify a composition $\ell = i_1+i_2+\ldots+i_{s-1}$, then from each A_k select a subset of size i_k, and then for each element of the chosen subset, decide which of $A_{k1}, \ldots A_{kk}$ the element belongs to. This gives that the number of ℓ-selections is

$$\sum_{i_1+\ldots+i_{s-1}=\ell} \prod_{k=1}^{s-1} k^{i_k} \binom{n-1}{i_k} = [x^\ell] \left((1+x)(1+2x)\ldots(1+(s-1)x)\right)^{n-1}.$$

Another way to count ℓ-selections is to first specify a composition $\ell = i_1+i_2+\ldots+i_{s-1}$. Then, select from A_{s-1} a subset of size i_1 to be $A_{(s-1)(s-1)}$. Next, select a subset of $(A_{s-2} \cup A_{s-1}) \setminus A_{(s-1)(s-1)}$ of size i_2; let its intersection with A_{s-1} be $A_{(s-1)(s-2)}$, and its intersection with A_{s-2} be $A_{(s-2)(s-2)}$. Next, select a subset of $(A_{s-3} \cup A_{s-2} \cup A_{s-1}) \setminus$

$(A_{(s-1)(s-1)} \cup A_{(s-1)(s-2)} \cup A_{(s-2)(s-2)})$ of size i_3; let its intersection with A_{s-1} be $A_{(s-1)(s-3)}$, its intersection with A_{s-2} be $A_{(s-2)(s-3)}$ and its intersection with A_{s-3} be $A_{(s-3)(s-3)}$. Continue in this manner, until finally we are selecting a subset of size i_1 of the as-yet-unselected elements of $A_1 \cup \ldots \cup A_{s-1}$; for each k, $1 \leq k \leq s-1$, let its intersection with A_k be A_{k1}. This gives that the number of ℓ-selections is

$$\sum_{i_1+i_2+\ldots+i_{s-1}=\ell} \binom{n-1}{i_1}\binom{2(n-1)-i_1}{i_2}\cdots\binom{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}{i_{s-1}}.$$

This completes the verification of (5).

6 Bijections involving increasing forests

Recall that for the word $w = (x^r D)^n$, we set $G_w = G(n,r)$ and $H_w = H(n,r)$; in this case $G(n,r)$ is the threshold graph obtained from the empty graph by iterating n times the operation of adding an isolated vertex followed by $r-1$ dominating vertices, and $H(n,r)$ is the indifference graph on vertex set $\{1,\ldots,rn\}$ with edges covered by cliques on vertices i through ir, for $1 \leq i \leq n$. In this section we will exhibit bijections between the set of partitions $G(n,r)$ (and $H(n,r)$) into $k+(r-1)n$ independent sets and the set of *decreasing* r-ary forests with n vertices and k components (where decreasing means that the labels go in decreasing order when read along any path starting from the root vertex of a tree). Decreasing forests turn out to be notationally a little easier to deal with than increasing forests, but of course via the order-reversing permutation $x \mapsto n+1-x$ there is a perfect correspondence between the two.

6.1 Bijection for $G(n,r)$

Here we provide a bijective proof that

$$\left\{ {G(n,r) \atop k+(r-1)n} \right\} = |F'(r,n,k)|,$$

where $F'(r,n,k)$ is the set of decreasing r-ary forests with n vertices and k components. It will be useful to exhibit the threshold graph $G(n,r)$ concisely. To that end, we will use '+' to indicate the addition of a dominating vertex and '$\bullet$' to indicate the addition of an isolated vertex, presenting the vertices from right to left (see Figure 3). In particular, +'s are adjacent to every vertex to their right and also every + to their left, while $\bullet$'s are only adjacent to the +'s to their left. In the concise representation of $G(n,r)$, label the vertices in increasing order from left to right; we will often abuse notation and identify a vertex with its label.

We also note in passing that $G(n,r)$ may be constructed from the half-graph (on $2n$ vertices $x_1,\ldots,x_n,y_1,\ldots,y_n$ with x_i joined to y_j when $j \leq i$), by adding edges of the form y_iy_j for $i \neq j$ and blowing up each y_i to a copy of K_{r-1} (see Figure 4).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
+	+	+	•	+	+	+	•	+	+	+	•	+	+	+	•	+	+	+	•

Figure 3: The concise notation for the graph $G(5,4)$, with labels.

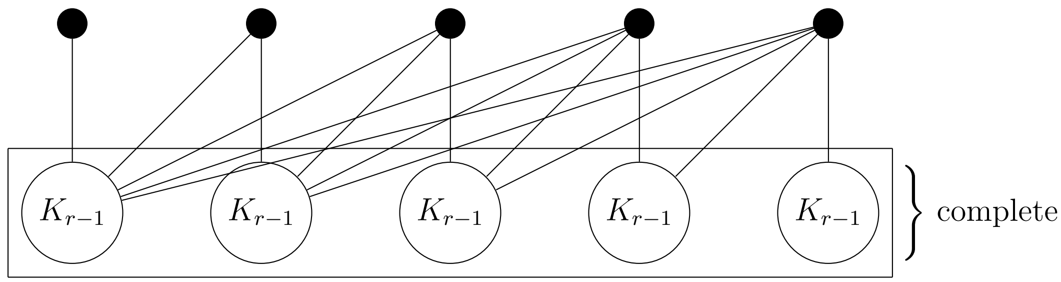


Figure 4: The graph $G(5,r)$. An edge from a vertex on top to a K_{r-1} below indicates that the vertex on top is joined to all vertices of K_{r-1}.

Let $\mathcal{P}$ be the set of partitions of $G(n,r)$ into $k+(r-1)n$ independent sets. We begin by describing a map from $\mathcal{P}$ to $F'(r,n,k)$. Fix $P \in \mathcal{P}$, which corresponds to a partition of $\{1,\dots,rn\}$ via the labels on the vertices. Note that the vertices in any subset of $V(G(n,r))$, and so in particular the independent sets that make up P, are totally ordered by the labeling of the vertices. We will iteratively build a decreasing r-ary forest from P.

Let M denote the set of vertices of the form ar for $1 \leq a \leq n$ (these are the $\bullet$'s from the concise representation of $G(n,r)$). The vertices not in M, together with vertex rn, form a clique in $G(n,r)$ and so must be in distinct independent sets in P. This means that there are k independent sets in P that contain only elements from M. For each of these k independent sets we place a root in the forest with label j, where jr is the largest label in the corresponding independent set.

We use the location of the remaining vertices in M to iteratively construct the rest of the forest, and we do so by considering these $n-k$ vertices in decreasing order. Suppose that xr is the largest vertex in M that has not been considered, meaning in particular that xr is not the largest vertex in an independent set consisting of elements of M. Among the elements of xr's independent set with label larger than xr, let z be the element with the smallest label. (Notice that such a z must exist: xr must be in an independent set with either a vertex from M with a larger label, or some vertex outside of M, which would necessarily have a larger label than xr since xr is adjacent to all vertices outside of M with smaller labels.)

There are unique positive integers $1 \leq a \leq n$ and $1 \leq p \leq r$ such that $z=(a-1)r+p$. We place a vertex labeled x in the forest in the slot reserved for the p^{th} child of the vertex labeled a. (Notice that a has indeed been placed in the forest already, since

$xr < z = (a-1)r + p \leq ar$ and we consider the elements of M in decreasing order.)

Since distinct z's are associated to distinct xr's, and the pair (a, p) is uniquely determined by z, we will never be forced to place distinct vertices in the same location as this iterative process goes on, and so the final result will indeed be an n-vertex, k-component, r-ary forest; and the fact that $x < a$ and x is placed below a in the forest means that the labeling is decreasing.

The whole process is reversible. We can see this by considering the vertices of the forest in order of increasing labels (from 1 to n). For each label x we can use p and a (which are known from the location of x in the forest) to obtain z, which is the next largest vertex in the independent set containing xr. In this way, we recover the partition of $G(n, r)$ into $k + (r-1)n$ independent sets that led to the forest. This shows that our map is indeed a bijection.

6.2 Bijection for $H(n, r)$

We now provide a bijective proof that

$$\left\{ \begin{matrix} H(n,r) \\ k + (r-1)n \end{matrix} \right\} = |F'(r, n, k)|,$$

where $H(n, r)$ is the graph on vertex set $\{1, \ldots, rn\}$ with edges covered by cliques on vertices i through ir, for $i = 1, \ldots, n$ (see Figure 5).

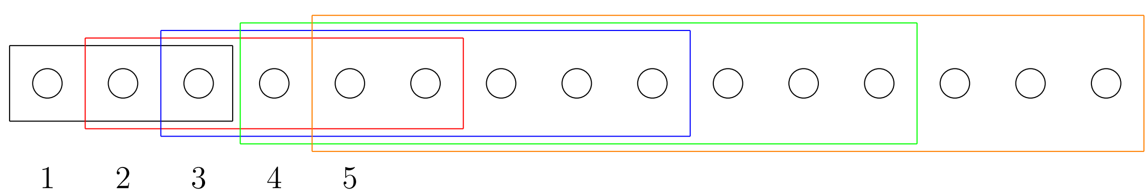


Figure 5: The graph $H(5, 3)$.

Let $\mathcal{P}'$ be the set of partitions of $H(n, r)$ into $k + (r-1)n$ independent sets. As before, we start by describing a map from $\mathcal{P}'$ to $F'(r, n, k)$. Fix $P' \in \mathcal{P}'$. This corresponds to a partition of $\{1, \ldots, rn\}$ via the labels on the vertices (we will again abuse notation and refer to the vertices by their labels). Place the vertices of each independent set into an increasing ordered list. We will iteratively build a decreasing r-ary forest from P'.

If $i \leq n$ is the largest label in an independent set, put a root in the forest with label i. This gives k roots for the $k + (r-1)n$ independent sets. We will consider the remaining vertices in $\{1, \ldots, n\}$ in decreasing order.

Suppose that x is the largest vertex that has not been placed in the forest, and suppose that x immediately precedes y in the list for some independent set (so $x < n$, and $y > xr$). Then find the unique positive integers $2 \leq a \leq n$ and $1 \leq p \leq r$ satisfying $y = (a-1)r + p$. Place a vertex labeled x in the p^{th} possible position (again scanning from left to right, say) below the vertex labeled a in the forest. Using the increasing order on the independent set assures that we do not try to place two labels in the same slot beneath a vertex in the forest, since each y will only be considered once and each y is in bijective correspondence with a pair (a, p).

Now x and y are in the same independent set (and no vertex with a label between that of x and y is in the independent set), so $rx < y = (a-1)r + p \leq ra$. This implies that $x < a$. Therefore this iterative procedure produces a decreasing r-ary forest with k-components. As with the map described in Section 6.1, this process is reversible. We see this by considering the vertices of a forest in order of increasing labels. For each x we can use p and a (which are known from the location of x in the forest) to obtain y, which is the next label in the independent set containing x. In this way, we recover the partition of $H(n, r)$ into $k + (r-1)n$ independent sets that led to the forest, showing that our map is a bijection.

7 q-analogs

The proof of Theorem 2.12, like that of Theorem 2.3, is based on two claims.

Claim 7.1. *Let $w = xw'D$ be an irreducible Dyck word in the q-deformed Weyl algebra, with w' non-empty. For all $k \geq 0$ we have*

$$S_w^q(k) = S_{w'}^q(k-1),$$

where $S_{w'}^q(-1)$ is interpreted as 0.

Proof. We have $w = \sum_{k \geq 0} S_w^q(k) x^k D^k$, but also

$$\begin{aligned} w &= xw'D \\ &= x\left(\sum_{k\geq 0} S_{w'}^q(k-1)x^{k-1}D^{k-1}\right)D \\ &= \sum_{k\geq 0} S_{w'}^q(k-1)x^k D^k. \end{aligned}$$

The claim now follows by the uniqueness of the representation in (8). □

For the second claim, we define the q-factorial (for integer a) via

$$[a]_q! = \begin{cases} [a]_q[a-1]_q[a-2]_q \dots [2]_q[1]_q & \text{if } a \geq 1, \\ 1 & \text{if } a = 0, \text{ and} \\ 0 & \text{if } a < 0, \end{cases}$$

where recall that for positive integer m, $[m]_q = 1 + q + \ldots + q^{m-1}$, and define the q-binomial coefficient (for integers a and b) via

$$\binom{a}{b}_q = \frac{[a]_q!}{[b]_q![a-b]_q!}.$$

Claim 7.2. *Let w_1, w_2 be Dyck words in the q-deformed Weyl algebra. For all $k \geq 0$ we have*

$$S^q_{w_1 w_2}(k) = \sum_{r,s\geq 0} S^q_{w_1}(r) S^q_{w_2}(s) q^{(r-k)(s-k)} \binom{r}{r+s-k}_q \binom{s}{r+s-k}_q [r+s-k]_q!$$

where $w_1 w_2$ is the concatenation of w_1 and w_2.

Proof. We begin by deriving the following analog of (11) in the q-deformed Weyl algebra:

$$x^r D^r x^s D^s = \sum_{j\geq 0} q^{(r-j)(s-j)} \frac{([r]_q)^{\underline{j}} ([s]_q)^{\underline{j}}}{[j]_q!} x^{r+s-j} D^{r+s-j}. \tag{18}$$

We prove this by induction on $r+s$. If $rs = 0$ then the result is trivial. So, consider a pair (r, s) with $r + s \geq 2$ and $rs > 0$. Replace the right-most Dx in $x^r D^r x^s D^s$ with $qxD + 1$ to obtain a sum of two terms, one of which still has $r + s$ x's and $r + s$ D's. In this term, again replace the right-most Dx with $qxD + 1$. Iterating this operation r times we obtain

$$x^r D^r x^s D^s = x \left(q^r x^r D^r x^{s-1} D^{s-1} + [r]_q x^{r-1} D^{r-1} x^{s-1} D^{s-1} \right) D.$$

By induction, the right-hand side above is the sum of

$$\sum_{j\geq 0} q^{(r-j)(s-1-j)+r} \frac{([r]_q)^{\underline{j}} ([s-1]_q)^{\underline{j}}}{[j]_q!} x^{r+s-j} D^{r+s-j}$$

and

$$\sum_{j\geq 0} [r]_q q^{(r-1-j)(s-1-j)} \frac{([r-1]_q)^{\underline{j}} ([s-1]_q)^{\underline{j}}}{[j]_q!} x^{r+s-j-1} D^{r+s-j-1}.$$

The coefficient of $x^{r+s} D^{r+s}$ in the normal order of $x^r D^r x^s D^s$ comes from the $j = 0$ term of the first sum above, and is q^{rs}, as required. For $j > 0$ the coefficient of $x^{r+s-j} D^{r+s-j}$ is

$$q^{(r-j)(s-1-j)+r} \frac{([r]_q)^{\underline{j}} ([s-1]_q)^{\underline{j}}}{[j]_q!} + [r]_q q^{(r-j)(s-j)} \frac{([r-1]_q)^{\underline{j-1}} ([s-1]_q)^{\underline{j-1}}}{[j-1]_q!}$$

which, after a little algebra, simplifies to

$$q^{(r-j)(s-j)} \frac{([r]_q)^{\underline{j}} ([s]_q)^{\underline{j}}}{[j]_q!},$$

completing the induction. Armed with (18), we complete the proof of the claim exactly as in the derivation of (12). □

We also need the following lemma. Let $K_{r,s}$ be the complete bipartite graph with partition classes $X = \{x_1, \ldots, x_r\}$ and $Y = \{y_1, \ldots, y_s\}$. Let $\mathcal{M}_k(K_{r,s})$ be the set of matchings of size k in $K_{r,s}$ (selections of k edges from $K_{r,s}$, no two sharing an endvertex).

Encode a matching $M \in \mathcal{M}_k(K_{r,s})$ by an r by s matrix (which we also call M) whose ij entry is 1 if $\{x_i, y_j\}$ is in the matching, and 0 otherwise, and let $f(M)$ be as in Definition 2.11; that is, $f(M)$ is the number of unmarked 0's in M after marking all 0's that occur below or to the right of a 1.

Lemma 7.3. *With the notation as above, for all $k \geq 0$ we have*

$$\sum_{M \in \mathcal{M}_k(K_{r,s})} q^{f(M)} = q^{(r-k)(s-k)} \binom{r}{k}_q \binom{s}{k}_q [k]_q!. \tag{19}$$

Notice that when $q = 1$, (19) reduces to $|\mathcal{M}_k(K_{r,s})| = \binom{r}{k}\binom{s}{k}(k)!$, which is evident: to specify a matching of size k one first chooses subsets of X and Y, each of size k (there are $\binom{r}{k}\binom{s}{k}$ ways to do this) and then chooses how to match them up (there are $k!$ ways to do this).

Proof. (Lemma 7.3) We use the following well-known interpretations of $\binom{r}{k}_q$ and $[k]_q!$ (see for example [1, Theorem 3.6] for a statement that encompasses both). First, if $\mathbb{Q}_r^{(k)}$ is the set of 0-1 strings of length r that have exactly k 1's, and for each such string σ, $\mathrm{zeros}(\sigma)$ is the sum, over all the 1's in σ, of the number of 0's to the left of that 1, then

$$\binom{r}{k}_q = \sum_{\sigma \in \mathbb{Q}_r^{(k)}} q^{\mathrm{zeros}(\sigma)}.$$

Second, if $\mathbb{P}_k$ is the set of permutations of $\{1, \ldots, k\}$ (written in one-line notation), and for each such permutation π, $\mathrm{inv}(\pi)$ counts the number of inversions in π (the number of pairs (i, j) with $i < j$ such that j appears before i in π), then

$$[k]_q! = \sum_{\pi \in \mathbb{P}_k} q^{\mathrm{inv}(\pi)}.$$

From these interpretations we see that $q^{(r-k)(s-k)}\binom{r}{k}_q\binom{s}{k}_q[k]_q!$ expands out to a sum of monomials of the form $q^{(r-k)(s-k)+\mathrm{zeros}(\sigma)+\mathrm{zeros}(\tau)+\mathrm{inv}(\pi)}$ where $\sigma \in \mathbb{Q}_r^{(k)}$, $\tau \in \mathbb{Q}_s^{(k)}$ and $\pi \in \mathbb{P}_k$, with exactly one such monomial for each triple (σ, τ, π). We will prove (19) by exhibiting a bijection φ from $\mathcal{M}_k(K_{r,s})$ to $\mathbb{Q}_r^{(k)} \times \mathbb{Q}_s^{(k)} \times \mathbb{P}_k$ with the property that if $\varphi(M) = (\sigma, \tau, \pi)$ then $f(M) = (r-k)(s-k) + \mathrm{zeros}(\sigma) + \mathrm{zeros}(\tau) + \mathrm{inv}(\pi)$.

The map φ is defined as follows. For each $M \in \mathcal{M}_k(K_{r,s})$, we let σ be the 0-1 string of length r that has a 1 in the i^{th} position exactly when row i of M has a 1 (note that there are exactly k 1's in σ) and we let τ be the 0-1 string of length s that has a 1 in the j^{th} position exactly when column j of M has a 1 (again, this is a vector with exactly k 1's). Finally, to define π, we let M' be the (unique) k by k submatrix of M that includes all k of the 1's. Relabeling the rows and columns of M' by 1 through k in the natural way, let the locations of the 1's in M' be $(i_1, 1), (i_2, 2), \ldots, (i_k, k)$. The permutation π is then $i_1 \ldots i_k$.

For example, suppose that $r = 7$, $s = 6$, $k = 4$ and the matching M consists of the edges $\{x_2, y_3\}$, $\{x_4, y_4\}$, $\{x_5, y_1\}$ and $\{x_6, y_6\}$ (see Figure 6; the marked entries in M are identified with a "$\star$"). We have $f(M) = 19$, $\sigma = 0101110$, $\tau = 101101$, $\pi = (3, 1, 2, 4)$, $(r-k)(s-k) = 6$, $\text{zeros}(\sigma) = 7$, $\text{zeros}(\tau) = 4$ and $\text{inv}(\pi) = 2$.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \star & \star & \star \\ 0 & 0 & \star & 0 & 0 & 0 \\ 0 & 0 & \star & 1 & \star & \star \\ 1 & \star & \star & \star & \star & \star \\ \star & 0 & \star & \star & 0 & 1 \\ \star & 0 & \star & \star & 0 & \star \end{pmatrix}$$

Figure 6: An example matrix M, with the marked 0's in M identified with a "$\star$."

We first note that φ is a bijection from $\mathcal{M}_k(K_{r,s})$ to $\mathbb{Q}_r^{(k)} \times \mathbb{Q}_s^{(k)} \times \mathbb{P}_k$, since the location of M' can be reconstructed from σ and τ, and the exact location of the 1's in M is then determined by π.

To see $f(M) = (r-k)(s-k) + \text{zeros}(\sigma) + \text{zeros}(\tau) + \text{inv}(\pi)$, first note that there are exactly $(r-k)(s-k)$ entries ij in M, all unmarked 0's, with no 1 in row i and no 1 in column j (these entries are labeled "a" in the matrix in Figure 7). Call this set of entries M_0.

$$\begin{pmatrix} b & a & b & b & a & b \\ d & c & \bullet & \bullet & \bullet & \bullet \\ b & a & \bullet & b & a & b \\ d & c & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & c & \bullet & \bullet & c & \bullet \\ \bullet & a & \bullet & \bullet & a & \bullet \end{pmatrix}$$

Figure 7: The unmarked 0's of M, partitioned into M_0 (a's), M_σ (b's), M_τ (c's), and M' (d's).

Next, consider all of the entries in M that lie above (and in the same column as) a 1, but are not in M' (these entries are labeled "b" in Figure 7). They are all unmarked 0's (being above a 1, each such entry is not below a 1, and being outside of M', it is not to the right of a 1), and there are exactly $\text{zeros}(\sigma)$ of them. Call this set of entries M_σ; note that M_0 and M_σ are disjoint.

Next, consider all of the entries in M that lie to the left (and in the same row as) a 1, but are not in M' (these entries are labeled "c" in Figure 7). They are all unmarked 0's (being to the left of a 1, each such entry is not to the right of a 1, and being outside of M', it is not below a 1), and there are exactly $\text{zeros}(\tau)$ of them. Call this set of entries

M_τ; note that M_0 and M_τ are evidently disjoint, and that also M_σ and M_τ are disjoint — an entry in the intersection of M_σ and M_τ would have to be both above a 1 and to the left of a 1, and so be in M'.

Note that $M_0 \cup M_\sigma \cup M_\tau$ is exactly the set of unmarked 0's outside of M', so we are done if we can show that the number of unmarked 0's in M' (entries labeled "d" in Figure 7) is $\mathrm{inv}(\pi)$. Consider such an unmarked 0, at position (i, j). It must be located above some 1, at position (i', j), say, and to the left of some other 1, at position (i, j'), say. This gives rise to a pair of 1's at positions (i', j) and (i, j'), with $j < j'$ and $i < i'$, which is an instance of an inversion in π; and conversely, any inversion in π is easily seen to correspond to an unmarked 0 in M'. □

To prove Theorem 2.12, we proceed by induction on the length of w. If w has length 2 then $w = xD$, $G_w = K_1$, and the result is trivial.

If w is irreducible and of length greater than 2, then $w = xw'D$ for some Dyck word w. By induction, we have that for each k,

$$\sum_{P' \in \mathcal{P}(w', k-1)} q^{\mathrm{wt}(P')} = S^q_{w'}(k-1).$$

There is a one-to-one correspondence between $\mathcal{P}(w', k-1)$ and $\mathcal{P}(w, k)$ (as discussed in Section 2.5), and for each $P' \in \mathcal{P}(w', k-1)$ with corresponding partition $P \in \mathcal{P}(w, k)$ we have, by definition, $\mathrm{wt}(P) = \mathrm{wt}(P')$. So, using Claim 7.1,

$$\begin{aligned} \sum_{P \in \mathcal{P}(w', k-1)} q^{\mathrm{wt}(P)} &= \sum_{P' \in \mathcal{P}'(w', k-1)} q^{\mathrm{wt}(P')} \\ &= S^q_{w'}(k-1) \\ &= S^q_w(k) \end{aligned}$$

as required.

There remains the case where $w = w_1 w_2 \dots w_\ell$ is reducible (with each w_i irreducible). We assume $\ell = 2$; this will not lose us any generality, but will allow us to simplify notation by writing "w_2" for "$w_2 \dots w_\ell$" everywhere. Combining Claim 7.2, the induction hypothesis, and Lemma 7.3, what we need to show is that, for each k,

$$\sum_{P \in \mathcal{P}(w_1 w_2, k)} q^{\mathrm{wt}(P)} = \sum_{r,s \geq 0,\ P_1 \in \mathcal{P}(w_1, r),\ P_2 \in \mathcal{P}(w_2, s),\ M \in \mathcal{M}_{r+s-k}(K_{r,s})} q^{\mathrm{wt}(P_1) + \mathrm{wt}(P_2) + f(M)}$$

(note that $(r - (r + s - k))(s - (r + s - k)) = (r - k)(s - k)$). But this is immediate: we obtain all partitions of $G_{w_1 w_2}$ into k non-empty independent sets (and nothing more) by selecting integers r and s, selecting a partition P_1 of G_{w_1} into r non-empty independent sets and a partition P_2 of G_{w_2} into s non-empty independent sets, and selecting a matching M of size $r + s - k$ from $K_{r,s}$ which determines how P_1 and P_2 are merged; and by definition the weight $\mathrm{wt}(P)$ of such a matching produced by selecting a particular r, s, P_1, P_2 and M is equal to $\mathrm{wt}(P_1) + \mathrm{wt}(P_2) + f(M)$.

References

[1] G. Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, 1976.

[2] M. Arik and D. Coon, Hilbert spaces of analytic functions and generalized coherent states, *J. Mathematical Phys.* **17** (1976), 524-527.

[3] W. Asakly, T. Mansour and M. Schork, Representing elements of the Weyl algebra by labeled trees, *J. Math. Phys.* **54** (2013), 023514.

[4] G. Birkhoff, A determinant formula for the number of ways of coloring a map, *Ann. of Math.* **14** (1912/13), 42–46.

[5] P. Blasiak, G. Duchamp, A. Horzela, K. Penson and A. Solomon, Heisenberg-Weyl algebra revisited: combinatorics of words and paths, *Journal of Physics A: Mathematical and Theoretical* **41** (2008), 415204.

[6] P. Blasiak and P. Flajolet, Combinatorial models of creation-annihilation, *Seminaire Lotharingien de Combinatoire* **65** (2011), Article B65c.

[7] P. Blasiak, A. Horzela, K. Penson, A. Solomon and G. Duchamp, Combinatorics and Boson normal ordering: A gentle introduction, *American Journal of Physics* **75** (2007), 639–646.

[8] P. Blasiak, K. Penson and A. Solomon, The general boson normal ordering problem, *Physics Letters A* **309** (2003), 198-205.

[9] P. Blasiak, K. Penson and A. Solomon, The Boson Normal Ordering Problem and Generalized Bell Numbers, *Annals of Combinatorics* **7** (2003), 127–139.

[10] L. Carlitz, On a class of finite sums, *The American Mathematical Monthly* **37** (1930), 472–479.

[11] L. Carlitz, On Arrays of Numbers, *Amer. J. Math.* **54** (1932), 739-752.

[12] L. Carlitz, q-Bernoulli numbers and polynomials, *Duke Math. J.* **15** (1948), 987-1000.

[13] L. Comtet, Une formule explicite pour les puissances successives de l'opérateur de dérivation de Lie, *C. R. Acad. Sci. Paris Sér. A-B* **276** (1973), A165-A168.

[14] P. Codara, O. M. D'Antona and P. Hell, A simple combinatorial interpretation of certain generalized Bell and Stirling numbers, *Discrete Math.* **318** (2014), 53–57.

[15] D. Gebhard and B. Sagan, A chromatic symmetric function in noncommuting variables, *J. Algebraic Combin.* **13** (2001), 227–255.

[16] J. Goldman, J. Joichi and D. White, Rook Theory III. Rook polynomials and the Chromatic structure of graphs, *J. Combin. Th. Ser. B* **25** (1978), 135–142.

[17] M. Griffiths, Sitting Down for Dinner at a Twin Convention, *Mathematics Magazine* **86** (2013), 26–38.

[18] W. Lang, On Generalizations of the Stirling Number Triangles, *Journal of Integer Sequences* **3** (2000), Article 00.2.4.

[19] W. Lang, Combinatorial Interpretation of Generalized Stirling Numbers, *Journal of Integer Sequences* **12** (2009), Article 09.3.3.

[20] S.-M. Ma, T. Mansour and M. Schork, Normal ordering problem and the extensions of the Stirling grammar, *Russ. J. Math. Phys.* **21** (2014), 242–255.

[21] M. Mendez, P. Blasiak and A. Penson, Combinatorial approach to generalized Bell and Stirling numbers and boson normal ordering problem, *J. Math. Phys.* **46** (2005), 083511.

[22] A. Navon, Combinatorics and fermion algebra, *Nuovo Cimento* **16** (1973) 324-330.

[23] H. Scherk, *De evolvenda functione* $(yd.yd.yd\ldots ydX)/dx^n$ *disquisitiones nonnullae analyticae*, Ph.D. thesis, University of Berlin, 1823.

[24] M. Schork, On the combinatorics of normal ordering bosonic operators and deformations of it, *J. Phys. A: Math. Gen.* **36** (2003), 4651-4665.

[25] J. Stirling, *Methodus differentialis, sive tractatus de summatione et interpolatione serierum infinitarum*, London, 1730.

[26] I. Tomescu, *Méthods combinatoires dans le théorie des automates finis*, Ph.D. thesis, Bucarest, 1971.

[27] A. Solomon, G. Duchamp, P. Blasiak, A. Horzela and K. Penson, Normal order: combinatorial graphs, *Proc. 3rd Int. Symp. on Quantum Theory and Symmetries (Cincinnati 2003)* (Singapore: World Scientific), 527-536.

[28] A. Varvak, Rook numbers and the normal ordering problem, *J. Combin. Theory Ser. A* **112** (2005), 292-307.

[29] J-H. Yan, J.-J. Chen and G. Chang, Quasi-threshold graphs, *Discrete Applied Mathematics* **69** (1996), 247–255.