Combinatorially interpreting generalized Stirling numbers

John Engbers*  David Galvin†  Justin Hilyard‡

August 13, 2013

Abstract

Let $w$ be a word in alphabet \{x, D\}. Interpreting “$x$” as multiplication by $x$, and “$D$” as differentiation with respect to $x$, the identity

$$wf(x) = x^{(#(x’s \text{ in } w) - #(D’s \text{ in } w))} \sum_k S_w(k) x^k D^k f(x),$$

valid for any smooth function $f(x)$, defines a sequence $(S_w(k))_k$, the terms of which we refer to as the Stirling numbers (of the second kind) of $w$. The nomenclature comes from the fact that when $w = (xD)^n$, we have $S_w(k) = \{n\}_{k}$, the ordinary Stirling number of the second kind.

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 retains the spirit of the familiar interpretation of $\{n\}_{k}$ as a count of partitions. Specifically, we associate to each $w$ a 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 also discuss some relatives of, and consequences of, our interpretation.

1 Introduction

The Stirling number of the second kind, $\{n\}_{k}$, counts the number of ways of partitioning a set of $n$ elements into $k$ non-empty classes. Among the many algebraic identities satisfied by the Stirling numbers is the following, probably first observed by Scherk in his 1823 thesis [17]. With $x$ interpreted as multiplication by $x$, and $D$ as differentiation with respect to $x$, for $n \geq 0$ and all infinitely differentiable functions $f(x)$ we have

$$(xD)^n f(x) = \sum_k \{n\}_{k} x^k D^k f(x)$$

---

* jengbers@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556.
† dgalvin1@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556.
‡ jhilyard@nd.edu; Department of Mathematics, University of Notre Dame, Notre Dame IN 46556.
More generally, for each word $w$ in alphabet $\{x, D\}$, with $w$ having $m$ $x$’s and $n$ $D$’s, we have a unique expansion of the form

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

with both sides being viewed as operators on a space of infinitely differentiable functions (one may easily verify (2) by induction on the length of the word $w$). The integer sequence $(S_w(k))_k$ that arises from (2) is what we call the Stirling sequence (of the second kind) of $w$. The study of $S_w(k)$, both for general $w$ and for special cases, has a long history; see for example Scherk [17], Carlitz [6, 7], Comtet [9], Lang [13, 14], Navon [16], Varvak, [21], Schork [18], and Blasiak and various coauthors [4, 5, 1, 2, 15, 19]. These last references all come from the quantum mechanics community, where the problem of determining the $S_w(k)$ is referred to as the normal ordering problem; see [3] for a discussion of this perspective.

A number of these references — including Navon [16] and Mendez et al. [15] — give combinatorial interpretations of $S_w(k)$. Navon’s interpretation is in terms of the placement of non-attacking rooks on a certain Ferrers board associated with $w$, while that of Mendez et al. is in terms of certain generalized tree structures.

Here, we give 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 $w = (xD)^n$, the graph $G_w$ turns out to be the empty graph on $n$ vertices and we recover the usual combinatorial interpretation of $\binom{n}{k}$. In the case $w = (x^sD^s)^n$ the graph $G_w$ turns out to be $nK_s$, the disjoint union of $n$ copies of the complete graph $K_s$, and we recover a recent result of Codara, D’Antona and Hell [10]. 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.

In Section 2 we describe the construction of $G_w$ and establish that $S_w(k)$ enumerates restricted partitions of the vertex set of $G_w$. In Section 3 we discuss some consequences and further combinatorial interpretations, with some proofs deferred to a forthcoming fuller version of this paper.

2 A new combinatorial interpretation in terms of the graph $G_w$

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:

- it has the same number of $x$’s as $D$’s, and
- reading the word from left to right, every initial segment has at least as many $x$’s as $D$’s.

For Dyck words of length at least 4, we say that word is irreducible if, on deletion of a leading $x$ and a trailing $D$, the resulting word is a Dyck word, and say that it is reducible otherwise.
Note that an irreducible word is of the form $xw'D$ where $w'$ is a Dyck word, and a reducible word may be written (in a unique way) as $w_1 \ldots w_\ell$, where each $w_i$ is either an irreducible Dyck word, or the word $xD$.

Recall that 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 $D$. Irreducible Dyck words correspond to Dyck paths (involving at least 4 steps) 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 a Dyck word $w$ a graph $G_w$ inductively, as follows:

1. For $w = xD$, $G_w = K_1$ (the isolated vertex).
2. If $w = xw'D$ is irreducible, then $G_w = G_{w'} + K_1$ (the graph obtained from $G_{w'}$ by adding a dominating vertex).
3. If $w = w_1 \ldots w_\ell$ is reducible, then $G_w = G_{w_1} \cup \ldots \cup G_{w_\ell}$ (the disjoint union of the $G_{w_i}$’s).

For example, if $w = (xD)^n$ then $G_w$ is the empty graph on $n$ vertices; if $w = (x^sD^s)^n$ then $G_w = nK_s$, the disjoint union of $n$ copies of the complete graph $K_s$, and if $w = xxDXDD$ then $G_w$ is a path on three vertices. We note in passing that $G_w$ is always an instance of a quasi-threshold, or trivially perfect graph (see e.g. [8]).

To state our main theorem, we generalize the symbol $\genfrac{[}{]}{0pt}{}{n}{k}$ to graphs. Given a graph $G$ and an integer $k$, we use $\genfrac{[}{]}{0pt}{}{G}{k}$ to denote 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 $G$. This is equivalent to partitioning $G$ into non-empty independent sets, that is, sets of mutually non-adjacent vertices; it is also equivalent to properly coloring the vertices of $G$ using $k$ colors, with the names of the colors not mattering, and with each color being used on at least one vertex. We refer to $\genfrac{[}{]}{0pt}{}{G}{k}$ as a graphical Stirling number; these numbers were first introduced (to the best of our knowledge) by Tomescu [20].

**Theorem 2.2.** Let $w$ be Dyck word in the alphabet $\{x, D\}$. For all integers $k$ we have

$$S_w(k) = \genfrac{[}{]}{0pt}{}{G_w}{k}.$$

Our proof of Theorem 2.2 will be direct, in the sense that we do not rely on previously known formulae or combinatorial interpretations for $S_w(k)$.

In the case $w = (x^tD^t)^n$ we get a particularly appealing combinatorial interpretation. Here we have $G_w = K_{t_1} \cup \ldots \cup K_{t_n}$, the disjoint union of cliques of various sizes. Theorem 2.2 gives that $S_w(k)$ in this cases counts the number partitions of this union of cliques into $k$ non-empty independent sets (the case when all $t_i = 2$ is observed in [2], in slightly different language, and the case of all $t_i = 1$ for general $t$ has appeared recently in [10]). Note that when all $t_i = t$, this may be whimsically interpreted as the number of ways of breaking up a group of $n$ sets of $t$-tuplets into $k$ non-empty groups, in such a way that no class 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 [12]; our work on these notes began after reading that paper.

Theorem 2.2 can easily be extended to give a combinatorial interpretation of $S_w(k)$ when $w$ is an arbitrary word. Indeed, given an arbitrary $w$ 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. 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, and that $w = x^{m-n} \sum_k S_w(k) x^k D^k$. We have

$$x^a w D^b = x^a \left( x^{m-n} \sum_k S_w(k) x^k D^k \right) D^b = x^{a+m-n} \sum_k S_w(k) x^k D^{k+b} = \sum_k S_w(k) x^{k+b} D^{k+b},$$

the last equality using $a + m = b + n$. Since also $x^a w D^b = \sum_k S_{x^a w D^b}(k) x^k D^k$, we get the identity $S_w(k) = S_{x^a w D^b}(k + b)$, and so the following is an immediate corollary of Theorem 2.2.

**Corollary 2.3.** Let $w$ be an arbitrary word in the alphabet $\{x, D\}$, and let $x^a w D^b$ be its associated Dyck word, as in the discussion above. For all integers $k$ we have

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

The proof of Theorem 2.2 depends on the following two claims:

**Claim 2.4.** Let $w$ be a word in the alphabet $\{x, D\}$, and let $G$ be a graph with the property that for all $k$,

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

Let $w' = x w D$ and let $G'$ be obtained from $G$ by adding a dominating vertex. For all $k$ we have

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

**Proof.** Using the fact that

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

we easily get (applying the above with $f(x)$ replaced by $f'(x)$)

$$x w D f(x) = \sum_{k \geq 0} \left\{ G \right\}_{k-1} x^k D^k f(x).$$

But now note that $\left\{ G' \right\}_k = \left\{ G \right\}_{k-1}$, for all $k$. \qed
Claim 2.5. 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 \( k \),

\[
S_{w_i}(k) = \left\{ \frac{G_i}{k} \right\}.
\]

For all \( k \) we have

\[
S_{w_1w_2}(k) = \left\{ \frac{G_1 \cup G_2}{k} \right\}
\]

where \( w_1w_2 \) 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 (4) we have, for arbitrary \( f \),

\[
w_1w_2 \ f(x) = w_1 \sum_{k_2} \left\{ \frac{G_2}{k_2} \right\} x^{k_2} D^{k_2} f(x)
\]

\[
= \sum_{k_1} \left\{ \frac{G_1}{k_1} \right\} x^{k_1} D^{k_1} \sum_{k_2} \left\{ \frac{G_2}{k_2} \right\} x^{k_2} D^{k_2} f(x)
\]

\[
= \sum_{k_1,k_2} \left\{ \frac{G_1}{k_1} \right\} \left\{ \frac{G_2}{k_2} \right\} x^{k_1} D^{k_1} x^{k_2} D^{k_2} f(x).
\]

(5)

Now by Leibnitz’ rule (for the iterated derivative of a product) we have

\[
x^{k_1} D^{k_1} x^{k_2} D^{k_2} \ f(x) = \sum_j \binom{k_1}{j} k_2^j x^{k_1+k_2-j} D^{k_1+k_2-j} f(x)
\]

\[
= \sum_k \binom{k_1}{k_1+k_2-k} k_2^{k_1+k_2-k} x^{k_1+k_2-k} D^k f(x).
\]

(Here we use \( x^\down{n} \) for the \( j \)th falling power of \( x \), that is, \( x(x-1) \ldots (x-j+1) \).) Inserting into (5) we get

\[
S_{w_1w_2}(k) = \sum_{k_1,k_2} \left\{ \frac{G_1}{k_1} \right\} \left\{ \frac{G_2}{k_2} \right\} \binom{k_1}{k_1+k_2-k} k_2^{k_1+k_2-k}
\]

\[
= \sum_{k_1,k_2} \left\{ \frac{G_1}{k_1} \right\} \left\{ \frac{G_2}{k_2} \right\} \binom{k_1}{k_1+k_2-k} \binom{k_2}{k_1+k_2-k} (k_1+k_2-k)!.\]

We claim that the right-hand side above is exactly \( \left\{ \frac{G_1 \cup G_2}{k} \right\} \). Indeed, one way to generate all partitions of \( G_1 \cup G_2 \) into \( k \) nonempty independent sets is to first fix a \( k_1, k_2 \), then partition, for each \( i \), \( G_i \) into \( k_i \) non empty independent sets (\( \left\{ \frac{G_i}{k_i} \right\} \) ways), then choose \( k_1+k_2-k \) of the classes from \( G_1 \) and \( k_1+k_2-k \) of the classes from \( G_2 \) (\( \binom{k_1}{k_1+k_2-k} \binom{k_2}{k_1+k_2-k} \) ways), then merge the chosen classes in pairs, one from \( G_1 \) and one from \( G_2 \) (\( (k_1+k_2-k)! \) ways), thereby creating \( k_1+k_2-(k_1+k_2-k) = k \) non-empty independent sets. □
To prove Theorem 2.2, 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, with $S_w'(k) = \{G_w'\}$. That $S_w(k) = \{G_w\}$ where $G_w$ is obtained from $G_w'$ by adding a dominating vertex follows from Claim 2.4.

If $w$ is reducible, and of length greater than 2, then $w = w_1w_2\ldots w_k$ for some Dyck words $w_i$ which (by induction) have associated graphs $G_{w_i}$, constructed as described, with $S_{w_i}(k) = \{G_{w_i}\}$. That $S_w(k) = \{G_w\}$ where $G_w$ is the disjoint union of the $G_{w_i}$'s follows from repeated applications of Claim 2.5.

3 Further results

3.1 An explicit expression for $S_w(k)$

Theorem 2.2 leads quickly to an explicit expression for $S_w(k)$. Let $w$ be any word, with, say, $m$ $x$'s, and let $x^aw^bD$ be its associated Dyck word. For the $i$th $x$ in $x^aw^bD$, let $a_i$ be the excess of $x$'s over $D$'s in the initial segment of $x^aw^bD$ that ends immediately prior to the $i$th $x$ (we refer to this as the height of the $i$th $x$).

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

Similar explicit expressions have appeared in, for example, [21] and [15]. Note that in the case $w = (xD)^n$, this immediately reduces to the familiar

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

Given Theorem 2.2, what we need to do to prove Theorem 3.1 is to understand $\{G_{x^aw^bD}\}$. For this we utilize the chromatic polynomial. Recall that associated with 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 $f : V \rightarrow \{1, \ldots, q\}$ satisfying $f(u) \neq f(v)$ whenever $uv \in E$. A key observation is that for all $G$, $\chi_G(q)$ determines $(\binom{G}{k})_{k \geq 0}$, and vice-versa. Indeed, on the one hand inclusion-exclusion gives

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

while on the other hand

$$\chi_G(q) = \sum_{k \geq 0} \binom{G}{k} q^k$$

(recall $q^k = q(q-1)\ldots(q-k+1)$). To see this second relation, note that given a palette of $q$ colors, for each $k$ there are $\binom{G}{k}$ ways to partition the vertex set into $k$ non-empty color classes, and $q^k$ ways to assign colors the classes. A particular consequence of (6) is that if $G$ and $G'$ are different graphs with the same chromatic polynomial, then for all $k$, $\binom{G}{k} = \binom{G'}{k}$. 

6
We claim that if \( w \) is a Dyck word, then there is an easy way to compute its chromatic polynomial, in terms of heights of \( x \)’s.

**Claim 3.2.** For any Dyck word \( w \), with \( m \) \( x \)’s having heights \( a_1, \ldots, a_m \),

\[
\chi_G(q) = \prod_{i=1}^{m} (q - a_i).
\]

From this and (6), Theorem 3.1 immediately follows.

**Proof.** (Claim 3.2) We proceed by induction on the length of \( w \), with length 2 trivial. Consider a word \( w \) of length at least 4. If \( w \) is reducible, say \( w = w_1 \ldots w_k \) with each \( w_i \) either irreducible or the word \( xD_i \), we note (using repeated applications of Claim 2.5) that \( G_w = G_{w_1} \cup \ldots \cup G_{w_k} \), so that \( \chi_{G_w}(q) = \prod_{i=1}^{k} \chi_{G_{w_i}}(q) \). The claim in this case follows by induction. If instead \( w = xw'D \) is irreducible then we note (using Claim 2.4) that \( G_w \) is obtained from \( G_{w'} \) by adding a dominating vertex, so that \( \chi_{G_w}(q) = q \chi_{G_{w'}}(q - 1) \). The claim in this case again follows by induction.

### 3.2 A closely related combinatorial interpretation

By combining results of Navon [16] and Goldman, Joichi and White [11], we find another graph \( H_w \) that can naturally be associated with a Dyck word \( w \), such that \( S_w(k) = \left\{ \frac{H_w}{k} \right\} \) 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 \( w \) with \( n \) \( x \)’s and \( n \) \( D \)’s, let \( W_w \) 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 \( w_1 = xDxDDD \) then \( W_{w_1} = \emptyset \), and if \( w_2 = xxDDxDDDD \) then \( W_{w_2} = \{(1,2), (2,3), (2,4), (3,4), (3,5), (4,5)\} \). Define a graph \( H_w \) on vertex set \( \{1, \ldots, n\} \) by putting an edge from \( i \) to \( j \) (\( i < j \)) if and only if \( (i, j) \in W_w \).

It is worth noting that \( H_w \) is determined by the places where the staircase path of \( w \) takes a step up followed by a step to the right. To make this precise, say that the staircase path of \( w \) turns around the unit square labeled \( (x, y) \) 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 (labels of) unit squares that the path of \( w \) turns around. 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, \( \mathcal{T}_{w_1} = \{(1,1), (2,2), (3,3)\} \) and the edge set of \( H_{w_1} \) is empty; while \( \mathcal{T}_{w_2} = \{(1,2), (2,4), (3,5)\} \), and the edge set of \( H_{w_2} \) is \( \{12\} \cup \{23, 24, 34\} \cup \{34, 35, 45\} \).

Note that the graphs \( H_w \) and \( G_w \) often coincide (for example, when \( w = \prod_{i=1}^{t} x^{t_i}D^{t_i} \) for arbitrary \( t_i \)’s), but not always (for example, when \( w = w_2 \) above).

**Theorem 3.3.** Let \( w \) be Dyck word in the alphabet \( \{x, D\} \). For all integers \( k \) we have

\[
S_w(k) = \left\{ \frac{H_w}{k} \right\}.
\]
To prove Theorem 3.3, 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 \emph{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. [11, Theorem 2 (page 137)] is that for all \( i, k \), \( r_k(B) = \{r_{n_k}(B)\} \). (Goldman et al. use the notation \( q_{n-k} \) for \( \{r_{n-k}(B)\} \).

To interpret this result in the language of the present paper, let \( w \) be a Dyck word with \( n \) x’s (and so \( n \) D’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 \emph{Ferrers board}). While \( \mathcal{F}_w \) does not form a proper \( n \)-board, it is easy to check that its reflection across the line \( x = y \) does, and that the graph \( \Gamma_n(B) \) is isomorphic to \( H_w \) (via the identity map on the labels). It is also clear that if \( B \) is the reflection of \( \mathcal{F}_w \) across \( x = y \), then \( r_k(B) = r_k(\mathcal{F}_w) \). Thus Goldman et al.’s result is that for all \( k \),

\[
\tag{7} r_k(\mathcal{F}_w) = \left\{ \frac{H_w}{n-k} \right\}.
\]

The second result we need is Navon’s combinatorial interpretation of \( S_w(k) \) from [16]. Our explanation follows that of Varvak [21]. For an arbitrary word \( w \) with \( m \) x’s and \( n \) D’s, form the staircase path associated with \( w \) 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 \([0, n] \times [0, m]\). As before, let \( r_k(B) \) be the number of ways of placing \( k \) non-attacking rooks on \( B \). Navon’s combinatorial interpretation of the numbers \( S_w(k) \), as stated (and reproofed) in [21, 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

\[
\tag{8} S_w(n-k) = r_k(\mathcal{F}_w)
\]

for all \( k \). Combining (7) and (8) we get Theorem 3.3.

It is also possible to give a direct proof (not using Navon’s interpretation) of the identity \( \left\{ \frac{H_w}{k} \right\} = \{G_w\} \) for all Dyck words \( w \) and integers \( k \) (and so also a direct proof of Theorem 3.3); we sketch the details here.

We have already calculated (in Claim 3.2) the chromatic polynomial of \( G_w \):

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

where \( w \) has \( m \) x’s (and so also \( m \) D’s, since we are assuming that \( w \) is a Dyck word), and \( a_i \) is the height of the \( i \)th \( x \) (the excess of \( x \)’s over \( D \)’s in the initial segment of \( w \) that terminates
immediately before the $ith$ $x$). If we can show that $H_w$ has the same chromatic polynomial, then we get $\{H_w\} = \{G_w\}$ using (6). We will use an equivalent formulation of the height of an $x$, that can be calculated from the Dyck path associated with $w$: 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 $b - a$.

To compute the chromatic polynomial of $H_w$, consider the set $T_w = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ of squares that the Dyck path of $w$ turns around. As discussed earlier, the edge set of $H_w$ is obtained by putting cliques on each of the consecutive integer segments $\{x_i, \ldots, y_i\}$, $1 \leq i \leq k$. 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 of the heights of $x$’s in a word, that this is the same as $\prod_{i=1}^{y_1}(q - a_i)$ (where $a_i$ is the height of the $ith$ $x$).

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, we get that indeed the number of proper $q$-colorings of $H_w$ is $\prod_{i=1}^{n}(q - a_i)$.

### 3.3 Increasing forests

An $r$-ary tree is a tree in which every vertex, including a designated root, has some number $i$ ($0 \leq i \leq r$) of neighbors, with the set of neighbors equipped with a bijection to some subset of a fixed set of size $r$ (when $r = 2$ this set is often taken to be \{left, right\}, for example, and for $r = 3$ it might be \{left, middle, 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 the tree.

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 $|F(1, n, k)| = \binom{n}{k}$; 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 [15, 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 $k$,

$$|F(r, n, k)| = S_{w(r, n)}(k). \quad (9)$$

Theorems 2.2 and 3.3 (or rather, Corollary 2.3 and its natural analog with $G_{x^awD^b}$ replaced $H_{x^awD^b}$) 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 empty graph 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 $1 \leq i \leq n$. We refer to these graphs as $G(n, r)$ and $H(n, r)$ respectively. The following identities follow immediately from Theorems 2.2 and
3.3, via (9):

\[
\begin{align*}
\{ G(n,r) \} &= |F(r,n,k)| \\
\{ H(n,r) \} &= |F(r,n,k)|.
\end{align*}
\]

In the full version of this paper, we give bijective proofs of both these identities.

3.4 A new summation formula for \( S_w(k) \) when \( w = (x^sD^s)^n \)

All explicit expressions for \( S_w(k) \) that have appeared in the literature have taken the form of alternating sums. Using Theorem 2.2, we can obtain a new expression for \( S_w(k) \), in the special case \( w = (x^sD^s)^n \), as a positive linear combination of ordinary Stirling numbers. In what follows we use \([x^\ell]p(x)\) for the coefficient of \( x^\ell \) in the polynomial \( p(x) \).

**Theorem 3.4.** Let \( w = (x^sD^s)^n \). For each \( k \) we have

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

where

\[
f(n,s,\ell) = \left\{ \begin{array}{l}
[x^\ell] ((1 + x)(1 + 2x) \ldots (1 + (s-1)x))^{n-1} \\
\sum_{i_1 + i_2 + \ldots + i_{s-1} = \ell} \binom{n-1}{i_1} \binom{2(n-1) - i_1 - i_2 - \ldots - i_{s-2}}{i_{s-1}}.
\end{array} \right.
\] (10)

The (unsigned) Stirling number of the first kind \([^a\atop b]\) 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, the first expression on the right-hand side of (10) immediately gives the following nice connection between generalized Stirling numbers of the second kind, and ordinary Stirling numbers of the first kind:

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

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 3.4 we easily obtain

\[
B((x^sD^s)^n) = \sum_{\ell=0}^{(s-1)(n-1)} f(n,s,\ell) B_{s(n-1)+1-\ell}.
\] (11)

In [5] the comment is made that the Bell number \( B((x^sD^s)^n) \) can be always expressed in terms of conventional Bell numbers and \( r \)-nomial (binomial, trinomial, \ldots) 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; (11) illustrates this comment explicitly for arbitrary \( s \).
We now turn to the proof of Theorem 3.4, which will occupy us until the end of the section. From Theorem 2.2 we already know that
\[ S_w(k) = \left\{ \frac{nK_s}{k} \right\} \]
in the case \( w = (x^sD^s)^n \), where \( nK_s \) is the disjoint union of \( n \) copies of \( K_s \). To understand \( \left\{ \frac{nK_s}{k} \right\} \), we use the observation (from Section 3.1) that if the chromatic polynomials of \( G \) and \( G' \) coincide, then \( \{G\}_{k} = \{G'\}_{k} \) for all \( k \).

If \( G \) consists of \( a \) disjoint copies of \( K_b \), together with \( r \) isolated vertices, then its chromatic polynomial is \( q^{r-a} \) (for the isolated vertices) times the \( a \)th power of \( q(q-1)(q-2)\ldots(q-(b-1)) \) (one for each of the copies of \( K_b \)); i.e., it is
\[ q^{r+a}(q-1)^a(q-2)^a\ldots(q-(b-1))^a. \]

On the other hand, if \( G \) consists of \( a \) disjoint copies of \( K_{b-1} \), together with a single vertex joined to all the vertices of each of the \( K_{b-1} \)'s (creating a fan of \( K_b \)'s), together further with \( r + a - 1 \) isolated vertices, then the chromatic polynomial is \( q^r \) (for the single vertex joined to everything in each of the \( K_{b-1} \)'s), times the \( a \)th power of \( (q - 1)(q - 2)\ldots(q - (b - 1)) \) (one for each of the copies of \( K_{b-1} \)), times \( q^{r+a-1} \) (for the isolated vertices), i.e., it is also
\[ q^{r+a}(q-1)^a(q-2)^a\ldots(q-(b-1))^a. \]

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

We use this observation first to obtain an expression for \( \left\{ \frac{nK_2}{k} \right\} \). The chromatic polynomial of \( nK_2 \) is \( q^n(q-1)^{n-1} \). This 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 \( \left\{ \frac{nK_2}{k} \right\} = \{G\}_{k} \), and we can find \( \left\{ \frac{nK_2}{k} \right\} \) easily: first, decide on a subset of size \( \ell \) of the isolated vertices (perhaps empty) to be in the same class as the center of the star (which of course can’t go with 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 decide how to partition these vertices into \( k - 1 \) classes. This leads to the formula
\[ \left\{ \frac{nK_2}{k} \right\} = \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \binom{2n-1-\ell}{k-1}. \]

Now we deal with the general case. Here it will be convenient to define the graph \( G_{a,b,c} \) as the graph that 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; note that this reduces to a star when \( b = 2 \)) together with \( c \) isolated vertices. For example, \( G_{n,2,n-1} \) is the graph that replaced \( nK_2 \) in the last paragraph.

The chromatic polynomial of \( nK_s \) is \( (q(q-1)(q-2)\ldots(q-(s-1)))^n \). This is the same as the chromatic polynomial of the graph \( G_{n,s,n-1} \) (as discussed above). To partition this graph into \( k \) non-empty independent sets, we first decide on a subset of size \( i_1 \) of the isolated vertices (perhaps \( i_1 = 0 \); certainly \( i_1 \leq n - 1 \)) to be in the same class as the center of the fan (which of course can’t go 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. By our corollary, we may replace this with \( G_{n,s-1,2(n-1)-i_1} \), and so

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

What is \( \left\{ \frac{G_{n,s-1,2(n-1)-i_1}}{k-1} \right\} \)? By first partnering the center vertex of the fan with some subset of size \( i_2 \) (\( 0 \leq i_2 \leq 2(n-1) - i_1 \)) from 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

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

and so

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

This process continues until eventually we reach

\[
\left\{ \frac{G_{n,2,(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}}{k-(s-2)} \right\} = \sum_{i_{s-1}=0}^{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}} \binom{2(n-1) - i_1 - i_2 - \ldots - i_{s-2}}{i_{s-1}} \left\{ \frac{G_{n,2,(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}}{k-(s-1)} \right\}.
\]

Putting it all together, we get

\[
\left\{ \frac{nK_s}{k} \right\} = \sum_{i_{s-1}=0}^{(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}} \binom{2(n-1) - i_1 - i_2 - \ldots - i_{s-2}}{i_{s-1}} \left\{ \frac{G_{n,2,(s-1)(n-1)-i_1-i_2-\ldots-i_{s-2}}}{k-(s-1)} \right\}.
\]

We re-write this by setting \( \ell = i_1 + i_2 + \ldots + i_{s-1} \). The range of possible values of \( \ell \) is from 0 to \( (s-1)(n-1) \), and for each \( \ell \) in this range the coefficient of \( \left\{ \frac{s(n-1)+1-\ell}{k-(s-1)} \right\} \) in the above expression 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 the second equality in (10). To establish the first equality, we need to show that the expression in (12) equals \[ (1 + x)(1 + 2x) \cdots (1 + (s - 1)x) \] for which we employ a counting argument. Let \( A_1, \ldots, A_{s-1} \) be \( s - 1 \) disjoint sets, each of size \( n - 1 \). An \( \ell \)-selection from \( A_1 \) through \( A_{s-1} \) is a specification of sets \( A_{11}, A_{21}, A_{22}, A_{31}, A_{32}, A_{33}, \ldots A_{(s-1)(s-1)}, \ldots, A_{(s-1)(s-1)} \), pairwise disjoint, with \( A_{11} \subseteq A_1, A_{21} \cup A_{22} \subseteq A_2, \ldots, A_{(s-1)(s-1)} \cup \ldots \cup A_{(s-1)(s-1)} \subseteq A_{s-1}, \) and \( |A_{11}| + |A_{21}| + \ldots + |A_{(s-1)(s-1)}| = \ell \).

To count the number of \( \ell \)-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 \( \ell \)-selections is

\[
\sum_{i_1 + i_2 + \ldots + i_{s-1} = \ell} \prod_{k=1}^{s-1} k^{i_k} \binom{n-1}{i_k} = [x^\ell](1 + x)(1 + 2x) \cdots (1 + (s - 1)x)^{n-1}.
\]

Another way to count \( \ell \)-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 \( A_{kk} \). This gives that the number of \( \ell \)-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 proof of Theorem 3.4.

References


