FOREST-LIKE PERMUTATIONS

MIREILLE BOUSQUET-MÉLOU AND STEVEN BUTLER

Abstract. Given a permutation $\pi \in S_n$, construct a graph $G_\pi$ on the vertex set $\{1, 2, \ldots, n\}$ by joining $i$ to $j$ if (i) $i < j$ and $\pi(i) < \pi(j)$ and (ii) there is no $k$ such that $i < k < j$ and $\pi(i) < \pi(k) < \pi(j)$. We say that $\pi$ is forest-like if $G_\pi$ is a forest. We first characterize forest-like permutations in terms of pattern avoidance, and then by a certain linear map being onto. Thanks to recent results of Woo and Yong, this shows that forest-like permutations characterize Schubert varieties which are locally factorial. Thus forest-like permutations generalize smooth permutations (corresponding to smooth Schubert varieties).

We compute the generating function of forest-like permutations. As in the smooth case, it turns out to be algebraic. We then adapt our method to count permutations for which $G_\pi$ is a tree, or a path, and recover the known generating function of smooth permutations.

1. Introduction

Take a permutation $\pi = \pi(1)\pi(2)\cdots\pi(n)$ in the symmetric group $S_n$. Let $G_\pi$ be the graph on the vertex set $\{1, 2, \ldots, n\}$ with an edge joining $i$ to $j$ if and only if (i) $i < j$ and $\pi(i) < \pi(j)$ and (ii) there is no $i < k < j$ with $\pi(i) < \pi(k) < \pi(j)$. An example is shown in Figure 1. We say that $\pi$ is forest-like if $G_\pi$ is a forest (i.e., has no cycle). Note that the edges of $G_\pi$ correspond to the edges of the Hasse diagram of the sub-poset of $N^2$ consisting of the points $(i, \pi(i))$ (Figure 1, left). This (sub-)poset is known to play a crucial role in the Robinson-Schensted correspondence.

Consider also the following construction, borrowed from [23]. Label $n$ columns by $1, 2, \ldots, n$, and place $n-1$ vertical dividers between the columns. Draw a horizontal bar between column $i$ and column $j$ if and only if there is an edge joining $i$ and $j$ in $G_\pi$. These bars are simply the horizontal projections of the edges of the Hasse diagram.

![Figure 1](image-url)

**Figure 1.** The permutation $\pi = 6 \ 4 \ 3 \ 7 \ 5 \ 1 \ 8 \ 2$, the associated graph $G_\pi$ and the corresponding collection of bars.

---

Date: 26th June 2006.

Key words and phrases. Schubert varieties. Pattern avoiding permutations.
We use this construction to define a linear map from $\mathbb{Z}^{n-1}$ to $\mathbb{Z}^{e(\pi)}$, where $e(\pi)$ is the number of horizontal bars in the diagram (also the number of edges in $G_\pi$). Choose a linear order on the bars, and associate variables $\alpha_i$ with the vertical dividers and $\beta_k$ with the horizontal bars. If the $k$th horizontal bar starts in column $i$ and goes to column $j$ then we set

$$\beta_k = \sum_{t=1}^{j-1} \alpha_t. \quad (1)$$

The map $L_\pi$ sends $(\alpha_1, \ldots, \alpha_{n-1})$ to $(\beta_1, \ldots, \beta_{e(\pi)})$. In the example above we have $\beta_1 = \alpha_1 + \alpha_2 + \alpha_3, \beta_2 = \alpha_2 + \alpha_3, \ldots, \beta_9 = \alpha_6 + \alpha_7$.

Our first result describes forest-like permutations in terms of the map $L_\pi$, and gives a characterization of these permutations in terms of pattern avoidance.

**Theorem 1.** For $\pi \in S_n$ the following are equivalent:

1. the graph $G_\pi$ is a forest;
2. the linear map $L_\pi : \mathbb{Z}^{n-1} \rightarrow \mathbb{Z}^{e(\pi)}$ is onto;
3. the permutation $\pi$ avoids the patterns $1324$ and $21354$.

We need to clarify the third point. A permutation $\pi$ avoids the pattern $1324$ if one cannot find indices $p < q < r < s$ such that $\pi(p) < \pi(q) < \pi(r) < \pi(s)$. Similarly, $\pi$ avoids the pattern $21354$ if every occurrence of the pattern $21354$ is a subsequence of an occurrence of $21354$. That is to say, for all indices $p < q < r < s$ such that $\pi(q) < \pi(p) < \pi(s) < \pi(r)$, there exists a $t$ such that $q < t < r$ and $\pi(p) < \pi(t) < \pi(s)$. The notation was introduced by J. West in his thesis [21], and appears, for instance, in [11]. There are several equivalent ways to describe the latter avoidance condition. In particular, it is easy to see that, in the terminology introduced by Woo and Yong [23], avoiding $21354$ is equivalent to avoiding $2143$ with Bruhat condition (1 ↔ 4). However, the first description is more symmetric, more clearly showing that $\pi$ avoids $21354$ if and only if $\pi^{-1}$ does.

Given that a linear map $\mathbb{Z}^{n-1} \rightarrow \mathbb{Z}^e$ is bijective if and only if it is onto and $e = n - 1$, we obtain the following result.

**Corollary 2.** The map $L_\pi$ is a bijection if and only if $G_\pi$ is a tree. In this case we say that $\pi$ is tree-like.

Our second result is the enumeration of forest-like permutations. We will show that their generating function is

$$F(x) = \frac{(1-x)(1-4x-2x^2)-(1-5x)\sqrt{1-4x}}{2(1-5x+2x^2-x^3)}. \quad (2)$$

We also enumerate several natural subclasses of forest-like permutations, such as tree-like permutations.

The original motivation for studying tree-like permutations came from a question of Woo and Yong related to Schubert varieties. It is known that Schubert varieties can be indexed by permutations [12], and various properties of Schubert varieties have been translated into properties of permutations. One famous example is that a variety is smooth if and only if the associated permutation avoids the patterns $1324$ and $2143$ [11]. A weakening of smoothness is the locally factorial property, an algebra-geometric condition which states that all local rings are unique factorization domains. Woo and Yong established a condition for being locally factorial which is equivalent to $L_\pi$ being onto [23, Prop. 2]. They conjectured that this holds if and only if $\pi$ is $1324$ and $21354$ avoiding [22]. Theorem 1 settles this conjecture.
We note that every smooth permutation (1324 and 2143 avoiding) is forest-like. Smooth permutations have been counted before \cite{Cortez:2013}, and their generating function is:
\[ S(x) = \frac{1 - 5x + 4x^2 + x\sqrt{1 - 4x}}{1 - 6x + 8x^2 - 4x^3}. \]
As Reference \cite{Cortez:2013} is not easily available, we will show how to adapt our proof of \cite{Cortez:2013} to enumerate smooth permutations. The series \( S(x) \) occurs in several other enumeration problems \cite{Manivel:2017}.

**Remark.** Results of Cortez \cite{Cortez:2013}, and independently Manivel \cite{Manivel:2017}, show that 1324 and 21354 avoidance is necessary and sufficient to characterize which Schubert varieties are generically locally factorial. Here generic has the following sense: the variety is smooth at almost all points but has a closed subset \( Y_\pi \) where it is not smooth, and in that closed subset it is factorial at almost all points.

We will proceed as follows. In Section 2 we prove Theorem 1. The proof involves a fourth condition, equivalent to those of Theorem 1, which uses a certain sorting procedure on the bars. In Section 3 we count forest-like permutations and several of their natural subclasses, such as tree-like permutations and smooth permutations. We conclude in Section 4 by describing several simple bijections related to some of our enumerative results, and state some open problems.

## 2. Characterization of forest-like permutations

The aim of this section is to prove Theorem 1. We begin with proving that (1) \( \Rightarrow \) (3) and (2) \( \Rightarrow \) (3) by proving the contrapositive: if \( \pi \) contains 1324 or 21354, then \( G_\pi \) contains a cycle and \( L_\pi \) is not onto.

### 2.1. Permutations containing 1324 or 21354

We first look at the structure found in the diagrams of permutations containing 1324 and 21354. We begin with a very simple lemma which follows from the definition of the diagram of bars (alternatively, from the definition of the Hasse diagram of a poset).

**Lemma 3.** Let \( \pi \in S_n \). If \( p < q \) and \( \pi(p) < \pi(q) \) then there is a sequence \( p = p_0 < p_1 < \cdots < p_k = q \) such that \( \pi(p_i) < \pi(p_{i+1}) \) and in the diagram for \( \pi \) there are horizontal bars from column \( p_i \) to column \( p_{i+1} \) for each \( i = 0, \ldots, k - 1 \).

**Lemma 4.** Given a permutation \( \pi \),

(a) if \( \pi \) contains the pattern 1324 then there are indices \( p < q < r < s \) such that \( \pi(p) < \pi(r) < \pi(q) < \pi(s) \) and in the diagram for \( \pi \) there are horizontal bars from \( p \) to \( r \) and from \( q \) to \( s \).

(b) if \( \pi \) contains the pattern 21354 then there are indices \( p < q < r < s \) such that \( \pi(q) < \pi(p) < \pi(s) < \pi(r) \) and in the diagram for \( \pi \) there are horizontal bars from \( p \) to \( s \) and from \( q \) to \( r \).

**Proof.** The general idea is the following. If we have an occurrence of the pattern that does not satisfy the requisite bar conditions, then we find a tighter occurrence that satisfies them.

For instance, start from an occurrence of the pattern 1324, that is, from a sequence \( p < q < r < s \) such that \( \pi(p) < \pi(r) < \pi(q) < \pi(s) \). Define \( p' := \max\{i < q : \pi(i) < \pi(r)\} \) and \( r' := \min\{j > q : \pi(r) \geq \pi(j) > \pi(p')\} \). Then \( p \leq p' < q < r' \leq r < s \), the sequence \( p', q, r', s \) corresponds to another occurrence of 1324, and there is a bar between columns \( p' \) and \( r' \).

The rest of the lemma is proved by similar arguments. \( \square \)
Remark. Point (b) in the above lemma shows that 21\bar{3}54 avoidance can also be described graphically as follows. Take a permutation $\pi$ and plot it as on the left of Figure 1. Represent by straight lines the edges of the Hasse diagram of the poset $\{(i, \pi(i))\}$. We thus obtain the natural embedding of $G_\pi$. Then $\pi$ avoids 21\bar{3}54 if and only if this embedding of $G_\pi$ is planar (no edges cross). This does not mean that avoiding 21\bar{3}54 is necessary for $G_\pi$ to be planar: for instance, the permutation $\pi = 2143$ contains 21\bar{3}54 but $G_\pi$ is planar (though its natural embedding is not).

Lemma 4 is illustrated in Figure 2, where the solid lines indicate a single bar and the dashed lines indicate a sequence of bars (coming from Lemma 3).

![Figure 2](image)

**Figure 2.** Patterns in the bar diagrams of permutations containing 1324 or 21354.

We now show that the occurrence of either of the two “forbidden” patterns implies the existence of cycles in $G_\pi$, and prevents $L_\pi$ from being onto. First, from Figure 2 we can read off cycles in $G_\pi$. For example in the 1324 case we have a cycle that starts at $p$, goes to $r$ then by a sequence of edges goes to $s$ then to $q$ and finally by another sequence of edges we return to $p$. This is a true cycle, as it contains the edge joining $p$ to $r$ only once. Similarly, in the 21354 case, there is a true cycle visiting $p, s, q, r$ in this order. Secondly, we also see that there are nontrivial linear dependencies among the $\beta_j$. In the 21354 case the sum of the solid bars equals the sum of the dashed bars, and a similar event happens in the 1324 case. This prevents $G_\pi$ from being onto.

So if the permutation contains 1324 or 21354 then $G_\pi$ has cycles and $L_\pi$ is not onto. Taking the contrapositive gives $(1) \Rightarrow (3)$ and $(2) \Rightarrow (3)$ in Theorem 1.

### 2.2. Sorting the horizontal bars

In this subsection, we define a new condition $(2')$ that clearly implies the surjectiveness condition $(2)$. We then prove that $(2')$ is implied by the pattern avoidance condition $(3)$, and finally that $(2')$ implies the acyclicity condition $(1)$. Combined with Section 2.1, this proves that the four conditions $(1)$, $(2)$, $(2')$ and $(3)$ are equivalent, and establishes Theorem 1. The structure of the proof is schematized below.

![Diagram](image)

In the construction of the diagram for a permutation we placed no condition on the ordering of the horizontal bars from top to bottom. We now describe a way to attempt to sort them. Create a second diagram with the same columns but no
horizontal bars. We now look for bars to move to the second diagram by scanning the vertical dividers from left to right, looking for any divider which is intersected by exactly one horizontal bar. As soon as we find such an intersection we move the corresponding horizontal bar to the second diagram and put it above any previously moved bar. We then repeat this scanning process, starting again from the leftmost divider, until no divider intersects exactly one horizontal bar. If at this stage all the horizontal bars are moved over, we say that

(2') the bars are fully sortable.

By construction, this can only happen when the number of edges satisfies $e(\pi) \leq n - 1$. An example of a fully sorted diagram is shown in Figure 3.

![Figure 3. The permutation $\pi = 1 6 7 5 3 4 2$ and the associated sorted diagram of bars.](image)

Assume the bars are fully sortable. In terms of the equations (1), this means that at least one new variable $\alpha_i$ occurs in each equation. More precisely, if $V_k$ denotes the set of variables $\alpha_i$ occurring in $\beta_1, \ldots, \beta_k$, then $V_k \subseteq V_{k+1}$. Hence, given $\beta \in \mathbb{Z}^{e(\pi)}$, the system (1) can be solved for $\alpha$ by backward substitution from the top equation to the bottom equation. Consequently, we have the following.

**Lemma 5.** If the bars are fully sortable then $L_{\pi}$ is onto.

In other words, (2') implies (2). We shall see below that the converse is also true. This will be a consequence of Theorem 1 and the following lemma, which proves (the contrapositive of) (3) $\Rightarrow$ (2').

**Lemma 6.** If the bars are not fully sortable then $\pi$ contains 1324 or 21354.

**Proof.** If we stopped before all the bars have been moved over then it must be the case that for what remains all the vertical dividers intersect either zero, or two or more horizontal bars. We will work with these remaining (i.e., unmoved, unmoved) horizontal bars.

Suppose that column $a$ is the leftmost column which has the start of a bar, then as noted above it must be the start of at least two bars (otherwise we would have moved the bar over). Let $c$ denote the column where the longest horizontal bar starting in column $a$ ends. Let $b$ be the rightmost column satisfying $a < b < c$ and $\pi(c) < \pi(b)$ (such a $b$ exists because the end of a second bar that starts in $a$ satisfies both conditions).

We now consider cases on how to cover the vertical divider to the right of column $b$ with a second horizontal bar.

**Case (1).** There is a horizontal bar that begins at $b$. This bar ends at some position $d$, which, by the choice of $b$, satisfies $d > c$. In this case we have that $a < b < c < d$ while $\pi(a) < \pi(c) < \pi(b) < \pi(d)$ and so $\pi$ contains the pattern 1324.

**Case (2).** There is a horizontal bar that begins at column $d$ where $d < b$ and crosses to some column $e$ where $e > b$. By the choice of $b$, we have $a < d$. Since $d$ lies between $a$ and $c$ we must have that $\pi(d) < \pi(a)$ or $\pi(d) > \pi(c)$ (if $\pi(d)$ were
in the interval \([\pi(a), \pi(c)]\), there would not be a bar from \(a\) to \(c\). So we consider subcases.

**Case (2a).** If \(\pi(d) < \pi(a)\) then we have that \(a < d < b < c\) and \(\pi(d) < \pi(a) < \pi(c) < \pi(b)\) and since there is a horizontal bar from \(a\) to \(c\), \(\pi\) contains the pattern 21354. (Note this includes the possibility that \(c = e\).)

**Case (2bi).** Suppose that not only \(\pi(d) > \pi(c)\), but also \(\pi(d) > \pi(b)\). Then we note that we have \(a < d < b < e\) and \(\pi(a) < \pi(b) < \pi(d) < \pi(e)\) and so \(\pi\) contains the pattern 1324.

**Case (2bii).** Suppose finally that \(\pi(c) < \pi(d) < \pi(b)\). By the choice of \(b\), we must have \(e > c\). Then we note that we have \(a < d < c < e\) and \(\pi(a) < \pi(c) < \pi(d) < \pi(e)\) and so \(\pi\) contains the pattern 1324.

Our final lemma proves that \((2') \Rightarrow (1)\).

**Lemma 7.** *If the bars are fully sortable, then \(G_\pi\) is a forest.*

**Proof.** Suppose on the contrary that \(G_\pi\) contains a cycle and we can fully sort the bars. Now consider the set \(B\) of bars that correspond to the edges of a cycle in \(G_\pi\). At some stage in the sorting procedure, a first bar \(b\) of \(B\) is moved over. At this stage, it is the only bar that crosses some vertical divider, say, the \(i\)th one. In particular, all the other bars involved in the cycle lie entirely to the right or entirely to the left of the \(i\)th divider. In terms of \(G_\pi\), this means that removing the edge corresponding to \(b\) has *disconnected* the cycle. This is of course impossible, so \(G_\pi\) cannot contain a cycle. \(\square\)

### 3. Generating functions for forest-like permutations

We now want to prove the enumerative result \(\mathbb{1}\). At the heart of this result is a recursive description of forest-like permutations, given in Proposition \(\mathbb{1}\). This decomposition is then translated into a functional equation defining the generating function of forest-like permutations (Proposition \(\mathbb{1}\)), which we solve using the *kernel method*.

The same decomposition can be recycled to count various subclasses of forest-like permutations. We will thus also obtain the generating functions of

1. **tree-like permutations,**
2. **rooted tree-like permutations** (the term *rooted* meaning that \(\pi(1) = 1\),
3. **path-like permutations** (\(G_\pi\) is a path),
4. **smooth** permutations (\(\pi\) avoids 1324 and 2143).

Note that every forest-like permutation satisfying \(\pi(1) = 1\) is actually tree-like (every vertex of \(G_\pi\) is connected to the vertex 1), and thus is a rooted tree-like permutation. Note also the following inclusions:

```
Path-like --|-- Tree-like
    |      |       |
    V      V       V
Rooted tree-like --|-- Smooth
```

For \(n \geq 1\), we denote by \(f_n\) (resp. \(t_n, r_n, p_n, s_n\)) the number of permutations \(\pi \in S_n\) of the above five types. We introduce the corresponding generating functions \(F(x)\) (resp. \(T(x), R(x), P(x), S(x)\)). In particular,

\[
F(x) = \sum_{n \geq 1} f_n x^n = x + 2x^2 + 6x^3 + 22x^4 + 89x^5 + 379x^6 + 1661x^7 + \cdots
\]

Our enumerative results are summarized in the following theorem.
Theorem 8. The five generating functions defined above are given by:

\[
\begin{align*}
F(x) & = \frac{(1-x)(1-4x-2x^2) - (1-5x)\sqrt{1-4x}}{2(1-5x+2x^2-x^3)}, \\
T(x) & = \frac{1-3x-6x^2-4x^3}{2(2-9x)}, \\
R(x) & = \frac{1-\sqrt{1-4x}}{2}, \\
P(x) & = x\frac{1-2x+2x^2}{(1-x)(1-2x)}, \\
S(x) & = x\frac{1-5x+4x^2+x\sqrt{1-4x}}{1-6x+8x^2-4x^3}.
\end{align*}
\]

From these generating functions it can be shown that there exists positive constants \( \kappa \) such that

\[
f_n \sim \kappa_f(4.61\ldots)^n, \quad t_n \sim \kappa_t(4.5)^n, \quad r_n = \frac{1}{n}\left(\frac{2n-2}{n-1}\right) \sim \kappa_r 4^{n-1}n^{-3/2},
\]

where the growth constants occurring in the asymptotics of \( f_n \) and \( s_n \) are respectively the real roots of the polynomial \( t^3 - 5t^2 + 2t - 1 \) and \( t^3 - 6t^2 + 8t - 4 \).

We note that \( r_n \) is the \((n-1)\)st Catalan number and has numerous combinatorial interpretations \([28]\, \text{Chap. 6}]. We give in Section 1.1.3 a bijective proof of this result, as well as another bijection explaining why the numbers \( p_n \) are so simple. The terms \( t_n \) have also occurred before and enumerate the number of \textit{stacked directed animals on a triangular lattice} \([1]\). No direct bijection between stacked directed animals and tree-like permutations is currently known.

The form of our decomposition of forest-like permutations will force us to take into account an additional statistic, namely the number of \( rl \)-minima for forest-like or tree-like permutations, and the length of the final ascent in smooth permutations. This is why we actually obtain bivariate generating functions that refine the above theorem (see \([13], [14], [17]\)). Other statistics, like the number of descents, could also be carried through our calculations.

3.1. Decomposing forest-like permutations

If \( \pi \in S_n \) we say that \( \pi \) has length \( n \), and write \(|\pi| = n \). We say that \( \pi(i) \) is an \( rl \)-minimum (right-to-left-minimum) if for all \( j > i \), we have \( \pi(j) > \pi(i) \). We denote by \( m(\pi) \) the number of \( rl \)-minima of \( \pi \). Finally, \( \pi \) is increasing if \( \pi = 12\ldots n \).

Let \( \pi \in S_n \) be forest-like. We decompose \( \pi \) by considering which element maps to 1. So suppose that \( i = \pi^{-1}(1) \) then there are two cases:

- **First case**: \( i = n = |\pi| \). Then the permutation \( \pi \in S_{n-1} \) defined by \( \tau(i) = \pi(i) - 1 \) is forest-like. Conversely, starting with a forest-like permutation \( \tau \in S_{n-1} \) we can construct a forest-like permutation \( \pi \in S_n \) by letting \( \pi(i) = \tau(i) + 1 \) for \( 1 \leq i < n-1 \) and \( \pi(n) = 1 \). Note that \( \pi \) is tree-like if and only if \( n = 1 \).
- **Second case**: \( i = \pi^{-1}(1) < n \). We now focus on this case, illustrated in Figure 3.

Let

\[
h = \min \left( \{ \pi(i+1) \} \cup \{ \pi(j) : j < i \} \right).
\]

(3)

So \( h \) is the smaller of the lowest value of \( \pi \) to the left of \( i \) or the value of \( \pi \) at \( i + 1 \).

First note that for all \( j \geq i + 1 \) we have \( \pi(j) \leq h \) or \( \pi(j) \geq \pi(i+1) \). If not, then for some \( j \) we have \( \pi(i) < h < \pi(j) < \pi(i+1) \) and \( \pi^{-1}(h) < i < i + 1 < j \), so the permutation contains the pattern 21354, and cannot be forest-like. Further,
if \( j, k \geq i + 1 \) with \( \pi(j) \geq \pi(i + 1) \) and \( \pi(k) < h \) then \( j < k \). If not, then \( i < i + 1 < k < j \) and \( \pi(i) < \pi(k) < \pi(i + 1) < \pi(j) \), so the permutation contains the pattern 1324, and cannot be forest-like.

The latter property implies that the last \( h - 2 \) terms of \( \pi = \pi(1) \pi(2) \cdots \pi(n) \) are 2, 3, \ldots, \( h - 1 \) in some order. Let \( \tau \) be the permutation obtained from \( \pi \) by retaining only its \( h - 1 \) smallest entries, i.e.,

\[
\tau = \pi(n - h + 3) \pi(n - h + 4) \cdots \pi(n).
\]

Then \( \tau \) is rooted and tree-like. Similarly, let \( \sigma \) be the permutation obtained by deleting these \( h - 1 \) smallest entries and subtracting \( h - 1 \) from the remaining entries:

\[
\sigma = (\pi(1) - h + 1) \cdots (\pi(i - 1) - h + 1)(\pi(i + 1) - h + 1) \cdots (\pi(n - h + 2) - h + 1).
\]

Then \( \sigma \) is forest-like. Moreover, \( \sigma(i) \) is an \( rl \)-minimum of \( \sigma \). If \( \sigma(i) \) is the \( k \)th \( rl \)-minimum of \( \sigma \) (read from right to left), define \( \Phi(\pi) = (\tau, \sigma, k) \). Observe that \( k = m(\sigma) \) if \( \sigma(i) = 1 \) (that is to say, \( h = \pi(i + 1) \)), and \( k = 1 \) if \( i = |\sigma| \).

\[\text{Figure 4. The structure of a forest-like permutation. The shaded areas show which regions of the embedding in } \mathbb{N}^2 \text{ may contain points.}\]

Conversely, starting from a 3-tuple \((\tau, \sigma, k)\) such that \( \tau \) is tree-like, \( \sigma \) is forest-like and \( k \leq m(\sigma) \), we can construct a (unique) forest-like permutation \( \pi \) satisfying \( \Phi(\pi) = (\tau, \sigma, k) \). If \( |\tau| = h - 1 \) and the \( k \)th \( rl \)-minimum of \( \sigma \) is \( \sigma(i) \), this is done by adding \( h - 1 \) to the entries of \( \sigma \), inserting 1 at position \( i \) and adding the other entries of \( \tau \) to the right of \( \sigma \), in the same order as in \( \tau \). By looking at the number of \( rl \)-minima of the resulting permutation \( \pi \), we obtain the following result.

**Proposition 9.** The map \( \Phi \) is a bijection between forest-like permutations \( \pi \) with \( \pi^{-1}(1) < |\pi| \) and 3-tuples \((\tau, \sigma, k)\) such that \( \tau \) is rooted tree-like, \( \sigma \) is forest-like, and \( 1 \leq k \leq m(\sigma) \). Moreover,

\[
|\pi| = |\tau| + |\sigma| \quad \text{and} \quad m(\pi) = \begin{cases} k + 1 & \text{if } \tau = 1, \\ m(\tau) & \text{otherwise.} \end{cases}
\]

(4)

In order to count the various sub-classes of forest-like permutations we have defined, we need the following result.

**Proposition 10.** Let \( \pi = \Phi(\tau, \sigma, k) \) be a forest-like permutation such that \( \pi^{-1}(1) < |\pi| \). Then

1. \( \pi \) is tree-like if and only if \( \sigma \) is tree-like,
2. \( \pi \) is rooted tree-like if and only if \( \sigma \) is rooted tree-like and \( k = m(\sigma) \),
(3) $\pi$ is path-like if and only if $\tau$ is increasing, $\sigma$ is path-like and its $k$th rl-minimum $\sigma(i)$ is such that $i$ has degree 1 in $G_{\sigma}$.

(4) $\pi$ is smooth if and only if $\sigma$ is smooth and either $k = m(\sigma)$ or $k \leq a(\sigma)$, where $a(\sigma)$ is the length of the final ascent of $\sigma$: if $|\sigma| = \ell$,

$$a(\sigma) = \max \{i : \sigma(\ell - i + 1) < \cdots < \sigma(\ell - 1) < \sigma(\ell)\},$$ (5)

Moreover,

$$a(\pi) = \begin{cases} 
    k + 1 & \text{if } \tau = 1 \text{ and } k \leq a(\sigma), \\
    a(\sigma) & \text{if } \tau = 1 \text{ and } k = m(\sigma) > a(\sigma), \\
    a(\tau) - 1 & \text{if } \tau \neq 1 \text{ is increasing,} \\
    a(\tau) & \text{otherwise.} 
\end{cases}$$ (6)

Proof. The first three results simply follow from the decomposition of Figure 4. The reader should look at Figures 6 and 9 to see this decomposition specialized to the rooted case and the path case, respectively.

Now let’s assume that $\pi$ is smooth. Since $\sigma$ and $\tau$ are obtained by deleting entries from $\pi$, they are smooth as well. This does not restrict the choice of $\tau$, since every rooted tree-like permutation is smooth. Conversely, when we construct $\Phi(\tau, \sigma, k)$ (assuming that $\sigma$ is smooth and $\tau$ rooted) we do not create any occurrence of 2143 if we insert 1 just before the smallest entry of $\sigma$. This corresponds to the case $k = m(\sigma)$.

However, if $k < m(\sigma)$, then the value $h$ defined by (3) satisfies $h < \pi(i + 1)$, and the final permutation $\pi$ contains 2143 if, and only if, there is a descent in $\sigma$ somewhere to the right of $\sigma(i)$. In other words, if $k < m(\sigma)$, then $\pi$ avoids 2143 if and only if 1 is inserted in the final ascent of $\sigma$, that is to say, $k \leq a(\sigma)$.

A case study finally provides the value of $a(\pi)$. □

3.2. Functional equations

We now translate Propositions 9 and 10 into enumerative terms. We first note that every pair $(\tau, \sigma)$ can be combined in $m(\sigma)$ different ways. To account for this we refine our generating functions by further distinguishing by the number of rl-minima. So let

$$F(u) \equiv F(x, u) = \sum_{n, \ell \geq 1} f_{n, \ell} x^n u^\ell = \sum_{\ell} F_\ell(x) u^\ell$$

where $f_{n, \ell}$ is the number of forest-like permutations of $S_n$ having $\ell$ rl-minima. Note that $F(x) = F(1)$. Define similarly the bivariate series $T(x, u), R(x, u), P(x, u)$.

The case of smooth permutation is a bit different: here, the crucial parameter is the length of the final ascent, defined by (5). We thus use a new indeterminate $v$ and define

$$S(v) \equiv S(x, v) = \sum_{n, \ell \geq 1} s_{n, \ell} x^n v^\ell = \sum_{\ell} S_\ell(x) v^\ell$$

where $s_{n, \ell}$ is the number of smooth permutations of $S_n$ having a final ascent of length $\ell$. We define similarly the series $R(x, v)$ that counts rooted tree-like permutations by the same statistics.
Proposition 11. The (bivariate) generating functions $\mathcal{F}(u)$, $\mathcal{T}(u)$, $\mathcal{R}(u)$ and $\mathcal{P}(u)$ satisfy:

$$\mathcal{F}(u) = xu + xu\mathcal{F}(1) + xu^2 \frac{\mathcal{F}(u) - \mathcal{F}(1)}{u-1} + (\mathcal{R}(u) - xu)\mathcal{F}'(1),$$

$$\mathcal{T}(u) = xu + xu^2 \frac{\mathcal{T}(u) - \mathcal{T}(1)}{u-1} + (\mathcal{R}(u) - xu)\mathcal{T}'(1),$$

$$\mathcal{R}(u) = xu + xu\mathcal{R}(u) + (\mathcal{R}(u) - xu)\mathcal{R}(1),$$

$$\mathcal{P}(1) = x + \frac{x^2}{(1-x)^2} + \frac{x}{1-x} (\mathcal{P}(1) - x),$$

where $\mathcal{F}'(1) = \frac{\partial \mathcal{F}}{\partial u}(x,1)$ and similarly for $\mathcal{T}'(1)$. Moreover,

$$\mathcal{F}(u) = \frac{\mathcal{T}(u)}{1 - \mathcal{T}(1)}.$$  (7)

For the smooth case,

$$\mathcal{S}(v) = xv(1-x) + x\mathcal{S}(v) + xv(1-x) \frac{v\mathcal{S}(v) - \mathcal{S}(1)}{v-1} + \left( \frac{\mathcal{R}(v) - xv(1-x)}{1-xv} \right) ((1-x)(\mathcal{S}'(1) + \mathcal{S}(1)) - x)$$

where

$$\mathcal{R}(v) = \frac{xv(1-x)}{1-xv} + x\mathcal{R}(v) + \left( \frac{\mathcal{R}(v) - xv(1-x)}{1-xv} \right) \mathcal{R}(1).$$

Proof. We first consider $\mathcal{F}(u)$. The terms $xu + xu\mathcal{F}(1)$ count forest-like permutations with $\pi^{-1}(1) = |\pi|$, which have only one $rl$-minimum. For the remaining forest-like permutations we use Proposition 10. The generating function of permutations $\sigma$ such that $\tau = 1$ is:

$$\sum_{\ell} \mathcal{F}_\ell(x) \sum_{k=1}^\ell xu^{k+1} = xu^2 \sum_{\ell} \mathcal{F}_\ell(x) \frac{u^\ell - 1}{u-1} = xu^2 \frac{\mathcal{F}(u) - \mathcal{F}(1)}{u-1},$$

while for the permutations such that $\tau \neq 1$ we obtain:

$$\sum_{\ell} \mathcal{F}_\ell(x) \sum_{k=1}^\ell (\mathcal{R}(u) - xu) = (\mathcal{R}(u) - xu)\mathcal{F}'(1).$$

Combining all cases gives the result for $\mathcal{F}(u)$.

The equation for $\mathcal{T}(u)$ is proved in a similar way (note that there is no counterpart to the term $xu\mathcal{F}(1)$ since this corresponds to forests where 1 is an isolated vertex).

For rooted tree-like permutations there is no choice in the way we merge $\tau$ and $\sigma$ and so we obtain a significantly simpler equation (see Figure 3).

The equation we have obtained for $\mathcal{F}(u)$ shows that the indeterminate $u$ is needed to exploit the decomposition of Proposition 10. This is not the case for path-like permutations, and this is why we will not take into account the number of $rl$-minima. If $\sigma$ is path-like, the graph $G_\sigma$ has exactly 2 vertices of degree 1, unless $\sigma = 1$. If $\sigma$ is increasing, both of these end vertices correspond to $rl$-minima. Otherwise, only the largest one does (Figure 3). The term $x^2/(1-x)^2$ in the equation corresponds to the case where $\sigma$ is increasing and $k = |\sigma|$. The term $x/(1-x)(\mathcal{P}(1) - x)$ corresponds to the case $k < |\sigma|$. 
The relationship (4) can be explained by noting that a forest-like permutation \( \pi \) is either tree-like, or is obtained by appending a tree-like permutation \( \tau \) to the beginning of another forest-like permutation \( \sigma \). More formally,

\[
\pi = (\tau(1) + h)(\tau(2) + h)\cdots(\tau(k) + h)\sigma(1)\sigma(2)\cdots\sigma(h),
\]

where \( \tau \) is tree-like and \( \sigma \) is forest-like. Note that \( m(\pi) = m(\sigma) \). In terms of generating functions, this gives \( F(u) = T(u) + T(1)F(u) \).

We now proceed with the smooth case. Let us first determine the generating function \( S_0(v) \) counting the smooth permutations \( \pi \) such that \( a(\pi) = m(\pi) \) if and only if \( \tau = 1 \) and \( k \leq a(\sigma) \). Hence

\[
S_0(v) = xv(1 + S(1)) + x\sum_{\ell} S_\ell(x)\sum_{k=1}^{\ell} v^{k+1} = xv + xv\frac{vS(v) - S(1)}{v - 1}.
\]

Combining this with (4), it follows that the smooth permutations \( \pi = \Phi(\tau, \sigma, k) \) such that \( \tau = 1 \) but \( k = m(\sigma) > a(\sigma) \) are counted by

\[
x(S(v) - S_0(v)).
\]

In the case where \( \tau \neq 1 \) is increasing, we obtain the series

\[
\frac{x^2v}{1-xv} (S'(1) + S(1) - S_0(1))
\]

while in the case where \( \tau \) is not increasing, we find:

\[
\left( \frac{\overline{R}(v) - xv}{1-xv} \right)(S'(1) + S(1) - S_0(1)).
\]

The series \( S(v) \) is the sum of \( S_0(v) \) and \( S(1) \). This gives the desired functional equation for \( S(v) \).

It remains to count rooted tree-like permutations by the length of the final ascent. We obtain an equation for \( \overline{R}(v) \) by specializing the above study to the rooted case, that is to say, to the case where \( \sigma \) is rooted and \( k = m(\sigma) \). The counterparts of the terms (4–5) are respectively

\[
\overline{R}_0(v) = \frac{xv}{1-xv}, \quad x(\overline{R}(v) - \overline{R}_0(v)), \quad \frac{x^2v}{1-xv} R(1) \quad \text{and} \quad \left( \frac{\overline{R}(v) - xv}{1-xv} \right) R(1).
\]

The sum of these four terms is \( \overline{R}(v) \), and this gives the desired equation. \( \square \)

### 3.3. Solution of the functional equations

We are finally going to solve the equations of Proposition 11 to obtain Theorem 8. Three of them do not raise any difficulty. Namely, the equation defining \( P(1) \) is readily solved, while the equations defining \( R(u) \) and \( \overline{R}(v) \) can be solved by first setting \( u = 1 \) (or \( v = 1 \)) to determine the value of these series at \( u = 1 \) (or \( v = 1 \)) and then using these preliminary results to compute the full series. In particular,

\[
R(u) = \frac{xu(2 - u - u\sqrt{1 - 4x})}{2(1 - u + xu^2)} = \frac{xu}{1 - uR(1)}.
\]

The other three equations (defining \( R, T \) and \( S \)) involve divided differences of the form

\[
\frac{A(u) - A(1)}{u - 1}
\]

and cannot be solved by setting \( u = 1 \). Instead, we will solve them by using the kernel method \( \overline{3} \overline{8} \). Consider for instance the equation for tree-like permutations.
This is a linear equation with one catalytic variable \((u)\) and two additional unknown functions \((T(1)\) and \(T'(1)\)). However, these two functions are not independent: by taking the limit as \(u\) goes to 1 in the equation we find

\[
T(1) = x + R(1)T'(1).
\]

(13)

The coefficient of \(T(u)\) in the equation defining \(T(u)\) is

\[
1 - \frac{xu^2}{u - 1} = \frac{u - 1 - xu^2}{u - 1},
\]

which vanishes for two values of \(u\). One of these values is a formal power series in \(x\), \(U \equiv U(x) = \frac{1 - \sqrt{1 - 4x}}{2x}\).

Replacing \(u\) by \(U\) in the functional equation gives a second linear relation between \(T(1)\) and \(T'(1)\):

\[
0 = xU - T(1) + (R(U) - xU)T'(1).
\]

(14)

One can now solve (13) and (14) for \(T(1)\) and \(T'(1)\), in terms of \(x, U, R(1)\) and \(R(U)\). Then the solution can be written as a pair of rational functions of \(U\) using:

- the expression of \(R(U)\) in terms of \(x, U\) and \(R(1)\) (see (12)),
- the fact that \(R(1) = xU\),
- the equation \(x = (U - 1)/U^2\).

Replacing the expressions of \(T(1)\) and \(T'(1)\) in the original functional equation gives an expression for \(T(x, u)\) in terms of \(x, u\) and \(U\), which can be rewritten as

\[
T(x, u) = xu \frac{(1 + V)^2(1 - 2V) - uV(1 - 2V - 2V^2)}{(1 - 2V)(1 + V - uV)^2}
\]

(15)

where

\[
V = U - 1 = \frac{1 - 2x - \sqrt{1 - 4x}}{2x}.
\]

We can use similar techniques to find \(F(u)\). However, it is easier to use (1) and what we have obtained for \(T\) to get

\[
F(x, u) = uV \frac{(1 + V)^2(1 - 2V) - uV(1 - 2V - 2V^2)}{(1 - 2V - V^2 - V^3)(1 + V - uV)^2}
\]

(16)

where \(V\) is given above.

The solution of the equation defining \(S(u)\) is similar to what we have done for \(T(u)\). One possible expression of the bivariate series that counts smooth permutations by the length and the length of the final ascent is

\[
S(x, u) = xu \frac{(1 + V)(1 - V^2 - V^3) - Vu (1 - V - V^2 - V^3)}{(1 + V - uV)(1 - V - V^2 - V^3)(1 - xu)}.
\]

(17)

Putting \(u = 1\) into equations (13), (16) and (17) and simplifying then gives the results of Theorem 3.

4. Final comments and open questions

We first show that several bijections are underlying the results presented in this paper. We then raise a number of questions of an enumerative or graph-theoretic nature.
4.1. Bijections

In what follows, we discuss three objects closely related to the graph $G_\pi$: first the graph itself, second its oriented version $\vec{G}_\pi$ (each edge is oriented from the vertex with the lower label to the vertex with the higher label), and finally its natural embedding in $\mathbb{N}^2$ (where the vertex $i$ is placed at position $(i, \pi(i))$ and the edges are represented by straight lines, as on the left of Figure 1).

4.1.1. The graph $G_\pi$. We first note that the map $\pi \mapsto G_\pi$ is injective. That is, one can recover $\pi$ from $G_\pi$. To see this, orient $G_\pi$ to obtain $\vec{G}_\pi$. Then, for every vertex $i$ in $\vec{G}_\pi$, let $a(i)$ be the number of vertices that can be reached from $i$ by a directed path. This is the number of $j \geq i$ such that $\pi(j) \geq \pi(i)$, and the sequence $\pi(1), \pi(2), \ldots, \pi(n)$ can be easily reconstructed (in this order) from the list $(a(1), \ldots, a(n))$. For instance, if $m$ vertices can be reached from 1 (that is, $a(1) = m$), then it means that $\pi(1) = n - m + 1$, and so on (we have assumed implicitly that $|\pi| = n$).

As noted at the beginning of the paper, $\vec{G}_\pi$ is the Hasse diagram of a certain poset $P$ on $[n] = \{1, 2, \ldots, n\}$. The underlying order is natural, meaning that if $i < j$ in $P$, then $i < j$ in $\mathbb{N}$. (We refer to [19, Chap. 3] for generalities on posets.) The $n!$ permutations of $S_n$ thus provide $n!$ distinct natural orders on $[n]$. Not all natural orders are obtained in that way: even for $n = 3$, there are 7 natural orders but only 6 permutations, and the poset in which the only relation is 1 < 3 is not obtained from any permutation (Figure 5). The posets that are actually obtained from the construction $\pi \mapsto \vec{G}_\pi$ are, by definition, the natural orders on $[n]$ of dimension 2 [19, Exercise 3.10].

Some graph properties of $G_\pi$ easily follow from the construction. For instance, $G_\pi$ is isomorphic to $G_{\pi^{-1}}$ (more precisely, $G_{\pi^{-1}}$ is obtained by relabelling the vertex $i$ by $\pi(i)$). The natural embedding of $G_{\pi^{-1}}$ is obtained by reflecting the embedding of $G_\pi$ through the main diagonal. Of course, $G_\pi$ is triangle free (every Hasse diagram is). In particular, $e(\pi) \leq \lfloor n^2/4 \rfloor$ (see [2]) and it is easy to construct a permutation showing that this bound is tight. The number of edges of $G_\pi$ can also be interpreted in terms of Bruhat order (see [19, Exercise 3.75], [6, Ch. 2]): it is the number of permutations covering (poset-wise) $\pi$ in the Bruhat order.

4.1.2. Rooted tree-like permutations. Here, we want to show a simple bijection between rooted tree-like permutations of size $n$ and plane trees with $n - 1$ edges. This explains why such permutations are counted by the Catalan number $C_{n-1}$. Recall that permutations $\pi$ avoiding 21354 are exactly those such that the natural embedding of $G_\pi$ is planar (see the remark following Lemma 3). This holds in particular for rooted tree-like permutations: the embedding of $G_\pi$ is thus a (rooted) plane tree. Then, observe that the decomposition of forest-like permutations illustrated in Figure 4, once specialized to rooted permutations, coincides with the standard
decomposition of plane trees (a left subtree joined to the root by an edge, and another plane tree, see Figure 6). This means that every plane tree is obtained from exactly one rooted tree-like permutation. This is illustrated in Figure 7 for permutations of length 4.

\[ \pi(1) = 1 \]

\[ \pi(2) \]

\[ \tau \]

\[ \emptyset \]

**Figure 6.** The decomposition of rooted tree-like permutations.

4.1.3. *Path-like permutations.* Consider a path-like permutation \( \pi \) of length at least 2. The graph \( G_\pi \) has two vertices of degree 1. Define a word \( W(\pi) \) on the alphabet \( \{U, D\} \) by following the path \( G_\pi \) from the vertex of degree 1 with the lowest label to the other vertex of degree 1, encoding each edge of this path by a letter \( U \) (like up) or \( D \) (like down) depending on how the labels of the vertices vary along this edge. Examples are shown in Figure 8. It turns out that the map \( W \) is a bijection from path-like permutations of length \( n \) to words of length \( n - 1 \) distinct from \( D^{n-1} \). In particular, this explains why the number of path-like permutations of length \( n \) is \( 2^{n-1} - 1 \).

\[ \{U, D\}^* - D^+ \]

**Figure 7.** The 5 rooted tree-like permutations of length 4 and the corresponding plane trees.

**Figure 8.** The 7 path-like permutations of length 4 and the corresponding words.

Again, this result follows from the decomposition of path-like permutations that led to the equation of Proposition 11. Indeed, this decomposition gives, for the noncommutative generating function defined by

\[ P = \sum_{\pi \text{ path-like}} W(\pi) \]

the following equation:

\[ P = \epsilon + U^+ + D^+U^+ + (P - \epsilon)D^*U^*, \]

where \( \epsilon \) denotes the empty word and we have used the standard notation \( D^+ = \sum_{i \geq 1} D^i \) and \( U^* = \sum_{i \geq 0} U^i \). It is easy to see that the solution of this equation is

\[ P = \{U, D\}^* - D^+. \]
That is to say, the non-empty words $W(\pi)$ are those containing at least one $U$, and each such word corresponds to a unique path-like permutation.

4.2. Open problems

4.2.1. Enumeration. In this paper, we have characterized and counted forest-like permutations and some of their natural subclasses. This work raises similar questions for several supersets of forest-like permutations. The most natural ones are probably the following two:

1. what is the number of plane permutations of $S_n$, that is to say, permutations avoiding 21354?
2. what is the number of permutations associated with a Gorenstein Shubert variety? These permutations generalize forest-like permutations, and have been characterized in [23].

We also recall that the enumeration of 1324 avoiding permutations is still an open problem [3, 18]. Permutations avoiding 2143 are called vexillary and are equinumerous with 1234 avoiding permutations [4, 21], which have been enumerated in [3].

Another natural question is to count permutations $\pi$ by their length and the number $e(\pi)$ of bars in their bar diagram (which is the number of permutations covering $\pi$ in the Bruhat order). To our knowledge, the bivariate series

$$E(t, x) = \sum_{n \geq 0} \frac{t^n}{n!} \sum_{\pi \in S_n} x^{e(\pi)}$$

is not known. However, the total number of edges in the bar diagrams of permutations of $S_n$ is known: if

$$e(n) = \sum_{\pi \in S_n} e(\pi),$$

then

$$e(n) = (n + 1)! (H(n + 1) - 2) + n!$$

where $H(n) = 1 + 1/2 + \cdots + 1/n$ is the $n$th harmonic number. Indeed, as communicated to us by David Callan, it is not hard to see that the number of permutations of $S_n$ having a bar going from $i$ to $j$, with $i < j$, is $n!/(j - i + 1)$, and the above result follows easily.

Note that $e(n)$ is also the number of edges in the Hasse diagram of the Bruhat order of $S_n$. The exponential generating function of the numbers $e(n)$ is

$$\sum_{n \geq 0} e(n) \frac{t^n}{n!} = \frac{\partial E}{\partial x}(t, 1) = \frac{1}{(1 - t)^2} \left( \log \frac{1}{1 - t} - t \right).$$
The average number of bars in a permutation of $S_n$ is
\[
\frac{e(n)}{n!} = \log(n) n + (-2 + \gamma) n + \log(n) + 1/2 + \gamma + O(1/n)
\]
where $\gamma$ is Euler’s constant. This can be compared to the average number of non-inversions, which is known to be $n(n+1)/4$. Related questions have recently been studied in [1].

4.2.2. Graph questions. We have seen that the labeled graphs obtained from the map $\pi \mapsto G_\pi$ are the Hasse diagrams of natural orders of dimension 2. One can also wonder which unlabelled graphs are obtained through our construction. Clearly, these graphs must be triangle free. However, this is not a sufficient condition. For example, by an exhaustive computer search one can verify that the triangle-free graph formed of the vertices and edges of a cube is not produced from any permutation in $S_8$.

Note that, by Section 4.1.2, all unlabelled trees (and thus all unlabelled forests) are obtained through our construction.

Acknowledgements. The authors would like to thank David Callan, Aurélie Cortez, Alexander Woo and Alexander Yong for their advice and also the anonymous referees whose comments on a former draft of this paper greatly improved its quality.

References


CNRS, LaBRI, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex, France

E-mail address: mireille.bousquet@labri.fr

Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112, USA

E-mail address: sbutler@math.ucsd.edu