Posets and Permutations in the Duplication-Loss Model: Minimal Permutations with d Descents.

Mathilde Bouvel, Elisa Pergola

Posets and Permutations in the Duplication-Loss Model: Minimal Permutations with d Descents.

Mathilde Bouvel [*]   Elisa Pergola [†]

June 9, 2008

Abstract

In this paper, we are interested in the combinatorial analysis of the whole genome duplication - random loss model of genome rearrangement initiated in [8] and [7]. In this model, genomes composed of n genes are modelled by permutations of the set of integers $[1..n]$, that can evolve through duplication-loss steps. It was previously shown that the class of permutations obtained in this model after a given number p of steps is a class of pattern-avoiding permutations of finite basis. The excluded patterns were described as the minimal permutations with $d = 2^p$ descents, minimal being intended in the sense of the pattern-involvement relation on permutations. Here, we give a local and simpler characterization of the set $\mathcal{B}_d$ of minimal permutations with d descents. We also provide a more detailed analysis - characterization, bijection and enumeration - of two particular subsets of $\mathcal{B}_d$, namely the patterns in $\mathcal{B}_d$ of size $d + 2$ and $2d$.

1 Pattern-avoidance in the duplication-loss model

The study of genome evolution has been the source of extensive research in computational biology in the last decades. Many models for genome evolution were defined, taking into account various biological phenomema (see [4], [9], [10] for recent examples in literature). Among them, the *tandem duplication - random loss model* represents genomes with permutations, that can evolve through *duplication-loss steps* representing the biological phenomenon that duplicates fragments of genomes, and then loses one copy of every duplicated gene. For the original biological motivations, we refer to [8]. In this first section, we describe the duplication-loss model, and recall some previous results obtained by other authors. We recall some definitions and properties on pattern-avoidance that are necessary to introduce the permutations that will arise from this model and on which we will focus in the rest of the paper.

---

[*]LIAFA, Université Paris Diderot - Paris 7, Case 7014, 75205 Paris Cedex 13, France, `mbouvel@liafa.jussieu.fr`

[†]Università degli Studi di Firenze, Dipartimento di Sistemi e Informatica, v.le Morgagni 65, 50134 Firenze, Italy, `elisa@dsi.unifi.it`

1.1 The tandem duplication - random loss model for genome evolution

A permutation of size n is a bijective map from $[1..n]$ to itself. We denote by S_n the set of permutations of size n. We consider a permutation $\sigma \in S_n$ as the word $\sigma_1 \sigma_2 \ldots \sigma_n$ of n letters on the alphabet $\{1, 2, \ldots, n\}$, containing exactly once each letter (we often prefer the word *element* instead of letter). For example, 346251 represents the permutation $\sigma \in S_6$ such that $\sigma_1 = 3, \sigma_2 = 4, \ldots, \sigma_6 = 1$.

In our model, permutations can be modified by *duplication-loss steps*. Each of these steps is composed of two elementary operations. Firstly, a fragment of consecutive elements of the permutation is duplicated, and the duplicated fragment is inserted immediately after the original copy: this is the *tandem duplication*. After this first operation, any duplicated element appears twice in the sequence of integers (that is no more a permutation at this stage). Then the *random loss* occurs: one copy of every duplicated element is lost, so that we get a permutation at the end of the step. For any duplication-loss step, we call its *width* the number of elements that are duplicated.

$$1\,2\,\overbrace{3\,4\,5\,6}\,7 \quad \rightsquigarrow \quad 1\,2\,\overbrace{3\,4\,5\,6}\,\overbrace{3\,4\,5\,6}\,7$$

$$\text{(tandem duplication)}$$

$$\rightsquigarrow \quad 1\,2\,\cancel{3}\,4\,5\,\cancel{6}\,3\,\cancel{4}\,\cancel{5}\,6\,7$$

$$\text{(random loss)}$$

$$\rightsquigarrow \quad 1\,2\,4\,5\,3\,6\,7$$

Figure 1: Example of one step of tandem duplication - random loss of width 4

Notice that the duplication-loss model is a particular case of the very general framework for transforming permutations defined in [1]: the *permuting machines*. A permuting machine takes a permutation in input and performs on it a transformation that satisfies the two properties of independance with respect to the values and of stability with respect to pattern-involvement (see [1] for more details). These two properties are satisfied by the duplication-loss transformation.

We will consider permutations that are obtained from an identity permutation $12\ldots n$ after a given number p of duplication-loss steps, that is to say that are the output of a combination in series of p permuting machines with input $12\ldots n$. The reason is that these permuations are the ones obtainable at a cost of at most p in the duplication-loss model with a particular *cost function*.

Indeed, various duplication-loss models can be defined depending on the cost function $c \in \overline{\mathbb{R}}^{\mathbb{N}}$ that is chosen. We will always assume that the cost $c(k)$ of a duplication-loss step is dependant only on the width k of this step. In the original model of Chaudhuri, Chen, Mihaescu and Rao [8], the *cost* of a duplication-loss step of width k is $c(k) = \alpha^k$, for a parameter $\alpha \geq 1$. In [7], we consider the cost function defined by $c(k) = 1$ if $k \leq K$, $c(k) = \infty$ otherwise, for a parameter $K \in \overline{\mathbb{N}} \setminus \{0, 1\}$. The model we will focus on in what follows has a very simple cost function, namely $c(k) = 1, \forall k$. It is a special case of both the

model of [8] (with $\alpha = 1$) and the model of [7] (with $K = \infty$). This particular model is called the *whole genome duplication - random loss model*: indeed, since any step has cost 1 no matter its width, we can assume *w.l.o.g* that the whole permutation is duplicated at any step.

As said before, we are now going to focus on permutations obtained from an identity permutation $12 \ldots n$ after a certain number p of duplication-loss steps in the whole genome duplication - random loss model, that is to say on permutations obtainable at a cost of at most p in this model. We will describe combinatorial properties of those permutations in Subsection 1.3, in terms of pattern-avoidance.

1.2 Previous results on the duplication-loss model

The permutations obtainable in at most p duplication-loss steps in the whole genome duplication - random loss model were implicitly characterized in [8], through Theorem 1:

Theorem 1. *Let $\sigma \in S_n$. In the whole genome duplication - random loss model, $\lceil \log_2 (\text{number of maximal increasing substrings of } \sigma) \rceil$ steps are necessary and sufficient to obtain σ from $12 \ldots n$.*

An *increasing substring* of σ is just a sequence of consecutive elements of σ that are in increasing order. An increasing substring is maximal if it can be extended neither on the left nor on the right.

Example 1. *For example, 698413725 contains 5 maximal increasing substrings that are $69, 8, 4, 137$ and 25.*

In [7], we reformulated Theorem 1 into Theorem 2, introducing, instead of the number of maximal increasing substrings, the number of descents which is a very well-known statistics on permutations.

Definition 1. *Given a permutation σ of size n, we say that there is a* descent *(resp.* ascent*) at position i, $1 \leq i \leq n-1$, if $\sigma_i > \sigma_{i+1}$ (resp. $\sigma_i < \sigma_{i+1}$). We indicate the number of descents of the permutation σ by $desc(\sigma)$.*

Example 2. *For example, $\sigma = 698413725$ has 4 descents, namely at positions $2, 3, 4, 7$.*

It is often convenient to see permutations through their grid reprensentation defined in [5] and described in Figure 2, especially because it gives a better view of descents and ascents.

Obviously, we have:

Remark 1. *The number of maximal increasing substrings of a permutation σ is $desc(\sigma) + 1$.*

More precisely, the positions of the descents and n indicate the positions of the last elements of the maximal increasing substrings of σ.

These definitions allow us to state Theorem 2:

Theorem 2. *The permutations that can be obtained in at most p steps in the whole genome duplication - random loss model are exactly those whose number of descents is at most $2^p - 1$.*
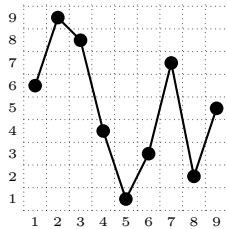
Figure 2: The grid representation of the permutation $\sigma = 698413725$

Proof. By Theorem 1, the permutations obtainable in at most p steps are exactly those having their number of maximal increasing substrings at most 2^p, that is to say having at most $2^p - 1$ descents by Remark 1. □

Generalizing a little, we will focus in the remaining of the paper on the set of permutations with at most d descents, without assuming that d is of the form $d = 2^p - 1$. We can notice that this corresponds to the set of permutations composed of $d + 1$ increasing sequences, separated either by ascents or by descents (a permutation may have more than one such decomposition). In [2] this set is denoted $W(e_1, \ldots, e_{d+1})$ with $\forall i, e_i = +$. In this paper, and as an application of their results, the authors are concerned with properties of $W(e_1, \ldots, e_{d+1})$ in terms of pattern-avoidance, and they prove that this set is a finitely based pattern-avoiding permutation class. Our work can be seen as a more detailed analysis of this particular result.

1.3 Pattern-avoidance in the duplication-loss model

We need to recall a few definitions on pattern-avoidance in permutations to proceed.

Definition 2. *A permutation $\pi \in S_k$ is a* pattern *of a permutation $\sigma \in S_n$ if there is a subsequence of σ which is order-isomorphic to π; i.e., if there is a subsequence $\sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_k}$ of σ (with $1 \le i_1 < i_2 < \ldots < i_k \le n$) such that $\sigma_{i_\ell} < \sigma_{i_m}$ whenever $\pi_\ell < \pi_m$.*
We also say that π is involved *in σ and call $\sigma_{i_1} \sigma_{i_2} \ldots \sigma_{i_k}$ an* occurrence *of π in σ.*

Example 3. *For example $\sigma = 142563$ contains the pattern $\pi = 1342$; and 1563, 1463, 2563 and 1453 are the occurrences of this pattern in σ. But σ does not contain the pattern 321 as no subsequence of size 3 of σ is isomorphic to 321, i.e., is decreasing.*

We write $\pi \prec \sigma$ to denote that π is a pattern of σ. We say that a set $\mathcal{C}$ of permutations is *stable for* $\prec$ if, for any $\sigma \in \mathcal{C}$, for any $\pi \prec \sigma$, then we also have $\pi \in \mathcal{C}$.

A permutation σ that does not contain π as a pattern is said to *avoid* π. The class of all permutations avoiding the patterns $\pi_1, \pi_2 \ldots \pi_k$ is denoted $S(\pi_1, \pi_2, \ldots, \pi_k)$. We say that $S(\pi_1, \pi_2, \ldots, \pi_k)$ is a class of pattern-avoiding permutations of *basis* $\{\pi_1, \pi_2, \ldots, \pi_k\}$. The basis of a class of pattern-avoiding

4

permutations may be finite or infinite. Pattern-avoiding permutation classes considered in the literature (see for example [6], [12] and their references) are often of finite basis.

Although it may sound a powerful statement, it is simple to understand that:

Proposition 1. *A set $\mathcal{C}$ of permutations that is stable for $\prec$ is a class of pattern-avoiding permutations. However, its basis might be infinite.*

Proof. Consider $\mathcal{C}$ a set of permutations that is stable for $\prec$. Define $\mathcal{B}$ to be the set of minimal permutations that do not belong to $\mathcal{C}$, minimal being intended in the sense of $\prec$. More formally, $\mathcal{B} = \{\sigma \notin \mathcal{C} : \forall \pi \prec \sigma$ with $\pi \neq \sigma, \pi \in \mathcal{C}\}$. We claim that $\mathcal{C} = S(\mathcal{B})$. Indeed, take $\sigma \notin S(\mathcal{B})$. Then there exists $\pi \in \mathcal{B}$ such that $\pi \prec \sigma$. Since $\pi \in \mathcal{B}$, $\pi \notin \mathcal{C}$ and considering that $\mathcal{C}$ is stable for $\prec$, we deduce from $\pi \prec \sigma$ that $\sigma \notin \mathcal{C}$ either. Conversely, if $\sigma \notin \mathcal{C}$, then either $\sigma \in \mathcal{B}$ (and consequently $\sigma \notin S(\mathcal{B})$) or there exists $\pi \prec \sigma$ with $\pi \neq \sigma$ such that $\pi \notin \mathcal{C}$. In this second case, by induction we obtain that $\sigma \notin S(\mathcal{B})$.

We conclude that the set $\mathcal{C}$ is a class of pattern-avoiding permutations whose basis $\mathcal{B} = \{\sigma \notin \mathcal{C} : \forall \pi \prec \sigma$ with $\pi \neq \sigma, \pi \in \mathcal{C}\}$ has no reason *a priori* to be finite. $\square$

In [7], we proved that classes of permutations defined in duplication-loss models, as the permutations obtained in at most a given number p of steps, are classes of pattern-avoiding permutations. We have not always been able to find the basis, even though we have proved in any case we considered that this basis is finite. In this paper, we take into consideration in particular the following result:

Theorem 3. *The class of permutations obtainable in at most p steps in the whole genome duplication - random loss model is a class of pattern-avoiding permutations whose basis $\mathcal{B}_d$ is finite and is composed of the minimal permutations with $d = 2^p$ descents, minimal being intented in the sense of $\prec$.*

The proof of Theorem 3 we gave in [7] is implicit, and for sake of clarity we give below an explicit proof of it.

Proof. Let us denote by $\mathcal{C}_p$ the class of permutations obtainable in at most p steps in the whole genome duplication - random loss model.

We first prove that $\mathcal{C}_p$ is stable for $\prec$. Consider $\sigma \in \mathcal{C}_p$ of size n and π of size $k \leq n$ such that $\pi \prec \sigma$. There is a sequence of at most p duplication-loss steps that transforms $12 \ldots n$ into σ. By definition, π has an occurrence in σ. In the duplication-loss scenario for σ, if you keep track only of the elements that form an occurrence of π, you obtain a sequence of duplication-loss steps moving from $12 \ldots k$ to π, of no more than p steps. This shows that $\pi \in \mathcal{C}_p$, and consequently that $\mathcal{C}_p$ is stable for $\prec$.

According to Proposition 1, $\mathcal{C}_p$ is a class of pattern-avoiding permutations whose basis is $\{\sigma \notin \mathcal{C}_p : \forall \pi \prec \sigma$ with $\pi \neq \sigma, \pi \in \mathcal{C}_p\}$. Following Theorem 2, we deduce that this basis $\mathcal{B}_d$ of excluded patterns is made of the minimal permutations with $d = 2^p$ descents, that is to say the permutations with 2^p descents that contain no pattern with 2^p descents, except themselves. What is left to prove is that this basis is finite.

It is sufficient to establish an upper bound on the size of the permutations in $\mathcal{B}_d$ to show that $\mathcal{B}_d$ is finite. We postpone this part of the proof to Proposition

2, where we show in particular that the permutations of $\mathcal{B}_d$ are of size at most $2d$. A consequence is that the basis $\mathcal{B}_d$ of excluded patterns of $\mathcal{C}_p$ is finite. □

In this paper, we focus on the basis $\mathcal{B}_d$ of excluded patterns appearing in Theorem 3. More generally, we do not assume that d is a power of 2 but rather wish to characterize and enumerate the set $\mathcal{B}_d$ of permutations that are the minimal ones in the sense of $\prec$ for the property of having d descents.

1.4 Outline of the paper

In this paper, we focus on the sets $\mathcal{B}_d$ of permutations that are the minimal ones in the sense of $\prec$ for the property of having d descents. For the cases $d = 2^p$, $\mathcal{B}_d$ is the basis of excluded patterns of the class of permutations obtainable in at most p steps in the whole genome duplication - random loss model.

The work that is presented hereafter is organized as follows. First, we give a local characterization of the permutations of $\mathcal{B}_d$. Indeed, the definition of these permutations as the minimal ones with respect to $\prec$ for the property of having d descents is not very easy to use. We will prove in Section 2 that the permutations of $\mathcal{B}_d$ are the permutations σ whose ascents satisfy a simple and local property: there is an ascent in $\sigma \in S_n$ at position i if and only if $2 \leq i \leq n - 2$ and $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ forms an occurrence of either the pattern 2143 or the pattern 3142.

This characterization is used to try and count the permutations in $\mathcal{B}_d$. Despite our effort, we did not succeed in this direction, and focused on simpler cases that can be seen as a first step in the enumeration of $\mathcal{B}_d$. First, as explained at the beginning of Section 2, Proposition 2, the size of the permutations in $\mathcal{B}_d$ is at least $d + 1$ and at most $2d$. Obviously there is only one permutation of $\mathcal{B}_d$ of size $d+1$, that is the reverse identity permutation $(d+1)d(d-1)\ldots321$. For any other size, there is no immediate result. Using a representation of permutations of $\mathcal{B}_d$ as posets (partially ordered sets), we could enumerate the permutations in $\mathcal{B}_d$ having size $2d$ and $d + 2$ respectively.

In Section 3, we prove that the permutations of $\mathcal{B}_d$ having size $2d$ (i.e. maximal size) are enumerated by the Catalan numbers: there are $C_d = \frac{1}{d+1}\binom{2d}{d}$ of them. We give two possible proofs of this result. We describe an *"ECO" generation* (see [3]) of the permutations of $\mathcal{B}_d$ of size $2d$ whose associated succession rule is known to correspond to the Catalan numbers. More directly, we could provide a simple bijection between Dyck paths of length $2d$ and an adequate representation of the permutations of size $2d$ in $\mathcal{B}_d$.

In Section 4, we consider permutations of size $d+2$ (minimal non-trivial case) in $\mathcal{B}_d$. After a combinatorial analysis and some computations, we obtain that there are $s_d = 2^{d+2} - (d+1)(d+2) - 2$ such permutations. The sequence (s_d) does not appear in the Online Encyclopedia of Integer Sequences [11]. However, we realized that the sequence $(\frac{s_d}{2})$ does. This sequence counts the number of non-interval subsets of the set $\{1, 2, \ldots, d + 1\}$. Section 4 also gives a bijective proof of the fact that there are twice as many permutations of size $d + 2$ in $\mathcal{B}_d$ as non-interval subsets of $\{1, 2, \ldots, d + 1\}$.

Section 5 summarizes some open problems in the study of the sets $\mathcal{B}_d$'s.

From here on, by minimal permutation with d descents, we mean a permutation that is minimal in the sense of the pattern-involvement relation $\prec$ for the

property of having d descents.

Example 4, which is illustrated on Figure 3, should clarify the notion of minimal permutation with d descents.

Example 4. *Permutation* $\sigma = 8613241\underline{1}195\underline{1}07$ *has* 6 *descents but is not minimal with* 6 *descents. Indeed, the elements* 1 *and* 4 *(that are circled on Figure 3) can be removed from* σ *without changing the number of descents.*

Doing this, we obtain permutation $\pi = 642197385$ *which is minimal with* 6 *descents: it is impossible to remove an element from it while preserving the number of descents equal to* 6.

However, π *is not of minimal* size *among the permutation with* 6 *descents:* π *has size* 9 *whereas permutation* 7654321 *has* 6 *descents but size* 7.
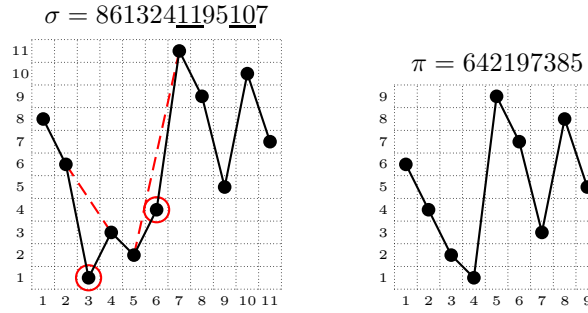

Figure 3: Permutations σ and π of Example 4.

2 A characterization for minimal permutations with d descents

The aim of this section is to provide a more practical characterization of minimal permutations with d descents, by finding necessary and sufficient conditions on permutations for being minimal with d descents. First, we provide a necessary condition on the size of those permutations with Proposition 2.

Proposition 2. *Let* σ *be a minimal permutation with* d *descents. Then every ascent of* σ *is immediately preceded and immediately followed by a descent, and the size* n *of* σ *satisfies* $d + 1 \leq n \leq 2d$.

Proof. Consider a permutation $\sigma \in \mathcal{B}_d$, and denote by n the size of σ. By minimality in the sense of $\prec$, σ has exactly d descents. To create a permutation with d descents, you need at least $d + 1$ elements, and with $d + 1$ elements, the only permutation with d descents you can create is $(d+1)d(d-1)\ldots 21$, which is minimal. Therefore, $n \geq d + 1$.

It is also easily seen that σ does neither start nor end with an ascent, otherwise the permutation obtained by removing the first or the last element of σ would have the same number d of descents, contradicting that σ is minimal with d descents. In the same way, σ cannot have two consecutive ascents $\sigma_{i-1}\sigma_i$ and $\sigma_i\sigma_{i+1}$, otherwise we would get the same contradiction removing σ_i, since this removal does not change the number of descents.

This proves that a minimal permutation with d descents is composed of non-empty sequences of descents, separated by isolated ascents. A longest possible permutation with d descents so obtained has d isolated descents, separated by $d-1$ isolated ascents, and consequently has $2d$ elements. We then get that the size of σ is at most $2d$: $n \leq 2d$. □
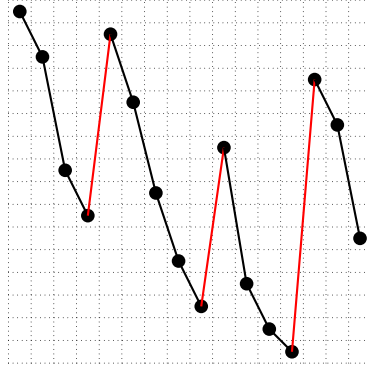

Figure 4: Decomposition of a minimal permutation with d descents into non-empty sequences of descents separated by isolated ascents

The decomposition of a minimal permutation with d descents into non-empty sequences of descents separated by isolated ascents that is described in the proof of Proposition 2 is illustrated in Figure 4. This decomposition can be carried further to give a necessary and sufficient condition on permutations for being minimal with d descents. This characterization is described in Theorem 4.

Theorem 4. *A permutation σ is minimal with d descents if and only if it has exactly d descents and its ascents $\sigma_i\sigma_{i+1}$ are such that $2 \leq i \leq n-2$ and $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ forms an occurrence of either the pattern 2143 or the pattern 3142.*

Proof. Let σ be a minimal permutation with d descents. In the decomposition of σ into non-empty sequences of descents separated by isolated ascents – illustrated in Figure 4 – it appears clearly that an ascent $\sigma_i\sigma_{i+1}$ is necessarily such that $2 \leq i \leq n-2$, with $\sigma_{i-1}\sigma_i$ and $\sigma_{i+1}\sigma_{i+2}$ being descents.

Now, consider an ascent $\sigma_i\sigma_{i+1}$. The previous remarks lead to $\sigma_{i-1} > \sigma_i$, $\sigma_i < \sigma_{i+1}$ and $\sigma_{i+1} > \sigma_{i+2}$.

Let us assume that $\sigma_{i-1} > \sigma_{i+1}$. Then the permutation obtained from σ by the removal of σ_i has as many descents as σ (and one ascent less), contradicting the minimality of σ. Consequently, $\sigma_{i-1} < \sigma_{i+1}$. Similarly, if $\sigma_i > \sigma_{i+2}$, the removal of σ_{i+1} from σ does not change the number of descents, contradicting the minimality of σ. So $\sigma_i < \sigma_{i+2}$ (see Figure 5).

At this point, we have the five following inequalities: $\sigma_{i-1} > \sigma_i$, $\sigma_i < \sigma_{i+1}$, $\sigma_{i+1} > \sigma_{i+2}$, $\sigma_{i-1} < \sigma_{i+1}$ and $\sigma_i < \sigma_{i+2}$. Thanks to them it is possible to check that the sequence $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ is an occurrence of either the pattern 2143 or the pattern 3142.
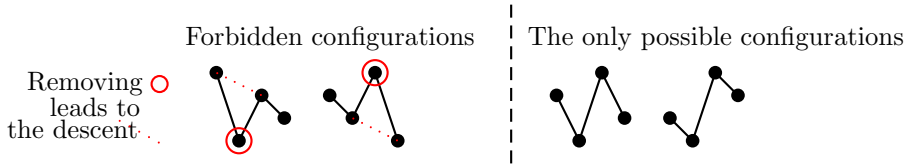
Figure 5: The elements $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ around an ascent $\sigma_i\sigma_{i+1}$ in a permutation σ which is minimal with d descents

Conversely, consider a permutation σ with d descents whose ascents $\sigma_i\sigma_{i+1}$ are such that $2 \leq i \leq n-2$ and $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ forms an occurrence of either the pattern 2143 or the pattern 3142. This implies that σ has the shape of non-empty sequences of descents separated by isolated ascents. And it is a simple matter to prove that the removal of any element of σ makes the number of descents decrease by one – there are three cases to consider: the removed element may be either the first element of an ascent, or the second element of an ascent, or it may be between two descents. This proves that σ is a minimal permutation with d descents. □

We thought this characterization could help us to enumerate the minimal permutations with d descents. Although we did not reach this goal, we still obtain partial results when we studied minimal permutation with d descents and of a given size n. For $n = d + 1$, we already proved that there is only one such permutation. For $n = d + 2$ and $n = 2d$, the next two sections describe the enumeration we obtained. In both cases, we will use a *partially ordered set* (or *poset*) representation of permutations, that comes directly from the characterization of minimal permutations with d descents in Theorem 4.

Representation of minimal permutations with d descents with posets
Consider a set of all the permutations of a given size n, that are minimal with d descents, and having their descents and ascents in the same positions. In all these permutations, the elements are locally ordered in the same way, even around the ascents, because of Theorem 4. We can give a representation of this whole set of permutations by a *partially ordered set* (or *poset*) indicating the necessary conditions on the relative order of the elements between them. For a descent, we just have a link from the first and greatest element to the second and smallest one. For any ascent $\sigma_i\sigma_{i+1}$, the elements $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ form a diamond-shaped structure with σ_{i+1} on the top, σ_i on the bottom, σ_{i-1} on the left and σ_{i+2} on the right. See Figure 6 for an example. By Theorem 4, any labelling of the elements of the poset respecting its ordering constraints is a minimal permutation with d descents.

We will say that a permutation σ satisfies the diamond property when each of its ascent $\sigma_i\sigma_{i+1}$ is such that $\sigma_{i-1}\sigma_i\sigma_{i+1}\sigma_{i+2}$ forms a diamond, that is to say is an occurrence of either 2143 or 3142.
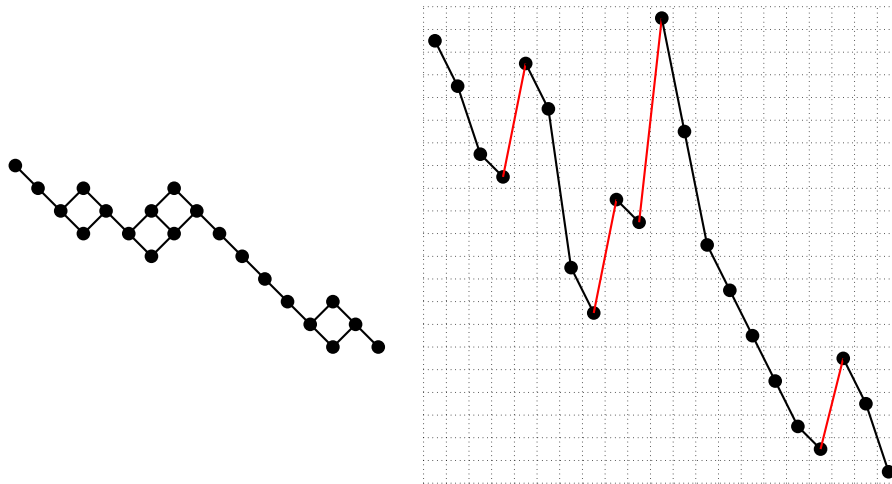
9

Figure 6: A poset representing a set of minimal permutations with 16 descents and 4 ascents (consequently of size 21) containing, among others the permutation 20 18 15 14 19 17 10 8 13 12 21 16 11 9 7 5 3 2 6 4 1 whose grid representation is also given

3 Enumeration of minimal permutations with d descents and of size $2d$

The minimal permutations with d descents that have size $2d$ are, because of minimality, of a very particular shape. Indeed, they cannot have two consecutive ascents as usual, but neither can they have two consecutive descents, otherwise it would be impossible to reach size $2d$. Consequently, they all result from of an alternation of isolated descents and isolated ascents, of course starting and ending with a descent. An example is given in Figure 7(a).
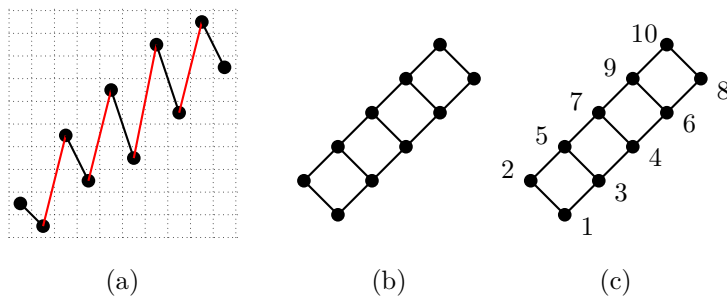

(a) (b) (c)

Figure 7: (a) A minimal permutation $\sigma = 2\ 1\ 5\ 3\ 7\ 4\ 9\ 6\ 10\ 8$ with $d = 5$ descents and of size $2d = 10$, (b) the poset representing the set of all minimal permutations with $d = 5$ descents and of size $2d = 10$ and (c) the authorized labelling of the subsequent poset associated with σ

A consequence is that all minimal permutations of size $2d$ with d descents

10

have their descents and ascents in the same position, so that a unique poset represents the set of all minimal permutations with d descents having size $2d$. This poset has the shape of a ladder with d steps: it is a sequence of $d-1$ diamonds, two consecutive diamonds being linked by an edge. These diamonds correspond to the ascents in the permutations, that are separated by one descent only in this case. See Figure 7(b) for an example.

The paragraph on poset representation at the end of Section 2 justifies Proposition 3:

Proposition 3. *The minimal permutations with d descents and of size $2d$ correspond exactly to the labellings of the ladder poset with d steps with the integers $\{1, 2, \ldots, 2d\}$ that respect its ordering constraints.*

An example of this correspondance is given in Figure 7(c).

The poset representation allows to see at once some properties of minimal permutations with d descents having size $2d$. For example, such a permutation always has 1 as its second element and $2d$ as its next to last element.

The main result of this section is :

Theorem 5. *The minimal permutations with d descents and of size $2d$ are enumerated by the Catalan numbers $C_d = \frac{1}{d+1}\binom{2d}{d}$.*

Proof of Theorem 5 by an analytical method A possible way to prove Theorem 5 is to use the ECO method, presented in details in [3]. In our case, the idea developed by this method is to build all the authorized labellings of the ladder poset with d steps from all the authorized labellings of the ladder poset with $d-1$ steps without creating twice the same labelling.

In its original form, the ECO method builds combinatorial objects of size d from those of size $d-1$, through a process of *local expansion*, whereby the objects are modified only by the addition of an elementary block of object. In our case, in order to get a labelling of size d, the *local expansions* might modify many labels in the labelling of size $d-1$, but the relative order of these labels between them will remain the same. In this sense, we can consider that the expansion is still local.

In the ECO method, the combinatorial objects (labellings of the ladder poset with d steps in our case) receive *labels*. The label of an object is the number of its *children*, that is to say the number of objects that are obtained from it in the local expansion process. Those children can again receive a label by the same method. The infinite tree in which any permutation is the father of its children is called the *generating tree* of the combinatorial class.

With the ECO labelling of the combinatorial objects, we derive a *succession rule* or *rewriting rule* that describes the production (in terms of labels) of the possible labels of these objects, together with a starting point. There is a simple succession rule that is associated with some combinatorial classes enumerated by the Catalan numbers (for example with Dyck paths [3]):

$$\begin{cases} (2) \\ (k) \rightsquigarrow (2)(3) \cdots (k)(k+1) \end{cases}$$

A possible way of proving that authorized labellings of the ladder poset with d steps are enumerated by the Catalan numbers is to find an ECO construction for this class whose associated succession rule is the one above.

11

The ECO labels that are given to authorized labellings of the ladder posets with d steps for this purpose are $(2d - \sigma_{2d} + 1)$, σ_{2d} being the label of the rightmost element of the poset. Notice also that $2d$ is the label of the uppermost element of the poset, and that this element is also the second rightmost one.

Consider an authorized labelling σ of the ladder poset with d steps that has ECO label (k). Its children are the labellings of the ladder poset with $d + 1$ steps obtained by adding a new step on the right, this new step of the ladder being labelled with $2d + 2$ for the top element, and i for the rightmost one, for $2d + 2 - k \le i \le 2d + 1$. The elements j in σ with $j \ge i$ are turned into to $j + 1$ to maintain both the relative order of the elements of σ and the property that these new labellings use all the integers of $[1..2d + 2]$ exactly once.

Since $k = 2d - \sigma_{2d} + 1$, it is easy to check that all the labellings obtained in this way are authorized, and that all of them are obtained. We can now focus on the ECO labels of the children (of size $d+1$) of an authorized labelling of size d with ECO label (k). There are of course k of them whose ECO labels are, by the above formula, $(2(d + 1) - i + 1)$ with $2d + 2 - k \le i \le 2d + 1$, that is to say the children of a labelling with ECO label (k) have labels $(2), (3), \ldots (k), (k+1)$.

The starting point for this ECO construction is the ladder poset with one step provided with its only authorized labelling 21, and whose ECO label is (2).

To sum up, the succession rule obtained for this ECO construction of authorized labelling of the ladder posets is

$$\begin{cases} (2) \\ (k) \rightsquigarrow (2)(3)\cdots(k)(k+1) \end{cases}$$

and this succession rule corresponds to combinatorial classes enumerated by Catalan numbers.

Figure 8 shows the beginning of the generating tree associated with this ECO construction. To improve the understanding of this tree, we do not represent labellings of ladder posets in its nodes, but rather the minimal permutations with d descents of size $2d$ associated with them.

Proof of Theorem 5 by bijection It is well known that Dyck paths of length $2d$ are enumerated by the Catalan numbers $C_d = \frac{1}{d+1}\binom{2d}{d}$. Let us recall the definition of Dyck paths.

Definition 3. *A Dyck path of length $2d$ is a path in $\mathbb{N} \times \mathbb{N}$ starting at $(0, 0)$ and ending at $(2d, 0)$, with steps going up (of coordinate $(1, 1)$) and steps going down (of coordinate $(1, -1)$).*

As it is a path in $\mathbb{N} \times \mathbb{N}$, a Dyck path never goes under the x-axis. We can also notice that a Dyck path has as many steps going up as those going down, and that any prefix of a Dyck path contains at least as many steps going up as those going down. This is actually a characterization of Dyck paths.

We provide a bijection between Dyck paths of length $2d$ and authorized labellings of the ladder poset with d steps with the integers $\{1, 2, \ldots, 2d\}$. The bijection is simple. Starting from a Dyck path $\mathcal{D}$ of length $2d$, we number its steps with the integers from 1 to $2d$, from left to right. Then, we label the lower line of the ladder with the numbers of the steps of $\mathcal{D}$ going up and its upper line with the numbers of the steps of $\mathcal{D}$ going down. An example is shown in Figure 9.
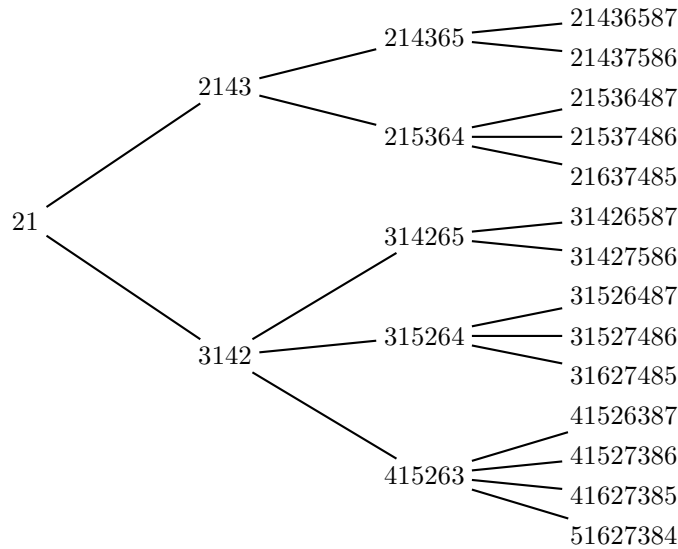
Figure 8: The first four levels of the generating tree associated with the ECO construction of authorized labellings of the ladder posets described above
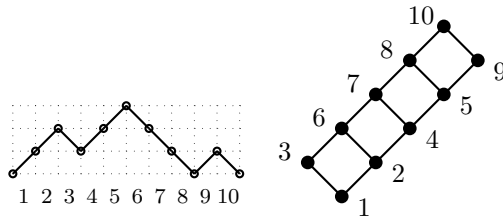

Figure 9: An example of the bijection between minimal permutations with d descents of size $2d$ (seen as authorized labellings of the ladder poset with d steps) and Dyck paths with $2d$ steps

The application we described is actually a bijection between Dyck paths and the authorized labellings of the ladder posets, corresponding to the permutations we are interested in. The reason is simple. It is sufficient to notice that a labelling of the ladder poset with d steps is authorized if and only if any i-th element x on the upper line has at least i smaller elements on the lower line (the element y on the lower line that is linked to x by a step on the ladder, and all the elements below y). See Figure 10 for a better understanding of this statement. In the same way, a path with d steps going up and d step going down is a Dyck path if and only if any i-th step going down has at least i steps going up before it.

4   Enumeration of minimal permutations with d descents and of size $d + 2$

In Section 3, we enumerated the minimal permutations with d descents and of size $2d$, that is to say of *maximal* possible size. We have already proved that the
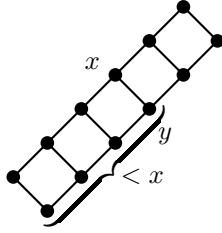
Figure 10: A condition for a labelling of the ladder poset to be authorized

minimal possible size for a minimal permutation with d descents is $d + 1$ and shown that there is only one such permutation, namely the reversed identity $(d+1)d \ldots 21$. In this section, we will focus on the minimal permutations with d descents and of size $d + 2$, *i.e.* the minimal non-trivial case, and give a closed formula for their enumeration through Theorem 6.

Theorem 6. *The minimal permutations with d descents and of size $d + 2$ are enumerated by the sequence (s_d) defined as follows: $s_d = 2^{d+2} - (d+1)(d+2) - 2$.*

We provide two possible proofs for Theorem 6. Both of them are based on the poset representation of minimal permutations with d descents and of size $d + 2$, that consequently have a unique ascent. The first one is straightforward with this decomposition, but implies rather complex computations. The second proof is more complicated but it does not involve such technicalities: it consists in a correspondance between non-interval subsets of $\{1, 2, \ldots, d+1\}$ and minimal permutations with d descents of size $d + 2$, each non-interval subset being associated with exactly two distinct permutations.

Proof of Theorem 6 by a computational method Let us recall that a minimal permutation σ with d descents and of size $d + 2$ has a unique ascent, between two sequences of descents, and that the elements surrounding the ascent are organized in a diamond in the poset representation of the permutation.

Let us denote by i and k the elements of the ascent, $i < k$, by j the element preceeding i in σ, and by h the element following k. In the permutation σ, the subsequence $jikh$ forms an occurence of either the pattern 2143 (if $j < h$) or the pattern 3142 (if $j > h$). This defines two types of minimal permutations with d descents of size $d + 2$. We denote by N_1 the number of those permutations for which $j < h$ and by N_2 the number of those having $j > h$.

We first compute N_1. In order to characterize a minimal permutation σ with d descents, of size $d + 2$, and having its diamond of the type 2143, you first need to establish the values of j, i, k and h satisfying the constraints $1 \leq i < j < h < k \leq d+2$. Then (see left part of Figure 11), the elements greater than h (except k) are necessarily placed before j, in decreasing order, forming the sequence of descents B. Similarily, the elements smaller than j (except i) have to come after h in σ, again in decreasing order, to form the sequence of descents A. The set C of elements between j and h must be partitioned into two parts C_1 and C_2, possibly empty, the elements of C_1 being placed in decreasing order between B and j, those of C_2 between h and A. There are $2^{card(C)} = 2^{h-j-1}$ such partitions of C into $C_1 \uplus C_2$.

14

To sum up, a minimal permutation with d descents, of size $d+2$, and having its diamond of the type 2143 is determined by the values of its i, j, h, and k, with $1 \le i < j < h < k \le d+2$, and a partition of the set C of elements between j and h into $C_1 \uplus C_2$. This characterization allows us to compute N_1:

$$
\begin{aligned}
N_1 &= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \sum_{h=j+1}^{d+1} \sum_{k=h+1}^{d+2} 2^{h-j-1} \\
&= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \sum_{h=j+1}^{d+1} (d+2-h)2^{h-j-1} \\
&= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \sum_{m=0}^{d-j} (d+1-m-j)2^m \\
&= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \left[\sum_{m=0}^{d-j} (d+1-j)2^m - \sum_{m=0}^{d-j} m2^m \right] \\
&= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \left[(d+1-j)(2^{d-j+1}-1) - 2^{d-j+1}(d-j-1) - 2 \right] \\
&= \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} 2^{d-j+2} - (d-j+2) - 1 \\
&= \sum_{i=1}^{d-1} \sum_{n=2}^{d-i+1} 2^n - n - 1 \\
&= \sum_{i=1}^{d-1} 2^{d-i+2} - \frac{(d-i+1)(d-i+2)}{2} - (d-i) - 3 \\
&= \sum_{p=3}^{d+1} 2^p - \frac{p(p-1)}{2} - (p-2) - 3 \\
&= \sum_{p=3}^{d+1} 2^p - \frac{1}{2}p^2 - \frac{1}{2}p - 1 \\
&= 2^{d+2} - \frac{1}{2}\frac{(d+1)(d+2)(2d+3)}{6} - \frac{1}{2}\frac{(d+1)(d+2)}{2} - d - 3 \\
&= 2^{d+2} - \frac{(d+1)(d+2)(d+3)}{6} - d - 3
\end{aligned}
$$

For the minimal permutations σ with d descents and of size $d+2$, whose diamond is of type 3142, the analysis is simpler (this case is illustrated on the right side of Figure 11). Indeed, following the previous notations, to characterize such a permutation, you must again choose i, j, h and k with the constraint that $1 \le i < h < j < k$, but not every such choice is acceptable. Namely, consider the set of elements between h and j. Those elements cannot be before j in σ, since they are smaller than j. But neither can they go after h since they are greater than h. Consequently, there cannot be any element between h and j, and $h = j - 1$. Now, once i, j and k are established, the permutation σ is completly characterized. The elements greater than j (except k) necessarily
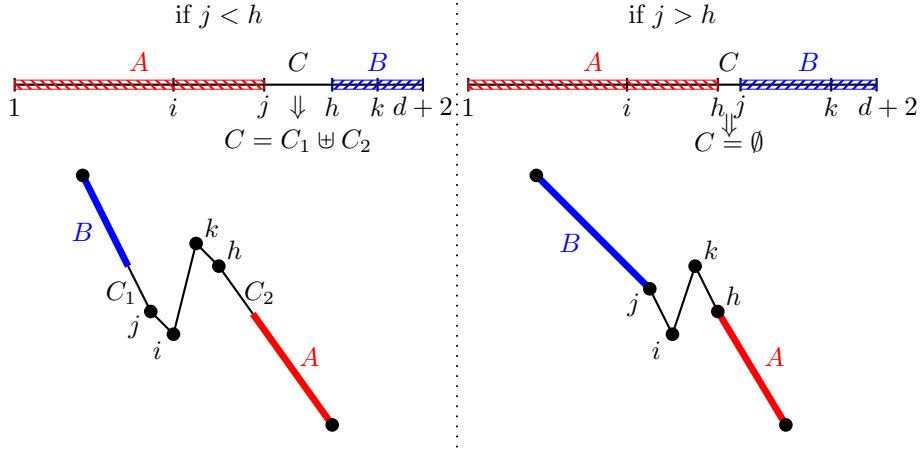
Figure 11: The two types of minimal permutations with d descents and of size $d + 2$, with the decomposition used for their enumeration

form a sequence B of descents before j, and those smaller than $j - 1$ (except i) form a sequence A of descents after $h = j - 1$. The computation of N_2 is then straigthforward:

$$
\begin{aligned}
N_2 &= \sum_{i=1}^{d-1} \sum_{j=i+2}^{d+1} \sum_{k=j+1}^{d+2} 1 \\
&= \sum_{i=1}^{d-1} \sum_{j=i+2}^{d+1} d + 2 - j \\
&= \sum_{i=1}^{d-1} \sum_{m=1}^{d-i} m \\
&= \sum_{i=1}^{d-1} \frac{(d-i)(d-i+1)}{2} \\
&= \sum_{n=1}^{d-1} \frac{n(n+1)}{2} \\[2mm]
&= \frac{1}{2} \left[\frac{d(d-1)(2d-1)}{6} + \frac{d(d+1)}{2} \right] \\
&= \frac{d(d-1)(d+1)}{6}
\end{aligned}
$$

The total number of minimal permutations with d descents of size $d + 2$ is

now simply obtained by the final computation:

$$
\begin{aligned}
N = N_1 + N_2 &= 2^{d+2} - \frac{(d+1)(d+2)(d+3)}{6} - d - 3 + \frac{d(d-1)(d+1)}{6} \\
&= 2^{d+2} - (d+1)^2 - d - 3 \\
&= 2^{d+2} - (d+1)(d+2) - 2
\end{aligned}
$$

This achieves the computational proof of Theorem 6. We now turn to a bijective proof of it.

Proof of Theorem 6 by bijection A *non-interval subset* of $\{1, 2, \ldots, d+1\}$ is a non-empty subset of $\{1, 2, \ldots, d+1\}$ that is not an interval. For example, the non-interval subsets of $\{1, \ldots, 4\}$ are: $\{1, 3\}$, $\{1, 4\}$, $\{2, 4\}$, $\{1, 2, 4\}$ and $\{1, 3, 4\}$. Non-interval subsets of $\{1, 2, \ldots, d+1\}$ are easy to enumerate, as shown in Proposition 4.

Proposition 4. *The number of non-interval subsets of the set* $\{1, 2, \ldots, d+1\}$ *is* $2^{d+1} - \frac{(d+1)(d+2)}{2} - 1$.

Proof. There are 2^{d+1} subsets of $\{1, 2, \ldots, d+1\}$, one being the empty set. So we only need to prove that there are $\frac{(d+1)(d+2)}{2}$ subsets of $\{1, 2, \ldots, d+1\}$ that *are* (non-empty) intervals. It is simple to see that there are i interval subsets of $\{1, 2, \ldots, d+1\}$ whose greatest element is i, namely the intervals $[j..i]$ for $1 \leq j \leq i$. And since $\sum_{i=1}^{d+1} i = \frac{(d+1)(d+2)}{2}$, the proof of Proposition 4 is completed. $\qquad\square$

Notice that the sequence $(2^{d+1} - \frac{(d+1)(d+2)}{2} - 1)_d$ is registered in the Online Encyclopedia of Integer Sequences [11] as [A002662]. To prove Theorem 6, we need to show that there are twice as many minimal permutations with d descents and of size $d+2$ as non-interval subsets of $\{1, 2, \ldots, d+1\}$. For this purpose, we partition the set of minimal permutations with d descents and of size $d+2$ into two subsets S^1 and S^2, and show bijections between S^1 (resp. S^2) and the set of non-interval subsets of $\{1, 2, \ldots, d+1\}$, denoted $\mathcal{NI}$.

The set S^1 contains the minimal permutations σ with d descents and of size $d+2$ such that (1) $d+2$ is the element at the top of the ascent of σ, and (2) the first sequence of descents of σ is not composed of elements that are consecutive. The set S^2 contains all the other minimal permutations with d descents and of size $d+2$. Figure 12 shows the shapes of the permutations in S^1 and in S^2.

We first describe the simple bijection Φ^1 between $\mathcal{NI}$ and S^1. Consider a non-interval subset s of $\{1, 2, \ldots, d+1\}$. Let us denote by w the set of "wholes" associated with s: $w = \{1, 2, \ldots, d+1\} \setminus s$. Now we set $\Phi^1(s)$ to be the permutation consisting of the elements of s in decreasing order, followed by $d+2$ and then by the elements of w in decreasing order. This definition is illustrated in Figure 13.

Proposition 5. *The application* Φ^1 *defines a bijection between* $\mathcal{NI}$ *and* S^1.

Proof. Let s be a non-interval subset of $\{1, 2, \ldots, d+1\}$, and let w be the associated set of wholes $w = \{1, 2, \ldots, d+1\} \setminus s$.

We start by proving that $\Phi^1(s) \in S^1$. Since $s \in \mathcal{NI}$, s contains at least two elements, and w at least one. Consequently, $\Phi^1(s)$ consists of two non-empty
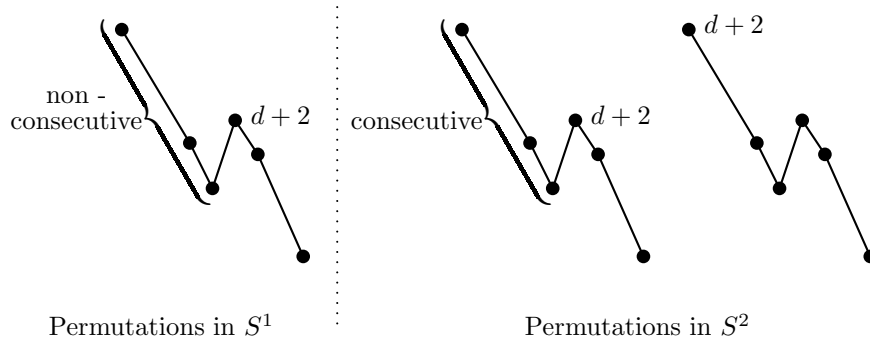
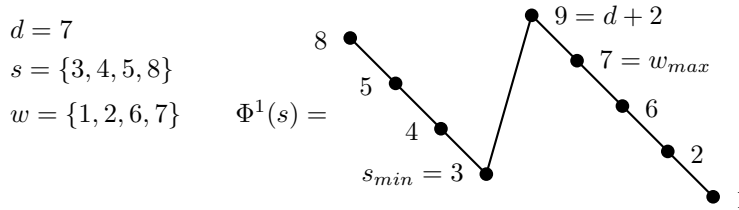Figure 12: The shapes of the permutations in the sets S^1 and S^2

Figure 13: Definition of the bijection Φ^1 on an example

sequences of descents separated by one ascent, and we just need to check the diamond property around its ascent to prove that $\Phi^1(s)$ is a minimal permutation with d descents and of size $d+2$. In our case, proving this diamond property is the same as showing that the smallest element s_{min} of s is smaller than the bigger element w_{max} of w. Since s is not an interval, there is at least one element of w that is bigger than s_{min}, and consequently $s_{min} < w_{max}$. Finally, considering again that s is not an interval, we get that $\Phi^1(s) \in S^1$.

Now – given that among the minimal permutations with d descents and of size $d+2$, the permutations of S^1 are defined as those whose elements in the first sequence of descents do not form an interval – it should now be clear that Φ^1 is a bijection between $\mathcal{NI}$ and S^1. $\qquad\square$

The bijection between $\mathcal{NI}$ and S^2 is less simple, and we will need to classify the permutations of S^2 by dividing them into types, from A to E. Those types are illustrated in Figure 14.

The permutations σ of type A are those of S^2 such that (1) $d+2$ is the second element of the ascent of σ, and (2) the first sequence of descents of σ contains only two elements, that are consecutive.

The permutations σ of type B are those of S^2 such that (1) $d+2$ is the second element of the ascent of σ, (2) the first sequence of descents of σ is composed of consecutive elements, and contains at least 3 elements, and (3) the second sequence of descents of σ has the form $(d+2)(d+1)r$, with r being either empty or a sequence of consecutive elements in decreasing order and whose smallest element is 1.
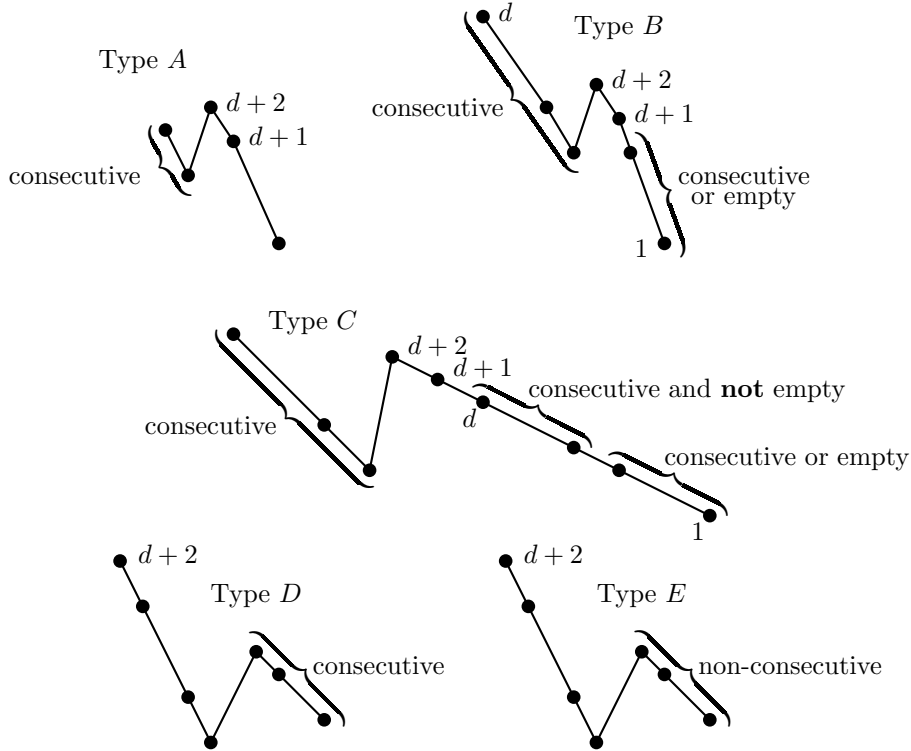
18

Figure 14: The classification of the permutations in S^2 into 5 types A to E

The permutations σ of type C are those of S^2 such that (1) $d+2$ is the second element of the ascent of σ, (2) the first sequence of descents of σ is made of consecutive elements, and contains at least 3 elements, and (3) the second sequence of descents of σ is of the form $(d+2)(d+1)r_1 r_2$ with r_1 being a sequence of consecutive elements in decreasing order and whose greatest element is d, and r_2 being either empty or a sequence of consecutive elements in decreasing order and whose smallest element is 1. Notice that r_1 cannot be empty.

The permutations of type D are those of S^2 such that (1) $d+2$ is the first element of σ, and (2) the elements of the second sequence of descents of σ are consecutive.

The permutations of type E are those of S^2 such that (1) $d+2$ is the first element of σ, and (2) the elements of the second sequence of descents of σ are not consecutive.

Given this classification, it is now easy to prove that:

Proposition 6. *Let σ be a permutation of S^2. Then σ is of one type exactly, among the types A to E.*

Proof. We distinguish two cases, according to the position of $d+2$ in σ: $d+2$ is either the first element of σ or the second element of the ascent of σ. In the first case, it is clear that σ is either of type D or of type E. Let us now assume that $d+2$ is the second element of the ascent of σ. Then, because $\sigma \in S^2$, the

19

elements of the first sequence of descents of σ are necessarily consecutive.

Let us consider the position of $d + 1$ in σ. If it is the first element of σ, and since the elements in the first sequence of descents of σ are consecutive, then the diamond property around the ascent of σ is not satisfied. Indeed, in such a situation, it is impossible for the rightmost element of the diamond to be greater than the lowest one. Consequently, the only possible position for $d + 1$ in σ is just after $d + 2$.

If there are only two elements in the first sequence of descents of σ, then σ is of type A. If there are at least three elements in the first sequence of descents of σ, then it is of type C if $d+1$ is followed by d, of type B otherwise. Because the elements in the first sequence of descents of σ are consecutive, the reader will easily understand that the second sequence of descents of σ is composed of consecutive element for σ of type B, and splits into two sequences of consecutive elements in case σ is of type C. $\qquad\square$

We are now able to define the application Φ^2 from $\mathcal{NI}$ to S^2, and to prove that it is a bijection.

Consider a non-interval subset s of $\{1, 2, \ldots, d+1\}$, and call w the associated set of wholes $w = \{1, 2, \ldots, d+1\} \setminus s$.

1. If w contains only one element x, then necessarily $x \neq 1$ and $x \neq d+1$, or s would be an interval. In this case, we set $\Phi^2(s)$ to the permutation of type A with $x(x-1)$ on its first descent. This permutation obviously satisfies the diamond property (see Figure 15).

$$s = \{1, 2, \ldots, d+1\} \setminus \{x\}$$
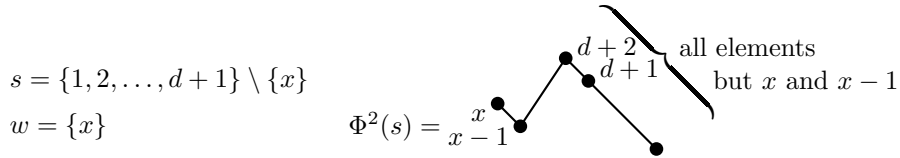$$w = \{x\}$$

Figure 15: Definition of the bijection Φ^2 for s such that $|w| = 1$

If w contains at least two elements, let us denote by n the cardinality of s and by m the cardinality of w increased by 1. Notice that $m \geq 3$ and $n \geq 2$. We will also call w_1 and w_2 the smallest and second smallest elements of w, and s_n and s_{n-1} the greatest and second greatest elements of s. We will associate to s a permutation of S^2 with m elements on its first sequence of descents and n on its second, according to the relative order of w_1, w_2, s_n and s_{n-1}.

Actually, there are few ways to order those 4 elements, since they must satisfy the conditions $w_1 < w_2$, $s_{n-1} < s_n$, and $w_1 < s_n$ (or s would be an interval). Namely there are five possible such orderings.

2. If $w_1 < w_2 < s_{n-1} < s_n$ or $w_1 < s_{n-1} < w_2 < s_n$, then $\Phi^2(s)$ is the permutation of type E obtained as follows: we start from $d+2$, then write the elements of w in decreasing order, and finally the elements of s in

decreasing order. Because of the conditions satisfied by w_1, w_2, s_n and s_{n-1}, this permutation satisfies the diamond property (see Figure 16).

$s = \{s_1, s_2, \ldots, s_n\}$ with $s_1 < s_2 < \ldots < s_n$

$w = \{w_1, w_2, \ldots, w_{m-1}\}$ with $w_1 < w_2 < \ldots < w_{m-1}$

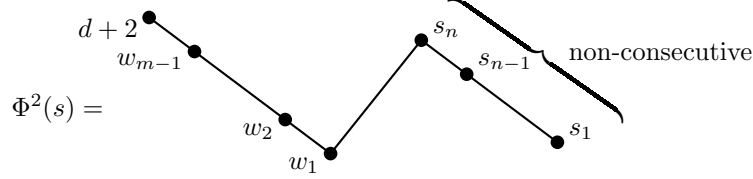$w_1 < w_2 < s_{n-1} < s_n$ or $w_1 < s_{n-1} < w_2 < s_n$

Figure 16: Definition of the bijection Φ^2 for s such that $w_1 < w_2 < s_{n-1} < s_n$ or $w_1 < s_{n-1} < w_2 < s_n$

3. If $s_{n-1} < w_1 < w_2 < s_n$, then the non-interval subset s is completly determined by knowing the cardinality n of s and the grestest element s_n of s. Indeed, it is necessary that $s = \{1, 2, \ldots, n-1\} \uplus \{s_n\}$ to satisfy the condition $s_{n-1} < w_1 < w_2 < s_n$. In this case, we associate to s a permutation of type D as follows. The first element of $\Phi^2(s)$ is $d+2$, the second sequence of descents of $\Phi^2(s)$ is made of n consecutive elements in decreasing order, the greatest of which is s_n, and the remaining elements are placed after $d+2$ in decreasing order to complete the first sequence of descents of $\Phi^2(s)$. To prove that this permutation is of type D, we must check that it belongs to S^2, that is to say that it satisfies the diamond property. It is simple to see that s_n has at least $n+1$ elements smaller than itself: the remaining $n-1$ elements of s, w_1 and w_2. Consequently, 1 and 2 cannot be in the second sequence of descents of $\Phi^2(s)$. Therefore, the first sequence of descents of $\Phi^2(s)$ ends with 2 1, and this is enough to prove the diamond property (see Figure 17).

4. If $w_1 < s_{n-1} < s_n < w_2$, the elements of $\{1, 2, \ldots, d+1\}$ are partitioned into $s \uplus w$ in the following way : $s = \{1, \ldots, w_1 - 1\} \uplus \{w_1 + 1, \ldots, n+1\}$ and $w = \{w_1\} \uplus \{w_2 = n+2, \ldots, d+1\}$. The non-interval s is then completly determined by knowing the cardinality n of s and the number $p = n + 1 - w_1$ of elements of s between w_1 and w_2. Let us notice that $p \geq 2$ (since s_{n-1} and s_n are between w_1 and w_2) and $p \leq n-1$ ($p = n$ would imply that s is an interval). In this case, we associate to s the permutation $\Phi^2(s)$ of type C as follows. The second sequence of descents of $\Phi^2(s)$ splits into two parts (the second one possibly empty). The first part contains $p+1$ elements (we can check that $3 \leq p+1 \leq n$) that are consecutive, and whose greatest element is $d+2$, of course written in decreasing order. The second part is composed of $n - p - 1$ consecutive elements in decreasing order, with 1 as minimal element. This construction leaves m consecutive elements unused so far: written in decreasing order, they will constitute the first sequence of descents of $\Phi^2(s)$. Now, it is easy

$$s_{n-1} < w_1 < w_2 < s_n$$

$$s = \{1, 2, \ldots, n-1\} \uplus \{s_n\}$$

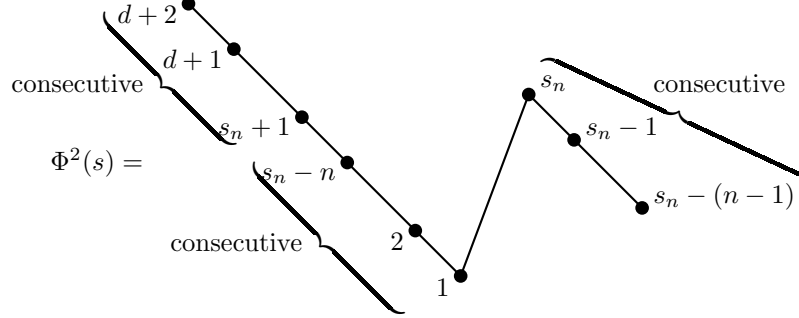$$w = \{n, n+1, \ldots, s_n - 1\} \uplus \{s_n + 1, \ldots, d + 1\}$$

Figure 17: Definition of the bijection Φ^2 for s such that $s_{n-1} < w_1 < w_2 < s_n$

to prove the diamond property, since the second sequence of descents of $\Phi^2(s)$ necessarily starts with $(d+2)(d+1)$. This remark completes the proof that the permutation $\Phi^2(s)$ we just defined is in S^2, and of type C (see Figure 18).

$$w_1 < s_{n-1} < s_n < w_2$$

$$s = \{1, \ldots, w_1 - 1\} \uplus \{w_1 + 1, \ldots, n + 1\}$$

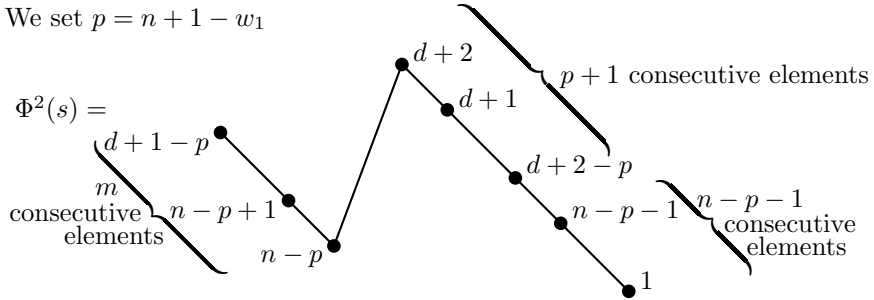$$w = \{w_1\} \uplus \{n + 2, \ldots, d + 1\}$$

We set $p = n + 1 - w_1$

Figure 18: Definition of the bijection Φ^2 for s such that $w_1 < s_{n-1} < s_n < w_2$

5. The last possible relative order of w_1, w_2, s_n and s_{n-1} is $s_{n-1} < w_1 < s_n < w_2$. This case is particularly simple since the cardinality n of s determines s completely. Indeed, it is necessary that $s = \{1, \ldots, n-1\} \uplus \{n+1\}$ to satisfy the conditions $s_{n-1} < w_1 < s_n < w_2$. The permutation $\Phi^2(s)$ is of type B, with the n elements on the second sequence of descents starting with $(d+2)(d+1)$ and then either nothing or consecutive numbers

22

in decreasing order and ending with 1. This leaves m consecutive numbers, with greatest element d, to fill in the first sequence of descents of $\Phi^2(s)$. Because the second sequence of descents starts with $(d+2)(d+1)$, $\Phi^2(s)$ clearly satisfies the diamond property, justifying that $\Phi^2(s)$ is a permutation of S^2 and of type B (see Figure 19).

$$s_{n-1} < w_1 < s_n < w_2$$

$$s = \{1, \ldots, n-1\} \uplus \{n+1\}$$

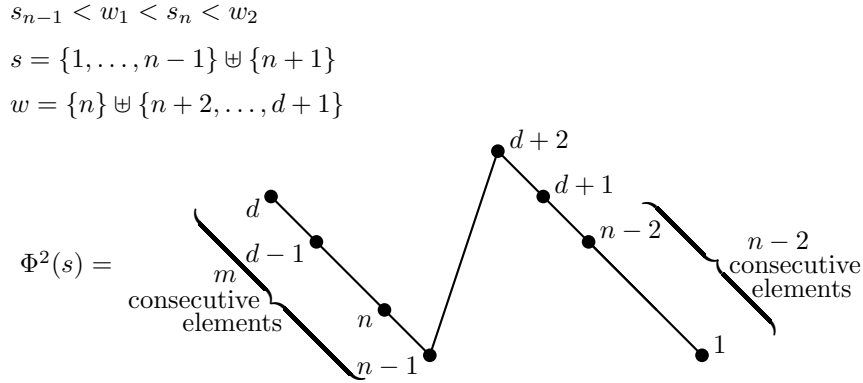$$w = \{n\} \uplus \{n+2, \ldots, d+1\}$$

Figure 19: Definition of the bijection Φ^2 for s such that $s_{n-1} < w_1 < s_n < w_2$

These different cases to define $\Phi^2(s)$ are exemplified in Figure 20.

This ends the definition of the application $\Phi^2 : \mathcal{NI} \to S^2$. Moreover, we have:

Proposition 7. *The application Φ^2 defines a bijection between $\mathcal{NI}$ and S^2.*

Proof. The inverse application of Φ^2, from S^2 to $\mathcal{NI}$, can easily be defined from the previous paragraphs, distinguishing cases according to the type (from A to E) of a permutation of S^2. The details are left to the reader. $\square$

Putting things all together, we have a partition of the set of minimal permutations with d descents and of size $d+2$ into $S^1 \uplus S^2$, and two bijections Φ^1 (resp. Φ^2) between S^1 (resp. S^2) and $\mathcal{NI}$. Combining this with the enumeration of non-interval subsets of $\{1, 2, \ldots, d+1\}$ obtained in Proposition 4, we get another proof of Theorem 6, by a bijective approach.

5 Conclusion and open problems

The goal pursued in this paper is the analysis (characterization, enumeration, . . .) of the permutations that are minimal for the property of having d descents, minimal being intended in the sense of the pattern-involvement relation. For $d = 2^p$, those permutations arise from the whole genome duplication - random loss model, defined in computational biology, where they appear as the excluded patterns defining the pattern-avoiding classes of permutations obtained in at most p steps in this model.

We first provided a local characterization of the minimal permutations with d descents, focusing only on the elements of the permutation surrounding its ascents. This characterization is easy to check: indeed, it provides a linear-time

Case for s	Example of s	$\Phi^2(s)$	Type
(1) with $w = \{x\}$	$s = \{1,2,4,5,6\}$ $w = \{3\}$ $x = 3$		A
(2) with $w_1 < w_2 < s_{n-1} < s_n$	$s = \{1,5,6\}$ $w = \{2,3,4\}$		E
(2) with $w_1 < s_{n-1} < w_2 < s_n$	$s = \{1,3,4,6\}$ $w = \{2,5\}$		E
(3) with $s_{n-1} < w_1 < w_2 < s_n$	$s = \{1,2,5\}$ $w = \{3,4,6\}$ $|s| = 3,\ s_n = 5$		D
(4) with $w_1 < s_{n-1} < s_n < w_2$	$s = \{1,3,4,5\}$ $w = \{2,6\}$ $|s| = 4,\ p = 3$		C
(4) with $w_1 < s_{n-1} < s_n < w_2$	$s = \{1,2,4,5\}$ $w = \{3,6\}$ $|s| = 4,\ p = 2$		C
(5) with $s_{n-1} < w_1 < s_n < w_2$	$s = \{1,2,3,5\}$ $w = \{4,6\}$ $|s| = 4$		B

Figure 20: Definition of $\Phi^2(s)$ for some non-interval subsets s of $\{1, 2, \ldots, 6\}$ ($d = 5$), illustrating all the possible cases in the construction of Φ^2

procedure for deciding whether a permutation is minimal with d descents or not.

The second step of our study was more about enumerating these permutations. We proved that a minimal permutation with d descents has size at least $d + 1$ and at most $2d$. We could not find the enumeration of all minimal permutations with d descents, but we were able to enumerate such permutations of size $d+1$, $d+2$ and $2d$. More precisely, there is only one of size $d+1$ (which is the reversed identity), there are $2^{d+2} - (d+1)(d+2) - 2$ minimal permutations with d descents of size $d+2$, and those of size $2d$ are enumerated by the Catalan numbers.

The enumeration of the minimal permutations with d descents and of size $n \in [(d+3)..(2d-1)]$ remains an open question. For $n = d+3$, we computed the first few terms of the enumerating sequence, and it seems not to appear in the Online Encyclopedia of Integer Sequences [11]. Notice however that the analytical technique used to enumerate the minimal permutations with d descents of size $d + 2$ could theoretically be applied to any other size $n \in [(d + 3)..(2d - 1)]$, but there would be many more cases to consider. Indeed, only for $n = d + 3$, there are more than eighty of them, instead of the two cases for $n = d + 2$. This combinatorial complexity suggests that to solve this enumerating problem, either other techniques or an automated examination of the numerous cases are needed.

References

[1] M.H. Albert, R.E.L. Aldred, M.D. Atkinson, H.P. Van Ditmarsch, C.C. Handley, D.A. Hotlon, and D.J. McCaughan. Compositions of pattern restricted sets of permutations. Technical report, 2004.

[2] M. Atkinson, M. Murphy, and N. Ruskuc. Partially well-ordered closed sets of permutations. *Order*, 2(19):101–113, 2002.

[3] E. Barcucci, A. Del Lungo, E. Pergola, and R. Pinzani. ECO: A methodology for the enumeration of combinatorial objects. *J. Difference Equ. Appl.*, 5:435–490, 1999.

[4] S. Bérard, A. Bergeron, C. Chauve, and C. Paul. Perfect sorting by reversals is not always difficult. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*, 4(1):4–16, 2007.

[5] A. Bernini, L. Ferrari, and R. Pinzani. Enumerating permutations avoiding three Babson-Steingrímsson patterns. *Annals of Combinatorics*, 9:137–162, 2005.

[6] M. Bousquet-Mélou. Four classes of pattern-avoiding permutations under one roof: Generating trees with two labels. *Electr. J. Comb.*, on(2), 2002.

[7] M. Bouvel and D. Rossin. A variant of the tandem duplication - random loss model of genome rearrangement. arXiv:0801.2524v1.

[8] K. Chaudhuri, K. Chen, R. Mihaescu, and S. Rao. On the tandem duplication-random loss model of genome rearrangement. SODA, pages 564 – 570, 2006.

[9] M.C. Chen and R.C.T. Lee. Sorting by transpositions based on the first increasing substring concept. In *BIBE '04: Proceedings of the 4th IEEE Symposium on Bioinformatics and Bioengineering*, page 553, Washington, DC, USA, 2004. IEEE Computer Society.

[10] A. Labarre. New bounds and tractable instances for the transposition distance. *IEEE/ACM Trans. Comput. Biology Bioinform*, 3(4):380–394, 2006.

[11] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences, 2007. published electronically at www.research.att.com/∼njas/sequences/.

[12] V. Vatter. Enumeration schemes for restricted permutations. *Comb. Probab. Comput.*, 17(1):137–159, 2008.