

# Right-jumps & pattern avoiding permutations

Cyril Banderier<sup>1\*</sup> Jean-Luc Baril<sup>2†</sup> Céline Moreira Dos Santos<sup>2‡</sup>

<sup>1</sup> LIPN UMR-CNRS 7030, Université de Paris Nord, France.

<sup>2</sup> LE2I UMR-CNRS 6306, Université de Bourgogne, France.

received 8<sup>th</sup> Dec. 2015, revised 8<sup>th</sup> Feb. 2017, accepted 8<sup>th</sup> Feb. 2017.

---

We study the iteration of the process of moving values to the right in permutations. We prove that the set of permutations obtained in this model after a given number of iterations from the identity is a class of pattern avoiding permutations. We characterize the elements of the basis of this class and enumerate it by giving their bivariate exponential generating function: we achieve this via a catalytic variable, the number of left-to-right maxima. We show that this generating function is a D-finite function satisfying a differential equation of order 2. We give some congruence properties for the coefficients of this generating function, and show that their asymptotics involves a rather unusual algebraic exponent (the golden ratio  $(1 + \sqrt{5})/2$ ) and some closed-form constants. We end by proving a limit law: a forbidden pattern of length  $n$  has typically  $(\ln n)/\sqrt{5}$  left-to-right maxima, with Gaussian fluctuations.

**Keywords:** Permutation pattern, left-to-right maximum, insertion sort, generating function, analytic combinatorics, D-finite function, supercongruence

---

## 1 Introduction

In computer science, many algorithms related to sorting a permutation have been analysed and shown to have behaviours linked to nice combinatorial properties (see e.g. Knuth (1998)). Their complexity can be analysed in terms of memory needed, or number of key operations (like comparisons or pointer swaps). An important family of algorithms, like the so-called insertion algorithms, or in situ permutations, are quite efficient in terms of the number of pointer swaps (but are not the fastest ones in terms of comparisons). Due to this higher cost, they have been much less studied than the faster stack sorting algorithms. Like for the stack algorithms, instead of seeing them as an input/output pair, we can see them like a *process*: input and set of intermediate steps. This opens a full realm of questions on such processes, and they often lead to nice links with other parts of mathematics (like the link between trees, birth and death processes, random walks in probability theory, or permutations and Young tableaux in algebraic combinatorics). Our article will investigate a link between a sorting algorithm, patterns in permutations, and their asymptotics counterparts.

---

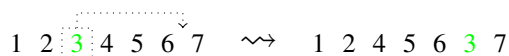
\*<http://lipn.fr/~banderier>

†<http://jl.baril.u-bourgogne.fr/>

‡<http://le2i.cnrs.fr/~Celine-Moreira-Dos-Santos->

Another motivation to analyse such processes comes from the field of bioinformatics. Indeed, in genomics, a crucial part of study is to estimate the similarity of two genomes. This consists in finding the length of a shortest path of evolutionary mutations that transforms one genome into another. Usually, the main operations used in the rearrangement of a genome are of three different types: substitutions (one gene is replaced with another), insertions (a gene is added) and deletions (a gene is removed). For instance, we refer to Jones and Pevzner (2004); Meidanis and Setubal (1997) for an explanation of these operations, and the notions of transposons or jumping genes.

As the problem is too hard in full generality, many simpler mathematical models of the genome are used (see Ewens et al. (1982)): one of them is using permutations of  $\{1, 2, \dots, n\}$  where each gene is assigned a number. Following the idea of transposition mutations (see Jones and Pevzner (2004)), our motivation is to find some combinatorial properties in terms of pattern avoiding permutations whenever one element is deleted and inserted in a position to its right. This operation will be called a right-jump.



**Figure 1:** A right-jump in the permutation  $\sigma = 1234567$ . In this article, we investigate the structure of permutations obtained after several iterations of such right-jumps.

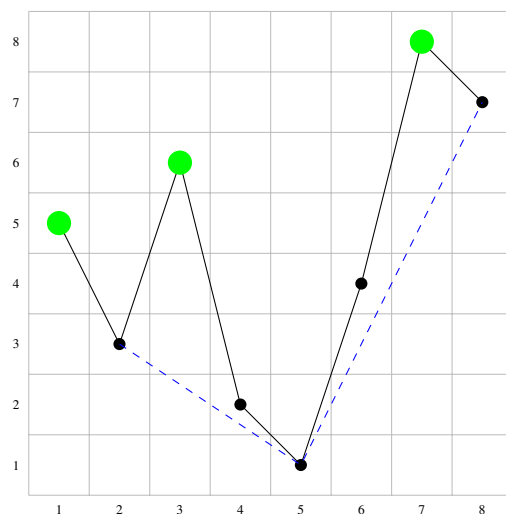
This operation is a variant of genome duplication, which consists of copying a part of the original genome inserted into itself, followed by the loss of one copy of each of the duplicated genes. In particular, it is comparable to the whole duplication-random loss model studied in Chaudhuri et al. (2006). Although there are many connections between these models, it is surprising that the behaviour of their combinatorial properties depends on different parameters: the right-jump model reveals some links with *left-to-right maximum* statistics (see Baril (2013); Bóna (2012)), while the whole duplication-random loss model reveals links with *descent* statistics (see Baril and Vernay (2010); Bouvel and Ferrari (2013); Bouvel and Pergola (2010); Bouvel and Rossin (2009); Chaudhuri et al. (2006); Mansour and Yan (2010)).

In the literature, such right-jumps in permutations are also found in the domain of sorting theory. Indeed, it corresponds (modulo a mirror symmetry) to the *insertion-sorting algorithms* on permutations (see Knuth (1998)). Since the seminal work of Knuth on this subject, many articles related to sorting with a stack exhibit links with pattern avoiding permutations. In contrast, for insertion sorting algorithms (also the subject of a vivid literature), only one study exhibits links with pattern avoiding permutations. Indeed, in his thesis, Magnússon (2013) proves that the set of permutations that can be sorted with one step of the insertion-sorting operator is the class of permutations avoiding the three patterns 321, 312 and 2143.

**Plan of the article.** In Section 2, we recall some basic facts on permutations patterns. In Section 3, we generalize the result of Magnússon by studying the iteration of right-jumps in terms of pattern avoiding permutations: we prove that the set of permutations obtained from the identity after a given number of right-jumps is the class of permutations avoiding some patterns, which we characterize. In Section 4, we enumerate these forbidden patterns by giving their bivariate exponential generating function (involving an additional parameter: the number of left-to-right maxima), and we give the corresponding asymptotics and limit law. We also give some modular congruences for our main enumeration sequence. In Section 5, we conclude with several possible extensions of this work.

## 2 Patterns in permutations

In this section, we give some classical definitions and properties on patterns in permutations. For any permutation  $\sigma \in \mathcal{S}_n$  (the set of permutations of length  $n$ ), the *graphical representation* of  $\sigma = \sigma_1\sigma_2 \dots \sigma_n$  is the set of points in the plane at coordinates  $(i, \sigma_i)$  for  $i \in \llbracket n \rrbracket$ <sup>(i)</sup>. For instance, the permutation 53621487 has the graphical representation illustrated in Figure 2. A *left-to-right maximum* of  $\sigma \in \mathcal{S}_n$  is a value  $\sigma_i$ ,  $1 \leq i \leq n$ , such that  $\sigma_j \leq \sigma_i$  for  $j \leq i$ . A value  $\sigma_i$  of  $\sigma$ ,  $1 \leq i \leq n$  which is not a left-to-right maximum will be called a *non-left-to-right-maximum* of  $\sigma$ . For instance, if  $\sigma = 53621487$  then the left-to-right maxima are 5, 6, 8 and the non-left-to-right-maxima are 1, 2, 3, 4, 7.



**Figure 2:** The graphical representation of  $\sigma = 53621487$ . We show an occurrence of a pattern 213 with a dashed line; big green points are the left-to-right maxima and small black points are non-left-to-right-maxima.

A permutation  $\pi$  of length  $k$ , is a *pattern* of a permutation  $\sigma \in \mathcal{S}_n$  if there is a subsequence of  $\sigma$  which is order-isomorphic to  $\pi$ , *i.e.*, if there is a subsequence  $\sigma_{i_1} \dots \sigma_{i_k}$  of  $\sigma$  with  $1 \leq i_1 < \dots < i_k \leq n$  and such that  $\sigma_{i_j} < \sigma_{i_\ell}$  whenever  $\pi_j < \pi_\ell$ . We write  $\pi \prec \sigma$  to denote that  $\pi$  is a pattern of  $\sigma$ . A permutation  $\sigma$  that does not contain  $\pi$  as a pattern is said to *avoid*  $\pi$ . For example,  $\sigma = 2413$  contains the patterns 231, 132, 213 and 312, but  $\sigma$  avoids the patterns 123 and 321. The set of all permutations avoiding the patterns  $\pi_1, \dots, \pi_m$  is denoted by  $\text{Avoid}(\pi_1, \dots, \pi_m)$ . We say that  $\text{Avoid}(\pi_1, \dots, \pi_m)$  is a *class* of pattern avoiding permutations with *basis*  $\{\pi_1, \dots, \pi_m\}$ . For instance, we refer to the book of Kitaev (2011) and Bóna (2012) to deepen these notions. A set  $\mathcal{C}$  of permutations is *stable* for the involvement relation  $\prec$  if, for any  $\sigma \in \mathcal{C}$ , for any  $\pi \prec \sigma$ , then we also have  $\pi \in \mathcal{C}$ .

Now, we formulate a definition that is crucial for the present study.

<sup>(i)</sup> In this article, we write  $\llbracket n \rrbracket$  for  $\{1, 2, \dots, n\}$ .

**Definition 1** (Permutation basis and basis permutations). *If a set  $\mathcal{C}$  of permutations is stable for the involvement relation  $\prec$ , then  $\mathcal{C}$  is a class of pattern avoiding permutations:  $\mathcal{C} = \text{Avoid}(\mathcal{B})$ . The basis  $\mathcal{B}$  of forbidden patterns is then given by  $\mathcal{B} = \{\sigma \notin \mathcal{C}, \forall \pi \prec \sigma \text{ with } \pi \neq \sigma, \pi \in \mathcal{C}\}$ . In other words, the basis  $\mathcal{B}$  is the set of minimal permutations  $\sigma$  that do not belong to  $\mathcal{C}$ , where minimal is intended in the sense of the pattern-involvement relation  $\prec$  on permutations, that is: if  $\pi \prec \sigma$  and  $\pi \neq \sigma$  then  $\pi \in \mathcal{C}$ . Notice that  $\mathcal{B}$  might be infinite. We call basis permutations the permutations belonging to  $\mathcal{B}$ .*

Equipped with these definitions, our mission consists now in giving a description of the permutations belonging to the basis  $\mathcal{B}_p$  (the permutations which are the minimal forbidden patterns) for the set  $\mathcal{C}_p$  of permutations at distance at most  $p$  from the identity, i.e., permutations obtained from the identity after at most  $p$  right-jumps.

### 3 Iteration of right-jumps in permutations: a structural description of the forbidden patterns

In this section we study the iteration of right-jumps in terms of pattern avoiding permutations. We establish that the set  $\mathcal{C}_p$  of permutations obtained from the identity after at most  $p$  right-jumps is a class of permutations avoiding some patterns that we characterize.

**Lemma 1** (Characterization of the distance). *A permutation obtained from the identity after  $p$  right-jumps contains at most  $p$  non-left-to-right-maxima.*

**Proof:** The result holds for  $p = 1$ ; indeed a right-jump transformation of the identity permutation creates the permutation  $1\ 2\ \dots\ (i-1)\ (i+1)\ \dots\ (j-1)\ i\ j\ \dots\ n$  for  $1 \leq i < j$ , where  $i$  is the only one non-left-to-right-maximum. Now, let us assume that each permutation  $\pi$  obtained from the identity after  $(p-1)$  right-jumps contains at most  $p-1$  non-left-to-right-maxima. Let  $\sigma$  be a permutation obtained from the identity after  $p$  right-jumps. Using the recurrence hypothesis,  $\sigma$  is obtained from a permutation  $\pi$  with at most  $p-1$  non-left-to-right-maxima by moving an element  $\pi_i$ ,  $1 \leq i < n$ , in a position to its right.

We distinguish two cases: (1)  $\pi_i$  is a non-left-to-right-maximum, and (2)  $\pi_i$  is a left-to-right maximum.

Case (1): Since  $\pi_i$  is a non-left-to-right-maximum, there exists  $j < i$  such that  $\pi_j$  is a left-to-right maximum satisfying  $\pi_j > \pi_i$ . Since we move  $\pi_i$  to its right,  $\pi_j$  remains on the left of  $\pi_i$  in  $\sigma$  which implies that  $\pi_i$  is a non-left-to-right-maximum in  $\sigma$ . Using the same argument, any non-left-to-right-maximum  $\pi_k$  in  $\pi$  remains a non-left-to-right-maximum in  $\sigma$ . Moreover, let  $\pi_k$  be a left-to-right maximum in  $\pi$ , i.e.,  $\pi_j < \pi_k$  for all  $j < k$ . Since the right-jump transformation moves to the right of a non-left-to-right-maximum, all values on the left of  $\pi_k$  in  $\sigma$  are lower than  $\pi_k$ , which proves that  $\pi_k$  remains a left-to-right maximum in  $\sigma$ . Therefore,  $\sigma$  contains at most  $p-1$  non-left-to-right-maxima (as  $\pi$  does).

Case (2):  $\pi_i$  is a left-to-right maximum, i.e.,  $\pi_j < \pi_i$  for all  $j < i$ . Since  $\pi_i$  is moved to its right, any left-to-right maximum located on the left of  $\pi_i$  in  $\pi$  remains a left-to-right maximum in  $\sigma$ . On the other hand, any left-to-right maximum located on the right of  $\pi_i$  in  $\pi$  is greater than  $\pi_i$  and thus, it remains a left-to-right maximum in  $\sigma$ . Therefore, the number of left-to-right maxima in  $\sigma$  is at least the number of left-to-right maxima in  $\pi$  minus one (we do not consider  $\pi_i$ ). This means that the number of non-left-to-right-maxima in  $\sigma$  is at most  $p$ .

Considering the two previous cases allows to complete the proof, by induction.  $\square$

We now derive our first enumeration result for the set  $\mathcal{D}_p$  of permutations at distance  $p$  from the identity, *i.e.*, the set of permutations reachable from the identity with  $p$  right-jumps, but that one cannot reach with less than  $p$  right-jumps:  $\mathcal{C}_p = \cup_{k \in \llbracket p \rrbracket} \mathcal{D}_k$ , and this union is disjoint.

**Theorem 1** (Permutations after  $p$  right-jumps). *The set  $\mathcal{D}_p$  of permutations at distance  $p$  from the identity is the set of permutations with exactly  $p$  non-left-to-right-maxima. Accordingly, the number  $d_{n,p}$  of permutations of length  $n$  in  $\mathcal{D}_p$  is counted by the Stirling numbers  $s(n, n-p)$ :*

$$d_{n,p} = s(n, n-p) = \sum_{0 \leq j \leq h \leq p} (-1)^j \binom{h}{j} \binom{n-1+h}{p+h} \binom{n+p}{p-h} \frac{(j-h)^{p+h}}{h!}.$$

**Proof:** After considering Lemma 1, it suffices to prove that any permutation  $\sigma$  with at most  $p$  non-left-to-right-maxima can be obtained from the identity after  $p$  right-jumps. Let  $\sigma$  be a permutation with  $p \geq 1$  non-left-to-right-maxima. Let us assume that the leftmost non-left-to-right-maximum is  $\sigma_i$  and let  $j < i$  be the position of the smallest left-to-right maximum  $\sigma_j$  such that  $\sigma_j > \sigma_i$ . Then we set  $\sigma' = \sigma_1 \dots \sigma_{j-1} \sigma_i \sigma_j \dots \sigma_{i-1} \sigma_{i+1} \dots \sigma_n$ . Since we have  $\sigma_j > \sigma_i$  and also  $\sigma_i > \sigma_{j-1}$  (if  $\sigma_{j-1}$  exists),  $\sigma_i$  becomes a left-to-right maximum in  $\sigma'$ . Thus,  $\sigma'$  contains exactly  $p-1$  non-left-to-right-maxima and by construction,  $\sigma$  can be obtained from  $\sigma'$  by a right-jump. This proves that permutations at distance  $p$  from the identity are exactly the permutations with  $n-p$  left-to-right-maxima, which are known to be counted by  $s(n, n-p)$ , the signless Stirling number of the first kind (see Flajolet and Sedgewick (2009) for the closed-form formula due to Schlämilch, and sequence OEIS A094638 in Sloane and collaborators (2016) for many occurrences of the corresponding triangular array).  $\square$

For instance, the values of  $d_{n,p}$  for  $n = 7$  and  $0 \leq p < 7$  are 1, 21, 175, 735, 1624, 1764, 720.

The following corollary says a little more on the lattice structure associated to our process "a particle jumps to the right".

**Corollary 1** (Changing the starting point and sorting algorithms). *For any permutation  $p$ , one denotes by  $t_p$  its number of non-left-to-right-maxima. Let  $\sigma$  and  $\pi$  be two permutations, then  $t_{\sigma^{-1} \cdot \pi}$  right-jumps are necessary to obtain  $\pi$  from  $\sigma$ . In particular,  $t_{\sigma^{-1}}$  right-jumps are necessary and sufficient to sort by insertion the permutation  $\sigma$  into the identity.*

**Proof:** Firstly,  $t_\sigma$  right-jumps are necessary and sufficient to obtain  $\sigma$  from the identity. Therefore,  $t_{\sigma^{-1} \cdot \pi}$  transformations are sufficient and necessary to obtain  $\sigma^{-1} \cdot \pi$  from the identity. We set  $t = t_{\sigma^{-1} \cdot \pi}$  and let  $Id = \chi_0, \chi_1, \dots, \chi_{t-1}, \chi_t = \sigma^{-1} \cdot \pi$  be a shortest path between the identity and  $\sigma^{-1} \cdot \pi$ . Now, let us prove that if a permutation  $\beta$  is obtained from  $\alpha$  by one right-jump, then for any permutation  $\gamma$ ,  $\gamma \cdot \beta$  is also obtained from  $\gamma \cdot \alpha$  by one right-jump. Indeed, if we have  $\alpha = \alpha_1 \alpha_2 \dots \alpha_n$  then  $\beta$  can be written as  $\beta = \alpha_1 \dots \alpha_{i-1} \alpha_{i+1} \dots \alpha_{j-1} \alpha_i \alpha_j \dots \alpha_n$ . Composing by a permutation  $\gamma$ , we obtain  $\gamma \cdot \alpha = \gamma(\alpha_1) \gamma(\alpha_2) \dots \gamma(\alpha_n)$  and  $\gamma \cdot \beta = \gamma(\alpha_1) \dots \gamma(\alpha_{i-1}) \gamma(\alpha_{i+1}) \dots \gamma(\alpha_{j-1}) \gamma(\alpha_i) \gamma(\alpha_j) \dots \gamma(\alpha_n)$  which proves that  $\gamma \cdot \beta$  is also obtained from  $\gamma \cdot \alpha$  by one right-jump. So if we compose by  $\sigma$  at each step of the above shortest path, then we obtain a shortest path of  $t_{\sigma^{-1} \cdot \pi}$  right-jumps from  $\sigma$  to  $\pi$ , which completes the proof.  $\square$

Since the set  $\mathcal{C}_p$  of permutations obtained after  $p$  right-jumps is stable for the relation  $\prec$ ,  $\mathcal{C}_p$  is also a class  $\text{Avoid}(\mathcal{B}_p)$  of pattern avoiding permutations where  $\mathcal{B}_p$  is the basis consisting of minimal permutations  $\sigma$  that are not in  $\mathcal{C}_p$  (see Definition 1). Theorem 2 gives the explicit description of these basis permutations.

**Theorem 2** (Structural description of the basis permutations). *A permutation  $\sigma \in \mathcal{S}_n$  belongs to the basis  $\mathcal{B}_p$  of forbidden patterns, if and only if the following conditions hold:*

- (i)  $\sigma$  contains exactly  $p + 1$  non-left-to-right-maxima.
- (ii)  $n - 1$  is a non-left-to-right-maximum.
- (iii)  $\sigma_2$  is a non-left-to-right-maximum.
- (iv) For any three left-to-right maxima,  $\sigma_i, \sigma_j$  and  $\sigma_k$  (with  $i < j < k$ ) such that there is no left-to-right maximum between them, there exists a non-left-to-right-maximum  $\sigma_t$  (with  $j < t < k$ ) satisfying  $\sigma_t > \sigma_i$ .

**Proof:** N.B.: Figure 3 on next page illustrates the different claims and notations of this theorem.

Let  $\sigma \in \mathcal{S}_n$  be a permutation belonging to the basis  $\mathcal{B}_p$ , i.e.,  $\sigma \notin \mathcal{C}_p$  and  $\pi \prec \sigma$  implies  $\pi \in \mathcal{C}_p$ . Throughout this proof, we refer to Figure 3 for an illustration of the three conditions (ii), (iii) and (iv).

- First, the deletion of a non-left-to-right-maximum in  $\sigma$  decreases the number of non-left-to-right-maxima by one exactly. Therefore, the minimality of  $\sigma$  implies that  $\sigma$  necessarily contains exactly  $p + 1$  non-left-to-right-maxima, which proves (i).

- For a contradiction, assume that (ii) is not satisfied, i.e.,  $n - 1$  is a left-to-right maximum. Since  $n$  is always a left-to-right maximum,  $n$  is on the right of  $n - 1$  in  $\sigma$ . Thus, the permutation  $\pi$  obtained by deleting  $n$  from  $\sigma$  also contains  $p + 1$  non-left-to-right-maxima (a non-left-to-right-maximum on the right of  $n$  in  $\sigma$  remains a non-left-to-right-maximum on the right of  $n - 1$  in  $\pi$ ). Therefore,  $\pi$  does not belong to  $\mathcal{C}_p$ ; this gives a contradiction with the minimality of  $\sigma$ .

- For a contradiction, assume that (iii) is not satisfied, i.e.,  $\sigma_2$  is a left-to-right maximum and thus,  $\sigma_1$  is smaller than  $\sigma_2$ . Thus, the permutation  $\pi$  obtained by deleting  $\sigma_1$  from  $\sigma$  also contains  $p + 1$  non-left-to-right-maxima. Indeed, a non-left-to-right-maximum  $\sigma_i$  in  $\sigma$  such that  $\sigma_i < \sigma_1$  becomes a non-left-to-right-maximum  $\sigma_i < \sigma_2 - 1$  in  $\pi$ . Moreover, a non-left-to-right-maximum  $\sigma_i$  in  $\sigma$  such that  $\sigma_i > \sigma_1$  (there is  $\ell, 2 \leq \ell < i$ , with  $\sigma_i < \sigma_\ell$ ) becomes a non-left-to-right-maximum  $\sigma_i - 1$  in  $\pi$  with  $\sigma_i - 1 < \sigma_\ell - 1$ . Therefore,  $\pi$  does not belong to  $\mathcal{C}_p$ ; this gives a contradiction with the minimality of  $\sigma$ .

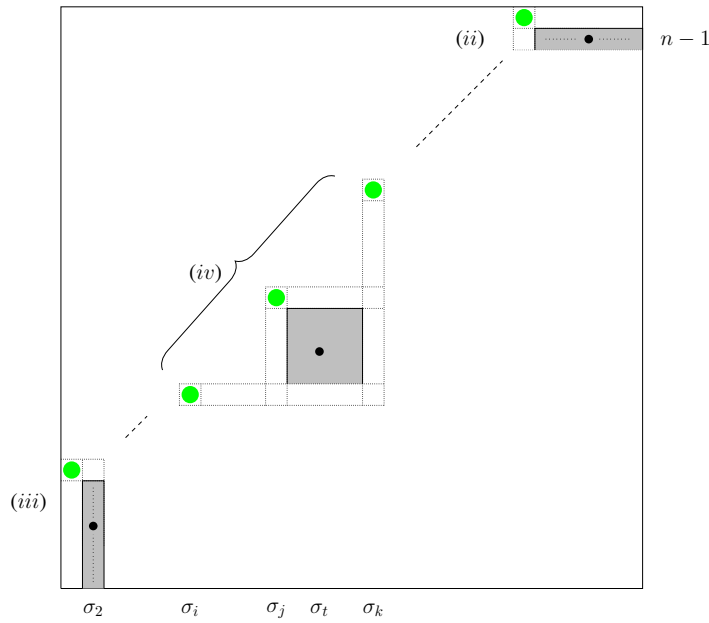
- For a contradiction, assume that (iv) is not satisfied; i.e., there are  $(i, j, k)$ ,  $1 \leq i < j < k \leq n$ , such that  $\sigma_i, \sigma_j$  and  $\sigma_k$  are three consecutive left-to-right maxima of  $\sigma$  (consecutive means that there is no other left-to-right maximum between  $\sigma_i$  and  $\sigma_j$  and between  $\sigma_j$  and  $\sigma_k$ ), and such that there is no non-left-to-right-maximum  $\sigma_\ell$ ,  $j < \ell < k$ , satisfying  $\sigma_i < \sigma_\ell$ . Let  $\pi$  be the permutation obtained from  $\sigma$  by deleting  $\sigma_j$ . It is clear that any non-left-to-right-maximum on the left of  $\sigma_j$  in  $\sigma$  remains a non-left-to-right-maximum in  $\pi$ . Let  $\sigma_\ell, \ell > k$ , be a non-left-to-right-maximum on the right of  $\sigma_k$  in  $\sigma$ . If  $\sigma_\ell < \sigma_j$  then  $\sigma_\ell < \sigma_j \leq \sigma_k - 1$  and  $\sigma_\ell$  remains a non-left-to-right-maximum in  $\pi$ . If  $\sigma_\ell > \sigma_j$  then there is  $\sigma_t, t \geq k$ , such that  $\sigma_\ell \leq \sigma_t \geq \sigma_k$ , and thus, there is  $\sigma_t - 1$  on the left of  $\sigma_k - 1$  in  $\pi$  with  $\sigma_t - 1 > \sigma_k - 1$  which means that  $\sigma_\ell - 1$  is a non-left-to-right-maximum in  $\pi$ . Let  $\sigma_\ell, j < \ell < k$ , be a non-left-to-right-maximum between  $\sigma_j$  and  $\sigma_k$  in  $\sigma$ . Assuming that (iv) is not satisfied, we deduce that  $\sigma_\ell < \sigma_i$ , and  $\sigma_\ell$

remains a non-left-to-right-maximum in  $\pi$ . Finally,  $\pi$  also contains  $p + 1$  non-left-to-right-maxima; this gives a contradiction with the minimality of  $\sigma$ .

Conversely, let  $\sigma$  be a permutation satisfying (i), (ii), (iii) and (iv) and  $\pi$  be a permutation obtained by deleting  $\sigma_i$ ,  $1 \leq i \leq n$ , from  $\sigma$ . Let us prove that  $\pi$  belongs to  $\mathcal{C}_p$ , that is,  $\pi$  contains at most  $p$  non-left-to-right-maxima.

- If  $\sigma_i$  is a non-left-to-right-maximum of  $\sigma$ , then  $\pi$  has  $p$  non-left-to-right-maxima and thus,  $\pi \in \mathcal{C}_p$ .
- Now, let us assume that  $\sigma_i$  is a left-to-right maximum of  $\sigma$ . If  $\sigma_i = n$ , then (ii) implies that  $n - 1$  is a non-left-to-right-maximum of  $\sigma$  and a left-to-right maximum in  $\pi$ ; this implies that  $\pi$  contains  $p$  non-left-to-right-maxima and thus,  $\pi \in \mathcal{C}_p$ . If  $\sigma_i = \sigma_1$ , then (iii) implies that  $\sigma_2$  is a non-left-to-right-maximum of  $\sigma$  and  $\sigma_2 - 1$  is a left-to-right maximum in  $\pi$ ; this implies that  $\pi$  contains  $p$  non-left-to-right-maxima and thus,  $\pi \in \mathcal{C}_p$ . If there exists  $(j, k)$ ,  $1 \leq j < i < k \leq n$ , such that  $\sigma_j$  and  $\sigma_k$  are left-to-right maxima (we choose  $j$  the greatest possible and  $k$  the lowest possible with this property). Then, (iv) implies that there is  $\sigma_\ell$ ,  $i < \ell < k$ , such that  $\sigma_j < \sigma_\ell < \sigma_i$  (we choose the lowest possible  $\ell > i$ ). Thus,  $\sigma_\ell$  is a non-left-to-right-maximum in  $\sigma$  and becomes a left-to-right maximum in  $\pi$ , which implies that  $\pi$  contains exactly  $p$  non-left-to-right-maxima, and thus  $\pi \in \mathcal{C}_p$ .

Finally, the permutation  $\pi$  necessarily belongs to  $\mathcal{C}_p$ , which completes the proof. □



**Figure 3:** An illustration of Theorem 2 that characterizes the basis permutations of  $\mathcal{B}_p$ . Condition (i) states such a basis permutation has  $p + 1$  non-left-to-right maxima (drawn with a small black point, while left-to-right maxima are drawn with a big green point), condition (ii) states that  $n - 1$  is not a left-to-right maximum, condition (iii) states that  $\sigma_2$  is not a left-to-right maximum, and condition (iv) states that there is a "higher" non-left-to-right maximum between 3 left-to-right maxima.

**Corollary 2** (Length of the forbidden patterns). *Permutations in  $\mathcal{B}_p$  have length  $\leq 2(p+1)$  and  $\geq p+2$ . As a consequence,  $\mathcal{B}_p$  is a finite set.*

**Proof:** Theorem 2 implies that the number of left-to-right maxima in a basis permutation is at most the number of non-left-to-right-maxima. Since a basis permutation of  $\mathcal{B}_p$  has  $p+1$  non-left-to-right-maxima, its length is at most  $2(p+1)$ .  $\square$

For instance, the basis for  $p = 0, 1, 2$  are respectively  $\mathcal{B}_0 = \{21\}$ ,  $\mathcal{B}_1 = \{312, 321, 2143\}$  (recovering the result of Magnússon), and  $\mathcal{B}_2 = \{4123, 4132, 4213, 4231, 4312, 4321, 21534, 21543, 31254, 32154, 31524, 31542, 32514, 32541, 214365\}$ .

## 4 Enumerative results for basis permutations

In order to obtain a recursive formula for the number  $b_{n,p}$  of permutations of length  $n$  in the basis  $\mathcal{B}_p$ , we present the following preliminary lemma.

**Lemma 2** (A recursive description). *Let  $\sigma \in \mathcal{S}_n$  be a basis permutation having  $p \geq 1$  non-left-to-right-maxima and such that  $\sigma_{k+1} = n$ ,  $k \geq 0$ . Let  $\alpha$  be the subsequence  $\sigma_1\sigma_2 \dots \sigma_k$  and  $\pi$  be the permutation in  $\mathcal{S}_k$  isomorphic to  $\alpha$ . Then,  $\pi$  is a basis permutation with  $p - n + k + 1$  non-left-to-right-maxima.*

**Proof:** Any permutation  $\sigma$  can be uniquely written as  $\sigma = \alpha n \beta$  where  $\alpha$  and  $\beta$  are two subsequences of  $\llbracket n-1 \rrbracket$ . Let  $k$  be the length of  $\alpha$  and let  $\pi = \pi_1\pi_2 \dots \pi_k$  be the permutation of  $\llbracket k \rrbracket$  that is isomorphic to the subsequence  $\alpha$ . Let us prove that  $\pi$  is minimal.

Since  $\sigma$  is minimal, it satisfies the three conditions (ii), (iii) and (iv) of Theorem 2. Since all elements in  $\beta$  are non-left-to-right-maxima in  $\sigma$ ,  $\pi$  contains exactly  $p - (n-1-k)$  non-left-to-right-maxima and thus,  $n-p-1$  left-to-right maxima.

- The condition (iii) of Theorem 2 on  $\sigma$  does not involve the part  $n\beta$ . Therefore,  $\pi$  satisfies (iii).

- The deletion of  $n\beta$  from  $\sigma$  preserves the condition (iv) on  $\pi$ . Thus,  $\pi$  satisfies (iv).

- Let  $\sigma_i, \sigma_j$  and  $\sigma_{k+1} = n$ ,  $1 \leq i < j \leq k$ , be the last three left-to-right maxima of  $\sigma$ . After the deletion of  $n\beta$ , the two left-to-right maxima of  $\sigma$ ,  $\sigma_i$  and  $\sigma_j$ , are respectively transported in  $\pi$  into  $\pi_i$  and  $\pi_j = k$ . Condition (iv) on  $\sigma$  ensures that there is  $\sigma_\ell$ , between  $\sigma_j$  and  $n$  such that  $\sigma_\ell > \sigma_i$ . The greatest value  $\sigma_\ell$  satisfying this property is then transported in  $\pi$  into  $k-1$ , which proves that  $k-1$  is on the right of  $k$  in  $\pi$ . Thus,  $\pi$  satisfies (ii).

Using Theorem 2, the permutation  $\pi$  is a basis permutation with  $p + \ell - n + 1$  non-left-to-right-maxima.  $\square$

**Theorem 3** (An infinite recursion). *The number  $b_{n,p}$  of basis permutations of length  $n$  in  $\mathcal{B}_p$  (or equivalently having exactly  $p+1$  non-left-to-right-maxima) is given by the following recurrence relation (for  $p < n-2$ ):*

$$b_{n,p} = \sum_{\ell=0}^{p-1} (\ell+1)! \cdot \binom{n-2}{\ell} \cdot b_{n-\ell-2, p-\ell-1}$$

anchored with  $b_{n,p} = 0$  if  $p < (n-2)/2$  or  $p > n-2$ , and  $b_{n,n-2} = (n-1)!$  for  $n > 1$ .



**Proof:** Any permutation  $\sigma$  of length  $n \geq 1$  contains at least one left-to-right maximum and thus, at most  $n - 1$  non-left-to-right-maxima which implies that  $b_{n,p} = 0$  for  $n \leq p + 1$ . Using the proof of Corollary 2, we also have  $b_{n,p} = 0$  for  $n > 2(p + 1)$ . Moreover, the basis permutations of length  $n$  with  $n - 1$  non-left-to-right-maxima are the permutations of the form  $n\alpha$  where  $\alpha \in \mathcal{S}_{n-1}$ . So, we have  $b_{n,n-2} = (n - 1)!$  for  $n > 1$ .

Now, let us prove the recursive relation. Let  $\sigma \in \mathcal{S}_n$  be a basis permutation with  $p + 1$  non-left-to-right-maxima. We consider its unique decomposition  $\sigma = \alpha n \beta$  where  $\alpha$  and  $\beta$  are some subsequences of  $\llbracket n - 1 \rrbracket$ . Let  $\ell + 2, \ell \geq 0$ , be the length of  $n\beta$  and let  $\pi$  be the permutation in  $\mathcal{S}_{n-\ell-2}$  isomorphic to  $\alpha$ . Using Lemma 2 with  $k = n - \ell - 2$ ,  $\pi$  is minimal with  $p - \ell - 1$  non-left-to-right-maxima. So, we can associate to  $\sigma = \alpha n \beta$  the pair  $(\pi, \gamma)$  where  $\pi \in \mathcal{S}_{n-\ell-2}$  is minimal with  $p - \ell - 1$  non-left-to-right-maxima and  $\gamma \in \mathcal{S}_{\ell+1}$  is isomorphic to  $\beta$ .

Conversely, let  $\pi$  be a basis permutation of length  $n - \ell - 2$  with  $p - \ell - 1$  non-left-to-right-maxima and  $\gamma \in \mathcal{S}_{\ell+1}$ . We construct a basis permutation  $\sigma$  of length  $n$  with  $p + 1$  non-left-to-right-maxima as follows. From  $\gamma \in \mathcal{S}_{\ell+1}$ , we construct a subsequence  $\beta$  of  $\llbracket n - 1 \rrbracket$  of length  $\ell + 1$  such that  $\beta$  contains the value  $n - 1$  and such that  $\beta$  is isomorphic to  $\gamma$ . Since  $n - 1$  belongs to  $\beta$ , its position in  $\beta$  also is the position of the greatest value of  $\gamma$ . So,  $\beta$  is characterized by the choice of  $\ell$  values among  $\llbracket n - 2 \rrbracket$ . Now, we define the unique subsequence  $\alpha$  of  $\llbracket n - 2 \rrbracket \setminus X$  isomorphic to  $\pi$  where  $X$  is the set of values used in  $\beta$ . This construction ensures that  $\sigma = \alpha n \beta$  is a basis permutation of length  $n$  with  $p + 1$  non-left-to-right-maxima, and so  $\sigma \in \mathcal{B}_p$ . So, there are  $\binom{n-2}{\ell}$  possibilities to choose the values of  $\beta$  and  $(\ell + 1)!$  possibilities to choose  $\gamma$  and  $b_{n-\ell-2,p-\ell-1}$  possibilities to choose a basis permutation  $\pi \in \mathcal{S}_{n-\ell-2}$  with  $p - \ell - 1$  non-left-to-right-maxima. Varying  $\ell$  from 0 to  $p - 1$ , we obtain the recursive formula.  $\square$

Theorem 3 allows us to find the bivariate exponential generating function for the number of basis permutations according to the number of non-left-to-right-maxima.

**Theorem 4** (Closed-form for the bivariate generating function). *Consider the bivariate exponential generating function  $B(x, y) = \sum_{n \geq 0, p \geq 0} b_{n,p} \frac{x^n y^p}{n!}$  where the coefficient of  $\frac{x^n y^p}{n!}$  is the number  $b_{n,p}$  of basis permutations of length  $n$  in  $\mathcal{B}_p$ . Then, we have*

$$B(x, y) = \frac{1}{2y} \left( 1 - \frac{1}{V} \right) (1 - xy)^{\frac{1}{2}(1+V)} + \frac{1}{2y} \left( 1 + \frac{1}{V} \right) (1 - xy)^{\frac{1}{2}(1-V)} - \frac{1}{y},$$

$$\text{where } V := \sqrt{1 + 4/y}.$$

**Proof:** Setting  $F_p(x) := \sum_{n \geq 0} b_{n,n-p} \frac{x^n}{n!}$  and  $F(x, y) := \sum_{p \geq 0} F_p(x) y^p$ , we have  $B(x, y) = F(xy, 1/y)$ . (We work with the generating function  $F_p(x)$  of the  $(b_{n,n-p})$ 's rather than the generating function of the  $(b_{n,p})$ 's because then the derivation of the proof is simpler to write).

Taking the second derivative of  $F(x, y)$  with respect to  $x$  gives

$$\partial_x^2 F(x, y) = \partial_x^2 \left( \sum_{p \geq 0} F_p(x) y^p \right) = \partial_x^2 F_0(x) + \partial_x^2 F_1(x) y + \partial_x^2 F_2(x) y^2 + \sum_{p \geq 3} \partial_x^2 F_p(x) y^p. \quad (1)$$

Now, the recursive relation of Theorem 3 for  $b_{n+2, n-p+2}$  implies for  $p \geq 3$ :

$$\begin{aligned}\partial_x^2 F_p(x) &= \sum_{n \geq 0} b_{n+2, n-p+2} \frac{x^n}{n!} = \sum_{n \geq 0} \frac{x^n}{n!} \sum_{\ell=0}^{n-p+1} (\ell+1)! \cdot \binom{n}{\ell} \cdot b_{n-\ell, (n-\ell)-(p-1)} \\ &= \sum_{n \geq 0} (n+1)! \frac{x^n}{n!} \cdot \sum_{n \geq 0} b_{n, n-p+1} \frac{x^n}{n!} = \frac{1}{(1-x)^2} F_{p-1}(x).\end{aligned}$$

Plugging this recurrence into the equation (1) (and using  $F_0(x) = F_1(x) = 0$ ) gives:

$$\partial_x^2 F(x, y) = \partial_x^2 F_2(x) y^2 + \sum_{p \geq 2} \frac{y}{(1-x)^2} F_p(x) y^p.$$

It remains to simplify  $F_2(x)$ ; the initial conditions of Theorem 3 ( $b_{n, n-2} = (n-1)!$  and  $b_{n, p} = 0$  for  $n < 2$ ) imply that

$$F_2(x) = \sum_{n \geq 2} b_{n, n-2} \frac{x^n}{n!} = \sum_{n \geq 2} \frac{(n-1)!}{n!} x^n = -\ln(1-x) - x.$$

This leads to the main differential equation:

$$\partial_x^2 F(x, y) = \partial_x^2 F_2(x) y^2 + \frac{y}{(1-x)^2} F(x, y) = \frac{y}{(1-x)^2} (y + F(x, y)). \quad (2)$$

First, by plug & prove, the solutions of  $\partial_x^2 F(x) = yF(x)/(1-x)^2$  are a linear combination of  $(1-x)^\alpha$ , with  $\alpha = (1 + \sqrt{1+4y})/2$ , or  $\alpha = (1 - \sqrt{1+4y})/2$ . Now,  $F(x, y) = -y$  is a trivial particular solution of the non-homogeneous differential equation (2), so the general solution of this differential equation is of the form:

$$F(x, y) = K(y) \cdot (1-x)^{\frac{1}{2}(1+\sqrt{1+4y})} + L(y) \cdot (1-x)^{\frac{1}{2}(1-\sqrt{1+4y})} - y.$$

Since  $F(0, y) = 0$  and  $\frac{\partial F(x, y)}{\partial x} \Big|_{x=0} = 0$ , we respectively deduce the two equations

$$K(y) + L(y) - y = 0 \quad \text{and} \quad -K(y) \cdot (1 + \sqrt{1+4y}) - L(y) \cdot (1 - \sqrt{1+4y}) = 0,$$

thus we obtain

$$K(y) = \frac{y \sqrt{1+4y} - 1}{2 \sqrt{1+4y}} \quad \text{and} \quad L(y) = \frac{y \sqrt{1+4y} + 1}{2 \sqrt{1+4y}}.$$

This gives

$$F(x, y) = \frac{y \sqrt{1+4y} - 1}{2 \sqrt{1+4y}} \cdot (1-x)^{\frac{1}{2}(1+\sqrt{1+4y})} + \frac{y \sqrt{1+4y} + 1}{2 \sqrt{1+4y}} \cdot (1-x)^{\frac{1}{2}(1-\sqrt{1+4y})} - y,$$

and therefore the theorem, as  $B(x, y) = F(xy, 1/y)$ .  $\square$

**Theorem 5** (Asymptotics). *The exponential generating function for the number of basis permutations with respect to their length is given by*

$$\begin{aligned} B(x) &= B(x, 1) = \sum_{n \geq 0} b_n \frac{x^n}{n!} = \frac{\sqrt{5}-1}{2\sqrt{5}} \cdot (1-x)^{\frac{1+\sqrt{5}}{2}} + \frac{\sqrt{5}+1}{2\sqrt{5}} \cdot (1-x)^{\frac{1-\sqrt{5}}{2}} - 1 \\ &= \frac{x^2}{2!} + 2\frac{x^3}{3!} + 7\frac{x^4}{4!} + 32\frac{x^5}{5!} + 179\frac{x^6}{6!} + 1182\frac{x^7}{7!} + 8993\frac{x^8}{8!} + O(x^9). \end{aligned}$$

*It is a D-finite transcendental function satisfying the following differential equation*

$$B(x) - (1-x)^2 \partial_x^2 B(x) + 1 = 0, \quad B(0) = B'(0) = 0. \quad (3)$$

*Equivalently, its coefficients  $b_n$  satisfy the recurrence*

$$b_{n+2} = 2nb_{n+1} + (1+n-n^2)b_n, \quad b_0 = b_1 = 0, b_2 = 1, \quad (4)$$

*and the asymptotics are given by*

$$\frac{b_n}{n!} \sim \frac{\phi}{\sqrt{5} \Gamma(\phi-1)} \frac{1}{n^{2-\phi}} (1 + o(1)),$$

where  $\phi$  is the golden ratio  $\phi = (1 + \sqrt{5})/2$ , and  $\Gamma(z) := \int_0^{+\infty} t^{z-1} \exp(-t) dt$  is the Euler gamma function.

Accordingly, a permutation of length  $n$  has a probability asymptotically 0 to be an element of the basis of forbidden patterns, however, this probability is not “very small” as it decays only polynomially:

$$\text{Prob}(s \in \mathcal{S}_n \text{ belongs to } \cup_{p \in \mathbb{N}} \mathcal{B}_p) = \frac{b_n}{n!} \approx 0.499/n^{0.381} (1 + o(1)).$$

**Proof:** Setting  $y = 1$  in the bivariate exponential generating function given in Theorem 4 gives  $B(x)$ . If a function  $B(z) = \sum b_n z^n$  is D-finite (it is satisfying a linear differential equation, with polynomial coefficients in  $z$ ), then its coefficients  $b_n$  are polynomially recursive (in short, P-recursive): they satisfy a linear recurrence, with polynomial coefficients in  $n$ . See e.g. Flajolet and Sedgewick (2009); Stanley (2012) for more on these two equivalent notions. Starting from the building blocks  $(1-x)^a$ , which are D-finite, and then using the closure properties of D-finite functions (by sum and product) gives the differential equation (3) (this is e.g. implemented in the Gfun Maple package, see Salvy and Zimmermann (1994)). The recurrence is obtained by extracting the coefficient of  $x^n$  on both sides of the differential equation. The asymptotics follows from a singularity analysis (see Flajolet and Sedgewick (2009)) on each term of the shape  $(1-x)^a$ , indeed, for any  $a \in \mathbb{R}$  which is not an integer, one has:

$$[x^n](1-x)^a = \frac{1}{\Gamma(-a)n^{1+a}} \left( 1 + \frac{1}{2}a(a+1)\frac{1}{n} + O\left(\frac{1}{n^2}\right) \right).$$

□

Note that  $B(x, y)$  is D-finite in the variable  $x$ :

$$1 + yB(x, y) - (1 - xy)^2 \partial_x^2 B(x, y) = 0 \quad \text{with } B(0, y) = (\partial_x B)(0, y) = 0,$$

but it is not D-finite in the variable  $y$ . This follows from a saddle point analysis on  $B(1, y) = \sum_n \beta_n y^n$ , indeed the asymptotics of  $\beta_n$  involve arbitrarily large  $(\ln n)^d$ , while the asymptotics of a D-finite function can only have a finite sum of such powers of log, see [Flajolet and Sedgewick \(2009\)](#). This argument is thus similar to a proof that  $(1 - y)^{1-y}$  is not D-finite.

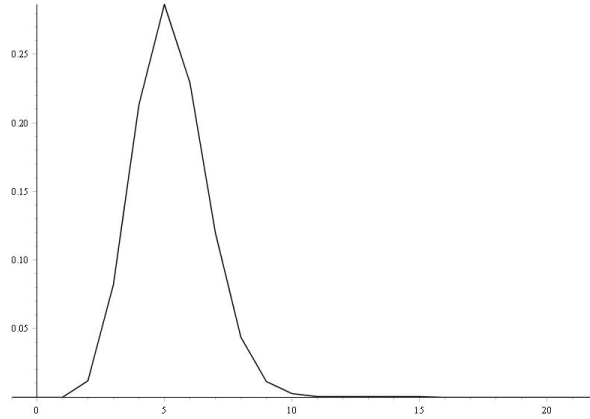
**Remark** [Irrational critical exponent]: It is very seldom that a combinatorial problem leads to some asymptotics involving an irrational number as exponent. In fact, in combinatorics and in statistical physics, most of the asymptotics of integer sequences are of the shape  $b_n \sim Cn^\alpha A^n$ , and the exponent  $\alpha$  which appears there is a key quantity: its value is often the signature of some universal phenomena (in physics, it is called a critical exponent). For D-finite sequences, the theory implies that it is an algebraic number, however, this exponent is very often  $-3/2$ , or a dyadic number (for the reasons explained in [Banderier and Drmota \(2015\)](#)), or a rational number (due to a result on G-functions). Indeed, a theorem (resulting from the works of Katz, André, Chudnovsky & Chudnovsky, see [Chambert-Loir \(2002\)](#)) states that G-functions (D-finite functions with integer coefficients and non-zero radius of convergence) have a rational critical exponent. Now, instead of considering the exponential generating function  $B(x) = \sum b_n x^n / n!$ , we may consider its inverse Borel transform, *i.e.*, the ordinary generating function  $\sum b_n x^n$ . It is also a D-finite function, because D-finite functions are closed by Hadamard product, and therefore the Borel transform (and the inverse Borel transform) of a D-finite function is D-finite (*i.e.*, if the sequence  $b_n$  is P-recursive, so are  $n!b_n$  and  $b_n/n!$ ). We have thus a new D-finite function with integer coefficients and irrational critical exponent (involving the golden ratio  $\phi$ ), but this is not contradicting the G-function theorem, because, due to the multiplication by  $n!$ , we now have a 0 radius of convergence. In conclusion, we have the pleasure to have here one of the few examples in combinatorics of a problem leading to an irrational critical exponent. Other examples are given via the KPZ formula in physics, or via quantities related to quadrees, see [Flajolet and Sedgewick \(2009\)](#).

Pushing further the asymptotics from [Theorem 5](#), we get the following limit law:

**Theorem 6** (Limit Law). *In the model where all permutations of length  $n$  are equidistributed, a random permutation of length  $n$  in  $\cup_{p \in \mathbb{N}} \mathcal{B}_p$  is typically a member of  $\mathcal{B}_p$ , for  $p \sim n - (\ln n) / \sqrt{5}$ , with Gaussian fluctuations. Equivalently, the average number of left-to-right-maxima in a random basis permutation is  $p \sim (\ln n) / \sqrt{5}$  with Gaussian fluctuations.*

**Proof:** This follows from the closed-expression for  $B(x, y)$ , or from a singularity analysis of the differential equation. Indeed, the average and standard deviation follow from the computation of  $\partial_y B(x, y)$  and  $\partial_y^2 B(x, y)$  at  $y = 1$ . The Gaussian limit law follows from the quasi-power theorem applied to a variable exponent perturbation or to our non-confluent differential equation (see [Theorem IX.11](#) and [Theorem IX.18](#) from [Flajolet and Sedgewick \(2009\)](#)).  $\square$

As a random permutation of  $\mathcal{S}_n$  has  $\ln n$  left-to-right maxima on average, the above theorem quantifies to what extent the right-jump process kills the left-to-right maxima when one starts from the identity permutation.



**Figure 4:** This histogram illustrates Theorem 6: the average number of left-to-right-maxima in a random basis permutation is  $p \sim (\ln n)/\sqrt{5}$  with Gaussian fluctuations. However the speed of convergence and the small amplitude of the variance makes that large values of  $n$  have to be considered to recognize clearly the bell curve of the Gaussian (above, this is a histogram for  $n = 4000$ ).

For the combinatorial structure  $\mathcal{B}_p$ , it could be possible that its complementary set has a nicer structure. Those permutations not in the basis are for sure counted by  $u_n = n! - b_n$ ; this sequence satisfies  $u_{n+3} = (n + 1)(n^2 - n - 1)u_n - (3n^2 + 3n - 1)u_{n+1} + 3(n + 1)u_{n+2}$ , which is still a nice recurrence but of order one more than the recurrence for  $b_n$ , so it is a heuristic confirmation that  $b_n$  is a more fundamental sequence than  $u_n$ .

The first values of  $b_n$  (the number of basis permutations of length  $n$ ) are 1, 2, 7, 32, 179, 1182, 8993, 77440 for  $2 \leq n \leq 9$ . We added this sequence to the On-line Encyclopedia of Integer Sequences (hereafter abbreviated OEIS), see Sloane and collaborators (2016):

$p \setminus n$	2	3	4	5	6	7	8	9	10	11	$\#\mathcal{B}_p$
0	1										1
1		2	1								3
2			6	8	1						15
3				24	58	18	1				101
4					120	444	244	32	1		841
5						720	3708	3104	700	50	8232
6							5040	33984	39708	13400	78732
$\Sigma$	1	2	7	32	179	1182	8993	77440	744425	7901410	

**Table 1:** Number  $b_{n,p}$  of basis permutations of length  $n$  (the "minimal forbidden patterns" of  $\mathcal{B}_p$ , or equivalently, with  $p + 1$  non-left-to-right-maxima) where  $2 \leq n \leq 11$  and  $0 \leq p \leq 6$  (OEIS A265163). The last column contains  $\beta_p := \sum_n b_{n,p}$  (OEIS A265164); the last line contains  $b_n := \sum_p b_{n,p}$  (OEIS A265165).

There is a vast literature in number theory analysing the modular congruences of famous sequences

(Pascal triangle, Fibonacci, Catalan, Motzkin, Apéry numbers, see [Deutsch and Sagan \(2006\)](#); [Xin and Xu \(2011\)](#); [Rowland and Zeilberger \(2014\)](#); [Kauers et al. \(2011\)](#)). The properties of  $b_n \bmod m$  are sometimes called "supercongruences" when  $m$  is the power of a prime number: many articles consider  $m = 2^r$ , or  $m = 3^r$ . We now give a result which holds for any  $m$  (not necessarily the power of a prime number).

**Theorem 7** (Supercongruences for D-finite functions).

Consider any  $P$ -recurrence of order  $r$ :

$$P_0(n)u_n = \sum_{i=1}^r P_i(n)u_{n-i}.$$

If the polynomial  $P_0(n)$  is ultimately invertible mod  $m$  (i.e.,  $\gcd(P_0(n), m) = 1$ , for all  $n$  large enough), then the sequence  $(u_n)$  is ultimately periodic<sup>(ii)</sup> mod  $m$ , and there is an algorithm to get this period.

In particular, recurrences such that  $P_0(n) = 1$  are periodic mod  $m$ .

Accordingly, our sequence  $b_n \bmod m$  (defined by recurrence (4)) is periodic for any  $m$ .

**Proof:** Indeed, as the leading term  $P_0$  is invertible, we can write:

$$u_n \bmod m = \sum_{i=1}^r \frac{P_i(n) \bmod m}{P_0(n) \bmod m} (u_{n-i} \bmod m),$$

in which each term has just a finite set of possible values. What is more, for any polynomial  $P(n)$  with integer coefficients,  $P(n) \bmod m$  is of period  $p$ , for some  $p|m$ . (This follows from the fact that the sum and the product is preserving periodicity mod  $m$ , as we did not require in the definition of "period" that  $m$  is the smallest  $m$  such that the sequence is  $m$  periodic). Therefore, one can then construct a Markov chain (an automaton): the states are all the possible  $2r + 1$ -tuples of values mod  $m$  for  $(P_0(n), \dots, P_r(n), u_{n-1}, \dots, u_{n-r})$ , and the recurrence dictates the transitions in this Markov chain. The pigeonhole principle implies that there is a loop in this finite graph, and our period is thus one of the divisors of the length of this loop.

Besides, the smallest period  $p$  and the first integer  $n_1$  satisfying  $u_n = u_{n+p} \bmod m$  for  $n \geq n_1$  are such that  $p + n_1$  is smaller than the number of states in the automaton, i.e. smaller than  $m^{2r+1}$ , or more precisely smaller than  $\phi(m)m^{2r}$  where  $\phi(m)$  is the Euler totient function (the number of invertible elements modulo  $m$ , as  $P_0(n)$  is ultimately invertible).  $\square$

This theorem explains the periodic behaviour of  $b_n \bmod m$ . By brute-force computation, we can get  $b_n \bmod m$ , for any given  $m$ . For example  $b_n \bmod 15$  is periodic of period 12: for  $n \geq 9$ , one has

$$b_n \bmod 15 = (10, 5, 10, 10, 0, 10, 5, 10, 5, 5, 0, 5)^\infty.$$

The period can be quite large, for example  $b_n \bmod 3617$  has period 26158144. We say more on this phenomenon for P-recursive sequences in [Banderier and Luca \(2018\)](#).

---

<sup>(ii)</sup> In the sequel, we will omit the word "ultimately": a periodic sequence of period  $p$  is thus a sequence for which  $u_{n+p} = u_n$  for all large enough  $n$ . Some authors use the terminology "eventually periodic" instead.

## 5 Conclusion

In this article, we analysed the iteration of the process "a particle jumps to the right" in a permutation, and we gave the typical properties of the patterns which are not reached after  $p$  moves. We expect our approach (introducing a catalytic variable and getting a D-finite function) to work in many other cases. However, we already know a nice permutation class for which the basis is not D-finite. Indeed, as an extension of this work, an interesting question is to consider a model in which both right-jumps and left-jumps are allowed: this is a very natural process, also related to sorting algorithms and bioinformatics processes. In a forthcoming work, we show that for this new model, the basis of forbidden minimal patterns for permutations obtained by  $p$  iterations of the process is related to Young Tableaux with 2 equally long first rows (but it is no longer D-finite, unlike the pure right-jump iteration process that we considered in this article).

Another natural question is: is it the case that using e.g. the correspondence between records and cycles in permutations, there is an elegant process corresponding to a "particle jumps to the right", with permutations at distance  $p$  from identity being counted in terms of cycles in the permutation? To get direct "bijective" proofs of our formulae is also an interesting question: as a credo, it cannot be the case that such nice formulae/recurrences are only reached by solving differential equations (like we did in this article). It may be the case that a generating tree approach leads to the simple recurrence (4) we get for  $b_n$  (see e.g. Banderier et al. (2002) or Barucci et al. (1999)). With respect to the asymptotics, it is noteworthy that the process analysed in the present article involves the golden ratio (this is very unusual in combinatorics to have this constant as *critical exponent*): is it the trace of some universality class? (Like it is sometimes the case for problems coming from statistical mechanics, see our remark after Theorem 5.)

Last but not least, we already mentioned a vast literature of publications in number theory analysing the modular congruences of famous sequences (Pascal triangle, Fibonacci, Catalan, Motzkin, Apéry numbers, . . . , Deutsch and Sagan (2006); Xin and Xu (2011)). It seems to us that our approach to tackle them at the level of D-finite functions is new (see also Kauers et al. (2011); Rowland and Zeilberger (2014)), and it would be worth analysing these properties in full generality. In this article, we proved by a mixture of Ansatz and brute force proof that  $b_n \bmod m$  (where  $m$  can be any integer) is a periodic function (of period bounded by a polynomial in  $m$ ). In fact, we prove in Banderier and Luca (2018) that this period and the values of  $b_n \bmod m$  for any given  $m$  can be made explicit.

**Acknowledgments:** We thank our laboratories (LE2I and LIPN) for funding this collaboration, and presentation of preliminary results at the conferences AofA'15 and Permutation Patterns'15. We also thank the two referees for their careful reading.

## References

- Cyril Banderier and Michael Drmota. Formulae and asymptotics for coefficients of algebraic functions. *Combin. Probab. Comput.*, 24(1):1–53, 2015. [[online pdf](#)].
- Cyril Banderier and Florian Luca. On the period mod  $m$  of polynomially-recursive sequences: a case study. *In preparation*, 2018. [[online pdf](#)].
- Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, Danièle Gardy, and Dominique Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Math.*, 246(1-3):29–55, 2002. URL [http://dx.doi.org/10.1016/S0012-365X\(01\)00250-3](http://dx.doi.org/10.1016/S0012-365X(01)00250-3). Formal power series and algebraic combinatorics (Barcelona, 1999).
- Elena Barucci, Alberto Del Lungo, Elisa Pergola, and Renzo Pinzani. ECO: a methodology for the enumeration of combinatorial objects. *J. Differ. Equations Appl.*, 5(4-5):435–490, 1999. URL <http://dx.doi.org/10.1080/10236199908808200>.
- Jean-Luc Baril. Gray code for permutations with a fixed number of left-to-right minima. *Ars Combin.*, 111:225–239, 2013. [[online pdf](#)].
- Jean-Luc Baril and Rémi Vernay. Whole mirror duplication-random loss model and pattern avoiding permutations. *Inform. Process. Lett.*, 110(11):474–480, 2010. URL <http://dx.doi.org/10.1016/j.ipl.2010.04.016>.
- Miklós Bóna. *Combinatorics of permutations*. Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, second (1st ed: 2004) edition, 2012. URL <http://dx.doi.org/10.1201/b12210>.
- Mathilde Bouvel and Luca Ferrari. On the enumeration of  $d$ -minimal permutations. *Discrete Math. Theor. Comput. Sci.*, 15(2):33–48, 2013. [[online pdf](#)].
- Mathilde Bouvel and Elisa Pergola. Posets and permutations in the duplication-loss model: minimal permutations with  $d$  descents. *Theoret. Comput. Sci.*, 411(26-28):2487–2501, 2010. URL <http://dx.doi.org/10.1016/j.tcs.2010.03.008>.
- Mathilde Bouvel and Dominique Rossin. A variant of the tandem duplication-random loss model of genome rearrangement. *Theoret. Comput. Sci.*, 410(8-10):847–858, 2009. URL <http://dx.doi.org/10.1016/j.tcs.2008.11.017>.
- Antoine Chambert-Loir. Théorèmes d’algébricité en géométrie diophantienne (d’après J.-B. Bost, Y. André, D. & G. Chudnovsky). *Astérisque*, pages 175–209, 2002. [[online pdf](#)].
- Kamalika Chaudhuri, Kevin Chen, Radu Mihaescu, and Satish Rao. On the tandem duplication-random loss model of genome rearrangement. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 564–570. ACM, New York, 2006. URL <http://dx.doi.org/10.1145/1109557.1109619>.



- Emeric Deutsch and Bruce E. Sagan. Congruences for Catalan and Motzkin numbers and related sequences. *J. Number Theory*, 117(1):191–215, 2006. [online pdf].
- Warren J. Ewens, Geoffrey A. Watterson, Thomas Eric Hall, and A. Morgan. The chromosome inversion problem. *Journal of Theoretical Biology*, 99:1–7, 1982. URL [http://dx.doi.org/10.1016/0022-5193\(82\)90384-8](http://dx.doi.org/10.1016/0022-5193(82)90384-8).
- Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. [online pdf].
- Neil C. Jones and Pavel A. Pevzner. *An Introduction to Bioinformatics Algorithms*. MIT Press, 2004.
- Manuel Kauers, Christian Krattenthaler, and Thomas W. Müller. A method for determining the mod- $2^k$  behaviour of recursive sequences, with applications to subgroup counting. *Electron. J. Combin.*, 18(2): Paper 37, 83p., 2011. [online pdf].
- Sergey Kitaev. *Patterns in permutations and words*. Monographs in Theoretical Computer Science. An EATCS Series. Springer, Heidelberg, 2011. URL <http://dx.doi.org/10.1007/978-3-642-17333-2>.
- Donald E. Knuth. *The art of computer programming. Vol. 3*. Addison-Wesley, Reading, MA, 1998. Sorting and searching, Second edition (1st ed: 1973).
- Hjalti Magnússon. *Sorting Operators and Their Preimages*. Thesis, Reykjavik University, 2013.
- Toufik Mansour and Sherry H. F. Yan. Minimal permutations with  $d$  descents. *European J. Combin.*, 31(5):1445–1460, 2010. ISSN 0195-6698. URL <http://dx.doi.org/10.1016/j.ejcb.2010.01.001>.
- Carlos Meidanis and João Carlos Setubal. *Introduction to Computational Molecular Biology*. PWS Publishing, 1997.
- Eric Rowland and Doron Zeilberger. A case study in meta-automation: automatic generation of congruence automata for combinatorial sequences. *J. Difference Equ. Appl.*, 20(7):973–988, 2014. URL <http://dx.doi.org/10.1080/10236198.2013.876422>.
- Bruno Salvy and Paul Zimmermann. Gfun: a Maple package for the manipulation of generating and holonomic functions in one variable. *ACM Transactions on Mathematical Software*, 20(2):163–177, 1994. URL <http://dx.doi.org/10.1145/178365.178368>.
- Neil J.A. Sloane and collaborators. On-Line Encyclopedia of Integer Sequences. Published electronically at <https://oeis.org/>, 2016.
- Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second (1st ed: 1997) edition, 2012.
- Guoce Xin and Jing-Feng Xu. A short approach to Catalan numbers modulo  $2^r$ . *Electron. J. Combin.*, 18(1):Paper 177, 12p., 2011. [online pdf].