Generating trees, a method for enumeration and local *(and scaling)* limits

Mathilde Bouvel
(Loria, CNRS, Univ. Lorraine)

Journées Aléa, March 2022

A bit of history

Generating trees were introduced in the nineties independently by

- J. West, for pattern-avoiding permutations;
- the Florentine combinatorics group (R. Pinzani, E. Barcucci, A. Del Lungo, . . .), for a variety of combinatorial objects, including pattern-avoiding permutations.
  (They use the name "ECO method".)

Once combined with the kernel method on functional equations for generating functions (as explained by M. Bousquet-Mélou), it is a general method that can be used to enumerate some families of discrete objects.

In this course, I present this method, illustrated by several examples.
I also discuss how generating trees can be used to establish local and scaling limit results for permutations (recent results of J. Borga).

A toy example:
312-avoiding permutations

Permutations

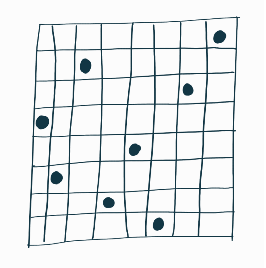
A permutation of size n is a sequence containing exactly once each symbol between 1 and n.

Ex: 5 3 7 2 4 1 6 8 is a permutation of size 8.

Notation: We usually write a permutation $\sigma = \sigma_1 \sigma_2 \ldots \sigma_n$.

We often represent a permutation by its diagram: the $n \times n$ grid which contains a dot in each column i, in row σ_i.
(Rows are numbered from bottom to top.)

Ex: The diagram of our example is



Patterns in permutations

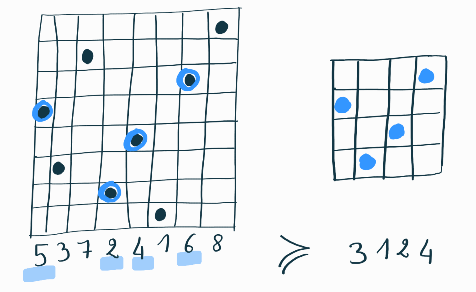
For σ a permutation of size n and π a permutation of size $k \leq n$, we say that σ contains π as a pattern when there exists $i_1 < i_2 < \cdots < i_k$ such that $\sigma_{i_a} < \sigma_{i_b}$ if and only if $\pi_a < \pi_b$.

This is written $\pi \preccurlyeq \sigma$.

The subsequence $\sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k}$ is an occurrence of π.

Ex: $\sigma = 5\,3\,7\,2\,4\,1\,6\,8$ contains the pattern $\pi = 3\,1\,2\,4$, an occurrence being $\sigma_1\sigma_4\sigma_5\sigma_7 = 5\,2\,4\,6$.

We can see patterns and occurrences on the diagrams:

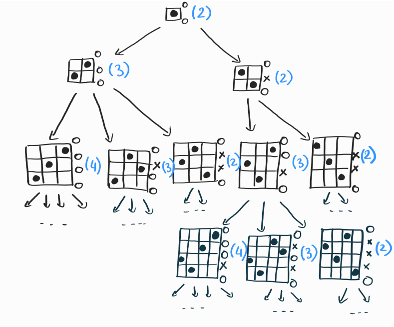


σ avoids π when π has no occurrence in σ.

For B any set of patterns, we denote by $Av(B)$ the set of permutations (of all sizes) avoiding all patterns in B.

Ex: $5\,3\,7\,2\,4\,1\,6\,8 \notin Av(3124)$, but $2\,5\,1\,3\,4\,7\,6\,8 \in Av(321)$.

In this first part, we consider the permutation class $Av(312)$, which we will enumerate using a generating tree.



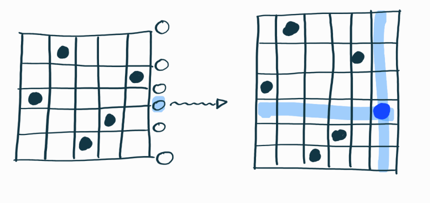
Letting permutations grow on the right

One way of building all permutations of size $n+1$:

- Start from all permutations σ of size n
- For each such σ, append to σ a new final value $a \in \{1, 2, \ldots, n+1\}$, adding 1 to any σ_i such that $\sigma_i \geq a$.

Ex: Appending 3 to 35124 gives 461253

On diagrams:



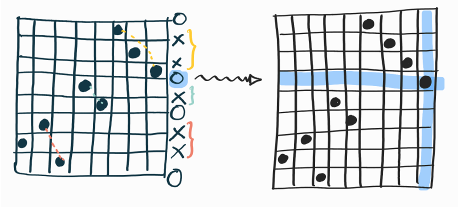
Restriction to $Av(312)$

To build all permutations of size $n+1$ in $Av(312)$, we can

- Start from all permutations of size n avoiding 312
- For each such σ, append a new final value as before, in all possible "places" which do not create an occurrence of 312.

Such places are called active sites ($\circ$), the others are inactive sites ($\times$).

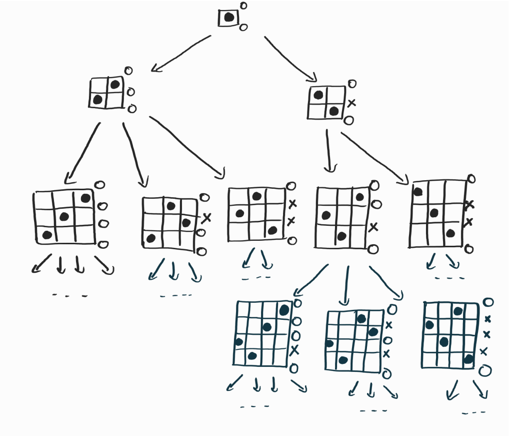
Ex:



Remark: For every family $Av(B)$ defined by the avoidance of (classical) patterns, it is possible to build permutations appending a new final value.

It is the infinite tree

- whose root is $\boxed{\bullet}$ (the permutation of size 1),
- and where the children of any permutation σ are the permutations obtained appending a new final value to σ, in all possible ways which do not create a pattern 312.



Remark: The nodes at level n are the 312-avoiding permutations of size n.

Generating tree for a class $\mathcal{C}$ of discrete objects

In general, a generating tree for a combinatorial class $\mathcal{C}$ is

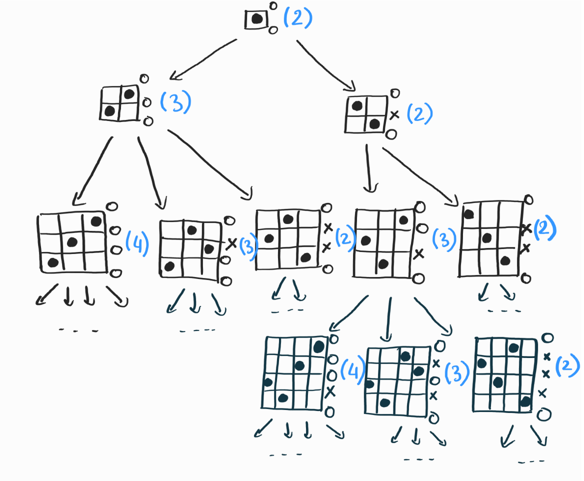
- an infinite tree,
- whose nodes are the elements of $\mathcal{C}$, each occurring exactly once,
- whose root is the element of size 1 in $\mathcal{C}$ (assumed to exist and be unique),
- and where the children of any node c are obtained from c by performing local expansions according to some prescribed rules.

These rules must be carefully chosen to ensure that every element of $\mathcal{C}$ appears, and does not appear multiple times.

Remarks:

- Objects of size n are at level n in the tree. Hence enumerating $\mathcal{C}$ amounts to counting the number of nodes at each level.
- There may be several generating trees for $\mathcal{C}$, depending on the "local expansion rule" which is chosen.

To each 312-avoiding permutation, assign a label: its number of active sites.

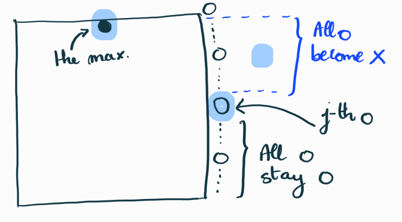


Conjecture:

If a permutation has label k, then its k children have labels $2, 3 \ldots, k+1$.

Proving the conjecture (active sites of $Av(312)$)

- Observe that the bottommost and topmost sites are always active.
- Observe that a site cannot become active it if was previously inactive.
- Number the active sites 1 to k from bottom to top.
- When inserting in the j-th active site for $j \neq k$, all active sites above it become inactive, except the topmost one.



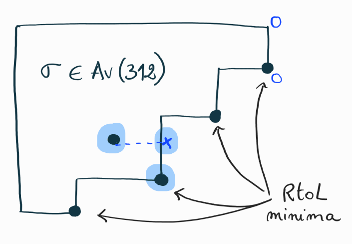
So insertion in active site j produces a permutation with label $j + 1$.

- Insertion in the topmost site produces a permutation with label $k + 1$.

Prop.: If σ has label k, then its children have labels $2, 3 \ldots, k + 1$. □

Detour: other characterization of the active sites/labels

- The top site is always active.
- A site immediately below a RtoL-minimum is always active.
- A site immediately below an element which is not a RtoL-minimum is always inactive.

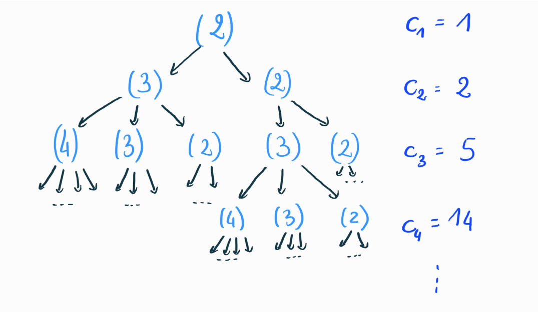


σ_j is a RtoL-minimum of σ if $\sigma_j < \sigma_i$ for all $i > j$, *i.e.*, if, when reading σ from right to left, σ_j is minimal among the elements already read.

Hence, active sites = top site + sites immediately below RtoL-minima, and therefore labels = 1 + number of RtoL-minima.

The simplified generating tree, and the rewriting rule

Keeping only the labels, the generating tree for $Av(312)$ is



$c_n = |Av_n(312)|$ is the number of nodes at level n.

This tree is completely described by the rewriting rule (or succession rule)

$$\Omega_{Cat} = \begin{cases} (2) & \\ (k) & \rightsquigarrow (2), \dots, (k), (k+1). \end{cases}$$

Labels and rewriting rules in general

For a generating tree to be useful in some way, we need to identify

- labels for the objects
- a rewriting rule describing the labels of the children of an object from just the label of that object.

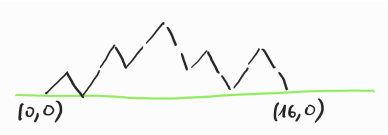
Labels can be integers, pairs of integers, or even tuples of integers of varying size!

Ex: For $Av(1342, \underline{31}42)$, from Dulucq, Gire and Guibert, we can use

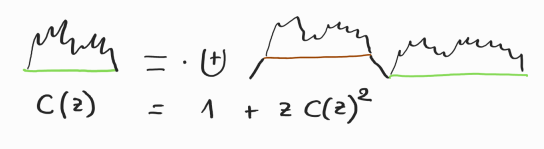
$$\begin{cases} (2,1,(1)) & \\ (x,k,(p_1,\dots,p_k)) & \rightsquigarrow \quad (2+p_j, j, (p_1,\dots,p_{j-1},i)) \\ & \qquad\qquad \text{for } 1 \leq j \leq k \text{ and } p_{j-1} < i \leq p_j \\ & \quad\;\; (x+1, k+1, (p_1,\dots,p_k,i)) \\ & \qquad\qquad \text{for } p_k < i \leq x \end{cases}$$

Enumerating $Av(312)$ from its generating tree: bijective version

A Dyck path of size n is a sequence of up steps $(1, 1)$ and down steps $(1, -1)$ starting at $(0, 0)$, ending at $(2n, 0)$, and staying (weakly) above the horizontal axis.



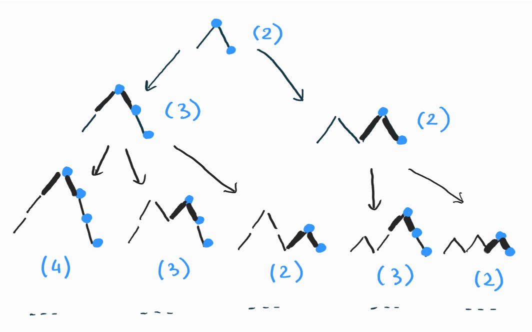
Prop.: Dyck paths are counted by the Catalan numbers: there are $c_n = \frac{1}{n+1}\binom{2n}{n}$ Dyck paths of size n.



A generating tree for Dyck paths

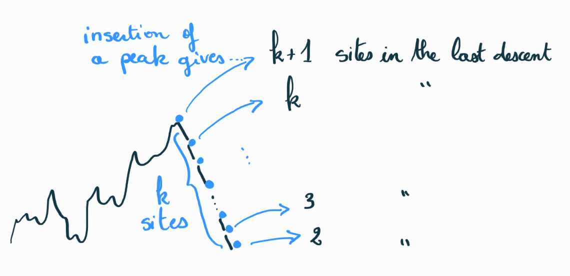
To build all Dyck paths of size $n+1$, we can

- Start from all Dyck paths of size n
- For each of them, insert a new peak ($\bigwedge$) at each point (•) in the last descent (=longest suffix of down steps).



Labels in this generating tree are number of children, or equivalently number of steps in the last descent $+1$.

Associated rewriting rule



Therefore, the propagation of labels in the generating tree for Dyck paths is given by:

$$\Omega_{Cat} = \begin{cases} (2) & \\ (k) & \rightsquigarrow (2), \ldots, (k), (k+1). \end{cases}$$

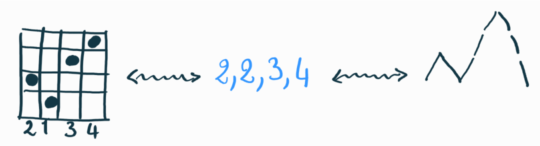
This is the same rule as the one obtained for $Av(312)$.

Isomorphic generating trees for Dyck paths and $Av(312)$

Keeping only the labels, the generating trees for Dyck paths and 312-avoiding permutations are the same.

Prop.: $Av(312)$ is counted by the Catalan numbers.

A node in the tree can be identified by the sequence of labels reaching it from the root. This induces an (implicit) size-preserving bijection between Dyck paths and 312-avoiding permutations.



This bijection records some statistics: it sends the number of RtoL-minima to the number of steps in the last descent (both being the label-1).

Isomorphic generating trees in general

Assume the following:

- A combinatorial class $\mathcal{C}_1$ admits a generating tree, with labels recording the value of the statistics stat_1.
- Same with $\mathcal{C}_2$ and stat_2.
- Both generating trees are described by the same rewriting rule.

Then, it holds that

- The generating trees are isomorphic.
- The isomorphism entails a size-preserving bijection between $\mathcal{C}_1$ and $\mathcal{C}_2$ which preserves statistics: it sends stat_1 to stat_2.

Remark: In general, the bijection is implicit.
I'm only aware of one example where it is turned into an explicit bijection: see the work of Bonichon, Bousquet-Mélou and Fusy on plane bipolar orientations and Baxter permutations.

Enumerating $Av(312)$ from its generating tree: generating function version

From the rewriting rule to a functional equation

Recall the rewriting rule for 312-avoiding permutations:

$$\Omega_{Cat} = \begin{cases} (2) & \\ (k) & \rightsquigarrow (2), \dots, (k), (k+1). \end{cases}$$

Let $c_{n,k}$ be the number of 312-avoiding permutations having size n and label k. Consider the bivariate generating function $C(x; y) = \sum_{n,k} c_{n,k} x^n y^k$.

Remark: $C(x; 1)$ is the generating function of 312-avoiding permutations.

The rewriting rule gives

$$\begin{aligned} C(x; y) &= xy^2 + \sum_{n,k} c_{n,k} x^{n+1} (y^2 + \dots y^k + y^{k+1}) \\ &= xy^2 + \sum_{n,k} c_{n,k} x^{n+1} y^2 \frac{1 - y^k}{1 - y}. \end{aligned}$$

Putting the equation in kernel form

$$\begin{aligned} C(x;y) &= xy^2 + \sum_{n,k} c_{n,k} x^{n+1} y^2 \frac{1-y^k}{1-y} \\ &= xy^2 + \frac{xy^2}{1-y} \left(\sum_{n,k} c_{n,k} x^n - \sum_{n,k} c_{n,k} x^n y^k \right) \\ &= xy^2 + \frac{xy^2}{1-y} \big(C(x;1) - C(x;y) \big) \end{aligned}$$

It follows that $(1-y+xy^2)C(x;y) = xy^2(1-y+C(x;1))$.

The coefficient of $C(x;y)$ is the kernel of this equation.

Remark: Putting $y=1$ in the equation gives no information on $C(x;1)$.

Solving the equation: the kernel method

Method:

- Find a formal power series $Y(x)$ which cancels the kernel.
- Substituting y by $Y(x)$ gives an equation for $C(x;1)$.

Our example:

- The equation is $(1-y+xy^2)C(x;y) = xy^2(1-y+C(x;1))$.
- The formal power series canceling the kernel is $Y(x) = \frac{1-\sqrt{1-4x}}{2x}$.
- Substitution gives $C(x;1) = Y(x) - 1$.
- It follows that there are $c_n = \frac{1}{n+1}\binom{2n}{n}$ 312-avoiding permutations of any size $n \geq 1$.

Remark: For "similar" generating trees with integer labels, the generating functions are always algebraic.
See the "Generating functions for generating trees" paper.

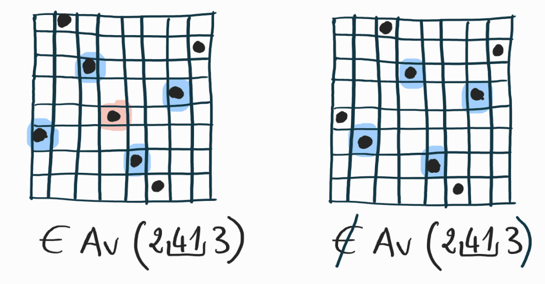
Generating trees where labels are pairs of integers: the case of $Av(2\,\underline{41}\,3)$

The pattern 2 41 3

A permutation $\sigma = \sigma_1\sigma_2\ldots\sigma_n$ contains the pattern 2413 if there exists indices $i < j < j+1 < k$ such that $\sigma_{j+1} < \sigma_i < \sigma_k < \sigma_j$ *i.e.* such that the subsequence $\sigma_i\sigma_j\sigma_{j+1}\sigma_k$ is an occurrence of 2413.

Otherwise σ avoids 2 41 3.

$Av(2\,41\,3)$ denotes the set of all permutations avoiding 2 41 3.

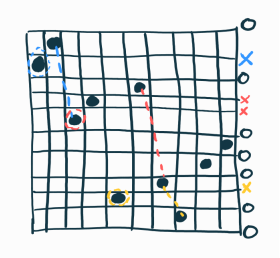


Letting permutations avoiding $2\underline{41}3$ grow on the right

Remark: If $\sigma_1 \dots \sigma_n \sigma_{n+1}$ avoids $2\underline{41}3$, then so does $\sigma_1 \dots \sigma_n$.
(Be careful! Not true for any element removed, *e.g.* 25314!)

Thus, to build all $2\underline{41}3$-avoiding permutations of size $n+1$, we can

- Start from all $2\underline{41}3$-avoiding permutations of size n
- For each such σ, append a new final value in all active sites (= the sites which do not create an occurrence of $2\underline{41}3$).

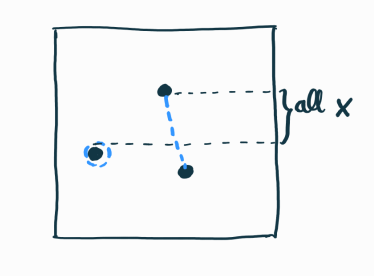


This induces a generating tree for $Av(2\underline{41}3)$.

Non-empty descents are the reason for sites to be inactive

We say that a non-empty descent of σ is an occurrence of the pattern $2\underline{31}$ in σ, *i.e.* a subsequence $\sigma_i \sigma_j \sigma_{j+1}$ (with $i < j$) such that $\sigma_j > \sigma_i > \sigma_{j+1}$.

A site is inactive if and only if it is above the 2 of a non-empty descent.

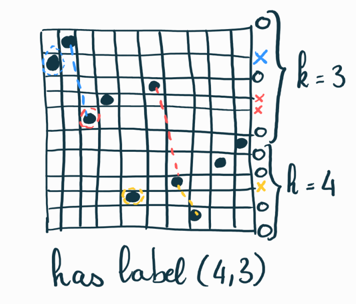


- The non-empty descents determine the active sites.
- Appending a new final value affects the set of non-empty descents of $\sigma = \sigma_1\sigma_2 \ldots \sigma_n$ differently if we insert below or above σ_n.

Thus, we record separately the active sites below and above σ_n.

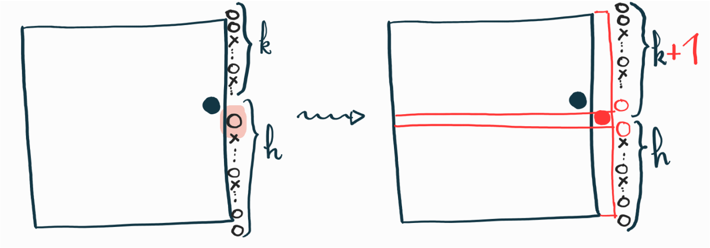
We take the label of σ of size n avoiding 2 41 3 to be (h, k)

- with h = number of active sites below σ_n
- and k = number of active sites above σ_n.



Remark: The site immediately below σ_n is always active.

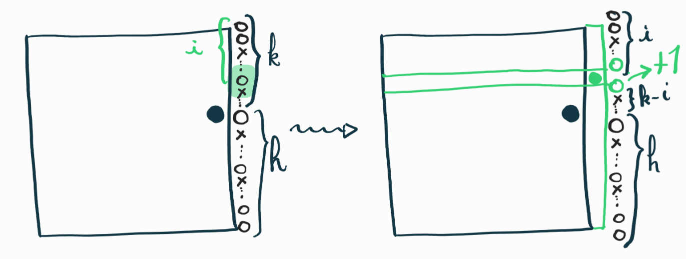
For σ of label (h, k), insertion in the site immediately below σ_n produces an empty descent. Hence, all active sites stay active.



The label of the corresponding child of σ is $(h, k+1)$.

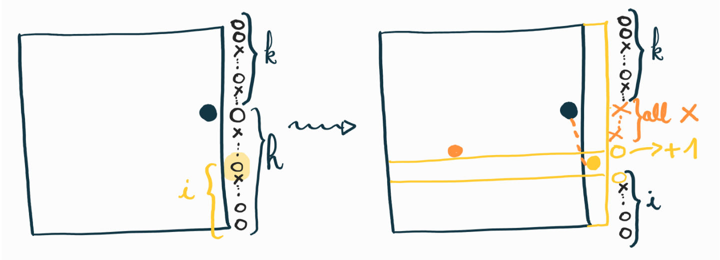
Labels of the children (2/3)

Insertion in an active site above σ_n produces an ascent.
Hence, all active sites stay active.



For insertion in the i-th such active site from the top, the label of the corresponding child of σ is $(h + k - i + 1, i)$.

Insertion in an active site below σ_n (and not immediately below) produces a non-empty descent. Hence, all active sites between σ_n and the new final element σ_{n+1} become inactive (except the site immediately above σ_{n+1}).



For insertion in the i-th such active site from the bottom, the label of the corresponding child of σ is $(i, k+1)$.

Rewriting rule

The generating tree for permutations avoiding $2\underline{41}3$ growing on the right is described by the following rewriting rule:

$$\Omega_{semi} = \begin{cases} (1,1) & \\ (h,k) \rightsquigarrow & (1,k+1),\ldots,(h,k+1) \\ & (h+k,1),\ldots,(h+1,k). \end{cases}$$

Next steps:

- Consider the trivariate generating function

$$S(y,z) = S(x;y,z) = \sum_{n,h,k} s_{n,h,k} x^n y^h z^k,$$

  where $s_{n,h,k}$ is the number of $2\underline{41}3$-avoiding permutations having size n and label (h,k).
- Translate Ω_{semi} into a functional equation for $S(y,z)$.
- Apply the (obstinate) kernel method to solve this equation.

The functional equation

Recall the rewriting rule $\Omega_{semi} = \begin{cases} (1,1) & \\ (h,k) \rightsquigarrow & (1,k+1),\dots,(h,k+1) \\ & (h+k,1),\dots,(h+1,k). \end{cases}$

Therefore,

$$\begin{aligned} S(y,z) &= xyz + \sum_{n,h,k\geq 1} s_{n,h,k} x^{n+1} \Big((y + y^2 + \dots + y^h) z^{k+1} \\ &\qquad\qquad + (y^{h+k} z + y^{h+k-1} z^2 + \dots + y^{h+1} z^k) \Big) \\ &= xyz + \sum_{n,h,k\geq 1} s_{n,h,k} x^{n+1} \left(\frac{1-y^h}{1-y} y\, z^{k+1} + \frac{1-\left(\frac{y}{z}\right)^k}{1-\frac{y}{z}} y^{h+1} z^k \right) \\ &= xyz + \frac{xyz}{1-y} \left(S(1,z) - S(y,z) \right) + \frac{xyz}{z-y} \left(S(y,z) - S(y,y) \right). \end{aligned}$$

Kernel form of the equation

We obtained

$$S(y,z) = xyz + \frac{xyz}{1-y}\left(S(1,z) - S(y,z)\right) + \frac{xyz}{z-y}\left(S(y,z) - S(y,y)\right) .$$

In kernel form, and substituting y with $1+a$, this is

$$\begin{aligned} K(a,z)S(1+a,z) = xz(1+a) - \frac{xz(1+a)}{a}S(1,z) \\ - \frac{xz(1+a)}{z-1-a}S(1+a,1+a), \end{aligned}$$

where the kernel is $K(a,z) = 1 - \dfrac{xz(1+a)}{a} - \dfrac{xz(1+a)}{z-1-a}$.

Notation: We denote with $R(x,a,z,S(1,z),S(1+a,1+a))$ the right-hand side of the equation in kernel form.

Canceling the kernel

Recall that the kernel is $K(a,z) = 1 - \frac{xz(1+a)}{a} - \frac{xz(1+a)}{z-1-a}$.

Solving for z the (quadratic) equation $K(a,z) = 0$ gives two solutions:

$$Z_+(a) = \frac{1}{2}\frac{a+x+ax-Q}{x(1+a)} = (1+a) + (1+a)^2 x + O(x^2),$$
$$Z_-(a) = \frac{1}{2}\frac{a+x+ax+Q}{x(1+a)} = \frac{a}{(1+a)x} - a - (1+a)^2 x + O(x^2),$$

where $Q = \sqrt{a^2 - 2ax - 6a^2x + x^2 + 2ax^2 + a^2x^2 - 4a^3x}$.

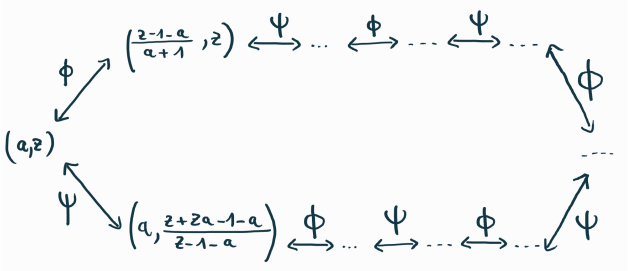
Substituting z for Z_+, we obtain an equation relating the formal power series $Z_+, S(1, Z_+)$ and $S(1+a, 1+a)$, namely:

$$R(x, a, Z_+, S(1, Z_+), S(1+a, 1+a)) = 0.$$

We would like to eliminate $S(1, Z_+)$ in order to find $S(1+a, 1+a)$.

Being obstinate in canceling the kernel

- Look for the transformations leaving the kernel unchanged.
  - Here observe that
    $K(a,z) = K(\frac{z-1-a}{1+a}, z)$ and $K(a,z) = K(a, \frac{z+za-1-a}{z-1-a})$.
  - Therefore, define the involutions
    $\Phi : (a,z) \to \left(\frac{z-1-a}{1+a}, z\right)$ and $\Psi : (a,z) \to \left(a, \frac{z+za-1-a}{z-1-a}\right)$.
- Examine the group generated by Φ and Ψ.
  - Here, they generate a group of order 10.



Being obstinate in canceling the kernel, continued

- Substituting z for Z_+, each element $(f_1(a,z), f_2(a,z))$ in this group cancels the kernel.
- Find the pairs $(f_1(a,z), f_2(a,z))$ such that $f_1(a,Z_+)$ and $f_2(a,Z_+)$ are formal power series in x.
  - Here, we obtain the following pairs:
    $[a,z] \overset{\Phi}{\leftrightarrow} \left[\frac{z-1-a}{1+a}, z\right] \overset{\Psi}{\leftrightarrow} \left[\frac{z-1-a}{1+a}, \frac{z-1}{a}\right] \overset{\Phi}{\leftrightarrow} \left[\frac{z-1-a}{az}, \frac{z-1}{a}\right] \overset{\Psi}{\leftrightarrow} \left[\frac{z-1-a}{az}, \frac{1+a}{a}\right].$
- Each such pair $(f_1(a,Z_+), f_2(a,Z_+))$ can be substituted in the kernel equation $K(a,z)S(1+a,z) = R(x,a,z,S(1,z),S(1+a,1+a))$.
  It results in an equation
  - involving only formal power series,
  - and where the kernel is 0.
- Therefore, each pair satisfies
  $R(x, f_1(a,Z_+), f_2(a,Z_+), S(1, f_2(a,Z_+)), S(1+f_1(a,Z_+), 1+f_1(a,Z_+))) = 0.$

Combining kernel equations

We obtain the following system, with 5 equations and 6 unknowns:

$$\begin{cases} 0 = R(x, a, Z_+, S(1, Z_+), S(1+a, 1+a)) \\ 0 = R\left(x, \frac{Z_+-1-a}{1+a}, Z_+, S(1, Z_+), S(1+\frac{Z_+-1-a}{1+a}, 1+\frac{Z_+-1-a}{1+a})\right) \\ 0 = R\left(x, \frac{Z_+-1-a}{1+a}, \frac{Z_+-1}{a}, S(1, \frac{Z_+-1}{a}), S(1+\frac{Z_+-1-a}{1+a}, 1+\frac{Z_+-1-a}{1+a})\right) \\ 0 = R\left(x, \frac{Z_+-1-a}{aZ_+}, \frac{Z_+-1}{a}, S(1, \frac{Z_+-1}{a}), S(1+\frac{Z_+-1-a}{aZ_+}, 1+\frac{Z_+-1-a}{aZ_+})\right) \\ 0 = R\left(x, \frac{Z_+-1-a}{aZ_+}, \frac{1+a}{a}, S(1, \frac{1+a}{a}), S(1+\frac{Z_+-1-a}{aZ_+}, 1+\frac{Z_+-1-a}{aZ_+})\right). \end{cases}$$

We eliminate all unknowns except $S(1+a, 1+a)$ and $S(1, \frac{1+a}{a})$.

A single resulting equation

We usually write $\bar{a} = a^{-1}$. Observe that $\frac{1+a}{a} = 1 + \bar{a}$.

Elimination from the previous system yields

$$S(1+a, 1+a) + \frac{(1+a)^2 x}{a^4} S(1, 1+\bar{a}) = P(a, Z_+),$$

where $P(a,z) = (z-1-a)(-za^4 + z^2a^4 - za^3 + z^2a^3 - z^3a^2 - 2a^2 + z^2a^2 + za^2 - 4a + 5az - 3az^2 + z^3a + 3z - z^2 - 2)/(za^4(z-1))$.

In this equation, we can separate powers of a:

- $S(1+a, 1+a)$ involves only powers of a that are ≥ 0.
- $\frac{(1+a)^2 x}{a^4} S(1, 1+\bar{a})$ involves only powers of a that are ≤ -2.

Therefore, $S(1+a, 1+a) = \Omega_{\geq}[P(a, Z_+)]$, where for

$G(x;a) = \sum_{n \geq 0} \sum_{i \in \mathbb{Z}} g_{n,i} x^n a^i$, we define $\Omega_{\geq}[G(x;a)] = \sum_{n \geq 0} \sum_{i \geq 0} g_{n,i} x^n a^i$.

Who is $P(a, Z_+)$?

Recall that Z_+ is the unique formal power series canceling the kernel $K(a,z) = 1 - \frac{xz(1+a)}{a} - \frac{xz(1+a)}{z-1-a}$.

Therefore $W = Z_+ - (1+a)$ is the unique formal power series solution of

$$W = x\bar{a}(1+a)(W+1+a)(W+a).$$

Then, $P(a, Z_+)$ can be expressed from W as $P(a, Z_+) = F(a, W)$ for

$$F(a,W) = (1+a)^2\, x + \left(\frac{1}{a^5} + \frac{1}{a^4} + 2 + 2a\right) x\, W \\ + \left(-\frac{1}{a^5} - \frac{1}{a^4} + \frac{1}{a^3} - \frac{1}{a^2} - \frac{1}{a} + 1\right) x\, W^2 + \left(\frac{1}{a^4} - \frac{1}{a^2}\right) x\, W^3.$$

This gives a more direct definition of $\Omega_{\geq}[P(a, Z_+)] = S(1+a, 1+a)$.

What about the number of $2\underbrace{41}3$-avoiding permutations?

- There are $a_n = [x^n]S(1,1)$ $2\underbrace{41}3$-avoiding permutations of size n.
- It holds that $[x^n]S(1,1) = [x^n a^0]S(1+a, 1+a)$.
- Since $S(1+a,1+a) = \Omega_{\geq}[F(a,W)]$, it follows that

$$[x^n]S(1,1) = [x^n a^0]S(1+a,1+a) = [x^n a^0]F(a,W).$$

- From the equation for W and the expression of $F(a,W)$, Lagrange inversion gives an (ugly) summation formula for $[x^n]S(1,1)$.
- The method of creative telescoping of Zeilberger produces a nice recursive formula for a_n:

$$a_n = \frac{11n^2+11n-6}{(n+4)(n+3)}a_{n-1} + \frac{(n-3)(n-2)}{(n+4)(n+3)}a_{n-2}.$$

- Nicer (previously conjectured) summation formulas for a_n then follow.

Here are the nicer formulas

The number a_n of $2\underbrace{41}3$-avoiding permutations of size $n \geq 2$ is

$$a_n = \frac{24}{(n-1)n^2(n+1)(n+2)} \sum_{j=0}^{n} \binom{n}{j+2}\binom{n+2}{j}\binom{n+j+2}{j+1}$$

$$= \frac{24}{(n-1)n^2(n+1)(n+2)} \sum_{j=0}^{n} \binom{n+1}{j+3}\binom{n+2}{j+1}\binom{n+j+3}{j}$$

$$= \frac{24}{(n-1)n^2(n+1)(n+2)} \sum_{j=0}^{n} \binom{n}{j+2}\binom{n+1}{j}\binom{n+j+2}{j+3}$$

$$= \frac{24}{(n-1)n(n+1)^2(n+2)} \sum_{j=0}^{n} \binom{n+1}{j}\binom{n+1}{j+3}\binom{n+j+2}{j+2}$$

What we saw and what comes next

- What is a generating tree, and its encoding by a rewriting rule.
- How to use them to prove existence of (implicit) bijections.
- How to turn a rewriting rule into a functional equation for the multivariate generating function.
- How to solve it with the (obstinate) kernel method.

- How to use generating trees to encode uniform random permutations in some pattern-avoiding families as conditioned random walks.
- How to use this encoding to derive local and scaling limit results for permutations (and objects used by this encoding).

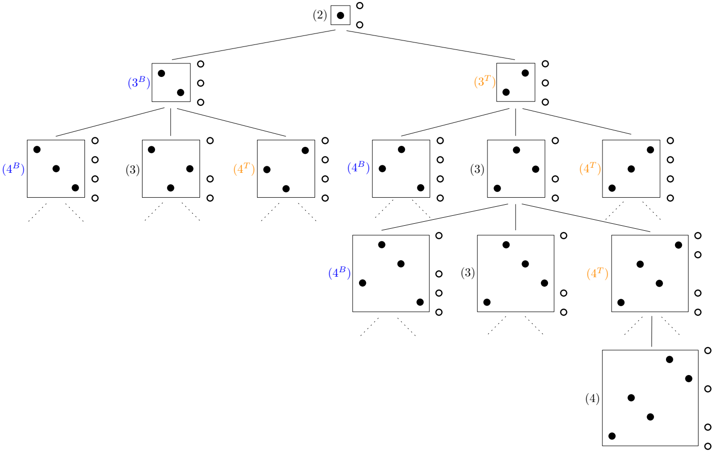
- Applications mostly on pattern-avoiding permutations (but also a bit on lattice paths).

Using generating trees with integer labels to establish CLT for consecutive patterns: the case of $Av(1423, 4123)$

Following J. Borga, *Asymptotic normality of consecutive patterns in permutations encoded by generating trees with one-dimensional labels*

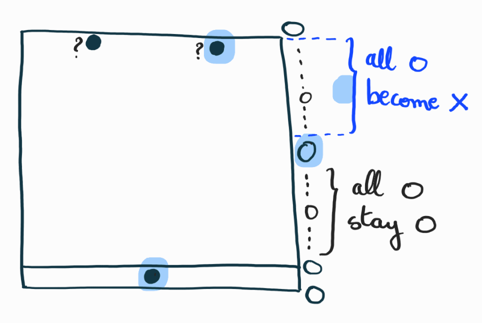
Growing permutations that avoid $\{1423, 4123\}$ on the right

As in previous examples, we obtain a generating tree for $Av(1423, 4123)$ by appending new final values in all possible ways which do not create occurrences of 1423 nor of 4123.



Active sites in $\sigma \in Av(1423, 4123)$

- The topmost site is always active.
- The two bottommost sites are always active.
- Insertion in every site different from the top- and bottommost deactivates all sites above it (except the topmost one).
- No site becomes inactive when inserting in the top- or bottommost site.



Labels and rewriting rule

We take labels to be the numbers of active sites.

The subsequent rewriting rule is

$$\Omega_{Sch} = \begin{cases} (2) & \\ (k) & \rightsquigarrow (k+1), (3), \ldots, (k), (k+1). \end{cases}$$

where the labels of the children from left to right correspond to insertions in the active sites from bottom to top.

We want to identify a permutation σ with the sequence of labels from the root to σ in this generating tree.

Problem: What is the permutation corresponding to $(2, 3)$?
$\sigma = 12$ and $\sigma = 21$ work...

Distinguishing labels by colors

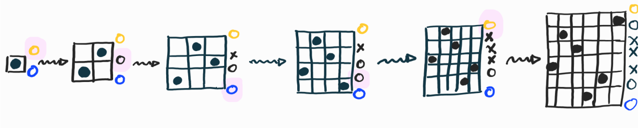
In the production $(k) \rightsquigarrow (k+1), (3), \ldots, (k), (k+1)$, the first $(k+1)$ corresponds to insertion in the bottommost site, and the second one to insertion in the topmost site. We use colors to distinguish these two cases:

$$\Omega^C_{Sch} = \begin{cases} (2) \\ (k) \rightsquigarrow (k+1)^B, (3), \ldots, (k), (k+1)^T. \end{cases}$$

Now, $\sigma = 12$ corresponds to $(2, 3^T)$ and $\sigma = 21$ corresponds to $(2, 3^B)$.

More generally, we now have a one-to-one correspondence between $Av(1423, 4123)$ and sequences of colored labels compatible with Ω^C_{Sch}.

Ex: $\sigma = 354126$ corresponds to $(2, 3^T, 3, 4^B, 3, 4^T)$ as seen from



From labels to jumps

- Rather than recording the sequence of labels, we can record the jumps between labels along a sequence.
- To have a bijective correspondence, we need colored jumps.

The jumps corresponding to the production
$(k) \rightsquigarrow (k+1)^B, (3), \dots, (k-1), (k), (k+1)^T$ are
$(+1)^B, (-k+3), \dots, (-1), (0)(+1)^T$

Ex (continued): The labels encoding $\sigma = 354126$ are $(2, 3^T, 3, 4^B, 3, 4^T)$, corresponding to the sequence of jumps $((+1)^T, 0, (+1)^B, -1, (+1)^T)$.

Remarks: The set of possible jumps is $\mathcal{J} = \mathbb{Z}_{\leq 0} \cup \{(+1)^B, (+1)^T\}$.
Not every sequence of such jumps corresponds to a permutation.

Next: We look for a distribution on the jumps that allows to see a uniform permutation in $Av(1423, 4123)$ as a conditioned random walk for this distribution on the jumps.

Finding the distribution on the jumps

Let $\phi(t) = \sum_{\substack{y \in \mathcal{J} \\ \text{with multiplicity}}} t^{-y} = \frac{1+1}{t} + (1 + t + t^2 + \dots) = \frac{2}{t} + \frac{1}{1-t}$.

Lemma [Janson]: For t such that $0 < \phi(t) < \infty$ and $\phi'(t) = 0$ (here $t = 2 - \sqrt{2}$), setting $p = \frac{1}{\phi(t)} (= 3 - 2\sqrt{2})$ and $q = \frac{1}{t} (= \frac{1}{2-\sqrt{2}})$, we have

$$p \cdot \sum_{\substack{y \in \mathcal{J} \\ \text{w. m.}}} q^y = 1, \qquad p \cdot \sum_{\substack{y \in \mathcal{J} \\ \text{w. m.}}} y \cdot q^y = 0, \qquad \text{and } p \cdot \sum_{\substack{y \in \mathcal{J} \\ \text{w. m.}}} y^2 \cdot q^y < \infty.$$

Consequence: Setting $\xi_y = p \cdot q^y$, $(\xi_y)_{y \in \mathcal{J}}$ is a centered probability distribution with finite variance.

(Here, $\xi_y = (3 - 2\sqrt{2})(2 - \sqrt{2})^{-y}$ for $y \leq 0$, and $\xi_{+1} = \xi_{+1} = \frac{2-\sqrt{2}}{2}$.)

Proposition: Take a uniform permutation of size n in $Av(1423, 4123)$ and encode it as a sequence of labels according to the generating tree.
It is distributed like the random walk with jump distribution $(\xi_y)_{y \in \mathcal{J}}$ conditioned to having length n and staying larger than the minimal label.

Two properties of conditioned random walks

Let $(\xi_y)_{y \in \mathcal{J}}$ be a centered probability distribution with finite variance (the distribution of the jumps).

Let $(X_i)_{i \geq 1}$ be the (colored) random walk with jump distribution (ξ_y).

(The X_i correspond to the labels in the generating tree.)

Proposition 1: Labels stay large.

Once conditioned to staying larger than the minimal label until step n, $(X_i)_{i \geq 1}$ satisfies that, for all $c > 0$, the probability that $X_i > c$ for all $i \leq n$ except at the beginning or end tends to 1 as n tends to infinity.

Proposition 2: Any factor of jumps is asymptotically normal.

Again, this holds for the conditioned version of $(X_i)_{i \geq 1}$.
For each fixed h, and for each fixed set A of h consecutive jumps, we count the number of positions $j \in [1..n]$ such that the sequence of jumps corresponding to $(X_i)_{i \in [j..j+h]}$ is an element of A.

From factors of jumps to consecutive patterns

Consider a permutation π of size h.
We say that π is a consecutive pattern occurring at position m in σ if $\sigma_m \sigma_{m+1} \ldots \sigma_{m+h-1}$ is an occurrence of π.

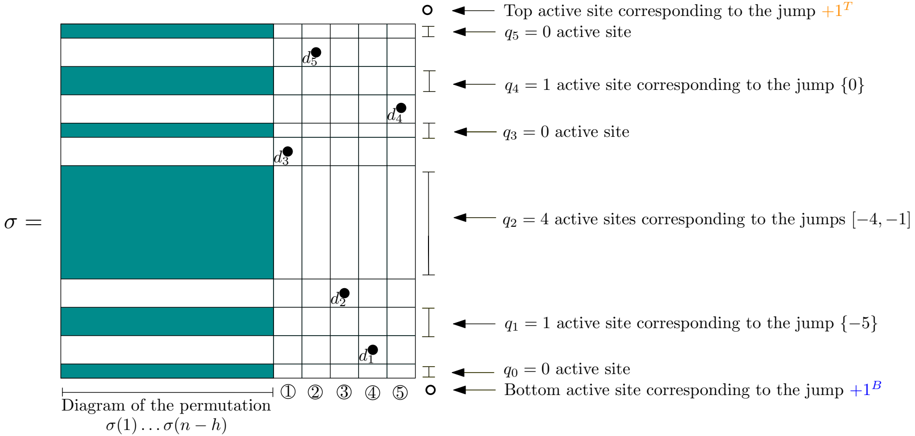
Proposition: A factor of jumps determines a consecutive pattern, provided labels are large enough.

- Let $\sigma \in Av(1423, 4123)$.
- Let $(k_i)_{i \in [1..n]}$ be the sequence of colored labels encoding σ in the generating tree.
- Assume that $k_i > h + 1$ for all $i \in [m, m + h - 1]$.

Then the consecutive pattern occurring at position m in σ is determined from the jumps describing $(k_i)_{[m,m+h-1]}$.

Proof by example and pictures: along the next few slides

Encoding the final consecutive pattern of size h in σ



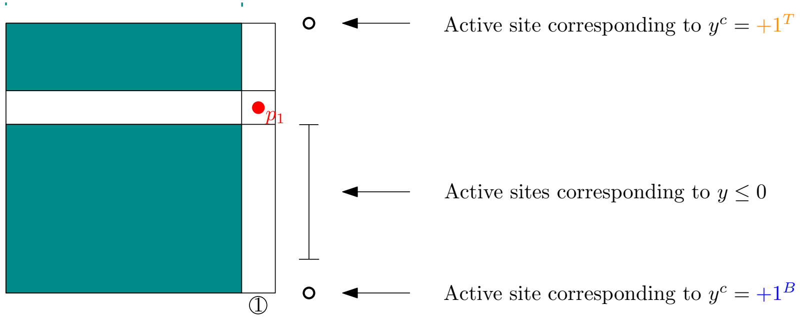
Reading the right part from bottom to top give

$$S = [④, \{-5\}, ③, [-4, -1], ①, ⑤, \{0\}, ②].$$

In particular, the final consecutive pattern of σ can be deduced from S.

Reading S (hence patterns) from the jumps

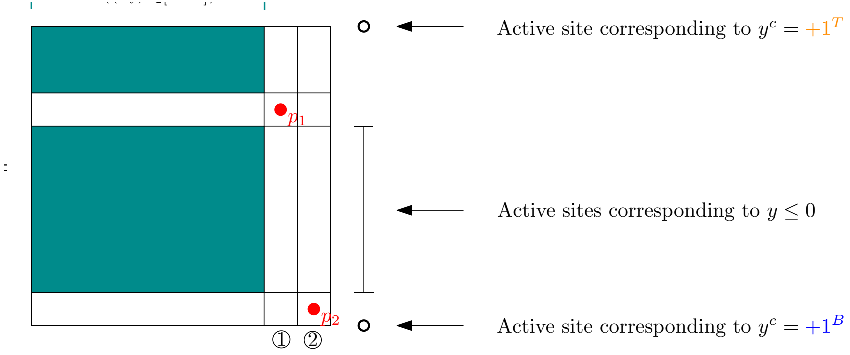
The pattern corresponding to the jumps $(-2, +1^B, +1^B, +1^T, +1^T, -7)$: step 1.



This situation is recorded by a truncated S: $S = [[?, 0], ①]$.

Reading S (hence patterns) from the jumps

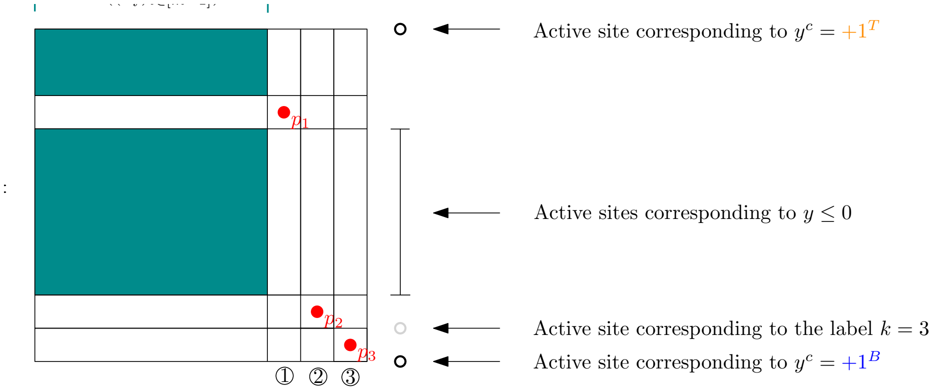
The pattern corresponding to the jumps $(-2, +1^B, +1^B, +1^T, +1^T, -7)$: step 2.



$S = [②, [?, 0], ①]$

Reading S (hence patterns) from the jumps

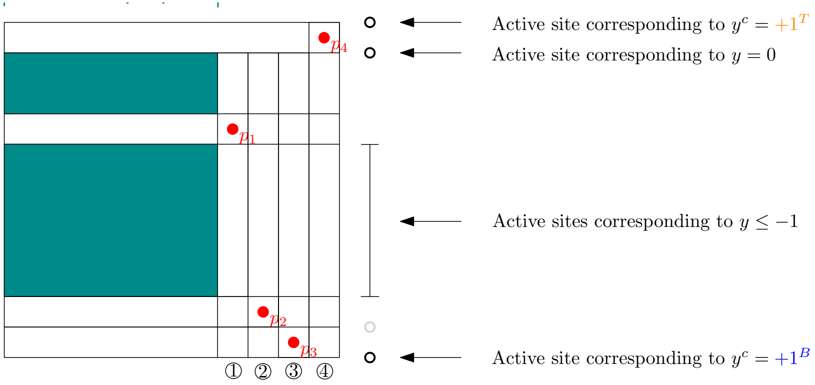
The pattern corresponding to the jumps $(-2, +1^B, +1^B, +1^T, +1^T, -7)$: step 3.



$S = [③, \{s\}, ②, [?, 0], ①]$ where the jump s yields a small label

Reading S (hence patterns) from the jumps

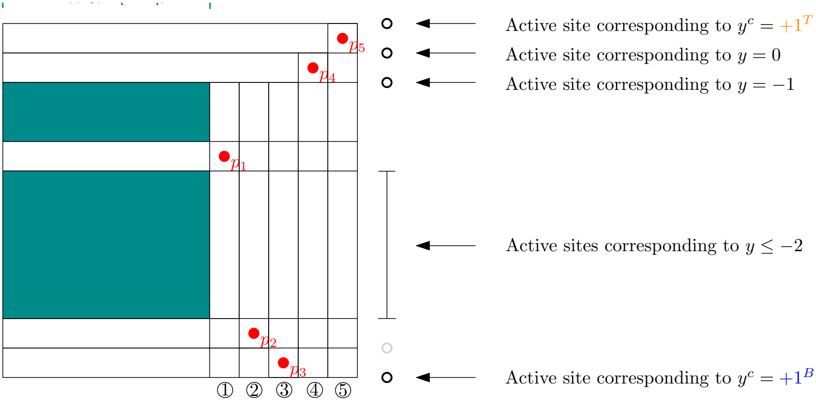
The pattern corresponding to the jumps $(-2, +1^B, +1^B, +1^T, +1^T, -7)$: step 4.



$S = [③, \{s\}, ②, [?, -1], ①, \{0\}, ④]$

Reading S (hence patterns) from the jumps

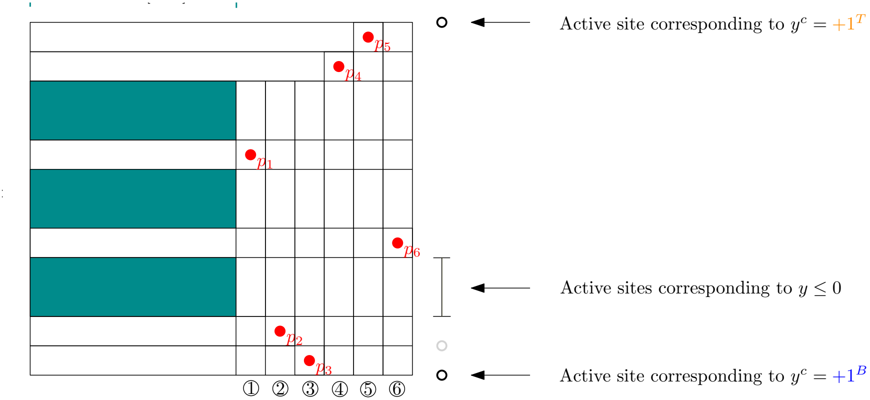
The pattern corresponding to the jumps $(-2, +1^B, +1^B, +1^T, +1^T, -7)$: step 5.



$S = [③, \{s\}, ②, [?, -2], ①, \{-1\}, ④, \{0\}, ⑤]$

Reading S (hence patterns) from the jumps

The pattern corresponding to the jumps $(-2, {\color{blue}+1^B}, {\color{blue}+1^B}, {\color{orange}+1^T}, {\color{orange}+1^T}, -7)$: step 6.



If labels are large, this gives $S = [③, \{s\}, ②, [?, -0], ⑥, ①, ④, ⑤]$

Putting everything together

- From their generating tree and associated rewriting rule, uniform permutations σ in $Av(1423, 4123)$ are encoded as conditioned random walks (X_i) for a certain distribution on the jumps.
- In these conditioned random walks (X_i), labels stay large.
- In these conditioned random walks (X_i), factors of jumps satisfy a CLT.
- Provided labels are large, a factor of jumps in (X_i) determines a consecutive pattern in σ.

$\Rightarrow$ Consecutive patterns in uniform permutations of $Av(1423, 4123)$ satisfy a CLT.

From CLT for consecutive patterns to local limit result

Local limits of permutations:

- Take a random permutation σ of size tending to ∞.
- Take a random point in it.
- Look, for each h, at the consecutive pattern π_h of size $2h+1$ around this point.

The distribution of π_h for all h characterizes the local limit of σ.

Theorem: The limiting distribution of the densities of occurrences of all **consecutive patterns** in σ characterizes the **local limit** of σ.

Consequence: Uniform random permutations in $Av(1423, 4123)$ have a local limit.

Remark: The local limit of uniform random permutations in $Av(1423, 4123)$ can be explicitly described (as a random total order on $\mathbb{Z}$, from the interpretation of factors of jumps as patterns).

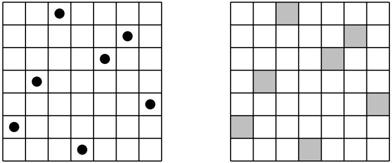
Local and permuton limits
for families of permutations
encoded by a generating tree
where labels are pairs of integers

Local limits versus permuton limits

- Local limits describe the limiting behavior around a random point.
- Permuton limits describe the global limiting behavior of a permutation. This is a scaling limit.

Permuton limits of permutations:

- Once properly rescaled, the diagram of a permutation gives a probability distribution on the unit square.



Local limits versus permuton limits

- Local limits describe the limiting behavior around a random point.
- Permuton limits describe the global limiting behavior of a permutation. This is a scaling limit.

Permuton limits of permutations:

- Once properly rescaled, the diagram of a permutation gives a probability distribution on the unit square.
- A permuton μ is a probability distribution on the unit square with uniform projections.
- A sequence of random permutations (σ_n) converges to μ if the diagrams of the σ_n converge to μ for the weak topology.

Theorem: The limiting distribution of the densities of occurrences of all **patterns** in σ characterizes the **permuton limit** of σ.

Back to $Av(2\underline{41}3)$, following J. Borga, arxiv:2112.00159

Letting these permutations grow on the right, we obtain a generating tree encoded by the following rewriting rule:

$$\Omega_{semi} = \begin{cases} (1,1) & \\ (h,k) \rightsquigarrow & (1,k+1),\ldots,(h,k+1) \\ & (h+k,1),\ldots,(h+1,k). \end{cases}$$

Shifting all labels by $(-1,-1)$, this is equivalent to

$$\Omega'_{semi} = \begin{cases} (0,0) & \\ (h,k) \rightsquigarrow & (0,k+1),\ldots,(h,k+1) \\ & (h+k+1,0),\ldots,(h+1,k). \end{cases}$$

As before, we can identify a permutation $\sigma \in Av(2\underline{41}3)$ with the sequence of labels from the root to σ in the generating tree.

Remark: All productions being distinct, we do not need to introduce colors.

Increments and encoding as conditioned random walks

Recall the rewriting rule:

$$\Omega'_{semi} = \begin{cases} (0,0) & \\ (h,k) \rightsquigarrow & (0,k+1),\dots,(h,k+1) \\ & (h+k+1,0),\dots,(h+1,k). \end{cases}$$

- The possible labels are (i,j) s.t. $i,j \geq 0$.
- The increments are $\mathcal{J} = \{(-i,+1) : i \geq 0\} \uplus \{(+i,-i+1) : i \geq 1\}$.

Can we see a uniform permutation in $Av(2\,\underline{41}\,3)$ as a 2-dimensional random walk starting at $(0,0)$ and conditioned to stay in the non-negative quadrant?

Yes, choosing the following distribution on the increments:

$$\mathbb{P}(-i,+1) = \alpha\gamma^i = \mathbb{P}(+i,-i+1) \text{ for } \alpha = \sqrt{5}-2 \text{ and } \gamma = \frac{\sqrt{5}-1}{2}.$$

An alternative to reading patterns in this encoding

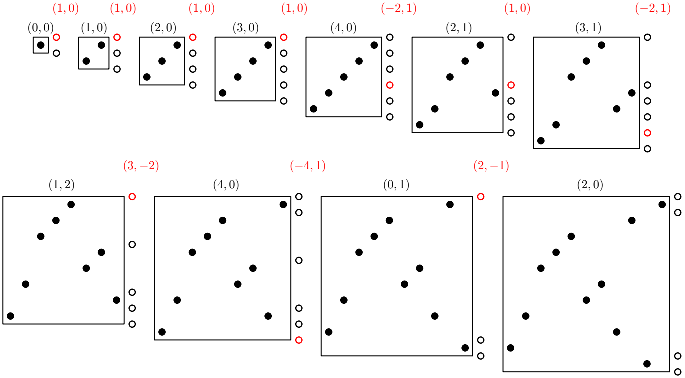
- Random walks in cones are well-understood (e.g. [Denisov, Wachtel]).
- Can we transfer results about "patterns in random walks" to patterns in permutations? And how?

We rather work with continuous objects, which are the limits of our discrete objects, and which "encode all pattern densities":

Permutation	Permuton
Walk	2-dimensional Brownian excursion
discrete coalescent-walk process	continuous coalescent-walk process

The coalescent-walk processes are built from the walks and we can read on them the permutation and its (non-consecutive) patterns.

The path from the root of the generating tree to $\sigma = 2\,4\,7\,8\,9\,5\,6\,3\,10\,1\,11 \in Av(2\,\underline{41}\,3)$ is



The walk encoding σ is shown in black.

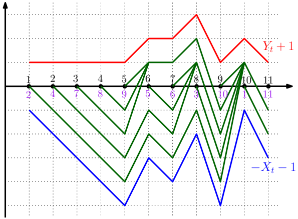
The sequence of increments encoding σ is shown in red.

Building a coalescent-walk process from a walk

Consider the (same) permutation $\sigma = 2\,4\,7\,8\,9\,5\,6\,3\,10\,1\,11 \in Av(2\,\underline{41}\,3)$ and the corresponding sequence of increments

$$(1,0),(1,0),(1,0),(1,0),(-2,1),(1,0),(-2,1),(3,-2),(-4,1),(2,-1).$$

The associated discrete coalescent-walk process is shown in green on the right.
It is a set of $n = 11$ walks $Z_n^{(i)}$ starting at $(t,0)$ for $i = 1$ to n.



Remarks:

- A permutation can be read on the coalescent-walk process.
- It coincides with the permutation we started from.
- A pattern in the permutation can be read from a finite number of green paths.

Scaling limit result for walks and coalescent-walk process

- Let W_n be a conditioned random walk as before.
- Let $\mathcal{W}_n(t)$ for $t \in [0,1]$ be the rescaled version (interpolated for t).
- Let $\mathcal{E}_\rho = (\mathcal{X}_\rho, \mathcal{Y}_\rho)$ be a two-dimensional Brownian excursion of correlation ρ in the non-negative quadrant, for $\rho = -\frac{1+\sqrt{5}}{4}$.

Theorem: $\mathcal{W}_n \quad \xrightarrow{(d)} \mathcal{E}_\rho$

Scaling limit result for walks and coalescent-walk process

- Let W_n be a conditioned random walk as before.
- Let $\mathcal{W}_n(t)$ for $t \in [0,1]$ be the rescaled version (interpolated for t).
- Let $\mathcal{E}_\rho = (\mathcal{X}_\rho, \mathcal{Y}_\rho)$ be a two-dimensional Brownian excursion of correlation ρ in the non-negative quadrant, for $\rho = -\frac{1+\sqrt{5}}{4}$.
- Let $(Z_n^{(i)})_{1 \le i \le n}$ be the coalescent-walk process associated with W_n.
- Let $\mathcal{Z}_n^{(u)}(t)$ be the rescaled version (interpolated for t and u).
- Let $(\mathcal{Z}_\rho^{(u)})_{u \in [0,1]}$ be the solution of the SDEs

$$d\mathcal{Z}_\rho^{(u)}(t) = \begin{cases} \mathbb{1}_{\{\mathcal{Z}_\rho^{(u)}(t)>0\}} d\mathcal{Y}_\rho(t) - \mathbb{1}_{\{\mathcal{Z}_\rho^{(u)}(t)<0\}} d\mathcal{X}_\rho(t) & t \in (u,1) \\ 0 & t \in [0,u]. \end{cases}$$

Theorem: $(\mathcal{W}_n, \mathcal{Z}_n^{(u_i)}) \xrightarrow{(d)} (\mathcal{E}_\rho, \mathcal{Z}_\rho^{(u_i)})$
for a sequence $(u_i)_{i \ge 0}$ of i.i.d. uniform random variables in $[0,1]$.

Scaling limit result for permutations

$\sigma \in Av(2\underbrace{41}3)$ encoded by a discrete coalescent-walk process $(Z_n^{(i)})_{1 \leq i \leq n}$ is described by $\sigma(i) < \sigma(j) \Leftrightarrow i \leqslant_Z j$ for the total order $\leqslant_Z$ defined by

$$i \leqslant_Z i; \quad i \leqslant_Z j \text{ if } i < j \text{ and } Z^{(i)}(j) \leq 0; \quad j \leqslant_Z i \text{ if } i < j \text{ and } Z^{(i)}(j) > 0.$$

$\sigma \in Av(2\underline{41}3)$ encoded by a discrete coalescent-walk process $(Z_n^{(i)})_{1\le i\le n}$ is described by $\sigma(i) < \sigma(j) \Leftrightarrow i \leqslant_Z j$ for the total order $\leqslant_Z$ defined by

$$i \leqslant_Z i; \quad i \leqslant_Z j \text{ if } i<j \text{ and } Z^{(i)}(j) \le 0; \quad j \leqslant_Z i \text{ if } i<j \text{ and } Z^{(i)}(j) > 0.$$

Similarly, we can define the permuton associated with a continuous coalescent-walk process $(\mathcal{Z}_\rho^{(u)})_{u\in[0,1]}$.

- The total order is replaced by the function

$$\varphi_{\mathcal{Z}_\rho}(t) = Leb(\{x \in [0,t) : \mathcal{Z}_\rho^{(x)}(t) \le 0\} \cup \{x \in [t,1] : \mathcal{Z}_\rho^{(t)}(x) > 0\}).$$

- The corresponding permuton μ_ρ is given by

$$\mu_\rho(\cdot) = (Id, \varphi_{\mathcal{Z}_\rho}) * Leb(\cdot) = Leb(\{t \in [0,1] : (t, \varphi_{\mathcal{Z}_\rho}(t)) \in \cdot\}).$$

Scaling limit result for permutations

$\sigma \in Av(2\underline{41}3)$ encoded by a discrete coalescent-walk process $(Z_n^{(i)})_{1\leq i\leq n}$ is described by $\sigma(i) < \sigma(j) \Leftrightarrow i \leqslant_Z j$ for the total order $\leqslant_Z$ defined by

$$i \leqslant_Z i; \quad i \leqslant_Z j \text{ if } i < j \text{ and } Z^{(i)}(j) \leq 0; \quad j \leqslant_Z i \text{ if } i < j \text{ and } Z^{(i)}(j) > 0.$$

Similarly, we can define the permuton associated with a continuous coalescent-walk process $(\mathcal{Z}_\rho^{(u)})_{u\in[0,1]}$.

- The total order is replaced by the function

$$\varphi_{\mathcal{Z}_\rho}(t) = Leb(\{x \in [0,t) : \mathcal{Z}_\rho^{(x)}(t) \leq 0\} \cup \{x \in [t,1] : \mathcal{Z}_\rho^{(t)}(x) > 0\}).$$

- The corresponding permuton μ_ρ is given by

$$\mu_\rho(\cdot) = (Id, \varphi_{\mathcal{Z}_\rho}) * Leb(\cdot) = Leb(\{t \in [0,1] : (t, \varphi_{\mathcal{Z}_\rho}(t)) \in \cdot\}).$$

Theorem: Uniform random permutations avoiding $2\underline{41}3$ converge in the permuton sense to μ_ρ.

The case of $Av(2\underline{41}3, 3\underline{14}2, 3\underline{41}2)$

Also in J. Borga, arxiv:2112.00159.

- These admit a generating tree where labels are pairs of integers.
- Uniform permutations σ can be encoded by conditioned random walks for some (explicit) distribution on the increments.
- From the increments, we can build a (similar but not identical) discrete coalescent-walk process, encoding σ and all its patterns.
- The scaling limit of the coalescent-walk process is the solution of the SDEs (with $\rho \approx -0.2151$ and $q \approx 0.3008$)

$$d\mathcal{Z}^{(u)}_{\rho,q}(t) = \begin{cases} \mathbb{1}_{\{\mathcal{Z}^{(u)}_{\rho,q}(t)>0\}} d\mathcal{Y}_\rho(t) - \mathbb{1}_{\{\mathcal{Z}^{(u)}_{\rho,q}(t)<0\}} d\mathcal{X}_\rho(t) + (2q-1) d\mathcal{L}^{\mathcal{Z}^{(u)}_{\rho,q}}(t) & t \in (u,1) \\ 0 & t \in [0,u]. \end{cases}$$

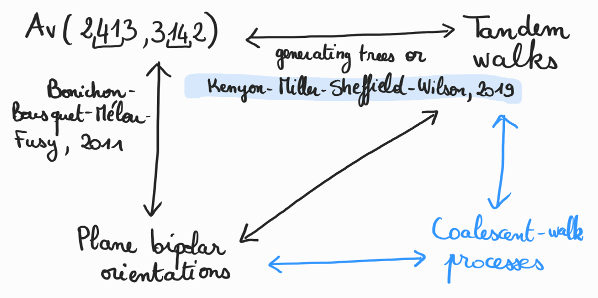
where $\mathcal{L}^{\mathcal{Z}^{(u)}_{\rho,q}}(t) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbb{1}_{\mathcal{Z}^{(u)}_{\rho,q}(s) \in [-\varepsilon,\varepsilon]} ds$

is the symmetric local-time process at zero of $\mathcal{Z}^{(u)}_{\rho,q}$.

- The permuton limit of uniform permutations avoiding $2\underline{41}3$, $3\underline{14}2$ and $3\underline{41}2$ is also obtained as the pushforward of the function $\varphi_{\mathcal{Z}_{\rho,q}}$.

The case of $Av(2\underline{41}3, 3\underline{14}2)$ (Baxter permutations)

A family of maps makes the picture richer. See J. Borga, M. Maazoun.



Results proved are

- local limit result, jointly for all four objects;
- scaling limit result, jointly for walks (associated with a map and with its dual, hence somehow for the maps), coalescent-walk processes, and permutations.

What we (hopefully) learned

- What is a generating tree, and its encoding by a rewriting rule.
- How to use them to prove existence of (implicit) bijections.
- How to turn a rewriting rule into a functional equation for the multivariate generating function.
- How to solve it with the (obstinate) kernel method.

- How to use generating trees to encode uniform random permutations in some pattern-avoiding families as conditioned random walks.
- How to use this encoding to derive local and scaling limit results for permutations (and objects used by this encoding).

- Applications mostly on pattern-avoiding permutations (but also a bit on lattice paths).

Merci beaucoup !!!