The shape of random combinatorial objects

Igor Pak, UCLA

(joint with Ted Dokos, Sam Miner)

MFO Workshop

March 6, 2014

Old Problem:

Find *nice* bijections between combinatorial objects.
Specifically, between 200+ counted by the *Catalan numbers*.

New Problem:

Explain why some objects have *super nice* (canonical) bijections while others do not (and what this all even means).

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{4^n}{\sqrt{\pi n^3}}\left(1 - \frac{9}{8n} + \frac{145}{128n^2} - \dots\right)$$

Plan:

1. Classical Catalan structures

2. Selected known results

3. Pattern avoidance

4. The results

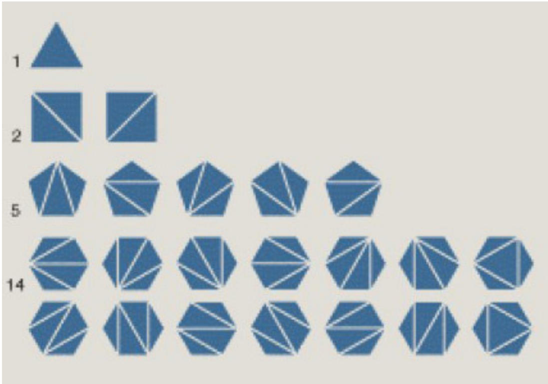
5. Connections to probability

6. Applications

7. Alternating and Baxter permutations

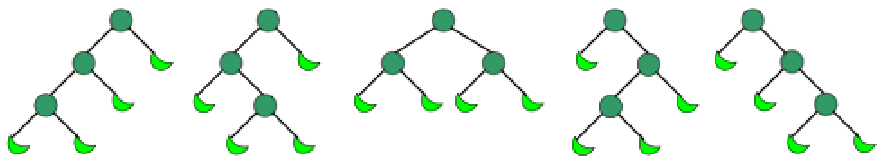
1. Classical Catalan structures:

1) C_n = number of triangulations of $(n+2)$-gon (Euler, 1756)



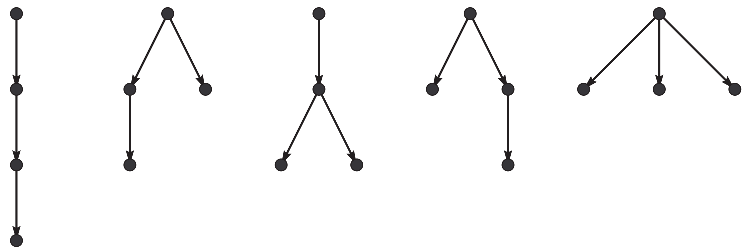
2) C_n = number of non-associative products of $(n+1)$ numbers (Catalan, 1836)

$$((ab)c)d \qquad (a(bc))d \qquad (ab)(cd) \qquad a((bc)d) \qquad a(b(cd))$$



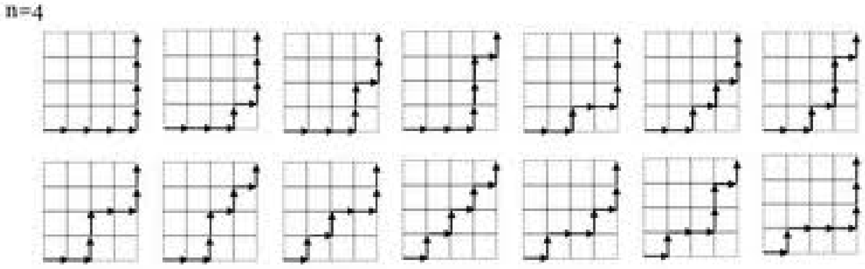
3) C_n = number of binary trees on $(2n+1)$ vertices

4) C_n = number of *plane trees* with $(n+1)$ vertices



5) C_n = number of *Dyck paths* of length $2n$

i.e. lattice paths $(0,0) \to (n,n)$ below $y = x$ line.



Canonical bijections:

Triangulations	$\longleftrightarrow$	Binary trees
Binary trees	$\longleftrightarrow$	Non-associative products
Binary trees	$\longleftrightarrow$	Plane trees
Plane trees	$\longleftrightarrow$	Dyck paths

These can be extremely useful for studying asymptotics of combinatorial statistics and more generally the *shape of combinatorial objects*.

2. Selected asymptotic results:

Theorem (Aldous, 1991; DFHNS, 1999)

The p.d.f. of the maximal chord-length in a random triangulation of regular n-gon

$$\text{converges to} \qquad \frac{3x-1}{\pi x^2(1-x)^2\sqrt{1-2x}}\,, \quad \frac{1}{3} < x < \frac{1}{2}\,, \quad \text{as } n \to \infty.$$

Theorem (DFHNS, 1999)

Δ_n = maximal degree of a random triangulation of n-gon. Then for all $c > 0$

$$P\big(|\Delta_n - \log_2 n| < c \log\log n\big) \to 1 \quad \text{as} \quad n \to \infty.$$

DFHNS = Devroye, Flajolet, Hurtado, Noy and Steiger.

Theorem: Let δ_n be the degree of a root in a random plane tree with n vertices.

$$P(\delta_n = r) \to \frac{r}{2^{r+1}}, \quad E[\tau] \to 3 \quad \text{as} \quad n \to \infty.$$

Theorem: Let h_n height of a random plane tree with n vertices, m_n the height of a random Dyck path of length $2n$. Then:

$$h_n,\, m_n \sim \sqrt{\frac{\pi n}{2}}$$

General References: Flajolet & Sedgewick, *Analytic Combinatorics*, 2009.
M. Drmota, *Random Trees*, 2009.

3. Pattern avoidance:

Permutation $\sigma \in S_n$ contains *pattern* $\omega \in S_n$ if matrix $M(\sigma)$ contains $M(\omega)$ as a submatrix. Otherwise, σ *avoids* ω.

Example

$\sigma = (2, 4, 5, 1, 3, 6)$ contains **132** but *not* **321**.

$$M(\sigma) = \begin{pmatrix} 0 & \textcircled{1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \textcircled{1} & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \textcircled{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad \textit{contains} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \textit{but not} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Patterns of length 3

$s_n(\omega) :=$ number of permutations $\sigma \in S_n$ avoiding ω

Theorem (MacMahon, 1915; Knuth, 1968)
$s_n(\omega) = C_n$ for all $\omega \in S_3$.

Two Observations:
$s_n(123) = s(321)$, $s_n(132) = s(231) = s_n(312) = s(213)$ via symmetries
[Kitaev]: Nine different bijections between **123**- and **132**-avoiding permutations.

Question: Can it be true that all nine and *nice*? How about canonical?

My Answer: No canonical bijection is possible. Here is why...

Simulations by Madras and Pehlivan

Monte Carlo simulation 1

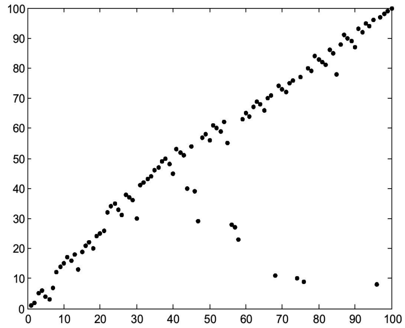


Figure: Randomly generated 312 avoiding permutation with N=100

Monte Carlo simulation 2

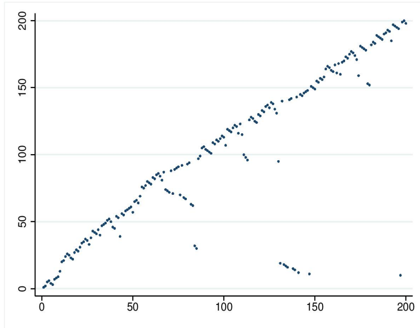


Figure: Randomly generated 312 avoiding permutation with N=200

4. Shape of random pattern avoiding permutations

$$P_n(i,j) := \frac{1}{C_n} \sum_{\sigma} M(\sigma)_{ij} \, ,$$

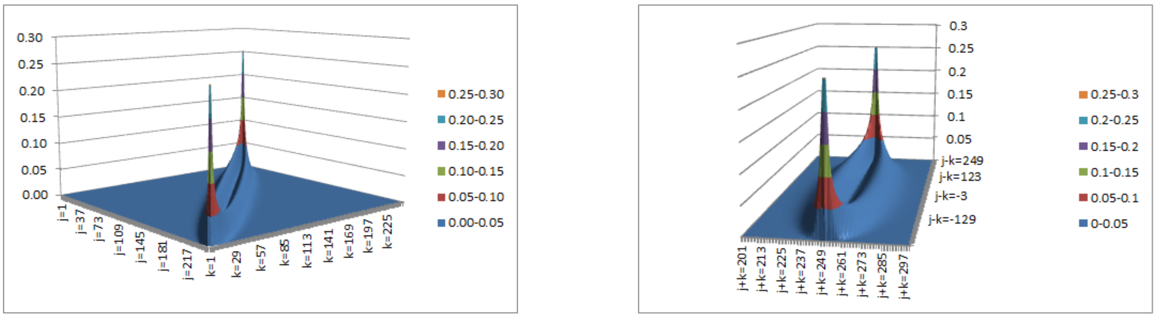
where the sum is over all **123**-avoiding permutations.

$$Q_n(i,j) := \frac{1}{C_n} \sum_{\sigma} M(\sigma)_{ij} \, ,$$

where the sum is over all **132**-avoiding permutations.

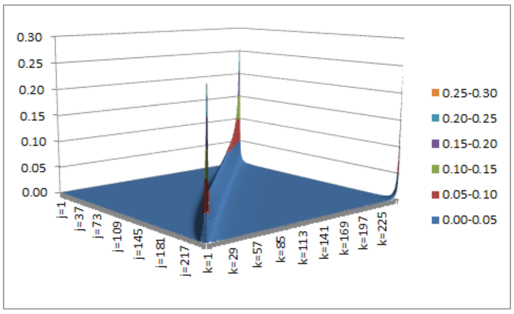
Main Question: What do $P_n(*,*)$ and $Q_n(*,*)$ look like, as $n \to \infty$?

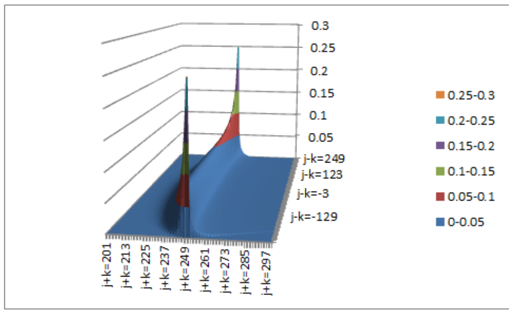
Shape of random 123-avoiding permutations (surface)



Surface $P_{250}(i, j)$ and the same surface in greater detail.

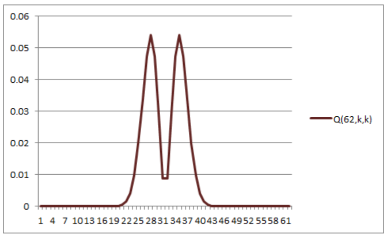
Shape of random 132-avoiding permutations (surface)

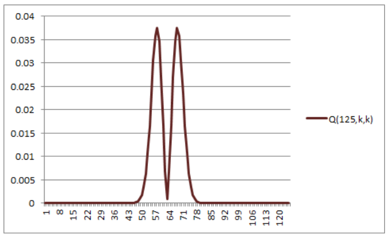


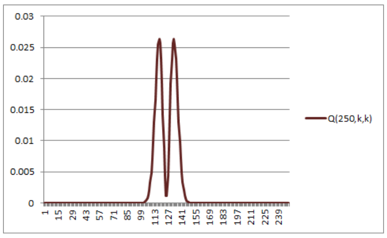


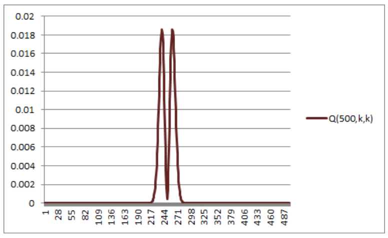
Surface $Q_{250}(i, j)$ and the same surface in greater detail.

Diagonal of $P_n(*,*)$ in details









Main Theorem for $P_n(*,*)$, [Miner-P.]

$$P_n(an, bn) < \varepsilon^n, \qquad a+b \neq 1, \quad \varepsilon = \varepsilon(a,b), \quad 0 < \varepsilon < 1$$

$$P_n(an - cn^\alpha, (1-a)n - cn^\alpha) < \varepsilon^{n^{2\alpha-1}}, \qquad \frac{1}{2} < \alpha < 1, \quad \varepsilon = \varepsilon(a,b,\alpha), \quad 0 < \varepsilon < 1$$

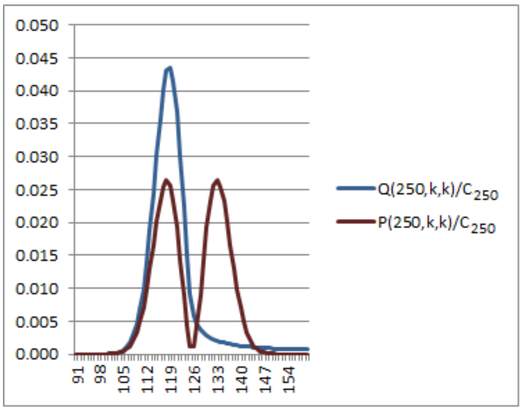
$$P_n(an - cn^\alpha, (1-a)n - cn^\alpha) \sim \eta(a,c)\,\varkappa(a,c)\,\frac{1}{\sqrt{n}}, \qquad \alpha = \frac{1}{2}, \quad c \neq 0$$

$$P_n(an - cn^\alpha, (1-a)n - cn^\alpha) \sim \eta(a,c)\,\frac{1}{n^{3/2-2\alpha}}, \qquad 0 < \alpha < \frac{1}{2}, \quad c \neq 0$$

where

$$\eta(a,c) = \frac{c^2}{\sqrt{\pi}(a(1-a))^{\frac{3}{2}}} \qquad \text{and} \qquad \varkappa(a,c) = \exp\left[\frac{-c^2}{a(1-a)}\right]$$

Diagonal of $Q_n(*,*)$ vs. $P_n(*,*)$



Main Theorem for $Q_n(*,*)$, macro picture:

$$Q_n(an, bn) < \varepsilon^n, \qquad 0 \le a+b < 1, \quad \varepsilon = \varepsilon(a,b), \quad 0 < \varepsilon < 1$$

$$Q_n(an, bn) \sim v(a,b)\,\frac{1}{n^{3/2}}\,, \qquad 1 < a+b < 2$$

$$Q_n(n,n) \sim \frac{1}{4}$$

where

$$v(a,b) = \frac{1}{\sqrt{32\pi}\,(2-a-b)^{\frac{3}{2}}\,(1-a-b)^{\frac{3}{2}}}$$

Main Theorem for $Q_n(*,*)$, micro picture:

$$Q_n(an-cn^\alpha,(1-a)n-cn^\alpha) < \varepsilon^{n^{2\alpha-1}}, \qquad \frac{1}{2}<\alpha<1,\ \varepsilon=\varepsilon(a,b,\alpha),\ 0<\varepsilon<1,\ c>0$$

$$Q_n(an-cn^\alpha,(1-a)n-cn^\alpha) \sim z(a)\,\frac{1}{n^{3/2-2\alpha}}, \qquad \frac{3}{8}<\alpha<\frac{1}{2},\quad c>0$$

$$Q_n(an-cn^\alpha,(1-a)n-cn^\alpha) \sim z(a)\,\frac{1}{n^{3/4}}, \qquad 0<\alpha<\frac{3}{8}$$

$$Q_n(an+cn^\alpha,(1-a)n+cn^\alpha) \sim y(a,c)\,\frac{1}{n^{3/4}}, \qquad \frac{3}{8}<\alpha<\frac{1}{2},\quad c>0$$

$$Q_n(an+cn^\alpha,(1-a)n+cn^\alpha) \sim w(c)\,\frac{1}{n^{3\alpha/2}}, \qquad \frac{1}{2}<\alpha<1,\quad c>0$$

$$Q_n(n-cn^\alpha,n-cn^\alpha) \sim w(c)\,\frac{1}{n^{3\alpha/2}}, \qquad 0<\alpha<1,\quad c>0$$

where

$$w(c)=\frac{1}{16c^{\frac{3}{2}}\sqrt{\pi}},\quad y(a,c)=\left(1+\frac{\zeta(\frac{3}{2})}{\sqrt{\pi}}\right)\frac{c^2}{\sqrt{\pi}a^{\frac{3}{2}}(1-a)^{\frac{3}{2}}},\quad z(a)=\frac{\Gamma(\frac{3}{4})}{2^{\frac{9}{4}}\pi a^{\frac{3}{4}}(1-a)^{\frac{3}{4}}}$$

Proof idea:

Lemma 1. For $j + k \leq n + 1$,

$$P_n(j,k) \;=\; B(n-k+1,j)\,B(n-j+1,k), \quad \text{where}$$

$$B(n,k) \;=\; \frac{n-k+1}{n+k-1}\binom{n+k-1}{n} \quad \text{are the } \textit{ballot numbers}$$

Lemma 2.

$$Q_n(j,k) \;=\; \sum_{r=\max\{0,j+k-n-1\}}^{\min\{j,k\}-1} B(n-j+1,k-r)\,B(n-k+1,j-r)\,C_r$$

Proof of the Main Theorem $=$ Lemmas $+$ Stirling's formula $+$ [details]

Bijective combinatorics:

123-avoiding permutations $\mapsto_{\text{RSK}}$ Pairs of SYT $\mapsto$ Dyck paths

Corollary: $P_n(i, j)$ = Probability that random Dyck path is at height j after $(i + j)$ steps

132-avoiding permutations $\mapsto$ Binary trees

5. Connections to Probability:

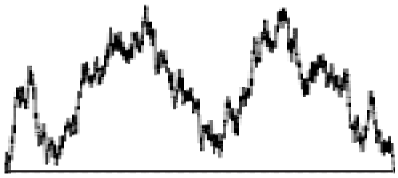
Random Dyck paths $\longrightarrow$ Brownian excursion

This explains everything!

Hint:

(1) heights in Dyck paths $\longleftrightarrow$ distances to anti-diagonal in **123**-av

(2) tunnels in Dyck paths $\longleftrightarrow$ distances to anti-diagonal in **132**-av



6. Applications

Corollary [Miner-P.]

Let $fp(\sigma)$ denote the number of fixed points in $\sigma \in S_n$.

$$\mathbb{E}[fp(\sigma)] \sim \frac{2\Gamma(\frac{1}{4})}{\sqrt{\pi}}\, n^{\frac{1}{4}}, \quad \text{as} \;\; n \to \infty.$$

where $\sigma \in S_n$ is a uniform random **231**-avoiding.

Note: For other patterns the expectations for the number of fixed points were computed by Elizalde (MIT thesis, 2004). Curiously, they are all $O(1)$.

Main theorem also gives asymptotics for a large number of other statistics, such as rank, λ-rank, lis, last, etc.

7. The mysterious Baxter surface

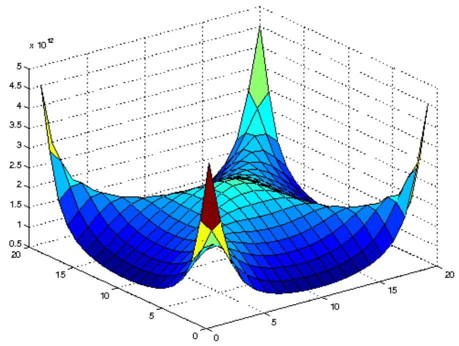
Baxter permutations: Permutations $\sigma \in S_n$ such that there are no indices $i < j < k$ with $\sigma(j+1) < \sigma(i) < \sigma(k) < \sigma(j)$ or $\sigma(j) < \sigma(k) < \sigma(i) < \sigma(j+1)$.

$$B_n = \sum_{k=1}^{n} \frac{\binom{n+1}{k-1}\binom{n+1}{k}\binom{n+1}{k+1}}{\binom{n+1}{1}\binom{n+1}{2}}$$

Note: They are connected to tilings (Korn), to plane bipolar orientations (Bonichon – Bousquet-Mélou – Fusy), and 3-tuples of non-intersecting paths (Dulucq – Guibert, Fusy – Poulalhon – Schaefer). They were introduced in analytic context by Glen Baxter (1964).

Open Problem: What is the the limit shape of Baxter permutations?

Note: The bijections allow uniform generation, but don't seem to be very helpful.



Note: Computation by Ted Dokos, UCLA.

Doubly alternating Baxter permutations

Theorem [Guibert–Linusson, 2000]

The number of Baxter permutations $\sigma \in S_{2n}$ (or S_{2n+1}), such that both σ and σ^{-1} are alternating, is the Catalan number C_n.

Denote by $\mathcal{B}_n$ the set of such permutations.

Question: What is the limit shape of permutations $\mathcal{B}_m$?

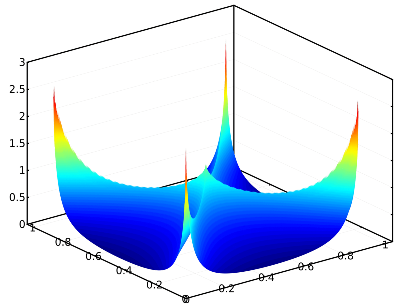
Let $P(m, i, j)$ denote the probability that a random $\sigma \in \mathcal{B}_{2m}$ has $\sigma(i) = j$.

Theorem [Dokos–P., 2014]
Let $0 < \alpha < \beta < 1 - \alpha$. We have:

$$P(m, \lfloor 2\alpha m \rfloor, \lfloor 2\beta m \rfloor) \sim \frac{\varphi(\alpha,\beta)}{m} \quad \text{as} \quad m \to \infty\,,$$

where

$$\varphi(\alpha,\beta) = \frac{1}{8\pi}\int_0^{\alpha}\int_0^{\alpha-y} \frac{dx\ dy}{[(x+y)(\beta-x)(1-\beta-y)]^{3/2}}\,.$$

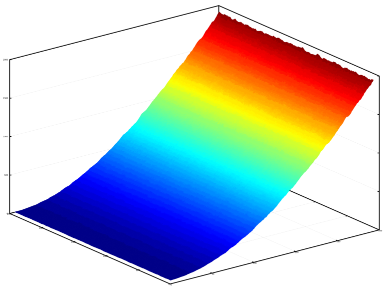


Final note: alternating permutations

Below is a plot of random $\sigma \in Alt_{500}$, i.e. $\sigma(1) > \sigma(2) < \sigma(3) > \sigma(4) < \ldots > \sigma(500)$. (only odd values are shown, boundary smoothened).

Right boundary is an inverted $\sin(x)$ curve, $0 < x < \pi/2$ [Diaconis–Matchett, 2012]

Conjecture: Limit shape of Alt_n is horizontally flat.



Thank you!


ROUSSILLON
ARIÈGE
AUDE
ANDORRA
HUESCA
LLEIDA
LÉRIDA
PYRÉNÉES-
ORIENTALES
CATALUNYA
CATALUÑA
GIRONA
GERONA
BARCELONA
TARRAGONA
Barcelona
0
50Km