Séminaire de Combinatoire Philippe Flajolet Institut Henri Poincaré 03/04/2014

Limit shapes of large Alternating Sign Matrices Filippo Colomo INFN, Florence

Joint works with:

Andrei Pronko (PDMI-Steklov, Saint-Petersbourg) Andrea Sportiello (CNRS - UPN Paris 13) Paul Zinn-Justin (CNRS - UPMC Paris 6) Limit shapes: a simple example 1-d random walk: x(t) $t \in \mathbb{N}$ $x \in \mathbb{Z}$ $x(t+1) = x(t) \pm 1$ x(0) = 0

It is well known that under the rescaling:



Now, instead, let us rescale: $x \to x' = \frac{x}{T}$ $t \to t' = \frac{t}{T}$ and send $t, x, T \to \infty$ with x', t' fixed (scaling limit) $\frac{1}{T}x(t/T)$ $\xrightarrow{\tau \to \infty} x' = Vt'$ straight line (non-random curve)

Limit shape of Young diagrams

Uniform measure [Temperley'52][Vershik'77]

Plancherel measure [Vershik-Kerov'77][Logan-Shepp'77]





$$f(x) = \frac{\sqrt{6}}{\pi} \log \left(2 \cosh \frac{\sqrt{6}}{\pi} x \right) \qquad f(x) = \begin{cases} \frac{2}{\pi} \left[x \arcsin(\frac{x}{2}) + \sqrt{4 - x^2} \right] & |x| \le 2 \\ |x| & |x| \ge 2 \end{cases}$$

Height function models and 2-d limit shapes





Rhombi tilings of an hexagon (a.k.a. Boxed plane partitions)

[Cohn-Larsen-Propp'98]

The Arctic Circle



And further...

...till considering more generic domains



Actually all these models are avatars of the same model, `dimer covering of regular planar bipartite lattices', a.k.a. `discrete free fermions', a.k.a `non-intersecting paths'. A beautiful unified theory has been provided for regular planar bipartite graphs with deep implications in algebraic geometry and algebraic combinatorics. [Kenyon, Sheffield, Okounkov, '03-'05]

Alternating Sign Matrices

[Mills-Robbins-Rumsey'82]

An ASM is an *N* by *N* matrix such that:

- entries $\in \{0, 1, -1\}$
- non-zero entries alternate in sign
- Sum of entries along each row and column is 1

ASMs generalize permutation matrices.

The seven ASMs of order 3:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ASMs enumeration:

$$A_N = 1$$
 2, 7, 42, 429, ..., $\prod_{j=0}^{N-1} \frac{(3j+1)!}{(N+j)!}$, ...

[Mills-Robbins-Rumsey'82] [Zeilberger'95] [Kuperberg'95]

Weighted enumeration:

$$A_N(\omega) := \sum_{A \in \mathcal{A}_N} \omega^{N(A)}$$

where N(A) is the number of -1 in matrix A, and A_N is the set of ASMs of order N

Nice round formulae for $\omega = 1, 2, 3$

[MRR'83][Propp *et al*'95][Kuperberg'96]

Why? Relation with classical Orthogonal Polynomials [FC-Pronko'2005]

NB: $\omega = 2$ also enumerates `domino tilings of the Aztec Diamond'. [Propp *et al*'95]

Weighted enumeration:

$$A_{N}(\omega) := \sum_{A \in \mathcal{A}_{N}} \omega^{N(A)}$$

where N(A) is the number of -1 in matrix A, and A_N is the set of ASMs of order N

Nice round formulae for $\omega = 1, 2, 3$

[MRR'83][Propp et al'95][Kuperberg'96]

Why? Relation with classical Orthogonal Polynomials[FC-Pronko'2005]NB: $\omega = 2$ also enumerates `domino tilings of the Aztec Diamond'.[Propp et al'95]

<u>Refined enumeration</u>, according to the position r of the only 1 of the first row

$$\mathcal{A}_{N,r}(\omega) := \sum_{\substack{\mathcal{A} \in \mathcal{A}_N \\ \mathcal{A}_{1,r} = 1}} \omega^{\mathcal{N}(\mathcal{A})}$$

Again, nice round formulae for $\omega = 1, 2, 3$

[MRR'83][Zeilberger'96][FC-Pronko'2005]

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$
$$A_{3,r} = \{2, 3, 2\}$$

Doubly.....Triply.....Quadruply refined enumerations

[Stroganov'04][Di Francesco 05][FC-Pronko-05][Behrend'13][Ayyer-Romik'13]...

but nothing more because...

...a matrix has only 4 edges!

In particular,

no enumeration with conditioning of entries away from the first/last rows/columns

Many interconnections and developments:

- <u>combinatorial objects</u>: plane partitions, <u>domino tilings</u>, monotone triangles, heigth function matrices, fully packed loops...
- <u>exactly solvable models of statistical mechanics</u>: six-vertex model, dense loop model, supersymmetric quantum spin chains...
- Razumov-Stroganov correspondence, Cantini-Sportiello theorem, ...

Alternating Sign Matrices and the six-vertex model [MRR'82] [Kuperberg'96]



Alternating Sign Matrices and the six-vertex model [MRR'82] [Kuperberg'96]



Note the peculiar boundary conditions (domain wall b.c.)

ASM enumeration:

weighted enumeration:

a=b=c=1

a=b=1 $\omega=c^2$

and, in general, two independent parameters.

A typical 10 x 10 ASM



Just a single 1 in first and last rows and columns
Non-zeroe entries stay away from corners

From now on, ASM pictures produced with an improved version of a C code kindly provided by Ben Wieland, and based on `Coupling from the past' algorithm [Propp-Wilson'96]

A typical 100 x 100 ASM



A typical 500 x 500 ASM



Summary and program

- We assume that ASMs have a definite Arctic curve as $N \to \infty$.
- To determine it we first need to define and evaluate a suitable `correlation function', i.e, a <u>sufficiently refined</u> ASM-enumeration, able to recognize the presence of non-zero entries in the matrices, away from the boundaries.
- Next, we shall evaluate its asymptotic behaviour in some `scaling limit' for $N \rightarrow \infty$, obtaining an analitic expression for the Arctic Curve.
- The results generalizes to arbitrary ω weights.
- Finally we provide an alternative derivation of the above results, that can be extended to ASMs of generic shape, that we call "Alternating Sign Arrays" (ASAs).

Emptiness Formation Probability (EFP)

$$egin{aligned} \mathcal{A}_{\mathcal{N},r,s}(\omega) &:= \sum_{\substack{A \in \mathcal{A}_{\mathcal{N}} \ \mathcal{A}_{i,j} = 0 \ ext{if} \ i \leq s \ ext{and} \ j \leq r}} \omega^{\mathcal{N}(A)} \end{aligned}$$

NB: r and s are horizontal and vertical coordinates, respectively.

 $F_{N}^{(r,s)}$



A 10 \times 10 ASM with only 0 entries in the top-left rectangle of size s = 3and r = 4 $F_N^{(r,s)}$

Emptiness Formation Probability (EFP)

[FC-Pronko'08]

$$A_{N,r,s}(\omega) := \sum_{\substack{A \in \mathcal{A}_N \\ A_{i,j} = 0 \text{ if } i \leq s \text{ and } j \leq r}} \omega^{N(A)}$$

NB: r and s are horizontal and vertical coordinates, respectively.



```
A 10 \times 10 ASM with only
0 entries in the top-left
rectangle of size s = 3
and r = 4
```



Emptiness Formation Probability (EFP) [FC-Pronko'08]



 $A_{10} \times 10$ ASM with only 0 entries in the top-left rectangle of size s = 3and r = 4

It is easily seen that

- $F_N^{(r,s)} \sim 1$ for small r, s, i.e. near the top left corner; $F_N^{(r,s)} \sim 0$ for large r, s, deeply inside the matrix;

If the Arctic Curve exists, in the scaling limit:

$$N, r, s \rightarrow \infty, \qquad \frac{r}{N} = x \qquad \frac{s}{N} = y \qquad x, y \in [0, 1]$$

then $F_N^{(r,s)}$ should have a stepwise behaviour, from 1 outside the Arctic Curve, to 0 inside it, with the unit jump occurring in correspondence of the Arctic Curve.

Of course only the top-left `quarter' of the Arctic Curve can be detected

Multiple Integral Representation for EFP [FC-Pronko'08]

Define the generating function for the refined enumeration:

$$h_N(z) := \frac{1}{A_N} \sum_{r=1}^N A_{N,r} z^{r-1}, \qquad h_N(1) = 1.$$

Now define, for $s = 1, \ldots, N$:

$$h_N^{(s)}(z_1,\ldots,z_s) := \prod_{1 \le j < k \le s} (z_k - z_j)^{-1} \det_{1 \le j,k \le s} \left[h_{N-s+k}(z_j)(z_j - 1)^{k-1} z_j^{s-k} \right]$$

The functions $h_N^{(s)}(z_1, \ldots, z_s)$ are totally symmetric polynomials of order N - 1 in z_1, \ldots, z_s .

Two important properties of $h_N^{(s)}(z_1, ..., z_s)$: $h_N^{(s)}(z_1, ..., z_{s-1}, 0) = h_N(0)h_{N-1}^{(s-1)}(z_1, ..., z_{s-1}),$

$$h_N^{(s)}(z_1,\ldots,z_{s-1},1)=h_N^{(s-1)}(z_1,\ldots,z_{s-1}).$$

Multiple Integral Representation for EFP [FC-Pronko'08]

The following Multiple Integral Representation is valid (r, s = 1, 2, ..., N):

$$F_{N}^{(r,s)} = \frac{(-1)^{s}A_{s}}{s!(2\pi i)^{s}} \oint_{C_{0}} \cdots \oint_{C_{0}} d^{s}z \prod_{j=1}^{s} \frac{1}{z_{j}^{r}(z_{j}-1)^{s}} \\ \times \prod_{\substack{j,k=1\\j\neq k}}^{s} \frac{z_{k}-z_{j}}{z_{j}z_{k}-z_{j}+1} h_{N,s}(z_{1},\ldots,z_{s})h_{s,s}(1-z_{1},\ldots,1-z_{s})$$

where $h_N(z)$ is known from [Zeilbeger'96]:

$$h_N(z) = {}_2F_1\left(\left. \begin{array}{c} -N+1, N\\ 2N \end{array} \right| 1-z \right)$$

- rigorous result
- the only one providing info about ASM entries away from boundaries
- a similar, more general formula, holds for generic values of ω

Ingredients:

- bijection of ASMs with the six-vertex model with domain-wall b.c.
- Quantum Inverse Scattering Method to obtain a recurrence relation, which is solved in terms of a determinant representation on the lines of Izergin-Korepin formula;
- Orthogonal Polynomial and Random Matrices technologies to rewrite it as a multiple integral.

Evaluate: $F(x, y) := \lim_{N \to \infty} F_N(xN, yN)$ $x, y \in [0, 1]$ in the limit: $N, r, s \to \infty$ $\frac{r}{N} = x$ $\frac{s}{N} = y$

using Saddle-Point method.

Saddle-point equations:

$$-\frac{s}{z_{j}-1} - \frac{r}{z_{j}} - \sum_{\substack{k=1\\k\neq j}}^{s} \left(\frac{z_{k}-1}{z_{j}z_{k}-z_{j}+1} + \frac{z_{k}}{z_{j}z_{k}-z_{k}+1} + \frac{2}{z_{k}-z_{j}} \right) + \frac{\partial \ln h_{N,s}(z_{1},\ldots,z_{s})}{\partial z_{j}} + \frac{\partial \ln h_{s,s}(1-z_{1},\ldots,1-z_{s})}{\partial z_{j}} = 0,$$

$$(j = 1, \ldots, s)$$

[FC-Pronko'10]

<u>NB1</u>:

- $s \times s$ Vandermonde determinant
- *s*-order pole at z = 1



Penner Random Matrix model [Penner'88]

<u>NB2</u>:

- By construction, in the scaling limit, EFP is 1 in the frozen region, and 0 in the disordered one, with a stepwise behaviour in correspondence of the Arctic curve.
- From the structure of the Multiple Integral Representation, such stepwise behaviour can be ascribed to the position of the SPE roots with respect to the pole at z = 1.
- The considered generalized Penner model allows for condensation of `almost all' SPE roots at z = 1. [Tan'92] [Ambjorn-Kristjansen-Makeenko'94]

[FC-Pronko'10]

<u>NB1</u>:

- $s \times s$ Vandermonde determinant
- s-order pole at z = 1



Penner Random Matrix model [Penner'88]

<u>NB2</u>:

- By construction, in the scaling limit, EFP is 1 in the frozen region, and 0 in the disordered one, with a stepwise behaviour in correspondence of the Arctic curve.
- From the structure of the Multiple Integral Representation, such stepwise behaviour can be ascribed to the position of the SPE roots with respect to the pole at z = 1.
- The considered generalized Penner model allows for condensation of `almost all' SPE roots at z = 1. [Tan'92] [Ambjorn-Kristjansen-Makeenko'94]

Condensation of `almost all'

SPE roots at z = 1

Arctic Curves

[FC-Pronko'10]

<u>NB1</u>:

- $s \times s$ Vandermonde determinant
- s-order pole at z = 1



Penner Random Matrix model [Penner'88]

<u>NB2</u>:

- By construction, in the scaling limit, EFP is 1 in the frozen region, and 0 in the disordered one, with a stepwise behaviour in correspondence of the Arctic curve.
- From the structure of the Multiple Integral Representation, such stepwise behaviour can be ascribed to the position of the SPE roots with respect to the pole at z = 1.
- The considered generalized Penner model allows for condensation of `almost all' SPE roots at z = 1. [Tan'92] [Ambjorn-Kristjansen-Makeenko'94]

Condensation of `almost all' SPE roots at z = 1

Arctic Curves

Mathematically, the condition of total condensation (i.e. the Arctic curve) is given by:

$$\frac{y}{z-1} - \frac{1-x+y}{z} + \lim_{N \to \infty} \frac{1}{N} \partial_z \ln h_N(z) = 0$$

must have two coinciding real roots in the interval: $z \in [1,\infty)$.

Evaluation of $h_N(z)$ ($\omega = 1$)

We have:

$$h_N(z) = {}_2F_1\left(\left. \begin{array}{c} -N+1, N\\ 2N \end{array} \right| 1-z \right)$$
 [Zeilberger'96]

Applying saddle-point method to the corresponding Euler integral representation we evaluate the large N behaviour:

og
$$h_N(z; \frac{1}{2}, 1) = N \log [4v(1-v)(1-v+zv)] + o(N)$$

where $v := \frac{2-z-\sqrt{z^2-z+1}}{3(1-z)}$

The `reduced SPE' thus read:

$$\frac{y}{z-1} - \frac{1-x+y}{z} + \frac{1-\sqrt{z^2-z+1}}{z(1-z)} = 0$$

Requiring this has two coinciding roots over the interval $[1, +\infty)$ gives:

$$x = 1 - rac{2z_0 - 1}{2\sqrt{z_0^2 - z_0 + 1}}, \qquad y = 1 - rac{z_0 + 1}{2\sqrt{z_0^2 - z_0 + 1}}, \qquad z_0 \in [1,\infty)$$

i.e.:

$$(2x-1)^2 + (2y-1)^2 - 4xy = 1, \qquad x, y \in \left[0, \frac{1}{2}\right].$$



Ben Wieland (January 2008)

http://www.math.brown.edu/~wieland

$$(2x-1)^2 + (2y-1)^2 - 4xy = 1, \qquad x, y \in \left[0, \frac{1}{2}\right].$$

ASMs: N=1500

10 samples

[FC-Pronko'10]



Ben Wieland (April 2008)

http://www.math.brown.edu/~wieland

Arbitrary ω ($\Delta = 1 - \omega/2$):

$$\begin{array}{lll} x & = & \displaystyle \frac{1}{\varPhi(\zeta + \lambda - \eta, 2\eta) \varPsi(\zeta, 2\eta) - \varPsi(\zeta + \lambda - \eta, 2\eta) \varPhi(\zeta, 2\eta)} \\ & & \displaystyle \times \left\{ \left[\varPsi(\zeta, \lambda - \eta) - \gamma^2 \varPsi(\gamma \zeta, \gamma (\lambda - \eta)) \right] \varPhi(\zeta, 2\eta) \\ & & \displaystyle - \left[\varPhi(\zeta, \lambda - \eta) - \gamma \varPhi(\gamma \zeta, \gamma (\lambda - \eta)) \right] \varPsi(\zeta, 2\eta) \right\}, \end{array} \\ y & = & \displaystyle \frac{1}{\varPhi(\zeta + \lambda - \eta, 2\eta) \varPsi(\zeta, 2\eta) - \varPsi(\zeta + \lambda - \eta, 2\eta) \varPhi(\zeta, 2\eta)} \\ & & \displaystyle \times \left\{ \left[\varPsi(\zeta, \lambda - \eta) - \gamma^2 \varPsi(\gamma \zeta, \gamma (\lambda - \eta)) \right] \varPhi(\zeta + \lambda - \eta, 2\eta) \\ & & \displaystyle - \left[\varPhi(\zeta, \lambda - \eta) - \gamma \varPhi(\gamma \zeta, \gamma (\lambda - \eta)) \right] \varPsi(\zeta + \lambda - \eta, 2\eta) \right\}. \end{array}$$

where

$$\begin{split} \varPhi(\mu) &:= \frac{\sin(2\eta)}{\sin(\mu + \eta)\sin(\mu - \eta)}, & \gamma := \frac{\pi}{\pi - \arccos \Delta} \\ \Psi(\zeta) &:= \cot \zeta - \cot(\zeta + \lambda - \eta) - \gamma \cot \gamma \zeta + \gamma \cot \gamma (\zeta + \lambda - \eta), \\ & \text{(Disordered regime, } -1 < \Delta < 1) \end{split}$$

or

NB:



Criticisms

 The present derivation of Arctic curves is based on an assumption (the `condensation hypothesis') which is rather bold and probably hard to prove.

Extension of the result to more generic domains



Use the theory provided by [Kenyon-C

[Kenyon-Okounkov-Sheffield'03-05]

Extension of the result to ASArrays ?

ASA:

the array is obtained from a $N \times N$ matrix by erasing the top-left $r \times s$ entries.



N = 1000r, s = 450

Criticisms

 The present derivation of Arctic curves is based on an assumption (the `condensation hypothesis') which is rather bold and probably hard to prove.

- Moreover the whole procedure is rather `ad hoc' and probably it can not be extended to more general situations.
- By `more general situations' we here mean arrays that are not anymore matrices, but still have rows and columns, and whose entries still satisfy the defining conditions of ASMs. We call such objects Alternating Sign Arrays (ASAs)

Alternative derivation and extension to generic ASA

Our previous result on the Arctic curve in a square domain can be rephrased as follows:

The arctic curve is the geometric caustic (envelope) of the family of straight lines:

$$\frac{1}{z}(x-1)+\frac{\omega}{(z-1)(z-1+\omega)}y+\lim_{N\to\infty}\frac{1}{N}\partial_z\ln h_N(z)=0\,,\qquad z\in[1,+\infty)$$

<u>Questions:</u>

- What is the geometrical meaning of this family of straight line?
 - why the constant term is determined by the refined enumeration (via $h_N(z)$)?
 - o what determines the angular coefficient of these lines?

Understanding this would provide:

- an alternative (geometrical) derivation of the Arctic curve;
- a geometrical strategy to attack the problem of Arctic curves of ASMs of generic shape.

Some numerical results

[FC-Sportiello, in prep.]

We here considere an `L-shaped' ASA, corresponding to an $N \times N$ ASM restricted by the condition that it should have only 0's in a top-left rectangular region of size $r \times s$.



N = 1000 *r*, *s* = 132













 $N = 1000 \\ r = 100 \\ s = 700$



⊢100 ⊣













N = 500N' = 499r = 400



 $N = 500 \\ N' = 499 \\ r = 450$

So the Arctic curve on the $N \times (N - 1)$ lattice with a bottom heavy edge at site r is the usual Arctic plus a straight tangent line crossing the boundary at r.

Similar phenomena are also observed in a variety of more general situation.



$$N = 500$$

 $N' = 499$
 $r = 450$







ASM refined at r, $A_{N,r}$

of directed path from r to Z

Maximizing the above probability with respect to r, one obtains a family of straight lines, parameterized by z:

$$rac{1}{z}(x-1)+rac{\omega}{(z-1)(z-1-\omega)}y+\lim_{N
ightarrow\infty}rac{1}{N}\partial_z\ln h_N(z)=0\,,\qquad z\in[1,+\infty)$$

which we immediately recognize! The point is that this `geometrical' construction interpretation holds for generic domains!

Note that the same procedure, applied to the most various situations always reproduce the above equation !



Maximizing the above probability with respect to r, one obtains a family of straight lines, parameterized by z:

$$rac{1}{z}(x-1)+rac{\omega}{(z-1)(z-1-\omega)}y+\lim_{N
ightarrow\infty}rac{1}{N}\partial_z\ln h_N(z)=0\,,\qquad z\in[1,+\infty)$$

which we immediately recognize! The point is that this `geometrical' construction interpretation holds for generic domains!

Note that the same procedure, applied to the most various situations always reproduce the above equation !

Thus on generic domains the problem of computing the Arctic curve is reduced to the evaluation of the generating function $h_N(z)$ of the corresponding refined enumeration.

Does this really work?

- Checking our recipe in two cases where the boundary correlation function $h_N(z)$ is available, we have reproduced:
 - the Arctic curve of the DW 6VM for generic values of Δ and t
 - the Arctic circle of the rhombus tiling of an hexagon (use the formula for Semi-strict Gelfand patterns to evaluate the refined enumeration you need, see [Cohn-Larsen-Propp '98])

What about new results? You need to know the corresponding refine enumeration!

s а C b S С

Consider the ASA built from three bundles crossing each other:

[Cantini-Sportiello '11]:

A corollary of the generalized R-S correspondence is that

$$A_{a,b,c} = A_{a+b+c} M_{a,b,c} ,$$

/ a

6

 $M_{a,b,c}$ counts rhombi tilings of the $a \times b \times c$ hexagon. where

But more is true: $A_{[a,b,c]}(r) = \sum_{r'} A_{a+b+c}(r-r')M_{a,b,c}(r')$

[Cantini-Sportiello '12]

$$x(b,c;p) = \frac{3-c}{2} - \frac{2-p}{2\sqrt{1-p+p^2}} - \frac{(1-c)(1-pb-pc+c)-2bpc}{2\sqrt{pb+pc-c}^2 - 2(pb-qc+c)+1}$$

$$y(b, c; p) = x(c, b; 1 - p)$$

where $p \in [0, 1]$ parametrizes the Arctic curve, and a + b + c = 1

$$(a = 20, b = 45, c = 70)$$

Conclusions?

There is a lot of work to do and many things to understand!