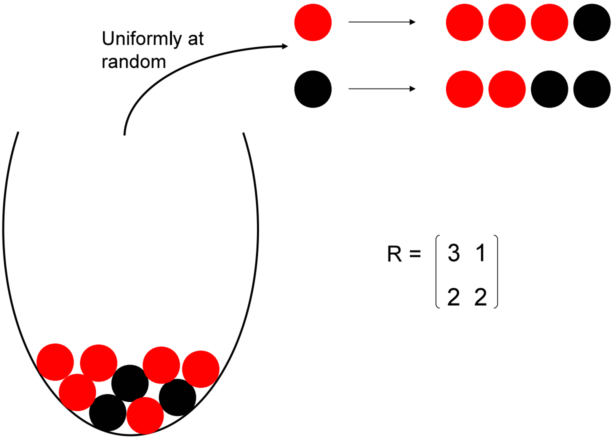
Large Pólya Urns and Smoothing Equations

Cécile Mailler
joint work with Brigitte Chauvin and Nicolas Pouyanne

Laboratoire de Mathématiques de Versailles
Université de Versailles-Saint-Quentin

May 7, 2013

What is a Pólya urn?



$$R = \begin{pmatrix} 3 & 1 \\ 2 & 2 \end{pmatrix}$$

Examples

- Pólya-Eggenberger urn (spread of epidemics)

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

- Friedman (adverse campaign) urn

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

- Ehrenfest gaz model

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

- 3-ary search tree

$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

Pólya-Eggenberger urn

Initial composition α red balls and β black balls.

Replacement matrix $\begin{pmatrix} S & 0 \\ 0 & S \end{pmatrix}$.

Composition vector at time n :

$$U(n) = \begin{pmatrix} \#\text{red balls in the urn at time } n \\ \#\text{black balls in the urn at time } n \end{pmatrix}.$$

Theorem ATHREYA:

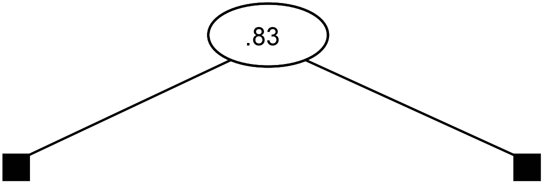
Asymptotically when n tends to infinity, almost surely,

$$\frac{U(n)}{nS} \to V,$$

where V is a Dirichlet random vector of parameter $\left(\frac{\alpha}{S}, \frac{\beta}{S}\right)$.

3-ary search tree

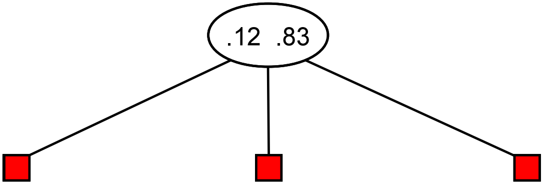
Some entries (say i.i.d. on $[0,1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

3-ary search tree

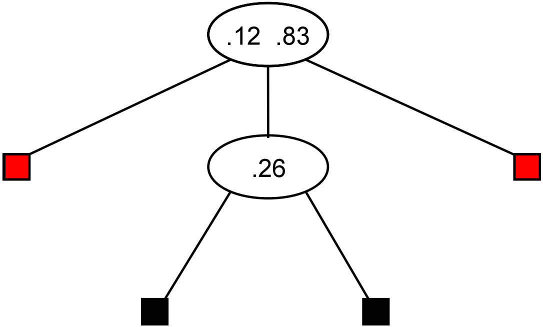
Some entries (say i.i.d. on $[0, 1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

3-ary search tree

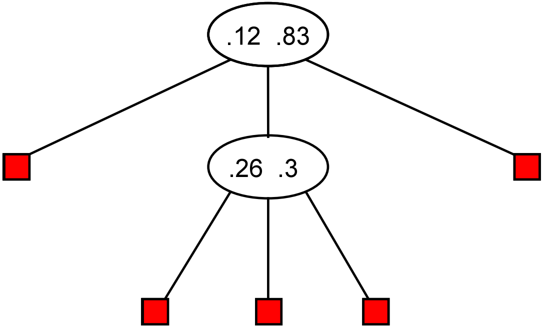
Some entries (say i.i.d. on $[0,1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

3-ary search tree

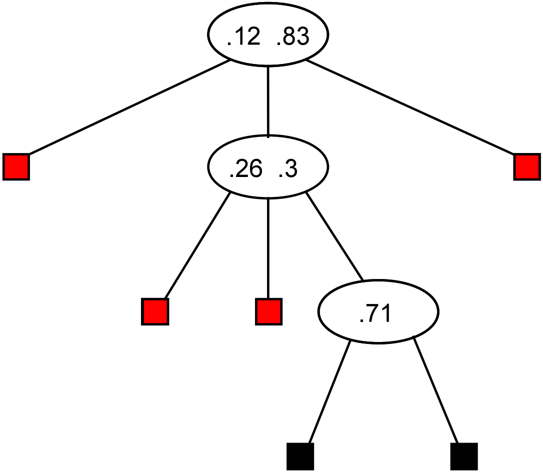
Some entries (say i.i.d. on $[0, 1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

3-ary search tree

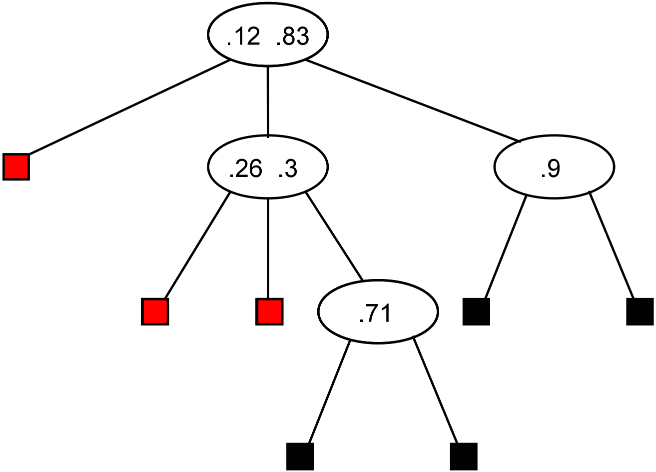
Some entries (say i.i.d. on $[0,1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

3-ary search tree

Some entries (say i.i.d. on $[0,1]$): .83 .12 .26 .3 .71 .9



$$\begin{pmatrix} -1 & 2 \\ 3 & -2 \end{pmatrix}$$

Literature

- Probabilistic approaches, via martingales, embedding in continuous time, branching processes

  ATHREYA ET AL. 60'S, GOUET 93, JANSON 05, ...

- Analytic combinatorics FLAJOLET ET AL. 05

- Algebraics POUYANNE 08

In general,

A two-colour Pólya urn is defined by:

- An initial composition:

$$U_0 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

- A replacement matrix:

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We assume:

- $a, b, c, d \geq 0$ (non extinction) and $bc \neq 0$,
- the urn is **balanced**, i.e. $a + b = c + d = S$

We denote by:

- m the second eigenvalue of R
- $\boxed{\sigma = \frac{m}{S}}$ the ratio of the two eigenvalues of R

Limit theorems

Composition vector:

$$U(n) = \begin{pmatrix} \#\text{ red balls at time } n \\ \#\text{ black balls at time } n \end{pmatrix}$$

Theorem ATHREYA, JANSON, ... :

- If $\sigma < \frac{1}{2}$

$$\frac{U(n) - nv_1}{\sqrt{n}} \xrightarrow[n\to+\infty]{(law)} \mathcal{N}(0, \Sigma^2).$$

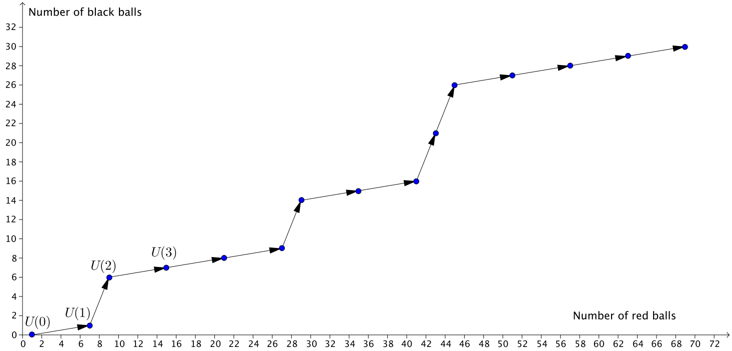
- If $\sigma > \frac{1}{2}$, then, a.s. and in all L^p ($p \geq 1$),

$$U(n) = nv_1 + n^{\sigma} W^{DT} v_2 + o(n^{\sigma}).$$

where (v_1, v_2) is a (well chosen) basis of ${}^tR$, and (u_1, u_2) its dual basis.

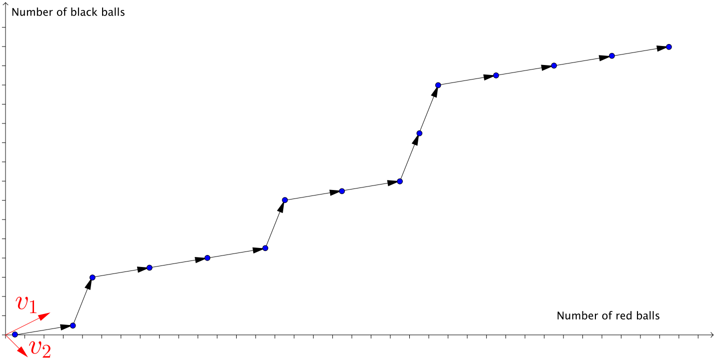
Example of a trajectory

$$U(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad R = \begin{pmatrix} 6 & 1 \\ 2 & 5 \end{pmatrix}$$



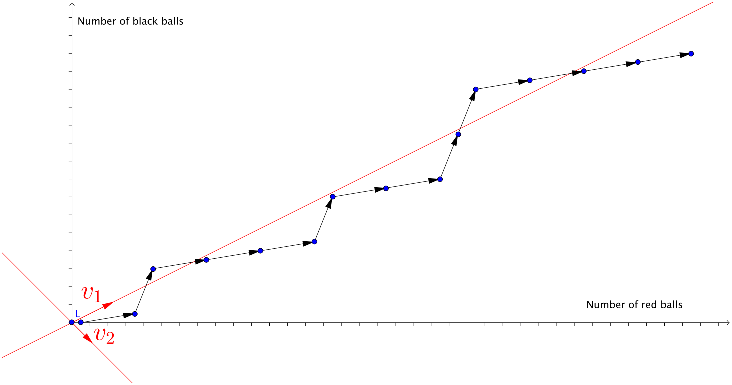
Example of a trajectory

$$U(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad R = \begin{pmatrix} 6 & 1 \\ 2 & 5 \end{pmatrix}$$



Example of a trajectory

$$U(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad R = \begin{pmatrix} 6 & 1 \\ 2 & 5 \end{pmatrix}$$



Embedding in continuous time

Each ball becomes a clock that rings after a random time of law $\mathcal{E}xp(1)$, **independently** from the others.

We denote by τ_n the date of the n^{th} ring,

$$(U(n))_{n\geq 0} \stackrel{(law)}{=} (U^{CT}(\tau_n))_{n\geq 0}.$$

Theorem ATHREYA, JANSON, ... **:**

If $\sigma > \frac{1}{2}$, asymptotically when n tends to $+\infty$, a.s. and in all L^p ($p \geq 1$),

$$U^{CT}(t) = e^{St}\xi v_1(1 + o(1)) + e^{mt} W^{CT} v_2(1 + o(1)),$$

where ξ follows a Gamma$\left(\frac{\alpha+\beta}{S}\right)$ law.

Connexions

We are interested in the properties of W

$$W^{DT}_{(\alpha,\beta)} = \lim_{n\to+\infty} u_2\left(\frac{U_{(\alpha,\beta)}(n)}{n^\sigma}\right) \quad \text{and} \quad W^{CT}_{(\alpha,\beta)} = \lim_{t\to+\infty} u_2\left(\frac{U^{CT}_{(\alpha,\beta)}(t)}{e^{mt}}\right)$$

We know that (embedding in continuous time) :

$$W^{CT}_{(\alpha,\beta)} \overset{(law)}{=} \xi^\sigma W^{DT}_{(\alpha,\beta)} \quad \text{et} \quad W^{DT}_{(\alpha,\beta)} \overset{(law)}{=} \xi^{-\sigma} W^{CT}_{(\alpha,\beta)},$$

where ξ and $W^{DT}_{(\alpha,\beta)}$ are independent in the left identity.

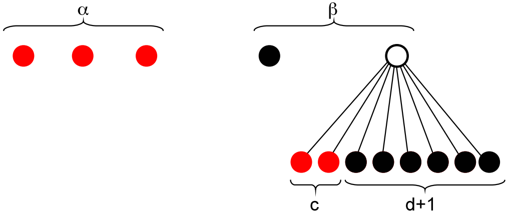
We also know Chauvin, Pouyanne, Sahnoun (among other results) that

- W^{DT} and W^{CT} admit a density on $\mathbb{R}$
- we know explicitely the Fourier transform of W^{CT}
- W^{DT} and W^{CT} have non symetric laws

Forest and urn

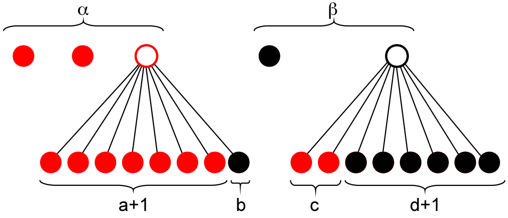
Composition of the urn = leaves in the forest.

Forest and urn



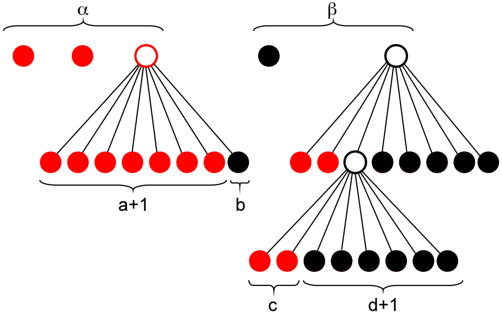
Composition of the urn = leaves in the forest.

Forest and urn



Composition of the urn = leaves in the forest.

Forest and urn



Composition of the urn = leaves in the forest.

$D_k(n)$ = number of leaves in the k^{th} tree of the forest at time n:

$$\frac{D_k(n) - 1}{S} = \text{ internal “time” in the } k^{th} \text{ tree}$$

Forest and urn

$$U_{(\alpha,\beta)}(n) \overset{(law)}{=} \sum_{k=1}^{\alpha} U_{(1,0)}^{(k)}\left(\frac{D_k(n)-1}{S}\right) + \sum_{k=\alpha+1}^{\alpha+\beta} U_{(0,1)}^{(k)}\left(\frac{D_k(n)-1}{S}\right)$$

Remark: $(D_1(n), \ldots, D_{\alpha+\beta}(n))$ is the composition vector of an urn with initial composition ${}^t(1, \ldots, 1)$ and with replacement matrix $SI_{\alpha+\beta}$.

Theorem Athreya :

$$\frac{1}{nS}(D_1(n), \ldots, D_{\alpha+\beta}(n)) \xrightarrow[n\to+\infty]{a.s.} (V_1, \ldots, V_{\alpha+\beta}),$$

where $(V_1, \ldots, V_{\alpha+\beta})$ is Dirichlet$\left(\frac{1}{S}, \ldots, \frac{1}{S}\right)$-distributed.

Forest and urn

$$U_{(\alpha,\beta)}(n) \overset{(law)}{=} \sum_{k=1}^{\alpha} U_{(1,0)}^{(k)}\left(\frac{D_k(n)-1}{S}\right) + \sum_{k=\alpha+1}^{\alpha+\beta} U_{(0,1)}^{(k)}\left(\frac{D_k(n)-1}{S}\right)$$

implies

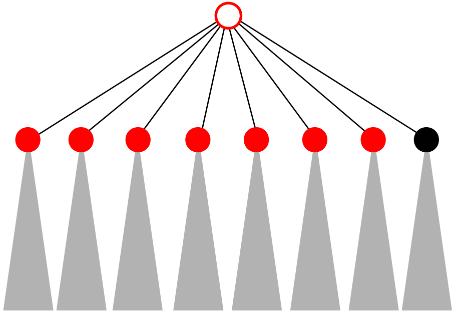
$$u_2\left(\frac{U_{(\alpha,\beta)}(n)}{n^\sigma}\right) \overset{(law)}{=} \sum_{k=1}^{\alpha} u_2\left(\frac{1}{n^\sigma}U_{(1,0)}^{(k)}\left(\frac{D_k(n)-1}{S}\right)\right) + \sum_{k=\alpha+1}^{\alpha+\beta} u_2\left(\frac{1}{n^\sigma}U_{(0,1)}^{(k)}\left(\frac{D_k(n)-1}{S}\right)\right).$$

and since

$$W_{(\alpha,\beta)}^{DT} = \lim_{n\to+\infty} u_2\left(\frac{U_{(\alpha,\beta)}(n)}{n^\sigma}\right),$$

$$\boxed{W_{(\alpha,\beta)} \overset{(law)}{=} \sum_{k=1}^{\alpha} V_k^\sigma W_{(1,0)}^{(k)} + \sum_{k=\alpha+1}^{\alpha+\beta} V_k^\sigma W_{(0,1)}^{(k)}}.$$

Let us study $W_{(1,0)}$ and $W_{(0,1)}$



$$W_{(1,0)} \overset{(law)}{=} \sum_{k=1}^{a+1} V_k^\sigma W_{(1,0)}^{(k)} + \sum_{k=a+2}^{S+1} V_k^\sigma W_{(0,1)}^{(k)}$$

$$W_{(0,1)} \overset{(law)}{=} \sum_{k=1}^{c} V_k^\sigma W_{(1,0)}^{(k)} + \sum_{k=c+1}^{S+1} V_k^\sigma W_{(0,1)}^{(k)}$$

Summary

- we reduced the study to $W_{(1,0)}$ and $W_{(0,1)}$
- $W_{(1,0)}$ and $W_{(0,1)}$ are solutions of a fixed point system:

$$\begin{cases} W_{(1,0)} \overset{(law)}{=} \sum_{k=1}^{a+1} V_k^{\sigma} W_{(1,0)}^{(k)} + \sum_{k=a+2}^{S+1} V_k^{\sigma} W_{(0,1)}^{(k)} \\ W_{(0,1)} \overset{(law)}{=} \sum_{k=1}^{c} V_k^{\sigma} W_{(1,0)}^{(k)} + \sum_{k=c+1}^{S+1} V_k^{\sigma} W_{(0,1)}^{(k)} \end{cases}$$

What information does the system give us?

Continuous time

We can use the same “tree” argument:

- we reduce the study to $W_{(1,0)}$ and $W_{(0,1)}$
- $W^{CT}_{(1,0)}$ and $W^{CT}_{(0,1)}$ are solutions of the following fixed point system:

$$\begin{cases} W_{(1,0)} \stackrel{(law)}{=} U^m \left(\sum_{k=1}^{a+1} W^{(k)}_{(1,0)} + \sum_{k=a+2}^{S+1} W^{(k)}_{(0,1)} \right) \\ W_{(0,1)} \stackrel{(law)}{=} U^m \left(\sum_{k=1}^{c} W^{(k)}_{(1,0)} + \sum_{k=c+1}^{S+1} W^{(k)}_{(0,1)} \right) \end{cases}$$

where U is uniformly distributed on $[0,1]$.

What information does the system give us?

Contraction

$\mathcal{M}_2(C)$ = space of square integrable probability distributions on $\mathbb{R}$ of mean C equipped with the Wasserstein distance:

$$d_2(\mu,\nu) = \min_{(X_1,X_2)} \left(\mathbb{E}(X_1 - X_2)^2\right)^{1/2}$$

is a complete metrix space.
For all $bC_1 + cC_2 = 0$, we define

$$\phi : \mathcal{M}_2(C_1) \times \mathcal{M}_2(C_2) \to \mathcal{M}_2(C_1) \times \mathcal{M}_2(C_2),$$

by

$$\phi(\mu,\nu) = \left(\mathcal{L}\left(U^m\left(\sum_{k=1}^{a+1} X^{(k)} + \sum_{k=a+2}^{S+1} Y^{(k)}\right)\right), \mathcal{L}\left(U^m\left(\sum_{k=1}^{c} X^{(k)} + \sum_{k=c+1}^{S+1} Y^{(k)}\right)\right)\right)$$

where the $X^{(k)} \sim \mu$ and the $Y^{(k)} \sim \nu$ ($k \geq 1$) are all independent copies of X (resp. Y).

Contraction

We prove that ϕ is $\sqrt{\frac{S+1}{2m+1}}$-Lipschitz:

Proposition:

If $\sigma > \frac{1}{2}$, both fixed point systems have a unique solution on $(\mathcal{M}_2(C_1) \times \mathcal{M}_2(C_2), d_2 \otimes d_2)$.

Remark: the means of the W are explicitely known.

Moments of W^{CT}

$$\begin{cases} W_{(1,0)} \overset{(law)}{=} U^m \left(\sum_{k=1}^{a+1} W_{(1,0)}^{(k)} + \sum_{k=a+2}^{S+1} W_{(0,1)}^{(k)} \right) \\ W_{(0,1)} \overset{(law)}{=} U^m \left(\sum_{k=1}^{c} W_{(1,0)}^{(k)} + \sum_{k=c+1}^{S+1} W_{(0,1)}^{(k)} \right) \end{cases}$$

Proposition Chauvin, Pouyanne, Sahnoun:

The Laplace transforms of $W_{(1,0)}^{CT}$ and $W_{(0,1)}^{CT}$ have a radius of convergence equal to 0: for all $C > 0$, for all large enough p,

$$C^p \leq \frac{\mathbb{E}|W^{CT}|^p}{p!}.$$

Moments of W^{CT}

$$\begin{cases} W_{(1,0)} \overset{(law)}{=} U^m \left(\sum_{k=1}^{a+1} W_{(1,0)}^{(k)} + \sum_{k=a+2}^{S+1} W_{(0,1)}^{(k)} \right) \\ W_{(0,1)} \overset{(law)}{=} U^m \left(\sum_{k=1}^{c} W_{(1,0)}^{(k)} + \sum_{k=c+1}^{S+1} W_{(0,1)}^{(k)} \right) \end{cases}$$

Theorem:

$W_{(1,0)}^{CT}$ and $W_{(0,1)}^{CT}$ admit all their moments and both sequences $\left(\frac{\mathbb{E}|W^{CT}|^p}{p! \ln^p p} \right)^{1/p}$ are bounded. The random variables $W_{(1,0)}^{CT}$ and $W_{(0,1)}^{CT}$ are thus determined by their moments.

Via martingale connexions,

$W_{(1,0)}^{DT}$ and $W_{(0,1)}^{DT}$ are determined by their moments.

Via the expression of $W_{(\alpha,\beta)}$ in terms of $W_{(1,0)}$ and $W_{(0,1)}$:

for all α, β, $W_{(\alpha,\beta)}$ is determined by its moments.

Densities:

The variables W all have a density on $\mathbb{R}$.

Let $\psi_W(t) = \mathbb{E}e^{itW}$ be the Fourier transform of W.

If the Fourier transform is invertible, the W admits a density, namely the inverse of the Fourier transform.

- If ψ_W is L^2, then it's ok. But ψ_W is not L^2...

Densities:

The variables W all have a density on $\mathbb{R}$.

Let $\psi_W(t) = \mathbb{E}e^{itW}$ be the Fourier transform of W.

If the Fourier transform is invertible, the W admits a density, namely the inverse of the Fourier transform.

- If ψ_W is L^2, then it's ok. But ψ_W is not L^2...
- If ψ_W is L^1, then it's ok. But ψ_W is not L^1...

Densities:

The variables W all have a density on $\mathbb{R}$.

Let $\psi_W(t) = \mathbb{E}e^{itW}$ be the Fourier transform of W.

If the Fourier transform is invertible, the W admits a density, namely the inverse of the Fourier transform.

- If ψ_W is L^2, then it's ok. But ψ_W is not L^2...
- If ψ_W is L^1, then it's ok. But ψ_W is not L^1...
- If ψ'_W is L^1 and if $t \mapsto \frac{\psi_W(t)}{t}$ is L^1 then it's ok. Phew!

Fourier analysis

We begin from the fixed point system verified by $W_{(1,0)}^{CT}$ and $W_{(0,1)}^{CT}$:

$$\begin{cases} X \overset{(law)}{=} U^m \left(\sum_{k=1}^{a+1} X^{(k)} + \sum_{k=a+2}^{S+1} Y^{(k)} \right) \\ Y \overset{(law)}{=} U^m \left(\sum_{k=1}^{c} X^{(k)} + \sum_{k=c+1}^{S+1} Y^{(k)} \right) \end{cases}$$

implies

$$\begin{cases} \psi_X(t) = \mathbb{E}\left[\psi_X(U^m t)^{a+1} \psi_Y(U^m t)^b \right] \\ \psi_Y(t) = \mathbb{E}\left[\psi_X(U^m t)^c \psi_Y(U^m t)^{d+1} \right] \end{cases}$$

We can derive the system, we get information on ψ_W and ψ'_W. We apply our ad hoc Fourier inversion theorem and prove the existence of a density for W.

Halfway conclusion

The underlying tree structure of the urn permits

- to reduce the study to very few initial composition vectors,
- to write fixed point systems that are verified by the Ws,
- to apply to these systems usual methods (cf. *smoothing equations* in literature Liu 90's, Durrett and Liggett 83, Biggins and Kyprianou 05, Knape and Neininger 13),
- to prove existence of densities of the Ws,
- and to study the moments of these variables.

Halfway conclusion

The underlying tree structure of the urn permits

- to reduce the study to very few initial composition vectors,
- to write fixed point systems that are verified by the Ws,
- to apply to these systems usual methods (cf. *smoothing equations* in literature LIU 90'S, DURRETT AND LIGGETT 83, BIGGINS AND KYPRIANOU 05, KNAPE AND NEININGER 13),
- to prove existence of densities of the Ws,
- and to study the moments of these variables.

Can we extend the results to d-colour urns?

Definitions

$$U_0 = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} a_{11} & \dots & a_{1d} \\ \vdots & & \vdots \\ a_{d1} & \dots & a_{dd} \end{pmatrix}$$

We assume

- $\forall i, j,\ a_{i,j} \geq 0$ (non-extinction)
- irreducibility
- the urn is **balanced**: $\forall i,\ \sum_{j=1}^{d} a_{ij} = S$

Composition vector:

$$U(n) = \begin{pmatrix} \#\text{ balls of colour 1 at time } n \\ \vdots \\ \#\text{ balls of colour } d \text{ at time } n \end{pmatrix}$$

Large eigenvalues

Let us write the Jordan decompoition of R: $R = \mathtt{diag}(J_1, \dots, J_r)$ where

$$J_i = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_i \end{pmatrix}$$

We choose a Jordan block associated to a large eigenvalue λ, i.e. an eigenvalue such that

$$\sigma = \frac{\mathrm{Re}\lambda}{S} > \frac{1}{2}.$$

- we denote by $\nu + 1$ the size of the Jordan block,
- E is the stable subspace associated to this Jordan block,
- $v \in E$ is a unitary eigenvector associated to λ,
- we denote by π_E the projection on E.

Limit theorems

Theorem POUYANNE**:**

There exists a random complex variable W^{DT} such that

$$\lim_{n\to+\infty} \frac{\pi_E(U(n))}{n^{\lambda/s}\ln^{\nu} n} = \frac{1}{\nu!} W^{DT} v.$$

Embedding in continuous time

Theorem JANSON**:**

There exists a random complex variable W^{CT} such that

$$\lim_{n\to+\infty} \frac{\pi_E(U(t))}{e^{\lambda t} t^{\nu}} = \frac{1}{\nu!} W^{CT} v.$$

Connexions:

$$W^{CT} \overset{(law)}{=} S^{\nu} \xi^{\lambda/s} W^{DT} \quad \text{and} \quad W^{DT} \overset{(law)}{=} S^{-\nu} \xi^{-\lambda/s} W^{CT}.$$

Tree structure

- We reduce the study to $(W_{\mathbf{e_i}})_{i=1..d}$, where $\mathbf{e_i}$ corresponds to the initial composition "one ball of colour i".
- The d variables $W_{\mathbf{e_i}}$ are solutions of a system of d equations in law.
- We consider $\times_{i=1}^{d} \mathcal{M}_2(m_i)$ equipped with the Wasserstein distance: the solution of the system is unique on this space.
- What information can we get on $W_{\mathbf{e_i}}$?
  - density ?
  - moments ?

Tree structure

- We reduce the study to $(W_{\boldsymbol{e_i}})_{i=1..d}$, where $\boldsymbol{e_i}$ corresponds to the initial composition "one ball of colour i".
- The d variables $W_{\boldsymbol{e_i}}$ are solutions of a system of d equations in law.
- We consider $\times_{i=1}^{d} \mathcal{M}_2(m_i)$ equipped with the Wasserstein distance: the solution of the system is unique on this space.
- What information can we get on $W_{\boldsymbol{e_i}}$?
  - density ?
  - moments ?

Thanks for your attention!