CIMPA Summer School 2014
Random structures, analytic and probabilistic approaches
University An Najah, Nablus (Palestine),
NICOLAS POUYANNE

Pólya urn models

— Lecture notes —

Contents

1 Pólya urn: first steps

Let R be a 2-dimensional square matrix having integral entries and U_0 a nonzero 2-dimensional (column) vector with nonnegative integral entries:

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad U_0 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

The *Pólya urn process* $(U_n)_{n\in\mathbb{N}}$ with *replacement matrix* R and initial *composition vector* U_0 is in an imaging way defined as follows. An urn contains red and black balls. At time 0, it contains α red balls and β black ones. A ball is drawn uniformly at random from the urn and its colour is checked. If the drawn ball is red, it is replaced into the urn together with a red balls and b black ones; if the drawn ball is black, it is replaced into the urn as well, together with c red balls and d black ones. One get in this way a new composition vector U_1. The random process $(U_n)_{n\in\mathbb{N}}$ is recursively defined by iterating this mechanism.

In this lecture, the following assumptions on R and U_0 are made:
(i) R est *balanced*, *i.e.* $a + b = c + d \geq 1$;

(ii) R is *"tenable"*, *i.e.* $\Big(b, c \geq 0\Big)$ and $\Big(a \leq -1 \Longrightarrow a|c \text{ and } a|\alpha\Big)$ and $\Big(d \leq -1 \Longrightarrow d|b \text{ and } d|\beta\Big)$.
The balance hypothesis guarantees that the same number of balls $S = a + b = c + d \geq 1$ is added at any step of time. Thanks to the tenability assumption, the process can never extinguish, which means that if a or d is negative, one can always respectively subtract $-a$ or $-d$ balls from the urn.

• In more rigorous terms,

$$(U_n)_{n\in\mathbb{N}} = \begin{pmatrix} U_n^{(1)} \\ U_n^{(2)} \end{pmatrix}_{n\in\mathbb{N}}$$

is the $\mathbb{N}^2 \setminus \{0\}$-valued discrete time Markov chain defined by the transition conditional probabilities

$$\begin{cases} \mathbf{P}\left(U_{n+1} = U_n + \begin{pmatrix} a \\ b \end{pmatrix} \Big| U_n\right) = \dfrac{U_n^{(1)}}{U_n^{(1)} + U_n^{(2)}} \;; \\ \\ \mathbf{P}\left(U_{n+1} = U_n + \begin{pmatrix} c \\ d \end{pmatrix} \Big| U_n\right) = \dfrac{U_n^{(2)}}{U_n^{(1)} + U_n^{(2)}}. \end{cases} \tag{1}$$

The balance assumption implies that $U_n^{(1)} + U_n^{(2)} = \alpha + \beta + nS$ for any n: at any time n, the composition of the urn is random but the total number of balls is deterministic.

• A complete definition of the Pólya urn process as a Markov chain is given by the family

$$\left(\mu_{\begin{pmatrix} x \\ y \end{pmatrix}}\right)_{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{N}^2\setminus\{0\}}$$

of probability measures on $\mathbb{N}^2 \setminus \{0\}$ defined by:

$$\forall \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{N}^2 \setminus \{0\}, \qquad \mu_{\begin{pmatrix} x \\ y \end{pmatrix}} = \frac{x}{x+y}\delta_{\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}} + \frac{y}{x+y}\delta_{\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}},$$

where δ_P denotes the Dirac measure at P. Notice that the tenability assumption guarantees that $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$ and $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}$ belong to $\mathbb{N}^2 \setminus \{0\}$ as soon as $\begin{pmatrix} x \\ y \end{pmatrix}$ does.

[Generalisation to any finite number of colour, to random replacement matrices. In the present lecture, we will restrict ourselves to non random replacement matrices.]

Notations (spectral decomposition of R)
Thanks to the balance assumption, S is an eigenvalue of ${}^tR$. By elementary considerations à la Perron-Frobenius, the second eigenvalue $m := a - c = d - b$ of ${}^tR$ is less than or equal to S. We denote

$$\sigma = m/S \leq 1$$

(note that σ may be negative).
When $(b,c)\neq(0,0)$, let

$$v_1=\frac{S}{b+c}\left(\begin{array}{c}c\\ b\end{array}\right)\quad\text{and}\quad v_2=\frac{S}{b+c}\left(\begin{array}{c}1\\ -1\end{array}\right).$$

The vectors v_1 and v_2 are eigenvectors of ${}^tR$, respectively associated with the eigenvalues S and m. The dual basis (u_1,u_2) of linear eigenforms is given by the formulae

$$u_1(x,y)=\frac{1}{S}(x+y)\quad\text{and}\quad u_2(x,y)=\frac{1}{S}(bx-cy).$$

These vectors and linear forms will be useful later on in the lecture.

Note that in dimension larger than 3, the matrix R is not necessarily diagonalizable, even on $\mathbb{C}$. This fact leads to some intricacy in the statement of the results but in a first approach, one can assume that R is diagonalizable.

2 The approach in analytic combinatorics

The approach by analytic combinatorics is due to Philippe Flajolet and his co-authors Philippe Dumas, Joaquim Gabarró, Helmut Pekari and Vincent Puyhaubert in the 2000's. There are two founding articles, namely [4] et [3].

The very first idea consists in coding the urn composition by a sequence $(W_n)_{n\in\mathbb{N}}$ of finite words written in the 2-letter alphabet $\{\mathtt{r},\mathtt{b}\}$ ($\mathtt{r}$ for *red*, $\mathtt{b}$ for *black*). The initial composition is coded by

$$W_0=\mathtt{rr}\ldots\mathtt{rbb}\ldots\mathtt{b}=\mathtt{r}^\alpha\mathtt{b}^\beta.$$

Drawing a ball in the urn amounts to choosing a letter in the word uniformly at random. When the chosen letter is an $\mathtt{r}$, it is replaced in the world by the subword $\mathtt{r}^{a+1}\mathtt{b}^b$; when the chosen letter is a $\mathtt{b}$, it is replaced by $\mathtt{r}^c\mathtt{b}^{d+1}$. Thus, the successive drawings give rise to a sequence of random words

$$W_0,W_1,W_2\ldots$$

Of course, at any time n, the composition vector U_n can be recovered by counting the number of $\mathtt{r}$'s and the number of $\mathtt{b}$'s in the word W_n.

Definition 1 (Histories of the process)
When n is a natural number, when $\binom{u_0}{v_0},\binom{u}{v}\in\mathbb{N}^2\setminus\{0\}$, a history of length n leading from $\binom{u_0}{v_0}$ to $\binom{u}{v}$ *is a sequence of words $W_0=\mathtt{r}^{u_0}\mathtt{b}^{v_0},W_1,W_2,\ldots,W_n$ produced in that way, for which W_n contains exactly u letters $\mathtt{r}$ et v letters $\mathtt{b}$.*

Of course, with this coding, because of the balance hypothesis, the word W_n always contains u_0+v_0+nS letters, whatever its history is. The key object of Flajolet's method is the number of these histories: denote by

$$\boxed{H_n\begin{pmatrix} u_0 & u \\ v_0 & v\end{pmatrix}}$$

the number of histories of length n leading from $\binom{u_0}{v_0}$ to $\binom{u}{v}$.

Exercise 1. When $R = \begin{pmatrix} 0 & 3 \\ 2 & 1\end{pmatrix}$, code and count all histories of length 2 leading from $\binom{2}{0}$ to $\binom{4}{4}$.

[One possible solution: start from $W_0 = \mathtt{r}^2$. One can draw a tree of all possibilities: $W_1 \in \{\mathtt{rb}^3\mathtt{r}, \mathtt{r}^2\mathtt{b}^3\}$, then $W_2 \in \{\mathtt{rb}^6\mathtt{r}, \mathtt{r}^3\mathtt{b}^4\mathtt{r}, \mathtt{rbr}^2\mathtt{b}^3\mathtt{r}, \mathtt{rb}^2\mathtt{r}^2\mathtt{b}^2\mathtt{r}, \mathtt{rb}^3\mathtt{rb}^3\}$ or $W_2 \in \{\mathtt{rb}^3\mathtt{rb}^3, \mathtt{r}^2\mathtt{b}^6, \mathtt{r}^4\mathtt{b}^4, \mathtt{r}^2\mathtt{br}^2\mathtt{b}^3, \mathtt{r}^2\mathtt{b}^2\mathtt{r}^2\mathtt{b}^2\}$. Amongst the ten histories of length 2, six of them lead to $\binom{4}{4}$ and four lead to $\binom{2}{6}$: starting from two red balls, the probability that the urn contains four red balls and four black ones after two drawings is $3/5$.
Beware: in the example, the configuration $\mathtt{rb}^3\mathtt{rb}^3$ is reached by two different histories. We count histories, not the different word that are potentially obtained.]

Exercise 2 (this urn is Pólya's original one in his article published in 1930)**.** Whenever $R = SI_2$, compute all numbers H_n, $n \geq 0$.
[This is elementary enumerative combinatorics. Make the picture of a path in $\mathbb{N}^2$ and count the histories that follow each of these paths. For any $(p, q) \in \mathbb{N}^2$ such that $p + q = n$, one gets

$$\begin{aligned} H_n\begin{pmatrix} \alpha & \alpha + pS \\ \beta & \beta + qS\end{pmatrix} &= \binom{n}{p}\alpha(\alpha+S)\dots(\alpha+(p-1)S)\beta(\beta+S)\dots(\beta+(q-1)S) \\ &= n!S^n\binom{\frac{\alpha}{S}+p-1}{p}\binom{\frac{\beta}{S}+q-1}{q}\ ; \end{aligned}$$

all others H_n vanish.]

Exercise 3. For any urn, if $N = \alpha + \beta$, show that the total number of histories of length n starting from $\binom{\alpha}{\beta}$ equals $N(N+S)(N+2S)\dots(N+(n-1)S) = n!S^n\binom{\frac{N}{S}+n-1}{n}$.

Generating series (or functions) are central tools in analytic combinatorics. In the case of 2-colour urns, the relevant one is the trivariate generating series of histories: the variable x counts the final number of red balls, the variable y counts the final number of black ones while the variable z counts the length of the history. Thus, the replacement matrix R being given, denote

$$\boxed{H\left(x, y, z \,\middle|\, \begin{matrix} u_0 \\ v_0\end{matrix}\right) = \sum_{u,v,n\in\mathbb{N}} H_n\begin{pmatrix} u_0 & u \\ v_0 & v\end{pmatrix} x^u y^v \frac{z^n}{n!}.}$$

Exercise 4. For any urn (*i.e.* for any R), $H\left(1, 1, z \,\middle|\, \begin{matrix} u_0 \\ v_0\end{matrix}\right) = \left(\frac{1}{1-Sz}\right)^{\frac{u_0+v_0}{S}}$.

Exercise 5. For the original urn ($R = SI_2$),

$$H\left(x, y, z \,\middle|\, \begin{matrix} u_0 \\ v_0\end{matrix}\right) = \frac{x^{u_0}y^{v_0}}{(1-Szx^S)^{\frac{u_0}{S}}(1-Szy^S)^{\frac{v_0}{S}}}.$$

[Computations on multivariate power series, based on the formula $\frac{1}{(1-X)^N} = \sum_{n\geq 0} \binom{N+n-1}{n} X^n$.]

[Commentary on papers by P. Flajolet *et al.*: pointing an object amounts to make a partial derivative on the generating series; proceeding to a replacement amounts to multiply the series by some appropriate monomial. Such considerations lead to the following “Basic isomorphism”, stated and proven in [3].]

Theorem 1 (Flajolet, Dumas, Puyhaubert, 2006)
Let x and y be complex numberes such that $xy \neq 0$. Let $X(t)$ and $Y(t)$ be the solutions of the Cauchy Problem (formal version or analytic version)

$$\begin{cases} \dfrac{dX}{dt} = X^{a+1}Y^b \\[2ex] \dfrac{dY}{dt} = X^cY^{d+1} \\[2ex] X(0) = x, Y(0) = y. \end{cases} \tag{2}$$

Then, for any initial composition (u_0, v_0), for any z in some small enough neighbour of the origin (analytic version),

$$H\left(x, y, z \,\middle|\, \begin{array}{c} u_0 \\ v_0 \end{array}\right) = X(z)^{u_0}Y(z)^{v_0}.$$

Example 1. Back to the original Pólya urn for which $R = SI_2$: the differential system writes $X' = X^{S+1}, Y' = Y^{S+1}$and can be solved. The solution of the Cauchy Problem is $X(t) = x(1 - Stx^S)^{-1/S}$, $Y(t) = y(1 - Sty^S)^{-1/S}$. Theorem 1 provides a second proof of exercice 5.

Proof of Theorem 1. Consider the following differential operator on 2-variable functions:

$$\mathcal{D} = x^{a+1}y^b\frac{\partial}{\partial x} + x^cy^{d+1}\frac{\partial}{\partial y}.$$

The action of $\mathcal{D}$ on monomials is related to urn histories *via* the formula

$$\begin{aligned} \mathcal{D}\left(x^{u_0}y^{v_0}\right) &= u_0x^{a+u_0}y^{b+v_0} + v_0x^{c+u_0}y^{d+v_0} \\ &= H_1\left(\begin{array}{cc} u_0 & u_0+a \\ v_0 & v_0+b \end{array}\right)x^{a+u_0}y^{b+v_0} + H_1\left(\begin{array}{cc} u_0 & u_0+c \\ v_0 & v_0+d \end{array}\right)x^{c+u_0}y^{d+v_0} \end{aligned}$$

which can also be written

$$\mathcal{D}\left(x^{u_0}y^{v_0}\right) = \sum_{u,v\geq 0} H_1\left(\begin{array}{cc} u_0 & u \\ v_0 & v \end{array}\right)x^uy^v$$

where only two terms of the infinite sum are nonzero. This implies by induction that for any $n \in \mathbb{N}$,

$$\mathcal{D}^n\left(x^{u_0}y^{v_0}\right) = \sum_{u,v\geq 0} H_n\left(\begin{array}{cc} u_0 & u \\ v_0 & v \end{array}\right)x^uy^v. \tag{3}$$

[Notice that the Markov property of the urn process is expressed in this induction.] Besides, if (X,Y) is a solution of the differential system $X' = X^{a+1}Y^b$, $Y' = X^cY^{d+1}$, then

$$\begin{aligned}\frac{d}{dt}\left(X(t)^{u_0}Y(t)^{v_0}\right) &= u_0X(t)^{a+u_0}Y(t)^{b+v_0} + v_0X(t)^{c+u_0}Y(t)^{d+v_0}\\ &= \mathcal{D}\left(x^{u_0}y^{v_0}\right)\Big|_{\substack{x=X(t)\\ y=Y(t)}}\end{aligned}$$

which extends to an analogous formula for the n-th derivative. Gathering these results leads successively to

$$\begin{aligned}H\left(X(t),Y(t),z\,\middle|\,\begin{array}{c}u_0\\ v_0\end{array}\right) &= \sum_{n\geq 0}\mathcal{D}^n\left(x^{u_0}y^{v_0}\right)\Big|_{\substack{x=X(t)\\ y=Y(t)}}\frac{z^n}{n!}\\ &= \sum_{n\geq 0}\frac{d^n}{dt^n}\left(X(t)^{u_0}Y(t)^{v_0}\right)\frac{z^n}{n!}.\end{aligned}$$

Thanks to Taylor Formula at the origin (analytic or formal version), one concludes by

$$H\left(X(t),Y(t),z\,\middle|\,\begin{array}{c}u_0\\ v_0\end{array}\right) = X(t+z)^{u_0}Y(t+z)^{v_0}.$$

The final result follows taking the value at the origin ($t=0$). ■

When the differential system can be solved, applying Theorem 1 leads to a close form of the H function. When this is possible, one gets very accurate probabilistic consequences on the distribution of the composition of the urn at finite time, or on the asymptotics of the process as well. We give hereunder a couple of examples, essentially drawn from [4] and [3].

Remark. 1- One gets immediately from Theorem 1 that

$$H\left(x,y,z\,\middle|\,\begin{array}{c}u_0\\ v_0\end{array}\right) = H\left(x,y,z\,\middle|\,\begin{array}{c}1\\ 0\end{array}\right)^{u_0} H\left(x,y,z\,\middle|\,\begin{array}{c}0\\ 1\end{array}\right)^{v_0}.$$

This formula evokes some (combinatoric) convolution property. It has to be related to the branching property of the continuous time corresponding urn process, that leads to a similar equation on the Fourier transforms of large urns limit laws. See [2]. A direct link between both properties remains an open question.

Example 2. Take the urn having $\begin{pmatrix}0&1\\1&0\end{pmatrix}$ as replacement matrix.. [Friedmann's urn. Talk about the propaganda campaign used by P. Flajolet.] The Cauchy Problem writes

$$\begin{cases}X' = XY\\ Y' = XY\\ X(0)=x, Y(0)=y\end{cases}$$

and can be easily solved. One finds

$$H\left(x,y,z\,\middle|\,\begin{array}{c}u_0\\ v_0\end{array}\right) = \left(\frac{x(x-y)}{x-ye^{z(x-y)}}\right)^{u_0}\left(\frac{y(y-x)}{y-xe^{z(y-x)}}\right)^{v_0}.$$

For example, when one starts with a sole red ball, the probability generating function of the number of red balls is

$$\mathbf{E}\left(x^{U_n^{(1)}}\right) = \left[\frac{z^n}{n!}\right]\sum_{n,k}\mathbf{P}(U_n^{(1)}=k)x^k\frac{z^n}{n!} = [z^n]\,H\left(x,1,z\,\middle|\,\begin{matrix}1\\0\end{matrix}\right),$$

since the total number of histories of length n starting from one red ball is $n!$ (see Exercise 3). Using the explicit expression of H, one gets

$$\mathbf{E}\left(x^{U_n^{(1)}}\right) = [z^n]\,\frac{x(x-1)}{x-e^{z(x-1)}}.$$

This function of the z-variable has a simple pôle at $z=\frac{\log x}{x-1}$ as unique singularity. Since this function of the x-variable is analytic at 1, singularity analysis shows that one can apply Hwang's Quasi-power Theorem: the mean and the variance of $U_n^{(1)}$ are both asymptotically proportional to n, and the number of red balls at time n (*i.e.* the random variable $U_n^{(1)}$) satisfies a Law of Large Numbers and a Central Limit Theorem as well (Gaussian distribution).

Exemple 3. This example is the central one in [4]. It deals with the urn process that models the leaves of a 2-3–tree, which is an important search tree algorithm. Its replacement matrix is $\begin{pmatrix}-2&3\\4&-3\end{pmatrix}$. Here, the Cauchy Problem writes

$$\left\{\begin{array}{l}X'=X^{-1}Y^3\\Y'=X^4Y^{-2}\\X(0)=x, Y(0)=y.\end{array}\right. \tag{4}$$

Pose $Z=X^2$; one gets successively $Z'=2Y^3$ and $Z''=6Z^2$. Multiply first the latter equation by Z' then integrate. This leads to show that Z is necessarily a solution of the Cauchy Problem

$$\left\{\begin{array}{l}Z'^2=4Z^3-g_3\\Z(0)=x^2\\Z'(0)=2y^3\end{array}\right. \tag{5}$$

where $g_3=4(x^6-y^6)$. This equation is solved using the famous and beautiful theory of elliptic functions. Quickly said, let $\wp(z)=\wp(z;0,-4)$ be the elliptic Weiestrass function, associated to the (so-called) invariants $g_2=0$ et $g_3=-4$: if one denotes

$$\omega=\frac{1}{2}B\left(\frac{1}{6},\frac{1}{3}\right)$$

(Euler Beta function) and if Λ denotes the hexagonal lattice

$$\Lambda=\omega\left(e^{i\pi/6}\mathbb{Z}+e^{-i\pi/6}\mathbb{Z}\right),$$

then $\wp$ is the meromorphic function of the complex plane defined on the complementary of the lattice Λ by

$$\wp(z)=\frac{1}{z^2}+\sum_{\lambda\in\Lambda\setminus(0)}\left[\frac{1}{(z+\lambda)^2}-\frac{1}{\lambda^2}\right].$$

The function $\wp$ has a double pôle at any point of Λ and is Λ-periodic (such complex functions are called *doubly periodic*). Modulo Λ, the zeroes of $\wp$ are exactly $\omega/3$ and $2\omega/3$. The theory of holomorphic functions shows that $\wp$ is a solution of (5). There is another way to describe this famous $\wp$: it is the inverse of the elliptic integral that underpins Equation (5). More precisely, if z and w are complex numbers one gets the equivalence

$$\wp(z) = w \Longleftrightarrow z = \int_{[w,\infty]} \frac{d\zeta}{2\sqrt{\zeta^3+1}},$$

where the symbol $[w, \infty]$ denotes any half-line having w as origin, and that do not contain any root of the polynomial $\zeta^3 + 1$ (the square root denotes here the determination defined by the split plane associated to this half-line). Note for example that the Weierstraß functions, even if they have been defined in the 1860's, are objects of recent interest because they give parametrizations of smooth plane cubics that are central in modern cryptography; here, the pair $(\wp, \wp')$ is a parametrization of the curve $Y^2 = 4X^3 + 4$.

Thus, the solutions of the differential system (4) can be expressed by means of elliptic functions on the hexagonal lattice. Take for instance an urn containing initially 2 red balls and no black ones. Let p_n be the probability that all balls are black at time n. In terms of H functions, this number writes

$$p_n = \frac{1}{n+1}[z^n]H\left(0, 1, z \,\middle|\, \begin{matrix} 2 \\ 0 \end{matrix}\right).$$

By solving the Cauchy Problem, one shows that

$$H\left(0, 1, z \,\middle|\, \begin{matrix} 2 \\ 0 \end{matrix}\right) = \wp\left(z - \frac{\omega}{3}\right).$$

One concludes by means of singularity analysis: check the pôles of $\wp$ and give an asymptotics of p_n as powers of $3/w \sim 0,7132$.

Remarks. 1- The monomial differential system (2) has a simple first integral: if X and Y are solutions, then $1/X^m - 1/Y^m$ is a (locally) constant function. Writing by this means Y as a function of X and reporting in the system, one gets the inverse abelian integrals described above. All "elliptic urns", *i.e.* all urns for which these abelian integrals are related to curves of genus 1 (*elliptic* curves) are classified in [4].

2- In the case of more than 3 colours, Theorem 1 remains valid. Nevertheless, the efficiency and the preciseness of the beautiful analytic method for urns is darkened by a theoretical obstruction: the monomial differential system is, in general, not integrable in dimension more than 3 (this is a difficult result of differential algebra and algebraic geometry, see final comments and note 11 in [3]).

3 The probabilistic approach

We first adopt two experimental approaches, where the effect of the famous phase transition on urns appears. Then, the results on urns asymptotics are stated. Finally, the methods of proving these asymptotics are evoked.

3.1 Introduction: an experimental computational approach

3.1.1 Distributions

As a first approach, for any urn, consider the probability generating function of the number of (say) red balls at time n, starting from the initial composition $\binom{u_0}{v_0}$:

$$p_n\left(x \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right) := \sum_{u\geq 0} \mathbf{P}_{(u_0,v_0)}\left(U_n^{(1)} = u\right) x^u = \mathbf{E}_{(u_0,v_0)}\left(x^{U_n^{(1)}}\right).$$

Since the total number af balls at time n is deterministic, this probability generating function describes the whole distribution of the urn composition at time n. This probability generating function can be expressed by means of H functions: denote by

$$H_n\left(x,y \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right) := \sum_{u,v\geq 0} H_n\left(\begin{matrix} u_0 & u \\ v_0 & v \end{matrix}\right) x^u y^v = n!\,[z^n]\,H\left(x,y,z \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$$

the generating series (it is a 2-variable polynomial) of histories of length n starting from $\binom{u_0}{v_0}$. Then, $H_n\left(1,1 \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$ is the total number of histories of length n starting from $\binom{u_0}{v_0}$ (see Exercise 3) and

$$p_n\left(x \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right) = \frac{H_n\left(x,1 \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)}{H_n\left(1,1 \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)}.$$

Thus, it suffices to compute $H_n\left(x,y \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$, or even $H_n\left(x,1 \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$ to get p_n. But, as shown in the proof of Theorem 1, the bivariate function $H_n\left(x,y \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$ satisfies Equation (3), namely

$$H_n\left(x,y \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right) = \mathcal{D}^n\left(x^{u_0}y^{v_0}\right).$$

As a matter of consequence, by means of computer algebra, starting from the monomial $x^{u_0}y^{v_0}$, it suffices to make an iteration of the operator $\mathcal{D}$ to get a symbolic expression of the entire function $H_n\left(x,y \,\middle|\, \begin{matrix} u_0 \\ v_0 \end{matrix}\right)$. The probability generating function p_n is then extracted by substitutions ($y = 1$ and $x = 1$). By this means, the distribution of red balls at given times can be graphically represented. This is done below for three particular urns and initial compositions.

3.1.2 Simulations of trajectories

Another approach consists in simulating the random successive compositions of an urn. One can by this means have a representation of *trajectories* of the composition vector, namely $\{(n, U_n),\ n = 0, 1, 2\ldots\}$

for different random drawings. Taking only the first coordinate of U_n leads to trajectories of the number of red balls, namely
$$\left\{ \left(n, U_n^{(1)}\right),\ n = 0, 1, 2 \dots \right\}.$$
This is done below for three particular urns and initial compositions.

3.1.3 Three urns

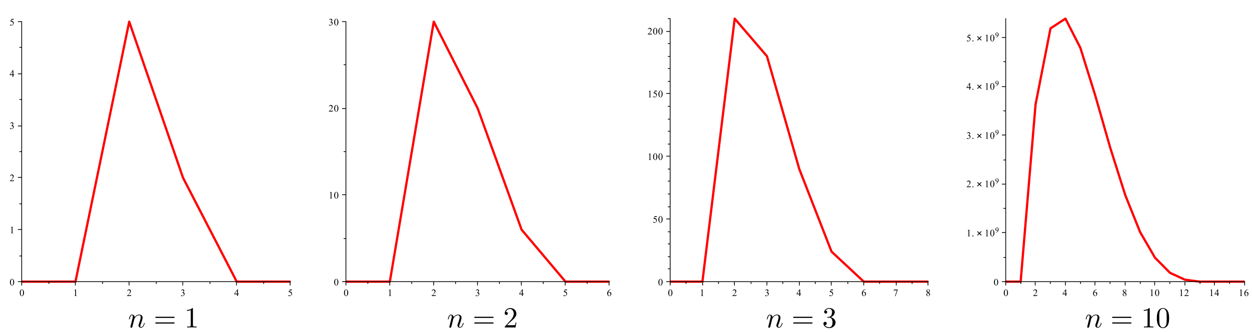
Consider the urn processes having respectively

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 1 & 12 \\ 11 & 2 \end{pmatrix} \text{ and } \quad R_2 = \begin{pmatrix} 12 & 1 \\ 2 & 11 \end{pmatrix}$$

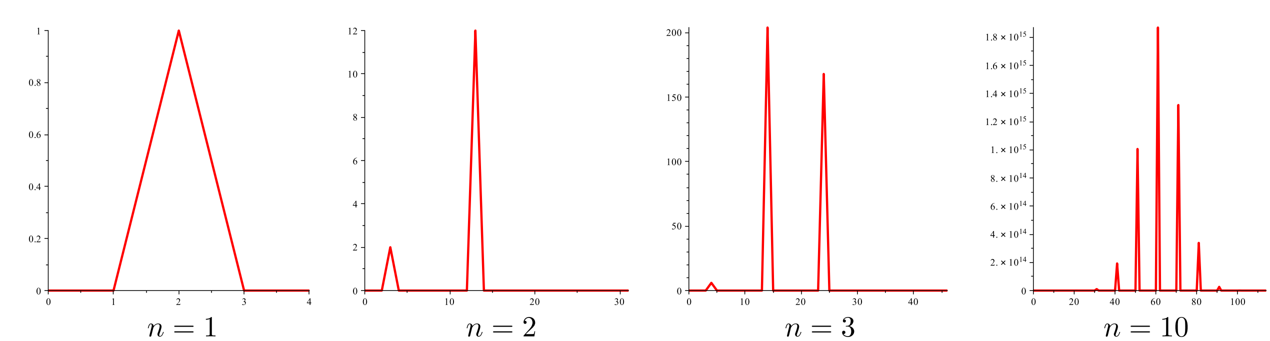
as matrix transitions. The drawings presented hereunder are made taking respectively $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$, $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ as initial composition. All graphics are different representations of the number of red balls contained in the urn.

1- Very first histograms

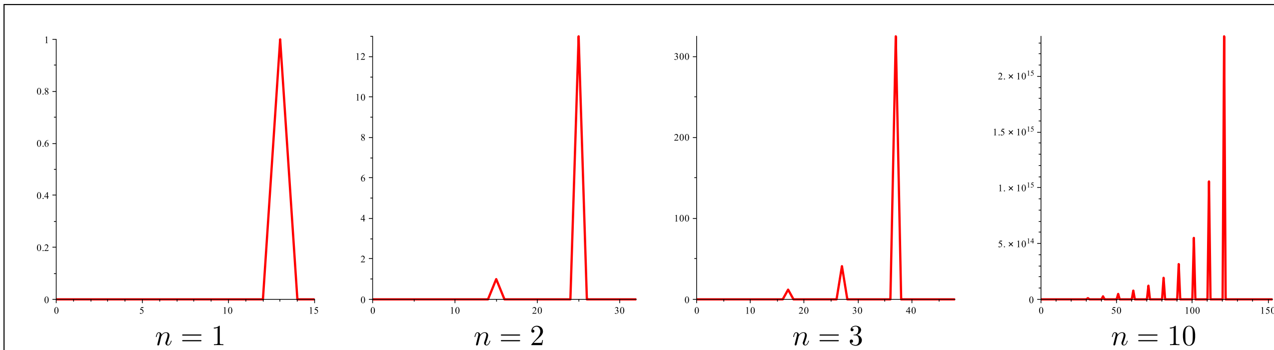
Any picture is made for a given number n of drawings in the urn. On the x-axis, the number of red balls in the urn after n drawings. On the y-axis, the number of histories of length n starting from the initial composition. Points at integer abscissae are related by line segments.



Original Pólya urn I_2 after n drawings with initial composition $(2, 5)$:
number of histories with respect to the number of red balls.



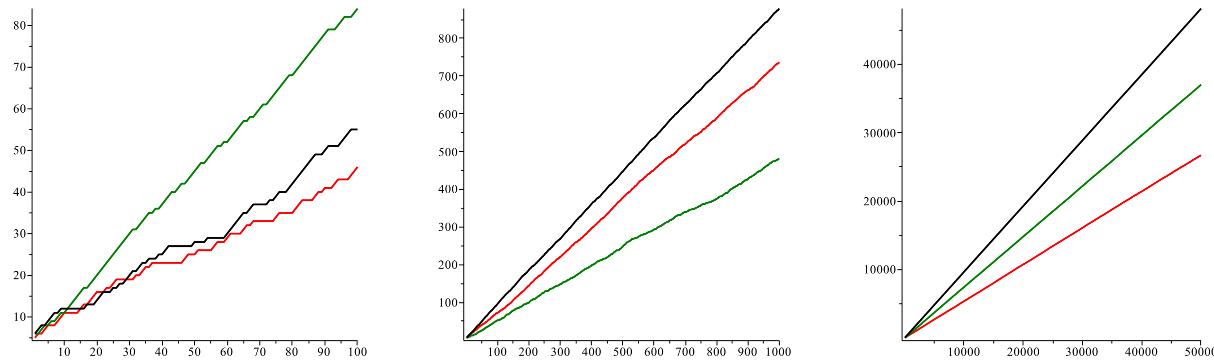
Urn R_1 after n drawings with initial composition $(1, 0)$:
number of histories with respect to the number of red balls.



Urn R_2 after n drawings with initial composition $(1, 0)$:
number of histories with respect to the number of red balls.

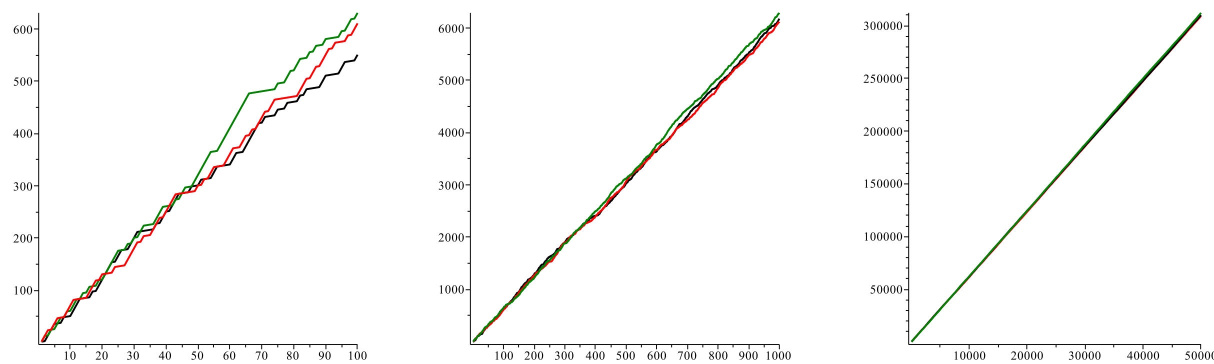
2- Very first trajectories

For a given urn, we draw three different trajectories, corresponding to three random sequences of drawings in the urn. On the x-axis, the number of drawings (discrete time); the maximal number of drawings is successively $N = 100,\ 1000,\ 50000$. On the y-axis, the number of red balls in the urn.



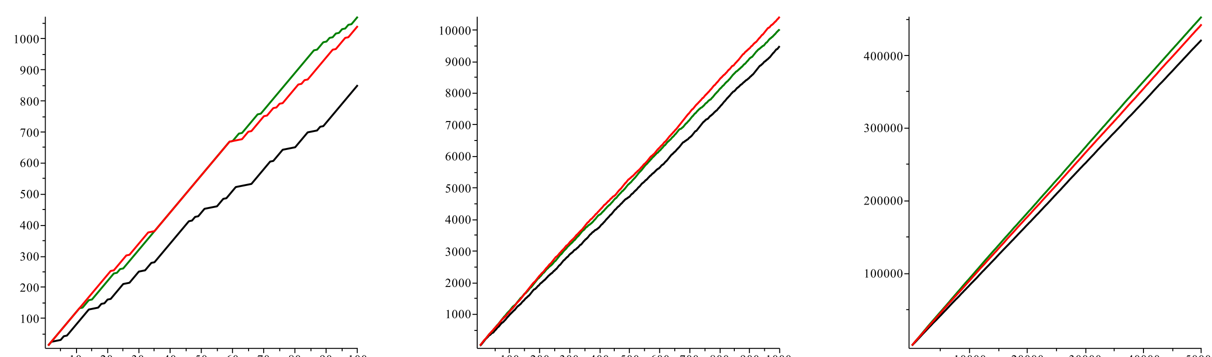
$N = 100$ $N = 1000$ $N = 50000$

Red balls in three sequences of N drawings in an original Pólya urn I_2, initial composition $(2, 5)$



$N = 100$ $N = 1000$ $N = 50000$

Red balls in three sequences of N drawings in an urn R_1, initial composition $(1, 0)$



$N = 100$ $N = 1000$ $N = 50000$

Red balls in three sequences of N drawings in an urn R_2, initial composition $(1, 0)$

3.2 Asymptotics of the composition vector, phase transition, figures

The composition vector U_n of a Pólya urn process has different asymptotics régimes when n tends to infinity, depending on the spectral decomposition of the replacement matrix R. In this section, we state, comment and illustrate these asymptotic results. All of them can be extended in higher dimension (any finite number of colours). Methods of proofs are introduced in Section 3.3.

Take a two-colour Pólya urn with replacement matrix $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and initial composition vector $U_0 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. We adopt the notations of Section 1, especially the *balance* $S = a + b = c + d$, the second eigenvalue $m = a - c = d - b$, the ${}^tR$-eigenvectors v_1, v_2 and, above all, the ratio

$$\sigma = m/S.$$

The original Pólya urn holds a particular place; its asymptotics is described in Theorem 2. The famous phase transition occurs at $\sigma = 1/2$. When $\sigma \leq 1/2$, the urn is said *small* and its composition vector satisfies a central limit theorem as stated in Theorem 3. When $\sigma \in]\frac{1}{2}, 1[$, the urn is said *large* and the centered composition vector admits, after a suitable normalisation, an almost sure random limit; this result is made precise in Theorem 4.

Theorem 2 (Pólya original urn)
Suppose that the urn is Pólya's original one, i.e. *that* $R = I_2$. *Then, as* n *tends to infinity,*

$$\frac{U_n}{Sn} \underset{n\to\infty}{\longrightarrow} D$$

almost surely and in any L^p, $p \geq 1$, *where* D *is a Dirichlet distributed* 2*-dimensional random vector with parameter* $\left(\frac{\alpha}{S}, \frac{\beta}{S}\right)$.

If u and v are two positive real numbers, a 2-dimensional Dirichlet distribution with parameter (u, v) is the measure on the simplex $\Sigma = \left\{(x, y) \in [0, 1]^2,\ x + y = 1\right\}$ that admits the function

$$(x, y) \mapsto \frac{\Gamma(u + v)}{\Gamma(u)\Gamma(v)} x^{u-1} y^{v-1}$$

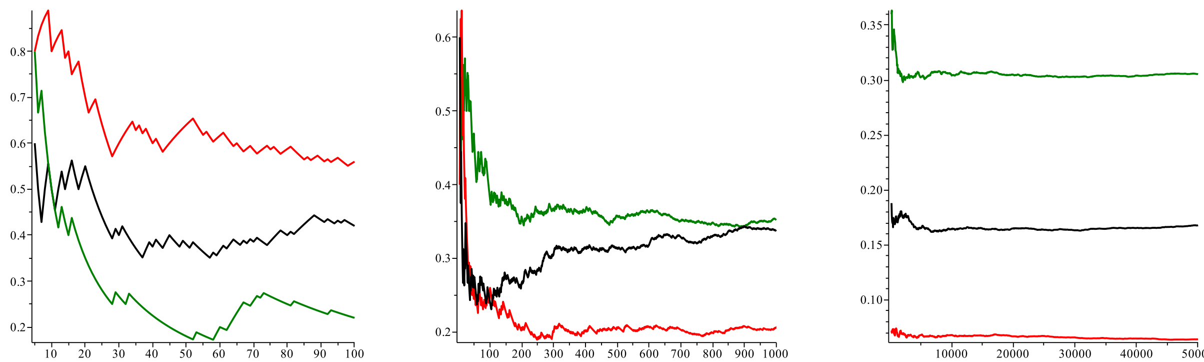
as density with regard to Lebesgue measure on Σ. In other words, if D is a Dirichlet distributed 2-dimensional random vector with parameter (u, v), then for any continuous function f on Σ,

$$\mathbf{E}\left(f(D)\right) = \frac{\Gamma(u + v)}{\Gamma(u)\Gamma(v)} \int_0^1 f(x, 1 - x) x^{u-1} \left(1 - x\right)^{v-1} dx.$$

In particular, if $D = (X, Y)$, then the *marginals* X and Y are (mutually dependent) Beta distributed random variables, X having parameter (u, v) and Y having parameter (v, u).

Firstly, the convergence is almost sure, which means that, with probability 1, a sequence of random drawings leads to the convergence of the vector U_n/Sn to some vector in the simplex Σ. Secondly,

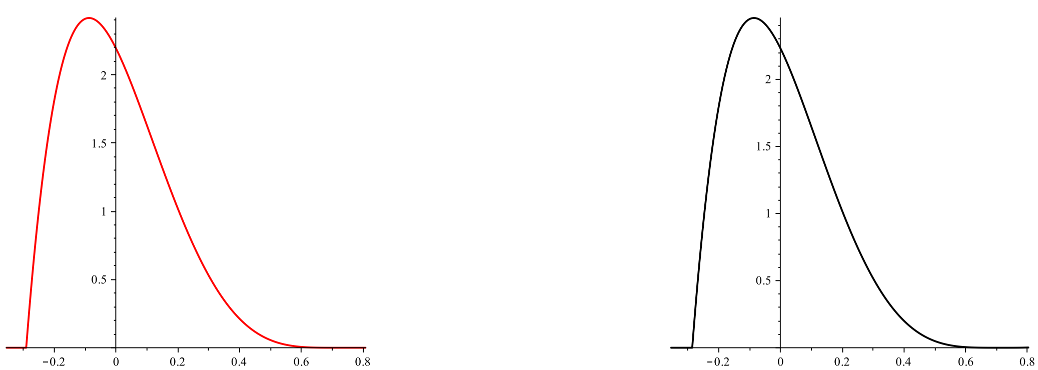
the limit D is random, which means that two different sequences of random drawings converge with probability 1 to two different vectors of Σ.
This almost sure random limit can be visualised on the above simulations: any trajectory gives rise to a (trembled) line, but the three slopes are different. We give hereunder new figures, where three normalised trajectories are represented, showing three different limits: on the x-axis, the number n of drawings up to $N = 100$, 1000 or 50000. On the y-axis, the normalised number of red balls $\frac{1}{n}U_n^{(1)}$.



$N = 100$ $\quad$ $N = 1000$ $\quad$ $N = 50000$

$\frac{1}{n}U_n^{(1)}$ in three sequences of N drawings in an original Pólya urn I_2, initial composition $(2, 5)$

One can also visualise the Beta distributed limit of the normalised number of red balls. Hereunder, the figure on the left represent the (exact) distribution of the normalised number of red balls in the urn after $n = 200$ drawings. On the x-axis, $\frac{1}{n}\left(U_n^{(1)} - \mathbf{E}U_n^{(1)}\right)$. On the y-axis, the probability; it has been computed from the probability generating function p_n introduced above. The figure on the right represents the graph of the density of the centered Beta distribution with parameter $(2, 5)$, namely the function $x \mapsto \frac{1}{B(2,5)}\,(x-\mu)^1\,(1-x+\mu)^4$ where $\mu = B(3,5)/B(2,5) = 2/7$ is the expectation.



$n = 200$

Normalised distribution of the number of red balls in an original Pólya urn I_2, initial composition $(2, 5)$

Density of a centered Beta $(2, 5)$ distribution

Theorem 3 (Small urns)
Suppose that the urn is small, *which means that* $\sigma < 1/2$. *Then as* n *tends to infinity,*

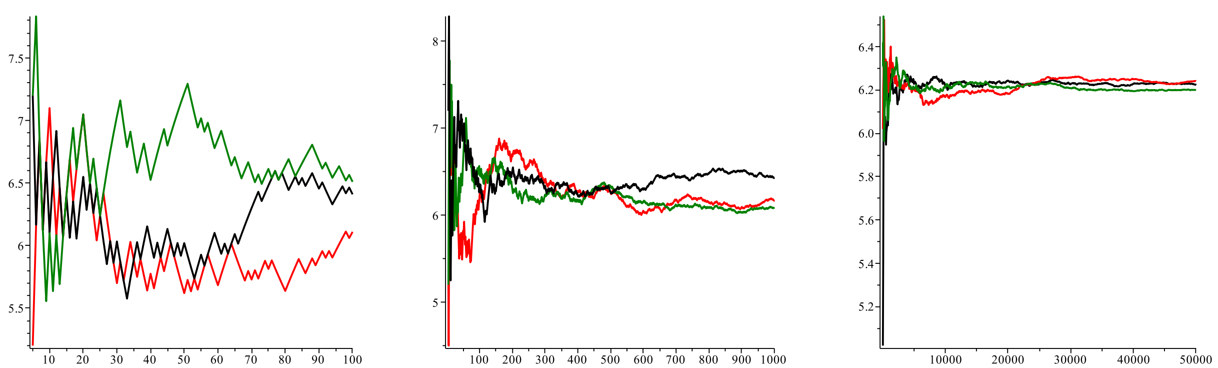
(i) $\dfrac{U_n}{n}$ *converges to* v_1, *almost surely and in any* L^p, $p \geq 1$;

(ii) assume further that R *is not triangular,* i.e. *that* $bc \neq 0$. *Then,* $\dfrac{U_n - nv_1}{\sqrt{n}}$ *converges in distribution to a centered gaussian vector with covariance matrix*

$$\frac{1}{1-2\sigma}\frac{bcm^2}{(b+c)^2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

[When $\sigma = 1/2$, one says also that the urn is small. In this case, assertion (i) holds as well whereas, when R is not triangular, assertion (ii) must be replaced by: $\frac{U_n - nv_1}{\sqrt{n \log n}}$ *converges in distribution to a centered Gaussian vector with covariance matrix* $\frac{1}{4}bc\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$.]

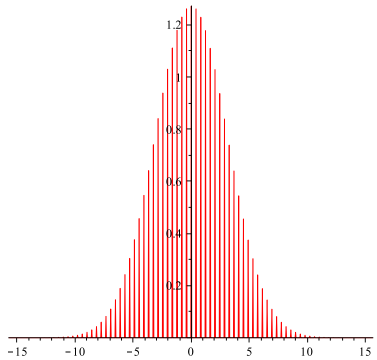
Here, the convergence of U_n/n is almost sure again, but the limit is deterministic: with probability 1, a sequence of random drawings leads to the convergence of the vector U_n/n, but the limit is now always the same one (namely, v_1). This phenomenon can be visualised on the trajectories for the urn R_1: the three asymptotic slopes are identical. When the normalised trajectories are drawn, one gets the following pictures. Here again, on the x-axis, the number n of drawings up to $N = 100$, 1000 or 50000; on the y-axis, the normalised number of red balls $\frac{1}{n}U_n^{(1)}$.



$N = 100$ $N = 1000$ $N = 50000$

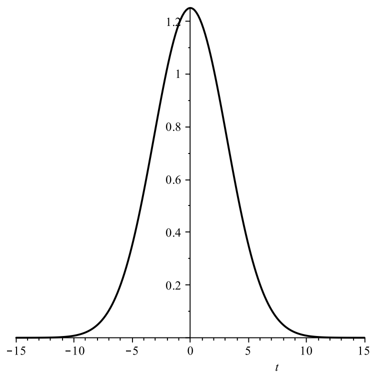
$\frac{1}{n}U_n^{(1)}$ in three sequences of N drawings in a small urn R_1, initial composition $(1, 0)$

The convergence in distribution stated in (ii) is of a radically different nature. It means that the distribution at finite time n converges to some given distribution when n tends to infinity. The limit distribution is here normal. As before, for the R_1-urn, with the help of the probability generating function, the (exact) distribution of the number $\frac{1}{\sqrt{n}}\left(U_n^{(1)} - \mathbf{E}U_n^{(1)}\right)$ is drawn on the leftside figure for $n = 600$. On the right, the graph of the density of the centered normal distribution with variance $\frac{1}{1-2\sigma}\frac{bcm^2}{(b+c)^2} = \frac{5200}{529}$.



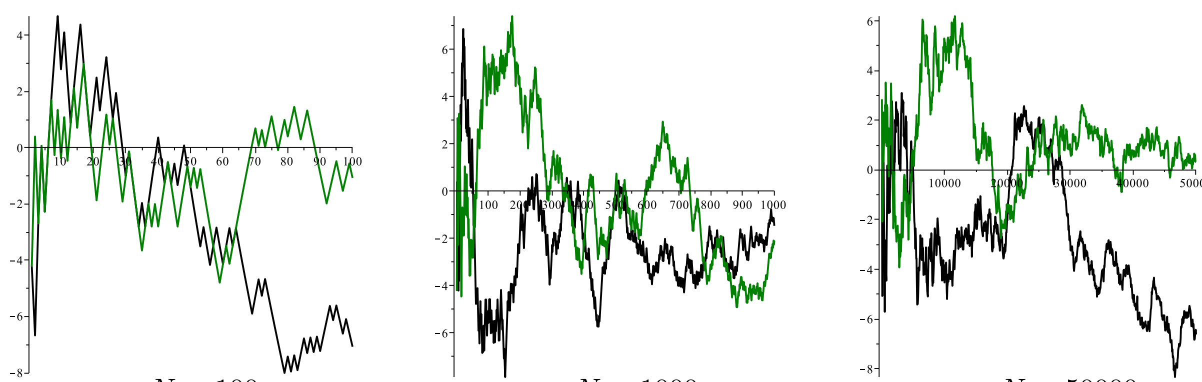
$n = 600$

Normalised distribution of the number of red balls in an small urn R_1, initial composition $(1,0)$



Density of a centered normal distribution with variance $\frac{5200}{529}$

The difference with almost sure convergence can be visualised on the following trajectory graphs. Even if the distribution at time n converges to a normal distribution, for a given sequence of random drawings, the number $\frac{1}{\sqrt{n}}\left(U_n^{(1)} - \mathbf{E}U_n^{(1)}\right)$ does not converge to a real number. The trajectory is erratic and looks like a brownian motion. On the figure hereunder, two different trajectories of the (completely) normalised number of red balls in a R_1-urn. On the x-axis, the number n of drawings; on the y-axis, $\frac{1}{\sqrt{n}}\left(U_n^{(1)} - nv_1^{(1)}\right)$, where $v_1^{(1)}$ is v_1 first coordinate.



$N = 100$ $\quad N = 1000$ $\quad N = 50000$

$\frac{1}{\sqrt{n}}\left(U_n^{(1)} - nv_1^{(1)}\right)$ in two sequences of N drawings in a small urn R_1, initial composition $(1,0)$

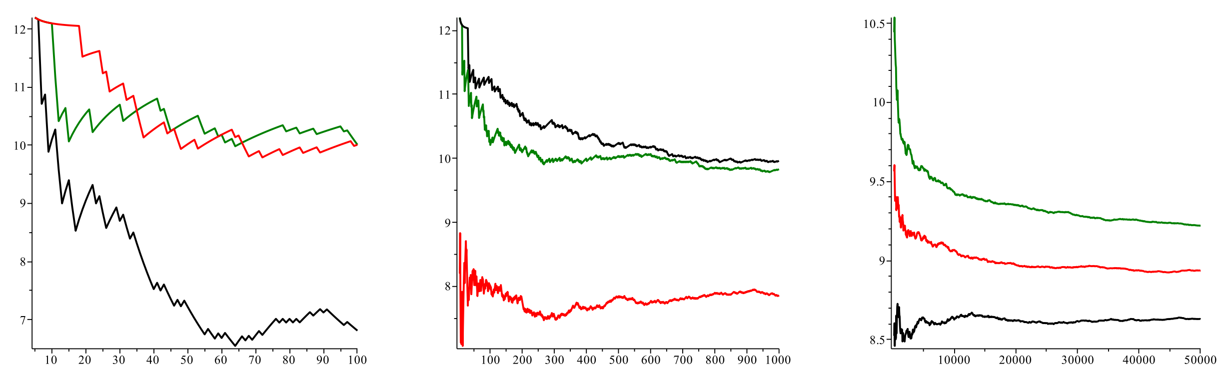
Theorem 4 (Large urns)
Suppose that the urn is large, *which means that* $1/2 < \sigma < 1$. *Then as n tends to infinity,*

(i) $\dfrac{U_n}{n}$ *converges to* v_1, *almost surely and in any* L^p, $p \geq 1$;

(ii) $\dfrac{U_n - nv_1}{n^\sigma}$ *converges almost surely and in any* L^p, $p \geq 1$ *to* Wv_2 *where* v_2 *is the (deterministic) eigenvector of* ${}^tR$ *defined in Section 1 and* W *is a real-valued random variable which admits a density and is supported by the whole real line. Besides, with the notations of Section 1,*

$$\mathbf{E}W = \frac{\Gamma\left(\frac{\alpha+\beta}{S}\right)}{\Gamma\left(\frac{\alpha+\beta}{S}+\sigma\right)} \frac{b\alpha - c\beta}{S}.$$

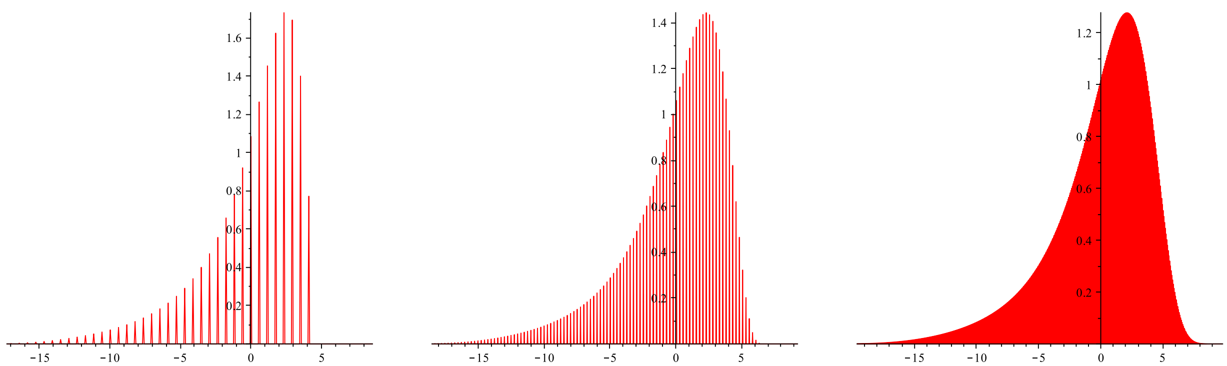
Assertion (i) is the same one as in the case of small urns. We make the same simulations as before for the urn R_2. The convergence to the (same) limit is visibly much slower, due to the second order term which grows like n^σ with $\sigma \simeq 0.77$ (instead of $\sqrt{n}$ for small urns). This second order term was already seeable on the trajectories of the number of red balls: the three slopes do not look not as similar as in the case of the small urn R_1 (but they really tend to a same one as N tends to infinity). Hereunder, again, on the x-axis, the number n of drawings up to $N = 100$, 1000 or 50000; on the y-axis, the normalised number of red balls $\frac{1}{n}U_n^{(1)}$.



$N = 100$ $N = 1000$ $N = 50000$

$\frac{1}{n}U_n^{(1)}$ in three sequences of N drawings in a large urn R_2, initial composition $(1, 0)$

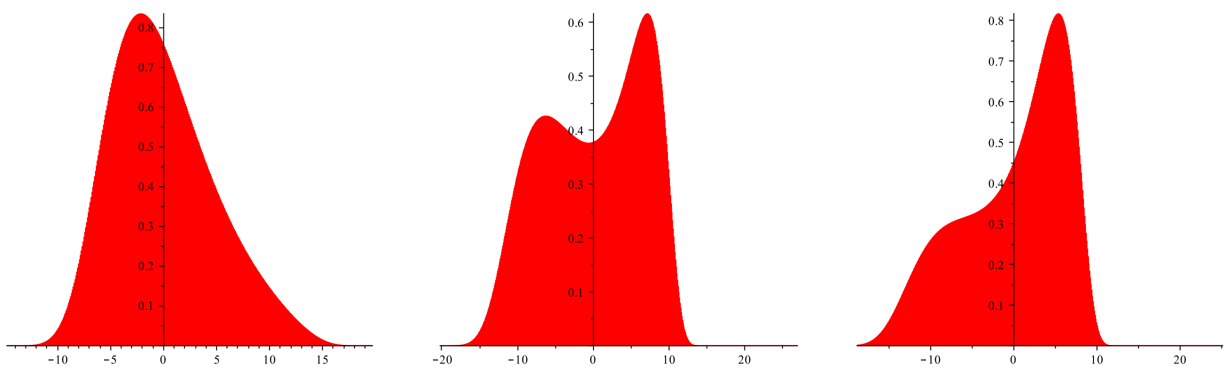
Almost sure convergence implies convergence in distribution. In particular, by formal computation of the probability generating function of red balls, the shape of W's density can be approached as already done (Beta function for the original Pólya urn, Gauss function for a small urn). The Fourier transform of W can be expressed in terms of the inverse of some suitable abelian integral (see [2]). Despite of this, very few is known about its density. The figure hereunder shows the graph of the density of $W - \mathbf{E}W$, approached by the (exact) distribution of $\frac{1}{n^\sigma}\left(U_n^{(1)} - \mathbf{E}U_n^{(1)}\right)$ for $n = 40$, 120 and 800.



$n = 40$ $\quad$ $n = 120$ $\quad$ $n = 800$

Normalised distribution of the number of red balls
in a large urn R_2 after n drawings, initial composition $(1, 0)$

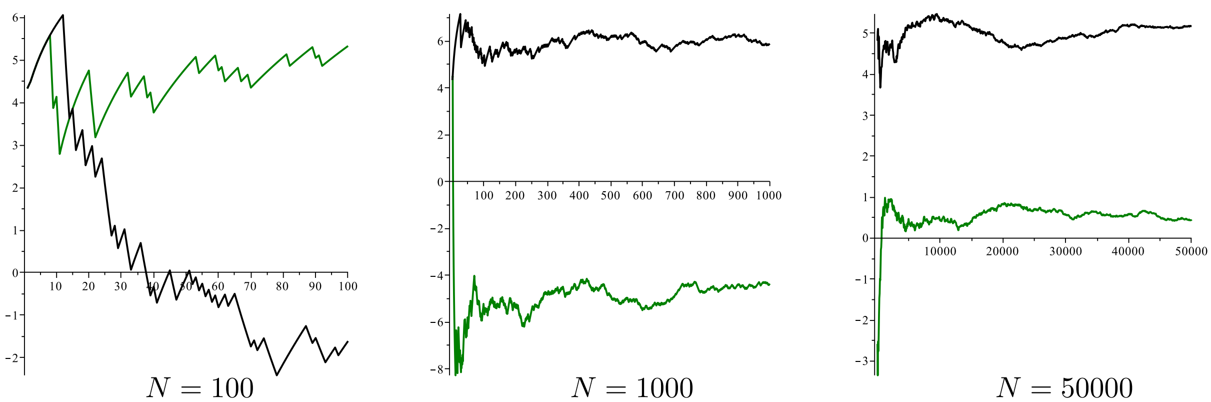
A remarkable fact: the distribution W depends on the initial composition of the urn, which does not happen for small urns. The graphs hereunder illustrate this property, representing $W - \mathbf{E}W$'s density for the large urn R_2 starting with respectively $(1, 0)$, $(1, 1)$ and $(2, 1)$ as initial composition vector.



$(\alpha, \beta) = (1, 0)$ $\quad$ $(\alpha, \beta) = (1, 1)$ $\quad$ $(\alpha, \beta) = (2, 1)$

Normalised distribution of the number of red balls in a large urn R_2
after 500 drawings, initial composition (α, β)

The last illustration concerns the second term order which has a random asymptotics. Two normalised trajectories of the number of red balls in an R_2-urn up to time $N = 100$, 1000 and 50000 are plotted. The convergence of $\frac{1}{n^\sigma}\left(U_n^{(1)} - nv_1^{(1)}\right)$ is here almost sure: for (almost) any sequence of random drawings in the large urn, this random variable converges to a (random) limit. The situation is very different from the small urn case, where a given trajectory do not give rise to the convergence of the second order normalised number of red balls. Here again, on the x-axis, the number n of drawings up to N; on the y-axis, the second order normalised number of red balls $\frac{1}{n^\sigma}\left(U_n^{(1)} - nv_1^{(1)}\right)$. Here again, $v_1^{(1)}$ denotes v_1 first coordinate



$\frac{1}{n^\sigma}\left(U_n^{(1)} - nv_1^{(1)}\right)$ in three sequences of N drawings in a large urn R_2, initial composition $(1,0)$

3.3 Hint of proof

All the proofs of these asymptotic results rely on martingale theory.
Historically, the first approach was made in the 70's by Athreya and Karlin who considered the composition vector process of an urn as a multitype branching process. They first embed the urn process into continuous time and make its study as a continuous-time branching process [1]. In his seminal article [5], Janson adapts the method in a complete study of an urn process under an irreducibility assumption. A direct discrete time approach based on moments is made in [6]. The arguments presented hereunder rely essentially on this latter approach.

The vector-valued Markov process $(U_n)_{n\in\mathbb{N}}$ is defined by the probability transitions (1) and the initial composition vector $U_0 = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$. In particular, if $f : \mathbb{R}^2 \to V$ is any function that takes its value in any real vector space V, the conditional expectation of U_{n+1} writes

$$\mathbf{E}\left(f\left(U_{n+1}\right)\Big|U_n\right) = \frac{U_n^{(1)}}{nS+\alpha+\beta}\, f\left(U_n + \begin{pmatrix}a\\ b\end{pmatrix}\right) + \frac{U_n^{(2)}}{nS+\alpha+\beta}\, f\left(U_n + \begin{pmatrix}c\\ d\end{pmatrix}\right).$$

Thanks to the deterministic relation $U_n^{(1)} + U_n^{(2)} = nS + \alpha + \beta$, this formula can be written the following way:

$$\mathbf{E}\left(f\left(U_{n+1}\right)\Big|U_n\right) = \left(\mathrm{Id} + \frac{\Phi}{nS+\alpha+\beta}\right)(f)\left(U_n\right) \tag{6}$$

where Φ denotes the operator defined, for any function f as above and any vector $v = \begin{pmatrix}v^{(1)}\\ v^{(2)}\end{pmatrix} \in \mathbb{R}^2$, by

$$\Phi\left(f\right)\left(v\right) = v^{(1)}\left[f\left(v + \begin{pmatrix}a\\ b\end{pmatrix}\right) - f\left(v\right)\right] + v^{(2)}\left[f\left(v + \begin{pmatrix}c\\ d\end{pmatrix}\right) - f\left(v\right)\right]. \tag{7}$$

A first consequence is the expectation of $f\left(U_n\right)$, obtained by recursion from Formula (6): if $f : \mathbb{R}^2 \to V$ is any function,

$$\mathbf{E}f\left(U_n\right) = \gamma_{n,\alpha+\beta}\left(\Phi\right)\left(f\right)\left(U_0\right) \tag{8}$$

where $\gamma_{n,\tau}$ is the real polynomial defined by

$$\gamma_{n,\tau}(X) = \prod_{k=0}^{n-1}\left(1+\frac{X}{kS+\tau}\right)$$

(τ is a non zero real number; if $n=0$, this empty product equals 1). Notice that, thanks to Stirling Formula, when z is any complex number, one gets the asymptotics

$$\gamma_{n,\tau}(z) = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau+z}{S}\right)} n^{\frac{z}{S}}\left(1+O\left(\frac{1}{n}\right)\right) \tag{9}$$

where Γ denotes Euler Gamma function. Formulae (6) and (8) are basic tools for the present proof. When $f \neq 0$ is an eigenvector of Φ related to the eigenvalue λ, *i.e.* when $\Phi(f) = \lambda f$, then $\gamma_{n,\tau}(\Phi)(f)(v) = \gamma_{n,\tau}(\lambda) \times f(v)$ so that Formula (9) gives immediately the asymptotics of $\mathbf{E}f(U_n)$ when n tends to infinity. With this elementary remark, one can evaluate the asymptotic joint moments of U_n's coordinates, leading to the proof of Theorem 4. Theorem 2 can also be proven with such tools. Classically, the proof of the small irreducible case (Theorem 3) is made by embedding the process into continuous time, and coming back to discrete time using some suitable random stopping-time. See [5] for a complete proof.

Exercise 6.
6.1- (Linear functions)
Show that if V is a real vector space and if $f : \mathbb{R}^2 \to V$ is linear, then

$$\Phi(f) = f \circ A$$

where $A : \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $A(v) = A\begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} := {}^tR\begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix}$.

6.2- (Vector-valued martingale)
Denote $\tau := \alpha+\beta$. Show that the process $\left(\gamma_{n,\tau}\left({}^tR\right)^{-1}(U_n)\right)_n$ is a martingale (with regard to the natural filtration) as soon as it is defined, *i.e.* as soon as all matrices $I_2 + \frac{1}{kS+\tau}R$, $k \in \mathbb{N}$ are invertible. Show that this martingale is not defined if, and only if $m \leq -1$ and S divides $m+\alpha+\beta$.

[Apply **6.1-** to $f = \mathrm{Id}$. This leads to $\mathbf{E}\left(U_{n+1}\middle|U_n\right) = \left(I_2 + \frac{1}{nS+\tau}A\right)(U_n)$. This implies all the answers, because A is diagonalizable, with eigenvalues S and m. For the martingale assertion, one can also refer to Brigitte Chauvin's course *Random trees and probability*, Proposition 3.7, where a similar argument is given.]

6.3- (Expectation)
Let u_1 and u_2 be the eigenforms defined in Section 1. Verify (or remember!) that $u_1 \circ A = Su_1$ and $u_2 \circ A = mu_2$. Show that for any $n \in \mathbb{N}$,

$$\mathbf{E}u_1(U_n) = n + \frac{\tau}{S}$$

and, when n tends to infinity,

$$\mathbf{E}u_2(U_n) = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau}{S}+\sigma\right)}\frac{b\alpha - c\beta}{S} n^{\sigma}\left(1+O\left(\frac{1}{n}\right)\right).$$

When $R \neq SI_2$, using that $v = u_1(v)v_1 + u_2(v)v_2$ for any vector $v \in \mathbb{R}^2$, show that, when n tends to infinity,

$$\mathbf{E}U_n \sim nv_1$$

[An induction using **6.1-** leads to $\mathbf{E}u_1\left(U_n\right) = \gamma_{n,\tau}(S) \times u_1\left(U_0\right) = \frac{nS+\tau}{\tau} \times \frac{\tau}{S}$. For u_2, apply Formula (9) to $\mathbf{E}u_2\left(U_n\right) = \gamma_{n,\tau}(m) \times u_2\left(U_0\right)$ with a O-remainder. The third assertion is obtained by addition of asymptotic developments.]

6.4- (Real-valued projected martingales)
Show that

$$\left(\frac{u_1\left(U_n\right)}{nS+\tau}\right)_n$$

is an almost surely bounded (thus convergent) martingale and compute its expectation. Show that

$$\left(\frac{u_2\left(U_n\right)}{\gamma_{n,\tau}\left(m\right)}\right)_n$$

is a martingale as well, as soon as $m \geq 0$ or $m+\tau$ is not a multiple of S.

[Using **6.1-** again, one gets $\mathbf{E}\left(u_1\left(U_{n+1}\right)\middle|U_n\right) = \left(1+\frac{S}{nS+\tau}\right) \times u_1\left(U_n\right)$, so that $\mathbf{E}\left(\frac{u_1(U_{n+1})}{(n+1)S+\tau}\middle|U_n\right) = \frac{u_1(U_n)}{nS+\tau}$, proving the martingale property. Same argument from $\mathbf{E}\left(u_2\left(U_{n+1}\right)\middle|U_n\right) = \left(1+\frac{m}{nS+\tau}\right) \times u_2\left(U_n\right)$.]

6.5- (Second moments)
Denote by P and Q the 2-variable polynomials defined by

$$P(x,y) = u_1(x,y)\Big(u_1(x,y)+1\Big) \quad \text{and} \quad Q(x,y) = \Big(u_1(x,y)+\sigma\Big)u_2(x,y).$$

Show that $\Phi(P) = 2SP$ and $\Phi(Q) = (S+m)Q$ and prove the asymptotics when n tends to infinity

$$\mathbf{E}P\left(U_n\right) = n^2\left(1+O\left(\frac{1}{n}\right)\right)$$

and

$$\mathbf{E}Q\left(U_n\right) = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau}{S}+\sigma\right)}\frac{b\alpha-c\beta}{S}n^{1+\sigma}\left(1+O\left(\frac{1}{n}\right)\right)$$

(if one feels depressed, one can just show that $Q\left(U_n\right) \in O\left(n^{1+\sigma}\right)$, :-)).

Suppose that $\sigma \neq 1/2$ and denote

$$R = u_2^2 - \frac{bc\sigma^2}{1-2\sigma}u_1 + (b-c)\,\sigma u_2.$$

Using (7), show that, in this case, $\Phi\left(R\right) = 2mR$ and that, when n tends to infinity,

$$\mathbf{E}R\left(U_n\right) = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau}{S}+2\sigma\right)}R\left(\alpha,\beta\right)n^{2\sigma}\left(1+O\left(\frac{1}{n}\right)\right).$$

Show that $(1, u_1, u_2, P, Q, R)$ is a basis of the vector space $\mathbb{R}_2[x,y]$ of polynomials of degree less than or equal to 2. Write x^2, xy and y^2 in this basis and compute the asymptotics of the co-moment matrix $\mathbf{E}\left[U_n{}^tU_n\right]$ and of the covariance matrix $\mathbf{E}\left[\left(U_n - \mathbf{E}U_n\right)^t\left(U_n - \mathbf{E}U_n\right)\right]$ (one has to discuss whether $\sigma < 1/2$ or $\sigma > 1/2$).

Check what happens when $\sigma = 1/2$ and do the same job using $T = u_2^2 + \frac{2b-m}{2}u_2$ instead of R.

[One gets $\Phi(P)$, $\Phi(Q)$ and $\Phi(R)$ by simple computation. Since $\Phi(P) = 2SP$, $\mathbf{E}P\left(U_n\right) = \gamma_{n,\tau}(2S) \times P\left(U_0\right)$ and the required asymptotics for $\mathbf{E}P\left(U_n\right)$ is obtained thanks to Formula (9). *Idem* for $\mathbf{E}Q\left(U_n\right)$ and $\mathbf{E}R\left(U_n\right)$. The remainder of the exercise is completely left to the reader.]

6.6- (For large urns, the second projected martingale is square-bounded)
Suppose that $\sigma > 1/2$. Expressing u_2^2 as a function of R, u_1 and u_2, show that the martingale $\left(\frac{u_2(U_n)}{\gamma_{n,\tau}(m)}\right)_n$ is bounded in L^2, thus convergent.

[$u_2^2 = R + \frac{bc\sigma^2}{1-2\sigma}u_1 - (b-c)\,\sigma u_2$, so that $\mathbf{E}u_2^2\left(U_n\right) = c_1 n^{2\sigma}\left(1 + O(1/n)\right) + c_2 n + c_3 n^{\sigma}\left(1 + O(1/n)\right)$ where c_1, c_2 and c_3 are constants. Since $\sigma > 1/2$, the principal term is the one in $n^{2\sigma}$, proving that the martingale is square bounded (use Formula (9) again to get the asymptotics of $\gamma_{n,\tau}(m)^2$).]

Exercise 7 (triangular urn).
Assume that $b = 0$, so that $R = \begin{pmatrix} S & 0 \\ S-m & m \end{pmatrix}$. Assume also that the initial number of black balls is non zero, *i.e.* that $\beta \neq 0$ (and check that $\beta = 0$ leads to a degenerate process). Let as above u_1 be the linear form $u_1(x,y) = \dfrac{x+y}{S}$ but let here u_2 be the linear form

$$u_2(x,y) = \frac{y}{S}.$$

For any $p \in \mathbb{N}^*$, let also A_p and B_p be the bivariate polynomials

$$A_p = u_1\left(u_1 + 1\right)\dots\left(u_1 + p - 1\right) = \frac{\Gamma\left(u_1 + p\right)}{\Gamma\left(u_1\right)}$$

and

$$B_p = u_2\Big(u_2 + \sigma\Big)\dots\Big(u_2 + (p-1)\sigma\Big) = \frac{\Gamma\left(u_2 + p\sigma\right)}{\Gamma\left(u_2\right)}.$$

Show that $\Phi\left(A_p\right) = pSA_p$ (as always, even if R is not triangular) and that $\Phi\left(B_p\right) = pmB_p$ for any $p \geq 1$. Deduce from this that, when n tends to infinity,

$$\mathbf{E}B_p\left(U_n\right) = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau}{S} + p\sigma\right)}\frac{\Gamma\left(\frac{\beta}{S} + p\sigma\right)}{\Gamma\left(\frac{\beta}{S}\right)}n^{p\sigma}\left(1 + O\left(\frac{1}{n}\right)\right).$$

• Assume that $m \geq 1$.
Using the inversion formula

$$u_2^p = \sum_{k=1}^{p}(-\sigma)^{p-k}\left\{\begin{matrix} p \\ k \end{matrix}\right\}B_k,$$

show that, for any $p \geq 1$,

$$\lim_{n\to\infty} \mathbf{E}\left(\frac{u_2(U_n)}{n^\sigma}\right)^p = \frac{\Gamma\left(\frac{\tau}{S}\right)}{\Gamma\left(\frac{\tau}{S}+p\sigma\right)} \frac{\Gamma\left(\frac{\beta}{S}+p\sigma\right)}{\Gamma\left(\frac{\beta}{S}\right)}. \tag{10}$$

so that the number of black balls $U_n^{(2)} = Su_2(U_n)$ converges in law to a random variable having the right side of Equality (10) as p-th moment (to make a complete proof of that fact, one has to check that a distribution having such a p-th moment is determined by its moments, which can be done by computing the asymptotics of (10) as p tends to infinity with the help of Stirling Formula). This law can be related to stable laws or to Mittag-Leffler ones.

• Assume that $m = 0$. Show that the process is deterministic (degenerate case).

• Assume that $m \leq -1$. Show that the number of black balls tends almost surely to zero (degenerate case again).

References

[1] K.B. Athreya and S. Karlin. Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist*, 39:1801–1817, 1968.

[2] B. Chauvin, N. Pouyanne, and R. Sahnoun. Limit distributions for large Pólya urns. *Annals Applied Prob.*, 21(1):1–32, 2011.

[3] Philippe Flajolet, Philippe Dumas, and Vincent Puyhaubert. Some exactly solvable models of urn process theory. In Philippe Chassaing, editor, *Fourth Colloquium on Mathematics and Computer Science*, volume AG of *DMTCS Proceedings*, pages 59–118, 2006.

[4] Philippe Flajolet, Joaquim Gabarró, and Helmut Pekari. Analytic urns. *Annals of Probability*, 33:1200–1233, 2005.

[5] S. Janson. Functional limit theorem for multitype branching processes and generalized Pólya urns. *Stochastic Processes and their Applications*, 110:177–245, 2004.

[6] N. Pouyanne. An algebraic approach to Pólya processes. *Annales de l'Institut Henri Poincaré*, 44:293–323, 2008.