ADAMA 2012
Ecole d'automne en Analyse d'Algorithmes et Modèles Aléatoires
Mahdia (Tunisie), 7 au 11 octobre 2012
Nicolas Pouyanne

Urnes de Pólya

— Notes de cours —

1 Urnes de Pólya : définition du processus

On se donne une matrice carrée de dimension 2 à coefficients entiers relatifs

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

et un vecteur-colonne non nul $U_0 = {}^t(\alpha, \beta)$ à coefficients entiers naturels.
A ces données, on associe le processus d'urne de Pólya $(U_n)_{n\in\mathbb{N}}$ défini de façon imagée par le mécanisme suivant : une urne contient des boules rouges et des boules noires (ou des olives noires et vertes). A l'origine des temps, elle contient α boules rouges et β boules noires. On tire une boule au hasard, toutes les boules ayant la même probabilité d'être choisies. Si la boule tirée est rouge, on la replace dans l'urne et on ajoute a boules rouges et b boules noires ; si la boule tirée est noire, ce sont c boules rouges et d boules noires que l'on ajoute dans l'urne. On obtient ainsi une nouvelle composition U_1. Le processus des vecteurs-composition $(U_n)_{n\in\mathbb{N}}$ est défini en itérant ce procédé.

Dans ce mini-cours, on fera les hypothèses suivantes sur la matrice de remplacement :
(i) R est *balancée*, *i.e.* $a + b = c + d \geq 1$;
(ii) R est *"tenable"*, *i.e.* $\Big(b, c \geq 0\Big)$ et $\Big(a \leq -1 \Longrightarrow a|c\Big)$ et $\Big(d \leq -1 \Longrightarrow d|b\Big)$.
L'hypothèse de balance assure que l'on ajoute au total le même nombre $S \geq 1$ de boules à chaque instant. L'hypothèse de tenabilité garantit que le processus ne s'éteigne pas, c'est-à-dire qu'à chaque tirage, si a ou d est négatif, on puisse enlever le nombre de boules requis par la règle de remplacement (l'hypothèse de tenabilité complète pose aussi des conditions arithmétiques sur la composition initiale de l'urne).

• En termes plus rigoureux,

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

est la chaîne de Markov à valeurs dans l'ensemble $\mathcal{O}$ des vecteurs-colonne non nuls de dimension 2 à coefficients entiers naturels, définie par les probabilités conditionnelles de transition

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

L'hypothèse de balance entraîne que $U_n^{(1)} + U_n^{(2)} = \alpha + \beta + nS$ pour tout n : à tout instant n, la composition de l'urne est aléatoire mais le nombre total de boules est déterministe.

• En termes encore plus rigoureux, la chaîne de Markov sur $\mathcal{O}$ à temps discret est définie par la famille

$$\left(\mu_{\binom{x}{y}}\right)_{\binom{x}{y} \in \mathcal{O}}$$

de mesures sur $\mathcal{O}$ donnée par : $\forall\ {}^t(x, y) \in \mathcal{O}$,

$$\mu_{\binom{x}{y}} = \frac{x}{x+y}\delta_{\binom{x}{y}+\binom{a}{b}} + \frac{y}{x+y}\delta_{\binom{x}{y}+\binom{c}{d}},$$

le symbole δ_P désignant la mesure de Dirac en P. Noter que l'hypothèse de tenabilité assure que $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$ et $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix}$ sont dans $\mathcal{O}$ dès lors que $\begin{pmatrix} x \\ y \end{pmatrix}$ y est.

[Généralisation à plusieurs couleurs (R est de dimension finie quelconque), aux urnes à remplacements aléatoires (les coefficients de R sont des variables aléatoires, travaux de S. Janson et R. Aguech par exemple). Dans ce cours, on se limite aux hypothèses restrictives décrites dans cette section.]

Notations (décomposition spectrale de R)
Grâce à l'hypothèse de balance, S est valeur propre de ${}^tR$. Par des considérations élémentaires à la Perron-Frobenius, la seconde valeur propre $m := a - c = d - b$ de ${}^tR$ est inférieure ou égale à S. On notera $\sigma = m/S \leq 1$.
Lorsque $(b, c) \neq (0{,}0)$, on note $v_1 = \dfrac{S}{b+c}\begin{pmatrix} c \\ b \end{pmatrix}$ et $v_2 = \dfrac{S}{b+c}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$, vecteurs propres de ${}^tR$ respectivement associés aux valeurs propres S et m. La base duale de formes linéaires propres correspondante est donnée par $u_1(x, y) = \dfrac{1}{S}(x + y)$ et $u_2(x, y) = \dfrac{1}{S}(bx - cy)$.
Ces données serviront plus bas.

Noter qu'en dimension supérieure, la matrice R n'est pas nécessairement diagonalisable, même sur $\mathbb{C}$, ce qui induit quelques difficultés supplémentaires mais surmontables.

[Commentaire : de l'utilité des modèles d'urnes pour l'analyse d'algorithmes.]

2 L'approche en combinatoire analytique

L'approche combinatoire analytique des urnes de Pólya est due à Philippe Flajolet et ses co-auteurs Dumas, Gabarró, Pekari et Puyhaubert au milieu des années 2000. Les deux articles fondateurs sont [3] et [2].

La première idée consiste à coder la composition de l'urne par une suite de mots finis dont les lettres sont prises dans l'alphabet $\{\mathtt{r},\mathtt{n}\}$ ($\mathtt{r}$ pour rouge, $\mathtt{n}$ pour noire). La composition initiale de l'urne est codée par le mot

$$W_0 = \mathtt{rr}\ldots\mathtt{rnn}\ldots\mathtt{n} = \mathtt{r}^\alpha\mathtt{n}^\beta.$$

La tirage dans l'urne revient à choisir uniformément une lettre du mot. Si la lettre choisie est un $\mathtt{r}$, on la remplace dans le mot par $\mathtt{r}^{a+1}\mathtt{n}^b$; si la lettre choisie est un $\mathtt{n}$, on la remplace par $\mathtt{r}^c\mathtt{n}^{d+1}$. La succession des tirages donne ainsi lieu à une suite de mots (aléatoires)

$$W_0, W_1, W_2\ldots$$

A l'instant n, on retrouve bien sûr le vecteur de composition U_n en comptant le nombre de lettres $\mathtt{r}$ et $\mathtt{n}$ dans le mot W_n.

Définition 1 (Histoires du processus)
Si n est un entier naturel, si $\binom{u_0}{v_0} \in \mathcal{O}$ et $\binom{u}{v} \in \mathcal{O}$, une histoire de longueur n menant de $\binom{u_0}{v_0}$ à $\binom{u}{v}$ *est une suite de mots $W_0 = \mathtt{r}^{u_0}\mathtt{n}^{v_0}, W_1, W_2, \ldots, W_n$ produits de la manière ci-dessus, pour lesquels W_n contient exactement u lettres $\mathtt{r}$ et v lettres $\mathtt{n}$.*

Naturellement, avec ces notations, à cause de l'hypothèse de balance, quelle que soit son histoire, le mot W_n a toujours $u_0 + v_0 + nS$ lettres. L'objet combinatoire central de la méthode Flajolet est le nombre de ces histoires : on note

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

le nombre d'histoires de longueur n menant de $\binom{u_0}{v_0}$ à $\binom{u}{v}$.

Exercice 1. Lorsque $R = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}$, coder et compter toutes les histoires de longueur 2 qui mènent de $\binom{2}{0}$ à $\binom{4}{4}$.
[Une solution : on part de $W_0 = \mathtt{r}^2$. On peut dessiner un arbre des possibles : $W_1 \in \{\mathtt{rn}^3\mathtt{r}, \mathtt{r}^2\mathtt{n}^3\}$, puis $W_2 \in \{\mathtt{rn}^6\mathtt{r}, \mathtt{r}^3\mathtt{n}^4\mathtt{r}, \mathtt{rnr}^2\mathtt{n}^3\mathtt{r}, \mathtt{rn}^2\mathtt{r}^2\mathtt{n}^2\mathtt{r}, \mathtt{rn}^3\mathtt{rn}^3\}$ ou $W_2 \in \{\mathtt{rn}^3\mathtt{rn}^3, \mathtt{r}^2\mathtt{n}^6, \mathtt{r}^4\mathtt{n}^4, \mathtt{r}^2\mathtt{nr}^2\mathtt{n}^3, \mathtt{r}^2\mathtt{n}^2\mathtt{r}^2\mathtt{n}^2\}$. Parmi les dix histoires de longueur 2, six mènent à $\binom{4}{4}$ et quatre mènent à $\binom{2}{6}$: partant de deux boules rouges, la probabilité que l'urne contienne quatre boules rouges et quatre boules noires au bout de deux tirages est 3/5. Attention aux craintes infondées : dans l'exemple, la configuration $\mathtt{rn}^3\mathtt{rn}^3$ est atteinte par deux histoires différentes. Ce sont bien les histoires que l'on compte, et non les différents mots auxquels on aboutit.]

Exercice 2 (c'est l'urne de l'article original de Pólya en 1930)**.** Lorsque $R = SI_2$, calculer tous les nombres H_n, $n \geq 0$.

[Combinatoire élémentaire, faire le dessin d'un chemin dans $\mathbb{N}^2$ et compter les histoires qui suivent chacun des chemins. Pour tout $(p,q)\in\mathbb{N}^2$ tel que $p+q=n$, on trouve

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

les autres H_n sont nuls.]

Exercice 3. Pour n'importe quelle urne, si $N=\alpha+\beta$, le nombre total d'histoires de longueur n partant de $\binom{\alpha}{\beta}$ est $N\big(N+S\big)\big(N+2S\big)\dots\big(N+(n-1)S\big)=n!S^n\binom{\frac{N}{S}+n-1}{n}$.

Comme presque toujours en combinatoire analytique, on travaille sur des séries génératrices. Ici, c'est de la série trivariée des histoires qu'il s'agit, les variables comptant respectivement le nombre de boules rouges ou noires à l'arrivée et la longueur. Ainsi, une matrice de remplacement étant donnée, on notera

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

Exercice 4. Pour n'importe quelle urne, $H\left(1,1,z\,\middle|\,\begin{array}{c} u_0 \\ v_0 \end{array}\right)=\left(\dfrac{1}{1-Sz}\right)^{\frac{u_0+v_0}{S}}$.

Exercice 5. Pour l'urne originale $R=SI_2$,

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

[Manipulations sur les séries entières multivariées à partir de $\frac{1}{(1-X)^N}=\sum_{n\geq 0}\binom{N+n-1}{n}X^n$.]

[Commentaire oral sur les papier de P. Flajolet *et al.* : pointer un objet combinatoire se traduit par une dérivée partielle sur les séries génératrices ; opérer à un remplacement a pour effet de multiplier par un monôme *ad hoc*. Ces considérations conduisent au résultat suivant, qualifié de "Basic isomorphism" dans [2].]

Théorème 1 (Flajolet, Dumas, Puyhaubert, 2006)
Soient x,y,z des nombres complexes tels que $xy\neq 0$. Soient $X(t)$ et $Y(t)$ les solutions du problème de Cauchy (formel ou analytique)

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

Alors, pour toute configuration initiale (u_0, v_0)*, pour tout* z *assez proche de l'origine (cadre analytique),*

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

Exemple 1. Retour à l'urne originale de Pólya pour laquelle $R = SI_2$: le système différentiel s'écrit $X' = X^{S+1}, Y' = Y^{S+1}$, qui se résout. Le problème de Cauchy a pour solution $X(t) = x(1-Stx^S)^{-1/S}$, $Y(t) = y(1-Sty^S)^{-1/S}$. Avec le théorème, on retrouve le calcul fait plus haut (exercice 5).

Preuve du théorème. On considère l'opérateur différentiel sur les fonctions de deux variables

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

L'effet de $\mathcal{D}$ sur les monômes est lié aux histoires de l'urne par la formule

$$\begin{aligned}\mathcal{D}\left(x^{u_0}y^{v_0}\right)&=u_0x^{a+u_0}y^{b+v_0}+v_0x^{c+u_0}y^{d+v_0}\\&=H_1\begin{pmatrix}u_0&u_0+a\\v_0&v_0+b\end{pmatrix}x^{a+u_0}y^{b+v_0}+H_1\begin{pmatrix}u_0&u_0+c\\v_0&v_0+d\end{pmatrix}x^{c+u_0}y^{d+v_0}\\&=\sum_{u,v\geq 0}H_1\begin{pmatrix}u_0&u\\v_0&v\end{pmatrix}x^uy^v\end{aligned}$$

et par récurrence, pour tout $n \in \mathbb{N}$,

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

[Noter que c'est dans cette récurrence que les propriétés markoviennes du processus interviennent.] Par ailleurs, si (X, Y) est une solution du système différentiel $X' = X^{a+1}Y^b$, $Y' = X^cY^{d+1}$, alors

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

avec une formule analogue pour la dérivée $n^{\text{ième}}$. En regroupant, on a successivement

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

en enfin, grâce à la formule de Taylor (analytique ou formelle),

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

Le théorème s'ensuit immédiatement en prenant la valeur à l'origine ($t = 0$). ∎

Ce théorème permet de trouver des expressions exactes des fonctions H et d'en tirer des conséquences probabilistes très précises sur la distribution de l'urne à temps fini, ou encore sur l'asymptotique du processus. Suivent quelques exemples, tirés pour l'essentiel de [3] et [2].

Remarque. Une conséquence du théorème est que

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

Cette formule évoque une propriété de convolution et doit être mise en perspective avec la propriété de branchement du processus de l'urne plongée en temps continu, qui conduit à une équation semblable sur les transformées de Fourier de lois limites des grandes urnes. Voir le cours de Brigitte Chauvin (ADAMA 2012, Mahdia) et l'article [1]. Le lien entre ces deux propriétés reste à établir.

Exemple 2. Il s'agit de l'urne dont la matrice de remplacement est $\begin{pmatrix}0&1\\1&0\end{pmatrix}$. [Parler de la métaphore de la campagne électorale, utilisée par P. Flajolet.] Le problème de Cauchy

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

se résout sans difficulté. On trouve

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

Par exemple, lorsque l'on part d'une unique boule rouge, la fonction génératrice de probabilité du nombre de boules rouges est

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

puisque le nombre total d'histoires de longueur n partant d'une boule rouge est $n!$ (calculé plus haut dans l'exercice 3). En remplaçant par l'expression de la fonction H, on obtient

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

Cette fonction de z a un pôle simple en $z=\frac{\log x}{x-1}$ comme unique singularité. Puisque cette fonction de x est analytique en 1, l'analyse des singularités permet d'appliquer le théorème des quasi-puissances de Hwang : la moyenne et la variance de $U_n^{(1)}$ sont asymptotiquement proportionnelles à n, et le nombre de boules rouges à l'instant n (*i.e.* la variable aléatoire $U_n^{(1)}$) vérifie une loi des grands nombres et un théorème central limite (loi gaussienne).

Exemple 3. Arbres $2-3$, exemple fétiche de l'article [3]. Il s'agit de l'urne $\begin{pmatrix}-2&3\\4&-3\end{pmatrix}$, qui modélise le processus des feuilles d'un arbre $2-3$, structure célèbre de l'algorithmique. Cette fois, le problème de Cauchy s'écrit

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

On pose $Z = X^2$ et on obtient successivement $Z' = 2Y^3$ et $Z'' = 6Z^2$. En multipliant par Z' puis en intégrant, on montre que Z est nécessairement solution du problème de Cauchy

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

où $g_3 = 4(x^6 - y^6)$. Cette équation précipite l'étude de cette urne dans la théorie des fonctions elliptiques. Rapidement dit, soit $\wp(z) = \wp(z; 0, -4)$ la fonction elliptique de Weiestraß associée aux invariants $g_2 = 0$ et $g_3 = -4$: si on note

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

(fonction Beta d'Euler) et si Λ désigne le réseau hexagonal

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

la fonction $\wp$ est la fonction méromorphe du plan complexe définie hors des points de Λ par

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

La fonction $\wp$ présente un pôle double en tout point de Λ et est Λ-périodique (on dit ainsi que $\wp$ est *doublement périodique*). Ses zéros modulo le réseau sont localisés en les points $\omega/3$ et $2\omega/3$. La raison pour laquelle cette fonction de Weierstraß intervient est qu'elle est solution de l'équation (3) ; cela se montre par la théorie élémentaire des fonctions holomorphes. Une autre manière voir $\wp$: elle inverse l'intégrale elliptique sous-jacente à l'équation (3). Plus précisément, si z et w sont des nombres complexes, on a l'équivalence

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

le symbole $[w, \infty]$ désignant n'importe quelle demi-droite du plan issue de w et qui ne rencontre aucune des racines du polynôme $\zeta^3 + 1$ (la racine carrée étant alors la détermination associée à la coupure du plan définie par cette demi-droite). Un regain récent de célébrité des fonctions de Weiertraß vient du fait qu'elles fournissent des paramétrisations des cubiques planes lisses qui sont un des objets phares de la cryptographie moderne ; ici, le couple $(\wp, \wp')$ paramétrise la courbe d'équation $Y^2 = 4X^3 + 4$.

Ainsi, les solutions du système différentiel (2) s'expriment en termes de fonctions elliptiques sur le réseau hexagonal. Prenons par exemple le cas de l'urne contenant initialement 2 boules rouges (et aucune boule noire). Soit p_n la probabilité que toutes les boules soient noires au temps n. En terme de fonction H, ce nombre s'écrit

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

La résolution du problème de Cauchy donne ici simplement

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

Au vu des pôles de la fonction $\wp$, l'analyse des singularités exprime l'asymptotique de ce nombre en termes de puissances de $3/w \sim 0,7132$.

Remarques. 1- Le système différentiel monomial (1) admet une intégrale première simple : si X et Y en sont solutions, alors $1/X^m - 1/Y^m$ est une fonction (localement) constante. En exprimant ainsi Y en fonction de X et en reportant dans le système, on tombe naturellement sur des intégrales abéliennes à inverser. Tous les cas d'urnes "elliptiques", c'est-à-dire des urnes pour lesquelles ces intégrales abéliennes sont liées à des courbes de genre 1 (courbes elliptiques) sont répertoriés dans [3].

2- En dimensions supérieures (au moins trois couleurs), le théorème liant les fonctions H à un système monomial reste valide. Par exemple, le cours de Brigitte Chauvin (ADAMA 2012, Mahdia) s'appuiera sur une urne de dimension trois dans laquelle les fonctions H sont explicites et permettent une description exacte des distributions de l'urne à temps fini. Cependant, une obstruction théorique ternit l'espoir de généraliser l'efficacité et surtout la précision de la méthode en dimension 2 : le système différentiel monomial général des urnes n'est pas intégrable en dimension supérieure ou égale à 3 (algèbre différentielle, mais aussi géométrie algébrique, voir la note 11 et les commentaires finaux de [2]).

3 Approche probabiliste et méthode de point fixe

3.1 Asymptotiques des grandes urnes et systèmes en lois

On se place sous les hypothèse de la section 1. On rappelle les résultats de l'asymptotique des processus d'urne en temps discret ou continu en dimension 2 (voir le cours de Brigitte Chauvin, ADAMA, Mahdia 2012). Ils se généralisent en dimension supérieure.

Théorème 2 (Asymptotique en temps discret)
(i) si $\sigma < 1/2$ (petite urne), alors $\frac{U_n - nv_1}{\sqrt{n}}$ *converge en loi vers un vecteur gaussien ;*

(ii) si $\sigma > 1/2$ (grande urne), alors

$$U_n = nv_1 + n^\sigma W^{DT} v_2 + o_{\text{fort}}(n^\sigma),$$

où W^{DT} est une variable aléatoire réelle.

[Rappel : (α, β) est la composition initiale, S la balance, $\sigma = m/S$ le rapport des valeurs propres de la matrice de remplacement dont les coefficients anti-diagonaux sont b et c. Commentaires sur “fort”, presque sûr et dans tous les espaces L^p, $p \geq 1$.]
En outre, on montre que

$$\mathbf{E}W^{DT} = \mathbf{E}\left(\lim_{n\to\infty} \frac{u_2(U_n)}{n^\sigma}\right) = \frac{\Gamma\left(\frac{\alpha+\beta}{S}\right)}{\Gamma\left(\frac{\alpha+\beta}{S}+\sigma\right)} \frac{b\alpha - c\beta}{S}.$$

Théorème 3 (Asymptotique en temps continu)
(i) si $\sigma < 1/2$, $e^{-St}U(t)$ converge presque sûrement vers ξv_1 où ξ est une variable aléatoire réelle distribuée selon la loi Gamma $\left(\frac{\alpha+\beta}{S}\right)$; en outre, $e^{-St/2}u_2\left(U(t)\right)$converge en loi vers $\sqrt{\xi}G$ où G est une loi gaussienne indépendante de ξ.

(ii) si $\sigma > 1/2$, alors $U(t) = e^{St}\xi v_1\left(1 + o_{\mathrm{fort}}(1)\right) + e^{mt}W^{CT}v_2\left(1 + o_{\mathrm{fort}}(1)\right)$ où ξ est une variable aléatoire réelle distribuée selon la loi Gamma $\left(\frac{\alpha+\beta}{S}\right)$ et où W^{CT} est une variable aléatoire réelle.

[Commentaires sur “fort”. Remarque sur l'indépendance linéaire de v_1 et v_2 sans laquelle les deux o_{fort} n'auraient pas de sens ! Point de vue géométrique : la courbe (aléatoire) d'équation $y = W^{CT}\xi^{-\sigma}x^\sigma$ est presque sûrement asymptote à la trajectoire de $U(t)$.]
Là encore, l'espérance s'exprime facilement :

$$\mathbf{E}W^{CT} = \mathbf{E}\left(\lim_{t\to\infty} \frac{u_2\left(U(t)\right)}{e^{mt}}\right) = \frac{b\alpha - c\beta}{S}. \tag{4}$$

La question est celle des distributions de W^{DT} et de W^{CT}. On voit dans le cours de Brigitte Chauvin que dans le cas général, ces lois sont des convolées (multiplicatives) des lois élémentaires pour lesquelles les compositions initiales ne contiennent qu'une seule boule. On note ces lois élémentaires

$$\boxed{\begin{array}{ll} X^{DT} = W^{DT}_{(1,0)} & \quad Y^{DT} = W^{DT}_{(0,1)} \\ & \\ X^{CT} = W^{CT}_{(1,0)} & \quad Y^{CT} = W^{CT}_{(0,1)}. \end{array}}$$

Ces lois sont reliées par les connexions $X^{CT} = \xi^\sigma X^{DT}$ et $Y^{CT} = \xi^\sigma Y^{DT}$, la variable ξ étant indépendante de X^{DT} et de Y^{DT}, distribuée selon la loi *Gamma* $\left(\frac{1}{S}\right)$ dont le $p^{\text{ième}}$ moment est $\Gamma(1/S + p)/\Gamma(1/S)$. Les lois X et Y sont toutes solutions de systèmes de points fixes.

Théorème 4 (Système en loi pour le temps discret)
Les distributions X^{DT} et Y^{DT} sont solutions du système en loi

$$\left\{\begin{array}{l} X \overset{\mathcal{L}}{=} \displaystyle\sum_{k=1}^{a+1} (V_k)^\sigma X^{(k)} + \sum_{k=a+2}^{S+1} (V_k)^\sigma Y^{(k)} \\ \\ Y \overset{\mathcal{L}}{=} \displaystyle\sum_{k=1}^{c} (V_k)^\sigma X^{(k)} + \sum_{k=c+1}^{S+1} (V_k)^\sigma Y^{(k)} \end{array}\right. \tag{5}$$

où (i) $V = (V_1, \dots, V_{S+1})$ *est une distribution de Dirichlet de paramètres* $(\frac{1}{S}, \dots, \frac{1}{S})$ *;*
(ii) les $X^{(k)}$ *et les* $Y^{(k)}$ *sont des copies respectives de* X *et* Y*, indépendantes entre elles et indépendantes de* V*.*

Noter que chaque V_k est distribuée selon la loi de U^S où U est une variable uniforme sur $[0, 1]$.

Théorème 5 (Système en loi pour le temps continu)
Les distributions X^{CT} *et* Y^{CT} *sont solutions du système en loi*

$$\left\{ \begin{array}{l} X \overset{\mathcal{L}}{=} V^{\sigma} \left(\displaystyle\sum_{k=1}^{a+1} X^{(k)} + \sum_{k=a+2}^{S+1} Y^{(k)} \right) \\ \\ Y \overset{\mathcal{L}}{=} V^{\sigma} \left(\displaystyle\sum_{k=1}^{c} X^{(k)} + \sum_{k=c+1}^{S+1} Y^{(k)} \right) \end{array} \right. \tag{6}$$

où
(i) V *est la puissance* S *d'une variable uniforme sur* $[0, 1]$ *;*
(ii) les $X^{(k)}$ *et les* $Y^{(k)}$ *sont des copies respectives de* X *et* Y*, indépendantes entre elles et indépendantes de* V*.*

3.2 Contractions dans des espaces de mesures

[Travail en cours, en commun avec Brigitte Chauvin, Quansheng Liu et Cécile Mailler.]
Considérons par exemple le système en lois pour le temps continu (6). Pour résoudre la question de l'existence de surtout de l'unicité de ses solutions, l'idée consiste à trouver un espace de mesures *ad hoc* et une transformation de cet espace dont le couple de lois (X^{CT}, Y^{CT}) apparaisse comme un point fixe. Si on arrive à faire en sorte que l'espace soit un métrique complet et que la transformation soit contractante, le théorème de point fixe de Banach donnera à la fois l'existence et l'unicité.

Si u est n'importe quel nombre réel, on désigne par $\mathcal{M}_2(u)$ l'ensemble des mesures de probabilité sur $\mathbb{R}$ de carré intégrable et qui admettent u comme espérance. On définit classiquement sur $\mathcal{M}_2(u)$ la distance de Wasserstein comme suit : si μ_1 et μ_2 sont dans $\mathcal{M}_2(u)$,

$$d_W(\mu_1, \mu_2) = \min_{(X_1, X_2)} \left(\mathbf{E}(X_1 - X_2)^2 \right)^{1/2}$$

où le minimum est pris parmi les vecteurs aléatoires (X_1, X_2) de $\mathbb{R}^2$ dont les marginales X_1 et X_2 suivent respectivement les lois μ_1 et μ_2. Que le minimum soit atteint est garanti par le théorème de Kantorovich-Rubinstein. En outre, il est de notoriété publique que, muni de cette distance, l'espace métrique $\mathcal{M}_2(u)$ est complet.
Pour travailler sur le système, c'est de couples de variables qu'on doit traiter. Si u et v sont des réels, on choisit sur $\mathcal{M}_2(u) \times \mathcal{M}_2(v)$ la distance du maximum

$$d\Big((\mu_1, \nu_1), (\mu_2, \nu_2) \Big) = \max \left\{ d_W\left(\mu_1, \mu_2\right), d_W\left(\nu_1, \nu_2\right) \right\}$$

qui fait bien sûr encore de ce produit un espace complet.

On définit sur $\mathcal{M}_2(u) \times \mathcal{M}_2(v)$ l'application $\mathcal{K}$ à valeurs dans les couples de mesures de probabilité sur $\mathbb{R}$ par

$$\mathcal{K}(\mu, \nu) = \left(\mathcal{L}\left\{ V^\sigma \Big([a+1]X + [b]Y \Big) \right\}, \mathcal{L}\left\{ V^\sigma \Big([c]X + [d+1]Y \Big) \right\} \right). \tag{7}$$

Si $\mathcal{X}$ et $\mathcal{Y}$ sont des variables aléatoires, la notation $[p]\mathcal{X} + [q]\mathcal{Y}$ désigne la somme de p copies indépendantes de $\mathcal{X}$ et de q copies indépendantes de $\mathcal{Y}$, les copies de $\mathcal{X}$ et de $\mathcal{Y}$ étant indépendantes entre elles. La notation $\mathcal{L}(\mathcal{X})$ désigne la loi de la variable $\mathcal{X}$. Dans la définition (7), X désigne une variable aléatoire qui suit la loi μ et Y une variable aléatoire qui suit la loi ν ; enfin, V est la puissance S d'une variable uniforme sur $[0,1]$, indépendante des variables X et des variables Y. L'application $\mathcal{K}$ est faite pour voir le couple $\left(X^{CT}, Y^{CT}\right)$ comme un de ses points fixes. L'affaire marche par la grâce du théorème suivant.

Théorème 6 (Contraction pour les lois en temps continu)
Si $cu + bv = 0$, l'application $\mathcal{K}$ stabilise $\mathcal{M}_2(u) \times \mathcal{M}_2(v)$. En outre, elle est contractante.

En fait, on montre qu'elle est $\frac{S+1}{2m+1}$-lipschitzienne. Comme l'urne est grande, $S < 2m$ ce qui assure la contraction. Comme les moyennes de X^{CT} et Y^{CT} sont respectivement, b/S et $-c/S$ (voir la formule ci-dessus qui donne la moyenne de W^{CT}), le théorème s'applique au couple $\left(X^{CT}, Y^{CT}\right)$.

Par la même méthode, on montre un théorème analogue pour la transformation naturelle – jumelle de $\mathcal{K}$ – associée au système en loi du temps discret vérifié par X^{DT} et Y^{DT}. Ainsi, les systèmes en loi caractérisent les lois limites des grandes urnes.

Corollaire 1 (Caractérisation des lois limites par les systèmes en loi)
(i) Le couple de variables aléatoires $\left(X^{DT}, Y^{DT}\right)$ est l'unique couple de variables aléatoires de carrés intégrables solutions de (5), dont les moyennes sont respectivement $\dfrac{\Gamma\left(\frac{1}{S}\right)}{\Gamma\left(\frac{m+1}{S}\right)}\dfrac{b}{S}$ *et* $\dfrac{-\Gamma\left(\frac{1}{S}\right)}{\Gamma\left(\frac{m+1}{S}\right)}\dfrac{c}{S}$.
(ii) Le couple de variables aléatoires $\left(X^{CT}, Y^{CT}\right)$ est l'unique couple de variables aléatoires de carrés intégrables solutions de (6), dont les moyennes sont respectivement $\dfrac{b}{S}$ *et* $-\dfrac{c}{S}$.

Nombreuses sont les conséquences que l'on peut tirer de ces caractérisations en ne considérant plus nos mesures qu'en tant que solutions des systèmes en loi. Sans entrer dans le détail des preuves, on montre par exemple, par analyse de Fourier, que ces mesures de probabilité sont des ***densités*** dont le ***support est*** $\mathbb{R}$ tout entier. On peut aussi en tirer des résultats sur les moments de ces lois.

3.3 Moments des lois limites des grandes urnes

Le théorème d'asymptotique des grandes urnes assure que les lois limites ont des moments de tous ordres. Se pose immédiatement la question de la caractérisation de ces lois par leurs moments. On a démontré dans [1] – il en sera question plus bas – que leurs séries génératrices exponentielles des moments ont un rayon de convergence nul. Pour en savoir davantage, on invoque les systèmes en loi qui fournissent des équations de récurrence sur les moments. Par exemple, pour le temps continu, si on note

$$a_p = \mathbf{E}\left(X^{CT}\right)^p \quad \text{et} \quad b_p = \mathbf{E}\left(Y^{CT}\right)^p,$$

on tire de (6) et de la formule du multinôme le système récursif

$$\begin{cases} a_p = \dfrac{1}{1+mp} \displaystyle\sum_{k_1+\cdots+k_{S+1}=p} \dfrac{p!}{k_1!\ldots k_{S+1}!} \prod_{j=1}^{a+1} a_{k_j} \prod_{j=a+2}^{S+1} b_{k_j} \\ \\ b_p = \dfrac{1}{1+mp} \displaystyle\sum_{k_1+\cdots+k_{S+1}=p} \dfrac{p!}{k_1!\ldots k_{S+1}!} \prod_{j=1}^{c} a_{k_j} \prod_{j=c+1}^{S+1} b_{k_j}. \end{cases}$$

Après analyse de cette récursion, on en déduit le résultat suivant.

Théorème 7 (Magnitude des moments des lois limites des grandes urnes) *Si Z est n'importe laquelle des variables aléatoires X^{DT}, Y^{DT}, X^{CT} ou Y^{CT}, la suite*

$$\left(\frac{1}{\log p}\left(\frac{\mathbf{E}\,|Z|^p}{p!}\right)^{1/p}\right)_p$$

est bornée.

Le critère de Carleman assure que si Z est une variable aléatoire réelle pour laquelle la série

$$\sum_p \left(\mathbf{E}Z^{2p}\right)^{-\frac{1}{2p}}$$

diverge, alors elle est déterminée par ses moments. Son application aux lois qui nous occupent est immédiate.

Corollaire 2 *Les lois de X^{DT}, Y^{DT}, X^{CT} et Y^{CT} sont déterminées par leurs moments.*

4 Système différentiel monomial et transformées de Fourier

On se place dans le cadre des lois limites X^{CT} et Y^{CT} d'une grande urne en temps continu. On note leurs transformées de Fourier

$$\mathcal{F}(x) = \mathbf{E}\left(e^{ixX^{CT}}\right) = \int_{-\infty}^{+\infty} e^{ixt}d\mu_{X^{CT}}(t) \quad \text{et} \quad \mathcal{G}(x) = \mathbf{E}\left(e^{ixY^{CT}}\right).$$

Puisque X^{CT} et Y^{CT} admettent des moments de tous ordres, $\mathcal{F}$ et $\mathcal{G}$ sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Le système en loi (6) combiné avec les valeurs des espérances (4) conduit à l'énoncé suivant.

Théorème 8 (Système différentiel sur les Fourier de l'urne continue)
Les fonctions $\mathcal{F}$ et $\mathcal{G}$ sont solutions du système différentiel

$$\left\{\begin{array}{l} \mathcal{F}(x) + mx\mathcal{F}'(x) = \mathcal{F}(x)^{a+1}\mathcal{G}(x)^b \\ \\ \mathcal{G}(x) + mx\mathcal{G}'(x) = \mathcal{F}(x)^c\mathcal{G}(x)^{d+1} \end{array}\right.$$

et vérifient les conditions au bord à l'origine

$$\left\{\begin{array}{l} \mathcal{F}(x) = 1 + i\frac{b}{S}x + O(x^2) \\ \\ \mathcal{G}(x) = 1 - i\frac{c}{S}x + O(x^2). \end{array}\right.$$

Pour le calcul de la preuve, conditionner par U. Noter que ce système différentiel – très simple – est singulier à l'origine. On en fait dans [1] une résolution en termes d'inverses d'intégrales abéliennes sur la courbe de Fermat de degré m, d'équation $x^m + y^m = z^m$. C'est cette par cette étude analytique qu'on a montré que les séries génératrices exponentielles des moments

$$\sum_p \frac{\mathbf{E}\left(X^{CT}\right)^p}{p!} z^p \quad \text{et} \quad \sum_p \frac{\mathbf{E}\left(Y^{CT}\right)^p}{p!} z^p$$

ont un rayon nul : elles définissent des fonctions de classe $\mathcal{C}^\infty$ sur la droite imaginaire (Fourier) mais ne sont pas analytiques à l'origine, ne définissant de fonctions de la variable réelle (Laplace) que sur des intervalles bornés de la forme $]0, \rho[$.

• Pour terminer, on montre par l'heuristique de la résolution de leur système différentiel comment ces transformées de Fourier sont reliées aux fonction H de l'approche combinatoire analytique de Flajolet et de ses co-auteurs. On procède au changement de variable et de fonctions

$$f(w) = x^{\frac{1}{m}}\mathcal{F}(x) \quad \text{et} \quad g(w) = x^{\frac{1}{m}}\mathcal{G}(x)$$

où les variables $x \in \mathbb{R}$ et $w \in \mathbb{C}$ sont liées par la relation $(-Sw)^m x^S = 1$ — tout cela mérite un peu de soin à cause des déterminations des racines complexes, naturellement. On obtient formellement un nouveau système monomial non singulier dans le champ complexe

$$\left\{\begin{array}{l} f' = f^{a+1}g^b \\ \\ g' = f^c g^{d+1} \end{array}\right.$$

avec des conditions de bord à l'infini. Ce système différentiel est exactement celui qui régit les fonctions H élémentaires $H\left(x, y, z \,\middle|\, \begin{array}{c} 1 \\ 0 \end{array}\right)$ et $H\left(x, y, z \,\middle|\, \begin{array}{c} 0 \\ 1 \end{array}\right)$ de la méthode combinatoire analytique. L'étude des urnes à deux couleurs conduit ainsi par deux chemins d'apparence éloignés au même système différentiel monomial, mais les fonctions concernées sont de natures très différentes. D'un côté, les séries (formelles) génératrices trivariées des histoires de l'urne en temps discret sont les solutions du problème de Cauchy générique associé au système. De l'autre, les transformées de Fourier des lois limites des *grandes* urnes sont les solutions du système qui vérifient certaines conditions aux limites à l'infini, qui elles-mêmes traduisent la valeur des espérances et l'existence de seconds moments. Un lien direct entre ces deux occurrences du même système différentiel monomial reste à établir ...

References

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

[2] 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.

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