Le modèle de diffusion de Ehrenfest, activités élève, éléments de réponse

- Le problème
- Les connaissances et compétences mises en jeu (il y a plusieurs niveaux d'approfondissement)
- Quelques idées d'activités pour l'élève
- Quelques éléments de réponse

A lire attentivement : dans ce document les vecteurs sont notés en ligne, donc le produit obtenu par UM correspond à ${}^tM{}^tU$ (obtenu avec les notations classiques d'algèbre linéaire avec des vecteurs en colonne)

Le problème

On considère 2 urnes A et B. On dispose de N boules numérotées de 1 à N avec N ≥ 1. Au départ (à l'étape 0), les N boules sont dans l'urne A. A chaque étape, on tire au hasard un nombre entre 1 et N (avec équiprobabilité) et l'on change d'urne la boule portant le numéro tiré: par exemple, si on a tiré le nombre 4, on change la boule numéro 4 d'urne (si la boule 4 est dans A, on la retire de A et on la met dans B ; si la boule 4 est dans B, on la retire de B et on la met dans A).
On note X_n le nombre de boules dans l'urne A à la fin de la $n$ième étape.
La loi de X_n est donnée par un le vecteur $E_n = (p_0, p_1, \dots, p_N)$, où p_i désigne la probabilité d'avoir i boules dans A à l'issue des n transferts aléatoires.

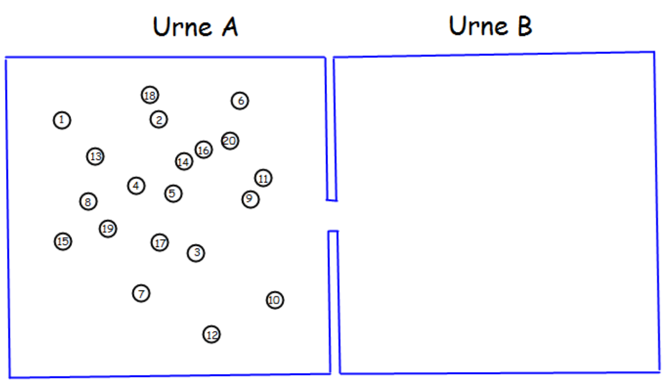


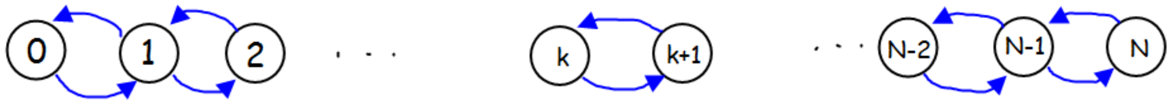
figure1

Les connaissances qui peuvent être mises en jeu (il y a plusieurs niveaux d'approfondissement)

Calcul sur les matrices, résolution de systèmes
Graphes probabilistes : matrice de transition, état stable
Etude de suites
Notions de probabilités : variable aléatoire, probabilités conditionnelles
TIC : tableur, algorithmique, calcul formel sur les matrices et sur les systèmes

Quelle activité mathématique pour l'élève ?

- Effectuer une simulation avec un **tableur** : observer, selon les valeurs de N et de n .Quelques pistes : valeur moyenne du nombre de boules dans A, stabilisation ou non de ce nombre, retours à l'état initial (toutes les boules dans A) de probabilité nulle ou non : sinon estimations de cette probabilité
- Ecrire et tester un **algorithme** de simulation, mêmes observations
- **Graphe probabiliste** :
  - porter les probabilités sur les arcs :



- donner la **matrice de transition** du graphe
- travailler sur les lois de probabilité des variables aléatoires X_n pour de petites valeurs de n, N étant quelconque
  Etude des lois de X_0, X_1 et X_2 , espérance, écart type

- Travailler sur **l'espérance de X_n** pour N fixé, et déterminer la limite de cette espérance quand n tend vers l'infini.
- Travailler sur **l'état stable** :

  Le déterminer pour de petites valeurs de N :

  Calculer l'état stable par résolution d'un système à la main ou avec le calcul formel

  Pour N quelconque : montrer ou admettre que l'état stable est la loi binomiale $B_{N,\frac{1}{2}}$

  Interpréter ce résultat : espérance, écart type, période estimée pour le retour à l'état initial.

Quelques éléments de réponse

- **Simulations avec tableur ou algorithmes et exploitations** : voir le document de B. Body et les fichiers Ehrenfest_1.xlsx et Ehrenfest2.xlsx et ehrenfest2.sce .
- **Pour les algorithmes**, on peut simuler les deux urnes dans une seule liste U remplie de 0 et de 1 :

Numéro de boule	1	2	3	4	5	6	7	8	9
urne	0	1	1	0	1	1	1	0	1

Si le $i$ème élément vaut 1, cela signifie que la $i$ème boule est dans A, sinon elle est dans B.

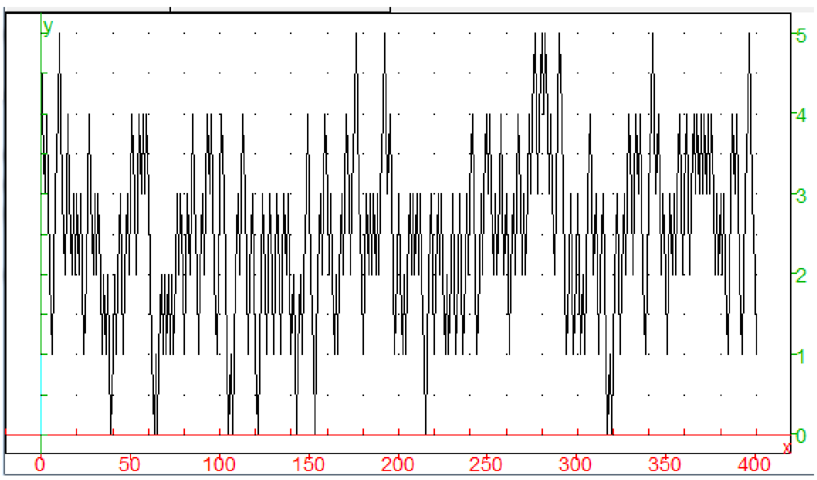
Ici, les boules 1,3 et 8 sont dans l'urne B

A chaque tirage aléatoire, la boule tirée numéro i change d'urne, donc le 1 devient 0 et vice et versa, cela se traduit par l'instruction : $U[i]$ prend la valeur $1 - U[i]$

Le nombre de boules dans A est alors modifié : si $U[i]$ valait 1, ce nombre est diminué de 1, sinon il est augmenté de 1.

Exemples de résultats :

Avec 5 boules et 400 transferts aléatoires entre les urnes A et B : on observe beaucoup de fluctuations et un certain nombre de retours à l'état initial (5 boules dans A)



Nombre de boules dans A en fonction du numéro du transfert aléatoire

Avec Xcas :

Ehrenfest1serie.xws

Ouvrir la fenêtre graphique DispG

Cliquer sur Auto

Zoommer « in » et augmenter l'échelle sur les ordonnées avec les flèches

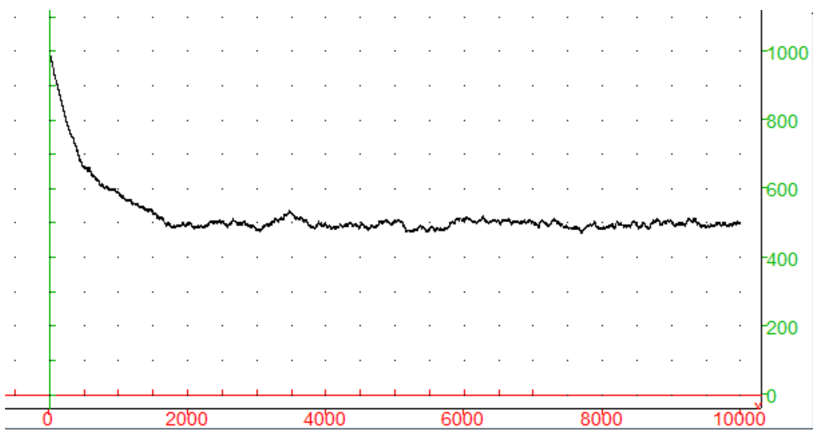
ou avec Scilab

ehrenfest1.sce

Figure 2

Avec 1000 boules et 10000 transferts aléatoires entre les urnes A et B : il semble qu'il y ait une stabilisation de la répartition des boules.

Figure 3



Il y a un certain **paradoxe** entre les deux graphiques : bien qu'il s'agisse du même problème, dans le premier on a des retours à l'état initial assez fréquents, dans le deuxième, on a du mal à penser que les N boules risquent de se retrouver une deuxième fois dans l'urne A, (au départ toutes les boules sont dans A).

Une question se pose alors : les retours à l'état initial sont-ils possibles (autrement dit, ont-ils une probabilité non nulle pour un N fixé), même si N est grand ?

Pour tenter d'y répondre à l'aide de simulations, on peut aussi observer le nombre moyen de transfert entre deux retours à l'état initial : avec N =2, puis N=3 etc., et l'on constate que cette amplitude vaut approximativement 4, puis 8, 16…

En abscisse : le nombre N de boules, en ordonnée le nombre moyen de transferts entre deux retours à l'état initial : fichier Xcas : Ehrenfest4RetoursEtatInitial

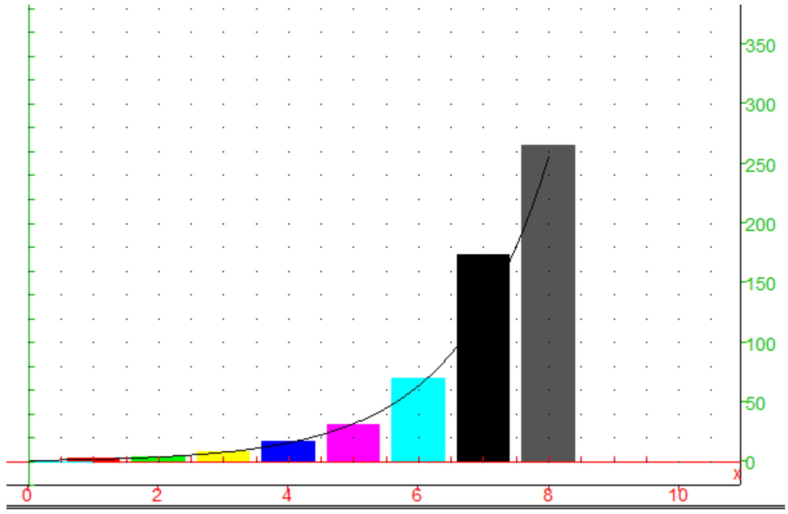


figure 4

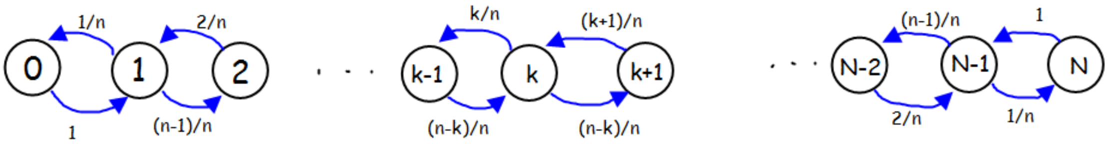
Quelques éléments d'étude théorique

Pour savoir à quoi s'en tenir par rapport au paradoxe observé, il est temps d'étudier théoriquement cette expérience aléatoire.

- **Graphe probabiliste**

Tout comme pour la matrice de transition, on peut commencer par de petites valeurs de N.

Dans ce graphe, chaque sommet représente le nombre possible de boules dans l'urne A



- **Matrice de transition** associée à ce graphe

On note X_n la variable aléatoire égale au nombre le nombre de boules dans l'urne A à la fin de la *n-ième* étape.

X_n prend ces valeurs dans {0 ;1 ;2… ;N}

La loi de X_n est donnée par un le vecteur $E_n = (p_0, p_1, \dots, p_N)$, où p_i désigne la probabilité d'avoir i boules dans A à l'issue des n transferts aléatoires.

L'état initial X_0 a pour loi de probabilité $E_0 = (0,0, \dots ,0,1)$, car au départ les n boules sont dans l'urne A.

Probabilité d'avoir 0 boule dans A au départ

Probabilité d'avoir N boules dans A au départ

La matrice M_N permet de connaître l'état X_{n+1} en fonction de X_n : $X_{n+1} = X_n M_N$

Avec : $M_N = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 & 0 & 0 \\ \frac{1}{N} & 0 & \frac{N-1}{N} & \dots & \dots & 0 & 0 & 0 \\ 0 & \frac{2}{N} & 0 & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{N} & \dots & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \dots & \frac{3}{N} & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & 0 & \frac{2}{N} & 0 \\ 0 & 0 & 0 & \dots & \dots & \frac{N-1}{N} & 0 & \frac{1}{N} \\ 0 & 0 & 0 & \dots & \dots & 0 & 1 & 0 \end{pmatrix}$

Cette matrice carrée est de dimension N+1. Elle s'interprète par ligne :

La première ligne donne les probabilités d'avoir 0, 1, …, ou N boules dans l'urne A à l'étape suivante, sachant qu'il y en a 0 dans A.
La deuxième ligne donne les probabilités d'avoir 0, 1, …, ou N boules dans l'urne A à l'étape suivante, sachant qu'il y en a 1 dans A, etc.
Avec $1 \leq k \leq N-1$, la kième ligne fournit les probabilités conditionnelles : Sachant qu'il y a k boules dans A (on est sur le sommet numéroté k du graphe): à l'étape suivante, il y en aura k-1 dans A avec la probabilité $\frac{k}{N}$, ou k+1 avec la probabilité $\frac{N-k}{N}$.

La somme des coefficients de chaque ligne vaut 1.

Ainsi, le produit matriciel donne bien la probabilité qu'il y ait k boules dans A à la (n+1)ème étape :

$$P(X_{n+1} = k) = P(X_n = k-1) \times \frac{N-(k-1)}{N} + P(X_n = k+1) \times \frac{(k+1)}{N}$$

- **Quelques exemples d'étude de lois de probabilité des variables aléatoires X_n**

On peut se reporter au [document de B. Body]

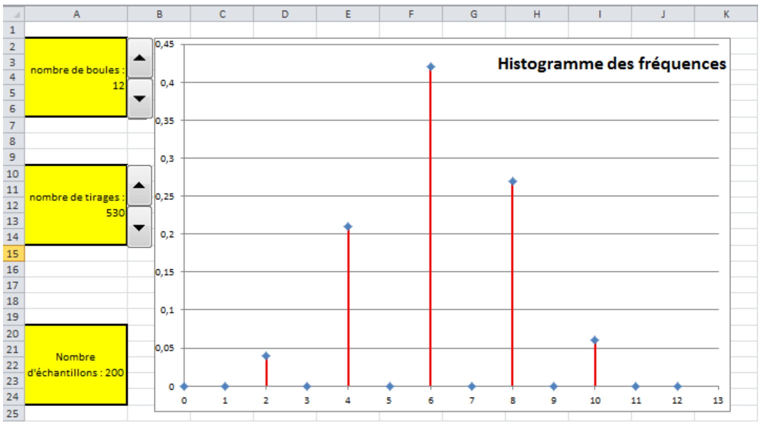


figure 5

Ehrenfest2.xlsx et ehrenfest2.sce .

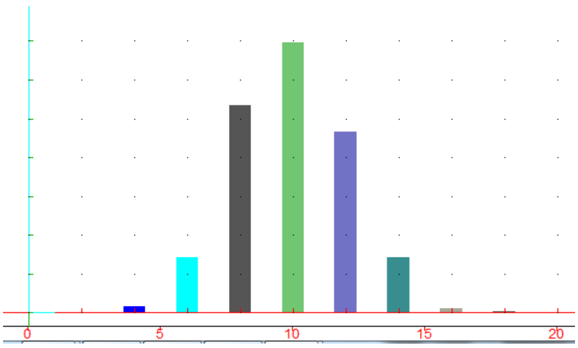
A noter qu'une fois définie la matrice de transition M : on peut obtenir la loi de probabilité de X_n dans le vecteur E_n **récursivement**, par exemple avec N=4 :

```
E0:=[0,0,0,0,1]
E(n):=si n=0 alors E0 sinon E(n-1)*M fsi;
```

ou de façon **itérative**, ce qui est beaucoup plus efficace pour la gestion des mémoires :

```
E0:=[0,0,0,0,1] ;
Etat :=E0 ;
pour k variant de 1 à n
faire
    Etat :=Etat*M
fin pour
```

ou encore avec le calcul formel : E0*M^n



Simulations avec Xcas : ouvrir la fenêtre DispG, et cliquer sur Auto, pour avoir des unités correctes

Fichier : Ehrenfest2EtudeXn.xws

Etude des espérances des variables aléatoires X_n

Pour un nombre N de boules fixées, on peut montrer la formule de récurrence reliant les espérances de X_n et X_{n+1} :

$$E(X_n) = \left(1-\frac{2}{N}\right)E(X_{n-1})+1 \quad , \quad \text{d'où} \quad E(X_n)-\frac{N}{2} = \left(1-\frac{2}{N}\right)^n \times \left(E(X_{n-1})-\frac{N}{2}\right)$$

Cette formule peut être admise en terminale, sa démonstration étant un peu technique, avec des manipulations d'indices dans des sommes. Elle peut être envisagée avec les élèves, avec de petites valeurs de N, par exemple N=4 dans le document de B. Body.

La suite des espérances est arithmético-géométrique, et il n'est pas difficile de montrer qu'elle converge exponentiellement vers $\frac{N}{2}$, la limite ne surprend pas l'intuition.

- **Travailler sur l'état stable**

1. On peut commencer par déterminer l'état stable pour de petites valeurs de N, en résolvant des systèmes linéaires « à la main ». voir le document de B. Body .
   On se rend compte que l'on a un système linéaire à N+1 inconnues $p_0, p_1 \dots p_N$. On dispose de N+2 équations :
   - N+1 équation provenant de l'égalité $E = E * M$
   - Une N+2 ième équation: $\sum_{i=0}^{i=N} p_i = 1$

   et en prenant par exemple p_0 pour paramètre, ce système est triangulaire sur les N premières équations, comme on le verra la N+1 ième équation est compatible avec les précédentes, mais donne un système indéterminé. Par contre la somme de ces probabilités vaut 1, d'où la valeur de p_0 et la résolution du système[1]. Cette démarche assure l'unicité d'une solution. L'existence et la détermination sont vues au paragraphe 2.

2. Détermination de l'état stable, un résultat fort, mais accessible :

$$\boxed{\text{l'état stable existe et suit la loi binomiale } B_{N,\frac{1}{2}}}$$

Pour N fixé, toujours en notant M la matrice de transition du graphe, chercher l'état stable $(p_0, p_1, \dots, p_N)$, s'il existe, consiste à résoudre le système à N+1 équations et N+1 inconnues $E = E * M$ avec comme équation supplémentaire : $\sum_{i=0}^{i=N} p_i = 1$. Comme on l'a vu précédemment, si ce système possède une solution, alors elle est unique.

On peut vérifier que la loi binomiale vérifie ce système ! (ceci nécessite la connaissance des $\binom{n}{p}$ à l'aide des factorielles).

Utilisation du calcul formel : résolution de systèmes

Compte tenu de ce qui précède, on peut résoudre le système avec un instrument de calcul ou un logiciel :

On garde les N premières équations, et prend comme N+1ième équation, celle qui donne la somme des probabilités

```
1  Prog  Edit  Ajouter   2   nxt                OK (F9)
resoudre_systeme_lineaire(
//  recherche de l'état stable avec N=4, liste des équations
   [    p1                =4*p0,
     -2*p1+  p2           =-2*p0,
      3*p1-4*p2+3*p3      =0,
             p2-2*p3+2*p4 =0, // fin de la partie triangulaire
            p0+p1+p2+p3+p4=1   ],// équation supplémentaire
// liste des inconnues
            [p0,p1,p2,p3,p4])
```

$$\left[\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\right]$$

Ehrenfest etat stable M4_2.xws

- Pour de petites valeurs de N, on peut résoudre le système donnant l'état stable avec une calculatrice.
- Pour des valeurs plus grandes de N : on peut utiliser un logiciel de calcul formel comme Xcas ou WxMaxima…
- Pour de grandes valeurs de N, en approfondissement, on pourrait envisager l'utilisation d'une matrice creuse (ici M comporte beaucoup de zéros) et de programmer la résolution du système.
- La suite des vecteurs (E_n) ne converge pas vers l'état stable, puisque selon la parité de n, un terme sur deux de E_n est nul, comme on a déjà pu l'observer sur les diagrammes en bâtons figure 5, et comme on pourrait le prouver en comparant les parités de n et N.

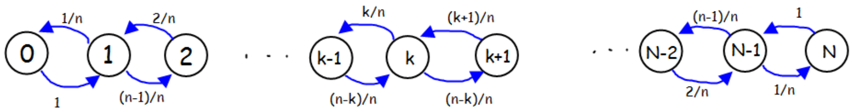
[1] En terme d'algèbre linéaire, le sous-espace propre associé à la valeur propre 1 de la matrice M est de dimension 1, mais ici on a une équation supplémentaire donnée par la somme des probabilités.

- **Retours à l'état initial, pour N fixé**

  - Si l'état initial est l'état stable, (autrement dit : au départ la distribution des boules dans les urnes A et B est aléatoire, de façon que la probabilité d'avoir k boules dans A suive $B_{N,\frac{1}{2}}$, ce qui est facile à obtenir en laissant tomber une bille sur une planche de Galton), alors la probabilité d'obtenir toutes les boules dans A est $\frac{1}{2^N}$, ce qui signifie qu'en moyenne, sur un grand nombre de transferts d'une urne à l'autre, il faut 2^N transferts de boules pour observer un retour à l'état initial[2]. Ce que l'on peut déjà observer sur la figure 4

Autrement dit l'impression de stabilisation observée à la figure 3 n'est qu'illusoire !

  - Si l'état initial n'est pas l'état stable, le résultat ci-dessus est encore vrai pour deux raisons :



. dans le graphe probabiliste : tout sommet peut être atteint à partir de n'importe quel état initial, autrement dit, si k est un entier entre 0 et N, la probabilité d'avoir k boules dans l'urne A au bout d'un certain nombre d'étapes est non nulle.
. La configuration des deux urnes A et B à l'étape n+1 ne dépend que de celle à l'étape n, et de la matrice de transition : on n'a pas besoin de connaitre ce qui s'est passé avant l'étape n : autrement dit, cette configuration à l'étape n aurait très bien pu provenir de l'état initial stable.

Grands nombres

Si l'on met une mole d'oxygène dans l'urne A, et que l'on fait le vide dans l'urne B, en admettant que les atomes se déplacent selon le modèle des urnes de Ehrenfest, il faudrait en moyenne $2^{6,02\times10^{23}}$ transferts avant d'observer de nouveau tous les atomes dans A. Même si les déplacements se faisaient sur quelques centimètres, l' un après l'autre et à la vitesse de la lumière, il faudrait un temps certain qui dépasse largement l'âge estimé de l'univers, 15 milliards d'années !

Kylie Ravera écrit dans un article de la revue Tangente (hors-série n°43, septembre 2011) :
« […] Au final, la superposition d'un très grand nombre d'états réversibles peut conduire à un état en pratique irréversible ».

[2] On peut généraliser à un nombre fixé k de boules dans A : avec l'état stable la probabilité d'avoir k boules dans A est $\binom{N}{k}\left(\frac{1}{2}\right)^n$ d'après la loi binomiale, donc en moyenne il faut $\frac{2^N(k!(N-k)!)}{N!}$ étapes pour retrouver k boules dans A.