Lycée Clemenceau **2025-26**
ECG 2

Corrigé du Devoir Maison de Mathématiques $n^o 5$

Corrigé de l'exercice 1. Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ la fonction définie pour tout $x \in \mathbb{R}$ par $f(x) = \ln\left(x + \sqrt{1+x^2}\right)$.

(1) (a) Calculons le développement limité de f à l'ordre 3 en 0. D'après le cours, on sait que :

$$\begin{aligned}\sqrt{1+y} &\underset{y\to 0}{=} 1 + \frac{1}{2}y + \frac{1}{2}\left(\frac{1}{2}-1\right)\frac{y^2}{2!} + \frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)\frac{y^3}{3!} + o(y^3) \\ &\underset{y\to 0}{=} 1 + \frac{y}{2} - \frac{y^2}{8} + \frac{3y^3}{48} + o(y^3).\end{aligned}$$

Si l'on pose $y = x^2$, alors on voit que y tend vers 0 quand x tend vers 0. Dès lors, par substitution puis élimination des termes de degré > 3, on obtient que :

$$\sqrt{1+x^2} \underset{x\to 0}{=} 1 + \frac{x^2}{2} - \frac{x^4}{8} + \frac{3x^6}{48} + o(x^3) \underset{x\to 0}{=} 1 + \frac{x^2}{2} + o(x^3).$$

Posons maintenant $z = x + \sqrt{1+x^2} - 1$. D'après le calcul précédent, on trouve que :

$$z = x + \sqrt{1+x^2} - 1 \underset{x\to 0}{=} x + \frac{x^2}{2} + o(x^3).$$

Comme $1 + z = x + \sqrt{1+x^2}$ et que de plus z tend vers 0 quand x tend vers 0, on obtient par substitution, puis par utilisation de la formule du binôme et enfin par élimination des termes de degré > 3 que :

$$\begin{aligned}f(x) &= \ln(1+z) \\ &\underset{z\to 0}{=} z - \frac{z^2}{2} + \frac{z^3}{3} + o(z^3) \\ &\underset{x\to 0}{=} \left(x + \frac{x^2}{2}\right) - \frac{1}{2}\left(x + \frac{x^2}{2}\right)^2 + \frac{1}{3}\left(x + \frac{x^2}{2}\right)^3 + o(x^3) \\ &\underset{x\to 0}{=} x + \frac{x^2}{2} - \frac{1}{2}\left(x^2 + \frac{x^4}{4} + x^3\right) + \frac{1}{3}\left(x^3 + \frac{3x^4}{4} + \frac{3x^5}{4} + \frac{x^6}{8}\right) + o(x^3) \\ &\underset{x\to 0}{=} x + \frac{x^2}{2} - \frac{1}{2}\left(x^2 + x^3\right) + \frac{1}{3}\left(x^3\right) + o(x^3) \\ &\underset{x\to 0}{=} x + \frac{x^2}{2} - \frac{x^2}{2} - \frac{x^3}{2} + \frac{x^3}{3} + o(x^3),\end{aligned}$$

d'où il s'ensuit que :

$$\boxed{f(x) \underset{x\to 0}{=} x - \frac{x^3}{6} + o(x^3).}$$

(b) D'après le résultat de la question précédente, on en déduit que l'équation de la tangente à la courbe de f en 0 est donnée par :

$$\boxed{y = x.}$$

De plus, comme $f(x) \underset{x\to 0}{=} x - \dfrac{x^3}{6} + o(x^3)$, il s'ensuit que $f(x) - x \underset{x\to 0}{=} -\dfrac{x^3}{6} + o(x^3)$, et donc :

$$f(x) - x \underset{x\to 0}{\sim} -\frac{x^3}{6}.$$

En particulier, le signe de $f(x) - x$ est donné par celui de $-x^3/6$ au voisinage de 0, c'est-à-dire $f(x) - x > 0$ pour $x < 0$ proche de 0, et $f(x) - x < 0$ pour $x > 0$ proche de 0. En d'autres termes, si $\mathcal{C}_f$ désigne la courbe de f et $\mathcal{T}_0$ sa tangente en 0, alors :

$$\boxed{\begin{array}{l}\mathcal{C}_f \text{ est située au dessus de } \mathcal{T}_0 \text{ pour } x < 0 \text{ proche de } 0, \text{en dessous de } \mathcal{T}_0 \\ \text{pour } x < 0 \text{ proche de } 0, \text{et } (0,0) \text{ est un point d'inflexion de } \mathcal{C}_f.\end{array}}$$

(2) (a) Calculons $f'(x)$ pour tout $x \in \mathbb{R}$. D'après les règles de dérivation classiques, on trouve que :

$$\begin{aligned} f'(x) &= \left(\ln(x+\sqrt{1+x^2})\right)' = \frac{(x+\sqrt{1+x^2})'}{x+\sqrt{1+x^2}} \\ &= \frac{1+\frac{2x}{2\sqrt{1+x^2}}}{x+\sqrt{1+x^2}} = \frac{1+\frac{x}{\sqrt{1+x^2}}}{x+\sqrt{1+x^2}} \\ &= \frac{\frac{\sqrt{1+x^2}+x}{\sqrt{1+x^2}}}{x+\sqrt{1+x^2}} = \frac{1}{\sqrt{1+x^2}} \times \frac{\sqrt{1+x^2}+x}{x+\sqrt{1+x^2}}, \end{aligned}$$

d'où il s'ensuit que :

$$\boxed{\forall x \in \mathbb{R}, \quad f'(x) = \frac{1}{\sqrt{1+x^2}}.}$$

De même, calculons $f''(x)$ pour tout $x \in \mathbb{R}$. D'après les règles de dérivation classiques et l'expression de $f'(x)$ trouvée ci-dessus, on obtient que :

$$\begin{aligned} f''(x) &= \left(\frac{1}{\sqrt{1+x^2}}\right)' = \left((1+x^2)^{-\frac{1}{2}}\right)' \\ &= 2x \times -\frac{1}{2} \times (1+x^2)^{-\frac{1}{2}-1} = -x(1+x^2)^{-\frac{3}{2}}, \end{aligned}$$

d'où il s'ensuit que :

$$\boxed{\forall x \in \mathbb{R}, \quad f''(x) = \frac{-x}{(\sqrt{1+x^2})^3}.}$$

(b) D'après la formule de Taylor-Young appliquée à la fonction f à l'ordre 2 en 1, on trouve que :

$$\begin{aligned} f(x) &\underset{x\to 1}{=} f(1) + f'(1)(x-1) + \frac{f''(1)}{2!}(x-1)^2 + o((x-1)^2) \\ &\underset{x\to 1}{=} \ln(1+\sqrt{1+1^2}) + \frac{1}{\sqrt{1+1^2}}(x-1) + \frac{-1}{2(\sqrt{1+1^2})^3}(x-1)^2 + o((x-1)^2), \end{aligned}$$

d'où il s'ensuit que :

$$\boxed{f(x) \underset{x\to 1}{=} \ln(1+\sqrt{2}) + \frac{1}{\sqrt{2}}(x-1) - \frac{1}{4\sqrt{2}}(x-1)^2 + o((x-1)^2).}$$

(c) D'après le résultat de la question précédente, on en déduit que l'équation de la tangente à la courbe de f en 1 est donnée par :

$$\boxed{y = \ln(1+\sqrt{2}) + \frac{1}{\sqrt{2}}(x-1).}$$

De plus, d'après la question (2)(b) et par les mêmes arguments qu'à la question (1)(c), on a :

$$f(x) - \ln(1+\sqrt{2}) - \frac{1}{\sqrt{2}}(x-1) \underset{x\to 1}{\sim} -\frac{1}{4\sqrt{2}}(x-1)^2.$$

En particulier, le signe de $f(x) - \ln(1+\sqrt{2}) - (x-1)/\sqrt{2}$ est donné par celui de $-(x-1)^2$ au voisinage de 1, c'est-à-dire $f(x) - \ln(1+\sqrt{2}) - (x-1)/\sqrt{2} \leq 0$ pour tout x proche de 1. En d'autres termes, si $\mathcal{T}_1$ désigne la tangente à la courbe de f en 1, alors :

$$\boxed{\mathcal{C}_f \text{ est située en dessous de } \mathcal{T}_1 \text{ pour tout } x \text{ proche de 1.}}$$

Corrigé de l'exercice 2. A l'aide des développements limités, calculons :

$$\lim_{x\to 0} \frac{xe^x - \ln(1+x) + 3x^2}{1-\cos(x)}.$$

Pour ce faire, on commence par calculer des développements limités à l'ordre 2 en 0 des numérateur et dénominateur de cette fraction. D'après les règles de calcul des développements limités, on trouve que :

$$\begin{aligned} xe^x - \ln(1+x) + 3x^2 & \underset{x\to 0}{=} x\left(1 + x + \frac{x^2}{2} + o(x^2)\right) - \left(x - \frac{x^2}{2} + o(x^2)\right) + 3x^2 \\ & \underset{x\to 0}{=} x + x^2 + \frac{x^3}{2} + o(x^2) - x + \frac{x^2}{2} + o(x^2) + 3x^2 \\ & \underset{x\to 0}{=} \frac{9x^2}{2} + o(x^2). \end{aligned}$$

De même, comme $\cos(x) \underset{x\to 0}{=} 1 - \frac{x^2}{2} + o(x^2)$, on voit que $1 - \cos(x) \underset{x\to 0}{=} \frac{x^2}{2} + o(x^2)$, et donc :

$$\frac{xe^x - \ln(1+x) + 3x^2}{1 - \cos(x)} \underset{x\to 0}{=} \frac{\frac{9x^2}{2} + o(x^2)}{\frac{x^2}{2} + o(x^2)} = \frac{9 + o(1)}{1 + o(1)}.$$

Comme les termes de la forme "$o(1)$" correspondent à des fonctions qui tendent vers 0 quand x tend vers 0, on en déduit que :

$$\boxed{\lim_{x\to 0} \frac{xe^x - \ln(1+x) + 3x^2}{1 - \cos(x)} = 9.}$$

Corrigé de l'exercice 3. Soit n un entier ≥ 2. On considère un endomorphisme f de $\mathbb{R}^n$, dont la matrice dans la base canonique de $\mathbb{R}^n$ est une matrice M de rang 1. On note C la première colonne de M et on suppose que C est non nulle.

(1) Donnons tout d'abord la dimension de $\ker(f)$. Comme M est la matrice de f dans la base canonique de $\mathbb{R}^n$, on voit que :

$$\mathrm{rg}(f) = \mathrm{rg}(M) = 1.$$

D'après le théorème du rang, ceci nous donne que :

$$\dim \ker(f) = n - \mathrm{rg}(f) = n - 1.$$

Dès lors, comme $n \geq 2$, on voit que $\dim \ker(f) = n - 1 > 0$. Par conséquent, on en déduit que :

$$\boxed{\dim \ker(f) = n - 1 \text{ et } 0 \text{ est valeur propre de } f.}$$

(2) (a) Montrons qu'il existe une matrice $L = \begin{pmatrix} 1 & l_2 & \ldots & l_n \end{pmatrix}$ appartenant à $\mathcal{M}_{1,n}(\mathbb{R})$ telle que $M = CL$. Comme la matrice M est de rang 1, ses colonnes $C_1, ..., C_n$ sont colinéaires entre elles. Dès lors, comme $C = C_1$ est non nulle, toutes les colonnes $C_2, ..., C_n$ sont colinéaires à C, et donc il existe des réels $l_2, ..., l_n$ tels que $C_i = l_i C$ pour tout $i \in [\![2, n]\!]$. En particulier, si $c_1, ..., c_n$ sont les composantes de C, alors la matrice M est de la forme :

$$M = \begin{pmatrix} C & | & l_2 C & | & \ldots & | & l_n C \end{pmatrix} = \begin{pmatrix} c_1 & l_2 c_1 & \ldots & l_n c_1 \\ c_2 & l_2 c_2 & \ldots & l_n c_2 \\ \vdots & \vdots & & \vdots \\ c_n & l_2 c_n & \ldots & l_n c_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \begin{pmatrix} 1 & l_2 & \ldots & l_n \end{pmatrix}.$$

Mais comme $C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$, on en déduit que :

$$\boxed{\text{il existe une matrice } L = \begin{pmatrix} 1 & l_2 & \ldots & l_n \end{pmatrix} \text{ telle que } M = CL.}$$

(b) Vérifions que $\mathrm{Tr}(M) = LC$. D'après la question précédente, on voit que :

$$\mathrm{Tr}(M) = \mathrm{Tr}\left(\begin{pmatrix} c_1 & l_2 c_1 & \ldots & l_n c_1 \\ c_2 & l_2 c_2 & \ldots & l_n c_2 \\ \vdots & \vdots & & \vdots \\ c_n & l_2 c_n & \ldots & l_n c_n \end{pmatrix}\right) = c_1 + l_2 c_2 + \ldots + l_n c_n.$$

En outre, on trouve avec le produit ligne-colonne que :

$$\begin{pmatrix} 1 & l_2 & \dots & l_n \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = 1 \times c_1 + l_2c_2 + \dots + l_nc_n.$$

Par conséquent, on en déduit que :

$$\boxed{\mathrm{Tr}(M) = LC.}$$

(c) Etablissons l'égalité $M^2 = \mathrm{Tr}(M)M$. D'après la question (2)(a), on voit par associativité du produit matriciel que :

$$M^2 = MM = CLCL = C(LC)L.$$

Comme LC est un réel, ceci entraine avec la question (2)(b) que :

$$M^2 = C(LC)L = (LC)CL = (LC)M = \mathrm{Tr}(M)M.$$

Par conséquent, on en déduit que :

$$\boxed{M^2 = \mathrm{Tr}(M)M.}$$

(3) Montrons que $\mathrm{Tr}(M)$ est une valeur propre de f. D'après la question (2)(c), on sait que $M^2 = \mathrm{Tr}(M)M$. Comme f est l'endomorphisme canoniquement associé à M, ceci entraine que $f^2 = \mathrm{Tr}(M)f$. A noter que, comme f est de rang 1 d'après la question (1), f n'est pas l'endomorphisme nul. En particulier, il existe un vecteur $x \in \mathbb{R}^n$ tel que $f(x) \neq 0$. Dès lors, ceci nous donne que :

$$f(f(x)) = f^2(x) = \mathrm{Tr}(M)f(x).$$

Comme $f(x) \neq 0$, il s'ensuit que $f(x)$ est un vecteur propre de f pour la valeur propre $\mathrm{Tr}(M)$. Par conséquent, on en déduit que :

$$\boxed{\mathrm{Tr}(M) \text{ est une valeur propre de } f.}$$

(4) On suppose que $\mathrm{Tr}(M) = 0$. Montrons que f n'est pas diagonalisable. Comme $M^2 = \mathrm{Tr}(M)M$ d'après la question (2)(c) et que $\mathrm{Tr}(M) = 0$ par hypothèse, on voit que $M^2 = 0$, et donc le polynôme $P : x \longmapsto x^2$ est annulateur de M. Comme f est l'endomorphisme canoniquement associé à M, ceci entraine que P est aussi annulateur de f. En particulier, on obtient que $\mathrm{Sp}(f) \subset \{0\}$. De plus, comme 0 est valeur propre de f d'après la question (1), il s'ensuit que $\mathrm{Sp}(f) = \{0\}$. Qui plus est, on voit avec la question (1) que :

$$\dim E_0(f) = \dim \ker(f) = n - 1 \neq \dim \mathbb{R}^n.$$

Par conséquent, on en déduit que :

$$\boxed{\text{si } \mathrm{Tr}(M) = 0, \text{ alors } f \text{ n'est pas diagonalisable.}}$$

(5) On suppose que $\mathrm{Tr}(M) \neq 0$. Déterminons tout d'abord les valeurs propres de f. D'après la question (2)(c), on sait que $M^2 = \mathrm{Tr}(M)M$. Comme f est l'endomorphisme canoniquement associé à M, ceci entraine que $f^2 = \mathrm{Tr}(M)f$, et donc le polynôme $P : x \longmapsto x^2 - \mathrm{Tr}(M)x$ est annulateur de f. En particulier, on obtient que $\mathrm{Sp}(f) \subset \{0, \mathrm{Tr}(M)\}$. Dès lors, comme 0 et $\mathrm{Tr}(M)$ sont valeurs propres de f d'après les questions (1) et (3), il s'ensuit que :

$$\boxed{\mathrm{Sp}(f) = \{0, \mathrm{Tr}(M)\}.}$$

A présent, montrons que f est diagonalisable. D'après la question (1), on sait déjà que :

$$\dim E_0(f) = \dim \ker(f) = n - 1.$$

De plus, comme $\mathrm{Tr}(M)$ est une valeur propre de f, on voit que $E_{\mathrm{Tr}(M)}(f)$ n'est pas réduit à $\{0\}$, et donc $\dim E_{\mathrm{Tr}(M)}(f) \geq 1$. En particulier, comme $\mathrm{Tr}(M) \neq 0$ par hypothèse, ceci nous donne que :

$$\dim E_{\mathrm{Tr}(M)}(f) + \dim E_0(f) \geq 1 + n - 1 = n.$$

Par ailleurs, on sait d'après le cours que $\dim E_{\mathrm{Tr}(M)}(f) + \dim E_0(f) \leq n$, et donc :

$$\dim E_{\mathrm{Tr}(M)}(f) + \dim E_0(f) = n.$$

Par conséquent, on en déduit que :

$$\boxed{\text{si } \mathrm{Tr}(M) \neq 0, \text{ alors } f \text{ est diagonalisable.}}$$

A présent, on se fixe trois réels a, b, c non nuls, et on considère l'endomorphisme g de $\mathbb{R}^3$ dont la matrice dans la base canonique de $\mathbb{R}^3$ est donnée par :

$$A = \begin{pmatrix} 1 & 1/a & 1/b \\ a & 1 & 1/c \\ b & c & 1 \end{pmatrix}.$$

Par la suite, on suppose que A n'est pas inversible.

(1) Ecrivons une fonction en Python qui, étant donnés trois réels a, b, c non nuls, construit et affiche la matrice A. Pour ce faire, on procède comme suit :

```
import numpy as np

def matrice(a,b,c):
    m=np.array([[1,1/a,1/b],[a,1,1/c],[b,c,1]])
    return m
```

(2) (a) Etablissons que $ac = b$. Pour ce faire, on raisonne par l'absurde et on suppose que $ac \neq b$. Considérons alors le système $AX = 0$, où X a pour composantes x, y, z. Alors ce système peut se réécrire sous la forme :

$$\left\{ \begin{array}{ccccccc} x & + & \frac{1}{a}y & + & \frac{1}{b}z & = & 0 \\ ax & + & y & + & \frac{1}{c}z & = & 0 \\ bx & + & cy & + & z & = & 0 \end{array} \right. .$$

En effectuant les opérations élémentaires $L_2 \leftarrow L_2 - aL_1$ et $L_3 \leftarrow L_3 - bL_1$, on trouve que :

$$\left\{ \begin{array}{ccccccc} x & + & \frac{1}{a}y & + & \frac{1}{b}z & = & 0 \\ & & & & \left(\frac{1}{c} - \frac{a}{b}\right) z & = & 0 \\ & & \left(c - \frac{b}{a}\right) y & & & = & 0 \end{array} \right. .$$

En effectuant l'opération élémentaire $L_2 \leftrightarrow L_3$, on obtient que :

$$\left\{ \begin{array}{ccccccc} x & + & \frac{1}{a}y & + & \frac{1}{b}z & = & 0 \\ & & \left(c - \frac{b}{a}\right) y & & & = & 0 \\ & & & & \left(\frac{1}{c} - \frac{a}{b}\right) z & = & 0 \end{array} \right. .$$

Comme b et c sont non nuls par hypothèse, on trouve en effectuant les opérations élémentaires $L_2 \leftarrow cL_2$ et $L_3 \leftarrow bcL_3$ que :

$$\left\{ \begin{array}{ccccccc} x & + & \frac{1}{a}y & + & \frac{1}{b}z & = & 0 \\ & & (ac - b)\, y & & & = & 0 \\ & & & & -(b - ac)\, z & = & 0 \end{array} \right. .$$

Comme A n'est pas inversible, le système $AX = 0$ admet une solution non nulle, ce qui n'est possible d'après ce qui précède que si $ac - b = 0$. Mais ceci est impossible car $ac \neq b$ par hypothèse. Par conséquent, on en déduit que :

$$\boxed{ac = b.}$$

(b) Déterminons le rang de A. Par définition, on voit que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1/a & 1/b \\ a & 1 & 1/c \\ b & c & 1 \end{pmatrix}.$$

En effectuant les opérations élémentaires $L_2 \leftarrow L_2 - aL_1$ et $L_3 \leftarrow L_3 - bL_1$, on trouve que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1/a & 1/b \\ 0 & 0 & 1/c - a/b \\ 0 & c - b/a & 0 \end{pmatrix}.$$

En effectuant l'opération élémentaire $L_2 \leftrightarrow L_3$, on obtient que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1/a & 1/b \\ 0 & c - b/a & 0 \\ 0 & 0 & 1/c - a/b \end{pmatrix}.$$

Comme b et c sont non nuls par hypothèse, on trouve en effectuant les opérations élémentaires $L_2 \leftarrow cL_2$ et $L_3 \leftarrow bcL_3$ que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1/a & 1/b \\ 0 & ac - b & 0 \\ 0 & 0 & b - ac \end{pmatrix}.$$

Comme $ac = b$ d'après la question précédente, ceci entraine que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 1/a & 1/b \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Comme a et b sont non nuls par hypothèse, on trouve en effectuant les opérations élémentaires $C_2 \leftarrow aC_2 - C_1$ et $C_3 \leftarrow bC_3 - C_1$ que :

$$\operatorname{rg}(A) = \operatorname{rg}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1.$$

Par conséquent, on en déduit que :

$$\boxed{\operatorname{rg}(A) = 1.}$$

(3) (a) Montrons que g est diagonalisable et donnons ses valeurs propres. Par définition de la matrice A, on voit que $\operatorname{Tr}(A) = 3 \neq 0$. Dès lors, comme $\operatorname{rg}(A) = 1$ d'après la question précédente, que la première colonne de A est non nulle et que g est l'endomorphisme canoniquement associé à A par définition, on obtient d'après la question (5) de la première partie que :

$$\boxed{g \text{ est diagonalisable et } \operatorname{Sp}(g) = \{0, 3\}.}$$

(b) Montrons par récurrence la propriété $\mathcal{P}$ définie pour tout $n \in \mathbb{N}^*$ par :

$$\mathcal{P}(n) : \text{"}A^n \in \operatorname{Vect}(A)\text{"}.$$

Tout d'abord, on voit que $\mathcal{P}(1)$ est vraie, car $A^1 = A$ appartient à $\operatorname{Vect}(A)$. A présent, supposons la propriété $\mathcal{P}$ vraie à l'ordre $n \geq 1$, et montrons-la à l'ordre $n+1$. Par hypothèse de récurrence, on sait que A^n appartient à $\operatorname{Vect}(A)$, et donc il existe un réel λ_n tel que $A^n = \lambda_n A$. Comme $\operatorname{rg}(A) = 1$ d'après la question (2)(b), que la première colonne de A est non nulle et que $\operatorname{Tr}(A) = 3$, on voit que $A^2 = 3A$ d'après la question (2)(c) de la première partie, et donc :

$$A^{n+1} = A^n A = \lambda_n A A = \lambda_n A^2 = 3\lambda_n A,$$

d'où il s'ensuit que $\mathcal{P}(n+1)$ est vraie. D'après le principe de récurrence, la propriété $\mathcal{P}$ est vraie à tout ordre $n \geq 1$. Par conséquent, on en déduit que :

$$\boxed{\forall n \in \mathbb{N}^*, \quad A^n \in \operatorname{Vect}(A).}$$

Corrigé du problème 1. Soit r un entier ≥ 2. Une urne contient r boules numérotées de 1 à r. On pioche indéfiniment les boules avec remise, chaque boule pouvant être piochée de façon équiprobable. Pour tout $i \in \{1, ..., r\}$, on désigne par Y_i le nombre de pioches nécessaires pour obtenir i boules distinctes. Par convention, on pose $Y_1 = 1$. De même, on désigne par X_r le nombre de pioches nécessaires pour obtenir les r boules numérotées. Il est clair que $X_r = Y_r$. Par exemple, en supposant que $r = 4$ et si les boules piochées portent les numéros : $3,3,3,1,1,1,1,2,3,2,4,1,...$, alors on voit que : $Y_1 = 1$, $Y_2 = 4$, $Y_3 = 8$, $Y_4 = X_4 = 11$.

(1) **Partie I : résultats préliminaires.**
On considère la suite $(u_n)_{n\geq 1}$ définie pour tout $n \geq 1$ par : $u_n = \left(\sum_{k=1}^{n} \frac{1}{k}\right) - \ln(n)$.

(a) Ecrivons une fonction en Python qui, étant donné un entier $n \geq 1$, calcule et affiche u_n. Pour ce faire, on procède comme suit :

```
import numpy as np

def prob(n):
    u=np.sum(1/np.arange(1,n+1))-np.log(n)
    return u
```

(b) A l'aide d'un développement limité, montrons que $u_n - u_{n+1} \underset{n\to+\infty}{\sim} \frac{1}{2n^2}$. Par des calculs simples et à l'aide des développements limités de $\ln(1+x)$ et de $\frac{1}{1+x}$ à l'ordre 2 en 0, on a :

$$\begin{aligned}
u_n - u_{n+1} &= \left(\sum_{k=1}^{n} \frac{1}{k}\right) - \ln(n) - \left(\sum_{k=1}^{n+1} \frac{1}{k}\right) + \ln(n+1) \\
&= \ln(n+1) - \ln(n) - \frac{1}{n+1} \\
&= \ln\left(1 + \frac{1}{n}\right) - \frac{1}{n} \times \frac{1}{1+\frac{1}{n}} \\
&\underset{n\to+\infty}{=} \frac{1}{n} - \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right) - \frac{1}{n}\left(1 - \frac{1}{n} + \frac{1}{n^2} + o\left(\frac{1}{n^2}\right)\right) \\
&\underset{n\to+\infty}{=} \frac{1}{2n^2} + o\left(\frac{1}{n^2}\right).
\end{aligned}$$

Par conséquent, on en déduit que :

$$\boxed{u_n - u_{n+1} \underset{n\to+\infty}{\sim} \frac{1}{2n^2}.}$$

(c) Déterminons la nature de la série $\sum_{n\geq 1}(u_n - u_{n+1})$. D'après la question précédente, on a :

$$|u_n - u_{n+1}| \underset{n\to+\infty}{\sim} \frac{1}{2n^2}.$$

Comme la série de Riemann $\sum \frac{1}{n^2}$ converge d'après le cours, la série $\sum \frac{1}{2n^2}$ converge aussi par linéarité. De plus, comme la série $\sum_{n\geq 1}|u_n - u_{n+1}|$ est à termes positifs, elle est convergente d'après le critère d'équivalence des séries à termes positifs. Dès lors, la série $\sum(u_n - u_{n+1})$ est absolument convergente. Mais comme toute série absolument convergente est convergente, il s'ensuit que :

$$\boxed{\text{la série } \sum_{n\geq 1}(u_n - u_{n+1}) \text{ converge.}}$$

En particulier, la suite des sommes partielles $(\sum_{k=1}^{n}(u_k - u_{k+1}))$ converge. Mais comme par télescopage $\sum_{k=1}^{n-1}(u_k - u_{k+1}) = u_1 - u_n$, il s'ensuit que la suite $(u_1 - u_n)$ converge, et donc :

$$\boxed{\text{la suite } (u_n)_{n\geq 1} \text{ converge.}}$$

(2) **Partie II : étude de la variable X_r.**

(a) Etude du cas $r = 3$.
Dans cette question, on suppose que $r = 3$, c'est-à-dire que l'urne contient 3 boules numérotées $1, 2, 3$ pouvant être piochées avec probabilité $\frac{1}{3}$. Pour tout $n \in \mathbb{N}^*$, on désigne par C_n l'événement "les n premières pioches fournissent des boules portant toutes le même numéro".

(i) Comparons les événements $[Y_2 > n]$ et C_n. Par définition, l'événement $[Y_2 > n]$ est réalisé si et seulement s'il faut $> n$ pioches pour obtenir 2 boules distinctes, c'est-à-dire si les n

premières pioches fournissent des boules portant toutes le même numéro, ou en d'autres termes si l'événement C_n est réalisé, et donc :

$$\boxed{[Y_2 > n] = C_n.}$$

A présent, calculons la probabilité $P(C_n)$. Pour tout $i \in \{1, ..., n\}$ et tout $j \in \{1, 2, 3\}$, on désigne par $C_{i,j}$ l'événement "On pioche la boule numéro j au i-ème tirage". Alors C_n est réalisé si et seulement s'il existe $j \in \{1, 2, 3\}$ tel que l'on tire n fois de suite la boule numéro j, et donc on voit que :

$$C_n = (C_{1,1} \cap ... \cap C_{n,1}) \cup (C_{1,2} \cap ... \cap C_{n,2}) \cup (C_{1,3} \cap ... \cap C_{n,3}).$$

Par incompatibilité des événements $C_{1,1} \cap ... \cap C_{n,1}, C_{1,2} \cap ... \cap C_{n,2}, C_{1,3} \cap ... \cap C_{n,3}$, on obtient que :

$$P(C_n) = P(C_{1,1} \cap ... \cap C_{n,1}) + P(C_{1,2} \cap ... \cap C_{n,2}) + P(C_{1,3} \cap ... \cap C_{n,3}).$$

D'après la formule des probabilités composées, on trouve que, pour tout $j \in \{1, 2, 3\}$:

$$\begin{aligned} P(C_{1,j} \cap ... \cap C_{n,j}) &= P(C_{1,j})P_{C_{1,j}}(C_{2,j})...P_{C_{1,j}\cap...\cap C_{n-1,j}}(C_{n,j}) \\ &= \frac{1}{3} \times ... \times \frac{1}{3} = \left(\frac{1}{3}\right)^n. \end{aligned}$$

Par conséquent, on en déduit que :

$$\boxed{P(C_n) = \left(\frac{1}{3}\right)^{n-1}.}$$

(ii) Déterminons la valeur de $P([Y_2 > n])$. Comme $[Y_2 > n] = C_n$, on en déduit que :

$$\boxed{P([Y_2 > n]) = \left(\frac{1}{3}\right)^{n-1}.}$$

A présent, donnons la loi de Y_2. Par définition, on voit que l'on ne peut obtenir deux boules distinctes qu'en effectuant au moins 2 tirages, et donc Y_2 ne peut prendre que des valeurs entières ≥ 2. De plus, pour tout entier $n \geq 2$, on voit aisément que $P([Y_2 > n - 1]) = P([Y_2 > n]) + P([Y_2 = n])$, et donc :

$$P([Y_2 = n]) = P([Y_2 > n - 1]) - P([Y_2 > n]) = \left(\frac{1}{3}\right)^{n-2} - \left(\frac{1}{3}\right)^{n-1}.$$

Par conséquent, on obtient après calculs que la loi de Y_2 est donnée par :

$$\boxed{\forall n \geq 2, \quad P([Y_2 = n]) = \frac{2}{3}\left(\frac{1}{3}\right)^{n-2}.}$$

(iii) Justifions que, pour tout $n \geq 1$, on ait : $P([Y_3 - Y_2 = n]) = \sum_{k=2}^{+\infty} P([Y_3 = n+k] \cap [Y_2 = k])$. Comme Y_2 ne peut prendre que des valeurs entières ≥ 2 d'après la question précédente, la famille $([Y_2 = k])_{k\geq 2}$ est un système complet d'événements. Dès lors, la formule des probabilités totales entraine que :

$$P([Y_3 - Y_2 = n]) = \sum_{k=2}^{+\infty} P([Y_3 - Y_2 = n] \cap [Y_2 = k])$$

Mais comme $[Y_3 - Y_2 = n] \cap [Y_2 = k]$ est réalisé si et seulement si $[Y_3 = n + k]$ et $[Y_2 = k]$ sont simultanément réalisés, c'est-à-dire si $[Y_3 = n + k] \cap [Y_2 = k]$ l'est, il s'ensuit que $[Y_3 - Y_2 = n] \cap [Y_2 = k] = [Y_3 = n + k] \cap [Y_2 = k]$, et donc :

$$\boxed{P([Y_3 - Y_2 = n]) = \sum_{k=2}^{+\infty} P([Y_3 = n + k] \cap [Y_2 = k]).}$$

(iv) Montrons que : $\forall n \geq 1, \forall k \geq 2, P([Y_3 = n + k] \cap [Y_2 = k]) = \frac{1}{3^{k-1}}\left(\frac{2}{3}\right)^n$. Pour ce faire, supposons que l'événement $[Y_2 = k]$ soit réalisé. Pour tout $i \in \{k+1, ..., n+k\}$, on désigne par D_i l'événement "On pioche l'une des deux boules déjà piochées au i-ème tirage". Alors $[Y_3 = n + k]$ est réalisé si et seulement si $D_{k+1}, ..., D_{n+k-1}, \overline{D_{n+k}}$ le sont, et donc :

$$[Y_3 = n + k] \cap [Y_2 = k] = [Y_2 = k] \cap D_{k+1} \cap ... \cap D_{n+k-1} \cap \overline{D_{n+k}}.$$

D'après la formule des probabilités composées, on obtient que, pour tout $j \in \{1,2,3\}$:

$$\begin{aligned} P([Y_3 = n+k] \cap [Y_2 = k]) &= P([Y_2 = k])P_{[Y_2=k]}(D_{k+1}) \times ... \\ &\quad ... \times P_{[Y_2=k]\cap D_{k+1}\cap ... \cap D_{n+k-1}}(\overline{D_{n+k}}) \\ &= \frac{2}{3}\left(\frac{1}{3}\right)^{k-2} \times \frac{2}{3} \times ... \times \frac{2}{3} \times \frac{1}{3}. \end{aligned}$$

Par conséquent, on en déduit après calculs que, pour tout $n \geq 1$ et tout $k \geq 2$:

$$\boxed{P([Y_3 = n+k] \cap [Y_2 = k]) = \frac{1}{3^{k-1}}\left(\frac{2}{3}\right)^n.}$$

(v) D'après les questions (2)(*a*)(*iii*) et (2)(*a*)(*iv*), on voit que, pour tout $n \geq 1$:

$$P([Y_3 - Y_2 = n]) = \sum_{k=2}^{+\infty} P([Y_3 = n+k] \cap [Y_2 = k]) = \sum_{k=2}^{+\infty} \frac{1}{3^{k-1}}\left(\frac{2}{3}\right)^n.$$

Par linéarité de la somme, on obtient que, pour tout $n \geq 1$:

$$P([Y_3 - Y_2 = n]) = 3\left(\frac{2}{3}\right)^n \sum_{k=2}^{+\infty} \frac{1}{3^k}.$$

On reconnaît alors à droite la somme des termes d'une suite géométrique de raison $\frac{1}{3}$. D'après le cours, il s'ensuit que, pour tout $n \geq 1$:

$$P([Y_3 - Y_2 = n]) = 3\left(\frac{2}{3}\right)^n \frac{1}{3^2}\frac{1}{1-\frac{1}{3}} = \left(\frac{2}{3}\right)^n \frac{1}{3}\frac{1}{\frac{2}{3}} = \frac{1}{2}\left(\frac{2}{3}\right)^n.$$

Remarquons au passage que $Y_3 - Y_2$ ne peut prendre que des valeurs entières ≥ 1. En effet, pour tirer une nouvelle boule non déjà tirée, il faut effectuer au moins un tirage supplémentaire par rapport aux précédents. Par conséquent, la loi de la variable aléatoire $Y_3 - Y_2$ est donnée par :

$$\boxed{\forall n \geq 1,\ \ P([Y_3 - Y_2 = n]) = \frac{1}{2}\left(\frac{2}{3}\right)^n.}$$

(b) <u>Loi de $Y_{i+1} - Y_i$ pour tout $i \in \{1, ..., r-1\}$.</u>

Dans toute la suite du problème, r désigne un entier ≥ 2.

(i) Justifions que $Y_i(\Omega) = \{i, i+1, ...\} = \mathbb{N} \setminus \{0, 1, ..., i-1\}$. Tout d'abord, pour obtenir i boules distinctes, il faut avoir effectué au moins i tirages au préalable, et donc $Y_i \geq i$. De plus, pour tout entier $k \geq i$, l'événement $[Y_i = k]$ est réalisé si l'on tire $(i-1)$ boules distinctes au cours des $(i-1)$ premiers tirages, si les boules tirées ensuite jusqu'au $(k-1)$-ème tirage font partie des boules déjà tirées et enfin si l'on tire une nouvelle boule jamais tirée au cours du k-ème tirage. Dès lors, la variable aléatoire Y_i peut prendre toute valeur entière $k \geq i$, et donc :

$$\boxed{Y_i(\Omega) = \{i, i+1, ...\} = \mathbb{N} \setminus \{0, 1, ..., i-1\}.}$$

Justifions à présent que $(Y_{i+1} - Y_i)(\Omega) = \mathbb{N}^*$. Tout d'abord, on peut remarquer que $Y_{i+1} - Y_i$ ne peut prendre que des valeurs entières ≥ 1. En effet, pour tirer une nouvelle boule non déjà tirée, il faut effectuer au moins un tirage supplémentaire par rapport aux précédents. De plus, pour tout $k \geq 1$, l'événement $[Y_{i+1} - Y_i = k]$ est réalisé si, après le i-ème tirage, on tire $(k-1)$ fois une des boules déjà tirées, puis si l'on tire une nouvelle boule jamais tirée au cours du tirage suivant. Dès lors, la variable aléatoire $Y_{i+1} - Y_i$ peut prendre toute valeur entière $k \geq 1$, et donc :

$$\boxed{(Y_{i+1} - Y_i)(\Omega) = \mathbb{N}^*.}$$

(ii) Montrons que : $\forall n \geq 1$, $\forall k \geq i$, $P_{[Y_i=k]}([Y_{i+1} - Y_i = n]) = \left(\frac{i}{r}\right)^{n-1}\left(1 - \frac{i}{r}\right)$. Pour ce faire, supposons que l'événement $[Y_i = k]$ soit réalisé. Pour tout $l \in \{k+1, ..., n+k\}$, on désigne par B_l l'événement "On pioche l'une des boules déjà piochées au l-ème tirage". Alors $[Y_{i+1} - Y_i = n]$ est réalisé si et seulement si $[Y_{i+1} = n+k]$ est réalisé, c'est-à-dire si les événements $B_{k+1}, ..., B_{n+k-1}, \overline{B_{n+k}}$ sont simultanément réalisés, et donc :

$$[Y_{i+1} = n+k] \cap [Y_i = k] = [Y_i = k] \cap B_{k+1} \cap ... \cap B_{n+k-1} \cap \overline{B_{n+k}}.$$

D'après la formule des probabilités composées, on obtient que, pour tout $j \in \{1,2,3\}$:

$$\begin{aligned} P([Y_{i+1} = n+k] \cap [Y_i = k]) &= P([Y_i = k]) P_{[Y_i=k]}(B_{k+1}) \times ... \\ &\quad ... \times P_{[Y_i=k]\cap B_{k+1}\cap ... \cap B_{n+k-1}}(\overline{B_{n+k}}) \\ &= P([Y_i = k]) \frac{i}{r} \times ... \times \frac{i}{r} \times \left(1 - \frac{i}{r}\right) \\ &= P([Y_i = k]) \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right). \end{aligned}$$

Par conséquent, on en déduit que, pour tout $n \geq 1$ et tout $k \geq 2$:

$$\boxed{P_{[Y_i=k]}([Y_{i+1} - Y_i = n]) = \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right).}$$

(iii) D'après la question précédente, on sait que Y_i ne peut prendre que des valeurs entières $\geq i$, et donc la famille $([Y_i = k])_{k \geq i}$ est un système complet d'événements. D'après la formule des probabilités totales, on obtient que :

$$\begin{aligned} P([Y_{i+1} - Y_i = n]) &= \sum_{k=i}^{+\infty} P([Y_{i+1} - Y_i = n] \cap [Y_i = k]) \\ &= \sum_{k=i}^{+\infty} P([Y_{i+1} = n+k] \cap [Y_i = k]). \end{aligned}$$

D'après le résultat de la question précédente, il s'ensuit que :

$$P([Y_{i+1} - Y_i = n]) = \sum_{k=i}^{+\infty} P([Y_i = k]) \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right).$$

Par linéarité de la somme, on obtient que :

$$P([Y_{i+1} - Y_i = n]) = \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right) \sum_{k=i}^{+\infty} P([Y_i = k]).$$

Mais comme la famille $([Y_i = k])_{k \geq i}$ est un système complet d'événements, la somme de la série $\sum_{k=i}^{+\infty} P([Y_i = k])$ est égale à 1. Dès lors, il s'ensuit que, pour tout $n \geq 1$:

$$P([Y_{i+1} - Y_i = n] = \left(\frac{i}{r}\right)^{n-1} \left(1 - \frac{i}{r}\right).$$

Comme $Y_{i+1} - Y_i$ ne peut prendre que des valeurs entières $n \geq 1$, on en déduit que :

$$\boxed{Y_{i+1} - Y_i \hookrightarrow \mathcal{G}\left(1 - \frac{i}{r}\right).}$$

D'après le cours, on sait qu'alors :

$$E(Y_{i+1} - Y_i) = \frac{1}{1 - \frac{i}{r}} \quad \text{et} \quad V(Y_{i+1} - Y_i) = \frac{\frac{i}{r}}{(1 - \frac{i}{r})^2}.$$

Par conséquent, on en déduit après simplification que :

$$\boxed{E(Y_{i+1} - Y_i) = \frac{r}{r-i} \text{ et } V(Y_{i+1} - Y_i) = \frac{ri}{(r-i)^2}.}$$

(c) <u>Espérance et variance de X_r.</u>

(i) Justifions l'égalité $X_r = 1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i})$. Comme $Y_1 = 1$, on obtient par télescopage que :

$$1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i}) = Y_1 + (Y_r - Y_{r-1}) + ... + (Y_2 - Y_1) = Y_r.$$

Mais comme $X_r = Y_r$, on en déduit que :

$$\boxed{X_r = 1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i}).}$$

Admettons que les variables $Y_2 - Y_1, Y_3 - Y_2, ..., Y_r - Y_{r-1}$ soient indépendantes. Comme $X_r = 1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i})$ et que les variables aléatoires $(Y_{r-i+1} - Y_{r-i})$ admettent toutes une espérance, X_r admet aussi un espérance. De plus, par linéarité de l'espérance et d'après la question $(2)(b)(iii)$, on obtient que :

$$\begin{aligned}
E(X_r) &= E\left(1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i})\right) \\
&= 1 + \sum_{i=1}^{r-1} E(Y_{r-i+1} - Y_{r-i}) \\
&= 1 + \sum_{i=1}^{r-1} \frac{r}{r-(r-i)} \\
&= \frac{r}{r} + r\sum_{i=1}^{r-1}\frac{1}{i}.
\end{aligned}$$

Par conséquent, on en déduit que :

$$\boxed{E(X_r) = r\sum_{i=1}^{r}\frac{1}{i}.}$$

De même, comme les variables $Y_2 - Y_1, Y_3 - Y_2, ..., Y_r - Y_{r-1}$ sont indépendantes et qu'elles admettent toutes une variance, la variable aléatoire X_r admet aussi une variance. De plus, d'après la question $(2)(b)(iii)$ et les propriétés de la variance, on obtient que :

$$\begin{aligned}
V(X_r) &= V\left(1 + \sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i})\right) \\
&= V\left(\sum_{i=1}^{r-1}(Y_{r-i+1} - Y_{r-i})\right) \\
&= \sum_{i=1}^{r-1} V\left(Y_{r-i+1} - Y_{r-i}\right) \\
&= \sum_{i=1}^{r-1} \frac{r(r-i)}{(r-(r-i))^2} \\
&= \frac{r(r-r)}{r^2} + \sum_{i=1}^{r-1}\frac{r(r-i)}{i^2} \\
&= \sum_{i=1}^{r}\frac{r(r-i)}{i^2}.
\end{aligned}$$

Par linéarité de la somme, on trouve que :

$$V(X_r) = \sum_{i=1}^{r}\frac{r^2 - ri}{i^2} = r^2\left(\sum_{i=1}^{r}\frac{1}{i^2}\right) - r\left(\sum_{i=1}^{r}\frac{i}{i^2}\right).$$

Par conséquent, on en déduit que :

$$\boxed{V(X_r) = r^2\left(\sum_{i=1}^{r}\frac{1}{i^2}\right) - r\left(\sum_{i=1}^{r}\frac{1}{i}\right).}$$

(ii) Ecrivons une fonction en Python qui, à partir d'un entier $r \geq 1$, calcule et affiche la matrice $M_r \in \mathcal{M}_{2,r}(\mathbb{R})$ dont les lignes sont $L_1 = (E(X_1), ..., E(X_r))$ et $L_2 = (V(X_1), ..., V(X_r))$, puis qui trace dans le plan l'ensemble des points $M_k = (E(X_k), V(X_k))$ pour $k \in [\![1, r]\!]$. Pour ce faire, on procède comme suit :

```
import numpy as np
import matplotlib.pyplot as plt

def prob2(r):
    m=np.zeros([2,r])
    for k in range(r):
        m[0,k]=(k+1)*np.sum(1/np.arange(1,k+1))
        m[1,k]=((k+1)**2)*np.sum(1/(np.arange(1,k+1)**2))-m[0,k]
    print(m)
    x=np.zeros(r)
    y=np.zeros(r)
    for k in range(r):
        x[k]=m[0,k]
        y[k]=m[1,k]
    plt.plot(x,y,'.')
    plt.show()
```

(iii) Montrons qu'il existe deux réels α, β tels que :

$$E(X_r) \underset{r\to+\infty}{=} r\ln(r) + \alpha r + o(r) \quad \text{et} \quad V(X_r) \underset{r\to+\infty}{\sim} \beta r^2.$$

D'après la question (1)(c), on sait que la suite (u_n) converge. Si α est la limite de (u_n), alors on voit que $u_r \underset{r\to+\infty}{=} \alpha + o(1)$ par définition. Mais comme $u_r = \left(\sum_{k=1}^{r} \frac{1}{k}\right) - \ln(r)$ pour tout $r \in \mathbb{N}^*$, on a :

$$u_r = \left(\sum_{k=1}^{r} \frac{1}{k}\right) - \ln(r) \underset{r\to+\infty}{=} \alpha + o(1),$$

d'où il s'ensuit que :

$$\sum_{k=1}^{r} \frac{1}{k} \underset{r\to+\infty}{=} \ln(r) + \alpha + o(1).$$

Par produit avec r, on trouve que :

$$r\sum_{k=1}^{r} \frac{1}{k} \underset{r\to+\infty}{=} r\ln(r) + \alpha r + o(r).$$

Par conséquent, on en déduit avec la question précédente que :

$$\boxed{E(X_r) \underset{r\to+\infty}{=} r\ln(r) + \alpha r + o(r).}$$

En outre, d'après la question (1)(d), on sait que la série $\sum \frac{1}{k^2}$ converge. Si β est la somme de cette série, alors on voit que $\sum_{i=1}^{r} \frac{1}{i^2} \underset{r\to+\infty}{=} \beta + o(1)$ par définition. De plus, comme cette série est à termes > 0, on voit que $\beta > 0$ et donc :

$$V(X_r) \underset{r\to+\infty}{=} r^2\left[\beta + o(1)\right] - r\left[\ln(r) + \alpha + o(1)\right].$$

Par des calculs simples, on trouve alors que :

$$V(X_r) \underset{r\to+\infty}{=} \beta r^2 + r\left[-\ln(r) - \alpha + o(r) - o(1)\right].$$

Comme $\ln(r) \underset{r\to+\infty}{=} o(r)$ par croissance comparée, que $\alpha \underset{r\to+\infty}{=} o(r)$ et que $o(1) \underset{r\to+\infty}{=} o(r)$, on voit que :

$$V(X_r) \underset{r\to+\infty}{=} \beta r^2 + r\left[o(r)\right] = \beta r^2 + o(r^2).$$

Mais comme $\beta \neq 0$, on en déduit que :

$$\boxed{V(X_r) \underset{r\to+\infty}{\sim} \beta r^2.}$$

(3) **Partie III : loi de X_r et déviation asymptotique par rapport à sa moyenne.**
Pour tout entier $m > 0$ et pour tout entier $k \in \{1, ..., r\}$, on considère les événements $A_{k,m}$: "le numéro k n'a pas été pioché durant les m premières pioches" et $B_{k,m}$: "k numéros fixés au départ n'ont pas été piochés durant les m premières pioches". Enfin, on admet la formule du crible pour n événements $A_1, ..., A_n$, à savoir :

$$P(A_1 \cup ... \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1 \leq i_1 < ... < i_k \leq n} P(A_{i_1} \cap ... \cap A_{i_k}) \right).$$

(a) Détermination de la loi de X_r.

(i) Calculons la probabilité de l'événement $A_{k,m}$. Fixons pour cela les entiers k, m. Pour tout $i \in [\![1, m]\!]$, on désigne par E_i l'événement "la boule numéro k n'a pas été pioché lors du i-ème lancer". Alors l'événement $A_{k,m}$ est réalisé si et seulement si les événements $E_1, ..., E_m$ sont simultanément réalisés, et donc :

$$A_{k,m} = E_1 \cap ... \cap E_m.$$

Par indépendance des E_i (vu que les tirages se font avec remise), on trouve que :

$$P(A_{k,m}) = P(E_1)...P(E_m) = \frac{r-1}{r} \times ... \times \frac{r-1}{r}.$$

Par conséquent, on en déduit que :

$$\boxed{P(A_{k,m}) = \left(\frac{r-1}{r}\right)^m.}$$

A présent, calculons la probabilité de l'événement $B_{k,m}$. Fixons pour cela les entiers k, m. Pour tout $i \in [\![1, m]\!]$, on désigne par F_i l'événement "les k numéros en question n'ont pas été piochés lors du i-ème lancer". Alors l'événement $B_{k,m}$ est réalisé si et seulement si les événements $F_1, ..., F_m$ sont simultanément réalisés, et donc :

$$B_{k,m} = F_1 \cap ... \cap F_m.$$

Par indépendance des F_i (vu que les tirages se font avec remise), on trouve que :

$$P(B_{k,m}) = P(F_1)...P(F_m) = \frac{r-k}{r} \times ... \times \frac{r-k}{r}.$$

Par conséquent, on en déduit que :

$$\boxed{P(B_{k,m}) = \left(\frac{r-k}{r}\right)^m.}$$

(ii) Justifions que $P([X_r > m]) = P(A_{1,m} \cup A_{2,m} \cup ... \cup A_{r,m})$. Par définition, l'événement $[X_r > m]$ est réalisé si et seulement si, au cours des m premières pioches, on n'a pas tiré les r boules, c'est-à-dire s'il existe un entier $k \in \{1, ..., r\}$ tel que le numéro k n'a pas été pioché durant les m premières pioches, ou en d'autres termes si l'un des $A_{k,m}$ est réalisé, et donc :

$$[X_r > m] = A_{1,m} \cup A_{2,m} \cup ... \cup A_{r,m}.$$

En passant aux probabilités, on en déduit que :

$$\boxed{P([X_r > m]) = P(A_{1,m} \cup A_{2,m} \cup ... \cup A_{r,m}).}$$

(iii) Montrons à l'aide de la formule du crible que : $P([X_r > m]) = \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^m$.
D'après la formule du crible appliquée à la question précédente, on trouve que :

$$P([X_r > m]) = \sum_{k=1}^{r} (-1)^{k-1} \left(\sum_{1 \leq i_1 < ... < i_k \leq r} P(A_{i_1,m} \cap ... \cap A_{i_k,m}) \right).$$

Par définition, l'événement $A_{i_1,m} \cap ... \cap A_{i_k,m}$ est réalisé si et seulement si les k numéros $i_1, ..., i_k$ n'ont pas été piochés, c'est-à-dire si $B_{k,m}$ est réalisé, et donc :

$$P(A_{i_1,m} \cap ... \cap A_{i_k,m}) = P(B_{k,m}) = \left(\frac{r-k}{r}\right)^m.$$

Dès lors, on trouve que :

$$P([X_r > m]) = \sum_{k=1}^{r} (-1)^{k-1} \left(\sum_{1 \leq i_1 < ... < i_k \leq r} \left(\frac{r-k}{r} \right)^m \right).$$

Déterminons maintenant le nombre de termes de la deuxième somme de droite. Par définition, cette somme comporte autant de termes qu'il y a de façons de choisir k termes en ordre croissant parmi $\{1, ..., r\}$. Pour obtenir k termes en ordre croissant parmi $\{1, ..., r\}$, on commence par choisir k termes parmi $\{1, ..., r\}$, sans ordre et sans répétition, ce qui fait $\binom{r}{k}$ possibilités. Ensuite, on les ordonne de la seule façon possible. Il y a donc $\binom{r}{k} \times 1$ façons de choisir k termes en ordre croissant dans l'ensemble $\{1, ..., r\}$, et donc la deuxième somme de droite comporte $\binom{r}{k}$ termes. Par conséquent :

$$\boxed{P([X_r > m]) = \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^m.}$$

(iv) Déterminons la loi de X_r. Par définition, on voit que l'on ne peut obtenir r boules distinctes qu'en effectuant au moins r tirages, et donc X_r ne peut prendre que des valeurs entières $\geq r$. De plus, pour tout $m \geq r$, on voit que $P([X_r > m-1]) = P([X_r > m]) + P([X_r = m])$, et donc on a par linéarité de la somme :

$$\begin{aligned}
P([X_r = m]) &= P([X_r > m-1]) - P([X_r > m]) \\
&= \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^m - \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m-1} \\
&= \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left[\left(1 - \frac{k}{r}\right)^m - \left(1 - \frac{k}{r}\right)^{m-1} \right] \\
&= \sum_{k=1}^{r} (-1)^{k-1} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m-1} \left[\left(1 - \frac{k}{r}\right) - 1 \right]
\end{aligned}$$

Par conséquent, on obtient après simplification que la loi de X_r est donnée par :

$$\boxed{\forall m \geq r, \;\; P([X_r = m]) = \sum_{k=1}^{r} (-1)^k \frac{k}{r} \binom{r}{k} \left(1 - \frac{k}{r}\right)^{m-1}.}$$

(b) Comportement asymptotique de X_r au delà de sa moyenne.
Pour tout réel $\varepsilon > 0$, on désigne par M_r la partie entière de $(1+\varepsilon) r \ln(r)$, c'est-à-dire l'unique entier relatif M_r tel que $M_r \leq (1+\varepsilon) r \ln(r) < M_r + 1$.

(i) Montrons par récurrence la propriété $\mathcal{P}$ définie pour tout $m \geq 1$ par :

$\mathcal{P}(m)$: "Pour tous événements $D_1, ..., D_m$, $P(D_1 \cup ... \cup D_m) \leq P(D_1) + ... + P(D_m)$".

Tout d'abord, on voit que $\mathcal{P}(1)$ est vraie, car $P(D_1) \leq P(D_1)$. A présent, supposons que $\mathcal{P}(m)$ soit vraie, et montrons que $\mathcal{P}(m+1)$ l'est aussi. Soient $D_1, ..., D_{m+1}$ des événements quelconques, et posons $E_m = D_m \cup D_{m+1}$. Alors on voit que :

$$D_1 \cup ... \cup D_{m+1} = D_1 \cup ... \cup D_{m-1} \cup E_m.$$

Par hypothèse de récurrence, on obtient que :

$$P(D_1 \cup ... \cup D_{m+1}) \leq P(D_1) + ... + P(D_{m-1}) + P(E_m).$$

ce que l'on peut réécrire sous la forme :

$$P(D_1 \cup ... \cup D_{m+1}) \leq P(D_1) + ... + P(D_{m-1}) + P(D_m \cup D_{m+1}) \quad (*).$$

Or, d'après la formule du crible, on sait que :

$$P(D_m \cup D_{m+1}) = P(D_m) + P(D_{m+1}) - P(D_m \cap D_{m+1}).$$

Comme toute probabilité est un réel ≥ 0, on voit que :

$$P(D_m \cup D_{m+1}) \leq P(D_m) + P(D_{m+1}).$$

Dès lors, cela entraine avec l'inégalité $(*)$ que :

$$P(D_1 \cup ... \cup D_{m+1}) \leq P(D_1) + ... + P(D_{m-1}) + P(D_m) + P(D_{m+1}),$$

d'où il s'ensuit que $\mathcal{P}(m+1)$ est vraie. D'après le principe de récurrence, la propriété $\mathcal{P}$ est vraie à tout ordre $m \geq 1$. Dès lors, pour tous événements $D_1, ..., D_m$, on a :

$$\boxed{P(D_1 \cup ... \cup D_m) \leq P(D_1) + ... + P(D_m).}$$

(ii) Démontrons que, pour tout réel x, on a : $e^x \geq 1 + x$. Pour ce faire, on peut remarquer que la fonction exponentielle est convexe. En effet, cette fonction est de classe $\mathcal{C}^2$, et de plus $(e^x)'' = e^x > 0$ pour tout $x \in \mathbb{R}$. En particulier, la courbe de l'exponentielle est située au dessus de sa tangente en 0, qui a pour équation $y = 1 + x$, et donc :

$$\boxed{\forall x \in \mathbb{R}, \quad e^x \geq 1 + x.}$$

A présent, montrons que : $\forall m \in \mathbb{N}^*$, $\forall k \in \{1, ..., r\}$, $P(A_{k,m}) \leq e^{-\frac{m}{r}}$. D'après la question $(3)(a)(i)$, on sait que :

$$P(A_{k,m}) = \left(\frac{r-1}{r}\right)^m = \left(1 - \frac{1}{r}\right)^m.$$

D'après l'inégalité ci-dessus, on trouve que $0 \leq 1 - \frac{1}{r} \leq e^{-\frac{1}{r}}$, et donc :

$$P(A_{k,m}) \leq \left(e^{-\frac{1}{r}}\right)^m.$$

Par conséquent, on en déduit que :

$$\boxed{\forall m \in \mathbb{N}^*, \quad \forall k \in \{1, ..., r\}, \quad P(A_{k,m}) \leq e^{-\frac{m}{r}}.}$$

(iii) Comparons les événements $[X > M_r]$ et $[X > (1+\varepsilon)r\ln(r)]$. Pour ce faire, supposons que l'événement $[X > M_r]$ soit réalisé. Comme $X > M_r$, on voit que $X \geq M_r + 1$ (car X ne prend que des valeurs entières). Mais comme $M_r + 1 > (1+\varepsilon)r\ln(r)$ par définition, il s'ensuit que $X > (1+\varepsilon)r\ln(r)$, que l'événement $[X > (1+\varepsilon)r\ln(r)]$ est réalisé, et donc :

$$[X > M_r] \subset [X > (1+\varepsilon)r\ln(r)].$$

Réciproquement, supposons que l'événement $[X > (1+\varepsilon)r\ln(r)]$ soit réalisé. Comme $X > (1+\varepsilon)r\ln(r)$ et que $(1+\varepsilon)r\ln(r) \geq M_r$ par définition de la partie entière, on voit que $X > M_r$, et donc :

$$[X > (1+\varepsilon)r\ln(r)] \subset [X > M_r].$$

Par double inclusion, on en déduit que :

$$\boxed{[X > M_r] = [X > (1+\varepsilon)r\ln(r)].}$$

A présent, montrons que $P([X_r > (1+\varepsilon)r\ln(r)]) \leq \frac{e}{r^\varepsilon}$. Partant de l'égalité ci-dessus, on obtient en passant aux probabilités et à l'aide des questions $(3)(b)(i)$ et $(3)(b)(ii)$:

$$\begin{aligned} P([X > (1+\varepsilon)r\ln(r)]) &= P([X > M_r]) \\ &= P(A_{1,M_r} \cup ... \cup A_{r,M_r}) \\ &\leq P(A_{1,M_r}) + ... + P(A_{r,M_r}) \\ &\leq e^{-\frac{M_r}{r}} + ... + e^{-\frac{M_r}{r}} \\ &\leq re^{-\frac{M_r}{r}} \quad (*). \end{aligned}$$

En outre, comme M_r est la partie entière de $(1+\varepsilon)r\ln(r)$, on obtient que :

$$\begin{aligned} M_r > (1+\varepsilon)r\ln(r) - 1 &\implies -\tfrac{M_r}{r} < -(1+\varepsilon)\ln(r) + \tfrac{1}{r} \\ &\implies e^{-\frac{M_r}{r}} < e^{-(1+\varepsilon)\ln(r) + \frac{1}{r}} \\ &\implies e^{-\frac{M_r}{r}} < \tfrac{1}{r^{1+\varepsilon}} e^{\frac{1}{r}}. \end{aligned}$$

Mais comme r est un entier > 0, on voit que $r \geq 1$, et donc $e^{\frac{1}{r}} \leq e^1 = e$. Dès lors, on a :

$$e^{-\frac{M_r}{r}} < \frac{e}{r^{1+\varepsilon}} \quad (**).$$

A l'aide des inégalités $(*)$ et $(**)$, il s'ensuit que :

$$P([X > (1+\varepsilon) r \ln(r)]) \leq r \times \frac{e}{r^{1+\varepsilon}},$$

d'où l'on déduit que :

$$\boxed{P([X > (1+\varepsilon) r \ln(r)]) \leq \frac{e}{r^{\varepsilon}}.}$$

(c) Distribution de X_r autour de sa moyenne.
Pour tout réel t fixé, on désigne par m_r la partie entière de $r\ln(r) + rt$, c'est-à-dire l'unique entier relatif m_r tel que $m_r \leq r\ln(r) + rt < m_r + 1$. De plus, on introduit la suite $(Z_r)_{r \geq 2}$ de variables aléatoires définies pour tout $r \geq 2$ par :

$$Z_r = \frac{X_r - r\ln(r)}{r}.$$

(i) Justifions l'existence d'un rang $r_0(t)$ tel que : $\forall r \geq r_0(t)$, $m_r \geq 1$. Par définition, on a :

$$r\ln(r) + rt - 1 = r\ln(r)\left[1 + \frac{t}{r} - \frac{1}{r\ln(r)}\right] < m_r.$$

Comme $\frac{t}{r}$ tend vers 0 quand r tend vers $+\infty$, et que $r\ln(r)$ tend vers $+\infty$ quand r tend vers $+\infty$, il s'ensuit que $r\ln(r) + rt - 1$ tend vers $+\infty$ quand r tend vers $+\infty$. D'après le théorème d'encadrement, on voit que m_r tend vers $+\infty$ quand r tend vers $+\infty$. Mais alors cela signifie que :

$$\forall K > 0, \quad \exists R_{K,t} > 0, \quad \forall r > R_{K,t}, \quad m_r \geq K.$$

En posant $K = 1$ et $r_0(t) = E(R_{1,t}) + 1$, on en déduit que :

$$\boxed{\exists r_0(t) \in \mathbb{N}^*, \quad \forall r \geq r_0(t), \quad m_r \geq 1.}$$

(ii) Etablissons l'égalité : $\forall r \geq r_0(t)$, $P([Z_r > t]) = P([X_r > m_r])$. Pour ce faire, supposons que l'événement $[Z_r > t]$ soit réalisé. Alors, on voit que :

$$\begin{aligned} Z_r > t \quad &\Longrightarrow \quad \tfrac{X_r - r\ln(r)}{r} > t \\ &\Longrightarrow \quad X_r - r\ln(r) > rt \\ &\Longrightarrow \quad X_r > r\ln(r) + rt \geq E(r\ln(r) + rt) \\ &\Longrightarrow \quad X_r > E(r\ln(r) + rt) \\ &\Longrightarrow \quad X_r > m_r, \end{aligned}$$

et donc l'événement $[X_r > m_r]$ est réalisé. En particulier :

$$[Z_r > t] \subset [X_r > m_r].$$

A présent, supposons que l'événement $[X_r > m_r]$ soit réalisé. Comme $X_r > m_r$ et que X_r ne prend que des valeurs entières, on voit que $X_r \geq m_r + 1$, et donc :

$$\begin{aligned} X_r \geq m_r + 1 \quad &\Longrightarrow \quad X_r > E(r\ln(r) + rt) + 1 \\ &\Longrightarrow \quad X_r > r\ln(r) + rt \\ &\Longrightarrow \quad X_r - r\ln(r) > rt \\ &\Longrightarrow \quad \tfrac{X_r - r\ln(r)}{r} > t \\ &\Longrightarrow \quad Z_r > t, \end{aligned}$$

et donc l'événement $[Z_r > t]$ est réalisé. En particulier :

$$[X_r > m_r] \subset [Z_r > t].$$

Dès lors, il s'ensuit par double inclusion que $[Z_r > t] = [X_r > m_r]$, et donc :

$$\boxed{P([Z_r > t]) = P([X_r > m_r]).}$$

(iii) Soit $k \in \mathbb{N}$. A l'aide d'un développement limité, montrons que :

$$m_r \ln\left(1-\frac{k}{r}\right) \underset{r\to+\infty}{=} -k\ln(r) - kt + o(1).$$

Comme m_r est la partie entière de $r\ln(r)+rt$, on voit que $r\ln(r)+rt-m_r$ appartient à $[0,1[$ pour tout $r \in \mathbb{N}^*$, et donc $r\ln(r)+rt-m_r \underset{r\to+\infty}{=} o(r)$, ce qui entraine que :

$$m_r \underset{r\to+\infty}{=} r\ln(r)+rt+o(r).$$

Comme $-\frac{k}{r}$ tend vers 0 quand r tend vers $+\infty$ et que $\ln(1+x) \underset{r\to+\infty}{=} x - \frac{x^2}{2} + o(x^2)$, on obtient par substitution que :

$$\ln\left(1-\frac{k}{r}\right) \underset{r\to+\infty}{=} -\frac{k}{r} - \frac{k^2}{2r^2} + o\left(\frac{1}{r^2}\right).$$

Par produit, on trouve alors que :

$$\begin{aligned} m_r\ln\left(1-\frac{k}{r}\right) & \underset{r\to+\infty}{=} (r\ln(r)+rt+o(r))\left(-\frac{k}{r}-\frac{k^2}{2r^2}+o\left(\frac{1}{r^2}\right)\right) \\ & \underset{r\to+\infty}{=} -k\ln(r) - \frac{k^2\ln(r)}{2r} + o\left(\frac{\ln(r)}{r}\right) - kt - \frac{k^2 t}{2r} \\ & \qquad +o\left(\frac{1}{r}\right) - o(k) - o\left(\frac{k^2}{2r}\right) + o\left(\frac{1}{r}\right). \end{aligned}$$

Comme $\frac{\ln(r)}{r}$ tend vers 0 quand r tend vers $+\infty$ par croissances comparées, et que $\frac{1}{r}$ tend vers 0 quand r tend vers $+\infty$, on en déduit que :

$$\boxed{m_r \ln\left(1-\frac{k}{r}\right) \underset{r\to+\infty}{=} -k\ln(r) - kt + o(1).}$$

(iv) Montrons que $\binom{r}{k} \underset{r\to+\infty}{\sim} \frac{r^k}{k!}$ pour tout $k \in \mathbb{N}$. Par définition, on sait que :

$$\binom{r}{k} = \frac{r(r-1)(r-2)...(r-k+1)}{k!}.$$

Pour k fixé, le numérateur de cette fraction est un polynôme en r, unitaire et de degré r. Comme tout polynôme est équivalent en $+\infty$ à son terme de plus haut degré, on a :

$$\boxed{\binom{r}{k} \underset{r\to+\infty}{\sim} \frac{r^k}{k!}.}$$

A présent, montrons que :

$$\lim_{r\to+\infty} \binom{r}{k}\left(1-\frac{k}{r}\right)^{m_r} = \frac{e^{-kt}}{k!}.$$

Par des calculs simples et d'après la question précédente, on trouve que :

$$\left(1-\frac{k}{r}\right)^{m_r} = e^{m_r\ln\left(1-\frac{k}{r}\right)} \underset{r\to+\infty}{=} e^{-k\ln(r)-kt+o(1)} \underset{r\to+\infty}{=} \frac{e^{-kt}e^{o(1)}}{r^k}.$$

Comme un $o(1)$ est une expression qui tend vers 0 quand r tend vers $+\infty$, on voit que $e^{o(1)}$ tend vers 1 quand r tend vers $+\infty$, et donc :

$$\left(1-\frac{k}{r}\right)^{m_r} \underset{r\to+\infty}{\sim} \frac{e^{-kt}}{r^k}.$$

D'après le calcul précédent et les règles de calculs des équivalents, on a :

$$\binom{r}{k}\left(1-\frac{k}{r}\right)^{m_r} \underset{r\to+\infty}{\sim} \frac{r^k}{k!} \times \frac{e^{-kt}}{r^k} \underset{r\to+\infty}{\sim} \frac{e^{-kt}}{k!}.$$

Par conséquent, on en déduit que :

$$\boxed{\lim_{r\to+\infty} \binom{r}{k}\left(1-\frac{k}{r}\right)^{m_r} = \frac{e^{-kt}}{k!}.}$$