Devoir sur Table 5

Durée : 4h

- Tous les documents sur papier sont interdits.
- Les calculatrices ne sont pas autorisées.
- Le matériel de géométrie (règle, compas, équerre) est autorisé.
- La notation des copies tiendra compte dans une large mesure de la qualité de la rédaction. Ceci implique que vous devez faire des raisonnements clairs, concis et complets, utiliser un langage mathématiques adapté et précis, être lisible et éviter les fautes d'orthographe et de grammaire.
- Si, au cours du devoir, vous repérez ce qui vous semble être une erreur d'énoncé, vous le signalez sur votre copie et poursuivez sa composition en expliquant les raisons des initiatives que vous avez été amené à prendre.
- Mettez en évidence vos résultats en les encadrant.
- Conformément au règlement de la Banque PT
  - Composer lisiblement sur les copies avec un stylo à bille à encre foncée : bleue ou noire.
  - L'usage de liquide de correction et dérouleur de ruban correcteur est interdit.

Le soin apporté à la copie fera l'objet d'une évaluation suivant les critères suivants :

- Mise en évidence des résultats
- Soin et lisibilité de la copie. En particulier les traits, y compris pour les ratures, devront être tracés à l'aide d'une règle
- Respect des consignes concernant le liquide de correction et le dérouleur de ruban correcteur
- Respect de la grammaire et de l'orthographe

Problème

(*adapté de CCINP PC 2017*)

Soit $p \in]0,1[$. On pose $q = 1 - p$.

On considère un automate qui génère successivement les lettres P ou T jusqu'à obtenir une certaine séquence prédéfinie.

On suppose que pour tout $n \in \mathbb{N}^*$, l'automate génère la n-ième lettre à l'instant n de façon indépendante de toutes les générations précédentes.

On suppose également qu'à chaque génération, les lettres P et T ont des probabilités p et q (respectivement) d'être générées.

Suivant les parties considérées, on définit différents niveaux que l'automate peut atteindre.

On considère dans tous les cas que l'automate est initialement au niveau 0.

On se propose alors d'étudier essentiellement l'existence de l'espérance et de la variance de la variable aléatoire correspondant au temps d'attente de la séquence prédéfinie à travers sa série génératrice.

Pour cette étude probabiliste, on mobilise diverses propriétés analytiques (surtout sur les séries entières) et quelques propriétés d'algèbre linéaire.

Dans les parties I, II et V, on examine le temps d'attente pour les séquences T puis TT, puis TPT et TTPPT. La partie II est indépendante de la partie I et traite de questions préliminaires sur les séries entières qui seront investies dans les parties III et V. La partie IV est indépendante des parties précédentes et traite les questions préliminaires d'algèbre linéaire qui servent exclusivement dans la partie V. La partie III ne dépend de la partie I que par la question 4. et de la partie II que par la question 10.. La partie V utilise seulement la question 11. de la partie II et la partie IV.

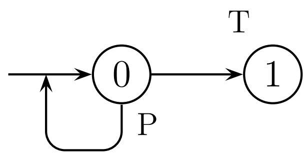
Pour $n \in \mathbb{N}^*$, on note

— P_n l'évènement « l'automate génère la lettre P à l'instant n »
— T_n l'évènement « l'automate génère la lettre T à l'instant n » .

Partie I — Étude d'un cas simple

Dans cette partie, on dit que l'automate passe du niveau 0 au niveau 1 dès qu'il génère la lettre T. Si, en revanche, il génère la lettre P, alors il reste au niveau 0. L'expérience s'arrête dès que l'automate a atteint le niveau 1. On résume l'expérience par la figure 1 suivante :

Figure 1



On note Y l'instant où, pour la première fois, l'automate atteint le niveau 1. On admet que Y est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, P)$ telle que $Y(\Omega) \subset \mathbb{N}^*$. On note G_Y la série génératrice de Y et R_Y son rayon de convergence.

On sait alors que $R_Y \geqslant 1$ et que :

$$\forall t \in]-R_Y, R_Y[, \qquad G_Y(t) = \mathbb{E}(t^Y) = \sum_{n=1}^{+\infty} \mathbb{P}(Y = n)t^n.$$

1. Reconnaître la loi de Y et préciser en particulier $\mathbb{P}(Y = n)$ pour $n \in \mathbb{N}^*$.
2. Montrer que $R_Y = \dfrac{1}{p} > 1$ et que : $\forall t \in \left]-\dfrac{1}{p}, \dfrac{1}{p}\right[$, $G_Y(t) = \dfrac{qt}{1-pt}$.
3. Montrer que G_Y est 2 fois dérivable en 1 et que $G'(1) = \dfrac{1}{q}$ et $G''(1) = \dfrac{2p}{q^2}$.
4. Donner les valeurs de $\mathbb{E}(Y)$ et de $\mathbb{V}(Y)$.

Partie II — Séries entières

Soit $z \in \mathbb{C}$ et $a \in \mathbb{C}^*$. Pour $n \in \mathbb{N}$, on pose $u_n(a) = -\dfrac{1}{a^{n+1}}$

5. Montrer que $\sum u_n(a)z^n$ est une série entière de rayon de convergence égal à $|a|$.
6. Montrer que si $|z| < |a|$, on a : $\dfrac{1}{z-a} = \displaystyle\sum_{n=0}^{+\infty} u_n(a)z^n$.

Soit a, b et λ des nombres complexes non nuls. Dans les questions 7. à 10., on suppose que $|a| < |b|$.

On définit alors, pour tout $n \in \mathbb{N}$,

$$v_n = \sum_{k=0}^{n} u_k(a)u_{n-k}(b)$$

et, pour tout réel t tel que $|t| < |a|$,

$$f(t) = \frac{\lambda t^2}{(t-a)(t-b)}$$

7. Montrer que l'on a :

$$v_n = \frac{1}{ab^{n+1}} \sum_{k=0}^{n} \left(\frac{b}{a}\right)^k = \frac{1}{b-a}\left(\frac{1}{a^{n+1}} - \frac{1}{b^{n+1}}\right)$$

8. Trouver un équivalent simple de v_n quand n tend vers $+\infty$.
9. En déduire que le rayon de convergence de $\sum v_n z^n$ est égal à $|a|$ et que si $|z| < |a|$, alors

$$\frac{1}{(z-a)(z-b)} = \sum_{n=0}^{+\infty} v_n z^n.$$

10. Justifier que f est développable en série entière au voisinage de 0 et que la série entière qui lui est associée possède un rayon de convergence R_f tel que $R_f = |a|$.

Soit a, b, c et λ des nombres complexes non nuls. On suppose que : $|a| \leqslant |b| \leqslant |c|$.

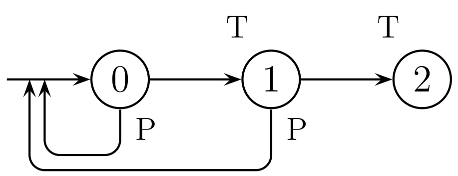
Pour tout réel t tel que $|t| < |a|$, on pose : $g(t) = \dfrac{\lambda t^3}{(t-a)(t-b)(t-c)}$.

11. Justifier que g est développable en série entière au voisinage de 0 et que la série entière qui lui est associée possède un rayon de convergence R_g tel que $R_g \geqslant |a|$.

Partie III — Étude d'un cas intermédiaire

Dans cette partie, on suppose que l'automate passe du niveau 0 au niveau 1 en générant la lettre T. De même, l'automate passe du niveau 1 au niveau 2 en générant la lettre T. Si, en revanche, il génère la lettre P, alors qu'il est au niveau 0 ou 1, il retombe au niveau 0. L'expérience s'arrête dès que l'automate a atteint le niveau 2, c'est-à-dire dès que l'automate aura généré la séquence TT. On résume l'expérience par la figure 2 suivante :

Figure 2



On note Z l'instant où, pour la première fois, l'automate atteint le niveau 2. Ainsi Z est le temps d'attente de la séquence TT.

On admet que Z est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ telle que $Z(\Omega) \subset \mathbb{N}^*$.

Pour tout $n \in \mathbb{N}^*$, on note $p_n = \mathbb{P}(Z = n)$.

On note G_Z la série génératrice de Z et R_Z son rayon de convergence. On rappelle que $R_Z \geqslant 1$.

12. Calculer p_1, p_2 et p_3.
13. Justifier que $(P_1, T_1 \cap P_2, T_1 \cap T_2)$ est un système complet d'évènements.
14. En déduire que pour tout $n \geqslant 3$, on a : $p_n = pp_{n-1} + pqp_{n-2}$.
15. En déduire que pour tout $t \in [-1,1]$, on a : $G_Z(t)(1 - pt - pqt^2) = q^2t^2$.

Pour $t \in \mathbb{R}$, on note $Q(t) = 1 - pt - pqt^2$, $\Delta = p^2 + 4pq > 0$, $a = \dfrac{\sqrt{\Delta} - p}{2pq}$ et $b = \dfrac{-\sqrt{\Delta} - p}{2pq}$.

16. Montrer que $Q(-1) = 1 + p^2 > 0$ et que $Q(1) = q^2 > 0$.
17. Montrer que, pour tout $t \in \mathbb{R}$, $Q(t) = -pq(t-a)(t-b)$.
18. Montrer que $1 < |a| < |b|$.

Pour tout réel t tel que $|t| < |a|$, on définit $f(t) = \dfrac{q^2t^2}{1 - pt - pqt^2}$.

19. Montrer à l'aide de la question 10. que f est développable en série entière au voisinage de 0, que sa série entière associée est G_Z et que $R_Z = |a|$.
20. Montrer que, pour tout $t \in]-|a|, |a|[$, on a : $G_Z(t) = \dfrac{q^2t^2}{1 - pt - pqt^2}$.
21. Montrer que Z admet une espérance et une variance puis que $\mathbb{E}(Z) = q^{-1} + q^{-2}$.
22. Vérifier, à l'aide des questions 4. et 21., que $\mathbb{E}(Z) \geqslant \mathbb{E}(Y) + 1$ où Y est la variable aléatoire définie en partie I.
23. Pouvait-on prévoir ce résultat ?

Partie IV — Algèbre linéaire

On considère les matrices $I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} p & 0 & p & 0 \\ q & q & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & q & 0 \end{pmatrix}$ et $L = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$.

Soit $t \in \mathbb{R}$. On note χ_A le polynôme caractéristique de A, i.e. $\chi_A(t) = \det(tI_4 - A)$.

24. Montrer que 0 est valeur propre de A et donner un vecteur propre de A associé à la valeur propre 0.
25. Trouver les réels α, β et γ tels que, pour tout $t \in \mathbb{R}$, $\chi_A(t) = t^4 - t^3 + \alpha t^2 + \beta t + \gamma$.

On dit que la matrice colonne $S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$ est solution de (E_t) lorsque $S = tAS + L$.

26. Montrer que, pour tout $t \in \mathbb{R}$, S est solution de (E_t) si et seulement si $(I_4 - tA)S = L$.

Pour tout $t \in \mathbb{R}$, on note $\psi_A(t)$ le déterminant de la matrice $I_4 - tA$.

27. Montrer que pour tout $t \in \mathbb{R}^*$, $\psi_A(t) = t^4 \chi_A\left(\dfrac{1}{t}\right)$.
28. Vérifier que pour tout $t \in \mathbb{R}$, $\psi_A(t) = -p^2qt^3 + pqt^2 - t + 1$.
29. En déduire que, pour t au voisinage de 0, l'équation (E_t) possède une unique solution S.

Pour tout $k \in [\![1,4]\!]$, on note U_k la k-ième colonne de $I_4 - tA$. On note $\mathscr{B}$ la base canonique de $\mathscr{M}_{4,1}(\mathbb{C})$ et on suppose que la matrice colonne $S = \begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix}$ est solution de (E_t).

30. Vérifier que $L = U_1S_0 + U_2S_1 + U_3S_2 + U_4S_3$.

31. En déduire que $\det_{\mathscr{B}}(U_1, U_2, U_3, L) = S_3 \cdot \det_{\mathscr{B}}(U_1, U_2, U_3, U_4) = S_3 \cdot \psi_A(t)$.
32. Montrer que, pour t au voisinage de 0, on a l'égalité :

$$S_3 = \frac{pq^2t^3}{-p^2qt^3 + pqt^2 - t + 1}.$$

On se propose de déterminer certaines propriétés des valeurs propres de A. On note λ une valeur propre complexe non nulle de A.

33. Montrer que λ est valeur propre de la matrice transposée de A.
34. En déduire qu'il existe trois complexes non tous nuls x_1, x_2 et x_3 tels que :

$$(\mathscr{H}) \qquad \left\{\begin{array}{ccccc} px_1 & + & qx_2 & = & \lambda x_1 \\ qx_2 & + & px_3 & = & \lambda x_2 \\ & & px_1 & = & \lambda x_3 \end{array}\right. .$$

On considère désormais trois complexes non tous nuls x_1, x_2 et x_3 qui vérifient le système $(\mathscr{H})$. On note alors $M = \max(|x_1|, |x_2|, |x_3|)$ et on remarque que l'on peut toujours se placer dans l'un des trois cas suivants :

(i) $M = |x_3|$

(ii) $M = |x_2|$ avec $M > |x_3|$;

(iii) $M = |x_1|$ avec $M > |x_2|$ et $M > |x_3|$.

35. Montrer, en distinguant ces trois cas, que $|\lambda| < 1$.
36. Montrer l'existence de nombres complexes λ_1, λ_2 et λ_3 tels que :

$0 < \ |\lambda_1| \leqslant |\lambda_2| \leqslant |\lambda_3| < 1 \quad$ et $\quad \forall t \in \mathbb{R}, \ \chi_A(t) = t(t - \lambda_1)(t - \lambda_2)(t - \lambda_3)$.

37. Montrer l'existence de nombres complexes μ, a, b et c tels que $\mu \neq 0$, $\qquad 1 < |a| \leqslant |b| \leqslant |c|$ et

$$\forall t \in \mathbb{R}, \qquad \psi_A(t) = \mu(t - a)(t - b)(t - c)$$

Partie V — Étude d'un dernier cas

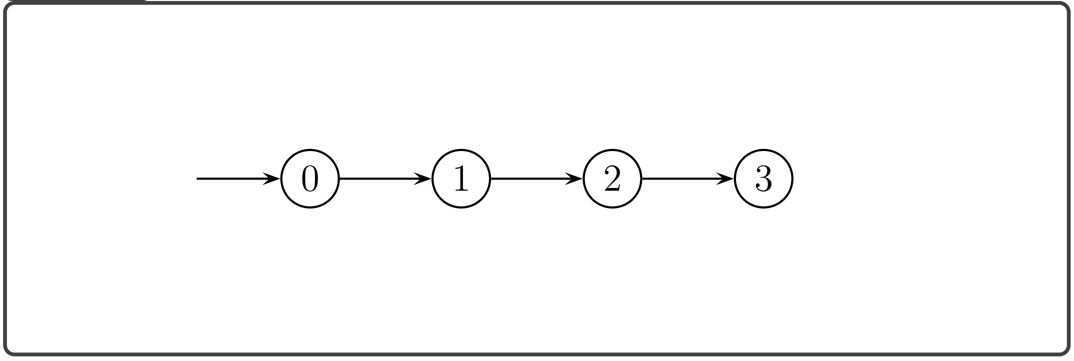
Dans cette partie, on suppose que :

- l'automate passe du niveau 0 au niveau 1 en générant la lettre T ;
- l'automate passe du niveau 1 au niveau 2 en générant la lettre P ;
- l'automate passe du niveau 2 au niveau 3 en générant la lettre T ;
- si l'automate est au niveau 0 ou 2 et qu'il génère la lettre P, alors il retombe au niveau 0 ;
- si l'automate est au niveau 1 et qu'il génère la lettre T, alors il reste au niveau 1.

L'expérience s'arrête dès que l'automate a atteint le niveau 3, c'est-à-dire dès que l'automate aura généré la séquence TPT.

38. Reproduire, sur votre copie, la figure 3 suivante en la complétant pour résumer l'expérience de cette partie V.

Figure 3



Pour $i \in [\![0,3]\!]$ et $n \in \mathbb{N}^*$, on note

— $E_{n,i}$ l'événement « après avoir généré la n-ième lettre, l'automate se trouve au niveau i »

— $E_{0,i}$ l'événement « l'automate se trouve initialement au niveau i »

On pose $p_{n,i} = \mathbb{P}(E_{n,i})$ et pour tout $t \in [-1,1]$, on définit $S_i(t) = \displaystyle\sum_{n=0}^{+\infty} p_{n,i} t^n$.

On note X l'instant où, pour la première fois, l'automate atteint le niveau 3.

On admet que X est une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ telle que $X(\Omega) \subset \mathbb{N}^*$.

On remarque que la série génératrice de X (notée G_X) est alors S_3 et on note R_X son rayon de convergence. On rappelle que $R_X \geqslant 1$.

39. Déterminer $p_{0,0}$, $p_{0,1}$, $p_{0,2}$ et $p_{0,3}$.
40. Montrer que pour tout $n \in \mathbb{N}^*$, on a : $\left\{\begin{array}{lllll} p_{n,0} & = & p \cdot p_{n-1,0} & + & p \cdot p_{n-1,2} \\ p_{n,1} & = & q \cdot p_{n-1,0} & + & q \cdot p_{n-1,1} \\ p_{n,2} & = & p \cdot p_{n-1,1} & & \\ p_{n,3} & = & q \cdot p_{n-1,2} & & \end{array}\right.$

Soit $t \in [-1,1]$. On note $S(t)$ la matrice colonne suivante : $S(t) = \begin{pmatrix} S_0(t) \\ S_1(t) \\ S_2(t) \\ S_3(t) \end{pmatrix}$.

41. Montrer que $\left\{\begin{array}{lllllll} S_0(t) & = & tp \cdot S_0(t) & + & tp \cdot S_2(t) & + & 1 \\ S_1(t) & = & tq \cdot S_0(t) & + & tq \cdot S_1(t) & & \\ S_2(t) & = & tp \cdot S_1(t) & & & & \\ S_3(t) & = & tq \cdot S_2(t) & & & & \end{array}\right.$.
42. Montrer que la matrice colonne $S(t)$ est solution de l'équation (E_t) définie en partie IV.
43. Montrer que
$$\forall t \in]-R_X, R_X[,\ G_X(t) = \frac{pq^2t^3}{-p^2qt^3 + pqt^2 - t + 1}$$
et montrer que $R_X > 1$.
44. Montrer que X admet une espérance et une variance.
45. Donner l'expression de $\mathbb{E}(X)$ en fonction de q seulement.
46. Proposer une méthode permettant de déterminer le temps d'attente moyen de la première réalisation par l'automate de la séquence TTPPT : on précisera notamment le schéma des six niveaux correspondants et la matrice analogue à A que l'on peut faire intervenir dans ce problème.

Corrigé

Réponse du problème

Partie I — Étude d'un cas simple

1. Y est le rang du premier succès dans une suite d'expériences de Bernoulli **indépendantes** : cette variable aléatoire suit donc une loi géométrique de paramètre q (la probabilité de générer T).

   $\boxed{Y \hookrightarrow \mathscr{G}(q) \text{ et, pour tout } n \in \mathbb{N}^*,\ \mathbb{P}(Y=n) = p^{n-1}q}$

2. La série génératrice de Y est donc $\displaystyle\sum_{n\geqslant 1} \frac{q}{p}(pt)^n$, de rayon de convergence $\dfrac{1}{p}$.

   La somme de cette série vaut $G_Y(t) = \dfrac{q}{p}\displaystyle\sum_{n=1}^{+\infty}(pt)^n = \frac{q}{p}\frac{pt}{1-pt}$.

   Ainsi $\boxed{R_Y = \dfrac{1}{p} > 1 \text{ et, pour tout } t \in \left]-\dfrac{1}{p}, \dfrac{1}{p}\right[\ G_Y(t) = \dfrac{qt}{1-pt}}$.

3. En tant que somme de série entière, on sait que G_Y est de classe $\mathcal{C}^\infty$ sur $]-R, R[$, les dérivées se calculant en dérivant terme à terme. Comme $R_Y > 1$ (car $0 < p < 1$) on en déduit que, en particulier, $\boxed{G_Y \text{ est deux fois dérivable en } 1.}$

   De plus pour tout $t \in \left]-\dfrac{1}{p}, \dfrac{1}{p}\right[$,

   $$G'_Y(t) = \frac{q(1-pt)+qtp}{(1-pt)^2} = \frac{q}{(1-pt)^2} \quad \text{et} \quad G''_Y(t) = \frac{2pq}{(1-pt)^3}$$

   D'où $\boxed{G'_Y(1) = \dfrac{1}{q} \text{ et } G''_Y(1) = \dfrac{2p}{q^2}.}$

4. On sait que $G'_Y(1) = \displaystyle\sum_{n=1}^{+\infty} n\mathbb{P}(X=n) = \mathbb{E}(Y)$, et $G''(Y) = \displaystyle\sum_{n=2}^{+\infty} n(n-1)\mathbb{P}(X=n) = \mathbb{E}(Y(Y-1))$

   Ainsi $\boxed{\mathbb{E}(Y) = G'_Y(1) = \dfrac{1}{q} \text{ et } \mathbb{V}(Y) = G''_Y(1) + G'_Y(1) - G'_Y(1)^2 = \dfrac{p}{q^2}.}$

Partie II — Séries entières

5. On est à nouveau face à une série entière géométrique de raison $\dfrac{1}{a}$; le cours ou l'argument vu plus haut nous donne son rayon de convergence :

   $\boxed{\text{Le rayon de convergence de la série entière } \sum\left(-\dfrac{1}{a^{n+1}}\right)z^n \text{ vaut } |a|.}$

6. Pour $|z| < |a|$, notre fine connaissance des séries géométriques nous donne :

   $$\sum_{n=0}^{+\infty}\left(-\frac{1}{a^{n+1}}\right)z^n = -\frac{1}{a}\sum_{n=0}^{+\infty}\left(\frac{z}{a}\right)^n = -\frac{1}{a}\cdot\frac{1}{1-\frac{z}{a}},$$

   Ainsi, $\boxed{\text{pour } |z| < |a|,\ \displaystyle\sum_{n=0}^{+\infty}\left(-\frac{1}{a^{n+1}}\right)z^n = \frac{1}{z-a}.}$

7. Rappelons que, pour $(x,y) \in \mathbb{R}^2$ et $p \in \mathbb{N}$ on a

$$x^p - y^p = (x-y)\sum_{k=0}^{p-1} x^k y^{p-1-k}$$

En particulier pour $x = \frac{1}{a}$, $y = \frac{1}{b}$ et $p = n+1$ on obtient

$$\frac{1}{a^{n+1}} - \frac{1}{b^{n+1}} = \left(\frac{1}{a} - \frac{1}{b}\right)\sum_{k=0}^{n} \frac{1}{a^k}\frac{1}{b^{n-k}} = \frac{b-a}{ab}\sum_{k=0}^{n} \frac{1}{a^k}\frac{1}{b^{n-k}}$$

D'où

$$v_n = \sum_{k=0}^{n} \frac{1}{a^{k+1}} \cdot \frac{1}{b^{n+1-k}} = \frac{1}{ab}\sum_{k=0}^{n} \frac{1}{a^k} \cdot \frac{1}{b^{n-k}} = \frac{1}{b-a}\left(\frac{1}{a^{n+1}} - \frac{1}{b^{n+1}}\right)$$

On obtient bien $\boxed{v_n = \frac{1}{b-a}\left(\frac{1}{a^{n+1}} - \frac{1}{b^{n+1}}\right).}$

8. On a $|a| < |b|$, donc $\frac{1}{|a|} > \frac{1}{|b|}$

Ainsi $\left(\frac{1}{b}\right)^{n+1} \underset{n\to+\infty}{=} o\left(\left(\frac{1}{a}\right)^{n+1}\right)$ et, par conséquent

$$\boxed{v_n \underset{n\to+\infty}{\sim} \frac{1}{(b-a)a^{n+1}}}$$

9. La suite $(v_n)_{n\in\mathbb{N}}$ est non-nulle à partir d'un certain rang.

De plus

$$\left|\frac{v_{n+1}}{v_n}\right| \underset{n\to+\infty}{\sim} \left|\frac{(b-a)a^{n+1}}{(b-a)a^{n+2}}\right| \underset{n\to+\infty}{\sim} \frac{1}{|a|}$$

Ainsi, par critère de D'Alembert, $\boxed{\text{Le rayon de convergence de } \sum v_n z^n \text{ vaut } |a|.}$

Pour $|z < |a$, les séries $\sum u_n(a)z^n$ et $\sum v_n(a)z^n$ sont absolument convergentes, donc leur produit de Cauchy également, la somme du produit de Cauchy étant égale au produit des sommes des deux séries, on en déduit que

$$\boxed{\text{Pour } |z < |a,\ \sum_{n=0}^{+\infty} v_n z^n = \frac{1}{(z-a)(z-b)}}$$

10. Pour $t \in]-a, a[$, on a $f(t) = \sum_{n=0}^{+\infty} v_n t^{n+2} = \sum_{k=2}^{+\infty} v_{k-2} t^k$.

Il s'agit bien d'un développement en série entière et le rayon de convergence de la série associée est le même que celui de $\sum v_n z^n$. $\boxed{\text{Ce qui est ce qu'on voulait montrer.}}$

11. On note que $g(t) = f(t) \cdot \frac{1}{t-c}$,

Par produit de Cauchy de deux séries de rayon de convergence respectifs $|a|$ et $|c|$. De plus $|a| \leqslant |c|$. Ainsi

$\boxed{g \text{ est développable en série entière au voisinage de 0, avec un rayon de convergence } R \geqslant \min(|a|,|c|) = |a|.}$

Produit de Cauchy

Les arguments sur les produits de Cauchy donnent seulement une minoration du rayon, la valeur exacte obtenue à la question 9. n'étant acquise que par la forme particulière de la série entière.

Partie III — Étude d'un cas intermédiaire

12. On a clairement $p_1 = 0$.

    Ensuite, $p_2 = \mathbb{P}(T_1 \cap T_2) = \mathbb{P}(T_1)\mathbb{P}(T_2)$ par indépendance donc $p_2 = q^2$.

    Enfin, la seule possibilité pour avoir $Z = 3$ est que les trois premières lettres soient PTT, donc, toujours par indépendance,$p_3 = \mathbb{P}(P_1 \cap T_2 \cap T_3) = \mathbb{P}(P_1)\mathbb{P}(T_2)\mathbb{P}(T_3) = pq^2$.

    Finalement $\boxed{p_1 = 0,\ p_2 = q^2 \text{ et } p_3 = pq^2.}$

13. On a $P_1 \cup (T_1 \cap P_2) \cup (T_1 \cap T_2) = P_1 \cup (T_1 \cap (P_2 \cup P_2)) = P_1 \cup T_1 = \Omega$

    De plus $P_1 \cap (T_1 \cap P_2) = P_1 \cap T_1 \cap P_2 = \emptyset$, $P_1 \cap (T_1 \cap T_2) = P_1 \cap T_1 \cap T_2 = \emptyset$ et $(T_1 \cap P_2) \cap (T_1 \cap T_2) = T_1 \cap P_2 \cap T_1 \cap T_2 = \emptyset$

    Ainsi $\boxed{(P_1, T_1 \cap P_2, T_1 \cap T_2) \text{ est un système complet d'événements.}}$

14. La formule des probabilités totales pour le système complet d'événements précédent nous donne, pour $n \in \mathbb{N}$

    $$p_n = \mathbb{P}(Z = n) = \mathbb{P}(Z = n|P_1)\mathbb{P}(P_1) + \mathbb{P}(Z = n|T_1 \cap P_2)\mathbb{P}(T_1 \cap P_2) + \mathbb{P}(Z = n|T_1 \cap T_2)\mathbb{P}(T_1 \cap T_2)$$

    Si les deux premières lettres sont T_1T_2, alors $Z = 2$ donc $Z \neq n$, ainsi $\mathbb{P}(Z = n|T81 \cap T_2) = 0$.

    Si la première lettre est P_1, alors l'automate est retourné dans l'état zéro, donc la probabilité pour que $n - 1$ lettres plus tard il soit dans l'état 2 vaut $\mathbb{P}(Z = n - 1)$ i.e $\mathbb{P}(Z = n|P_1) = \mathbb{P}(Z = n - 1) = p_{n-1}$.

    De même, si après 2 lettres on est dans l'état 0 alors la probabilité de se retrouver dans l'état 2 après les $n - 2$ lettres suivantes est $\mathbb{P}(Z = n|T_1 \cap P_2) = \mathbb{P}(Z = n - 2) = p_{n-2}$.

    Enfin, $\mathbb{P}(P_1) = p$ et, par indépendance, $\mathbb{P}(T_1 \cap P_2) = \mathbb{P}(T_1)\mathbb{P}(P_2) = pq$.

    Ainsi $\boxed{\text{pour tout } n \geqslant 3,\ p_n = p\,p_{n-1} + pq\,p_{n-2}.}$

    > **Explication**
    > La première étape nous amène à rester sur place. Le lemme des coalitions nous disant que l'étape 1 est indépendante du processus à partir du temps 2 tout se passe comme si la première étape n'avait pas eu lieu.

15. Soit $t \in [-1, 1]$, on a alors

    $$\forall n \geqslant 3, \qquad p_n t^n = p\,p_{n-1}t^n + pq\,p_{n-2}t^n$$

    On somme pour $n \in [\![3, +\infty[\![$ ce qui est licite car les séries en jeu sont toutes absolument convergentes (les termes généraux sont positifs et majorés par $p_n = \mathbb{P}(Z = n)$, terme général d'une série absolument convergente). On obtient ainsi

    $$\sum_{n=3}^{+\infty} p_n t^n = pt \sum_{n=3}^{+\infty} p_{n-1}t^{n-1} + t^2 pq \sum_{n=3}^{+\infty} p_{n-2}t^{n-2} \qquad (R)$$

    Par décalages d'indices

    $$\sum_{n=3}^{+\infty} p_{n-1}t^{n-1} = \sum_{i=2}^{+\infty} p_i t^i = G_Z(t) - (p_0 + p_1 t) = G_Z(t)$$

    et $\displaystyle\sum_{n=3}^{+\infty} p_{n-2}t^{n-2} = G_Z(t)$.

    Puisque le membre de gauche de (R) vaut $G_Z(t) - p_2t^2 = q^2t^2$, on obtient $G_Z(t) - q^2t^2 = (pt + pqt^2)G_Z(t)$

    Ainsi $\boxed{\text{pour tout } t \in [-1, 1],\ G_Z(t)(1 - pt - pqt^2) = q^2t^2.}$

16.

    $$Q(-1) = 1 + p - pq = 1 + p - p(1 - p) = 1 + p^2 \quad \text{et} \quad Q(1) = 1 - p - pq = q - pq = q(1 - p) = q^2$$

    Ainsi $\boxed{Q(-1) = 1 + p^2 > 0 \text{ et } Q(1) = q^2 > 0}$

17. Le polynôme $Q = 1 - pX - pqX^2$ est de degré 2, coefficient dominant $-pq$ et possède deux racines réelles distinctes a et b. On peut donc le factoriser sous la forme $Q = -pq(X-a)(X-b)$.

    Ainsi $\boxed{\text{pour tout } t \in \mathbb{R},\ Q(t) = -pq(t-a)(t-b).}$

18. L'application Q vérifie $Q(-1) = 1 > 0$, $\lim\limits_{t \to -\infty} Q(t) = -\infty$ et est **continue**, donc le théorème des valeurs intermédiaires nous assure qu'elle s'annule au moins une fois dans $]-\infty, -1[$. De même, elle s'annule au moins une fois $]1, +\infty[$.

    Par ailleurs, les deux seules racines de Q sont a et b avec $b < a$, donc $b < -1 < 0 < 1 < a$.

    Il s'ensuit que
    $$|b| = -b = \frac{\sqrt{\Delta} + p}{2pq} > \frac{\sqrt{\Delta} - p}{2pq} = a = |a| > 1$$
    et ainsi $\boxed{1 < |a| < |b|.}$

19. D'après les questions précédentes, pour $t \in]-|a|, |a|[$, $f(t) = \dfrac{-\frac{q}{p}t^2}{(t-a)(t-b)}$,

    On applique la question 10. s'applique avec $\lambda = -\dfrac{q}{p}$ (la condition $|a| < |b|$ étant bien vérifiée).

    Ainsi $\boxed{f \text{ est développable en série entière, de série entière associée } G_Z \text{ ayant un rayon de convergence } R_Z = |a|}$

20. On a montré dans ce qui précède que la relation $G_Z(t) = \dfrac{q^2 t}{1 - pt - pqt^2}$ est valable pour $t \in]1, 1[$.

    Il s'agit donc de l'étendre à $]-|a|, |a|[$. Or on sait que
    $$\forall t \in]-1, 1[, \qquad \sum_{n=0}^{+\infty} p_n t^n = \sum_{n=0}^{+\infty} v_n t^n$$
    Donc **par unicité du développement en série entière**, on a, pour tout $n \in \mathbb{N}$, $p_n = v_n$. La relation s'étend alors bien à $]-|a|, |a|[$.

    Ainsi $\boxed{\text{pour tout } t \in]-|a|, |a|[,\ G_Z(t) = \dfrac{q^2 t^2}{1 - pt - pqt^2}.}$

    Subtilité

    Cette question est plus intéressante et fine qu'il n'y parait. Il nous faut ici montrer que deux séries entières qui coïncident sur un disque contenant 0 sont en fait égales partout.

21. Une somme de série entière de rayon de convergence $R > 0$ est de classe $\mathscr{C}^\infty$ sur $]-R, R[$.

    Ici, $R_Z = |a| > 1$, donc G_Z est deux fois dérivable en 1. Ainsi $\boxed{Z \text{ possède une espérance et une variance.}}$

    De plus
    $$\forall t \in]-R_Z, R_Z[, \qquad G_Z'(t) = \frac{2tq^2(1 - pt - pqt^2) + q^2t^2(p + 2pqt)}{(1 - pt - pqt^2)^2}$$
    Puis
    $$\mathbb{E}(Z) = G_Z'(1) = 2 + \frac{p}{q^2} + 2\frac{p}{q} = 2 + \frac{1-q}{q^2} + 2\frac{1-q}{q}$$
    D'où $\boxed{\mathbb{E}(Z) = \dfrac{1}{q^2} + \dfrac{1}{q}.}$

22. Il s'agit d'établir que $\dfrac{1}{q} + \dfrac{1}{q^2} \geqslant 1 + \dfrac{p}{q^2}$

    Or
    $$\begin{aligned} \frac{1}{q} + \frac{1}{q^2} \geqslant 1 + \frac{p}{q^2} &\Leftrightarrow q + 1 \geqslant q^2 + p = q^2 + 1 - q \\ &\Leftrightarrow q^2 \leqslant 2q \\ &\Leftrightarrow q \leqslant 2 \end{aligned}$$

Cette dernière inégalité est vérifiée, ce qui prouve l'inégalité demandée : $\boxed{\mathbb{E}(Z) \geqslant \mathbb{E}(Y) + 1.}$

23. La première occurrence de TT est évidemment **strictement** précédée par la première occurrence de T, donc on a toujours $Z \geqslant Y + 1$. Par croissance de l'espérance, on on déduit : $\mathbb{E}(Z) \geqslant 1 + \mathbb{E}(Y)$.

    $\boxed{\text{Le résultat était donc prévisible.}}$

Partie IV — Algèbre linéaire

24. En notant $X_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, on a $AX_0 = 0 = 0.X_0$.

    Comme $X_0 \neq 0$, ceci prouve que $\boxed{0 \text{ est valeur propre de } A, \text{ un vecteur propre associé étant } \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}$

25. Soit $t \in \mathbb{R}$.

$$\begin{aligned} \chi_A(t) &= \begin{vmatrix} t-p & 0 & -p & 0 \\ -q & t-q & 0 & 0 \\ 0 & -p & t & 0 \\ 0 & 0 & -q & t \end{vmatrix} \\ &= t \begin{vmatrix} t-p & 0 & -p \\ -q & t-q & 0 \\ 0 & -p & t \end{vmatrix} \\ &= t\left((t-p)(t-q)t + q(-(-p)(-p))\right) \\ &= t\left(t^3 - \underbrace{(p+q)}_{=1} t^2 + pqt - qp^2\right) \end{aligned}$$

    Ainsi $\boxed{\text{pour tout } t \in \mathbb{R},\ \chi_A(t) = t^4 - t^3 + \underbrace{pq}_{\alpha} t^2 \underbrace{-qp^2}_{\beta} t + \underbrace{0}_{\gamma}}$

26. S est solution de (E_t) si et seulement si $A - tAS = L$, c'est-à-dire si et seulement $(I_4 - tA)S = L$. $\boxed{\text{Ce qui prouve le résultat demandé.}}$ Oui cette question est triviale.

27. On rappelle que lorsque $\lambda \in \mathbb{K}$ et $M \in \mathcal{M}_n(\mathbb{K})$, la linéarité du déterminant **par rapport à chacune de ses colonnes** implique : $\det(\lambda M) = \lambda^n \det(M)$, donc ici, pour $t \neq 0$,

$$\psi_A(t) = \det(I_4 - tA) = \det\left(t\left(\frac{1}{t}I_4 - A\right)\right) = t^4 \det\left(\frac{1}{t}I_4 - A\right)$$

    Ainsi, $\boxed{\text{pour } t \in \mathbb{R}^*,\ \psi_A(t) = t^4 \chi_A\left(\frac{1}{t}\right).}$

28. D'après ce qui précède et la question 25., on a, pour $t \neq 0$,

$$\psi_A(t) = t^4\left(\frac{1}{t^4} - \frac{1}{t^3} + pq\frac{1}{t^2} - p^2q\frac{1}{t}\right) = 1 - t + pqt^2 - p^2qt^3$$

    Pour $t = 0$, $\psi_A(0) = \det(I_4) = 1 = -p^2q0^3 + pq0^2 - 0 + 1$.

    D'où, $\boxed{\text{pour tout } t \in \mathbb{R},\ \psi_A(t) = -p^2qt^3 + pqt - t + 1.}$

29. Puisque ψ_A est continue et non nulle en 0, il existe un voisinage de 0 (i.e. un intervalle de la forme $]-\varepsilon, \varepsilon[$ avec $\varepsilon > 0$) sur lequel ψ_A ne s'annule pas.

    Sur ce voisinage, $I_4 - tA$ est alors inversible, ainsi l'équation (E_t), qui est équivalente à $(I_4 - tA)S = L$, possède une unique solution (accessoirement, c'est $S = (I_4 - tA)^{-1}L$).

    $\boxed{\text{Donc, pour } t \text{ au voisinage de } 0,\ (E_t) \text{ possède une unique solution } S.}$

30. Le calcul par bloc de $(I_4 - tA)S$ donne

⚠Attention

les U_k sont des colonnes et les S_k des scalaires

$$(I_4 - tA)S = \big(U_1|U_2|U_3|U_4\big)\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = S_0U_1 + \cdots + S_3U_4.$$

Mais par ailleurs on a supposé que $(I_4 - tA)S = L$. Ainsi $\boxed{L = S_0U_1 + S_1U_1 + S_2U_3 + S_3U_4}$

31. Le caractère multilinéaire et alterné du déterminant nous assure que :

$$\begin{aligned}
\det_{\mathcal{B}}(U_1, U_2, U_3, L) &= \det_{\mathcal{B}}(U_1, U_2, U_3, S_0U_1 + S_1U_1 + S_2U_3 + S_3U_4) \\
&= S_0 \underbrace{\det_{\mathcal{B}}(\mathbf{U_1}, U_2, U_3, \mathbf{U_1})}_{0} + S_1 \underbrace{\det_{\mathcal{B}}(U_1, \mathbf{U_2}, U_3, \mathbf{U_2})}_{0} + S_2 \underbrace{\det_{\mathcal{B}}(U_1, U_2, \mathbf{U_3}, \mathbf{U_3})}_{0} + S_3 \det_{\mathcal{B}}(U_1, U_2, U_3, U_4) \\
&= S_3 \det_{\mathcal{B}}(U_1, U_2, U_3, U_4) \\
&= S_3 \det(I_4 - tA) \\
&= S_3\psi_A(t)
\end{aligned}$$

Ainsi $\boxed{\det_{\mathcal{B}}(U_1, U_2, U_3, L) = S_3 \det(U_1, U_2, U_3, U_4) = S_3\psi_A(t).}$

32. Au voisinage de 0, on a $\psi_A(t) \neq 0$, et $S_3 = \dfrac{\det_{\mathcal{B}}(U_1, U_2, U_3, L)}{\psi_A(t)}$.

    Or le déterminant au numérateur se calcule très simplement en développant par rapport à la dernière colonne :

$$\det_{\mathcal{B}}(U_1, U_2, U_3, L) = \begin{vmatrix} 1-pt & 0 & -pt & 1 \\ -qt & 1-qt & 0 & 0 \\ 0 & -pt & 1 & 0 \\ 0 & 0 & -qt & 0 \end{vmatrix} = -\begin{vmatrix} -qt & 1-qt & 0 \\ 0 & -pt & 1 \\ 0 & 0 & -qt \end{vmatrix} = pq^2t^3$$

D'où $\boxed{\text{au voisinage de } 0,\ S_3 = \dfrac{pq^2t^3}{-p^2qt^3 + pqt^2 - t + 1}.}$

33. On sait A et $A^\top$ ont le même spectre. Puisque λ est valeur propre de A, le résultat suit, $\boxed{\lambda \text{ est valeur propre de } A^\top.}$

34. D'après la question précédente, il existe $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$ **non nul** tel que ${}^tAX = \lambda X$, c'est-à-dire :

$$\begin{cases} px_1 + qx_2 = \lambda x_1 \\ qx_2 + px_3 = \lambda x_2 \\ px_1 + qx_4 = \lambda x_3 \\ 0 = \lambda x_4 \end{cases}$$

Puisque $\lambda \neq 0$, la dernière équation fournit $x_4 = 0$, donc x_1, x_2 et x_3 ne sont pas tous nuls. Par ailleurs la troisième équation devient $px_1 = \lambda x_3$.

Ainsi $\boxed{\text{le système } (\mathcal{H}) \text{ possède une solution non nulle.}}$

35. Dans le cas où $M = |x_3| > 0$ (la solution est non nulle, donc le coefficient de plus gros module est non nul), la troisième équation fournit $\lambda = p\dfrac{x_1}{x_3}$, donc $|\lambda| = p\dfrac{|x_1|}{|x_3|}$.

    Puisque $0 < p < 1$ et $\dfrac{|x_1|}{|x_3|} \leqslant 1$, on obtient bien $|\lambda| < 1$.

    Dans le second cas, on regarde la deuxième équation. Par inégalité triangulaire

    $$|\lambda||x_2| \leqslant q|x_2| + p|x_3|$$

    Donc $|\lambda| \leqslant q + p\dfrac{|x_3|}{|x_2|}$.

    Mais $\dfrac{|x_3|}{|x_2|} < 1$ et $p > 0$ donc $p\dfrac{|x_3|}{|x_2|} < p$ puis $|\lambda| < p + q = 1$.

    Le dernier cas se traite de la même façon, en s'intéressant à la première équation de $(\mathcal{H})$.

    Finalement, dans tous les cas $\boxed{|\lambda| < 1.}$

36. χ_A est un polynôme de degré 4, unitaire, et possédant 0 comme racine.

    On sait alors qu'il est scindé sur $\mathbb{C}$ et que plus précisément il se factorise sous la forme $X(X-\alpha_1)(X-\alpha_2)(X-\alpha_3)$, où, a priori, les α_i peuvent être égaux entre eux et éventuellement nuls.

    Or, comme $\chi_A'(0) = 1 \neq 0$, χ_A possède 0 comme racine SIMPLE donc les α_i sont tous non nuls.

    Il ne reste plus à les réordonner pour que les modules soient croissants (et la question précédente nous assure que ces modules sont tous strictement plus petits que 1.

    $$\boxed{\exists \lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}; \quad 0 < |\lambda_1| \leqslant |\lambda_2| \leqslant |\lambda_3| < 1 \quad \text{et} \quad \forall t \in \mathbb{R},\ \chi_A(t) = t(t-\lambda_1)(t-\lambda_2)(t-\lambda_3)}$$

37. On a, pour $t \neq 0$:

    $$\begin{aligned}
    \psi_A(t) &= t^4 \chi_A\left(\frac{1}{t}\right) \\
    &= t^4 \frac{1}{t}\left(\frac{1}{t} - \lambda_1\right)\left(\frac{1}{t} - \lambda_2\right)\left(\frac{1}{t} - \lambda_3\right) \\
    &= t(1-\lambda_1 t)(1-\lambda_2 t)(1-\lambda_3 t) \\
    &= -\lambda_1\lambda_2\lambda_3 t\left(t - \frac{1}{\lambda_1}\right)\left(t - \frac{1}{\lambda_2}\right)\left(t - \frac{1}{\lambda_3}\right)
    \end{aligned}$$

    Puisque $0 < |\lambda_1| \leqslant |\lambda_2| \leqslant |\lambda_3| < 1$, on a $1 < \left|\dfrac{1}{\lambda_3}\right| \leqslant \left|\dfrac{1}{\lambda_2}\right| \leqslant \left|\dfrac{1}{\lambda_1}\right|$, ce qui nous incite à poser $a = \dfrac{1}{\lambda_3}$, $b = \dfrac{1}{\lambda_2}$, $c = \dfrac{1}{\lambda_1}$ et $\mu = -\lambda_1\lambda_2\lambda_3 (= -p^2 q) \neq 0$.

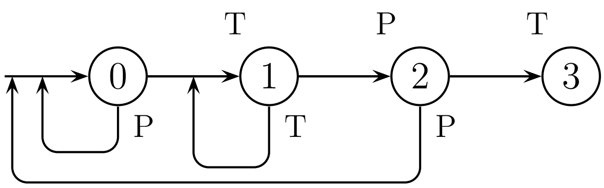
    Il reste à noter que la relation prouvée pour $t \neq 0$ s'étend à $t = 0$ (les deux membres sont égaux à 1).

    Ainsi $\boxed{\text{il existe } (\mu, a, b, c) \in \mathbb{C}^4 \text{ tel que } \mu \neq 0,\ 1 < |a| \leqslant |b| \leqslant |c| \text{ et}}$

    $$\boxed{\forall t \in \mathbb{R}, \qquad \psi_A(t) = \mu(t-a)(t-b)(t-c)}$$

Partie V — Étude d'un dernier cas

Figure 4



38. L'idée intuitive est de s'assurer de passer à l'état 3 au premier moment où on aura rencontré le mot TPT.

39. On commence à l'état 0, d'où $\boxed{S_0(0) = \begin{pmatrix} p_{0,0} \\ p_{0,1} \\ p_{0,2} \\ p_{0,3} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}$

40. On ne détaille que la première équation, les autres étant de même nature. Le principe est le même qu'à la question 14 : on conditionne l'événement $E_{n,0}$ selon le système complet d'événements $(E_{n-1,0}, E_{n-1,1}, E_{n-1,2}, E_{n-1,3})$, sachant que la modélisation du problème nous donne les probabilités conditionnelles $\mathbb{P}_{E_{n-1,i}}(E_{n,0})$: elles sont données par les flèches entrantes dans l'état 0 :

$$\mathbb{E}_{n,0} = \underbrace{\mathbb{P}(E_{n-1,0})}_{p_{n-1,0}} \underbrace{\mathbb{P}_{E_{n-1,0}}(E_{n,0})}_{p} + \underbrace{\mathbb{P}(E_{n-1,1})}_{p_{n-1,1}} \underbrace{\mathbb{P}_{E_{n-1,0}}(E_{n,0})}_{0} + \underbrace{\mathbb{P}(E_{n-1,2})}_{p_{n-1,2}} \underbrace{\mathbb{P}_{E_{n-1,2}}(E_{n,0})}_{p} + \underbrace{\mathbb{P}(E_{n-1,3})}_{p_{n-1,3}} \underbrace{\mathbb{P}_{E_{n-1,3}}(E_{n,0})}_{0},$$

ce qui nous donne exactement la première équation demandée.

Une lecture attentive du graphe nous donne les autres équations, basées sur le même principe.

On obtient bien que,

$$\boxed{\forall n \in \mathbb{N}^*, \qquad \left\{ \begin{array}{lllll} p_{n,0} & = & p \cdot p_{n-1,0} & + & p \cdot p_{n-1,2} \\ p_{n,1} & = & q \cdot p_{n-1,0} & + & q \cdot p_{n-1,1} \\ p_{n,2} & = & p \cdot p_{n-1,1} & & \\ p_{n,3} & = & q \cdot p_{n-1,2} & & \end{array} \right.}$$

41. On reprend le raisonnement de la question 15.

Pour $n \in \mathbb{N}^*$ et $t \in [-1, 1]$ on a

$$p_{n,0} t^n = p p_{n-1,0} t^n + p p_{n-1,2} t^n$$

On somme ensuite pour n allant de 1 à $+\infty$ (toutes les séries étant absolument convergentes).

On obtient alors $S_0(t) - p_{0,0} = tpS_0(t) + tpS_1(t)$, i.e. $S_0(t) = tpS_0(t) + tpS_1(t) + 1$.

Le même principe sur les autres relations nous fournit les autres équations

On obtient bien

$$\boxed{\left\{ \begin{array}{lllllll} S_0(t) & = & tp \cdot S_0(t) & + & tp \cdot S_2(t) & + & 1 \\ S_1(t) & = & tq \cdot S_0(t) & + & tq \cdot S_1(t) & & \\ S_2(t) & = & tp \cdot S_1(t) & & & & \\ S_3(t) & = & tq \cdot S_2(t) & & & & \end{array} \right.}$$

42. On a juste à calculer

$$tAS(t) + L = \begin{pmatrix} tpS_0(t) + tpS_1(t) + 1 \\ tqS_0(t) + tqS_1(t) \\ tpS_1(t) \\ tq \cdot S_2(t) \end{pmatrix} = S(t)$$

$\boxed{\text{Ainsi } S(t) \text{ est solution de l'équation } (E_t).}$

43. La fonction génératrice $G_X = S_3$ vaut, d'après la question 32, pour t au voisinage de 0, $S_3(t) = \dfrac{pq^2t^3}{-p^2qt^3 + pqt^2 - t + 1}$.

    Mais le raisonnement fait en fin de partie II (questions 19 et 20) nous assure qu'en fait, S_3 est de rayon de convergence au moins égal au module de la plus petite racine du dénominateur, et que la relation s'étend du voisinage de 0 à $]-R, R[$.

    La question 37 nous assure que ces racines sont de module strictement plus grand que 1, ce qui prouve que $R_X > 1$.

    Ainsi $\boxed{R_X > 1 \text{ et pour tout } t \in]-R_X, R_X[,\ G_X(t) = \dfrac{pq^2t^3}{-p^2qt^3 + pqt^2 - t + 1}}$.

44. G_X est de classe $\mathcal{C}^\infty$ sur $]-R_X, R_X[$ qui contient $[-1, 1]$, donc G_X est deux fois dérivable en 1.

    Ainsi $\boxed{X \text{ possède une espérance et une variance.}}$

45. Pour $t \in]-R_X, R_X[$ on a

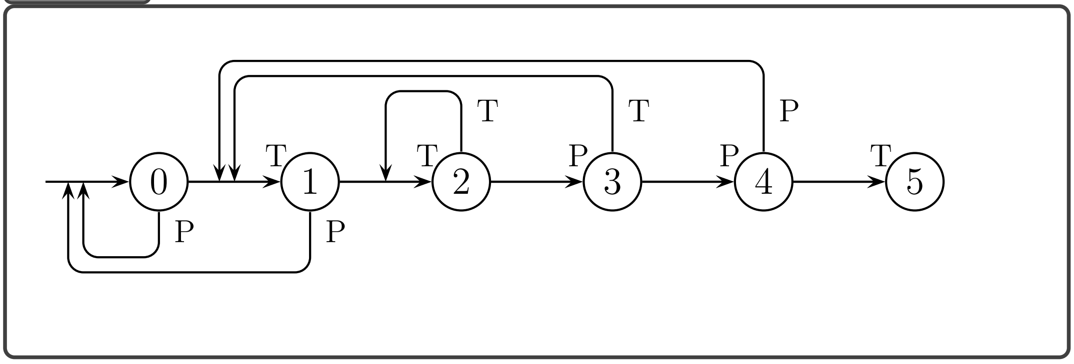
$$G'_X(t) = \frac{pq^2t^2(3 - 2t + pqt^2)}{(-p^2qt^3 + pqt^2 - t + 1)^2}$$

    D'où $G'_X(1) = \dfrac{(1 + pq)}{pq^2} = \dfrac{1 + q - q^2}{q^2(1 - q)}$

    Ainsi $\boxed{\mathbb{E}(X) = \dfrac{1 + q - q^2}{q^2(1 - q)}}$.

46. Commençons par construire l'automate associé à la recherche de ce motif. C'est la même chose que pour la question 18 : quel est le plus gros préfixe de TTPPT que je viens de rencontrer ? Si c'est par exemple TTP, j'aimerais lire P (le plus gros préfixe devient TTPP, je passe à l'état 4) ; mais si c'est T, alors je viens de lire TTPT : le plus gros préfixe de TTPPT en cours est T et je passe dans l'état 1.

    Figure 5



    La matrice associée à cet automate se construit en plaçant, en position (i, j), la probabilité de passer de l'état j à l'état i (les numérotations partent de 0)

$$A = \begin{pmatrix} p & p & 0 & 0 & p & 0 \\ q & 0 & 0 & q & 0 & 0 \\ 0 & q & q & 0 & 0 & 0 \\ 0 & 0 & p & 0 & 0 & 0 \\ 0 & 0 & 0 & p & 0 & 0 \\ 0 & 0 & 0 & 0 & q & 0 \end{pmatrix}$$

    Si on reprend les méthodes de la partie V on obtient, en notant U le temps d'attente du motif $TTPPT$,

$$G_U(t) = S_5(t) = \frac{p^2q^3t^5}{1 - t + p^2q^2t^4 - p^3q^2t^5}$$

Puis, $\mathbb{E}(U) = G'_U(1) = \dfrac{1 + q^2(1-q)^2}{q^3(1-q)^2}$.

Un brin d'ADN peut être vu comme un mot (sur un alphabet à 4 lettres) de longueur de l'ordre de quelques centaines de millions (ou moins, ou plus !). Les recherches de propriétés de ce mot sont très consommatrices « d'algorithmes du texte »(un peu plus élaborés que la simple recherche d'un motif, mais qui restent souvent de nature proche).

L'automate associé à un motif donné m permet de parcourir un grand texte (disons de taille n) et repérer les occurrences de m sans revenir en arrière (ce qu'on ferait avec un algorithme naïf). Le temps de recherche du motif m passe alors de $n|m|$ à n, ce qui est une amélioration considérable, surtout si m a une longueur de l'ordre de n (ce qui arrive dans la vraie vie).

Mais tout ceci nécessite d'avoir construit l'automate préalablement. L'algorithme naïf (chercher le plus gros suffixe qui soit un préfixe...) est de complexité $|m|^3$, ce qui est rédhibitoire si m contient de l'ordre du milliard (ou même seulement du million) de caractères. En 1970, Knuth, Morris et Pratt ont conçu un algorithme très intelligent pour construire cet automate en temps linéaire en $|m|$. Le très belle idée consiste (mais ça constituerait un sujet complet d'informatique de MPI ou d'option info en MP) grosso-modo à construire l'automate en utilisant à la volée ce qui a déjà été construit jusque là.

Sans ces algorithmes efficaces, le séquençage de l'ADN aurait probablement été sans objet : un informaticien donne du travail à des milliers de biologistes.