Dans *l'origine des espèces*, Charles Darwin explique que sans contraintes extérieures, la tendance que les membres d'une espèce ont à se reproduire conduirait à une surpopulation du milieu :

> *«De tous les animaux connus, l'éléphant, pense-t-on, est celui qui se reproduit le plus lentement. J'ai fait quelques calculs pour estimer quel serait le taux minimum de son augmentation en nombre. On peut, sans crainte de se tromper, admettre qu'il commence à se reproduire à l'âge de trente ans, et qu'il continue jusqu'à quatre-vingt-dix ; dans l'intervalle, il produit six petits. Dans 500 ans, il y aurait 15 millions d'éléphants vivants, tous descendants du premier couple».*
>
> Charles Darwin , dans *L'origine des espèces*,1872, chapitre 3.

Comment Charles Darwin a-t-il obtenu ces résultats ?

Le problème se décompose en deux grandes parties ***I*** *et* ***II****. Les thèmes du programme sur lesquels repose chaque partie sont indiqués ci-dessous. Dans une assez large mesure, les parties sont indépendantes.*

— *Dans la partie* ***I.*** *on analyse un modèle déterministe de dynamique d'une population d'éléphants afin de retrouver les résultats de Darwin.*

  A. *Étude de fonction polynomiale.*

  B. *Informatique : programmation de suites.*

  C. *Diagonalisation des matrices.*

— *Dans la partie* ***II.*** *on analyse un modèle probabiliste et on s'intéresse à la probabilité d'extinction de la population.*

  D. *Informatique : simulation de variables aléatoires.*

  E. *Calcul des probabilités : étude de la suite des probabilités d'extinction.*

  F. *Analyse : détermination de la limite d'une suite récurrente.*

I Modèle linéaire déterministe

Mettant de côté les mécanismes de reproduction, on suppose qu'à la date $\tau_0 = 0$, on dispose d'un éléphant qui se reproduit suivant les règles énoncées par Darwin.

Ensuite, l'effectif total de la population d'éléphants est évalué tous les 30 ans. En notant alors, pour tout entier n positif, t_n cet effectif au bout de $30n$ années (ainsi $t_0 = 1$), on partitionne la population d'éléphants en différentes classes d'âge comme suit :

— Les éléphants *immatures* (de 0 à 29 ans) : Ceux-ci ne sont pas en âge de se reproduire. On note i_n l'effectif de cette classe d'âge au bout de $30n$ années.

— Les éléphants *matures* (de 30 à 59 ans) : Ceux-ci se reproduisent et chaque éléphant mature donne naissance à 3 éléphants dans cette période. On note m_n l'effectif de cette classe d'âge au bout de $30n$ années.

— Les éléphants *déclinants* (de 60 à 89 ans) : Ceux-ci se reproduisent également, et, tout comme les éléphants matures, chacun d'eux donne naissance à 3 éléphanteaux dans cette période. On note d_n l'effectif de cette classe d'âge au bout de $30n$ années.

— Les éléphants *séniles* (au delà de 90 ans) : Ceux-là ne donnent naissance à aucun petit et disparaissent. On note s_n l'effectif de cette classe d'âge au bout de $30n$ années.

Le modèle est alors défini par le système suivant :

$$(S)\quad \left\{\begin{array}{rclr} (i_0, m_0, d_0, s_0) & = & (0,1,0,0) & \\ \forall n \in \mathbf{N} \quad i_{n+1} & = & 3m_n + 3d_n & (1) \\ m_{n+1} & = & i_n & (2) \\ d_{n+1} & = & m_n & (3) \\ s_{n+1} & = & d_n & (4) \end{array}\right.$$

A. Préliminaire : Étude d'une fonction polynomiale

Soit Q le polynôme donné par $Q = X^3 - 3X - 3$.

1. **a)** Dresser rapidement le tableau de variations de Q en indiquant les limites en $\pm\infty$.

b) En déduire que Q possède une unique racine réelle, notée r qu'on ne cherchera pas à calculer.

c) Vérifier que $r \in]2, 3[$.

d) Montrer les autres racines de Q sont deux racines non réelles conjuguées, notées z et $\overline{z}$. *On ne cherchera pas à calculer les valeurs de ces racines.*

2. **a)** Donner la forme factorisée de Q dans $\mathbf{C}$.

b) Calculer $Q(0)$ de deux façons différentes et en déduire que $r|z|^2 = 3$.

c) En déduire que $|z| < r$.

B. Simulation numérique du modèle

3. Exprimer, pour tout entier $n \geq 0$, l'effectif total t_n en fonction de i_n, m_n, d_n, et s_n.

4. Justifier brièvement que l'équation (1) du système (S) est en adéquation avec la description de Darwin.

5. Compléter sur la copie le script `Python` de la fonction suivante, prenant en entrée un entier n, et renvoyant en sortie le tuple (I, M, D, S) des listes des termes `[i0,..,in]`,..., `[s0,..,sn]` des suites calculés suivant le modèle (S) :

```
def effectifs(n):
    i,m,d,s = 0,1,0,0
    I,M,D,S = [i],[m],[d],[s]
    for k in range(n):
```

```python
        # Partie à compléter sur la copie
    return I,M,D,S
```

6. En déduire une fonction Python, `total(n)`, prenant en entrée un entier `n`, et renvoyant en sortie la valeur de t_n correspondante.

7. La commande `total(17)` renvoie `507627`. Confronter cette valeur produite par le modèle avec l'affirmation de Darwin, et commenter le modèle.

C. Approche matricielle

Cette partie a pour but d'estimer la croissance de l'effectif total t_n par des méthodes de diagonalisation.

On reprend le modèle (S), et on pose pour tout entier naturel n, $X_n = \begin{pmatrix} i_n \\ m_n \\ d_n \\ s_n \end{pmatrix}$.

8. **a)** Déterminer une matrice $A \in \mathscr{M}_4(\mathbf{C})$ indépendante de n telle que le modèle (S) s'écrive :

$$\forall n \in \mathbf{N} \quad X_{n+1} = AX_n.$$

b) Donner une relation simple entre X_n, A, et X_0, et valable à tout rang n (on ne demande pas de justification).

9. **a)** Montrer que 0 est valeur propre de A.

b) Soit λ un nombre complexe non nul. Montrer que λ est valeur propre de A si et seulement si λ est racine du polynôme Q défini dans la partie **A. Préliminaires**.

c) En déduire le spectre de A.

d) Conclure que A est diagonalisable dans $\mathscr{M}_4(\mathbf{C})$ et donner la dimension des sous-espaces propres de A (on ne cherchera pas de base de ces sous-espaces propres). Expliciter une matrice D diagonale semblable à A.

10. Montrer qu'il existe trois matrices M_1, M_2, M_3 de $\mathscr{M}_4(\mathbf{C})$ *indépendantes* de n, dont on ne cherchera pas à calculer les coefficients, et telles que :

$$\forall n \in \mathbf{N}^\star \quad A^n = r^n M_1 + z^n M_2 + \overline{z}^n M_3.$$

On pourra introduire une matrice P inversible indépendante de n, et pour tout entier $n > 0$, trois matrices diagonales $D_1(n)$, $D_2(n)$, et $D_3(n)$ judicieusement choisies telles que $P^{-1}A^nP = D_1(n) + D_2(n) + D_3(n)$.

11. **a)** En déduire qu'il existe une constante $K \geq 0$ indépendante de n telle que pour tout entier $n > 0$, on a la majoration $|t_n| \leq Kr^n$.
On pourra noter pour alléger la démonstration, pour toute colonne U de $\mathscr{M}_{4,1}(\mathbf{R})$, s_U la somme des modules de ses coefficients.

Un calcul explicite des matrices M_j de la question **10.** donnerait une valeur explicite de la constante K définie en **11. a)**, et permettrait même d'aboutir au fait que $t_n \underset{\infty}{\sim} Cr^n$ pour un réel $C > 0$.

b) Interpréter ce dernier résultat en termes d'évolution et d'extinction de la population.

II Modèle probabiliste de Galton-Watson (1874)

Le modèle probabiliste de Galton-Watson est un modèle général décrivant l'évolution d'une lignée issue d'un individu pour lequel on s'intéresse aux probabilités d'extinction de la lignée au bout d'un nombre de générations données. Les hypothèses du modèle sont les suivantes :

(H1). À chaque génération n ($n \in \mathbf{N}$), le nombre total d'éléphants présents est une variable aléatoire T_n à valeurs dans $\mathbf{N}$.

(H2). Si à une génération n donnée, on observe $[T_n = 0]$, alors les variables T_k sont supposées nulles à toute génération $k \geq n$.

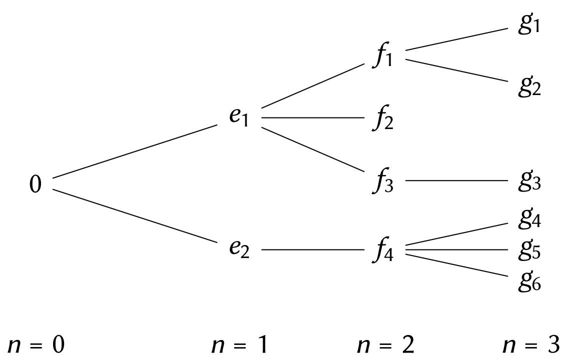
(H3). On part d'un unique éléphant à la génération $n = 0$. La loi de T_0 est donc une loi certaine de paramètre 1, c'est-à-dire que : $P(T_0 = 1) = 1$.

(H4). À toute génération n, les variables aléatoires qui comptent le nombre d'éléphanteaux issus d'un éléphant donné de cette génération sont des variables aléatoires mutuellement indépendantes et identiquement distribuées suivant une même loi appelée *loi de reproduction de modèle*, indépendante de n, et qui sera spécifiée au cas des éléphants en **partie F.** C'est en particulier la loi de la variable T_1.

D. Modèle de Galton-Watson et simulation

Voici une illustration du processus de Galton-Watson sur 4 générations :

partant d'un individu unique à la génération $n = 0$, on voit sur cet exemple qu'à la génération $n = 1$, il y a 2 individus e_1, e_2, c'est-à-dire que l'évènement $[T_1 = 2]$ est observé. Au bout d'une génération, l'individu e_1 a trois descendants, *etc.*



12. Quelles sont les réalisations des variables aléatoires T_2 et T_3 sur cet exemple ?

13. **a)** On suppose qu'on dispose d'une fonction `Python`, `repro()` qui simule une variable aléatoire X qui suit la loi de reproduction du modèle. Son appel renvoie donc un entier tiré suivant cette loi. On a en particulier :

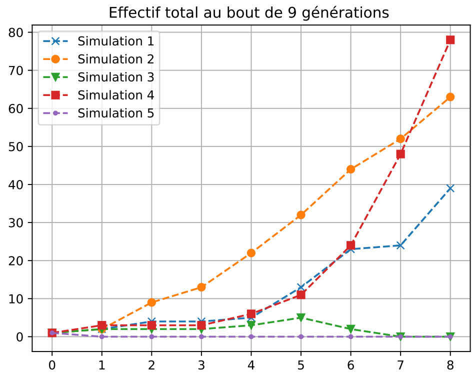
$$\forall k \in \mathbf{N} \quad T_{k+1} = \sum_{i=1}^{T_k} X_i,$$

les X_i étant des variables mutuellement indépendantes de même loi que X. Dans le script ci-après, la fonction `processus` prend en entrée un entier `n` et renvoie en sortie la liste `L` des valeurs observées des variables $T_0, \dots, T_n$. Compléter sur la copie les lignes 5 et 6 du script ci-dessous.

```
1 def processus(n):
2     L=[1]
3     for k in range(1,n+1):
4         T = 0
```

```
        for i in range(________): # compléter sur la copie
            X = ________          # idem
            T+=X
        L.append(T)
    return L
```

b) Le graphique suivant montre 5 réalisations de L, obtenues à partir d'une variable aléatoire X qui suit la loi de reproduction du modèle :



5 simulations de $(T_n)_{0\leq n\leq 8}$ sur 9 générations

Par rapport aux propriétés établies en **11.b)** du modèle déterministe de la partie **I**, quelle(s) différence(s) essentielle(s) constate-t-on avec le modèle de Galton-Watson ?

E. Étude de la suite des probabilités d'extinction à une génération donnée

Pour tout entier naturel n, on note $q_n = P(T_n = 0)$ la *probabilité d'extinction de la population à la génération n*. On appelle *fonction de reproduction* du modèle (voir **(H4)**) la fonction f définie par :

$$f(t) = \sum_{k=0}^{+\infty} P(T_1 = k)t^k. \tag{1}$$

14. Soit $t \in [0,1]$. Montrer que la série de terme général $P(T_1 = k)t^k$ converge.

Ce dernier résultat prouve que la fonction $t \mapsto f(t)$ est définie sur $[0,1]$.

15. Que vaut $f(0)$? Montrer que $q_1 = f(q_0)$.

16. Soit $\in \mathbf{N}$. On admet que pour tout entier naturel k : $P_{[T_1=k]}(T_{n+1} = 0) = q_n^k$.
En remarquant que si $[T_1 = k]$ est observé, le processus de Galton-Waston consiste en la juxtaposition de k processus de Galton-Watson indépendants de génération initiale 1, déduire que :

$$q_{n+1} = f(q_n).$$

17. Soit Q l'évènement : «La population d'éléphants s'éteint un jour». Le but de cette question est de relier la probabilité $P(Q)$ à la suite (q_n).
Pour alléger l'écriture, on pose, pour tout entier $k \in \mathbf{N}^\star$, $Q_k = [T_k = 0]$, si bien que

$$Q = \bigcup_{k=0}^{+\infty} Q_k.$$

On introduit pour tout entier naturel k les évènements Z_k définis par : «La population totale est d'effectif nul pour la première fois à la génération k».

a) Justifier que pour tout entier $k \geq 0$: $Q_k \subset Q_{k+1}$.

b) En déduire que la suite $(q_n)_{n\in\mathbf{N}}$ est convergente. On notera q sa limite, et on ne cherchera pas à la calculer.

c) Exprimer, pour tout entier $k \geq 1$, l'évènement Z_k en fonction de Q_k et Q_{k-1}.

d) Justifier que $Q = \displaystyle\bigcup_{k=1}^{+\infty} Z_k$ et en déduire que $P(Q) = q$.

F. Calcul de la probabilité d'extinction de la population

On se place dans le cas particulier où la loi de reproduction f du modèle, et définie par la relation (1) de la **partie E.** est une loi de Poisson de paramètre $\mu > 1$. On pose donc, pour tout réel t donné :

$$f(t) = e^{-\mu} \sum_{k=0}^{+\infty} \frac{\mu^k}{k!} t^k,$$

et on s'intéresse à la suite des probabilités d'extinction (q_n), qui vérifie, d'après la question **16.** :

$$q_0 = P(T_0 = 0) = 0 \quad \text{et} \quad \forall n \in \mathbf{N} \quad q_{n+1} = f(q_n).$$

On rappelle que cette suite est convergente vers un réel q d'après la question **17. b)**. On souhaite prouver dans ce qui suit que la probabilité d'extinction $P(Q)$, qui vaut q d'après **17. d)**, vérifie $q \in]0, 1[$.

18. Prouver que $q \geq q_1$. L'évènement Q de la question **17.** est-il quasi-impossible ?

19. Vérifier que :

$$\forall t \in \mathbf{R} \quad f(t) = e^{\mu(t-1)}.$$

20. Justifier que le réel q vérifie $f(q) = q$.

21. On pose pour tout réel t dans $]0, 1]$: $\varphi(t) = \mu(t-1) - \ln t$, et $g(t) = f(t) - t$.

a) Calculer $f(1)$.

b) Montrer que : $\forall t \in]0, 1] \quad g(t) = 0 \Leftrightarrow \varphi(t) = 0$.

c) On admet que pour tout réel $u > 1$, $\ln u < u - 1$. Déduire que dans l'intervalle $]0, 1[$, l'équation $f(x) = x$ possède une solution unique notée ρ, et que $\rho < \dfrac{1}{\mu}$.

22. **a)** Montrer par récurrence que : $\forall n \in \mathbf{N} \quad q_n \leq \rho$.

b) Déduire de ce qui précède la valeur de la limite q de la suite (q_n).

c) Expliquer pourquoi dans le problème de Darwin, il est cohérent de choisir pour la fonction f la valeur $\mu = 6$.

d) L'évènement Q de la question **17.** est-il quasi-certain ? Donner deux réels $a > 0$, et $b < 1$ tels que $q \in [a, b]$.