Un sujet pour l'épreuve B
(modélisation et informatique)

Présentation

Le texte proposé ci-après est conçu pour l'épreuve B, portant plus particulièrement sur la modélisation et l'informatique (l'épreuve B dure 3 heures 30).
Les compétences mobilisées dans ce problème sont essentiellement les suivantes :
▷ Engager une recherche, définir une stratégie : questions III.3.c, IV.2, IV.3, IV.4.
▷ Modéliser un phénomène à l'aide du langage mathématique : question I.4, IV.2.b, IV.3.
▷ Représenter, changer de registre : questions IV.1.c, IV.4.
▷ Traduire un algorithme dans un langage de programmation : questions III.2.b, III.2.d, III.3.b.
▷ Mobiliser des connaissances scientifiques pertinentes : questions II.2.a, III.2.a.
▷ Identifier un problème sous différents aspects : questions I.3.d, IV.4.e.
▷ Critiquer ou valider un modèle ou un résultat : questions I.1, I.3, IV.1.
▷ Communiquer à l'écrit et à l'oral : compétence présente, par nature, dans l'ensemble des questions.

On rappelle que l'emploi d'une calculatrice est autorisé dans cette épreuve.

Énoncé

Contexte et but de l'étude

Dans ce sujet, on veut analyser mathématiquement et interpréter certains modèles de dynamique des populations. Dans tout le problème, on notera $N(t)$ le nombre d'individus d'une espèce donnée présents à l'instant $t \in \mathbf{R}^+$. La répartition spatiale des individus ne sera pas considérée ici. Les modèles considérés seront des équations différentielles ordinaires de la forme

$$\begin{cases} \frac{dN}{dt}(t) = f(N(t)), \qquad t \in \mathbf{R}^{+*}, \\ N(t=0) = N_0, \end{cases} \tag{A}$$

où $N_0 \geqslant 0$ représente le nombre d'individus à l'instant initial et $f \in \mathscr{C}^\infty(\mathbf{R})$ une fonction de réaction qui sera spécifique à chaque modèle dans la suite.
L'expression de la fonction f devra contenir certaines caractéristiques biologiques naturelles pour une population. L'énoncé du problème proposera différents modèles et il sera demandé au candidat de commenter qualitativement ces choix au regard des impératifs biologiques. De même, dès lors que certains résultats seront obtenus, le candidat devra les confronter aux possibilités biologiques.
On veut, étant donné une fonction f, étudier le comportement en temps long de la fonction N. On sera alors amené à étudier des solutions particulières appelées états stationnaires, ainsi que leur stabilité. On donne pour cela les définitions suivantes :

Definition 1 (**État stationnaire**). *On appelle **état stationnaire** toute solution constante de* (A).

Definition 2 (**Stabilité d'un état stationnaire**). *On dit qu'un état stationnaire N_∞ de* (A) *est **stable** lorsque toute solution de* (A) *ayant pour condition initiale un N_0 proche de N_∞ tend vers N_∞ lorsque t tend vers l'infini. Un état stationnaire qui n'est pas stable au sens de cette définition sera dit **instable**.*

La partie I étudie un premier modèle important, le modèle de Verhulst, qui permet d'introduire par l'exemple la partie suivante. La partie II permet de caractériser les états stationnaires et leur stabilité. On attire l'attention du candidat sur le fait que les conclusions de cette partie seront utilisées dans la partie IV. La troisième partie consiste en une étude informatique du modèle de la partie I, à partir de données expérimentales. Il est signalé au candidat que la partie III peut être traitée indépendamment des autres parties du problème. Dans la quatrième et dernière partie, on étudie plus particulièrement un modèle de dynamique des populations ou l'on prend en compte un effet de prédation. Le comportement des états stationnaires en fonction de paramètres du modèle permet d'illustrer l'intérêt de la prédation dans la prévention d'invasions par des parasites.

I Le modèle logistique de Verhulst

Soit $K \in \mathbf{R}^+$ et $N_0 \in \mathbf{R}^+$. On considère maintenant le modèle suivant :

$$\begin{cases} \frac{dN}{dt}(t) = rN(t)\left(1 - \frac{N(t)}{K}\right), \qquad t \in \mathbf{R}^{+*}, \\ N(t=0) = N_0. \end{cases} \tag{V}$$

On admet que la solution de (V) existe en tant que fonction définie sur $\mathbf{R}_+$, et qu'elle est unique.

1. *Justification du modèle*
   (a) Identifier la fonction f. Que représente la fonction ρ définie par $\rho(N) = r\left(1 - \frac{N}{K}\right)$?
   (b) Préciser le signe de $\rho(N)$ en fonction des valeurs de N.
   (c) La constante K est appelée « capacité d'accueil ». Commenter cette appellation.
2. *Résolution explicite*
   (a) Montrer que 0 et K (en tant que fonctions constantes) sont les états stationnaires de (V).
   (b) Montrer que si $N_0 > 0$ alors la solution de (V) peut s'écrire
   $$\forall t \in \mathbf{R}^+, \qquad N(t) = \frac{K}{1 + \left(\frac{K}{N_0} - 1\right)e^{-rt}}.$$
   (c) Vérifier que $N(t) > 0$ dès lors que $N_0 > 0$.
3. *Étude qualitative*
   (a) Soit $N_0 > 0$. Donner la limite de $N(t)$ lorsque $t \to +\infty$. Commenter qualitativement. Converge-t-on vers un état stationnaire de (V) ?
   (b) Donner le sens de variation de $N(t)$ en fonction de la valeur de N_0. Commenter.
   (c) Montrer que lorsque K tend vers l'infini (N_0 étant constant), on a
   $$N(t) \underset{K \to +\infty}{\sim} N_0 e^{rt},$$
   à $t > 0$ fixé. Commenter.
   (d) En figure 1 (voir à la fin de ce document), on a tracé deux courbes $\mathscr{C}_1$ et $\mathscr{C}_2$ représentatives de solutions de (V). Pour chacune des deux courbes, identifier les paramètres r, K, N_0 conduisant à ces solutions. On détaillera la méthode suivie.
4. *Vers la stabilité des états stationnaires*
   (a) À l'issue de la question 3, que peut-on dire de la stabilité des deux états stationnaires ? Donner les signes des nombres dérivés $f'(0)$ et $f'(K)$.
   (b) Établir une conjecture entre la stabilité d'un état stationnaire N_s et le signe du nombre dérivé $f'(N_s)$.

II Dynamique en temps long et stabilité des états stationnaires

On revient, dans cette partie, au cas d'une fonction f quelconque dans l'équation (A).

1. *Convergence vers un état stationnaire*
   Montrer que N_∞ est un état stationnaire de (A) si et seulement si $f(N_\infty) = 0$.

On admettra alors la proposition suivante, qui pourra être utilisée sans démonstration, mais en y faisant explicitement référence :

Proposition 3 (Convergence vers un état stationnaire). *On suppose que N est une solution de* (A) *et que $N(t)$ converge lorsque t tend vers $+\infty$. On appelle N_∞ la limite de N. Alors N_∞ est un état stationnaire de* (A).

2. *Stabilité des états stationnaires*
   Dans cette question, on souhaite montrer le critère de stabilité suivant, que l'on pourra utiliser par la suite :

   Proposition 4 (Critère de stabilité). *Soit N_∞ un état stationnaire de* (A) *tel que $f'(N_\infty) < 0$. Alors N_∞ est un état stationnaire stable (au sens de la définition 2).*

   (a) Justifier qu'il existe $a > 0$ et $r > 0$ tels que

   $$\forall u \in [N_\infty - r, N_\infty + r], \qquad f'(u) \leqslant -a < 0.$$

   On suppose que $N(t)$ est une solution de (A) avec pour condition initiale N_0 tel que $|N_0 - N_\infty| \leqslant r$. On admettra qu'alors pour tout $t \in \mathbf{R}^+$, $|N(t) - N_\infty| \leqslant r$.
   Enfin, on définit la fonction g par $g(t) = (N(t) - N_\infty)^2$.

   (b) On fixe $t \in \mathbf{R}^+$. On définit alors pour $s \in [0,1]$, $G(s) = N_\infty + s(N(t) - N_\infty)$.
      (i) Calculer $G'(s)$, pour $s \in [0,1]$.
      (ii) Montrer qu'alors

      $$f(N(t)) = f(N_\infty) + \int_0^1 f'(G(s))\,(N(t) - N_\infty)\,\mathrm{d}s.$$

      En déduire alors que $N'(t) = \left(\int_0^1 f'(G(s))\,\mathrm{d}s\right)(N(t) - N_\infty)$.
      (iii) Montrer que

      $$\forall s \in [0,1], \quad G(s) \in [N_\infty - r, N_\infty + r].$$

      En déduire alors que $\forall t \in \mathbf{R}^+$, $g'(t) \leqslant -2ag(t)$.

   (c) Montrer alors que $\forall t \in \mathbf{R}^+, \quad g(t) \leqslant g(0)e^{-2at}$.

   (d) En déduire que $\lim_{t\to\infty} N(t) = N_\infty$. Conclure.

   On admettra de même que si N_∞ un état stationnaire de (A) tel que $f'(N_\infty) > 0$ alors N_∞ est un état stationnaire instable.

III Étude informatique du modèle logistique de Verhulst

Les estimations et statistiques concernant la population des Etats-Unis de 1790 à 1950 donnent les chiffres suivants (exprimés en millions d'individus) :

dates t_i	1790	1800	1810	1820	1830	1840	1850	1860	1870
population y_i	3,929	5,308	7,240	9,638	12,866	17,069	23,192	31,443	38,558

dates t_i	1880	1890	1900	1910	1920	1930	1940	1920
population y_i	50,156	62,498	75,995	91,972	105,711	122,775	131,669	150,697

Nous noterons dans la suite : $n = 17$. Nous supposons que cette population a suivi une loi logistique entre 1790 et 1920. Pour tout temps t compris entre 1790 et 1920, la valeur théorique $y(t)$ (exprimée en millions d'individus) du nombre d'habitants des USA l'instant t est donc donnée par la formule :

$$\forall t \in [1790, 1920], \qquad y(t) = \frac{K}{1 + C\exp(-r\,t)}$$

où K, C et r sont des constantes réelles strictement positives. Le but de cette partie est de déterminer ces constantes à l'aide des données expérimentales ci-dessus.

1. Montrer, en justifiant les différentes étapes que :

$$\forall t \in [1790, 1950], -r\,t + \ln(C) = \ln\left(\frac{K - y(t)}{y(t)}\right).$$

   Dans la suite, nous noterons : $\begin{cases} \alpha = -r \\ \beta = \ln(C) \end{cases}$.

2. Dans cette question, nous supposons connue la population limite K.

   Nous allons déterminer les constantes α et β à l'aide de la méthode des moindres carrés. Plus précisément, nous allons calculer les valeurs de α et β telles les écarts entre les valeurs mesurées $z_i = \ln\left(\frac{K - y_i}{y_i}\right)$ et les valeurs théoriques $\ln\left(\frac{K - y(t_i)}{y(t_i)}\right)$ aux mêmes instants t_i, $i \in [\![0, n-1]\!]$ soient les plus petits possibles.
   Nous allons donc déterminer les valeurs α et β telles que la quantité

$$E(\alpha, \beta) = \sum_{i=0}^{n-1} (\alpha\,t_i + \beta - z_i)^2$$

   soit la plus petite possible.

   (a) Donner les formules de α^* et β^* qui minimisent la quantité $E(\alpha, \beta)$, pour $(\alpha, \beta) \in \mathbf{R}^2$.

   (b) Les données sont stockées dans un tableau P à deux entrées, la première ligne contenant les dates, la seconde contenant le nombre d'habitants (en millions d'unités), les colonnes sont indexées de 0 à $n-1$. Par exemple, $P[0,4] = 1830$, $P[1,9] = 50{,}156$ est la population des USA en millions d'habitants en 1880. Écrire un algorithme qui calcule les valeurs de α^* et β^*. L'algorithme sera rédigé au moyen d'une transcription en langage Python.

   (c) Donner les valeurs numériques de α^* et β^* pour $K = 195$. En déduire alors les valeurs de r et C.

   (d) Dans la suite, nous noterons : $\mathrm{Err}(\ell) = E(\alpha^*, \beta^*)$, c'est une évaluation de l'erreur globale faite en fonction de la population limite ℓ. Écrire un algorithme qui P et ℓ étant donnés, calcule la valeur de $\mathrm{Err}(\ell)$.

3. Malheureusement, dans la plupart des situations, nous ne connaissons pas la population limite K. Ici, nous savons seulement qu'elle est supérieure à $150{,}697$ millions d'habitants. Notre idée pour déterminer la population limite va être de minimiser l'erreur globale faite en fonction de K que nous avons notée $\mathrm{Err}(K)$, et nous adopterons alors comme population limite, la valeur de K qui minimise $\mathrm{Err}(K)$.
   Pour cela, nous allons partir d'une valeur arbitraire de K_0 ($K_0 = 160$ serait une valeur plausible), et calculer successivement $K_j = K_0 + \frac{j}{10}$ et $\mathrm{Err}(K_j)$, $j \in \mathbf{N}$, et nous stopperons les itérations dès que :

$$\mathrm{Err}(K_j) < \mathrm{Err}(K_{j+1}).$$

   Nous estimons alors que la population limite K est comprise entre K_j et K_{j+1} , nous pourrons améliorer notre approximation par dichotomie.

   (a) Le critère d'arrêt suffit-il à justifier que le minimum de la fonction Err est compris entre K_j et K_{j+1} ?

(b) Écrire un algorithme qui détermine la plus petite valeur de j telle que : $\mathrm{Err}(K_j) < \mathrm{Err}(K_{j+1})$ et renvoie K_j.

(c) La valeur de K_j étant déterminée, mettre en place un algorithme permettant d'améliorer notre approximation de la population limite K avec une précision ϵ choisie à l'avance. On pourra notamment considérer l'algorithme suivant :

Algorithme de dichotomie

On considère une fonction f continue sur un segment $[a, b]$; les nombres a et b sont les données initiales de l'algorithme.

On suppose que f s'annule exactement une fois sur $[a, b]$, en un point que l'on note α.

On définit les suites $(a_k)_{k\geqslant 0}$ et $(b_k)_{k\geqslant 0}$ de la façon suivante :

▷ $a_0 = a$ et $b_0 = b$.

▷ Pour tout entier naturel k, on note $c_k = \dfrac{a_k + b_k}{2}$ et :

$$\text{si} \quad f(a_k)f(c_k) \leqslant 0, \qquad \begin{array}{lll} \text{alors} & a_{k+1} = a_k & \text{et} \quad b_{k+1} = c_k \\ \text{sinon} & a_{k+1} = c_k & \text{et} \quad b_{k+1} = b_k \end{array}$$

On sait alors que les suites (a_k) et (b_k) convergent toutes les deux vers α, en vérifiant :

$$\forall\, k \in \mathbf{N},\ a_k \leqslant \alpha \leqslant b_k \qquad \text{et} \qquad \forall\, k \in \mathbf{N},\ b_k - a_k = \frac{b-a}{2^k}$$

On peut alors montrer que si l'entier k est tel que $\dfrac{b-a}{2^k} \leqslant \varepsilon$, alors a_k et b_k sont des valeurs approchées à ε-près de α. Le nombre $\frac{a_k+b_k}{2}$ est la valeur retournée par l'algorithme.

IV Étude des états stationnaires pour un modèle d'évolution avec prédation.

La tordeuse des bourgeons de l'épinette, *Choristoneura fumiferana* (Clemens), est l'insecte le plus destructeur des peuplements de conifères de l'Amérique du Nord. On trouve ce défoliateur indigène dans toutes les provinces canadiennes, de la Colombie-Britannique à Terre-Neuve. Au Québec, la tordeuse des bourgeons de l'épinette consomme principalement le feuillage annuel du sapin baumier, de l'épinette blanche ainsi que, à un degré moindre, de l'épinette rouge et de l'épinette noire. Vers la fin d'avril ou le début de mai, les jeunes chenilles sortent de leur hibernation. Attirées par la lumière, elles se dirigent vers les extrémités des branches, s'y installent et s'y nourrissent jusqu'à leur sixième et dernier stade larvaire, soit jusqu'à la fin de juin. C'est à ce moment que leurs dégâts sont les plus apparents. La plupart du temps, le nombre de chenilles reste à un niveau assez faible. Néanmoins, régulièrement leur nombre explose et provoque une large épidémie dans les forêts canadiennes.

Dans cette partie, on considère une population N de *Choristoneura fumiferana* installée dans les arbres d'une forêt et subissant une prédation de la part d'oiseaux environnants. On souhaite étudier comment la prédation permet de contenir ou non la population de chenilles. On décrit l'évolution de N de la manière suivante

$$\begin{cases} \dfrac{dN}{dt}(t) = rN(t)\left(1 - \dfrac{N(t)}{K}\right) - \dfrac{AN(t)^2}{B^2 + N(t)^2}, & t \in \mathbf{R}^{+*}, \\ N(t=0) = N_0, \end{cases} \tag{IV.1}$$

où r, K, A, B, N_0 sont des paramètres réels positifs.

1. *Commentaires sur le modèle*

(a) Identifier la fonction f.

(b) On souhaite mettre la fonction f sous la forme $f(N) = c(N) + p(N)$, où la partie $c(N)$ représente la croissance propre de la population d'insectes, et le terme $p(N)$ décrit la prédation par les oiseaux. Donner l'expression de $c(N)$ et de $p(N)$ en fonction de N, en expliquant ce choix.

(c) Tracer l'allure de $p(N)$ sur $[0,10]$, pour $A = 2$ et $B = 1$. Justifier qualitativement la pertinence du choix d'une telle fonction.

2. *Modèle adimensionné*

   (a) Donner deux réels α et β tels que la fonction u telle que $u(t) = \alpha N(\beta t)$ vérifie l'équation différentielle

$$\begin{cases} \dfrac{du}{dt}(t) = \mu u(t)\left(1 - \dfrac{u(t)}{q}\right) - \dfrac{u(t)^2}{1+u(t)^2}, & t \in \mathbf{R}^{+*}, \\ u(t=0) = u_0, \end{cases} \tag{IV.2}$$

   pour deux paramètres μ et q que l'on précisera. On explicitera la valeur de u_0.

   (b) Comment peut-on interpréter une augmentation de A, les autres paramètres étant inchangés ? On examinera ce que deviennent μ et q, et on interprètera pour les populations observées.

3. *Un cas particulier : Absence de saturation par manque de ressources*
   Dans cette question, on suppose que la capacité d'accueil K est infinie.

   (a) Montrer qu'alors le modèle (IV.2) se réduit en

$$\begin{cases} \dfrac{du}{dt}(t) = \mu u(t) - \dfrac{u(t)^2}{1+u(t)^2}, & t \in \mathbf{R}^{+*}, \\ u(t=0) = u_0, \end{cases} \tag{IV.3}$$

   (b) Montrer que $u = 0$ est un état stationnaire instable de (IV.3).

   (c) Montrer que (IV.3) a deux états stationnaires non nuls positifs et distincts si et seulement si $\mu < \dfrac{1}{2}$. On les notera $U_1(\mu) \leqslant U_2(\mu)$ et on précisera leur expressions en fonction de μ.

   (d) Donner un équivalent de $U_2(\mu)$ lorsque $\mu \to 0$.

   (e) Démontrer que :

$$\forall t \in \mathbf{R}+, \qquad \frac{du}{dt}(t) > \left(\mu - \frac{1}{2}\right) u(t).$$

   (f) On suppose que $\mu > \dfrac{1}{2}$. Quelle est alors la limite de $u(t)$ lorsque $t \to \infty$? Commenter qualitativement l'efficacité de la prédation dans ce cas.

   (g) On se place dorénavant dans le cas $\mu \leqslant \dfrac{1}{2}$. Étudier la stabilité des états stationnaires non nuls.

   (h) On admet le résultat suivant :

   si $u_0 > U_2(\mu)$, alors nécessairement $u(t) > U_2(\mu)$ pour tout $t > 0$.

   En déduire que u croit vers $+\infty$. Commenter qualitativement.

   (i) On admet maintenant le résultat suivant :
   ▷ si $u_0 < U_1(\mu)$, alors nécessairement $u(t) < U_1(\mu)$ pour tout $t > 0$.
   ▷ si $U_1(\mu) < u_0 < U_2(\mu)$, alors nécessairement $U_1(\mu) < u(t) < U_2(\mu)$ pour tout $t > 0$.
   On suppose maintenant que $u_0 < U_2(\mu)$. Quelle est la limite de u lorsque $t \to +\infty$? La prédation a-t-elle un intérêt dans ce cas ? On précisera la monotonie de $u(t)$ en fonction de u_0.

4. *Recherche des états stationnaires* On revient au cas où K est fini.

   (a) Montrer que $U = 0$ est un état stationnaire de (IV.2).

   (b) Montrer que U est un état stationnaire *non nul* de (IV.2) si et seulement si

$$\mu\left(1 - \frac{U}{q}\right) = \frac{U}{1+U^2}. \tag{IV.4}$$

   (c) Justifier que l'équation (IV.4) a 1 ou 3 solutions réelles selon les paramètres μ et q.
   Attention, on ne demande pas d'expliciter les paramètres correspondant à telle ou telle situation.

(d) En traçant sur un même graphique la fonction $x \mapsto \dfrac{x}{1+x^2}$ ainsi que la fonction $x \mapsto \mu\left(1-\dfrac{x}{q}\right)$ pour deux couples de paramètres (q, μ) convenablement choisis, illustrer graphiquement les deux possibilités de la question précédente.

(e) Dans le plan (q, μ), on a tracé en figure 2 les zones correspondant à 1 ou 3 états stationnaires non nuls. Identifier, en le justifiant, les deux zones.

(f) On appelle $\mathscr{C}$ la courbe frontière entre les deux zones (voir figure 2). Donner l'équation des deux asymptotes à $\mathscr{C}$ lorsque $q \to \infty$.

5. *Étude de la stabilité des états stationnaires*

(a) Montrer que l'état stationnaire 0 est toujours instable.

(b) On se place dans un cas ou il existe trois états stationnaires non nuls distincts. On les note $U_1 < U_2 < U_3$. Déterminer lesquels sont stables et lesquels sont instables.

Indication : On pourra écrire $f(u) = -\dfrac{\mu}{q}\dfrac{u}{1+u^2}(u-U_1)(u-U_2)(u-U_3)$ *pour calculer simplement les nombres dérivés nécessaires.*

Figures

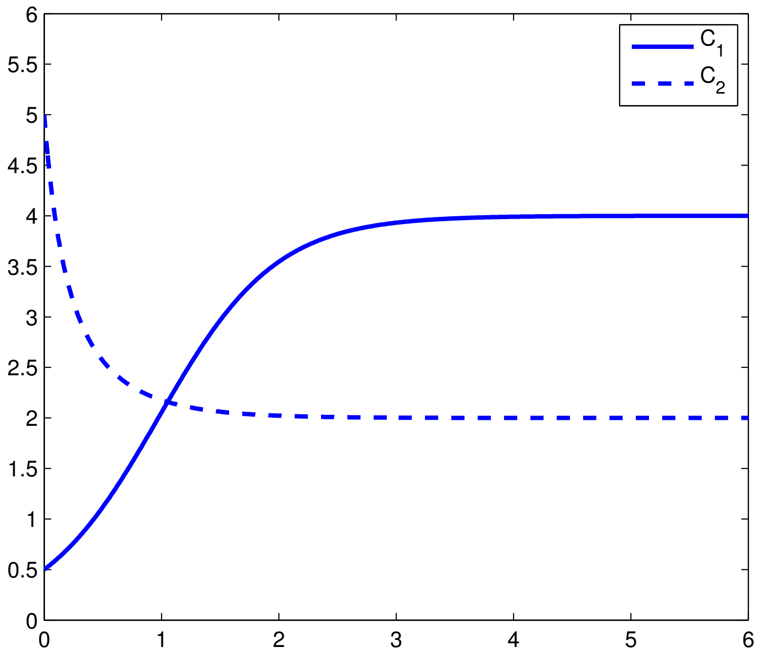

FIGURE 1 – Courbes $\mathscr{C}_1$ (en trait plein) et $\mathscr{C}_2$ (en trait pointillé) donnant deux solutions de (V) pour $t \in [0;6]$.

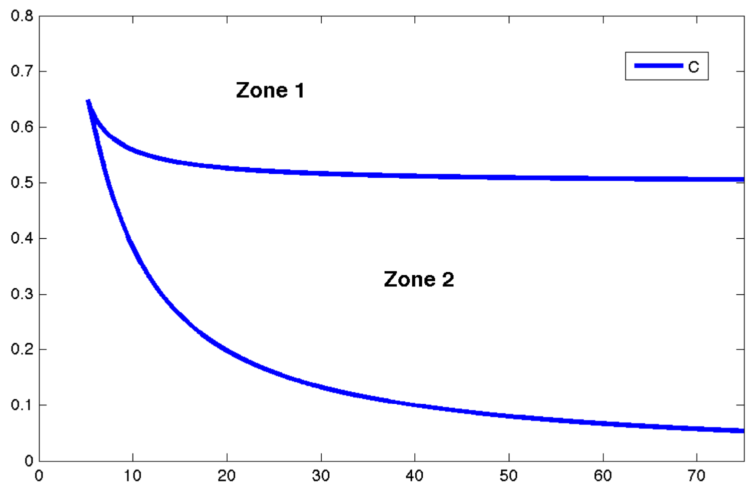

FIGURE 2 – Dans le plan (q, μ), nombre d'états stationnaires de (IV.2).