Avant-propos

Cet ouvrage collectif est bâti autour de textes proposés à l'épreuve d'oral "Modélisation" du concours d'agrégation externe de mathématiques, session 2001. Ces textes présentent des exemples de modélisation mathématique, suffisamment simples pour pouvoir être compris et exploités par les candidats en quatre heures. Conçus par les membres du jury, il offrent de nombreuses opportunités pour illustrer les mathématiques au programme du concours dans une démarche de modélisation mathématique appliquée. Pour le lecteur qui, disposant de temps et de recul, voudra aller plus au fond de l'étude proposée, ces textes constitueront souvent une introduction à des domaines actifs de la recherche.

L'introduction en 1999 d'une épreuve orale de Modélisation a constitué une réforme importante du programme du concours, qui doit beaucoup à la clairvoyance et à la détermination de C. Ruget, Présidente du jury. Depuis, le jury constate que les candidats sont de mieux en mieux préparés année après année, ce qui témoigne également du travail accompli dans les préparations universitaires.

A travers les programmes et les modalités des épreuves de concours, le jury cherche à stimuler les candidats, et les inciter à développer les qualités requises pour enseigner les mathématiques dans les classes et les universités. L'épreuve de modélisation vise donc à approfondir le lien entre les mathématiques, souvent abstraites, et leur utilisation dans les autres disciplines théoriques ou appliquées. La puissance des outils mathématiques permet des contributions essentielles dans leurs grandes applications qui continuent à se diversifier. La démarche de modélisation, quant à elle, constitue une médiation entre le domaine de l'application et celui des mathématiques. Le spécialiste de l'application y vient par un effort de créativité et d'abstraction, et conserve toute la responsabilité de l'interprétation critique des résultats. Le mathématicien se trouve confronté très souvent à des problèmes ouverts, voire à des questions pour lesquelles il faut construire un cadre mathématique adapté.

Nous espérons donc que nos candidats, dont l'effort pour comprendre la modélisation mathématique aura été soutenu par ceux de leurs préparateurs, se verront récompensés par une perception plus complète de la place des mathématiques dans les domaines scientifiques et techniques. Ils pourront également y trouver des satisfactions professionnelles, qu'ils décident finalement d'enseigner ou de mettre en pratique dans la recherche appliquée.

Pour aider plus encore les candidats, leurs préparateurs et les enseignants qui souhaitent enrichir leur cours en développant des exemples de modélisation mathématique, cet ouvrage a un format différent de celui publié l'an passé

sous la direction de C.Ruget. En effet, à coté d'une sélection de textes issus du concours 2001, on trouve des rubriques :

- **commentaire de l'auteur** : Donne plus de détails sur le contexte dont le texte est extrait, les prolongements appliqués, les méthodes mathématiques permettant l'étude à divers niveaux. L'auteur y explicite ses attentes et les possibilités d'exploitation qu'il a ménagées dans le texte à l'intention des candidats. Le plus souvent, les futurs candidats verront que des développements abordables sont possibles en complétant les preuves et développements présentés dans le texte. L'explication de la démarche de modélisation les conduira à se poser de multiples questions. Ils comprendront aussi que les textes ont été conçus non pas pour les limiter à une exposition linéaire, mais bien pour fournir la matière à un exposé original.
- **exploitation dans le cadre de l'épreuve** : Ecrits par des membres du jury qui ne sont pas l'auteur, il s'agit d'exemples et de propositions d'exploitation du texte dans le cadre d'un exposé. Certains explorent toute la latitude que les modalités de l'épreuve offrent au candidat de sélectionner et de compléter les éléments fournis. Ces propositions ont été écrites tantôt par des spécialistes du domaine, tantôt par des non spécialistes qui siègent le plus souvent dans les épreuves d'analyse ou d'algèbre. Il ne s'agit pas de figure de style, l'épreuve de modélisation touche un débat de fond sur la place des mathématiques qui aura concerné tout le jury 2001.
- **exploitation dans un cadre TIPE/TPE** : Ecrit par des membres du jury enseignant en classe préparatoire, il s'agit d'indications pour exploiter le texte dans le cadre des TPE et TIPE. Dans ce cas, il faut d'une part simplifier et rendre les aspects techniques compréhensibles aux élèves visés, d'autrepart profiter de la durée du projet pour leur permettre d'approfondir par une recherche documentaire originale. On conçoit donc ici le texte comme un support pour l'équipe enseignante qui devra proposer une démarche active convenant aux élèves.

Je tiens à remercier tous les membres du jury qui ont animé l'épreuve de modélisation depuis sa mise en place en 1999, conçu des textes, et enfin ceux qui ont participé à cet ouvrage collectif. Ma gratitude va également à C. Ruget, qui a courageusement donné sa place à la modélisation dans le concours d'agrégation externe. Notre effort vise avant tout à permettre au plus grand nombre de s'intéresser à l'application des mathématiques et à leur apport dans le monde moderne. Les enseignants que nous formons auront un rôle essentiel et les initiatives prises par les pouvoirs publics avec les TIPE et les TPE en fournissent le cadre.

Le travail nécessaire pour la préparation de cet ouvrage, qui comporte aussi bien l'organisation de la relecture croisée des contributions que l'intégration des contributions en TEX et l'harmonisation de leur présentation a été réalisé par Alain Chillès. Qu'il trouve ici l'expression de toute ma reconnaissance.

Enfin je voudrais remercier Jean-Michel Ghidaglia, qui a accepté d'accueillir ce livre dans la collection SCOPOS, aussi bien pour son aide à faire aboutir ce travail dans un délai raisonnable, que pour son engagement à mettre en valeur les mathématiques et leurs applications.

Paris le 7 Novembre 2001

Première partie

Localisation par GPS

Chapitre 1
Localisation par GPS

§1. Introduction

Le GPS (Global Positioning System) est une application particulièrement spectaculaire des progrès réalisés récemment dans le domaine des télécommunications.

Ce système de repérage par satellite a été développé à l'origine pour des applications militaires par le DoD américain. Depuis Mai 2000 le brouillage sélectif a été supprimé, ouvrant la voie à de très nombreuses applications : localisation de véhicules, guidage, mesure de dérive des continents etc.

Le GPS est constitué d'un ensemble de 24 satellites placés sur des orbites situées à environ 28 000 km du centre de la terre. Les orbites des satellites GPS sont connues à chaque instant avec une précision extrême. De plus les horloges atomiques internes de chaque satellites sont parfaitement synchronisées entre elles.

Le principe de localisation d'un récepteur GPS est le suivant : on peut mesurer le temps τ_j mis par un signal émis par le satellite j pour parvenir jusqu'au récepteur GPS. On en déduit la distance d_j à laquelle se trouve le satellite (le signal se propage à la vitesse de la lumière c, constante connue, et $d_j = c\tau_j$). Le récepteur doit donc se trouver sur une sphère centrée sur le satellite j et de rayon d_j. En principe, il suffirait de connaitre la position de trois satellites pour obtenir (au moins génériquement) deux positions possibles du récepteur (intersection de trois sphères) dont l'une serait a priori aberrante, d'où la position du récepteur. Le problème vient du fait que la mesure de la distance d_j suppose que le récepteur soit parfaitement synchronisé avec l'horloge des satellites, ce qui n'est pas le cas. On mesure en fait des «pseudo-distances» ψ_j qui sont :

$$\psi_j = d_j + c\Delta t$$

où Δt est le décalage entre l'horloge du récepteur et l'horloge des satellites GPS, quantité inconnue. En se plaçant dans un repère ECEF («Earth centered-Earth Fixed») et en négligeant tous les bruits de mesures, le satellite j fournit l'équation :

$$\sqrt{(X_j - x)^2 + (Y_j - y)^2 + (Z_j - z)^2} + c\Delta t = \hat{\psi}_j \tag{1}$$

Dans cette équation les inconnues sont (x, y, z), coordonnées du récepteur, et Δt (ou de façon équivalente, (x, y, z) et $w = c\Delta t$) tandis que les valeurs (X_j, Y_j, Z_j), coordonnées du satellite, et $\hat{\psi}_j$ sont des mesures calculées à partir de données transmises par le satellite j. Dans la suite on notera ξ le vecteur des inconnues (appelé aussi l'état du système) :

$$\xi = (x, y, z, w)$$

Il est donc nécessaire que le récepteur puisse capter les signaux d'au moins quatre satellites. Dans la suite on notera X^T le transposé du vecteur (ou de la matrice) X.

§2. Le problème statique

Le premier problème à résoudre est celui de trouver la position du récepteur GPS. Pour cela on doit résoudre le système d'équations non linéaires (1) pour $j = 1, \ldots, N$, $N \geq 4$. On prendra ici $N = 4$, ce qui est aussi fait en pratique, le récepteur GPS sélectionnant les 4 satellites fournissant les meilleures données.

En posant

$$\psi_j(\xi) = \sqrt{(X_j - x)^2 + (Y_j - y)^2 + (Z_j - z)^2} + c\Delta t \tag{2}$$

l'équation (1) s'écrit $\psi(\xi) = \hat{\psi}$. Pour trouver ξ, on peut utiliser une méthode de Newton basée sur une linéarisation des équations (1). On peut supposer qu'on dispose d'une bonne approximation (x_0, y_0, z_0) de la vraie position (x, y, z) du récepteur GPS et on peut prendre $\Delta t_0 = w_0 = 0$ comme point de départ ξ_0 de l'algorithme de Newton.

Soit $H(\xi)$ la matrice $N \times 4$ dérivée de ψ au point ξ. La j-*ème* ligne de $H(\xi)$ vaut

$$\begin{aligned} H_j(\xi) &= \begin{bmatrix} \dfrac{\partial \psi_j}{\partial x}(\xi) & \dfrac{\partial \psi_j}{\partial y}(\xi) & \dfrac{\partial \psi_j}{\partial z}(\xi) & \dfrac{\partial \psi_j}{\partial w}(\xi) \end{bmatrix} \\ &= \begin{bmatrix} -\dfrac{X_j - x}{d_j(\xi)} & -\dfrac{Y_j - y}{d_j(\xi)} & -\dfrac{Z_j - z}{d_j(\xi)} & 1 \end{bmatrix} \end{aligned}$$

où $d_j(\xi) = \sqrt{(X_j - x)^2 + (Y_j - y)^2 + (Z_j - z)^2}$.

Pour la première itération de Newton on doit résoudre le système linéaire :

$$H(\xi_0)(\xi_1 - \xi_0) = \hat{\psi} - \psi(\xi_0)$$

et de manière générale, pour $N = 4$, en désignant par ξ l'estimation courante, on doit résoudre :

$$\begin{bmatrix} \dfrac{\partial \psi_1}{\partial x}(\xi) & \dfrac{\partial \psi_1}{\partial y}(\xi) & \dfrac{\partial \psi_1}{\partial z}(\xi) & 1 \\ \dfrac{\partial \psi_2}{\partial x}(\xi) & \dfrac{\partial \psi_2}{\partial y}(\xi) & \dfrac{\partial \psi_2}{\partial z}(\xi) & 1 \\ \dfrac{\partial \psi_3}{\partial x}(\xi) & \dfrac{\partial \psi_3}{\partial y}(\xi) & \dfrac{\partial \psi_3}{\partial z}(\xi) & 1 \\ \dfrac{\partial \psi_4}{\partial x}(\xi) & \dfrac{\partial \psi_4}{\partial y}(\xi) & \dfrac{\partial \psi_4}{\partial z}(\xi) & 1 \end{bmatrix} \begin{pmatrix} \delta x \\ \delta y \\ \delta z \\ \delta w \end{pmatrix} = \begin{pmatrix} \hat{\psi}_1 \\ \hat{\psi}_2 \\ \hat{\psi}_3 \\ \hat{\psi}_4 \end{pmatrix} - \begin{pmatrix} \psi_1(\xi) \\ \psi_2(\xi) \\ \psi_3(\xi) \\ \psi_4(\xi) \end{pmatrix} \tag{3}$$

En général, on constate que quelques itérations de Newton suffisent pour avoir une bonne approximation. On peut aussi utiliser un calcul ad-hoc, donné en annexe, pour résoudre le système non-linéaire (1). En pratique, le problème doit être résolu dans un cadre dynamique puisque toutes les variables sont des fonctions du temps. Notons que l'équation (3) peut s'interpréter comme suit : une variation $\delta\xi$ autour du point de référence ξ donne une variation $\delta\psi = H(\xi)\delta\xi$ du vecteur des pseudo-distances. En posant $\delta\psi = \hat{\psi} - \psi(\xi)$, valeur mesurée (correspondant à la vraie valeur $\xi + \delta\xi$) moins valeur prédite par le modèle (correspondant à ξ) et en supposant $\delta\psi$ petit on peut calculer l'erreur $\delta\xi$ correspondant à l'observation $\delta\psi$ en résolvant (3) i.e.

$$H(\xi)\delta\xi = \delta\psi \tag{4}$$

§3. Le problème dynamique

En pratique, le récepteur GPS reçoit en temps réel des signaux transmis par les satellites qu'il a repérés et il doit ajuster à chaque instant l'estimation de son état $\xi(t)$ en fonction des signaux reçus. Le principe de calcul est le suivant : l'équation (4) devient

$$H(\xi(t))\delta\xi(t) = \delta\psi(t) + \epsilon(t) \tag{5}$$

où $\xi(t)$ est l'état de référence à l'instant t (on suppose connaitre une trajectoire de référence pour ξ), $\delta\xi(t)$ est une variation par rapport à $\xi(t)$ qu'on cherche à estimer et $\delta\psi(t)$ est une observation faite à l'instant t liée à $\delta\xi(t)$ par (5). $\epsilon(t)$ est un bruit de mesure. L'algorithme qui réalise cette opération de suivi en temps réel s'appelle un filtre de Kalman et il est lui-même basé sur la méthode des moindres carrés rappelée ci-dessous.

3.1. Moindres carrés. Considérons un système d'équations linéaires qu'on cherche à résoudre aux moindres carrés :

$$y = Hx + \epsilon \tag{6}$$

où y représente le vecteur des observations, x est le vecteur d'état du système qu'on cherche à estimer et ϵ est un «bruit de mesure» aléatoire de moyenne nulle ($\mathbb{E}\epsilon = 0$) et dont on connait la matrice de covariance $R = \mathbb{E}\epsilon\epsilon^T$. (On note $\mathbb{E}$ l'espérance mathématique). On supposera R définie positive. En général, on dispose de plus d'observations que d'inconnues de sorte que le système précédent est surdéterminé. On suppose H injective. L'estimation de x aux moindres carrés consiste à minimiser une norme de l'erreur ϵ, soit :

$$\min_x \|Hx - y\|^2_{R^{-1}}$$

(On note $\|\epsilon\|^2_{R^{-1}} = \epsilon^T R^{-1}\epsilon$). La solution est donnée par

$$\hat{x} = PH^T R^{-1} y \qquad \text{où} \qquad P = (H^T R^{-1} H)^{-1} \tag{7}$$

L'estimateur $\hat{x}$ obtenu est clairement sans biais ($\mathbb{E}\hat{x} = x$) et on vérifie que sa matrice de covariance est P ($\mathbb{E}((\hat{x}-x)(\hat{x}-x)^T) = P$). La matrice P mesure la précision de l'estimation : en particulier si le bruit de mesure ϵ est Gaussien, on peut construire facilement un intervalle (un ellipsoide) de confiance pour x. Dans la suite, on appellera solution du problème de moindres carrés associé à (6), le couple $(\hat{x}, P)$.

Calcul récursif. En pratique, la dimension du vecteur des observations y dans (6) augmente à mesure que de nouvelles observations apparaissent et il devient rapidement impossible de résoudre (7) (ou même de stocker la matrice H). On doit donc procéder de manière récursive, c'est-à-dire calculer $\hat{x}$ en fonction d'une estimation précédente et de nouvelles données. On va donc supposer que la matrice H est partitionnée en lignes $H = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}$, que le vecteur y est lui aussi partitionné en $y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$, et que les bruits $\begin{pmatrix} \epsilon_0 \\ \epsilon_1 \end{pmatrix}$ sont indépendants ce qui se traduit par une matrice de covariance bloc diagonale, soit $R = \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix}$.
L'estimation de x réalisée au vu de la première observation

$$y_0 = H_0 x + \epsilon_0 \tag{8}$$

donne

$$\hat{x}_0 = P_0 H_0^T R_0^{-1} y_0 \qquad \text{avec} \qquad P_0 = (H_0^T R_0^{-1} H_0)^{-1} \tag{9}$$

L'estimation $\hat{x}$ calculée avec le vecteur y complet s'écrit :

$$\begin{aligned} \hat{x} &= P \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix}^{-1} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = P\left(H_0^T R_0^{-1} y_0 + H_1^T R_1^{-1} y_1\right) \\ &= P\left(H_0^T R_0^{-1} H_0 \hat{x}_0 + H_1^T R_1^{-1} y_1\right) \end{aligned}$$

soit

$$\hat{x} = P\left(P_0^{-1}\hat{x}_0 + H_1^T R_1^{-1} y_1\right)$$

avec

$$\begin{aligned} P &= \left\{ \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix}^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \right\}^{-1} = \left(H_0^T R_0^{-1} H_0 + H_1^T R_1^{-1} H_1\right)^{-1} \\ &= \left(P_0^{-1} + H_1^T R_1^{-1} H_1\right)^{-1} \end{aligned}$$

ou

$$P^{-1} = P_0^{-1} + H_1^T R_1^{-1} H_1 \tag{10}$$

ce qui donne la formule recherchée

$$\hat{x} = \hat{x}_0 + P H_1^T R_1^{-1}(y_1 - H_1 \hat{x}_0) \tag{11}$$

La matrice $K = P H_1^T R_1^{-1}$ s'appelle la matrice de gain de Kalman et peut s'écrire sous la forme :

$$K = (P_0^{-1} + H_1^T R_1^{-1} H_1)^{-1} H_1^T R_1^{-1} = P_0 H_1^T (R_1 + H_1 P_0 H_1^T)^{-1} \tag{12}$$

(Cette dernière égalité matricielle se vérifie aisément). On remarquera que la formule (11) de mise à jour de l'estimation de x ne fait intervenir que le couple $(\hat{x}_0, P_0)$: les données particulières y_0, H_0 et R_0 de la première observation sont inutiles pour calculer la solution complète $(\hat{x}, P)$. Le terme $y_1 - H_1 \hat{x}_0$ dans (11) peut s'interpréter comme une erreur de prédiction sur y_1. On notera aussi que la formule (10) de mise à jour de P peut aussi s'écrire en fonction de K sous la forme :

$$P = (I - K H_1) P_0 \tag{13}$$

On a donc montré :

Lemme 1 *Si la matrice R est bloc diagonale, la solution $(\hat{x}, P)$ du problème de moindres carrés (6) donnée par (7) est aussi donnée de manière récursive par*

$$\hat{x} = \hat{x}_0 + K(y_1 - H_1 \hat{x}_0) \tag{14}$$

où $(\hat{x}_0, P_0)$ est la solution du problème de moindres carrés (8) donnée par (9), la matrice K est donnée par $K = P_0 H_1^T (R_1 + H_1 P_0 H_1^T)^{-1}$ et $P = (I - K H_1) P_0$.

Un cas particulier. Supposons maintenant que la matrice H du problème de moindres carrés (6) soit bloc triangulaire et plus précisément soit de la forme $H = \begin{bmatrix} H_{00} & 0 \\ -H_{10} & I \end{bmatrix}$. Ici, x est partitionné en (x_0, x_1) et on sait exprimer explicitement x_1 en fonction de x_0 et y_1. On s'intéresse donc au problème de moindres carrés :

$$\min_x \|Hx - y\|^2_{R^{-1}} \tag{15}$$

sous la contrainte

$$y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} H_{00} & 0 \\ -H_{10} & I \end{bmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} + \begin{pmatrix} \epsilon_0 \\ \epsilon_1 \end{pmatrix} \tag{16}$$

avec $R = \begin{bmatrix} R_0 & 0 \\ 0 & R_1 \end{bmatrix}$. On a alors :

Lemme 2 *Si la matrice R est bloc diagonale, la solution* $(\hat{x} = \begin{pmatrix} \hat{x}_0 \\ \hat{x}_1 \end{pmatrix}, P)$ *du problème de moindres carrés (15) est donnée par :*

$$\hat{x}_0 = P_0 H_{00}^T R_0^{-1} y_0 \qquad \text{où} \qquad P_0 = (H_{00}^T R_0^{-1} H_{00})^{-1} \tag{17}$$

$$\hat{x}_1 = H_{10}\hat{x}_0 + y_1$$

la matrice de covariance de $\hat{x}$ étant

$$P = \begin{bmatrix} P_0 & P_0 H_{10}^T \\ H_{10} P_0 & H_{10} P_0 H_{10}^T + R_1 \end{bmatrix} \tag{18}$$

<u>Preuve</u> Le problème de moindres carrés s'écrit :

$$\min_{x_0, x_1} \ (\|H_{00}x_0 - y_0\|^2_{R_0^{-1}} + \| - H_{10}x_0 + x_1 - y_1\|^2_{R_1^{-1}}) \qquad \text{sous (16)}$$

d'où on trouve immédiatement $\hat{x}_0$ et $\hat{x}_1$ en fonction de $\hat{x}_0$. (On peut aussi calculer $\hat{x}$ par la formule (7)). Pour la matrice de covariance de $\hat{x}$, il vient :

$$P^{-1} = H^T R^{-1} H = \begin{bmatrix} P_0^{-1} + H_{10}^T R_1^{-1} H_{10} & -H_{10}^T R_1^{-1} \\ -R_1^{-1} H_{10} & R_1^{-1} \end{bmatrix}$$

d'où on déduit

$$P = \begin{bmatrix} P_0 & P_0 H_{10}^T \\ H_{10} P_0 & H_{10} P_0 H_{10}^T + R_1 \end{bmatrix}$$

3.2. Equations du filtre. On peut appliquer les formules des lemmes précédents au problème dynamique suivant : on observe à chaque instant de discrétisation t un vecteur $\zeta(t)$ lié au vecteur d'état du système $X(t)$ par une équation linéaire du type

$$\zeta(t) = H(t)X(t) + \epsilon(t) \tag{19}$$

où la matrice $H(t)$ est connue et où $\epsilon(t)$ est un bruit de mesure de matrice de covariance connue $R(t)$, supposée inversible. (Les bruits $\epsilon(t)$ sont de plus supposés indépendants entre eux). Dans le cas le plus simple (statique), $X(t)$ est un vecteur constant x (comme dans (6)) sur lequel on dispose d'observations données par (19). Par example, l'équation (19) pourrait être (3) avec $X(t) = x = \delta\xi$, $H(t) = H(\xi(t))$ et $\zeta(t) = \delta\psi(t)$. On pourrait alors estimer à chaque instant le vecteur constant $x = \delta\xi$ aux moindres carrés en appliquant les formules de mises à jour du lemme 1.

On suppose en fait que l'évolution de l'état du système $X(t)$ est donnée par un modèle linéaire dynamique du type

$$X(t+1) = F(t)X(t) + \eta(t) \tag{20}$$

où $\eta(t)$ est un bruit indépendant de $\epsilon(t)$, de matrice de covariance connue $Q(t)$. (Les bruits $\eta(t)$ sont aussi supposés indépendants entre eux et indépendants

des $\epsilon(t)$). On cherche à estimer aux moindres carrés l'état $X(t)$ en fonction des observations passées $\zeta(s)$, $\quad s = 0, \dots, s = t$.
En notant $\hat{X}(t|t-1)$ et $P(t|t-1)$ l'estimation de $X(t)$ et sa matrice de covariance, avant l'observation de $\zeta(t)$ (donnée par 19) et $\hat{X}(t|t)$ l'estimation de $X(t)$ tenant compte de $\zeta(t)$, c'est-à-dire de l'équation (19), on a par les formules de mise à jour du lemme 1 :

$$K(t) = P(t|t-1)H(t)^T[H(t)P(t|t-1)H(t)^T + R(t)]^{-1}$$

$$\begin{cases} \hat{X}(t|t) = & \hat{X}(t|t-1) + K(t)[\zeta(t) - H(t)\hat{X}(t|t-1)] \\ P(t|t) = & [I - K(t)H(t)]P(t|t-1) \end{cases} \tag{21}$$

Entre deux instants de mesure, l'évolution de l'état est donnée par (20) d'où les formules de mise à jour données par le lemme 2 :

$$\begin{cases} \hat{X}(t+1|t) = F(t)\hat{X}(t|t) \\ P(t+1|t) = F(t)P(t|t)F^T(t) + Q(t) \end{cases} \tag{22}$$

On obtient ces dernières formules en appliquant le lemme 2 au couple (x_0, x_1) avec $x_0 = (X(0), \dots, X(t))$, $x_1 = X(t+1)$ et $y_1 = 0$ (et en remarquant que x_1 ne dépend que de la dernière composante de x_0, c'est-à-dire $X(t)$). Les formules (21) et (22) sont faciles à implémenter sur ordinateur. Elles constituent l'algorithme du filtre de Kalman, qui possède de très nombreuses applications.

§4. Application

En négligeant les perturbations, l'état du récepteur GPS, $\xi(t) = (x(t), y(t), z(t), w(t))$, est une fonction du temps, qui peut être la solution d'une équation différentielle $\dot{\xi}(t) = f(t, \xi(t))$, ce qui donne une trajectoire de référence $\{\xi(s), s \in [0,T]\}$ (supposée connue), la trajectoire effective étant $\{\xi(s) + \delta\xi(s), s \in [0,T]\}$. On prend en fait une modélisation en position-vitesse dans laquelle $\xi(t)$ est la composante «position» d'un vecteur d'état «position-vitesse» $\Xi(t) = (\xi(t), \dot{\xi}(t))$. Dans les modèles les plus simples de GPS, on suppose que le vecteur $\Xi(t)$ est solution d'une équation différentielle linéaire $\dot{\Xi}(t) = A\Xi(t) + e(t)$, et plus précisément, en explicitant :

$$\begin{pmatrix} \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \\ \dot{\xi}_4 \\ \ddot{\xi}_1 \\ \ddot{\xi}_2 \\ \ddot{\xi}_3 \\ \ddot{\xi}_4 \end{pmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \\ \dot{\xi}_1 \\ \dot{\xi}_2 \\ \dot{\xi}_3 \\ \dot{\xi}_4 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \end{pmatrix}$$

où les variables $e_i(t), i = 5, \ldots, 8$ sont des «bruits blancs» (on peut donner un sens précis à l'équation différentielle précédente). La trajectoire de référence $\{\Xi(s), s \in [0,T]\}$ est obtenue par l'équation précédente en prenant $e_i = 0, i = 5, \ldots, 8$. On dispose à chaque instant de discrétisation $t, t+h, \ldots$ (de l'ordre d'une fraction de seconde) d'une observation $\zeta(t)$ qui est liée à l'état du système $X(t) = \delta\Xi(t)$ par une équation linéaire du type (5). L'équation donnant la trajectoire de référence $\{\Xi(s), s \in [0,T]\}$ étant linéaire, l'équation donnant la variation $\{\delta\Xi(s), s \in [0,T]\}$ est identique.

En prenant un schéma d'Euler au premier ordre, on peut réécrire les équations d'état (20) et (19) suivantes :

$$\begin{pmatrix} X_1(t+h) \\ X_2(t+h) \\ X_3(t+h) \\ X_4(t+h) \\ X_5(t+h) \\ X_6(t+h) \\ X_7(t+h) \\ X_8(t+h) \end{pmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & h & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & h & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & h & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & h \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ X_4(t) \\ X_5(t) \\ X_6(t) \\ X_7(t) \\ X_8(t) \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ U_1(t) \\ U_2(t) \\ U_3(t) \\ U_4(t) \end{pmatrix}$$

$$\begin{pmatrix} \zeta_1(t) \\ \zeta_2(t) \\ \zeta_3(t) \\ \zeta_4(t) \end{pmatrix} = \begin{bmatrix} \frac{\partial\psi_1}{\partial x} & \frac{\partial\psi_1}{\partial y} & \frac{\partial\psi_1}{\partial z} & 1 & 0 & 0 & 0 & 0 \\ \frac{\partial\psi_2}{\partial x} & \frac{\partial\psi_2}{\partial y} & \frac{\partial\psi_2}{\partial z} & 1 & 0 & 0 & 0 & 0 \\ \frac{\partial\psi_3}{\partial x} & \frac{\partial\psi_3}{\partial y} & \frac{\partial\psi_3}{\partial z} & 1 & 0 & 0 & 0 & 0 \\ \frac{\partial\psi_4}{\partial x} & \frac{\partial\psi_4}{\partial y} & \frac{\partial\psi_4}{\partial z} & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \\ X_4(t) \\ X_5(t) \\ X_6(t) \\ X_7(t) \\ X_8(t) \end{pmatrix} + \begin{bmatrix} V_1(t) \\ V_2(t) \\ V_3(t) \\ V_4(t) \end{bmatrix}$$

Ici, les variables U_i et V_i sont des bruits blancs discrets et les dérivées partielles sont évaluées en $\xi(nh)$. En utilisant Scilab ou Matlab[1], on peut implémenter les deux fonctions `update` (formules (21)) et `propag` (formules (22)) ci-dessous, qui permettent le calcul récursif de $\hat{X}(t)$ et de sa covariance $P(t|t)$. Etant donnée l'estimation courante $(\hat{X}(t|t-1), P(t|t-1))$ notée (`x0,P0`)), on met à jour (`x0,P0`) au vu de l'observation $y(t)$ (équation (19)) ce qui donne (`x,P`) par la fonction `update` (formules (21). On calcule ensuite l'estimée de $X(t+h)$ et la covariance associée par `propag` (formules (22), conséquence de l'équation (20), soit (`x1,P1`). A l'instant suivant, on recommence : appel de `update` avec (`x1=x0,P1=P0`) puis appel de `propag` avec le nouveau (`x,P`) calculé par `update`.

```
function [x,Q]=update(x0,Q0,H,y,R)
//estimation x0 et covariance Q0 avant observation
//+ observation y = H x + e , cov(e)=R
//==>estimation x et covariance Q apres observation
```

[1] Le code proposé ci-dessous est donné en Scilab (ou Matlab en adaptant les commentaires). On pourra bien entendu utiliser Maple, les opérations à effectuer sont très simples : produit de matrices, transposition, inversion ...

```
K=Q0*H'*inv(H*Q0*H' + R);
x=x0+K*(y-H*x0);
Q=(eye()-K*H)*Q0;

function [x1,Q1]=propag(x,Q,F,Qe)
x1 = F*x;
Q1 = F*Q*F'+Qe;
```

§5. Annexe : Un calcul explicite

En développant l'équation (1), il vient pour $k = 1, \dots 4$:

$$X_k^2 + Y_k^2 + Z_k^2 - \hat{\psi}_k^2 - 2(X_k x + Y_k y + Z_k z - \hat{\psi}_k w) = -(x^2 + y^2 + z^2 - w^2)$$

On introduit le produit scalaire de deux vecteurs à 4 composantes par $\langle g, h\rangle = g^T L h$ où $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$

Le système d'équations précédent peut alors s'écrire :

$$\alpha - 2BL\xi + \langle \xi, \xi\rangle e = 0 \tag{23}$$

où $B = \begin{bmatrix} X_1 & Y_1 & Z_1 & \hat{\psi}_1 \\ X_2 & Y_2 & Z_2 & \hat{\psi}_2 \\ X_3 & Y_3 & Z_3 & \hat{\psi}_3 \\ X_4 & Y_4 & Z_4 & \hat{\psi}_4 \end{bmatrix}$, $\alpha_k = \langle \begin{pmatrix} X_k \\ Y_k \\ Z_k \\ \hat{\psi}_k \end{pmatrix}, \begin{pmatrix} X_k \\ Y_k \\ Z_k \\ \hat{\psi}_k \end{pmatrix} \rangle$ et $e = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. De (23), on tire $L\xi = \frac{1}{2}(B^{-1}\alpha + \langle \xi, \xi\rangle e)$, puis en formant le crochet du premier membre de (23) avec lui même :

$$\langle B^{-1}\alpha, B^{-1}\alpha\rangle + 2\langle \xi, \xi\rangle(\langle B^{-1}\alpha, B^{-1}e\rangle - 1) + \langle \xi, \xi\rangle^2 \langle B^{-1}e, B^{-1}e\rangle = 0$$

On en déduit $< \xi, \xi >$ en résolvant cette équation du second degré (et en éliminant la solution aberrante), puis on résout l'équation (23) pour trouver ξ. Les données numériques qui suivent

$$B = \begin{bmatrix} X_1 & Y_1 & Z_1 & \hat{\psi}_1 \\ X_2 & Y_2 & Z_2 & \hat{\psi}_2 \\ X_3 & Y_3 & Z_3 & \hat{\psi}_3 \\ X_4 & Y_4 & Z_4 & \hat{\psi}_4 \end{bmatrix}$$

$$= \begin{bmatrix} \texttt{-11716227.778} & \texttt{-10118754.628} & \texttt{21741083.973} & \texttt{22163882.029} \\ \texttt{-12082643.974} & \texttt{-20428242.179} & \texttt{11741374.154} & \texttt{21492579.823} \\ \texttt{14373286.650} & \texttt{-10448439.349} & \texttt{19596404.858} & \texttt{21492492.771} \\ \texttt{10278432.244} & \texttt{-21116508.618} & \texttt{-12689101.970} & \texttt{25284588.982} \end{bmatrix}$$

donnent la solution ξ = `[ 595025.05, -4856501.2, 4078330., 147.33078 ]` avec c = `299792458` .

Développements possibles suggérés pour le commentaire

Les points suivants du texte pourront être détaillés :

- Résolution les équations (1) par une méthode de Newton (ou en appliquant la procédure (cf annexe) donnant le calcul exact).
- Exposé de méthodes numériques classiques (Householder, ...) de résolution des moindres carrés.
- Mise à jour des calculs lorsqu'on rajoute une ligne ou une colonne à la matrice avec ces méthodes.
- Expliciter comment les équations du filtre de Kalman (21) et (22) sont obtenues à partir des lemmes 1 et 2.
- Illustrer sur machine un exemple simplifié de filtre de Kalman. On pourra estimer la trajectoire (position et vitesse) d'un mobile se déplaçant à vitesse constante sur une droite à partir des mesures des distances du mobile à de deux stations fixes (modèle simplifié de GPS sur le plan constitué des variables X_1, X_2, X_5, X_6).
- Expliciter comment on définit un modèle linéarisé autour d'une trajectoire.
- Discuter les limitations de la solution proposée au problème.

Chapitre 2
Commentaire de l'auteur

Ce texte sur le GPS est un exemple concret de problème de moindres carrés. La principale difficulté du texte vient de sa longueur. On demandera donc d'abord au candidat de faire la synthèse du texte en expliquant le problème posé, la modélisation retenue et les contraintes posées pour sa résolution. Le candidat devrait être capable d'expliquer en quoi le problème posé est différent du problème des moindres carrés linéaire habituel. Il y a deux points particuliers dans l'algorithme des moindres carrés présenté (qui s'appelle un filtre de Kalman dans la littérature anglo-saxone). D'une part, on doit ici traiter les données au fur et à mesure de leur réception, ce qui nécessite un calcul récursif (on doit mettre à jour en temps réel l'estimation courante). D'autre part on cherche à estimer à l'instant t seulement la partie $X(t)$ du vecteur $[X(0), X(1), \ldots, X(t)]$.

La partie modélisation GPS est simple : on cherche à déterminer 4 inconuues qui sont les coordonnées du récepteur et le décalage entre les horloges GPS et l'horloge interne du récepteur. Le modèle linéarisé (3) est celui de la méthode de Newton. Le calcul ad-hoc donné en appendice pour résoudre le système non linéaire (1) n'est pas très intéressant (d'un point de vue mathématique) et ne mérite pas que l'on s'y attarde.

L'essentiel du texte se trouve dans l'exposition du modèle dynamique. Le modèle général considéré est donné par (19) et (20). On cherche $\hat{X}(t|t)$ une estimation aux moindres carrés de $X(t)$ au vu des équations (19) et (20) pour $t = 0, \ldots, t = t$. La solution est donnée par les formules (21) et (22).

Dans un premier temps on pourrait pour simplifier l'exposé, supposer que la variable d'état $X(t)$ est une constante X. On a alors affaire à un problème de moindres carrés classique. Le candidat pourrait alors exposer rapidement les grandes lignes d'une méthode de résolution classique et exposer la méthode donnée dans le Lemme 1. Il n'est sans doute pas nécessaire de refaire la démonstration de ce lemme qui est exposée en détail dans le texte. En revanche, il conviendra d'expliquer clairement pourquoi on cherche une formule de type (14), comment elle s'interprête et pourquoi elle est intéressante. On pourra l'illuster sur machine en utilisant la fonction `update`.

Le point suivant est de montrer comment on doit adapter la méthode des moindres caarés lorsque l'état à estimer $X(t)$ dépend du temps. On doit connaitre la matrice de covariance $Q(t|t-1)$ de l'état $X(t)$ à l'instant t, mais avant l'observation faite sur $X(t)$ (c'est-à-dire sans tenir compte de l'équation (19)). Le Lemme 2 permet de calculer cette covariance donnée par (22). Il n'est sans doute pas non plus nécessaire de refaire en détail le calcul du Lemme 2.

L'implémentation de l'algorithme consiste à appliquer successivement les deux fonctions `update` lorsqu'on dispose d'une observation et `propag` entre deux instants d'observation. Le candidat pourra donner une simulation à par-

tir du modèle linéarisé proposé. Il pourra aussi expliquer la signification mathématique d'un modèle linéarisé.

Deuxième partie

Modèle simplifié de mélange liquide-gaz

Chapitre 3
Modèle simplifié de mélange liquide–gaz

§1. Modélisation

1.1. Modèle physique. On cherche à modéliser la propagation d'ondes dans un mélange liquide–gaz. On suppose que le liquide est incompressible, et que le gaz, compressible, se trouve dans le liquide sous la forme de petites bulles qui contiennent le même nombre de molécules. On s'intéressera ici à une modélisation mono-dimensionnelle de cette situation : on considère, sur le segment $[0, L]$, une alternance de gaz et de liquide, comme l'illustre la figure 1.

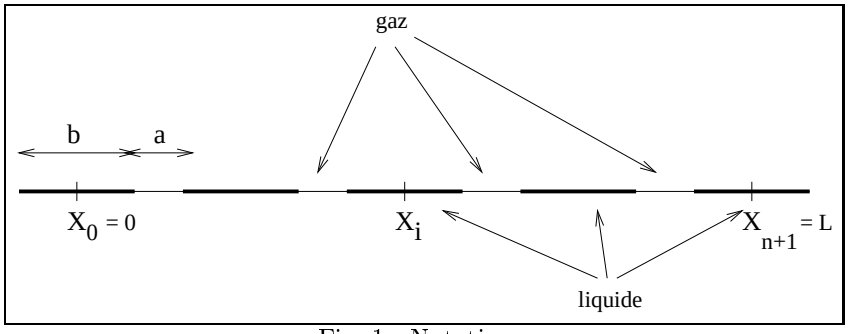


Fig. 1 : Notations.

On note ρ la masse linéique du liquide, de telle sorte que la masse d'un "segment" S_i centré en x_i est ρb. Le gaz piégé entre deux segments successifs est supposé obéir à la loi des gaz parfaits à température constante, c'est–à–dire

$$\text{(pression)} \times \text{(volume)} \quad = \quad \text{Constante} \times \text{(nbre de molécules de gaz)}$$

soit

$$PV = kN, \tag{1}$$

où le "volume" représente en fait une longueur dans le cadre de ce modèle mono–dimensionnel. On suppose les deux extrémités fixes (x_0 et x_{n+1} positionnés en 0 et L, respectivement, sont fixes), et on prend un nombre de molécules de gaz dans chaque poche constant égal à N. On considérera que N s'écrit $\beta(a+b)$, où β est le nombre de molécules de gaz par unité de longueur du mélange. Ce système liquide/gaz admet un état d'équilibre représenté schématiquement sur la figure 1, pour lequel toutes les bulles ont même taille b. Les positions des centres des segments fluides sont (en notant $h = a + b$)

$$X_0 = 0\,, \quad X_1 = h\,, \quad \ldots\,, \quad X_i = ih\,, \quad \ldots\,, \quad X_{n+1} = (n+1)h = L. \tag{2}$$

1.2. Equation d'évolution. On note $x_i = X_i + u_i$ la position du centre de S_i ($i^{\text{ème}}$ segment), de telle sorte que l'état d'équilibre correspond à $u_i = 0$ pour tout i. Le bilan F_i des forces s'exerçant sur un élément de liquide centré en x_i est la somme des pressions exercées à droite et à gauche par le gaz, c'est–à–dire d'après le modèle (1)

$$F_i = k\beta h \left(\frac{1}{a + u_i - u_{i-1}} - \frac{1}{a + u_{i+1} - u_i} \right). \tag{3}$$

La masse de S_i étant ρb, la relation fondamentale de la dynamique s'écrit donc

$$\rho b \ddot{u}_i = F_i. \tag{4}$$

On note $\mathbf{u} = (u_1, \ldots, u_n)$ le vecteur des inconnues, $v_i = \dot{u}_i$ la vitesse de S_i, $\mathbf{v}$ le vecteur des v_i et $\mathbf{F}$ le vecteur des forces F_i. On peut se ramener de cette façon à une équation différentielle ordinaire dans $\mathbb{R}^{2n}$:

$$\left\{ \begin{array}{rcl} \dfrac{d\mathbf{u}}{dt} & = & \mathbf{v} \\ \\ \rho b \dfrac{d\mathbf{v}}{dt} & = & \mathbf{F}(\mathbf{u}). \end{array} \right. \tag{5}$$

§2. Vitesse de propagation

On cherche à estimer la vitesse de propagation associée au modèle décrit ci–dessus, pour des ondes de faible amplitude ($u_i << a + b$). La force F_i peut s'exprimer

$$\begin{aligned} F_i &= k\beta h \left(\frac{1}{a + u_i - u_{i-1}} - \frac{1}{a + u_{i+1} - u_i} \right) \\ &= k\beta h \frac{u_{i+1} - 2u_i + u_{i-1}}{(a + u_i - u_{i-1})(a + u_{i+1} - u_i)} \\ &= k\beta h \, \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \, \frac{1}{\left(\alpha + \dfrac{u_i - u_{i-1}}{h}\right)\left(\alpha + \dfrac{u_{i+1} - u_i}{h}\right)}, \end{aligned}$$

avec $h = a + b$ comme précédemment, et $\alpha = a/(a + b)$. On considère maintenant que les u_i sont les valeurs en X_i d'une fonction u régulière, et l'on identifie les expressions faisant intervenir les u_i à des dérivées en espace de u :

$$\frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} \approx u''(X_i) \quad , \quad \frac{u_i - u_{i-1}}{h} \approx \frac{u_{i+1} - u_i}{h} \approx u'(X_i),$$

où u' (respectivement u'') est la dérivée (respectivement la dérivée seconde) partielle par rapport à la variable d'espace. On aboutit donc, formellement, à l'équation aux dérivées partielles sur u (en remarquant que $b/h = 1 - \alpha$)

$$\rho(1-\alpha)\ddot{u} - \frac{k\beta}{(\alpha + u')^2} u'' = 0. \tag{6}$$

Si l'on fait l'hypothèse supplémentaire que u' reste petit devant α, on obtient, toujours formellement, une équation des ondes linéaire

$$\ddot{u} - c^2 u'' = 0 \ \text{ avec } \ c^2 = \frac{k\beta}{\rho(1-\alpha)\alpha^2}, \tag{7}$$

à laquelle sont associées les vitesses de propagation $+c$ et $-c$.

§3. Simulation numérique

On se propose de résoudre numériquement le système (5) pour modéliser la propagation d'une onde dans le milieu. Pour perturber le système, on va agir sur la position du point x_0, qu'on ne supposera donc plus fixe, mais dont la position va être imposée au cours du temps. On introduit une discrétisation de l'intervalle en temps $[0, T]$:

$$t_0 = 0 < t_1 = \delta t < t_2 = 2\delta t < \cdots < t_M = M\delta t = T,$$

et on note u_n^m (respectivement v_n^m) l'approximation de u_n (respectivement v_n) au temps t_m. On discrétise le système (5) de la façon suivante :

$$\left\{ \begin{array}{rcl} \dfrac{u_n^m - u_n^{m-1}}{\delta t} & = & v_n^m \\[2ex] \rho b \, \dfrac{v_n^{m+1} - v_n^m}{\delta t} & = & k\beta h \left(\dfrac{1}{a + u_n^m - u_{n-1}^m} - \dfrac{1}{a + u_{n+1}^m - u_n^m} \right). \end{array} \right. \tag{8}$$

ce qu'on peut écrire en éliminant les v :

$$u_n^{m+1} = 2u_n^m - u_n^{m-1} + \delta t^2 \mu \left(\frac{1}{a + u_n^m - u_{n-1}^m} - \frac{1}{a + u_{n+1}^m - u_n^m} \right). \tag{9}$$

avec $\mu = k\beta h/\rho b$. Les conditions aux limites en espace s'écrivent

$$u_0^m = \lambda(t) \ , \quad u_N^m = 0 \quad \forall m = 0, \ldots, M, \tag{10}$$

où $\lambda(t)$ est donnée, et les conditions initiales en temps

$$u_n^0 = U_n^0 \ , \quad u_n^{-1} = U_n^{-1} \ , \qquad \forall n = 1, \ldots, N, \tag{11}$$

où les U_n^0 et U_n^{-1} sont des données qui permettent de représenter l'état du système au temps initial.

Les relations (9) et les conditions (10) et (11) permettent donc de calculer les u_n^m pour $n = 1, \ldots, N$, et $m = 1, \ldots, M$. On présente un résultat de calcul basé sur les paramètres suivants

$$\rho = 1 \ , \quad N = 99 \ , \quad L = 1.0 \quad (\Rightarrow h = 10^{-2}) \ , \quad \alpha = 0.2 \ ,$$
$$\beta = 1 \ , \quad k = 1.0 \ , \quad \delta t = 10^{-3} \ , \quad T = 0.12.$$

La perturbation λ est simplement $\lambda(t) =$ constante $= h/10$, et les conditions initiales correspondent à l'état d'équilibre statique $U_n^0 = U_n^{-1} = 0$ pour tout $n = 1, \ldots, N$. Les approximations des u_i au temps $T = 0.12$ sont représentés sur la figure 2.

Dans un second temps, on s'intéresse au comportement de l'onde dans le cas où le milieu n'est plus homogène. Il s'agit d'étudier la possibilité d'utiliser le phénomène de propagation mis en évidence pour détecter la présence dans un mélange fluide–bulles de matériaux plus lourds que l'eau, et si possible les localiser.

On va modéliser numériquement une situation parfaitement analogue à la précédente, dans le cas où l'un des segments incompressibles de masse linéique ρ est remplacé par un segment de masse linéique $\rho' > \rho$. On choisit de modifier la masse volumique du segment correspondant à u_{50} (au milieu de la zone modélisée), et on prend $\rho' = 2\rho$. La figure 3 représente comme précédemment l'approximation des u_i au temps T. On remarque la présence (sur la moitié gauche) d'une onde réfléchie du fait de la présence de la partie plus lourde au centre de la configuration.

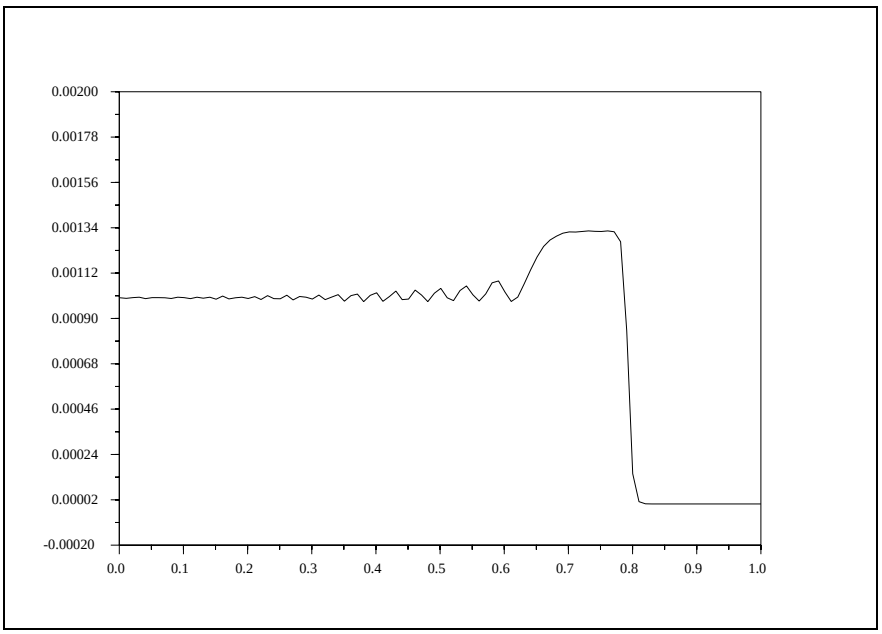


Fig. 2 : Milieu homogène.

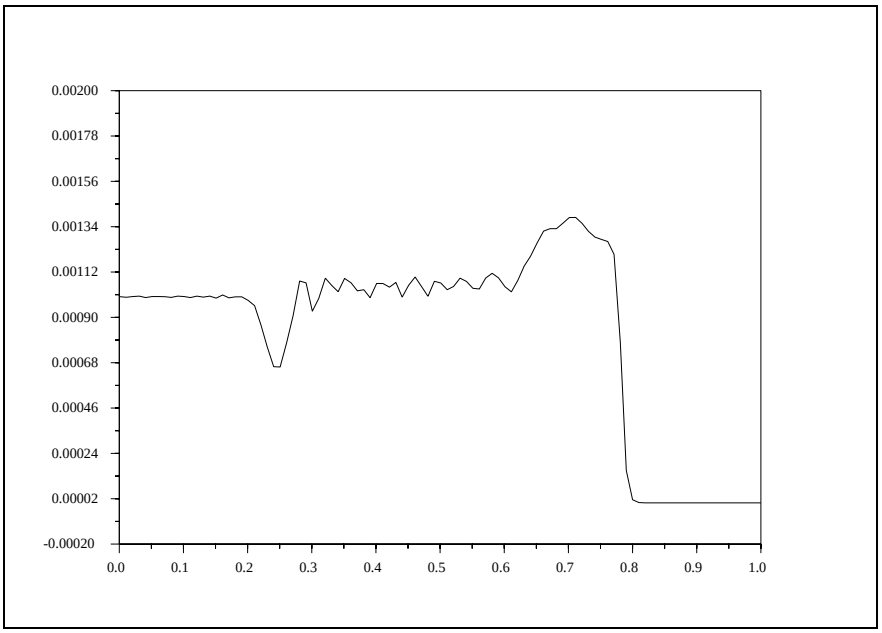


Fig. 3 : Milieu non–homogène.

Suggestions de développement

- Proposer d'autres méthodes pour résoudre numériquement le système différentiel (5).
- Étudier numériquement la validité de la vitesse de propagation c obtenue dans la section 2.
- Étudier la stabilité de l'état d'équilibre du système (5), et du système obtenu par ajout d'un terme de dissipation

$$\begin{cases} \dfrac{d\mathbf{u}}{dt} &= \mathbf{v} \\[2ex] \rho b\,\dfrac{d\mathbf{v}}{dt} &= \mathbf{F}(\mathbf{u}) - \kappa\mathbf{v}. \end{cases}$$

- Envisager l'application de la démarche proposée au cas où la température n'est plus supposée constante. L'équation (1) est alors remplacée par

$$PV = NRT, \tag{12}$$

où R est une constante, et T est la température, que l'on suppose connue sur le domaine. Etudier l'influence d'une température variable (par exemple variant de façon affine sur le domaine) sur les phénomènes mis en évidences.

– Appliquer la démarche proposée à d'autres situations (bulles de tailles différentes, juxtaposition de deux milieux de caractéristiques différentes, ...).

Chapitre 4
Commentaires de l'auteur

L'étude du comportement de mélanges fluide–gaz est un problème difficile qui fait l'objet à l'heure actuelle d'un grand nombre de projets de recherche. Ce texte ne prétend aucunement se faire le reflet de l'avancée de ces travaux, mais illustre la possibilité d'utiliser des outils mathématiques et numériques simples pour obtenir des informations, au moins qualitatives, sur des sytèmes complexes. L'élaboration du modèle est basée sur le principe suivant : comment, disposant d'un formalisme et d'outils de résolution numériques donnés (en l'occurence le programme de l'Agrégation externe), améliorer sa compréhension d'une certaine réalité physique (ici l'observation d'ondes se propageant dans les milieux diphasiques) ?

La considération préliminaire du problème physique sans arrière pensée de faisabilité, qui constitue la première étape de la démarche de modélisation, consiste à identifier les paramètres ou phénomènes au sein d'un mélange fluide–gaz qui sont susceptibles d'avoir une influence sur la propagation d'ondes. On est amené à considérer

1. la proportion de gaz dans le liquide, et la taille (pas nécessairement uniforme) des bulles,
2. la disposition relative des bulles de gaz dans le fluide (elles n'ont pas de raison d'être disposées de façon régulières),
3. la compressibilité du gaz,
4. la compressibilité du liquide (on sait que des ondes peuvent se propager dans le liquide seul, même s'il n'est que faiblement compressible),
5. la tension surfacique (force agissant au niveau de l'interface gaz–liquide, qui tend à donner aux bulles une forme sphérique), et conjointement la forme des bulles et son évolution,
6. l'écoulement du liquide, ainsi que celui du gaz dans les bulles, tous deux susceptibles de dissiper de l'énergie sous forme de chaleur,
7. la densité du liquide,
8. la densité du gaz.

La volonté de se ramener à la dimension 1 élimine d'emblée la prise en compte des points 2, 5, et 6. La compressibilité du liquide, en général très inférieure à celle du gaz, incite à penser que la vitesse de transmission de l'information dans le mélange dépendra essentiellement de la compressibilité du gaz, ce qui explique le choix de prendre un liquide incompressible. Remarquons d'une part que ce choix n'est probablement plus valide dans le cas d'un mélange très dispersé, mais que d'autre part la compressibilité du liquide pourrait être prise en compte dans le cadre d'un modèle analogue à celui proposé. Pour ce qui est des densités, le rapport de 1 à 1000 pour les gaz et

liquides courants suggère de "négliger" la densité du gaz. Ici encore ce choix n'a qu'une zone de validité limitée : si la fraction volumique de liquide est de l'ordre de 1/1000, alors les masses de liquide et de gaz (et par conséquent les énergies cinétiques, puisque les vitesses sont les mêmes) sont du même ordre de grandeur, et le modèle n'est plus justifié.

La démarche suivante consiste à quantifier le point essentiel (le 3), ce que l'on choisit ici de faire par utilisation du modèle de gaz le plus simple. Il est à noter ici encore que la faisabilité joue un rôle dans le choix du modèle : le passage d'une onde dans un tel milieu se traduit par une augmentation de la température, qui n'est donc plus constante. Une appréhension plus fine du phénomène devrait être basée sur un modèle non pas isotherme mais adiabatique (on néglige les échanges de chaleur par diffusion thermique, très lents par rapport aux phénomènes de propagation considérés). Le modèle résultant (qui fait intervenir la constante γ, qui dépend elle–même de la structure atomique du gaz) peut être traité numériquement, mais l'approche asymptotique menée à la section 2 est plus délicate.

Malgré son caractère extrèment grossier, le modèle résultant de cette démarche présente certaines des caractéristiques que l'on doit attendre d'un modèle satisfaisant :

- mise en évidence d'une onde qui se déplace à vitesse finie,
- reproduction numérique du phénomène de réflexion partielle d'une onde par une singularité
- possibilité d'étudier la dépendance de l'observable essentielle (la vitesse de propagation) en fonction des paramètres du milieu,

auxquelles on peut ajouter des propriétés non abordées qui pourraient donner lieu à des développements

- le système différentiel admet une intégrale première, qui correspond à la conservation de l'énergie totale du mélange,
- les contacts entre segments liquides sont exclus (ou, de façon équivalente, le volume d'une poche de gaz ne s'annule pas).

Chapitre 5
Commentaire d'un non spécialiste

§1. Un problème d'interaction liquide-gaz très simplifié

L'objectif du texte proposé est l'étude d'un mélange de deux fluides, l'un incompressible (en l'occurrence un liquide, ou la juxtaposition de deux milieux liquides de masses volumiques différentes), l'autre compressible (en l'occurrence un gaz, présent dans le mélange sous forme de bulles de taille *a priori* diverses à un instant $t = 0$ initial, réparties arbitrairement au sein du mélange). La compressibilité du gaz est responsable de forces de pression exercées suivant la normale extérieure au point courant de la frontière de chaque bulle B, forces dont l'intensité p_B se trouve régie par la loi des gaz parfaits à température constante (si l'on suppose en première approximation la température du milieu constante, soit $p_B \ \mathrm{vol}_B = \mathbf{k} M_B$, $\mathbf{k}$ désignant une constante liée au milieu et plus précisément à la température, vol_B le volume de la bulle et M_B le nombre de molécules de gaz présentes au sein de la bulle B) ou, si l'on envisage – ce qui est plus naturel compte tenu du fait que la propagation d'une onde dans un milieu engendre une augmentation de la température– un milieu adiabatique plutôt qu'isotherme, par les principes de la thermodynamique, soit $p_B \ \mathrm{vol}_B = \gamma M_B w_B$, w_B désignant la température (supposée uniforme au sein de cette bulle) à laquelle B est soumise et γ une constante absolue. Devraient être considérées aussi les forces de tension surfacique générées suivant le principe de moindre action par les forces visant à redonner à chaque bulle une forme sphérique (minimalité de l'aire du bord à volume prescrit). Un tel modèle se trouve outrageusement simplifié ici par le fait qu'il est transposé dans le cadre mono-dimensionnel, ce qui permet de s'affranchir des questions de nature géométrique intervenant de manière évidente dans l'étude des forces de tension surfacique ; c'est ce modèle linéique que nous décrirons dans notre première section (section 2), modèle permettant de ramener l'étude du problème (le nombre $N+2$ de segments fluides, indexés de 0 à $N+1$ étant fixé) à l'étude d'un système différentiel ordinaire (et résoluble),

$$u' = v \ , \ v' = F(u,t) \ , \quad u = (u_1, ..., u_N) \ , \ v = (v_1, ..., v_N) \tag{1}$$

système présenté dans le texte comme l'équation d'évolution sous-jacente au modèle ; $u_1, ..., u_N$ mesurent ici les écarts des positions des centres des segments fluides numérotés de 1 à N avec leur position initiale –les segments fluides extrêmes sont exclus car le dernier est supposé fixe tandis que le premier est soumis dans le temps à une perturbation imposée, responsable par ailleurs de la dépendance en t de F dans le système différentiel (1) – et les v_n, $n = 1, ..., N$, sont les dérivées (comme fonction du temps) des u_i. Nous justifierons (grâce à des considérations basées sur le caractère lipschitzien de

$F(\cdot, t)$) la stabilité numérique du schéma à d'Euler utilisé par la suite pour la simulation numérique de l'évolution au départ de la position d'équilibre statique ($u_n(0) = v_n(0) = 0$, $n = 1, ..., N$) du phénomène que génère un signal perturbateur initial $t \mapsto \mathbf{U}_0(t)$ (perturbation du centre du premier segment liquide S_0).

Dans la section suivante (section 3), nous envisagerons la résolution numérique du système (1) suivant les méthodes classiques (la méthode d'Euler à un pas est dans le texte la méthode proposée, mais nous pourrions envisager d'autres méthodes numériques du type Runge-Kutta par exemple). L'algorithme (écrit sous l'environnement MATLAB suivant le principe du schéma d'Euler à un pas proposé dans la section 3 du texte) nous permettra d'intégrer soit la juxtaposition de deux milieux fluides, soit la présence, au sein du milieu, d'un segment fluide de paramètres différents des autres. Il permettra de mettre en évidence un phénomène de propagation d'onde à vitesse constante (dans chacun des milieux fluides juxtaposés).

Puis, dans la section 4, nous tenterons de valider numériquement (dans le cadre d'un milieu fluide homogène ou de la juxtaposition de deux milieux fluides homogènes, les segments liquides incompressibles ayant tous même longueur $\mathbf{b}$ et les bulles même diamètre en position d'équilibre statique) le fait que, si les ondes sont de faible amplitude, le phénomène se trouve régi par <u>l'équation des ondes</u> en dimension spatiale un

$$\frac{\partial^2 u}{\partial t^2} = c^2(x)\,\frac{\partial^2 u}{\partial x^2}\,, \tag{2}$$

la fonction c^2 étant une fonction constante par morceaux, t désignant la variable de temps, x la variable d'espace (discrétisée suivant la position des centres des segments fluides à l'état initial) ; les discontinuités de c^2 rendront compte des points où s'opère la juxtaposition des divers milieux fluides, tandis que la valeur de $c^2(x)$ sera, quant-à-elle, <u>inversement proportionnelle</u> à la masse linéique $\rho(x)$ du segment fluide dont le centre est en position x à l'instant initial. Remarquons que, si le milieu liquide est homogène (ce qui implique par voie de conséquence $c^2(x) = c^2$), le fait que u obéisse à l'équation des ondes (2) implique via le changement de variable $(t, x) \mapsto (x+t/c, x-t/c)$ (si u est supposée régulière, disons de classe C^2) que u s'écrive

$$u(x,t) = \varphi(x - t/c) + \psi(x + t/c)\,,$$

où φ et ψ sont des fonctions de classe C^2 d'une seule variable, ce qui justifie le qualificatif de $\pm 1/c$ comme vitesses de propagation de l'onde ; nous y reviendrons dans la section suivante.

Nous envisagerons enfin (dans la section 5) l'évolution du phénomène (numériquement car il s'agit d'un système hautement <u>non-linéaire</u>) en présence d'un terme de dissipation, ce qui transforme l'EDO (1) en

$$u' = v\ ,\ \ v' = F(u,t) - \lambda v \tag{3}$$

(λ désignant un paramètre) ou sous l'hypothèse d'adiabaticité en place de celle qui consiste à supposer le milieu isotherme. Nous reviendrons en particulier dans cette dernière section sur les questions de stabilité relative à la position d'équilibre statique $u_n(0) = v_n(0) = 0$, $n = 1, ..., N$, lorsque la perturbation $\mathbf{U}_0$ est constante en étudiant le spectre de la matrice intervenant dans le problème linéarisé.

Nous proposerons enfin une conclusion succinte.

§2. L'équation d'évolution et ses variantes

Soient $S_0, ..., S_{N+1}$ les $N+2$ "segments" liquides consécutifs (supposés incompressibles et de même longueur $\mathbf{b}$) ; on ne supposera pas ici *a priori* le milieu liquide homogène (chaque segment liquide l'étant néanmoins) et l'on notera ρ_n, $n = 0, ..., N+1$, la masse linéique du liquide constituant le segment S_n (la masse de ce segment vaut donc $\rho_n b$). Notons que l'on conviendra de désigner par des caractères majuscules, des lettres grecques ou des caractères gras les paramètres initiaux du système (tels que N, les ρ_n, $n = 0, ..., N+1$, γ, $\mathbf{b}$) et par des caractères italiques les quantités appelées éventuellement à varier au cours du temps. Les $N+2$ segments liquides "emprisonnent" donc $N+1$ bulles de gaz (elles, compressibles, et numérotées de 0 à N) ; si, pour n entre 0 et N, M_n désigne le nombre de molécules de gaz présent dans la n-ème bulle et w_n la température à laquelle est soumise cette bulle (température que pour simplifier on supposera uniforme sur la bulle, cette simplification se trouvant justifiée par le fait que le milieu est assujetti à un champ de température variant éventuellement au cours du temps, mais de variation lente en espace), la relation fondamentale de la dynamique classique (loi de Newton) implique, si l'on suppose que la pression exercée par chaque bulle d'air B_n se trouve régie par la loi des gaz parfaits (à température constante w_n), que pour tout n entre 1 et N

$$\rho_n \mathbf{b} \frac{d^2 u_n}{dt^2} = \gamma \Big(\frac{M_{n-1}\, w_{n-1}}{\mathbf{a}_{n-1} + u_n - u_{n-1}} - \frac{M_n\, w_n}{\mathbf{a}_n + u_{n+1} - u_n} \Big) , \tag{4}$$

γ désignant une constante absolue (liée au gaz). Ici u_n, $n = 0, ..., N$, désigne la variation (dépendant, elle, du temps) entre le centre du segment liquide S_n à l'instant courant et sa position initiale et $\mathbf{a}_n$, $n = 0, ..., N$, représente la longueur initiale de la n-ème bulle de gaz. Le $N+2$-ème segment liquide est, quant à lui, supposé fixe (de centre le point d'abscisse $L = \mathbf{a}_0 + \cdots + \mathbf{a}_N + (N+1)\mathbf{b}$) tandis que le centre du segment liquide S_0 est assujetti à une perturbation temporelle $\mathbf{U}_0(t)$ avec $|\mathbf{U}_0(t)| < \mathbf{a}_0$ fixée.

Si le milieu est supposé isotherme, on écrira le système (4) sous la forme (1), où, si $u = (u_1, ..., u_N)$, pour tout $n \in \{1, ..., N\}$,

$$F_n(u,t) = \frac{\mathbf{k}}{\rho_n \mathbf{b}} \Big(\frac{M_{n-1}}{\mathbf{a}_{n-1} + u_n - u_{n-1}} - \frac{M_n}{\mathbf{a}_n + u_{n+1} - u_n} \Big) , \ n = 1, ..., N ,$$

$\mathbf{k}$ désignant une constante (liée au gaz et à la température) ; notons que seul F_1 dépend en fait de t et s'écrit

$$F_1(u,t) = \frac{\mathbf{k}\beta\mathbf{h}}{\rho_1\mathbf{b}}\Big(\frac{1}{\mathbf{a}_0 + u_1 - \mathbf{U}_0(t)} - \frac{1}{\mathbf{a}_1 + u_2 - u_1}\Big)$$

tandis que l'on a, puisque le dernier segment liquide reste fixe

$$F_N(u,t) = F_N(u) = \frac{\mathbf{k}\beta\mathbf{h}}{\rho_N\mathbf{b}}\Big(\frac{1}{\mathbf{a}_{N-1} + u_N - u_{N-1}} - \frac{1}{\mathbf{a}_N - u_N}\Big)\,.$$

Dans le cas particulier où toutes les bulles ont, à l'instant initial, même volume, ce qui ici signifie même longueur, et contiennent le même nombre de molécules $M_n = \beta\mathbf{h}$ avec $\mathbf{h} = \mathbf{a} + \mathbf{b}$, l'expression des F_n se simplifie encore en

$$F_n(u,t) = \frac{\mathbf{k}\beta\mathbf{h}}{\rho_n\mathbf{b}}\Big(\frac{1}{\mathbf{a} + u_n - u_{n-1}} - \frac{1}{\mathbf{a} + u_{n+1} - u_n}\Big)\,, \tag{5}$$

$n = 1, ..., N$, toujours avec les conventions $u_{N+1} \equiv 0$ et $u_0 \equiv \mathbf{U}_0(t) < \mathbf{a}$.

Dans un milieu liquide homogène (tous les ρ_n, $n = 0, ..., N+1$, sont égaux à ρ) et lorsque les $\mathbf{a}_n$ sont tous égaux à $\mathbf{a}$, comme les $\mathbf{M}_n$ à M, nous verrons (à la section suivante) que le phénomène généré par une perturbation $\mathbf{U}_0$ constante dans le temps de la position de S_0 consiste en la propagation d'une onde à vitesse constante ; on peut penser intuitivement que le passage de l'onde, en provoquant une agitation des molécules de gaz à l'intérieur de chaque bulle, provoque une élévation de la température ; si l'on fait l'hypothèse grossière (on la justifiera dans la section suivante) selon laquelle l'indice $n(t)$ du segment liquide où $u_{n(t)}(t)$ est maximal varie, tant que l'onde issue du segment S_0 ne s'est pas encore réfléchie au niveau du segment S_{N+1}, de manière sensiblement linéaire avec t, il est naturel, si l'on veut se placer dans le cadre adiabatique et non plus isotherme, de faire l'approximation suivant laquelle, dans l'intervalle de temps correspondant,

$$w_n = \kappa_0 - \kappa_1 n\mathbf{h}\,,$$

où les constantes κ_0 et κ_1 dépendent des caractéristiques du mélange, κ_0 étant liée à la vitesse de propagation de l'onde dans le milieu (plus la bulle est loin du point S_0 où l'onde a été initiée, plus la température à laquelle est soumise cette bulle diminue). Cette hypothèse nous permet d'expliciter encore dans ce cas l'équation d'évolution sous la forme d'un système différentiel du type (1), où cette fois

$$F_n(u,t) = \gamma\frac{\beta\ bfh}{\rho\mathbf{b}}\Big(\frac{\kappa_0 - \kappa_1(n-1)\mathbf{h}}{\mathbf{a} + u_n - u_{n-1}} - \frac{\kappa_0 - \kappa_1 n\mathbf{h}}{\mathbf{a} + u_{n+1} - u_n}\Big)\,, \tag{6}$$

$n = 1, ..., N$, toujours avec les deux conventions $u_{N+1} \equiv 0$ et $u_0 \equiv \mathbf{U}_0(t)$ avec $\|\mathbf{U}_0\|_\infty < \mathbf{a}$.

Notons que, même dans le cas du modèle le plus simple (en l'occurrence le cas où l'évolution est régie par le système différentiel résoluble (1), F étant

donné par (5)) le système est évidemment un système non linéaire. Si l'on regarde le système (1) dans ce cas, on voit que bien sûr le théorème de Cauchy-Lipschitz s'applique au voisinage de $(0, \underline{0}, \underline{0})$; plus précisément, si $t \in [0, T]$ et si

$$\begin{aligned} \sup_{2 \le j \le N-1} |u_{j+1} - u_j| &\le a/2 \\ \sup_{2 \le j \le N-1} |\tilde{u}_{j+1} - \tilde{u}_j| &\le a/2\,, \end{aligned}$$

alors, pour tout n entre 2 et $N-1$,

$$|F_n(u) - F_n(\tilde{u})| \le 16\,\frac{\mathbf{k}\beta\mathbf{h}}{\mathbf{a}^2} \sup_{1 \le j \le N} |u_j - \tilde{u}_j|\,;$$

de plus, si $\sup(|u_N|, |\tilde{u}_N|) \le a/2$ et $\sup(|u_1|, |\tilde{u}_1|) \le (a - \|\mathbf{U}_0\|_{[0,T]})/2$, alors, pour tout $t \in [0, T]$,

$$\|F(u,t) - F(\tilde{u},t)\| \le 16\,\frac{\mathbf{k}\beta\mathbf{h}}{(\mathbf{a} - \|\mathbf{U}_0\|_{[0,T]})^2}\,\|u - \tilde{u}\|\,,$$

(la norme $\|\cdot\|$ étant ici le sup des valeurs absolues des coordonnées), ce qui fournit une estimation de la constante de Lipschitz utile pour justifier la stabilité de la méthode d'Euler à un pas qui sera proposée au paragraphe suivant. Le système (1) admet (théorème de Cauchy-Lipschitz) une unique solution $t \mapsto (u(t), v(t))$ telle que :

$$u_1(0) = \cdots = u_N(0) = 0$$

et que le vecteur $(v_1(0), ..., v_N(0))$ soit égal à un vecteur vitesse "initial" $(v_1^{\text{init}}, ..., v_N^{\text{init}})$ donné. C'est cette solution que l'on va approcher dans la section suivante ; nous envisagerons aussi une perturbation des "conditions initiales" correspondant à la position d'équilibre où tous les $u_n(0)$ et tous les v_n^{init}, $n = 1, ..., N$, sont nuls (toujours en suivant l'évolution temporelle lorsque le $N+2$-ème segment liquide est fixe, tandis que le centre du premier est lui, assujetti à la variation donnée $t \mapsto \mathbf{U}_0(t)$ de norme uniforme petite devant la constante $\mathbf{a}$). Le théorème de Cauchy-Lipschitz assure encore l'existence d'une solution dont on souhaite (ce sera la clause de stabilité) qu'elle reste bornée dans le temps lorsque la perturbation des données à l'équilibre statique que sont $(\underline{0}, \underline{0})$ est assez petite.

§3. Le schéma numérique d'Euler et quelques variantes

Nous avons simulé numériquement dans cette section l'équation d'évolution (1) dans le cadre d'un milieu isotherme (les F_n sont alors donnés par

les relations (5) ou adiabatique (les F_n sont donnés par les relations (6)). Comme dans le texte, et faute de temps pour développer d'autres approches numériques telles que celles du type Runge-Kutta, c'est le modèle numérique d'Euler que nous avons illustré.

Dans le premier cas, ce modèle se traduit ainsi : le temps (considéré en abscisse) est échantillonné avec un pas Δt entre $t = 0$ et $t = M\,\Delta t$, le signal temporel $\mathbf{U}_0$ est modélisé par un vecteur ligne de longueur $M+1$ (noté encore $\mathbf{U}_0$ pour simplifier) dont la m-ème composante, $m = 1, ..., M+1$, correspond à la valeur de son échantillon en $(m-1)\Delta t$, tandis que les conditions initiales $u(0)$ et $v(0)$ sont précisées à partir de deux vecteurs colonnes $\mathbf{V}_0$ et $\mathbf{V}_{-1}$ selon les règles

$$\begin{aligned} u_n(0) &= \mathbf{V}_0(n)\,, \quad n = 1, ..., N \\ v_n(0) &= \frac{\mathbf{V}_0(n) - \mathbf{V}_{-1}(n)}{\Delta t}\,, \quad n = 1, ..., N\,. \end{aligned}$$

Le jeu d'équations récurrentes à résoudre, si le modèle numérique suivi est celui de la méthode d'Euler, où l'on a choisi de modéliser l'opérateur de dérivation continue par l'opérateur de dérivation discrète à gauche, est alors

$$\begin{aligned} U(n,m) &= 2U(n,m-1) - U(n,m-2) \\ &\quad + \frac{[\Delta t]^2\,\mu_n}{\mathbf{a} + U(n,m-1) - U(n-1,m-1)} \\ &\quad - \frac{[\Delta t]^2\,\mu_n}{\mathbf{a} + U(n+1,m-1) - U(n,m-1)}\,, \end{aligned}$$

avec $n = 1, ..., N$, $m = 1, ..., M+1$ (où $\mu_n = \mathbf{k}\beta\mathbf{h}/(\rho_n\mathbf{b})$) et les conditions initiales

$$U(0\,,\,:) = \mathbf{U}_0\ ,\ U(N+1\,,\,:) = (0, ..., 0)$$

et

$$U(:\,,\,0) = \mathbf{V}_0\ ,\ U(:\,,\,-1) = \mathbf{V}_{-1}\,.$$

On suppose que $(N+1)\mathbf{h} = L$ (distance entre le centre du segment S_0 à l'instant $t = 0$ et celui du segment S_{N+1}) et que α est le paramètre tel que $\mathbf{a} = \alpha\mathbf{h}$, soit $\mathbf{b} = (1-\alpha)\mathbf{h}$. Le programme (écrit sous **MATLAB**) est fourni en annexe.

Pour rendre compte du phénomène de propagation d'onde généré par cette équation d'évolution, nous avons choisi de l'illustrer en prenant comme données numériques $\Delta t = 10^{-4}$ (notons que ce pas de temps est plus précis que celui proposé dans l'illustration numérique du texte), $\alpha = .2$, $\rho_n = 1$ pour tout $n = 1, ..., N$, β, $\mathbf{k}$ et L étant tous normalisés à la valeur 1 ; nous avons pris, comme dans le texte, $N = 99$, ce qui reste raisonnable (bien que cela soit un peu limite car les bulles sont alors peu dispersées dans le mélange) à la fois pour supposer l'incompressibilité des segments liquides et pour avoir avec les valeurs $(n-1)\mathbf{h}$, $n = 1, ..., N+2$ un échantillonnage suffisamment

fin de l'intervalle spatial afin de pouvoir rendre compte (comme nous le ferons à la section 4) du fait que le phénomène peut, tant que les u_n restent en module assez petits, se présenter comme régi, en fonction des variables de temps et d'espace, par l'équation des ondes (2) en dimension spatiale un. Dans un premier temps, nous avons pris $\mathbf{V}_0(n) = \mathbf{V}_{-1}(n) = 0$, $n = 1, ..., N$, ce qui signifie que nous avons supposées nulles les conditions initiales à l'instant $t = 0$ dans le problème de Cauchy lié au système différentiel (1) ; par contre, pour voir apparaître un phénomène non trivial, nous avons prescrit une donnée $\mathbf{U}_0 = (\mathbf{h}, ..., \mathbf{h})/10$; comme $\mathbf{a} = .002$ et $\mathbf{h} = L/100 = .01$, on a bien $\|\mathbf{U}_0\|_\infty = .001 < \mathbf{a}$ comme convenu pour valider les hypothèses d'application du théorème de Cauchy-Lipschitz au moins au voisinage de la position d'équilibre statique. Sur la figure 1, nous avons représenté les valeurs de $u_0, ..., u_{N+1}$ aux instants $t = .12$, $t = .17$, $t = .18$, $t = .25$ afin de bien mettre en évidence le premier phénomène observé, à savoir un phénomène de propagation d'onde, l'onde se trouvant amortie, puis réfléchie au moment où elle atteint l'extrémité droite du segment spatial ; le phénomène se poursuit ainsi (la figure est à lire de haut en bas, puis de gauche à droite).

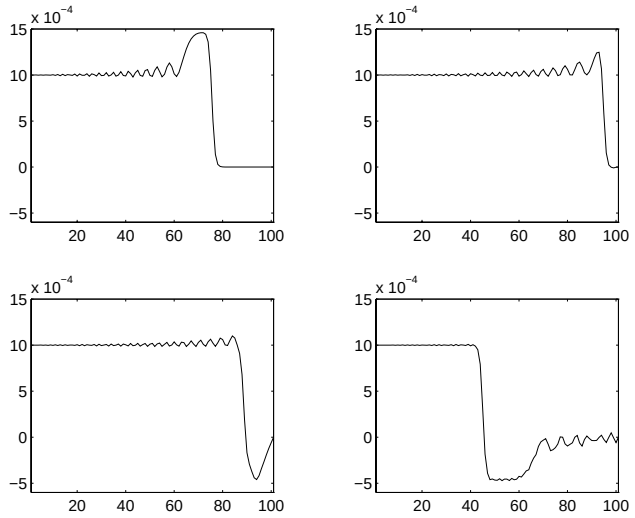


FIG. 1 – $U(:, t)$, $t = .12$, $t = .17$, $t = .18$, $t = .2$ sous conditions initiales $u(0) = v(0) = 0$ (équilibre statique)

Toujours en supposant $\mathbf{U}_0$ inchangé, perturbons les données initiales $u_1, ..., u_N$ en supposant cette fois (par exemple) que $\mathbf{V}_0 = 10^{-4}\,(1, ..., 1)$ et $\mathbf{V}_0 - \mathbf{V}_{-1} = .01\Delta t\,(1, ..., 1)$ (ce qui correspond à une perturbation uniforme de $u(0) = \underline{0}$ en $\tilde{u}(0)$ en 10^{-4} et de $v(0) = \underline{0}$ en $\tilde{v}(0)$ en 10^{-2} (on imprime une pente initiale de .01 aux fonctions $u_1, ..., u_N$ en $t = 0$). On observe cette fois encore (figure

2, à lire comme la figure 1) un phénomène de propagation d'onde ; l'écart entre la nouvelle solution et la solution générée sous les conditions initiales correspondant à la position d'équilibre statique pour les segments $S_1, ..., S_N$ ($u(0) = v(0) = \underline{0}$) reste bornée, ce qui semblerait traduire (au moins qualitativement) la stabilité de cet état d'équilibre statique, et ce même en présence d'une perturbation beaucoup plus significative $\mathbf{U}_0$ pour le centre du segment S_0 ; nous reviendrons dans la section 5 sur cette question. Si l'on imprime une

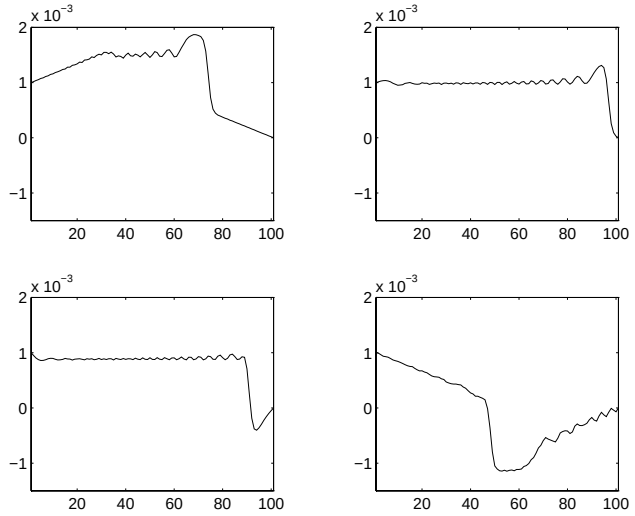


FIG. 2 – $U(:,t)$, $t = .12$, $t = .17$, $t = .18$, $t = .2$ sous conditions initiales perturbées (vitesse initiale positive)

pente initiale de $-.01$ aux fonctions $u_1, ..., u_N$ par rapport à la position d'équilibre statique, l'évolution du phénomène est identique, comme le reflètent les instantanés visualisés sur la figure 3.

Afin de conforter l'idée selon laquelle le phénomène décrit par l'équation d'évolution en milieu isotherme correspond à un phénomène de propagation d'onde à une vitesse en relation avec le milieu fluide, nous nous sommes placés dans le cadre d'un milieu non-homogène, où cohabitent 101 segments liquides de deux types différents, les 51 premiers ayant une masse linéique triple de la masse linéique des 50 premiers, valant, elle, 1. Sur la figure 4 se trouve l'image obtenue comme le module du laplacien de l'image discrète

$$(n,m) \mapsto U(n,m)\,, \quad n = 1, ..., 101,\ m = 0, ..., 5000$$

lorsque $M = 4999$, $(u(0), v(0)) = (\underline{0}, \underline{0})$, $\mathbf{U}_0 = 10^{-3}(1, ..., 1)$, $\Delta t = 10^{-4}$ (l'évolution du phénomène a donc été suivie pendant la durée $T = .5$), $\alpha = .2$, tous les autres paramètres $\mathbf{k}$, β, L, étant normalisés à la valeur 1) ; Le laplacien,

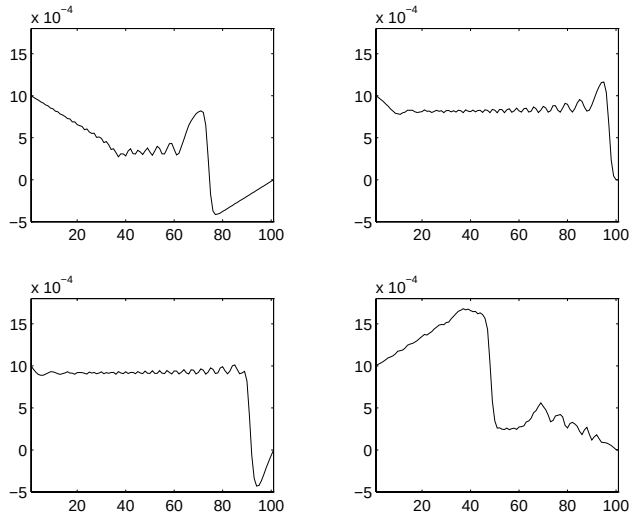


FIG. 3 – $U(:,t)$, $t = .12$, $t = .17$, $t = .18$, $t = .2$ sous conditions initiales perturbées (vitesse initiale négative)

de par la formule de Green-Ostrograski, permet en effet de mettre en évidence les "lignes de contraste" d'une image 2D, ce que nous entendons exploiter ici. À la lecture de cette image (le temps est en abscisse, le numéro des segments fluides en ordonnée), on constate à la fois les phénomènes de propagation d'onde, de réflexion, et le fait que la vitesse de propagation reste uniforme dans chacun des deux milieux. Sur la figure 5, on a échangé l'ordre des deux milieux ; les lignes brisées du tracé rendent compte de ce phénomène. Le calcul des diverses pentes montre qu'elles sont proportionnelles à la racine carrée de la masse linéique du milieu, ce que nous validerons plus rigoureusement dans la section suivante. Il est à noter qu'un certain nombre de phénomènes d'interférence (dont on peut se demander s'ils correspondent à une réalité physique ou plutôt à des avatars numériques) apparaissent après la première réflexion au moment du changement de milieu.

Le phénomène de réflexion est aussi intéressant à analyser lorsque seulement l'un des segments liquides a une masse linéique différente (disons par exemple ici double) de tous les autres, qui, elle, est normalisée à la valeur 1 (pour reprendre l'exemple suggéré dans le texte) ; cet "obstacle" que constitue ce segment génère, au passage de l'onde principale, une onde "réfléchie" partant en sens inverse de la propagation de la première ; c'est ce qu'illustre la figure 6, ou nous avons, toujours en supposant que $u_1, ..., u_N$ partent de la position d'équilibre statique et que $\mathbf{U}_0$ est uniforme, de valeur .001 tandis que $\mathbf{a} = .002$ (mêmes paramètres $\mathbf{k} = L = 1$ et toujours choix de $\Delta t = 10^{-4}$), suivi le déplacement de l'onde, l'apparition de l'onde réfléchie et son évolution.

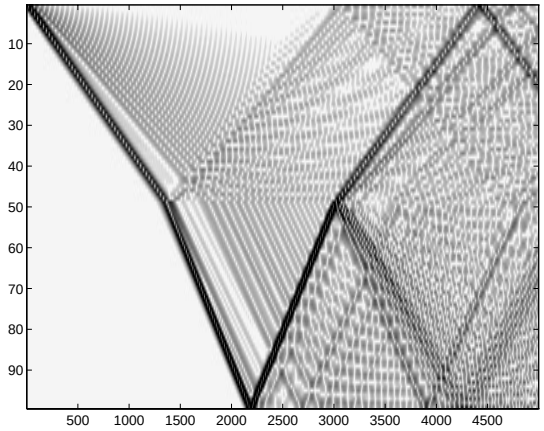


FIG. 4 – Module du laplacien de U en milieu non-homogène (1)

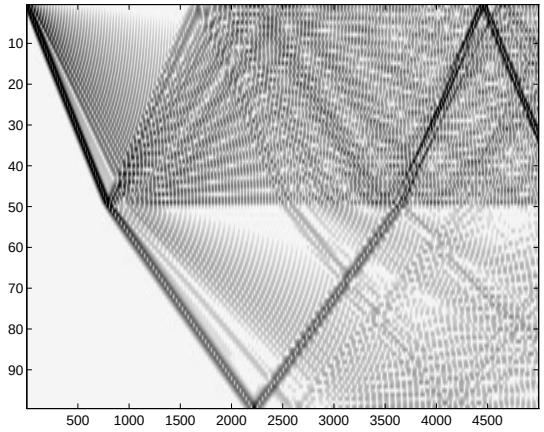


FIG. 5 – Module du laplacien de U en milieu non-homogène (2)

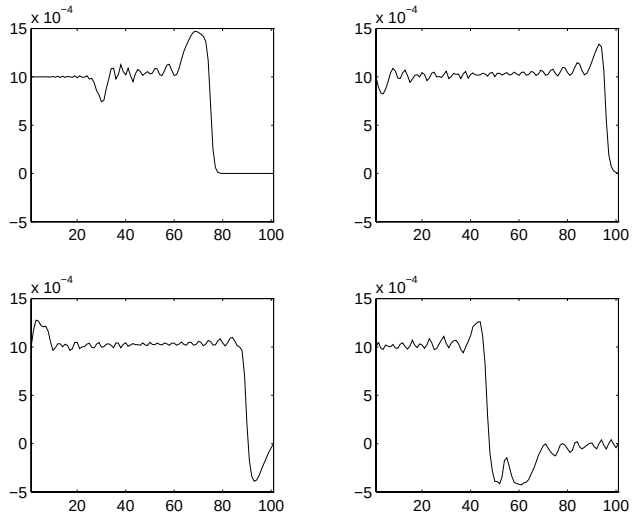


FIG. 6 – $U(:,t)$, $t = .12$, $t = .17$, $t = .18$, $t = .2$, obstacle en $n = 50$

Le laplacien de l'image (évolution temporelle de 0 à .5), dont l'intensité du module apparaît sur la figure 7, illustre aussi les divers phénomènes que nous venons d'évoquer.

§4. Validation de l'approche asymptotique

Dans le cadre d'un milieu isotherme (où l'équation d'évolution a été résolue numériquement dans la section 3 tout au moins sous l'hypothèse que toutes les bulles aient même diamètre **a** en position d'équilibre statique), l'approche asymptotique décrite dans la section 2 du texte proposé et reposant sur le fait que les écarts u_n, $n = 1, ..., N$ restent petits devant **h** devrait se trouver validée (tout au moins numériquement), puisque, dans les divers tests numériques proposés, tandis que $\mathbf{h} = .01$, le maximum des $|u_n|$ évolue toujours entre $-.002$ et $.002$ (voir les diverses figures présentées dans la section précédente). Évoluant dans le milieu fluide de masse linéique $\rho = 1$, l'onde devrait se plier à l'équation des ondes (2) avec ici (d'après le développement proposé dans la section 2 du texte)

$$c^2 = \frac{\mathbf{k}\beta}{\rho(1-\alpha)\alpha^2} = \frac{1}{(1-\alpha)\alpha^2},$$

ce qui, avec notre donnée $\alpha = .2$, correspond à $c^2 = 31.25$. Non seulement ce fait peut être validé, mais il est intéressant de noter qu'il peut l'être encore

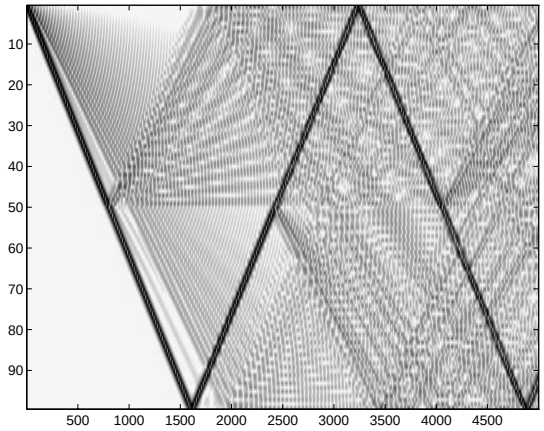


FIG. 7 – Module du laplacien de U en présence d'un obstacle en $n = 50$

dans un milieu non-homogène, ce que nous avons tenté de faire sous l'hypothèse où, sur les 101 segments liquides, les 50 premiers auraient par exemple une masse linéique triple de celle des autres (normalisée, elle, à 1). Si l'on saisit un instantané de l'onde au temps $T = .18$ (soit peu de temps avant la première réflection en L), on constate (figure 8) que les versions discrètes des fonctions $\partial_t^2 u$ (en trait plein) et de $c^2\partial_x^2 u$ (en pointillé) coïncident sur la portion d'espace correspondant au milieu liquide de masse linéique $\rho = 1$ (valeurs de n entre 51 et 100), mais non dans la partie gauche de l'espace (valeurs de n entre 1 et 50) ; les seuls "décrochages" apparaissant dans cette partie droite de l'image se situent près du segment liquide à la jonction des deux milieux ou dans les zones où ces dérivées sont maximales (notons que la position des extrémums est cependant la même pour les deux fonctions. Il apparaît le même phénomène si l'on confronte (figure 9) les traçés de $\partial_t^2 u$ (en trait plein) et $c^2\partial_x^2 u/3$ (en pointillés) à gauche de la jonction $n = 50$, c'est à dire dans la zone d'espace correspondant au milieu liquide de masse linéique précisément $\rho' = 3\rho = 3$.

La conclusion pratique que l'on peut retenir de cette étude qualitative est qu'il est possible d'extraire (de l'étude de la solution numérique du problème d'évolution) des informations relatives à la composition du milieu fluide (milieu homogène, juxtaposition de deux milieux homogènes de masses linéiques différentes,...) et aux valeurs des paramètres (en l'occurrence ici la masse linéique) des divers composants du milieu fluide.

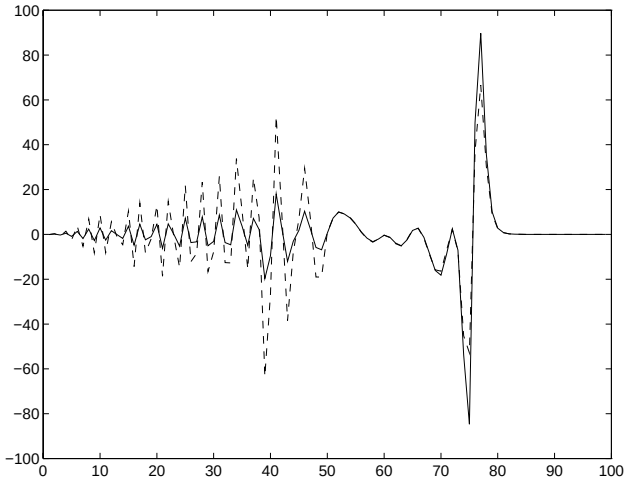


FIG. 8 – $\partial_t^2 u$ (en $-$) et $c^2\partial_x^2 u$ (en $- -$) en $T = .18$

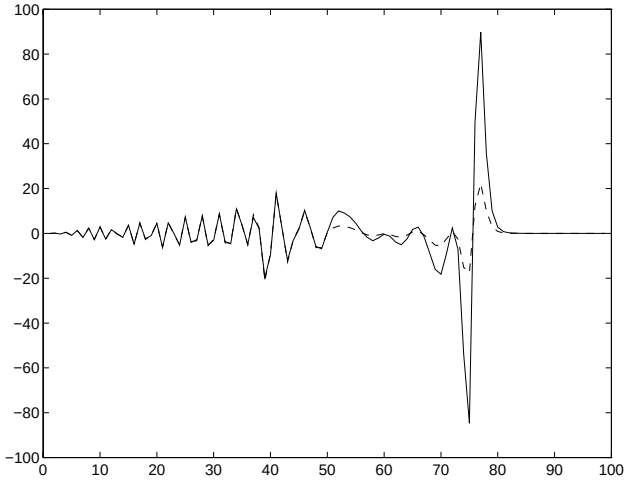


FIG. 9 – $\partial_t^2 u$ (en $-$) et $c^2\partial_x^2 u/3$ (en $- -$) en $T = .18$

§5. Évolution en milieu adiabatique ou en présence d'un terme de dissipation

Dans cette section, nous étudions l'équation d'évolution (1) en milieu homogène, tous les segments liquides étant de même longueur $\mathbf{b}$, toutes les bulles de gaz ayant même diamètre $\mathbf{a}$ en position d'équilibre statique ($u(0) = v(0) = 0$). On travaillera toujours avec $N = 99$, $L = 1$, $\mathbf{a} = .002$, $\mathbf{b} = .008$, $\rho = 1$. La perturbation $\mathbf{U}_0$ sera supposée constante dans le temps, avec de toutes façons $|\mathbf{U}_0| < \mathbf{a}$.

Si la température n'est plus constante, mais que l'on suppose le milieu adiabatique (pas d'échanges de chaleur par diffusion thermique) le fait que l'augmentation de température soit liée au passage de l'onde nous autorise à remplacer la constante $\mathbf{k}$ intervenant dans le cadre homogène (système (5)) par $\mathbf{k}_{init} - \kappa_1 n\mathbf{h}$; on supposera donc que $\mathbf{k}$ varie linéairement en espace de la valeur initiale $\mathbf{k}_{\text{init}}$ jusqu'à la valeur finale $\mathbf{k}_{N+1} = 1$. La figure 10 représente l'évolution d'une onde constante (égale à $\mathbf{U}_0 = .001$) impulsée au segment S_0 à partir de la position d'équilibre statique pour les segments $S_1, ..., S_N$ (on a pris ici $\mathbf{k}_{\text{init}} = 2$). On peut remarquer nuériquement que l'équation (2) avec $c(x)$ variant linéairement (comme $\mathbf{k}$) en fonction de la variable x reste approximativement satisfaite, ce qui laisse penser que l'approche asymptotique proposée dans la section 2 du texte pourrait être dans ce cas reprise, bien que plus délicate ; le fait que $\mathbf{k}$ varie avec une pente faible en fonction de la variable d'espace x permet ici de supposer que la fonction c est localement constante, ce qui pourrait expliquer le fait que l'évolution de l'onde se plie encore à l'équation des ondes à coefficients fonctions cette fois affines (et non plus constants par morceaux comme c'était le cas dans le contexte des milieux fluides non homogènes).

Pour clôre cette section, revenons sur les questions de stabilité de la position d'équilibre $u(0) = v(0) = 0$ en milieu isotherme (les $N+2$ bulles ayant même diamètre $\mathbf{a}$ à cet état d'équilibre, tous les segments liquides ayant eux même diamètre $\mathbf{b}$), l'onde perturbatrice impulsée au niveau du segment initial S_0 étant une onde constante, identiquement égale à $\mathbf{U_0}$ avec $|\mathbf{U}_0| < \mathbf{a}$. L'équation d'évolution est alors du type

$$X' = \Phi(X)$$

avec $X = (u, v)$, la fonction Φ (dépendant de la constante $\mathbf{U}_0$) étant une fonction de classe C^∞ au voisinage de $(\underline{0}, \underline{0})$, à valeurs dans $\mathbf{R}^{2N}$. La linéarisation du système (au voisinage de la position d'équilibre statique $(t, X) = (0, \underline{0}, \underline{0})$) conduit au système $X' = \Psi X$ où la matrice Ψ est la matrice

$$\begin{pmatrix} 0_N & I_N \\ A & 0_N \end{pmatrix}$$

avec

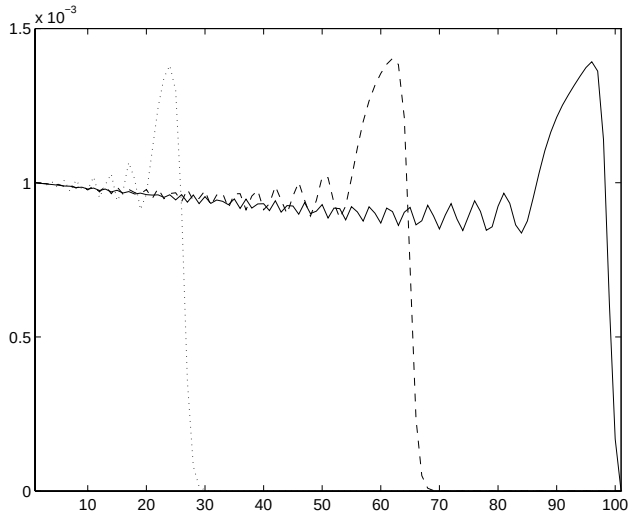


FIG. 10 – $U(:,t)$, $t = .03$, $t = .08$, $t = .13$, en milieu adiabatique et sous conditions initiales d'équilibre statique pour $S_1, ..., S_N$

$$A = \frac{\mathbf{k}\beta\mathbf{h}}{\rho\mathbf{b}\mathbf{a}^2}\begin{pmatrix} -1-\theta & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \vdots & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

et $\theta = \mathbf{a}^2/(\mathbf{a}-\mathbf{U}_0)^2$; en présence d'un terme dissipatif de paramètre κ (comme cela est suggéré parmi les développements proposés), il convient de remplacer (si l'on linéarise la nouvelle équation d'évolution) la matrice Ψ par la matrice

$$\Psi_\kappa = \begin{pmatrix} 0_N & I_N \\ A & -\frac{\kappa}{\rho\mathbf{b}} I_N \end{pmatrix}.$$

Une approche numérique (conduite avec la routine **eig** de l'environnement **MATLAB**) montre que les valeurs propres de la matrice Ψ sont imaginaires pures (un calcul permettrait aussi de le justifier *a posteriori* théoriquement car le polynôme caractéristique de Ψ s'écrit $P(X^2)$, les racines de P étant strictement négatives), ce qui empêche de ramener dans ce cas l'étude de la stabilité (au sens de Liapounov) de la position d'équilibre à celle de cette même position vis-à-vis du système linéarisé. Une étude plus soigneuse (faisant intervenir le développement de Taylor au second ordre de l'application Φ en $(\underline{0},\underline{0})$) demanderait à être conduite ; notons toutefois qu'il n'existe pas à notre connaissance de critère exhaustif de stabilité au sens de Liapounov.

En revanche, dès que le paramètre de dissipation κ est strictement positif, on voit numériquement, par exemple en utilisant la routine **eig** sous l'environnement **MATLAB**, que toutes les valeurs propres de Ψ_κ sont de partie réelle strictement négative, ce qui implique (voir par exemple [Arn] V. Arnold, *Équations différentielles ordinaires*, Éditions MIR, Moscou, 1974, chapitre 3, paragraphe 23) la stabilité asymptotique de la position d'équilibre statique (pour le système linéarisé), donc aussi la stabilité asymtotique de la position d'équilibre statique pour le système

$$u' = v\,, \quad v' = F(u) - \frac{\kappa}{\rho \mathbf{b}} v$$

(voir par exemple [Arn], 3-23). Sur la figure 11, nous avons représenté l'évolution de l'onde générée par une perturbation constante affectant la position de X_0 ($\mathbf{U_0} = .001$) lorsque les données initiales $u(0)$ et $v(0)$ sont deux vecteurs de composantes toutes égales respectivement à .0005 et .1 ; le paramètre κ a été choisi égal à .285 de manière à ce que (lorsque $\mathbf{a} = .002$, $\mathbf{b} = .008$, $\mathbf{k} = L = \beta = 1$) le maximum des parties réelles des valeurs propres de la matrice Ψ_κ soit le plus négatif possible. On voit nettement, en suivant l'évolution (entre $t = 0$ et $t = .5$) de u_{20}, u_{40}, u_{60}, u_{80} (en pointillés) sous les conditions initiales $(u(0), v(0))$, confrontée avec celle des mêmes variables sous l'hypothèse initiale ($u(0) = v(0) = \underline{0}$) d'équilibre statique (en trait plein), le phénomène de stabilité asymptotique.

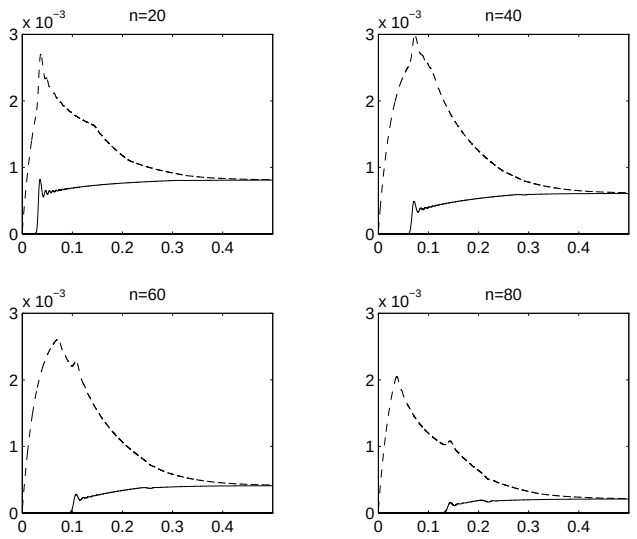


FIG. 11 – Validation numérique de la stabilité asymptotique de la position d'équilibre statique en présence d'un terme de dissipation

Il est intéressant de confronter les résultats obtenus provenant de l'étude conduite avec ceux obtenus avec les mêmes données numériques, mais en l'absence du terme de dissipation ($\kappa = 0$) ; les résultats obtenus numériquement (voir la figure 12, où l'on a représenté en traits pleins l'évolution sous conditions initiales d'équilibre statique, et en pointillés l'évolution sous les conditions initiales perturbées, semblent indiquer que la position d'équilibre statique est dans ce cas stable (mais non asymptotiquement stable) au sens de Liapounov ; la raison majeure en est probablement le fait que le centre du segment liquide S_{N+1} soit considéré comme "gelé".

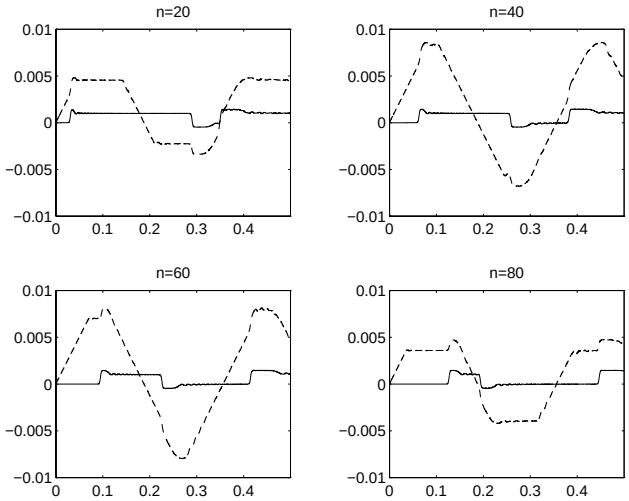


FIG. 12 – Validation numérique de la stabilité (?) de la position d'équilibre statique lorsque $\kappa = 0$

§6. En guise de conclusion

La conclusion majeure que l'on puisse extraire de l'étude de ce modèle est que le résultat de l'étude d'une modélisation (même simpliste) d'un phénomène physique permet déjà d'en extraire un certain nombre de paramètres (homogénéité ou non-homogénéité du milieu, calculs des masses linéiques des fluides en jeu, distinction milieu isotherme – milieu adiabatique, etc...) Il est clair cependant qu'ici les vrais problèmes (de nature plus géométrique cette fois, en relation par exemple avec la théorie des surfaces minimales ou la géométrie des ensembles convexes) se posent plutôt en dimension 2 ou 3 lorsque les forces de tension surfacique sont appelées à intervenir. Néanmoins, ce modèle montre la capacité de l'analyse numérique à nous guider (de manière

expérimentale) vers une compréhension (puis au delà, une modélisation plus poussée intégrant cette compréhension) du phénomène physique décrit par le jeu d'équations d'évolution proposé.

§7. Annexe (calcul de $U, \partial_t^2 U, \partial_x^2 U$)

```
h=L/(N+1);
a=alpha*h;
b=(1-alpha)*h;
mu=k*beta*h/(rho*b);
param=mu*((delta)^2);
VVinit=[0;VVinit; 0];
Vinit=[Uinit(1);Vinit; 0];
W=[VVinit Vinit];
for k=3:M+2
   V=zeros(N+2,1);
   V(1)=Uinit(k-2);
   for l=2:N+1
      xi=1/(a+W(l,k-1)-W(l-1,k-1));
      eta=1/(a+W(l+1,k-1)-W(l,k-1));
      V(l)=2*W(l,k-1)-W(l,k-2)+param*(xi-eta);
   end;
   W=[W V];
end;
u=W(:,2:M+2);
u_tt=[];
u_xx=[];
for k=2:M+1
   X=(W(:,k+1) + W(:,k-1) -2*W(:,k))/(delta^2);
   u_tt=[u_tt X];
end;
for n=2:N+1
   Y=(W(n+1,:)+W(n-1,:)-2*W(n,:))/(h^2);
   u_xx=[u_xx ; Y];
end;
u_tt=u_tt(2:N+1,:);
u_xx=u_xx(:,1:M);
```

Chapitre 6
modèle simplifié de mélange liquide-gaz

§1. Modélisation

1.1. Modèle physique. On cherche à modéliser la propagation d'ondes dans un mélange liquide-gaz.On suppose que le liquide est incompressible, et que le gaz , compressible, se trouve dans le liquide sous forme de petites bulles qui contiennent le même nombre de molécules.On s'interessera ici à une modélisation mono-dimensionnelle de cette situation : on considère sur le segment $[0, L]$,une alternance de gaz et de liquide, comme l'illustre la figure 1.
ftbpF5.3774in1.9882in0ptpbb.epsOn note ρ la masse linéique du liquide, de telle sorte que la masse d'un "segment" S_i centré en x_i est ρb.Le gaz piégé entre deux segments successifs est supposé obéir à la loi des gaz parfaits à température constante, c'est à dire

$$(\text{pression}) \times (\text{volume}) = \text{constante} \times (nbre \text{ de molécules de gaz})$$

soit

$$PV = kN \tag{1}$$

où le "volume" représente en fait une longueur dans le cadre de ce modèle mono-dimensionnel. On suppose les deux extrémités fixes (x_0 et x_{n+1} positionnés en 0 et L respectivement, sont fixes), et on prend un nombre constant de molécules de gaz dans chaque poche, nombre égal à N. On considèrera que N s'écrit $\beta(a+b)$, où β est le nombre de molécules de gaz par unité de longueur du mélange. Ce système liquidegaz admet un état d'équilibre représenté shématiquement sur la figure 1, pour lequel les bulles ont même taille b.Les positions des centres des segments fluides sont (en notant $h = a + b$)

$$X_0 = 0,\ X_1 = h,\ ..., X_i = ih,\ ...X_{n+1} = (n+1)h = L \tag{2}$$

1.2. Equation d'évolution. On note $x_i = X_i + u_i$ la position du centre de S_i (i^{ème}segment) , de telle sorte que l'état d'équilibre correspond à $u_i = 0$ pour tout i. Le bilan F_i des forces s'exerçant sur un élément centré en x_i est la somme des pressions exercées à droite et à gauche par le gaz, c'est à dire d'après le modèle (1) :

$$F_i = k\beta h\left(\frac{1}{a+u_i-u_{i-1}} - \frac{1}{a+u_{i+1}-u_i}\right) \tag{3}$$

La masse de S_i étant ρb, la relation fondamentale de la dynamique s'écrit donc

$$\rho b\ddot{u}_i = F_i \tag{4}$$

.On note $u = (u_1, \ldots, u_n)$ le vecteur des inconnues, $v_i = \dot{u}_i$ la vitesse de S_i, v le vecteur des v_i et F le vecteur des forces F_i. On peut se ramener de cette façon à une équation différentielle ordinaire dans $\mathbb{R}^{2n}$:

$$\left\{\begin{array}{l} \dfrac{du}{dt} = v \\ \rho b\dfrac{dv}{dt} = F(u) \end{array}\right. \tag{5}$$

§2. Vitesse de propagation

On cherche à estimer la vitesse de propagation associée au modèle décrit ci–dessus, pour des ondes de faible amplitude ($u_i << a+b$). La force F_i peut s'exprimer

$$\begin{aligned} F_i &= k\beta h\left(\frac{1}{a+u_i-u_{i-1}} - \frac{1}{a+u_{i+1}-u_i}\right) \\ &= k\beta h\frac{u_{i+1}-2u_i+u_{i-1}}{(a+u_i-u_{i-1})(a+u_{i+1}-u_i)} \\ &= k\beta h\,\frac{u_{i+1}-2u_i+u_{i-1}}{h^2}\,\frac{1}{\left(\alpha+\dfrac{u_i-u_{i-1}}{h}\right)\left(\alpha+\dfrac{u_{i+1}-u_i}{h}\right)}, \end{aligned}$$

avec $h = a+b$ comme précédemment, et $\alpha = a/(a+b)$. On considère maintenant que les u_i sont les valeurs en X_i d'une fonction u de classe $\mathcal{C}^2$, et l'on identifie les expressions faisant intervenir les u_i à des dérivées en espace de u :

$$\frac{u_{i+1}-2u_i+u_{i-1}}{h^2} \approx u''(X_i) \quad , \quad \frac{u_i-u_{i-1}}{h} \approx \frac{u_{i+1}-u_i}{h} \approx u'(X_i),$$

où u' (respectivement u'') est la dérivée (respectivement la dérivée seconde) partielle par rapport à la variable d'espace. On aboutit donc, formellement, à l'équation aux dérivées partielles sur u (en remarquant que $b/h = 1-\alpha$)

$$\rho(1-\alpha)\ddot{u} - \frac{k\beta}{(\alpha+u')^2}u'' = 0. \tag{6}$$

Si l'on fait l'hypothèse supplémentaire que u' reste petit devant α, on obtient, toujours formellement, une équation des ondes linéaire

$$\ddot{u} - c^2u'' = 0 \text{ avec } c^2 = \frac{k\beta}{\rho(1-\alpha)\alpha^2}, \tag{7}$$

à laquelle sont associées les vitesses de propagation $+c$ et $-c$.

§3. Simulation numérique

3.1. Discrétisation. On se propose de résoudre numériquement le système (3) pour modéliser la propagation d'une onde dans le milieu. Pour perturber le système, on va agir sur la position du point x_0, qu'on ne supposera donc plus fixe, mais dont la position va être imposée au cours du temps. On introduit une discrétisation de l'intervalle en temps $[0, T]$:

$$t_0 = 0 < t_1 = \delta t < t_2 = 2\delta t < \cdots < t_M = M\delta t = T,$$

et on note u_n^m (respectivement v_n^m) l'approximation de u_n (respectivement v_n) au temps t_m. On discrétise le système (3) de la façon suivante :

$$\left\{ \begin{array}{rcl} \dfrac{u_n^m - u_n^{m-1}}{\delta t} & = & v_n^m \\ \rho b \dfrac{v_n^{m+1} - v_n^m}{\delta t} & = & k\beta h \left(\dfrac{1}{a + u_n^m - u_{n-1}^m} - \dfrac{1}{a + u_{n+1}^m - u_n^m} \right) \end{array} \right. \tag{8}$$

ce qu'on peut écrire en éliminant les v :

$$u_n^{m+1} = 2u_n^m - u_n^{m-1} + \delta t^2 \mu \left(\frac{1}{a + u_n^m - u_{n-1}^m} - \frac{1}{a + u_{n+1}^m - u_n^m} \right). \tag{9}$$

avec $\mu = k\beta h/\rho b$. Les conditions aux limites en espace s'écrivent

$$u_0^m = \lambda(t)\ ,\ u_N^m = 0 \quad \forall m = 0, \ldots, M \tag{10}$$

où $\lambda(t)$ est donnée, et les conditions initiales en temps

$$u_n^0 = U_n^0\ ,\ u_n^{-1} = U_n^{-1}\ , \qquad \forall n = 1, \ldots, N \tag{11}$$

où les U_n^0 et U_n^{-1} sont des données qui permettent de représenter l'état du système au temps initial.

Les relations (9) et les conditions (10) et(11) permettent donc de calculer les u_n^m pour $n = 1, \ldots, N$, et $m = 1, \ldots, M$.

3.2. Programme de calcul en maple. On présente un programme de calcul basé sur les paramètres suivants

$$\rho = 1\ ,\ N = 99\ ,\ L = 1.0 \quad (\Rightarrow h = 10^{-2})\ ,\ \alpha = 0.2\ ,$$

$$\beta = 0.2\ ,\ k = 1.0\ ,\ \delta t = 10^{-2}\ ,\ .$$

La perturbation λ est simplement $\lambda(t) =$ constante $= h/10$, et les conditions initiales correspondent à l'état d'équilibre statique $U_n^0 = U_n^{-1} = 0$ pour tout $n = 1, \ldots, N$.

On considère d'abord le milieu homogène :. On calcule l'approximation des u_i jusqu'au temps $T = 6$.On remarquera , au temps $T = 3$, que l'onde provoquée par la déformation constante $h/10$ au point x_0 n'est pas encore parvenue au point x_{100}.Ce point ne bouge pas (u_{100} =constante = 0) On observe la propagation de cette onde en divers temps et en particulier jusqu'à ce qu'il y ait "retour" de l'information .
On peut en particulier vérifier sur cette simulation que la vitesse de propagation reste constante $\simeq$ $c = 0.25$, que le maximun des u_i reste petit devant $a + b$,et que les bulles de gaz restent disjointes .

Dans un second temps le milieu n'est plus homogène :. Il s'agit d'étudier la possibilité d'utiliser le phénomène de propagation mis en évidence pour détecter la présence dans un mélange fluide–bulles de matériaux plus lourds que l'eau, et si possible les localiser.On va modéliser numériquement une situation parfaitement analogue à la précédente, dans le cas où l'un des segments incompressibles de masse linéique ρ est remplacé par un segment de masse linéique $\rho' > \rho$. On choisit de modifier la masse volumique du segment correspondant à u_{50} (au milieu de la zone modélisée), et on prend $\rho' = 2\rho$.
On calcule comme précédemment l'approximation des u_i au temps $T = 3$. On remarque la présence (sur la moitié gauche) d'une onde réfléchie du fait de la présence de la partie plus lourde au centre de la configuration.On observe la propagation de ces ondes en plusieurs temps .

Troisième partie

Logarithmes discrets et cryptographie

Chapitre 7
Logarithme discret et cryptographie

Indications aux candidats

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Les candidat(e)s sont laissé(e)s libres d'organiser leur discussion comme ils ou elles l'entendent.

Des suggestions de développement, largement indépendantes les une des autres, sont proposées, le candidat n'est pas tenu de les suivre. Le jury appréciera que cette discussion soit accompagnée d'exemples effectivement traités sur ordinateur.

§1. Introduction

Il y a à peine quelques années, le problème de la sécurité des transmissions de données semblait être l'apanage des seuls militaires. Ce n'est plus le cas, avec l'essor des techniques numériques dans le commerce (Internet, les cartes de crédit, les télécommunications, les décodeurs de télévision, etc.)

Les méthodes empiriques traditionnelles se sont révélées trop fragiles ; elles reposaient souvent sur le schéma suivant : on choisit un livre, une fois pour toutes, et on considère la permutation des vingt-six lettres de l'alphabet définie par les vingt-six premières lettres de ce livre.

Le codage d'un texte consiste alors à appliquer cette permutation σ au texte à coder, et le décodage à appliquer la permutation réciproque σ^{-1} au texte à décoder.

En numérique, si par exemple le message à transmettre est codé sur 64 bits, on peut employer cette technique en considérant une permutation $\sigma \in S_{64}$. C'est ce genre d'idées qui est sous-jacent au procédé DES, dû à IBM, et qui est encore le plus utilisé en informatique.

Le talon d'Achille de ce genre de cryptosytème est le suivant : si quelqu'un sait coder, il sait aussi décoder, car il est très facile de trouver la permutation réciproque d'une permutation sur 26, 64, 128 ou même 256 lettres.

C'est pourquoi l'apparition de cryptosystèmes dits *à clé publique*, à la fin du siècle dernier, a fait sensation.

Ces cryptosystèmes, comme le nom l'indique, sont tels que le *codage* est public : tout le monde connaît l'algorithme pour coder le message. Mais on ne peut pas en déduire le décodage.

En fait, cela revient à construire une permutation σ d'un ensemble à N éléments, mais ici N est gigantesque (de l'ordre de 10^{500}, par exemple.) On ne peut même pas écrire la liste de ces éléments, et la méthode habituelle

pour trouver la permutation réciproque d'une permutation donnée ne peut plus s'appliquer.

Dans le paragraphe suivant, nous décrivons le cryptosystème le plus célèbre de ce genre : la méthode RSA, du nom des ses inventeurs Rivest, Shamir et Adelman.

§2. RSA

Cette méthode est basée sur le théorème suivant :

Théorème.- *Soit G un groupe fini d'ordre N et m un entier premier à N. L'application $f_m : G \to G$ définie par $x \mapsto x^m$ est bijective, et son application réciproque est l'application f_k définie par $x \mapsto x^k$, où k est l'inverse de m dans le groupe multiplicatif $(\mathbf{Z}/N\mathbf{Z})^*$.*

Si l'on connaît N et m, on en déduit donc rapidement k, par exemple par l'algorithme d'Euclide étendu, qui permet d'obtenir en $O(\log N)$ opérations élémentaires des entiers k et l tels que $mk + lN = 1$.

Par ailleurs, pour $x \in G$, le calcul de x^k se fait lui aussi rapidement, en $O(\log k)$ opérations dans le groupe G, en écrivant k en binaire. L'algorithme RSA consiste à fournir au codeur un groupe G dans lequel il sait calculer, mais dont il ne connaît pas l'ordre N, et un certain entier m. Le codage consiste alors, à partir d'un message x identifié à un élément de G, à calculer x^m. Le décodage nécessite de connaître k, inverse de m dans $(\mathbf{Z}/N\mathbf{Z})^*$. Dans les cas utilisés dans la pratique, on ne sait pas comment trouver k sans connaître N.

Donnons l'exemple habituel de groupe G utilisé : soient deux nombres premiers p et q, et soient $M = pq$, et $G = (\mathbf{Z}/M\mathbf{Z})^*$. Il suffit de connaître M pour savoir calculer dans G (et rapidement : les additions et multiplications se font en $O(\log^2 M)$ opérations), mais on ne connaît l'ordre de G, à savoir $(p-1)(q-1)$ que si l'on connaît p et q, c'est-à-dire la décomposition en facteurs premiers de M.

On ne peut bien sûr pas choisir $m = 2$ pour coder. Par contre, si $m = 3$, et si l'on a choisi p et q tels que $p \equiv q \equiv 2 \mod 3$, $x \mapsto x^m$ est bien une permutation de G, et on trouve immédiatement $k = \dfrac{2(p-1)(q-1)+1}{3}$.

La force de la méthode RSA réside dans le fait qu'à l'heure actuelle on ne connaît pas de méthode rapide de factorisation d'un entier : il faut actuellement des mois et plusieurs centaines de stations de travail travaillant en parallèle pour factoriser un entier de l'ordre de 150 chiffres. **Quelques remarques.**

- Supposons que le mot à coder soit un entier $x < \sqrt[3]{M}$. On retrouve alors rapidement x à partir de x^3, sans rien savoir de la factorisation de M.
- Supposons que l'on ait un même entier x à coder pour, par exemple, trois destinataires différents. On pourrait penser qu'il est plus prudent

de considérer trois entiers M_1, M_2 et M_3 distincts (et $> x$) pour le codage RSA. En fait, il n'en est rien : une personne mal intentionnée qui connaît les trois messages codés $x^3 \mod M_1, x^3 \mod M_2$ et $x^3 \mod M_3$ reconstitue immédiatement x par le théorème chinois !

– Au lieu de prendre $m = 3$, on peut prendre tout m premier à $(p-1)(q-1)$. L'intérêt de prendre m petit est que le codage est d'autant plus rapide, mais k est alors grand, et le décodage est plus lent. Par ailleurs, comme les deux remarques précédentes le montrent, m petit a des effets pervers. C'est pourquoi il a été proposé de prendre d'autres valeurs de m, même très grandes, mais pour lesquelles le codage reste rapide, par exemple $m = 65537$.

§3. Signatures sécurisées

Un problème voisin du précédent consiste à s'assurer de l'identité d'un interlocuteur : par exemple, si A introduit une carte bancaire dans un lecteur de cartes, celui-ci doit s'assurer que A est bien le possesseur de la carte en lui demandant son code secret. Ou encore : si A veut se connecter à un ordinateur, celui-ci s'assure de son identité en lui demandant son mot de passe. De même, un décodeur de télévision satellite ou cable s'assure de l'identité d'un utilisateur à l'aide d'un code secret, etc.

La méthode la plus utilisée est la suivante : sur la carte, figure le code secret *codé* par un algorithme à clé publique.

Lorsque le possesseur de la carte tape le nombre à quatre chiffres correspondant à son code sur le clavier du lecteur, le lecteur code ce nombre, et vérifie que le résultat est identique à celui qui est sur la carte.

De même, si, depuis chez lui, un utilisateur d'un réseau informatique veut se connecter sur l'un des ordinateurs du réseau, il envoie son mot de passe, l'ordinateur le code, et vérifie qu'il est bien égal au mot de passe codé qu'il possède.

Une alternative à ce type de codage est la suivante : supposons que A connaisse la clé publique P d'un algorithme à clé publique, et que B soit le seul à en connaître la clé privée S.

Bien sûr, $S \circ P$ est l'application identique, mais il en est de même de $P \circ S$. Par suite, A a une méthode simple pour s'assurer de l'identité de B : supposons par exemple que B s'appelle Bernard ; B envoie alors à A le message $X = S(Bernard)$ Pour vérifier que B est bien Bernard, A calcule alors $P(X) = P \circ S(Bernard) = Bernard$.

Comme on le voit, il n'est pas du tout nécessaire que A connaisse la clé privée P pour vérifier la signature de B !

§4. Le logarithme discret

Soit G un groupe cyclique d'ordre N , et soit g l'un de ses générateurs. Le *logarithme discret* de G en base g est l'application $\log_g : G \to \mathbf{Z}/N\mathbf{Z}$ définie par $\log_g(g^m) = m$. C'est donc un isomorphisme de groupes, d'application réciproque $m \mapsto g^m$.

Il y a des cas où le calcul du logarithme discret est trivial, par exemple si $G = \mathbf{Z}/n\mathbf{Z}$ et $g = 1$. Par contre, si $G = (\mathbf{Z}/p\mathbf{Z})^*$, où p est un nombre premier, trouver un générateur de G n'est déjà pas facile, et le calcul du logarithme discret semble très difficile.

Ceci a donné l'idée d'algorithmes sécurisés dits symétriques : supposons que A et B veuillent s'assurer de leur identité respective. Ils se donnent un groupe G cyclique de générateur g ; A (resp. B) choisit au hasard un entier a (resp. b), et envoie à B (resp. A) l'élément g^a (resp. g^b) ; A et B s'assurent alors de leur identités respectives en calculant $g^{ab} = (g^a)^b = (g^b)^a$.

§5. Preuves sans apport d'information

Le principal talon d'Achille de la méthode de signature sécurisée décrite dans le troisième paragraphe réside dans le fait qu'à un moment donné, la personne qui doit prouver son identité doit transmettre son code en clair au lecteur de cartes ou à l'ordinateur, par exemple, afin que ce dernier puisse vérifier l'identité de son interlocuteur. Il est donc possible à quelqu'un de mal intentionné d'intercepter ce code (par exemple, dans le cas d'un distributeur de billets, si quelqu'un, derrière vous, regarde les chiffres que vous tapez). Plus généralement, supposons que A possède un secret (ici un code secret), et que B veuille s'assurer que A possède ce secret. Est-ce possible sans que A donne à B une quelconque information sur ce secret ? Il existe des algorithmes (de type probabiliste : B saura avec une très forte probabilité que A possède bien le secret en question) résolvant ce problème. Nous en donnons deux ci-dessous :

I.- Soit $(G, \times)$ un groupe dans lequel on sait calculer rapidement, mais dans lequel le calcul des racines carrées est difficile (par exemple $(\mathbf{Z}/pq\mathbf{Z})^*$, où p et q sont de grands nombres premiers disctincts).

Le secret de A est une racine carrée x d'un élément $a \in G$; B connaît a, et veut s'assurer que A connaît x.

Il demande alors à A de choisir $y \in G$, et de lui envoyer $b = y^2$. Il aura alors le choix entre deux questions qu'il pourra poser à A :

1) Lui demander y (Il vérifie alors que $b = y^2$)

2) Lui demander $c = xy$ (Il vérifie alors que $ab = c^2$).

Supposons que A ne connaisse pas x.

a) S'il parie que B va lui poser la question 2), il choisit un élément quelconque $z \in G$, et envoie $b = a/z^2$ à B. Si B lui pose effectivement la question

2), le bluff aura marché, car A enverra $c = a/z$. Par contre, si B pose la question 1), B sera piégé, car il ne connaît pas de racine carrée de a/z^2 (sinon il connaîtrait une racine carrée de a).

b) S'il parie que B va poser la question 1), il joue le jeu en choisissant bien un $y \in G$ et en envoyant y^2 à B.

Mais il se fait piéger si B pose la question 2) !

Dans le cas par exemple du groupe $(\mathbf{Z}/pq\mathbf{Z})^*$, il est facile de voir qu'en $O(\log pq)$ questions, on a une probabilité de $1/(pq)$ d'être trompé par A. Or, les informations recueillies par B pendant son questionnaire ne lui donnent pas d'information susceptible de trouver x.

II.- Une méthode assez voisine repose sur le logarithme discret.

Soit G un groupe cyclique de générateur g, dans lequel le problème du logarithme discret est difficile.

Ici, le secret de A est un entier x, et B connaît $a = g^x$. Il doit s'assurer que A connaît bien x.

Il demande à A de choisir un entier y, et de lui envoyer $b = g^y$. Il a alors le choix entre deux questions :

1) demander y à A (Il vérifie alors que $b = g^y$)

2) demander $z = x + y$ à A (Il vérifie alors que $g^z = ab$.)

omme précédemment, la probabilité que A ne connaisse pas x au bout de $O(\log N)$ questions est de l'ordre de $1/N$, et les informations recueillies par B durant son questionnaire ne lui donnent pas d'informations sur x.

Indications pour le candidat

Le candidat pourra choisir de développer les points qu'il désire sur le texte. En ce qui concerne le passage sur machine, il pourra programmer la méthode RSA, pour M et m quelconque.

Plusieurs points du texte sont affirmés sans démonstration. Le candidat est invité à les expliciter.

Le candidat pourra par exemple démontrer le théorème de la section 2, et détailler les remarques concluant cette section. Par ailleurs, pourquoi le connaissance de l'ordre de G implique-t-elle la connaissance de p et q ? Pourquoi ne peut-on pas choisir $m = 2$?

Le candidat pourra également programmer le calcul du logarithme discret dans le cas du groupe $(\mathbf{Z}/n\mathbf{Z}, +)$, puis dans le cas du groupe $(\mathbf{Z}/p\mathbf{Z})^*, \times)$, où p est un nombre premier, avec l'algorithme de son choix, ou encore programmer une méthode de preuve sans apport d'information proposée dans le texte.

Chapitre 8
Commentaire de l'auteur

Ce texte, "Logarithme discret et cryptographie", se rapporte au thème d'application intitulé "Télécommunications". Il est destiné à être compris par un candidat ne connaissant a priori rien au sujet. L'introduction est destinée à familiariser le lecteur avec la problématique de la cryptographie, notamment l'opposition clé secrète-clé publique.

Le premier paragraphe décrit la méthode RSA, qui semble incontournable dans le sujet, autant pour des raisons historiques que pour son efficacité. Cela peut permettre au candidat de la programmer ; cela doit lui permettre de dégager une première difficulté : autant l'élévation au cube est aisée, même modulo un entier M de plusieurs centaines de chiffres (en utilisant par exemple Maple), autant l'application réciproque $x \mapsto x^k$ devient impossible lorsque k est de l'ordre de grandeur de M, si on l'effectue naïvement. C'est pourquoi certains candidats se sont limités à des M de l'ordre de 15, ce qui est bien sûr sans beaucoup d'intêret. Pour des M de taille cryptographique, c'est-à-dire ayant plusieurs centaines de chiffres, la méthode en $2\log_2(k)$ multiplications (utilisant la décomposition binaire de k) est incontournable. Notons dans Maple la commande $a\&^b modc$ qui calcule l'exponentiation par cette méthode.

Il est alors impressionnant de voir avec quelle rapidité Maple calcule $x \mapsto x^k$ mod M.

Toujours pour permettre au candidat de se mettre "dans le bain", la démonstration, facile, du théorème du paragraphe 2 était proposée dans les suggestions de développement.

De la même manière, le texte est délibérement parsemé d'affirmations sans indication, afin de donner au candidat l'occasion d'y réfléchir. Par exemple dans les remarques de la fin du paragraphe 2 (pourquoi $m = 65537$ est-il astucieux ?)

Le paragraphe suivant est consacré au problème des signatures sécurisées, et de l'application de RSA à ce problème. J'ai été particulièrement bref ici, en éludant les multiples difficultés de ce type de signature. Répétons-le : ce texte s'adresse à un lecteur qui ne connaît rien au sujet, et il me semble impossible en quatre heures de préparation, avec la tension qu'engendre une situation d'examen, de pouvoir entrer dans les subtilités du domaine. Ce texte est un survol, le plus élémentaire possible, de certains thèmes qui me semblent séduisants en cryptographie.

Le paragraphe suivant introduit l'idée de logarithme discret sur un groupe cyclique, et de son application dans le problème des signatures symétriques. Ce paragraphe, bien que court, peut donner matière à développement au candidat : par exemple, comment trouver un générateur de $(\mathbf{Z}/p\mathbf{Z})^*$? Pourquoi le calcul du logarithme discret dans le cas de $\mathbf{Z}/nZ$ est-il trivial (même si le générateur choisi n'est pas $g = 1$?)

Le dernier paragraphe traite des preuves sans apport d'information (au fond assez proches des QCM, mais dans un sens particulièrement strict : à la première réponse fausse, on est rejeté !) Il s'agissait pour finir le texte de réunir le logarithme discret et un algorithme proche de RSA (l'él/'evation au carré modulo M) dans une même application, qui touche aussi bien aux probabilités qu'à la logique et bien sûr à l'arithmétique.

Ici encore, cela donnait du grain à moudre au candidat, qui pouvait par exemple écrire le programme correspondant au logarithme discret, prendre la place du bluffeur qui dit connaître x et envoie g^x, et voir au bout de combien de tests en moyenne l'ordinateur le piégera.

Il n'était bien entendu pas possible à un candidat de considérer tout le texte. Il pouvait se contenter du paragraphe d'introduction et du paragraphe sur RSA, ou du paragraphe sur le logarithme discret, et de celui sur les preuves sans apport d'information. Il s'agissait de montrer au jury quelles sont les problématiques décrites dans le texte, quels sont les outils mathématiques qui entrent en jeu, et d'illustrer cela avec un petit programme, en expliquant les difficultés algorithmiques rencontrées (principalement l'élévation à une grande puissance).

Notons pour finir un bon ouvrage en français d'introduction à la cryptographie :

G. Zémor, Cours de Cryptographie, Cassini 2000.

Chapitre 9
Logarithme discret et cryptographie

Ecrit par un membre du jury non lui-même spécialiste ni d'algèbre finie, ni d'arithmétique ni a fortiori de cryptographie, ce commentaire sur le texte *Logarithme discret et cryptographie* vise à indiquer des pistes pour effectuer un exposé argumenté et mathématiquement pertinent, à l'intention d'un candidat ne possédant que les bases d'arithmétique et d'algèbre figurant au programme général de l'agrégation.

Comme pour de nombreux textes, le candidat avait intérêt à mettre en évidence le recul qui pouvait être le sien face au thème choisi, en évitant de suivre pas à pas le plan du document (et il est opportun de rappeler ici qu'il n'était en aucune façon tenu de le traiter intégralement).

Deux attitudes principales pouvaient être admissibles face à ce texte : ou bien centrer l'exposé sur la technique de la cryptographie à clef publique et sa réalisation sur ordinateur, en se limitant éventuellement à deux ou trois des variantes présentées dans le texte (et faire alors intervenir les propriétés d'algèbre ou d'arithmétique en appui de cette présentation technique), ou bien privilégier les énoncés et démonstrations d'algèbre et d'arithmétique, en faisant intervenir les méthodes du type RSA, avec leur mise en œuvre informatique, comme des sous-produits.

Dans un cas comme dans l'autre, une démarche intellectuelle (facile mais parfois omise par les candidats) était indispensable pour assurer que l'exposé relève bien des mathématiques appliquées : se mettre "dans la peau" des opérateurs de cryptographie, en pensant aux trois acteurs potentiels (l'expéditeur, le destinataire, l'espion , ou bien les deux partenaires qui veulent s'identifier et l'espion) et bien mettre en évidence :
1. comment procède un opérateur qui veut se "mettre sur le réseau" en fabriquant sa clef publique (2 nombres, M $(= pq)$ et m) et sa clef secrète (k) ;
2. comment, lors de l'envoi d'un message, procèdent l'expéditeur et le destinataire, et pourquoi l'espion ne peut pratiquement pas décoder le message.
Un exposé "vivant" , mettant en évidence, par exemple à l'aide d'organigrammes, la succession de ces opérations, était parfaitement réalisable sans connaissance préalable du thème, à partir des descriptions données dans le texte, et ce sujet se prêtait donc particulièrement à la manifestation de talents d'exposition chez le candidat. Si la mise en œuvre informatique était indispensable (et ne nécessitait qu'une connaissance limitée de MAPLE ou d'un autre logiciel de calcul scientifique) et devait bien sûr porter sur de grands nombres, des exemples non réalistes, mais présentés "à la main" avec de petits nombres, pouvaient être éclairants (et permettre de montrer pourquoi c'est justement grâce à l'emploi de grands entiers que la possibilité de

"casser" le code est extrêmement faible).

En ce qui concerne les connaissances mathématiques à mettre en évidence, le candidat avait sans doute avantage à adopter le point de vue algébrique général adopté par le texte, mais il pouvait aussi, plus modestement, se concentrer plus sur l'aspect proprement arithmétique, tel qu'il est utilisé dans les versions standards de RSA. Ceci lui permettait d'expliciter la phrase du texte : *un groupe G dans lequel il sait calculer mais dont il ne connait pas l'ordre* ; il s'agissait là de bien clarifier ce qui est sous-jacent dans la description du texte, à savoir les techniques de calcul de puisances modulaires, leur facilité et leur rapidité ; en particulier, en s'en tenant à un niveau d'arithmétique assez élémentaire, il semblait naturel de mettre en évidence :
- pour la fabrication du couple (clef publique, clef secrète) l'usage de l'algorithme d'Euclide étendu (autrement dit la fabrication d'une identité de Bezout) aisé à rappeler et à "faire tourner" devant le jury,
- pour l'échange des messages le rôle central de l'indicateur d'Euler $\phi(M)$ (nombre d'entiers premiers avec M et strictement inférieurs à M), son mode de calcul (et la trivialité de celui-ci si M est le produit de deux entiers premiers p et q, auquel cas $\phi(M) = (p-1)(q-1)$) et l'utilité de l'identité d'Euler (si x et M sont premiers entre eux, $x^{\phi(M)} \equiv 1$ modulo M).

Les évaluations de complexité des calculs données dans le texte pouvaient perturber un peu des candidats non habitués à ce type de considération. Les considérer avec calme permettait en général d'en comprendre rapidement la signification avec un peu de réflexion et d'en faire état devant le jury (par exemple le fait que le calcul de x^k peut se faire avec un nombre d'opérations de l'ordre du logarithme binaire de k)

Bien sûr le candidat ayant déjà des connaissances de cryptographie pouvait y recourir, en particulier confronter les informations apportées ici à son savoir, mais ces commentaires voudraient persuader le lecteur de la possibilité de faire valoir à partir de ce texte, sans telles connaisances préalables, des qualités de bon sens, des capacités pédagogiques et une culture basique en arithmétique.

Chapitre 10
Commentaire d'un non spécialiste

Je ne crois pas qu'il est possible pour un étudiant d'être spécialiste en tout et en particulier en cryptographie, même si le système R.S.A. est fréquemment mentionné dans les leçons d'Algèbre comme *Congruences dans* $\mathbb{Z}$, *Anneau* $\mathbb{Z}/n\mathbb{Z}$, *applications* et *Nombres premiers et applications*.

Dans les commentaires, je ne parlerai qu'en annexe des méthodes récentes faisant appel à des connaissances débordant largement le cadre du programme de l'agrégation. Cet exercice ne s'adresse pas aux collègues mais à des étudiants titulaires d'une maîtrise de mathématiques.

Il convient de rappeler quelques points matériels qui entrent dans la stratégie de travail :

. la préparation dure quatre heures,

. au tirage on ne connaît que le titre du texte, ce dernier n'est découvert que dans la *salle de travail*,

enfin

. la bibliothèque est à la disposition des candidats, mais n'est pas à proximité immédiate du lieu de composition.

Ce dernier point est lié au lieu de passage du concours mais a son importance pratique car l'accès à la bibliothèque nécessite un déplacement, l'accompagnement par un appariteur et implique une consommation de temps. Il convient donc de bien choisir sa documentation au moment du tirage. Il y a par conséquent un travail de connaissance de la bibliothèque qui doit avoir été entrepris au cours de l'année de préparation.

Je vais donc expliquer comment je procéderais si, dans la position de l'étudiant je découvrais le titre :

Logarithme discret et cryptographie

et que je tente de le préparer. Je commencerais par faire mon choix dans les livres, je crois que je me limiterais à deux ouvrages :

COURS D'ALGÈBRE, PRIMALIT, DIVISIBILIT, CODES de M. Demazure

COURS DE CRYPTOGRAPHIE de G. Zémor.

peut-être trois : si ma culture me le permet, j'emprunterai aussi

A COURSE IN COMPUTATIONAL ALGEBRAIC NUMBER THEORY de H. Cohen

Le dernier livre ne figure pas dans la bibliothèque de l'Agrégation mais se trouve accessible dans beaucoup de *malles*. Les deux premiers ouvrages sont ciblés sur le sujet, le premier offre des algorithmes en CAML, des points proposés en exercice dans l'un sont résolus dans l'autre, des algorithmes sont également disponibles dans le troisième.

D'autres choix sont sans doute possibles, j'assume ceux là dans ma position de candidat sur le tard.

Pour éviter les problèmes de droits d'auteurs je m'abstiendrai, dans ce qui suit, de recopier et renverrai à [D], [Z], [C] en précisant à chaque fois la page.

La seconde étape, découverte du texte, se termine en page 6 par des indications pour le candidat. Passons en revue ces suggestions :

Plusieurs points du texte sont énoncés sans démonstration :

C'est le cas en particulier pour le second pararaphe. Les détails sur le coût du calcul du pgcd sont développés dans [D] (pages 20 à 28, en liaison avec la suite de Fibonacci) où sont également exposées les diverses variantes de l'algorithme d'Euclide (en particulier l'algorithme binaire, utile pour ceux qui ont accès à la programmation en langage machine), les réponses sur les points de la méthode R.S.A. donnés sans détails dans le texte peuvent se trouver dans [Z] (pages 84 à 87) avec des considérations intéressantes (page 87) sur les risques d'utilisation de nombres facilitant les calculs.

La méthode rapide d'exponentation est exposée dans [D] (pages 10 à 14), on remarquera que l'exposant proposé page 3 vaut $2^{16}+1$ et se prête particulièrement bien à l'exploitation.

Connaissant $n = pq$ et $\varphi(n) = (p-1)(q-1) = n+1-(p+q)$, on a la somme et le produit de p et q, donc ces deux nombres.

Le texte nous invite à programmer la méthode R.S.A., la description qui est donnée permet facilement de construire l'algorithme, une fois choisis p et q. Pour le choix des nombres premiers, on peut piocher dans [D] (pages 70, 77), ou utiliser les fonctions de MAPLE, tout en précisant qu'il ne s'agit pas d'une réalisation efficace (est-il réaliste de demander à un étudiant découvrant le texte de créer en quatre heures un projet utilisable ?), mais d'une illustration.

Pour celui qui veut vraiment mettre en œuvre la méthode R.S.A. (comme on y est invité) plusieurs problèmes pratiques vont se poser qui peuvent être objets de commentaires, suivant le point de vue où on se place :

1) Pour tous les acteurs (le concepteur, l'utilisateur, l'indiscret), les calculs doivent se faire sur des grands nombres il semble possible de faire un développement :

. Sur le calcul modulaire, à partir du moment où on connaît une borne des nombres avec lesquels on doit travailler. C'est encore une application du théorème des restes chinois (cette technique est conseillée dans [Z], page 86, pour lire le message crypté).

. Sur la transformée de Fourier rapide, le chapitre 4 (pages 93 à 106) de [D] convient pour cet objectif ; il faut toutefois s'attendre, si on fait ce choix à des questions sur ses autres applications.

2) On peut bien évidemment se placer du point de vue du constructeur du système, il faut trouver deux nombres premiers de grande taille. Il est conseillé ([Z], page 85) de choisir des nombres au hasard et de tester leur primalité. Cela peut, bien sur, conduire à un développement sur les tests de primalité. C'est l'occasion de mettre en évidence la différence de difficulté (en l'état actuel de l'art) entre primalité et factorisation, point négligé, à juste titre, dans les petites classes. On pourra, par exemple, s'intéresser au :

i) Critère de Miller-Rabin [D] (page 62 à 64) avec le théorème de Rabin [D] (page 64, démonstration pages 90 à 92), ou [Z] (pages 104 et 105), ce test est *probabiliste*, il ne certifie pas que le nombre étudié est premier mais en donne une probabilité proche de un.

ii) Procédé de fabrication de nombres premiers (critère à la Lehmer dans [D] pages 72 à 75), théorème de Lucas dans [Z] (page 105) qui présente l'avantage de fournir un générateur du groupe multiplicatif, aisément programmable, ce procédé est critiqué dans la mesure où une méthode systématique de construction de nombres premiers peut présenter des failles exploitables par un cryptoanalyste, une référence sur ces dernières est donnée dans [Z] (page 95) qui cite également (page 115) les inconvénients du choix de deux nombres premiers trop proches l'un de l'autre. Au sujet de la détermination d'une racine primitive modulo p, on notera qu'il n'existe pas, en général, d'autre méthode que la recherche exhaustive [D] (pages 69 à 71) ; expérimentalement, il y a toujours une racine primitive de petite taille (la table de [D] page 70 est parlante), c'est d'ailleurs une conséquence de l'Hypothèse de Riemann Généralisée.

3) On peut aussi essayer de se placer du point de vue de l'indiscret et faire un développement sur la factorisation en partant de l'idée de Fermat : bases de factorisation, utilisation des fractions continues, cribles quadratiques, crible quadratique multiple ([Z] pages 106 à 112).

Les tests de primalité développés ci-dessus sont de nature probabiliste, la probabilité est proche de 1, suffisamment pour créer la classe des "nombres premiers commerciaux", si le candidat utilise, pour son application informatique, le logiciel Maple il remarquera que c'est là sa situation comme le fait remarquer l'aide en ligne.

Des procédés, plus récents, soit de primalité, soit de factorisation sont disponibles dans la littérature (voir [C] chapitres 9 et 10) ; mais qu'il s'agisse des tests de primalité (par les sommes de Jacobi, les courbes elliptique) ou bien des méthodes de factorisation de Shanks (par la réduction des formes quadratiques, probablement peu exploitable dans le cas des nombres de 150 chiffres) de Lenstra (par le crible des corps de nombres), ou par les courbes elliptiques je crains d'une part qu'elles ne dépassent de beaucoup le programme officiel de l'agrégation et d'autre part que le temps de préparation ne soit trop court pour les programmer (il n'en irait pas de même dans le cadre d'un TER – Travail d'Etude et de Recherche – en maîtrise), le candidat à même de présenter un exposé sur ce thème n'a sans doute pas grand besoin de lire ces lignes. On remarquera qu'en contrepartie de leur difficulté, les tests de primalité évoqués ci-dessus, ne sont pas probabilistes.

Une fois répondu aux questions de l'auteur du texte, il peut venir à l'esprit du candidat que s'il est bon de savoir crypter, décrypter encore faut-il que le message reçu soit conforme à celui expédié, des altérations dues à des problèmes de transmission pouvant le rendre incompréhensible. Il me semble qu'une vue d'ensemble du procédé nécessite de parler de code correcteurs d'er-

reurs, mais je ne me hasarderais pas, dans la position du candidat, à en faire le centre de mon développement et de mon application informatique.

Voilà donc une liste de points que j'aurais eu envie de développer, il y a là, ce me semble, matière à plusieurs exposés et encore n'ai-je pas exploité complètement le texte proposé.

Chapitre 11
Commentaire TIPE du texte

Les thèmes de TIPE passés (système dynamique/mesure, terre et espace, transformations et fonctions, composition décomposition) ont donné lieu à de nombreux exposés sur la cryptographie.

Comment utiliser avec profit le texte d'agrégation *cryptographie et logarithme discret* comme source de sujets adaptés à l'épreuve de TIPE des concours d'entrée aux grandes écoles ? Tel sera le fil directeur de ce bref commentaire[1].

• Le texte commence par la présentation de la problématique de la cryptographie à clé publique. Ce système de chiffrement utilise le concept de *fonction à sens unique avec trappe* : une bijection facile à calculer mais difficile à inverser sans information supplémentaire (la trappe). Dans le cas de RSA, le chiffrement est réalisé par exponentiation modulo un grand entier n et la connaissance de la factorisation de n permet la détermination de la clé de déchiffrement (l'exposant de décodage). Le sens «unique» de cette fonction de chiffrement est «garanti» par la difficulté mathématique de la factorisation d'un grand entier= =2E

Il peut être intéressant pour un étudiant d'étudier d'autres systèmes de cryptage à clé publique en dégageant à chaque fois la fonction à sens unique avec trappe employée. Voici une brève[2] présentation de quelques cryptosystèmes abordables au niveau des CPGE :

- Le cryptage par *sac à dos*. Cette méthode est basée sur le problème mathématiquement difficile des partitions d'entier encore appelé problème du sac à dos : il faut trouver un sous-ensemble d'objets dont la somme des poids vaut une valeur donnée (le poids final du sac). La suite de bits à crypter est considérée comme la suite des coefficients d'une combinaison linéaire d'entiers donnés. La somme de cette combinaison linéaire est le message chiffré. Si la suite d'entiers en question n'a pas de propriété particulière, seule une recherche exhaustive (et donc exponentielle) permet de retrouver les coefficients c'est-à-dire le message en clair. Lorsque la suite d'entiers est supercroissante[3], le décodage s'effectue en temps polynomial. L'astuce consiste ainsi à dissimuler (de façon modulaire) la supercroissance de la suite de codage : c'est la clé secrète. Hormis l'abord

[1] Je remercie N. Sendrier du projet CODES de l'INRIA pour sa vue synthétique des choses ainsi que les étudiants que j'ai encadrés en TIPE, pour leur travail et la modeste expérience des TIPE qu'ils m'ont apportée.

[2] Évidemment, dans un commentaire de quatre pages, il m'est impossible de développer convenablement chaque système présenté (ce qui est plutôt frustrant !) ; je renvoie le lecteur à la bibliographie finale.

[3] Le n-uplet d'entiers positifs $(a_1, a_2, \ldots, a_n)$ vérifie pour tout i : $a_i > \sum_{j<i} a_j$.

aisé pour un étudiant de CPGE d'un tel cryptosystème, le sac à dos présente l'attrait supplémentaire d'avoir été l'objet de nombreuses attaques. Un étudiant peut par exemple étudier et implémenter l'attaque de Lagarias/Odlyzko (basée sur la recherche de petits vecteurs dans un réseau).
- Le système *HFE* (pour Hidden Field Equations). À nouveau, on chiffre une suite de bits en les considérant comme les coefficients d'un polynôme cette fois. Celui-ci (une composée plus exactement) est projeté modulo un polynôme donné irréductible à coefficient dans un corps fini : le résultat constitue le message chiffré. Le déchiffrement, sans la connaissance du polynôme, impose la résolution de systèmes d'équations quadratiques : c'est le problème mathématiquement difficile. À l'instar de RSA, la difficulté des attaques[4] menées contre ce cryptosystème garantit sa sécurité.
- Le cryptage par *code correcteur d'erreur*. De façon simplifiée, le principe du cryptage est d'introduire volontairement des erreurs dans le message à envoyer ; le décodage se fait alors à l'aide d'un code correcteur d'erreur, la trappe étant la structure algébrique du code. Le système le plus répandu est celui de Mac Eliece. La clé privée est la matrice génératrice du code, la clé publique est son produit par une matrice aléatoire et de permutation. À la différence des systèmes du type RSA/log discret qui utilisent pour crypter l'exponentiation dans un corps fini, les systèmes à base de code correcteur sont basés sur de l'algèbre linéaire : ce qui constitue, par l'utilisation du cours de mathématiques de spéciales, un attrait supplémentaire pour l'étudiant de CPGE.
- Le cryptage par *automate fini*. L'étudiant de CPGE suivant l'option informatique peut mettre à profit son cours sur les automates finis. Les données sont chiffrées en les passant dans un automate fini (la clé publique) et déchiffrées en les passant à travers les inverses de ses composants : la détermination de l'inverse de l'automate est l'opération difficile.

• Un autre aspect de développement pour les TIPE du texte *Logarithme discret et cryptographie* consiste à revenir sur le système RSA. Ce cryptosystème a certes déjà été choisi à de maintes reprises et, par souci d'originalité, il convient d'éviter de réduire son exposé de TIPE à sa simple présentation. Mais on peut à mon avis insister avec profit sur le cœur du système qu= 'est le problème mathématiquement difficile de la factorisation de grands entiers. À nouveau, voici quelques pistes abordables par un étudiant de CPGE :
- L'implémentation des méthodes classiques (détaillée dans les ouvrages cités en fin d'article) où l'étudiant peut expérimenter la répartition des nombres premiers (et l'indicateur d'Euler), la méthode de Fermat (pour trouver des facteurs premiers proches de $\sqrt{n}$), la méthode de Legendre (qui est en $O(n^{1/3})$) etc.
- La *méthode* ρ de Pollard. L'idée basique consiste ici à constater qu'une suite récurrente évoluant dans un ensemble fini de valeurs devient pério-

[4] L'attaque menée par J. Patarin et N. Courtois a permis de montrer les avantages de HFE par rapport à RSA.

dique. On exploite alors l'égalité entre deux termes d'indices différents pour déterminer (à l'aide d'un calcul de pgcd) un facteur premier de l'entier considéré. Pollard a poursuivi son étude sur la factorisation pour mettre au point le *crible algébrique* (ou NFS : Number Field Sieve), algorithme extrêmement efficace à l'origine de récents records[5] de factorisation.

– *Les courbes elliptiques.* Il s'agit sans doute d'un domaine à la limite de compétence d'un étudiant moyen de CPGE. Une courbe elliptique est l'ensemble des points à coordonnées dans un corps fini vérifiant une équation du type $y^2 = 3Dax^3 + bx + c$. Le point important est que ces points forment un groupe additif. On peut transposer le problème du logarithme discret[6] dans ce groupe (on parle d'exponentielle elliptique) et l'étude des courbes elliptiques fournit à la fois des méthodes de chiffrement et de factorisation. Le lecteur désireux d'approfondir le sujet se reportera à la bibliographie.

Pour conclure sur ces points, j'ai volontairement choisi de ne pas parler de cryptographie quantique ni de factorisation quantique[7]. L'interdisciplinarité est certes louable en TIPE, mais la mécanique quantique n'est pas enseignée en CPGE !

• Le texte se poursuit par la présentation de notions connexes à la cryptographie. Chacune de ces notions (leur compréhension[8] accompagnée d'une réalisation concrète) constitue un TIPE intéressant. Notez que l'interdépendance de ces notions participe à la difficulté de leur compréhension rigoureuse.

– *La signature.* Cette fonction permet de garantir à la fois l'intégrité d'un message[9] et doit convaincre le destinataire que le signataire a délibérément signé le document. Il faut pouvoir éviter la falsification d'une signature ainsi que la répudiation de celle-ci par son auteur. Le texte présente l'utilisation de RSA comme méthode de signature. Il existe bien d'autres méthodes que le lecteur trouvera dans les références citées en fin d'article.

– *L'identification* c'est-à-dire prouver son identité (et donc empêcher un tiers de se faire passer pour qui il n'est pas). C'est le rôle par exemple du

[5] La barre psychologique des 512 bits de clé RSA est ainsi tombée en 1999.

[6] C'est-à-dire la difficulté mathématique à inverser une exponentiation modulaire.

[7] En 1993, Peter Shor, des laboratoires AT&T, a conçu un algorithme permettant de factoriser un nombre en un temps polynomial. Ceci constitue un argument de poids en faveur des cryptosystèmes basés sur les codes correcteurs d'erreurs par rapport à ceux basés sur RSA/log discret. Seul problème (temporaire ?) : on ne dispose pas encore d'ordinateur quantique pour implémenter l'algorithme de Shor !

[8] Ce qui est moins simple qu'il n'y paraît, et constitue un excellent exercice d'exposé pour un étudiant.

[9] Dans le cadre de transactions bancaires, par exemple, il est impératif de pouvoir vérifier que les montants d= es transactions ainsi que les numéros de compte n'ont pas été altérés.

code secret de la carte bancaire. Le texte présente un tel protocole basé sur le logarithme discret (Diffie-Hellman). L'algorithme de chiffrement librement distribué (open-)PGP permet de réaliser des petites communautés où les utilisateurs certifient, en les signant, les clés des utilisateurs en qui ils ont confiance (mécanisme appelé *Web of trust*).

– *Les preuves sans apport d'information* ou *zero-knowledge*. Il s'agit, comme le dit le texte, de convaincre un interlocuteur que l'on dispose d'un secret en lui révélant cependant trop peu d'informations pour éventuellement le reconstituer. L'approche probabiliste[10] des choses (la probabilité pour que l'émetteur ne détienne pas le secret décroît très rapidement vers 0 et à la convenance du récipiendaire) constitue aussi un thème de TIPE intéressant. L'étudiant peut la rapprocher des tests probabilistes de primalité (Test de =46ermat -> nombres de Carmichael etc.) facilement programmables.

Le lecteur trouvera dans la bibliographie quantité[11] d'autres problèmes analogues (et pourtant toujours différents) développables en TIPE par un étudiant de CPGE.

Ma conclusion sera d'insister sur ce qui constitue à mon avis un des objectifs majeurs du TIPE : l'illustration du caractère «vivant» et expérimental des mathématiques. Il convient d'encourager le contact avec les chercheurs et l'utilisation immodérée de l'outil informatique (et notamment des logiciels de calcul formel). Finalement, j'espère que le lecteur pourra, avec les quelques pistes présentées dans ce commentaire de *logarithme discret et cryptographie*, apporter sa propre contribution aux «mathématiques en situation».

Luc ALBERT

À nouveau, j'ai choisi de présenter en priorité des références bibliographiques abordables en CPGE. J'ai notamment privilégié les ouvrages en français. Le lecteur désireux d'approfondir les notions y trouvera de toute façon les références vers des articles plus spécialisés.

– M. Mignotte. *Mathématiques pour le calcul formel.* PUF.
– P. Naudin, C. Quitté. *Algorithmique algébrique.* Masson.
– G. Robin. *Algorithmique et cryptographie.* SMAI, mathématiques et applications 8, 1991.
– B. Schneier. *Cryptographie appliquée.* International Thomson Publishing.
– G. Zémor. *Cours de cryptographie.* Cassini.

[10] bien que les probabilités ne soient pas au programme des CPGE scientifiques.

[11] Citons le problème du partage d'un secret par une communauté, de la signature aveugle, de la signature simultanée d'un contrat, de la partie de poker jouée à distance etc.

Quatrième partie

Configuration centrale de quatre corps dans le plan

Chapitre 12
Configurations centrales de quatre corps dans le plan

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Les candidat(e)s sont laissé(e)s libres d'organiser leur discussion comme ils ou elles l'entendent. Des suggestions de développement, largement indépendantes les unes des autres, sont proposées à la fin. Le candidat n'est pas tenu de les suivre. Le jury appréciera que la discussion soit accompagnée d'exemples effectivement traités sur ordinateur.

Résumé

On étudie les configurations de quatre corps de même masse soumis à un potentiel gravitationnel et suivant un mouvement homothétique. Ce sont des configurations planes qui minimisent l'énergie. Utilisant les propriétés algébriques de la coplanarité, on décrit explicitement les configurations en résolvant un système d'équations polynomiales.

§1. Introduction

La notion de *configuration centrale* apparaît naturellement quand on recherche des solutions simples des équations différentielles qui régissent le mouvement de plusieurs masses ponctuelles soumises aux lois de Newton. Un tel mouvement ne subit ni déformation, ni rotation, et change simplement de taille au cours du temps.

Les configurations centrales trouvent tout leur intérêt

1. dans la recherche de configurations particulières observées dans l'univers, par exemple le Soleil, Jupiter et les astéroïdes troyennes.
2. dans les problèmes de changements de topologie des variétés intégrales du problème général des n corps.

Les configurations centrales se caractérisent de façon simple. Nous allons expliciter toutes ces configurations dans le cas de quatre masses ponctuelles soumises à leur force de gravitation (potentiel newtonien). La résolution repose sur une étude de propriétés algébriques de configurations de quatre points coplanaires puis, admettant une propriété de symétrie, sur la résolution d'un système d'équations polynomiales.

§2. Configurations centrales de quatre corps

Considérons 4 particules de masses $m_1, \dots, m_4$, réels strictement positifs dont la somme est notée M, soumises à leur force de gravitation. Une origine ayant été fixée, notons r_i le vecteur position de la particule i. Définissons le potentiel newtonien (énergie potentielle)

$$U = \sum_{1 \le i < j \le n} \frac{m_i m_j}{\|r_i - r_j\|}.$$

Les équations du mouvement deviennent

$$m_i \ddot{r}_i = \frac{\partial U}{\partial r_i}, \; i = 1, \dots, n.$$

Le centre de gravité du système correspond à $r_c = \dfrac{1}{M}(m_1 r_1 + \cdots + m_n r_n)$ et suit un mouvement rectiligne uniforme. Nous choisissons donc pour toute la suite l'origine comme le centre de gravité du système.

Nous posons $\overrightarrow{r_{ij}} = \overrightarrow{r_i} - \overrightarrow{r_j}$. Après renormalisation, on se donne une fonction réelle Φ et les deux fonctions

$$U = \sum_{1 \le i,j \le 4} m_i m_j \Phi(r_{ij}^2), \quad I = \frac{1}{M} \sum_{1 \le i,j \le 4} m_i m_j r_{ij}^2.$$

appelées *fonction de forces* et *moment d'inertie par rapport au centre de masse*. Dans le cas du potentiel newtonien, nous avons $\Phi(r^2) = 1/r$.

On utilise une caractérisation simple des configurations centrales : **ce sont les points critiques de la fonction U restreinte aux configurations planes de moment d'inertie I fixé.** Le problème est formulé en terme des 6 distances mutuelles.

Pour cela nous utilisons le fait que la coplanarité de quatre points de distances mutuelles respectives r_{ij} est équivalent à la nullité du déterminant de Cayley

$$P = \begin{vmatrix} 0 & r_{12}^2 & r_{13}^2 & r_{14}^2 & 1 \\ r_{12}^2 & 0 & r_{23}^2 & r_{24}^2 & 1 \\ r_{13}^2 & r_{23}^2 & 0 & r_{34}^2 & 1 \\ r_{14}^2 & r_{24}^2 & r_{34}^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{vmatrix}. \tag{1}$$

En utilisant une propriété de symétrie, nous allons expliciter les configurations centrales.

§3. Déterminants de Cayley

Soit $A = (x_1, y_1)$, $B = (x_2, y_2)$ et $C = (x_3, y_3)$ trois points dans le plan euclidien $\mathbb{R}^2$. L'aire orientée du triangle (A, B, C) est donnée par

$$S(A,B,C) = \frac{1}{2}\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}. \tag{2}$$

En posant $r_{ij}^2 = (x_i - x_j)^2 + (y_i - y_j)^2$, un simple calcul montre l'égalité

$$\begin{vmatrix} 0 & r_{12}^2 & r_{13}^2 & 1 \\ r_{12}^2 & 0 & r_{23}^2 & 1 \\ r_{13}^2 & r_{23}^2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix} = -16S^2(A,B,C). \tag{3}$$

Considérons à présent quatre points M_1, M_2, M_3, M_4 dans l'espace. Posons $\overrightarrow{r_{ij}} = \overrightarrow{M_iM_j}$.

Lemma 3.1. *Le déterminant P (équation 1) est nul si et seulement si M_1, M_2, M_3, M_4 sont coplanaires.*

Preuve. — Considérons la fonction

$$X(M) = \alpha MM_1^2 + \beta MM_2^2 + \gamma MM_3^2 + \delta MM_4^2,$$

où $\alpha, \beta, \gamma, \delta$ sont quatre nombres réels de somme nulle. Les surfaces de niveau $X(M) = \lambda$ sont des plans perpendiculaires au vecteur $\alpha\overrightarrow{r_{14}} + \beta\overrightarrow{r_{24}} + \gamma\overrightarrow{r_{34}}$. Choisissons $\alpha, \beta, \gamma, \delta$ de sorte que le plan en question porte $M_1M_2M_3$. On note h la hauteur issue de M_4. Alors nous obtenons

$$\begin{aligned} \lambda &= \beta r_{12}^2 + \gamma r_{13}^2 + \delta r_{14}^2 = \alpha r_{12}^2 + \gamma r_{23}^2 + \delta r_{24}^2 \\ &= \alpha r_{13}^2 + \beta r_{23}^2 + \delta r_{34}^2 = \alpha r_{14}^2 + \beta r_{24}^2 + \gamma r_{34}^2 - \delta h^2. \end{aligned}$$

On déduit alors que pour tout ω, nous avons

$$A \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \omega \end{pmatrix} = \begin{pmatrix} \lambda + \omega \\ \lambda + \omega \\ \lambda + \omega \\ \lambda + \omega - 2\delta h^2 \\ 0 \end{pmatrix}, \text{ où } A = \begin{pmatrix} 0 & r_{12}^2 & r_{13}^2 & r_{14}^2 & 1 \\ r_{12}^2 & 0 & r_{23}^2 & r_{24}^2 & 1 \\ r_{13}^2 & r_{23}^2 & 0 & r_{34}^2 & 1 \\ r_{14}^2 & r_{24}^2 & r_{34}^2 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \tag{4}$$

Le choix $\omega = -\lambda$ donne, en résolvant le système pour δ,

$$\delta \det(A) = -2\delta h^2 \begin{vmatrix} 0 & r_{12}^2 & r_{13}^2 & 1 \\ r_{12}^2 & 0 & r_{23}^2 & 1 \\ r_{13}^2 & r_{23}^2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix}$$

soit $P = \det A = 288V^2$ où V est le volume du tétraèdre $M_1M_2M_3M_4$. □

§4. Configurations planes

Considérons à présent quatre points M_1, M_2, M_3, M_4 dans le plan, non alignés. Posons

$$\Delta_1 = S(M_2, M_3, M_4), \Delta_2 = -S(M_1, M_3, M_4) = S(M_4, M_3, M_1) \tag{5}$$
$$\Delta_3 = S(M_1, M_2, M_4), \Delta_4 = -S(M_1, M_2, M_3) = S(M_3, M_2, M_1). \tag{6}$$

Des déterminants

$$\begin{vmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ 1 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ 1 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} y_1 & y_2 & y_3 & y_4 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ 1 & 1 & 1 & 1 \end{vmatrix} = 0,$$

on déduit que $\sum_{i=1}^{4} \Delta_i = 0$ et $\sum_{i=1}^{4} \Delta_i \overrightarrow{r_i} = \overrightarrow{0}$. Ici, on a posé $\overrightarrow{r_i} = \overrightarrow{OM_i}$ où O est une origine du repère.

Ces propriétés ne sont rien d'autres que l'expression des coordonnées barycentriques d'un point dans un triangle. On en déduit alors que l'expression

$$\sum_{i=1}^{4} \Delta_i \left\| \overrightarrow{r_i} - \overrightarrow{r} \right\|^2$$

est constante et en particulier, si $r_{ij} = \left\| \overrightarrow{r_i} - \overrightarrow{r_j} \right\|$, que

$$\sum_{j \neq i} \Delta_j r_{ij}^2 = \lambda, \ i = 1, \ldots, 4. \tag{7}$$

On déduit également le fait que

$$\sum \Delta_i \Delta_j r_{ij}^2 = 0. \tag{8}$$

L'équation (7) montre que le noyau de A (définie en (4)) est engendré par le vecteur $(\Delta_1, \Delta_2, \Delta_3, \Delta_4, -\lambda)$. On déduit le

Lemma 4.1. *Lorsque les M_i, $1 \leq i \leq 4$, sont coplanaires, il existe une constante σ telle que*

$$\frac{\partial P}{\partial r_{ij}^2} = 2\sigma \Delta_i \Delta_j. \tag{9}$$

Preuve. — Soit l'application $\phi : X \mapsto \det X$. Notons X^* la comatrice de X, c'est-à-dire la transposée de la matrice des cofacteurs de X. On a $\dfrac{\partial \phi}{\partial X_{i,j}} = X^*_{i,j}$.

Soit A une matrice de rang $n-1$ et $(V_1, \dots, V_n)$ un vecteur non nul du noyau de A. On a $A^*A = AA^* = 0$. On déduit alors que les colonnes de A^* sont dans le noyau de A donc proportionnelles à $(V_1, \dots, V_n)$.

Si A est de plus une matrice symétrique, alors A^* est symétrique et on déduit aussi que les lignes de A^* sont proportionnelles au même vecteur et donc $A^*_{ij} = \sigma V_i V_j$.

Dans le cas qui nous préoccupe, considérant la composée de fonctions

$$\psi : (r_{12}^2, \dots, r_{34}^2) \mapsto A \mapsto \det(A),$$

on déduit, en utilisant que $(\Delta_1, \Delta_2, \Delta_3, \Delta_4, -\lambda)$ engendre $\ker A$, que

$$\frac{\partial \psi}{\partial r_{ij}^2} = A^*_{ij} + A^*_{ji} = 2A^*_{ij} = 2\sigma \Delta_i \Delta_j$$

Remarque. — Considérant que $A^*_{1,1} = -16\Delta_1^2$, on déduit $\sigma = -16$.

On retrouve d'ailleurs, l'équation (8) en utilisant l'homogénéité de la fonction ϕ. □

§5. Détermination des configurations centrales

Nous recherchons les configurations centrales dans le cas d'un potentiel newtonien ($\Phi(r^2) = 1/r$) et de quatre masses égales.

De la caractérisation des configurations centrales, nous déduisons que les dérivées partielles

$$\frac{\partial \Phi}{\partial r_{ij}^2}, \frac{\partial I}{\partial r_{ij}^2}, \frac{\partial P}{\partial r_{ij}^2},$$

sont liées et puisque les quatre points sont supposés non alignés, nous déduisons l'existence, pour chaque configuration, de deux réels u et v tels que

$$\Phi'(r_{ij}^2) = u + v\Delta_i\Delta_j. \tag{10}$$

La détermination des configurations centrales de quatre corps coplanaires équivaut à la résolution du système

$$\begin{cases} \exists \lambda \in \mathbb{R} \quad ; \quad \sum\limits_{j \neq i} \Delta_j r_{ij}^2 = \lambda,\ i = 1, \dots, 4 \\ \qquad\qquad\qquad \sum\limits_{1 \leq i \leq 4} \Delta_i = 0. \end{cases} \tag{11}$$

$$\exists u, v \in \mathbb{R} \quad ; \quad \Phi'(r_{ij}^2) = u + v\Delta_i\Delta_j,\ 1 \leq i < j \leq 4, \tag{12}$$

Nous allons de plus faire une hypothèse de symétrie : *il existe un axe de symétrie contenant deux des corps.* Cette hypothèse a en fait été démontrée en

toute généralité pour quatre corps si on suppose que la fonction Φ est convexe et de dérivée concave. C'est le cas par exemple pour $\Phi(r) = r^{-1/2}$.

La preuve repose d'une part sur la distinction entre configurations convexes (c'est-à-dire aucun des sommets n'est dans l'enveloppe convexe des trois autres) et non convexes et sur le fait que les équations (11,12) entraînent des équations entre déterminants, lesquels ont un signe imposé par l'hypothèse de concavité.

Pour résoudre le système, nous allons donc supposer que la configuration possède un axe de symétrie contenant les corps 1 et 2. Cette hypothèse entraîne que $r_{13}^2 = r_{14}^2 = b$ et que $r_{23}^2 = r_{24}^2 = d$. Ces conditions entraînent que $\Delta_3 = \Delta_4$. Posant $r_{12}^2 = a$, $r_{34}^2 = f$, on obtient

$$\Delta_2 a + 2\Delta_3 b = \Delta_1 a + 2\Delta_3 d = \Delta_1 b + \Delta_2 d + \Delta_3 f, \tag{13}$$

$$\Delta_1 + \Delta_2 + 2\Delta_3 = 0 \tag{14}$$

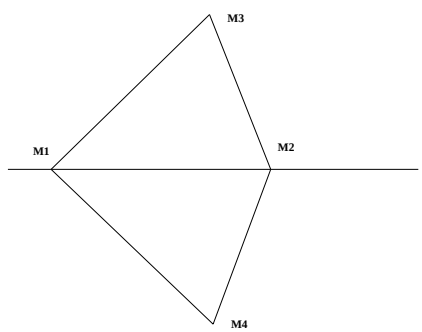


FIG. 1 – Symétrie de la configuration centrale

§6. Résolution des équations

Indépendamment du potentiel Φ, on traite d'abord la coplanarité. Le système (14) devient

$$\left\{ \begin{array}{lcl} \Delta_1 = \Delta_3(-1-t) & , & \Delta_2 = \Delta_3(-1+t) \\ 4b = f + a(1-t)^2 & , & 4d = f + (1+t)^2 a. \end{array} \right. \tag{15}$$

et le système (12)

$$\left\{ \begin{array}{lcl} \Phi'(a) & = & u + v\Delta_3^2(1-t^2), \\ \Phi'(b) & = & u - v\Delta_3^2(1+t), \\ \Phi'(f) & = & u + v\Delta_3^2, \\ \Phi'(d) & = & u - v\Delta_3^2(1-t). \end{array} \right. \tag{16}$$

On pose ici $a = 1$, qui est un choix de normalisation. On peut poser $\Delta_3^2 = 1$, ce qui revient à remplacer en fait v par v/Δ_3^2. Enfin, posons $f = z^2$ (avec $z \geq 0$). Le nouveau système à traiter devient

$$\left\{ \begin{array}{rcl} \Phi'(1) & = & u + v(1-t^2), \\ \Phi'(b) & = & u - v(1+t), \\ \Phi'(z^2) & = & u + v, \\ \Phi'(d) & = & u - v(1-t). \end{array} \right. \tag{17}$$

6.1. Potentiel newtonien. Ici $\Phi'(z) = -z^{-3/2}/2$. On obtient tout de suite (quitte à multiplier u et v par -2), en résolvant la première et le troisième équation de (17)

$$v = \frac{1-z^3}{z^3t^2},\ u = \frac{(t^2-1)+z^3}{z^3t^2}. \tag{18}$$

Ici, $t = 0$ fournit le carré comme configuration centrale.

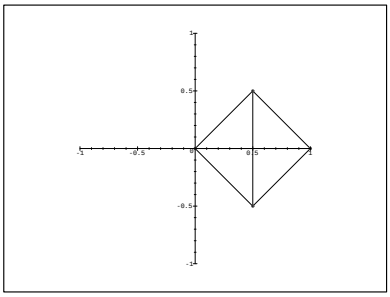


FIG. 2 – carré

Une fois u et v imposés, on a

$$b^{-3} = (1/4(z^2 + (1-t)^2))^{-3} = (u - v(1+t))^2, \tag{19}$$
$$d^{-3} = (1/4(z^2 + (1+t)^2))^{-3} = (u - v(1-t))^2. \tag{20}$$

L'équation (19) fournit donc une équation polynomiale $P(z,t) = 0$ tandis que l'équation (20) fournit $P(z,-t)$. On déduit alors deux équations

$$A(z,t^2) = \frac{1}{4t}(P(z,t) - P(z,-t)) = 0,\ B(z,t^2) = \frac{1}{2}(P(z,t) + P(z,-t)) = 0.$$

Ces deux derniers polynômes, s'écrivent, en posant $t^2 = T$,

$$\begin{aligned}A = &-(z^3-4)\,T^4 + (z^6 - 5\,z^3 - 3\,z^5 + 9\,z^2 + 14)\,T^3\\ &- (-6\,z^5 - 6\,z^4 + 3\,z^7 + 8\,z^6 - 41\,z^3 + 30)\,T^2\\ &- (z^3-1)\,(3\,z^7 + z^6 + 6\,z^5 - 12\,z^4 - 13\,z^3 - 27\,z^2 + 2)\,T\\ &- 2\,(z^2-5)\,(z^3-1)^2\,(z^2+1)^2\end{aligned}$$

$$\begin{aligned}B = &-T^5 - (-10\,z^3 + z^6 + 24 + 3\,z^2)\,T^4\\ &- (3\,z^8 - 5\,z^6 - 18\,z^5 + 3\,z^4 + 30\,z^3 - 10 + 33\,z^2)\,T^3\\ &- (-18\,z^8 + 3\,z^{10} + 84\,z^5 - 52 - 63\,z^2 - 68\,z^6 + 6\,z^4 - 6\,z^7 + 58\,z^3)\,T^2\\ &- (z^3-1)\,(z^2+1)\,(z^7 - 10\,z^5 + 3\,z^4 + 37\,z^3 + 18\,z^2 - 33)\,T\\ &- 4\,(z^3-1)^2\,(z^2+1)^3\end{aligned}$$

6.2. Recherche des solutions. Dans un premier temps, nous recherchons les valeurs de z pour lesquelles, A et B ont des solutions en T communes. Ce sont les zéros du résultant $S(z) = \mathrm{Res}_T(A, B)$. On obtient

$$S(z) = z^{12}\,(z^2-3)(z^2+1)^3\,(z^3-1)^4 S_{37}(z)$$

où

$$\begin{aligned}S_{37}(z) = &\, z^{37} - 61\,z^{34} + 336\,z^{33} - 240\,z^{32} + 2052\,z^{31} - 12120\,z^{30}\\ &+ 8400\,z^{29} - 30456\,z^{28} + 175113\,z^{27} - 88548\,z^{26} + 241040\,z^{25}\\ &- 1364385\,z^{24} + 338994\,z^{23} - 1081984\,z^{22} + 6241506\,z^{21} + 642162\,z^{20}\\ &+ 2319507\,z^{19} - 15790278\,z^{18} - 12287376\,z^{17} + 1386909\,z^{16}\\ &+ 11212992\,z^{15} + 55894536\,z^{14} - 19889496\,z^{13} + 53738964\,z^{12}\\ &- 128353329\,z^{11} + 44215308\,z^{10} - 172452240\,z^{9} + 160917273\,z^{8}\\ &- 42764598\,z^{7} + 217615248\,z^{6} - 115440795\,z^{5} + 17124210\,z^{4}\\ &- 139060395\,z^{3} + 39858075\,z^{2} + 39858075\end{aligned}$$

Pour $z = 1$, on obtient un carré à nouveau tandis que pour $z = \sqrt{3}$, nous obtenons un triangle équilatéral.

La méthode de Sturm fournit pour S_{37} trois solutions réelles, dont les valeurs approchées sont

$$z_1 = -1.4142,\ z_2 = 1.0469,\ z_3 = 1.7140.$$

Nous excluons z_1 qui est négative, z_2 qui fournit une racine T négative. Il reste z_3, pour laquelle la solution commune T vaut approximativement $T = 4.4722$.

On obtient alors une dernière configuration (triangle)

En résumé, les trois configurations obtenues correspondent aux solutions :

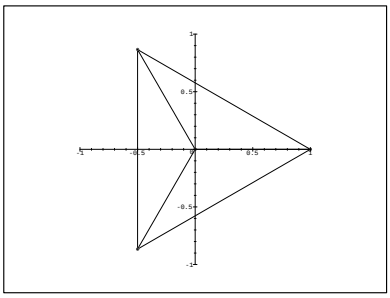


FIG. 3 – triangle équilatéral

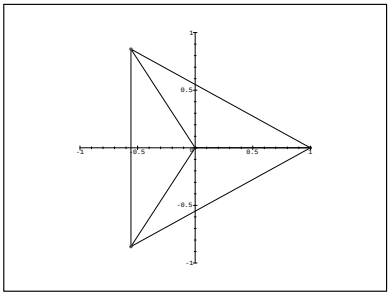


FIG. 4 – triangle

Configurations			
	Triangle équilatéral	Triangle	Carré
r_{12}	1	1	1
r_{13}	1	$\simeq 1.0223$	$1/\sqrt{2}$
r_{14}	1	$\simeq 1.0223$	$1/\sqrt{2}$
r_{23}	$\sqrt{3}$	$\simeq 1.7776$	$1/\sqrt{2}$
r_{24}	$\sqrt{3}$	$\simeq 1.7776$	$1/\sqrt{2}$
r_{34}	$\sqrt{3}$	$\simeq 1.7140$	1

Configurations centrales de quatre corps dans le plan

Suggestions de développement

Le texte a été découpé en quatre parties qui peuvent être étudiées indépendamment.

Dans la dernière partie qui concerne la mise en forme du système (10), le candidat remarquera que le système (15) exprime la coplanarité et est donc utilisable pour tout potentiel Φ. En ce qui concerne l'expression des polynômes dont on cherche des zéros, le candidat n'essaiera pas de les recopier mais de les recalculer avec un logiciel de calcul formel.

Le candidat pourra répondre à tout ou partie des questions suivantes.

1. Justifier brièvement l'équation (2).
2. Vérifier par le calcul et l'aide d'un logiciel la nullité de P.
3. Vérifier par le calcul et l'aide d'un logiciel le lemme 4.1.
4. Justifier l'équation (10)
5. Trouver les configurations centrales dans le cas d'un potentiel logarithmique $\Phi(r^2) = \log r$. Le candidat pourra suivre la même méthode que pour le potentiel newtonien.

Chapitre 13
Commentaire de l'auteur

Le problème que nous étudions est tiré de la thèse d'habilitation d'Alain Albouy : *recherches sur le problème des n corps* dont je cite des passages introductifs :
Le thème de cette étude est un vieux problème : l'étude de plusieurs massses ponctuelles soumises à l'attraction newtonienne. [...] La notion de configuration centrale *apparâit naturellement quand on recherche des solutions simples des équations différentielles qui régissent le mouvement de plusieurs masses ponctuelles soumises aux lois de Newton. Dans un mémoire lu en 1765, Euler fut le premier à expliciter de telles solutions. Voulant avertir ses lecteurs de la difficulté du problème des trois corps, il simplifia les conditions initiales pour se ramenr à un système différentiel d'ordre 3, puis lorsque l'énergie est nulle à un système d'ordre 2. Il pouvait intégrer des solutions très particulières du système d'ordre 3, les mouvements honmothétiques. C'est la configuration d'un tel mouvement qu'on appelle* centrale.

Le système d'équations associé aux lois de Newton s'écrit

$$\ddot{x} = -\nabla U(x)$$

Les configurations centrales vérifient pour un certain nombre réel λ,

$$2\lambda x = \nabla U(x).$$

La configuration $x(t)$ *reste alors au cours du temps proportionnelle à la configuration* $x(0)$ *grâce à l'homogénéïté du potentiel* U.

La propriété dynamique qui caractérise les configurations centrales, c'est qu'elles donnent naissance à un mouvement homothétique, si l'on choisit les vitesses convenablement. On appelle mouvement homothétique *un mouvement tel que la configuration ne subit pas de rotation, et change simplement de taille au cours du temps. Si l'on accepte également les rotations, on parlera de* mouvement homographique *et si l'on accepte les rotations mais pas les les changements de taille, nous parlerons de* mouvement rigide.

La détermination de telles configurations se fait en deux étapes. La première, vise à éliminer les vitesses pour se ramener à des équations sur la seule configuration (les distances mutuelles), la seconde, beaucoup plus ardue est la résolution de ces dernières équations.

Depuis la publication des travaux d'Alain Albouy, la seconde étape a intéressé des spécialistes de calcul formel car elle fournit des exemples non triviaux de systèmes d'équations polynomiales. Dans le cas présenté ici de quatre masses ponctuelles égales, la résolution peut se faire par éliminations successives grâce. à la propriété de symétrie démontrée par Albouy, utilisant des propriétés de convexité.

Les techniques utilisées pour l'étude de tels systèmes sont les bases de Gröbner. Elles permettent la résolution effective (sous réserve d'une complexité pas trop importante) du problème de la position des zéros (Nullstellensatz effectif). On peut consulter à ce sujet ...

Notons d'ailleurs que le calcul formel ne parvient pas à démontrer la symétrie dans le cas de quatre masses égales.

Dans une version optimale de ce texte, j'attendais du candidat qu'il procède de la façon suivante :

1. faire une présentation (résumé) du texte : le sujet, les méthodes employées et les résultats obtenus.
2. choisir de développer une partie (pas nécessairement d'aillaurs une suggestion proposée par l'auteur). Ici on pouvait choisir de parler des déterminants de Cayley et des liens avec le calcul barycentrique ou de la résolution d'un système de deux équations à deux inconnues (utilisation du résultant) ou de la résolution numérique d'une équation polynomiale (suites de Sturm ou autre) ou encore de justifier le système d'équations obtenu (extrema liés).
3. Proposer une illustration sur ordinateur : soit reprendre les résolutions faites dans le texte ou suggérées, soit (mais c'est peut-être plus audacieux), illustrer des développements personnels en promlongement du texte.

Quelques pistes

On peut définir en toute généralité le déterminant de Cayley par un procédure. En posant $R_{i,j} = r_{i,j}^2$,

```
for i from 1 to 5 do R.i.5:=1 od:
f := (i,j) -> if i=j then 0 elif i<j then R.i.j else R.j.i fi:
A:=matrix(5,5,f);
```

On obtient alors $A = \begin{pmatrix} 0 & R_{12} & R_{13} & R_{14} & 1 \\ R_{12} & 0 & R_{23} & R_{24} & 1 \\ R_{13} & R_{23} & 0 & R_{34} & 1 \\ R_{14} & R_{24} & R_{34} & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$ et l'expression du déterminant P qui n'est pas nul. Ensuite posant

```
long :={seq(seq(R.i.j=(x.i-x.j)^2+(y.i-y.j)^2,
        j.=i+1..4),i=1..3)};
```

on obtient bien 0 par la commande

```
factor(subs(long,P));
```

Signalons qu'à partir de la matrice A, on peut répondre à la question : *il n'exsite pas de configuration de quatre points dans le plan qui soient tous à des distances impaires les uns des autres.* Il suffit alors de remplacer les R_{ij} par $(1 + 2b_{ij})^2$ et de remarquer alors que $P = 4[8]$ ne peut être nul. Le

lecteur intéressé remarquera que MAPLE échoue à calculer le déterminant de P modulo 8 (`Det(P) mod 8`) car un des pivots obtenus est un diviseur de 0. Cette question est tiré de l'article ([?])

Pour fabriquer les Δ_i, on peut fabriquer la matrice

```
B := matrix(4,4,[0$4,1$4,x.1,x.2,x.3,x.4,y.1,y.2,y.3,y.4]);
```

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{pmatrix}$$

puis

```
delta := i -> (-1)^(i+1)/2*factor(det(minor(B,1,i)));
```

On vérifie ensuite que la fonction

```
dP := (i,j) -> factor(subs(long,diff(P,R.i.j))
               +32*delta(i)*delta(j));
```

est nulle, c'est-à-dire, que $\dfrac{\partial P}{\partial r_{ij}^2} = -32\Delta_i\Delta_j$. L'équation (10) résulte des extrema liés. On aurait en fait $w\Phi'(r_{ij}^2) = u + v\Delta_i\Delta_j$ Si $w = 0$, alors on a $u + v\Delta_i\Delta_j$. Mais $\Delta_i\Delta_j$ est de signe constant ce qui est impossible sauf si les Δ_i sont nuls ce qui est exclu.

Dans le cas $\Phi(r^2) = \log(r)$, on obtient avec $a = 1$ et $f = z^2$, en suivant le même chemin que pour le potentiel newtonien,

$$\left\{\begin{array}{lcl} \Delta_1 = \Delta_3(-1-t) & , & \Delta_2 = \Delta_3(-1+t) \\ 4b = f + a(1-t)^2 & , & 4d = f + (1+t)^2 a. \end{array}\right. \quad \left\{\begin{array}{rcl} 1/a = 1 & = & u + v\Delta_3^2(1-t^2), \\ 1/b & = & u - v\Delta_3^2(1+t), \\ 1/f & = & u + v\Delta_3^2, \\ 1/d & = & u - v\Delta_3^2(1-t). \end{array}\right.$$

Puis $u = (t^2 + z - 1)/zt^2, v = (1-z)/zt^2$. On obtient alors, en posant $T = t^2$,

$$A = -(-2\,f + t^2 + 2)\,(-f + t^2 - 1), B = -(-3 + f)\,(-1 + f + t^2)$$

et

$$S = \mathrm{Res}_T(A, B) = -6\,(-3 + f)^2\,(-1 + f)\,f$$

On ne trouve que le carré ($f = 1$) et le triangle ($f = 3$) équilatéral comme solutions.

Petite bibliographie

[1] ALBOUY, A., Symétrie des configurations centrales de quatre corps, C. R. Acad. Sci., Paris **320** (1995), pp. 217-220

[2] ALBOUY, A., Recherches sur le problème des n corps, Notes scientifiques et techniques du Bureau des Longitudes (1997), 112 p.

[**3**] Ph. Dumas, B. Salvy, Maple and the Putnam Competition, RT 0179, (1995), INRIA (Rocquencourt, Rennes, Sophia, Nancy, Grenoble), disp. sur `//algo.inria.fr/dumas/DuSa95/Gif/`

[**4**] Dziobek O. Über einen merkwürdigen Fall des Vielkörperproblems, Astron. Nach. **152** (1900), pp. 33-46

[**5**] J.- C.Faugère, I. Kotsireas, Symmetry theorems for the newtonian 4- and 5-body problems with equal masses, Computer Algebra in Scientific Computing CASC'99, (Munich), Springer-Verlag, Ganzha, Mayr, Vorozhtsov eds. pp. 81-92.

[**6**] I. Kotsireas, D. Lazard, Central Configurations of the 5-body problem with equal masses in three-dimensional space, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 1999 (258), pp. 292-317.

[**7**] Moeckel, R., Generic finiteness for Dziobek configurations. Trans. Amer. Math. Soc. **353** (2001), 11, pp. 4673–4686

Chapitre 14
Commentaire d'un spécialiste

L'exposé doit montrer que le candidat a compris le plan du texte et doit contenir un développement d'une des quatre parties. Dans la suite, un développement possible est proposé.
Le texte considéré permet au candidat de faire nombreux commentaires introductifs sur le problème des n corps, d'un point de vue historique et mathématique : l'exemple de 2 corps (soleil et terre) est probablement connu par les étudiants, et peut être très agréable à écouter (par exemple la première loi de Kepler (voir [HS] page 23-26, ou les premières trente pages de [MH] pour une introduction plus générale)).
Une lecture rapide suffit au candidat pour s'apercevoir que le but du texte est de trouver des équations algébriques qui déterminent les configurations centrales des quatre corps. Le lecteur identifie les inconnues du problème, qui sont les distances relatives des quatre particules. Le problème est ramené à la recherche des points critiques de la fonction des forces restreinte aux configurations planes de moment d'inertie fixé. On traduit la condition de coplanarité à l'aide du déterminant de Cayley, qui exprime des conditions algébriques sur les distances entre les particules. On obtient donc un problème d'extréma liés avec des contraintes algébriques. On se limite au cas de quatre masses égales et l'on écrit explicitement les équations algébriques pour un potentiel quelconque. On suppose qu'il existe un axe de symétrie contenant deux des corps, ce qui réduit le nombre des inconnues. On impose alors que le potentiel soit newtonien et l'on obtient un système, explicitement écrit dans le texte, de deux équations polynomiales à deux inconnues. On cherche les racines du système avec la méthode du résultant, et l'on obtient trois configurations.
On pourra demander au candidat d'être critique par rapport aux simplifications du texte (nombre des particules, masses égales, symétrie).
Un choix possible pour le développement d'une partie du texte est la dernière question proposée par l'auteur. On écrit un programme maple pour déterminer les configurations centrales dans le cas du potentiel newtonien et ensuite on le modifie en choisissant un potentiel logarithmique. Ce choix permettra au candidat de parler de résolution exacte ou approchée de systèmes (par exemple : résultant, suite de Sturm, Newton) et de montrer ses connaissances dans la mise en place d'un programme (par exemple : resultant, sturmseq, fsolve, solve). On remarquera que le choix du potentiel logarithmique simplifie beaucoup les calculs. On donne un exemple de programme :

```
Premier potentiel phi(z^2)=1/z
entree du potentiel
> phi:=z**(-1/2):
> phi1:=diff(phi,z):
expression de b et d en fonction de z et t d'apres (15)
```

```
> solbd:={b=1/4*(z**2+(1-t)**2),d=1/4*(z**2+(1+t)**2)}:
entree des 4 equations (17)
>
>eq17:=[subs(z=1,phi1)-(u+v*(1-t**2)),subs(z=b,phi1)-(u-v*(1+t)),subs(z=z**2,phi1
>)-(u+v),subs(z=d,phi1)-(u-v*(1-t))]:
expression de u et v en fonction de z et t d'apres les equations 1 et 3 de (17)
> soluv:=solve({eq17[1],eq17[3]},{u,v}):
reecriture des equations 2 et 4 de (17)
> eq17bis2:=b^(-3)-(solve(subs(b**(-3/2)=b32,eq17[2]),b32))**2:
> eq17bis4:=d^(-3)-(solve(subs(d**(-3/2)=d32,eq17[4]),d32))**2:
remplacement de u,v,b et d par leurs valeurs en fonction de z et t dans les
equation 2 et 4 de (17)
> Pzt:=collect(numer(subs(solbd union soluv,eq17bis2)),t,factor):
> Pzmt:=collect(numer(subs(solbd union soluv,eq17bis4)),t,factor):
> simplify(subs(t=-t,Pzmt)-Pzt):
polynomes A et B
>
> A:=subs(t=T^(1/2), collect(1/4/t*(Pzt-Pzmt),t,factor));
> B:=subs(t=T^(1/2), collect(1/2*(Pzt+Pzmt),t,factor));
resultant
> S:=factor(resultant(A,B,T)):
> S37:=sort(op(4,S),z):
> {fsolve(S)}:
> {fsolve(S37)}:
> subs(z=1.714000326,A):fsolve(%):subs(z=1.714000326,B):fsolve(%):
> gcd(subs(z=1.714000326,A),subs(z=1.714000326,B)):
> {fsolve(subs(z=1,A))},{fsolve(subs(z=1,B))}:
> {fsolve(subs(z=sqrt(3),A))},{fsolve(subs(z=sqrt(3),B))}:
Deuxieme potentiel: phi(z^2)=ln(z)
entree du potentiel
> phi:=(1/2*ln(z)):
> phi1:=diff(phi,z):
expression de b et d en fonction de z et t
> solbd:={b=(z**2+(1-t)**2)/4,d=(z**2+(1+t)**2)/4}:
entree des equations (17)
>
>eq17:=eval([subs(z=1,phi1)-(u+v*(1-t**2)),subs(z=b,phi1)-(u-v*(1+t)),subs(z=z**2
>,phi1)-(u+v),subs(z=d,phi1)-(u-v*(1-t))]):
expression de u et v en fonction de z et t d'apres equations 1 et 3 de (17)
> soluv:=(solve({eq17[1],eq17[3]},{u,v})):
calcul de Pzt en remplacant u,v,b et d par leurs valeurs dans les equations
2 et 4 de (17)
> Pzt:=collect(numer(subs([op(soluv),op(solbd)],eq17[2])),t,factor):
> Pzmt:=collect(numer(subs([op(soluv),op(solbd)],eq17[4])),t,factor):
> simplify(subs(t=-t,Pzmt)-Pzt):
```

```
calcul de A et B
> A:=subs(t=T^(1/2), collect(1/4/t*(Pzt-Pzmt),t,factor));
> B:=subs(t=T^(1/2), collect(1/2*(Pzt+Pzmt),t,factor));

resultant de A et B
> S:=factor(resultant(A,B,T)):
> {fsolve(subs(z=1,A))},{fsolve(subs(z=1,B))}:
> {fsolve(subs(z=sqrt(3),A))},{fsolve(subs(z=sqrt(3),B))}:
> {fsolve(subs(z=0,A))},{fsolve(subs(z=0,B))}:
```

Références : M.W.Hirsch and S. Smale, Differential Equations, Dynamical Systems, and Linear Algebra. Academic Press. 1974.
K.R. Meyer and G.R. Hall, Introduction to Hamiltonian Dynamical Systems and the N-Body Problem. AMS 90, Springer Verlag. 1992.

Cinquième partie

Modélisation du système ALOHA

Chapitre 15
Modélisation du système ALOHA

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Les candidat(e)s sont laissé(e)s libres d'organiser leur discussion comme ils ou elles l'entendent. Des suggestions de développement, largement indépendantes les unes des autres, sont proposées. Le candidat n'est pas tenu de les suivre. Le jury appréciera que la discussion soit accompagnée d'exemples traités sur ordinateur.

Résumé

On présente un système de gestion des communications radio entre un ensemble d'utilisateurs et un satellite à travers un canal unique. Ce système élémentaire est fondé sur l'émission aléatoire des informations par les utilisateurs. Le modèle est une chaîne de Markov homogène, instable, mais dont le temps de fonctionnement normal est cependant très long. On décrit succintement des méthodes de simulation possibles.

Mots clés : chaîne de Markov, transience, récurrence, estimation du paramètre d'une loi de Bernoulli.

1. Présentation d' ALOHA.

Dans les années 70, les ordinateurs étant reliés à un nombre croissant de terminaux à des distances pouvant être très importantes, il est devenu nécessaire (pour des raisons de coût et de complexité des réseaux) de trouver une alternative aux réseaux à câbles.

Les communications radio avec un satellite relais constituent une alternative. Cependant l'encombrement du réseau hertzien exige une occupation optimale de chaque canal, en cherchant simultanément à minimiser les risques d'attentes longues. La nature des communications le permettant, chaque canal est partagé entre un très grand nombre d'utilisateurs. La transmission se fait par "paquets d'information" de longueur fixée. Il se peut que de nombreux utilisateurs envoient simultanément des "paquets d'information", ce qui crée des interférences nuisibles. Il est donc nécessaire d'avoir une politique d'émission des "paquets" et de résolution des "conflits".

Le système *ALOHA*, étudié pour un grand nombre de terminaux sans beaucoup de mémoire interne (par exemple des consoles), est basé sur l'accès aléatoire des paquets non transmis pour cause de conflit. Un algorithme de priorité de type déterministe éviterait certes toute perte de message, mais au prix d'une forte réduction de débit.

Fonctionnement du système ALOHA discrétisé.

- Le temps de référence (temps canal) est découpé en tranches égales à la durée d'un paquet (les paquets étant supposés de taille constante), chaque tranche est de 10^{-6} seconde environ.

- Les débuts de transmission des paquets sont synchronisés de façon à n'être permis qu'au début d'une tranche.

- Chaque fois que deux paquets ou plus sont émis simultanément, ils interfèrent et ils sont perdus. Ils doivent alors être réémis aléatoirement par les stations concernées, dans l'espoir d'éviter un nouveau conflit. Le choix de la politique de réémission est essentiel : il faut limiter l'engorgement des mémoires-tampons où sont stockés provisoirement les paquets (limiter les débordements de "buffer").

2. Un modèle mathématique très simple : ALOHA non stabilisé.

Nous considérons un modèle avec M usagers indépendants (M est très grand). A un instant (supposé entier par discrétisation) donné, chaque usager a deux états possibles :

– actif : il peut émettre un paquet avec une loi de probabilité de Bernoulli de paramètre σ ; σ caractérise le taux d'émission des usagers. A ce stade il n'y a pas de tirage aléatoire de régulation : le paquet est directement envoyé.

– bloqué : il vient d'émettre (ou il a déjà émis auparavant) et le paquet qu'il a déjà émis est bloqué ; il le réémet à chaque instant suivant une loi de Bernoulli de paramètre p (constant dans ce modèle) ; p caractérise la politique (élémentaire ici) de réémission.

On suppose M, σ, p indépendants du temps.

Soit $N(t)$ le nombre de terminaux bloqués à l'instant t. terminaux actifs à Le processus $N(t)$ caractérise l'état du système, et c'est une chaîne de Markov. En distinguant les différents cas possibles (aucune nouvelle émission avec ou sans réémission d'un ou plusieurs terminaux bloqués, une nouvelle émission sans ou avec réémission d'un terminal bloqué, etc.) on obtient les probabilités de transition suivantes (notons que le canal ne transportant qu'un seul message à la fois, le débloquage se fait un par un) : question que tu

$$p_{ij} = \text{Prob}\big(N(t+1) = j | N(t) = i\big) \qquad (\text{pour } i = 0, ..., M)$$

$$\begin{array}{rcl}
= & 0 & \text{si } j \leq i-2 \\
= & ip(1-p)^{i-1}(1-\sigma)^{M-i} & \text{si } j = i-1 \\
= & (1-p)^i (M-i)\sigma(1-\sigma)^{M-i-1} & \\
 & +\big[1 - ip(1-p)^{i-1}\big](1-\sigma)^{M-i} & \text{si } j = i \\
= & (M-i)\sigma(1-\sigma)^{M-i-1}\big[1-(1-p)^i\big] & \text{si } j = i+1 \\
= & C_{M-i}^{j-i}\,\sigma^{j-i}(1-\sigma)^{M-j} & \text{si } j \geq i+2
\end{array}$$

Cette chaîne de Markov est récurrente et lorsque M est grand on constate que le système finit par se stabiliser aux alentours de M : la plupart des terminaux sont bloqués.
Puisque dans notre problème M est naturellement très grand, on est conduit à considérer le modèle "infini" dans lequel $M \to +\infty$ et $\sigma \to 0$ de telle sorte que $M\sigma \to \alpha$ fini. Les probabilités de transition deviennent (pour $i \geq 0$, $j \geq 0$) :

$$\begin{array}{rcl}
p_{ij} = & 0 & \text{pour } j \leq i-2 \\
= & ip(1-p)^{i-1}\,e^{-\alpha} & \text{pour } j = i-1 \\
= & (1-p)^i \alpha\, e^{-\alpha} + \big[1 - ip(1-p)^{i-1}\big]e^{-\alpha} & \text{pour } j = i \\
= & \alpha\, e^{-\alpha}\big[1-(1-p)^i\big] & \text{pour } j = i+1 \\
= & \frac{\alpha^{j-i}}{(j-i)!}\,e^{-\alpha} & \text{pour } j \geq i+2
\end{array}$$

Cela correspond à un flux d'émissions fraîches Poissonnien de paramètre α.

Dans ce cas, la chaîne de Markov $N(t)$ est transiente pour toutes valeurs de α et p. Autrement dit le canal n'est pas "stable", et s'engorge irrémédiablement (c'est-à-dire que le nombre de terminaux bloqués tend vers l'infini). Le temps

avant que ce phénomène soit observé dépend de la valeur des paramètres. A cet égard l'étude du **comportement transitoire** du système est intéressante. Les simulations faites montrent que suivant les valeurs des paramètres, $N(t)$ reste très longtemps "stable" autour d'une valeur n_0, avant de croître assez rapidement vers l'infini (le canal se **sature**) dès qu'il dépasse une certaine valeur n_c.

Les points n_0 et n_c sont des points d'équilibre en moyenne, l'un stable, l'autre instable, définis par l'équation

$$E\big(N(t+1) - N(t)|N(t) = n\big) = 0.$$

Remarquons que ces points ne sont pas nécessairement entiers. Cependant, quitte à les remplacer par l'entier le plus proche, nous les supposons entiers.

Par exemple on obtient les valeurs de n_0, n_c suivantes pour quelques cas :

p	α	n_0	n_c
0.095	0.270	2	18
0.049	0.270	3	35
0.024	0.270	6	70
0.095	0.300	2	15
0.049	0.300	4	30
0.024	0.300	8	60
0.095	0.330	3	13
0.049	0.330	6	25
0.024	0.330	11	49
0.095	0.360	4	10
0.049	0.360	8	18
0.024	0.360	17	36

Posons $b(n) = E\big(N(t+1) - N(t)|N(t) = n\big)$ l'accroissement moyen du nombre de bloqués si il y en a déjà n. On a $b(n_0) \simeq 0$ et $b(n_c) \simeq 0$, mais aussi $0 < b(n)$ si $0 \leq n < n_0$ ou $n > n_c$, et $b(n) < 0$ si $n_0 < n < n_c$. L'équilibre n_0 est stable tandis que n_c est instable (voir figure 1).

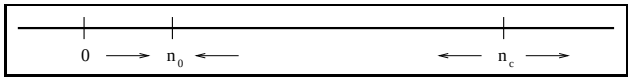


Figure 1.

Une mesure de la stabilité du système est donnée par le temps moyen d'atteinte de n_c partant du point n_0. Ce temps peut être très grand pour des valeurs bien choisies des paramètres, c'est **le temps de fonctionnement** du système.

3. Estimation par simulation.

On cherche donc a obtenir une valeur du temps moyen avant blocage, qui est le temps moyen avant de dépasser n_c. On suppose partir d'une situation de fonctionnement correct, on choisit n_0. On décrit ici une simulation.

Plusieurs approches sont possibles. Relativement au temps : ou bien on fixe un temps maximum de simulation ou bien on fixe un nombre prédéterminé de passages par n_c. Dans le second cas le temps de simulation est aléatoire. Relativement à la précision : on se fixe une erreur (relative) et une valeur de la confiance pour cette erreur.

• Si on fixe l, le nombre de passages, on simule aisément ces l passages obtenant un l-échantillon de la loi du temps de passage partant de n_0, ce qui fournit un estimateur de l'espérance ainsi qu'un intervalle de confiance asymptotique de niveau désiré.

• Si on fixe le nombre de pas de la simulation, c'est-à-dire qu'on effectue au total t tirages aléatoires, alors le nombre de passages qu'on observe est une variable aléatoire L_t. On dispose donc de $(T_i)_{1\leq i\leq L_t}$ et on construit un estimateur $\hat{T}_t = \frac{\sum_1^{L_t} T_i}{L_t}$ qui, lorsque t tend vers l'infini, converge $p.s.$ vers le temps moyen cherché.

En effet, L_t est un processus de renouvellement et nous savons que :

i) $\frac{L_t}{t}$ converge $p.s.$ vers $1/ET_1$. D'où l'on déduit aisément la convergence $p.s.$ de $\hat{T}_t$.

ii) $\sqrt{t}(\frac{L_t}{t} - \frac{1}{ET_1})$ converge en loi vers une gaussienne centrée de variance $var(T_1)/ET_1^3$. D'où l'on déduit la convergence en loi de $\sqrt{t}(\hat{T}_t - ET_1)$ vers une gaussienne centrée de variance $ET_1 \times var(T_1)$ et un intervalle de confiance pour ET_1.

Cependant on observe que dans des circonstances raisonnables le temps de passage est très grand, ce qui ne permet pas d'obtenir une précision acceptable compte tenu de la qualité des générateurs de nombres aléatoires. Cela est lié à la notion d'événement rare et nécessite des techniques de simulation rapide.

Indications, sans obligation

1. Précisez en détail comment sont obtenues les probabilités de transition p_{ij} dans les cas fini et infini.

2. A titre d'exemple donnez la matrice de transition p_{ij} dans les cas $M = 2$ et $M = 3$ ou plus. Indiquez la nature de la chaîne : est-elle irréductible, admet-elle une loi invariante, etc. ?

3. Donnez une interprétation de la figure 1.

4. Détaillez les calculs pour obtenir les intervalles de confiance pour les différents choix de simulation.

5. Proposez une méthode analytique pour obtenir le temps moyen avant blocage dans le cas M fini ou infini.

6. Réalisez une simulation du système dans le cas fini pour des valeurs de M de votre choix.

7. Réalisez cette simulation dans le cas infini. Est-il possible de calculer par simulation le temps moyen avant blocage ?

Chapitre 16
Commentaire de l'auteur

Le texte propose un modèle simple de transmission de "paquets" d'information sur un canal. Ce modèle conduit à une chaîne de Markov à espace d'états fini. Lorsque le nombre d'utilisateurs du canal est très grand une modèlisation par une chaîne de Markov à espace d'états dénombrable est plus commode. L'introduction décrit le fonctionnement du système. A partir de cette description il est aisé de comprendre que la chaîne de Markov pertinente est $N(t)$ le nombre de "terminaux" souhaitant envoyer un paquet d'information mais n'ayant pas pu le faire à l'instant t.
On passe du cas fini au cas dénombrable en approchant la loi des arrivées des nouvelles demandes d'émission qui est une loi binômiale $B(M,p)$, par une loi de Poisson de paramètre α. Cette approximation se justifie par le résultat suivant bien connu : si X_n suit une loi $B(n,\alpha/n)$, X_n converge en loi vers une loi de Poisson de paramètre α. Lorsque M est grand et p petit on applique avec $\alpha = Mp$.

Dans les deux cas (fini et infini dénombrable) les lois de transition sont obtenues en remarquant qu'au plus un terminal peut émettre, ce qui conduit à $P_{ij} = 0$ si $j \leq i-2$, puis en distinguant 4 cas :

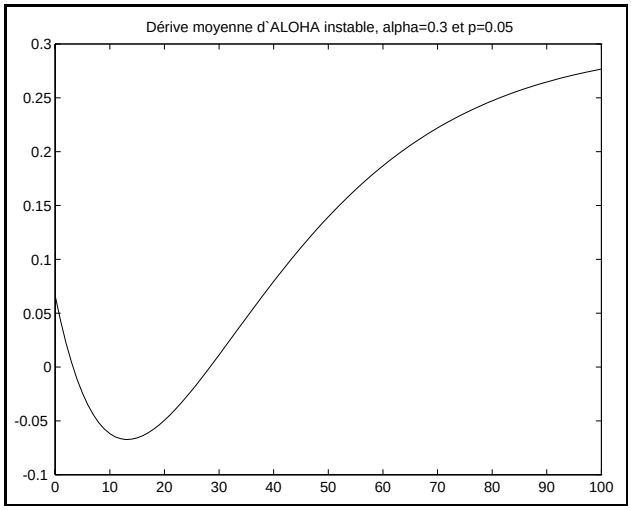
1. Un seul terminal bloqué émet et aucun nouveau message ne se présente, ce qui fait passer de $N(t)$ à $N(t+1) = N(t) - 1$.
2. Un seul nouveau message arrive et aucun terminal bloqué ne choisit d'émettre, auquel cas $N(t) = N(t+1)$. Cela se produit aussi si aucun nouveau message n'arrive mais que plusieurs terminaux bloqués cherchent à émettre.
3. Un nouveau message arrive alors qu'un terminal bloqué a choisi d'émettre ;
4. Enfin, plusieurs nouveaux messages arrivent.

Finie ou infinie, la chaîne est irreductible car tous les états (k) communiquent avec 0 (en k étapes de temps) et zéro communique avec tous les états $k \leq 2$ (en une étape) et avec 1 (en deux étapes).
Dans le cas fini il s'ensuit que la chaîne est apériodique récurrente positive.
Dans le cas dénombrable cela est moins évident et le texte donne un argument intuitif pour montrer la transience :

Considérons la quantité $b(n) = E(N(t+1) - N(t)|N(t) = n$, cette fonction décrit "l'évolution moyenne" de $N(t)$ d'un instant à l'autre.
Le calcul de b est assez facile puisqu'à deux termes prés il s'agit d'une espérance de loi de Poisson (égale au paramètre). On obtient

$$b(n) = \alpha - e^{-\alpha}(1-p)^{n-1}((1-p)\alpha + np).$$

Prenons (chiffres voisins de ceux du texte) $p = 0,05$ et $\alpha = 0,3$, la courbe représentative de b est la suivante :



On en déduit bien que pour $4 < n < 30$ $N(t)$ diminue en moyenne, et augmente sinon. Donc la chaîne de Markov (toujours en moyenne) se rapproche de 4 tant que $n < 30$ et augmente (toujours en moyenne) au delà de 30. Cette description qualitative est en réalité une comparaison à l'équation differentielle (virtuelle) $\dfrac{dx}{dt} = b(x)$ qui posséde un point déquilibre localement stable (proche de 4), dont le bassin d'attraction est à peu près $[0, 30]$ et un point d'équilibre instable (proche de 30).
On en tire la description qualitative du fonctionnement du système qui suit : le nombre de terminaux bloqués oscille autour de 4 puis vient "par hasard" à dépasser 30 ce qui conduit à l'engorgement.

Le calcul de $b(t)$ est un argument suffisant pour montrer la transience en faisant appel à un résultat du type théorème de Foster, mais cela n'était pas demandé du tout. On peut aussi, pour montrer la non récurrence, supposer l'existence d'une distribution de probabilité stationnaire et faire apparaître une contradiction, par exemple sur la fonction génratrice de la distribution stationnaire, en utilisant (si $(\pi_j, j \geq 0)$ est la distribution stationnaire) la relation

$$\pi_j = \sum_{k=0}^{j} \pi_{j-k} \frac{\alpha^k e^{-\alpha}}{k!} - \pi_{j-1} \alpha e^{-\alpha} (1-p)^{j-1} + \pi_j \alpha e^{-\alpha} (1-p)^j$$
$$-\pi_j e^{-\alpha} jp(1-p)^{j-1} + \pi_{j+1} e^{-\alpha} (j+1) p (1-p)^j.$$

La fonction génératrice $\pi(s)$ vérifie alors : $\pi(s) = \pi(s) e^{\alpha(s-1)} + \alpha e^{-\alpha} \dfrac{1-s}{1-(1-p)s} + e^{-\alpha} \dfrac{p}{(1-(1-p)s)^2}$ et la limite à gauche $\pi(1\_)$ est infinie.

La suite du texte s'intéresse à la simulation du système pour éventuellement estimer le temps moyen avant engorgement, que l'on assimile au temps de

dépassement de l'équilibre instable n_c. En effet si ce temps moyen est grand un système aussi simple peut être utilisé avec une "remise à zéro" en cas d'engorgement.
Dans cette partie on utilise des méthodes et des résultats bien connus pour l'essentiel.

Dans le cas où le nombre d'engorgements est fixé à l'avance le temps de simulation devient une variable aléatoire, mais un intervalle de confiance du temps moyen cherché est immédiatement obtenu. Soient $T_1, ..., T_l$ les temps observés, $\bar{T}_l$ leur moyenne empirique et S_l^2 leur variance empirique un intervalle de confiance pour ET_1 au risque asymptotique de 5% est $\left[\bar{T}_l - 1,96\sqrt{\frac{S_l^2}{l}}, \bar{T}_l + 1,96\sqrt{\frac{S_l^2}{l}}\right]$.

Si le temps de simulation est fixé, c'est le nombre d'engorgements observés qui est une variable aléatoire, L_t. Les instants $T_1, T_1 + T_2, ..., T_1 + ... + T_{L_t}$ sont les instants de saut d'un processus de renouvellement pour lequel nous disposons d'un théorème de la limite centrale brièvement énoncé dans le texte. On écrit alors :

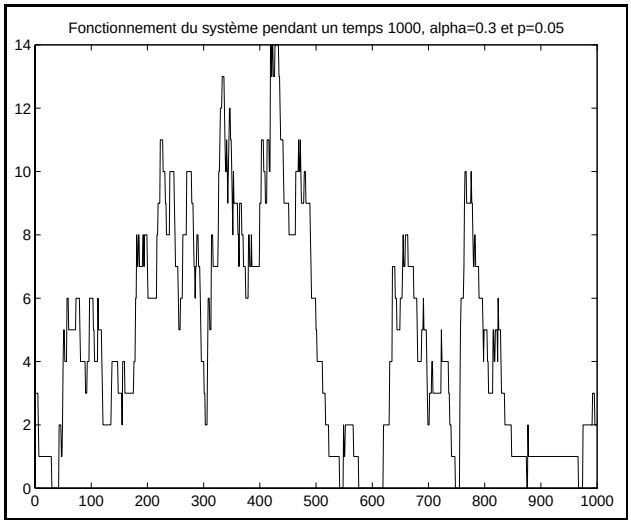
$$\sqrt{t}(\hat{T}_t - ET_1) = \sqrt{t}(\frac{\sum_{i=1}^{L_t} T_i}{L_t} - ET_1) = \frac{\sqrt{t}}{L_t}(\sum_{i=1}^{L_t} T_i - t) + \sqrt{t}(\frac{t}{L_t} - ET_1).$$

Dans le membre de droite le premier terme tends $p.s.$ vers 0 car $\frac{t}{L_t} - ET_1$ tend $p.s.$ vers 0, $|\sum_{i=1}^{L_t} T_i - t| \leq T_{L_t+1}$ et $\frac{T_{L_t+1}}{t^a}$ tend vers 0 en probabilité pour tout $a > 0$. Quant au second terme on peut l'écrire :

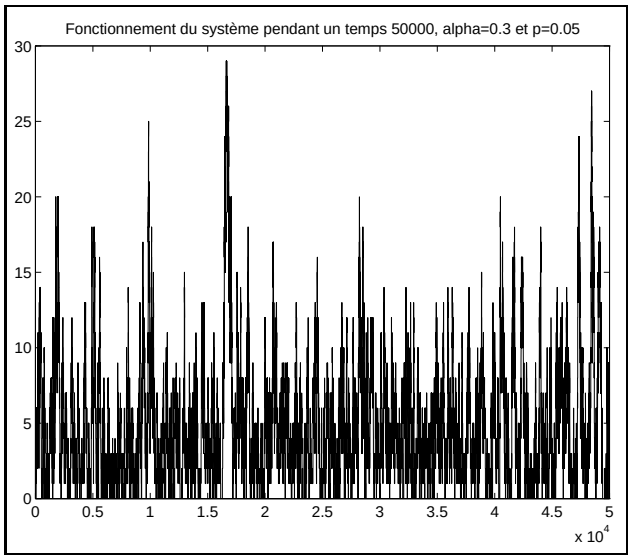
$$\sqrt{t}(\frac{t}{L_t} - ET_1) = \sqrt{t}(\frac{\frac{1}{ET_1} - \frac{L_t}{t}}{\frac{L_t}{t}\frac{1}{ET_1}}).$$

Comme $\sqrt{t}(\frac{L_t}{t} - \frac{1}{ET_1})$ converge en loi vers une $N(0, \frac{var(T_1)}{ET_1{}^3})$ et $\frac{L_t}{t}$ converge $p.s.$ vers $\frac{1}{ET_1}$, $\sqrt{t}(\frac{t}{L_t} - ET_1)$ converge en loi vers une $N(0, \frac{var(T_1)ET_1{}^4}{ET_1{}^3})$. Un intervalle de confiance au risque asymptotique 5% est donc $\left[\hat{T}_t - 1,96\sqrt{\frac{\hat{T}_t S_t^2}{t}}, \hat{T}_t + 1,96\sqrt{\frac{\hat{T}_t S_t^2}{t}}\right]$, où S_t^2 est la variance empirique des L_t temps observés.

La réalisation d'une simulation dans le cas infini est faisable et on obtient ($\alpha = 0,3$ et $p = 0,05$) le genre de trajectoire ci-dessous, ici pour un temps 1000.



La même simulation avec un temps 50.000 ne fait toujours pas apparaître d'engorgement.



Le temps moyen recherché est donc très long et une accelération de la simulation est souhaitable.

Chapitre 17
Commentaire d'un spécialiste

Le commentaire qui suit est très complet. Il n'est pas demandé aux candidats à l'agrégation d'en faire autant lors de l'épreuve : le temps de préparation ne le permettrait probablement pas.

Un modèle simple d'émission d'information est celui d'utilisateurs émettant à chaque instant de façon aléatoire et indépendante des paquets d'information. Dans le cas d'un canal de transmission unique, si plusieurs paquets sont émis simultanément, ils interfèrent et l'information est perdue. Ces paquets sont alors bloqués et doivent être réémis ultérieurement.

§1. Les probabilités de transition

Si n utilisateurs émettent à un instant donné indépendamment avec une probabilité $0 < \sigma < 1$, il passe dans le canal un nombre aléatoire Y de paquets qui suit une loi binômiale $\mathcal{B}(n, \sigma)$. On attend donc en moyenne $\mathbb{E}Y = n\sigma$ paquets par instant. Pour que le système fonctionne efficacement, il est nécessaire que $\mathbb{P}(Y = 1) = n\sigma(1-\sigma)^{n-1}$ soit grand, puisque $\{Y = 1\}$ est le seul évènement permettant une émission (si l'on ne tient pas compte de la réémission des paquets bloqués). Cette fonction de n et σ est maximale lorsque $1 = n\sigma (= \mathbb{E}Y)$ et n tend vers l'infini (ce qui impose que σ tende vers zéro). Cette asymptotique est celle de l'approximation binômiale-Poisson et il y a convergence étroite de $\mathcal{B}(n, \sigma)$ vers la loi $\mathcal{P}(1)$ de Poisson de paramètre 1. En pratique on essaiera donc d'allouer un canal pour un nombre d'utilisateurs de l'ordre de $1/\sigma$. Notons que faire tendre σ vers zéro revient à faire décroître la durée d'inter-émission.

Avec la politique d'Aloha de réémettre les paquets bloqués à chaque instant $t = 0, 1, \ldots$ indépendamment avec la probabilité $0 < p < 1$, le nombre N_t d'utilisateurs bloqués (c'est-à dire en possession d'un paquet bloqué) est une chaîne de Markov d'ensemble d'états $\{0, 1, \ldots, M\}$ où M est le nombre d'utilisateurs. Sa matrice de transition est notée $P = (P(i,j); i, j \in \{0, 1, \ldots, M\})$. Nous interprétons par exemple les termes diagonaux de la première matrice stochastique du texte par

$$P(i,i) = \mathbb{P}(\{X = 0, Y = 1\} \cup \{X \neq 1, Y = 0\}) \tag{1}$$

où X et Y sont des variables aléatoires indépendantes telles que $X \sim \mathcal{B}(i, p)$ représente le nombre de paquets bloqués qui sont réémis et $Y \sim \mathcal{B}(M - i, \sigma)$

représente celui de paquets actifs (non bloqués) qui sont émis à un instant donné.

La seconde matrice stochastique du texte correspond au cas limite $M = \infty$ dans l'asymptotique de l'approximation binômiale-Poisson où $\lim_{M\to\infty} M\sigma = \alpha$. Les $e^{-\alpha}$ proviennent de la limite $\lim_{M\to\infty}(1-(\sigma+o(1))/M)^M = e^{-\alpha}$. Les termes de cette matrice s'interprètent comme pour (2.1) à la différence que cette fois Y suit la loi limite $\mathcal{P}(\alpha)$ de Poisson de paramètre $\mathbb{E}Y = \alpha$.

Nous présentons maintenant des résultats liés à la simulation de la chaîne finie (à $M < \infty$ utilisateurs). Nous travaillons avec Matlab, les textes des programmes sont donnés en appendice. Les figures que nous présentons plus bas correspondent à des calculs effectués avec les paramètres

$$M = 30, \quad \sigma = 0.012, \quad p = 0.2 \tag{2}$$

ce qui donne $M\sigma = 0.36$ de l'ordre de l'unité et inférieur à 1 pour éviter un engorgement trop rapide.

Le programme `aloha1.m` (voir la Section 6.2) permet d'implémenter la première matrice stochastique P et de calculer la Figure 1 où sont donnés les représentations graphiques de P en 3D, de sa diagonale $i \mapsto P(i,i)$, de sa sous-diagonale $i \mapsto P(i,i-1)$, de sa sur-diagonale $i \mapsto P(i,i+1)$ et de la masse restante $i \mapsto P(i, j \geq i+2)$. Grâce à cette dernière indication, on constate que pour des grandes valeurs de M avec $M\sigma$ de l'ordre de l'unité, la matrice P est "essentiellement" tridiagonale. On remarque aussi que l'essentiel de la masse est porté par la diagonale ; une forte diagonale ralentit l'évolution : la chaîne aime faire du surplace.

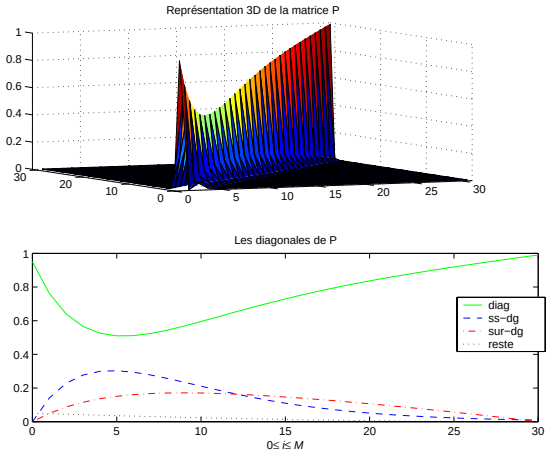


FIG. 1 – La matrice stochastique

§2. Dynamique et équilibres

La dynamique moyenne de la chaîne $(N_t; t \geq 0)$ est décrite par le “champ de force" b donné par

$$b(i) = \mathbb{E}(N_{t+1} - N_t \mid N_t = i),\ 0 \leq i \leq m. \tag{3}$$

Ce qui signifie que $N_{t+1} - N_t = b(N_t) + \Delta M_t$, où ΔM_t est l'accroissement d'une martingale (pour la filtration naturelle). Autrement dit, (N_t) résulte d'une perturbation aléatoire de la suite récurrente (n_t) dont l'évolution est spécifiée par

$$n_{t+1} - n_t = b(n_t),\ t = 0, 1, \ldots \tag{4}$$

Le programme `aloha1.m` calcule aussi le champ b ainsi que le potentiel V dont “dérive" b, soit pour tout $0 \leq i \leq M$,

$$b(i) = \sum_{0 \leq j \leq M} P(i,j)(j-i) \tag{5}$$

et

$$V(i) = -\sum_{0 \leq j \leq i} \frac{b(j-1)+b(j)}{2} \tag{6}$$

dont l'interprétation graphique est plus directe que celle de b. Les représentations graphiques de b et de V sont réalisées par le programme et données à la Figure 2.

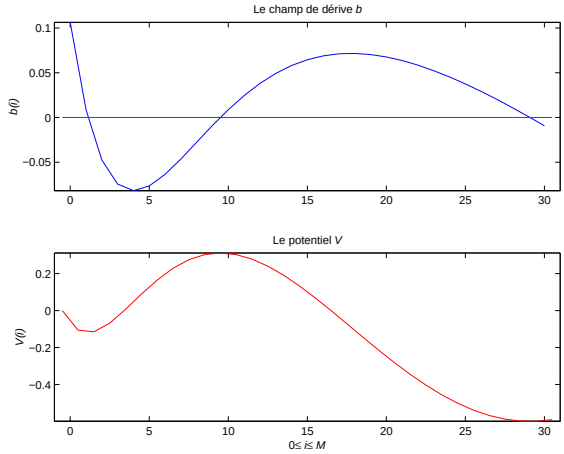


FIG. 2 – Dérive et potentiel

En faisant varier les paramètres M, σ et p, on constate qu'il arrive, comme c'est le cas pour notre jeu de paramètres, que V peut admettre un maximum local en $0 < n_c < M$ encadré par deux minima locaux en n_0 et n_M : $0 \leq n_0 < n_c < n_M \leq M$. Dans le cas de la Figure 2, nous avons en arrondissant à la valeur entière la plus proche : $n_0 = 1, n_c = 10$ et $n_M = 29$. On dit qu'on a un double puits de potentiel. Les valeurs n_0 et n_M correspondent à des équilibres stables de l'évolution moyenne (4), alors que n_c est une valeur d'équilibre instable. L'indice c de n_c souligne le fait que n_c est une *valeur critique* : une fois cette valeur atteinte et légèrement dépassée, la chaîne (N_t) aura tendance à dévaller la pente du potentiel V pour se stabiliser vers n_M. Un paysage de potentiel en double puits est donc avantageux pour Aloha lorsque n_0 est petit et n_c est "éloigné" de n_0 au sens où la hauteur $V(n_c) - V(n_0)$ de la barrière de potentielle est grande. Dans une telle situation et en partant de $N_0 \simeq n_0$, la chaîne mettra en moyenne un temps grand pour atteindre n_c, puis la saturation au voisinage de $n_M \simeq M$.

Puisque le nombre d'états : $M+1$, est fini et que pour tout i, $P(i, i-1) > 0$ et $P(i, i+1) > 0$, (lorsque ces quantités sont définies) : la chaîne est récurrente. On sait donc qu'irrémédiablement (presque sûrement) N_t finira par atteindre la saturation M. Mais aussi qu'elle finira par retourner (presque sûrement) dans le voisinage de 0 en oscillant autour de n_0. Et ainsi de suite. La chaîne admet une unique probabilité invariante que nous noterons $\mu = (\mu(i); 0 \leq i \leq M)$. Lorsque le paysage de potentiel est un double puits, μ admet deux modes approximativement centrés en n_0 et n_M, le mode le plus fort étant celui correspondant au puits le plus profond.

§3. Une simulation

Le programme `aloha2.m` (voir la Section 6.3) éxécute une simulation de T itérations de la chaîne $(N_t)_{0 \leq t \leq T}$. La représentation graphique d'une trajectoire est donnée à la Figure 3.

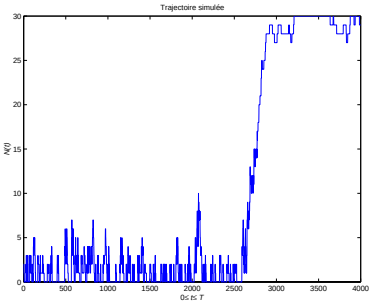


FIG. 3 – Histoire d'une saturation

D'après le théorème ergodique, on s'attend à ce que pour des T grands, le système visite avec préférence les voisinages des équilibres stables n_0 et n_M : la proportion du temps passé en chacun des états sera proche des probabilités de ces états sous la loi invariante μ. C'est pourquoi nous calculons μ à l'aide du programme `aloha3.m` (voir la Section 6.4) et nous effectuons à la Figure 4 les histogrammes d'une part des fréquentations de la trajectoire simulée et d 'autre part de la loi théorique μ.

La différence spectaculaire entre les deux histogrammes de la Figure 4 s'explique de la façon suivante. En $T = 4000$ itérations, la trajectoire issue de 0 a passé beaucoup de temps au voisinage de n_0 avant d'entreprendre une excursion au voisinage de n_M. Si l'on avait poursuivi la simulation, il est très probable que la trajectoire serait restée beaucoup plus longtemps au voisinage de n_M qui est plus stable que n_0 comme l'indique les profondeurs relatives des puits de V, mais surtout l'histogramme de la mesure invariante μ. L'histogramme observé se serait alors approché de celui de μ. Nous invitons le lecteur à faire cette simulation avec une condition initiale proche de $n_M = 29$.

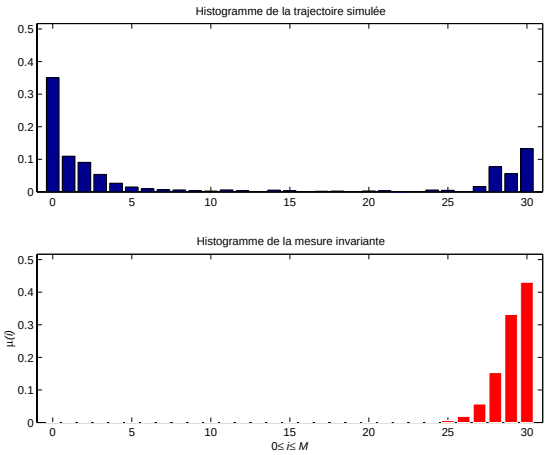


FIG. 4 – Histogrammes observé et théorique

§4. Temps moyen de saturation

Un indicateur d'efficacité de la chaîne est le temps moyen d'atteinte de n_c en partant de zéro ou bien de n_0 puisque c'est à peu prés le temps moyen de saturation. On peu l'évaluer à l'aide de simulations répétées comme le texte le suggère. Une méthode dite de simulation accélérée sera alors souhaitable. Ces méthodes, dont l'ingrédient théorique majeur est la théorie des grandes

déviations, sont effectivement mises en pratique pour des systèmes dont la complexité ne permet pas d'espérer des résultats par une approche numérique de type analytique (sans simulation). Dans la situation qui est présentée ici, nous sommes en fait en mesure d'évaluer le temps moyen d'atteinte d'un domaine par une méthode analytique. Nous nous éloignons donc du texte pour présenter cette méthode simple.

Soit $C \subset \{0, \ldots, M\}$ un ensemble d'états : le domaine à atteindre, et i la condition initiale de la chaîne. On note $\tau(i)$ le temps moyen d'atteinte de C en partant de i, c'est-à-dire

$$\tau(i) = \mathbb{E}_i(S)$$

où S est le temps aléatoire d'atteinte de C : $S = \inf\{t \geq 0; N_t \in C\}$ et $\mathbb{E}_i(S)$ est une notation compacte pour $\mathbb{E}(S \mid N_0 = i)$. On a clairement pour tout i dans C, $\tau(i) = 0$. Regardons ce qu'il se passe lorsque i est hors de C. Nous conditionnons par la variable N_1, état de la chaîne à l'instant $t = 1$: $\tau(i) = \mathbb{E}_i(S) = \mathbb{E}_i\mathbb{E}_i(S \mid N_1)$. Si $j \in C$, $\mathbb{E}_i(S \mid N_1 = j) = 1$. Et si $j \notin C$, $\mathbb{E}_i(S \mid N_1 = j) = 1 + \mathbb{E}_j(S) = 1 + \tau(j)$, grâce à la propriété de Markov. Donc pour tout $i \notin C$, nous avons

$$\tau(i) = \sum_{j \in C} P(i,j) + \sum_{j \notin C} P(i,j)(1 + \tau(j)) = 1 + \sum_{j \notin C} P(i,j)\tau(j). \tag{7}$$

C'est un système d'équations linéaires dont les inconnues sont $(\tau(i); i \notin C)$. En notant P_{C^c} la restriction de la matrice P aux indices dans l'ensemble C^c complémentaire de C, I_{C^c} la matrice identité correspondante, τ le vecteur colonne $(\tau(i); i \notin C)$ et 1 le vecteur colonne correspondant rempli de 1, ce système s'écrit matriciellement

$$(P_{C^c} - I_{C^c})\tau = -1. \tag{8}$$

Même avec un grand cardinal de C^c, sa résolution ne pose pas de problème à Matlab. Le script correspondant est écrit à la Section 6.5, il calcule τ pour un domaine de la forme $C = [n, M]$. La représentation graphique de τ est donnée à la Figure 5 pour $n = 25$. Bien que n_c soit égal à 10, nous préférons présenter des résultats correspondant à une saturation forte. En effectuant le calcul du temps moyen d'atteinte de [10,30], nous obtenons par exemple en partant de $0 : \tau(0) \simeq 500$, et en partant de $7 : \tau(7) \simeq 300$.

§5. Nombre infini d'utilisateurs

La question de la simulation de la chaîne lorsque $M = \infty$ se résoud bien sûr de manière approximative, en tronquant la seconde matrice du texte (avec les $e^{-\alpha}$). On peut choisir M suffisament grand pour que les $P(i, i-1)$ pour des i proches de M soient inférieurs à un certain seuil ε de l' ordre de, par

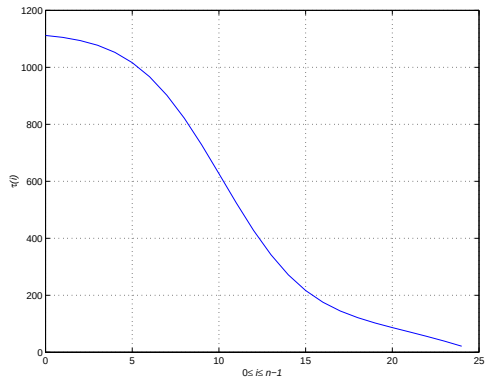


FIG. 5 – Temps moyen d'atteinte de [25,30]

exemple 0.1 ou moins, ce qui impose en général un champ b assez positif (le lecteur pourra critiquer le flou de cette affirmation). Bien sûr, on complète la dernière colonne de sorte que les sommes par lignes soient égales à 1. Par

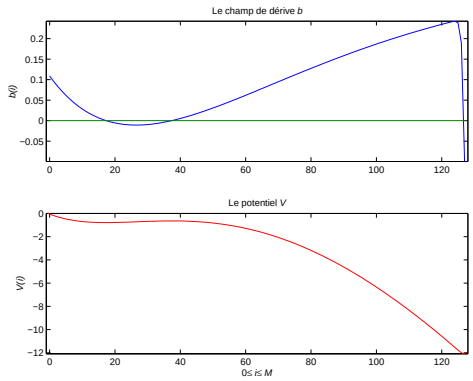


FIG. 6 – Un toboggan de potentiel

exemple avec les paramètres du texte

$$\alpha = 0.36, \quad p = 0.024 \tag{9}$$

et $\varepsilon = 0.1$, on obtient $M = 127$. On peut alors utiliser les programmes de calcul précédents. Le paysage de potentiel correspondant à ce jeu de paramètres est donné à la Figure 6. On constate que l'on retrouve à peu près les valeurs d'équilibre du texte, puisque nous avons $n_0 = 17$ et $n_c = 37$. Mais surtout, il apparaît que la descente de potentiel pour $i \geq n_c$ "semble être inéxorable". Ceci serait bien sûr à vérifier théoriquement.

§6. Appendice : Les programmes Matlab

Nous commençons par donner le script principal, puis les programmes des fonctions appelées dans ce script et finalement le script du calcul du temps moyen de saturation.

6.1. Script principal.

```
clear
% VARIABLES GLOBALES
 global p                % = probabilite de transmission d'un paquet bloque
 global sigma        % = probabilite de transmission d'un paquet libre
 p=.2
 sigma=.012
 M=30 % Le nombre d'utilisateurs est M

% Calcul de P, b et V
 [P,b,V] = aloha1(M);

% Representation graphique de la matrice
 y0=zeros(M+1,1); ymoins=zeros(M+1,1); yplus=zeros(M+1,1);
 y0(1)=P(1,1); y0(M+1)=P(M+1,M+1); ymoins(M+1)=P(M+1,M);
 yplus(1)=P(1,2); for i=2:M
   y0(i)=P(i,i);
   ymoins(i)=P(i,i-1);
   yplus(i)=P(i,i+1);
 end
 reste=1-ymoins-y0-yplus; x=0:M;
 figure(1)
 subplot(2,1,1); surf(P) title('Representation 3D de la matrice P')
 axis([0 M+1 0 M+1 0 1])
 subplot(2,1,2);
 plot(x,y0,'g',x,ymoins,'--b',x,yplus,'-.r',x,reste,':k')
 title('Les diagonales de P') xlabel('0\leq {\iti}\leq {\itM}')
 axis([0 M 0 1])

% Representation graphique de la DERIVE b et du POTENTIEL V
 figure(2)
 x=0:M; y=-0.5:1:M+0.5; z=zeros(M+2,1);
 subplot(2,1,1); plot(x,b,y,z,'-') title('Le champ de derive {\itb}')
 ylabel('{\itb(i)}') axis([-1 M+1 min(b) max(b)])
 subplot(2,1,2); plot(y,V,'r') title('Le potentiel {\itV}')
 xlabel('0\leq{\iti}\leq {\itM}') ylabel('{\itV(i)}')
 axis([-1 M+1 min(V) max(V)])
```

```
% SIMULATION et representation graphique de la trajectoire simulee
% T = nombre d'iterations, x0 = etat initial.
 figure(3)
 T =500; x0=2;
 X = aloha2(P,T,x0);

% MESURE INVARIANTE
 mu = aloha3(P);
 figure(4)
 s = 0:M; [n,Xout]=hist(X,s);pmax=1.2*max([max(n/T),max(mu(1,:))]);
 subplot(2,1,1);  bar(Xout,n/T)
 title('Histogramme de la trajectoire simulee') axis([-1 M+1 0 pmax])
 subplot(2,1,2); bar(0:M,mu(1,:)) h = findobj(gca,'Type','patch');
 set(h,'FaceColor','r','EdgeColor','w')
 title('Histogramme de la mesure invariante')
 xlabel('0\leq {\iti}\leq {\itM}') ylabel('{\it\mu(i)}')
 axis([-1 M+1 0 pmax])
```

6.2. Fonction `aloha1.m`.

```
 function [P,b,V] = aloha1(M)
% Implementation de la matrice stochastique d'ALOHA : P
% Calcul de la derive b et du potentiel V associe
 global sigma global p
 P = zeros(M+1); b = zeros(M+1,1);
 V=zeros(M+2,1); pp = 1-p; s = sigma; ss = 1-sigma;
 for i = 0:M
   ii=i+1;
   for j = 0:M
      jj=j+1;
      if j==(i-1)
               P(ii,jj) = i*p*(pp^(i-1))*(ss^(M-i));
      elseif j==i
               P(ii,jj) =  (pp^i)*(M-i)*s*(ss^(M-i-1)) ...
                            + (1-i*p*(pp^(i-1)))*(ss^(M-i));
      elseif j==(i+1)
               P(ii,jj) = (M-i)*s*(ss^(M-i-1))*(1-pp^i);
      elseif j>=(i+2)
               P(ii,jj) = factorial(M-i)/(factorial(j-i)...
                          *factorial(M-j))*(s^(j-i))*(ss^(M-j));
      end
      b(ii)=b(ii)+(j-i)*P(ii,jj);
   end
 end
 V(1)=0; V(2)=-b(1);
```

```
 for i=3:M+2
    V(i)=V(i-1)-b(i-1);
 end
```

6.3. Fonction aloha2.m.

```
 function X = aloha2(P,T,x0)
% ALOHA2 simule la chaine de Markov de matrice P.
% T = longueur de la chaine, x0 = etat initial
 M=size(P,1)-1;
% Calcul de la matrice Q des probabilites cumulees associees a P.
 Q = zeros(M+1,M+2); for i=0:M
   ii=i+1;
   Q(ii,2)=P(ii,1);
   for j=1:M
          Q(ii,j+2)=Q(ii,j+1)+P(ii,j+1);
   end
 end
% Simulation
 X = zeros(1,T+1); U = rand(1,T);
 X(1) = x0; % (etat initial)
 for n=2:T+1
   for j=0:M
      if ((Q(X(n-1)+1,j+1)<U(n-1))&(U(n-1)<=Q(X(n-1)+1,j+2)))
            X(n)=j;
         end
   end
 end
% Representation graphique de la trajectoire simulee
 t = 0:T; plot(t,X)
 title('Trajectoire simulee')
 xlabel('0\leq {\itt}\leq {\itT}') ylabel('{\itN(t)}')
```

6.4. Fonction aloha3.m.

```
function mu = aloha3(P)
% Calcul de la mesure invariante de la chaine ALOHA a M utilisateurs
% P : la matrice stochastique. Sortie: histogramme de la mesure invariante.
 d=size(P,1); Q = P'-eye(d); c = null(Q);
 mu = inv(diag(sum(c)))*(c');
```

6.5. Script du calcul du temps de saturation.

```
% Script du calcul du TEMPS MOYEN d'atteinte de C=[n,M]
% La matrice P doit être en mémoire
 n=25   % n est la borne inf de l'ensemble à atteindre: [n,M]
```

```
t=zeros(n,1);
Q=P(1:n,1:n);   %extraction de la sous-matrice Q=P_{C^c}
t=-inv(Q-eye(n))*ones(n,1);
figure(5) x=0:n-1; plot(x,t) grid on
xlabel('0\leq {\iti}\leq {\itn-1}') ylabel('{\it\tau(i)}')
```

Chapitre 18
Utilisation en TIPE

En théorie des probabilités, un des outils de modélisation les plus efficaces, mais aussi les plus simples, est la "chaîne de Markov". Il s'agit en gros d'une suite aléatoire $N(t)$, à valeurs dans un ensemble d'états au plus dénombrable, représentant l'état d'un système lorsque le temps t varie dans $\mathbf{N}$ (on parle de temps discret). La propriété fondamentale d'une chaîne de Markov est que l'évolution du système après le temps t ne dépend que du "passé immédiat" : la loi de $N(t+1)$ connaissant $N(0), N(1), \ldots, N(t)$ ne dépend en fait que de $N(t)$. Introduisons un peu de vocabulaire concernant les chaînes de Markov :

- Lorsque la chaîne de Markov est *homogène*, cela veut dire que l'évolution du système après le temps t ne dépend que de $N(t)$ (son état au temps t) mais pas de t [1]. En termes de probabilités conditionnelles, $P(N(t+1) = \beta | N(t) = \alpha)$ ne varie pas au cours du temps. Cette probabilité, notée $p_{\alpha\beta}$ est appelée *probabilité de transition* de l'état α vers l'état β. Lorsque α est fixé, la famille $(p_{\alpha\beta})_\beta$ est une loi de probabilité ; en particulier, elle est sommable et on a $\sum_\beta p_{\alpha\beta} = 1$.
- On dit que la chaîne de Markov est *irréductible* si pour tout couple d'états (α, β), on peut passer de α à β, quitte à y passer du temps : on peut trouver t tel que $P(N(t) = \beta | N(0) = \alpha) > 0$. Elle est en plus *apériodique* si la dernière propriété est vérifiée pour tout t à partir d'un certain temps t_0 (dépendant bien sûr de α et β)[2].

Le résultat principal sur les CMIHA (Chaînes de Markov Irréductibles Homogènes Apériodiques) concerne leur évolution asymptotique. Il établit que la chaîne de Markov oublie son passé, ce qui revient à dire que le comportement asymptotique (pour t grand) ne dépend pas de la situation à $t = 0$, qui est éventuellement aléatoire. Plus précisément, le comportement de la chaîne de Markov est forcément dans l'un des trois types suivants :

- La chaîne est *transiente* (ou instable) : pour tout état α, alors p.s.[3], la chaîne de Markov ne passera par l'état α qu'un nombre fini de fois.

[1] Ceci est à rapprocher, dans le cadre déterministe, des systèmes dynamiques $u_{n+1} = f(u_n)$ ou bien des équations différentielles du type $y' = f(y)$.

[2] Si les propriétés d'homogénéité et d'irréductibilité sont courantes, la propriété d'apériodicité n'est pas toujours vérifiée. Par exemple, l'évolution de la fortune d'un joueur de pile ou face est bien représentée par une chaîne de Markov homogène et irréductible à valeurs dans $\mathbf{Z}$, mais pas apériodique, pour des raisons de parité.

[3] p.s. : presque sûrement. Une propriété est vérifiée presque sûrement si sa probabilité d'apparition est de 1. Cela ne signifie pas forcément qu'elle est *certaine* : considérer par exemple sur une infinité de lancers d'un dé non truqué, le fait de sortir au moins une fois le 6.

– La chaîne est *récurrente nulle* : pour tout α, la chaîne de Markov passera une infinité de fois par l'état α, mais "de moins en moins souvent"[4] : presque sûrement,
$$\frac{\mathrm{Card}\{\tau \leq t, N(\tau)=\alpha\}}{t} \longrightarrow 0$$
lorsque t tend vers $+\infty$.
– La chaîne est *récurrente positive* (ou ergodique) : pour tout α,
$$\frac{\mathrm{Card}\{\tau \leq t, N(\tau)=\alpha\}}{t} \longrightarrow p_\alpha$$
et la famille $(p_\alpha)_\alpha$ définit une loi de probabilité sur l'ensemble des états, appelée *loi limite*. La chaîne de Markov passera une infinité de fois par α, et ce de façon "'régulière".

Il est intéressant de remarquer à ce sujet que dans le cas particulier où l'espace d'états est fini, alors la chaîne de Markov est forcément ergodique. La démonstration de ce résultat est parfaitement accessible à l'élève de classes préparatoires. Elle utilise simplement des techniques de réduction de matrices, et le bagage probabiliste demandé est très faible (à vrai dire, l'enseignement de terminale suffit). L'argument principal de la démonstration est le suivant : on a
$$P(N(t+1)=\beta | N(t)=\alpha) = \sum_\alpha P(N(t)=\alpha) p_{\alpha\beta}$$
Donc le vecteur ligne $\pi_t = (P(N(t)=\beta)_\beta$ qui représente la loi du système au temps t vérifie la récurrence $\pi_{t+1} = \pi_t P$, où P est la matrice des probabilités de transition. Le comportement de la suite π_t dépend donc de l'étude des vecteurs et valeurs propres de P.

Revenons au texte. Celui-ci est parfaitement adapté aux contraintes des TIPE, puisqu'il présente une application des mathématiques à un problème concret de communications. On peut le développer suivant plusieurs axes, comme par exemple :

– Détailler l'étude des chaînes de Markov à espace d'états discrets. On peut alors donner un sens précis, et illustrer la phrase "Le système finit par se stabiliser autour de M : la plupart des terminaux sont bloqués". Cet aspect se prête bien à une expérimentation informatique. On prend une valeur de M "raisonnable" (par exemple $M = 100$). On calcule alors la matrice P, puis on part de $\pi_0 = (1, 0, \ldots, 0)$ (aucun terminal n'est bloqué au temps $t = 0$) et on calcule par récurrence π_t, jusqu'à ce que π_t ait convergé[5]. On étudie alors la forme de la loi limite suivant les paramètres p et σ. Cette loi limite décrit bien les tendances générales du système.

[4] En fait, l'espérance du temps de retour en α (partant de α) est infinie, même si cette variable est p.s. finie.

[5] On pourra considérer que π_t a convergé lorsque $\sum_\beta |\pi_{t,\beta} - \pi_{t-1,\beta}| \leq 10^{-6}$

– Tracer la “courbe" de l'espérance conditionnelle

$$E(N(t+1) - N(t)|N(t) = n)$$

en fonction de n, et calculer expérimentalement les valeurs de u_0 et u_c. La valeur de cette espérance conditionnelle s'écrit en réalité comme une somme de série, dont n est un paramètre. On obtient

$$E(N(t+1) - N(t)|N(t) = n) = \sum_{k=0}^{+\infty} (k-n) P_{n,k}$$

la série pouvant d'ailleurs commencer à l'indice $k = n-1$, les probabilités $p_{n,k}$ étant nulles dans les autres cas. On peut alors calculer numériquement les valeurs de cette espérance conditionnelle pour diverses valeurs de n, de façon à faire apparaître les constantes n_0 et n_c.

– Enfin, on peut produire des arguments qualitatifs de nature probabiliste pour expliquer les phénomènes en jeu, même avec une approche simplifiée.
  - Le paramètre $\alpha = M\sigma$ correspond au nombre moyen de terminaux qui tentent d'émettre un paquet au temps t. Si $\alpha > 1$, il y a à chaque instant “en moyenne" au moins un terminal de plus qui tente d'émettre ; aucun de ceux qui sont déjà bloqués ne peut se débloquer et la suite $N(t)$ est “globalement" croissante.
  - Dans le cas où $M\sigma <\leq 1$, la probabilité qu'un terminal se débloque en l'absence de tentative d'émission est de $\delta = N(t)p(1-p)^{N(t)-1}$. Si $\alpha \geq \delta$ (ce qui est le cas pour $N(t)$ grand, mais aussipour $N(t)$ petit), la suite $N(t)$ a tendance à croître. Comme la chaîne de Markov est irréductible, il y a toujours un moment où on aura $N(t)$ suffisamment grand pour que $N(t)$ devienne croissante. Les deux solutions de l'équation $xp(1-p)^{x-1} = \alpha$, d'inconnue x sont certainement de bonnes approximations de n_0 et n_c.

– En revanche, l'étude statistique des estimateurs du temps moyen de fonctionnement, et de leur convergence, nécessite des notions de statistiques difficiles à maîtriser dans le cadre de l'épreuve de TIPE.

En conclusion, ce qui est intéressant pour un étudiant en classe préparatoires, c'est le fait que l'investissement en termes de théorie des probabilités est relativement faible. De plus, les possibilités d'extension du texte lui permettent d'être traité au choix suivant un axe “mathématiques", “modélisation", ou bien “simulation", ou bien encore un panaché des trois. L'étudiant qui voudrait aller plus loin dans l'étude des protocoles probabilistes de communication aura intérêt à lire le texte “fonctions de Lyapounov et stabilité de réseaux de télécommunications", présenté plus loin dans le recueil.

Sixième partie

Détection d'un signal en télécommunication

Chapitre 19
Détection d'un signal en télécommunications

§1. Le texte proposé

Résumé Une antenne reçoit des ondes électromagnétiques qui peuvent contenir un signal ou seulement du bruit de fond. Dans les deux situations, on modélise les signaux reçus par des variables aléatoires gaussiennes d'amplitudes différentes. Une procédure de détection du signal, basée sur un théorème de statistique inférentielle, est proposée. Les propriétés asymptotiques de cette procédure sont étudiées.
Mots clés : Tests statistiques, Théorème de la limite central

1.1. Modèle bruit et signal plus bruit. Une antenne est réglée pour recevoir des signaux à une fréquence donnée. A chaque instant $t \in [0, T]$, elle reçoit une onde d'amplitude $f(t)$. Cette amplitude est générée soit par le bruit électromagnétique ambiant soit par un signal émis par une source S auquel vient s'ajouter le bruit ambiant. Nous concentrons notre analyse dans la situation où l'on ne dispose que de données echantillonnées. Pour $N \in \mathbb{N}^*$, on suppose donc que l'on ne connaît la fonction f qu'aux instants $T/N, 2T/N, \ldots, T$. On pose

$$X_j = f\left(\frac{jT}{N}\right), \quad j = 1, \ldots, N.$$

On propose alors le modèle stochastique suivant :

- Si S n'a pas émis de signal, alors l'antenne ne reçoit que du bruit et on a pour, $j = 1, \ldots, N$, $X_j = \varepsilon_j$, où la suite (ε_j) est indépendante et identiquement distribuée (i.i.d.), de loi normale centrée de variance σ^2 (notée $\mathcal{N}(0, \sigma^2)$).
- Si S émet un signal, alors à la date jT/N $(j = 1, \ldots, N)$ l'amplitude de ce signal est modélisée par une variable aléatoire Y_j. On suppose que la suite (Y_j) est i.i.d. de loi $\mathcal{N}(0, \xi^2)$. Le signal reçu à l'instant jT/N a été dégradé par un bruit additif indépendant de l'onde émise par S. On a finalement
$$X_j = Y_j + \varepsilon_j, \quad j = 1, \ldots, N.$$
L'indépendance entre le bruit et le signal conduit à supposer que les suites (Y_j) et (ε_j) sont indépendantes.

On suppose que les paramètres de dispersion σ^2 et ξ^2 sont connus. Le but de cette étude est de construire une méthode qui détecte, à partir de l'observation $X_1, \ldots, X_N$, si la source S a émis une onde ou non. En termes statistiques, cela

signifie que l'on souhaite faire un test d'hypothèses dans le modèle suivant. On observe un N-échantillon (N variables aléatoires i.i.d.), $X_1, X_2, \dots, X_N$, de loi F inconnue. Deux hypothèses sont possibles pour F :

H_0 : F est la loi normale centrée de variance σ^2 (bruit).

H_1 : F est la loi normale centrée de variance $\sigma^2 + \xi^2$ (signal plus bruit).

Il s'agit donc de construire une fonction mesurable T_N sur $\mathbb{R}^N$ à valeurs dans $\{0, 1\}$ (une fonction test). Celle-ci conduit ensuite à choisir à partir de $X_1, X_2, \dots, X_N$ l'hypothèse $H_{T_N(X_1, X_2, \dots, X_N)}$. La décision rendue est aléatoire. On se propose d'étudier ses propriétés statistiques. On étudiera principalement les problèmes suivants :

a) Pour une fonction de test T_N donnée, quantifier les probabilités de fausse détection :

$$\alpha(T_N) = \mathbb{P}(T_N \text{ prend la décision } H_1 \text{ à tort}) = \mathbb{P}(T_N = 1 | H_0 \text{ est vraie}),$$

et de non détection :

$$\beta(T_N) = \mathbb{P}(T_N \text{ prend la décision } H_0 \text{ à tort}) = \mathbb{P}(T_N = 0 | H_1 \text{ est vraie}).$$

$\alpha(T_N)$ (resp. $\beta(T_N)$) s'appelle l'erreur de première espèce (resp. seconde espèce) associée à la fonction de test T_N. En général, on construit la procédure de test en fixant le niveau de l'erreur de première espèce (on prend souvent 5%).

b) Existence d'une fonction de test optimale (elle minimise l'erreur de seconde espèce à erreur de première espèce fixée).

c) Evaluation du nombre d'observations nécessaire pour réaliser un certain objectif (erreurs de première et seconde espèce fixées).

On s'intéresse dans un premier temps au problème b) que l'on étudie dans un cadre abstrait très général.

1.2. Le test de Neyman-Pearson. Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité et $(E, \mathcal{B})$ un espace mesurable. On considère une application mesurable Z sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(E, \mathcal{B})$. On note G la loi de Z :

$$G(B) = \mathbb{P}(Z \in B), \quad B \in \mathcal{B}.$$

Soit μ_0, μ_1, μ des probabilités sur $(E, \mathcal{B})$. On suppose qu'il existe des fonctions mesurables f_0, f_1 sur $(E, \mathcal{B})$ à valeurs dans $\mathbb{R}^+$ avec

$$\mu_i = f_i \mu, \quad (i = 0, 1).$$

Dans ce cadre, supposons que l'on observe Z et que l'on s'intéresse au test d'hypothèses :

$$H_0 \quad G = \mu_0 \quad | \quad H_1 \quad G = \mu_1.$$

Le Théorème suivant permet de construire la fonction de test optimale.

Théorème 1 (Neyman et Pearson) *Supposons*

$$\forall t \in \mathbb{R}^+,\ \mu_0\left(\{f_1 = tf_0\}\right) = 0. \tag{1}$$

Soit $\alpha^* \in]0,1[$. *Alors*

i) Il existe $C_{\alpha^*} > 0$ *avec* $\mu_0(\{f_1 > C_{\alpha^*} f_0\} = \alpha^*$.

ii) Soit $\Phi_{\alpha^*}(Z) = \mathbb{1}_{\{f_1(Z) > C_{\alpha^*} f_0(Z)\}}$. *On a,*

a) Φ_{α^*} *est une fonction test, son erreur de première espèce vaut* α^*.

b) Soit Φ *une fonction test dont l'erreur de première espèce vaut* α^*. *Alors, son erreur de seconde espèce est minorée par celle de* Φ_{α^*}.

Remarque : On peut se passer de l'hypothèse (1) à condition de travailler avec des fonctions de test dites randomisées. C'est-à-dire dont les valeurs appartiennent à $[0,1]$.

Donnons quelques éléments qui permettent de démontrer ce Théorème.

i) Considérons la variable aléatoire $U = f_1(Z_0)/f_0(Z_0)$ où Z_0 a la loi μ_0. D'après l'hypothèse (1) cette variable aléatoire possède, lorsque Z a la loi μ_0, une fonction de répartition M continue. On en déduit alors qu'il existe un réel C_{α^*} vérifiant $M(C_{\alpha^*}) = 1 - \alpha^*$.

ii) Par construction, l'erreur de première espèce de Φ_{α^*} vaut α^*. Pour montrer que l'erreur de seconde espèce de Φ est minorée par celle de Φ_{α^*} il suffit d'exprimer leur différence sous la forme d'une intégrale par rapport à la probabilité μ_1 :

$$\beta(\Phi_{\alpha^*}) - \beta(\Phi) = -\int (\Phi_{\alpha^*} - \Phi) f_1 \, d\mu.$$

On décompose cette intégrale en considérant l'ensemble des valeurs de Φ_{α^*} puis on la majore par 0.

1.3. Construction du test optimal pour le problème de détection. Le Théorème de Neyman-Pearson donne la règle de décision suivante (pour une erreur de première espèce égale à $\alpha^* \in]0,1[$) :

$$\Phi_{\alpha^*}(X_1, \ldots, X_N) = \mathbb{1}_{\{\sum_{j=1}^N X_j^2 > K_{\alpha^*}\}}, \tag{2}$$

où $K_{\alpha^*} > 0$ est l'unique réel satisfaisant

$$\mathbb{P}_{H_0}\left(\sum_{j=1}^N X_j^2 > K_{\alpha^*}\right) = \mathbb{P}\left(\sum_{j=1}^N \varepsilon_j^2 > K_{\alpha^*}\right) = \alpha^*. \tag{3}$$

On se propose de calculer K_{α^*} et l'erreur de seconde espèce associée $(\beta(\Phi_{\alpha^*}))$. Soit H une loi à densité sur $\mathbb{R}$. On rappelle que le quantile supérieur de niveau $\alpha^* \in]0,1[$ de H est l'unique réel h_{α^*} qui vérifie $H(]-\infty, h_{\alpha^*}]) = 1 - \alpha^*$. D'autre part, soit p un entier naturel non nul. La loi du Khi deux à p degrés de liberté est la loi de la somme des carrés de p variables i.i.d. de loi $\mathcal{N}(0,1)$. On note

$\chi^2(p)$ cette loi. Pour $\alpha^* \in]0,1[$ $\chi^2_{\alpha^*,p}$ désigne son quantile supérieur de niveau α^* et ψ_p sa fonction de répartition. Le quantile supérieur de niveau α^* de la loi normale centrée réduite est noté $\mathcal{N}_{\alpha^*}$ et sa fonction de répartition φ. On a alors les résultats suivants :

i) $K_{\alpha^*} = \sigma^2 \chi^2_{\alpha^*,N}$,

ii) $\beta(\Phi_{\alpha^*}) = \psi_N\left(\frac{K_{\alpha^*}}{\sigma^2+\xi^2}\right)$.

D'un point de vue pratique, les résultats précédents sont difficiles à utiliser. En effet, la loi $\chi^2(N)$ n'est tabulée que pour les petites valeurs de N. La propriété suivante est basée sur le théorème limite central (T.L.C.). Lorsque N est grand, elle permet un calcul approché des quantiles et de la fonction de répartition de la loi $\chi^2(N)$.

Propriété 1 *Soit $\alpha^* \in]0,1[$ et $t \in \mathbb{R}$. On a*

$$\lim_{N\to\infty} \frac{\chi^2_{\alpha^*,N} - N}{\sqrt{2N}} = \mathcal{N}_{\alpha^*} \quad \textit{et} \quad \lim_{N\to\infty} \psi_N(N + \sqrt{2N}t) = \varphi(t).$$

La proposition précédente nous permet d'écrire le développement

$$K_{\alpha^*} = \sigma^2(N + \mathcal{N}_{\alpha^*}\sqrt{2N}) + \sqrt{N}o(1). \tag{4}$$

On propose donc d'utiliser la valeur approchée de K_{α^*} : $\tilde{K}_{\alpha^*} = \sigma^2(N + \mathcal{N}_{\alpha^*}\sqrt{2N})$. Pour cette valeur approchée du seuil K_{α^*}, on obtient la valeur approchée de l'erreur de seconde espèce :

$$\tilde{\beta}\left(\Phi_{\alpha^*}\right) = \varphi\left(\frac{1}{\xi^2+\sigma^2}\left[\sigma^2\mathcal{N}_{\alpha^*} - \xi^2\sqrt{\frac{N}{2}}\right]\right). \tag{5}$$

Ainsi, si notre objectif est d'égaliser les erreurs de première et seconde espèce à α^*, l'effectif d'échantillonnage N nécessaire devrait être approximativement supérieur à la partie entière par excès de

$$2\left(\mathcal{N}_{\alpha^*}\left[1 + \frac{2\sigma^2}{\xi^2}\right]\right)^2. \tag{6}$$

1.4. Quelques simulations numériques. Le programme suivant permet d'estimer empiriquement, pour différentes valeurs de N, l'erreur de seconde espèce. Celle-ci est comparée graphiquement à l'approximation donnée dans (5).

```
%Programme d'estimation de l'erreur de 2nd espèce
%
clear
%Initialisation des constantes
%nombre de simulations d'échantillon
nsimu=10000;
rac2=sqrt(2);
```

```
%Intervalle d'étude pour la taille d'échantillon 1..Nmax
Nmax=50;
%variance du bruit et du signal
si2=0.1;
xi2=1;
vartot=si2+xi2;
sigtot=sqrt(vartot)
rap1=si2/vartot;
rap2=xi2/vartot;
%erreur de 1ère espèce et quantile gaussien correspondant
alpha=0.05;
Nalpha=rac2*erfinv(1-2*alpha);
%simulation du signal reçu par l'antenne
x=sigtot*randn(Nmax,nsimu);
%calcul des carrés des signaux
x2=x.*x;
%calcul des sommes glissantes des carrés des signaux
sx2=cumsum(x2);
%Construction d'un masque contenant les seuils pour effectuer le test
i=1 :Nmax;
Ktilde=si2*(i+rac2*sqrt(i)*Nalpha);
un=ones(1,nsimu);
masq=Ktilde'*un;
%On teste avec le seuil Ktilde
phi=(sx2>masq);
%calcul de l'erreur de 2nd espèce observée
betachap=mean((1-phi)');
%calcul de l'erreur de 2nd espèce approchée
t=Nalpha*rap1-rap2*sqrt(i)/rac2;
betatilde=0.5*(1+erf(t/rac2));
%Affichage des graphiques
subplot(1,2,1);
plot(betachap);
title('erreur de seconde espèce empirique');
ylabel('probabilité');
xlabel('taille de l''échantillon');
subplot(1,2,2);
plot(betatilde);
title('erreur théorique approchée');
ylabel('probabilité');
xlabel('taille de l''échantillon');
```

1.5. Suggestions. Pour traiter ce sujet, on suggère de répondre à certaines des questions suivantes :

1) Commenter le modèle proposé au paragraphe 1.1.
2) Montrer une partie du Théorème de Neyman-Pearson.
3) Montrer que le test optimal, pour une erreur de première espèce égale à α^*, est bien donné par les équations (2) et (3). Calculer les valeurs de K_{α^*} et de $\beta(\Phi_{\alpha^*})$.
4) Montrer la Propriété 1. Etablir les formules d'approximation (4), (5), (6).
5) Commenter les graphiques obtenus après exécution du programme de simulation.
6) Compléter le programme pour pouvoir valider numériquement la formule (6).

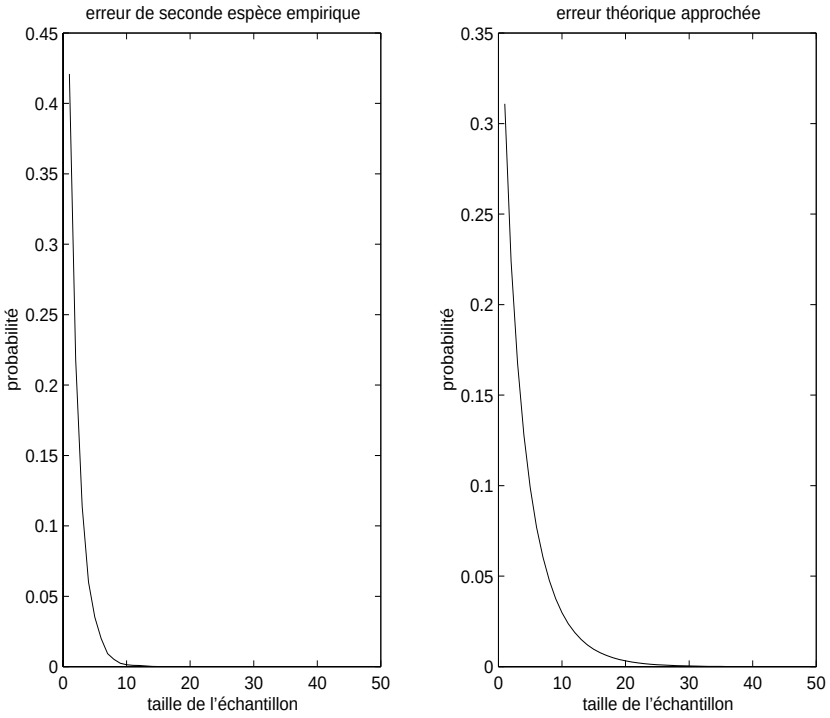


FIG. 1 – Comparaison de l'erreur de seconde espèce

Chapitre 20
Commentaire de l'auteur

§1. Commentaire

Le plan de notre commentaire s'articule sur des suggestions proposées dans le texte.

1.1. Sur le modèle proposé. Pour simplifier, nous supposerons que l'intervalle d'émission du signal est $[0,1]$ ($T = 1$). Notons g le signal émis par la source (g est une fonction sur $[0,1]$). Par convention, on pose $g = 0$ si la source n'émet pas. Le signal reçu par l'antenne peut donc être décomposé en une partie *signal pur* (g) et du bruit (B) créé par le milieu où transite le signal :

$$f(t) = g(t) + B(t) \quad (t \in [0,1]). \tag{1}$$

Le signal reçu est échantillonné. C'est-à-dire que l'on observe la fonction f qu'aux instants j/N, $j = 1, \dots, N$. Posons $X_j = f(j/N)$. La relation (1) devient

$$X_j = g(j/N) + B(j/N), \quad j = 1, .., N.$$

Les bruits $B(j/N)$ sont supposés être indépendants et de même loi gaussienne $\mathcal{N}(0, \sigma^2)$. Cette hypothèse est *raisonnable physiquement*. Elle signifie que les perturbations du signal à des instants distincts n'ont pas de cohérence. De plus, le caractère gaussien des loi marginales peut-être physiquement justifié à l'aide du théorème de la limite central (bruit ambiant obtenu par superposition de bruits infinitésimaux). Dans le cas où la source émet, on modélise l'amplitude du signal émis à l'instant j/N par une variable gaussienne Y_j indépendante du bruit. On suppose de plus que les amplitudes $Y_1, \dots, Y_N$ sont indépendantes. Remarquons que l'hypothèse d'indépendance entre le signal émis et le bruit de fond est physiquement réaliste. En effet, le bruit est créé par le milieu où le signal transite. L'hypothèse sur le caractère gaussien et i.i.d. des amplitudes du signal émis par la source est un *artifice mathématique* qui permet de traiter le problème sans que l'on ait besoin de connaître la forme de la fonction *déterministe* g. Pour illustrer cette remarque, prenons deux exemple de signaux :

1) $g_1(t) = \sin(2\pi t)$,
2) $g_2(t) = (t - 1/2)^3$.

Pour $N = 10^4$, le programme suivant permet de calculer les valeurs échantillonnées du signal émis. Le programme donne également des informations sur les répartitions des séries statistiques $(g_i(j/N))_{i=1,2,j=1,\dots,N}$.

```
% Calcul de j/N et des fonctions g1 et g2 échantillonnee
t=0 :0.0001 :1 ;
g1=sin(2*pi*t) ;
g2=(t-0.5).^ 3 ;
% Affichage des histogrammes des 2 séries statistiques
subplot(1,2,1) ;
histo(g1) ;
title('g1(t)=sin t') ;
subplot(1,2,2) ;
histo(g2) ;
title('g2(t)=(t-0.5)^ 3 ;') ;
% Calcul et affichage des moyennes et écart-types des 2 séries
m1=mean(g1) ;
s1=std(g1) ;
m2=mean(g2) ;
s2=std(g2) ;
disp('moyenne de g1') ;
m1
disp('écart type de g1') ;
s1
disp('moyenne de g2') ;
m2
disp('écart type de g2') ;
s2
```

L'exécution du programme donne alors :

```
moyenne de g1
  m1 =
```

$-4.2080e-017$

```
  ecart type de g1

  s1 =
```

0.7071

```
  moyenne de g2

  m2 =
```

$-7.5488e-019$

```
  ecart type de g2

  s2 =
```

0.0473

La figure 2 donne une représentation graphique sous forme d'histogramme des séries statistiques.

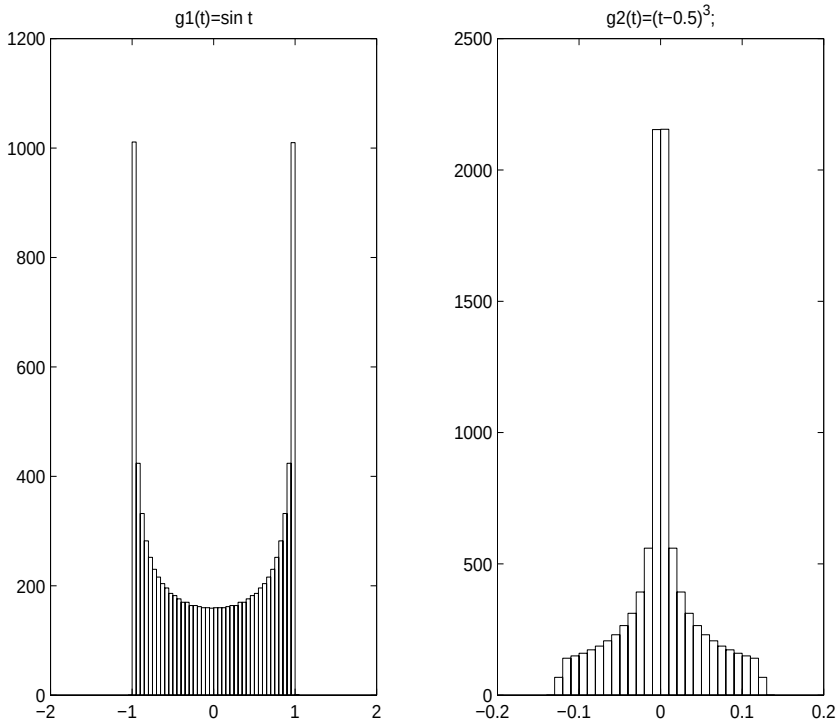


FIG. 1 – Répartitions des valeurs des 2 fonctions

On peut réaliser dans l'exemple 1 (g_1) que l'on est fort loin d'une distribution gaussienne. En assimilant les valeurs échantillonnées du signal émis à la réalisation de variables indépendantes et de même loi, on pourrait aussi mettre en oeuvre le test de Kolmogorov-Smirnov pour tester le caractère gaussien de la loi marginal. A partir de maintenant, on se place dans le cadre statistique où $X_1 \dots X_N$ sont des variables i.i.d. de loi F. On désire tester

H_0 : F est la loi normale centrée de variance σ^2 (bruit).

H_1 : F est la loi normale centrée de variance $\sigma^2 + \xi^2$ (signal plus bruit).

Ce test est un test d'hypothèses simples. C'est-à-dire que les ensembles de probabilités vérifiant H_0 ou H_1 sont des singletons. Dans le paragraphe suivant, on prouve un théorème général qui permet de construire, dans le cas d'un test d'hypothèses simples, un test optimal.

1.2. Le test de Neyman-Pearson.

Démonstration du théorème de Neyman-Pearson

i) Pour $t \geq 0$, soit

$$\begin{aligned} H(t) = \mu_0(\{f_1 > tf_0\}) &= \mu_0\left(\left\{\tfrac{f_1}{f_0} > t\right\}\right) \quad \text{car } \mu_0(\{f_0 = 0\}) = 0 \\ &= 1 - \mu_0\left(\left\{\tfrac{f_1}{f_0} \leq t\right\}\right) = 1 - M(t), \end{aligned}$$

où M est la fonction de répartition de la variable $f_1(Z_0)/f_0(Z_0)$ et Z_0 suit la loi μ_0. La fonction M est croissante et continue sur $\mathbb{R}^+$ (grâce à (1)). Cela entraîne que la fonction H est continue et décroissante sur $\mathbb{R}^+$. De plus, $H(0) = 1$ et $\lim_{t\to+\infty} H(t) = 0$. Donc H est injective de $\mathbb{R}^+$ sur $]0,1]$ et l'on peut conclure.

ii) a) Par définition Φ_{α^*}, est bien une fonction test. De plus

$$\mathbb{P}(\text{décider } H_1 \text{ à tort}) = \mu_0(\Phi_{\alpha^*} = 1) = \mu_0(f_1 > C_{\alpha^*} f_0) = \alpha^*.$$

Donc l'erreur de première espèce de Φ_{α^*} est bien α^*.

b) Soit Φ une fonction test de niveau α^*. On a

$$\begin{aligned} \beta(\Phi) - \beta(\Phi_{\alpha^*}) = \int (\Phi_{\alpha^*} - \Phi) f_1 d\mu &= \int_{f_1 > C_{\alpha^*} f_0} (1 - \Phi) f_1 d\mu - \int_{f_1 \leq C_{\alpha^*} f_0} \Phi f_1 d\mu \\ &\geq C_{\alpha^*}\left(\int_{f_1 > C_{\alpha^*} f_0} (1 - \Phi) f_0 d\mu - \int_{f_1 \leq C_{\alpha^*} f_0} \Phi f_0 d\mu\right) \\ &= \int (\Phi_{\alpha^*} - \Phi) f_0 d\mu = 0. \end{aligned}$$

Finalement l'erreur de seconde espèce de Φ_{α^*} est moindre que celle de Φ.

1.3. Construction du test optimal pour le problème de détection. Avec les notations précédentes on a $E = \mathbb{R}^n$ et $Z = (X_j)_{j=1,\dots,N}$. La probabilité μ_0 est le produit tensoriel de gaussiennes centrées de variance σ^2 et μ_1 est le produit tensoriel de gaussiennes centrées de variance $\sigma^2 + \xi^2$. On a alors

$$\frac{f_1(Z)}{f_0(Z)} = \frac{\sigma^N \exp - \left(\frac{\sum_{j=1}^N X_j^2}{2(\sigma^2+\xi^2)}\right)}{(\sqrt{\sigma^2+\xi^2})^N \exp - \left(\frac{\sum_{j=1}^N X_j^2}{2\sigma^2}\right)} = \left(\frac{\sigma}{\sqrt{\sigma^2+\xi^2}}\right)^N \exp\left(\frac{\xi^2}{2\sigma^2(\sigma^2+\xi^2)} \sum_{j=1}^N X_j^2\right).$$

La dernière expression est une fonction croissante de $\sum_{j=1}^N X_j^2$. Le théorème de Neyman-Pearson permet alors de conclure que le test optimal (pour une erreur de première espèce $\alpha^* \in]0,1[)$est :

$$\Phi_{\alpha^*}(Z) = \mathbb{1}_{\{\sum_{j=1}^N X_j^2 > K_{\alpha^*}\}} \quad \text{avec} \quad \mu_0\left(\left\{\sum_{j=1}^N X_j^2 > K_{\alpha^*}\right\}\right) = \mathbb{P}\left(\sum_{j=1}^N \varepsilon_j^2 > K_{\alpha^*}\right) = \alpha^*.$$

Remarquons que la variable aléatoire $\sum_{j=1}^{N} \varepsilon_j^2/\sigma^2$ suit la loi du χ^2 à N degrés de liberté. C'est une loi à densité et donc K_{α^*} est unique. De plus, $K_{\alpha^*} = \sigma^2 \chi^2_{\alpha^*,N}$. Calculons l'erreur de seconde espèce associée à cette procédure de test. On a

$$\begin{aligned}
\beta(\Phi_{\alpha^*}) = \mu_1\left(\{\Phi_{\alpha^*} = 0\}\right) &= \mathbb{P}(\sum_{j=1}^{N}(Y_j + \varepsilon_j)^2 \leq K_{\alpha^*}) \\
&= \mathbb{P}\left(\frac{\sum_{j=1}^{N}(Y_j+\varepsilon_j)^2}{\sigma^2+\xi^2} \leq \frac{K_{\alpha^*}}{\sigma^2+\xi^2}\right) \\
&= \psi_N\left(\frac{K_{\alpha^*}}{\sigma^2+\xi^2}\right) = \psi_N\left(\chi^2_{\alpha^*,N}\left[1 - \frac{\xi^2}{\sigma^2+\xi^2}\right]\right)
\end{aligned}$$

1.4. Approximation de K_{α^*} et de $\beta(\Phi_{\alpha^*})$. Par définition, la loi $\chi^2(N)$ est la loi de la somme des carrés de N variables de loi $\mathcal{N}(0,1)$ indépendantes. De plus si $W \sim \mathcal{N}(0,1)$, on vérifie facilement que $\mathbb{E}(W^2) = 1$ et $\mathrm{var}(W^2) = 2$. Soit U_N une variable aléatoire de loi $\chi^2(N)$, le théorème de la limite central permet de déduire de ce qui précède que $\frac{U_N - N}{\sqrt{2N}}$ converge en loi, quand N tend vers l'infini, vers la loi $\mathcal{N}(0,1)$. La Propriété 1 en résulte clairement. En effet, on a

$$\frac{\chi^2_{\alpha^*,N} - N}{\sqrt{2N}} = \mathcal{N}_{\alpha_*} + o(1),$$

qui entraîne

$$\chi^2_{\alpha^*,N} = N + \sqrt{2N}\mathcal{N}_{\alpha_*} + o(\sqrt{N}).$$

Cette dernière égalité donne un développement asymptotique pour K_{α^*} :

$$K_{\alpha^*} = \sigma^2 \chi^2_{\alpha^*,N} = \sigma^2(N + \sqrt{2N}\mathcal{N}_{\alpha_*}) + o(\sqrt{N}).$$

D'autre part, on a

$$\begin{aligned}
\beta(\Phi_{\alpha^*}) = \psi_N\left(\chi^2_{\alpha^*,N}\left[1 - \frac{\xi^2}{\sigma^2+\xi^2}\right]\right) &= \psi_N\left(\frac{\sigma^2(N+\mathcal{N}_{\alpha_*}\sqrt{2N})}{\xi^2+\sigma^2} + o(\sqrt{N})\right) \\
&\approx \psi_N\left(N + \frac{\sqrt{2N}\mathcal{N}_{\alpha_*}\sigma^2}{\xi^2+\sigma^2} - \xi^2\sqrt{\frac{N}{2}}\sqrt{2N}(2)\right) \\
&= \psi_N\left(N + \sqrt{2N}\left[\frac{\sigma^2\mathcal{N}_{\alpha_*} - \xi^2\sqrt{\frac{N}{2}}}{\xi^2+\sigma^2}\right]\right) \\
&\approx \varphi\left(\frac{\sigma^2\mathcal{N}_{\alpha_*} - \xi^2\sqrt{\frac{N}{2}}}{\xi^2+\sigma^2}\right) = \tilde{\beta}\left(\Phi_{\alpha^*}.\right) \qquad (3)
\end{aligned}$$

Les formules d'approximation (2) et (3) sont, sous la forme précédente, mathématiquement abusives. Cependant, nous verrons plus bas sur des simulations numériques, qu'elles permettent de donner un ordre de grandeur *raisonnable* pour $\beta(\Phi_{\alpha^*})$.

Evaluons maintenant la taille de l'échantillon nécessaire pour égaliser l'erreur de seconde espèce à α^*. En utilisant les approximations précédentes on obtient l'équation

$$\varphi\left(\frac{\sigma^2\mathcal{N}_{\alpha^*}-\xi^2\sqrt{\frac{N}{2}}}{\xi^2+\sigma^2}\right)=\alpha^*.$$

Cette égalité est satisfaite dès que

$$N\geq 2\left(\mathcal{N}_{\alpha^*}\left[1+\frac{2\sigma^2}{\xi^2}\right]\right)^2.$$

1.5. Commentaires sur le programme de simulation proposé. Comme nous l'avons vu, les formules d'approximation (2) et (3) n'ont pas été mathématiquement justifiées. Les graphiques générés par le programme proposé semblent indiquer que la formule d'approximation (3) donne une bonne idée de l'ordre de grandeur de l'erreur de seconde espèce. Pour pouvoir effectuer une meilleure comparaison visuelle, on modifie comme suit la partie du programme qui permet de tracer les courbes de comparaison des erreurs de seconde espèce :

```
%
% Calcul de l'erreur relative
%
err=2*(betachap-betatilde)./(betachap+betatilde);
% Affichage des graphiques
% premiere partie de N=1 à 10 subplot(2,3,1);
plot(1 :10,betachap(1 :10));
title('seconde espèce empirique');
ylabel('probabilité')
subplot(2,3,2);
plot(1 :10,betatilde(1 :10));
title('théorique approchée');
subplot(2,3,3);
plot(1 :10,err(1 :10));
title('erreur relative');
% seconde partie de N=11 à Nmax
subplot(2,3,4);
plot(11 :Nmax,betachap(11 :Nmax));
ylabel('probabilité');
xlabel('taille de l''échantillon');
subplot(2,3,5);
plot(11 :Nmax,betatilde(11 :Nmax));
xlabel('taille de l''échantillon');
subplot(2,3,6);
plot(11 :Nmax,err(11 :Nmax));
xlabel('taille de l''échantillon');
```

Après exécution du programme, on obtient les courbes reportées en figure 3. Il apparaît sur ces figures que, pour les paramètres d'études, la formule d'ap-

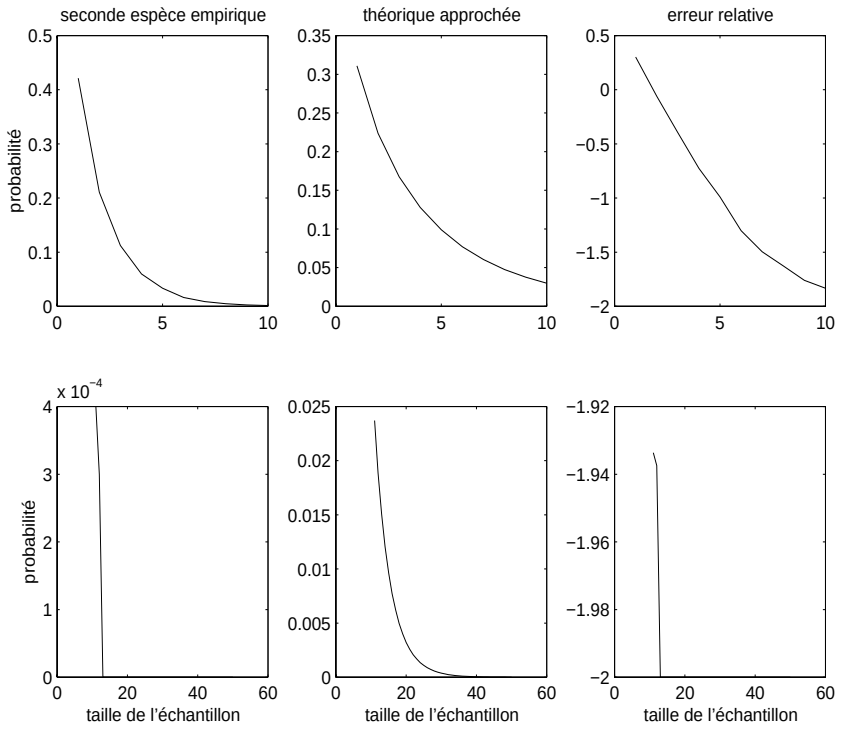


FIG. 2 – Comparaison de l'erreur de seconde espèce

proximation (3) surestime toujours l'erreur de seconde espèce (à partir d'un effectif d'une dizaine de points d'échantillonnage).

1.6. Vers d'autres développements.

1) Les données obtenues à partir des fonctions g_1 et g_2 peuvent être utilisées pour valider le modèle de détection proposé. En particulier, on peut dans ces exemples comparer l'erreur de première espèce α^* théorique utilisée dans la procédure de test avec l'erreur de première espèce effective à l'aide de simulations numériques.
2) L'évaluation du seuil K_{α^*} et de l'erreur de seconde espèce relève plus d'un calcul de grandes déviations que du théorème de la limite central. On peut approcher ces quantités en utilisant le théorème de grandes déviations de Cramer.

Chapitre 21
Commentaire d'un spécialiste

§1. Introduction

Pour introduire la problématique du texte, considérons l'exemple suivant. On recueille une observation, x issue d'une loi de probabilité $\mathcal{L}$. Sous l'hypothèse nulle ($\mathbf{H}_0$), la loi est uniforme sur l'intervalle $[1,3]$: $\mathcal{L} = \mathcal{U}_{[1,3]}$. Sous l'hypothèse alternative ($\mathbf{H}_1$), la loi est uniforme sur l'intervalle $[2,6]$: $\mathcal{L} = \mathcal{U}_{[2,6]}$. Si $1 \leq x < 2$, il faut clairement privilégier $\mathbf{H}_0$; de même, l'hypothèse alternative $\mathbf{H}_1$ sera acceptée si $3 < x \leq 6$. On pourra représenter les densités des deux lois pour s'en convaincre. Le doute subsiste si x appartient à l'intervalle $[2,3]$ commun aux deux distributions. Le texte présente une méthode de résolution de ce problème, qui possède en outre des propriétés d'optimalité qui seront détaillées.

L'exposé sera organisé comme suit. Une première partie rappelle les définitions de test et de test uniformément plus puissant. La partie suivante est consacrée à la démonstration du lemme de Neyman-Pearson. La dernière partie considère une version randomisée du résultat précédent.

§2. Les définitions

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité et $(E, \mathcal{B})$ un espace mesurable. On considère une variable aléatoire (v.a.), Z, définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $(E, \mathcal{B})$. On suppose que la loi image de $\mathbb{P}$ (notée G) appartient à une famille $\mathcal{P}$ de lois de probabilité sur $(E, \mathcal{B})$. Soit $\mathcal{P}_0 \subset \mathcal{P}$ et $\mathcal{P}_1 = \{G \in \mathcal{P} : G \notin \mathcal{P}_0\}$ une partition de $\mathcal{P}$. Un test d'hypothèses se propose de décider laquelle des deux hypothèses

$$\mathbf{H}_0 : G \in \mathcal{P}_0 \qquad \text{ou} \qquad \mathbf{H}_1 : G \in \mathcal{P}_1 \tag{1}$$

est vraie.

Définition 2.1 *Un test est une fonction mesurable $T : E \to [0,1]$.*

Remarque 2.1 *Le test T conduit à rejeter $\boldsymbol{H}_0$ avec probabilité $T(Z)$ et à accepter $\boldsymbol{H}_0$ avec probabilité $1 - T(Z)$. D'où deux risques d'erreurs corrélatifs à la décision prise, dont on retire la :*

Définition 2.2 *(•) L'erreur de première espèce est la probabilité de rejeter $\boldsymbol{H}_0$ quand $G \in \mathcal{P}_0$:*

$$\alpha_T(G) = \mathbb{E}(T(Z)) \tag{2}$$

Le niveau de signification du test est $\sup_{G \in \mathcal{P}_0} \alpha_T(G)$.

(•) *L'erreur de seconde espèce est la probabilité de rejeter* $\boldsymbol{H}_1$ *quand* $G \in \mathcal{P}_1$. *Sa contre probabilité est la puissance du test,* $\beta_T(G)$, *définie par le terme de droite dans (2).*

Définition 2.3 *Soit* $\alpha \in]0,1[$. *Un test* T^* *de niveau* α *est uniformément plus puissant (U.P.P.) si tout test* T *de niveau inférieur ou égal à* α *vérifie*

$$\beta_T(G) \leq \beta_{T^*}(G), \qquad G \in \mathcal{P}_1. \tag{3}$$

§3. Le test de Neyman et Pearson

3.1. Le cas non randomisé. On se propose ici de compléter la démonstration du Théorème 1 du texte, en reprenant les indications fournies.

Proposition 3.1 *On pose* $V = f_1(Z)/f_0(Z)$. *Si* $G = \mu_0$, *alors* V *est une v.a.* $\mathbb{P}$*-presque sûrement finie. De plus, pour tout* $\alpha \in]0,1[$, *sa fonction de répartition, possède un unique quantile supérieur de niveau* α *(noté* C_α*).*

Démonstration. La mesurabilité de V est immédiate.

Puisque f_0 est une densité, l'événement $\{f_0(Z) > 0\}$ est quasi certain.

Puis

$$\begin{aligned} 1 = \int_E f_1(z)d\mu(z) &\geq n \int_{\{f_1 > n\}} d\mu(z) \\ &= n \int_{\{f_1 > n, f_0 > 0\}} \tfrac{1}{f_0(z)} d\mu(z), \end{aligned}$$

entraîne par le lemme de Beppo Lévi que l'événement $\{f_1(Z) = \infty\}$ est quasi impossible. La v.a. V est donc $\mathbb{P}$-p.s. finie. Pour tout nombre réel t, il vient que $F(t) = \mathbb{P}(f_1(Z) \leq t f_0(Z))$.

La fonction F est continue car

$$\begin{aligned} F(t) - F(t-) &= \mathbb{P}(f_1(Z) = t f_0(Z)) \\ &= \mu_0 \{z \in E : f_1(z) = t f_0(z)\}, \end{aligned}$$

est nul par hypothèse.

Enfin, $F(0) = \mathbb{P}(f_1(Z) < 0)) = 0$. Le théorème de convergence dominée de Lebesgue établit que $F(t) \to \mathbb{P}(f_0(Z) = 0)) = 0$, lorsque $t \to \infty$. La fonction F réalise une bijection de $]0, +\infty[$ sur $]0,1[$, ce qui prouve l'existence et l'unicité de C_α $\square$

Il est aisé de calculer l'erreur de première espèce du test Φ_{α^*}.

$$\alpha_{\Phi_{\alpha^*}}(\mu_0) = \mathbb{E}(\Phi_{\alpha^*}(Z)) = \mu_0 \{z \in E : f_1(z) > C_{\alpha^*} f_0(z)\} = \alpha^*. \tag{1}$$

Proposition 3.2 *Soit* $\alpha^* \in]0.1[$. *Alors* $\beta_{\Phi_{\alpha^*}}(\mu_1) \leq \beta_\Phi(\mu_1)$ *pour tout test* Φ *de niveau inférieur ou égal à* α^*.

Démonstration.

$$\begin{aligned}\beta_{\Phi_{\alpha^*}}(\mu_1) - \beta_{\Phi}(\mu_1) &= \int_E \{\Phi_{\alpha^*}(z) - \Phi(z)\} f_1(z) d\mu(z) \\ &= \int_{\{\Phi_{\alpha^*}=\Phi\}} + \int_{\{\Phi_{\alpha^*}>\Phi\}} + \int_{\{\Phi_{\alpha^*}<\Phi\}} (\Phi_{\alpha^*}(z) - \Phi(z)) f_1(z) d\mu(z).\end{aligned}$$

La première intégrale est nulle. Le deuxième ensemble d'intégration est inclus dans $\{\Phi_{\alpha^*} = 1\} = \{f_1 > C_{\alpha^*} f_0\}$. En particulier,

$$(f_1(z) - C_{\alpha^*} f_0(z))(\Phi(z) - \Phi_{\alpha^*}(z)) \leq 0. \tag{2}$$

Par un raisonnement analogue, on prouve que l'inégalité (2) est encore valable sur le dernier ensemble d'intégration. Finalement,

$$\begin{aligned}\beta_{\Phi_{\alpha^*}}(\mu_1) - \beta_{\Phi}(\mu_1) &= \int_E \{\Phi(z) - \Phi_{\alpha^*}(z)\}\{f_1(z) - C_{\alpha^*} f_0(z)\} d\mu(z) + C_{\alpha^*} \int_E \{\Phi(z) - \Phi_{\alpha^*}(z)\} d\mu_0(z) \\ &\leq 0 + C_{\alpha^*}(\alpha_{\Phi}(\mu_0) - \alpha_{\Phi_{\alpha^*}}(\mu_0)) \\ \beta_{\Phi_{\alpha^*}}(\mu_1) - \beta_{\Phi}(\mu_1) &\leq 0 \quad \square\end{aligned}$$

3.2. Le test de détection. On revient au modèle bruit et signal plus bruit de la première partie du texte.

Proposition 3.3 *Soit $\alpha^* \in]0.1[$. Alors le théorème de Neyman et Pearson conduit au test (pour une erreur de première espèce α^* et un échantillon $Z = (X_1, \ldots, X_n)$) :*

$$\Phi_{\alpha^*}(Z) = \mathbb{1}_{\left\{\sum_{j=1}^n X_j^2 > K_{\alpha^*}\right\}} \tag{3}$$

avec $K_{\alpha^*} = Var(\varepsilon_1)\chi^2_{\alpha^*,n} = \sigma^2 \chi^2_{\alpha^*,n}$. De plus, ce test est de puissance $\psi_n\left(\frac{K_{\alpha^*}}{\sigma^2+\xi^2}\right)$.

Démonstration. Sous l'hypothèse nulle, la v.a.r. la loi X_1 suit la loi normale $\mathcal{N}(0, \sigma^2)$. La v.a. Z a donc pour densité (par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$)

$$f_0(z) = (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{\sum_{j=1}^n x_j^2}{2\sigma^2}\right\}, \qquad z = (x_1, \ldots, x_n) \tag{4}$$

La densité f_1 de la v.a. Z sous l'hypothèse alternative s'exprime analoguement, en changeant dans (4) le terme de variance σ^2 par $\sigma^2 + \xi^2$.

De plus, pour tout nombre réel positif t, l'ensemble $\{z \in \mathbb{R}^n : f_1(z) = t f_0(z)\}$ est soit vide, soit une sphère. Sa mesure est nulle, pour la mesure de Lebesgue, puis par absolue continuité, pour la loi μ_0 de densité (3). D'après le lemme de Neyman et Pearson, le test $\Phi_{\alpha^*}(Z)$ vaut 1 sur l'ensemble

$$\{z \in \mathbb{R}^n : f_1(z) > C_{\alpha^*} f_0(z)\} = \left\{ z = (x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{j=1}^{n} x_j^2 > C(\alpha^*) \right\} \tag{5}$$

La constante $C(\alpha^*)$ est déterminée en relation avec la formule (1). Puisque sous $\mathbf{H}_0$ on a $X_1 = \varepsilon_1$,

$$\begin{aligned} \alpha &= \alpha_{\Phi_{\alpha^*}}(\mu_0) = \mathbb{E}\left(\mathbb{I}_{\left\{\sum_{j=1}^n \varepsilon_j^2 > C(\alpha^*)\right\}}\right) \\ &= \mathbb{P}\left(\frac{\sum_{j=1}^n \varepsilon_j^2}{\sigma^2} > \frac{C(\alpha^*)}{\sigma^2}\right) \\ \alpha &= 1 - \psi_n\left(\frac{C(\alpha^*)}{\sigma^2}\right) \end{aligned}$$

Ce qui établit que $C(\alpha^*) = \sigma^2 \chi^2_{\alpha^*,n}$.

La puissance du test se calcule en relation avec la formule (2). Sous l'hypothèse $\mathbf{H}_1$, on a $X_1 = Y_1 + \varepsilon_1$, donc

$$\begin{aligned} 1 - \beta_{\Phi_{\alpha^*}}(\mu_1) &= \mathbb{P}\left(\frac{\sum_{j=1}^n (Y_j + \varepsilon_j)^2}{\xi^2 + \sigma^2} > \frac{C(\alpha^*)}{\sigma^2}\right) \\ &= 1 - \psi_n\left(\frac{C(\alpha^*)}{\xi^2 + \sigma^2}\right) \quad \square \end{aligned}$$

Remarque 3.1 *Le test Φ_{α^*} ne dépend pas de ξ. On peut établir qu'il est U.P.P. (au sens de la définition 2.3) pour la famille de lois $\mathcal{P} = \{\mathcal{N}(0, \theta^2) : \theta > 0\}$ si l'échantillon est de taille un) en vue de tester*

$$\boldsymbol{H}_0 : \theta = \sigma \qquad \textit{contre} \qquad \boldsymbol{H}_1 : \theta \neq \sigma \tag{6}$$

Etablissons la propriété 1 du texte.

Soit une suite de v.a. $(U_i)_{i \in \mathbb{N}}$ i.i.d. de loi $\mathcal{N}(0,1)$. Pour tout entier $n \geq 1$ la v.a. $W_n = \sum_{i=1}^n U_i^2$ suit la loi $\chi^2(n)$. De plus, $\mathbb{E}(W_n) = n$ et $\mathrm{Var}(W_n) = \sum_{i=1}^n \mathrm{Var}(U_i^2) = 2n$. Le théorème de limite centrale de Paul Lévy assure que, quel que soit $t \in \mathbb{R}$, si $n \to \infty$

$$\mathbb{P}\left(\frac{W_n - n}{\sqrt{2n}} \leq t\right) \to \varphi(t) \tag{7}$$

Pour établir la première partie de la propriété, on pose $t = t_n = \frac{\chi^2_{\alpha^*,n} - n}{\sqrt{2n}}$. Alors le terme de gauche est égal à $1 - \alpha^*$. Soit une valeur d'ahérence de la suite (t_n) : $\ell = \lim_{k \to \infty} t_{n_k}$. On note $V_k = \left(\frac{W_{n_k} - n_k}{\sqrt{2n_k}}\right)$.

Lemme 3.1 *Lorsque $k \to \infty$, $\left(t_{n_k}^{-1} - \ell^{-1}\right) V_k \to 0$ (en probabilité).*

Soit en effet $\delta, \epsilon > 0$. Il existe $M > 0$ tel que $\varphi(M) - \varphi(-M) \geq 1 - \delta$. Pour k suffisamment grand ($k \geq k_0$), $\left|t_{n_k}^{-1} - \ell^{-1}\right| \leq \epsilon/M$. Pour n suffisamment grand ($n \geq N_1$) le théorème de la limite centrale assure que

$$\mathbb{P}\left(\left|\frac{W_n-n}{\sqrt{2n}}\right| \geq M\right) \geq 1-2\delta \tag{8}$$

Alors pour tout k suffisamment grand ($n_k \geq \max(N_{k_0}, N_1)$)

$$\begin{aligned}\mathbb{P}\left(\left|\left(t_{n_k}^{-1}-\ell^{-1}\right)V_k\right|>\epsilon\right) &\leq \mathbb{P}\left(\left|\left(t_{n_k}^{-1}-\ell^{-1}\right)V_k\right|>\epsilon, |V_k|\geq M\right)+\mathbb{P}\left(\left|\left(t_{n_k}^{-1}-\ell^{-1}\right)V_k\right|>\epsilon, |V_k|>M\right)\\ &\leq 2\delta+0\end{aligned}$$

La valeur δ est arbitraire. Le lemme est démontré □

Il résulte du lemme (3.1) et du lemme de Slutsky que

$$t_{n_k}^{-1}V_k = \ell^{-1}V_k + \left(t_{n_k}^{-1}-\ell^{-1}\right)V_k \overset{\text{loi}}{\to} \mathcal{N}\left(0,\ell^{-2}\right) \tag{9}$$

Donc

$$\begin{aligned}1-\alpha^* &= \mathbb{P}\left(V_k \leq t_{n_k}\right)\\ &\to \varphi(\ell)\end{aligned}$$

La suite (t_n) possède une unique valeur d'adhérence : $\ell = \mathcal{N}_{\alpha^*}$. Elle converge, et sa limite est $\mathcal{N}_{\alpha^*}$. La propriété 1 est établie □

Les formules (4), (5) et (6) en découlent aisément.

3.3. Application informatique.

```
write(%io(2), 'taille de l échantillon')
// Début de saisie des donnees
nmax=read(%io(1),1,1) ;
write(%io(2), 'nombre de trajectoires simulées'), nsimul=read(%io(1),1,1) ;
write(%io(2), 'variance du signal'), xi2=read(%io(1),1,1) ;
write(%io(2), 'variance du bruit'), si2=read(%io(1),1,1) ;
varx=si2+xi2 ; sigx=sqrt(varx) ; rap1=si2/varx ; rap2=xi2/varx ;
write(%io(2), 'erreur alpha'), alpha=read(%io(1),1,1) ;
nalpha=cdfnor("X",0,1,1-alpha,alpha) ; //Fin de saisie des donnees
x=sigx*rand(nmax,nsimul,'n') ;
x2=x.*x ; sommex2=cumsum(x2,'r') ;
un=ones(1,nsimul) ; z=1 :1 :nmax ; //L'estimation du seuil K
ktilde=si2*(z+(nalpha*sqrt(2*z))) ;
masque=(ktilde)'*un ;
phi=bool2s(sommex2 > masque) ;
ktildeb=(si2/2)*((nalpha+sqrt(2*z-1))^2) ; //L'estimation de Fisher du seuil
K
masqueb=(ktildeb)'*un ;
phib=bool2s(sommex2 > masqueb) ;
//Les erreurs de seconde espece empiriques
betest=mean((1-phi)','r') ; betestb=mean((1-phib)','r') ;
// Vecteurs ligne des moyennes-colonnes théoriques
t=(nalpha*rap1)-(rap2*sqrt(z/2)) ; moy=zeros(1,nmax) ;
ect=ones(1,nmax) ; betheo=cdfnor("PQ",t,moy,ect) ;
```

```
tb=(sqrt(rap1)*(nalpha+sqrt(2*z-1)))-sqrt(2*z-1) ;
betheob=cdfnor("PQ",tb,moy,ect) ;
//Les graphiques
xsetech([0,0,0.5,0.4]) ; plot2d(z',betest',3,'151','fig.1a',[0,0,nmax,1])
xstring(nmax/2,0.5,["Beta " ; "empirique"])
xsetech([0,0.6,0.5,0.4]) ; plot2d(z',betheo',2,'151','fig.2a',[0,0,nmax,1])
xstring(nmax/2,0.5,["Beta " ; "approche"])
xset("default') ; xsetech([0.5,0,0.5,0.4]) ;
plot2d(z',betestb',3,'151','fig.1b',[0,0,nmax,1])
xstring(nmax/2,0.5,["Beta " ; "empirique Fisher"])
xsetech([0.5,0.6,0.5,0.4]) ; plot2d(z',betheob',2,'151','fig.2b',[0,0,nmax,1])
xstring(nmax/2,0.5,["Beta " ; "approche Fisher"])
```

Ce programme rédigé sous Scilab reprend l'essentiel des fonctionnalités de celui présenté dans le texte. Il y ajoute une estimation du seuil K_{α^*} du test optimal du problème de détection basée sur l'approximation de Fisher

$$\sqrt{2W_n} - \sqrt{2n-1} \quad \overset{\text{loi}}{\to} \quad \mathcal{N}(0,1) \tag{10}$$

Nous renvoyons à [1]. Elle conduit à une seconde estimation du seuil K_{α^*}

$$K_{\alpha^*} \approx Kb_{\alpha^*} = 2\sigma^2 \left(\frac{\mathcal{N}_{\alpha^*}+\sqrt{2n-1}}{2}\right)^2, \tag{11}$$

ainsi qu'à une erreur de seconde espèce approchée (dite de Fisher dans la légende du programme).

§4. Le test randomisé de Neyman et Pearson

4.1. L'exemple introductif. Revenons à l'exemple introductif de détection d'un signal de loi uniforme sur l'intervalle $[1,3]$ contre un second signal de loi uniforme sur l'intervalle $[2,6]$. On recueille une observation, x issue d'une loi de probabilité $\mathcal{L}$. Le signal x sera certainement issu de la loi de l'hypothèse nulle $(\mathbf{H}_0)$ si $1 \leq x < 2$; en revanche, l'hypothèse alternative $\mathbf{H}_1$ sera acceptée si $3 < x \leq 6$. Le doute qui subsiste si x appartient à l'intervalle $[2,3]$ commun aux deux distributions est levé par la procédure suivante : attribuer une probabilité $\gamma_\alpha(G)$ de rejeter l'hypothèse nulle $(\mathbf{H}_0)$ de façon que l'erreur de première espèce

$$\mu_0\left(\{x : 3 < x \leq 6\}\right) + \gamma_\alpha(\mu_0)\ \mu_0\left(\{x : 2 < x \leq 3\}\right) = 0 + \tfrac{\gamma_\alpha(\mu_0)}{2}, \tag{1}$$

soit égale à α, soit $\gamma_\alpha(\mu_0) = 2\alpha$. D'où le test

$$\Phi_\alpha(z) = \begin{cases} 0 & \text{si } 1 \leq z < 2, \\ 2\alpha & \text{si } 2 \leq z < 3, \\ 1 & \text{si } 3 \leq z \leq 6. \end{cases} \tag{2}$$

Ce test n'est pas une conséquence du théorème 1 du texte :

$$\mu_0\left(\left\{x : f_1(x) = \frac{f_0(x)}{2}\right\}\right) = \frac{1}{2}. \tag{3}$$

Il illustre l'extension naturelle du test de Neyman et Pearson (pour des hypothèses simples) et possède les mêmes propriétés d'optimalité (U.P.P.). Le lecteur intéressé est renvoyé à la consultation des ouvrages [4] et [3].

4.2. Application à un test de proportion. On désire tester l'hypothèse d'un signal de loi de Bernoulli de paramètre $p \in]0,1[$. Plus précisément,

$$\mathbf{H}_0 : p = p_0 \qquad \text{contre} \qquad \mathbf{H}_1 : p = p_1 \tag{4}$$

avec $p_1 > p_0$. On considère un échantillon Z de taille n ; on définit alors la v.a.r. $Y = \sum_{i=1}^{n} X_i$. En tirant parti de la croissance de la vraisemblance V (définie dans la proposition 3.1) par rapport à Y, on établit la

Proposition 4.1 *Soit $\alpha^* \in]0,1[$. Pour une erreur de première espèce α^* il existe un unique entier k et une unique probabilité $\gamma(\alpha^*)$ tels que le test*

$$\Phi_{\alpha^*}(Z) = \begin{cases} 1 & si\ Y > k, \\ \gamma(\alpha^*) & si\ Y = k, \\ 0 & si\ Y < k, \end{cases} \tag{5}$$

possède une erreur de première espèce égale à α^*. De plus, ce test est U.P.P.

Remarque 4.1 *Le lecteur peut consulter [4] pour une démonstration de ce résultat. Il y trouvera également des extensions à d'autres familles de lois paramétriques très générales. De même que dans le théorème 1, on peut remarquer que le test ne dépend pas de la valeur p_1 du paramètre p sous l'hypothèse alternative $\boldsymbol{H}_1$; il est encore U.P.P. lorsque cette dernière est remplacée par l'hypothèse composite $\boldsymbol{H}_1 : p > p_0$.*

Les constantes m et $\gamma(\alpha^*)$ se calculent en exprimant l'erreur de première espèce du test

$$\alpha^* = \sum_{j=k+1}^{n} \mathrm{C}_n^j p_0^j (1-p_0)^{n-j} + \gamma(\alpha^*)\, \mathrm{C}_n^k p_0^k (1-p_0)^{n-k}.$$

Le tableau suivant indique les valeurs obtenues sous Scilab, pour $(k, \gamma(\alpha^*))$ avec $p_0 = 0.40$ et $\alpha^* = 0.05$, puis $\alpha^* = 0.01$.

	$\alpha = 0.01$		$\alpha = 0.05$	
	k	γ	k	γ
$n = 10$	8	0.7839	7	0.8879
$n = 50$	28	0.2831	26	0.7169
$n = 100$	52	0.9988	48	0.3583
$n = 1000$	436	0.3109	426	0.9632

1. H. Cramér, (1999). *Mathematical methods of statistics,* 19ème édition. Princeton University Press.
2. D. Dacuhna-Castelle, M. Duflo, M. (1983). *Probabilité et statistique.* Masson.
3. E.L. Lehmann, (1986). *Testing statistical hypotheses,* 2ème édition. Springer Verlag.
4. J. Shao, (1998). *Mathematical statistics.* Springer Verlag.

Septième partie

Fonctions de Lyapounov et stabilité de réseaux de télécommunication

Chapitre 22
Fonctions de Lyapounov et stabilité de réseaux de télécommunications

Résumé

Nombre de réseaux de télécommunications ou de transport se laissent modéliser de façon assez réaliste à l'aide de chaînes de Markov à espace d'états dénombrable. Un problème important est alors de déterminer les conditions d'existence d'un régime stationnaire pour ces systèmes. Plus formellement, on parlera de *conditions de stabilité*. À cette fin, une méthode particulièrement puissante consiste à construire des *fonctions de Lyapounov* idoines. Après avoir énoncé les principaux théorèmes nécessaires, le texte propose trois exemples d'applications à l'évaluation de protocoles (algorithmes) probabilistes, utilisés pour gérer l'accès à une ressource (bus, canal radio ou satellite, etc.) partagée entre un grand nombre d'utilisateurs.

Quelques mots clés : Chaîne de Markov, martingale, protocole, récurrence, transience, accès multiple, arbre.

Indications préliminaires

L'analyse des modèles proposés dans la section 2 s'appuie sur les fondements théoriques de la section 1. Par souci de complétude, on donne dans l'appendice 3 le vocabulaire et les notions de base, en soulignant que ces préalables doivent être *admis sans démonstration*, leur lecture pouvant ainsi être relativement *rapide*.

§1. Critères de classification des chaînes de Markov à espace d'états dénombrable

Soit $\{X_t\}_{t\in\mathbb{N}}$ une chaîne de Markov (CM) temporellement homogène (le temps t est discret), à espace d'états dénombrable $\mathcal{A}$, irréductible et apériodique (CMIA). L'espace probabilisé sous-jacent est noté $(\Omega, \mathcal{F}, P)$ et on désigne par $\{\mathcal{F}_t\}_{t\geq 0}$ une famille croissante de tribus formant l'histoire de la CM, avec $\lim_{t\to\infty} \mathcal{F}_t = \mathcal{F}$. La matrice stochastique de transition (non nécessairement finie) est indicée par les points de $\mathcal{A}$ et écrite $\mathbf{P} = \{p_{\alpha\beta}\}_{\alpha,\beta\in\mathcal{A}}$.

Dans la suite, on va s'intéresser aux conditions d'existence des solutions probabilistes de l'équation

$$\pi = \pi\mathbf{P}, \tag{1}$$

où $\pi = (\pi_\alpha,\ \alpha \in \mathcal{A})$ est le vecteur (ligne) inconnu, communément appelé *mesure invariante (ou stationnaire)*.

En général, il y a peu d'espoir d'obtenir une solution explicite de (1) mathématiquement *exploitable*, sauf dans quelques situations bien identifiées (processus de naissance et de mort, certains réseaux de files d'attente, etc). Par contre, il existe des méthodes efficaces pour déterminer les conditions d'ergodicité, sans avoir à calculer la mesure invariante : l'idée repose sur la construction de *fonctions de Lyapounov*, qui sont en fait des sur-martingales ou des sous-martingales.

Bien sûr, lorsque $\mathcal{A}$ est fini et sous les hypothèses faites au début de cette section, la CM est ergodique. Cependant, dans certains de modèles issus des applications, le nombre d'états n'est pas fini et la recherche des conditions précitées n'est pas évidente. On donne ci-après des critères généraux, qu'on propose d'appliquer dans la section 2 à divers exemples empruntés au domaine des télécommunications. Ces critères sont puissants et certains d'entre eux fournissent des conditions nécessaires et suffisantes (CNS). Le formalisme est celui des martingales et plusieurs démonstrations (omises ici) s'appuient sur les théorèmes de convergence des martingales positives. Le résultat le plus important est sans doute celui de Foster (1953), qui fut le point de départ de nombreux travaux.

Theorem 1.1. [Foster]
Une CNS pour que la chaîne de Markov $\{X_t\}_{t\geq 0}$ soit ergodique est qu'il existe

une fonction positive $f : \mathcal{A} \to \mathbb{R}^+$, *un nombre* $\epsilon > 0$ *et un ensemble fini* U, *tels que*

$$\begin{cases} \mathbb{E}[f(X_{t+1}) - f(X_t)|X_t = \alpha] \leq -\epsilon \quad \forall \alpha \notin U, \\ \mathbb{E}[f(X_{t+1})|X_t = \alpha] < \infty, \quad \forall \alpha \in U. \end{cases} \tag{2}$$

Corollary 1.1. [Lemme de Pakes]
Si une CM $\{X_t\}_{t\geq 0}$ *définie sur* $\mathcal{A} = \mathbb{N}$ *vérifie,* $\forall t \geq 0$ *et* $\forall i \in \mathcal{A}$,

$$\begin{cases} \limsup_{i\uparrow\infty} \mathbb{E}[X_{t+1} - X_t|X_t = i] < -\epsilon, \\ \mathbb{E}[X_{t+1}|X_t = i] < \infty, \end{cases} \tag{3}$$

alors elle est ergodique.

Le corollaire 1.1 se déduit immédiatement du théorème 1.1.

Theorem 1.2. *La CM* $\{X_t\}_{t\geq 0}$ *est transiente si et seulement si il existe une fonction positive* $f : \mathcal{A} \to \mathbb{R}^+$ *et un ensemble* U *(non nécessairement fini), tels que les inégalités suivantes soient satisfaites :*

$$\begin{cases} \mathbb{E}[f(X_{t+1}) - f(X_t)|X_t = \alpha] \leq 0, \quad \forall \alpha \notin U, \\ f(\beta) < \inf_{\gamma\in U} f(\gamma), \quad \textit{pour au moins un état } \beta \notin U. \end{cases} \tag{4}$$

Typiquement, dans les applications où $\mathcal{A} = \mathbb{N}$, la fonction f du théorème 1.2 tendra vers zéro à l'infini. En fait, la question de la transience est délicate, car souvent liée à la construction de sous-martingales. On donne ci-dessous une condition suffisante *naturelle*, dans laquelle intervient la bornitude inférieure des sauts de la fonction.

Theorem 1.3. *Pour que la CM* $\{X_t\}_{t\geq 0}$ *soit transiente, il suffit qu'il existe une fonction positive* $f : \mathcal{A} \to \mathbb{R}^+$, *des constantes positives* d, ϵ, C, *telles que l'ensemble* $U \stackrel{\text{def}}{=} \{x : f(x) > C\}$ *soit non vide et que*

$$\begin{cases} \mathbb{E}[f(X_{t+1}) - f(X_t)|X_t = \alpha] \geq \epsilon, \quad \forall \alpha \in U, \\ f(X_{t+1}) - f(X_t) \geq -d, \quad \forall t \quad \textit{presque sûrement.} \end{cases} \tag{5}$$

§2. Application à la stabilité de certains protocoles de communications

2.1. Généralités. Vers la fin des années soixante, d'importants moyens furent mis en œuvre pour concevoir des réseaux à accès multiple, de grande

taille et d'utilisation aisée. Il s'agissait principalement de faire communiquer un grand nombre de sites (station, utilisateur) par l'intermédiaire d'un medium partagé unique (canal radio ou satellite, bus, fibre optique). De là naquit en 1968 à l'université d'Hawaï, sous l'impulsion de l'ingénieur N. Abramson, le plan du premier réseau informatique radio ALOHA, dont le nom signifie « bienvenue » en Hawaïen. Il s'agissait d'une alternative aux liaisons point-à-point classiques et son principe en était fort simple : chaque station émet dès qu'elle le souhaite, de façon parfaitement décentralisée.

Le nœud Gordien du système est évidemment la gestion des conflits d'accès (*collisions*) qui se traduisent par un brouillage : chaque station impliquée dans un conflit retransmet son message selon un certain algorithme, qu'on appelle *protocole*. Cette problématique fut le point de départ d'un vaste champ d'activités, intéressant des physiciens, des informaticiens et des mathématiciens.

Modèle du canal partagé.

– Chaque utilisateur émet des *paquets* de taille fixe.

– Le temps canal t est discret, découpé en tranches élémentaires (en anglais *slot*), $t = 0, 1, 2, \ldots$, dont la longueur représente la durée de transmission d'un paquet. En outre les stations sont synchronisées, une émission ne pouvant être lancée qu'au début d'une tranche.

– À tout instant les utilisateurs peuvent écouter le canal et seuls les états suivants sont observables :

(**0**) : Aucune transmission, le canal est silencieux.

(**1**) : Un seul paquet a été émis ; transmission avec succès.

(**C**) : Au moins deux utilisateurs ont tenté d'émettre simultanément et il y a brouillage. Les messages impliqués dans la collision sont alors *bloqués* et devront être retransmis.

Soit A_t le nombre de paquets nouveaux générés par l'ensemble des stations à l'instant t. On suppose que la suite $\{A_t\}_{t\geq 0}$ est formée de variables i.i.d., avec la distribution

$$\mathbb{P}(A_t = i) = a_i, \quad i \geq 0, \tag{6}$$

dont la moyenne $\lambda \stackrel{\text{def}}{=} \mathbb{E}(A_t)$ représente le trafic sur le canal. Pour éviter les situations triviales, on supposera toujours

$$0 < a_0 < 1 \quad \text{et} \quad \mathbb{P}(A_t > 1) > 0. \tag{7}$$

On décrit ci-après trois protocoles de nature probabiliste (ALOHA, ETHERNET et le protocole en ARBRE), dont certaines versions sont actuellement utilisées dans les réseaux locaux.

2.2. Le protocole Aloha. – Lorsqu'une station a un nouveau paquet à transmettre, elle émet.

– Si la transmission a échoué, la politique est de retransmettre à chaque instant avec une probabilité fixe $f \in [0, 1]$, identique pour chaque station, et ce jusqu'au succès final.

Soit X_t le nombre de message bloqués à l'instant t. D'après (7), $\{X_t\}_{t\geq 0}$ est une CMIA, dont la matrice de transition satisfait, en posant $c_i = if(1-f)^{i-1}$ et $b_i = (1-f)^i$,

$$p_{ij} = \begin{cases} a_0 c_i & \text{si } j = i-1, \\ a_0(1-c_i) + a_1 b_i & \text{si } j = i, \\ a_1(1-b_i) & \text{si } j = i+1, \\ a_{j-i} & \text{si } j \geq i+2. \end{cases} \tag{8}$$

Proposition 2.1. *Le protocole* ALOHA *est instable, au sens où la chaîne* $\{X_t\}$ *est transiente.*

En fait, on peut constater et prouver que le système explose exponentiellement lentement.

Aloha contrôlé. Au vu de la proposition 2.1, il est normal de chercher à contrôler le canal, en agissant sur la politique de retransmission, c'est-à-dire sur f. En admettant que les stations à l'écoute du trafic sur le canal connaissent à tout instant le nombre X_t de messages en attente, on considérera des politiques de retransmission markoviennes de la forme $f(X_t)$. On conserve les notations précédentes en remplaçant implicitement f par $f(i)$.

En introduisant la quantité $d_i \stackrel{\text{def}}{=} a_0 c_i + a_1 b_i$, $\forall i \geq 0$ (le *débit conditionnel* du canal), on a alors le résultat suivant, qui englobe la proposition 2.1 :

Proposition 2.2.

- *Lorsque* $\lambda < \liminf_i d_i$, $\{X_t\}$ *est ergodique.*
- *Lorsque* $\lambda > \limsup_i d_i$, $\{X_t\}$ *est transiente.*

Si on suppose en outre l'existence de la quantité $p \stackrel{\text{def}}{=} \lim_{i\to\infty} if(i)$, *la CNS d'ergodicité s'écrit sous la forme*

$$\lambda < e^{-p}[pa_0 + a_1]. \tag{9}$$

En particulier, le système est toujours transient si $a_0 \leq a_1$. Lorsque le processus des arrivées est de Poisson, i.e. $a_i = e^{-\lambda}\dfrac{\lambda^i}{i!}$, $\forall i \geq 0$, le débit maximum du canal contrôlé vaut $e^{-1} = 0.3678$.

2.3. Le protocole Ethernet. Ce protocole avait été proposé par B. Metcalfe [1976]. La seule différence avec ALOHA est que la probabilité de retransmission f_k d'un message dépend du nombre d'échecs k qu'il a déjà subis.
L'état du système est décrit par un vecteur $Z_t = (Z_1(t), Z_2(t), ...)$ dont le nombre de composantes à l'instant t est fini, $Z_i(t)$ représentant le nombre de messages retransmis i fois sans succès. Donc $\{Z_t\}_{t\geq 0}$ est une CMIA. Lorsque $f_k = f$ pour tout $k \geq 1$, on retrouve exactement ALOHA.

Proposition 2.3. *Lorsque* $\liminf_k f_k > 0$, $\{Z_t]_{t\geq 0}$ *est transiente et* ETHERNET *est instable.*

Une démonstration possible consiste à appliquer le théorème 1.3 à la fonctionnelle $Y_t = \sum_{l=0}^{\infty} Z_l(t)$, qui vérifie

$$\mathbb{E}[Y_{t+1}|Z_t] = Y_t + \lambda - D_t, \tag{10}$$

en posant $D_t = P_t[a_1 + a_0 S_t]$, avec

$$P_t = \prod_{l=1}^{\infty}(1-f_l)^{Z_l(t)}, \quad S_t = \sum_{l=1}^{\infty} \frac{f_l Z_l(t)}{1-f_l}. \tag{11}$$

Lorsque $\liminf_l f_l = 0$, le problème de la stabilité est toujours *ouvert*. En fait, le protocole original, parfois appelé *exponential back-off* (EBO), correspondait au choix $f_l = 2^{-l}$, l étant le nombre de collisions déjà subies par un message. En 1987, Aldous a établi le résultat suivant, dont la preuve est difficile :

Theorem 2.1. *Pour tout $\lambda > 0$, le protocole EBO est instable. De plus, en désignant par $N(t)$ le nombre de messages transmis avec succès sur [0,t], on a $N(t)/t \to 0$ lorsque $t \to \infty$, presque sûrement.*

2.4. Le protocole en arbre. En 1979, Capetanakis d'une part, Tsybakov et Mikhailov d'autre part, ont de façon indépendante proposé un protocole *intrinsèquement stable*, basé sur un partitionnement récursif des conflits, du genre « diviser pour régner ».

Lorsqu'un groupe de S stations entrent en collision à l'instant t, chacune tire à pile ou face et tous les tirages sont indépendants. Il se forme ainsi deux sous-groupes $S = S_1 + S_2$ selon un partage binômial : S_1 est le sous-groupe des stations ayant tiré pile (probabilité p) et qui ont la permission de retransmettre à $t+1$; le sous-groupe S_2 des stations ayant tiré face (probabilité $q = 1-p$) ne transmet que lorsque toutes les conflits du groupe S_1 auront été résolus. On a ainsi un arbre de résolution : partant d'une S-collision qui joue le rôle de racine, il y a création d'un sous-arbre gauche (S_1) et d'un sous-arbre droit (S_2). On exécute toujours le sous-arbre gauche en premier.

L'implémentation de l'algorithme est aisée à l'aide d'un mécanisme de *pile*. Chaque station s, lorsqu'elle a un message à transmettre, observe le canal (cf. section 2.1) et maintient un compteur local $H_s(t)$ qui évolue comme suit.

La permission d'émettre n'est donnée que si $H_s(t) = 0$. Dans ce cas, $H_s(t+1) = 1$ avec probabilité p [resp. $H_s(t+1) = 0$ avec probabilité q].

Par contre, dès que $H_s(t) \geq 1$, le message est bloqué en attente : sur une collision, $H_s(t+1) = H_s(t) + 1$, mais sur un succès ou un silence, $H_s(t+1) = H_s(t) - 1$.

Soit L_n la variable aléatoire représentant l'intervalle de temps nécessaire pour résoudre une collision de multiplicité n – on parlera de session. On s'intéresse au protocole dit à *arrivées bloquées* : les nouveaux messages sont retenus tant que la session en cours n'est pas résolue.

La v.a. L_n possède quelque propriétés fascinantes, qui s'obtiennent en travaillant sur la relation

$$L_n \stackrel{\mathcal{D}}{=} 1 + L_I + L_{n-I}, \quad n \geq 2, \quad L_0 = L_1 = 1, \tag{12}$$

avec $\mathbb{P}(I = k) = \binom{n}{k} p^k q^{n-k}$, l'égalité dans (12) étant à lire en distribution. Soit $\alpha_n \stackrel{\text{def}}{=} \mathbb{E}[L_n]$.

Theorem 2.2.

$$\alpha_n = \frac{-2n}{p \log p + q \log q} F(\log_p n) + o(n), \tag{13}$$

où la fonction $F(x)$ possède les propriétés suivantes :

- *$F(x) = 1$ si $\dfrac{\log p}{\log q}$ est irrationnel.*
- *Lorsque $\dfrac{\log p}{\log q}$ est rationnel, $F(x)$ est une série de Fourier dont les fluctuations autour de sa valeur moyenne 1 sont de faible amplitude (10^{-5} pour $p = q = 0.5$).*

La preuve du théorème 2.2 repose sur des méthodes fines d'analyse complexe, via la transformée de Mellin qui est un outil utilisé dans la théorie des séries de Dirichlet.

Proposition 2.4. *le protocole en arbre à arrivées bloquées est ergodique pour des taux d'arrivées satisfaisant*

$$0 < \lambda < \lambda_{inf} \stackrel{\text{def}}{=} -\frac{p \log p + q \log q}{2} - \epsilon, \tag{14}$$

où typiquement $0 < \epsilon < 10^{-5}$.
Lorsque $p = q = 1/2$, il vient $\lambda_{inf} \sim 0.34657$.

Indication de démonstration. Soit S_r le nombre de stations bloquées au début de la r-ième session et $B[t] \stackrel{\text{def}}{=} \sum_{i=0}^{t} A_i$, le nombre d'arrivées extérieures cumulées sur $[0, t]$. On voit aisément que la suite $\{S_r\}_{r \geq 0}$ forme une chaîne de Markov, dont l'évolution est régie par l'équation

$$S_{r+1} = B\big[L_{S_r}\big], \quad r \geq 0. \tag{15}$$

Ensuite on utilise le théorème 2.2 et le corollaire 1.1.

Suggestions pour le développement de la section 2

Soulignons d'emblée qu'il s'agit d'un *menu à la carte*. Vous pouvez choisir d'étudier *un ou plusieurs* protocoles, de façon plus ou moins fouillée, la philosophie générale étant de s'appuyer sur les outils théoriques présentés dans la section 1.

Tous les *Théorèmes* peuvent être admis. Par contre, on appréciera les démonstrations – parfois déjà esquissées dans le texte – des *Propositions*. Selon votre choix :

Section 2.2 Prouvez les propositions 2.1 et 2.2 en utilisant par exemple le corollaire 1.1 et le théorème 1.3. Vérifiez les affirmations qui suivent la proposition 2.2. En outre, si le système est stable, alors $p \neq 0$ et $p \neq \infty$.

Section 2.3 Après avoir vérifié que $\{Z_t\}$ est une CM à espace d'états dénombrable, la proposition 2.3 se montre à partir de l'équation (10) et du théorème 1.3.

Section 2.4 Il est souhaitable d'expliquer l'égalité distributionnelle (12). Si vous avez des idées probabilistes conduisant à un encadrement $K_1 < \alpha_n/n < K_2$, où K_1 et K_2 sont des constantes positives [cf. equation (13)], elles sont bienvenues. Vous pouvez achever la démonstration de la proposition 2.4.

Simulation La plupart des résultats théoriques évoqués se laissent valider par simulation, suivant une méthode classique. On conseille l'implémentation du protocole en ARBRE, qui est assez simple sous sa forme de *pile* avec les compteurs $H(t)$.

§3. Appendice sur la classification des chaînes de Markov à espace d'états dénombrable

Les propriétés des CMIA énoncées ci-dessous sont fondamentales et classiques. On les admettra sans autre forme de procès.

Notation. Un $\mathcal{F}_t$-temps d'arrêt est une variable aléatoire τ à valeurs dans $\mathbb{N} \cup \infty$, telle que l'événement $\{\tau = t\}$ soit $\mathcal{F}_t$-mesurable, $\forall t \geq 0$. Le temps de premier retour à l'état $\alpha \in \mathcal{A}$ est noté

$$\tau_\alpha \stackrel{\text{def}}{=} \inf\{t \geq 1; X_t = \alpha\}, \tag{16}$$

où $\tau_\alpha = \infty$ si $X_t \neq \alpha \, \forall t \geq 1$, et c'est de fait un temps d'arrêt. On écrira $\mathbb{P}_\alpha$ (resp. $\mathbb{E}_\alpha$) pour signifier la probabilité (resp. l'espérance) partant de l'état $\alpha, \forall \alpha \in \mathcal{A}$.

Definition 3.1. *Récurrence et transience.*
L'état α est dit récurrent *si $\mathbb{P}_\alpha(\tau_\alpha < \infty) = 1$, autrement il est dit* transient. *Un état récurrent est dit* récurrent positif *si $\mathbb{E}_\alpha[\tau_\alpha] < \infty$, sinon il est* récurrent nul.

Theorem 3.1.

1. *Tous les états de la chaîne sont de même nature, i.e. récurrents positifs, récurrents nuls, ou transients. Le mot* ergodique *sera utilisé comme synonyme de « récurrent positif ».*
2. *La CM est transiente si $\mathbb{P}_\beta(\tau_\alpha = \infty) > 0$, pour un couple (α, β).*

3. *En notant* $p_{\alpha\beta}^{(t)}$ *les éléments de* $\mathbf{P}^t$*, les limites*

$$v_\beta = \lim_{t\to\infty} p_{\alpha\beta}^{(t)}, \quad \forall \beta \in \mathcal{A}, \tag{17}$$

existent et sont indépendantes de l'état initial α*. Lorsque la CM n'est pas* ergodique*,* $v_\beta = 0, \forall \beta$*. Lorsque la CM est ergodique, on a*

$$v_\beta > 0, \quad \sum_\beta v_\beta = 1, \quad v_\beta = \sum_\alpha v_\alpha p_{\alpha\beta}, \tag{18}$$

i.e. le vecteur v *satisfait (1).*

Theorem 3.2. *Les conditions suivantes sont équivalentes :*

1. *la chaîne de Markov est ergodique ;*
2. *l'équation (1) admet une solution dans* l^1 *(l'espace des suites sommables), unique à un facteur multiplicatif près ;*
3. *Il existe une seule distribution stationnaire* $(\pi_\alpha, \alpha \in \mathcal{A})$*, i.e. une solution de (1) telle que* $\pi_\alpha \geq 0$ *et* $\sum_\alpha \pi_\alpha = 1$*. Alors,* $\forall \alpha \in \mathcal{A}$*,*

$$\pi_\alpha = \frac{1}{\mathbb{E}_\alpha[\tau_\alpha]} = \lim_{t\to\infty} p_{\gamma\alpha}^{(t)} > 0, \quad \forall \gamma \in \mathcal{A}. \tag{19}$$

Chapitre 23
Commentaire d'un spécialiste

§1. Plan

Comme je n'ai pas eu le temps de comprendre comment le protocole en arbre pouvait être implémenté uniquement à l'aide de compteurs locaux, je n'en parlerai pas. Je ferai seulement quelques remarques sur (10) à la fin, s'il reste du temps. J'ai choisi de me concentrer sur les protocoles ALOHA et ALOHA contrôlé. Avant d'exposer les résultats mathématiques que j'ai étudiés en détail, je vais préciser rapidement quelques points relatifs à ALOHA et ALOHA contrôlé qui ne sont pas explicités dans l'article. Ensuite, j'établirai les formules (7) et les propositions 2.1 et 2.2, que je discuterai. Dans chaque cas, j'exposerai des méthodes de simulation numérique et les commenterai. Enfin, s'il me reste du temps, je ferai quelques remarques sur (9) et (10).

§2. Formulation du problème étudié

Il s'agit d'abord de reformuler le problème posé de manière claire, précise, intuitive et sans paraphraser l'article, avant d'indiquer les questions qu'on va aborder.

Imaginons que des voitures arrivent sur une immense esplanade et qu'elles ne puissent passer que par une route à une seule voie. Une seule voiture peut donc passer à la fois. Si plusieurs voitures (au moins deux) tentent de passer au même moment, c'est-à-dire pendant la même brève fenêtre de temps entre t et $t+1$ (*slot*), alors il y a collision et aucune ne passe. Peu à peu, de nouvelles voitures arrivent sur l'esplanade. On suppose — et c'est là une hypothèse de modélisation — que le nombre de voitures qui arrivent pendant chaque fenêtre de temps est une variable aléatoire de loi définie par la suite $a_0, a_1, \ldots$ de réels ≥ 0 telle que $\sum_{i=0}^{\infty} a_i = 1$, avec $0 < a_0 < 1$ et $\sum_{i=2}^{\infty} a_i > 0$. Cette loi est la même pour toutes les fenêtres. De plus, on suppose que les variables aléatoires associées à diverses fenêtre de temps sont indépendantes. Enfin, on suppose que ces variables aléatoires positives sont d'espérance finie $\lambda = \sum_{i=0}^{\infty} i a_i$.

Comment faire passer ces voitures qui tendent à s'accumuler ? On est tenté de répondre qu'il suffit de ranger les voitures dans leur ordre d'arrivée et de les faire passer une à une. Pourquoi ne procède-t-on pas ainsi ? C'est qu'il faudrait pour cela une piste où aligner les voitures et un agent de police pour distribuer les numéros d'arrivée et faire respecter l'ordre de passage. Or, c'est contraire à l'architecture des réseaux, où les stations sont décentralisées. D'où la nécessité de définir un protocole de gestion des conflits qui ne fasse pas appel à une instance régulatrice centrale. C'est ce que font les protocoles ALOHA

et ETHERNET. Comme je n'ai pas eu le temps de comprendre comment le protocole en arbre pouvait être implémenté uniquement à l'aide de compteurs locaux installés dans chaque station, et sans avoir recours à un système central, je n'en parlerai pas.

A l'instant 0, $X_0 = 0$. Nous sommes maintenant à l'instant $t > 0$ et $X_t = i$ voitures attendent de passer. Si $i = 0$ et qu'aucune voiture n'arrive ($A_t = 0$), alors $X_{t+1} = 0$. Si $i = 0$ et qu'une voiture arrive ($A_t = 1$), alors il n'y a pas de collision, elle passe et $X_{t+1} = 0$. Enfin, si $i = 0$ et qu'au moins deux voitures arrivent ($A_t \geq 2$), alors il y a collision et $X_{t+1} = A_t$. Supposons maintenant que $i \geq 1$. La seule possibilité pour qu'une voiture passe est qu'aucune voiture n'arrive ($A_t = 0$) et que parmi celles qui attendent, une seule obtienne par tirage au hasard le droit de passer. Dans ce cas et dans ce cas uniquement, $X_{t+1} = i - 1$. A quelles conditions obtient-on $X_{t+1} = i$ (conservation du nombre de voitures en attente entre t et $t+1$) ? Cela arrive ou bien si aucune voiture n'arrive ($A_t = 0$) et si parmi celles qui attendent, soit aucune ne tire le droit de passer soit au moins deux tirent le droit de passer, d'où collision ; ou bien si une voiture arrive ($A_t = 1$) et si parmi celles qui attendent, aucune n'a tiré le droit de passer ; il s'agit donc de la réunion disjointe de deux événements. A quelle condition obtient-on $X_{t+1} = i + 1$? Cela ne peut arriver que si une voiture arrive, mais ne peut pas passer en raison d'une collision avec au moins une des voitures qui attendent. En effet, si $k \geq 2$ voitures arrivent simultanément alors il y a collision, donc aucune ne peut passer et par suite $X_{t+1} = i + k$.

On voudrait savoir si les protocoles ALOHA ou ALOHA contrôlé, qui sont représentés, sous les hypothèses modélisatrices faites, par une chaîne de Markov $(X_t)_{t\geq 0}$ dont les états sont l'ensemble des entiers naturels $\mathbf{N}$, a un comportement transient (c'est bien le terme "transient" qui est utilisé, et non pas le terme "transcient" — mot très *New Age* qui ne semble figurer dans aucun dictionnaire !) ou au contraire récurrent. En cas de récurrence, s'agit-il de récurrence positive ? La transience cela signifie ici que X_t tend p.s. vers l'infini quand $t \to \infty$, c'est-à-dire que le système finit par "exploser". La récurrence positive (on dira aussi ergodicité) signifie que X_t converge en loi vers une mesure de probabilité π sur $\mathbf{N}$ quand $t \to \infty$. Si $\pi(\{0\}) > 0$, X_t reprend la valeur 0 une infinité de fois p.s., et l'espérance de la durée de l'intervalle de temps entre deux retours en 0 est finie et égale à $1/\pi(\{0\})$. C'est-à-dire que le système peut être viable, à condition tout de même que $1/\pi(\{0\})$ prenne des valeurs acceptables, c'est-à-dire pas trop grandes.

§3. Résultats mathématiques

3.1. Probabilités de transition. Il faudrait d'abord démontrer que $(X_t)_{t\geq 0}$ est bien une chaîne de Markov. C'est facile et je ne développe pas ce point, mais suis prêt à l'expliquer si on me le demande. Notons p_{ij} la probabilité de transition de l'état i vers l'état j. Notons, comme dans l'article,

$b_0 = 1$, $b_i = (1-f(i))^i$ supposé > 0 pour $i \geq 1$, $c_0 = 0$ et $c_i = if(i)(1-f(i))^{i-1}$ supposé > 0 pour $i \geq 1$, $f(i)$ pouvant être la constante f (ALOHA) ou une fonction de i (ALOHA contrôlé). D'après les explications détaillées fournies ci-dessus et en raison de l'indépendance supposée des variables aléatoires A_t et des variables aléatoires représentant les tirages au hasard successifs nécessités par le protocole, on trouve que si $i = 0$,

$$p_{0\,j} = \begin{cases} a_0 + a_1 & \text{si } j = 0, \\ 0 & \text{si } j = 1, \\ a_j & \text{si } j \geq 2, \end{cases} \tag{1}$$

tandis que si $i \geq 1$,

$$p_{i\,j} = \begin{cases} a_0\, c_i & \text{si } j = i-1, \\ a_0\,(1-c_i) + a_1 b_i & \text{si } j = i, \\ a_1\,(1-b_i) & \text{si } j = i+1, \\ a_{j-i} & \text{si } j \geq i+2. \end{cases} \tag{2}$$

3.2. Irréductibilité et apériodicité. Dire que la chaîne de Markov $(X_t)_{t\geq 0}$ est irréductible, c'est dire que partant d'un état i à l'instant t (donc $X_t = i$), il est possible d'atteindre tout autre état j en un temps fini, plus ou moins long (donc $\exists k \geq 1$, $X_{t+k} = j$). Par exemple, ce serait impossible si on avait $a_\ell = 0$ pour tout $\ell \geq 2$: en effet, 0 serait alors un point absorbant. Dès que $a_\ell > 0$ au moins pour $\ell = 0, 1, 2$ et que $f(i) \in]0, 1[$ pour tout $i \geq 0$, les formules (1)–(2) montrent que partant d'un état quelconque $i \geq 1$ à l'instant t, la probabilité d'atteindre chacun des états voisins $i-1$, i et $i+1$ est > 0. Partant de $i = 0$, même chose pour les états 0 et 1. C'est une condition suffisante d'irréductibilité. C'est aussi une condition suffisante d'apériodicité. Je suis prêt à expliquer tout cela plus formellement si on me le demande, en faisant intervenir les probabilités de transition en k coups, $p_{i\,j}^{(k)}$, qui sont définies par récurrence pour $k \geq 1$ par $\forall i, j$, $p_{i\,j}^{(1)} = p_{i\,j}$, et par les formules

$$\forall i, j, \quad p_{i\,j}^{(k+1)} = \sum_{\ell=0}^{\infty} p_{i\,\ell}^{(k)}\, p_{\ell\,j}\,.$$

3.3. Démonstration de la Proposition 2.1. La Proposition 2.1 est une conséquence directe du Théorème 1.3, dans lequel on prend pour fonction de Lyapounov, notée f dans l'énoncé du théorème, la fonction $f(i) = i$ pour tout $i \in \mathbf{N}$, avec $d = -1$ dans (5). Il faut alors calculer l'espérance conditionnelle

$$E\{X_{t+1} - X_t | X_t = i\} = E\{X_{t+1} - i | X_t = i\} = \sum_{j=0}^{\infty} (j-i)\, p_{i\,j}$$

pour tout état $i \geq 1$ assez grand. Un calcul simple à partir de (1)–(2) montre que pour tout $i \geq 1$,

$$E\{X_{t+1} - X_t | X_t = i\} = -(a_0 c_i + a_1 b_i) + \sum_{\ell=0}^{\infty} \ell a_\ell = \lambda - d_i \,. \tag{3}$$

Or, puisque f est constante ici et que $0 < f < 1$, on a $\lim_{i\to\infty} b_i = \lim_{i\to\infty} c_i = 0$, donc $\lim_{i\to\infty} d_i = 0$. Par conséquent, il existe un entier $C > 0$ tel que $d_i < \lambda/2$ pour tout $i > C$, c'est-à-dire sur l'ensemble non vide U de l'énoncé du théorème. La condition (5) est alors vérifiée pour $\epsilon = \lambda/2$, et le résultat est établi.

3.4. Démonstration de la Proposition 2.2. On reprend le schéma de la démonstration de la Proposition 2.1 avec $f = f(i)$, en utilisant maintenant le Corollaire 1.1 en même temps que le Théorème 1.3. Cela permet de démontrer les deux premières assertions directement, après avoir rappelé la définition de lim inf et lim sup.

On suppose maintenant que $\exists p \in \mathbf{R}_+^*$, $\lim_{i\to\infty} i f(i) = p > 0$, avec $0 < f(i) < 1$ pour tout i. Tout revient à étudier le comportement asymptotique de la suite $(d_i)_{i\geq 1}$ dans ce cas. Or, les hypothèses impliquent que

$$\lim_{i\to\infty} (1 - f(i))^i = \lim_{i\to\infty} e^{i \ln(1 - f(i))} = \lim_{i\to\infty} e^{i \ln(1 - p/i + o(1/i))} = e^{-p} \,.$$

De même,

$$\lim_{i\to\infty} i f(i)(1 - f(i))^{i-1} = p\, e^{-p} \,.$$

Donc,

$$\lim_{i\to\infty} d_i = e^{-p}\,(p a_0 + a_1) \;,$$

d'où (8).

3.5. Discussion du choix de $f(i)$. J'aimerais ajouter ceci sur le choix de $f(i)$ pour tout $i \geq 2$. Supposons que $a_0 > a_1$. En raison de (3), il semble que l'on ait intérêt à maximiser $a_0 c_i + a_1 b_i$ pour chaque $i \geq 2$, afin de "rappeler en moyenne" X_{t+1} vers les petites valeurs aussi fortement que possible. Il s'agit donc de maximiser la fonction

$$x \in]0,\,1[\mapsto a_0 i x (1-x)^{i-1} + a_1 (1-x)^i = (1-x)^{i-1}\,(a_1 + (i a_0 - a_1)\,x) \,.$$

Pour $i \geq 1$, annuler la dérivée sur $]0,\,1[$ revient à résoudre une équation du second degré. Pour i assez grand, on trouve dans $]0,\,1[$ une racine $\approx a_0 \,/\, a_1 i$. Par exemple, dans le cas d'une loi de Poisson de paramètre $\lambda < 1$, cela conduirait à choisir $p = 1/\lambda > 1$. (Rappelons que le paramètre positif p peut être choisi dans $\mathbf{R}_+^*$.) Pour $i = 1$, il semble qu'il faudrait prendre $f(1)$ voisin de 1, c'est-à-dire privilégier fortement la réémission de l'unique paquet en attente, quel que soit le nombre d'arrivées.

Une autre démarche serait de maximiser $a_0 c_i$ pour tout $i \geq 2$, d'où $f(i) = 1/i$ et $p = 1$.

La question du choix optimal de $f(i)$ reste donc ouverte du point de vue théorique. En fait, des essais numériques effectués avec les programmes ci-dessous suggèrent de choisir plutôt p entre 0.5 et 1.

§4. Expérimentations numériques

4.1. Cas $f(i) = f$. Nous choisissons d'abord $f(i)$ constante, $f(i) = f \in]0, 1[$, avec $a_i = \exp(-\lambda)\, \lambda^i / i\,!$, $i \geq 0$, c'est-à-dire que les variables aléatoires A_t suivent une loi de Poisson de paramètre λ. (Cette hypothèse pourrait être expliquée en termes de processus de Poisson sous-jacent, mais je ne le ferai pas pour des raisons de temps.) Il s'agit de simuler M trajectoires mutuellement indépendantes $X_t^{(m)}$, $1 \leq m \leq M$, de la chaîne de Markov de probabilités de transition (1)–(2) entre les instants $t = 0$ et $t = T$. Les paramètres à entrer sont donc $\lambda > 0$, T et M (entiers > 0).

On peut commencer par simuler une trajectoire de X_t entre les instants $t = 0$ et $t = T$. En pratique, le plus simple est de choisir un entier $k_{\max}$ de taille moyenne au-delà duquel on considère que $a_\ell = 0$, c'est-à-dire qu'on pose $a_\ell = 0$ pour tout $\ell > k_{\max}$. Donc lorsque $X_t = i \geq 1$, X_{t+1} ne peut prendre que les valeurs $i-1, i, \ldots, i + k_{\max}$. Lorsque $X_t = 0$, X_{t+1} ne peut prendre que les valeurs $0, \ldots, k_{\max}$. Vu la forme des a_i et l'ordre de grandeur des valeurs de λ (en général ≤ 1), il semble qu'il pourrait suffire de choisir $k_{\max} \geq 10$. On part toujours de $X_0 = 0$. Pour simuler une variable aléatoire X_{t+1} lorsque $X_t = i \geq 1$, on utilise une variable aléatoire U uniforme sur $[0, 1]$ et indépendante de $X_t, X_{t-1}, \ldots, X_0$, puis

$$X_{t+1} = \begin{cases} i-1 & \text{si } 0 \leq U < a_0 c_i, \\ i & \text{si } a_0 c_i \leq U < a_0 + a_1 b_i, \\ i+1 & \text{si } a_0 + a_1 b_i \leq U < a_0 + a_1, \\ i + k \ \ (2 \leq k < k_{\max}) & \text{si } \sum_{\ell=0}^{k-1} a_\ell \leq U < \sum_{\ell=0}^{k} a_\ell, \\ i + k_{\max} & \text{si } \sum_{\ell=0}^{k_{\max}-1} a_\ell \leq U \leq 1. \end{cases} \tag{4}$$

Pour $X_t = 0$,

$$X_{t+1} = \begin{cases} 0 & \text{si } 0 \leq U < a_0 + a_1, \\ k \ (2 \leq k < k_{\max}) & \text{si } \sum_{\ell=0}^{k-1} a_\ell \leq U < \sum_{\ell=0}^{k} a_\ell, \\ k_{\max} & \text{si } \sum_{\ell=0}^{k_{\max}-1} a_\ell \leq U \leq 1. \end{cases} \tag{5}$$

Cela pose déjà des questions importantes ; a-t-on encore une chaîne de Markov irréductible et apériodique ? Les conditions d'ergodicité ou de transience sont-elles les mêmes ? Sans entrer dans les détails, on a encore irréductibilité et apériodicité dès que partant d'un état i à l'instant t, on peut atteindre les états voisins $i-1$, i et $i+1$ (ceci pour $i \geq 0$, les états 0 et 1 si $i = 0$). Les conditions restent les mêmes, à condition de recalculer λ, qui change très peu. Autre question : si la suite $(a_i)_{i \geq 0}$ tend lentement vers 0 quand $i \to \infty$ (typiquement, en $1/i^\gamma$ avec $\gamma > 2$) et que la CNS d'ergodicité (8) est vérifiée (elle ne porte que sur λ, $p > 0$, a_0 et a_1), ne risque-t-on pas de modifier qualitativement la forme de la mesure de probabilité invariante π en tronquant la suite $(a_k)_{k \geq 0}$ au-delà de l'indice $k_{\max}$? C'est très vraisemblable. Il faut donc de ce point de vue là rester extrêmement prudent sur l'allure approchée de π

que nous allons obtenir par simulations de Monte-Carlo. (Notons qu'il serait possible de dépasser au moins partiellement, dans les limites du calcul numérique par ordinateur, la limitation due à l'introduction de $k_{\max}$; cependant, le programme deviendrait alors plus délicat à écrire.)

Pour implémenter la simulation d'une trajectoire selon (4)–(5), il est conseillé de définir une matrice représentant les sommes cumulées (i indice de ligne, k indice de colonne) de la forme $\sum_{j=0}^{k} p_{i,j}$. Nous avons alors une limitation sur les indices i ($1 \leq i \leq i_{\max}$) et une limitation sur les indices k. Il faut choisir $i_{\max}$ grand, de l'ordre de quelques centaines. (Je n'insiste pas. Ici aussi, il serait possible de dépasser cette limitation sur i en réévaluant la matrice si nécessaire, ce qui permettrait en pratique d'atteindre de beaucoup plus grandes valeurs de i. Cependant, on resterait ici aussi limité par les capacités de la mémoire requise pour stocker la matrice et par l'existence d'un entier maximal représentable par la machine. Voir ci-dessous l'exemple du programme de simulation selon une loi de Poisson.) Si l'indice 0 n'est pas accepté (par exemple en MATLAB), il faut tout décaler de 1.

Une fois simulé un tableau de valeurs de $X_t^{(m)}$, où $0 \leq t \leq T$ indice de ligne et $1 \leq m \leq M$ indice de colonne, avec typiquement $T = 500$ (plus si possible) et $M = 100$:

EXPLOITATION

- On aimerait représenter quelques trajectoires particulières pendant un temps assez long, afin de mettre en évidence leur tendance à s'éloigner de l'état 0. (On pourrait même envisager une animation.) Ci-dessous, on commence par représenter l'histogramme relatif aux états visités par une seule trajectoire assez longue.
- On est ici dans le cas transient, ce que l'on veut illuster. On doit mettre en évidence la dégénerescence des lois π_t des $X_t^{(m)}$ quand t augmente à partir du tableau de simulations. Pour cela, il faudrait repreésenter pour des valeurs de t prises assez régulièrement entre 0 et T, l'histogramme des $X_t^{(m)}$, $1 \leq m \leq M$. On pourrait aussi tracer la courbe représentant l'approximation de Monte-Carlo $M^{-1} \sum_{m=1}^{M} X_t^{(m)}$ de l'espérance de X_t en fonction de t, $0 \leq t \leq T$. (On pourrait aussi, éventuellement, représenter l'écart-type correspondant.)

4.2. Cas $f(i) = 1 - \exp(-p/i)$, $i \geq 1$. Nous allons choisir $p > 0$, $f_0 = 1/2$ (par exemple) et $f(i) = 1 - \exp(-p/i)$ pour $i \geq 1$, qui appartient à $]0, 1[$ pour tout $i \geq 1$ et vérifie bien $\lim_{i\to\infty} if(i) = p$. Nous choisissons encore $a_i = \exp(-\lambda)\lambda^i/i\,!$, $i \geq 0$. Il s'agit à nouveau de simuler M trajectoires mutuellement indépendantes entre les instants 0 et T. Les paramètres à entrer sont cette fois $\lambda > 0$, $p > 0$, T et M (entiers > 0).

On doit distinguer les cas

$$\lambda < (p+\lambda)\, e^{-(p+\lambda)} \quad \text{(ergodicité)} \quad \text{et} \quad \lambda > (p+\lambda)\, e^{-(p+\lambda)} \quad \text{(transience)}\,.$$

Il faudrait représenter le graphe de la fonction $\lambda \mapsto (p+\lambda)e^{-(p+\lambda)}$ et calculer l'unique racine positive de l'équation $\lambda = (p+\lambda)e^{-(p+\lambda)}$.

On conjecture que pour $\lambda = (p+\lambda)\exp(-(p+\lambda))$, la chaîne de Markov est récurrente nulle.

Exploitation

- Dans le cas ergodique, on met en évidence l'allure de la loi commune π_t des $X_t^{(m)}$ pour certaines valeurs de t, en formant l'histogramme des $X_t^{(m)}$, $1 \leq m \leq M$. Il faut toujours se demander dans un tel cas quel est l'ordre de grandeur, en fonction de M, de la précision avec laquelle un tel histogramme approche π_t. (Le résultat est donné par le théorème central de la limite.) On pourrait aussi représenter la fonction de répartition empirique des $X_t^{(m)}$, $1 \leq m \leq M$, et calculer pour chacune des valeurs de t retenues la moyenne et l'écart-type des valeurs de $X_t^{(m)}$, $1 \leq m \leq M$.
- Dans le cas transient, on peut faire comme pour f constant.
- On aimerait explorer le cas $\lambda = (p+\lambda)\exp(-(p+\lambda))$.
- Dans les deux cas on aimerait représenter d'abord quelques trajectoires particulières.

4.3. Remarques sur π dans le cas ergodique.

- Question de la vitesse de la convergence de π_t vers π quand $t \to \infty$, en définissant de manière appropriée un écart entre π_t et π. Sur un espace d'états fini, on sait que cette vitesse est d'ordre géométrique : elle est liée aux valeurs propres différentes de un de la matrice des transitions. Ceci n'est plus nécessairement vrai sur un espace d'états infini.
- Relation avec les suites récurrentes linéaires d'ordre fini dans le cas ergodique, en raison de $k_{\max}$. Dès que i est assez grand, on a $f(i) = 1 - \exp(-p/i) \approx p/i$, de sorte qu'alors, approximativement, c_i et d_i ne dépendent plus de i, et $p_{i,j}$ ne dépend plus que de $j-i$. Si on écrit les équations traduisant que la mesure de probabilité π est invariante, on obtient approximativement pout tout j assez grand une équation de récurrence linéaire d'ordre fini (ordre fonction de $k_{\max}$) pour $\pi\{j\}$. Si toutes les racines du polynôme caractéristique sont distinctes, on en déduit la décroissance géométrique des $\pi\{j\}$ quand $j \to \infty$. On peut aussi étudier le cas où ces racines ne sont pas toutes distinctes. Cela pourrait faire éventuellement l'objet d'un développement pédagogique.
- Par contre, si a_i tend lentement vers 0 quand $i \to \infty$, par exemple si le moment d'ordre deux $\sum_{i=0}^{\infty} i^2 a_i$ est infini, alors dans le cas ergodique donné par (8) on peut conjecturer que $\pi(\{j\})$ tend lentement vers 0 quand $j \to \infty$, ce qui pourrait empêcher le système de fonctionner de manière satisfaisante. Une exploration numérique soigneuse et précise pourrait donner des indications sur ce point.

§5. Compléments

5.1. Remarques sur Ethernet. À partir des formules de P_t et S_t que l'on admet, on peut commencer par étudier, au moins dans le cas où la suite $(f_k)_{k\geq 1}$ tend vers une limite finie non nulle, le comportement asymptotique quand Z_t tend vers l'infini (ceci à t fixé) de D_t. On trouve que, à t fixé, P_t et le produit P_tS_t tendent vers 0. On déduit alors de (9) que la chaîne est transiente.

5.2. Remarques sur (10). Il s'agit de l'égalité des lois des membres de gauche et de droite. En conditionnant par $I = k$ pour toutes les valeurs possibles de k, en exprimant par linéarité l'espérance conditionnelle de L_n en fonction de celles de L_k et de L_{n-k} à partir de (10) puis en déconditionnant (c'est-à-dire en sommant en k avec les coefficients binomiaux), on trouve une relation linéaire entre α_n et les α_k, $k < n$. On peut en déduire par récurrence sur n qu'il existe des constantes positives K_1 et K_2 telles que $K_1 < \alpha_n/n < K_2$ pour tout $n \geq 1$.

Une dernière remarque (à ne faire que si on est sûr de soi !) : notons la parenté de la relation

$$S_{r+1} = B\left[L_{S_r}\right]$$

avec un processus de branchement comme ceux qu'on rencontre en dynamique des populations, par exemple. En admettant le Théorème 2.2, on peut en première approximation considérer que $L_n \approx \mathrm{cte}\, n$, auquel cas la relation récursive devient $S_{r+1} \approx B[\mathrm{cte}\, S_r]$. Si on avait cte $= 1$, on retrouverait alors le processus de Galton-Watson.

§6. Conclusion

On attend du (de la) candidat(e), dans la mesure du possible, une brève conclusion. Il (elle) peut récapituler très rapidement ce qu'il (elle) a établi et ce qu'il (elle) a mis en évidence au cours de ses expérimentations numériques et graphiques. Puis, l'idéal serait qu'il ou elle rappelle ou indique les questions qui se sont posées, vis-à-vis des procédures ou modèles étudiés, et au cours des démonstrations mathématiques ou des expérimentations, en indiquant éventuellement des pistes possibles pour des investigations futures. Ici, nous avons évoqué de tels points au fur et à mesure : question sur l'utilisation d'une loi de Poisson pour les A_t (est-ce raisonnable, pourquoi ?, que pourrait-il se passer avec d'autres lois ?) ; question d'un choix optimal de $f(i)$; questions liées à l'introduction de $k_{\max}$ et de $i_{\max}$ dans les programmes de simulation ; question de la décroissance de $\pi\{j\}$ vers 0 quand $j \to \infty$, dans le cas ergodique, en relation avec la décroissance de a_i vers 0 quand $i \to \infty$; question du cas $\lambda = (p+\lambda)\exp\left(-(p+\lambda)\right)$; question de la vitesse de la convergence de π_t vers π

quand $t \to \infty$; enfin, éventuellement, questions sur l'intérêt de ETHERNET par rapport à ALOHA contrôlé (je ne comprends toujours pas pourquoi choisir ETHERNET !), et même sur celui du protocole en arbre (il est indiqué dans l'article que ce dernier est stable par construction, ce qui n'est pas le cas de ALOHA contrôlé, mais lequel serait le plus efficace en situation ergodique ?).

§7. Programme Matlab pour expérimenter

Je remercie très vivement Didier Chauveau, qui m'a beaucoup aidé à mettre au point ce programme. Je remercie aussi Myriam Garrido. Pour faire comprendre le mécanisme, voici d'abord un programme qui simule n variables aléatoires indépendantes selon une loi de Poisson. Ce n'est qu'un exemple, car il existe une commande MATLAB qui permet de simuler directement selon une loi de Poisson. Ce programme recalcule la table des F_i si une réalisation dépasse la valeur $imax$ fixée au début. Les valeurs que peuvent prendre les variables aléatoires simulées ne sont donc ici limitées que par les capacités de la machine.

```
%% SIMULATION SELON LA LOI DE POISSON PAR INVERSION
%% avec stockage d'une table de repartition
clear
clf
n = input('n : ');        % nombre de v.a. simulees
L = input('lambda : ');   % parametre lambda
m = floor(7*L);           % borne max pour histogrammes
imax = m;                 % borne imax initiale
%% Construction de la table des F_i
%% F(i+1) = P(X <= i), i = 0, ..., imax
P = exp(-L);
F(1) = P;                 % proba de 0
for i=1:imax,             % calcul de la proba de i
        P = P*L/i;            % pourrait etre fait matriciellement
        F(i+1) = F(i) + P;
end;
FMAX = F(imax+1);               % proba[X <= imax]
PMAX = F(imax+1) - F(imax);     % proba[X  = imax]
%% ALGORITHME DE SIMULATION PAR INVERSION %%%
for i=1:n,
    U = rand(1)
            if (U <= FMAX),    % choix dans la table
                X = 0;
                while (U > F(X+1)),
                        X = X+1;
                end;
        else,                  % sinon recalcul de la table
                X = imax;          % en vue de l'etendre
                FR = FMAX;
                while (U > FR),
                        X = X+1;
                        P = P*L/X;
                        FR = FR + P;
                end;  % (while)
        end; % (if else)
        Y(i) = X;
end;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Trace de l'histogramme sur les entiers (bins entiers)
% Les bins sont les milieux des classes
% Construction de la table des frequences empiriques
bins = 0:m;
freq = hist(Y,bins);
freq = freq/n;
% Calcul de l'histogramme theorique
```

```
% et representation sur le meme histogramme
P = exp(-L);
PT(1) = P; % proba de 0
for i=1:m, % calcul de proba de i
        P = P*L/i;
        PT(i+1) = P; % proba de i
end;
z = [freq ; PT]';
bar(bins,z)
title('Poisson - Histogrammes empirique et theorique');
%%%%%%%%%%%%% Stat Descriptives
disp('     MIN      MAX      MEAN     STD');
disp([min(Y) max(Y) mean(Y) std(Y)])
```

Nous passons maintenant au programme pour ALOHA. N'hésitez pas à l'expérimenter et à l'améliorer !

```
%% SIMULATION DE MONTE-CARLO DU PROTOCOLE ALOHA CONTROLE
%%
%% Simulation de la chaine de Markov discrete par inversion avec
%% stockage d'une table de fonctions de repartition pour les valeurs
%% les plus probables.
%%
%% a(i) contient a_{i-1}, i=1, ..., imax+2
%% idem pour f, b et c
%%
%% F(i,j) contient la fonction de repartition du saut (i-1) -> (i-3+k)
%% donc si le test U < F(i,j) repond vrai pour j = k
%% alors poser X_{t+1} = X_t - 2 + k
%%
clear
% clf
n = input('n : ');     % n correspond a T (nombre de pas de la chaine)
m = input('nb chaines paralleles : '); % nombre de chaines paralleles
p = input('ALOHA CONTROLE, p > 0 : '); % exemple : p=0.5; lbd = 0.364;
lbd = input('lambda : ');              %
erg = (p+lbd)*exp(-p-lbd)
if (lbd < erg)
        disp('cas ergodique')
else
        disp('cas non ergodique');
end;
imax = 200;   % a regler
kmax = 20;    % a regler
%% NB : le pgm "se plante" tout seul si ces bornes sont depassees lors
%% de la simulation, d'ou un controle rustique du bon deroulement...
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
disp('Construction des parametres');
%% Construction des a(i), b(i), c(i) f(i) jusqu'à imax
%% Ceci devrait etre ecrit avec des operations matricielles
f(1) = 0.5; % f_0
for i=2:imax+2
        f(i) = 1 - exp(-p/(i-1)); % f_{i-1} stocke en f(i)
end;
a(1) = exp(-lbd);   % a_0
for i=2:imax+2
        a(i) = lbd*a(i-1)/(i-1);  % a_{i-1} stocke en a(i)
end;
b(1) = 1;   % b_0
for i=2:imax+2
        b(i) = (1-f(i))^(i-1);    % b_{i-1} stocke en b(i)
end;
c(1) = 0;   % c_0
for i=2:imax+2
        c(i) = (i-1)*f(i)*(1-f(i))^(i-2); % c_{i-1} dans c(i)
end;
%% [a' b' c']      % affichage pour controle
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% Construction de la table des F(i,j)
%% F(i+1, .) code la la fonction de repartition sachant Xt = i
%%
%% Cas Xt = 0
F(1,1) = 0;             % 0 -> -1 (homogene avec les autres)
F(1,2) = a(1) + a(2);   % 0 -> -1 ou 0
F(1,3) = a(1) + a(2);   % 0 -> -1 ou 0 ou 1
for j=4:kmax
        F(1,j) = F(1,j-1) + a(j);  % On devrait utiliser cumsum
end;
%%
%% Cas Xt = i > 0
for i=2:imax
        F(i,1) = a(1)*c(i+1);          % i -> i - 1
        F(i,2) = a(1) + a(2)*b(i+1);   % i -> i - 1 ou i
        F(i,3) = a(1) + a(2);          % i -> i - 1 ou i ou i + 1
        for j=4:kmax
                F(i,j) = F(i,j-1) + a(j-1); % Ici aussi et surtout : cumsum
    end;                            % car des boucles imbriquees sont
end;                                % a eviter absolument !
%%%%%%%%%%%%% SIMULATION CHAINES PARALLELES %%%%%%
disp('Simulation...');
X(1:m,1) = 0;           % m valeurs de depart X_0 = 0
```

```
U = rand(m,n);          % tout l'alea d'un coup
for k=1:m
        for t=1:n-1,
                z = 1;
                while ( U(k,t) > F(X(k,t) +1 , z) ),
                        z = z+1;
                end;
                X(k,t+1) = X(k,t) - 2 + z;
        end;
end;
%%%%%%%%%% STATISTIQUES SUR LA CHAINE 1 %%%%%%%%%%%
%% Moyenne et frequences de passage en chaque point
Xmean=mean(X(1,1:n));     % moyenne le long de la trajectoire
maxbar = floor(max(max(X))/2);         % modifiable, a regler
fprintf('mean X = %f, maxbar = %d\n', Xmean,maxbar);
for k=0:maxbar
        Fp(k+1) = sum(X(1,1:n) == k);     % passages en k
end;
% disp('statistiques chaine 1');
pe = Fp/n;      % matrice des frequences
maxY = max(pe); % quantite pour regler l'echelle commune des graphes
%% Moyennes empiriques successives chaine 1
XB = cumsum(X(1,1:n));
for i=1:n
        XB(i) = XB(i)/i;  % ceci pourrait etre effectue matriciellement
end;
%% TRACES GRAPHIQUES /  CHAINE 1
clf
subplot(3,1,1), stairs(X(1,1:n)); % trace des etats de la chaine 1
title('Trajectoire chaine 1');
b = 0:maxbar;
subplot(3,1,2), bar(b,pe), axis([-0.5 maxbar+0.5 0 maxY]);
title('loi empirique chaine 1');
subplot(3,1,3), plot(XB);
title('moyenne empirique chaine 1'); % trace des moyennes / chaine 1
pause
%%%%%%%%%%% STATISTIQUES PARALLELES %%%%%%%%%%%%%
nbt = 5;       % nombre de traces d'histogrammes paralleles
step = floor((n-1)/nbt);      % ecarts entre instants t
subplot(nbt+1,1,1), bar(b,pe);axis([-0.5 maxbar+0.5 0 0.8]);
title('loi empirique chaine 1');
for i = 0:nbt-1
        t = 20 + i*step;
        [freq, b] = hist(X(1:m,t),b);
        freq = freq/m;
```

```
        subplot(nbt+1,1,i+2), bar(b,freq);
                axis([-0.5 maxbar+0.5 0 0.8]);
        time = num2str(t);
        titre = strcat('loi empirique chaines paralleles à t = ',time);
        title(titre);
end;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Chapitre 24
Exploitation en TIPE

Pour l'étudiant de classes préparatoires, ce texte complète et généralise le texte "Modélisation du système ALOHA", présenté dans le même recueil. Il introduit des techniques mathématiques (fonctions de Lyapounov) destinées à garantir la stabilité ou l'instabilité de systèmes aléatoires[1] dans le cadre de protocoles de communication.

Un protocole probabiliste est bien adapté aux problèmes de communication et d'utilisation d'un canal commun (par exemple le partage des ressources offertes par un satellite). En effet, le problème principal est que si deux terminaux tentent d'émettre un message en même temps, les deux messages sont brouillés et aucun n'est transmis. Les deux terminaux doivent réémettre, mais s'ils réagissent de façon déterministe à l'information "le message est bloqué", ils vont encore le réémettre en même temps. Il faut donc qu'ils différencient leur politique de réémission, par exemple en introduisant un facteur aléatoire dans celle-ci[2].

La technique des fonctions de Lyapounov est généralement utilisée dans le cadre des équations différentielles. Elle consiste, lors de l'étude d'un système $y' = f(y)$, à introduire une fonction d'état E, telle que $E(y(t))$ soit décroissante (strictement si possible) le long d'une trajectoire. Cette propriété sert notamment à garantir le caractère borné (et/ou stable) des solutions de l'équation différentielle, et leur convergence éventuelle vers un minimum local de E.

Dans le cadre des chaînes de Markov, il est intéressant de voir qu'ici, on ne peut *garantir* la décroissance d'une fonction, du fait du caractère aléatoire de la suite, mais on exploite cependant une forme de décroissance "en moyenne", ce qui permet d'obtenir des informations sur le comportement "moyen" et asymptotique d'une chaîne de Markov.

La théorie des martingales qui sous-tend la démonstration des théorèmes de la partie 1 est difficile pour l'étudiant de classes préparatoires. Néanmoins les résultats exposés supportent une compréhension intuitive. Une approche accessible en TIPE consiste à étudier le principe d'un des trois protocoles présentés, puis à détailler l'application des théorèmes de Foster-Pakes pour obtenir un résultat concernant l'asymptotique du protocole envisagé. Une expérimentation informatique (qui est d'ailleurs réalisable à l'aide du logiciel de

[1] Pour une présentation plus complète des chaînes de Markov, le lecteur peut se reporter à la page 114

[2] Il est difficile, dans un cadre déterministe, de donner un ordre de priorité à chacun des terminaux. En effet, il faudrait qu'un terminal puisse signaler à l'autre qu'il est prioritaire, or les canaux de communications sont déjà encombrés !

calcul formel MAPLE) permettrait d'illustrer de façon adéquate les résultats, par exemple en obtenant des tracés de la suite X_t en fonction du temps t.

- Considérons le protocole ALOHA. Ce protocole est exposé plus en détail dans le texte "Modélisation du système ALOHA" présenté plus avant dans le recueil. La fonction $f(x) = \frac{1}{x+1}$ permet au processus X_t de satisfaire au théorème 1.3. En effet,

$$\begin{aligned} E(f(X_{t+1})|X_t = \alpha) &= \sum_{k=\alpha-1}^{+\infty} \tfrac{1}{k+1} p_{\alpha,k} \\ &\leq \tfrac{1}{\alpha} p_{\alpha,\alpha-1} + \tfrac{1}{\alpha+1} p_{\alpha,\alpha} + \tfrac{1}{\alpha+2}(1 - p_{\alpha,\alpha-1} - p_{\alpha,\alpha}) \end{aligned}$$

  Et ainsi $E(f(X_{t+1})|X_t = \alpha) \leq \frac{1}{\alpha+1}$ dès que

$$(1 - p_{\alpha,\alpha-1} - p_{\alpha,\alpha}) \geq \tfrac{\alpha+2}{\alpha} p_{\alpha,\alpha-1}.$$

  Cette inégalité supporte un passage à la limite. Comme on a

$$\begin{aligned} p_{\alpha,\alpha-1} &= a_0 \alpha f (1-f)^{\alpha-1} \\ p_{\alpha,\alpha} &= a_0(1 - \alpha f(1-f)^{\alpha-1}) + a_1(1-f)^{\alpha}, \end{aligned}$$

  le terme de gauche de l'inégalité tend vers $1 - a_0 > 0$ et celui de droite tend vers 0 lorsque α tend vers $+\infty$. L'inégalité est donc vérifiée pour $\alpha \geq M$. Il suffit donc de considérer $U = \{0, 1, \ldots, M\}$ pour que le théorème 1.3 puisse s'appliquer. Ce résultat est à rapprocher de l'étude des entiers n_0 et n_c caractéristiques du système, présentés dans le texte "ALOHA".
- Considérons le protocole ALOHA controlé. On pourrait ici vérifier que dans le cas $\lambda < e^{-p}(pa_0 + a_1)$ on se trouve bien dans les hypothèses du lemme de Pakes. Mais il est judicieux, dans le cadre d'un TIPE, d'expliquer plus qualitativement le rôle de la quantité $e^{-p}(pa_0 + a_1)$. Le terme $e^{-p}pa_0$ est approximativement (pour i grand) la probabilité qu'à la fois il n'y ait pas de nouveau message émis, et que l'un des terminaux bloqués arrive à émettre. Le terme $e^{-p}a_1$ est la probabilité qu'un nouveau message soit émis, et avec succès. En résumé, la valeur $e^{-p}(pa_0 + a_1)$ correspond au nombre moyen de messages transmis par unité de temps, lorsque le système est "chargé". Par ailleurs, la quantité λ correspond à la demande moyenne de l'ensemble des terminaux en nombre de messages à envoyer. L'inégalité $\lambda < e^{-p}(pa_0 + a_1)$ traduit le fait que, même chargé, le réseau a tendance à "désaturer".
- Le protocole Ethernet est une variante intéressante du système ALOHA. Les résultats mathématiques invoqués sont déjà plus difficiles à appréhender ; l'étudiant de classes préparatoires pourra plutôt chercher à effectuer une simulation informatique du processus. Il en est de même du dernier protocole observé, le protocole en arbre, qui offre d'ailleurs une application des principes des algorithmes "diviser pour régner" étudiés dans le cadre de l'option informatique.

Comme on vient de le voir, ce texte peut servir de support à plusieurs sujets de TIPE, avec trois approches possibles : une approche plus théorique, avec l'étude des méthodes mathématiques liées à l'asymptotique des chaînes de Markov ; une approche plus qualitative, dans laquelle le travail consistera à commenter de façon pertinente les concepts de stabilité ; et enfin une approche expérimentale, avec la simulation informatique des protocoles de communication. Malgré une lecture qui peut sembler difficile au premier abord, l'étudiant des classes préparatoires trouvera dans ce texte un bel exemple d'interaction entre des mathématiques théoriques, de l'expérimentation, et de la modélisation.

Huitième partie

Le coussin financier

Chapitre 25
Le coussin financier

1. Introduction

La méthode originelle du coussin, dite CPPI (*Constant Proportion Portfolio Insurance*), est une des méthodes fondamentales d'assurance du portefeuille financier : à tout instant, la valeur du portefeuille reste constamment au-dessus d'une valeur plancher de référence. Cette dernière représente une garantie, c'est à dire le montant minimal que l'investisseur désire détenir à chaque instant. Cette stratégie dynamique d'allocation d'actifs financiers est relativement simple et aisément applicable. Elle consiste à conserver une proportion constante d'exposition au risque : le montant total investi sur l'actif risqué doit être proportionnel au coussin, c'est à dire à la différence entre la valeur du portefeuille et le plancher garanti. Ce coefficient de proportion est en fait le paramètre clef de ce type de stratégie. C'est lui qui détermine la rentabilité et le risque associé du portefeuille considéré.
Nous examinons cette méthode dans ce qui suit. Dans un premier temps, l'analyse est menée en temps discret. Elle est étendue par la suite au cas du temps continu, notamment par des arguments de convergence.

2. Notions et résultats

2.1 Le marché financier

Considérons un marché financier composé (pour simplifier) de deux actifs, l'un considéré sans risque noté par B (placement à un taux fixe sans risque r), l'autre supposé "risqué" noté S (par exemple le cours d'une action ou d'un indice boursier). Nous supposons que $[0,1]$ représente l'intervalle de temps pour ce marché financier.
En temps discret, nous considérons pour chaque entier n non nul, la partition de l'intervalle de temps $[0,1[$ en une suite d'intervalles $[t_{n,k}, t_{n,k+1}[$ avec :

$$\forall k \in \{0, \dots, n\}, \quad t_{n,k} = \frac{k}{n}$$

Les dates $t_{n,k}$ représentent les instants de variations des prix des actifs. Pour chaque n, nous considérons l'espace de probabilité $(\Omega_n, \mathcal{F}_{0,n}, I\!P_n)$. Ω_n représente l'ensemble des aléas sur le prix de l'actif risqué S. $\mathcal{F}_{0,n}$ désigne la σ-algèbre initiale construite sur Ω_n et qui constitue "l'information initiale" disponible sur le marché. $I\!P_n$ est la probabilité construite sur Ω_n.
Nous notons également $Y_{n,k}$ le taux d'accroissement aux instants $t_{n,k}$ de l'actif risqué :

$$\begin{cases} B_{n,k} &= B_{n,k-1}(1+\frac{r}{n}) \\ S_{n,k} &= S_{n,k-1}(1+Y_{n,k}) \end{cases} \tag{1}$$

Pour des raisons de non-arbitrage sur le marché financier, il est toujours supposé que le rendement de l'actif risqué a non seulement une probabilité

non nulle d'être supérieur au rendement sans risque (profit supérieur en ce cas) mais aussi une probabilité non nulle d'être inférieur à celui-ci (possibilité de perdre sur l'actif risqué). Ceci se traduit par les inégalités suivantes :

$$I\!P_n[Y_{n,k} > \tfrac{r}{n}] > 0 \text{ et } I\!P_n[Y_{n,k} < \tfrac{r}{n}] > 0.$$

Considérons un cas particulier très important pour décrire le marché financier : *le modèle de Cox-Ross-Rubinstein*, encore appelé “modèle binomial".

Dans ce modèle, à chaque date $t_{n,k}$, le taux d'accroissement de l'actif risqué $Y_{n,k}$ ne peut prendre que deux valeurs. Plus précisément :

$$Y_{n,k} = \tfrac{\mu}{n} + vZ_{n,k} \tag{2}$$

où $(Z_{n,k})_{1\leq k\leq n}$ est une suite de variables aléatoires réelles indépendantes et de même loi telle que

$$I\!P_n[Z_{n,k} = \tfrac{-1}{\sqrt{n}}] = I\!P_n[Z_{n,k} = \tfrac{1}{\sqrt{n}}] = \tfrac{1}{2} \tag{3}$$

r, μ et v (appelée "volatilité") sont des constantes non négatives avec $\mu > r$.

Nous supposons dans la suite que les valeurs initiales (à $t = 0$) des actifs B et S ne dépendent pas de n. Nous les notons respectivement B_0 et S_0.

2.2 La méthode du coussin en temps discret

Notons $V_{n,k}$ la valeur du portefeuille à l'instant $t_{n,k}$. Désignons par V_0 la valeur du portefeuille initiale c'est à dire le montant total investi sur le marché financier à $t = 0$.
$P_{n,k}$ désigne la valeur du plancher à chaque instant $t_{n,k}$. Celui-ci est en général considéré comme un placement de référence sur l'actif sans risque B et donc rémunéré au taux r. P_0 est la valeur initiale du plancher qui est choisie en général de manière à récupérer un pourcentage fixé du capital investi V_0, à l'échéance $T = 1$ et cela quelle que soit l'évolution de l'actif risqué. P_0 est évidemment, par définition même, inférieure à V_0. Comme la valeur du plancher est supposée évoluer au taux sans risque r, nous obtenons :

$$P_{n,k} = P_{n,k-1}(1 + \tfrac{r}{n})$$

$C_{n,k}$ représente la valeur du coussin qui est égale à la différence $V_{n,k} - P_{n,k}$ entre la valeur du portefeuille $V_{n,k}$ et celle du plancher $P_{n,k}$. Nous avons donc

$$C_0 = V_0 - P_0$$

Notons également $e_{n,k}$ le montant total investi sur l'actif risqué S. $e_{n,k}$ est encore appelé l'*exposition* (au risque).

La méthode du coussin standard consiste à prendre $e_{n,k} = mC_{n,k}$ où m est une constante strictement positive (comprise en pratique entre 5 et 10) appelée usuellement le *multiple*.

On en déduit :
1) L'évolution de la valeur du portefeuille :

$$V_{n,k} = V_{n,k-1} + e_{n,k-1}Y_{n,k} + (V_{n,k-1} - e_{n,k-1})\tfrac{r}{n} \tag{4}$$

2) La dynamique de la valeur du coussin :

$$C_{n,k} = C_{n,k-1} + mC_{n,k-1}Y_{n,k} + (C_{n,k-1} + P_{n,k-1} - mC_{n,k-1})\tfrac{r}{n} - (P_{n,k} - P_{n,k-1}) \tag{5}$$

La valeur du coussin à chaque instant est donc donnée par :

$$C_{n,k} = C_0 \prod_{1\leq l\leq k} \left(1 + mY_{n,l} + (1-m)\tfrac{r}{n}\right) \tag{6}$$

En conséquence, la garantie est satisfaite (valeur du portefeuille au-dessus du plancher) dès lors que le multiple m est choisi de telle sorte que les sauts relatifs $Y_{n,k}$ de l'actif risqué vérifient

$$Y_{n,k} \geq (-1/m)[1 + (1-m)(r/n)] \tag{7}$$

Le modèle en temps continu sans saut

Nous l'obtenons ici par passage à la limite quand la durée de temps $\frac{1}{n}$ entre deux dates consécutives tend vers 0. Nous considérons pour cela en temps discret le modèle de Cox-Ross-Rubinstein :

Notation : $X_n \overset{\mathcal{L}}{\Longrightarrow} Y$ désigne la convergence en loi de la suite de variables aléatoires $(X_n)_n$ vers Y.
Sous les hypothèses précédentes, nous déduisons la convergence en loi des prix des actifs S et B :

Proposition 1
i) $B_{n,n} \overset{\mathcal{L}}{\Longrightarrow} B_1$ *avec* $B_1 = B_0 \exp(r)$ *(il s'agit ici de la convergence d'une suite réelle)*
ii) $S_{n,n} \overset{\mathcal{L}}{\Longrightarrow} S_1$ *avec* $S_1 = S_0 \exp((\mu - \frac{1}{2}v^2) + vW)$
où W *est est une variable aléatoire de loi gaussienne centrée réduite.*

Démonstration :
Elle repose sur la convergence des tableaux triangulaires (ici, suite double $(X_{n,k})_{n,k}$ avec $1 \leq k \leq n$) de variables aléatoires réelles indépendantes qui se déduit d'une version du *théorème de Lindeberg-Feller* suivant (admis ici) : (notation : $I\!E_{I\!P_n}$ désigne l'espérance sous la probabilité $I\!P_n$).

Si pour tout entier n fixé, $(X_{n,k})_{1\leq k\leq n}$ *est une suite de variables aléatoires indépendantes, de même loi et vérifiant :*

- $\forall \epsilon > 0, \lim_{n\to+\infty} \left[\sum_{1\leq k\leq n} I\!E_{I\!P_n}[X_{n,k}^2 I_{|X_{n,k}|\geq\epsilon}]\right] \longrightarrow 0$ *(notation :* I_A *est l'indicatrice de l'ensemble A).*

– $\sum_{1\leq k\leq n} I\!E_{I\!P_n}[X_{n,k}] \longrightarrow a$
– $\sum_{1\leq k\leq n} I\!E_{I\!P_n}[(X_{n,k} - E[X_{n,k}])^2] \longrightarrow s^2$
où a et s sont des constantes, s étant strictement positive.

Alors, quand n tend vers l'infini, $\left[\sum_{1\leq k\leq n} X_{n,k}\right] \overset{\mathcal{L}}{\Longrightarrow} W$ *où W est une variable aléatoire réelle de loi gaussienne $\mathcal{N}(a,s)$.*

De même, la valeur du plancher converge vers $P_t = P_0 e^{rt}$. Notons V_t la valeur du portefeuille à l'instant t et C_t la valeur du coussin qui désigne toujours la différence $V_t - P_t$ entre la valeur du portefeuille et celle du plancher P_t. Notons également e_t le montant total investi sur les actifs risqués.

La méthode du coussin en temps continu consiste à prendre $e_t = mC_t$ (m désignant toujours le multiple).

On obtient par convergence :

$$C_1 = C_0 e^{mvW + [r + m(\mu - r) - \frac{m^2 v^2}{2}]} \tag{8}$$

Le coussin suit donc une loi lognormale dont la densité, notée g, est donnée par

$$g(x) = \frac{I_{x\geq 0}}{x\sqrt{2\pi v^2 m^2}} e^{-\frac{1}{2v^2m^2}\left(ln\left[\frac{x}{C_0}\right] - (r + m(\mu - r) - \frac{m^2v^2}{2})\right)^2} \tag{9}$$

Le modèle en temps continu avec sauts

Considérons maintenant le cas d'une évolution du prix de l'actif risqué S en temps continu mais avec des sauts possibles à des instants aléatoires. Ces dates forment une suite croissante d'instants aléatoires T_n. Notons X_n les sauts relatifs de S à ces instants T_n. On a donc :

$$X_n = \frac{S_{T_n} - S_{T_n^-}}{S_{T_n^-}} \tag{10}$$

où $S_{T_n^-}$ désigne la limite à gauche $\lim\limits_{t<T_n, t\to T_n} S_t$.

Entre deux instants de sauts, le prix suit une trajectoire qui est une fonction continue du temps t, du type de celle du modèle sans saut. On peut alors montrer (résultat admis) que la valeur du coussin à la date t est donnée par :

$$C_t = C_0 \exp(rt + m(\mu - r)t - 1/2m^2v^2t + mv\sqrt{t}W) \times \prod_{0\leq T_n\leq t} (1 + mX_n) \tag{11}$$

Modèle avec sauts de type Poisson :

Supposons par exemple que les variables X_n soient indépendantes et identiquement distribuées de loi $K(dx)$ vérifiant $K(\{0\}) = 0$, d'espérance commune égale à b et de moment d'ordre 2 égal à c. ($N_t = \sum_{0\leq T_n\leq t} X_n$ est alors un processus de Poisson composé).

La suite des durées $(T_{n+1} - T_n)_n$ est une suite de variables également indépendantes et identiquement distribuées de loi exponentielle dont le paramètre

est noté λ. La variable aléatoire $\mathcal{N}_t$ qui compte le nombre de sauts entre 0 et t suit une loi de Poisson de paramètre λt.

La densité g_s du coussin C_1 en présence de sauts est alors déterminée de la manière suivante : (g est la densité du coussin quand il n'y a pas de saut)

$$g_s(x) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{(\lambda)^n}{n!} \int_{I\!R^n} \frac{1}{\prod_{i \leq n}(1+my_i)} g\left(x \frac{1}{\prod_{i \leq n}(1+my_i)}\right) \otimes_i K(dy_i) \tag{12}$$

Celle du portefeuille est immédiatement déduite :

$$f_{V_1}(x) = g_s(x - P_0 e^r) \tag{13}$$

Quelques propriétés et performances de ce modèle

Rentabilité/Risque : La rentabilité est définie comme l'espérance du rapport $\frac{V_1}{V_0}$ et le risque est défini comme la variance de ce rapport. On peut noter qu' ils sont des fonctions croissantes du multiple m et décroissantes par rapport au plancher initial P_0. Cependant, il n'est pas possible d'optimiser en m en utilisant le critère standard de Markowitz, c'est à dire la minimisation de la variance à un niveau d'espérance fixée L. En fait, si l'on fixe L alors le plancher initial P_0 est une fonction du multiple m :

$$P_0(m) = \frac{L - V_0 e^{r + m(\mu + b\lambda - r)}}{e^r[1 - e^{m(\mu + b\lambda - r)}]} \tag{14}$$

ce qui implique :

$$Var[V_1] = (V_0 e^r - L)^2 \frac{[e^{m^2(v^2 + \lambda c)} - 1]}{[1 - e^{-m(\mu + b\lambda - r)}]^2} \tag{15}$$

En fait, l'espérance de gain n'est pas l'objectif premier : le niveau de garantie c'est à dire le plancher est fondamental. Ce choix étant fait, le multiple permet d' ajuster la part de profit que l'on anticipe en cas de hausse des cours, l'espérance du rendement du portefeuille étant une fonction croissante de m.

Gestion par quantile : Pour atténuer cette contrainte, il est possible de définir une condition de garantie plus faible, basée sur une approche de type Value-at-Risk : on peut envisager par exemple de définir un plancher auxiliaire au-dessus du plancher contractuel et de se garantir que la valeur du portefeuille reste au-dessus de ce plancher à un certain niveau de confiance fixé. Examinons le cas du modéle avec sauts de type Poisson.

Ceci conduit à une condition du type :

$$I\!P[\forall t \leq 1, C_t \geq 0] \geq 1 - \epsilon \tag{16}$$

où ϵ est relativement faible. Considérons $K(dx)$ la loi des sauts et F sa restriction au support de la loi, qu'on supposera ici strictement croissante. Soit alors $F^{(-1)}$ sa réciproque. La condition

$$I\!P[\forall t \leq 1, C_t \geq 0] \geq 1 - \epsilon$$

est alors équivalente à la condition suivante sur m :

$$m \leq \frac{-1}{F^{(-1)}\left(\frac{1}{\lambda} ln(\frac{1}{1-\epsilon})\right)} \tag{17}$$

Cette condition qui fixe une nouvelle borne sur le multiple permet de le calibrer sur la distribution des amplitudes de sauts. On peut observer également que ce majorant est une fonction décroissante de l'intensité du nombre des sauts λ. Ainsi, conformément à l'intuition, si la fréquence des sauts augmente, le multiple doit être réduit en conséquence.

Développements possibles suggérés pour le commentaire

1) Justifier les équations d'évolution (4) et (5) des valeurs du portefeuille et du coussin pour le modèle en temps discret.
2) Retrouver la formule (6) donnant la valeur du coussin.
3) Calculer l'espérance du rendement du portefeuille, à savoir

$$I\!E\left[\frac{V_{n,n}}{V_0}\right]$$

pour le modèle binomial et examiner la en tant que fonction du multiple m. Rapprocher ce résultat de la contrainte de garantie (7)et commenter ces deux résultats.
4) Compléter la démonstration de la proposition 1 et montrer la convergence en loi de $C_{n,n}$ vers la valeur C_1 donnée dans (9).
5) Illustrer cette convergence par simulation en prenant par exemple les valeurs suivantes pour les paramètres : $\mu = 0.1$, $r = 0.05$, $v = 0.2$, $S_0 = 1$, $V_0 = 100$, $P_0 = 90$, $m = 5$.
6) Vérifier les expressions (12) et (13) des densités g_s et f_{V_1}.
7) Simuler des valeurs du coussin avec sauts pour les valeurs des paramètres suivantes : $\mu = 0.1$, $r = 0.05$, $v = 0.2$, $S_0 = 1$, $V_0 = 100$, $P_0 = 90$,$\lambda = 2$, $m = 5$.
8) Examiner le problème de Markowitz dans le modèle continu avec sauts.
9) Retrouver la condition sur le multiple dans la gestion par quantile.
10) Illustrer numériquement cette condition en fonction de λ pour $\epsilon = 0.05$ et pour une loi $K(dx)$ de type gaussienne centrée en $-0,125$ mais tronquée sur $[-0,2; -0,05]$.

Références

- Black F. et A.R. Perold,(1992), Theory of constant proportion portfolio insurance, *J. Econ.Dynamics Control*, 16, 403-426.
- Jacod J. et A. N. Shiryaev,(1987), Limit theorems for stochastic processes, *Springer-Verlag*.
- Poncet P. et R. Portait,(1997), Assurance de Portefeuille, dans *Encyclopédie des Marchés Financiers, Economica*, 140-141.

Chapitre 26
Commentaire de l'auteur

Les questions (1), (2) et (3) ont essentiellement pour but de se familiariser avec le modèle et de s'assurer de la bonne compréhension de celui-ci. La question (4) est une mise en oeuvre relativement immédiate du théorème de Lindeberg-Feller sur la convergence en loi des tableaux triangulaires, du type "limite centrale".
Les questions (5), (7) et (10) suggèrent des développements informatiques. Les valeurs des paramètres correspondent à des données classiques, calibrées sur une année. La question (10) permet de mettre en évidence le rôle de la fréquence des chutes de l'actif risqué. On peut vérifier que, conformément à l'intuition, plus le paramètre λ (qui correspond ici à l'espérance du nombre de sauts sur la période $[0,1]$) est élevé, plus la borne supérieure sur le multiple m est faible. A la limite, quand λ tend vers $+\infty$, on retrouve approximativement la borne inférieure du support de la loi K (si n est grand, $\frac{r}{n}$ est négligeable, et cette borne est donc égale ici à $\frac{1}{-(-0,2)} = 5$, qui permet de se prémunir contre une chute brutale de -20%.

Les questions (6), (8) et (9) doivent permettre au candidat de développer son savoir faire mathématique (calculs de densité, conditionnement de variables aléatoires...) tout en exposant les propriétés importantes du modèle.

Question 1 :
L'équation donnant la valeur des $V_{n,k}$ est obtenue de la manière suivante : à l'instant $t_{n,k-1}$, l'investisseur détient le capital $V_{n,k-1}$ sur le marché financier. Il répartit celui-ci sur les deux actifs disponibles sur le marché sous la forme d'un montant $e_{n,k-1}$ investi sur l'actif risqué S et donc le restant $V_{n,k-1} - e_{n,k-1}$ placé sur l'actif sans risque B. Quand les prix varient à l'instant $t_{n,k}$, ces deux montants ont des rendements respectifs $(1+Y_{n,k})$ et $(1+\frac{r}{n})$, c'est à dire ceux des deux actifs S et B. La nouvelle valeur $V_{n,k}$ du portefeuille à l'instant $t_{n,k}$ étant la somme de ces deux nouveaux montants, elle est nécessairement solution de l'équation (4).
La relation (5), donnant la valeur du coussin, s'en déduit en utilisant d'une part la relation

$$V_{n,l} = C_{n,l} + P_{n,l}, \text{ pour } l = k-1 \text{ et } l = k$$

et, d'autre part, le fait que l'exposition à l'instant $t_{n,k-1}$ vérifie $e_{n,k-1} = mC_{n,k-1}$.

Question 2 :
Pour obtenir la formule (6) donnant la valeur du coussin, il suffit de simplifier l'équation (5) en remarquant que, d'après l'hypothèse faite sur l'évolution du plancher (taux sans risque r), on a :

$$P_{n,k} - P_{n,k-1} = P_{n,k-1}\frac{r}{n}$$

La relation (5) s'écrit alors

$$C_{n,k} = C_{n,k-1}\left(1 + mY_{n,k} + (1-m)\frac{r}{n}\right)$$

qui, par itération, conduit immédiatement à la relation (6).

Question 3 :

L'espérance du rendement du portefeuille $I\!E[\frac{V_{n,n}}{V_0}]$ est donnée par :

$$I\!E[\frac{V_{n,n}}{V_0}] = \frac{C_{n,n}+P_{n,n}}{V_0}$$

Or, d'une part, comme nous avons la relation $P_{n,k} = P_{n,k-1}\left(1+\frac{r}{n}\right)$, nous obtenons : $P_{n,n} = P_0\left(1+\frac{r}{n}\right)^n$. D'autre part, par indépendance des variables $Y_{n,k}$ dans le modèle binomial, nous avons :

$$I\!E[C_{n,k}] = C_0 \prod_{1\leq k\leq n} I\!E\left[\left(1 + mY_{n,k} + (1-m)\frac{r}{n}\right)\right]$$

En conclusion, pour le modèle binomial, l'espérance du rendement du portefeuille est donc

$$I\!E[\frac{V_{n,n}}{V_0}] = \frac{C_0\left(1+\frac{r}{n}+m(\frac{\mu}{n}-\frac{r}{n})\right)^n + P_0\left(1+\frac{r}{n}\right)^n}{V_0}$$

C'est une fonction croissante du multiple m puisque que les constantes μ et r vérifient : $\mu > r$. On retrouve donc le fait que, conformément à l'intuition, plus on investit sur l'actif risqué S c'est à dire plus m est grand, plus la rentabilité (en espérance) est grande. Notons que la condition de garantie (7) impose une borne supérieure sur m : en effet, d'après la condition de non-arbitrage sur le marché financier, la borne inférieure des supports des $Y_{n,k}$ est strictement inférieure à $\frac{r}{n}$. Nous en déduisons que la relation (7) est équivalente à :

$$m \leq \frac{(1+\frac{r}{n})}{(\frac{r}{n}-\inf Y_{n,k})}$$

Ceci limite de fait les grandes valeurs de m et amène bien souvent l'investisseur à prendre des valeurs de ce multiple inférieures à 10, correspondant approximativement à se prémunir contre une chute subite de l'actif risqué de l'ordre de moins dix pour cent.

Question 4 :

L'assertion (i) est basée sur la propriété classique de convergence de $(1+\frac{r}{n})^n$ vers e^r, quand n tend vers $+\infty$. Celle-ci se démontre en passant par le logarithme. Utilisant ce même argument pour étudier la convergence de la valeur du coussin $C_{n,n}$ donnée dans (6), la démonstration de l'assertion (ii) de la proposition 1 est une application du théorème de Lindeberg-Feller. Il suffit

pour cela d'utiliser un développement à l'ordre 2 du logarithme au voisinage de 1 (voir plus loin la preuve alternative du théorème de convergence).

Question 5 :

Illustrons cette convergence par simulation (programme MAPLE).

Les paramètres du marché financier

```
mu :=0.1;
r :=0.05;
v :=0.2;
S0 :=1.0;
V0 :=100.0;
P0 :=90.0;
m :=5.0;
lambda :=2.0;
```

Chargement des packages statistiques

```
with(stats) :
with(stats[statplots]) :
```

(simulation d'une valeur du coussin $(V_0 - P_0)\prod_{1\leq l\leq k}\left(1+mY_{n,k}+(1-m)\frac{r}{n}\right)$)

```
Co :=array(1..n);
for i from 1 to n do
Co[i] :=1+m*(mu/n+(v/sqrt(n))*(2*stats[random, binomiald[1,0.5]](1,'default','auto')-1))
+(1-m)*r/n :
od :
(V0-P0)*product(Co[ii],ii=1..n) :
evalf(%) :
i :='i';
n :='n';
n;
```

Histogramme des valeurs du coussin (la validité du résultat dépend évidemment du nombre de tirages, celui-ci étant bien sûr limité par le temps disponible pour l'exécution du programme). Dans ce qui suit, n désigne le nombre de tirages pour simuler une valeur du coussin et a représente le nombre des valeurs du coussin générées.

```
a :=100;n :=1000;
Coussin :=array(1..a);
for j from 1 to a do
Co :=array(1..n);
for i from 1 to n do
Co[i] :=1+m*(mu/n+(v/sqrt(n))*(2*stats[random,binomiald[1,0.5]](1,'default','auto')-1))+(1-m)*r/n :
od :
Coussin[j] :=evalf((V0-P0)*product(Co[ii],ii=1..n));
evalf(%) :
i :='i';ii :='ii';
od :
histogram(convert(Coussin,'list'));
```

Question 6 :
La densité du coussin en présence de sauts de type "Poisson" peut être déterminée de la manière suivante : on commence par évaluer, pour x réel, $I\!P[C_1 \leq x]$. Pour cela, on conditionne par le nombre n de sauts sur la période de temps $[0,1]$ et par les valeurs de ces sauts "$X_i = y_i$". On obtient alors la relation suivante, par indépendance de W, des X_n et des T_l :

$$I\!P[C_1 \leq x] = \sum_{n=0}^{\infty} e^{-\lambda} \frac{(\lambda)^n}{n!} \int_{I\!R^n} I\!P[C_0 e^{mvW+[r+m(\mu-r)-\frac{m^2v^2}{2}]} \prod_{i\leq n}(1+my_i) \leq x] \otimes K(dy_i)$$

Il suffit alors d'utiliser la densité g du coussin sans saut puis de dériver par rapport à x.
La relation $V_1 = C_1 + P_1$ où $P_1 = P_0 e^r$ est une constante montre immédiatement que

$$f_{V_1}(x) = g_s(x - P_0 e^r)$$

Question 7 :

(on garde les mêmes valeurs des paramètres)

```
mu :=0.1;
r :=0.05;
v :=0.2;
S0 :=1.0;
V0 :=100.0;
P0 :=90.0;
m :=5.0;
lambda :=2.0;
t :=1;
```

(simulation d'une valeur du coussin avec sauts)

```
with(stats) :
with(stats[statplots]) :
stats[random, poisson[lambda*t]](1,'default','auto');

nombre :=stats[random, poisson[lambda*t]](1,'default','auto');print(nombre);
if nombre=0 then x :=1 else for i from 1 to nombre do
sauts[i] := 1+m*stats[random, uniform[-0.1,-0.05]](1,'default','auto') :
od :
x :=product(sauts[ii],ii=1..nombre);
fi;
C :=x*(V0-P0)*
*exp(r*t+m*(mu-r)*t-0.5*m^2*v^2*t+m*v*sqrt(t)*stats[random, normald](1,'default','auto'));
evalf(%) :
i :='i';
n :='n';
```

Question 8 :
Le problème de Markowitz, qui est le fondement de la théorie du choix de portefeuille, consiste à faire un choix entre la rentabilité (espérance mathématique du rendement du portfeuille) et le risque (assimilé en finance à la variance de ce rendement). En général, on fixe la rentabilité à un certain niveau L donné (par exemple +10%) et on détermine ensuite tous les portefeuilles qui ont cette rentabilité. Parmi ceux-ci, on détermine celui qui a le risque le moins élevé (il existe toujours sous des conditions standards). Ici, la méthode du coussin restreint les stratégies possibles et, par un simple calcul, on retrouve les assertions (14) et (15) qui montrent qu'à un niveau de rentabilité fixée L, le plancher P_0 et le multiple m sont liés, ce qui entraîne également qu'il n'y ait plus qu'une seule valeur possible pour le risque (la variance) si on fixe m (ou de manière équivalente ici P_0).

Question 9 :
La gestion par quantile s'est développée assez récemment. Elle consiste à s'assurer, à un certain niveau de probabilité donnnée (en général, au moins 95%) que la valeur du portefeuille est au-dessus d'une certaine valeur de référence. Ici, elle est appliquée à la positivité du coussin. Pour trouver l'inégalité (17), on utilise l'expression (11) qui donne la valeur du coussin en particulier à l'instant $t = 1$. On en déduit que cette condition s'écrit encore :

$$I\!P[\forall n \text{ tel que } T_n \leq 1, 1 + mX_n \geq 0] \geq 1 - \epsilon$$

Conditionnant par le nombre de sauts durant la période $[0, 1]$ et utilisant l'indépendance des variables X_n et T_l, on obtient la condition équivalente suivante :

$$\sum_{n=0}^{\infty} e^{-\lambda} \frac{(\lambda)^n}{n!} I\!P[1 + mX_1 \geq 0, ..., 1 + mX_n \geq 0] \geq 1 - \epsilon$$

ce qui permet encore d'écrire :

$$\sum_{n=0}^{\infty} e^{-\lambda} \frac{(\lambda)^n}{n!} (1 - F)^n (-\tfrac{1}{m}) \geq 1 - \epsilon$$

Le calcul de la somme de la série du terme de gauche et une transformation logarithmique donne le résultat.

Question 10 :

Illustrons numériquement cette condition en fonction de λ pour $\epsilon = 0.05$ et pour une loi $K(dx)$ de type gaussienne tronquée sur $[-0,2; -0,05]$ centrée sur cet intervalle. On peut alors observer qu'il s'agit d'une fonction décroissante, dont la limite pour λ tendant vers l'infini, est la borne inférieure du support de la loi des sauts. Ceci correspond bien à l'intuition : quand la fréquence des sauts augmente, il faut se prémunir contre un risque plus important donc réduire le multiple ; quand cette fréquence tend vers l'infini, tout se passe

comme si au moins un saut a une probabilité égale à 1 de se produire sur tout intervalle de temps $[0, t]$, $t > 0$ d'où la nécessité de majorer le multiple par la borne inférieure du support des sauts (cette borne pour les sauts relatifs de l'actif risqué S vaut ici $1/(0{,}2)$ soit 5).

```
  with(stats) :
with(statplots) :
a :=-0.2;
b :=-0.05;
sigma :=0.1;
with(plots) :
statevalf[cdf,normald](0.0);
F :=x->(statevalf[cdf,normald]((x-(a+b)/2)/sigma)-statevalf[cdf,normald]((a-(a+b)/2)/sigma))
/(statevalf[cdf,normald]((b-(a+b)/2)/sigma)-statevalf[cdf,normald]((a-(a+b)/2)/sigma));
epsilon :=0.05;
(-1)/fsolve(F(x)=(1/100)*log(1/(1-epsilon)),x);
g :=lambda->(-1)/fsolve(F(x)=(1/lambda)*log(1/(1-epsilon)),x);
n :=20;
L :=seq(g(i),i=1..n);
listplot([L]);
```

Chapitre 27
Commentaire d'un spécialiste : autour du théorème de Lindeberg

§1. Théorème de Lindeberg

Nous énonçons une variante très élémentaire de cet énoncé.

Théorème de Lindeberg. *Soient $X_1, X_2, \dots, X_n$ des variables aléatoires indépendantes centrées et possédant des moments d'ordre 3. On considère n variables aléatoires gaussiennes centrées, indépendantes et vérifiant*

$$\mathbb{E}Y_j = \mathbb{E}X_j, \quad \mathbb{E}Y_j^2 = \mathbb{E}X_j^2 \quad \textit{pour } j = 1, \dots, n.$$

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe C^3 telle que $|f^{(p)}(x)| \leq c_p$ pour chaque $x \in R$ et pour $p = 0, 1,$ ou 3

On a alors :

$$|\mathbb{E}f(X_1 + \cdots + X_n) - f(Y_1 + \cdots + Y_n)| \leq \tfrac{c_3}{3} \sum_{j=1}^{n} \mathbb{E}|X_j|^3.$$

Preuve. Si l'espace de probabilité sous-jacent est suffisamment riche, quitte à redéfinir le un vecteur aléatoire gaussien de même loi que $(Y_1, \dots, Y_n)$, on suppose le vecteur aléatoire $(Y_1, \dots, Y_n)$ indépendant du vecteur aléatoire $(X_1, \dots, X_n)$. Posons $Z_j = X_1 + \cdots + X_{j-1} + Y_{j+1} + \cdots + Y_n$ pour $j = 2, \dots, n-1$ et $Z_1 = Y_2 + \cdots + Y_n$, alors

$$\mathbb{E}\left(f(X_1 + \cdots + X_n) - f(Y_1 + \ldots + Y_n)\right) = \sum_{j=1}^{n-1} \mathbb{E}\left(f(Z_j + X_j) - f(Z_j + Y_j)\right).$$

Notons que les triplets (X_j, Y_j, Z_j) sont indépendants. Pour simplifier les preuves nous omettons aussi l'indice j, alors la formule de Taylor montre que

$$\left|f(Z + X) - f(Z) - Xf'(Z) - \tfrac{1}{2}X^2 f''(Z)\right| \leq \tfrac{c_3}{6}|X|^3$$

Par suite $|\mathbb{E}\left(f(Z_j + X_j) - f(Z_j + Y_j)\right)| \leq \frac{c_3}{6}(\mathbb{E}|X|^3 + \mathbb{E}|Y|^3)$. A présent l'inégalité de Hölder implique que $\mathbb{E}|Y|^3 = \mathbb{E}|N|^3 \left(\mathbb{E}X^2\right)^{3/2} < \mathbb{E}|X|^3$ si N désigne une gaussienne standard car

$$\mathbb{E}|N|^3 = \tfrac{2}{\sqrt{2\pi}} \int_0^\infty u^3 e^{-\frac{u^2}{2}} du = \tfrac{1}{\sqrt{2\pi}} \int_0^\infty v e^{-v/2} dv < 1.$$

Donc $|\mathbb{E}\left(f(Z_j + X_j) - f(Z_j + Y_j)\right)| \leq \frac{c_3}{3}\mathbb{E}|X|^3$ et le résultat suit.

L'énoncé précédent implique le théorème de limite centrale pour des tableaux de variables aléatoires $X_{j,n}$, pour $j = 1, \ldots, n$ et $n = 1, 2, \ldots$ possédant un moment d'ordre 3.

On renormalise un tel tableau en notant $X_j = X_{j,n}/\sqrt{\sum_{j=1}^n \mathbb{E}|X_{j,n}|^2}$. Les sommes indépendantes correspondantes ont ainsi une variance constante valant 1. On en déduit qu'une condition de convergence est ici

$$\lim_{n\to\infty} \frac{\sum_{j=1}^n \mathbb{E}|X_{j,n}|^3}{\left(\sum_{j=1}^n \mathbb{E}|X_{j,n}|^2\right)^{3/2}} = 0$$

On peut utiliser le théorème précédent pour deux propos (au moins !).

§2. Approximation d'une binomiale

Une première utilisation du théorème de Lindeberg est la validation de l'approximation gaussienne d'une loi binomiale de paramètres n et p. Une telle variable Z_n s'exprime bien comme somme de variables de Bernoulli indépendantes de même paramètre $p \in [0,1]$. Quitte à recentrer ces variables (par p) on approche une telle somme (renormalisée) $\frac{Z_n - np}{\sqrt{np(1-p)}}$ par une gaussienne T_n, de loi $N(0,1)$ alors le théorème précédent montre que

$$\left| \mathbb{E} f\left(\frac{Z_n - np}{\sqrt{np(1-p)}}\right) - f(T_n) \right| \leq \frac{c_3}{3\sqrt{np}}.$$

Par suite la qualité de l'approximation gaussienne est mesurée par la grandeur np où $p = p(n)$ peut dépendre de n. On sait que lorsque ce produit converge vers une constante λ alors la suite des lois de Z_n a pour limite une loi de Poisson de paramètre λ.

Une utilisation de l'énoncé est aussi valide pour le test du χ^2. Soient $(p_1, \ldots, p_k)$ les probabilités théoriques d'apparition de chacune des k classes considérées. Alors l'application du théorème central limite vectoriel conduisant à ce test n'est valide que si elle est valide coordonnée par coordonnée. Une règle heuristique $n \min_i p_i > 5$ semble déduite de ces considérations ; plus sérieusement il faut, pour valider l'approximation gaussienne, que $n \min_i p_i$ soit grand en un sens à préciser. Ceci fait l'objet d'une simulation informatique. Conjointement avec les histogrammes de loi binomiales $B(n,p)$ recentrées et renormalisées, nous traçons ici celui d'une loi normale.

Programme Matlab (toolbox : statistics) :

```
function bcompar(n, p, N) ;
x =binornd(n, p, [1, N]) ;
x=(x − n ∗ p)/sqrt(n ∗ p ∗ (1 − p)) ;
k=floor(N^(1/3))
h = hist(x, k)
P = −3 + 6 ∗ [1 : k]/k
```

```
YY =normcdf(P)
M =ones(k,k)
T =tril(M)
c = T * h'/N
c = c'; plot(c,[1:k],'b',YY,[1:k],'r') ;
pause ; clf ; newplot
normplot(x) ;
```

§3. Le coussin financier

Une application financière (le coussin financier) est développée par le papier de référence. Un financier dispose de deux actifs, l'un d'entre eux, S est risqué et pas l'autre, B. On conserve une proportion constante d'exposition au risque : pour y parvenir le montant total investi sur l'actif financier est proportionnel à celui du coussin. La méthode est une stratégie dynamique permettant d'obtenir un portefeuille strictement plus avantageux qu'un plancher donné.

Précisément, nous discrétisons le temps (entre 0 et 1 par une convention aussi arbitraire que modifiable) en posant $t_{n,k} = \frac{k}{n}$, pour chaque entier $k = 1, 2, 3, \ldots, n-1$. L'entier n définit donc le pas de cette discrétisation : il est appelé à tendre vers $+\infty$. Ainsi $B_{n,k}$ et $S_{n,k}$ désignent le prix de l'actif non risqué et celui de l'actif risqué au temps $t_{n,k}$. Ceux-ci sont naturellement régis par les équations de récurrence

$$B_{n,k} = B_{n,k-1}\left(1 + \frac{r}{n}\right), \quad S_{n,k} = S_{n,k-1}\left(1 + Y_{n,k}\right). \tag{1}$$

Dans l'équation (1), r désignera un taux d'intérêt (déterministe) tandis que la famille de taux aléatoires $Y_{n,k}, k = 1, \ldots n-1$ sera supposée indépendante. De plus $B_{n,0}, S_{n,0}$ sont déterministes, et ne dépendent pas de n : on les note B_0, S_0. Pour des raisons naturelles, on ramène ces variables à ne prendre que 2 valeurs. Si ce n'était pas le cas on pourrait en effet raffiner la discrétisation du temps. De plus, pour des raisons de modélisation financière, on supposera qu'il existe des constantes $\mu > r$ et $v \geq 0$ est la volatilité du modèle, et des variables aléatoires indépendantes $Z_{n,k}$ telles que

$$Y_{n,k} = \frac{\mu}{n} + vZ_{n,k} \quad \text{avec} \quad \mathbb{P}\left(Z_{n,k} = \frac{-1}{\sqrt{n}}\right) = \mathbb{P}\left(Z_{n,k} = \frac{1}{\sqrt{n}}\right) = \frac{1}{2}. \tag{2}$$

La valeur du plancher évolue au taux sans risque donc $P_{n,k} = P_{n,k-1}\left(1 + \frac{r}{n}\right)$. Notant $V_{n,k}$ la valeur du portefeuille à l'instant $t_{n,k}$, on appelle coussin la différence

$$C_{n,k} = V_{n,k} - P_{n,k}.$$

Le montant total investi sur l'actif risqué est noté $e_{n,k}$. La méthode du coussin consiste à choisir $e_{n,k} = mC_{n,k}$ proportionnel au coussin (m est appelé

multiple). L'équation $V_{n,k} = V_{n,k-1} + e_{n,k-1}Y_{n,k} + (V_{n,k-1} - e_{n,k-1})\frac{r}{n}$ s'en déduit (voir le corrigé du texte). Par la relation $C_{n,k} = V_{n,k} - P_{n,k}$, il vient

$$\begin{aligned} C_{n,k} &= C_{n,k-1} + mC_{n,k-1}Y_{n,k} + (C_{n,k-1} + P_{n,k-1} - e_{n,k-1})\tfrac{r}{n} - (P_{n,k} - P_{n,k-1}) \\ &= C_{n,k-1} + mC_{n,k-1}Y_{n,k} + (C_{n,k-1} + -mC_{n,k-1})\tfrac{r}{n} \\ &= C_{n,k-1}\left(1 + mY_{n,k} + (1-m)\tfrac{r}{n}\right). \end{aligned}$$

Donc, en posant $C_0 = V_0 - P_0$,

$$C_{n,k} = C_0 \prod_{\ell=1}^{k} \left(1 + mY_{n,\ell} + (1-m)\tfrac{r}{n}\right) \tag{3}$$

Notons que le produit précédent a des facteurs tous positifs si $Y_{n,k} \geq ((m-1)r/n - 1)/m$.

Dans la suite nous nous focaliserons sur l'époque $t \in [0,1]$. Pour celà nous choisissons l'asymptotique $n \to \infty$, $k = k_n \to \infty$ et $t_{n,k} = \frac{k}{n} \to t$. Nous notons aussi $k_n = nt_{n,k}$. Pour simplifier les notations posons, $B_n^{(t)}, S_n^{(t)}, P_n^{(t)}, V_n^{(t)}, C_n^{(t)}$, pour les expressions $B_{n,k_n}, S_{n,k_n}, P_{n,k_n}, V_{n,k_n}, C_{n,k_n}$.

Alors $P_n^{(t)} = \left(1 + \frac{r}{n}\right)^{k_n} \to P_0 e^{rt}$ lorsque $n \to t$. De plus $\log C_n^{(t)} = \log C_0 + K_n^{(t)}$, où $K_n^{(t)} = \sum_{\ell=1}^{k} \log\left(1 + mY_{n,\ell} + (1-m)\frac{r}{n}\right)$. On obtient après un développement limité et l'utilisation du théorème de Lindeberg (qui pouvait être remplacé par le TLC dans ce cas)

$$K_n^{(t)} \Longrightarrow_{n\to\infty} K_t = t(m(\mu - r) + r) - v^2m^2t/2 + vm\sqrt{t}W \quad \text{si } W \sim \mathcal{N}(0,1).$$

La convergence suivante s'en déduit :

$$C_n^{(t)} \Longrightarrow_{n\to\infty} C_t = C_0 e^{\sqrt{t}mvW + [t(r+m(\mu-r)) - \frac{tm^2v^2}{2}]}.$$

Notons quand même que le théorème de Lindeberg permet de préciser une vitesse de convergence dans les énoncés précédents. En effet la série qui définit le développement limité de $u \mapsto \log(1+u)$ (quand $u \to 0$) est alternée par suite $|\log(1+u) - u + u^2/2| \leq u^3/3$ et posant $C = m^3v^3/3 + m^2r^2$, nous avons

$$\left| K_n^{(t)} - \left(t(m(\mu - r) + r) - \tfrac{v^2m^2t}{2} + \sum_{\ell=1}^{k_n} Z_{n,k} \right) \right| \leq \tfrac{C}{\sqrt{n}} \quad \text{si } n \geq 1.$$

Nous en déduisons ainsi (avec l'inégalité des accroissements finis) la borne significative

$$\left| \mathbb{E}\left(f(K_n^{(t)}) - f(K_t) \right) \right| \leq \left(\tfrac{c_3}{3} + Cc_1\right) \tfrac{1}{\sqrt{n}}.$$

Nous traçons l'histogramme de la loi de $K_n^{(t)}$ conjointement avec celui de sa limite théorique (voir le corrigé du texte) avec $\mu = 0.05, v = 0.2, S_0 = 100, V_0 = 100, P_0 = 90, \lambda = 2, m = 5$.

Chapitre 28
Commentaire d'un non spécialiste

[*Le point de vue que je prends, est celui d'un candidat candide. Je choisis donc de rester au plus près du texte. J'ai essayé de suivre le texte jusqu'au bout - en y mettant le temps. Il est évident que le candidat pouvait s'arrêter bien avant...*]

Ce texte décrit une stratégie de gestion de portefeuille, appelée CPPI (*Constant Proportion Portfolio Insurance*). Dans cette méthode, on fixe d'une part un plancher P au dessous duquel la valeur V du portefeuille ne doit pas descendre, et un nombre m appelé le multiple. On appelle *coussin* la différence $C = V - P$. La stratégie utilisée consiste à placer dans un placement à risque une somme égale à $m(V - P)$. Cela sous-entend que l'on peut placer une somme $> V$, en empruntant, ou à l'aide de placements à terme j'imagine. A cause de cela, on aurait pu tout aussi bien limiter les paramètres en posant $V_0 = 0$. Dans ce cas, évidemment, on a $P < 0$ et $-P$ représente juste la somme que l'on accepte de perdre.

On utilisera des approximations du temps continu par le temps discret (et *vice versa*). Dans le modèle discret, le temps est découpé en une suite d'intervalles de longueur $1/n$. Il me semble que le modèle discret avec n grand est plus réaliste que le modèle continu : c'est en ce sens que je dis que le temps continu approxime le temps discret.

Le tout est compliqué par l'évolution dans le temps du placement sûr. *On supposera dans la suite que* $r = 0$ - on peut, évidemment, dans les calculs se ramener à ce cas en multipliant les quantités V, P, S et C par un facteur e^{-rt} (en temps continu - un facteur $(1 + r/n)^{-k}$ en temps discret). Par rapport au texte, on procédera ainsi à quelques ajustements : en particulier, l'espérance μ du taux de la valeur 'risquée' sera remplacée par $\mu - r$. Dans le traitement mathématique de notre problème, nous n'avons plus à nous occuper que d'un seul actif, l'actif risqué, dont la valeur est notée S.

En temps discret

La valeur $S_{n,k}$ de S au temps k/n évolue par une équation

$$S_{n,k} = S_{n,k-1}(1 + Y_{n,k}),$$

où $Y_{n,k}$ s'appelle le *taux d'accroissement* au temps k/n.

On note $V_{n,k}$ la valeur du portefeuille au temps k/n et $C_{n,k} = V_{n,k} - P$ celle du coussin. Notre *exposition*, le montant investi dans la valeur à risque est $e_{n,k} = mC_{n,k}$.

Au temps $(k-1)/n$, on a donc investi (dans la valeur à risque) la somme $e_{n,k-1}$ qui donnera $e_{n,k-1}(1 + Y_{n,k})$ au temps k/n et on a gardé (investie dans la valeur sans risque) la somme $V_{n,k-1} - e_{n,k-1}$.

On a donc $V_{n,k} = e_{n,k-1}(1 + Y_{n,k}) + V_{n,k-1} - e_{n,k-1}$, donc

$$V_{n,k} = V_{n,k-1} + e_{n,k-1}Y_{n,k} = V_{n,k-1} + mC_{n,k-1}Y_{n,k}.$$

En particulier, $C_{n,k} = V_{n,k} - P = (1 + mY_{n,k})C_{n,k-1}$, soit

$$C_{n,k} = C_0 \prod_{j=1}^{k} (1 + mY_{n,j}).$$

Nous utilisons, pour l'évolution de la valeur S, le modèle de Cox-Ross-Rubinstein. C'est un modèle de type 'pile ou face' : on pose

$$Y_{n,k} = \tfrac{\mu}{n} + \tfrac{v}{\sqrt{n}}\varepsilon_{n,k}$$

où μ et v sont des constantes strictement positives et $\varepsilon_{n,k}$ est une suite de variables aléatoires de Rademacher indépendantes.

En temps continu

On a

$$\ln S_{n,n} = \ln S_0 + \sum_{k=0}^{n-1} \ln(1 + Y_{n,k}).$$

Remarquons que $\ln(1 + Y_{n,k})$ prend deux valeurs avec probabilité $1/2$: on a

$$\ln(1 + Y_{n,k}) = a_n + b_n\varepsilon_{n,k}$$

où on a posé

$$a_n = \tfrac{1}{2}\left(\ln\left(1 + \tfrac{\mu}{n} + \tfrac{v}{\sqrt{n}}\right) + \ln\left(1 + \tfrac{\mu}{n} - \tfrac{v}{\sqrt{n}}\right)\right)$$

$$\text{et} \quad b_n = \tfrac{1}{2}\left(\ln\left(1 + \tfrac{\mu}{n} + \tfrac{v}{\sqrt{n}}\right) - \ln\left(1 + \tfrac{\mu}{n} - \tfrac{v}{\sqrt{n}}\right)\right).$$

On a donc

$$\ln S_{n,n} = \ln S_0 + na_n + b_n \sum_{k=0}^{n-1} \varepsilon_{n,k}.$$

Par le théorème de limite centrale, la suite $n^{-1/2}\sum_{k=0}^{n-1}\varepsilon_{n,k}$ converge en loi vers une variable gaussienne centrée réduite W. Comme $a_n = \frac{\mu - \frac{1}{2}v^2}{n} + O(n^{-2})$ et $b_n = \frac{v}{\sqrt{n}} + O(1/n)$, on en déduit :

Proposition 1. *La variable aléatoire $S_{n,n}$ converge en loi vers une variable*

$$S_1 = S_0 \exp\left(\mu - \tfrac{1}{2}v^2 + vW\right)$$

où W est une variable aléatoire de loi gaussienne centrée réduite.

L'évolution du coussin en temps discret est celle d'une valeur S' évoluant de manière identique à S. On doit juste remplacer $Y_{n,k}$ par $mY_{n,k}$. Cela donne donc

$$\ln C_{n,n} = \ln C_0 + na'_n + b'_n \sum_{k=0}^{n-1} \varepsilon_{n,k},$$

où l'on a posé

$$a'_n = \tfrac{1}{2} \ln \left(\left(1 + \tfrac{m\mu}{n}\right)^2 - \tfrac{m^2 v^2}{n} \right) \quad \text{et}$$

$$b'_n = \tfrac{1}{2} \left(\ln \left(1 + \tfrac{m\mu}{n} + \tfrac{mv}{\sqrt{n}}\right) - \ln \left(1 + \tfrac{m\mu}{n} - \tfrac{mv}{\sqrt{n}}\right) \right).$$

Remarquons que la suite na'_n converge vers $m\mu - \frac{1}{2}m^2v^2$ et la suite $\sqrt{n}b'_n$ converge vers mv. On en déduit :

Proposition 2. *La variable aléatoire $C_{n,n}$ converge en loi vers une variable*

$$C_1 = C_0 \exp \left(m\mu - \tfrac{1}{2}m^2v^2 + mvW \right)$$

(où W est la variable aléatoire de loi gaussienne centrée réduite ci-dessus).

En fait, la valeur du coussin en temps continu est :

$$C_1 = C_0 \left(\tfrac{S_1}{S_0}\right)^m \exp \tfrac{(m-m^2)v^2}{2}.$$

Remarquons de plus que les valeurs S_t et C_t de S et du coussin au temps t, s'obtiennent en prenant la limite de $S_{n,[nt]}$ et $C_{n,[nt]}$, où $[nt]$ désigne la partie entière de nt. Toujours par le théorème de limite centrale, on trouve :

Proposition 3. *Les variables $(S_{n,[nt]})_t$ et $(C_{n,[nt]})_t$ convergent respectivement en loi vers des variables*

$$S_t = S_0 \exp \left((\mu - \tfrac{v^2}{2})t + vW_t \right) \quad \textit{et} \quad C_t = C_0 \exp \left((m\mu - \tfrac{m^2v^2}{2})t + mvW_t \right)$$

où W_t est une variable de loi $\mathcal{N}(0,t)$.

En particulier C_t suit une loi log normale de densité g_t donnée (pour $x \in \mathbf{R}_+^*$) par

$$g_t(x) = \tfrac{1}{xmv\sqrt{2\pi t}} \exp - \frac{\left(\ln \frac{x}{C_0} - m\mu t + \frac{1}{2}m^2v^2t \right)^2}{2m^2v^2t}. \tag{2}$$

Pour illustrer informatiquement, on prend $\mu = 0,1$, $v = 0,2$, $C_0 = 10$, $t = 1$ et on représente sur le même graphique g pour les cinq valeurs de $m = 1,2,3,4,5$.

Rentabilité/risque

Il s'agit ici de trouver la stratégie qui permet de minimiser la variance à espérance donnée. Pour cela, on fait varier les paramètres P (le plancher) et m (le multiple) et on cherche l'équation qui, à espérance donnée, les relie.

Par la formule d'intégration d'une exponentielle sous une gaussienne, on calcule facilement l'espérance et la variance du coussin. On a :

$$\mathbb{E}(C_t) = C_0 e^{tm\mu} \quad \text{et} \quad \mathbb{E}\big((C_t)^2\big) = C_0^2 \exp t\big(2m\mu + m^2v^2\big),$$

donc

$$Var(C_t) = \mathbb{E}(C_t)^2 \Big(\exp(tm^2v^2) - 1\Big).$$

L'espérance de la valeur du portefeuille est donc (en fonction du plancher P)

$$\mathbb{E}(V_t) = \mathbb{E}(C_t) + P = (V_0 - P)\exp(tm\mu) + P.$$

En particulier, si l'on veut que cette espérance vaille L en $t = 1$, on doit avoir

$$P = \tfrac{V_0 \exp m\mu - L}{\exp m\mu - 1}.$$

Remarquons que $C_1 - C_0 = V_1 - V_0$. On peut donc exprimer $\mathbb{E}(C_1)$ en fonction de L :

$$\mathbb{E}(C_1) = \tfrac{L - V_0}{1 - \exp(-m\mu)} \qquad (1).$$

La variance de V_t est égale à celle de C_t. Par la formule (1), on a

$$Var(V_1) = (L - V_0)^2 \ \frac{\exp\left(m^2v^2\right) - 1}{\Big(1 - \exp(-m\mu)\Big)^2}.$$

Prenons par exemple $\mu = v = 0,2$. Posant $x = \frac{m}{5}$, on doit minimiser

$$\tfrac{e^{x^2} - 1}{(1 - e^{-x})^2}.$$

La commande 'plot' de maple permet de se convaincre que cette fonction est croissante. On a écrit :
plot (exp(x ^2)-1/(1-exp(-x))^2,x=0..infinity)

Pour le vérifier, écrivons

$$Var(V_1) = \tfrac{(L - V_0)^2 v^2}{\mu^2} \tfrac{g(m^2v^2)}{g(-m\mu)^2}$$

où l'on a posé

$$g(x) = \tfrac{e^x - 1}{x}.$$

Comme g est une fonction positive croissante, la variance est fonction croissante de m. Le minimum de la variance est atteint pour $m \to 0$, donc $P \to -\infty$. Cette limite correspond à une exposition constante

$$e = \lim_{m \to 0} mC_t = \tfrac{L - V_0}{\mu},$$

ce qui est en un sens le contraire même de la stratégie du coussin. Remarquons que l'on peut minimiser la variance à espérance constante en prenant des valeurs *négatives* de m. Le minimum est en fait atteint pour $m = -\frac{\mu}{v^2}$. Dans ce cas, le plancher est remplacé par un plafond et le coussin est négatif à tout instant !

Cependant, dans notre stratégie, le plancher est primordial. Une fois celui-ci fixé, on vise à augmenter l'espérance de gains, au prix d'une forte augmentation de la variance. Comment choisir alors un multiple convenable ? Proposons deux critères :

1. On veut une probabilité d'augmenter notre capital $\geq p$. Cela donne $I\!P(C_t \geq C_0) \geq p$, ce qui se lit avec des quantiles d'une gaussienne : $I\!P(W_t \geq \frac{2\mu - mv^2}{2v} t) \geq p$.
2. La médiane reste supérieure ou égale à celle obtenue avec la stratégie simple $m = 1$: cela donne $m\mu - \frac{m^2 v^2}{2} \geq \mu - \frac{v^2}{2}$, soit $\mu \geq \frac{(m+1)v^2}{2}$: on prend $m = 2\mu v^{-2} - 1$.

Le coussin en temps continu avec sauts

On veut à présent modéliser des "accidents boursiers".

Pour cela, on se donne en plus une suite croissante de temps aléatoires $(T_j)_j$ en lesquels la valeur S subit une modification brutale. En d'autres termes on a

$$S_t = S_0 \exp\Big((\mu - \tfrac{1}{2}v^2)t + vW_t\Big) \prod_{0 \leq T_j \leq t} (1 + X_j)$$

où X_j est le saut au temps T_j.

Pour calculer la valeur du coussin, on peut à nouveau utiliser une approximation par un temps discret. Cela revient à remplacer $1 + Y_{n,k}$ par $(1 + Y_{n,k})(1 + X_j)$ si $k = [nT_j]$.

L'évolution du coussin se déduit de celle de S en remplaçant Y par mY. On en déduit aisément que, dans notre modèle, on a

$$C_t = C_0 \exp\Big((m\mu - \tfrac{1}{2}m^2 v^2)t + mvW_t\Big) \prod_{0 \leq T_j \leq t} (1 + mX_j).$$

Pour plus de commodité, notons C_t^0 la valeur du coussin sans saut, de sorte que la valeur du coussin avec sauts est

$$C_t = C_t^0 \prod_{0 \leq T_j \leq t} (1 + mX_j).$$

Pour que le portefeuille reste au-dessus du plancher fixé, il faut supposer que, presque sûrement, le coussin reste positif, soit $X_j \geq -1/m$ pour tout j tel que $T_j \leq t$.

On suppose maintenant que la suite des durées $T_{j+1} - T_j$ est une suite de variables aléatoires indépendantes (et indépendantes de $(W_t)_t$) qui suivent

une loi exponentielle de même paramètre λ. Dans ce cas, le nombre de sauts avant le temps t suit une loi de Poisson de paramètre λt.

On suppose de plus que les variables aléatoires X_j sont indépendantes (avec W_t, les T_j et entre elles).

L'espérance du coussin sachant qu'il y a n sauts de 0 à t est l'espérance de la variable $C_t^n = C_t^0 \prod_{j=1}^n (1 + mX_j)$. On a

$$\mathbb{E}(C_t^n) = \mathbb{E}(C_t^0) \prod_{j=1}^n (1 + m\mathbb{E}(X_j)) = C_0 e^{tm\mu}(1 + mb)^n$$

supposant que les X_j ont même espérance b. On trouve

$$\mathbb{E}(C_t) = C_0 e^{tm\mu} e^{-\lambda t} \sum_{n=0}^{+\infty} \frac{(t\lambda)^n (1+mb)^n}{n!} = C_0 \exp(tm(\mu + b\lambda)).$$

L'espérance du carré de C_t se calcule de même, supposant de plus que les X_j ont même variance σ^2. On trouve

$$\begin{aligned}
\mathbb{E}\big((C_t)^2\big) &= C_0^2 e^{t\left(2m\mu + m^2 v^2\right)} e^{-\lambda t} \sum_{n=0}^{+\infty} \frac{(t\lambda)^n (1+2mb+m^2(b^2+\sigma^2))^n}{n!} \\
&= C_0^2 \exp t\big(2m\mu + m^2 v^2 + 2\lambda mb + \lambda m^2 (b^2 + \sigma^2)\big) \\
&= \mathbb{E}(C_t)^2 \exp tm^2\big(v^2 + \lambda(b^2 + \sigma^2)\big).
\end{aligned}$$

En d'autres termes, pour tout ce qui concerne uniquement variance et espérance (par exemple pour la discussion *rentabilité-risque*), on se ramène au cas sans saut en remplaçant μ par $\mu + \lambda b$ et v par $\sqrt{v^2 + \lambda(b^2 + \sigma^2)}$.

Calculons à présent la densité de la loi de C_t.

Notons F_j la loi de X_j et g_t^n la densité de la variable C_t sachant que le nombre de $T_j \leq t$ est n. On a $C_t^n = C_t^{n-1}(1 + mX_n)$, donc

$$g_t^n(x) = \int_0^{+\infty} g_t^{n-1}\left(\frac{x}{1+my}\right) \frac{dF_n(y)}{1+my}.$$

La formule pour g^0 est celle du coussin sans saut donnée ci dessus (formule (2)). On en déduit par récurrence sur n que

$$g_t^n = \int_{\mathbf{R}^n} g_t^0 \left(\frac{x}{\prod_{j\leq n}(1+my_j)}\right) \prod_{j=1}^n \frac{dF_j(y_j)}{1+my_j}.$$

La densité g_t de C_t est donnée par

$$g_t(x) = \sum_{n=0}^{+\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} g_t^n(x).$$

Nous proposons deux illustrations informatiques.

On prend $\mu = 0,1,\ v = 0,2,\ C_0 = 10,\ m = 5,\ t = 1$ et on représente sur le même graphique g pour les cinq valeurs de $\lambda = 0$, $\lambda = k/2$, où $k = 0,1,2,3,4$ dans les deux cas suivants :

1. On prend X_j équidistribué sur l'intervalle $[-(0,1),0]$. On tronque le calcul donnant g_1 à $n = 6$. En effet, pour $\lambda \leq 2$, la probabilité qu'il y ait 7 sauts ou plus est $e^{-\lambda}\sum_{n=7}^{\infty}\frac{\lambda^n}{n!} < 1\%$.
2. On prend les X_j constants égaux à une valeur $b = -0,05$. Dans ce cas on a

$$g_t(x) = \sum_{n=0}^{+\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!} g_t^0\Big(\frac{x}{(1+mb)^n}\Big)(1+mb)^{-n}.$$

Gestion par quantile

On suppose ici que les sauts X_j peuvent descendre sous la valeur fatidique $-1/m$, mais avec une petite probabilité p indépendante de j. Dans ce cas, la probabilité de rester au dessus du plancher après n sauts est $(1-p)^n$. Donc la probabilité de rester au dessus du plancher après un temps t est

$$e^{-\lambda t}\sum_{n=0}^{\infty}\frac{(\lambda t)^n(1-p)^n}{n!} = e^{-p\lambda t}.$$

Si l'on veut que la probabilité de rester au dessus du plancher après un temps $t = 1$ soit $\geq 1-\varepsilon$, il faut donc que $e^{-p\lambda} \geq 1-\varepsilon$ soit que $p \leq -\frac{\ln(1-\varepsilon)}{\lambda}$. Ce qui se lit, en supposant que les X_j obéissent à la même loi F :

$$F(-\tfrac{1}{m}) = \mathbb{P}(X_j < -\tfrac{1}{m}) \leq -\tfrac{\ln(1-\varepsilon)}{\lambda}.$$

Si F est strictement croissante (sur le support de X_j) cette inégalité peut aussi s'écrire $-\frac{1}{m} \leq F^{(-1)}(-\frac{\ln(1-\varepsilon)}{\lambda})$ soit

$$m \leq \frac{-1}{F^{(-1)}(-\frac{\ln(1-\varepsilon)}{\lambda})}.$$

Chapitre 29
Le coussin financier : TIPE

Préliminaire.

Ce texte se place du point de vue d'un étudiant de seconde année de premier cycle ou d'un élève de spéciales ; par exemple dans l'optique de la préparation à l'épreuve d'Analyse de Document Scientifique proposée par les concours au titre des Travaux d'Initiative Personnels Encadrés. Les prérequis en probabilités sont donc contenus dans le tronc commun des nouveaux programmes de Terminale Scientifique (BO hors série n°4 du 30 aout 2001) en vigueur à la rentrée 2002. Les libertés ad hoc que je m'autoriserai en matière d'espace mesuré seront assez légères au sens où les seules parties de l'intervalle $[0,1]$ que l'on aura véritablement besoin de mesurer seront des intervalles ou, au pire, des réunions d'un nombre fini d'intervalles.

1. Introduction.

Je me propose, à partir des paragraphes 2.1 et 2.2 (formules (1) à (7)) du texte portant sur le Coussin Financier, de donner un exemple concret de réalisation mathématique de variables indépendantes et de même loi : $(Z_{n,k})_{k\in\{1,\ldots,n\}}$ et d'étudier , pour cet exemple, le comportement asymptotique $(n\to\infty)$ de la variable aléatoire $C_{n,n}$ qui modélise la valeur du coussin à échéance c'est à dire l'écart entre la valeur du portefeuille à la fin du processus financier et celle du plancher fixée de façon déterministe puisqu' elle évolue exactement comme l'investissement sans risque.
On a clairement : $C_{n,n}=V_{n,n}-P_0\left(1+\frac{r}{n}\right)^n$et la connaissance asymptotique $(n\to\infty)$ de $C_{n,n}$ qui peut être déduite de (6) contient donc celle de $V_{n,n}$ qui est l'un des premiers objectifs de l'ensemble du document à analyser.
Le texte tentera ensuite de justifier, aux yeux d'un non-probabiliste, l'expression en $\frac{v}{\sqrt{n}}$ choisie pour mesurer l'amplitude de l'oscillation aléatoire du taux "d'accroissement" dans le modèle de Cox-Ross-Rubinstein : $Y_{n,k}=\frac{\mu}{n}+\frac{v}{\sqrt{n}}R_{n,k}$ où $R_{n,k}=\sqrt{n}Z_{n,k}$ est une variable dite de Rademacher qui prend les valeurs 1 et -1 avec la même probabilité :

$$P\left(R_{n,k}=1\right)=P\left(R_{n,k}=-1\right)=\tfrac{1}{2}.$$

2. Construction de variables $(Z_{n,k})_{k\in\{1,\ldots,n\}}$.

On construit d'abord une suite $(X_k)_{k\in\mathbb{N}}$ de variables aléatoires indépendantes sur l'univers $\Omega=[0,1[$ de la manière suivante : si ω est un événement élémentaire i.e. un réel de $[0,1[$, on appelle $X_k(n)$, le $k^{ième}$ chiffre après la virgule de son écriture binaire propre autrement dit :

$$X_k\left(\sum_{n=1}^{\infty}\frac{a_n}{2^n}\right)=X_k\left(0,a_1a_2\cdots\right)=a_k \qquad \text{où } a\in\{0,1\}^{\mathbb{N}^*} \tag{1}$$

On considère ici la probabilité uniforme sur Ω : $P([a,b])=b-a$ pour tout intervalle $[a,b]\subset[0,1[$.

L'indépendance de la suite $X = (X_k)_{k\in N}$ sera prouvée si on montre que $(\forall n \in \mathbb{N}^*)\,(\forall\,(i_1, i_2, ..., i_n) \in \{0,1\}^n)$:

$$P\left((X_1, X_{2,}..., X_n) = (i_1, i_2, ..., i_n)\right) = \prod_{k=1}^{n} P\left(X_k = i_k\right)$$

Question 1 : Pour chaque couple $(n,k) \in \mathbb{N}^2$ avec $1 \leq k \leq n$, décrire la partie de Ω : $\varpi\,(n, i_k) = 1 \leq k \leq n \bigcap X_k^{-1}\,(i_k) = \{\omega \in \Omega;\ \forall k \in [1,n]\ \ X_k = i_k\}$. En déduire l'indépendance de la suite X.
On pose maintenant, pour $(n,k) \in \mathbb{N}^2$ avec $1 \leq k \leq n$:

$$Z_{n,k} = \tfrac{2X_k - 1}{\sqrt{n}} \overset{déf}{=} \tfrac{R_k}{\sqrt{n}}.$$

Question 2 : Montrer que $Z_{n,1}, Z_{n,2}, ..., Z_{n,n}$ sont indépendantes et qu'elles suivent la loi de probabilité spécifiée à la formule (3) : celle du coefficient aléatoire de la volatilité du taux d'évolution de l'actif risqué.

Remarque : Jusqu'ici, nous avons modélisé rigoureusement la situation "concrète" suivante : n est fixé, c'est le nombre des instants auxquels est décidée la valeur du taux d'évolution de l'actif risqué. Un "Dieu du marché" détermine ledit taux de la manière suivante : avant le début des transactions, il tire au hasard un réel $\omega \in [0,1[$ puis, à chaque instant $\frac{k}{n}$ $(k \in [1,n])$, il observe la $k^{ième}$ décimale de l' écriture binaire de ω.
Si c'est un 0, $Z_{n,k}\,(\omega) = -\frac{1}{\sqrt{n}}$ et, au grand dam de l'investisseur, l'actif risqué passe de

$$S_{n,k}\,(\omega)\ \text{à}\ S_{n,k+1}\,(\omega) = S_{n,k}\,(\omega)\left(1 + \tfrac{\mu}{n} - \tfrac{v}{\sqrt{n}}\right). \tag{2}$$

Si c'est un 1, $Z_{n,k}\,(\omega) = \frac{1}{\sqrt{n}}$ et l'actif risqué passe de

$$S_{n,k}\,(\omega)\ \text{à}\ S_{n,k+1}\,(\omega) = S_{n,k}\,(\omega)\left(1 + \tfrac{\mu}{n} + \tfrac{v}{\sqrt{n}}\right). \tag{3}$$

Le choix de la probabilité uniforme sur Ω rend cette modélisation raisonnable au regard de la complexité réelle du marché. L'ordre de grandeur en $\frac{1}{\sqrt{n}}$ choisi pour $|Z_{n,k}\,(\omega)|$ s'explique profondément par le calcul qui va suivre.

3. Comportement asymptotique de $T_n = \ln(C_{n,n})$.
La formule (6) donne :

$$T_n = \ln(C_0) + 1 \leq k \leq n \ln\left(1 + mY_{k,n} + (1-m)\,\tfrac{r}{n}\right).$$

On va étudier le comportement de chacun des termes de cette somme. On constate tout d'abord que

$$mY_{k,n} + (1-m)\,\tfrac{r}{n} = \tfrac{m\mu + (1-m)r}{n} + \tfrac{mv}{\sqrt{n}} R_k, \tag{4}$$

et comme $|R_k| = |2X_k - 1| = 1$, on a clairement :

$\forall \omega \in \Omega$, $mY_{k,n}(\omega) + (1-m)\frac{r}{n}n \to \infty \to 0$ ce qui justifie le développement limité à l'ordre 2 suivant où il est tenu compte du fait que $R_k^2(\omega) = 1$:

$$\begin{aligned} \ln\left(1 + mY_{k,n}(\omega) + (1-m)\frac{r}{n}\right) &= \frac{m\mu + (1-m)r}{n} + \\ &\quad \frac{mv}{\sqrt{n}} R_k(\omega) - \frac{m^2 v^2}{2n} + \frac{1}{n\sqrt{n}} b_{n,k}(\omega) \end{aligned} \tag{5}$$

et où $b_{n,k}$ désigne une variable aléatoire (ne prenant que 2 valeurs comme le montre la formule ci-dessus)

Question 3 : Montrer que $(b_{n,k})_{n\in[N]^*, k\in[1,n]}$ est uniformément bornée i.e. que :

$$(\exists\, B \in \mathbb{R})\,;\, (\forall\, n \in \mathbb{N}^*)\,(\forall\, k \in [1,n])\,(\forall\, \omega \in [0,1[)\, |b_{n,k}(\omega)| \leq B. \tag{6}$$

Question 4 : En déduire que

$$T_n = \ln(C_0) + m\mu + (1-m)\,r - \frac{m^2 v^2}{2} + mv\Xi_n + \frac{B_n}{\sqrt{n}} \tag{7}$$

où Ξ_n est la variable aléatoire $\frac{1}{\sqrt{n}} \sum_{1\leq k\leq n} R_k$ et où $(B_n)_{n\in\mathbb{N}^*}$ est uniformément bornée.

Question 5 : Démontrer que

$$C_{n,n} = C_0 e^{m\mu + (1-m)r - \frac{m^2 v^2}{2}}\; e^{mv\Xi_n}\; U_n \tag{8}$$

où $(U_n)_{n\in\mathbb{N}^*}$ *est* une suite de variables aléatoires qui converge uniformément vers la variable aléatoire constante : 1.

La description asymptotique de $C_{n,n}$, qui est notre but, se ramène donc à celle de

$$\Xi_n = \frac{1}{\sqrt{n}} 1 \leq k \leq nR_k. \tag{9}$$

Pour attaquer cette question, nous allons introduire ici un instrument familier des probabilistes et qui ressortit à l'analyse de Fourier.

4. Fonction caractéristique d'une variable aléatoire. Si W est une variable aléatoire sur Ω, on définit la fonction caractéristique de W :

$$\Phi_W : \begin{cases} \mathbb{R} \to \mathbb{C} \\ t \to E\left(e^{itW}\right) \end{cases} \tag{10}$$

Si W est à valeurs dans un ensemble fini : $\Phi_W(t) = w \in W(\Omega) P(W = w)\, e^{itw}$.
Si W admet une densité w continue, $\Phi_W(t) =_{-\infty}^{+\infty} w(x)\, e^{itx} dx$.
Les probabilistes savent démontrer que Φ_W **caractérise** la loi de W. L'argument général réside dans une formule *d'inversion de Fourier.*

Question 6 : Démontrer cela dans le cas où l'ensemble $W(\Omega)$ est fini.

Nous pouvons donc caractériser la loi de Ξ_n en évaluant Φ_{Ξ_n}. Nous espérons ensuite pouvoir déduire, du comportement asymptotique de Φ_{Ξ_n}, celui de Ξ_n, mais là encore, nous devrons faire appel à un résultat admis. Voici comment on va procéder.

Nous allons construire une variable aléatoire W telle que Φ_W soit la limite simple de la suite de fonctions caractéristiques : $(\Phi_{\Xi_n})_{n\in\mathbb{N}^*}$:

$$\Phi_{\Xi_n} n \to \infty \overset{cvs}{\to} \Phi_W. \tag{11}$$

Le fait que, dans ces conditions, $(\Xi_n)_{n\in\mathbb{N}^*}$ converge en un certain sens vers W dépasse le niveau de ce texte et je me contenterai de décrire cette convergence héritée de celle des fonctions caractéristiques et que les probabilistes appellent la **convergence en loi.**

On dit que $(\Xi_n)_{n\in\mathbb{N}^*}$ converge en loi vers W et on écrit : $\Xi_n \overset{L}{\to} W$ pour exprimer que la connaissance de la moyenne de toute fonction de W contient l'information du comportement asymptotique de la moyenne de la même fonction de Ξ_n. Précisément :

$$\Xi_n \overset{L}{\to} W \Leftrightarrow \left(\forall\, f \in C_b^0(\mathbb{R},\mathbb{R})\right)\ E(f(\Xi_n))\, n \to \infty \to E(f(W)) \tag{12}$$

où $C_b^0(\mathbb{R},\mathbb{R})$ désigne l'ensemble des fonctions numériques, continues et bornées sur $\mathbb{R}$. Notons au passage que ce mode de convergence est le moins exigeant de ceux qu'on rencontre en probabilités, c'est donc celui qu'on rencontre le plus souvent.

Evaluons donc $\Phi_{\Xi_n}(t) = E\left(e^{it\Xi_n}\right) = E\left(\prod_{k=1}^{n} e^{it\frac{R_k}{\sqrt{n}}}\right) = E\left(\prod_{k=1}^{n} e^{itZ_{n,k}}\right)$.

Pour progresser, il nous faut nous souvenir de l'indépendance des variables $(Z_{n,k})_{k\in\{1,\dots,n\}}$ et aussi utiliser le résultat suivant, très facile à obtenir dans le cas où nous sommes (i.e. dans le cas où les variables prennent un nombre fini de valeurs) : si Z_1, Z_2 sont deux variables indépendantes $E(Z_1Z_2) = E(Z_1)E(Z_2)$ et, si $f \in C_b^0(\mathbb{R},\mathbb{C})$, alors $f(Z_1)$ et $f(Z_2)$ restent indépendantes.

Question 7 : Calculer $E\left(e^{itZ_{n,k}}\right)$, en déduire $\Phi_{\Xi_n}(t)$ puis $n \to \infty \lim \Phi_{\Xi_n}(t)$ pour $t \in \mathbb{R}$.

5. Densité gaussienne.

On considère désormais une variable W sur $\Omega = [0,1[$ dont la densité sur $\mathbb{R}$ est $w : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \to \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \end{cases}$ c'est à dire précisément que, pour tout intervalle $I \subset \mathbb{R}$ on a $P(W \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-\frac{x^2}{2}} dx$.

On rappelle que c'est bien une densité puisque :

$$\begin{aligned} {}_{\mathbb{R}}e^{-\frac{x^2}{2}} dx &= 2{}_0^{\infty} e^{-\frac{x^2}{2}} dx \overset{u=\frac{x^2}{2}}{=} 2{}_0^{\infty} \frac{e^{-u}}{\sqrt{2u}} \\ &= \sqrt{2}{}_0^{\infty} e^{-u} u^{\frac{1}{2}-1} du = \sqrt{2}\Gamma\left(\tfrac{1}{2}\right) = \sqrt{2\pi}. \end{aligned}$$

Question 8 : Montrer que Φ_W est de classe C^1 sur $\mathbb{R}$ et qu'elle est solution d'une équation différentielle linéaire du premier ordre. En déduire son calcul et que $\Xi_n \overset{L}{\to} W$.
Question 9 : Décrire le comportement asymptotique de $(C_{n,n})_{n\in\mathbb{N}}$ qui peut se déduire de l'ensemble des résultats précédents.

6. Le théorème de la limite centrale comme argument et comme motivation.

Le point crucial du calcul qui assure la convergence simple de $(\Phi_{\Xi_n})_{n\in\mathbb{N}^*}$ vers la fonction caractéristique d'une loi gaussienne lorsque $\Xi_n = \frac{1}{\sqrt{n}} 1 \leq k \leq nR_k$ est **la présence de $\sqrt{n}$ au dénominateur de cette somme de variables indépendantes et de même loi** (de support fini). Le résultat général qu'il illustre est le "théorème de la limite centrale" qui assure que si $(R_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires réelles indépendantes et de même loi alors :

$$\frac{1}{\sqrt{n}} 1 \leq k \leq nR_k \overset{L}{\to} W \tag{13}$$

où W est une variable aléatoire réelle de densité gaussienne : $w : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \to \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \end{cases}$

On constate donc que : ce qui assure la convergence vers la loi gaussienne qui est l'une des lois les plus présentes dans la nature en raison du théorème de la limite centrale, c'est précisément le choix fait à la formule (3) du texte sur le coussin financier. Cette formule consiste en effet à attribuer la valeur $\pm\frac{v}{\sqrt{n}}$ à l'oscillation autour de $\frac{\mu}{n}$ pour l'évolution aléatoire du taux de l'actif risqué ce qui permet de rendre compte de deux choses :
a) Même dans un marché stable, la tendance (modélisée par μ/n) doit rester négligeable devant l'oscillation des actifs, qui méritent ainsi leur nom de "risqués" et que $v/\sqrt{n}$ (où v est la volatilité) modélise.
b) La convergence lorsque $n \to \infty$ du modèle discret est un fait de la nature ou plus exactement de la perception que nous en avons et la seule façon de "normaliser" la somme $1 \leq k \leq nR_k$ pour la faire "converger" (ne serait-ce qu'en loi) c'est de la diviser par $\sqrt{n}$ et c'est exactement ce que font les formules : $P\left(Z_{n,k} = \frac{1}{\sqrt{n}}\right) = P\left(Z_{n,k} = \frac{1}{\sqrt{n}}\right) = \frac{1}{2}$ du modèle "binomial" de Cox-Ross-Rubinstein.

Ainsi l'ordre de grandeur en $1/\sqrt{n}$ rend-t-il à la fois compte de la convergence du modèle discret et de la prédominance du risque sur la tendance.

7. Réponses.
Question 1 : Pour chaque couple $(n,k) \in \mathbb{N}^2$ avec $1 \leq k \leq n$, décrire la partie de Ω :

$$\varpi(n, i_k) = 1 \leq k \leq n \bigcap X_k^{-1}(i_k) = \{\omega \in \Omega;\ \forall k \in [1,n],\ X_k = i_k\}.$$

En déduire l'indépendance de la suite X.
On pose maintenant, pour $(n,k) \in \mathbb{N}^2$ avec $1 \leq k \leq n$:

$$Z_{n,k} = \frac{2X_k - 1}{\sqrt{n}} \overset{déf}{=} \frac{R_k}{\sqrt{n}}.$$

$\varpi(n, i_k)$ est l'ensemble des réels dont l'écriture binaire commence par $\overline{i_1 i_2 \cdots i_n}$, il s'agit des réels du type $\frac{i_1}{2} + \frac{i_2}{4} + \cdots + \frac{i_n}{2^n} + \frac{\alpha}{2^{n+1}}$ *où* $\alpha \in [0, 2[$, il s'agit de l'intervalle

$$\left[\frac{i_1}{2} + \frac{i_2}{4} + \cdots + \frac{i_n}{2^n}, \frac{i_1}{2} + \frac{i_2}{4} + \cdots + \frac{i_n}{2^n} + \frac{2}{2^{n+1}}\right[\tag{14}$$

sa longueur est :

$$\begin{aligned} P((X_1, X_2, ..., X_n) = (i_1, i_2, ..., i_n)) &= P(\varpi(n, i_k)) \\ = \frac{1}{2^n} &= 1 \leq k \leq n \prod \frac{1}{2} = 1 \leq k \leq n \prod P(X_k = i_k). \end{aligned}$$

La dernière égalité provient du fait que :
Pour tout $k \in [1, n]$ $P(X_k = i_k) = \frac{1}{2}$. Voici pourquoi :
Méthode 1. les événements complémentaires $\{X_k = i_k\}$ et $\{X_k = 1 - i_k\}$ sont "clairement" équiprobables donc leurs probabilités valent bien $\frac{1}{2}$.
Méthode 2. Décrivons l'événement

$$\{X_k = i_k\} = \bigcup_{(j_1, j_2, ..., j_{k-1}) \in \{0,1\}^{k-1}} \left[\bigcup_{\alpha \in [0,2[} \left\{ \frac{j_1}{2} + \frac{j_2}{2^2} + \cdots + \frac{j_{k-1}}{2^{k-1}} + \frac{i_k}{2^k} + \frac{\alpha}{2^{k+1}} \right\} \right] \tag{15}$$

Pour chacune des 2^{n-1} valeurs de $(j_1, j_2, ..., j_{k-1}) \in \{0, 1\}^{k-1}$, l'ensemble :

$$\bigcup_{\alpha \in [0,2[} \left\{ \frac{j_1}{2} + \frac{j_2}{2^2} + \cdots + \frac{j_{k-1}}{2^{k-1}} + \frac{i_k}{2^k} + \frac{\alpha}{2^{k+1}} \right\} \tag{16}$$

est un intervalle de longueur $\frac{1}{2^k}$ disjoint de tous les autres puisque les extrémités gauches de tous ces intervalles sont au moins distantes de $\frac{1}{2^{k-1}}$. On en déduit que

$$P(X_k = i_k) =_{(j_1, j_2, ..., j_{k-1}) \in \{0,1\}^{k-1}} \frac{1}{2^k} = \frac{2^{k-1}}{2^k} = \frac{1}{2}. \tag{17}$$

Question 2 : Montrer que $Z_{n,1}, Z_{n,2}, ..., Z_{n,n}$ sont indépendantes et qu'elles suivent la loi de probabilité spécifiée à la formule (3) : celle du coefficient aléatoire de la volatilité du taux d'évolution de l'actif risqué.
On a évidemment :

$$\left\{ Z_{n,k} = \pm \frac{1}{\sqrt{n}} \right\} = \{2X_k - 1 = \pm 1\} = \left\{ X_k = \frac{1 \pm 1}{2} \right\}$$

d'où l'indépendance et la conformité à la formule (3) de la loi commune des $(Z_{n,k})_{1 \leq k \leq n}$.
Question 3 : Montrer que $(b_{n,k})_{n \in \mathbb{N}^*,\, k \in [1,n]}$ est uniformément bornée i.e. que :

$$(\exists\, B \in \mathbb{R})\,;(\forall\, n \in \mathbb{N}^*)\,(\forall\, k \in [1,n])\,(\forall\, \omega \in [0,1[)\,|b_{n,k}(\omega)| \leq B. \tag{18}$$

Posons $u = \frac{1}{\sqrt{n}}\left(\frac{a}{\sqrt{n}} + bR_k(\omega)\right) = \frac{m\mu+(1-m)r}{n} + \frac{mv}{\sqrt{n}}R_k(\omega)$, on a clairement $|u| \leq \frac{a+b}{\sqrt{n}}$ d'où pour n assez grand, $|u| < 1$ et les résultats sur les séries alternées fournissent, pour ces u, la majoration : $\left|\ln(1+u) - u + \frac{u^2}{2}\right| \leq \left|\frac{u^3}{3}\right|$ qui, dans notre situation, assure :

$$\left|\ln\left(1 + mY_{k,n}(\omega) + (1-m)\frac{r}{n}\right) - \frac{m\mu+(1-m)r}{n} - \frac{mv}{\sqrt{n}}R_k(\omega) + \frac{m^2v^2}{2n}\right| =$$
$$\left|\ln(1+u) - u + \frac{b^2}{2n}\right| \leq \left|\ln(1+u) - u + \frac{u^2}{2}\right| + \left|\frac{u^2}{2} - \frac{b^2}{2n}\right| \leq$$
$$\left|\frac{u^3}{3} + \frac{a^2}{2n^2} + \frac{2ab}{n\sqrt{n}}\right| \leq \frac{(a+b)^3+a^2+2ab}{n\sqrt{n}} \overset{déf}{=} \frac{B}{n\sqrt{n}}.$$

Question 4 : En déduire que

$$T_n = \ln(C_0) + m\mu + (1-m)r - \frac{m^2v^2}{2} + mv\Xi_n + \frac{B_n}{\sqrt{n}} \tag{19}$$

où Ξ_n est la variable aléatoire $\frac{1}{\sqrt{n}}1 \leq k \leq nR_k$ et où $(B_n)_{n\in\mathbb{N}^*}$ *est uniformément bornée.*
Il suffit de sommer et de constater que

$$|B_n(\omega)| = \left|\frac{1}{n\sqrt{n}}{}_{k=1}^{n} b_{n,k}(\omega)\right| \leq \frac{1}{n\sqrt{n}}{}_{k=1}^{n} B = \frac{B}{\sqrt{n}}. \tag{20}$$

Question 5 : Démontrer que

$$C_{n,n} = C_0 e^{m\mu+(1-m)r-\frac{m^2v^2}{2}}\; e^{mv\Xi_n}\; U_n \tag{21}$$

où $(U_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires qui converge uniformément vers la variable aléatoire constante : 1.
On prend bien-sûr l'exponentielle dans la formule obtenue à la question 4 et

$$\left|e^{\frac{B_n(\omega)}{\sqrt{n}}} - 1\right| \leq 2\left|\frac{B_n(\omega)}{\sqrt{n}}\right| \leq 2\frac{B}{\sqrt{n}} \textit{ assure que } U_n \; n \overset{cvu}{\to} \infty \to 1 \tag{22}$$

Question 6 : Démontrer que si W est à valeurs dans un ensemble fini, la connaissance de la fonction : Φ_W, définie par la formule :

$$\Phi_W(t) = w \in W(\Omega) P(W = w)\, e^{itw} \tag{23}$$

caractérise la loi de W dans le cas où l'ensemble $W(\Omega)$ est fini.
C'est clair puisque la famille des fonctions $\left(t \to e^{itw}\right)_{w\in W(\Omega)}$ est libre dans l'espace vectoriel qu'elle engendre sur $\mathbb{R}$. Ainsi la connaissance de Φ_W assure-t-elle celle de ses coordonnées dont la liste est la loi de W.
Question 7 : Calculer $E\left(e^{itZ_{n,k}}\right)$, en déduire $\Phi_{\Xi_n}(t)$ puis $n \to \infty \lim \Phi_{\Xi_n}(t)$ pour $t \in \mathbb{R}$.

Très facilement :

$$\begin{aligned} E\left(e^{itZ_{n,k}}\right) &= E\left(e^{it\frac{R_k}{\sqrt{n}}}\right) = e^{\frac{it(-1)}{\sqrt{n}}} P\left(R_k=-1\right) + e^{\frac{it(1)}{\sqrt{n}}} P\left(R_k=1\right) \\ &= \frac{e^{-\frac{it}{\sqrt{n}}}+e^{\frac{it}{\sqrt{n}}}}{2} = \cos\left(\frac{t}{\sqrt{n}}\right) \end{aligned}$$

La propriété rappelée juste avant l'énoncé de la question 7 assure que :

$$\begin{aligned} \Phi_{\Xi_n}(t) &= \cos^n\left(\frac{t}{\sqrt{n}}\right) = e^{nLn\left(1-\frac{t^2}{2n}+0\left(\frac{1}{n}\right)\right)} \\ &= e^{n\left(-\frac{t^2}{2n}+0\left(\frac{1}{n}\right)\right)} = e^{-\frac{t^2}{2}+0(1)} n \to \infty \to e^{-\frac{t^2}{2}}. \end{aligned}$$

Question 8 : Montrer que Φ_W est de classe C^1 sur $\mathbb{R}$ et qu'elle est solution d'une équation différentielle linéaire du premier ordre. En déduire son calcul et que $\Xi_n \overset{L}{\to} W$

On a $\Phi_W(t) = \frac{1}{\sqrt{2\pi}} {}_{-\infty}^{+\infty} e^{itx-\frac{x^2}{2}} dx$ appelons f l'intégrande :

$$\begin{cases} \mathbb{R}^2 \to \mathbb{C} \\ (x,t) \to e^{itx-\frac{x^2}{2}} \end{cases}$$

$$\text{Alors : } \left|\frac{\partial f}{\partial t}(x,t)\right| = \frac{|x|}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \overset{déf}{=} \phi(x). \tag{24}$$

ϕ est clairement intégrable sur $\mathbb{R}$.

Ainsi a-t-on : $\Phi_W \in C^1(\mathbb{R},\mathbb{C})$ et $\Phi'_W(t) = \frac{1}{\sqrt{2\pi}} {}_{-\infty}^{+\infty} ixe^{itx-\frac{x^2}{2}} dx$ donc

$$\begin{aligned} \frac{\Phi'_W(t)}{i} &= \frac{1}{\sqrt{2\pi}} {}_{-\infty}^{+\infty} xe^{itx-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} {}_{-\infty}^{+\infty} -(it-x)\, e^{itx-\frac{x^2}{2}} dx + \frac{it}{\sqrt{2\pi}} {}_{-\infty}^{+\infty} e^{itx-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \left[e^{itx-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} + it\Phi_W(t) = it\Phi_W(t). \end{aligned}$$

On a donc montré que Φ_W est solution de l'équation différentielle : $y' = -ty$ donc $\Phi_W(t) = \Phi_W(0)\, e^{-\frac{t^2}{2}} = e^{-\frac{t^2}{2}}$ selon le résultat rappelé. On en déduit la convergence simple de $(\Phi_{\Xi_n})_{n\in\mathbb{N}}$ vers $\left(t \to e^{-\frac{t^2}{2}}\right) = \Phi_W$ d'où la convergence en loi attendue via le théorème admis juste après la question 6.

Question 9 : Décrire le comportement asymptotique de $(C_{n,n})_{n\in\mathbb{N}}$ qui peut se déduire de l'ensemble des résultats précédents.

Comme $\Xi_n \overset{L}{\to} W$, on a clairement $e^{mv\Xi_n} \overset{L}{\to} e^{mvW}$, et aussi facilement :

$$C_{n,n} \xrightarrow{L} C_0 e^{m\mu+(1-m)r-\frac{m^2v^2}{2}} e^{mvW} \overset{déf}{=} K. \tag{25}$$

On dit que K **suit une loi log-normale** pour exprimer le fait que $\ln(K) \overset{déf}{=} a+bW$ suit une loi gaussienne. Posons $(a,b) = \left(m\mu + (1-m)\,r - \frac{m^2v^2}{2}, mv\right)$. Sa densité se déduit facilement de la densité w de W. Elle vaut :

$$w_{a,b} \begin{cases} \mathbb{R} \to \mathbb{R} \\ t \to \frac{1}{b} w\left(\frac{t-a}{b}\right) = \frac{e^{-\frac{(t-a)^2}{2b^2}}}{b\sqrt{2\pi}} \end{cases} \tag{26}$$

Neuvième partie

Trafic routier

Chapitre 30
Trafic Routier

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Les candidat(e)s sont laissé(e)s libres d'organiser leur discussion comme ils ou elles l'entendent. Des suggestions de développement, largement indépendantes les unes des autres, sont proposées à la fin. Le candidat n'est pas tenu de les suivre. Le jury appréciera que la discussion soit accompagnée d'exemples effectivement traités sur ordinateur.

Après avoir indiqué les données dont on dispose, nous modélisons le trafic routier sur une portion de route sans intersections. On obtient une équation hyperbolique du type loi de conservation. L'étude des solutions est entreprise, montrant l'existence d'ondes de chocs, qui se traduisent en pratique par des embouteillages. L'équation se prête également à une étude numérique par des schémas aux différences.

§1. Modéliser la circulation automobile

1.1. Lois empiriques. On considère un tronçon de route à sens unique, qui peut être long, sans entrée ni sortie de véhicules au milieu du tronçon. Les caractéristiques du tronçon peuvent varier, x désigne la distance d'un point du tronçon au départ et t représente le temps.

Le trafic automobile s'écoule librement lorsque la vitesse de chaque véhicule est celle que désire son conducteur, sans être affectée par la présence des autres véhicules, et en tenant compte des caractéristiques du véhicule et de la route. Ceci n'est possible que lorsqu'il n'y a que très peu de véhicules et à condition qu'il y ait suffisamment de files libres pour que les dépassements puissent se faire sans délai. Lorsque le trafic se fait plus dense, les conducteurs doivent modérer leur vitesse, car ils suivent plus fréquemment un véhicule lent sans pouvoir le dépasser immédiatement. On dit que le trafic est partiellement contraint. Si la densité augmente encore, le trafic s'écoule en groupes de vitesse identique : il est contraint.

La modélisation du trafic se fait à partir de lois empiriques liant les principales quantités en jeu :

$V(x,t)$	vitesse moyenne des véhicules au point x et à l'instant t
$\phi(x,t)$	flux moyen instantané : nombre de véhicules passant au point x par unité de temps
$\chi(x,t)$	densité moyenne instantanée : nombre de véhicules par unité de longueur

Ces quantités ne sont pas indépendantes, puisque l'on a évidemment

$$\chi V = \phi \tag{1}$$

Les observations permettent d'obtenir une relation empirique entre les quantités ϕ et χ :

$$\phi = f(\chi) \tag{2}$$

où f est une fonction de $[0, \chi_{\max}]$ dans $\mathbf{R}$, concave, régulière, vérifiant $f(0) = f(\chi_{\max}) = 0$. Les deux lois (1) et (1.1) se résument dans le diagramme fondamental en figure 1.1.

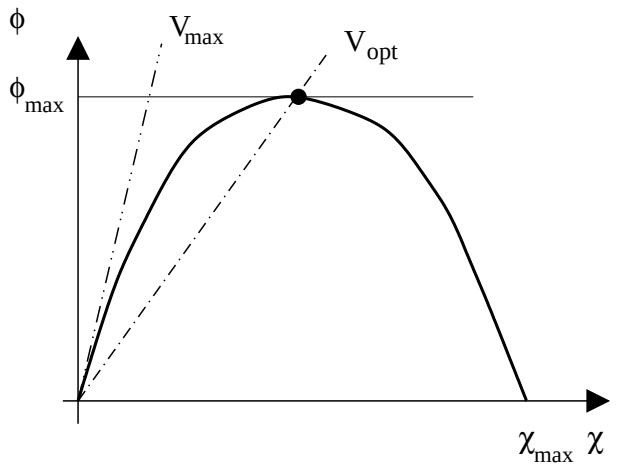


FIG. 1 – La relation fondamentale lie le débit ϕ en véhicules/heure et la χ densité en véhicules/km. Chaque point de cette courbe caractérise une situation de trafic, et la pente de la droite qui le joint à l'origine donne la vitesse moyenne des véhicules.

Pour les besoins de nos applications numériques, la fonction $\chi \to \phi$ sera approchée par un polynôme. On dispose de tables de données expérimentales caractérisant chaque portion de route. Une manière efficace de le déterminer est d'ajuster un polynôme, éventuellement sous des contraintes exprimant la concavité, par une méthode de moindres carrés.

Route A			Route B	
10	1200		10	550
18	1850		18	825
25	2500		25	900
38	2700		38	1000
50	2700		50	1100
60	2050		60	1050
80	1550		80	1060
92	1000		92	1000
100	500		100	500

TAB. 1 – Mesures densité/débit

1.2. Un modèle mathématique. On suppose maintenant que les quantités V, Φ et χ varient suffisamment régulièrement avec x et t pour que les calculs ci-aprés soient justifiés. On considère l'évolution du trafic sur un petit tronçon $[x,\, x+\Delta x]$ pendant un intervalle $[t,\, t+\Delta t]$. Il y a bien évidemment conservation du nombre de véhicules, ce qui s'exprime, en notant N_{Ve} le nombre de véhicules sur le segment :

$$N_{Ve}(\text{final}) - N_{Ve}(\text{initial}) = N_{\text{entrés}} - N_{\text{sortis}} \tag{3}$$

$$\int_x^{x+\Delta x} \left(\chi\left(\xi,\, t+\Delta t\right) - \chi\left(\xi,\, t\right)\right)\, \mathrm{d}\xi = \int_t^{t+\Delta t} \left(\phi\left(x,\tau\right) - \phi\left(x+\Delta x,\tau\right)\right)\, \mathrm{d}\tau \tag{4}$$

On en déduit, sous réserve que les fonctions soient assez régulières :

$$\int_x^{x+\Delta x}\int_t^{t+\Delta t} \frac{\partial\chi}{\partial t} + \frac{\partial\phi}{\partial x}\, \mathrm{d}\xi\mathrm{d}\tau = 0 \quad \text{donc} \quad \frac{\partial\chi}{\partial t} + \frac{\partial\phi}{\partial x} = 0 \tag{5}$$

On remarque que la relation empirique (1.1) fait que $\phi = f\left(\chi\right)$ et donc :

$$\frac{\partial\phi}{\partial x} = f'(\chi)\, \frac{\partial\chi}{\partial x} \tag{6}$$

si bien que :

$$\frac{\partial\chi}{\partial t} + f'(\chi)\, \frac{\partial\chi}{\partial x} = 0 \tag{7}$$

ou, sous la forme d'une loi de conservation :

$$\frac{\partial\chi}{\partial t} + \frac{\partial}{\partial x} f\left(\chi\right) = 0 \tag{8}$$

On peut ainsi calculer l'évolution des caractéristiques de la circulation sur une portion de route en partant de l'évolution de la densité qui est donnée par la solution du problème aux limites :

$$\frac{\partial\chi}{\partial t} + \frac{\partial}{\partial x} f\left(\chi\right) = 0 \quad \forall(x,t)\text{dans} \quad]0,L[\, \times\,]0,T[\tag{9}$$

$$\chi(x,0) = \chi_0(x) \quad \text{C.Initiale} \quad \forall x\,\text{dans}\,]0,L[\tag{10}$$

$$\chi(0,t) = \chi_{\text{Entrée}} \quad \text{C.Limite} \quad \forall t\,\text{dans}\,]0,T[\tag{11}$$

On obtient ensuite ϕ à l'aide de l'équation (1.1) et la vitesse moyenne des véhicules avec (1). D'un point de vue pratique, il n'est pas toujours très rationnel d'utiliser la densité pour spécifier l'état à l'entrée de la portion de route, mais il est possible de la déduire à l'aide du flux ϕ qui s'obtient par un simple comptage.

1.3. Limitations de la méthode des caractéristiques. L'étude de l'équation (9) s'effectue en cherchant les courbes caractéristiques $t \to (x(t), t)$ le long desquelles χ est constant. On obtient ainsi l'équation :

$$\frac{\mathrm{d}}{\mathrm{d}t}\chi\left(x\left(t\right), t\right) = 0 \quad \Rightarrow \frac{\mathrm{d}x(t)}{\mathrm{d}t} \;=\; f'(\chi) \tag{12}$$

Les caractéristiques sont donc des droites dont la pente est donnée par l'expression (12). Les caractéristiques de densité, débit, vitesse moyenne se propagent donc à la vitesse $f'(\chi)$, qui est distincte de la vitesse moyenne des véhicules $f(\chi)/\chi$. La figure 1.1 peut être utilisée pour représenter graphiquement ces deux quantités.

Cette méthode permet de déterminer χ tant que par tout point il ne passe qu'une unique courbe caractéristique. Ceci n'est valable, dans la meilleure hypothèse, que pendant un temps court qui dépend des données initiales et aux limites. Deux situations peuvent se présenter, qui mettent en défaut ce procédé de construction :

1. deux caractéristiques se rencontrent en (x_c, t_c) correspondant à deux valeurs distinctes $\chi_\pm$ de χ .
2. par un point (x, t) il ne passe aucune caractéristique.

Note 1 *Pour progresser vers la résolution de ce type de problèmes, nous allons voir que les situations ci-dessus imposent de traiter les solutions qui présentent des discontinuités. Cela se fait en définissant des solutions faibles construites dans des espaces fonctionnels adéquats, puis en montrant la convergence d'un procédé d'approximation. Ces aspects dépassent largement le cadre de cet exposé.*

Dans le premier cas, l'examen du voisinage du point (x_c, t_c) montre que deux courbes caractéristiques convergent lorsque $t \overset{-}{\to} t_c^-$ par valeurs inférieures, correspondant à des valeurs distinctes $\chi_\pm$ de χ . Il est possible de trouver des "solutions" de (9) qui propagent une discontinuité au dela de $t = t_c$ le long d'une courbe régulière. En fait ces "solutions" ne sont régulières qu'en dehors de Γ et ne vérifient l'équation aux dérivées partielles (9) qu'en dehors de cette courbe. L'exemple particulièrement simple, indiqué ci-dessous, montre qu'il n'y a pas unicité.

Example 1.1. On prend, pour fixer les idées, $f(\chi) = \chi\,(100 - \chi)$. On suppose qu'en $t = t_c$ on a :

$$\begin{cases} \chi = & 40 \quad 30 \leq x \leq 100 = L \\ \chi = & 10 \qquad 0 \leq x < 30 \end{cases} \tag{13}$$

Pour $\chi = 40$, les caractéristiques ont une pente $\frac{\mathrm{d}x}{\mathrm{d}t} = 20$, pour $\chi = 40$ $\frac{\mathrm{d}x}{\mathrm{d}t} = 80$. Il s'ensuit que pour toute valeur $20 \leq \alpha \leq 80$, la fonction

$$\chi_\alpha(x,t) = \begin{cases} 10 & \text{pour } 0 < x < 30 + \alpha t \\ 40 & \text{pour } 30 + \alpha t < x < 100 \end{cases} \tag{14}$$

est solution de (10) en dehors de la droite $x = 30 + \alpha t$. Pour juger de la pertinence de cette famille de solutions, nous pouvons essayer de vérifier le bilan (3) au temps Θ :

$$(10 - 40)\,\alpha\Theta = (900 - 2400)\Theta \quad \text{pour } 30 + \alpha\Theta \leq 100 = L \tag{15}$$

Il s'ensuit que la conservation (des véhicules) n'est vérifiée qu'à la condition $\alpha = 50$.

Note 2 *Il est facile de modifier les conditions initiales de (1.1) au voisinage de $x = 30$ pour obtenir une donnée initiale C^∞ pour laquelle la solution évoluera nécessairement vers une discontinuité.*

Ceci nous convainc qu'il n'est pas possible de se contenter de travailler sur un ensemble de solutions régulières de (9). Par contre il est souhaitable d'exprimer l'équation de conservation (3) par un autre procédé que l'équation aux dérivées partielles. Sinon, d'une part on ne peut espérer l'unicité, d'autre part les solutions ne conservent pas le nombre de véhicules en circulation, ce qui n'est pas acceptable.

1.4. Solutions discontinues. Nous étudions maintenant de manière plus détaillée une "solution" présentant une discontinuité le long d'une courbe donnée dans le plan (x,t) par le paramétrage $t \to (x_c(t), t)$. Notons $\Gamma = \{(x_c(t), t) \mid t \in [0,T]\}$. Plus précisément, nous supposons

1. la fonction est $\chi\,(x,t)$ est régulière en dehors de Γ et admet des limites à gauche et à droite de Γ notées respectivement $\chi_\pm$,
2. la fonction $t \to x_c(t)$ également régulière,
3. l'équation aux dérivée partielle (9) est vérifiée en dehors de la courbe Γ ainsi que ses conditions initiales et aux limites :

$$\tfrac{\partial\chi}{\partial t} + \tfrac{\partial}{\partial x} f\,(\chi) = 0 \quad \forall (x,t) \text{ dans }]0,L[\,\times\,]0,T[\,\setminus\,\Gamma \tag{16}$$

Fixons τ et examinons le bilan de conservation (3) dans $[x_c(\tau - \delta_t),\ x_c(\tau + \delta_t)] \times [\tau - \delta_t,\ \tau + \delta_t]$ il vient , lorsqu'on fait $\delta_t \to 0$, avec l'approximation $\delta_x \simeq x_c'(\tau)\delta_t$:

$$\tfrac{f(\chi_-) - f(\chi_+)}{\chi_- - \chi_+} = \tfrac{\delta_x}{\delta_t} = x_c'\,(\tau) \tag{17}$$

Par analogie avec la mécanique des fluides, la ligne de discontinuité est dite "onde de choc" ; à sa traversée, les conditions de trafic sont brutalement

modifiées. La relation (1.3) lie sa vitesse de propagation x'_c et la valeur de la discontinuité qui y est observée. On la désigne par "condition de saut" ou de Rankine-Hugoniot.

Complétée par (1.3), l'étude de (9) par la méthode des caractéristiques permet de traiter de manière satisfaisante divers problèmes de trafic, en particulier lorsque les conditions sont constantes par morceaux. On peut alors mettre à profit les remarques suivantes sur le diagramme fondamental (figure 1.4) pour une construction "géométrique" des solutions.

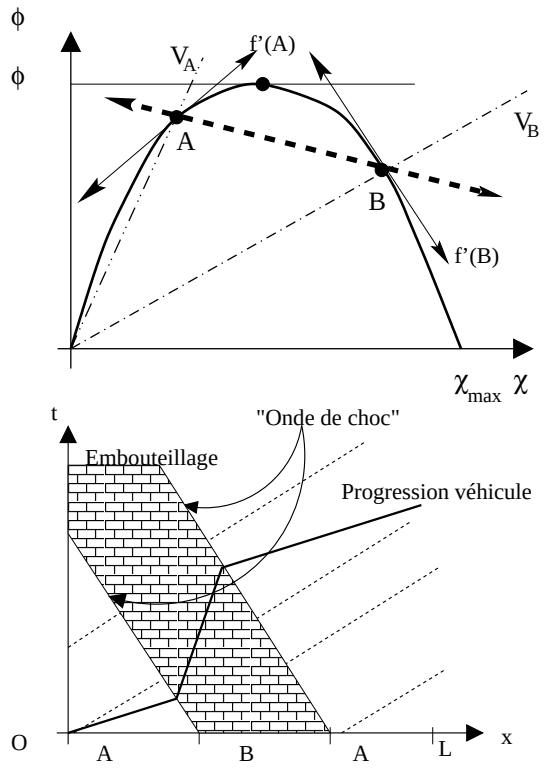


FIG. 2 – Construction Géométrique des solutions avec discontinuité. **Haut** : diagramme fondamental faisant apparaître deux régimes A et B . La pente de V_X donne la vitesse moyenne des véhicules dans le régime X , celle de $f(X)$ la pente des caractéristiques. Enfin la pente du vecteur AB est la vitesse de propagation de la discontinuité entre les régimes. Bas : évolution à partir d'une situation initiale comportant une zone B qui contient initialement une forte densité de véhicules roulant lentement. On distingue la vitesse des véhicules des vitesses de propagation des courbes caractéristiques et de chocs.

§2. Exploitation du modèle

2.1. Prédire l'évolution de la circulation routière. Avec la condition de saut, il devient possible de calculer l'évolution du trafic routier. Une difficulté à prendre en compte est l'apparition et la disparition de discontinuité. Une méthode est de calculer l'évolution en prenant en compte la propagation de toutes les discontinuités, l'apparition ou la disparition de ces discontinuités. On parle souvent de méthode d' "ajustement de chocs" ("shock fitting").

Malgré la difficulté technique, cette méthodologie permet d'analyser explicitement des situations simples telles que l'apparition d'embouteillages (cf. fig. 2), leur résorption,.... Il devient également possible de comprendre les modes de régulation du trafic (contrôle des flux d'entrée, limitation de la vitesse). Il est également important pour la pratique de vérifier à quelle vitesse les modifications de condition de trafic apparaissent au conducteurs : ceci peut être évalué à partir des données présentes sur le schéma (2) et être très supérieur à la vitesse à laquelle se déplace le véhicule.

2.2. Une modélisation purement numérique. Devant les difficultés techniques de la prédiction et du suivi des chocs dans les méthodes précédentes, il est souvent plus intéressant d'appliquer un schéma numérique à l'équation (9), qui soit capable de détecter les discontinuités. La théorie requise pour justifier cette approche est celle des solutions faibles et de la régularisation par "viscosité artificielle", et dépasse le cadre de cet exposé. Nous allons développer ci-après un premier pas dans cette direction est l'étude de schémas numériques pour l'équation (9) .

Nous discrétisons en espace en subdivisant $[0, L]$ en N segments de longueur Δx et en désignant par x_j le point de coordonnées $j\,\Delta x$. L'évolution en temps se fait par pas Δt, on note $t^n = n\,\Delta t$. Une solution approchée est définie aux points de la grille engendrée par les points $(t^n, x_j) = (n\Delta t, j\Delta x)$; la valeur approchée d'une fonction ψ en (t^n, x_j) est notée ψ_j^n. Deux schémas numériques intéressants à considérer sont :

Schéma de Lax-Friedrichs

$$\chi_j^{n+1} = \tfrac{1}{2}\left(\chi_{j+1}^n + \chi_{j-1}^n\right) - \tfrac{\Delta t}{2\,\Delta x}\left(f\left(\chi_{j+1}^n\right) - f\left(\chi_{j-1}^n\right)\right) \tag{18}$$

Schéma saute-mouton

$$\chi_j^{n+1} = \chi_j^{n-1} - \tfrac{\Delta t}{\Delta x}\left(f\left(\chi_{j+1}^n\right) - f\left(\chi_{j-1}^n\right)\right) \tag{19}$$

Ces deux schémas doivent être complétés par des conditions initiales et aux limites adaptées. Les conditions aux limites sont particulièrement délicates à spécifier, pour garder une signification : on peut imposer le flux de véhicules à l'entrée, ou ce qui revient presque au même, du fait de (1.1) la densité. A la sortie, on ne souhaite rien imposer, mais simplement observer le flot de véhicules issus de la portion de route. Nous éluderons ces difficultés potentielles en

nous plaçant dans le cadre d'une portion de route périodique pour l'étude des propriétés numériques. L'étude préliminaire de ces schémas sur des problèmes linéarisés[1] montre que :

Théorème 1 *Le schéma "Saute Mouton" (19) est stable si* $f(\chi_0)\frac{\Delta t}{\Delta x} \leq 1$ *pour le problème linéarisé en posant* $f(\chi) = f(\chi_0) + (\chi - \chi_0) f'(\chi_0)$. *En supposant la solution suffisamment régulière, il est consistant d'ordre* $\mathbf{O}(\Delta \mathbf{t}^2) + \mathbf{O}(\Delta \mathbf{x}^2)$.

La consistance peut être montrée par utilisation de la formule de Taylor. La stabilité s'étudie à partir de la base de $\mathbf{C}^{\mathbf{N}}$ des vecteurs $\left(\omega^{[k]}\right)_{k=0,\ldots,N-1}$ donnés par :

$$\omega_j^{[k]} = e^{i\,k\,j\,\frac{2\pi}{N}} \tag{20}$$

On cherche une base de l'espace des solutions de l'itération linéaire (19) de la forme :$\varphi^n\,\omega^{[k]}$, où $\varphi^n \in \mathbf{C}$. Il vient, en notant $c = f'(\chi_0)$:

$$\varphi^{n+1} = \varphi^{n-1} - 2i\,\frac{c\Delta t}{\Delta x}\sin\left(\frac{2k\pi}{N}\right)\,\varphi^n \tag{21}$$

récurrence linéaire dont la solution générale est de la forme :

$$\varphi^n = \tau_1\rho_+^n + \tau_2\rho_-^n \tag{22}$$

$$\rho_\pm = i\left(-\frac{c\Delta t}{\Delta x}\sin(\xi) \pm \sqrt{\left(\frac{c\Delta t}{\Delta x}\sin(\xi)\right)^2 - 1}\right) \tag{23}$$

en notant $\xi = \frac{2k\pi}{N}$. La condition de stabilité requiert que les solutions de (22) restent bornées, ce qui n'est possible que si $|\rho_\pm(\xi)| \leq 1$. Ceci est toujours vrai si $\left|\frac{c\Delta t}{\Delta x}\sin(\xi)\right| \leq 1$. Dans le cas $\left|\frac{c\Delta t}{\Delta x}\sin(\xi)\right| > 1$, on constate que $|\rho_\pm| > 1$ pour certaines valeurs de ξ . Exprimée de façon indépendante de k et de N, qui tend vers $+\infty$ lorsqu'on approche la solution, la condition de stabilité est finalement : $\frac{c\Delta t}{\Delta x} < 1$.

On montre que la méthode de Lax-Friedrichs est aussi stable pour le problème linéarisé.

Suggestions de développement

Modélisation : détailler le rôle dans la modélisation de la loi empirique (1.1), qui provient de mesures sur les caractéristiques du trafic, et celui des lois exprimant la conservation des véhicules.

Propriétés qualitatives : calculer des solutions correspondant à une brusque élévation du nombre de véhicules entrant sur la portion de route. par exemple, pour le cas de la route A, après avoir été longtemps constant à un

[1] C'est à dire en remplaçant $f(\chi)$ par $f(\chi_0) + (\chi - \chi_0) f'(\chi_0)$.

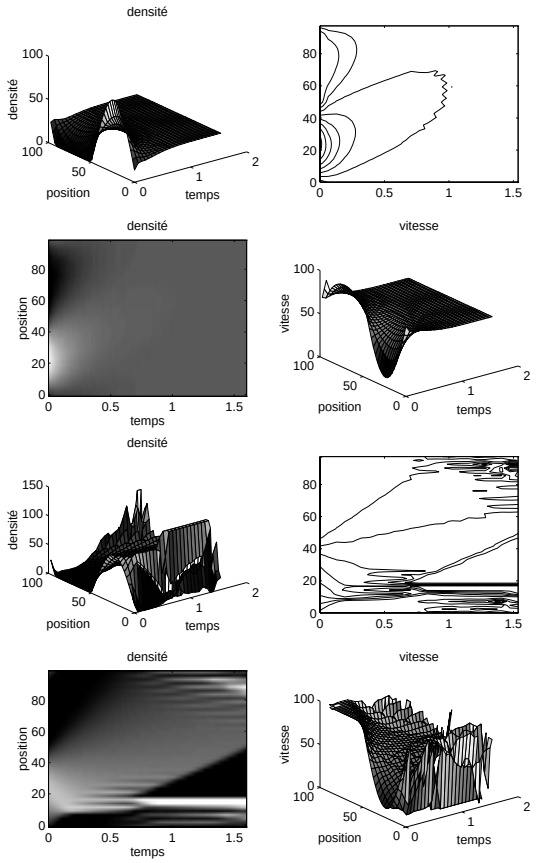


FIG. 3 – **Haut** : résultats avec méthode Lax Friedrichs. Les résultats ne sont pas satisfaisants, la solution étant rapidement lissée de manière non réaliste. **Bas** :résultats avec méthode saute mouton. Les fronts correspondant aux chocs sont ici bien visibles. Leur vitesse de propagation est réaliste.

débit de 1850 véhicule/heure (densité 18 véhicule/km), le débit d'entrée monte brusquement à 2700 v/h (densité 50 v/km). Calculer l'évolution d'un embouteillage résultant de l'introduction d'une barrière sur la route pendant 10 minutes.

Expliquer sur un exemple comment une solution régulière peut évoluer vers une solution discontinue en vous appuyant sur la méthode des caractéristiques.

Analyse numérique : mettre en oeuvre le schéma de votre choix sur le problème linéarisé. Préciser les notions de stabilité et de convergence dans le cas du problème linéarisé.

Chapitre 31
Commentaire de l'auteur

Ce texte entre dans le thème "Propagations d'Ondes" et a plusieurs objectifs pédagogiques :

- décrire une modélisation simple du trafic routier, comprenant des aspects mathématiques et pragmatiques,
- montrer que cette modélisation permet une prévision de la qualité de trafic et des temps de trajet, explicite dans des situations simples,
- montrer que les phénomènes bien connus des automobilistes de "blocage" provoquant des accidents sont en fait conséquence de la perte de régularité des solutions de notre modèle.
- aborder l'étude des schémas numériques sur un problème linéaire voisin, afin d'en faire percevoir les avantages et les difficultés.

Les données numériques fournies avec le texte permettaient d'exploiter le modèle en déduisant des résultats qualitatifs et quantitatifs suffisamment précis pour décrire plusieurs situations concrètes.

§1. La modélisation

La modélisation fait intervenir trois quantités moyennes liées : la densité de trafic, la vitesse et le flux (nombre de véhicules passant par unité de temps). Une loi empirique permet de prendre en compte aussi bien les caractéristiques de la portion de route que la nature du trafic routier : c'est la relation :

$$\phi = f(\chi) \tag{1.1}$$

Les conditions sous lesquelles cette relation peut être mesurée expérimentalement sont discutées dans Leutzbach paragraphe II.2. La dernière loi qui intervient dans notre modélisation est celle de la conservation des véhicules.

La principale loi spécifique au phénomène est (1.1) ; il est naturel que l'on s'attache à déterminer graphiquement comment elle traduit les conditions extrémales du phénomène : vitesse maximum, densité maximum, débit maximum,.... Vers la fin du texte, on montre également comment y lire les vitesses de propagation et de résorption des fronts d'embouteillages ce qui fournit un outil simple de planification approximative.

La seconde partie de la construction du modèle est mathématique : sous des hypothèses de régularité qui paraissent justifiées pour l'instant, on traduit la loi de conservation sous la forme d'une équation aux dérivées partielles. En prenant en compte les données initiales et aux limites, on aboutit au problème aux limites :

$$\begin{cases} \frac{\partial \chi}{\partial t} + \frac{\partial f(\chi)}{\partial x} = 0, \forall (x,t) \in [0,L] \times \mathbb{R}_+^*, \\ \chi(0,x) = \chi_0(x), \forall x \in [0,L], \\ \chi(t,0) = \chi_{\text{Entrée}}, \forall t \in \mathbb{R}_+^*, \end{cases} \tag{1.2}$$

Il s'agit d'une équation hyperbolique non linéaire, et l'application de la méthode des caractéristiques permet d'une part de compléter l'équation par des conditions aux limites adéquates, d'autre part d'obtenir des solutions du problème de Cauchy, au moins sur un court intervalle de temps.

Concernant ces déductions de nature mathématique, les candidats ont la possibilité de pousser plus ou moins loin la justification technique. Il nous semble que la justification des conditions aux limites à imposer pour que le problème admette une solution unique est certainement le point le plus intéressant. Ces conditions pouvaient ensuite être interprétées. L'absence de condition à la sortie du segment de route s'explique bien par les caractéristiques, mais signifie en fait que le monde extérieur a la capacité de recueillir le flux maximum sans ralentissement. Il n'influence pas autrement le trafic. Ici aussi les candidats cultivés ont la possibilité de montrer l'analogie avec d'autres phénomènes physiques présentant des "cônes de dépendance"....

Le texte n'abordait pas les difficultés inhérentes à la modélisation et à l'obtention de la relation fondamentale empirique. Il n'était donc pas demandé au candidat d'aborder les considérations relatives à la mesure d'une loi mettant en relation des quantités moyennes.

1.1. Limitations du modèle : les chocs. La première limitation de notre modèle est mathématique : si les conditions intiales et aux limites varient des discontinuités vont se créer en temps fini. L'objectif du texte, heureusement contraint par le programme de l'épreuve, n'est pas de généraliser la notion de solution pour obtenir des solutions quelques soient les données. Il est de montrer que l'étude peut être poursuivie en supposant que les solutions sont formées par raccordement de solutions régulières le long de courbes de discontinuité également régulières. Bien entendu, la compréhension de ce cas est un préliminaire à l'étude du cas général.

Il apparaît, d'abord à travers un exemple, puis dans le cas d'une discontinuité générale, que la loi de conservation des véhicules impose une relation de saut :

$$\frac{f(\chi_-) - f(\chi_+)}{\chi_- - \chi_+} = \frac{\delta_x}{\delta_t} = x_c'(\tau) \tag{1.3}$$

déterminant la pente de la courbe Γ le long de laquelle la discontinuité se propage. La situation la plus simple que nous conseillons alors d'étudier est celle ou la solution est constante par morceaux. Les courbes de discontinuités sont alors des droites, dont les pentes peuvent être obtenues facilement sur le diagramme fondamental. On pouvait espérer que les candidats consacrent un peu de temps à décrire la construction "graphique" des solutions constantes par morceaux, y compris la détermination graphique de la pente des courbes

de choc décrite par (1.3). Ceci fournit une méthode "empirique" utilisable par les praticiens et est décrite à ce titre dans [Leutzbach].

A ce point les candidats peuvent décrire complètement divers cas particulier importants : apparition et disparition d'embouteillage, vitesse de propagation des fronts les délimitant. Il est également utile de calculer la vitesse à laquelle un automobiliste vient à rencontrer l'embouteillage, qui peut être à un instant donné très supérieure à celle du trafic. La description des situations extrêmes peut se faire à l'aide de calculs simples.... et a la vertu d'inciter à la prudence sur la route.

Une situation que le texte ne détaillait pas était celle des ondes de détente, décrivant le comportement au moment ou un embouteillage se résorbe. (Voir [Serre, page 50].)

1.2. Extension du modèle. Les considérations que nous développons ici ne correspondent ni au programme de l'épreuve, ni aux exigences vis à vis des candidats. Elles permettent de situer le texte par rapport aux techniques requises pour étudier ces équations aux dérivées partielles non-linéaires. La théorie est développée dans D.Serre.

La première idée est que le modèle associé à un système physique doit être **bien posé** : il doit posséder une solution unique, définie sur un intervalle de temps suffisant, et dépendant continûment des paramètres. Le cadre doit, en outre, permettre d'évaluer le terme non linéaire.

Ceci conduit à considérer des solutions faibles dans $L^1_{loc}\left(]0,L[\times]0,T[\right)$, vérifiant l'équation au sens des distributions :

$$\int\int\left(u\,\frac{\partial\phi}{\partial t}\,+\,f\left(u\right)\frac{\partial\phi}{\partial x}\right)\,dx\,dt\;\;+\;\;\int u_0\left(x\right)\phi\left(x,0\right)\,dx\;\;=0\;\;\forall\phi\in\left\{D\right\}\;\;(1.4)$$

Il est déjà intéressant de noter qu'une solution faible présentant comme seule non régularité des discontinuités sur des courbes régulières vérifie l'équation où elle est régulière et la condition de saut en les points de discontinuité. Ceci ne suffit pas à rendre le problème de Cauchy bien posé : il faut encore imposer une condition d'entropie. On obtient alors un cadre satisfaisant, dans lequel les solutions sont de plus limite des solutions de problèmes régularisés par l'adjonction d'une viscosité artificielle.

En dehors de son intérêt théorique, ce cadre permet d'étudier complètement les solutions de notre équation sans avoir à faire d'hypothèse sur la nature des discontinuités. La condition d'entropie permet de se limiter à des solutions qui ont un sens physique. En modélisation appliquée, on s'attachera donc à l'approximation numérique de telles solutions faibles.

1.3. Etude mathématique. On va maintenant préciser ces points dans le cas où il est possible de se ramener à un problème sur tout $\mathbb{R}$. Il est clair que l'ensemble des développements de ce paragraphe sont très au delà du

programme de l'agrégation, mais certaines remarques qui contribuent à la solution étaient abordables pour les candidats.

En supposant que le flux entrant est constant dans le temps, nous nous ramenons à un problème sur tout $\mathbb{R}$:

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0, \forall (x,t) \in \mathbb{R} \times \mathbb{R}_+^*, \\ u(0,x) = u_0(x), \forall x \in \mathbb{R}, \end{array} \right. \tag{1.5}$$

où

$$u_0(x) = \left\{ \begin{array}{l} \chi_{\text{Entrée}}, \forall x \leq 0, \\ \chi_0(x), \forall x \in]0, L], \\ \chi_0(L), \forall x \geq L. \end{array} \right.$$

Alors, on a directement que la solution χ de (1.2) est la restriction de la solution u de (1.5) dans le cas où les deux problèmes sont bien posés. Examiner cette équivalence dans le cas de solutions régulières consiste essentiellement à étudier les conditions aux deux extrémités du segment, ce qui peut se faire par la méthode des caractéristiques.

Existence de solutions fortes du problème de Cauchy. Pour conclure sur l'existence de solutions fortes et éventuellement sur leur temps maximum d'existence, on utilise le résultat :

Theorem 1.1 (Temps d'explosion). *Soit u_0 dans $\mathcal{C}^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ tel que $\inf_{\mathbb{R}}(\frac{d}{dx} f' \circ u_0) > -\infty$. Alors, il existe $T^* \geq 0$ tel que (1.5) possède une unique solution dans $\mathcal{C}^1(\mathbb{R} \times [0, T^*[)$. T^*, temps d'explosion, est tel que :*

$$T^* = +\infty \text{ si } f' \circ u_0 \text{ est croissante au sens large,}$$
$$T^* = -(\inf_{\mathbb{R}}(\tfrac{d}{dx} f' \circ u_0))^{-1} \text{ sinon.}$$

Quand le temps d'explosion est fini, il est donc nécessaire de définir des solutions faibles car il n'est pas possible de prolonger les solutions fortes. Dans le problème du traffic routier , si la condition initiale est $\mathcal{C}^1(\mathbb{R})$, la solution sera de classe $\mathcal{C}^1$ jusqu'au temps T^* défini précédemment. Par exemple, pour

$$f(\chi) = V_{max}\chi(1 - \frac{\chi}{\chi_{max}}),$$

et u_0 dans $\mathcal{C}^1(\mathbb{R})$: si u_0 est une fonction décroissante au sens large, alors $T^* = +\infty$ et il n'y a pas de formation de "bouchons". Al'inverse si u_0 n'est pas décroissante, alors $T^* < +\infty$, il y a formation de "bouchons" qui apparaitront au temps T^*.Ces propriétés sont intuitives : les embouteillages ne surviendront que si un trafic rapide rattrappe un trafic plus lent. La démonstration est possible dans le cadre du programme.

Définition de solutions faibles. Pour définir les solutions faibles, il faut donner un sens aux dérivées de u dans l'équation (1.5). Pour celà, on pose :

$$\mathcal{C}_0^1(\mathbb{R}^2) = \{f \in \mathcal{C}^1(\mathbb{R}^2) \text{à} \quad \text{support} \quad \text{compact} \quad \text{dans } D^+ = \mathbb{R} \times \mathbb{R}^+\},$$

Alors, on définit les solutions faibles de (1.5) comme suit :

Definition 1.1 (Solutions faibles). *Soit u_0 dans $L^1_{loc}(\mathbb{R})$, on dira que u élément de $L^1_{loc}(D^+)$ est solution faible de (1.5) dans D^+ si :*

$$\int_{\mathbb{R}^2} (u\frac{\partial\varphi}{\partial t} + f(u)\frac{\partial\varphi}{\partial x})dx\ dt + \int_{\mathbb{R}} u_0(x)\varphi(x,0)dx,\ \forall\varphi \in \mathcal{C}_0^1(\mathbb{R}^2).$$

Cette définition des solutions assure l'existence mais non l'unicité des solutions, ainsi que de la relation de conservation dite de Rankine-Hugoniot (1.3) (Voir [Serre, page]). Pour obtenir l'unicité, c'est une notion d'entropie qui permet de sélectionner les solutions. La difficulté entraînant la non unicité est illustrée par l'existence de solutions de type "onde de détente".

1.4. Choix des chocs. L'un des moyens de discriminer les solutions est de considérer les solutions de (1.5) limites singulières des solutions d'un problème sans chocs, dit régularisé par viscosité :

$$\begin{cases} \frac{\partial u_\varepsilon}{\partial t} + \frac{\partial f(u_\varepsilon)}{\partial x} = \varepsilon\frac{\partial^2 u_\varepsilon}{\partial x^2}, \forall(x,t) \in \mathbb{R} \times \mathbb{R}_+^*, \\ u_\varepsilon(0,x) = u_0(x), \forall x \in \mathbb{R}, \end{cases} \tag{1.6}$$

où ε est un réel strictement positif.
On peut alors démontrer le lemme suivant :

Theorem 1.2. *Pour u_0 de classe $\mathcal{C}^1$ et bornée. Si $u_\varepsilon(x,t)$ solution de (1.6) tend vers $u(x,t)$ pour presque tout (x,t) dans $\mathbb{R}\times]0,T[$ quand ε tend vers 0, alors u est solution faible du problème (1.5).*

On peut démontrer que le problème (1.6) admet une unique solution forte vérifiant le principe du maximum. Alors, en combinant ce résultat avec le lemme précédent, on obtient un critère de choix des solutions du problème (1.5). On nomera ces solutions de (1.5) les solutions entropiques. Leur existence et unicité est donnée par le théorème suivant [Kružkov]

Theorem 1.3. *Pour toute fonction mesurable bornée u_0 sur $\mathbb{R}$, il existe une et une seule solution entropique de (1.5) dans $L^\infty(\mathbb{R}\times]0,T[)\cap\mathcal{C}([0,T[; L^1_{loc}(\mathbb{R}))$. Elle vérifie le principe du maximum :*

$$||u||_{L^\infty(\mathbb{R}\times]0,T[)} = ||u_0||_{L^\infty(\mathbb{R}\times]0,T[)}.$$

On montre alors qu'une solution de classe $\mathcal{C}^1$ par morceaux dans $\mathbb{R} \times [0,T[$ et vérifiant la condition de Rankine-Hugoniot sur les courbes ou elle n'est pas classe $\mathcal{C}^1$ est une solution entropique par conséquent unique en vertu de ce qui a été dit plus haut. La condition de Rankine Hugoniot combinée avec de Lax peut-être déduite de (1.6) en passant à la limite, mais c'est une démonstration technique très difficile pour laquelle on pourra consulter [Smoller].

§2. Modèle Numérique

Le calcul numérique des solutions avec discontinuité est difficile à mettre en oeuvre dans la pratique, en dehors de situations simples. En effet, il faut simultanément déterminer la solution et les courbes de discontinuité. Une autre approche est celle de rechercher la solution par un schéma numérique dont on espère qu'il converge en un sens adéquat vers une solution faible. L'étude générale dépasse largement le niveau du texte, ne serait-ce que par la nécéssité de trouver un cadre fonctionnel pour décrire des solutions non continues.

Le texte prend le parti de se limiter à une équation linéarisée et de la traiter à l'aide de deux schémas classiques aux différences finies. Après avoir étudié rapidement la stabilité de deux schémas classiques, on montre les résultats numériques obtenus. Ceci donne aux candidats l'occasion de montrer aussi bien leur aptitude à compléter les preuves de stabilité, qu'à s'essayer à la mise en oeuvre de ces schémas.

Le texte montre les solutions obtenues numériquement pour le système non linéaire : seule la méthode saute-mouton permet de percevoir les chocs. Ici aussi le filtrage rapide des discontinuités par la méthode de Lax-Friedrichs peut être lié aux propriétés spectrales de la matrice du schéma. Ceci permettait aux candidats d'ajouter des remarques mettant en relation ces propriétés spectrales et le comportement attendu ou effectif des schémas numériques.

2.1. Classification des schémas numériques. Il est possible d'utiliser deux types de schémas pour résoudre ce type problèmes : des schémas dispersif ou des schémas dissipatif. Les deux exemples de schémas figurant dans le texte illustrent ces deux catégories.

Shémas dissipatifs : ces schémas sont en réalité une approximation du système (1.6) ; ils vérifient le principe du maximum discret et donnent une bonne approximation de la forme des solutions. Le schéma de Lax-Friedrichs appartient à cette catégorie.

Schémas dispersifs : ils ne vérifiens pas le pricipe du maximum mais donne une meilleure approximation que les schémas précédents de la position du choc. Le schéma saute-mouton est de ce type.

2.2. Perspectives sur l'approche numérique. L'étude des méthodes numériques pour de telles équations montre de grandes difficultés. Il faut aussi bien se garder des instabilités que des méthodes numériques qui ne permettent pas un bon rendu des chocs. Les critères qu'il est possible d'appliquer pour construire ou sélectionner des schémas comprennent :

- stabilité et convergence,
- qualité de l'approximation des chocs,
- satisfaction de la loi de conservation, de la condition d'entropie,
- borne de la variation totale des solutions à t fixé,

– nature de la viscosité artificielle introduite par le schéma,
– efficacité (coût des calculs)

Une discussion approfondie est faite dans J.W. Thomas. Il s'agit d'un domaine extrèmement actif de l'analyse numérique moderne en raison de l'importance pratique des systèmes de lois de conservation.

§3. Bibliographie

1. R.Courant and K.O. Friedrichs, Supersonic Flow and Shock Waves, Interscience, 1948
2. Fritz John, Partial Differential Equations,Applied Math. Sciences, Springer, 1971
3. Wilhelm Leutzbach, Introduction to the Theory of Trafic Flow, Springer, 1988
4. Denis Serre, Systèmes de lois de conservation, vol 1,
5. J.W. Thomas, Numerical Partial Differential Equations, Conservation Laws and Elliptic Equations, TAM Vol 33, Springer, 1999
6. J.W. Thomas, Numerical Partial Differential Equations, Finite Difference Methods, TAM Vol 22, Springer, 1995
7. S. N. Kružkov. Generalized solutions of the Cauchy problem in the large for nonlinear equations of first order. Dokl. Akad. Nauk SSSR, 187 :29-32, 1969
8. J.Smoller, Shock Waves and Reaction Diffusion Equations, Springer ,1983,

Chapitre 32
Exploitation du texte pour une Analyse de Texte Scientifique (ADS)

§1. Plan

J'ai pris le parti dans ce commentaire d'utiliser ce texte en tant que texte d'ADS, c'est-à-dire d'analyse d'un document scientifique, telle que définie pour les concours d'entrée dans les écoles d'ingénieurs. Rappelons que les étudiants des classes préparatoires qui passent cette épreuve disposent en général d'un peu plus de deux heures pour préparer leur exposé oral, ainsi que de transparents et d'un rétroprojecteur pour l'exposition, mais bien évidemment d'aucun outil informatique.

§2. Proposition d'exposition de la part du candidat

Pendant le temps de préparation on peut suggérer au candidat de préparer l'exposé de deux des trois propositions suivantes.

a) *Présentation de la modélisation.* Le candidat introduit avec soin les différentes notions : vitesse moyenne des véhicules au point x et à l'instant t, flux moyen instantané et densité moyenne instantanée, ainsi que la relation entre ces trois quantités, éventuellement à l'aide d'un schéma, puis la loi empirique liant le flux et la densité. Ensuite, il pourra établir l'équation aux dérivées partielles vérifiée par la fonction densité.

b) *Présentation de la méthode des caractéristiques.* Le candidat expose la méthode des caractéristiques de densité, en détaillant la recherche des équations des courbes et montre comment on peut, moyennant certaines hypothèses sur les courbes caractéristiques, revenir à la fonction densité.

c) *Présentation de la partie numérique.* Le candidat explique la notion de discrétisation, puis détaille les deux schémas numériques. Il conviendrait donc d'ajouter au document une annexe présentant les notions classiques de stabilité, consistance et convergence d'un schéma, que ne connaît pas un étudiant de classe préparatoire.

§3. Propositions de questions

Le jury doit vérifier que le candidat maîtrise le vocabulaire qu'il emploie, qu'il comprend le lien entre le problème exposé et la mécanique des fluides,

même si celle-ci ne figure pas au programme de physique de toutes les classes préparatoires ; il l'interroge ensuite sur ses choix d'exposition en ne perdant pas de vue le fait que cet article reste difficile pour un étudiant de classe préparatoire.

§4. Conclusion

Dans l'hypothèse d'une ADS, l'article peut être repris soit en ôtant le paragraphe sur les solutions discontinues, soit en ôtant le paragraphe sur les méthodes numériques. Dans les deux cas, il conviendrait de préciser davantage à l'aide d'une annexe au moins soit la question des méthodes numériques, soit la question des caractéristiques. On peut aussi imaginer, dans la mesure où le sujet de cet article pourrait un jour relever du thème[1] des TIPE (si celui-ci évoquait les équations aux dérivées partielles, par exemple), de se servir de cet article tel quel comme point de départ d'un travail personnel. Il serait alors possible de construire une bibliographie pour éclairer les points délicats le concernant, puis de développer la partie numérique sur ordinateur.

[1] Rappelons que le thème des TIPE est imposé et change tous les deux ans.

Chapitre 33
Commentaire d'un non spécialiste

§1. Position du problème

L'objet de l'étude est de modéliser le trafic routier sur une portion de route sans intersection, modélisée par un intervalle $[0, L]$. On introduit une loi de conservation et une relation empirique entre le flux de véhicules φ et leur densité χ. Grâce à ce modèle, on peut donner une équation aux dérivées partielles vérifiée par la densité de véhicules. Il devient alors possible d'estimer une vitesse optimale sur la route considérée et d'étudier la vitesse de propagation d'un bouchon (dû à un accident par exemple) en fonction de la densité de véhicules à l'instant de l'évènement.

Dans cette étude, on va surtout s'intéresser à la modélisation mathématique du phénomène plus qu'aux estimations qui en dérivent (celles qu'on vient d'évoquer). Le premier élément à modéliser est la relation empirique $\varphi = f(\chi)$ décrivant le comportement des conducteurs confrontés à une certaine densité de véhicules, i.e. décrivant la vitesse moyenne (φ/χ) adoptée lors d'un certain trafic (χ). Cette relation est obtenue à partir d'observations ponctuelles, sujettes à erreur, et il est raisonnable d'adopter une approximation de f par la méthode des moindres carrés. Néanmoins on a besoin d'un certain nombre de propriétés de f afin de refléter la réalité et de garantir le comportement de $\chi(x, t)$ déduit de l'équation aux dérivées partielles. Ici x est la variable d'espace ($x \in [0, L]$) et t la variable temporelle ($t \in \mathbf{R}_+$).

La densité est un réel positif, inférieur à une certaine densité χ_{max} correspondant à une impossibilité de mouvement. La fonction f est donc définie sur l'intervalle $[0, \chi_{max}]$, nulle aux bornes : sans véhicule ou lorsqu'ils se touchent tous, il n'y a pas de circulation.

De plus φ/χ représente la vitesse instantanée des véhicules (au point x et à l'instant t). C'est *a priori* une fonction décroissante de χ. Puisque φ/χ est la pente entre le point $(0, f(0))$ et le point $(\chi, f(\chi))$, il semble raisonnable, afin de garantir cette décroissance, d'imposer la concavité de f.

Au final on cherche donc, à partir de données observées, à reconstituer une fonction sur un intervalle $[0, \chi_{max}]$ qui soit dérivable, nulle aux bornes et concave. La méthode des moindres carrés pour donner une approximation polynomiale rencontre alors quelques problèmes : la concavité est facile à obtenir en degré petit, mais la nullité aux bornes ne s'obtient qu'en degré élevé. En particulier, plus l'approximation est bonne, plus on est sûr de ne pas obtenir une approximation concave.

§2. Implémentation du schéma numérique

Pour obtenir une équation vérifiée par χ, on écrit la conservation du nombre de véhicules. Celle-ci s'écrit, en considérant le tronçon de route $[x, x+dx]$ pendant l'intervalle de temps $[t, t+dt]$: le nombre de véhicules sur le tronçon est $\chi.dx$ et sa variation pendant l'intervalle de temps est donc $\partial\chi/\partial t.dx.dt$. Cet accroissement est la différence entre le nombre de véhicules entrant et sortant. Le nombre de véhicules passant en un point pendant l'intervalle de temps est $\chi.dt$ et donc la différence entre le nombre de véhicules entrant (en x) et sortant (en $x+dx$) est $-\partial\chi/\partial x.dt.dx$. Il vient donc

$$\frac{\partial\chi}{\partial t} + \frac{\partial\varphi}{\partial x} = 0$$

ou encore, en écrivant $\varphi = f(\chi)$,

$$\frac{\partial\chi}{\partial t} + \frac{\partial f(\chi)}{\partial x} = 0\ .$$

Pour implémenter un schéma numérique, on fait un maillage de pas Δx (diviseur entier de L) en espace et de pas Δt en temps. On impose la densité initiale pour $x = 0$ et la densité limite pour $t = 0$. Le schéma choisi est, en notant, pour j entier entre 0 et $j_{max} = L/\Delta x$ et pour n entier positif, $\chi(n\Delta t, j\Delta x) = \chi_j^n$,

$$\Delta x\left(\chi_j^{n+1} - \chi_j^{n-1}\right) + \Delta t\left(f(\chi_{j+1}^n) - f(\chi_{j-1}^n)\right) = 0$$

justifié par la formule de Taylor : pour une fonction g dérivable en un point a, on a $g(a+\Delta a) - g(a-\Delta a) \sim 2\Delta a.g'(a)$.

On se donne χ_0^n la densité entrante et χ_j^0 la densité initiale. Numériquement, on s'intéresse à l'évolution en temps, pas par pas, et ce en calculant la densité à chaque point du maillage, en partant de l'entrée, pas par pas.

Pour résoudre les problèmes limites, on amende le schéma numérique de la façon suivante (justifiée également par la formule de Taylor) :

$$\left\{\begin{array}{l} \chi_j^1 = \chi_j^0 - \frac{\Delta t}{2\Delta x}\left(f(\chi_{j+1}^0) - f(\chi_{j-1}^0)\right) \\ \chi_{j_{max}}^{n+1} = \chi_{j_{max}}^{n-1} - \frac{2\Delta t}{\Delta x}\left(f(\chi_{j_{max}}^n) - f(\chi_{j_{max}-1}^n)\right) \end{array}\right. .$$

Pour le traitement numérique, on a choisi $L = 1$, $\Delta x = \Delta t = 10^{-2}$,

$$\chi_{max} = 110\ , \qquad \chi^0(x) = 70e^{x-L} \quad \text{et} \quad \chi_0(t) = 50\left(1 + \sin\left(2\pi t\right)\right)\ .$$

Il reste donc à modéliser f. Ce sera le centre de notre étude.

§3. Approximation polynomiale

Le comportement d'une route est modélisé par une réponse à une densité donnée par un certain flux. On l'étudie en se donnant un certain nombre

d'observations. Ainsi on a m observations $(\chi_k, \varphi_k)_{1 \leq k \leq m}$ et on veut en déduire une relation $\varphi = f(\chi)$ où f est une fonction dérivable sur $[0, \chi_{max}]$ ayant les propriétés suivantes :

1. $f(0) = f(\chi_{max}) = 0$,
2. f est concave sur $[0, \chi_{max}]$.

Une façon de procéder est d'approcher le plus possible les données par des polynômes ou des fonctions polynomiales par morceaux (ici il faudrait utiliser la méthode d'Hermite pour garantir la dérivabilité aux points de jonction). On peut traduire les deux conditions précédentes par

1. P est divisible par $X(\chi_{max} - X)$,
2. P'' est négative sur l'intervalle $[0, \chi_{max}]$.

Ces contraintes définissent un ensemble convexe inclus dans l'espace vectoriel $\mathbf{R}[X]$ des polynômes à coefficients réels. On peut, si on le souhaite, se placer dans un espace de dimension finie en bornant le degré.

Se pose la question de savoir comment on mesure la pertinence de l'approximation. Le procédé des moindres carrés consiste à minimiser la somme des carrés des écarts verticaux entre la fonction cherchée (ici un polynôme ayant un certain nombre de propriétés) et les données observées, i.e. à minimiser

$$\sum_{k=1}^{m} (\varphi_k - P(\chi_k))^2 \ .$$

Cette démarche consiste à supposer que les observations φ_k sont obtenues à partir des données réelles (et déterministes) χ_k par des perturbations gaussiennes homoscédatiques (de même variance) et à adopter le principe du maximum de vraisemblance.

Soit ν la semi-norme définie sur $C([0, \chi_{max}], \mathbf{R})$, espace des fonctions continues à valeurs réelles définies sur $[0, \chi_{max}]$, par

$$\nu(f) = \sqrt{\sum_{i=1}^{m} f(\chi_i)^2} \ .$$

Le problème est donc de minimiser la semi-distance des données (vues comme les valeurs d'une fonction) à un polynôme. Il n'y a évidemment pas, en général, unicité du point où la semi-distance est minimale.

Si on se restreint à l'espace E_m des polynômes de degré inférieur strictement à m, ν induit une norme et le problème admet une unique solution, donnée par la méthode des moindres carrés. De plus tout polynôme P admet une unique «projection» sur E_m au sens qu'il existe un unique polynôme dans E_m dont la semi-distance à P est nulle.

Soit maintenant n un entier naturel (sans relation avec m a priori) et E_n l'espace vectoriel des polynômes à coefficients réels de degré strictement inférieur à n. On note C_n le convexe de E_n formé des polynômes vérifiant les

deux conditions énoncées précédemment et on introduit Γ_n la projection de C_n dans E_m. Si n est inférieur ou égal à m, on a $C_n = \Gamma_n = \Gamma_m \cap E_n$ mais dans le cas contraire Γ_n peut contenir strictement Γ_m.

En tout cas Γ_n est un convexe fermé de E_m et on peut considérer la projection de P sur Γ_n, où P est l'unique polynôme de E_m tel que $P(\chi_k) = \varphi_k$ pour tout entier k entre 1 et m. Ce polynôme admet un représentant ayant les deux propriétés souhaitées et tout tel représentant est une solution du problème posé. On ne s'occupera pas ici du cas $n > m$, mais il est intéressant, contrairement au cas de l'approximation sans contrainte.

§4. Approximation polynomiale avec contraintes

L'espace vectoriel E_m muni de la norme ν est un espace euclidien et la condition $P(0) = P(\chi_{max}) = 0$ y définit un sous-espace H_m de co-dimension 2. Pour la suite de l'exposé, on supposera que 0 et χ_{max} font partie des données observées. Ainsi le plan orthogonal à H_m est le plan défini par $P(\chi_k) = 0$ pour χ_k distinct de 0 et χ_{max}. Autrement dit la projection d'un polynôme P sur H_m n'est rien d'autre que l'unique polynôme de E_m s'annulant en 0 et χ_{max} et prenant les mêmes valeurs que P en tous les autres points de contrôle.

La condition P concave sur $[0, \chi_{max}]$ détermine un convexe X de E_m. Pour simplifier on va ici traiter les deux conditions séparément et décrire X plutôt que Γ_m, tout en indiquant la procédure à suivre pour cette dernière étude.

Afin de mieux comprendre la projection en terme de la norme ν, il convient de décrire X dans une base orthonormée pour ν. Pour cela on orthonormalise la base canonique de E_m (ou $X(\chi_{max} - X)$ fois la base canonique de E_{m-2} dans le cas de Γ_m) et on écrit donc tout polynôme de E_m sous la forme

$$P = \sum_{k=0}^{m-1} a_k G_k$$

où $(G_0, \ldots, G_{m-1})$ est une suite étagée et orthonormée de polynômes à coefficients dominants positifs.

Soit Q un polynôme. La condition Q négatif sur $[0, \chi_{max}]$ peut s'écrire avec un nombre fini de conditions. En effet cette condition est équivalente au fait que Q est négatif aux bornes et en tous ses extrema intérieurs. Autrement dit

$$\forall x \in [0, \chi_{max}]\ Q(x) \leq 0 \Leftrightarrow \begin{cases} Q(0) \leq 0 \\ Q(\chi_{max}) \leq 0 \\ \forall x \in [0, \chi_{max}]\ Q'(x) = 0 \Rightarrow Q(x) \leq 0\ . \end{cases}$$

Remarque : ceci exhibe X et Γ_m comme des convexes fermés semi-algébriques.

On peut également formuler le problème de façon légèrement différente, en utlisant les suites de Sturm, bien adaptées au calcul formel. En effet un

polynôme est strictement négatif sur un intervalle si et seulement s'il est strictement négatif en un point et n'admet aucune racine sur l'intervalle. Cette dernière condition se teste grâce au théorème de Sturm.

Traitons un exemple : soit $Q(X) = aX^2 + bX + c$ sur $[0, 1]$, le premier point de vue s'écrit

$$c \leq 0 \ , \qquad a + b + c \leq 0 \qquad \text{et} \quad 0 < -\tfrac{b}{2a} < 1 \Rightarrow c - \tfrac{b^2}{4a} \leq 0; \ ,$$

tandis que le second s'écrit

$$c \leq 0 \qquad \text{et} \qquad (c, -b, \tfrac{b^2}{4a} - c) \simeq (a + b + c, -2a - b, \tfrac{b^2}{4a} - c) \ ,$$

où le symbole $\simeq$ signifie que les deux suites changent autant de fois de signe.

Par parité du nombre de changements de signe, la seconde condition impose $a+b+c$ négatif et il n'y a pas d'autre condition si $b^2/4a - c > 0$. Si au contraire $b^2/4a - c < 0$, la seconde condition est équivalente à ce que b et $b + 2a$ soient de même signe. Comme $c \leq 0$ et $b^2/4a - c < 0$ imposent que a soit négatif, b et $b + 2a$ sont de même signe si et seulement si b est négatif ou supérieur à $-2a$. On trouve donc la contraposée de l'implication donnée par le premier point de vue : les suites de Sturm montrent que Q est négatif sur $[0, 1]$ si et seulement si

$$c \leq 0 \ , \qquad a + b + c \leq 0 \quad \text{et} \quad \tfrac{b^2}{4a} - c < 0 \Rightarrow (b \leq 0 \text{ ou } b \geq -2a) \ .$$

Si l'on veut être rigoureux, pour traiter la concavité au sens large, il faut écrire : P est négatif sur $[0, \chi_{max}]$ si et seulement si P est négatif ou nul en m points fixés (les χ_k ou les $k\chi_{max}/(m-1)$) de façon à garantir la stricte négativité en au moins un point, et P n'a que des zéros d'ordre pair sur $[0, \chi_{max}]$. Pour tester cette dernière condition, on cherche les racines multiples de P (grâce au résultant de P et P') et on teste leur parité. Si tout va bien, on considère P divisé par tous ses facteurs multiples et on demande que ce quotient soit sans zéro sur $]0, \chi_{max}[$ grâce aux suites de Sturm.

Si l'on veut obtenir la meilleure approximation polynomiale concave en degré n (par exemple 3 ou 4), il faut donc procéder ainsi :

1. Ecrire le polynôme approximateur au sens des moindres carrés, de degré n.
2. Ecrire ses coefficients dans la base orthonormée pour ν obtenue par le procédé de Gram-Schmidt à partir de la base canonique de E_n.
3. Ecrire la condition de concavité en fonction des coefficients d'un polynôme écrit dans cette base orthonormée.
4. Projeter.

C'est a priori la projection le plus difficile. En degré 3, la dérivée seconde est de degré 1 et on obtient une condition simple portant sur les deux coefficients de plus haut degré, imposant qu'ils soient dans l'intersection de deux demi-plans. Le projeté d'un point extérieur à ce convexe est alors soit le projeté sur

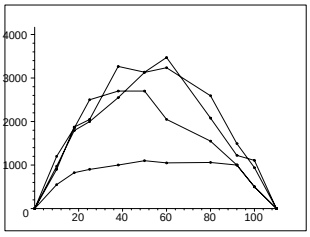
l'une des deux droites frontière, soit sur le point d'intersection de ces deux droites.

En degré 4, la condition porte sur trois coefficients et les calculs précédents montrent que les faces extérieures du convexe sont deux plans et un paraboloïde hyperbolique. La projection sur un tel convexe n'est pas simple et on peut se contenter d'une méthode algorithmique.

§5. Code Maple (Maple 6)

5.1. Mise en place des données de modélisation.

```
> with(linalg) : with(plots) :
> perturb := proc(a) 1+(rand(1001)-500)/(500*a) ; end proc :
> χ := [0, 10, 18, 25, 38, 50, 60, 80, 92, 100, 110] :
> m := nops(χ) ;
> route1 := [0, 1200, 1850, 2500, 2700, 2700, 2050, 1550, 1000,
  500, 0] :
> route2 := [0, 550, 825, 900, 1000, 1100, 1050, 1060, 1000, 500,
  0] :
> p :=perturb(10) : route3 := [seq(χ[i]*(110-χ[i])*p(), i=1..m)] :
> p :=perturb(3) : route4 := [seq(χ[i]*(110-χ[i])*p(), i=1..m)] :
> nb := 4 ;
> ChiMax := max(op(χ)) : xrange := 0..ChiMax ;
> yrange := seq(0..1.2*max(op(route||k)),k=1..nb) ;
> dup := proc(m,s) local i,d ; d := floor(m/nops(s))+1 ; [seq(op(s),
  i=1..d)] ; end proc :
> couleur := blue, red, yellow, green : ligne := 1 :
> clist := dup(m,[couleur]) : slist :=dup(m,[ligne]) :
> liste := seq( [seq( [χ[i],route||k[i]], i=1..m) ], k=1..nb) :
> dplot := [seq( plot([liste[k],liste[k]], xrange, yrange[k],
  symbol=circle, style=[point, line], colour=clist[k]), k=1..nb)] :
> display(dplot) ;
```



5.2. Calcul des approximations, sans contrainte.

```
> MC := proc(k,n) global χ ; local A, Y, a, P, i ; if n
  >= m then n := m-1 ; end if ; A := matrix(m,n+1, [seq(
  [seq(χ[i]ĵ, j=0..n)], i=1..m)]) ; Y := vector(m, route||k) ; a :=
  linsolve(transpose(A)&*A, transpose(A)&*Y) ; P := sum(a[i+1]*xî,
  i=0..n) ; end proc :
> for n from 0 to m-1 do PolMC||n := seq(MC(k,n), k=1..nb) end do :
> for n from 0 to m-1 do MCplot||n := seq(display([dplot[k],
  plot(PolMC||n[k], x=xrange, yrange[k], colour=clist[n+1])]),
  k=1..nb) end do :
> anime := proc(k) local n ; display( [seq(MCplot||n[k], n=0..m-1)],
  insequence=true) ; end proc :
> anime(1) : anime(2) : anime(3) : anime(4) :
> ε :=10 ;
> test_0 := proc(k) local t,n,e ; if nargs = 1 then e := ε else
  e := args[2] end if ; t := NULL ; for n from 0 to m-1 do if
  evalb(subs(x=0, PolMC||n[k]) <= e) then t := t, n ; end if ; end
  do ; [t] ; end proc :
> seq(test_0(k), k=1..nb) ;
```

$$[4, 5, 7, 8, 9, 10], [4, 5, 6, 7, 8, 9, 10],$$

$$[2, 3, 5, 6, 7, 8, 9, 10], [2, 3, 4, 6, 7, 8, 9, 10]$$

```
> test_max := proc(k) local t, n, e ; if nargs = 1 then e := ε
  else e :=args[2] end if ; t := NULL ; for n from 0 to m-1 do if
  evalb(subs(x=ChiMax, PolMC||n[k]) <= e) then t := t, n ; end if ;
  end do ; [t] ; end proc :
> seq(test_max(k), k=1..nb) ;
```

$$[2, 4, 5, 6, 7, 8, 9, 10], [4, 5, 6, 7, 8, 9, 10],$$

$$[2, 3, 4, 7, 8, 9, 10], [2, 3, 7, 8, 9, 10]$$

```
> test_concave := proc(P,x,a,b) local Q, n, t, i ; if degree(P,x)
  <= 1 then RETURN(true) ; end if ; Q := diff(P, x £2) ;
  evalb((subs(x=a, Q)< 0) and (sturm(sturmseq(Q, x), x, a, b) =
  0)) ; end proc :
> test_PolMC := proc(k) local t, n, pb ; t := NULL ; pb := -1 ; for n
  from 0 to m-1 do if test_concave(PolMC||n[k], x, 0, ChiMax) then
  t := t, n else if pb = -1 then pb := n ; end if ; end if ; end do ;
  [t], pb ; end proc :
> resultat := seq(test_PolMC(k), k=1..nb) : concave :=
  seq(resultat[2*k-1], k=1..nb) ; pb := seq(resultat[2*k],
  k=1..nb) ;
```

$$concave := [0, 1, 2, 4], [0, 1, 2, 3, 4], [0, 1, 2, 3, 4], [0, 1, 2, 3]$$

$$pb := 3, 5, 5, 4$$

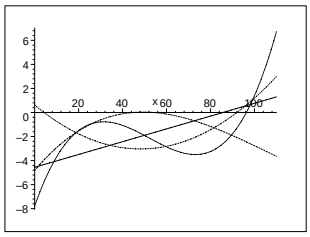
La concavité est proche de celle des données (qui le sont presque !) en bas degré. Plus on est exigeant en degré et moins la concavité est préservée. C'est un phénomène important dans l'approximation polynomiale qui se retrouve par exemple dans le phénomène de Runge.

```
> for k to nb do P||k := PolMC||(pb[k])[k] end do :
> Digits := 2 : for k to nb do evalf(P||k) end do ; Digits := 10 :
```

$$11. + 140.\, x - 2.3\, x^2 + .0089\, x^3$$

$$-4.5 + 78.\, x - 2.4\, x^2 + .037\, x^3 - .00024\, x^4 + .45\, 10^{-6}\, x^5$$
$$-13. + 140.\, x - 3.9\, x^2 + .087\, x^3 - .00099\, x^4 + .38\, 10^{-5}\, x^5$$
$$-24. + 98.\, x + .31\, x^2 - .025\, x^3 + .00013\, x^4$$

```
> plot([seq(diff(P||k, x £2), k=1..nb)], x=xrange, linestyle=slist,
  colour=clist) ;
```



5.3. Obtention d'une approximation concave.

Polynômes orthogonaux.

```
> prod_scal := proc(P, Q, x) global χ ; local i ; sum(subs(x=χ[i],
  P*Q), i=1..m) end proc :
> norme := proc(P, x) prod_scal(P, P, x)^(1/2) end proc :
```

On peut ainsi évaluer les résidus des approximations au sens des moindres carrés :

```
> Digits :=2 : seq( [seq( evalf( norme(PolMC||n[k]-PolMC||(m-1)[k],
  x)), n=0..m-1)], k=1..n b) ; Digits := 10 :
```

[3000., 3000., 1000., 420., 370., 330., 290., 240., 230., 230., 0.],

[1400., 1300., 420., 420., 170., 170., 160., 130., 8.9, 6.5, 0.],

[3600., 3600., 450., 440., 400., 280., 280., 170., 170., 3.1, 0.],

[3700., 3600., 830., 750., 620., 610., 550., 480., 460., 460., 0.]

```
normé :=proc(P,x) P/norme(P,x) end proc :
```

La procédure base(x,n) renvoie l'orthonormalisée de Gram-Schmidt de la base canonique des polynômes en x de degré inférieur ou égal à n. Si on rajoute un argument (comme "f" pour "flottant"), la procédure renvoie un polynôme à coefficients flottants.

```
> base := proc(x,n) local schmidt, i, j, P, ps ; for i from 0 to
  n do P := xî ; if nargs = 3 then P := evalf(P) end if ; if i >
  0 then for j from 0 to i-1 do ps := prod_scal(P, schmidt||j,
  x) ; P := P - ps*schmidt||j end do ; end if ; schmidt||i :=
  simplify(normé(P, x)) ; end do ; seq(schmidt||i, i=0..n) end proc :
> base(x,1) ;
```

$$\frac{1}{11}\sqrt{11},\ \frac{1}{14658}\,(x-53)\,\sqrt{14658}$$

```
> base(x,2,"f") ;
```

$$.3015113446,\ .008259668953\,x-.4377624546,$$

$$.0002901955036\,x^2+.5037442855-.03218117274\,x$$

On procède de même avec les polynômes s'annulant aux bornes, en partant de la base $(\chi_{max}-X)X^{i+1}$

```
> base_lim := proc(x, n) local schmidt, i, j, P, ps ; for i from 0
  to n do P := (ChiMax-x)*x^(i+1) ; if nargs = 3 then P := evalf(P)
  end if ; if i > 0 then for j from 0 to i-1 do ps := prod_scal(P,
  schmidt||j, x) ; P := P-ps*schmidt||j end do ; end if ; schmidt||i :=
  simplify(normé(P, x)) ; end do ; seq(schmidt||i, i=0..n) end proc :
> base_lim(x,0) ;
```

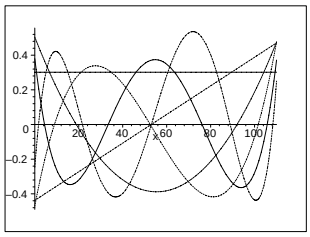
$$-\frac{1}{43245993}\,(-110+x)\,x\,\sqrt{43245993}$$

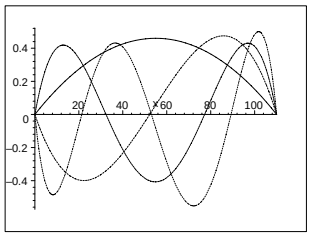
```
> base_lim(x,1,"f") ;
```

$$-.0001520642286\,(-110.+x)\,x,$$

$$.001106310225\,x^2-.6818357786\,10^{-5}\,x^3-.03919199559\,x$$

```
> GS := [base(x, 5)] : plot(GS, x=xrange, color=clist,
  linestyle=slist) ;
> GS_lim := [base_lim(x, 3)] : plot(GS_lim, x=xrange, color=clist,
  linestyle=slist) ;
```





Meilleure approximation concave dans le cas de la route A, en degré 3..

```
> for k to nb do c||k := seq(prod_scal(P||k, GS[j], x),
  j=1..nops(GS)) ;
> c_lim||k := seq(prod_scal(P||k, GS_lim[j], x), j=1..nops(GS_lim)) ;
  end do :
> Q3 := sum(a[i]*GS[i], i=1..4) : Q3_lim := sum(b[i]*GS_lim[i],
  i=1..2) :
> R3 := diff(Q3, x
   2) : R3_lim := diff(Q3_lim, x£2) :
      d1 := subs(x=0, R3) : d2 := subs(x=ChiMax, R3) : evalf(d1) ;
      evalf(d2) ;
```

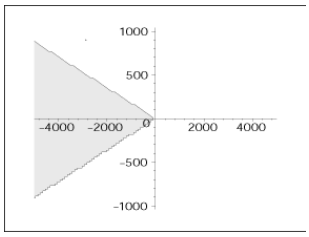
$.0005803910069\,a_3 - .003210991445\,a_4 .0005803910069\,a_3 + .003285173652\,a_4$

```
> d1_lim := subs(x=0, R3_lim) : d2_lim := subs(x=ChiMax, R3_lim) :
  evalf(d1_lim) ; evalf(d2_lim) ;
```

$$-.0003041284572\,b_1 + .002212620451\,b_2$$

$$-.0003041284572\,b_1 - .002287495689\,b_2$$

```
> test_ccv := proc(x, y) local u, v ; u := evalf(subs({a[3]=x,
  a[4]=y}, d1)) ; v := evalf(subs({a[3]=x, a[4]=y}, d2)) ; if
  sign(u) <= 0 and sign(v) <= 0 then 1 else -1 end if ; end
  proc :
> ccv := implicitplot(test_ccv, -5000..5000, -1000..1000,
  contours=[0], filled=true, colouring=[white, gray], grid=[100,
  100]) :
```

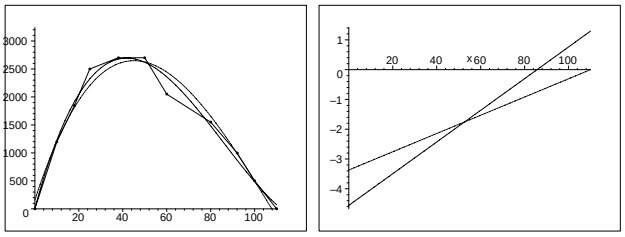


Il en résulte que le projeté de P1 sur le domaine de concavité, i.e. la meilleure approximation concave des données de la route 1 en degré 3, est la projection sur d2. De sorte que le projeté cherché sera a1.G1 + a2.G2 + U.G3 + V.G4 où (U,V) est le projeté de (u,v) sur d2.

```
> proj := proc(d, a, b, u, v) local A, B ; A := subs({a=1,b=0},
  d) ; B := subs({a=0,b=1}, d) ; u-A*(A*u+B*v)/(A^2+B^2),
  v-B*(A*u+B*v)/(A^2+B^2) ; end proc :
> C1 := [c1[1], c1[2], proj(d2, a[3], a[4], c1[3], c1[4])] :
> Proj1 := sum(C1[i]*GS[i], i=1..4) ;
```

$$Proj1 := \frac{638215794589780}{4073153358839} + \frac{1476796145449537}{12219460076517}\,x - \frac{6862981991470}{4073153358839}\,x^2 + \frac{62390745377}{12219460076517}\,x^3$$

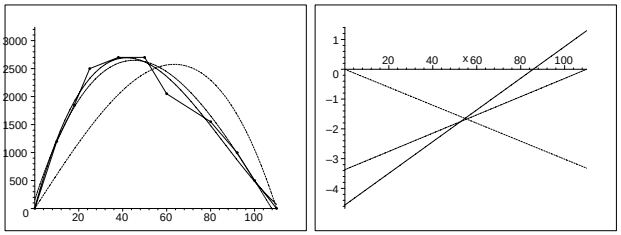
```
> display({MCplot3[1], plot(Proj1, x=xrange, linestyle=3,
  colour=black)}) ;
> plot([diff(P1, x£2), diff(Proj1, x£2)], x=xrange, colour=clist,
  linestyle=[1,3]) ;
```




```
> proj_lim := proc(u, v) local U, V ; U, V := proj(d1_lim,
  b[1], b[2], u, v) ; if sign(evalf(subs({b[1]=U, b[2]=V},
  d2_lim))) > 0 then U, V := proj(d2_lim, b[1], b[2], u, v) ; if
  sign(evalf(subs({b[1]=U2, b[2]=V2}, d1_lim))) > 0 then U, V :=0,
  0 end if ; end if ; U, V ; end proc :
> C1_lim := [proj_lim(c_lim1[1], c_lim1[2])] :
> Proj_lim1 := expand(sum(C1_lim[i]*GS_lim[i], i=1..2)) ;
```

$$Proj\_lim1 := \frac{23532265130000}{386676220659}\,x - \frac{1944815300}{386676220659}\,x^3$$

```
> display({MCplot3[1], plot(Proj_lim1, x=xrange, linestyle=2,
  colour=black), plot(Proj1, x=xrange, linestyle=3,
  colour=black)}) ;
> plot([diff(P1, x£2), diff(Proj_lim1, x£2), diff(Proj1, x£2)],
  x=xrange, colour=clist, linestyle=[1,2,3]) ;
```



5.4. Schéma numérique.

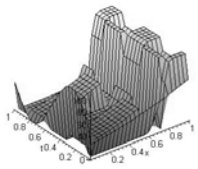
Implémentation du schéma numérique.

```
> trafic := proc(f,χ_i, χ_e, dx, dt, L, T) global ChiMax ; local
  j, n, jmax, d, c_x, c_t, e ; jmax := floor(L/dx) ; n := 0 ;
  for j from 0 to jmax do d[0][j] := evalf(χ_i(j*dx)) ; end do ;
  while n*dt < T do d[n+1][0] := evalf(χ_e(n*dt)) ; for j to
  jmax do if n = 0 then c_t := 1 else c_t := 2 ; end if ; if j =
  jmax then c_x := 1 else c_x := 2 ; end if ; e := d[n+1-c_t][j]
  - (c_t*dt)/(c_x*dx)*(f(d[n][j-1+c_x]) - f(d[n][j-1])) ;
  d[n+1][j] := min(max(0, e), ChiMax) ; end do ; n := n+1 ; end do ;
  d ; end proc :
> interpol := proc(d, x, t, dx, dt) local n, j ; j := round(x/dx) ;
  n := round(t/dt) ; d[n][j] ; end proc :
> tracé := proc(f, χ_i, χ_e, dx, dt, L, T) local d, h ; d :=
  trafic(f,χ_i,χ_e, dx, dt, L, T) ; h := (x, t) → interpol(d,
  x, t, dx, dt) ; plot3d(h(x, t), x=0..L, t=0..T, axes=NORMAL) ; end
  proc :
```

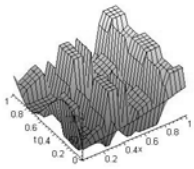
Impact de la modélisation.

```
> dx := 10^(-2) : dt := 10^(-2) : L := 1. : T := 1. : χ_i := x →
  70*exp(x-L) : χ_e := t → 50*(1+sin(2*π*t)) :
```

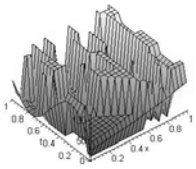
Les quatre premières figures montrent l'effet d'une perturbation sur la fonction f. Les polynômes de degré 2 étant concaves, on voit qu'une perturbation, même atténuée par la méthode des moindres carrés a un impact sur la résorption des bouchons. Les cinq dernières figures ne s'intéressent qu'à la route A (données du texte). D'abord avec la précision maximale, puis en degré 2 et enfin en degré 3 : sans contrainte, avec concavité, avec concavité et nullité aux bornes. On constate avec ces exemples que la modélisation n'est pas neutre et qu'il faut donc valider les hypothèses faites avec vigilance.



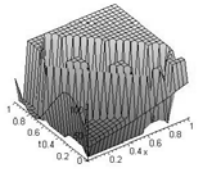
$\chi \to \chi * (110 - \chi)$



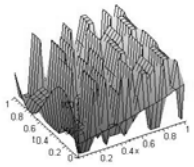
`PolMC10[3]`



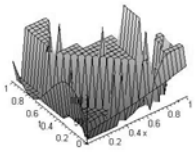
`PolMC2[3]`



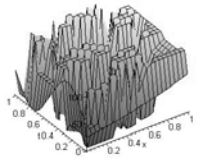
`PolMC2[4]`



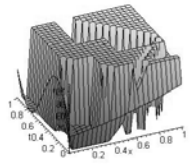
`PolMC10[1]`



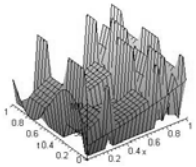
`PolMC2[1]`



`P1`



`Proj1`



`Proj_lim1`

§6. Commentaire

Le développement suggéré ici n'est bien sûr que le support de la prestation orale. Il pourra être étayé dans deux directions différentes. La première consiste à fournir des éclaircissements de nature théorique : le texte est un prétexte pour présenter une séquence pédagogique et on peut insérer dans le commentaire des outils en liaison avec une leçon classique. Ainsi, dans le format proposé on peut penser à fournir (dans le cadre de l'exposé oral ou en vue de réponses à des questions du jury) des détails sur les points suivants : méthode des moindres carrés d'un point de vue statistique, approximation polynomiale au sens des moindres carrés d'un point de vue d'algèbre linéaire, approximation au sens de Hermite et/ou polynomiale par morceaux, théorème de projection sur les convexes fermés par exemple en relation avec de la géométrie effective, méthodes de quadratures et recherche de polynômes orthogonaux, problèmes de convergence dans l'approximation polynomiale (les théorèmes, le phénomène de Runge etc.), la recherche de racines pour un polynôme, le théorème de Sturm etc.

La seconde direction est au niveau de l'implémentation sur machine. On peut choisir d'illustrer le propos tenu dans ce développement (en suivant sa démarche et en l'implémentant, par exemple pour en retrouver les résultats), compléter le tableau brossé ou regarder des cas particuliers, montrer au jury

ses connaissances dans les points non traités durant l'exposé oral etc. Ici on pourrait évoquer des schémas numériques pour les EDP ou la méthode des caractéristiques, traiter le cas linéaire (f linéaire), évoquer la stabilité des solutions ou le principe du maximum etc. D'une façon générale, le texte fournit une situation et il convient de se laisser porter par elle, de se laisser dériver au gré de ses idées pour montrer des phénomènes de modélisation. Le but n'est pas de se laisser contraindre par le texte, de le résumer et de ne jamais ni le dépasser ni, faute de temps, l'approfondir. On peut bien au contraire choisir un passage où on a du matériel à présenter et l'approfondir, tout en acceptant de répondre éventuellement à quelques questions plus centrées sur le texte. Notamment pour l'épreuve commune au calcul symbolique et au calcul numérique, il est possible de traiter plus «géométriquement» des sujets très numériques et réciproquement. Ainsi dans ce texte on aurait pu présenter une méthode algorithmique de projection en la reliant à des problèmes de minimisation et, par exemple, à une méthode de gradient.

Table des Matières

Table des matières

Partie I. Localisation par GPS

Chapitre 1. Localisation par GPS 7

§1. Introduction 7
§2. Le problème statique 8
§3. Le problème dynamique 9
§4. Application 13
§5. Annexe : Un calcul explicite 15

Chapitre 2. Commentaire de l'auteur 17

Partie II. Modèle simplifié de mélange liquide-gaz

Chapitre 3. Modèle simplifié de mélange liquide–gaz 21

§1. Modélisation 21
§2. Vitesse de propagation 22
§3. Simulation numérique 23

Chapitre 4. Commentaires de l'auteur 27

Chapitre 5. Commentaire d'un non spécialiste 29

§1. Un problème d'interaction liquide-gaz très simplifié 29
§2. L'équation d'évolution et ses variantes 31
§3. Le schéma numérique d'Euler et quelques variantes 33
§4. Validation de l'approche asymptotique 39
§5. Évolution en milieu adiabatique ou en présence d'un terme de dissipation 42
§6. En guise de conclusion 45
§7. Annexe (calcul de $U, \partial_t^2 U, \partial_x^2 U$) 46

Chapitre 6. modèle simplifié de mélange liquide-gaz 47

§1. Modélisation 47
§2. Vitesse de propagation 48
§3. Simulation numérique 49

Partie III. Logarithmes discrets et cryptographie

Chapitre 7. Logarithme discret et cryptographie 53

§1. Introduction 53
§2. RSA 54
§3. Signatures sécurisées 55
§4. Le logarithme discret 56
§5. Preuves sans apport d'information 56

Chapitre 8. Commentaire de l'auteur 58

Chapitre 9. **Logarithme discret et cryptographie** **60**

Chapitre 10. Commentaire d'un non spécialiste 62

Chapitre 11. Commentaire TIPE du texte 66

Partie IV. Configuration centrale de quatre corps dans le plan

Chapitre 12. Configurations centrales
de quatre corps dans le plan 73

§1. Introduction 73
§2. Configurations centrales de quatre corps 74
§3. Déterminants de Cayley 74
§4. Configurations planes 76
§5. Détermination des configurations centrales 77
§6. Résolution des équations 78

Chapitre 13. Commentaire de l'auteur 84

Chapitre 14. Commentaire d'un spécialiste 88

Partie V. Modélisation du système ALOHA

Chapitre 15. Modélisation du système ALOHA 93

Chapitre 16. Commentaire de l'auteur 99

Chapitre 17. Commentaire d'un spécialiste 103

§1. Les probabilités de transition 103
§2. Dynamique et équilibres 105
§3. Une simulation 106
§4. Temps moyen de saturation 107

§5. Nombre infini d'utilisateurs 108
§6. Appendice : Les programmes Matlab 110

Chapitre 18. Utilisation en TIPE 114

Partie VI. Détection d'un signal en télécommunication

Chapitre 19. Détection d'un signal en télécommunications 119

§1. Le texte proposé 119

Chapitre 20. Commentaire de l'auteur 125

§1. Commentaire 125

Chapitre 21. Commentaire d'un spécialiste 132

§1. Introduction 132
§2. Les définitions 132
§3. Le test de Neyman et Pearson 133
§4. Le test randomisé de Neyman et Pearson 137

Partie VII. Fonctions de Lyapounov et stabilité de réseaux de télécommunication

Chapitre 22. Fonctions de Lyapounov et stabilité de réseaux de télécommunications 143

§1. Critères de classification des chaînes de Markov à espace d'états dénombrable 144
§2. Application à la stabilité de certains protocoles de communications . 145
§3. Appendice sur la classification des chaînes de Markov à espace d'états dénombrable 150

Chapitre 23. Commentaire d'un spécialiste 152

§1. Plan 152
§2. Formulation du problème étudié 152
§3. Résultats mathématiques 153
§4. Expérimentations numériques 156
§5. Compléments 159
§6. Conclusion 159
§7. Programme Matlab pour expérimenter 160

Chapitre 24. Exploitation en TIPE 166

Partie VIII. Le coussin financier

Chapitre 25. Le coussin financier 171

Chapitre 26. Commentaire de l'auteur 177

Chapitre 27. Commentaire d'un spécialiste : autour du théorème de Lindeberg 183

§1. Théorème de Lindeberg 183
§2. Approximation d'une binomiale 184
§3. Le coussin financier 185

Chapitre 28. Commentaire d'un non spécialiste 187

Chapitre 29. Le coussin financier : TIPE 194

Partie IX. Trafic routier

Chapitre 30. Trafic Routier 205

§1. Modéliser la circulation automobile 205
§2. Exploitation du modèle 211

Chapitre 31. Commentaire de l'auteur 215

§1. La modélisation 215
§2. Modèle Numérique 220
§3. Bibliographie 221

Chapitre 32. Exploitation du texte pour une Analyse de Texte Scientifique (ADS) 222

§1. Plan 222
§2. Proposition d'exposition de la part du candidat 222
§3. Propositions de questions 222
§4. Conclusion 223

Chapitre 33. Commentaire d'un non spécialiste 224

§1. Position du problème 224
§2. Implémentation du schéma numérique 225
§3. Approximation polynomiale 225
§4. Approximation polynomiale avec contraintes 227
§5. Code Maple (Maple 6) 229

§6. Commentaire 236