

CHAPITRE II

ADAPTATION DES DIFFERENTES MODELISATIONS SOUS FORME SINGULIEREMENT PERTURBEE POUR LA SIMPLIFICATION DES CHAINES DE MARKOV

I. Introduction

L'utilisation pratique des chaînes de Markov se trouve souvent limitée suite au nombre très important d'états à prendre en considération. En effet, chaque état du modèle représente une combinaison d'états de tous les éléments qui composent le système. Nous obtenons une explosion combinatoire du nombre de situations possibles, donc la dimension du modèle devient très importante. Ceci a été mis en évidence par de nombreuses applications des chaînes de Markov dans des domaines comme les processus décisionnels [BEL, 57], [DER, 70], [WHI, 85, 88, 89], les systèmes de production manufacturiers [SHA, 88] [ZHA, 94], la détection de pannes [HU, 93], [COR-a, 94], [NIE, 92], [SIN, 83], les files d'attente [COU, 77], [HO, 83] ainsi que la gestion des systèmes hydroélectriques [DEL, 78].

L'idée d'appliquer les techniques de *Perturbations* pour simplifier et réduire les chaînes de Markov a déjà fait l'objet des travaux de PHILLIPS [PHP-a, 81], SCHWEITZER [SCHW, 86], DELEBEQUE [DEL, 81, 83]. Les techniques développées se basent uniquement sur l'existence de faibles interactions entre les sous-systèmes du modèle initial.

Notre étude porte sur l'utilisation des perturbations singulières pour la simplification et la réduction des chaînes de Markov dont l'évolution des probabilités d'état se fait à des vitesses différentes. Pour cela, nous définissons tout d'abord la propriété de double échelle de temps d'une chaîne de Markov. Ensuite, nous adaptons les différentes modélisations sous forme singulièrement perturbée aux chaînes de Markov en vue de les découpler en deux sous-systèmes : un lent et l'autre rapide. Enfin, nous analysons les différentes méthodes développées pour mettre en évidence les techniques les plus efficaces pour la simplification des chaînes de Markov.

II. Utilisation des perturbations singulières pour la simplification des chaînes de Markov

Nous proposons de simplifier les chaînes de Markov irréductibles qui correspondent à

une classe finale, en utilisant la méthode des perturbations singulières, basée sur l'existence de la propriété de double échelle de temps. Pour définir cette propriété dans le cas des chaînes de Markov, nous devons analyser les facteurs qui influencent l'évolution d'un tel système. Par conséquent, nous proposons une analyse de la distribution des probabilités d'état, en fonction des propriétés spectrales de la matrice de transition.

II.1. Distribution des probabilités d'une chaîne de Markov ergodique irréductible

Considérons une chaîne de Markov finie irréductible et ergodique, représentée par sa formule fondamentale :

$$\mathbf{P}(n+1) = \mathbf{P}(n)\dagger \quad (\text{II.1})$$

avec $\mathbf{P}(n)$ á $\hat{e}^{1 \times r}$ le vecteur des probabilités d'états de composantes $P_i(n)$, $i = 1, \dots, r$, et \dagger á $\hat{e}^{r \times r}$ la matrice de transition d'éléments p_{ij} , $i, j = 1, \dots, r$.

Nous nous proposons d'analyser la distribution des probabilités d'état de ce système en fonction des propriétés du spectre et des vecteurs propres de la matrice de transition.

II.1.1. Cas général - écriture de la matrice de transition à l'aide de la forme de Jordan

Dans le cas général, la matrice \dagger peut s'écrire sous la forme suivante :

$$\dagger = \mathbf{V}^* \mathbf{J} \mathbf{F}^* \quad (\text{II.2})$$

avec \mathbf{V}^* la matrice des vecteurs propres généralisés, \mathbf{F}^* la matrice des formes propres généralisées et \mathbf{J} la forme de Jordan associée, telles que :

$$\mathbf{V}^* = [\mathbf{v}_1^* \mathbf{v}_2^* \dots \mathbf{v}_r^*] , \quad \mathbf{F}^* = \mathbf{V}^{*-1} = \begin{bmatrix} \mathbf{f}_1^* \\ \mathbf{f}_2^* \\ \vdots \\ \mathbf{f}_r^* \end{bmatrix} , \quad \mathbf{J} = \begin{pmatrix} \mathbf{J}_1 & 0 & \dots & 0 \\ 0 & \mathbf{J}_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \mathbf{J}_m \end{pmatrix} \quad (\text{II.3})$$

où nous avons noté par m ($m \leq r$) le nombre de valeurs propres distinctes de \dagger .

Les vecteurs propres généralisés \mathbf{v}_i^* et les formes propres généralisées \mathbf{f}_i^* de la matrice \dagger s'écrivent :

$$\mathbf{v}_i^* = \text{Erreur !} \text{ respectivement } \mathbf{f}_i^* = \text{Erreur !}, \quad (i = 1, \dots, r)$$

La forme de Jordan \mathbf{J} á $\hat{a}^{r \times r}$ est constituée par les m blocs de Jordan \mathbf{J}_i á $\hat{a}^{k_i \times k_i}$ ($i = 1, \dots, m$) correspondant chacun à une valeur propre distincte λ_i ($i = 1, \dots, m$), d'ordre de multiplicité k_i :

$$\mathbf{J}_i = \begin{pmatrix} \lambda_i \alpha_i & 0 & \dots & 0 & 0 \\ 0 & \lambda_i \alpha_i & \dots & 0 & 0 \\ 0 & 0 & \lambda_i & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_i \alpha_i \\ 0 & 0 & 0 & \dots & 0 & \lambda_i \end{pmatrix} \quad \text{avec} \quad \begin{cases} \alpha_i = 0 & \text{si les vecteurs propres associés à } \lambda_i \text{ sont indépendants} \\ \alpha_i = 1 & \text{si les vecteurs propres associés à } \lambda_i \text{ sont dépendants} \end{cases} \quad (\text{II.4})$$

respectivement

$\mathbf{J}_i = \lambda_i$, si λ_i est une valeur propre simple ($k_i = 1$).

Nous avons l'égalité : $\left(\sum_{i=1}^m k_i \right) = r$.

Nous considérons d des m valeurs propres distinctes de la matrice de transition \dagger comme étant des valeurs propres simples ($d \leq m$). Dans ce cas, la forme de Jordan associée s'écrit sous la forme développée suivante :

$$\mathbf{J} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_d & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 \\ \hline 0 & 0 & \dots & 0 & \lambda_{d+1} \alpha_{d+1} & \dots & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & \lambda_{d+1} & \dots & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & \lambda_{d+1} \alpha_{d+1} & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \lambda_{d+1} & \dots & 0 & 0 \\ \hline \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & \lambda_m \alpha_m & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \lambda_m & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & \lambda_m & \alpha_m \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \lambda_m \end{pmatrix} \quad (\text{II.5})$$

La formule qui détermine la distribution des états de la chaîne de Markov en fonction de la matrice de transition et des conditions initiales, est :

$$\mathbf{P}(n) = \mathbf{P}(0) \dagger^n \quad (\text{II.6})$$

Lorsque nous remplaçons \dagger par l'expression (II.2), il nous vient :

$$\mathbf{P}(n) = \mathbf{P}(0) \begin{pmatrix} v_1^* & v_2^* & \dots & v_r^* \\ v_1^* & v_2^* & \dots & v_r^* \\ \dots & \dots & \dots & \dots \\ v_1^* & v_2^* & \dots & v_r^* \end{pmatrix} \mathbf{J}^n \begin{pmatrix} f_1^* & f_2^* & \dots & f_r^* \\ f_1^* & f_2^* & \dots & f_r^* \\ \dots & \dots & \dots & \dots \\ f_1^* & f_2^* & \dots & f_r^* \end{pmatrix} \quad (\text{II.7})$$

avec :

$$\mathbf{J}^n = \begin{pmatrix} (\lambda_1)^n & 0 & \dots & 0 & [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] \\ 0 & (\lambda_2)^n & \dots & 0 & [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & (\lambda_d)^n & [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] \\ [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] & (\mathbf{J}_{d+1})^n & [\mathbf{o}] & \dots & [\mathbf{o}] \\ [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] & [\mathbf{o}] & (\mathbf{J}_{d+2})^n & \dots & [\mathbf{o}] \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ [\mathbf{o}] & [\mathbf{o}] & \dots & [\mathbf{o}] & [\mathbf{o}] & [\mathbf{o}] & \dots & (\mathbf{J}_m)^n \end{pmatrix} \quad (\text{II.8})$$

Pour $n \geq k_i$ ($i = d+1, \dots, m$) nous pouvons écrire (voir *annexe I*) :

$$\mathbf{J}_i^n = \begin{pmatrix} \lambda_1^n & \alpha_i C_n^{n-1} \lambda_1^{n-1} & \dots & \alpha_i^{k_i-1} C_n^{n-k_i+1} \lambda_1^{n-k_i+1} \\ 0 & \lambda_i^n & \dots & \alpha_i^{k_i-2} C_n^{n-k_i+2} \lambda_i^{n-k_i+2} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda_i^n \end{pmatrix} \quad (\text{II.9})$$

où $C_n^i = n! / [(n-i)! i!]$ représente le nombre de combinaisons de i parmi n .

En introduisant l'expression (II.8) de \mathbf{J}^n dans (II.7), nous obtenons :

$$\begin{aligned} \mathbf{P}(n) = & (\mathbf{P}(0)\mathbf{v}_1^*) (\lambda_1)^n \mathbf{f}_1^* + (\mathbf{P}(0)\mathbf{v}_2^*) (\lambda_2)^n \mathbf{f}_2^* + \dots + (\mathbf{P}(0)\mathbf{v}_d^*) (\lambda_d)^n \mathbf{f}_d^* + (\mathbf{P}(0)\mathbf{v}_{d+1}^*) (\lambda_{d+1})^n \mathbf{f}_{d+1}^* + \\ & + \{[\mathbf{P}(0)\mathbf{v}_{d+2}^*](\lambda_{d+1})^n + [\mathbf{P}(0)\mathbf{v}_{d+1}^*] \alpha_{d+1} C_n^{n-1} (\lambda_{d+1})^{n-1}\} \mathbf{f}_{d+2}^* + \dots + \{(\mathbf{P}(0)\mathbf{v}_{d+k_{d+1}}^*) (\lambda_{d+1})^n + \\ & + \dots + \{(\mathbf{P}(0)\mathbf{v}_{d+k_{d+1}-1}^*) (\alpha_{d+1})^i C_n^{n-i} (\lambda_{d+1})^{n-i}\} \mathbf{f}_{d+k_{d+1}}^* + \dots + (\mathbf{P}(0)\mathbf{v}_{r-k_m+1}^*) (\lambda_m)^n \mathbf{f}_{r-k_m+1}^* + \dots \\ & + \{(\mathbf{P}(0)\mathbf{v}_r^*) (\lambda_m)^n + \dots + \{(\mathbf{P}(0)\mathbf{v}_{r-i}^*) (\alpha_m)^i C_n^{n-i} (\lambda_m)^{n-i}\} \mathbf{f}_r^* \end{aligned} \quad (\text{II.10})$$

qui peut s'écrire sous une forme condensée :

$$\begin{aligned} \mathbf{P}(n) = & (\lambda_1)^n \psi_1 \mathbf{f}_1^* + (\lambda_2)^n \psi_2 \mathbf{f}_2^* + \dots + (\lambda_d)^n \psi_d \mathbf{f}_d^* + (\lambda_{d+1})^n \psi_{d+1} \mathbf{f}_{d+1}^* + \Phi_{d+2}(n, \lambda_{d+1}) \mathbf{f}_{d+2}^* + \dots \\ & + \Phi_{d+k_{d+1}}(n, \lambda_{d+1}) \mathbf{f}_{d+k_{d+1}}^* + \dots + (\lambda_m)^n \psi_m \mathbf{f}_{r-k_m+1}^* + \dots + \Phi_r(n, \lambda_m) \mathbf{f}_r^* \end{aligned} \quad (\text{II.11})$$

avec ψ_i ($i = 1, \dots, m$) des constantes et $\Phi_j(n, \lambda)$ ($j = d+2, \dots, r$) des fonctions de la valeur propre correspondante et du temps n .

II.1.2. Cas particulier - matrice de transition diagonalisable

Lorsque la matrice de transition de la chaîne de Markov est irréductible et possède des vecteurs propres linéairement indépendants, elle peut s'écrire sous la forme décomposée suivante

:

$$\dot{\mathbf{P}} = \mathbf{V} \mathbf{\Lambda} \mathbf{F} \quad (\text{II.12})$$

$$\mathbf{V} = [\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_r], \quad \mathbf{F} = \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_r \end{bmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_r \end{pmatrix}$$

avec

$$(\text{II.13})$$

où $\mathbf{v}_i = \begin{pmatrix} v_i^1 \\ \vdots \\ v_i^r \end{pmatrix}$ et $\mathbf{f}_i = (f_i^1 \ f_i^2 \ \dots \ f_i^r)$ ($i=1, \dots, r$) représentent les vecteurs propres respectivement les formes propres de la matrice $\dot{\mathbf{P}}$.

En écrivant la formule fondamentale pour n instants successifs, nous obtenons :

$$\mathbf{P}(n) = \mathbf{P}(0) (\mathbf{V} \mathbf{\Lambda} \mathbf{F})^n = \mathbf{P}(0) \mathbf{V} \mathbf{\Lambda}^n \mathbf{F}, \quad (\text{II.14})$$

ou sous une forme explicite :

$$\mathbf{P}(n) = \mathbf{P}(0) \begin{pmatrix} v_1^1 & v_1^2 & \dots & v_1^r \\ v_2^1 & v_2^2 & \dots & v_2^r \\ \vdots & \vdots & \ddots & \vdots \\ v_r^1 & v_r^2 & \dots & v_r^r \end{pmatrix} \begin{pmatrix} (\lambda_1)^n & 0 & \dots & 0 \\ 0 & (\lambda_2)^n & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\lambda_r)^n \end{pmatrix} \begin{pmatrix} f_1^1 & f_1^2 & \dots & f_1^r \\ f_2^1 & f_2^2 & \dots & f_2^r \\ \vdots & \vdots & \ddots & \vdots \\ f_r^1 & f_r^2 & \dots & f_r^r \end{pmatrix} \quad (\text{II.15})$$

Le développement du calcul, donne un vecteur des probabilités d'état $\mathbf{P}(n)$ s'écrivant en fonction des valeurs propres de la matrice $\dot{\mathbf{P}}$ comme suit :

$$\mathbf{P}(n) = (\lambda_1)^n \mathbf{v} \cdot \mathbf{f}_1 + (\lambda_2)^n (\mathbf{P}(0) \cdot \mathbf{v}_2) \cdot \mathbf{f}_2 + \dots + (\lambda_r)^n (\mathbf{P}(0) \cdot \mathbf{v}_r) \cdot \mathbf{f}_r \quad (\text{II.16})$$

ou encore, sous une deuxième forme :

$$\mathbf{P}(n) = (\lambda_1)^n \mathbf{v} \cdot \mathbf{f}_1 + (\lambda_2)^n \varpi_2 \cdot \mathbf{f}_2 + \dots + (\lambda_r)^n \varpi_r \cdot \mathbf{f}_r \quad (\text{II.17})$$

$$\text{avec } \varpi_i = \mathbf{P}(0) \cdot \mathbf{v}_i, \quad i = 1, 2, \dots, r.$$

L'analyse des expressions (II.11) et (II.17), nous montre que la distribution des probabilités d'une chaîne de Markov irréductible finie dépend de *toutes* les valeurs propres de la matrice de transition. Ces valeurs propres jouent le rôle d'*évolutions élémentaires propres* de la chaîne de Markov. Elles symbolisent les tendances naturelles du système et influencent l'allure de l'évolution des probabilités d'état.

L'existence de valeurs propres de modules différents est à l'origine des évolutions des probabilités d'état à des vitesses différentes. Le regroupement des valeurs propres selon l'ordre de leur module permet de faire apparaître des groupes de probabilités d'état à évolutions différentes. Dans ce cas, les perturbations singulières peuvent être appliquées.

Parmi les deux types (ergodique et périodique) de chaînes de Markov irréductibles finies, celles qui présentent intérêt pour notre étude sont les chaînes de Markov ergodiques. Dans ce qui suit, nous définissons la propriété de double échelle de temps dans le cas des chaînes de Markov irréductibles ergodiques.

II.2. Définition de la double échelle de temps des chaînes de Markov ergodiques

Nous présentons la définition de la double échelle de temps d'une chaîne de Markov/ processus de Markov ergodique.

Définition

Un processus de Markov homogène fini et ergodique / une chaîne de Markov homogène finie irréductible et ergodique d'équation fondamentale

$$\mathbf{P}\ddot{\mathbf{E}} = \mathbf{P} \ddot{\ddagger} \ddot{\mathbf{Y}} / \mathbf{P}(n+1) = \mathbf{P}(n) \ddot{\ddagger} \quad (\text{II.18})$$

avec le vecteur des probabilités d'état $\mathbf{P} = [\mathbf{P}_1 \dots \mathbf{P}_r] \hat{\mathbf{e}}^{1 \times r}$ de conditions initiales $\mathbf{P}_0 / \mathbf{P}(n) = [\mathbf{P}_1(n) \dots \mathbf{P}_r(n)] \hat{\mathbf{e}}^{1 \times r}$ de conditions initiales $\mathbf{P}(0)$ et le générateur $\ddot{\ddagger} \ddot{\mathbf{Y}} \hat{\mathbf{e}}^{r \times r} /$ la matrice de transition $\ddot{\ddagger} \hat{\mathbf{e}}^{r \times r}$, présente la propriété de double échelle de temps si il / elle peut être décomposé(e) en deux sous-systèmes disjoints :

$$\begin{bmatrix} \dot{\mathbf{P}}_1 & \dot{\mathbf{P}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_r \end{bmatrix} \begin{pmatrix} \ddot{\ddagger}_1 & 0 \\ 0 & \ddot{\ddagger}_r \end{pmatrix} / \begin{bmatrix} \mathbf{P}_1(n+1) & \mathbf{P}_r(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1(n) & \mathbf{P}_r(n) \end{bmatrix} \begin{pmatrix} \ddot{\ddagger}_1 & 0 \\ 0 & \ddot{\ddagger}_r \end{pmatrix} \quad (\text{II.19})$$

avec $\mathbf{P} = [\mathbf{P}_1 \mathbf{P}_r] \hat{\mathbf{e}}^{1 \times r} / \mathbf{P}(n) = [\mathbf{P}_1(n) \mathbf{P}_r(n)] \hat{\mathbf{e}}^{1 \times r}$ la décomposition de la distribution du système initial en ses variables lentes $\mathbf{P}_1 \hat{\mathbf{e}}^{1 \times r_1}$ et rapides $\mathbf{P}_r \hat{\mathbf{e}}^{1 \times r_2}$ ($r = r_1 + r_2$), de telle façon que les valeurs propres des matrices $\ddot{\ddagger} \ddot{\mathbf{Y}}_1$ et $\ddot{\ddagger} \ddot{\mathbf{Y}}_r / \ddot{\ddagger}_1$ et $\ddot{\ddagger}_r$ satisfassent :

$$\Im \lambda_{\max}(\ddot{\ddagger} \ddot{\mathbf{Y}}_1) \Im \ll \Im \lambda_{\min}(\ddot{\ddagger} \ddot{\mathbf{Y}}_r) \Im / \Im \lambda_{\min}(\ddot{\ddagger}_1) \Im \gg \Im \lambda_{\max}(\ddot{\ddagger}_r) \Im \quad (\text{II.20})$$

$\ddot{\ddagger} \ddot{\mathbf{Y}}_1 / \ddot{\ddagger}_1$ et $\ddot{\ddagger} \ddot{\mathbf{Y}}_r / \ddot{\ddagger}_r$ regroupent les modes lents et les modes rapides de la matrice initiale $\ddot{\ddagger} \ddot{\mathbf{Y}} / \ddot{\ddagger}$.

□

La condition (II.20) traduit une évolution plus rapide de $(\ddot{\ddagger}_r)^n$ par rapport à $(\ddot{\ddagger}_1)^n$. A partir de cette relation, nous pouvons exprimer la propriété de double échelle de temps par des conditions portant sur les normes de matrices.

Toute matrice carrée $\ddot{\ddagger}$ satisfait la relation suivante :

$$\Im \lambda_{\max}(\ddot{\ddagger}) \Im \leq \|\ddot{\ddagger}\|$$

quelle que soit la norme de matrices utilisée. Si $\ddot{\ddagger}^{-1}$ existe, nous pouvons écrire aussi :

$$\Im \lambda_{\min}(\ddot{\ddagger}) \Im^{-1} \leq \|\ddot{\ddagger}^{-1}\|$$

Par conséquent, la propriété de double échelle de temps (II.20) existe si :

$$\|\ddot{\ddagger}_r\|^{-1} \gg \|\ddot{\ddagger}_1\|^{-1}. \quad (\text{II.21})$$

Le rapport des échelles de temps peut être estimé par la valeur d'un coefficient μ tel que :

$$\mu = \frac{\mathfrak{S}\lambda_{\max}(\dot{\mathfrak{T}}_r)\mathfrak{S}}{\mathfrak{S}\lambda_{\min}(\dot{\mathfrak{T}}_l)\mathfrak{S}} \quad \text{avec } \mu \in]0, 1] \quad (\text{II.22})$$

II.3. Méthodes pratiques de mise en évidence des dynamiques d'une chaîne de Markov irréductible ergodique

Pour pouvoir détecter les dynamiques et identifier les états lents et rapides d'une chaîne de Markov à double échelle de temps, nous pouvons utiliser deux types de méthodes : analytiques et géométriques.

Les méthodes *analytiques* de mise en évidence des dynamiques d'une chaîne de Markov utilisent la bloc diagonalisation de la matrice de transition du modèle. Elles consistent à résoudre deux équations matricielles de type Ricatti et Lyapunov.

Les *méthodes géométriques* de mise en évidence des dynamiques d'un système consistent à localiser les valeurs propres de la matrice de transition dans le plan complexe. Le but de ces méthodes est de faire apparaître deux ou plusieurs domaines disjoints dans le plan complexe et de pouvoir ainsi identifier les différentes dynamiques du système. Ce sont ces méthodes que nous allons développer dans la suite de notre étude.

II.3.1. Méthodes géométriques de mise en évidence des dynamiques d'une chaîne de Markov irréductible ergodique

Il existe de nombreuses méthodes géométriques de mise en évidence des dynamiques d'un système à double échelle de temps. Parmi celles-ci, les disques de Gershgorine, les régions de Gudkov et les ovales de Cassini [CHA, 71], [VAR, 82], [BOUK, 91], [BOUA, 90] sont les plus importantes. Le tracé des régions de Gudkov et des ovales de Cassini est très complexe et difficile à appliquer. Vu sa simplicité, la méthode des disques de Gershgorine nous paraît très intéressante. Nous nous proposons de l'appliquer à la détection des dynamiques des chaînes de Markov.

La définition des disques de Gershgorine, comme moyen de localisation des valeurs propres, provient de la singularité de la matrice de l'équation caractéristique du système. Dans ce sens, nous pouvons énoncer le théorème suivant :

Théorème de Gershgorine [VAR, 62], [BOR, 92]

Les valeurs propres de la matrice de transition $\dot{\mathfrak{T}}[p_{ij}]$ (de dimension $r \times r$) d'une chaîne de Markov appartiennent au domaine complexe V_1 formé par l'union des disques (appelés disques de Gershgorine) de centres p_{ii} et rayons respectifs R_i , définis par l'équation :

$$|x - p_{ii}| = \sum_{j=1, j \neq i}^r p_{ij} = R_i, \quad x \in \mathbb{C}, \quad i = 1, \dots, r \quad (\text{en ligne}) \quad (\text{II.23})$$

et aussi au domaine V_2 du plan complexe formé par l'union des disques de centres p_{ii} et de rayons respectifs R'_i définis par l'équation :

$$\left| x - p_{ii} \right| = \sum_{j=1, j \neq i}^r p_{ji} = R_i', \quad x \in \hat{\Delta}, \quad i = 1, \dots, r \quad (\text{en colonne}) \quad (\text{II.24})$$

L'intersection $V = V_1 \cap V_2$ définit la région du plan où se situent les valeurs propres de la matrice $\hat{\dagger}$.

Si " s " des disques $D_i(p_{ii}, R_i)$ forment un domaine séparé des autres disques, alors il y a exactement " s " valeurs propres de $\hat{\dagger}$ situées dans ce domaine. \square

Une variante de ce théorème (appelée théorème d'Ostrowski [PAR, 52]) fait appel à des disques de centre p_{ii} et de rayon $R_i^\alpha R_i^{1-\alpha}$, avec $\alpha \in [0, 1]$. À partir de ces résultats, nous pouvons énoncer la proposition suivante :

Proposition 1

Si il existe un réel $\alpha \in [0, 1]$ tel que les disques de Gershgorine $D_i(p_{ii}, R_i^\alpha R_i^{1-\alpha})$ et $D_k(p_{kk}, R_k^\alpha R_k^{1-\alpha})$ vérifient la relation :

$$p_{ii} - p_{kk} \gg R_i^\alpha R_i^{1-\alpha} + R_k^\alpha R_k^{1-\alpha} \quad (\text{II.25})$$

À $i \in I_i$ et À $k \in I_k$, avec $I_i \cap I_k = \emptyset$, $I_i \cup I_k = \{1, 2, \dots, r\}$,

alors la matrice $\hat{\dagger}$ possède deux ensembles de valeurs propres séparées, de modules très différents. Par conséquent, la chaîne de Markov présente la propriété de double échelle de temps.

- Démonstration -

Soit la matrice $\hat{\dagger}$ d'éléments p_{ij} ($i, j = 1, \dots, r$). Une configuration des disques de Gershgorine correspondant à (II.25) est illustrée dans la figure suivante :

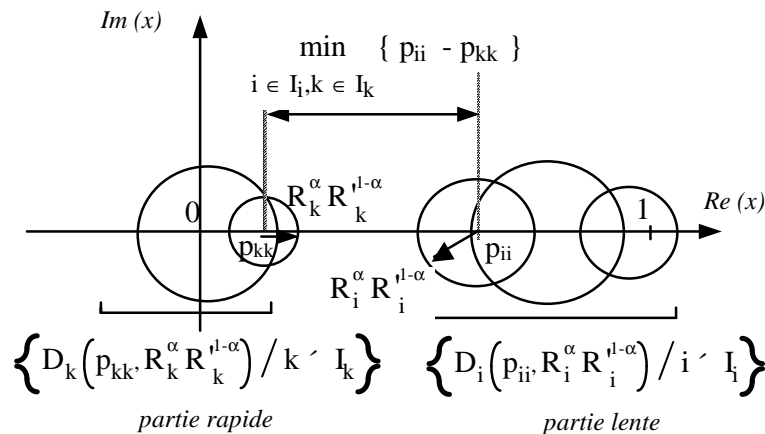


Fig. II.1. Disques de Gershgorine pour un système à double échelle de temps.

L'existence de deux ensembles disjoints de disques de Gershgorine D_i et D_k nous permet de conclure (suite au théorème de Gershgorine) à l'existence de deux groupes de valeurs propres de modules différents et par conséquent, à l'existence de la double échelle de temps.

□

Le cas $\alpha = 1$ (application sur les lignes) et $\alpha = 0$ (application sur les colonnes) de (II.25) présentent un grand intérêt et pratique, suite à la simplicité de construction des disques de Gershgorine. Vu la structure particulière de la matrice de transition d'une chaîne de Markov, la proposition 1 ne pourra pas être satisfaite pour $\alpha = 1$. En effet, suite à la propriété de stochasticté du système, tous les disques construits sur les lignes sont tangents en $(1, 0)$:

Ceci ne permet pas de mettre directement en évidence les dynamiques du système et nécessite par conséquent un conditionnement de la matrice. Pour cela, nous utilisons une méthode de *calibrage* [BOUA, 90]. Cette méthode est basée sur la minimisation de la norme infinie de la matrice formée par les éléments hors diagonaux de la matrice \ddagger . Elle utilise un algorithme d'homogénéisation des rayons des disques de Gershgorine. Pour cela, on utilise un changement de base diagonal Φ , représentant un produit de matrices diagonales de la forme suivante :

$$\Phi = \prod_{j=0} D_j, \text{ avec } D_0 = I_r \text{ (matrice identité) et } D_j = \begin{pmatrix} d_1^j & 0 & \dots & 0 \\ 0 & d_2^j & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_r^j \end{pmatrix} \text{ pour } j = 1, 2, \dots$$

(II.26)

Chaque matrice D_j se calcule par itération. Pour l'itération j , d_i^j ($i = 1, 2, \dots, r$) est la solution (unique) comprise entre 0 et 1, de l'équation :

$$\left| a_{mi}^{j-1} \right| (d_i^j)^2 + \left[\sum_{k=1, k \neq i}^r \left| a_{mk}^{j-1} \right| d_i^j - \sum_{k=1, k \neq i}^r \left| a_{ik}^{j-1} \right| \right] d_i^j = 0$$

(II.27)

avec a_{ik}^{j-1} ($i, k = 1, \dots, r$) représentant les éléments de la matrice $(D_0 D_1 \dots D_{j-1})^{-1} \ddagger (D_0 D_1 \dots D_{j-1})$ et "m" l'indice de la ligne caractérisée par la plus grande somme des modules de ses éléments hors diagonaux. L'itération s'arrête quand les rayons des disques de Gershgorine de la matrice $\Phi^{-1} \ddagger \Phi$ sont égaux (avec une variation admissible arbitrairement petite ou fonction de la précision du moyen de calcul).

La matrice calibrée $\Phi^{-1} \ddagger \Phi$ confère un conditionnement optimal pour obtenir une séparation des disques de Gershgorine. Dans le cas de l'apparition des deux domaines séparés, nous obtenons une configuration du type représenté dans la figure II.3 :

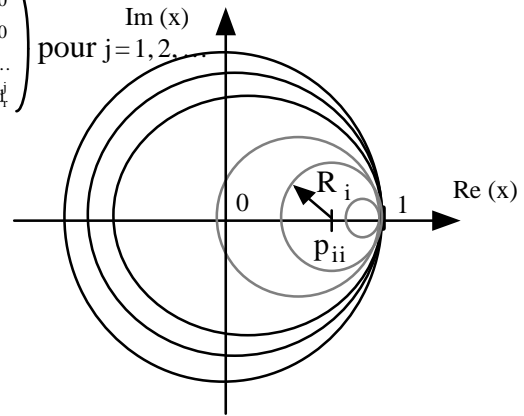


Fig. II.2: Disques de Gershgorine sur les lignes pour une chaîne de Markov

Nous pouvons conclure ainsi que la double échelle de temps existe et en plus, nous pouvons associer à chaque groupe de dynamiques différentes, un groupe d'états de la chaîne de Markov étudiée.

Un algorithme similaire peut être appliqué dans le cas de l'utilisation des disques de Gershgorine sur les colonnes de la matrice de transition. Le calibrage obtenu est le même.

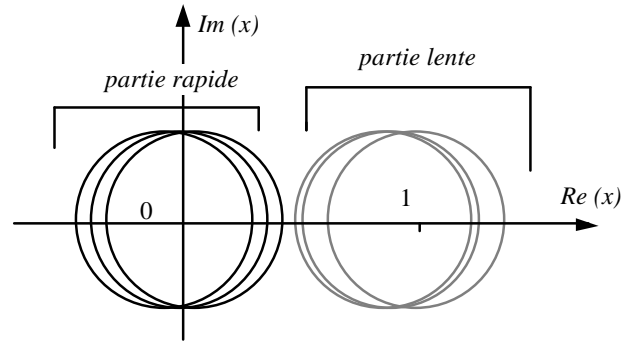


Fig. II.3. Disques de Gershgorine après calibrage de la matrice \dagger .

III. Application des perturbations singulières pour la simplification des chaînes de Markov à double échelle de temps

III.1. Définition des systèmes étudiés

Les systèmes que nous nous proposons d'étudier sont des chaînes de Markov irréductibles ergodiques d'équation fondamentale :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \dagger$$

possédant la propriété de double échelle de temps. L'équation fondamentale d'un tel système peut s'écrire sous une forme partitionnée :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \dagger_{11} & \dagger_{12} \\ \dagger_{21} & \dagger_{22} \end{pmatrix} \quad (\text{II.28})$$

avec

$\mathbf{L}(n) = [P_1(n) \ P_2(n) \ \dots \ P_{r_1}(n)]$ à $\hat{e}^{1 \times r_1}$ le vecteur des probabilités d'état de la partie lente,

$\mathbf{R}(n) = [P_{r_1+1}(n) \ P_{r_1+2}(n) \ \dots \ P_r(n)]$ à $\hat{e}^{1 \times r_2}$ la distribution de la partie rapide, $r_1 + r_2 = r$,

et $\mathbf{L}(0)$, $\mathbf{R}(0)$ les conditions initiales.

Dans ce qui suit nous adaptons les différentes méthodes de perturbations singulières introduites dans le premier chapitre pour simplifier les chaînes de Markov étudiées.

III.2. Utilisation de la modélisation de Phillips

La modélisation de Phillips présente un caractère général et trouve son application quelle que soit la nature des couplages entre les dynamiques du système Markovien à double échelle de temps.

III.2.1. Forme singulièrement perturbée

La forme singulièrement perturbée correspondant à l'équation fondamentale (II.28) s'écrit à l'aide de deux paramètres μ et j :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \dot{\dagger}_{11} & \mu^j \dot{\dagger}_{12}^* \\ \mu^{1-j} \dot{\dagger}_{21}^* & \mu \dot{\dagger}_{22}^* \end{pmatrix} \quad (\text{II.29})$$

avec $\dot{\dagger}_{21}^* = \dot{\dagger}_{21} / \mu^{1-j}$, $\dot{\dagger}_{12}^* = \dot{\dagger}_{12} / \mu^j$ et $\dot{\dagger}_{22}^* = \dot{\dagger}_{22} / \mu$.

La matrice $\dot{\dagger}_{11}$ est supposée inversible.

Le premier paramètre μ mesure la séparation des dynamiques du système et peut être choisi comme étant le rapport :

$$\mu = \|\dot{\dagger}_{22}\| / \|\dot{\dagger}_{11}\|, \quad \mu \in]0, 1] \quad (\text{II.30})$$

Le deuxième paramètre $j \in [0, 1]$ affecte les matrices de couplage entre les états lents et rapides.

III.2.2. Découplage des variables par bloc diagonalisation

Cette méthode consiste à effectuer deux changements de variables, appliqués sur l'équation fondamentale. Le premier changement de variable est

$$\mathbf{R}_r(n) = \mathbf{L}(n) \mathbf{R} + \mathbf{R}(n) \quad (\text{II.31})$$

avec \mathbf{R} la matrice de dimension $r_2 \times r_1$, solution de l'équation matricielle de type Riccati :

$$\dot{\dagger}_{12} + \dot{\dagger}_{11} \mathbf{R} - \mathbf{R} \dot{\dagger}_{21} \mathbf{R} - \mathbf{R} \dot{\dagger}_{22} = 0. \quad (\text{II.32})$$

Ce changement de variable permet d'obtenir la forme bloc triangulaire suivante :

$$[\mathbf{L}(n+1) \quad \mathbf{R}_r(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}_r(n)] \begin{pmatrix} \dot{\dagger}_{11} - \mathbf{R} \dot{\dagger}_{21} & 0 \\ \dot{\dagger}_{21} & \dot{\dagger}_{22} + \dot{\dagger}_{21} \mathbf{R} \end{pmatrix} \quad (\text{II.33})$$

Le second changement de variable sera :

$$\mathbf{L}_l(n) = \mathbf{L}(n) - \mathbf{R}_r(n) \mathbf{L} \quad (\text{II.34})$$

avec \mathbf{L} une matrice de dimension $r_1 \times r_2$, solution de l'équation matricielle de type Lyapunov :

$$\dot{\dagger}_{21} - (\dot{\dagger}_{22} + \dot{\dagger}_{21} \mathbf{R}) \mathbf{L} + \mathbf{L} (\dot{\dagger}_{11} - \mathbf{R} \dot{\dagger}_{21}) = 0. \quad (\text{II.35})$$

Ceci donne la forme bloc-diagonale recherchée :

$$[\mathbf{L}_l(n+1) \ \mathbf{R}_r(n+1)] = [\mathbf{L}_l(n) \ \mathbf{R}_r(n)] \begin{pmatrix} \dot{\gamma}_{11} - \mathbf{R} \dot{\gamma}_{21} & 0 \\ 0 & \dot{\gamma}_{22} + \dot{\gamma}_{21} \mathbf{R} \end{pmatrix} \quad (\text{II.36})$$

L'algorithme itératif de résolution de (II.32) et (II.35) utilise respectivement les suites suivantes :

$$\dot{\gamma}_{11}^{-1} (\mathbf{R}_i \dot{\gamma}_{22} + \mathbf{R}_i \dot{\gamma}_{21} \mathbf{R}_i - \dot{\gamma}_{12}) = \mathbf{R}_{i+1} \quad (\text{II.37})$$

et

$$[(\dot{\gamma}_{22} + \dot{\gamma}_{21} \mathbf{R}) \mathbf{L}_i - \dot{\gamma}_{21} + \mathbf{L}_i \mathbf{R} \dot{\gamma}_{21}] \dot{\gamma}_{11}^{-1} = \mathbf{L}_{i+1} \quad (\text{II.38})$$

Les solutions à l'ordre zéro en μ des équations (II.32) et (II.35) sont respectivement \mathbf{R}_0 et \mathbf{L}_0 telles que :

$$\mathbf{R}_0 = -\dot{\gamma}_{11}^{-1} \dot{\gamma}_{12} \quad (\text{II.39})$$

$$\mathbf{L}_0 = -\dot{\gamma}_{21} \dot{\gamma}_{11}^{-1}$$

En tenant compte des égalités (II.39) et en faisant l'approximation $\mathbf{I} - \mathbf{R}_0 \mathbf{L}_0 = \mathbf{I}$, nous obtenons le système découplé :

$$[\mathbf{L}_l(n+1) \ \mathbf{R}_r(n+1)] = [\mathbf{L}_l(n) \ \mathbf{R}_r(n)] \begin{pmatrix} \dot{\gamma}_l & 0 \\ 0 & \dot{\gamma}_r \end{pmatrix} \quad (\text{II.40})$$

avec

$$\dot{\gamma}_l = \dot{\gamma}_{11} + \dot{\gamma}_{11}^{-1} \dot{\gamma}_{12} \dot{\gamma}_{21}$$

$$\dot{\gamma}_r = \dot{\gamma}_{22} - \dot{\gamma}_{21} \dot{\gamma}_{11}^{-1} \dot{\gamma}_{12}$$

et

$$\mathbf{R}_l(n) = \mathbf{L}_l(n) \dot{\gamma}_{11}^{-1} \dot{\gamma}_{12}. \quad (\text{II.41})$$

La distribution \mathbf{R} des probabilités des états rapides est approchée par $\mathbf{R} = \mathbf{R}_r + \mathbf{R}_l$. La distribution des états lents \mathbf{L} est approchée par la distribution du système lent découplé $\mathbf{L} = \mathbf{L}_l$.

Les conditions initiales d'évolution des systèmes découplés sont données par les relations :

$$\mathbf{L}_l(0) = \mathbf{L}(0) \text{ et } \mathbf{R}_r(0) = \mathbf{R}(0) - \mathbf{L}(0) \dot{\gamma}_{11}^{-1} \dot{\gamma}_{12}$$

III.2.3. Analyse des systèmes découplés obtenus

a) Le sous-système lent

La matrice $\dot{\gamma}_l$ correspondant au système lent découplé est une matrice stochastique à l'ordre zéro en μ .

-Démonstration-

1°. La somme de chaque ligne de la matrice est égale à 1

Nous notons par $\mathbf{1}_{r_1}$ et respectivement $\mathbf{1}_{r_2}$ les vecteurs de dimension r_1 respectivement r_2 suivants :

$$\mathbf{1}_{r_1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{r_1 \times 1} \quad \text{et} \quad \mathbf{1}_{r_2} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{r_2 \times 1} \quad (\text{II.42})$$

En considérant l'expression (II.41) de la matrice $\dot{\mathbf{T}}_1$, nous obtenons :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \left(\dot{\mathbf{T}}_{11} + \dot{\mathbf{T}}_{11}^{-1} \dot{\mathbf{T}}_{12} \dot{\mathbf{T}}_{21} \right) \mathbf{1}_{r_1} = \dot{\mathbf{T}}_{11}^{-1} \left(\dot{\mathbf{T}}_{11} \dot{\mathbf{T}}_{11} \mathbf{1}_{r_1} + \dot{\mathbf{T}}_{12} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} \right) \quad (\text{II.43})$$

La propriété de stochasticté de la matrice initiale $\dot{\mathbf{T}}$ nous permet d'écrire la relation :

$$\dot{\mathbf{T}}_{11} \mathbf{1}_{r_1} = \left(\mathbf{1}_{r_1} - \dot{\mathbf{T}}_{12} \mathbf{1}_{r_2} \right) \quad (\text{II.44})$$

qui remplacée dans l'égalité (II.43), nous donne l'expression suivante :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \dot{\mathbf{T}}_{11}^{-1} \left[\dot{\mathbf{T}}_{11} \left(\mathbf{1}_{r_1} - \dot{\mathbf{T}}_{12} \mathbf{1}_{r_2} \right) + \dot{\mathbf{T}}_{12} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} \right] \quad (\text{II.45})$$

ou sous une autre forme :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \mathbf{1}_{r_1} + \dot{\mathbf{T}}_{11}^{-1} \left(\dot{\mathbf{T}}_{12} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} - \dot{\mathbf{T}}_{11} \dot{\mathbf{T}}_{12} \mathbf{1}_{r_2} \right) \quad (\text{II.46})$$

En écrivant cette égalité en fonction des sous-matrices normalisées, nous obtenons la forme équivalente :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \mathbf{1}_{r_1} + \dot{\mathbf{T}}_{11}^{-1} \left[\left(\mu^j \dot{\mathbf{T}}_{12}^* \right) \left(\mu^{1-j} \dot{\mathbf{T}}_{21}^* \right) \mathbf{1}_{r_1} - \dot{\mathbf{T}}_{11} \left(\mu^j \dot{\mathbf{T}}_{12}^* \right) \mathbf{1}_{r_2} \right] \quad (\text{II.47})$$

et par la suite :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \mathbf{1}_{r_1} + \mu \dot{\mathbf{T}}_{11}^{-1} \left(\dot{\mathbf{T}}_{12}^* \dot{\mathbf{T}}_{21}^* \mathbf{1}_{r_1} - \mu^{j-1} \dot{\mathbf{T}}_{11} \dot{\mathbf{T}}_{12}^* \mathbf{1}_{r_2} \right) \quad (\text{II.48})$$

A l'ordre zéro en μ et pour $j \neq 0$, nous avons la relation :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \mathbf{1}_{r_1} \quad (\text{II.49})$$

2°. Les éléments de la matrice sont non-négatifs

En partant de l'expression de la matrice $\dot{\mathbf{T}}_1$, nous pouvons la réécrire sous la forme :

$$\dot{\mathbf{T}}_1 = \left(\mathbf{I} + \dot{\mathbf{T}}_{11}^{-1} \dot{\mathbf{T}}_{12} \dot{\mathbf{T}}_{21} \dot{\mathbf{T}}_{11}^{-1} \right) \dot{\mathbf{T}}_{11}, \quad \text{où } \mathbf{I} \text{ est la matrice identité, (II.50)}$$

et par conséquent :

$$\dot{\mathbf{T}}_1 = \left(\mathbf{I} + \mu \dot{\mathbf{T}}_{11}^{-1} \dot{\mathbf{T}}_{12}^* \dot{\mathbf{T}}_{21}^* \dot{\mathbf{T}}_{11}^{-1} \right) \dot{\mathbf{T}}_{11} \quad (\text{II.51})$$

Le terme $\mu \dot{\mathbf{T}}_{11}^{-1} \dot{\mathbf{T}}_{12}^* \dot{\mathbf{T}}_{21}^* \dot{\mathbf{T}}_{11}^{-1} = \mathbf{R}_0 \mathbf{L}_0$ est de l'ordre de μ . A l'ordre $0(\mu)$, la matrice $\dot{\mathbf{T}}_1$ est approchée par $\dot{\mathbf{T}}_{11}$. Par conséquent, la matrice $\dot{\mathbf{T}}_1$ a des éléments non-négatifs.

En conclusion, la matrice de transition du système lent est une matrice stochastique à l'ordre zéro en μ . Le système lent est représenté par une chaîne de Markov irréductible indépendante. Ceci nous permet de résoudre ce sous-système en utilisant les méthodes de résolution caractéristiques aux chaînes de Markov.

□

b) Système rapide

La matrice $\dot{\mathbf{T}}_r$ correspondant au système rapide découplé n'est pas une matrice stochastique. Ceci ne nous permet pas de résoudre le sous-système rapide en utilisant les méthodes de résolution caractéristiques aux chaînes de Markov.

III.3. Simplification des chaînes de Markov en utilisant la modélisation d'El Moudni

III.3.1. Forme singulièrement perturbée

En reprenant la forme partitionnée (II.28) de l'équation fondamentale d'une chaîne de Markov irréductible ergodique présentant la propriété de double échelle de temps, la forme singulièrement perturbée de cette équation s'écrit :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mu \mathbf{R}(n)] \begin{pmatrix} \dot{\mathbf{T}}_{11} & \dot{\mathbf{T}}_{12} \\ \dot{\mathbf{T}}_{21}^* & \dot{\mathbf{T}}_{22}^* \end{pmatrix} \quad (\text{II.52})$$

avec : $\dot{\mathbf{T}}_{21}^* = \dot{\mathbf{T}}_{21} / \mu$, $\dot{\mathbf{T}}_{22}^* = \dot{\mathbf{T}}_{22} / \mu$ et $\dot{\mathbf{T}}_{11}$ matrice régulière.

Le coefficient μ reflète le rapport entre les dynamiques du système, et peut être défini comme étant :

$$\mu = \|\dot{\mathbf{T}}_{22}\| / \|\dot{\mathbf{T}}_{11}\|, \quad \mu \in [0, 1] \quad (\text{II.53})$$

Les chaînes de Markov correspondant à cette modélisation sont caractérisées par une influence réduite des variables rapides au cours de l'évolution du système. L'introduction du paramètre μ dans la formule fondamentale perturbée (II.52) correspond à une homogénéisation des matrices $\dot{\mathbf{T}}_{21}$ et $\dot{\mathbf{T}}_{22}$ par rapport aux matrices $\dot{\mathbf{T}}_{11}$ et $\dot{\mathbf{T}}_{12}$.

III.3.2. Découplage des dynamiques

a) Partie lente

En faisant $\mu=0$ dans l'équation (II.52), nous obtenons le système lent approché :

$$\begin{cases} \mathbf{L}_1(n+1) = \mathbf{L}_1(n) \dot{\mathbf{T}}_{11} \\ \mathbf{R}_1(n+1) = \mathbf{L}_1(n) \dot{\mathbf{T}}_{12} \end{cases} \quad \text{ou sous la forme} \quad \mathbf{R}_1(n) = \mathbf{L}_1(n) \dot{\mathbf{T}}_{11}^{-1} \dot{\mathbf{T}}_{12} \quad (\text{II.54})$$

avec les conditions initiales :

$$\mathbf{L}_1(0) = \mathbf{P}(0) \quad \text{et} \quad \mathbf{R}_1(0) = \mathbf{P}(0) \begin{matrix} \dagger \\ \dagger \end{matrix}_{11}^{-1} \begin{matrix} \dagger \\ \dagger \end{matrix}_{12}$$

b) Partie rapide

Pour l'étude de la partie rapide, les distributions des probabilités reviennent à leurs composantes rapides :

$$\mathbf{L}_r(n) = \mathbf{L}(n) - \mathbf{L}_1(n) \quad \text{et} \quad \mathbf{R}_r(n) = \mathbf{R}(n) - \mathbf{R}_1(n). \quad (\text{II.55})$$

dont les équations d'évolution sont :

$$\begin{cases} \mathbf{L}_r(n+1) = \mathbf{L}_r(n) \begin{matrix} \dagger \\ \dagger \end{matrix}_{11} + \mu \mathbf{R}_r(n) \begin{matrix} \dagger \\ \dagger \end{matrix}_{21}^* \\ \mathbf{R}_r(n+1) = \mathbf{L}_r(n) \begin{matrix} \dagger \\ \dagger \end{matrix}_{12} + \mu \mathbf{R}_r(n) \begin{matrix} \dagger \\ \dagger \end{matrix}_{22}^* \end{cases} \quad (\text{II.56})$$

Les équations qui décrivent le sous système rapide sont définies en supposant que la composante rapide de la partie lente s'éteint dans un temps très court. Ceci revient à poser dans (II.56) la condition suivante :

$$\mathbf{L}_r(n+1) = 0 \quad (\text{II.57})$$

Nous déduisons alors l'équation de la partie rapide :

$$\mathbf{R}_r(n+1) = \mathbf{R}_r(n) \left(\begin{matrix} \dagger \\ \dagger \end{matrix}_{22} - \begin{matrix} \dagger \\ \dagger \end{matrix}_{21} \begin{matrix} \dagger \\ \dagger \end{matrix}_{11}^{-1} \begin{matrix} \dagger \\ \dagger \end{matrix}_{12} \right) \quad (\text{II.58})$$

avec :

$$\mathbf{R}_r(0) = \mathbf{R}(0) - \mathbf{L}(0) \begin{matrix} \dagger \\ \dagger \end{matrix}_{11}^{-1} \begin{matrix} \dagger \\ \dagger \end{matrix}_{12}$$

En regroupant les résultats obtenus, nous obtenons le système découplé suivant :

$$\begin{cases} \begin{bmatrix} \mathbf{L}_1(n+1) & \mathbf{R}_r(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{L}_1(n) & \mathbf{R}_r(n) \end{bmatrix} \begin{pmatrix} \begin{matrix} \dagger \\ \dagger \end{matrix}_1 & 0 \\ 0 & \begin{matrix} \dagger \\ \dagger \end{matrix}_r \end{pmatrix} \\ \mathbf{R}_1(n) = \mathbf{L}_1(n) \begin{matrix} \dagger \\ \dagger \end{matrix}_{11}^{-1} \begin{matrix} \dagger \\ \dagger \end{matrix}_{12} \end{cases} \quad (\text{II.59})$$

où :

$$\begin{aligned} \begin{matrix} \dagger \\ \dagger \end{matrix}_1 &= \begin{matrix} \dagger \\ \dagger \end{matrix}_{11} \\ \begin{matrix} \dagger \\ \dagger \end{matrix}_r &= \begin{matrix} \dagger \\ \dagger \end{matrix}_{22} - \begin{matrix} \dagger \\ \dagger \end{matrix}_{21} \begin{matrix} \dagger \\ \dagger \end{matrix}_{11}^{-1} \begin{matrix} \dagger \\ \dagger \end{matrix}_{12} \end{aligned} \quad (\text{II.60})$$

III.3.3. Analyse des systèmes découplés

Dans le cas de ce découplage, les matrices $\begin{matrix} \dagger \\ \dagger \end{matrix}_1$ et $\begin{matrix} \dagger \\ \dagger \end{matrix}_r$ ne sont pas stochastiques. Les sous-systèmes lent et rapides ne peuvent pas être résolus directement à l'aide des méthodes connues pour les chaînes de Markov.

III.4. Application de la méthode de découplage de Bennis

III.4.1. Forme singulièrement perturbée

Dans la démarche de son étude, Bennis n'utilise pas de modélisation sous forme singulièrement perturbée.

Reprenons l'équation fondamentale sous forme partitionnée (II.28) :

$$\begin{cases} \mathbf{L}(n+1) = \mathbf{L}(n) \dot{\dagger}_{11} + \mathbf{R}(n) \dot{\dagger}_{21} \\ \mathbf{R}(n+1) = \mathbf{L}(n) \dot{\dagger}_{12} + \mathbf{R}(n) \dot{\dagger}_{22} \end{cases} \text{ de conditions initiales } \mathbf{L}(0), \mathbf{R}(0) \quad (\text{II.61})$$

III.4.2. Découplage des dynamiques

a) Partie lente

Les probabilités des états de la partie rapide de la chaîne de Markov tendent rapidement vers leur régime permanent. Dans ce cas, nous pouvons écrire les approximations suivantes :

$$\mathbf{R}(n+1) = \mathbf{R}(n) = \mathbf{R}_1(n), \quad (\text{II.62})$$

ce qui nous donne l'évolution des variables en régime quasi-permanent (variables lentes) :

$$\begin{cases} \mathbf{L}_1(n+1) = \mathbf{L}_1(n) \dot{\dagger}_{11} + \mathbf{R}_1(n) \dot{\dagger}_{21} \\ \mathbf{R}_1(n) = \mathbf{L}_1(n) \dot{\dagger}_{12} + \mathbf{R}_1(n) \dot{\dagger}_{22} \end{cases} \quad (\text{II.63})$$

La matrice $(\mathbf{I} - \dot{\dagger}_{22})$ est inversible, par conséquent, l'évolution de la partie lente du système est approchée par un sous système lent décrit par l'équation suivante :

$$\mathbf{L}_1(n+1) = \mathbf{L}_1(n) \left(\dot{\dagger}_{11} + \dot{\dagger}_{12} (\mathbf{I} - \dot{\dagger}_{22})^{-1} \dot{\dagger}_{21} \right) \quad (\text{II.64})$$

avec les conditions initiales $\mathbf{L}_1(0) = \mathbf{L}(0)$,

La composante lente de la distribution rapide $\mathbf{R}_1(n)$ peut être calculée par l'expression :

$$\mathbf{R}_1(n) = \mathbf{L}_1(n) \dot{\dagger}_{12} (\mathbf{I} - \dot{\dagger}_{22})^{-1} \quad (\text{II.65})$$

avec les conditions initiales $\mathbf{R}_1(0) = \mathbf{L}(0) \dot{\dagger}_{12} (\mathbf{I} - \dot{\dagger}_{22})^{-1}$

b) Partie rapide

Les transitoires rapides des probabilités d'état sont obtenues par les différences :

$$\mathbf{R}_r(n) = \mathbf{R}(n) - \mathbf{R}_1(n) \quad (\text{II.66})$$

avec $\mathbf{L}_1(n)$ et $\mathbf{R}_1(n)$ données par (II.64) et (II.65).

En introduisant la relation (II.66) dans le système initial (II.61) nous pouvons écrire :

$$\mathbf{R}_l(n+1) + \mathbf{R}_r(n+1) = \mathbf{L}(n) \dot{\dagger}_{12} + (\mathbf{R}_l(n) + \mathbf{R}_r(n)) \dot{\dagger}_{22}$$

En faisant par la suite la différence entre ce système et le système lent (II.63),

$$\mathbf{R}_l(n+1) = \mathbf{L}_1(n) \dot{\dagger}_{12} + \mathbf{R}_l(n) \dot{\dagger}_{22} \quad (\text{II.67})$$

nous obtenons, en considérant l'hypothèse $\mathbf{L}(n) = \mathbf{L}_1(n)$, le sous système rapide du modèle (II.61) qui s'écrit :

$$\mathbf{R}_r(n+1) = \mathbf{R}_r(n) \dot{\dagger}_{22} \quad (\text{II.68})$$

avec :

$$\mathbf{R}_r(0) = \mathbf{R}(0) - \mathbf{L}(0) \dot{\dagger}_{12} \left(\mathbf{I} - \dot{\dagger}_{22} \right)^{-1}$$

La chaîne de Markov découplée se présente comme suit :

$$\begin{cases} \begin{bmatrix} \mathbf{L}_l(n+1) & \mathbf{R}_r(n+1) \end{bmatrix} = \begin{bmatrix} \mathbf{L}_l(n) & \mathbf{R}_r(n) \end{bmatrix} \begin{pmatrix} \dot{\dagger}_1 & 0 \\ 0 & \dot{\dagger}_r \end{pmatrix} \\ \mathbf{R}_l(n) = \mathbf{L}_1(n) \dot{\dagger}_{12} \left(\mathbf{I} - \dot{\dagger}_{22} \right)^{-1} \end{cases} \quad (\text{II.69})$$

où :

$$\dot{\dagger}_1 = \dot{\dagger}_{11} + \dot{\dagger}_{12} \left(\mathbf{I} - \dot{\dagger}_{22} \right)^{-1} \dot{\dagger}_{21} \quad (\text{II.70})$$

$$\dot{\dagger}_r = \dot{\dagger}_{22} \quad (\text{II.71})$$

III.4.3. Analyse des systèmes découplés

a) Système lent

La matrice $\dot{\dagger}_1$ correspondant au système lent découplé est stochastique.

-Démonstration-

1°. La somme de chaque ligne de la matrice $\dot{\dagger}_1$ est égale à 1

La matrice de transition initiale $\dot{\dagger}$ étant stochastique, nous pouvons écrire, en gardant les notations (II.42), la relation suivante :

$$\mathbf{1}_{r_2} - \dot{\dagger}_{22} \mathbf{1}_{r_2} = \dot{\dagger}_{21} \mathbf{1}_{r_1} \quad (\text{II.72})$$

En multipliant à gauche par $\dot{\dagger}_{12}$, il nous vient :

$$\dot{\mathbf{T}}_{12} \mathbf{1}_{r_2} = \dot{\mathbf{T}}_{12} \left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} \quad (\text{II.73})$$

En rajoutant $\mathbf{1}_{r_1}$ aux deux membres de l'égalité, nous obtenons :

$$\mathbf{1}_{r_1} - \dot{\mathbf{T}}_{12} \mathbf{1}_{r_2} + \dot{\mathbf{T}}_{12} \left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} = \mathbf{1}_{r_1} \quad (\text{II.74})$$

La propriété de stochasticté de la matrice initiale nous permet d'écrire :

$$\dot{\mathbf{T}}_{11} \mathbf{1}_{r_1} + \dot{\mathbf{T}}_{12} \left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} \dot{\mathbf{T}}_{21} \mathbf{1}_{r_1} = \mathbf{1}_{r_1} \quad (\text{II.75})$$

donc :

$$\dot{\mathbf{T}}_1 \mathbf{1}_{r_1} = \mathbf{1}_{r_1} \quad (\text{II.76})$$

La somme des éléments de chacune des lignes de $\dot{\mathbf{T}}_1$ est égale à 1.

2°. Les éléments de la matrice $\dot{\mathbf{T}}_1$ sont non-négatifs

Nous avons $\|\dot{\mathbf{T}}_{22}\| < 1$, ce qui nous permet d'effectuer un développement en série de la matrice $(\mathbf{I} - \dot{\mathbf{T}}_{22})^{-1}$. Nous obtenons par conséquent :

$$\left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} = \mathbf{I} + \dot{\mathbf{T}}_{22} + \dot{\mathbf{T}}_{22}^2 + \dot{\mathbf{T}}_{22}^3 + \dots = \sum_{p=0}^{\infty} \dot{\mathbf{T}}_{22}^p \quad (\text{II.77})$$

Comme $\dot{\mathbf{T}}_{22}$ est une matrice non-négative, la série sera positive, donc nous pouvons écrire

$$\left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} \geq 0 \quad (\text{II.78})$$

Comme les matrices $\dot{\mathbf{T}}_{11}$, $\dot{\mathbf{T}}_{12}$ et $\dot{\mathbf{T}}_{21}$ sont non-négative, il en sera de même pour $\dot{\mathbf{T}}_1$:

$$\dot{\mathbf{T}}_1 = \dot{\mathbf{T}}_{11} + \dot{\mathbf{T}}_{12} \left(\mathbf{I} - \dot{\mathbf{T}}_{22} \right)^{-1} \dot{\mathbf{T}}_{21} \geq 0 \quad (\text{II.79})$$

Par conséquent, la matrice de transition du sous-système lent est une matrice stochastique.

□

b) Système rapide

La matrice $\dot{\mathbf{T}}_r$ correspondant au système rapide découplé n'est pas une matrice stochastique. Par conséquent, la résolution de la partie rapide ne peut pas se faire en utilisant les méthodes caractéristiques aux chaînes de Markov.

III.5. Utilisation des perturbations singulières en continu pour la simplification des chaînes de Markov ergodiques

Les techniques de perturbations singulières en continu présentent l'avantage de l'unicité de la modélisation. Pour pouvoir simplifier les chaînes de Markov par l'application des perturbations singulières en continu, nous proposons l'utilisation des transformations homographiques. Cette transformation permet d'associer à la chaîne de Markov à temps discret et état discret, un processus de Markov à temps continu et état discret. Ce dernier conserve toutes

les propriétés initiales du système étudié, en particulier les dynamiques ; par conséquent, nous pouvons lui appliquer les perturbations singulières en continu.

III.5.1. Etude de la transformation homographique appliquée aux chaînes de Markov

a) Définitions

a.1. Transformation homographique directe

Soit l'équation fondamentale d'une chaîne de Markov finie :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \mathbf{\dagger} \quad (\text{II.80})$$

Les changements de variables associées aux transformations homographiques directes sont les suivantes :

$$\mathbf{P}(n+1) - \mathbf{P}(n) \tilde{\mathbf{P}} \quad \text{et} \quad \frac{\mathbf{P}(n+1) + \mathbf{P}(n)}{2} \tilde{\mathbf{P}} \quad (\text{II.81})$$

Il nous vient alors le système continu :

$$\dot{\tilde{\mathbf{P}}} = \tilde{\mathbf{P}} \mathbf{\dagger} \ddot{\mathbf{Y}} \quad \text{où} \quad \mathbf{\dagger} \ddot{\mathbf{Y}} = \mathbf{H}(\mathbf{\dagger}) = \left(\frac{\mathbf{\dagger} + \mathbf{I}}{2} \right)^{-1} (\mathbf{\dagger} - \mathbf{I}) \quad (\text{II.82})$$

avec $\mathbf{\ddot{Y}}$ le vecteur de probabilités d'état et $\mathbf{\dagger} \ddot{\mathbf{Y}}$ la matrice de transition du processus continu.

a.2. Transformation homographique inverse

Inversement, nous pouvons associer à un système continu de type

$$\dot{\tilde{\mathbf{P}}} = \tilde{\mathbf{P}} \mathbf{\dagger} \ddot{\mathbf{Y}} \quad (\text{IV.83})$$

un système discret, en effectuant les transformations homographiques inverses :

$$\tilde{\mathbf{P}} + \frac{\tilde{\mathbf{P}}}{2} \tilde{\mathbf{P}} \mathbf{\dagger} \ddot{\mathbf{Y}} \quad \text{et} \quad \tilde{\mathbf{P}} - \frac{\tilde{\mathbf{P}}}{2} \tilde{\mathbf{P}} \mathbf{\dagger} \ddot{\mathbf{Y}} \quad (\text{II.84})$$

Nous obtenons ainsi l'équation suivante :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \mathbf{\dagger} \quad \text{où} \quad \mathbf{\dagger} = \mathbf{H}^{-1}(\mathbf{\dagger} \ddot{\mathbf{Y}}) = \left(\mathbf{I} - \frac{\mathbf{\dagger} \ddot{\mathbf{Y}}}{2} \right)^{-1} \left(\mathbf{I} + \frac{\mathbf{\dagger} \ddot{\mathbf{Y}}}{2} \right) \quad (\text{II.85})$$

avec la notation $\mathbf{P}(n)$ pour le vecteur d'état à l'instant n .

a.3. Etude graphique de la transformation homographique appliquée aux chaînes de Markov

La transformation homographique fait correspondre au disque unité du plan complexe discret z , le demi-plan gauche du plan complexe continu w (figure II.4).

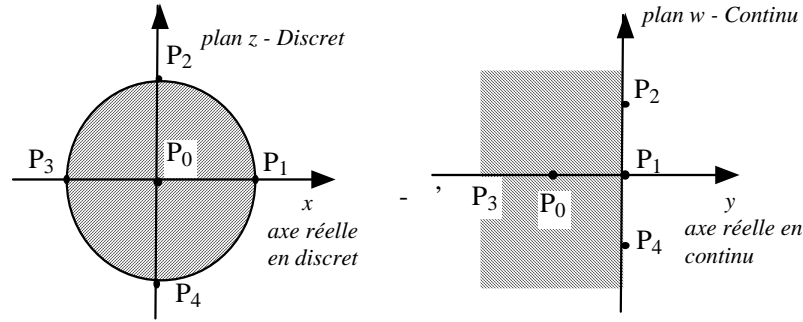


Fig. II.4. Transformation homographique sur les domaines complexes.

Le passage entre les deux plans complexes est donné par la relation [ELM, 85]:

$$w = 2(z-1)/(z+1) \quad (\text{II.86})$$

Un disque de Gershgorine $D_1(C_1, R_1)$ du système discret (figure II.5) se transforme en un disque $D_2(C_2, R_2)$ du système continu. Les limites de ces domaines circulaires se transformeront elles, suivant la loi $l_2 = 2(l_1-1)/(l_1+1)$, avec l_1 la limite du domaine de Gershgorine en discret, et l_2 la limite correspondante en continu.

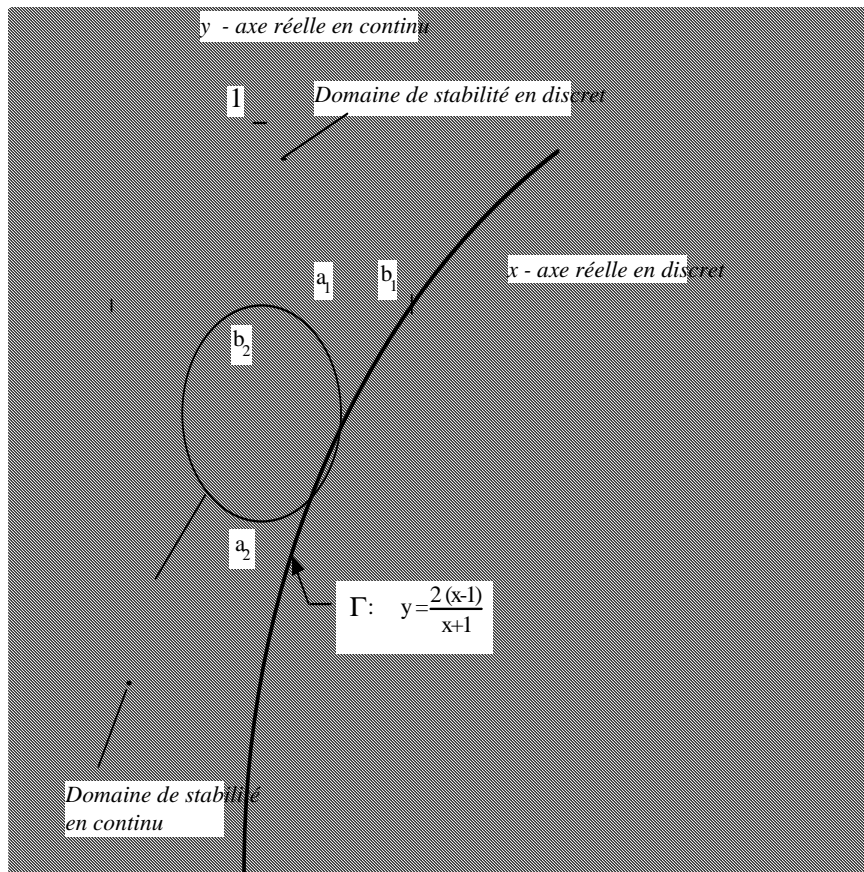


Fig. II.5. Limites des domaines des valeurs propres après transformation homographique.

Nous allons noter par Γ la fonction correspondante :

$$\Gamma(x) = \frac{2(x-1)}{x+1} \quad (\text{II.87})$$

b) Propriétés

Nous allons présenter une étude des éléments de l'équation (II.82) du système continu à espace d'état discret obtenu après l'application des transformations homographiques sur une chaîne de Markov homogène finie (II.80) de dimension r .

Théorème 1

La transformation homographique directe (inverse) fait correspondre à une chaîne de Markov (processus de Markov), un processus de Markov (respectivement une chaîne de Markov).

-Démonstration-

1°. L'application de la transformation homographique sur une chaîne de Markov, donne un processus de Markov.

Ceci revient à démontrer que la matrice $\ddot{\Upsilon}$ du système continu (II.82) présente les propriétés d'un générateur :

1.1. La somme des éléments de chaque ligne de la matrice $\ddot{\Upsilon}$ est égale à 0.

Suite à la stochasticté de la matrice $\dot{\Upsilon}$, nous obtenons :

$$\dot{\Upsilon} \mathbf{1}_r = \left(\frac{\dot{\Upsilon} + \mathbf{I}}{2} \right)^{-1} (\dot{\Upsilon} - \mathbf{I}) \mathbf{1}_r = \left(\frac{\dot{\Upsilon} + \mathbf{I}}{2} \right)^{-1} (\dot{\Upsilon} \mathbf{1}_r - \mathbf{I} \mathbf{1}_r) = \left(\frac{\dot{\Upsilon} + \mathbf{I}}{2} \right)^{-1} (\mathbf{1}_r - \mathbf{1}_r) = \mathbf{0}_r \quad (\text{II.88})$$

ce qui donne la relation suivante :

$$\ddot{\Upsilon} \mathbf{1}_r = \mathbf{0}_r \quad (\text{II.89})$$

1.2. Les éléments hors diagonaux de la matrice $\ddot{\Upsilon}$ sont non-négatifs

La courbe Γ (figure II.5) transforme le domaine $[-1,1]$ des parties réelles des valeurs propres de la matrice $\dot{\Upsilon}$ en l'intervalle $[-\infty,0]$, dans lequel se trouvent les parties réelles des valeurs propres de la matrice $\ddot{\Upsilon}$.

Les centres des disques de Gershgorine de la chaîne de Markov sont donnés par les probabilités de transition de type p_{ii} ($i = 1, \dots, r$). Nous avons :

$$p_{ii} \in [0, 1], \quad \forall i = 1, \dots, r.$$

Par conséquent, les éléments p_{ij} de la matrice de transition $\ddot{\Upsilon} = [p_{ij}]_{i=1, \dots, r}$ du processus continu vérifient :

$$p_{ii} \in [-2, 0], \quad \forall i = 1, \dots, r.$$

Comme le point (1,0) appartient à tous les disques de Gershgorine construits à partir des lignes de \tilde{P} , le point (0,0) appartiendra aux disques similaires associés à la matrice $\tilde{P} - \tilde{P}$. Nous obtenons ainsi :

$$R_i + \tilde{p}_{ii} = \sum_{j=1, i \neq j}^r |\tilde{p}_{ij}| + \tilde{p}_{ii} = 0, \quad \forall i = 1, 2, \dots, r. \quad (\text{II.90})$$

Sachant que les sommes des lignes de la matrice $\tilde{P} - \tilde{P}$ sont égales à 0 (relation (II.89)), nous pouvons écrire l'égalité suivante :

$$\sum_{i=1, i \neq j}^r |\tilde{p}_{ij}| = \sum_{i=1, i \neq j}^r \tilde{p}_{ij}, \quad \forall i = 1, 2, \dots, r. \quad (\text{II.91})$$

Par conséquent, tous les éléments hors diagonaux de $\tilde{P} - \tilde{P}$ sont non-négatifs. La matrice de transition du système continu est un *générateur*. Ce système représente donc un processus de Markov.

2°. L'application des transformations homographiques inverses sur un processus de Markov donne une chaîne de Markov.

Ceci revient à démontrer que la matrice \tilde{P} du système discret (II.85) est une matrice stochastique.

2.1. La somme des éléments de chaque ligne de \tilde{P} est égale à 1.

Nous utilisons les propriétés de générateur de la matrice $\tilde{P} - \tilde{P}$ (de dimension $r \times r$) :

$$\tilde{P} \mathbf{1}_r = \mathbf{0}_r \quad (\text{II.92})$$

A partir de cette propriété, nous pouvons écrire

$$\left(\mathbf{I} - \frac{\tilde{P}}{2} \right) \mathbf{1}_r = \left(\mathbf{I} + \frac{\tilde{P}}{2} \right) \mathbf{1}_r \quad (\text{II.93})$$

Vu la définition (II.85) de la transformation homographique inverse, la matrice $\mathbf{I} - \tilde{P}$ est régulière et nous obtenons :

$$\left(\mathbf{I} - \frac{\tilde{P}}{2} \right)^{-1} \left(\mathbf{I} + \frac{\tilde{P}}{2} \right) \mathbf{1}_r = \mathbf{1}_r \quad (\text{II.94})$$

ou, selon les relations de la transformation homographique :

$$\tilde{P} \mathbf{1}_r = \mathbf{1}_r \quad (\text{II.95})$$

La somme des lignes de la matrice de transition \tilde{P} du système discret est égale à 1.

2.2. Les éléments de la matrice \tilde{P} sont non-négatifs

Les transformations homographiques inverses des limites des domaines géométriques se font suivant la même courbe Γ (figure II.5). Cette fois la transformation se fera du processus markovien continu (y) vers le processus correspondant en discret (x).

En effectuant un raisonnement similaire à 1.2, nous concluons que les éléments diagonaux de la matrice $\ddagger = [p_{ij}]_{i,j=1,\dots,r}$ (centres des disques de Gershgorine) vérifient la condition :

$$p_{ii} \in]0, 1] , \quad \forall i = 1, 2, \dots, r. \quad (\text{II.96})$$

En prenant en compte le fait que les disques de Gershgorine du système sont tangents (figure II.5), nous obtenons :

$$\sum_{i=1, i \neq j}^r |p_{ij}| = \sum_{i=1, i \neq j}^r p_{ij} , \quad \forall i = 1, 2, \dots, r. \quad (\text{II.97})$$

ce qui démontre que tous les éléments de la matrice \ddagger sont non-négatifs.

Par conséquent, compte tenu des résultats de 2.1, la matrice \ddagger est stochastique. \square

Une méthode de localisation des valeurs propres, basée sur la séparation disques de Gershgorine de la matrice $\ddagger\ddot{Y}$ est applicable aussi dans le cas des systèmes continus. Par conséquent, nous pouvons énoncer le théorème suivant :

Théorème 2

La transformation homographique conserve la nature des dynamiques d'une chaîne de Markov à double échelle de temps.

-Démonstration-

Pour démontrer ce théorème, nous utilisons le graphe de la transformation homographique (figure II.6). Les deux disques D_1 et D'_1 matérialisent la localisation des modes rapides et respectivement lents de la chaîne de Markov. Ces disques correspondent aux disques D_2 et D'_2 du système continu obtenu par transformations homographiques.

Les éléments de la matrice de transition d'une chaîne de Markov sont toujours positifs. Au cas où la propriété de double échelle de temps existe, les domaines des disques de Gershgorine seront caractérisés par des limites qui vérifient la relation $b_1 < a'_1$. b_1 représente la limite inférieure du domaine rapide D_1 et a'_1 , la limite supérieure du domaine lent D'_1 .

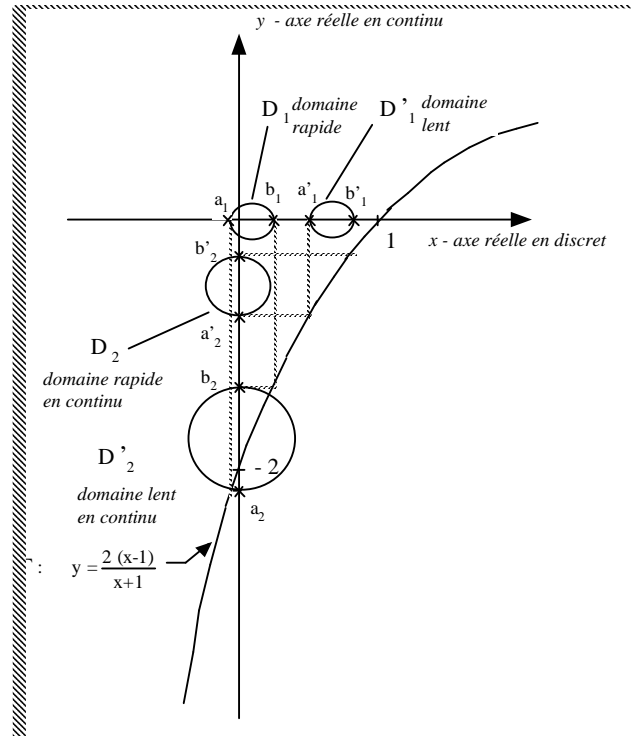


Fig.II.6. Domaines lent et rapide après transformation homographique.

La conservation de la nature des dynamiques nous vient de la croissance de la fonction qui caractérise la transformation homographique : $y = 2(x-1)/(x+1)$. Nous aurons donc :

$$b_2 = 2(b_1 - 1)/(b_1 + 1) < 2(a'_1 - 1)/(a'_1 + 1) = a'_2 \quad (\text{II.98})$$

ce qui implique que le domaine rapide transformé sera le domaine rapide en continu, et le domaine lent transformé sera le domaine lent en continu. La double échelle de temps de la chaîne de Markov (caractérisée par $b_1 \ll a'_1$) sera toujours présente dans le système continu correspondant ($b_2 \ll a'_2$).

□

Remarque

Les transformations homographiques conservent les conditions initiales [ELM, 85]. Si les conditions initiales de la chaîne de Markov initiales sont $\mathbf{P}(0)$, ceci implique les conditions initiales $\mathbf{P}_0 = \mathbf{P}(0)$ pour le processus de Markov correspondant. De même, pour un processus markovien continu ayant les conditions initiales \mathbf{P}_0 , nous obtenons une chaîne de Markov de conditions initiales $\mathbf{P}(0) = \mathbf{P}_0$.

III.5.2. Utilisation conjointe des transformations homographiques et des perturbations singulières en continu pour le découplage des chaînes de Markov à double échelle de temps

a) Principe de la méthode

La démarche proposée est la suivante :

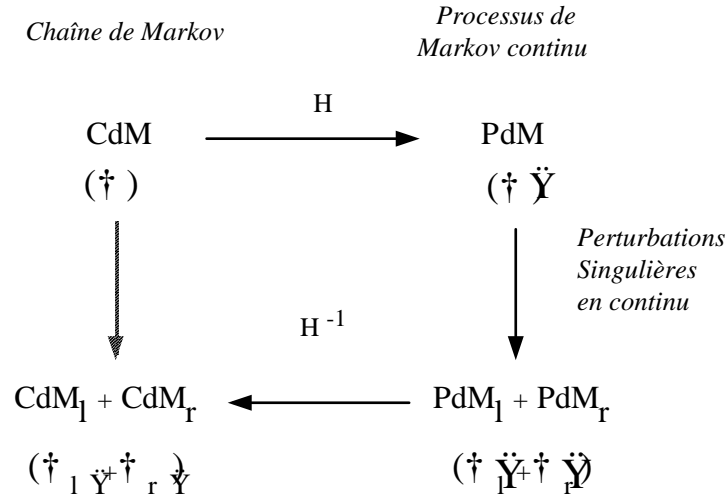


Fig. II.7. Application des perturbations singulières conjointement aux transformations homographiques pour les chaînes de Markov.

où : \ddagger représente la matrice de transition de la chaîne de Markov,
 $\ddagger \ddot{Y}$ représente la matrice de transition du processus de Markov après Tr.Homographiques,
 $\ddagger \ddot{Y}_l$ est la matrice de transition de la partie lente du processus de Markov,
 $\ddagger \ddot{Y}_r$ la matrice de transition de la partie rapide du processus de Markov,
 $\ddagger_l \ddot{Y}$ matrice de transition de la partie lente de la chaîne de Markov après Tr. H. inverses
 $\ddagger_r \ddot{Y}$ matrice de transition de la partie rapide de la chaîne de Markov après Tr.H.inverses

Par application de la transformation homographique directe sur la forme partitionnée de l'équation fondamentale de la chaîne de Markov à double échelle de temps :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \ddagger_{11} & \ddagger_{12} \\ \ddagger_{21} & \ddagger_{22} \end{pmatrix} \quad (\text{II.99})$$

nous obtenons le processus de Markov à double échelle de temps, représenté par :

$$[\tilde{\mathbf{L}} \quad \tilde{\mathbf{R}}] = [\tilde{\mathbf{L}} \quad \tilde{\mathbf{R}}] \begin{pmatrix} \tilde{\ddagger}_{11} & \tilde{\ddagger}_{12} \\ \tilde{\ddagger}_{21} & \tilde{\ddagger}_{22} \end{pmatrix} \quad (\text{II.100})$$

Dans ce cas, nous écrivons ce système sous la forme singulièrement perturbée suivante :

$$[\tilde{\mathbf{L}} \quad \varepsilon \tilde{\mathbf{R}}] = [\tilde{\mathbf{L}} \quad \tilde{\mathbf{R}}] \begin{pmatrix} \tilde{\ddagger}_{11} & \tilde{\ddagger}_{12}^* \\ \tilde{\ddagger}_{21} & \tilde{\ddagger}_{22}^* \end{pmatrix} \quad (\text{II.101})$$

où $\tilde{\ddagger}_{12}^* = \varepsilon \tilde{\ddagger}_{12}$, $\tilde{\ddagger}_{22}^* = \varepsilon \tilde{\ddagger}_{22}$, $\varepsilon = \|\tilde{\ddagger}_{22}\| / \|\tilde{\ddagger}_{11}\|$, $\varepsilon \in [0, 1]$.

L'application des techniques de perturbations singulières en continu exige que la matrice $\ddagger \ddot{\mathbf{Y}}_{22}$ soit régulière et permet d'obtenir les sous-systèmes lent et rapide découplés :

$$\begin{cases} \begin{bmatrix} \tilde{\mathbf{L}}_1 & \tilde{\mathbf{R}}_r \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{L}}_1 & \tilde{\mathbf{R}}_r \end{bmatrix} \begin{pmatrix} \tilde{\ddagger}_1 & 0 \\ 0 & \tilde{\ddagger}_r \end{pmatrix} \\ \tilde{\mathbf{R}}_1 = -\tilde{\mathbf{L}}_1 \ddagger_{12} \ddagger_{22}^{-1} \end{cases} \quad (\text{II.102})$$

avec :

$$\begin{aligned} \ddagger \ddot{\mathbf{Y}}_1 &= \ddagger \ddot{\mathbf{Y}}_{11} - \ddagger \ddot{\mathbf{Y}}_{12} \ddagger \ddot{\mathbf{Y}}_{22}^{-1} \ddagger \ddot{\mathbf{Y}}_{21} \\ \ddagger \ddot{\mathbf{Y}}_r &= \ddagger \ddot{\mathbf{Y}}_{22} \end{aligned}$$

et les conditions initiales $\mathbf{L} \ddot{\mathbf{Y}}_1(0) = \mathbf{L} \ddot{\mathbf{Y}}(0)$, $\mathbf{R} \ddot{\mathbf{Y}}_r(0) = \mathbf{R} \ddot{\mathbf{Y}}(0) - \mathbf{L} \ddot{\mathbf{Y}}(0) \ddagger \ddot{\mathbf{Y}}_{12} \ddagger \ddot{\mathbf{Y}}_{22}^{-1}$.

Pour obtenir les chaînes de Markov lentes et rapides correspondant aux systèmes découplés (II.102), nous leur appliquons les transformations homographiques inverses :

$$\begin{cases} \begin{bmatrix} \tilde{\mathbf{L}}_1^{(n+1)} & \tilde{\mathbf{R}}_r^{(n+1)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{L}}_1^{(n)} & \tilde{\mathbf{R}}_r^{(n)} \end{bmatrix} \begin{pmatrix} \tilde{\ddagger}_1^{(n)} & 0 \\ 0 & \tilde{\ddagger}_r^{(n)} \end{pmatrix} \\ \tilde{\mathbf{R}}_1^{(n)} = -2 \tilde{\mathbf{L}}_1^{(n)} \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \ddagger_{12} \left(\tilde{\ddagger}_r^{(n)} - \mathbf{I} \right)^{-1} \end{cases} \quad (\text{II.103})$$

avec :

$$\begin{aligned} \tilde{\ddagger}_1^{(n)} &= \left(\mathbf{I} + \ddagger_{12} \overline{\ddagger}_{22} \ddagger_{21} \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \right)^{-1} \left(\ddagger_{11} - \ddagger_{12} \overline{\ddagger}_{22} \ddagger_{21} \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \right) \\ \overline{\ddagger}_{22} &= \left(\left(\ddagger_{22} - \mathbf{I} \right) - \ddagger_{21} \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \ddagger_{12} \right)^{-1} \\ \tilde{\ddagger}_r^{(n)} &= \ddagger_{22} - \ddagger_{21} \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \ddagger_{12} \end{aligned}$$

et les conditions initiales :

$$\begin{aligned} \mathbf{L}_1 \ddot{\mathbf{Y}}(0) &= \mathbf{L}(0) \\ \tilde{\mathbf{R}}_r^{(0)} &= \mathbf{R}(0) + 2 \mathbf{L}(0) \left(\ddagger_{11} + \mathbf{I} \right)^{-1} \ddagger_{12} \left(\tilde{\ddagger}_r^{(0)} - \mathbf{I} \right)^{-1} \end{aligned}$$

Remarque

L'utilisation du découplage en continu nous donne une partie lente caractérisée par une grande précision. Ceci est dû à la prise en compte de l'évolution de la partie rapide dans le découplage de cette partie. Après l'application des transformations homographiques inverses, la partie lente obtenue en discret sera elle aussi très précise.

□

b) Analyse des systèmes découplés

Dans un processus de Markov à double échelle de temps, le système lent obtenu par application des perturbations singulières en continu représente un processus de Markov.

-Démonstration-

1°. La somme des éléments de chaque ligne de la matrice $\ddagger\ddot{\mathbf{Y}}_1$ est égale à 0

La matrice de transition du processus markovien continu est une matrice dont la somme des lignes est égale à zéro. Nous pouvons écrire par conséquent :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_{11} \mathbf{1}_1 + \begin{matrix} \sim \\ \ddagger \end{matrix}_{12} \mathbf{1}_r = \mathbf{0}_1 \quad (a) \quad (II.104)$$

et $\begin{matrix} \sim \\ \ddagger \end{matrix}_{21} \mathbf{1}_1 + \begin{matrix} \sim \\ \ddagger \end{matrix}_{22} \mathbf{1}_r = \mathbf{0}_r \quad (b)$

Vu que la matrice $\ddagger\ddot{\mathbf{Y}}_{22}$ est régulière (condition imposée pour appliquer les perturbations singulières en continu), nous pouvons écrire (II.104)-a sous la forme suivante :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_{11} \mathbf{1}_1 + \begin{matrix} \sim \\ \ddagger \end{matrix}_{12} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22}^{-1} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22} \mathbf{1}_r = \mathbf{0}_1 \quad (II.105)$$

ou encore :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_{11} \mathbf{1}_1 - \begin{matrix} \sim \\ \ddagger \end{matrix}_{12} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22}^{-1} \left(\mathbf{0}_r - \begin{matrix} \sim \\ \ddagger \end{matrix}_{22} \mathbf{1}_r \right) = \mathbf{0}_1 \quad (II.106)$$

En utilisant l'égalité (II.104) - b, nous obtenons la relation suivante :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_{11} \mathbf{1}_1 - \begin{matrix} \sim \\ \ddagger \end{matrix}_{12} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22}^{-1} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22} \begin{matrix} \sim \\ \ddagger \end{matrix}_{21} \mathbf{1}_1 = \mathbf{0}_1 \quad (II.107)$$

soit :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_1 \mathbf{1}_1 = \mathbf{0}_1 \quad (II.108)$$

2°. Les éléments hors - diagonaux de la matrice $\ddagger\ddot{\mathbf{Y}}_1$ sont non-négatifs

La matrice du sous-système lent continu $\ddagger\ddot{\mathbf{Y}}_1$ est donnée par la relation :

$$\begin{matrix} \sim \\ \ddagger \end{matrix}_1 = \begin{matrix} \sim \\ \ddagger \end{matrix}_{11} - \begin{matrix} \sim \\ \ddagger \end{matrix}_{12} \begin{matrix} \sim \\ \ddagger \end{matrix}_{22}^{-1} \begin{matrix} \sim \\ \ddagger \end{matrix}_{21} \quad (II.109)$$

Les matrices $\ddagger\ddot{\mathbf{Y}}_{11}$ et $\ddagger\ddot{\mathbf{Y}}_{22}$ sont des sous matrices de la matrice $\ddagger\ddot{\mathbf{Y}}$. Les éléments de leurs diagonales principales sont négatifs et leurs éléments hors diagonaux sont non-négatifs. Une telle structure correspond à la classe des M matrices et l'inverse d'une telle matrice (en occurrence $\ddagger\ddot{\mathbf{Y}}_{22}^{-1}$) est une matrice positive (théorème 6 [FAN, 62]). Nous concluons que tous les éléments diagonaux de $\ddagger\ddot{\mathbf{Y}}_1$ sont négatifs.

Pour démontrer que ses éléments hors-diagonaux sont non-négatifs, nous démontrons que $\ddagger\ddot{\mathbf{Y}}_1 + 2 \mathbf{I}$ est positive. La relation (II.109) prend la forme :

$$\tilde{\ddagger}_1 + 2\mathbf{I} = \tilde{\ddagger}_{11} + 2\mathbf{I} + \tilde{\ddagger}_{12} \left[2\mathbf{I} - \left(2\mathbf{I} + \tilde{\ddagger}_{22} \right) \right]^{-1} \tilde{\ddagger}_{21} = \mathbf{E}_{11} + \tilde{\ddagger}_{12} (2\mathbf{I} - \mathbf{E}_{22})^{-1} \tilde{\ddagger}_{21}$$

avec \mathbf{E}_{11} et \mathbf{E}_{22} des matrices positives ayant la propriété que la somme des éléments de chaque ligne est plus petite que 2. Nous pouvons donc écrire :

$$\tilde{\ddagger}_1 + 2\mathbf{I} = \mathbf{E}_{11} + 2\tilde{\ddagger}_{12} \left(\mathbf{I} - \frac{\mathbf{E}_{22}}{2} \right)^{-1} \tilde{\ddagger}_{21} \quad (\text{II.110})$$

Comme la norme infinie de $\frac{\mathbf{E}_{22}}{2}$ est inférieure à 1, nous pouvons développer (II.110) sous la forme suivante :

$$\tilde{\ddagger}_1 + 2\mathbf{I} = \mathbf{E}_{11} + 2\tilde{\ddagger}_{12} \left(\mathbf{I} + \frac{\mathbf{E}_{22}}{2} + \frac{(\mathbf{E}_{22})^2}{2^2} + \frac{(\mathbf{E}_{22})^3}{2^3} + \dots \right) \tilde{\ddagger}_{21} \quad (\text{II.111})$$

Nous déduisons ainsi que $\ddagger\ddot{\mathbf{Y}}_1 + 2\mathbf{I} \geq 0$. Par conséquent, la matrice $\ddagger\ddot{\mathbf{Y}}_1$ est elle aussi un *générateur*. La partie lente en continu regroupe toutes les propriétés essentielles du processus continu initial. □

Dans une chaîne de Markov à double échelle de temps, le sous-système lent obtenu par application des transformations homographiques conjointement aux perturbations singulières en continu représente une chaîne de Markov.

- Démonstration -

Par application des transformations homographiques directes sur les chaînes de Markov, nous obtenons un processus markovien continu avec la même nature des dynamiques. Suite à la décomposition singulièrement perturbée du système continu, la partie lente représente à elle toute seule un processus de Markov. Par conséquent, conformément au *théorème 1*, le système lent discret obtenu après transformations homographiques inverses représente une chaîne de Markov. □

III.6. Résolution des systèmes découplés obtenus par les différentes méthodes

L'application des différentes techniques de perturbations singulières nous a permis d'obtenir des systèmes découplés représentés par les parties lente et rapide qui s'écrivent sous la forme suivante :

$$\begin{aligned} \text{partie lente : } & \begin{cases} \mathbf{L}'_1(n+1) = \mathbf{L}'_1(n) \ddagger_1 \\ \mathbf{R}'_1(n) = \mathbf{L}'_1(n) \ddagger_{\text{infl}} \end{cases} \\ \text{partie rapide : } & \mathbf{R}'_r(n+1) = \mathbf{R}'_r(n) \ddagger_r \end{aligned} \quad (\text{II.112})$$

Suivant la méthode utilisée, \ddagger_1 , \ddagger_r et \ddagger_{infl} présentent des formes différentes comme le montre le tableau suivant :

	\ddagger_1	\ddagger_r	\ddagger_{infl}
<i>Phillips</i>	$\ddagger_{11} + \ddagger_{11}^{-1} \ddagger_{12} \ddagger_{21}$	$\ddagger_{22} - \ddagger_{21} \ddagger_{11}^{-1} \ddagger_{12}$	$\ddagger_{11}^{-1} \ddagger_{12}$
<i>Bennis</i>	$\ddagger_{11} + \ddagger_{12} (\mathbf{I} - \ddagger_{22})^{-1} \ddagger_{21}$	\ddagger_{22}	$\ddagger_{12} (\mathbf{I} - \ddagger_{22})^{-1}$
<i>El Moudni</i>	\ddagger_{11}	$\ddagger_{22} - \ddagger_{21} \ddagger_{11}^{-1} \ddagger_{12}$	$\ddagger_{11}^{-1} \ddagger_{12}$
<i>TH+PS en continu</i>	$\ddagger_1 = (\mathbf{I} + \ddagger_{12} \ddagger_{22} \ddagger_{21} (\ddagger_{11} + \mathbf{I})^{-1})^{-1} \times (\ddagger_{11} - \ddagger_{12} \ddagger_{22} \ddagger_{21} (\ddagger_{11} + \mathbf{I})^{-1})$ avec $\ddagger_{22} = (\ddagger_{22} - \mathbf{I} \ddagger_{21} (\ddagger_{11} + \mathbf{I})^{-1} \ddagger_{12})^{-1}$	$\ddagger_{22} - \ddagger_{21} (\ddagger_{11} + \mathbf{I})^{-1} \ddagger_{12}$	$-2 (\ddagger_{11} + \mathbf{I})^{-1} \ddagger_{12} (\ddagger_r - \mathbf{I})^{-1}$

Tab. II. 1. Résultats des méthodes de perturbations singulières étudiées.

En fonction des systèmes découplés obtenus, nous proposons deux méthodes de résolution. La première technique de résolution concerne les sous-systèmes lents découplés qui sont stochastiques. Dans cette situation, le système lent forme à lui tout seul une chaîne de Markov irréductible, capable d'évoluer de façon indépendante. La deuxième méthode correspond à la situation où le sous-système lent découplé ne présente pas la propriété de stochasticité.

III.6.1. Technique de résolution dans le cas où le système lent découplé représente une chaîne de Markov

Ce cas concerne les trois méthodes relatives à Phillips, Bennis et l'application des perturbations singulières en continu. Pour ces trois cas, la matrice de la partie lente est stochastique et nous pouvons appliquer la méthode directe de résolution.

a) Résolution de la partie lente

La résolution du système lent repose sur l'équation fondamentale de la chaîne de Markov irréductible qui lui correspond :

$$\mathbf{L}'(n+1) = \mathbf{L}'(n) \ddagger_1 \quad (\text{II.113})$$

avec $\mathbf{L}'(n) = [\mathbf{L}'_1(n) \mathbf{L}'_2(n) \dots \mathbf{L}'_{r_1}(n)]$ la distribution des probabilités des états du système lent à l'instant n . Nous considérons que les conditions initiales d'évolution du système lent réduit sont celles correspondant aux états lents du système initial.

$$\mathbf{L}'(0) = [\mathbf{P}_1(0) \mathbf{P}_2(0) \dots \mathbf{P}_{r_1}(0)]$$

Pour l'étude de l'évolution permanente des probabilités d'état du système lent, nous allons utiliser la méthode directe de résolution des chaînes de Markov. Nous obtenons ainsi :

$$\begin{cases} \mathbf{L}'(0) \left(\mathbf{\ddagger}_1 - \mathbf{I}_{r_1} \right) = 0 \\ \mathbf{L}'_i(0) = \mathbf{L}'_i(0) \end{cases}_{i=1}^{r_1} \quad (\text{II.114})$$

Nous obtenons ainsi la distribution limite du système lent $\mathbf{L}'(\infty) = [\mathbf{L}'_1(\infty) \mathbf{L}'_2(\infty) \dots \mathbf{L}'_{r_1}(\infty)]$.

b) Résolution de la partie rapide

L'évolution du système rapide découplé est donnée par l'équation :

$$\mathbf{R}'(n) = \mathbf{R}'_l(n) + \mathbf{R}'_r(n) \quad (\text{II.115})$$

où

$$\mathbf{R}'_r(n+1) = \mathbf{R}'_r(n) \mathbf{\ddagger}_r \text{ de conditions initiales } \mathbf{R}'_r(0) = [\mathbf{P}_{r_1+1}(0) \dots \mathbf{P}_r(0)] - \mathbf{L}'(0) \mathbf{\ddagger}_{\text{infl}}$$

et $\mathbf{R}'_l(n) = \mathbf{L}'(n) \mathbf{\ddagger}_{\text{infl}}$ avec les conditions initiales $\mathbf{L}'(0)$.

Le système de couche limite étant non stochastique, le comportement asymptotique du système rapide sera donné uniquement par l'influence de la partie lente :

$$\mathbf{R}'_l(\infty) = \mathbf{L}'(\infty) \mathbf{\ddagger}_{\text{infl}} \quad (\text{II.116})$$

c) Calibrage des distributions des sous-systèmes

Le vecteur des probabilités d'état $\mathbf{P}(n)$ de la chaîne de Markov initiale est approchée par les distributions $[\mathbf{L}'(n) \mathbf{R}'(n)]$ des sous-systèmes lent et rapide, avec des précisions qui dépendent des méthodes utilisées. Pour pouvoir retrouver aussi la propriété de stochasticité du vecteur $[\mathbf{L}'(n) \mathbf{R}'(n)]$ ainsi obtenu, il est nécessaire d'effectuer un calibrage des distributions des systèmes découplés. Ce calibrage se fait selon les formules suivantes :

$$\mathbf{L}(n) = \alpha(n) \mathbf{L}'(n) \text{ et } \mathbf{R}(n) = \alpha(n) \mathbf{R}'(n) \quad (\text{II.117})$$

où α est appelé *facteur de calibrage* et peut être calculé comme suit :

$$\alpha(n) = \frac{1}{\sum_{i=1}^{r_1} \mathbf{L}'_i(n) + \sum_{i=1}^{r_2} \mathbf{R}'_i(n)} \quad (\text{II.118})$$

Le principe de calibrage reste le même pour l'étude des distributions limites, en considérant $n = \infty$ dans (II.117) et (II.118).

III.6.2. Méthode de résolution dans le cas où la matrice de transition du système lent n'est pas une matrice stochastique

La modélisation correspondant à cette situation est la modélisation El Moudni. En effet, la matrice lente obtenue par ce découplage ne présente pas la propriété de stochasticité. Le

système ne peut donc plus être traité de façon indépendante.

Pour pallier à cette difficulté, nous proposons une méthode de résolution des sous-systèmes en utilisant une technique de redondance appliquée sur la partie lente.

a) Technique de redondance pour la résolution de la partie lente

Soit l'équation de la partie lente d'une chaîne de Markov finie homogène et ergodique

$$\mathbf{L}(n+1) = \mathbf{L}(n) \dot{\mathbf{T}}_1 \quad \text{avec} \quad \dot{\mathbf{T}}_1 = \dot{\mathbf{T}}_{11} = [p_{ij}]_{i,j=1,\dots,r_1} \quad (\text{II.119})$$

avec la distribution $\mathbf{L}(n) = [L_1(n) \ L_2(n) \ \dots \ L_{r_1}(n)]$ de condition initiale $\mathbf{L}(0)$. Nous rajoutons une ligne et une colonne à la matrice $\dot{\mathbf{T}}_1$, de façon à avoir la propriété de stochasticité. La ligne supplémentaire représente une moyenne des lignes de la matrice $\dot{\mathbf{T}}_{21}$, tandis que la colonne rajoutée est obtenue par sommation des colonnes de $\dot{\mathbf{T}}_{12}$.

La méthode de redondance proposée, introduit une mesure du couplage entre la partie lente et la rapide. Cette technique correspond à l'agrégation de toute la partie rapide en un seul état, que nous allons rajouter à la partie lente (état e_ξ).

Nous obtenons ainsi un système lent redondant, caractérisé par l'équation :

$$\mathbf{L}'(n+1) = \mathbf{L}'(n) \dot{\mathbf{T}}_1' \quad (\text{II.120})$$

avec les conditions initiales :

$$\mathbf{L}'(0) = [L_1(0) \ L_2(0) \ \dots \ L_{r_1}(0) \ L_\xi(0)] = \begin{bmatrix} P_1(0) & P_2(0) & \dots & P_{r_1}(0) & P_i(0) \\ & & & & \end{bmatrix}_{\substack{i=r_1+1 \\ \dots \\ i=r_1+1}}$$

où la nouvelle matrice $\dot{\mathbf{T}}_1'$:

$$\dot{\mathbf{T}}_1' = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1r_1} & p_{1\xi} \\ p_{21} & p_{22} & \dots & p_{2r_1} & p_{2\xi} \\ \dots & \dots & \dots & \dots & \dots \\ p_{r_1 1} & p_{r_1 2} & \dots & p_{r_1 r_1} & p_{r_1 \xi} \\ p_{\xi 1} & p_{\xi 2} & \dots & p_{\xi r_1} & p_{\xi \xi} \end{pmatrix} \quad (\text{II.121})$$

est caractérisée par :

$$p_{\xi j} = \frac{\sum_{i=r_1+1}^r p_{ij}}{r_2}, \quad \forall j \in \{1, 2, \dots, r_1\} \quad (\text{II.122})$$

$$p_{i\xi} = 1 - \sum_{j=1}^{r_1} p_{ij}, \quad \forall i \in \{1, 2, \dots, r_1, \xi\}$$

La matrice $\mathbf{\tilde{T}}'_1$ présente la propriété de stochasticité :

$$0 \leq p_{ij} \leq 1, \quad \forall i, j \in \{1, 2, \dots, r_1, \xi\} \quad (\text{II.123})$$

$$p_{i\xi} + \sum_{j=1}^{r_1} p_{ij} = 1, \quad \forall i \in \{1, 2, \dots, r_1, \xi\}.$$

et par conséquent, le système lent redondant peut être résolu comme une chaîne de Markov indépendante.

L'étude de l'évolution permanente de ce système se fait à l'aide de la méthode directe :

$$\begin{cases} [L_1(\cdot) \quad L_2(\cdot) \quad \dots \quad L_{r_1}(\cdot) \quad L_\xi(\cdot)] \mathbf{D}' = 0 \\ L_1(\cdot) + L_2(\cdot) + L_3(\cdot) + \dots + L_{r_1}(\cdot) + L_\xi(\cdot) = 1 \end{cases} \quad (\text{II.124})$$

avec $\mathbf{D}' = \mathbf{\tilde{T}}'_1 - \mathbf{I}_{r_1+1}$ (\mathbf{I}_{r_1+1} est la matrice unité $(r_1+1) \times (r_1+1)$).

La distribution limite de ce système sera :

$$\mathbf{L}'(\cdot) = [L_1(\cdot) \quad L_2(\cdot) \quad \dots \quad L_{r_1}(\cdot) \quad L_\xi(\cdot)] \quad (\text{II.125})$$

Remarque

Dans la résolution du système lent réduit, grâce à l'application de la redondance, nous obtenons des informations sur l'évolution *globale* du système rapide. Dans le cas où l'évolution des probabilités des états de la partie rapide nous intéresse particulièrement, nous pouvons lui appliquer des algorithmes de résolution spécifiques.

□

b) Résolution de la partie rapide

La distribution de la partie rapide se fait de la même manière que dans le cas précédent. Par contre, le calibrage utilisé est différent et se fait uniquement sur la partie rapide.

c) Calibrage des solutions de la partie rapide

Pour retrouver la somme égale à 1 des éléments de la distribution limite approchée, nous effectuons un calibrage. Comme les probabilités de la partie lente ont été obtenues en considérant les relations avec la partie rapide (état redondant ξ), nous allons calibrer seulement la distribution rapide :

$$\mathbf{R}(n) = \mathbf{R}'(n) \frac{L_{\xi}(n)}{r_2 \sum_{i=1}^{\infty} R'_i(n)} \quad \text{respectivement à la limite} \quad (II.126)$$

IV. Comparaison des différentes méthodes de perturbations singulières appliquées aux chaînes de Markov

Un résumé des résultats obtenus par application des différentes techniques de perturbations singulières sur les chaînes de Markov est présenté dans la figure suivante :

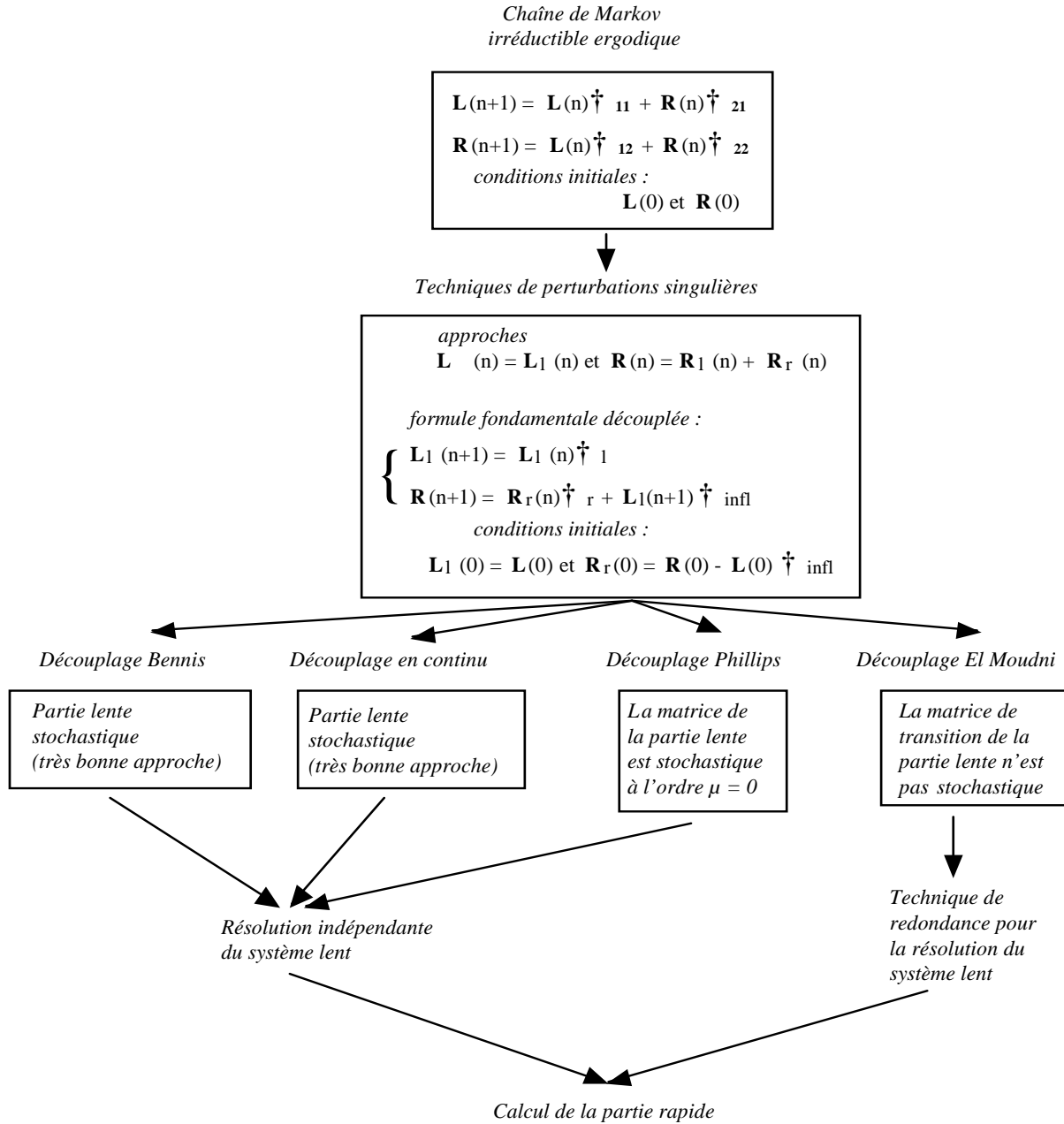


Fig. II.8. Analyse des méthodes développées.

IV.1. Etude de la structure des chaînes de Markov à double échelle de temps

De toutes les méthodes de perturbations singulières en discret que nous avons utilisé pour la simplification des chaînes de Markov, c'est celle de Bennis qui apparaît la mieux adaptée. Ceci est prouvé par la qualité du système découplé lent d'une part, et d'autre part par la structure de la chaîne de Markov à double échelle de temps, particulièrement adaptée à cette modélisation.

En effet, les chaînes de Markov à double échelle de temps sont caractérisées par une faible influence des états lents sur les états rapides par rapport aux influences des états rapides sur les états lents. Ceci est démontré par le théorème suivant :

Théorème 3

Dans une chaîne de Markov irréductible finie à double échelle de temps, le rapport des dynamiques et le rapport des couplages entre ces dynamiques sont du même ordre de grandeur.

-Démonstration-

Soit les normes de vecteurs N_1 et N_2 définies comme suit :

$$\|x\|_{N_1} = \sum_{j=1}^q |x_j|, \quad \text{À } x = [x_1 \dots x_q] \in \mathbb{R}^{1 \times q}; \quad \|\cdot\|_{N_1} : \mathbb{R}^{1 \times q} \rightarrow \mathbb{R}$$

$$\|y\|_{N_2} = r \cdot \max_{i=1}^r |y_i|, \quad \text{À } y = [y_1 \dots y_r] \in \mathbb{R}^{1 \times r}; \quad \|\cdot\|_{N_2} : \mathbb{R}^{1 \times r} \rightarrow \mathbb{R}$$

Nous allons noter par $\|\cdot\|_{N_1 N_2}$ la norme de matrices induite par les normes de vecteurs N_1 et N_2 .

Elle se définit comme étant (voir annexe 2) :

$$\|A\|_{N_1 N_2} = \frac{1}{r} \max_{i=1}^r \sum_{j=1}^q |a_{ij}|, \quad \text{À } A = [a_{ij}]_{\substack{i=1, \dots, r \\ j=1, \dots, q}} \in \mathbb{R}^{r \times q} \quad (\text{II.127})$$

Soit aussi la forme partitionnée de l'équation fondamentale d'une chaîne de Markov ergodique à double échelle de temps :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \dagger_{11} & \dagger_{12} \\ \dagger_{21} & \dagger_{22} \end{pmatrix} \quad (\text{II.128})$$

avec $\mathbf{L} \in \mathbb{R}^{1 \times r_1}$ la distribution de la partie lente et $\mathbf{R} \in \mathbb{R}^{1 \times r_2}$ la distribution de la partie rapide.

La chaîne de Markov étudiée est irréductible. Par conséquent, les normes des sous matrices de la forme partitionnée (II.128) sont strictement positives. Ceci nous permet de définir un coefficient de perturbation μ de la chaîne à double échelle de temps (II.128), en utilisant en particulier la norme (II.127) :

$$\mu = \frac{\|\dagger_{22}\|_{N_1 N_2}}{\|\dagger_{11}\|_{N_1 N_2}} \ll 1. \quad (\text{II.129})$$

L'utilisation de la propriété de stochasticité de la matrice de transition nous permet d'écrire l'égalité suivante :

$$\|\dagger_{11}\|_{N_1 N_2} + \|\dagger_{12}\|_{N_1 N_2} = \|\dagger_{21}\|_{N_1 N_2} + \|\dagger_{22}\|_{N_1 N_2} \quad (\text{II.130})$$

en tenant compte de (II.129) il nous vient :

$$\|\dagger_{21}\|_{N_1 N_2} \gg \|\dagger_{12}\|_{N_1 N_2}$$

En posant le coefficient de couplage entre les dynamiques :

$$\theta = \frac{\|\dot{\dagger}_{12}\|_{N_1 N_2}}{\|\dot{\dagger}_{21}\|_{N_1 N_2}} \ll 1 \quad (\text{II.131})$$

et en reportant (II.131) dans (II.130) d'une part et (II.129) dans (II.130) d'une autre part (voir *annexe 3*), l'approximation $\mu\theta \approx 0$ nous permet d'obtenir l'égalité suivante :

$$\frac{\|\dot{\dagger}_{22}\|_{N_1 N_2}}{\|\dot{\dagger}_{11}\|_{N_1 N_2}} = \frac{\|\dot{\dagger}_{12}\|_{N_1 N_2}}{\|\dot{\dagger}_{21}\|_{N_1 N_2}} \quad \text{donc} \quad \mu = \theta. \quad (\text{II.132})$$

□

Le théorème précédent nous permet de conclure que l'équation fondamentale d'une chaîne de Markov à double échelle de temps peut s'écrire sous la forme suivante :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}(n) \quad \mathbf{R}(n)] \begin{pmatrix} \dot{\dagger}_{11} & \mu \dot{\dagger}^*_{12} \\ \dot{\dagger}_{21} & \mu \dot{\dagger}^*_{22} \end{pmatrix} \quad (\text{II.133})$$

où nous avons introduit le facteur de séparation des dynamiques $\mu = \|\dot{\dagger}_{22}\|_{N_1 N_2} / \|\dot{\dagger}_{11}\|_{N_1 N_2}$ et

les notations $\dot{\dagger}^*_{12} = \frac{\dot{\dagger}_{12}}{\mu}$, $\dot{\dagger}^*_{22} = \frac{\dot{\dagger}_{22}}{\mu}$.

Cette forme de l'équation fondamentale correspond à la modélisation proposée par Phillips, dans le cas $j = 1$. Cette modélisation se montre bien adaptée, sous la réserve que la stochasticité de la matrice lente obtenue après découplage est approximative. Par conséquent, une meilleure solution est représentée par le découplage de Bennis qui correspond à la même structure (II.133) et en plus donne une stochasticité exacte du sous-système lent.

L'application des transformations homographiques sur la chaîne de Markov (II.133) nous donne le processus de Markov d'équation suivante :

$$[\tilde{\mathbf{L}} \quad \tilde{\mathbf{R}}] = [\tilde{\mathbf{L}} \quad \tilde{\mathbf{R}}] \begin{pmatrix} \tilde{\dagger}_{11} & \tilde{\dagger}_{12} \\ \varepsilon \tilde{\dagger}^*_{21} & \varepsilon \tilde{\dagger}^*_{22} \end{pmatrix} \quad (\text{II.134})$$

avec le coefficient de découplage des dynamiques $\varepsilon = \|\dot{\dagger}\ddot{\mathbf{Y}}_{22}\|_{N_1 N_2} / \|\dot{\dagger}\ddot{\mathbf{Y}}_{11}\|_{N_1 N_2}$

et les notations $\dot{\dagger}\ddot{\mathbf{Y}}^*_{21} = \frac{\dot{\dagger}\ddot{\mathbf{Y}}_{21}}{\mu}$, $\dot{\dagger}\ddot{\mathbf{Y}}^*_{22} = \frac{\dot{\dagger}\ddot{\mathbf{Y}}_{22}}{\mu}$.

En effet, les transformations homographiques conservent les dynamiques du système, et le principe de découplage en continu correspond à la structure (II.134) du processus de Markov obtenu. Suite à cette structure, l'utilisation conjointe des perturbations singulières en continu et des transformations homographiques donne de très bons résultats pour la simplification des chaînes de Markov. La stochasticité de la partie lente découplée le prouve. La méthode présente l'intérêt d'offrir un terme de comparaison pour l'appréciation des résultats obtenus à l'aide des techniques spécifiques aux systèmes discrets.

IV.2. Interprétation des résultats obtenus

La stochasticité de la matrice de transition de la partie lente obtenue suite au découplage par perturbations singulières, est le résultat du regroupement des modes dominants du système initial au niveau de cette partie. L'essentiel des caractéristiques du système initial est donc concentré dans cette partie, qui constitue le moteur de l'évolution des sous systèmes découplés. En effet, l'évolution du système initial est réduite à l'évolution indépendante du sous-système lent, et à l'évolution mixte du sous-système rapide. Dans l'évolution de la partie rapide, nous retrouvons une évolution propre au système (les transitoires rapides) et une évolution obtenue par l'influence du système lent.

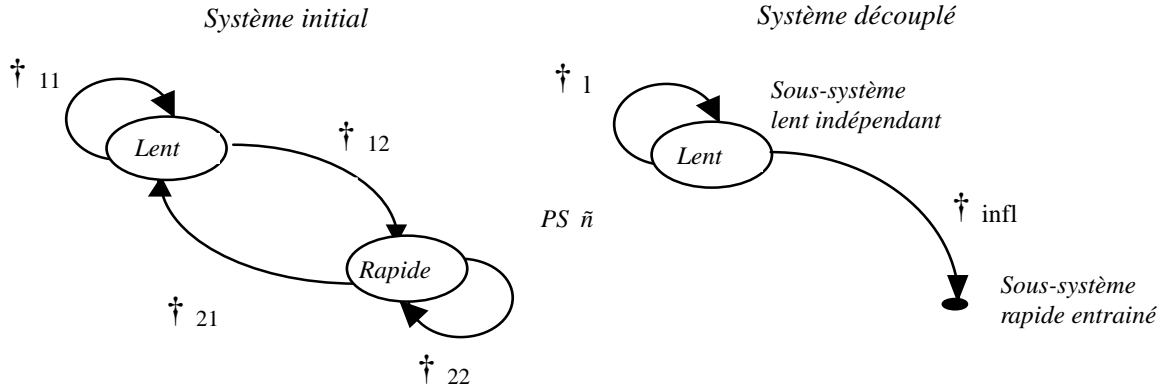


Fig. II.9. Simplification de l'étude d'une chaîne de Markov finie à l'aide des perturbations singulières.

Dans toutes les situations, les systèmes découplés s'écrivent :

$$[\mathbf{L}(n+1) \quad \mathbf{R}(n+1)] = [\mathbf{L}_l(n) \quad \mathbf{R}_r(n)] \begin{pmatrix} \dagger_1 & \dagger_l & \dagger_{\text{infl}} \\ \mathbf{0} & \dagger_r \end{pmatrix} \quad (\text{II.135})$$

avec \mathbf{L} la distribution de la partie lente, \mathbf{R} la distribution de la partie rapide, \dagger_1 la matrice de la partie lente, \dagger_r la matrice de la partie rapide, et \dagger_{infl} la matrice d'influence de la partie lente sur la partie rapide.

V. Calcul des probabilités de la chaîne de Markov réductible initiale

Reprenons la chaîne de Markov réductible initiale :

$$\mathbf{P}(n+1) = \mathbf{P}(n) \mathcal{M} \quad (\text{II.136})$$

dont la matrice de transition \mathcal{M} a été décomposée de la façon suivante :

$$\mathbf{M} = \left(\begin{array}{c|cccc} \mathbf{T} & & & & \\ \hline & \dagger_1 & 0 & 0 & \dots & 0 \\ & 0 & \dagger_2 & 0 & \dots & 0 \\ \mathbf{O} & 0 & 0 & \dagger_3 & \dots & 0 \\ & \dots & \dots & \dots & \dots & \dots \\ & 0 & 0 & 0 & \dots & \dagger_c \end{array} \right)$$

Entre les classes \dagger_i , nous avons choisi celles qui ont la propriété d'ergodicité et la double échelle de temps. Nous avons découplé les dynamiques de chaque tel type de classe finale (qui constitue en elle même une sous-chaîne de Markov), et déterminé les distributions limites de chaque partie. Le travail se fait donc sur des équations de dimension réduite. Les classes finales qui ne présentent pas la double échelle de temps doivent être traitées de façon classique, comme une chaîne entière.

Dans le cas où le système commence son évolution dans un état persistant, les probabilités limites concernent seulement les états de la même classe finale que l'état initial, et auront les valeurs déjà calculées.

Une situation spéciale intervient quand le système se retrouve dans un état initial transitoire. Dans ce cas, les probabilités limites des états persistants de la chaîne dépendent des probabilités d'absorption par la classe finale correspondante. Le calcul de ces probabilités d'absorption se fait en utilisant la méthode présentée dans le paragraphe II.5 du chapitre I.

Les probabilités d'absorption par la classe finale Fin_j associée à la matrice \dagger_j à partir de chaque état transitoire de la chaîne, sont regroupées dans un vecteur α^j et sont obtenues comme solutions du système :

$$\alpha^j = (\mathbf{I} - \mathbf{T})^{-1} \chi^j \quad (\text{II.137})$$

avec \mathbf{T} la matrice de transition entre les états transitoires, \mathbf{I} la matrice identité de même dimensions que \mathbf{T} , et χ^j le vecteur des sommes χ_i^j des probabilités de passage de l'état transitoire e_i à tous les états de la classe finale Fin_j .

Après avoir calculé les probabilités d'absorption par les classes finales, il nous reste à multiplier les résultats obtenus pour chaque classe finale ergodique avec la probabilité d'absorption de la classe. Nous déterminons ainsi les probabilités limites de tous les états persistants de la chaîne, dans le cas où l'évolution commence dans un état transitoire de celle-ci.

Si la chaîne de Markov a comme état initial un état transitoire e_i , la probabilité limite de l'état persistant lent e_m de la classe finale ergodique \dagger_j peut s'écrire sous la forme suivante :

$$P_m(0) = \alpha_i^j L_m(0) \quad , \quad e_m \in Fin_j \quad (\text{II.138})$$

Dans le cas où l'état persistant en cause est un état rapide de sa classe, sa probabilité limite

s'écrit, dans les mêmes conditions initiales d'évolution (état initial transitoire e_i) sous la forme suivante :

$$P_m(\infty) = \alpha_i^j R_m(\infty) , e_m \text{ ' } Fin_j \quad (II.139)$$

De cette façon, la chaîne de Markov réductible initiale est traitée globalement, avec l'avantage d'avoir à calculer seulement des systèmes de taille réduite.

VI. Conclusion

Dans ce chapitre, nous avons développé des méthodes de simplification des chaînes de Markov irréductibles ergodiques, basées sur les techniques de perturbations singulières. Dans ce sens, nous avons introduit la propriété de double échelle de temps des chaînes de Markov, les techniques de perturbations singulières et les découplages correspondants.

Parmi les méthodes des perturbations singulières ainsi étudiées, celle de Bennis correspond le mieux à la simplification des chaînes de Markov. Elle représente la seule méthode en discret capable de donner une stochasticité exacte du sous-système lent après le découplage. Ce résultat est dû à la structure particulière des chaînes de Markov à double échelle de temps, caractérisée par une très faible influence des états lents sur les rapides.

L'utilisation des transformations homographiques conjointement aux perturbations singulières en continu donne aussi de très bons résultats sur le système lent. La précision du système lent ainsi que sa propriété de stochasticité est la conséquence de la modélisation du continu, qui est unique indépendamment des couplages entre les dynamiques du système.

Un problème qui présente un grand intérêt pour l'étude d'un système aléatoire, est celui des valeurs des probabilités d'état en régime permanent. Une détection des états caractérisés par des valeurs grandes ou petites des probabilités limites est très utile pour élaborer une commande efficace des systèmes aléatoires modélisés. Pour cela, dans le chapitre suivant nous exposons des techniques de réduction des chaînes de Markov en utilisant une séparation des états selon l'importance des valeurs du régime permanent.