50376
1987
145

50376
1987
145

N° d'ordre: 732

# THESE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE
FLANDRES ARTOIS

pour obtenir le grade de

**Docteur es Sciences Physiques**

par

**Frédéric ROTELLA**
Ingénieur IDN
Docteur-Ingénieur

BIBLIOTHEQUE INTERUNIVERSITAIRE
SECTION DE SCIENCES
LILLE

# METHODES ALGEBRIQUES ET ANALYTIQUES POUR LA SIMPLIFICATION ET LA COMMANDE DE SYSTEMES BILINEAIRES A DEUX DYNAMIQUES

Soutenue le 26 juin 1987 devant la Commision d'Examen:

| | | |
|---|---|---|
| M. | P.BORNE | Président, Rapporteur |
| Mme. | G.DAUPHIN-TANGUY | Rapporteur |
| MM. | S.G.TZAFESTAS | Rapporteur |
| | A.FOSSARD | Examinateur |
| | G.GILLES | Examinateur |
| | G.JACOB | Examinateur |

# AVANT - PROPOS

*Le travail que nous présentons a été effectué sous la direction de Madame DAUPHIN-TANGUY, dans le cadre du Laboratoire d'Automatique et d'Informatique Industrielle de l'Institut Industriel du Nord.*

*Avant toutes choses, nous tenons à remercier Madame DAUPHIN-TANGUY pour l'accueil qu'elle nous a réservé au sein de son équipe de recherche ainsi que de la confiance qu'elle nous a témoignée. Instigatrice de ce travail, elle a su à tous moments se montrer disponible pour soutenir de nombreuses discussions très fructueuses. Ces échanges de vues, ainsi qu'un soutien moral constant, font que les résultats obtenus lui doivent beaucoup. Nous lui exprimons toute notre amitié et notre reconnaissance.*

*Malgré ses lourdes charges, Monsieur le Professeur BORNE, Directeur et Animateur du Laboratoire, Directeur Scientifique de l'I. D. N., s'est intéressé en permanence à nos recherches et nous a prodigués de nombreux conseils avisés. Nous sommes sensibles à l'honneur qu'il nous fait en acceptant de présider notre Jury de Thèse, nous l'en remercions et lui témoignons de notre plus vive reconnaissance.*

*Nous exprimons à Monsieur le Professeur TZAFESTAS nos plus grands remerciements pour l'honneur qu'il nous fait en acceptant de juger ce travail et de participer au Jury de Thèse.*

*Par des conseils judicieux et bienveillants, fondés sur sa connaissance très étendue de l'Automatique, Monsieur le Professeur FOSSARD nous a apporté une aide précieuse dans la mise au point définitive de ce mémoire. Nous le remercions et nous sommes honorés de sa participation à notre Jury.*

*Nous adressons à Monsieur le Professeur GILLES l'expression de notre gratitude pour l'extrême attention avec laquelle il a envisagé nos résultats de recherche. Sa présence dans ce Jury montre l'intérêt qu'il porte à nos travaux et nous l'en remercions vivement.*

*Que Monsieur le Professeur JACOB reçoive ici nos très sincères remerciements pour s'être intéressé à ce travail ainsi que pour la cordialité et la gentillesse qu'il nous a toujours témoignées. Nous lui exprimons notre reconnaissance pour sa présence dans ce Jury de Thèse.*

*Nous tenons à remercier tout particulièrement Madame BOURDREL, qui a bien voulu se charger de la réalisation matérielle de ce mémoire avec toute la compétence, le dévouement, la patience et la gentillesse qui la caractérise. Nous remercions également Madame FERRAR et Monsieur CARUZZI qui se sont chargés de l'impression.*

# INTRODUCTION GENERALE

**NOTE :** Les numérotations des formules et des paragraphes sont propres à chaque chapitre. Elles sont rappelées dans le texte par leur numéro, si elles sont définies dans le même chapitre, auquel on adjoint le numéro du chapitre de définition dans le cas contraire.

# INTRODUCTION GENERALE

Le thème général des travaux présentés dans ce mémoire est l'utilisation de méthodes de réduction de dimensionnalité, notamment les perturbations singulières, pour l'analyse et la commande des systèmes bilinéaires. En effet, l'élaboration de commandes dans le cas de processus complexes demande souvent l'utilisation d'un modèle plus simple. La méthode des perturbations singulières s'avère être un outil puissant de réduction d'ordre par découplage des dynamiques.

Pourquoi les systèmes bilinéaires ? C'est la question à laquelle nous nous sommes efforcés de répondre dans une première partie, en rappelant, tout d'abord, que leur structure est suffisamment simple pour rappeler les modèles linéaires et leur comportement suffisamment variable pour modéliser un système non linéaire. D'autre part, de nombreux travaux ont montré qu'une méthodologie développée dans le cadre de systèmes bilinéaires pouvait être adaptée à des systèmes plus généraux, comme, par exemple, les systèmes à commande affine. Dans ce premier chapitre, nous insistons particulièrement sur l'outil que constituent les séries formelles et nous proposons une méthode pratique de bilinéarisation d'un système non linéaire donné.

Dans un deuxième chapitre, nous rappelons les méthodes classiques de réduction de dimensionnalité des systèmes en distinguant celles qui sont propres aux systèmes linéaires, basées sur la fonction de transfert, de celles qui sont généralisables au cas des systèmes non linéaires, car basées sur l'équation d'état. Le point important de ce chapitre concerne la présentation, pour les système linéaires, d'une nouvelle méthode de découplage dans l'utilisation de la méthode des perturbations singulières. Cette nouvelle approche permet d'affiner les modèles réduits lents et rapides par une prise en compte plus nette des petits paramètres qui interviennent dans le modèle initial.

Les deux premières parties avaient pour objectif de présenter les systèmes étudiés et la méthode de réduction utilisée. La suite du mémoire est donc constituée de l'application de la méthode des perturbations singulières sur les systèmes bilinéaires. Ceci sera mené en deux temps : d'abord, dans le chapitre III, nous proposerons des méthodes de mise en évidence de dynamiques différentes ; puis, au cours du chapitre IV, nous découplerons les parties lentes et rapides des systèmes bilinéaires singulièrement perturbés. Les points nouveaux concernant ces deux chapitres se situent à tous les niveaux de l'analyse et dans les diverses méthodes généralement employées. On peut citer rapidement : détermination de la matrice L de bloc diagonalisation, cercles de Gershgorine par blocs, méthode du conditionnement, découplage dans le cas des systèmes bilinéaires, bilinéarisation de la partie lente. Les méthodes exposées dans ces deux chapitres permettent de construire deux modèles réduits bilinéaires lents et rapides à partir d'un modèle bilinéaire initial à deux dynamiques.

Enfin, dans une cinquième et dernière partie, nous abordons le problème de la simplification de la commande optimale d'un système bilinéaire singulièrement perturbé. Avant d'étudier ce problème, nous montrons, d'une part, que dans le cas linéaire, l'utilisation de la méthode de découplage permet d'améliorer les performances obtenues et, d'autre part, que dans le cas des systèmes non linéaires, une commande analytique par retour d'état est envisageable. Un paragraphe important est donc dévolu au calcul itératif de cette commande. Le cas des systèmes bilinéaires n'est plus alors qu'un cas particulier ; de plus, si l'on est dans le cas singulièrement perturbé, nous proposons l'utilisation des méthodes classiques propres aux systèmes linéaires : décomposition de l'équation de Riccati ou calcul d'une commande composite. Ces deux approches, comme nous le montrons sur un exemple, s'avèrent particulièrement performantes.

# CHAPITRE I

# STRUCTURE DES SYSTEMES

# INTRODUCTION

L'étude des processus réels, dans un but de synthèse, passe nécessairement par une phase de modélisation. La prise en compte de systèmes de plus en plus complexes, comme par exemple des manipulateurs à six (ou plus) degrés de liberté, infirme souvent l'hypothèse d'un modèle linéaire stationnaire valable pour tous les points de fonctionnement. Il est donc nécessaire de complexifier les modèles ; ce qui, sans passer au niveau des modèles non linéaires en l'état ou en l'entrée, peut être fait en considérant la classe des modèles bilinéaires. Ces modèles ont la propriété d'être suffisamment "adaptatifs" grâce à la présence de produits de variables d'état par des variables d'entrée pour pouvoir modéliser avec précision un nombre important de systèmes réels.

L'objet de ce chapitre est donc de situer les systèmes bilinéaires parmi les catégories classiques de modèles de système ; puis, d'en rappeler quelques propriétés ; et enfin, de présenter une méthode pratique d'obtention d'un modèle bilinéaire à partir de la description d'un système non linéaire.

## I. STRUCTURE DES SYSTEMES CONTINUS

La notion de système physique continu est, avant tout, reliée à la relation entrée-sortie entre des grandeurs qui dépendent d'une variable privilégiée réelle positive : t, le temps. Parmi les grandeurs d'entrée, on peut distinguer |FAU,84| :

* les signaux utiles ou commandes ;
* les signaux nuisibles ou perturbations ;

ces signaux pouvant être aléatoires ou déterministes.

Les grandeurs de sortie sont constituées par des variables accessibles du système.

Mais à un instant t donné, il n'est pas possible, compte tenu des informations d'entrée à cet instant, de pouvoir prédire l'évolution future des sorties. Il est donc nécessaire d'introduire un troisième ensemble de grandeurs qui constitue l'état du système. Cet état, variable dépendant du temps, peut être ou ne pas être mesurable et représente la mémoire du système. Le passé du système influence son évolution future par l'intermédiaire de cette variable.

Le comportement du système peut être décrit par une équation d'état, un des modèles mathématiques du système :

$$(S) \quad \begin{cases} \dot{x}(t) = f(x(t), u(t), t) \\ y(t) = g(x(t), u(t), t) \ ; \ x(o) = x_o \end{cases}$$

où $t \in \mathbb{R}^+$

$u(t)$ est le vecteur entrée, $u(t) \in \mathbb{R}^m$

$y(t)$ est le vecteur sortie, $y(t) \in \mathbb{R}^n$

$x(t)$ est un vecteur état, $x(t) \in \mathbb{R}^q$

$f(.)$ et $g(.)$ sont des fonctions

de $\mathbb{R}^q \times \mathbb{R}^m \times \mathbb{R}^+$ dans $\mathbb{R}^q$ et $\mathbb{R}^n$, respectivement.

Le système S est donné soit par la connaissance des lois physiques décrivant le système, soit par une solution du problème de réalisation d'un système à partir de son comportement entrée-sortie. Dans ce dernier cas, l'un des principaux problèmes réside dans la détermination de la dimension correcte de l'état, c'est-à-dire celle qui donne au modèle les propriétés de commandabilité et d'observabilité |HER,77| ; (S) est alors une réalisation minimale.

## II. CLASSES DE SYSTEMES

Dans toute la suite, on identifiera les notions de système et de modèles et par abus de langage une équation telle que (S) sera appelée système.

La forme particulière des fonctions non linéaires intervenant dans (S), principalement f(.), permet de définir plusieurs catégories de systèmes. Le système (S) représente, de façon générale, la très large classe des systèmes non-linéaires continus, à l'intérieur de laquelle on peut distinguer :

. les systèmes linéaires stationnaires.
. les systèmes linéaires non stationnaires.
. les systèmes bilinéaires.
. les systèmes à commande affine.

Par la suite, nous verrons que les systèmes bilinéaires forment une classe de systèmes "presque linéaires", donc permettant l'extension de certaines méthodes propres aux systèmes linéaires, tout en étant des systèmes qui approchent avec une bonne précision les systèmes non-linéaires.

### II.1. LES SYSTEMES LINEAIRES

Lorsque f(.) et g(.) sont des applications linéaires en l'état et en l'entrée, le système (S) est linéaire. L'équation d'état peut se mettre sous la forme

$$(1)\qquad \begin{aligned} \dot{x}(t) &= A(t)\,x(t) + B(t)\,u(t) \\ y(t) &= C(t)\,x(t) + D(t)\,u(t) \end{aligned}$$

où $t \in \mathbb{R}^+$, $A(t)$, $B(t)$, $C(t)$, $D(t)$ sont respectivement des matrices de $\mathbb{R}^{qxq}$, $\mathbb{R}^{qxm}$, $\mathbb{R}^{nxq}$, $\mathbb{R}^{nxm}$

Dans le cas où les matrices $A(t)$, $B(t)$, $C(t)$, $D(t)$ ne dépendent pas explicitement du temps le système est linéaire stationnaire ; nous l'appellerons système linéaire. Dans (1), $A(t)$ est la matrice système, $B(t)$ : la matrice de commande, $C(t)$ : la matrice d'observations, et $D(t)$ : la matrice de transmission directe.

### II.1.1. Forme externe des systèmes linéaires

L'utilisation de la transformée de Laplace permet de définir la fonction de transfert d'un système linéaire, à conditions initiales nulles :

$$(2)\qquad F(p) = C\,(pI-A)^{-1}\,B + D,\ p \text{ opérateur de dérivation ;}$$

Dans le cas de systèmes monovariables, i.e. $m=n=1$, $F(p)$ est une fraction rationnelle en $p$ que l'on écrit sous la forme :

$$(3)\qquad F(p) = \frac{\sum_{i=0}^{q-1} b_i\, p^i}{\sum_{i=0}^{q} a_i\, p^i} + d \quad , \qquad \text{où } a_q = 1,$$

et la sortie du système est donnée par la formule de convolution, sous forme analytique (pour une entrée $u(t)$ et des conditions initiales nulles) :

$$(4)\qquad t>0 \qquad y(t) = \int_0^t h(t-\tau)\,u(\tau)\,d\tau$$

où $h(t)$ est la réponse impulsionnelle du système.

### II.1.2. Forme interne des systèmes linéaires

L'équation d'état (1) constitue une réalisation de la relation entrée-sortie (4). Mais cette réalisation n'est pas unique, et même lorsqu'elle est minimale elle est définie à un changement de base près. Il existe cependant des formes canoniques, qui seront aussi valables pour les systèmes non-linéaires [LUE,67] bien que la notion d'une matrice A(.) dépendant de l'état soit délicate à définir.

Si on se place dans le cas monovariable, F(p) conduit de façon classique aux formes canoniques suivantes :

* la forme commandable ou compagne

$$\dot{x}(t) = A_c\, x(t) + B_c\, u(t)$$

$$y(t) = C_c\, x(t) + du(t)$$

avec $A_c = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -a_0 & \cdots & & -a_{q-1} \end{bmatrix}$, $B_c = \begin{bmatrix} 0 \\ | \\ | \\ . \\ 0 \\ 1 \end{bmatrix}$

$$C_c = [b_0, \text{———}, b_{q-1}]$$

* la forme observable ou Frobenius

$$\dot{x}(t) = A_c^T\, x(t) + C_c^T\, u(t)$$

$$y(t) = B_c^T\, x(t) + du(t)$$

où $^T$ désigne la transposition de matrices.

Dans le cas où le système est observable et commandable, on peut passer de l'une à l'autre par changement de base constant.

Par l'utilisation de l'algèbre linéaire ou des méthodes algébriques, les systèmes linéaires sont bien connus du point de vue mathématique mais ils traduisent parfois difficilement les comportements réels de certains systèmes.

## II.2. LES SYSTEMES A COMMANDE AFFINE

### II.2.1. Formule fondamentale

Bien que représentant un cas particulier de modèles non-linéaires, ils peuvent modéliser un très large ensemble de systèmes. Ils sont définis à partir de (S) par : f(.) fonction affine des variables d'entrée ; et g(.) ne dépendant que de l'état ; c'est-à-dire par une équation d'état de la forme :

$$\begin{cases} \dot{x}(t) = X_0(x(t)) + \sum_{i=1}^{m} u_i(t)\, X_i(x(t)) \\ y(t) = g(x(t)) \end{cases} \qquad (5)$$

où les $X_i(x)$, $i = 0,\ldots,m$, sont des fonctions de l'état de composantes $\theta_i^k(x)$, $k = 1,..,q$.

Dans le cas où les $\theta_i^k(x)$ sont des séries formelles en les composantes de l'état $x_1,..,x_q$ le système est dit analytique.

Les méthodes d'étude de ces systèmes peuvent être regroupées en trois classes : la plus ancienne est basée sur les séries de Volterra ; |BRO,76a ; LES,78 ; KU,83| ; la deuxième utilise la géométrie différentielle et les algèbres de Lie |LOB,76|, dont un exposé des récents développements se trouve dans |GAU,84| ; la troisième approche, plus algébrique, utilise les travaux de Fliess sur les séries formelles |FLI,76|. Bien que ces trois approches soient parallèles et fassent appel à certaines notions communes, la dernière apparaît plus maniable, notamment Fliess a montré que la sortie du système analytique (5) est une fonctionnelle causale analytique qui admet comme développement |FLI,80| :

$$(6)\quad y(t) = g|_{x(o)} + \sum_{v\geq o} \sum_{j_o,..,j_v=o}^{m} (X_{j_o}(...(X_{j_v}(g(x)))..))|_{x(o)} I_{j_v...j_o}$$

où $|_{x(o)}$ indique une évaluation en x(o) ;

$$\forall i\in\{o,..,m\}, \forall h\ (x) \in \mathbb{R}^n, X_i(h(x)) = \sum_{k=1}^{q} \theta_i^k\ (x)\ \frac{\partial h}{\partial x_k}(x)$$

$$= \frac{\partial h}{\partial x}.X_i(x)$$

et $I_{j_v...j_o}$ représente

$\int_o^t d\xi_{j_v} \ldots d\xi_{j_o}$ intégrale itérée au sens de Chen définie par récurrence par :

$$\xi_o(\tau) = \tau, \quad \xi_i(\tau) = \int_o^\tau u_i(\sigma)d\sigma\ ,\ i = 1,..,m$$

$$\int_o^t d\xi_j = \xi_j(t)$$

$$\int_o^t d\xi_{j_v}\ d\xi_{j_{v-1}}\cdot\cdot\ d\xi_{j_o} = \int_o^t d\xi_{j_v} \int_o^\tau d\xi_{j_{v-1}} \ldots d\xi_{j_o}$$

La formule (6) est appelée formule fondamentale. Par un codage de l'intégrale itérée à l'aide d'un nombre fini de symboles $\{\alpha_o,\alpha_1,..,\alpha_m\}$ :

$$(7)\qquad \int_o^t d\xi_j \ldots d\xi_{jo} \equiv \alpha_{jv} \ldots \alpha_{jo}$$

on associe à y(t), une série formelle, dite série génératrice, en les variables non commutatives $\alpha_o,..,\alpha_m$ (car de façon générale $\alpha_i\alpha_j \neq \alpha_j\alpha_i$ pour $i\neq j$) :

$$(8)\qquad s = g|_{x(o)} + \sum_{v\geq o} \sum_{j_o..j_v=o}^{m} X_{j_o}\ldots X_{j_v}g|_{x(o)}\ \alpha_{j_v}\ldots\alpha_{j_o}$$

Il a également montré l'équivalence entre cette série formelle et le développement en série de Volterra du système (5) |FLI,79|.

### II.2.2. Calcul symbolique non commutatif

De façon à déterminer la série génératrice d'un système analytique plus rapidement que par l'utilisation directe de la formule fondamentale, Lamnabhi |LAM,80| a proposé une méthode de calcul symbolique basée sur le produit mélange. Le mélange, ou produit de Hurwitz, de deux séries génératrices est l'opération linéaire, $\sqcup\!\sqcup$, définie par récurrence, par :

$$(9)\qquad \forall\,(\mu,\nu)\in\mathbb{N}^2\ \forall\,(j_0,\dots,j_\nu,i_0,\dots,i_\mu)\in\{0,\dots,m\}^{\nu+\mu}$$

$$\underline{1}\sqcup\!\sqcup\underline{1}=\underline{1},\quad \alpha_{j_0}\dots\alpha_{j_\nu}\sqcup\!\sqcup\underline{1}=\underline{1}\sqcup\!\sqcup\alpha_{j_0}\dots\alpha_{j_\nu}=\alpha_{j_0}\dots\alpha_{j_\nu}$$

$$\alpha_{j_0}\dots\alpha_{j_\nu}\sqcup\!\sqcup\alpha_{i_0}\dots\alpha_{i_\mu}=\alpha_{j_0}(\alpha_{j_1}\dots\alpha_{j_\nu}\sqcup\!\sqcup\alpha_{i_0}\dots\alpha_{i_\mu})+\alpha_{i_0}(\alpha_{j_0}\dots\alpha_{j_\nu}\sqcup\!\sqcup\alpha_{i_1}\dots\alpha_{i_\mu})$$

où $\underline{1}$ est le symbole vide.

Il vient alors le résultat suivant :

<u>Théorème</u> : Le produit de deux fonctionnelles causales analytiques est une fonctionnelle causale analytique dont la série génératrice est le mélange de leurs séries génératrices.

En conséquence de ce théorême, à la fonction analytique (10a) correspondra la série génératrice (10b) :

$$(10a)\qquad a(x_1,\dots,x_n)=\sum_{i_1,\dots,i_n\geq 0} a_{i_1,\dots,i_n}(x_1)^{i_1}\dots(x_n)^{i_n}$$

$$(10b)\qquad \overset{\sqcup\!\sqcup}{a}(s_1,\dots,s_n)=\sum_{i_1,\dots,i_n\geq 0} a_{i_1,\dots,i_n}(s_1)^{\sqcup\!\sqcup i_1}\dots(s_n)^{\sqcup\!\sqcup i_n}$$

où $s_i$ est la série génératrice associée à $x_i$, et $s_j^{\sqcup\!\sqcup i_j}$, $j=1,..,n$, représente : $\underbrace{s_j \sqcup\!\sqcup s_j \sqcup\!\sqcup \ldots \sqcup\!\sqcup s_j}_{i_j \text{ fois}}$.

Ainsi, si l'on intègre l'équation d'état (5), on obtient :

$$(11)\quad \begin{cases} x_k(t) = x_k(o) + \int_o^t \theta_o^k (x(\tau))\, d\tau + \sum_{i=1}^{m} \int_o^t u_i(\tau)\, \theta_i^k (x(\tau))\, d\tau \\ k = 1,..,q \\ y = g\,(x(t)) \end{cases}$$

ce qui donne, par codage de l'intégrale et en notant $s_k$ la série génératrice associée à la $k^{ième}$ composante d'état :

$$(12)\quad \begin{cases} s_k = x_k\,(o) + \sum_{i=o}^{m} \alpha i\, {}^{\sqcup\!\sqcup k}\theta_i\ (s_1,..,s_q) \\ k = 1,..,q \\ s = {}^{\sqcup\!\sqcup}g\,(s^1,..s^n) \end{cases}$$

Les séries génératrices $s^k$ sont donc solution d'une équation algébrique et les termes successifs de ces séries, donc de s, peuvent être calculés par itérations. L'intérêt de ce calcul est d'une part d'être systématique, d'autre part de permettre par la connaissance des premiers termes de la série formelle s (resp. $s_i$) de pouvoir estimer la transitoire de la sortie (resp. $x_i$).

### II.2.3 Estimation du transitoire

La série génératrice traduit algébriquement la relation entrée-sortie du système. La méthode proposée dans |LAM,80 ; FLI,83 ; LAM,85| permet de construire la série en t associée à la variable

étudiée. Il suffit de remplacer dans la série génératrice de la sortie, chaque lettre $\alpha_i$, i=1,..,m, par l'opérateur $\alpha_0(g_{u_i} \sqcup\!\sqcup)$, i=1,..,m, où $g_{u_i}$ représente la transformée de Laplace-Borel de la $i^{ème}$ entrée définie par :

$$\text{si } u_i(t) = \sum_{n\geq 0} u_{i(n)} \frac{t^n}{n!} \text{ alors}$$

(13)

$$g_{u_i} = \sum_{n\geq 0} u_{i(n)} \alpha_0^{\,n}$$

A titre d'illustration, considérons le système défini par

(14) $\quad \overset{\circ}{y}(t) + y(t) + y^2(t) = u(t)$ avec $y(o) = o$ ;

l'intégration de ce système fournit l'équation algébrique formelle :

(15) $\quad s + \alpha_0 s + \alpha_0 s \sqcup\!\sqcup s = \alpha_1$

Pour une entrée $u(t) = 1$, dont la transformée de Laplace-Borel est 1, $\alpha_1$ doit être remplacé par $\alpha_0$, ce qui donne l'équation

(15') $\quad s + \alpha_0 s + \alpha_0 s \sqcup\!\sqcup s \equiv \alpha_0$

Posons $s = \sum_{i\geq 0} s(i) \alpha_0^{\,i}$ ; or $\alpha_0^{\,i} \sqcup\!\sqcup \alpha_0^{\,j} = C^i_{i+j} \alpha_0^{\,i+j}$, il vient donc :

(16)
$$\begin{cases} s(o) = \ ; \ s(1) = 1 \ ; \\ \\ k > 1, \ s(k+1) = - s(k) - \sum_{i=0}^{k} C^i_k \ s(i)s(k-i) \end{cases}$$

ce qui donne les premiers termes de la sortie :

(17) $\quad y(t) = t - \frac{t^2}{2} - \frac{t^3}{3!} \ldots\ldots + s(k) \frac{t^k}{k!} + \cdots$

Cette méthode permet donc d'obtenir un développement analytique autour de t=o d'une variable quelconque dont on a une équation algébrique formelle, en généralisant le calcul symbolique d'Heaviside. Ceci peut être utilisé dans un but de comparaison de transitoires, donc de dynamiques, de variables différentes.

### II.2.4 Forme commandable des systèmes non-linéaires

Dans ce paragraphe, nous rappelons les résultats de |HUN,83| qui généralisent aux systèmes non-linéaires à commande affine, la représentation commandable des systèmes linéaires (II.1.2).

Considérons le système à commande affine (5) où m=1, et les opérateurs de dérivation suivants, où f et g sont deux fonctions vectorielles de l'état de dimensions convenables :

$$< df, g > = \frac{\partial q}{\partial x} \cdot f$$

$$[ f, g ] = < dg, f > - < df, g > \qquad (18)$$

$$\text{et } \forall k \in \mathbb{N}^* \ (ad^k f, g) = [ f, (ad^{k-1} f, g) ]$$

$$\text{avec } (ad^0 f, g) = g$$

Il vient alors le résultat suivant |HUN,83| :

Théorème : Si la matrice de commandabilité $\mathcal{C}$, définie par

$$\mathcal{C} = | X_0(x), (ad^1 X_0(x), X_1(x)), \ldots, (ad^{q-1} X_0(x), X_1(x))|$$

est régulière, i.e. $\forall x \in \mathbb{R}^q$ $rg\, \mathcal{C} = q$ :

alors le changement de variables $x^* = T(x)$

où

$$T = \begin{bmatrix} T_1(x) \\ < dT_1, X_0(x) > \\ < d < dT_1, X_0(x) >, X_0(x) > \\ \vdots \\ \vdots \\ \vdots \end{bmatrix}$$

où $T_1(x)$ est défini par l'équation

$$\frac{\partial T_1}{\partial x} = \mu(x)\, l(x)$$

(où : $l(x)$ dernière ligne de $\mathcal{C}^{-1}$, et $\mu(x)$ est une fonction scalaire permettant l'intégration).

transforme (5) avec m=1 en la forme commandable

$$(19) \qquad \frac{d}{dt} \begin{bmatrix} x_1^* \\ \cdot \\ \cdot \\ \cdot \\ \\ x_q^* \end{bmatrix} = \begin{bmatrix} x_2^* \\ \cdot \\ \cdot \\ \cdot \\ x_{q-1} \\ v(x^*) \end{bmatrix} + \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ \mu(x^*) \end{bmatrix} u$$

Cette forme correspond donc à une forme matricielle compagne généralisée.

### II.2.5. Exemple de système à commande affine

Considérons le système monovariable de type Lurie'Postnikov décrit figure 1. La non-linéarité séparable est de classe $\phi$ et la partie linéaire est représentée par une fonction de transfert non dégénérée |POP,73|.

$$\text{classe } \phi \quad \begin{cases} f : \mathbb{R} \longrightarrow \mathbb{R} \\ \qquad \varepsilon \longmapsto f(\varepsilon) = f^*(\varepsilon)\,.\,\varepsilon \\ f^*(\varepsilon) \in [\underline{L}, \bar{L}] \subset \mathbb{R},\ \underline{L} \leqslant \bar{L} \end{cases}$$

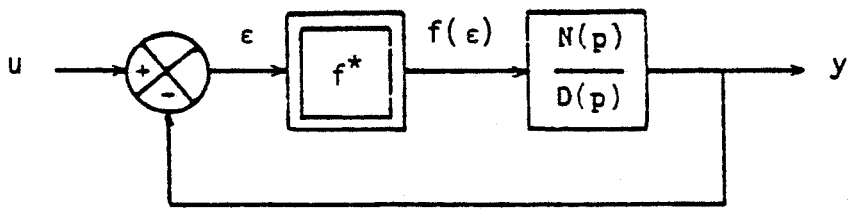



FIGURE 1

Soit (A, B, C) une réalisation de la partie linéaire, le système complet peut être modélisé par l'équation d'état

$$(20)\quad \begin{cases} \dot{x} = A\,x + B\,.\,f(\varepsilon) & , x \in \mathbb{R}^q, \text{ vecteur état} \\ y = C^T x & \\ f(\varepsilon) = f^*(\varepsilon)\,.\,\varepsilon\,,\ \varepsilon = u - y & ,\ u \in \mathbb{R}\,,\ y \in \mathbb{R} \end{cases}$$

$f^*$ est le gain équivalent de la non-linéarité défini par :

$$\text{si } \varepsilon \neq 0 \quad f^*(\varepsilon) = \frac{f(\varepsilon)}{\varepsilon}$$

$$\text{si } \varepsilon \doteq 0 \quad f^*(\varepsilon) = \left[\frac{df(\varepsilon)}{d\varepsilon}\right]_{\varepsilon = 0}$$

Compte tenu des hypothèses, on sait que l'on peut toujours se placer dans une base d'état telle que $\varepsilon$ soit une variable d'état (e.g.$x_q$). On obtient dans ce cas :

$$(21)\quad \dot{x} = \underbrace{\left[A - B\,C^T f^*(x_q)\right]}_{X_0(x)}\,x + \underbrace{\left[B\,f^*(x_q)\right]}_{X_1(x)}\,u$$

ce qui est la forme d'un système à commande affine. Dans le cas où f est analytique, f* l'est aussi, ce qui entraîne la même propriété pour le système présenté. Ces remarques demeurent valables pour des systèmes multivariables du type Lurie' Postnikov |GRU,75|.

## II.3. LES SYSTEMES BILINEAIRES

### II.3.1. Structure et définitions

Ces systèmes ont une structure intermédiaire entre les systèmes linéaires et les systèmes à commande affine puisqu'ils sont régis par une équation d'état (5) où les champs de vecteur $X_i(x)$ sont affines en l'état et la fonction de sortie g est une application linéaire. C'est-à-dire que l'on a :

$$(22)\qquad \begin{cases} \dot{x}(t) = A_0 x(t) + \sum_{i=1}^{m} u_i(t) \{A_i x(t) + B_i\} \\ y(t) = C^T x(t) \end{cases}$$

où $A_0$ ; $A_i$, i=1,..,m ; $B_i$, i=1,..,m ; $C^T$ sont des matrices constantes de dimensions convenables.

$u_i(t)$, i=1,..,m est supposé appartenir à un sous-ensemble de l'ensemble des fonctions mesurables localement bornées de $[0, \infty[$ dans $\mathbb{R}$.

La matrice non constante $(A_0 + \sum u_i A_i)$ est appelée matrice-système |BRA,81|. Certains auteurs utilisent également la notation condensée |BRU,74| :

$$(23)\qquad \begin{cases} \dot{x}(t) = A_0 x(t) + u(t) A x(t) + B u(t) \\ y(t) = C^T x(t) \end{cases}$$

où $u(t) A x(t) \equiv \sum_{i=1}^{m} u_i(t) A_i x(t)$

$B u(t) = \sum_{i=1}^{m} B_i u_i(t)$

ce qui correspond à la structure (figure 2) mettant en évidence la différence entre un système linéaire et un système bilinéaire:

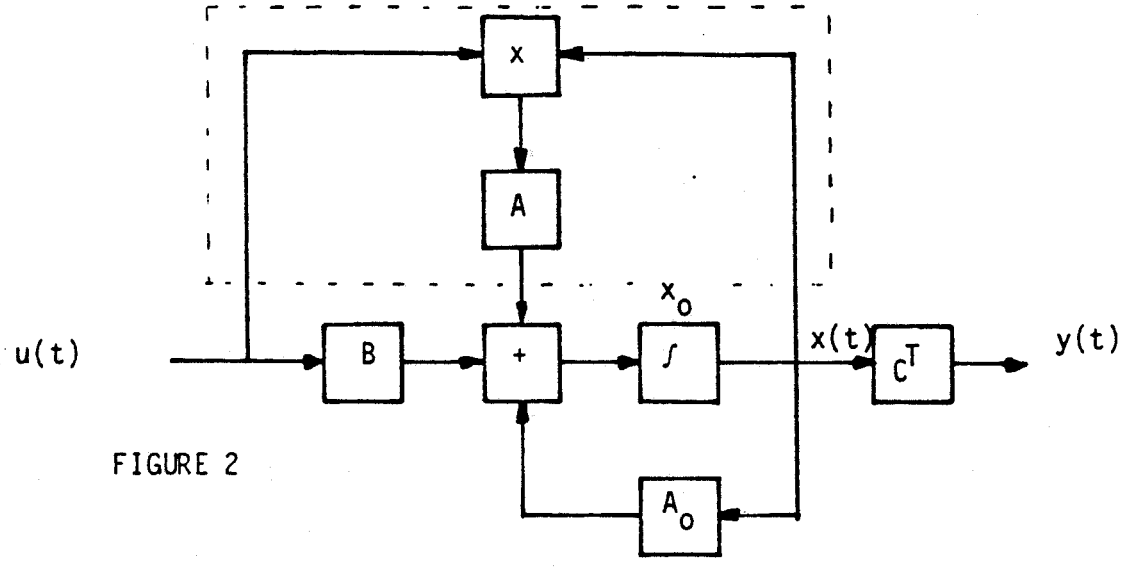



FIGURE 2

De même que pour les systèmes linéaires on peut considérer les systèmes bilinéaires non stationnaires pour lesquels les matrices $A_o$ et $A_i$ dépendent explicitement du temps. Mais par un changement de variables d'entrée dans lesquelles on distinguera deux groupes : celui concernant des perturbations déterministes ou aléatoires de paramètres et celui concernant les signaux de commande , on peut se ramener à un système bilinéaire stationnaire que nous appellerons bilinéaire par la suite.

Suivant la forme de l'équation (22) on peut définir plusieurs classes de systèmes bilinéaires |BRU,74|.

a) si $B = 0$ le système est dit homogène en l'état ou régulier ;

b) si $A_o = 0$ le système est dit homogène en l'entrée.

c) si $A_o$ et B sont nuls le système est dit strictement bilinéaire.

Il est à noter que ces définitions ne caractérisent pas le système bilinéaire mais la forme de l'équation de l'état, car un simple changement des variables d'entrée ou d'état permet de passer de l'une à l'autre de ces formes. Il suffit de montrer que (22) se met sous l'une des trois formes. En effet :

- considérons l'entrée fictive $u_o(t) = 1$, de (22) il vient avec $B_o = 0$ :

$$\dot{x}(t) = \sum_{i=0}^{m} u_i(t) \{A_i x(t) + B_i\}$$

qui est de la forme b), et le vecteur d'entrée $U(t) = [1, u_1(t), \ldots, u_m(t)]$ est sous la forme homogène.

- de même si on considère l'état homogène

$X(t) = [x_1(t), .., x_q(t), 1]^T$ alors (15) donne

$$\dot{X}(t) = A'o\, X(t) + \sum_{i=1}^{m} u_i(t)\, A'i\, X(t)$$

$$\text{où } A'o = \begin{bmatrix} & & 0 \\ & A_o & \vdots \\ & & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{bmatrix} \text{ et } A'i = \begin{bmatrix} & & \\ A_i & \vdots & B_i \\ & & \\ \cdots & \cdots & \cdots \\ 0 & \cdots & 0 \end{bmatrix}$$

- Ainsi en coordonnées homogènes pour l'état et l'entrée, il vient :

$$(24) \qquad \dot{X}(t) = \sum_{i=o}^{m} u_i(t)\, A'_i\, X(t)$$

qui correspond à une forme strictement bilinéaire qui peut être notée de façon condensée

$$\dot{X}(t) = U(t)\, A'\, X(t)$$

Comme nous venons de le voir, la structure de l'équation d'état d'un système bilinéaire permet de généraliser un certain nombre de structures simples comme les systèmes linéaires en ayant des systèmes "presque linéaires". Mais cette complexité n'est pas une vue de l'esprit et correspond à la réalité de certains systèmes physiques.

### II.3.2. Intérêt des systèmes bilinéaires

Il y en a, en effet, de nombreux motifs théoriques et pratiques qui justifient l'importance accordée aux systèmes bilinéaires depuis quelques années |LEV,83|. Depuis les travaux de Mohler qui a mis en évidence de nombreuses classes de systèmes gouvernés par des modèles bilinéaires |MOH,70a ; MOH,73|, de nombreuses études sur la commandabilité, la commande optimale ou l'identification de ces systèmes ont été publiés |MOH,80|. Concernant la modélisation de systèmes par des modèles bilinéaires, on peut citer par exemple :

* parmi les systèmes biologiques : la thermorégulation humaine |IBE,60 ; HAR,61|, la régulation du dioxyde de carbone dans le système respiratoire |GRO,63|.

* parmi les processus physiques : la fission nucléaire et les échangeurs de chaleur |MOH,70b|, les colonnes de distillation |ESP,78|.

Les premiers travaux à propos des systèmes bilinéaires ont porté sur l'étude de la commandabilité et il a été montré |RIN,68 ; IMB,75| qu'un système bilinéaire à entrées bornées peut être commandable alors que le système linéaire, obtenu en éliminant les produits des variables d'entrée par les variables d'état, ne l'est pas.

Cependant, l'avantage fondamental de l'utilisation des modèles bilinéaires provient des études de la réalisation des systèmes non linéaires |KRE,75|. Sussman |SUS,76| a démontré le théorème d'approximation qui établit que toute relation d'entrée-sortie F peut être approchée avec une erreur arbitrairement petite par une réalisation bilinéaire si F vérifie des conditions de causalité et de continuité. Quand les conditions initiales du système sont fixées, on peut associer à chaque sortie une fonction $F_u$ de [0, T] dans $\mathbb{R}$ pour chaque entrée u définie sur [0, T] ; alors F satisfait les propriétés :

* <u>de continuité</u> si les fonctions $F_{u_k}$ convergent uniformément vers $F_u$ pour toute suite $u_k$ d'entrées convergeant vers u.

* <u>de causalité</u> si les fonctions $F_u$ et $F_v$ coïncident sur |0,T| alors les entrées u et v coïncident sur |0,T'| pour tout $T' < T$.

<u>Théorême d'approximation</u> : Soit F une fonction arbitraire qui à chaque entrée u, définie de |0,T| dans $\mathbb{R}^m$ et bornée en module par un réel positif M, fait correspondre une relation de |0,T| dans $\mathbb{R}$. Si, de plus, F vérifie les conditions de causalité et de continuité, alors pour tout $\varepsilon > 0$, il existe un système bilinéaire dont la relation d'entrée-sortie $\Phi$ est telle que :

$$\forall\, t \in [0, T],\ \forall\, u : |0,T| \to \mathbb{R}^m,\ \|u\| < M$$

$$|F_u(t) - \Phi_u(t)| < \varepsilon$$

Remarquons que ce théorême ne prouve que l'existence d'un système bilinéaire approchant un système non linéaire donné et ne donne pas de procédé de construction. Nous verrons dans la dernière partie de ce chapitre des méthodes permettant d'obtenir un modèle bilinéaire approché .

### II.3.3. <u>Modèles mathématiques</u>

Comme nous l'avons mentionné, les systèmes bilinéaires ont une structure proche de celle des systèmes linéaires. De plus, Bruni |BRU,71| a montré que le système bilinéaire monovariable :

$$(25)\qquad \begin{cases} \dot{x} = Ax + Nxu + Bu, \quad x(o) = x_o \\ y = Cx \end{cases}$$

pouvait être exprimé par la séquence de systèmes linéaires à vecteurs de commande non stationnaires :

$$(26)\qquad \begin{aligned} &\dot{x}\,|o| = Ax|o| + Bu \;;\; x|o|(o) = x_o \\ &\dot{x}\,|i| = Ax|i| + Nx|i-1|u + Bu \;;\; x|i|(o) = x_o,\ i \in \mathbb{N}^* \end{aligned}$$

qui converge uniformément vers x, solution de (25) ; i.e. :

$$(27)\qquad \forall\, t \in \mathbb{R}^+ \quad \lim_{i \to \infty} \| x|i|(t) - x(t) \| = 0$$

Cette remarque permet d'écrire la sortie sous la forme d'une série de Volterra :

$$(28)\qquad y(t) = Ce^{At}x_o + \sum_{i=1}^{\infty} \int_o^t \int_o^{\sigma_1} \cdots \int_o^{\sigma_{i-1}} C\, e^{A(t-\sigma_1)} N e^{A(\sigma_1-\sigma_2)} N \cdots e^{A(\sigma_{i-1}-\sigma_i)} \left\{ N e^{A\sigma_i} x_o + B \right\} u(\sigma_i) \cdots u(\sigma_1)\, d\sigma_1 \cdots d\sigma_i$$

Cette expression met en évidence la partie de la sortie qui dépend des conditions initiales et celle qui dépend du vecteur de commande B. En particulier, la première partie est la réponse du système en régime libre, qui correspond bien à l'évolution du système linéaire pour u = o.

**II.3.3.1** Fonction de transfert multivariable

A partir de l'expression (28), prise pour des conditions initiales nulles et après changement de variables, il a été montré |MIT,79| que l'on pouvait définir une fonction de transfert multivariable pour le système bilinéaire (25). En utilisant la tranformée de Laplace à plusieurs variables définie par :

$$(29)\qquad F(p_1,\ldots,p_n) = \int_o^{\infty} \cdots \int_o^{\infty} f(\tau_1,\ldots,\tau_n) e^{-p_1\tau_1} \cdots e^{-p_n\tau_n}\, d\tau_1 \cdots d\tau_n$$

chacun des termes de (28) fournit la fonction de transfert

$$(30) \qquad F_n(p_1\ldots,p_n) = C\,(p_nI-A)^{-1}N(p_{n-1}\,I-A)^{-1}N\ldots N(p_1I-A)^{-1}B.$$

Ainsi, la fonction de transfert multivariable d'un système bilinéaire est une suite infinie de fonctions de transfert multivariables définie par |LEJ,85| :

$$(31) \qquad \hat{F}(p_1,..,p_\infty) = [F_1(p_1),\ F_2(p_1,p_2),\ \ldots,\ F_n(p_1,..,p_n),\ldots]$$

Cette expression généralise de façon évidente la fonction de transfert d'un système linéaire.

**II.3.3.2.** Codage des systèmes bilinéaires

L'application de la formule fondamentale (6) dans le cas du système bilinéaire régulier conduit à l'expression de la série génératrice de la sortie sous la forme |FLI,83|.

$$(32) \qquad s = C^Tx_o + \sum_{v>o} \sum_{j_o,..,j_v=o}^{m} C^T A_{j_v}\ldots A_{j_o}\,x_o\,\alpha_{j_v}\cdots\alpha_{j_o}$$

En effet, on a ici :

$$(33) \qquad i = o,..,m : X_i(x) = A_ix \text{ et } g(x) = C^Tx$$

donc d'après (6) :

$$(34) \qquad i=o,..,m\ X_i(g(x)) = \frac{\partial C^Tx}{\partial x} A_ix = C^TA_ix$$

Le terme général de la série génératrice s'écrit donc :

$$(35) \qquad (X_{j_o}(X_{j_i}(..(X_{j_v}(g(x))))))\Big|_{x(o)} = C^T A_{j_v}\ \ldots\ A_{jv}x_o$$

Il est à noter dans (35) le renversement de l'ordre des indices.

Cette formule peut être obtenue directement par l'utilisation de la formule de Peano Baker sur le système (22) |JAC,82|.

**II.3.3.3.** Fonction de transfert codée.

De façon a obtenir une formulation plus compacte de (32), intégrons le système bilinéaire (22), il vient :

$$(36)\qquad \begin{cases} x(t) = x_0 + A_0 \int_0^t x(\tau)\,d\tau + \sum_{i=1}^{m} A_i \int_0^t u_i(\tau)\,x(\tau)\,d\tau \\ \qquad\qquad + \sum_{i=1}^{m} B_i \int_0^t u_i(\tau)\,d\tau \\ y(t) = C^T x(t) \end{cases}$$

Si on reprend les notations introduites dans la partie du calcul symbolique non commutatif, il vient en notant S la série génératrice vectorielle de x, $(S = (s_1,\ldots,s_2)^T)$ :

$$(37)\qquad \begin{cases} S = x_0 + A_0\alpha_0 S + \sum_{i=1}^{m} A_i\,\alpha_i\,S + \sum_{i=1}^{m} B_i\alpha_i \\ s = C^T S \end{cases}$$

ce qui donne l'expression

$$(38)\qquad s = C^T \left[ I - \sum_{i=0}^{m} A_i\alpha_i \right]^{-1} \left[ x_0 + \sum_{i=1}^{m} B_i\alpha_i \right]$$

qui est une expression équivalente à (32), mais qui a la forme d'une fonction de transfert. En effet, dans le cas d'un système linéaire (i.e. i = 1,..., m : $A_i \equiv 0$), il viendrait, à conditions initiales nulles

$$(39)\qquad s = \sum_{i=1}^{m} C^T \left[ I - A_0\alpha_0 \right]^{-1} B_i\alpha_i$$

la $i^{\text{ème}}$ fonction de transfert est donc représentée par $C^T(I-A_0\alpha_0)^{-1}B_i$, ce qui, par un simple changement de notation, peut être ramené à la forme classique en p (2) de la fonction de transfert d'un système linéaire | FLI,81 |.

### II.3.3.4 Matrices de Hankel

Un autre outil, utilisé dans l'étude des systèmes linéaires | KAI,80 | et étendu aux systèmes bilinéaires | ARB,80 | est constitué par la notion de matrices de Hankel. Elles sont définies à partir de la série génératrice (6) du système étudié, de la façon suivante | BER,54 ; FLI,74 | :

$$(40)\qquad H(s) = \left[ H(s)_{\underbrace{\alpha_{j_0}\cdots\alpha_{j_\nu}}_{\text{indice de ligne}},\ \underbrace{\alpha_{k_0}\cdots\alpha_{k_\mu}}_{\text{indice de colonne}}} \right] \quad \nu \geq 0 \quad \mu \geq 0$$

$$j_0\ldots j_\nu,\ k_0\ldots k_\mu \in \{0,\ldots,m\}^{\nu+\mu+2}$$

$$(41)\qquad H(s)_{\alpha_{j_o}\ldots\alpha_{j_\nu},\,\alpha_{k_o}\ldots\alpha_{k_\mu}} = \left(X_{k_\mu}\left(..\left(X_{k_o}\left(X_{j_\nu}\left(..\left(X_{j_o}(g(x)..\right)\right)\right|_{x(o)}$$

Dans le cas des systèmes linéaires, les éléments de cette matrice sont les paramètres de Markov du système | KAI,80 | et dans le cas d'un système bilinéaire régulier, d'après (32), la matrice de Hankel a la structure suivante | JAC,81 |

$$(42)\qquad H_o = \left[H(s)_{\alpha_{j_o}\ldots\alpha_{j_\nu},\,\alpha_{k_o}\ldots\alpha_{k_\mu}}\right]_{\substack{\nu\geqslant o\\ \mu\geqslant o\\ j_o..j_\nu,\ k_o..k_\mu\in\{0,\ldots,m\}^{\nu+\mu+2}}}$$

$$H(s)_{\alpha_{j_o}..\alpha_{j_\nu},\,\alpha_{k_o}..\alpha_{k_\mu}} = C^T A_{j_o}..A_{j_\nu}A_{k_o}..A_{k_\mu}x_o$$

Cette matrice est notée avec un indice o car elle dépend de conditions initiales. Dans le cas non régulier, on peut se ramener à un système régulier par un vecteur d'état homogène, ou bien définir une matrice de Hankel associée à chaque entrée et définie, suivant (38) par

$$(43)\qquad i = 1\ldots m \qquad H_i = \left[H_i(s)_{\alpha_{j_o}\ldots\alpha_{j_\nu},\alpha_{k_o}..\alpha_{k_\mu}}\right]_{\substack{\nu\geqslant o\\ \mu\geqslant o\\ j_o..j_\nu,\ k_o\ldots k_\mu\in\{o\ldots m\}^{\nu+\mu+2}}}$$

$$\text{avec } H_i(s)_{\alpha_{j_o}\ldots\alpha_{j_\nu},\alpha_{k_o}\ldots\alpha_{k_\mu}} = C^T A_{j_o}\ldots A_{j_\nu}A_{k_o}\ldots A_{k_\mu}B_i$$

On peut remarquer que si on définit les matrices de dimensions infinies :

$$(44)\qquad \begin{cases} \mathcal{C} = \left[\begin{array}{c}\vdots\\ \cdots\cdots\\ C^T A_{j_o}\ldots A_{j_\nu}\\ \cdots\cdots\\ \vdots\end{array}\right]_{\substack{\nu\geqslant o\\ j_o,\ldots,j_\nu\in\{0,\ldots,m\}^{\nu+1}}} \\ i = 1,\ldots,m\quad \mathcal{B}_i = \left[\ldots\vdots(A_{k_o}\ldots A_{k_\mu}B_i)\vdots\ldots\right]_{\substack{\mu\geqslant o\\ k_o,\ldots,k_\mu\in\{0,\ldots,m\}^{\mu+1}}} \end{cases}$$

il vient pour tout i de 1,..., m

$$(45)\qquad H_i = \mathcal{C}.\mathcal{B}_i$$

ce qui généralise le fait que, dans le cas des systèmes linéaires, la matrice de Hankel est le produit de la matrice d'observabilité par la matrice de commandabilité | KAI,80 |.

Dans le cas des systèmes non linéaires, la matrice de Hankel associée à la série génératrice de la sortie permet de décider de la réalisation bilinéaire du système | ISI,73 |. En effet, le système étudié est bilinéaire si et seulement si sa matrice de Hankel est de rang fini q | FLI,74 | et un procédé de réalisation d'un système à partir de sa matrice de Hankel sous la forme d'un système bilinéaire régulier a été proposé dans | JAC,81 |.

## III. BILINEARISATION DES SYSTEMES NON LINEAIRES

De façon à approcher un système non linéaire par un modèle bilinéaire, plusieurs méthodes ont été proposées : certaines sont basées sur la théorie géométrique des algèbres de Lie pour déterminer des conditions d'existence de transformations non linéaires | KRE,73 ; TIN,75 | ou des retours d'état | CLA,85 ; ZAK,86 ; KRE,83 | pour linéariser globalement le système étudié ; une autre utilise les séries formelles | HES,85 |. L'inconvénient de ces méthodes est leur difficulté d'application dans le cas de systèmes de grandes dimensions.

Un autre principe de méthode, à partir de l'idée originelle de Carleman | CAR,32 |, utilise les puissances des composantes de l'état. Celles-ci ont été employées par divers auteurs pour bilinéariser les systèmes non linéaires | SVO,80 ; KRE,74 ; BRO,76a ; ARM,81 ; CEB,84 ; BRO,76b |, mais sans obtenir de méthode systématique de détermination du modèle bilinéaire. Par l'utilisation du produit tensoriel de l'espace d'état, nous présentons un algorithme récursif qui permet de construire directement le modèle cherché.

### III.1 NOTATIONS ET PROPRIETES

On supposera que le système à bilinéariser est sous la forme (5) analytique. On peut donc écrire :

$$(46) \quad X_0(x) = \sum_{i=1}^{\infty} F^i x^{[i]} \quad , \quad X_i(x) = \sum_{j=0}^{\infty} g_i^j x^{[j]}, \; i = 1,\ldots,m$$

où $x^{[i]}$ représente un vecteur redondant de dimension $q^i$ de formes i-homogènes en $x_1, \ldots, x_q$, composantes de x. Les $x^{[i]}$, introduits par | VET,70 |, pour généraliser les séries de Taylor sont définis par :

$$(47) \quad \begin{cases} x^{[0]} = 1 \\ \forall\, i \in \mathbb{N}^* \quad x^{[i]} = x^{[i-1]} \otimes x^{[1]} \end{cases}$$

où $\otimes$ est le produit tensoriel ou de Kronecker défini par :
si $A = [a_{ij}]$ et $B = [b_{ij}]$ sont deux matrices de dimensions respectives n x p et m x q

$$(48) \quad A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1p}B \\ \vdots & & \vdots \\ a_{n1}B & \cdots & a_{np}B \end{bmatrix}$$

$A \otimes B$ est une matrice de dimension nm x pq

Ainsi, si $x = [x_1, \ldots, x_q]^T$ et $y = [y_1, \ldots, y_p]^T$ alors

(49)

$$x \otimes y = [x_1y_1, x_1y_2, \ldots, x_1y_p, x_2y_1, x_2y_2, \ldots x_2y_p, \ldots, x_{q-1}y_p, x_qy_1, \ldots, x_qy_p]^T$$

et dans le cas où x = y, $x \otimes x$ est une forme redondante puisque $x_1x_2$, par exemple, est présent deux fois dans ce vecteur.

Concernant le produit de Kronecker utilisé, on a les propriétés suivantes (d'autres sont indiquées dans | BRE,78 |) :

Propriétés : Si x et y sont des vecteurs de dimensions p et q, A et B sont des matrices de dimensions (nxp) et (mxq), on a :

a) $(Ax) \otimes y = [A \otimes I_q]\,(x \otimes y)$
b) $x \otimes (By) = [I_p \otimes B]\,(x \otimes y)$
c) $(Ax) \otimes (By) = [I_n \otimes B]\,[A \otimes I_q]\,(x \otimes y)$
$= [A \otimes I_m]\,[I_p \otimes B]\,(x \otimes y)$

où $I_r$ est la matrice identité d'ordre r (r=p,n,q,m)

d) $y \otimes x = P^{qp}(x \otimes y)$

où $P^{qp}$ est une matrice pq x pq de permutation :

$$P^{qp} = \begin{bmatrix} P_{11} & P_{12} \cdots\cdots\cdots\cdots & P_{1p} \\ P_{21} & & \\ \vdots & & \\ \vdots & & \\ P_{q1} & \cdots\cdots\cdots\cdots\cdots & P_{qp} \end{bmatrix}$$

$P_{ij}$, $(i,j) \in [1,\ldots q] \times [1,\ldots,p]$, est une pq-matrice composée de 1 à l'intersection de sa $i^{ème}$colonne et de sa $j^{ième}$ ligne et 0 partout ailleurs.

e) $y \otimes x = (y \otimes I_p)x$

f) $\frac{d}{dt}(x \otimes y) = \frac{dx}{dt} \otimes y + x \otimes \frac{dy}{dt}$

## III.2 METHODE D'APPROXIMATION

Cette méthode est basée sur l'approximation par troncature | CAR,32 | ; en considérant le vecteur de dimension infinie :

$$(50) \qquad X = \left[ x^{[1]^T}, x^{[2]^T}, \ldots, x^{[i]^T}, \ldots \right]^T,$$

le modèle bilinéaire sera obtenu en écrivant l'équation d'état vérifiée par le vecteur $x^{[i]}$et en supprimant les termes de degré supérieur à n, avec

$$(51) \qquad X_n = \left[ x^{[1]^T}, x^{[2]^T}, \ldots, x^{[n]^T} \right]^T$$

Il vient la proposition suivante :

"le vecteur-redondant $X_n$(51) est régi par l'équation d'état bilinéaire

$$(52a) \qquad \overset{\circ}{X}_n = F_n X_n + \sum_{i=1}^{m} u_i(t)\, G_{i_n} X_n + G_n u(t)$$

où

$$F_n = \begin{bmatrix} F^1 & F^2 \cdots\cdots F^n \\ 0 & F^2_{[2]} \cdots F^n_{[2]} \\ \vdots & \vdots \\ 0 & 0 \quad F^n_{[n]} \end{bmatrix} \qquad G_{i_n} = \begin{bmatrix} g_i^1 & g_i^2 \cdots\cdots g_i^n \\ g_i^1{}_{[2]} & g_i^2{}_{[2]} \quad g_i^2{}_{[2]} \\ 0 & \vdots \\ \vdots & \vdots \\ 0 & 0 \quad g_i^{n-1}{}_{[n]} \quad g_i^n{}_{[n]} \end{bmatrix}$$

$i = 1,\ldots,m$

$$G_n = \begin{bmatrix} g_1^o \vdots g_2^o \vdots \cdots\cdots \vdots g_m^o \\ 0 \cdots\cdots\cdots\cdots 0 \\ \vdots \qquad\qquad \vdots \\ 0 \cdots\cdots\cdots\cdots 0 \end{bmatrix}$$

et on a pour $X_n(o) = \left[x^{[1]^T}(o), \ldots, x^{[n]^T}(o)\right]^T$

$$\lim_{n \to \infty} \left[\underbrace{1 \ldots 1}_{q}, o \text{ —— } o\right] X_n = x(t) \ ;$$

* les $F^i_{[j]}$ et $g^k_{i\,[j]}$ sont définis par les relations de récurrence suivantes :

(52b)

$$\begin{aligned} & r = 2, \ldots, n \\ & \quad * \ i = r, .., n \quad F^i_{[r]} = I_q \otimes F^{i-1}_{[r-1]} + F^{i-r+1} \otimes I_q(r-1) \\ & \quad * \ i = 1, \ldots, m \\ & \qquad j = r-1, \ldots, n \\ & \qquad g^j_{i\,[r]} = I_q \otimes g^{j-1}_{i\,[r-1]} + g^{j-r+1}_{[i]} \otimes I_q(r-1) \end{aligned}$$

avec les conditions initiales

$$\begin{aligned} & * \ i = 1, \ldots, n-1 : F^i_{[1]} = F^i \\ & * \ i = 1, \ldots, m \ ; \ j = 1, \ldots, n-1 : g^j_{i\,[1]} = g^j_i \end{aligned}$$

Ces relations peuvent être démontrées par récurrence | ROT,86a |. Si on suppose que , pour $r \in [2, \ldots, n]$ :

$$(53) \qquad \overset{\circ}{x}{}^{[r-1]} = \sum_{i=r-1}^{n} F^i_{[r-1]} x^{[i]} + \sum_{i=1}^{m} u_i(t) \left\{ \sum_{j=r-1}^{n} g^j_{i\,[r-1]} \, x^{[j]} \right\}$$

l'expression de la dérivée de $x^{[r]}$ est donnée, en éliminant les termes du degré supérieur à n, par :

$$(54) \quad \overset{\circ}{x}{}^{[r]} = \sum_{i=1}^{n} (F^i x^{[i]}) \otimes x^{[r-1]} + \sum_{i=1}^{m} u_i(t) \left[ \sum_{j=o}^{n} (g^j_i \, x^{[j]}) \otimes x^{[r-1]} \right]$$

$$+ \sum_{i=r-1}^{n} x^{[1]} \otimes (F^i_{[r-1]} x^{[i]}) + \sum_{i=1}^{m} u_i(t) \left[ \sum_{j=r-2}^{n} x^{[1]} \otimes (g^j_{i\,[r-1]} x^{[j]}) \right]$$

L'utilisation des propriétés énoncées du produit de Kronecker permet d'écrire :

$$(55)\qquad \overset{\circ}{x}[r] = \sum_{i=r}^{n} \underbrace{\left(F^{i-r+1} \otimes I_q(r-1) + I_q \otimes F_{[r-1]}^{i-1}\right)}_{F_{[r]}^{i}} x^{[i]}$$

$$+ \sum_{i=1}^{m} u_i(t) \left[ \sum_{j=r-1}^{n} \underbrace{\left(g_i^{j-r+1} \otimes I_q(r-1) + I_q \otimes g_{i\,[r-1]}^{j-1}\right)}_{g_{i\,[r]}^{j}} x^{[j]} \right]$$

La proposition énoncée permet de construire un système bilinéaire approchant un système non-linéaire analytique pour n (degré d'approximation) fixé. Le passage d'un modèle de degré d'approximation n à n+1 ne nécessite que le calcul de $(2n+m)$ matrices supplémentaires et le modèle de degré d'approximation 1 correspond au système linéarisé. Les modèles bilinéaires apparaissent donc comme une généralisation "naturelle" des modèles linéaires.

## III.3 ELIMINATION DE LA REDONDANCE

L'état redondant $X_n$ défini dans la partie précédente a une dimension de $\sum_{i=1}^{n} q^i$. Cette dimension peut être diminuée en éliminant les variables qui se répètent dans $X_n$ ; définissons la forme non-redondante de $X_n$ par :

$$(56)\qquad \tilde{X}_n = \left[ \tilde{x}[1]^T , \tilde{x}[2]^T , \ldots , \tilde{x}[n]^T \right]^T$$

$$\tilde{x}[1] = x[1] = x, \text{ et pour } r \text{ dans } [2,..,n]$$

$$\tilde{x}[r] = \left[ x_1^r, x_1^{r-1}x_2, \ldots, x_1^{r-1}x_q, x_1^{r-2}x_2^2, \right.$$

$$\left. x_1^{r-2}x_2x_3, .., x_1^{r-2}x_2x_q, \ldots, x_1^{r-2}x_q^2, x_1^{r-3}x_2^3, \ldots, x_q^r \right]^T$$

on a, de plus, les relations :

$$(57)\qquad \forall\, r \in [1, \ldots, n] \quad \exists!\, T_r \in \mathbb{R}^{q^r \times \binom{q+r-1}{r}}$$

$$x[r] = T_r \tilde{x}[r]$$

où $\binom{\ }{\ }$ désigne le coefficient du binôme

Ainsi, on peut écrire :

$$(58) \qquad X_n = {}_nT\tilde{X}_n \text{ avec } {}_nT = \underset{i=1,..n}{\text{diag}} \{T_i\}$$

et la relation inverse

$$(59) \qquad \overset{\circ}{\tilde{X}}_n = {}_nT^+ X_n \text{ avec } {}_nT^+ = \underset{i=1,..,n}{\text{diag}} \{T_i^+\}$$

où $T_i^+$ est la pseudo-inverse au sens de Moore et Penrose | BEN,73 | de $T_i$, $i = 1,...,n$.

De la proposition d'approximation (52), on déduit que $\tilde{X}_n$ est décrit par le modèle bilinéaire :

$$(60) \qquad \overset{\circ}{\tilde{X}} = \tilde{F}_n\tilde{X}_n + \sum_{i=1}^{n} u_i(t)\, \tilde{G}_{i_n}\tilde{X}_n + \tilde{G}^n u(t)$$

$$\text{où } \tilde{F}_n = {}_nT^+F_{nn}T \text{ , } \tilde{G}^n = {}_nT^+G^n$$

$$i = 1,..., m: \tilde{G}_{i_n} = {}_nT^+G_{i_n}{}_nT$$

ce qui, compte tenu de la structure de ${}_nT$ et ${}_nT^+$, assure que les structures bloc-triangulaire et bloc-Heissenberg de $F_n$ et $G_{i_n}$ sont conservées pour $\tilde{F}_n$ et $\tilde{G}_{i_n}$. Avec des notations évidentes, on a, de plus,

$$\begin{aligned} &i = 1,..., n \;: \tilde{F}^i = F^iT_i \\ &\qquad\qquad j = 2,..., n \;: \tilde{F}^i_{[j]} = T^+_j F^i_{[j]}\, T_i \\ &(61) \qquad i = 1,..., m \;: j = 1,..., n \;: k = 2,..., n \\ &\qquad\qquad \tilde{g}^j_i = g^j_i\, T_j \\ &\qquad\qquad \tilde{g}^j_{i[k]} = T^+_k g^j_{i[k]} T_j \\ &\qquad \tilde{G}^n = G^n \end{aligned}$$

La dimension de $\tilde{X}_n$ est alors de : $\sum_{i=1}^{n} \binom{n+i-1}{i}$

## III.4 EXEMPLE D'APPLICATION

Considérons le système défini par l'équation non linéaire

(62), étudié dans | FLI,83 ; HES,85 |, et qui représente l'équation de fonctionnement d'un schéma électrique, figure 3, à résistance non constante.

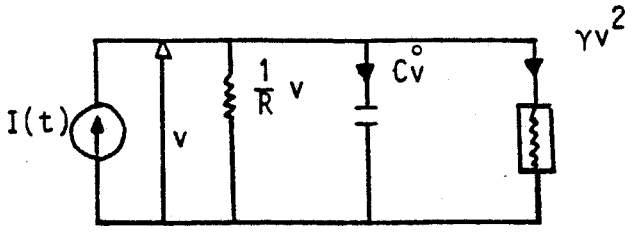



FIGURE 3

$$(62) \qquad \dot{v}(t) + k_1 v(t) + k_2 v^2(t) = i(t)$$
$$\text{avec } v(o) = \alpha,\ k_1 = \frac{1}{RC},\ k_2 = \frac{\gamma}{C},\ u(t) = \frac{I(t)}{C}$$

L'équation algébrique qui permet de déterminer par itération la série formelle s associée à v est | FLI,83 | :

$$(63) \qquad s = -k_1 \alpha_o s - k_2 \alpha_o s \sqcup\!\sqcup s + \alpha_1 + \alpha$$

L'application de la méthode exposée dans le paragraphe précédent conduit, si l'on définit le vecteur état tronqué :

$$(64) \qquad V_n = \left[ v(t), v^2(t), \ldots, v^n(t) \right]^T$$

à une équation bilinéaire

$$(65) \qquad \dot{V}_n(t) = \begin{bmatrix} -k_1 & -k_2 & & 0 \\ & -2k_1 & -2k_2 & \\ & & \ddots & -(n-1)k_2 \\ 0 & & & -nk_1 \end{bmatrix} V_n(t)$$
$$+ u(t) \begin{bmatrix} 0 & & & \\ & 2 & & 0 \\ & & \ddots & \\ 0 & & & n \end{bmatrix} V_n(t) + \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(t)$$
$$V_n(o) = \left[ \alpha, \alpha^2, \ldots, \alpha^n \right]$$

Les résultats de simulation montrent que, pour n fixé, les modèles obtenus par | HES,85 | et par (65) sont équivalents ; mais, dans | HES,85 |, la matrice-système, d'une part, dépend des conditions initiales et, d'autre part, est plus délicate à obtenir. Dans le cas de fonctionnement autonome, la méthode de bilinéarisation conduit à un modèle linéaire valide de façon globale ; quelques applications de ces réalisations infinies, notamment dans l'étude de la stabilité des systèmes non linéaires, ont été présentés dans | ROT,86b |.

## CONCLUSION

Nous avons dégagé, au cours de cette partie, un certain nombre de points de vue utiles dans la modélisation des systèmes non linéaires .

D'une part, nous avons insisté tout particulièrement sur l'outil des séries formelles qui, lorsqu'on considère un système non linéaire, permet de pouvoir tirer algébriquement quelques enseignements (notamment au niveau des transitoires).

D'autre part, nous avons rappelé que la structure des systèmes bilinéaires leur permettait de pouvoir bénéficier de certaines notions, comme la fonction de transfert, propres aux systèmes linéaires. Ceci constitue un des premiers intérêts de la prise en compte de ce type de modèles.

Le deuxième point vient du fait que la plupart des systèmes non linéaires, lorsqu'ils se présentent sous la forme de systèmes à commande affine (ce qui est un cas relativement général | BAL,76 |), peuvent être approchés avec une précision arbitraire par des modèles bilinéaires.

Concernant ce dernier problème, on peut, soit utiliser une des méthodes qui ont été mises au point pour identifier directement un système réel sous la forme d'un modèle bilinéaire | DAD,85;MOH,76;CHE,78;KAR,78 | ; soit, à partir d'un modèle non-linéaire, employer une méthode de bilinéarisation.

Nous avons, en fin de ce chapitre, présenté une méthode pratique basée sur un algorithme récursif qui permet de déterminer directement un modèle bilinéaire approché.

# CHAPITRE II

# SIMPLIFICATION DE MODELES

# INTRODUCTION

La description précise de systèmes réels conduit à des modèles de plus en plus complexes ou/et de dimensions élevées. De ce fait, l'analyse et l'application des résultats de la théorie de la commande optimale à de tels processus sont difficiles : le volume et la complexité des calculs nécessaires s'avérant rapidement prohibitifs. Ces dernières années ont donc vu croître l'intérêt porté aux méthodes de simplification de modèles afin d'élaborer les stratégies de commandes sous-optimales | BOS, 72 | .

Dans un premier temps, nous rappellons les différentes techniques de réduction de modèles. Elles peuvent être regroupées en trois classes : troncatures, agrégations et perturbations.

La méthode des perturbations singulières, que nous détaillons dans la deuxième partie, présente l'intérêt d'être applicable aux cas des systèmes non-linéaires. Avant de l'utiliser sur les systèmes bilinéaires, nous présentons une nouvelle méthode de découplage des parties lentes et rapides d'un système linéaire singulièrement perturbé. L'intérêt de cette modélisation sera mis en évidence dans le chapitre traitant de la commande optimale.

# I. METHODES DE SIMPLIFICATION DE MODELES

La simplification du modèle peut recouvrir plusieurs notions très différentes : en effet, on peut simplifier un modèle par une diminution de la complexité de sa structure par découplage ou décomposition ; ou bien, on peut l'approcher par un modèle d'ordre inférieur par réduction de dimensionnalité. Les différentes techniques utilisées dans le dernier point de vue peuvent se regrouper en plusieurs classes de méthodes | BIN,78 ; FOS,86 | :

* la réduction optimale consistant à minimiser une fonctionnelle de l'erreur entre le système et un modèle choisi "a priori",

* la méthode par approximation des fonctions de transfert,

* les techniques d'agrégation consistant à déterminer un système d'ordre inférieur dont le vecteur état soit une bonne approximation des composantes représentatives du vecteur état initial,

* les méthodes de perturbations qui normalisent, par un petit paramètre, des termes de grandeurs différentes ; le passage à la limite, permet de déterminer un modèle approché.

Les deux premières catégories seront plus brièvement rappelées que les deux suivantes et nous ne donnerons dans chaque cas que quelques références significatives.

## I.1. REDUCTION OPTIMALE

Le problème de la réduction optimale peut s'énoncer de la façon suivante : calculer les paramètres d'un modèle linéaire d'ordre inférieur au modèle initial qui minimiserait une fonctionnelle de l'erreur entre les sorties des deux modèles pour une entrée donnée appliquée simultanément au modèle initial et au modèle réduit. Ce

problème peut être abordé à partir de représentation par fonctions de transfert, on utilise alors une approche fréquentielle | AIG,49 ; MEI,67 | qui conduit à la résolution d'un système d'équations non linéaires pour déterminer les paramètres du modèle réduit. L'approche temporelle part, quant à elle, de la représentation d'état des modèles initial et réduit ; | WIL,74 |, en raisonnant sur le système élargi dont l'état est la somme directe des états de chacun des systèmes, a donné des conditions d'optimalité sur le système réduit, ces conditions nécessitent la résolution d'équations de Lyapounov Il est possible dans cette approche d'imposer également une contrainte égalité sur les gains statiques | DEC,76 |.

Les inconvénients de ce type de méthode sont, d'une part, de nécessiter des algorithmes de résolution d'équations qui, dans certains cas, peuvent s'avérer ne pas être convergents et, d'autre part, que la signification physique du modèle simplifié n'existe généralement plus, ce modèle étant de toute façon limité à une classe d'entrées particulières.

## I.2 APPROXIMATION DE FONCTIONS DE TRANSFERT

Les méthodes présentées ici sont issues de théories générales d'approximation des fonctions autour d'un point à l'aide de leur développement en séries de Taylor. Plus ou moins sophistiquées, elles permettent de construire un modèle réduit qui possède quelques caractéristiques du modèle initial, mais pas toutes.

Soit F(p), la fonction de transfert d'un système linéaire qui admet comme réalisation le triplet (A,B,C) ; on peut définir les moments, $m_i$, ou les paramètres de Markov, $p_i$ :

$$(1) \qquad i = 1,\ldots,\infty \quad m_i = -CA^{-i}B, \; p_i = CA^{i-1}B$$

qui permettent de construire, en conservant les 2q premiers, des modèles réduits d'ordre q | ASH,82 ; GUT,82 |. Dans le cas où l'on utilise les moments, on obtient une approximation de F(p) autour de p = 0 ou approximant de Padé | LAL,74 ; ZAK,73 | qui a la propriété,

si les deux modèles sont stables, d'assurer la même erreur statique que le modèle initial pour toute entrée du type $\sum_{i=0}^{2q} \alpha_i t^i$ .On peut approcher une fonction de transfert autour de n'importe quel point | SHA 74 ; COM,77 |, mais dans la pratique, seule l'approximation autour de p = ∞ a été utilisée. On emploie alors dans ce cas les paramètres de Markov ; l'identification par cette méthode permet d'obtenir une meilleure qualité de transitoire et de mieux reproduire les qualités de stabilité du modèle initial sans toutefois qu'il y ait équivalence.

Utilisées depuis longtemps en théorie des circuits, notamment pour la synthèse, les décompositions en fractions continues se sont révélées être un outil commode de simplification des fonctions de transfert | SHA,74 |. CHEN | CHE,61 ; CHE,70 | a proposé pour la décomposition la forme :

$$(2) \qquad F(p) = \cfrac{1}{h_1 + \cfrac{1}{\cfrac{h_2}{p} + \cfrac{1}{h_3 + \ldots\ldots\ldots\ldots}}}$$

qui conduit à représenter F(p) par une suite de chaînes directes et de réactions imbriquées. Cette structure peut être étendue au cas multivariable | CHE,74 |. Plusieurs formes mixtes de fractions continues ont été proposées par | CHU,70 ; KHA,80 | ou | GOL,77 | pour la forme mixte de Cauer qui conduisent à des modèles réduits d'ordre m dont les m moments et m paramètres de Markov sont identiques à ceux de F(p).

L'inconvénient de ces méthodes est de ne pas traduire, dans le modèle réduit, la stabilité ou l'instabilité du modèle initial. La méthode de Routh | HUT,75a | conserve les premiers éléments du tableau de Routh, ce qui garantit la stabilité du système réduit si le système initial est stable. Le modèle obtenu est une approximation de F(p) autour de p = ∞ (hautes fréquences) ; pour se ramener autour de p = 0, Hutton utilise la transformation réciproque :

(3) $$F(p) \longrightarrow \tilde{F}(p) = \frac{1}{p} F(\frac{1}{p}) \; ;$$

l'algorithme de réduction se déroule donc comme suit | HUT,75b | :

(4) $$F(p) \xrightarrow[\text{récip}]{\text{transf}} \tilde{F}(p) \xrightarrow[\text{Routh}]{\text{réduc}} \tilde{\tilde{F}}_r(p) \xrightarrow[\text{récip}]{\text{transf}} \hat{F}(p)$$

modèle initial — modèle réduit

Toujours dans le but de préserver la stabilité par réduction d'ordre, on peut utiliser conjointement plusieurs de ces méthodes | SHA,75 ; APP,78 |.

Les systèmes bilinéaires et non linéaires ne possédant pas de fonction de transfert, il n'est pas possible de leur appliquer les méthodes précédentes. Cependant, nous avons vu, au chapitre précédent, que la relation d'entrée-sortie de ces systèmes peut être décrite de façon symbolique par une série génératrice qui est une série formelle en plusieurs indéterminées non commutatives. Cette série s'exprimant simplement à l'aide des paramètres de Markov du système (Cf.I:(8)), il serait alors possible, par l'utilisation des approximants de Padé à plusieurs variables | CHI,73 ; CHI,74 ; BOS,80 | restreints au cas non commutatifs, de construire un système approché du système initial. Ce système réduit serait bilinéaire car, par définition d'un approximant de Padé, sa série génératrice est obtenue par une suite finie d'opérations élémentaires (somme, produit, inversion) sur des polynomes formels | SCH,61 |. La réalisation du système réduit peut alors être envisagée par la méthode proposée dans | JAC,75 |. Considérons, par exemple, le système bilinéaire monovariable :

(5) $$\begin{aligned} \dot{x} &= A_0 x + u A_1 x + B u \\ y &= C x \end{aligned}$$

où $x \in \mathbb{R}^2$, $y \in \mathbb{R}$, $u \in \mathbb{R}$

dont la série génératrice est :

(6) $$s = CB + CA_0B\alpha_0 + CA_1B\alpha_1 + CA_0^2B\alpha_0^2 + CA_0A_1B\alpha_0\alpha_1 + \ldots$$

que l'on désire approcher par un système réduit d'ordre 1, selon la méthode des approximants de Padé. La matrice de Hankel (Cf. I : II.3.3.4) doit donc être de la forme :

$$(7) \qquad H_r = \begin{bmatrix} (CB) & (CA_0B) & (CA_1B)\ldots \\ (CA_0B) & \alpha & \beta \ \ldots \\ (CA_1B) & \beta & \gamma \ \ldots \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}$$

où $\alpha, \beta, \gamma$ sont définis par la contrainte que tout mineur d'ordre 2 extrait de $H_r$ soit singulier. On a donc, si $CB \neq 0$,

$$(8) \qquad \alpha = \frac{(CA_0B)^2}{CB} \quad , \beta = \frac{(CA_1B)\ (CA_0B)}{CB} \quad , \gamma = \frac{(CA_1B)^2}{CB}$$

et la réalisation du système d'ordre 1 associé à la matrice $H_r$ donne :

$$(9) \qquad \overset{\circ}{x}_r = \frac{CA_0B}{CB} x_r + \frac{CA_1B}{CB} x_r + CBu$$

$$y_r = x_r \text{ où } x_r \in \mathbb{R}$$

La série génératrice de la sortie de ce système ayant les mêmes premiers termes que la série génératrice du système initial, cette méthode fournit, comme dans le cas linéaire, un système permettant d'approcher le transitoire du système initial.

Nous avons utilisé le conditionnel dans ce paragraphe pour indiquer que cette question est ouverte, mais déborde largement du cadre de notre exposé.

## I.3. TECHNIQUES D'AGREGATION

Les méthodes d'agrégation pour la réduction de modèles

ont été introduites en Automatique par | AOK,68 |. Elles consistent à substituer au modèle initial un modèle de taille réduite par projection du vecteur état. Le principe en est le suivant :

Si $(\Sigma)$, le modèle initial, et $(\hat{\Sigma})$ le modèle réduit sont représentés par :

$$(10a) \qquad (\Sigma) \begin{cases} \dot{x} = Ax + Bu & x \in \mathbb{R}^q,\ u \in \mathbb{R}^n \\ y = Cx & y \in \mathbb{R}^n \end{cases}$$

$$(10b) \qquad (\hat{\Sigma}) \begin{cases} \dot{z} = Fz + Gu & z \in \mathbb{R}^{q'} \text{ avec } q' < q \\ \hat{y} = Hz & \hat{y} \in \mathbb{R}^n \end{cases}$$

alors $(\hat{\Sigma})$ est un modèle agrégé de $(\Sigma)$ si z et x vérifient la relation d'agrégation linéaire :

$$(11) \qquad z = Lx,\ L \in \mathbb{R}^{q' \times q} \qquad \forall x \text{ si } z_0 = Lx_0$$

Il est, bien sûr, concevable d'envisager une relation d'agrégation non linéaire. Compte tenu de (10a-b), l'existence de L est assurée si :

$$(12) \qquad \begin{array}{l} (a)\ FL = LA \\ (b)\ G = LB \\ (c)\ HL = C \end{array}$$

Le spectre de F étant inclus dans celui de A | BER,79 |, une approche développée par| DAV,68 ; MAR,66 ; CHI,67 ; FOS,70 ; MIT,69 ; WIL,72 ; HIC,75 | consiste à choisir comme modes du modèle réduit les modes dominants du système réel. Les différents modèles ainsi obtenus constituent des cas particuliers de modèles agrégés où L est de la forme | BER,76 | :

$$(13) \qquad L = M\ [I_{q'}, 0\ ]\ T^{-1}$$

où T est la matrice régulière bloc diagonalisant A et M une matrice arbitraire régulière de dimension q'.

Ces méthodes sont utilisées pour construire des régulateurs sous-optimaux du système initial | AOK,68 ; MIC,79 |, mais la question importante est de savoir quels modes sont à retenir dans le modèle réduit. Une réponse peut être apportée en minimisant une fonctionnelle de l'erreur entre la sortie du modèle initial et la sortie du modèle réduit | COM,78 | ou en utilisant une mesure de dominance | LIT,81; MIC,86 |.

Dans tous les cas, pour un choix a priori de L, on peut définir un modèle réduit approché ($\Sigma^+$) défini par | KHR,85 | :

$$\begin{aligned} \dot{z} &= LAL^+ z + LB\, u \\ \hat{y} &= CL^+ z \quad ; z(o) = Lx(o) \end{aligned} \tag{14}$$

où $L^+$ est une pseudo-inverse à droite de L telle que $LL^+ = I_{q'}$

et il est à noter que le spectre de $LAL^+$ n'est plus inclus dans celui de A. Ce point de vue permet de définir des classes de méthodes de réduction d'ordre par agrégation et approximation | PIN,81 | où les modes propres du système réduit peuvent être choisis arbitrairement.

Ces techniques d'agrégation ont été étendues au cas des systèmes bilinéaires | LAS,84 | ; nous en citerons ici les principaux résultats.

Soient un modèle bilinéaire $(\Sigma_b)$ et un modèle réduit $(\hat{\Sigma}_b)$, définis par :

$$\begin{aligned} &(a)\ (\Sigma_b) : \dot{x} = Ax + \sum_{i=1}^{m} u_i\, A_i x + Bu \\ &(b)\ (\hat{\Sigma}_b) : \dot{z} = Fz + \sum_{i=1}^{m} u_i\, F_i z + Gu \end{aligned} \tag{15}$$

$$\text{où } x \in \mathbb{R}^q,\ z \in \mathbb{R}^{q'},\ q' < q,\ u^T = [u_1, \ldots, u_m] \in \mathbb{R}^m$$

L'agrégation étant réalisée par l'application linéaire (11), (15b) est un système agrégé de (15a) si sont vérifiées les relations (12-a), (12-b) et

$$\forall\, i \in \{1, \ldots, m\} \qquad F_i L = LA_i \tag{16}$$

Soit T la matrice des vecteurs propres de A , les conditions de compatibilité du système (12a) (16) sont alors données par le résultat suivant :

**Théorème | LAS,84 |** : Le système d'équation (16) est résoluble si et seulement si :

$$(17) \qquad \forall\, i \in \{1,\ldots,m\} \quad T^{-1}A_iT = \begin{bmatrix} S_{11i} & 0 \\ S_{21i} & S_{22i} \end{bmatrix}$$

où $S_{11i} \in \mathbb{R}^{q' \times q'}$

Lorsque ce théorême est vérifié, la matrice d'agrégation a la forme (13) et les matrices $F_i$ intervenant dans le système agrégé (15b) sont données par :

$$(18) \qquad \forall\, i \in \{1,\ldots,m\} \quad F_i = MS_{11i}M^{-1}$$

Dans le cas où les conditions du théorême ne sont pas vérifiées, on peut construire un modèle réduit approché bilinéaire $\hat{\Sigma}_b$ où :

$$(19) \qquad F = LAL^+,\ G = LB$$
$$\forall\, i \in \{1,\ldots,m\}\ F_i = LA_iL^+$$
$$\text{où } LL^+ = I_{q'}$$

Bien que ne reflétant plus les dynamiques des modèles initiaux, les modèles réduits (14) et (19), qui ne sauraient être considérés comme agrégés, peuvent être utiles pour la détermination de loi de bouclage quasi-optimale | BER,85 |. Nous en verrons une application particulière au cours du chapitre V.

## I.4. METHODES DE PERTURBATION

Basées sur la présence de petits paramètres dans le modèle initial, les méthodes de perturbation consistent à obtenir une solution approchée d'un problème complexe à partir de la solution d'un problème voisin plus simple. Les modèles réduits sont obtenus par passage à la limite ; couramment utilisées en Mathématiques | KAT,76 ; REI,82 ; ROU,73 |, leur emploi a été développé dans d'autres

domaines, comme par exemple, la Mécanique Celeste | ROT,79 |, mais surtout pour l'Analyse des Systèmes, à paramètres localisés | SAN,78 ; BER,82 | ou répartis | LIO,73 ; LIO,78 |. Dans ces applications, deux classes de perturbations apparaissent | SAN,78 | :

* les perturbations régulières correspondant à des couplages faibles de sous-systèmes ;

* les perturbations singulières correspondant à des couplages forts entre sous-systèmes. Ces dernières sont utilisées lorsque, dans un système, coexistent des variables de dynamiques très différentes.

### I.4.1. Perturbations régulières

Dans le cas général d'un système non linéaire, les équations de couplage faible sont de type :

$$(20)\qquad \begin{aligned} \dot{x}_1 &= f(x_1, \varepsilon x_2, u) \\ \dot{x}_2 &= g(\varepsilon x_1, x_2, u) \end{aligned}$$

où $x_1 \in \mathbb{R}^{q_1}$, $x_2 \in \mathbb{R}^{q_2}$, $u \in \mathbb{R}^m$ et $\varepsilon$ est un petit paramètre : $0 < \varepsilon << 1$

ce qui, lorsque $\varepsilon \longrightarrow 0$, fournit deux systèmes découplés

$$(21)\qquad \begin{aligned} \dot{\bar{x}}_1 &= f(\bar{x}_1, o, u) \\ \dot{\bar{x}}_2 &= g(o, \bar{x}_2, u) \end{aligned}$$

L'approximation du système global ne conduit pas à un système d'ordre inférieur mais à deux sous-systèmes complètement indépendants. Dans le cas des systèmes linéaires, on considère des équations de type :

$$(22)\qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} u$$

La question de déterminer alors une loi de bouclage optimale dans un problème linéaire quadratique peut être résolue de façon itérative | KOK, 72 ; TIT, 79 | : l'équation de Riccati associée peut être écrite en puissances de $\varepsilon$, on peut alors connaître les premiers termes de la décomposition analytique, autour de $\varepsilon = 0$, de la solution. Cette méthode permet, également, dans le cas de systèmes non linéaires, de résoudre le problème de la commande optimale | BEN,82 |.

### I.4.2. Perturbations singulières

On modélise un système singulièrement perturbé sous la forme suivante :

$$(23)\quad \begin{cases} (a) & \overset{\circ}{x} = f(x, z, u, \varepsilon, t) \quad x \in \mathbb{R}^{q_1},\ x(t_o) = x_o \\ (b) & \varepsilon\overset{\circ}{z} = g(x, z, u, \varepsilon, t) \quad z \in \mathbb{R}^{q_2},\ z(t_o) = z_o \end{cases}$$

où $\varepsilon$ est un petit paramètre : $0 < \varepsilon << 1$

Dans (23), les composantes de x sont dites variables-lentes et celles de z sont appelées rapides ; d'autre part, la perturbation est singulière car $g^{(\cdot)}/\varepsilon$ n'est pas définie pour $\varepsilon = 0$. Les problèmes et les applications de ce genre de modélisation ont donné lieu à une importante littérature | KOK,76 ; SAK,84 | dans de nombreux domaines de l'Automatique. L'un des problèmes les plus délicats est la mise en évidence du coefficient $\varepsilon$. Il n'est, en effet, pas défini a priori mais dépend, pour le système étudié, de la représentation adoptée :

* il peut correspondre au rapport entre les échelles de temps de la partie lente (t) et de la partie rapide ($\tau$) $\varepsilon = \frac{t - to}{\tau}$ où to est l'instant initial que nous prendrons dorénavant égal à 0.

* il peut avoir une signification physique comme le rapport de deux masses, deux constantes de temps électriques, deux constantes de raideur,...

Dans le cas d'un système linéaire caractérisé par l'équation d'état :

$$(24) \quad \begin{cases} \dot{x} = Ax + Bu & x \in \mathbb{R}^q,\ u \in \mathbb{R}^m \\ y = Cx & y \in \mathbb{R}^n \end{cases}$$

on peut définir la notion de double échelle de temps, donc $\varepsilon$, par le rapport des valeurs propres extrêmes de chaque partie :

$$\varepsilon = \frac{\left|\lambda_{\max} \text{ (partie lente)}\right|}{\left|\lambda_{\min} \text{ (partie rapide)}\right|}$$

Le système (24) possède la propriété de double échelle des temps si | CHO,76 ; DAU,85 | :

* il peut être décomposé en deux sous-systèmes :

$$(25) \quad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_r \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_r \end{bmatrix} u$$

$$y = \begin{bmatrix} C_1 & C_r \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix}$$

$$x_1 \in \mathbb{R}^{q_1},\ x_r \in \mathbb{R}^{q_2},\ q_1 + q_2 = q$$

$$(26) \quad \left|\lambda_{\max}(A_1)\right| \ll \left|\lambda_{\min}(A_r)\right|$$

Remarques : * la condition (26) peut être remplacée par la condition équivalente

$$\| A_r^{-1} \| \ll \| A_1 \|^{-1}$$

* la notion de double échelle de temps a été étendue sous une forme équivalente aux systèmes linéaires non stationnaires | O'MA,82| et à certaines classes de systèmes non linéaires | KOK,80 |.

L'utilisation des méthodes de perturbations conduit donc à une simplification de structures par découplage. Dans le cas des perturbations singulières, un des sous-systèmes est associé à des variables à transitoire rapide, ce qui permet, si la partie rapide est stable, d'approcher le modèle initial au bout d'un certain temps, par l'ensemble des variables lentes : il y a donc réduction de dimensionnalité. Nous détaillons plus particulièrement cette technique dans la prochaine partie.

## II - QUELQUES RESULTATS FONDAMENTAUX CONCERNANT LES SYSTEMES SINGULIEREMENT PERTURBES

Des articles de synthèse | KOK,76 ; SAK,84 ; FOS,86 | ont mis en évidence l'importance de la méthode des perturbations singulières dans l'analyse et la synthèse de systèmes complexes. On peut citer en particulier les applications sur la stabilité | GRU, 79a, 79b, 81 ; KHA,79 |, le calcul de commandes optimales | KOK,83 |, les systèmes à paramètres distribués | VAN,84 |, les systèmes discrets | ELM,85 ; DAU,82a |, les systèmes stochastiques | BEN,81 |. Les travaux cités ici ne sauraient être utilisés que comme des points de repères : la bibliographie que l'on peut trouver dans | SAK,84 | comporte, en effet, pour les seules années 1976-1983, plus de 350 références.

### II.1. CAS GENERAL DES SYSTEMES NON LINEAIRES

#### II.1.1. Réduction d'ordre et approximation | KOK,76 |

Considérons le système singulièrement perturbé (23) ; cette forme est appelée explicite | PEP,82 ; CHO,86 | ; elle est dite sous forme implicite si (23 - a) n'existe pas.

En posant dans (23) $\varepsilon = 0$, (23 - b) se réduit à une équation algébrique :

(27) $\quad 0 = g(\bar{x}, \bar{z}, \bar{u}, 0, t)$

Si cette équation admet une racine de la forme :

$$\bar{z} = h(\bar{x}, \bar{u}, t), \tag{28}$$

la substitution de (28) dans (23 - a) conduit au modèle simplifié caractérisant les variables lentes

$$\begin{cases} a) & \overset{\circ}{\bar{x}} = f(\bar{x}, h(\bar{x}, \bar{u}, t), \bar{u}, 0, t) = \bar{f}(\bar{x}, \bar{u}, t) \\ & \text{avec } \bar{x}(0) = x_0 \\ b) & \bar{z} = h(\bar{x}, \bar{u}, t) \end{cases} \tag{29}$$

Le modèle (29) est appelé modèle lent, dans lequel la partie rapide de z a disparu. A t = 0, $\bar{z}$ est imposé à priori par la condition initiale sur x, donc $\bar{z}$ ne peut pas être une approximation uniforme de z, alors que l'on peut approcher x par $\bar{x}$ sur toute l'évolution du système par

$$x(t,\varepsilon) = \bar{x}(t) + 0(\varepsilon) \tag{30}$$

Pour obtenir la composante rapide de z, un changement de l'échelle des temps $\tau = \frac{t}{\varepsilon}$ conduit à :

$$\begin{cases} a) & \frac{dx}{d\tau} = \varepsilon f(x, z, u, \varepsilon, \varepsilon\tau) \\ b) & \frac{dz}{d\tau} = g(x, z, u, \varepsilon, \varepsilon\tau) \end{cases} \tag{31}$$

ce qui, en posant de nouveau $\varepsilon = 0$, permet de dire que la partie rapide de x est constante ; or $\bar{x}(0) = x_0$. Cela valide l'approximation (30) pour x. L'équation (31 - b) donne dans les mêmes conditions :

$$\frac{dz_r}{d\tau} = g(\bar{x}, z_r, \bar{u}, 0, 0) \tag{32}$$

$$\text{avec } z_r(0) = z(0) - \bar{z}(0)$$

Ce modèle réduit caractérisant l'évolution des régimes transitoires des composantes rapides est appelé système couche-limite et l'un des problèmes majeurs de la théorie des perturbations

singulières est le raccordement de la solution de (32) définie à l'intérieur de la couche limite et la solution de (29 - b) entraînée par les variables lentes à l'extérieur | MOR,84 |. Certains auteurs (SAK,84) substituent dans (32) $x_0$ à $\bar{x}$ et $u_0$ à $\bar{u}$ ; mais nous verrons, dans le cas des systèmes linéaires, que ce point de vue permet une approximation moins fine de la partie rapide car elle est alors indépendante du comportement lent du système. Une autre approche adoptée principalement dans les problèmes de commande | FOS,83 | est de faire l'hypothèse que le modèle lent est excité par les composantes lentes, $u_l$, de la commande, et que le système couche-limite subit les composantes rapides, $u_r$, de la commande en ayant, bien sûr, $u = u_l + u_r$.

L'approximation du modèle initial par les modèles lents et couche limite est fondée sur le théorème de Tihonov | TIH,52 | :

<u>Théorème</u> : si dans (23) f(.) et g(.) sont deux fois différentiables en x, z et t,

si $z_r = 0$ est un équilibre asymptotiquement stable de (32) et $z_r(0)$ appartient à son domaine d'attraction,

si les valeurs propres de $\frac{\partial g}{\partial z}$, pour $\bar{x}$, $\bar{z}$ et $\bar{u}$, pour tout t dans $|0,T|$, ont des parties réelles strictement plus petites qu'un nombre fixé négatif,

alors les approximations :

$$(33)\quad \begin{cases} a) \quad x(t,\varepsilon) = \bar{x}(t) + \sigma(\varepsilon) \\ b) \quad z(t,\varepsilon) = \bar{z}(t) + z_r(\frac{t}{\varepsilon}) + \sigma(\varepsilon) \end{cases}$$

sont valides pour tout t dans $[0,T]$

Un autre choix, dans les conditions d'approximation, serait d'imposer dans (32), la condition initiale : $z_r(0) = z(0)$ ; l'équation (33 - b) serait alors remplacée par

$$(34)\quad z(t,\varepsilon) = \bar{z}(t) + z_r(\frac{t}{\varepsilon}) - \bar{z}(0) + \sigma(\varepsilon)$$

### II.1.2. Développement asymptotique

Lorsque l'approximation à l'ordre 0 est insuffisante, dans le cas où le paramètre n'est pas "très petit", une méthode habituelle | O'MA,74 ; HOP,74 ; ARD,82 | consiste à déterminer les premiers termes des expressions analytiques en $\varepsilon$ de x et z, solutions de (23). En supposant, dans (23), f(.) et g(.) analytiques en $\varepsilon$ :

$$(35)\qquad \begin{aligned} f(.) &= \sum_{i=0}^{\infty} f_i(x, z, u, t)\varepsilon^i \\ g(.) &= \sum_{i=0}^{\infty} g_i(x, z, u, t)\varepsilon^i \end{aligned}$$

on détermine par identification les expressions analytiques en $\varepsilon$ de x et z sous la forme :

$$(36)\qquad \begin{cases} x = x_i(\tau) \text{ puis } x_e(t) \\ z = z_i(\tau) \text{ puis } z_e(t) \end{cases}$$

avec

$$a)\begin{cases} x_e(t) = \sum_{k=0}^{\infty} \varepsilon^k x_{ek}(t) \\ z_e(t) = \sum_{k=0}^{\infty} \varepsilon^k z_{ek}(t) \end{cases}$$

$$b)\begin{cases} x_i(\tau) = \sum_{k=0}^{\infty} \varepsilon^k x_{ik}(\tau) \\ z_i(\tau) = \sum_{k=0}^{\infty} \varepsilon^k z_{ik}(\tau) \end{cases}$$

où $x_e$ (t) et $z_e$ (t) représentent les solutions extérieures (à la couche limite) et $x_i$ ($\tau$) et $z_i$ ($\tau$) les solutions intérieures. Ces solutions sont obtenues en considérant, soit le système intérieur (31) partant des conditions initiales $x_0$ et $z_0$, soit le système extérieur (23) dont les conditions initiales sont fixées de façon à assurer la continuité des solutions à la fin de la couche limite.

Concernant le développement asymptotique de f(.) et g(.), on a :

* pour la solution interne :

$$(37\text{ - a})\begin{cases} f(x_e, z_e, u, t, \varepsilon) = \sum_{k=0}^{\infty} f_{ek}\,\varepsilon^k \\ g(x_e, z_e, u, t, \varepsilon) = \sum_{k=0}^{\infty} g_{ek}\,\varepsilon^k \end{cases}$$

* pour la solution externe :

$$(37\text{ - b})\begin{cases} f(x_i, z_i, u, \varepsilon\tau, \varepsilon) = \sum_{k=0}^{\infty} f_{ik}\,\varepsilon^k \\ g(x_i, z_i, u, \varepsilon\tau, \varepsilon) = \sum_{k=0}^{\infty} g_{ik}\,\varepsilon^k \end{cases}$$

où pour $k \neq 0, f_{ek}, g_{ek}$ (resp $f_{ik}, g_{ik}$) sont des fonctions linéaires en les arguments $x_{ek}, z_{ek}$ (resp $x_{ik} z_{ik}$) | ARD,82|. L'identification termes à termes dans les systèmes extérieur et intérieur conduit aux systèmes :

$$(38)\quad \text{a)}\begin{cases} \dot{x}_{eo} = f_{eo} \;;\; o = g_{eo} \\ \dot{x}_{eo} = f_{ej} \;;\; \dot{z}_{e,j-1} = g_{e,j} \quad (j > 0) \end{cases}$$

$$\text{b)}\begin{cases} \dot{x}_{i,o} = o \;;\; \dot{z}_{i,o} = g_{i,o} \;;\; x_{i,o}(o) = x^o ;\; z_{io}(o) = z^o \\ \dot{x}_{ij} = f_{i,j-1} \;;\; \dot{z}_{i,j} = g_{ij} \;;\; x_{i,j}(o) = o \;;\; z_{i,j}(o) = o \;;(j>0) \end{cases}$$

Chacun de ces problèmes, hormis pour j = 0, est linéaire ; les conditions initiales sur les systèmes donnant la solution extérieure sont fixées par la loi de limite du mélange :

$$(39)\qquad \lim_{\substack{\varepsilon \to 0 \\ t \to 0 \\ \tau \to \infty}} \{ x_e(\varepsilon,t) - x_i(\varepsilon,\tau) \} = 0$$

La détermination d'une solution sur tout l'horizon considéré peut être faite en utilisant en plus un développement intermédiaire | BOR,87|.

### II.1.3. Cas du retour à grand gain

Considérons un système monovariable à commande affine. On sait(I : II.2.4) qu'une très large classe de ces processus peut être représentée, après difféomorphisme, sous la forme globalement commandable :

$$(40)\qquad \begin{bmatrix} \overset{\circ}{x}_1^* \\ \vdots \\ \overset{\circ}{x}_{q-1}^* \\ \overset{\circ}{x}_q^* \end{bmatrix} = \begin{bmatrix} x_2^* \\ \vdots \\ x_q^* \\ f(x^*) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mu(x^*) \end{bmatrix} u(t)$$

où $f(.)$ et $\mu(.)$ sont définis dans (I : II.2.4).

Si on boucle le système à travers un grand gain g avec retour d'état, la commande est de la forme

$$(41)\qquad u(t) = g.k^T x^*$$

$$\text{où } k^T = [k_1,\ldots,k_q]$$

qui entraîne pour le système bouclé :

$$(42)\qquad \begin{bmatrix} \overset{\circ}{x}_1^* \\ \vdots \\ \overset{\circ}{x}_{q-1}^* \\ \overset{\circ}{\varepsilon x}_q^* \end{bmatrix} = \begin{bmatrix} x_2^* \\ \vdots \\ x_q^* \\ \varepsilon f(x^*) + \mu(x^*)\, k^T x^* \end{bmatrix}$$

$$\text{avec } \varepsilon = \frac{1}{g}$$

Pour g grand, (42) est sous la forme singulièrement perturbée. Le modèle lent s'obtient à l'aide de l'équation algébrique :

$$(43)\qquad g = \infty : \mu(\bar{x}^*).k^T\bar{x}^* = 0$$

$$\text{dont la solution est : } \bar{x}_q = -\frac{k_1\bar{x}_1^* + \ldots + k_{q-1}\,\bar{x}_{q-1}^*}{k_q}$$

$$(\text{car } \mu(x^*) \neq 0)$$

Le système lent qui s'écrit alors :

$$(44)\qquad \begin{bmatrix} \dot{\bar{x}}_1^* \\ \vdots \\ \dot{\bar{x}}_{q-1}^* \end{bmatrix} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -\frac{k_1}{k_q} & \cdots\cdots & & -\frac{k_{n-1}}{k_q} \end{bmatrix} \begin{bmatrix} \bar{x}_1^* \\ \vdots \\ \bar{x}_{q-1}^* \end{bmatrix}$$

a la propriété d'être linéaire et de ne dépendre que de la loi de retour d'état (41). Ce point de vue généralise la notion de retour à grand gain pour les systèmes linéaires | YOU,77 | ou les systèmes non linéaires de type Lurie' Postnikov | DAU,83 a |. Les dynamiques du système lent peuvent être fixées arbitrairement par les coefficients $k_i$, et les non linéarités sont toutes reportées sur le système couche limite.

## II.2. CAS DES SYSTEMES LINEAIRES

Dans cette partie, nous allons résumer les résultats sur la modélisation des systèmes linéaires de grande dimension par la méthode des perturbations singulières.

### II.2.1. Modèles réduits lent et couche limite | FOS,82 |

Considérons le système (24) sous la forme singulièrement perturbée :

$$(45)\qquad \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = \left[\begin{array}{c:c} A_{11} & A_{12} \\ \hdashline A_{21} & A_{22} \end{array}\right] \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \,, \qquad \begin{matrix} x \in \mathbb{R}^{q_1} \\ z \in \mathbb{R}^{q_2} \end{matrix}$$

$$y = \left[\begin{array}{c:c} C_1 & C_2 \end{array}\right] \begin{bmatrix} x \\ z \end{bmatrix} \,, \quad x(0) = x_0,\ z(0) = z_0$$

où les normes des matrices $A_{ij}$, $(i,j)$ dans $\{0,1\}\times\{0,1\}$, et $B_i$, $i$ dans $\{0,1\}$ sont supposées du même ordre de grandeur et $\varepsilon$ est un petit paramètre positif proche de 0.

### II.2.1.1. Etude de la partie lente

Pour $\varepsilon = 0$, les variables x, z, u et y sont approchées par $\bar{x}$, $\bar{z}$, $\bar{u}$ et $\bar{y}$ décritespar l'équation d'état

$$(46)\quad \begin{cases} a)\ \overset{\circ}{\bar{x}} = A_{11}\bar{x} + A_{12}\bar{z} + B_1\bar{u} \\ b)\ 0 = A_{21}\bar{x} + A_{22}\bar{z} + B_2\bar{u} \\ c)\ \bar{y} = C_1\bar{x} + C_2\bar{z} \end{cases}$$

L'équation (46-b) décrit le mouvement d'entraînement des composantes lentes de z par les variables lentes $\bar{x}$. Comme le cas où $A_{22}$ admet une valeur propre nulle n'est pas cohérent avec la notion de variable rapide, cette équation se résoud en :

$$(47)\quad \bar{z} = -A_{22}^{-1}(A_{21}\bar{x} + B_2\bar{u})$$

ce qui, replacé dans les équations (46 - a) et (46 - c) fournit le modèle réduit lent :

$$(48)\quad \begin{cases} a)\ \dot{\bar{x}} = (A_{11} - A_{12}A_{22}^{-1}A_{21})\bar{x} + (B_1 - A_{12}A_{22}^{-1}B_2)\bar{u} \\ b)\ \bar{y} = (C_1 - C_2A_{22}^{-1}A_{21})\bar{x} - C_2A_{22}^{-1}B_2\bar{u} \end{cases}$$

avec $\bar{x}(0) = x_0$.

Ce modèle réduit lent, d'ordre $q_1$, qui est une approximation de x durant toute l'évolution du processus, présente une transmission directe entrée-sortie.

### II.2.1.2. Etude de la partie rapide

Pour mettre en évidence le système rapide, il suffit d'isoler les composantes rapides sous la forme :

$$(49)\quad \begin{cases} x_r = x - \bar{x} = 0 \\ z_r = z - \bar{z} \\ u_r = u - \bar{u} \\ y_r = y - \bar{y} \end{cases}$$

ce qui, replacé dans (45) fournit :

$$(50) \qquad \varepsilon \overset{\circ}{z}_r + \varepsilon \overset{\circ}{\bar{z}} = A_{21}\bar{x} + A_{22}\bar{z} + B_2\bar{u} + A_{22}z_r + B_2u_r$$

Or, d'après (46) : $\varepsilon \overset{\circ}{\bar{z}} = A_{21}\bar{x} + A_{22}\bar{z} + B_2\bar{u} = 0$, il vient l'équation de couche limite définissant le transitoire de z :

$$(51) \qquad \begin{cases} \varepsilon \overset{\circ}{z}_r = A_{22}z_r + B_2u_r \\ y_r = C_2z_r \end{cases}$$

avec $z_r(0) = z_0 + A_{22}^{-1}A_{21}x_0$

Le système global apparaît donc découplé temporellement. Mais, si l'utilisation d'une commande composite est naturelle dans le cas d'une commande par retour d'état, elle l'est moins du point de vue de la modélisation. En vue de la simulation, comment séparer a priori une consigne donnée en deux composantes ? Une première solution est d'utiliser une commutation dont le principe serait :

- au début du mouvement, toute l'entrée est sur le système rapide,

- après la couche limite, la commande est envoyée sur le système lent.

L'inconvénient de cette méthode réside dans la difficulté de déterminer avec précision la durée de la couche limite. Une deuxième solution, que nous proposons dans le paragraphe suivant, consiste à découpler les parties lente et rapide : chaque sous-système agissant respectivement comme filtre passe-haut et passe-bas.

### II.2.2. Modélisation par découplage

Soit le système linéaire singulièrement perturbé (45). L'approximation des variables lentes, avec $\varepsilon = 0$, est donnée par (48 -a). Nous allons donc la conserver en prenant pour $\bar{u}$ la commande complète u, cela donne pour une approximation du système complet :

$$(52)\quad \begin{cases} \begin{bmatrix} \overset{\circ}{\bar{x}} \\ \cdots \\ \overset{\circ}{\tilde{z}} \end{bmatrix} = \begin{bmatrix} A_1 & \vdots & 0 \\ \cdots & \vdots & \cdots \\ A_{21}^* & \vdots & A_{22}^* \end{bmatrix} \begin{bmatrix} \bar{x} \\ \cdots \\ \tilde{z} \end{bmatrix} + \begin{bmatrix} B_1 \\ \cdots \\ B_2^* \end{bmatrix} u \\ \tilde{y} = C_1\bar{x} + C_2\tilde{z} \quad , \ \bar{x}(0) = x_0, \ \tilde{z}(0) = z_0 \end{cases}$$

$$\text{avec} \begin{cases} A_1 = A_{11} - A_{12}A_{22}^{-1}A_{21} \text{ et } B_1 = B_1 - A_{12}A_{22}^{-1}B_2 \\ A_{21}^* = A_{21}/\varepsilon \ , \ A_{22}^* = A_{22}/\varepsilon \ , \ B_2^* = B_2/\varepsilon \end{cases}$$

La notation $\tilde{}$ indique que l'on a une approximation des variables initiales en ayant remplacé x par $\bar{x}$. Si l'on fait le changement de variables :

$$(53)\quad \begin{bmatrix} \bar{x} \\ z_r \end{bmatrix} = \begin{bmatrix} Iq_1 & 0 \\ L & Iq_2 \end{bmatrix} \begin{bmatrix} \bar{x} \\ \tilde{z} \end{bmatrix}$$

avec $L \in \mathbb{R}^{q_2 \times q_1}$

(52) est transformé en

$$(54)\quad \begin{cases} \begin{bmatrix} \overset{\circ}{\bar{x}} \\ \overset{\circ}{z_r} \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ R(L) & A_{22}^* \end{bmatrix} \begin{bmatrix} \bar{x} \\ z_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_r^* \end{bmatrix} u \\ \tilde{y} = C_1\bar{x} + C_2 z_r \end{cases}$$

$$\text{avec} \quad \begin{cases} R(L) = LA_1 - A_{22}^*L + A_{21}^* \\ B_r^* = B_2^* + LB_1 \\ C_1 = C_1 - C_2L \\ z_r(0) = z_0 + Lx_0 \end{cases}$$

Le système (52) est complètement découplé si R(L) = 0. Or, cette équation de Lyapounov admet, ici, toujours une racine | GAN,60 | car $A_1$ et $A_{22}^*$ ont des valeurs propres distinctes (sinon il n'y aurait pas de séparation de dynamiques). Si on écrit cette

équation sous la forme vectorielle :

(55) $$A_{22}L - \varepsilon L A_1 = A_{21} \iff$$

$$\left[A_{22} \otimes I_{q_1} - \varepsilon I_{q_2} \otimes A_1\right] \mathrm{vec}(L) = \mathrm{vec}(A_{21})$$

où $\otimes$ représente le produit de Kronecker (I : 48) et vec ( . ) est la fonction vectorielle d'une matrice | NEU,69 | définie par : si A est une p x q matrice,

$$A = \left[A_1, A_2, \ldots, A_q\right] \quad \text{alors} \quad \mathrm{vec}(A) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_q \end{bmatrix}$$

où $\forall\, i \in \{1,\ldots q\}\ \ A_i \in \mathbb{R}^p$

Comme les valeurs propres de A $\otimes$ B sont les produits de valeurs de propres de A par celles de B, la solution de (55) est donnée par :

(56) L telle que $\mathrm{vec}(L) = \left[A_{22} \otimes I_{q_1} - \varepsilon I_{q_2} \otimes A_1\right]^{-1} \mathrm{vec}\,(A_{21})$

Il est à noter que la solution à l'ordre 0 ($\varepsilon = 0$) est donnée par $L_o = - A_{22}^{-1} A_{21}$.

Lorsque L a la valeur donnée par (56), les modèles réduits lent et rapide sont découplés et y est la superposition d'une réponse lente $y_1$ due au passage de u à travers le système lent et d'une réponse rapide $y_2$ due à la traversée du système rapide (Fig.1).

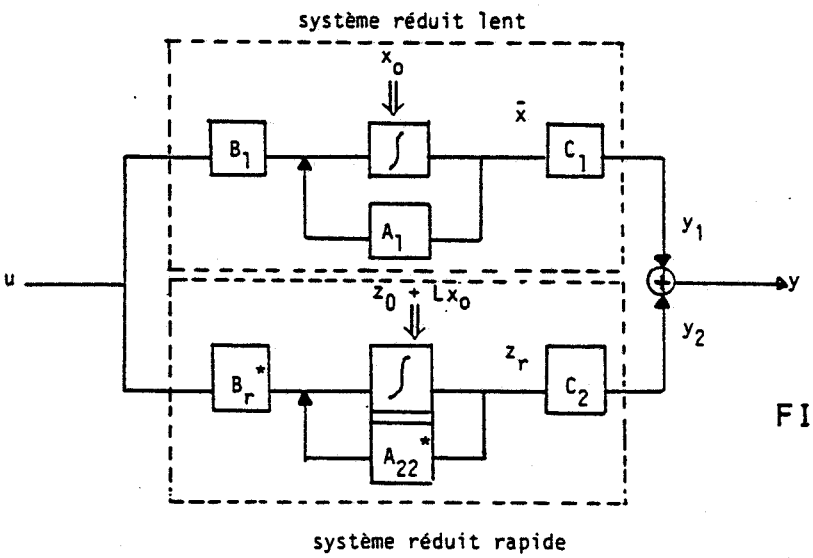



FIGURE 1

Le système réduit rapide ayant, par hypothèse, un transitoire rapide, son régime d'équilibre atteint après quelques instants de fonctionnement est donné par :

$$(57) \qquad y_{21} = - C_2 A_{22}^{-1} (B_2 + \varepsilon L B_1) u$$

ce qui est une expression voisine, à un o ($\varepsilon$) près, de la transmission directe trouvée dans (48 - b). Le système "dégénère" alors en un système d'ordre réduit $q_1$, figure 2,

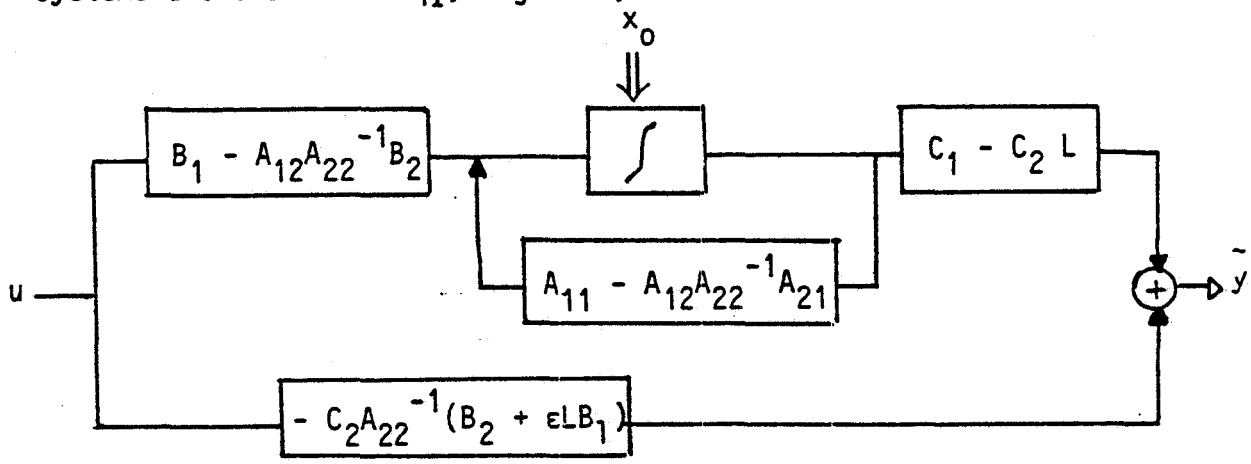



FIGURE 2 système lent dégénéré

qui pour $\varepsilon = 0$, donne les résultats (48). La variable z peut être approchée à partir de (54) et (53) par

$$(58) \qquad \begin{cases} \tilde{z} = z_r - L\bar{x} \\ \text{avec } \bar{z}(0) = z_0 \end{cases}$$

Cette approche supprime le problème de la commutation dans la modélisation et rejoint sensiblement les résultats de l'analyse fréquentielle présentés dans | FOS,83 |. Nous en verrons, au chapitre V, une application dans la détermination de régulateurs quasi-optimaux.

### II.2.3. Utilisation du développement asymptotique

Nous allons montrer que l'utilisation de la méthode du développement asymptotique, afin d'améliorer la précision du système rapide, permet d'écrire le système initial comme un ensemble de systèmes mis en parallèle.

Mettons le système (45), après changement de l'échelle des temps, $\tau = \frac{t}{\varepsilon}$, sous la forme :

$$(59)\quad \begin{cases} x' = \varepsilon A_{11}x + \varepsilon A_{12}z + \varepsilon B_1 u & ; x(o) = x_o \\ z' = A_{21}x + A_{22}z + B_2 u & ; z(o) = z_o \\ y = C_1 x + C_2 z \end{cases}$$

où $x' = \frac{dx}{d\tau}$ et $z' = \frac{dz}{d\tau}$

L'écriture analytique de x et z sous la forme :

$$(60)\quad \begin{cases} x = \sum_{i=0}^{\infty} \varepsilon^i x_{(i)} \\ z = \sum_{i=0}^{\infty} \varepsilon^i z_{(i)} \end{cases}$$

conduit à décomposer (59) en sous-systèmes définis par :

$$(61)\quad \begin{cases} (\Sigma_o) & \begin{cases} x'(o) = 0 \\ z'(o) = A_{21}x(o) + A_{22}z(o) + B_2 u \end{cases} \\ (\Sigma_1) & \begin{cases} x'(1) = A_{11}x(o) + A_{12}z(o) + B_1 u \\ z'(1) = A_{21}x(1) + A_{22}z(1) \end{cases} \\ (\Sigma_i)\ \ i \geqslant 2 & \begin{cases} x'(i) = A_{11}x(i-1) + A_{12}z(i-1) \\ z'(i) = A_{21}x(i) + A_{22}z(i) \end{cases} \end{cases}$$

et organisés en cascade suivant le schéma figure 3.

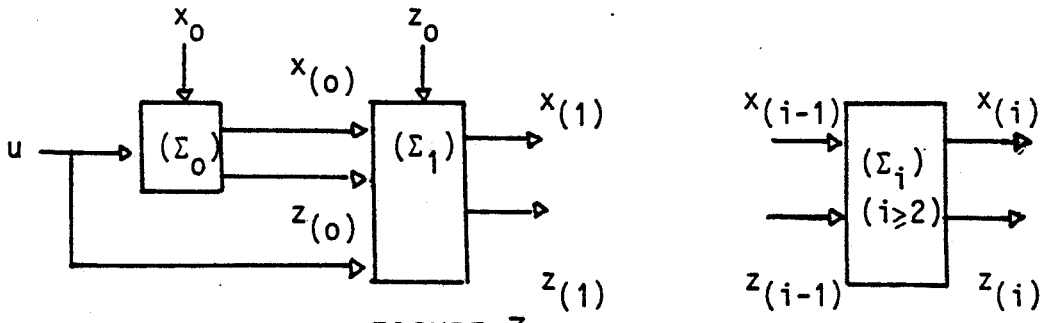



FIGURE 3

D'après (60), la sortie du système s'écrivant :

$$(62)\quad y = \sum_{i=0}^{\infty} \varepsilon^i y_{(i)} \quad \text{où } \forall i \in \mathbb{N}\ \ y_{(i)} = C_1 x_{(i)} + C_2 z_{(i)}$$

La matrice de transfert F(p) de (59) se mettra, en conservant la notation usuelle de l'opérateur de dérivation, sous la forme :

$$(63) \qquad F(p) = \sum_{i=0}^{\infty} F_i(p)\varepsilon^i$$

où $F_i(p)$ est définie par $Y_i(p) = F_i(p)\, U(p)$.

En supposant les conditions initiales nulles de (61), on tire les relations de transfert :

$$(64) \quad \begin{cases} (\Sigma_0) \begin{bmatrix} x_{(0)}(p) \\ z_{(0)}(p) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ (pI_{q2} - A_{22})^{-1} B_2 \end{bmatrix}}_{G_0(p)} U(p) \\ (\Sigma_1) \begin{bmatrix} x_{(1)}(p) \\ z_{(1)}(p) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{p}\left[B_1 + A_{12}(pI_{q2} - A_{22})^{-1} B_2\right] \\ \frac{1}{p}(pI_{q2} - A_{22})^{-1} A_{21}\left[B_1 + A_{12}(pI_{q2}-A_{22})^{-1}B_2\right] \end{bmatrix}}_{G_1(p)} U(p) \\ (\Sigma_i) \begin{bmatrix} x_{(i)}(p) \\ z_{(i)}(p) \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{A_{11}}{p} & \frac{A_{12}}{p} \\ (pI_{q2}-A_{22})^{-1}\frac{A_{21}A_{11}}{p} & (pI_{q2}-A_{22})^{-1}\frac{A_{21}A_{12}}{p} \end{bmatrix}}_{M(p)} \begin{bmatrix} x_{(i-1)}(p) \\ z_{(i-1)}(p) \end{bmatrix} \end{cases}$$

Ainsi ,par récurrence, on obtient :

$$(65) \qquad \forall i \geqslant 1 \quad \begin{bmatrix} x_{(i)}(p) \\ z_{(i)}(p) \end{bmatrix} = M^{i-1}(p)\, G_1(p)\, U(p)$$

Compte tenu de ces expressions, il vient pour la fonction de transfert globale, en notant $[\, C_1 \mid C_2 \,]$ par C :

$$(66) \qquad F(p) = CG_0(p) + \varepsilon C \underbrace{\left[\sum_{i=0}^{\infty} \varepsilon^i M^i(p)\right]}_{\left[I_q - \varepsilon M(p)\right]^{-1}} G_1(p)$$

et on peut, suivant la figure 3, modéliser le régime transitoire du système. De façon à étudier le comportement asymptotique de cette modélisation, faisons p ——>0, ce qui fournit :

$$\begin{cases} (67a) \quad \left[I_q - \varepsilon M(p)\right]_{p=0} \simeq \left\{ I_q - \varepsilon/p \begin{bmatrix} I_{q_1} \\ -A_{22}^{-1}A_{21} \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \right\} \\ (67b) \quad G_1(p)_{p=0} \simeq \frac{1}{p} \begin{bmatrix} I_{q_1} \\ -A_{22}^{-1}A_{21} \end{bmatrix} B_1 \end{cases}$$

avec $B_1 = B_1 - A_{12} A_{22}^{-1} B_2$

ainsi $\left[I_q' - \varepsilon M(p)\right]^{-1}$ peut être approché par

$$(68) \quad \left[I_q - \varepsilon M(p)\right]^{-1} \simeq I_q + \varepsilon \begin{bmatrix} I_{q1} \\ -A_{22}^{-1}A_{21} \end{bmatrix} (pI_{q1} - \varepsilon A_1)^{-1} \begin{bmatrix} A_{11} & A_{12} \end{bmatrix}$$

où $A_1 = A_{11} - A_{12} A_{22}^{-1} A_{21}$

Par cette méthode, la fonction de transfert du système initial peut être approchée autour de p = 0 par une somme de deux fonctions de transfert définies par :

$$(69) \quad F(p) = F_1(p) + \varepsilon F_2(p)$$

avec $F_1(p) = C_2(pI_{q_2} - A_{22})^{-1} B_2$

$F_2(p) = C_1(pI_{q_1} - \varepsilon A_1)^{-1} B_1$

où $C_1 = C_1 - C_2 A_{22}^{-1} A_{21}$

La décomposition analytique de F(p) est indiquée dans | SIL,84 |, où l'utilisation de la théorie des perturbations des opérateurs linéaires conduit à de lourds calculs.

### II.2.4 Utilisation du système réciproque

Si le but de l'étude concerne les transitoires rapides ou le comportement en hautes fréquences du système, la précision obtenue par résolution du système (51) peut s'avérer insuffisante. En effet, l'influence de la partie lente sur les dynamiques de la partie rapide

n'apparaît pas, même dans la solution du découplage (II.2.2). L'emploi simultané de la transformation réciproque et de la méthode des perturbations singulières améliore le résultat | DAU,83a |. Elle donne au système rapide réduit une précision semblable à celle obtenue pour le système lent découplé.

### II.2.4.1. Transformation réciproque

La transformation réciproque utilisée dans | DAU,82 ; DAU,83a,83b | a été introduite dans | HUT,75 | ; elle est définie par (3) et a la propriété d'inverser les dynamiques d'un signal. En effet, F(p) est la transformée de Laplace de la réponse impulsionnelle du système linéaire ; donc, la transformation réciproque peut être définie comme une application de l'espace des fonctions du temps sur lui-même.

$$\mathcal{R} : s(t) \xrightarrow{\mathcal{L}} S(p) \xrightarrow{R} \tilde{S}(p) \xrightarrow{\mathcal{L}^{-1}} \tilde{S}(t)$$

où $\mathcal{L}$ est la transformée de Laplace (d'inverse $\mathcal{L}^{-1}$)

R la transformation réciproque qui à S(p) fait correspondre $\tilde{S}(p) = \frac{1}{p} S(\frac{1}{p})$.

Si on considère le système linéaire présentant une chaîne de transmission directe :

$$(70) \quad \begin{cases} \dot{x} = Ax + Bu & x(0) = x_o \\ y = Cx + Du & x \in \mathbb{R}^q,\ u \in \mathbb{R}^m,\ y \in \mathbb{R}^n \end{cases}$$

et $F(p) = C(pI_q - A)^{-1}B + D$ sa fonction de transfert

D'après la définition précédente, la fonction de transfert du système réciproque associé est donnée par

$$(71) \quad \tilde{F}(p) = \frac{1}{p}\left[C(\frac{I_q}{p} - A)^{-1}B + D\right]$$

$$= C(I_q - Ap)^{-1}B + \frac{D}{p}$$

Considérons le cas particulier du système où A est régulière, alors il existe deux matrices, R, Q, de $\mathbb{R}^n$, régulières, telles que :

$$(72) \quad A = RQ$$

alors

$$(73) \qquad \tilde{F}(p) = - CQ^{-1} (pI_q - R^{-1}Q^{-1})^{-1} R^{-1}B + \frac{D}{p}$$
$$= - CA^{-1}R(pI_q - R^{-1}A^{-1}R)^{-1}R^{-1}B + \frac{D}{p}$$

Le système réciproque est décrit par l'équation d'état (à un changement de variable près) :

$$(74) \qquad \begin{cases} \overset{\circ}{w} = \tilde{A}w + \tilde{B}v \\ \tilde{y} = \tilde{C}w + D \int_0^t v \, dt \end{cases}$$

avec $\tilde{A} = R^{-1}A^{-1}R = R^{-1}Q^{-1}$
$\tilde{B} = R^{-1}B$
$\tilde{C} = - CA^{-1}R = - CQ^{-1}$

Or, F(p) est définie par : Y(p) = F(p) U(p) ; le réciproque du signal de sortie de (70) est donc défini par :

$$(75) \qquad \tilde{Y}(p) = \frac{1}{p} F(\frac{1}{p}) U(\frac{1}{p}) = \tilde{F}(p) \, p\tilde{U}(p),$$

la signification des variables introduites dans (74) apparaît alors comme $\tilde{y} = \mathcal{R}(y(t))$ et $v = \frac{d}{dt} \mathcal{R}(u(t))$

Concernant les conditions initiales à imposer au système réciproque, d'après la fonction de transfert de (70) en régime libre, on obtient celle de (74) en régime libre :

$$\begin{cases} (76a) \qquad \tilde{F}_0(p) = - CA^{-1}R(pI_q - R^{-1}Q^{-1}R)^{-1}R^{-1} \, x_o \\ (76b) \qquad \tilde{F}_0(p) = \tilde{C}(pI_n - \tilde{A})^{-1} w_o \end{cases}$$

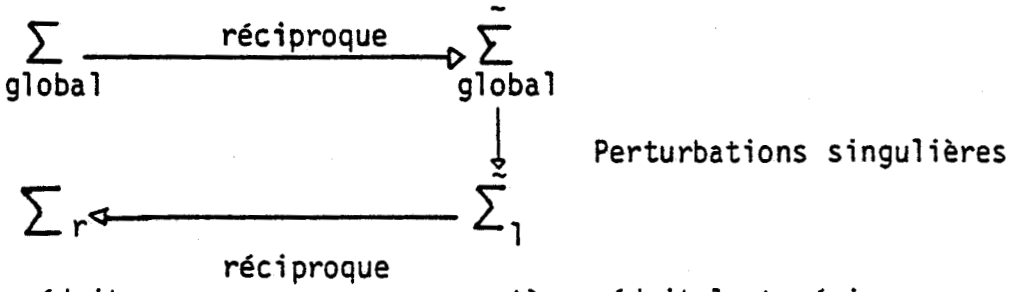
Ainsi, par identification, on a :

$$(77) \qquad w_o = R^{-1}x_o$$

Il est à noter que la décomposition de (A) par (72) intervient comme un changement de variable au niveau de la définition du système réciproque (74) et permet de regrouper en une formulation unique les deux cas distingués dans | DAU,83a |.

### II.2.4.2. Transformation réciproque etperturbations singulières

La transformation réciproque inverse les dynamiques : la démarche proposée dans | DAU 83a | pour déterminer plus précisément les variables rapides est la suivante :




Dans le cas d'un système singulièrement perturbé, les matrices intervenant dans (67) ont la forme (45) :

$$(78)\quad \begin{cases} A = I_\varepsilon \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} ; \quad B = I_\varepsilon \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} ; \; C = \begin{bmatrix} C_1, C_2 \end{bmatrix} ; \; D = 0 \\ \text{où } I_\varepsilon = \begin{bmatrix} I_{q_1} & 0 \\ 0 & I_{q_2}/\varepsilon \end{bmatrix} \end{cases}$$

où la décomposition (72) apparaît de façon naturelle. Mais, plusieurs choix sont possibles ; définissons tout d'abord, en supposant $A_{22}$ régulière :

$$(79)\quad \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{A_{11}} & \overline{A_{12}} \\ \overline{A_{21}} & \overline{A_{22}} \end{bmatrix}$$

avec, si $\Delta = A_{11} - A_{12} A_{22}^{-1} A_{21}$,

$$\overline{A_{11}} = \Delta^{-1}$$
$$\overline{A_{21}} = - A_{22}^{-1} A_{21} \Delta^{-1}$$
$$\overline{A_{12}} = -\Delta^{-1} A_{12} A_{22}^{-1}$$
$$\overline{A_{22}} = A_{22}^{-1}(I + A_{21}\Delta^{-1} A_{12} A_{22}^{-1})$$

Posons $R = I_\varepsilon$, il vient le système réciproque

$$(80)\qquad \begin{cases} \begin{bmatrix} \overset{\circ}{w}_1 \\ \overset{\circ}{w}_2 \end{bmatrix} = \begin{bmatrix} \overline{A_{11}} & \overline{A_{12}} \\ \varepsilon\overline{A_{21}} & \varepsilon\overline{A_{22}} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} v \\ \tilde{y} = -\left[ C_1\overline{A_{11}} + C_2\overline{A_{21}} \;;\; C_1\overline{A_{12}} + C_2\overline{A_{22}} \right] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \end{cases}$$

$$\text{avec} \begin{cases} w_1(0) = x_0 \\ w_2(0) = \varepsilon z_0 \end{cases}$$

De façon à normaliser l'équation d'état (80), introduisons un petit coefficient $\mu$, $0 < \mu \ll 1$, et réécrivons la première ligne de (80) sous la forme :

$$(81)\qquad \mu\overset{\circ}{w}_1 = \overline{A_{11}}^{*} w_1 + \overline{A_{12}}^{*} w_2 + B_1^{*} v$$

où $\overline{A_{11}}^{*} = \mu\overline{A_{11}}$, $\overline{A_{12}}^{*} = \mu\overline{A_{12}}$ et $B_1^{*} = \mu B_1$

sont des matrices dont les normes sont du même ordre de grandeur que les normes des matrices $\varepsilon\overline{A_{21}}$ et $\varepsilon\overline{A_{22}}$.

Pour $\mu = 0$, l'équation algébrique donnée par (81) fournit :

$$(82)\qquad \overline{w_1} = -\overline{A_{11}}^{-1}(\overline{A_{12}}\overline{w_2} + B_1 v)$$

ce qui replacé dans (80) donne

$$(83)\qquad \overset{\circ}{\overline{w_2}} = \varepsilon\left[\overline{A_{22}} - \overline{A_{21}A_{11}}^{-1}\,\overline{A_{12}}\right]\overline{w_2} + (B_2 - \varepsilon\overline{A_{21}A_{11}}^{-1} B_1)v$$

Cette expression, tous calculs faits, donne :

$$(84)\qquad \overset{\circ}{\overline{w_2}} = \varepsilon A_{22}^{-1}\,\overline{w_2} + (B_2 + \varepsilon A_{22}^{-1}A_{21}B_1)v \quad \text{avec } \overline{w_2}(0) = \varepsilon z_0$$

De façon à retrouver la même condition initiale ($z_0$) sur le système initial et sur le système réduit, on est conduit au système réciproque de (84) avec $R = \varepsilon I_{q_2}$ ; on obtient alors :

$$(85)\qquad \varepsilon\overset{\circ}{z}_{\tilde{r}} = A_{22}z_{\tilde{r}} + (B_2 + \varepsilon A_{22}^{-1}A_{21}B_1)u \text{ avec } z_{\tilde{r}}(0) = z_0$$

On retrouve ainsi les résultats exposés dans | DAU,83a |. Les vecteurs de commande et d'observation ont été modifiés (résultats analogues à ceux obtenus par la méthode de découplage pour la partie rapide), mais bien que les dynamiques de la partie rapide soient indépendantes de celles de la partie lente, cela donne une meilleure approximation de la partie rapide de z que (51). La sortie rapide du système est donnée par | DAU,83a | :

$$(86) \qquad y_{\tilde{r}} = C_2 x_{\tilde{r}} + \int_0^t (C_1 - C_2 A_{22}^{-1} A_{21}) B_1 u dt$$

Ce point de vue a permis de résoudre les problèmes de commandes optimales singulières | DAU 83a, DAU 85 |.

### II.2.4.3. Utilisation du système réciproque et du découplage

De façon à éliminer le problème de l'intégration de l'entrée dans l'expression finale (86), on peut utiliser sur le système réciproque la méthode de découplage (II.2.2.). Considérons l'approximation de $w_2$ par (84), alors une approximation de $w_1$ est donnée par :

$$(87) \qquad \begin{bmatrix} \overset{o}{\overline{\overline{w_1}}} \\ \overset{o}{\overline{w_2}} \end{bmatrix} = \begin{bmatrix} \overline{A_{11}} & \overline{A_{12}} \\ 0 & \varepsilon A_{22}^{-1} \end{bmatrix} \begin{bmatrix} \overline{\overline{w_1}} \\ \overline{w_2} \end{bmatrix} + \begin{bmatrix} B_1 \\ \overline{B_2} \end{bmatrix} v$$

avec $\overline{B_2} = B_2 + \varepsilon A_{22}^{-1} A_{21} B_1$

$\overline{\overline{w_1}}(0) = x_o$ ; $\overline{w_2}(0) = \varepsilon z_o$

En utilisant le changement de variable :

$$(88) \qquad \begin{bmatrix} \overline{w_r} \\ \overline{w_2} \end{bmatrix} = \begin{bmatrix} I & \bar{M} \\ 0 & I \end{bmatrix} \begin{bmatrix} \overline{\overline{w_1}} \\ \overline{w_2} \end{bmatrix}$$

où $\hat{M}$ est solution de l'équation de Lyapounov:

$$\overline{A_{12}} + \varepsilon M A_{22}^{-1} - \overline{A_{11}} M = 0$$

on obtient le système découplé réciproque

$$(89) \qquad \begin{bmatrix} \overline{w_r}^{\circ} \\ \overline{w_2}^{\circ} \end{bmatrix} = \begin{bmatrix} \overline{A_{11}} & 0 \\ 0 & \varepsilon A_{22}^{-1} \end{bmatrix} \begin{bmatrix} \overline{w_r} \\ \overline{w_2} \end{bmatrix} + \begin{bmatrix} \overline{B_1} \\ \overline{B_2} \end{bmatrix} v$$

$$\text{avec } \overline{B_1} = B_1 + \hat{M} B_2$$

$$\overline{w_r}(0) = x_o + \varepsilon \hat{M} z_o$$

$$\text{et } \bar{y} = -(\bar{C}_r \bar{w}_r + \bar{C}_2 \bar{w}_2) = \bar{y}_r + \bar{y}_2$$

$$\text{avec } \bar{C}_r = C_1 \overline{A_{11}} + C_2 \overline{A_{21}}$$

$$\bar{C}_2 = C_2(\overline{A_{22}} - \overline{A_{21}}\hat{M}) - \varepsilon C_1 \hat{M} A_{22}^{-1}$$

Si l'on prend le réciproque de chacun de ces deux sous-systèmes, on aboutit à :

$$(90) \qquad \begin{aligned} \tilde{x}_1^{\circ} &= (A_{11} - A_{12} A_{22}^{-1} A_{21})\tilde{x}_1 + \overline{B_1} u \\ \varepsilon \tilde{x}_2^{\circ} &= A_{22} \tilde{x}_2 + \overline{B_2} u \end{aligned}$$

$$\text{avec } \tilde{x}_1(0) = x_o + \varepsilon \hat{M} z_o$$

$$\tilde{x}_2(0) = z_o$$

$$s_1 = \left[ C_1 - C_2 A_{22}^{-1} A_{21} \right] \tilde{x}_1$$

$$s_2 = \left[ C_2(I + \varepsilon A_{22}^{-1} A_{21} \hat{M}) - \varepsilon C_1 \hat{M} \right] \tilde{x}_2$$

$$s = s_1 + s_2$$

Ceci nous fournit donc un système découplé modélisant le système pendant la période de couche limite, plus précisément que (54). Le tableau suivant permet de comparer les approximations à l'ordre 0 des systèmes obtenus.

| | SOUS-SYSTEME LENT | SOUS-SYSTEME RAPIDE | |
|---|---|---|---|
| INTERIEUR COUCHE LIMITE | $B_{l_i} = B_1 + \hat{M}B_2$<br>$C_{l_i} = C_1 - C_2A_{22}^{-1}A_{21}$ | $B_{r_i} = (B_2 + \varepsilon A_{22}^{-1}A_{21}B_1)/\varepsilon$<br>$C_{r_i} = C_2 - \varepsilon(C_1 - C_2A_{22}^{-1}A_{21})\tilde{M}$ | $\hat{M}$ solution de<br>$\overline{A_{12}} + \varepsilon M A_{22}^{-1} - \overline{A_{11}}M = 0$ |
| EXTERIEUR COUCHE LIMITE | $B_{l_e} = B_1 - A_{12}A_{22}^{-1}B_2$<br>$C_{l_e} = C_1 - C_2L$ | $B_{r_e} = (B_r + \varepsilon L B_{l_e})/\varepsilon$<br>$C_{r_e} = C_2$ | $\hat{L}$ solution de<br>$\varepsilon L A_1 - A_{22}L + A_{21} = 0$ |
| MATRICE D'EVOLUTION | $A_l = A_{11} - A_{12}A_{22}^{-1}A_{21}$ | $A_r = A_{22}/\varepsilon$ | |



BU LILLE

Ces méthodes (découplage + réciprocité) accroissent donc la précision des modèles réduits en faisant intervenir le paramètre $\varepsilon$. On peut remarquer que :

$$\lim_{\varepsilon \longrightarrow 0} \begin{bmatrix} B_{1_i} \\ C_{1_e} \\ C_{r_i} \\ B_{r_i} \end{bmatrix} = \begin{bmatrix} B_{1_e} \\ C_{1_i} \\ C_{r_e} \\ B_{r_e} \end{bmatrix}$$

et l'on retrouve pour $\varepsilon = 0$, les résultats classiques de la méthode des perturbations singulières

### II.2.5. Systèmes singuliers singulièrement perturbés

Les systèmes linéaires singulièrement perturbés tels que (45) sont exprimés sous la forme explicite. Dans le cas où le petit paramètre $\varepsilon$ apparaît sur toutes les dérivées des variables d'état, on a alors une forme non explicite ou un système singulier singulièrement perturbé | PEP,82 ; CHO 86 |. On a alors une équation d'état sous la forme, en régime libre :

$$(91) \qquad \varepsilon \overset{\circ}{w} = \left[ A_0 + \varepsilon A_1(\varepsilon) \right] w, \; w(0) = w_0$$

dans le cas où (91) vérifie la condition :

$$(92) \qquad \dim \operatorname{Ker} A_0 = \nu \geqslant 1 \text{ et } \dim \operatorname{Im} A_0 = p \geqslant 1$$

Il est possible de se ramener à une forme explicite par le théorême suivant :

**Théorème | CHO,86 |** :

soit Q une matrice dont les lignes engendrent Im $(A_0)$

soit P une matrice dont les lignes engendrent Ker$(A_0)$

$$T = \begin{bmatrix} P \\ \hline Q \end{bmatrix} \quad , \quad T^{-1} = \begin{bmatrix} V \\ \hline W \end{bmatrix}$$

le changement de variable $\begin{bmatrix} x \\ \\ z \end{bmatrix} = Tw$

transforme (91) en :

$$\begin{cases} \dot{x} = As(\varepsilon)\, x + Asf(\varepsilon) z \\ \varepsilon \dot{z} = \varepsilon Afs(\varepsilon)\, x + Af(\varepsilon) z \end{cases}$$

avec $As(\varepsilon) = PA_1(\varepsilon)V$

$Asf(\varepsilon) = PA_1(\varepsilon)W$

$Afs(\varepsilon) = QA_1(\varepsilon)V$

$Af(\varepsilon) = QA_0W + \varepsilon QA_1(\varepsilon)W$

Ce résultat a été généralisé dans le cas non linéaire | PEP,82 |. Des formes plus particulières de (91) peuvent être distinguées | CHO,86 | :

* forme séparée rapide

$$(93) \qquad \varepsilon \begin{bmatrix} \dot{x}_1 \\ \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \varepsilon A_{11} & A_{12} \\ \\ \varepsilon A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ \\ x_2 \end{bmatrix}$$

par opposition à la forme explicite ou forme séparée lente

* forme à connection forte

$$(94) \qquad \varepsilon \begin{bmatrix} \dot{x}_1 \\ \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} + \varepsilon A'_{11} & \varepsilon A_{12} \\ \\ \varepsilon A_{21} & A_{22} + \varepsilon A'_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ \\ x_2 \end{bmatrix}$$

Dans ce dernier cas, le changement de variables :

$i = 1,2 \quad y_i = P_i x_i$ et $z_i = Q_i x_i$

où $P_i$ et $Q_i$ engendrent Im $A_{ii}$ et Ker $A_{ii}$,

transforme le système (94) en une forme explicite séparée lente-rapide.

## CONCLUSION

Nous avons donc rappelé les résultats classiques sur la réduction de dimensionnalité dans l'analyse des systèmes. Nous nous sommes longuement étendus sur le cas des systèmes linéaires car les notions utilisées peuvent être appliquées à l'étude des systèmes bilinéaires ; nous avons déjà mentionné ce fait pour les techniques d'agrégation. Nous avons, au passage, développé un certain nombre de nouveaux points de vue comme le découplage lent-rapide ou le développement analytique, en fonction d'un petit paramètre, de la fonction de transfert.

Dans un prochain chapitre, nous développerons plus particulièrement la méthode des perturbations singulières dans le cas des systèmes bilinéaires, en mettant en lumière les problèmes soulevés.

# CHAPITRE III

# DETECTION DE DYNAMIQUES DIFFERENTES

# INTRODUCTION

Pour la réduction de dimensionnalité, l'emploi de la méthode des perturbations singulières est conditionnée par l'existence de petits paramètres intervenant sur les dérivées. Lorsque les systèmes ne sont pas "a priori" sous forme singulièrement perturbée, il est nécessaire de détecter des différences marquées dans les dynamiques des variables d'état.

En ce qui concerne les systèmes linéaires, les dynamiques étant définies par les valeurs propres, de nombreux travaux ont été consacrés à cette question. Nous les rappelerons dans une première partie ; il est possible de les grouper en deux catégories : les méthodes algébriques qui consistent à déterminer un changement de variables mettant en évidence les variables rapides et les variables lentes ; et les méthodes géométriques qui localisent des domaines du plan complexe contenant les valeurs propres.

Dans le cas des systèmes bilinéaires, les valeurs propres de la matrice-système dépendent de façon continue, mais non simple, des variables d'entrée. Pour simplifier le problème, nous serons conduits à utiliser certains des résultats précédents ou d'autres plus particuliers, comme celui utilisant la notion de conditionnement, notion déjà utilisée dans les problèmes d'identification | DEL,86 |.

# I. MISE EN EVIDENCE DES DYNAMIQUES DANS LE CAS DES SYSTEMES LINEAIRES

Un système linéaire possède la propriété de double échelle de temps s'il peut être découplé par changement de variable en deux sous-systèmes, un lent et un rapide, tel que les valeurs propres du modèle lent soient, en module, très inférieures à celles du modèle rapide (Cf II : I.4.2). Les valeurs propres étant conservées par changement de variables, on peut chercher à bloc-diagonaliser la matrice d'évolution du système initial , on utilise alors une méthode algébrique, ou bien à localiser les valeurs propres de cette matrice par une méthode géométrique.

## I.1. METHODES ALGEBRIQUES

### I.1.1. Bloc diagonalisation

Soit le système linéaire :

$$(1)\quad \begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ y = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad ; \; x_1 \in \mathbb{R}^{q_1},\; x_2 \in \mathbb{R}^{q_2},\; u \in \mathbb{R}^m,\; y \in \mathbb{R}^n \\ q_1 + q_2 = q \end{cases}$$

Pour bloc-diagonaliser ce système, on utilise un changement de variables| CHA,72 | qui, dans une première étape, bloc-triangularise (1), puis dans une deuxième le bloc-diagonalise :

* bloc-triangularisation

la transformation

$$(2)\quad \begin{bmatrix} x_1 \\ x_r \end{bmatrix} = \begin{bmatrix} I_{q_1} & 0 \\ L & I_{q_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ où } L \in \mathbb{R}^{q_2 \times q_1}$$

amène le système (1) sous la forme :

$$(3)\qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & A_{12} \\ R(L) & A_r \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_r \end{bmatrix} u$$

$$y = [\, C_1 \quad C_2 \,] \begin{bmatrix} x_1 \\ x_r \end{bmatrix}$$

avec $A_1 = A_{11} - A_{12}L$

$A_r = A_{22} + LA_{12}$

$R(L) = A_{21} + LA_{11} - LA_{12}L - A_{22}L$

$B_r = B_2 + LB_1$

$C_1 = C_1 - C_2L$

Le système (3) est bloc-triangularisé si L est solution de l'équation de Riccati : $R(L) = 0$.

* bloc-diagonalisation

la deuxième transformation

$$(4)\qquad \begin{bmatrix} x_1 \\ x_r \end{bmatrix} = \begin{bmatrix} I_{q_1} & -M \\ 0 & I_{q_2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix} \quad \text{où } M \in \mathbb{R}^{q_1 \times q_2}$$

modifie le système (3), avec $R(L) = 0$, sous la forme

$$(5)\qquad \begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_r \end{bmatrix} = \begin{bmatrix} A_1 & R(M) \\ 0 & A_r \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix} + \begin{bmatrix} B_1 \\ B_r \end{bmatrix} u \\ y = [\, C_1 \quad C_r \,] \begin{bmatrix} x_1 \\ x_r \end{bmatrix} \end{cases}$$

$$\begin{cases} \text{avec } R(M) = A_{12} + A_1M - MA_r \\ B_1 = B_1 - MB_r \\ C_r = C_2 + C_1M \end{cases}$$

Ce système apparaît alors sous la forme bloc-diagonalisée si M est solution de l'équation de Lyapounov: $R(M) = 0$.

Remarque : Dans le cas où le système est déjà sous forme singulièrement perturbée, c'est-à-dire :

$$A_{21} = \frac{A_{21}^*}{\varepsilon} \quad , \quad A_{22} = \frac{A_{22}^*}{\varepsilon} \quad , \quad B_2 = \frac{B_2^*}{\varepsilon} \quad ,$$

cette méthode est encore applicable et une approximation à l'ordre 0 fournit :

$$L_o = A_{22}^{-1} A_{21}$$
$$A_{1o} = A_{11} - A_{12} A_{22}^{-1} A_{21}$$
$$A_{ro} = A_{22}$$
$$M_o = A_{12} A_{22}^{-1}$$
$$B_{ro} = B_2$$
$$C_{1o} = C_1 - C_2 A_{22}^{-1} A_{21}$$
$$B_{1o} = B_1 - A_{12} A_{22}^{-1} B_2$$
$$C_{ro} = C_2$$

### I.1.2. Méthodes de détermination des matrices L et M

L'existence et le calcul des solutions des équations de Riccati, R(L) = 0 et de Lyapounov, R(M) = 0, ont été étudiées par de nombreux auteurs. Le principal problème à résoudre est, bien sûr, la résolution de l'équation de Riccati ; la résolution de l'équation de Lyapounov n'offre pas de difficultés (Cf. II : II.2.2.). Pour déterminer L, solution de R(L) = 0, plusieurs méthodes ont été proposées : celles nécessitant la connaissance des vecteurs propres et des valeurs propres de la matrice initiale | AND,78 ; AVR,79 ; SYR,82 | ; celles utilisant le polynome annulateur | MAG,81 | qui ont l'avantage par rapport aux précédentes de ne pas nécessiter la connaissance "a priori" des valeurs et/ou vecteurs propres ; enfin, celles résolvant, par un algorithme numérique l'équation de Riccati | KOK,75 ; CHO,76 |.

I.1.2.1. Utilisation du polynome annulateur

Soient $\psi$, le polynome caractéristique de A, | MAG,81 | part du polynome annulateur, $\psi_1$,$\psi_2$,deux polynomes définis par :

(6) $\psi = \psi_1\psi_2$ ; $d°\psi_1 = q_1$, $d°\psi_2 = q_2$, $q_1 + q_2 \doteq q$

où les racines de $\psi_1$ et $\psi_2$ appartiennent respectivement à $C_1$ et $C_2$ tels que

$C_1 \cap C_2 = \emptyset$ $C_1$ regroupe les valeurs propres lentes

$C_1 \cup C_2 = Sp(A)$ $C_2$ regroupe les valeurs propres rapides

De façon à éviter tout calcul de vecteur propre, la méthode proposée par | MAG,81 | consiste à partir de la matrice

(7) $$N = \begin{bmatrix} \psi_1(A) \\ I_q \end{bmatrix} \quad \text{où } A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

puis par des combinaisons linéaires des $q_2$ dernières colonnes avec les $q_1$ premières de (7), on la transforme en :

(8) $$\begin{array}{c} q \\ q_1 \\ q_2 \end{array} \left[\begin{array}{c:c} 0 & X \\ \hdashline I_{q_1} & 0 \\ \hdashline -L & I_{q_2} \end{array}\right] \quad \begin{array}{l} \\ \text{colonnes : } \underbrace{\phantom{-L}}_{q_1}\ \underbrace{\phantom{I_{q_2}}}_{q_2} \end{array}$$

où apparaît explicitement la matrice L.

Le choix de $q_1$ peut être fixé en posant d'une part |FOS,81| :

(9) $\psi_1(p) = p^{q_1}$

puis, en explorant pour des valeurs successives de $q_1$, les solutions L obtenues dans (8). Elles fournissent une valeur à $||R(L)||$ qui passe par un minimum lorsque la séparation des valeurs propres est la plus importante. Une méthode de séparation de polynomes peut être

utilisée à ce niveau de l'étude pour améliorer le choix de $q_1$ et de L.

a) Algorithme de séparation de polynomes.

Considérons un polynome D(p) que l'on veut écrire sous la forme :

$$(10) \qquad D(p) = A(p)\,B(p)$$

$$\text{avec} \quad D(p) = p^q + \sum_{i=0}^{q-1} d_i p^i$$

$$A(p) = \sum_{i=0}^{n} a_i p^i \; ; \; a_n = 1$$

$$B(p) = \sum_{i=0}^{m} b_i p^i \; ; \; b_m = 1$$

$$m + n = q$$

L'égalité (10) nous conduit au système d'équations:

$$(11) \qquad i = 0,..,q-1 \qquad d_i = \sum_{j=0}^{q-i} a_{n-q+i+j} b_{m-j}$$

avec $a_i = 0$ si $i \notin \{0,n\}$

et $b_i = 0$ si $i \notin \{0,m\}$

qui se met sous les formes matricielles

$$(12) \qquad \text{a)} \quad d = \bar{a} + \mathcal{A} b$$

$$\text{b)} \quad d = \bar{b} + \mathcal{B} a$$

$$\text{où } d = \begin{bmatrix} d_{q-1} \\ \vdots \\ d_1 \\ d_0 \end{bmatrix} \; ; \quad a = \begin{bmatrix} a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \\ a_0 \end{bmatrix} \; ; \quad b = \begin{bmatrix} b_{m-1} \\ b_{m-2} \\ \vdots \\ b_1 \\ b_0 \end{bmatrix} \; ;$$

$$\bar{a} = \begin{bmatrix} a \\ 0 \\ \vdots \\ 0 \end{bmatrix} \; ; \quad \bar{b} = \begin{bmatrix} b \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

et

$$\mathcal{A} = \begin{bmatrix} 1 & 0 \\ a_{n-1} & 1 \\ a_{n-2} & a_{n-1} \\ \vdots & \vdots \\ a_1 & \\ a_0 & \\ 0 & a_0 \end{bmatrix} \quad \mathcal{B} = \begin{bmatrix} 1 & 0 \\ b_{m-1} & 1 \\ b_{m-2} & b_{m-1} \\ \vdots & \vdots \\ b_1 & \\ b_0 & b_0 \end{bmatrix}$$

m col. n col.

A partir de (12), l'utilisation de la pseudo-inverse au sens de Moore-Penrose, donne :

(13) a) $\bar{b} = \mathcal{A}^{+}(d - \bar{a})$ où $\mathcal{A}^{+} = (\mathcal{A}^T\mathcal{A})^{-1}\mathcal{A}^T$

b) $\bar{a} = \mathcal{B}^{+}(d - \bar{b})$ où $\mathcal{B}^{+} = (\mathcal{B}^T\mathcal{B})^{-1}\mathcal{B}^T$

L'algorithme de séparation de polynomes est donc défini de la façon suivante :

(14) pour $k \geqslant 0$ $\quad b_{(k)} = \mathcal{A}^{+}_{(k)} \, |d - \bar{a}_{(k)}| \; ; \; a_{(k+1)} = \mathcal{B}^{+}_{(k)} \, |d - \bar{b}_{(k)}|$

Si l'on prend la condition initiale, $a_{(0)} = 0$, c'est-à-dire, $A(p) = p^n$, l'algorithme converge vers les coefficients des polynomes A(p) et B(p), tels que les racines de A(p) soient plus petites (en module) que celles de B(p). Si on tente une séparation impossible (e.g. un polynome du second degré à racines complexes séparé en deux polynomes du premier degré), soit l'algorithme diverge, soit le produit des polynomes obtenus ne donne pas le polynome de départ.

b) Application à la séparation de dynamiques

Le but est de trouver la décomposition (6), donc $q_1$, qui assure comme précédemment le minimum de $||R(L)||$. Nous proposons l'algorithme suivant :

initialisation

$\overset{\circ}{\psi}_2(p) := \psi(p)$

$\overset{\circ}{\psi}_1(p) := 1$

$i := 1$

par l'algorithme (14)

Détermination de $\psi_1^i(p)$ et $\psi_2^i(p)$ tels que

$\psi_2^{i-1}(p) = \psi_1^i(p)\ \psi_2^i(p)$

avec $\psi_1^i(p) = p^2 + a_1^i p + a_0^i$

$a_1^{i2} - 4a_0^i > 0$

$\psi_1^i(p) = (p + \lambda_1)(p + \lambda_2)$

$|\lambda_1| < |\lambda_2|$

$\psi_1^i(p) := (p + \lambda_1)$

$\psi_2^i(p) := (p + \lambda_2)\ \psi_2^i(p)$

$\psi_1^i(p) := \prod_{j=0}^{i} \psi_1^j(p)$

$i := i + 1$

par l'algorithme(8)

détermination de L

$||\ R(L)\ ||$ minimum ?

FIN

BU LILLE

Afin d'utiliser les algorithmes présentés avec la meilleure efficacité, il est nécessaire de reordonner les variables d'état pour que la nouvelle matrice d'évolution ait des termes diagonaux décroissants (Cf. exemple présenté dans | FOS,83 |).

### I.1.2.2. Algorithmes numériques

Le principe de ces méthodes est la résolution numérique itérative de l'équation de Riccati :

$$A_{21} + LA_{11} - LA_{12}L - A_{22}L = 0 \tag{15}$$

Plusieurs algorithmes ont été proposés pour définir une suite de matrices convergeant vers la solution de (15). | KOK,75 ; CHO,76 | proposent d'écrire (15) sous la forme :

$$\begin{aligned} L_{k+1} &= A_{22}^{-1}\,(A_{21} + L_kA_{11} - L_kA_{12}L_k) \\ L_o &= A_{22}^{-1}A_{21} \end{aligned} \tag{16}$$

qui converge vers une racine réelle bornée de (15) si les conditions suivantes (très conservatives) sont réalisées :

$$\begin{aligned} &a)\ A_{22} \text{ inversible} \\ &b)\ \|A_{22}^{-1}\| > \tfrac{1}{3}\,\|A_o\| + \|A_{12}\|\,\|L_o\|)^{-1} \\ &\quad \text{où } A_o = A_{11} - A_{12}L_o \end{aligned} \tag{17}$$

Les conditions (17) assurent, de plus :

$$\forall\, k \in \mathbb{N} \qquad \| L_{k+1} - \hat{L} \| < \| L_k - \hat{L} \| \tag{18}$$

où $\hat{L}$ est la solution cherchée

Un autre algorithme utilisant une transformation homographique de matrices est proposé dans | AND,81 | et défini par :

$$\begin{cases} L_{k+1} = (A_{22} + L_kA_{12})^{-1}\,(A_{21} + L_kA_{11}) \\ L_o = A_{22}^{-1}\,A_{21} \end{cases} \tag{19}$$

Les conditions de convergence de cet algorithme sont similaires à (17-b) de l'algorithme précédent, bien que pout tout k, on doive avoir $(A_{22}+L_kA_{12})$ régulière. L'avantage de cet algorithme est d'assurer une convergence plus rapide vers la solution. Comme le montre l'exemple présenté dans | KOK,75 |, les conditions (17), obtenues par majoration, sont trop fortes et L peut souvent être déterminée, même si le système ne vérifie pas strictement (17-b). Dans tous les cas, on a une convergence si les valeurs propres de $A_1$ sont très inférieures en modules à celles de $A_r$. De façon à éviter cet inconvénient, nous proposons l'algorithme suivant qui substitue à l'équation de Riccati, une équation de Lyapounov :

Considérons (15) écrit sous la forme itérative :

$$A_{21} + L_{k+1}A_{11} - L_{k+1} A_{12} L_k - A_{22}L_{k+1} = 0 \tag{20}$$

cela conduit à l'algorithme :

$$L_o = A_{22}^{-1}A_{21}$$

$$\text{(21) a)}\quad B_k = A_{11} - A_{12} L_k$$

$$\text{b)}\quad A_{21} + L_{k+1} B_k - A_{22}L_{k+1} = 0$$

L'équation (21 - b) est de Lyapounov et peut être résolue à chaque pas par la mise sous forme vectorielle | BAR,70 | :

$$\text{vec}\,(L_{k+1}) = |\, I_{q_1} \otimes A_{22} - B_k \otimes I_{q_2} \,|^{-1} \text{vec}\,(A_{21}) \tag{22}$$

(où vec(.) est la fonction vectorielle de matrice introduite en (II : 55).

La solution de (22) existe si $B_k$ et $A_{22}$ ont des valeurs propres distinctes. Afin de montrer l'avantage de cette méthode, considérons l'exemple comparatif suivant :

* soit à bloc-diagonaliser la matrice

$$A = \left[\begin{array}{c:cc} 0 & -1 & -1 \\ \hdashline 1 & -2 & -1 \\ 1 & -1 & -2 \end{array}\right]$$ par le changement de variables $$T = \left[\begin{array}{c:cc} 1 & 0 & 0 \\ \hdashline 1 & 1 & 0 \\ & 0 & 1 \end{array}\right]$$

les résultats obtenus sont les suivants :

| ALGORITHME | $TAT^{-1}$ |
|---|---|
| KOKOTOVIC, CHOW (16) | $\begin{bmatrix} -0,8148 & -1 & -1 \\ 0,1097 & -1,5926 & -0,5926 \\ 0,1097 & -0,5926 & -1,5926 \end{bmatrix}$ |
| ANDERSON, HALLAUER (19) | $\begin{bmatrix} -0,8572 & -1 & -1 \\ 0,0816 & -1,5714 & -0,5714 \\ 0,0816 & -0,5714 & -1,5714 \end{bmatrix}$ |
| ROTELLA (21) | $\begin{bmatrix} -1 & -1 & -1 \\ 0 & -1,5 & -0,5 \\ 0 & -0,5 & -1,5 \end{bmatrix}$ |

L'algorithme (21) est plus couteux en temps de calcul, mais apparaît plus efficace.

Pour déterminer M, solution de l'équation de Lyapounov :

(23) $$A_{12} + (A_{11} - A_{12}\bar{L})M - M(A_{22} + \bar{L}A_{12}) = 0$$

où $\bar{L}$ est solution de (15)

| KOK 75, CHO 76 | proposent l'algorithme récursif :

(24) $$M_{k+1} = \left[(A_{11} - A_{12}\bar{L})M_k - M_k \bar{L}A_{12}\right] A_{22}^{-1} + A_{12}A_{22}^{-1}$$

qui converge vers une solution si (17) est vérifiée. Nous pensons qu'il est plus facile de résoudre directement (23) qui est une équation similaire à (21 - b) ; la solution est donc donnée sous forme vectorielle par :

(25) $$\mathrm{vec}(\bar{M}) = \left[ I_{q_1} \otimes (A_{22} + \bar{L}A_{12}) - (A_{11} - A_{12}\bar{L}) \otimes I_{q_2}\right]^{-1} \mathrm{vec}(A_{12})$$

Dans le cas où $R(L) \neq 0$, la transformation (4) amène alors la matrice d'évolution initiale sous la forme :

$$(26) \quad \begin{bmatrix} A_{11} - A_{12}L - MR(L) & R(M)^* \\ R(L) & A_{22} + LA_{12} + R(L)M \end{bmatrix}$$

avec

a) $R(L) = A_{21} + LA_{11} - LA_{12}L - A_{22}L$

b) $R(M)^* = A_{12} + (A_{11} - A_{12}L)M - M(A_{22} + LA_{12}) - MR(L)M$

Cette forme est triangularisée si dans (26 b) $R(M) = 0$, ce qui correspond de nouveau à une équation de Riccati résoluble par les méthodes précédentes.

L'inconvénient de ces méthodes est de nécessiter, comme dans la partie précédente, un réarrangement de l'ordre des variables d'état et même parfois un conditionnement préalable. L'utilisation des méthodes géométriques permet d'estimer les ordres de grandeurs des dynamiques des variables, donc d'effectuer des groupements de variables afin de faciliter l'utilisation des méthodes précédentes.

## I.2 METHODES GEOMETRIQUES

Le principe des méthodes géométriques est l'utilisation des théorèmes de localisation des valeurs propres d'une matrice constante. Ces méthodes ont été développées dans | DAU,87a | et permettent un partionnement du système initial en un ensemble de variables d'état lentes et un autre regroupant les variables d'état rapides. L'autre avantage est l'utilisation possible des méthodes géométriques dans le cas de systèmes non-linéaires alors que les méthodes précédentes ne peuvent être employées.

Différents domaines ont été définis pour la localisation des valeurs propres d'une matrice dans le plan complexe : les cercles de Gershgorine | DEI,82 | dont une version généralisée a été proposée

dans | SOL,80 | ; les ovales de Cassini | BRA,64 | ou les régions de Gudkov | CHA,71 | qui fournissent des localisations plus fines mais plus difficiles à obtenir. Certaines de ces méthodes ont été utilisées par | DAU,83a | pour les cercles de Gershgorine et par | MOR,84 | pour les ovales de Cassini pour mettre en évidence les différentes dynamiques d'un système. Une autre classe de méthodes consiste à utiliser une homotopie du plan complexe | GUT,81 | ; en particulier, l'utilisation de la transformation en W et du théorème de Routh permet de séparer les valeurs propres à petits modules de celles à grands modules | BEN,86 |. Nous nous limiterons dans nos rappels au cercles de Gershgorine qui, étant les plus faciles à obtenir, seront les plus faciles d'utilisation dans le cas bilinéaire.

### I.2.1. Utilisation des cercles de Gershgorine

Notons $a_{ij}$, $i, j = 1,\ldots,q$, les éléments de la matrice A et définissons les quantités :

$$(27)\qquad \begin{cases} l_i = \sum_{\substack{j=1 \\ j\neq i}}^{q} [a_{ij}], \\ c_i = \sum_{\substack{j=1 \\ j\neq i}}^{q} [a_{ji}], \end{cases} \qquad i = \{1,\ldots,q\}$$

Le théorème suivant permet une localisation rapide des valeurs propres de la matrice A :

Théorème de Gershgorine : Les valeurs propres de A appartiennent au domaine formé par l'union des ensembles du plan complexe définis par :

$$\forall \alpha \in [0,1]$$
$$i = \{1,\ldots, q\} \quad | z - a_{ii} | < R_i = l_i^{\alpha} c_i^{1-\alpha}$$

Le domaine de Gershgorine est donc formé de la réunion de q disques centrés sur les termes diagonaux et toute séparation en sous-ensembles disjoints implique une séparation équivalente au

niveau des valeurs propres. Ce résultat a permis de montrer :

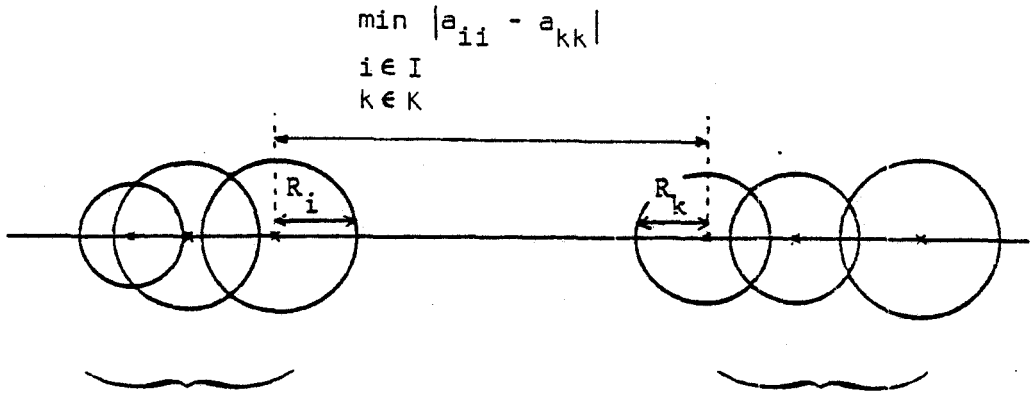
Théorème 1| DAU 83a, 85 | : si les cercles de Gershgorine associés à une matrice A sont tels que :

$$\exists I, K \; ; \; I \cap K = \emptyset \; ; \; I \cup K = \{ 1, \ldots, q \}$$

$$\forall i \in I, \forall k \in K \quad \| a_{ii} - a_{kk} \| >> (R_i + R_k)$$

alors A possède deux ensembles de valeurs propres séparées de modules très différents.

La vérification de ce théorème permet de relever la situation suivante :




qui met en évidence un système à dynamiques différentes ; le rapport des dynamiques, donc la précision du résutat dépend du paramètre

$$(28) \qquad \mu = \max_{\substack{i \in I \\ k \in K}} \frac{R_i + R_k}{\| a_{ii} - a_{kk} \|}$$

qui, si le théorème est vrai, est nécessairement très inférieur à 1.

L'application de cette méthode permet une mise en évidence directe du petit paramètre nécessaire à la méthode des perturbations singulières ; l'autre avantage est de grouper les variables d'état

en lentes et rapides. L'inconvénient est de nécessiter quelquefois un préconditionnement de la matrice à étudier par une transformation diagonale. Nous allons voir dans la partie suivante que l'utilisation de la version par blocs du théorème de Gershgorine permet, parfois, de simplifier cette étape préalable.

### I.2.2. Utilisation du théorème de Gershgorine par blocs

Ces résultats qui généralisent le théorème de Gershgorine ont été énoncés par | FEI,62 ; KOV,74 ; POW 76 | et utilisés en Automatique par | SOL,80 |.

Soit A une matrice q x q complexe bloc-partitionnée sous la forme :

$$(29) \qquad A = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \cdots & A_{NN} \end{bmatrix}$$

où $\forall i \in \{1,\ldots, N\}$ $A_{ii} \in \mathbb{R}^{q_i \times q_i}$ ; $\sum_{i=1}^{N} q_i = q$

et des normes matricielles subordonnées à des normes vectorielles par :

$$(30) \qquad \forall\, i, j \in \{1,\ldots, N\} \quad ||A_{ij}|| = \sup_{\substack{x \neq 0 \\ x \in \mathbb{R}^{q_j}}} \frac{||A_{ij}x||_i}{||x||_j}$$

et si $A_{ii}$ régulière

$$\forall\, i \in \{1,\ldots, N\} \quad ||A_{ii}^{-1}|| = \inf_{\substack{x \neq 0 \\ x \in \mathbb{R}^{q_i}}} \frac{||A_{ii}x||_i}{||x||_i}$$

Remarque : si l'on utilise dans (30) la norme euclidienne, on a :

$||A_{ij}|| = \rho(A_{ij}A_{ij}^T)$ où $\rho(.)$ désigne le rayon spectral

et dans le cas où $A_{ii}$ est singulière $(\| A_{ii}^{-1} \|)^{-1} = 0$

On a alors le théorème suivant :

Théorème 2 : les valeurs propres de A (29) appartiennent au domaine formé par l'union des ensembles $F_i$ du plan complexe défini par

$$\forall \alpha \in [0,1]$$

$$i = \{1, \ldots, N\} F_i : \quad \| (A_{ii} - zI_i)^{-1} \|^{-1} \leqslant \mathcal{R}_i = \mathcal{L}_i^{\alpha} \mathcal{C}_i^{1-\alpha}$$

$$\text{où } \mathcal{L}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \|A_{ij}\|$$

$$\mathcal{C}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \|A_{ij}\|$$

Remarque : si N = q, on retrouve le théorème de Gershgorine. De même en ce qui concerne le théorème 3.

Théorème 3 :

$$\text{Si } H = \bigcup_{j=1}^{m} F_{p_j}, \quad 1 \leqslant p_j \leqslant N,$$

est disjoint des N-m

autres ensembles de Gershgorine pour la matrice partitionnée, alors H contient précisément $\sum_{j=1}^{m} q_{p_j}$ valeurs propres de A.

Ces théorèmes sont intéressants car ils permettent d'utiliser les notions de systèmes pseudo-majorants et de normes vectorielles

déjà utilisées pour l'analyse et la synthèse de systèmes non-linéaires | GEN, 72,76 ; GRU,76,79| pour détecter les systèmes à plusieurs dynamiques. Cependant, le point délicat est la détermination de la quantité $||(A_{ii} - zI_i)^{-1}||^{-1}$ ; il peut être contourné à l'aide du résultat suivant :

Théorème 4 : si A, partitionnée suivant (29) est telle que :

$\forall\, i \in \{1,\ldots, N\}$ $A_{ii}$ est normale (i.e. $A_{ii}^T A_{ii} = A_{ii} A_{ii}^T$)

alors l'utilisation de la norme euclidienne implique que chaque ensemble de Gershgorine $F_i$ est la réunion de $q_i$, $i = 1,\ldots, N$, disques.

Le principe de la méthode de localisation utilisant le théorème 3 est donc de partitionner la matrice initiale de telle façon que chaque matrice de la bloc-diagonale principale soit normalisable par changement de variable. A cette fin, rappelons qu'une matrice de dimension 2 est normale si elle est de la forme :

$$(31) \quad a) \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \text{ ou } b) \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}$$

où $\alpha$, $\beta$, $\gamma$ sont des réels

ce qui signifie qu'une matrice de dimension 2 diagonalisable dans $\mathbb{C}$ est toujours normalisable dans $\mathbb{R}$. De façon générale, on peut proposer la méthode suivante :

1) choix du partitionnement de A (29) tel que

$\forall i \in \{1,\ldots, N\}$ $A_{ii}$ diagonalisable

2) diagonalisation des $A_{ii}$ et transformation de A par

$$A_{ij} \longrightarrow A'_{ij} = P_i^{-1} A_{ij} P_j$$

où Pi est la transformation diagonalisant $A_{ii}$ dans $\mathbb{C}$

$$A'_{ii} = P_i^{-1} A_{ii} P_i = \operatorname{diag}_{j=1}^{q_i} \{ \lambda_j^i \} \qquad \lambda_j^i \in \mathbb{C}$$

3) tracé des ensembles de Gershgorine à l'aide de :

$$\| (A'_{ii} - zI_i)^{-1} \|^{-1} = \left\{ |z - \lambda^i_j|,\ j \in \{1,\ldots,q_i\} \right\} i \in \{1,\ldots,N\}$$

Cette méthode peut permettre une détermination plus précise des valeurs propres d'une matrice que l'utilisation du théorème de Gershgorine. En effet, considérons l'exemple d'un système défini par la matrice d'évolution :

$$A = \begin{vmatrix} -4 & 2 & .5 & 0 \\ 2 & -4 & 0 & .5 \\ .5 & 0 & -.4 & .2 \\ 0 & .5 & .2 & -.4 \end{vmatrix} \tag{32}$$

L'utilisation du théorème de Gershgorine répartit les valeurs propres de A à l'intérieur des cercles $C_1(-4,2.5)$ et $C_2(-.4,.7)$. Le théorème 1 indique donc une séparation en deux dynamiques correspondant à un coefficient $\mu = .89$.

Si l'on partitionne A(32) par :

$$A_{11} = \begin{vmatrix} -4 & 2 \\ 2 & -4 \end{vmatrix}$$

$$A_{12} = A_{21} = \begin{vmatrix} .5 & 0 \\ 0 & .5 \end{vmatrix} \tag{33}$$

$$A_{22} = \begin{vmatrix} -.4 & .2 \\ .2 & -.4 \end{vmatrix}$$

chacun des blocs de la diagonale principale étant normal, l'utilisation de la norme euclidienne fournit :

$$\begin{aligned} &\| A_{12} \| = \| A_{21} \| = .5 \\ &\| (A_{11} - zI_2)^{-1} \|^{-1} = \min \{ | -6-z |, |-2-z | \} \\ &\| (A_{22} - zI_2)^{-1} \|^{-1} = \min \{ | -.6-z |, | -.2-z | \} \end{aligned} \tag{34}$$

Le théorème de Gershgorine par blocs conduit donc aux régions

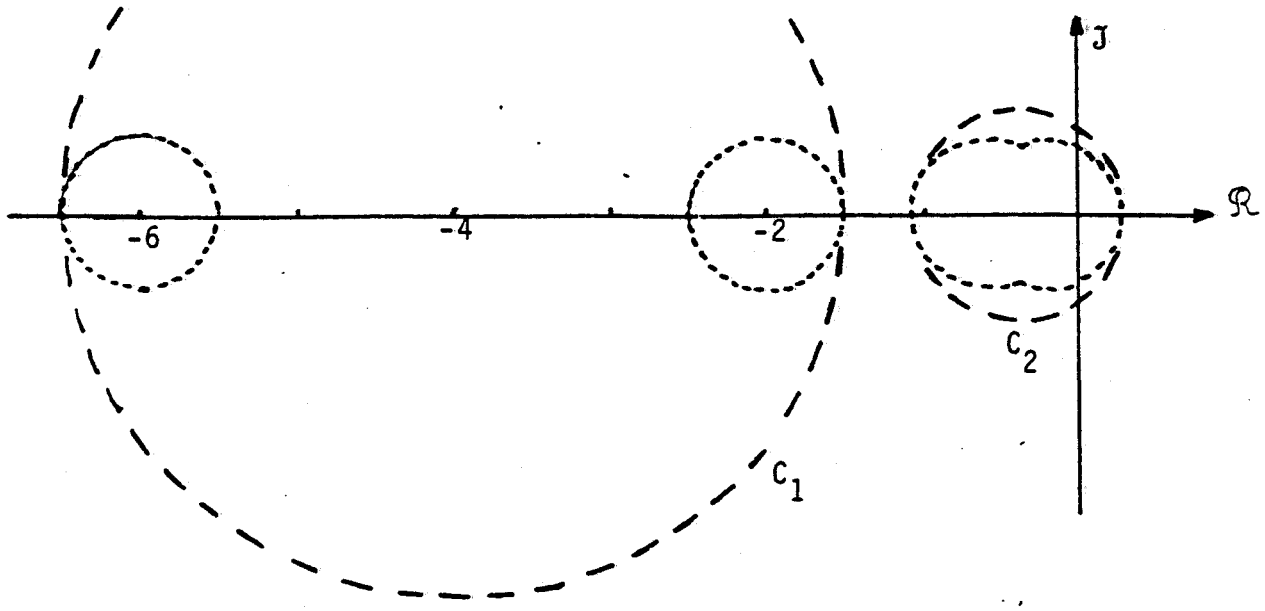
du plan complexe définies par

$$\min \{ | -6 - z |, | -2 - z | \} < .5$$

(35)

$$\min \{ | -.6 - z |, | -.2 - z | \} < .5$$

qui localisent les valeurs propres de A à l'intérieur de quatre cercles (figure 1)




_ _ _ Théorème de GERSHGORINE

_ _ _ _ _ Théorème par blocs

$Sp(A) = \{ -6.0459, -2.1296, -.5541, -.0704 \}$

FIGURE 1

Cette localisation indique une séparation en 3 dynamiques correspondant aux coefficients $\mu_1 = .71$ et $\mu_2 = .25$. Le changement de variable, défini par la matrice :

(36)

$$P = \begin{vmatrix} .7071 & .7071 & 0 & 0 \\ -.7071 & .7071 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

diagonalisant le bloc $A'_{11}$, permet de distinguer une variable très rapide, une variable rapide et un groupe de deux variables lentes. Ainsi, cette méthode, dans le cas de systèmes de grandes dimensions, s'avère être un intermédiaire entre la diagonalisation complète et la localisation des valeurs propres par le théorème de Gershgorine.

## II. APPLICATION AUX SYSTEMES BILINEAIRES

Les systèmes bilinéaires étant régis par une équation d'état de la forme :

$$(37)\quad \begin{cases} \dot{x}(t) = \left[ A_o + \sum_{i=1}^{m} u_i(t)\, A_i \right] x(t) + \sum_{i=1}^{m} B_i\, u_i(t) \\ y(t) = C\, x(t) \quad \text{où } x(t) \in \mathbb{R}^q,\ y(t) \in \mathbb{R}^n \end{cases}$$

les valeurs propres de la matrice d'évolution sont non constantes et dépendent de façon continue, mais non linéaire, des variables d'entrée $u_i$. Le problème de la détermination de dynamiques distinctes dans les variables d'état peut être abordé de deux façons différentes :

* soit globalement, indépendamment des entrées,

* soit localement, la séparation définie n'étant valable que dans un certain domaine de variations des entrées.

La première de ces deux démarches sera effectuée à partir des matrices constantes $A_i$, $i = 0, \ldots, m$ ; alors que la deuxième sera menée sur la matrice d'évolution $A_o + \sum_{i=1}^{m} u_i A_i$, en exploitant les résultats présentés dans la première partie de ce chapitre.

## II.1. SYSTEMES LINEAIRES NON STATIONNAIRES A DEUX ECHELLES DE TEMPS

Pour les systèmes linéaires non stationnaires définis par :

$$(38) \qquad \dot{x}(t) = A(t)\, x(t) + B(t)\, u(t),\ x(t) \in \mathbb{R}^q$$

| O'MA,82 | a proposé une définition de double échelle de temps qui généralise celle adoptée pour les systèmes linéaires.

Définition : Le système (38) est à double échelle de temps sur $[0,T]$ si le spectre de $A(t)$, pour tout $t$ dans $[0,T]$ peut être partitionné en deux ensembles disjoints $\mathcal{L}(t)$ et $\mathcal{R}(t)$ tels que :

$$\forall t \in [0,T] \quad \max_{l_i(t) \in \mathcal{L}(t)} |l_i(t)| = \tilde{l}(t) \ll \tilde{r}(t) = \min_{r_i(t) \in \mathcal{R}(t)} |r_i(t)|$$

Un indicateur de séparation de dynamique est donné par le rapport :

$$(39) \qquad \mu = \max_{0<t<T} \frac{\tilde{l}(t)}{\tilde{r}(t)}$$

qui doit être très inférieur à 1 si l'on veut qu'il y ait effectivement séparation de dynamique.

Afin de mettre en évidence dans le système initial deux sous-systèmes de dynamiques différentes, il proposent, à partir d'un partionnement de la matrice d'évolution de (38) sous la forme :

$$(40) \qquad A(t) = \begin{bmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{bmatrix} \quad ; \quad B(t) = \begin{bmatrix} B_1(t) \\ B_2(t) \end{bmatrix}$$

où $\forall t \in |0,T|$ , $B_i(t) \in \mathbb{R}^{q_i}$ , $A_{ij}(t) \in \mathbb{R}^{q_i \times q_j}$ , $(i, j) \in \{1,2\}$

$$q_1 + q_2 = q$$

d'utiliser la transformation :

$$\forall t \in |0,T|\ L(t) \in \mathbb{R}^{q2 \times q1},\ M(t) \in \mathbb{R}^{q_1 \times q_2}$$

$$T(t) = \begin{bmatrix} I - M(t)\,L(t) & -M(t) \\ L(t) & I \end{bmatrix} \tag{41}$$

Cette transformation est la version non stationnaire de la transformation utilisée pour les systèmes linéaires (I.1.1) ; le système (38) est alors découplé sous la forme :

$$\begin{bmatrix} \overset{\circ}{x}_1(t) \\ \overset{\circ}{x}_r(t) \end{bmatrix} = \begin{bmatrix} A_1(t) & 0 \\ 0 & A_r(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_r(t) \end{bmatrix} + \begin{bmatrix} B_1(t) \\ B_r(t) \end{bmatrix} u(t) \tag{42}$$

$$\text{avec } A_1(t) = A_{11}(t) - A_{12}(t)\,\hat{L}(t)$$

$$A_r(t) = A_{22}(t) + \hat{L}(t)\,A_{12}(t)$$

$$B_1(t) = B_1(t) - \hat{M}(t)B_r(t)$$

$$B_r(t) = B_2(t) + \hat{L}(t)B_1(t)$$

si $\hat{L}(t)$ est solution de l'équation de Riccati différentielle :

$$\overset{\circ}{L}(t) = A_{22}(t)\,L(t) - L(t)\,A_{11}(t) + L(t)\,A_{12}(t)L(t) - A_{21}(t) \tag{43}$$
$$\text{avec } \overset{\circ}{L}(0) = 0$$

et si $\hat{M}$ est solution de l'équation de Lyapounov différentielle :

$$-\overset{\circ}{M}(t) = |A_{11}(t) - A_{12}(t)\hat{L}(t)|M(t) - M(t)\,|A_{22}(t) + \hat{L}(t)A_{12}(t)| + A_{12}(t) \tag{44}$$
$$\text{avec } \overset{\circ}{M}(T) = 0$$

Cette démarche est applicable aux systèmes bilinéaires qui sont une écriture particulière de systèmes non stationnaires. Ce qui, avec un partitionnement défini par :

$$(45)\qquad \forall i \in \{0,\ldots,m\} \qquad A_i = \begin{bmatrix} A_{11}^{\,i} & A_{12}^{\,i} \\ A_{21}^{\,i} & A_{22}^{\,i} \end{bmatrix}$$

$$\text{où } A^i_{kl} \in \mathbb{R}^{q_k \times q_l} \quad ,(k,\ l)\in \{1,2\}^2$$

donne pour l'équation (43)

$$(46)\qquad \begin{aligned} \dot{L}(t) &= \sum_{i=0}^{m} (A_{22}^{\,i} L(t) - L(t)A_{11}^{\,i} + L(t)A_{12}^{\,i}L(t) - A_{21}^{\,i})u_i(t) \\ &\text{où } u_o(t) = 1 \text{ et } L(o) \text{ solution de l'équation de} \\ &\text{Riccati initiale} \end{aligned}$$

$$\sum_{i=0}^{m} A_{21}^{\,i} u_i(0) + L \sum_{i=0}^{m} A_{11}^{\,i} u_i(0) - \sum_{i=0}^{m} A_{22}^{\,i} u_i(0)L - L \sum_{i=0}^{m} A_{12}^{\,i} u_i(0)L = 0$$

Après avoir résolu l'équation de Riccati initiale, on peut intégrer numériquement l'équation différentielle de Riccati(46) par des procédes classiques tels que la méthode de Runge-Kutta ou la méthode de prédicteur-correcteur utilisant les formules d'Adams | NOU,83 |. La lourdeur des calculs impliqués rend cette bloc-diagonalisation peu applicable en temps réel lorsque les entrées ne sont pas a priori connues. C'est pour cette raison que nous nous proposons d'utiliser les méthodes propres aux systèmes linéaires qui, à défaut de donner des sous-systèmes parfaitement découplés, donneront plus rapidement une solution approchée.

## II.2 METHODE ALGEBRIQUE DE PSEUDO-DIAGONALISATION | ROT,85 |

Appliquons la transformation constante (2) au système bilinéaire (37-45) où de plus :

$$(47)\qquad \begin{cases} x^T = |x_1^T,\ x_2^T| \ ;\ x_1 \in \mathbb{R}^{q_1},\ x_2 \in \mathbb{R}^{q_2},\ q_1 + q_2 = q \\ \forall i \in \{1,\ldots,m\} \qquad B_i^T = |B_1^{iT}\ B_2^{iT}| \\ C = |C_1\ C_2| \end{cases}$$

cela donne :

$$(48)\quad \begin{bmatrix} \overset{\circ}{x}_1 \\ \\ \overset{\circ}{x}_r \end{bmatrix} = \left| \tilde{A}_0 + \sum_{i=1}^{m} \tilde{A}_i u_i \right| \begin{bmatrix} x_1 \\ \\ x_r \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} B_1^i \\ \\ B_r^i \end{bmatrix} u_i$$

$$y = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_r \end{bmatrix}$$

$$\text{où } \forall\, i \in \{0,\ldots,m\} \quad \tilde{A}_i = \begin{bmatrix} A_1^i & A_{12}^i \\ \\ R_i(L) & A_r^i \end{bmatrix}$$

$$\text{où } R_i(L) = A_{21}^i + LA_{11}^i - A_{22}^i L - LA_{12}^i L$$

$$A_1^i = A_{11}^i - A_{12}^i L$$

$$A_r^i = A_{22}^i + LA_{12}^i$$

$$\forall i \in \{1,\ldots,m\} \quad B_r^i = B_2^i + LB_1^i$$

$$C_1 = C_1 - LC_2$$

La matrice-système de (48) est sous forme bloc-triangulaire si :

$$(49)\quad \forall i \in \{0,\ldots,m\} \quad R_i(L) = 0$$

Ce système ne pouvant être réalisé pour tout i, sauf cas particuliers, nous allons résoudre ce système de m+ 1 équations de Riccati au sens des moindres carrés de façon à ce que L satisfasse : $\sum_{i=0}^{m} ||R_i(L)||$ minimal. Introduisons la notation :

$$(50)\quad \forall (k,l) \in \{1,2\}^2, \quad \mathcal{A}_{kl} = \begin{bmatrix} A_{kl}^0 \\ A_{kl}^1 \\ \vdots \\ A_{kl}^m \end{bmatrix}$$

alors le système (49) s'écrit sous la forme :

$$(51)\quad \mathcal{A}_{21} + L \underline{\otimes} [\mathcal{A}_{11} - \mathcal{A}_{12}L] = \mathcal{A}_{22}L$$

où $\underline{\otimes}$ représente le produit de Kronecker de matricespartionnées:

$$(\text{e.g. : } L \underline{\otimes} \mathcal{A}_{11} = \begin{bmatrix} LA_{11}^{0} \\ LA_{11}^{1} \\ \vdots \\ LA_{11}^{m} \end{bmatrix})$$

La façon la plus simple de résoudre ce système est d'utiliser parmi les trois algorithmes présentés en I.1.2.2., celui de Chow-Kokotovic qui conduit à :

$$(52)\quad \begin{cases} k \in \mathbb{N} \\ L_0 = \mathcal{A}_{22}^{+}\mathcal{A}_{21} \\ L_{k+1} = \mathcal{A}_{22}^{+}\left[\mathcal{A}_{21} + L_k \underline{\otimes} [\mathcal{A}_{11} - \mathcal{A}_{12}L_k]\right] \\ \text{où } \mathcal{A}_{22}^{+} = (\mathcal{A}_{22}^{+}\mathcal{A}_{22})^{-1}\mathcal{A}_{22}^{T} \end{cases}$$

Pour assurer la convergence de l'algorithme ci-dessus, on peut montrer (la preuve est analogue à celle utilisée pour (17)) :

Si $\mathcal{A}_{22}$ est de rang plein et que

$$(53)\quad \|\mathcal{A}_{22}^{+}\| < \tfrac{1}{3}(\|\mathcal{A}_0\| + \|\mathcal{A}_{12}\|\,\|L_0\|)^{-1}$$
$$\text{avec } \mathcal{A}_0 = \mathcal{A}_{11} - \mathcal{A}_{12}L_0$$

où la norme choisie est telle que

$$\| \operatorname{diag}_{i=1}^{q} \{Mi\} \| = \sup_{i \in \{1,\ldots,q\}} \| M_i \|$$

alors (52) converge vers une racine réelle $L = L_0 + D$ avec

$$\| D \| < \frac{2\|t_0\| \, \| L_0\|}{\|t_0\| \, \|t_{12}\| \, \|L_0\|}$$

De façon à faciliter les calculs dans (52), on pourra remplacer $t_{22}^{+}$ par $t_{22}^{*} = \frac{1}{m}[A_{22}^{0^{-1}}, A_{22}^{1^{-1}}, \ldots, A_{22}^{m^{-1}}]$ qui est une inverse généralisée à gauche de $t_{22}$, $L_0$ aurait alors pour valeur simple :

$$\frac{1}{m} \sum_{i=0}^{m} A_{22}^{i^{-1}} \, A_{21}^{i}$$

mais $\sum_{i=0}^{m} \|R_i\|$ ne serait plus minimale.

Dans un deuxième temps, l'application de la transformation (4) sur (48) conduit, d'après (26) à :

$$(54) \quad \begin{bmatrix} \overset{\circ}{x}_1 \\ \overset{\circ}{x}_r \end{bmatrix} = \left[ \tilde{A}_0 + \sum_{i=1}^{m} \tilde{A}_i \, u_i \right] \begin{bmatrix} x_1 \\ x_r \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} B_1^i \\ B_r^i \end{bmatrix} u_i$$

$$y = [C_1 \; C_r] \begin{bmatrix} x_1 \\ x_r \end{bmatrix}$$

$$\text{où } \forall i \in \{0,\ldots,m\} \quad \tilde{A}_i = \begin{bmatrix} A_1^i - MR_i(L) & Ri(M) \\ R_i(L) & A_r^i + R_i(L)M \end{bmatrix}$$

$$\text{où } R_i(M) = A_{12}^i + A_1^i M - MA_r^i - MR_i(L)M$$

$$\forall i \in \{1,\ldots,m\} \quad B_1^i = B_1^i - MB_r^i$$

$$C_r = C_2 + MC_1$$

La matrice-système est de nouveau bloc-triangulaire si

(55) $\forall i \in \{0,\ldots,m\} \quad R_i(M) = 0$

qui est un ensemble d'équations de Riccati semblable à (49). Pour résoudre ce système au sens des moindres carrés, on peut introduire les notations supplémentaires :

(56) $$\tilde{\mathcal{A}}_{12} = \begin{bmatrix} A_{12}^{0\,T} \\ \vdots \\ A_{12}^{m\,T} \end{bmatrix}, \tilde{\mathcal{A}}_1 = \begin{bmatrix} A_1^{0T} \\ \vdots \\ A_1^{mT} \end{bmatrix}, \tilde{\mathcal{A}}_r = \begin{bmatrix} A_r^{0T} \\ \vdots \\ A_r^{mT} \end{bmatrix}, \mathcal{R}(L) = \begin{bmatrix} R_0^T(L) \\ \vdots \\ R_m^T(L) \end{bmatrix}$$

(55) s'écrit alors après transposition :

(57) $$\tilde{\mathcal{A}}_{12} + M^T \,\underline{\circledast}\, \left[\tilde{\mathcal{A}}_1 - \mathcal{R}(L)M^T\right] = \tilde{\mathcal{A}}_r M^T$$

Cette écriture permet d'utiliser l'algorithme (53) :

(58) $$\begin{array}{l} k \in \mathbb{N} \\ M_0^T = \tilde{\mathcal{A}}_r^{+}\, \tilde{\mathcal{A}}_{12} \\ M_{k+1}^T = \mathcal{A}_r^{+} \left[\tilde{\mathcal{A}}_{12} + M_k^T \,\underline{\circledast}\, \left[\tilde{\mathcal{A}}_1 - \mathcal{R}(L)M_k^T\right]\right] \end{array}$$

et le résultat concernant la convergence de (53) peut être transposé ici.

Lorsque les matrices M et L vérifient les équations de Riccati (51) et (57), on obtient une matrice d'évolution mieux conditionnée pour mettre en évidence des dynamiques différentes. La forme obtenue ne sera pas toujours quasi-diagonale (Cf. exemple) mais surement quasi-triangulaire pour un écart de dynamiques nettement marqué ; les matrices d'évolution de chaque régime peuvent être approchées par :

* pour le système "lent"

$$\sum_{i=0}^{m} (A_{11}^i - A_{12}^i L - M R_i(L))\, u_i(t) \qquad \text{où } u_0(t) = 1$$

* pour le système "rapide"

$$\sum_{i=0}^{m} (A_{22}^i + L A_{12}^i + R_i(L) M)\, u_i(t)$$

Considérons, par exemple, le système bilinéaire défini par :

$$(59)\quad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \left\{ u_1(t) \begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix} + u_2(t) \begin{bmatrix} 2 & 1 \\ 6 & 10 \end{bmatrix} \right\} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u_2(t)$$

l'application de la première transformation , $\begin{bmatrix} 1 & 0 \\ l & 1 \end{bmatrix}$,

conduit au système d'équations :

$$(60)\quad \begin{bmatrix} 6 \\ 6 \end{bmatrix} + l \otimes \left[ \begin{bmatrix} 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} l \right] = \begin{bmatrix} 8 \\ 10 \end{bmatrix} l$$

dont la solution au sens des moindres carrés est $l = .7$, ce qui donne $r_1(l) = .12$ et $r_2(l) = -.09$

La deuxième transformation $\begin{vmatrix} 1 & -m \\ 0 & 1 \end{vmatrix}$ conduit aux équations

$$(61)\quad \begin{bmatrix} 2 \\ 1 \end{bmatrix} + m \otimes \left[ \begin{bmatrix} -.4 \\ 1.3 \end{bmatrix} - \begin{bmatrix} .12 \\ -.09 \end{bmatrix} m \right] = \begin{bmatrix} 9.4 \\ 10.7 \end{bmatrix} m$$

dont la solution au sens des moindres carrés est $m = .153$, ce qui donne $r_1(m) = .5$ et $r_2(m) = -.44$. Le changement de variable d'état défini par :

$$(62)\quad \begin{bmatrix} x_l \\ x_r \end{bmatrix} = \begin{bmatrix} .89 & -.153 \\ .7 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

conditionne donc le système (59) sous la forme :

$$(63)\quad \begin{bmatrix} \dot{x}_l \\ \dot{x}_r \end{bmatrix} = \left[ u_1(t) \begin{bmatrix} -.42 & .5 \\ .12 & 9.42 \end{bmatrix} + u_2(t) \begin{bmatrix} 1.29 & -.44 \\ -.09 & 10.69 \end{bmatrix} \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} .58 \\ 2.7 \end{bmatrix} u_2$$

Il est à noter ici que l'exemple présenté ne vérifie pas les conditions de convergence des algorithmes (53) et (59) : ce ne sont en effet que des conditions nécessaires.

## II.3 UTILISATION DE METHODES GEOMETRIQUES

La localisation des valeurs propres à l'aide des méthodes géométriques va être utilisée pour les systèmes bilinéaires à deux niveaux :

* au niveau global, pour la détermination des groupements de variables d'état lentes et de variables d'état rapides,

* au niveau local, pour déterminer les variations admissibles sur les variables d'entrée assurant une certaine séparation dans les dynamiques des variables d'état.

### II.3.1 Niveau global

L'utilisation et le tracé des cercles de Gershgorine, scalaire ou par blocs (Cf. I.2), sur chacune des matrices $A_i$, $i = 0,...m$, permet de localiser l'ensemble des valeurs propres. Si, pour toutes les matrices $A_i$, $i = 1,..., m$, le tracé indique une séparation en deux dynamiques et que les variables lentes et les variables rapides sont dans tous les cas les mêmes, alors le système sera globalement à deux dynamiques ; on pourra alors, connaissant les variables d'état lentes et les variables d'état rapides, utiliser la méthode de quasi diagonalisation de façon à mieux conditionner le système.

Si l'on reprend l'exemple présenté dans la partie précédente, on avait :

$$A_1 = \begin{bmatrix} 1 & 2 \\ 6 & 8 \end{bmatrix} \quad \text{et } A_2 = \begin{bmatrix} 2 & 1 \\ 6 & 10 \end{bmatrix}$$

ce qui, après changement de base diagonal, (1 ; 1.732) pour $A_1$ et (1 ; 2.45) pour $A_2$, conduit au tracé

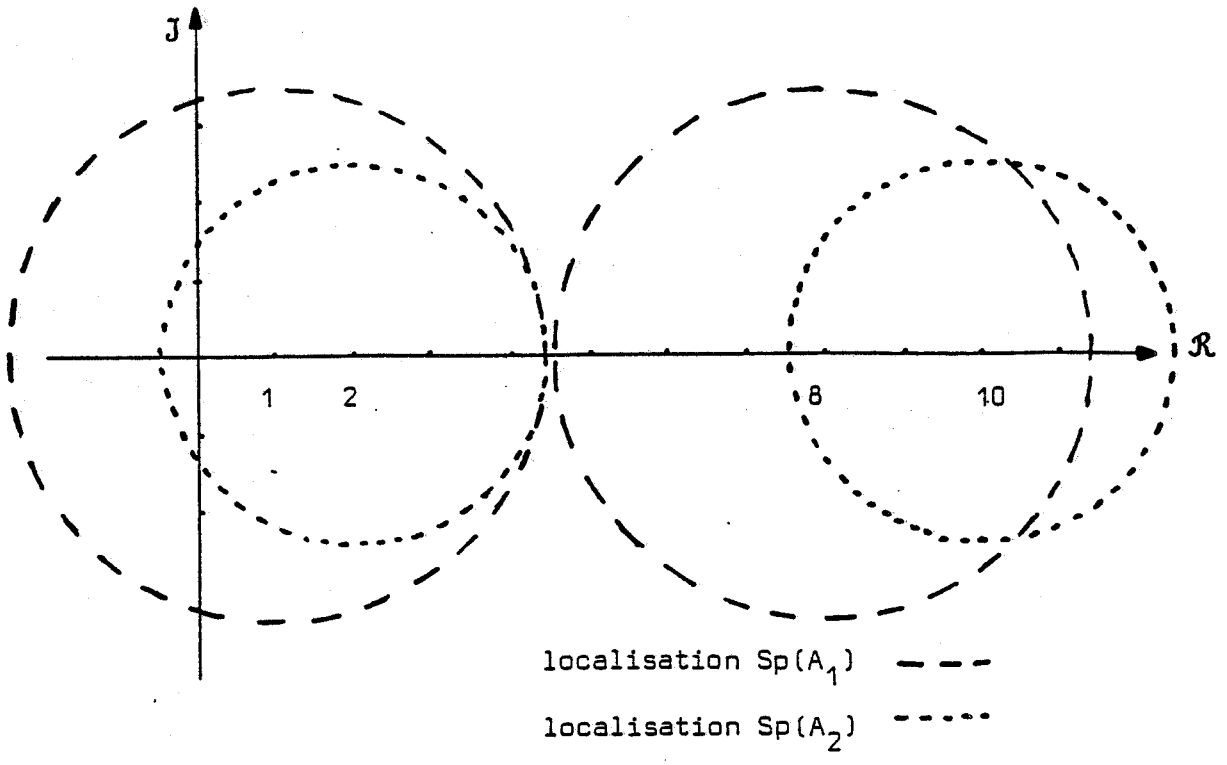



FIGURE 2

ce qui donne les coefficients de séparation de dynamique , $\mu(A_1)$ = .989 et $\mu(A_2)$ = .619, qui sont tous les deux inférieurs à 1, il y a donc bien des dynamiques différentes.

Il est à noter que, si l'on applique les cercles de Gershgorine sur le système (63) obtenu après transformations à partir de (59), on obtient maintenant $\mu'(A_1)$ = .0496 et $\mu'(A_2)$ = .0423, ce qui indique bien que les deux approches géométriques et algébriques sont complémentaires l'une de l'autre, la première permettant de déterminer $q_1$ et $q_2$.

### II.3.3 Niveau local

Lorsque la méthode précédente ne permet pas de trouver des séparations de dynamiques pour toutes les matrices avec les mêmes groupements de variables lentes-rapides, on peut rechercher les domaines de variations des entrées assurant des dynamiques différentes. Pour ceci, on peut utiliser deux méthodes :

- les cercles de Gershgorine dynamiques,

- la méthode du conditionnement.

La séparation de dynamiques sera alors valable dans un intervalle de temps où les contraintes sur les entrées seront vérifiées.

### II.3.2.1. Cercles de Gersghgorine

Si l'on écrit la matrice-système de (38) sous la forme :

$$(64) \qquad A(u) = \{a^0_{kl} + \sum_{i=1}^{m} a^i_{kl}\, u_i(t)\} \quad k,l \in \{1,\ldots,n\}$$

l'utilisation du théorème de Gershgorine (I.2.1) donne une localisation des valeurs propres instantanées de cette matrice. Ces valeurs propres sont situées à l'intérieur d'un ensemble de n cercles définis par :

$$\forall \alpha \in [0,1]$$

$$\text{si } l_k(u) = \sum_{\substack{l=1 \\ l\neq k}}^{q} \left| a^0_{kl} + \sum_{i=1}^{m} a^i_{kl}\, u_i \right|$$

$$(65) \qquad \text{et } c_k(u) = \sum_{\substack{l=1 \\ l\neq k}}^{q} \left| a^0_{lk} + \sum_{i=1}^{m} a^i_{lk}\, u_i \right|$$

$$k=1,\ldots,q \quad \left| z - a^0_{kk} - \sum_{i=1}^{m} a^i_{kk}\, u_i \right| < l_k^{\alpha}(u)\, c_k(u)^{1-\alpha} = R_k(u)$$

et l'on peut donner le résultat suivant :

Théorème 5 : si les cercles de Gershgorine associés à A(u) sont tels que

$$\exists\, \mathcal{U}_i = |\underline{u}_i, \bar{u}_i|,\ \underline{u}_i \leqslant \bar{u}_i\ ,\mathcal{U}_i \subset \mathbb{R}\ ,\ i = 1,\ldots,m$$

$$\exists\ K,L\ ,\ K \cap L = \emptyset\ ,\ K \cup L = \{1,\ldots,n\}$$

$$\forall l \in L,\ \forall k \in K,\ \forall i \in \{1,\ldots,m\}\ ,\ \forall\, u_i \in \mathcal{U}_i$$

$$\left[(a^0_{kk} + \sum_{i=1}^{m} a^i_{kk}\, u_i) - (a^0_{ll} + \sum_{i=1}^{m} a^i_{ll}\, u_i)\right] < R_l(u) + R_k(u)$$

alors dans $\mathcal{U} = \bigoplus_{i=1}^{m} \mathcal{U}_i$ , le système est à deux échelles de temps.

Un indicateur de séparation de dynamiques est alors donné par le paramètre :

$$(66) \qquad \mu = \inf_{u \in \mathcal{U}} \max_{\substack{l \in L \\ k \in K}} \frac{R_l(u) + R_k(u)}{\left| (a^0_{kk} + \sum_{i=1}^{m} a^i_{kk} u_i) - (a^0_{ll} + \sum_{i=1}^{m} a^i_{ll} u_i) \right|}$$

qui est inférieur à 1 s'il y a deux domaines distincts de localisations de valeurs propres. Les bornes inférieures et supérieures des domaines de variations admissibles $\mathcal{U}_i$, i = 1,...,m, seront donc obtenues pour $\mu = 1$.

Par exemple, considérons un système dont la matrice d'évolution est donnéepar :

$$(67) \qquad A(u) = \begin{bmatrix} -1 & 1 \\ 1 & -10 \end{bmatrix} + u \begin{bmatrix} -10 & 1 \\ 1 & -1 \end{bmatrix}$$

Il est évident, pour ce système, que si $u \simeq 0$, la première variable d'état est lente alors que pour $|u| \simeq \infty$ ,elle est devenue rapide. De façon à préciser quantitativement les variations admissibles, utilisons les cercles de Gershgorine dynamiques qui sont donnés par :

$$(68) \qquad \begin{aligned} C_1(u) &= \{-1-10u \; ; \; |1 + u|\} \\ C_2(u) &= \{-10-u \; ; \; |1 + u|\} \end{aligned}$$

En se limitant, pour simplifier l'étude, à $u \geqslant 0$, on obtient les résultats suivants :

- $u \in |0, .6364|$ $x_1$ est lente et $x_2$ est rapide
- $u \in |.6364, 1.5714|$ les deux composantes d'état ont des dynamiques semblables
- $u \in |1.5714, \infty|$ $x_2$ est lente et $x_1$ est rapide

Dans le cas où le nombre des variables d'entrée intervenant

dans la matrice système est important, il est possible de majorer les quantités $l_k(u)$ et $c_k(u)$ par :

$$(69)\quad \begin{cases} l_k(u) \leqslant l_k^0 + \sum_{i=1}^{m} l_k^i \mid u_i \mid = \overline{l_k}(u) \\ c_k(u) \leqslant c_k^0 + \sum_{i=1}^{m} c_k^i \mid u_i \mid = \bar{c}_k(u) \end{cases}$$

où les $c_k^i$ et $l_k^i$, $i = 0,\ldots,m$ sont les rayons des cercles de Gershgorine de chacune des matrices de $A_i$.
$R_k(u)$ est alors majoré par $\bar{c}_k^{1-\alpha}(u)\ \overline{l}_k^{\alpha}(u), \alpha \in [0,1]$.

De plus, lorsqu'il est possible de partitionner les matrices $A_i$ sous la forme particuière

$$(70)\quad A_i = \{A_{kl}^i\}\ (k,l) \in \{1,\ldots,N\}^2,\ i = 0,\ldots,m$$
$$\text{où } \forall i, j \in \{0,\ldots,m\}\ \forall k \in \{1,\ldots,N\}$$
$$A_{kk}^i \text{ diagonalisable et } A_{kk}^i\ A_{kk}^j = A_{kk}^j\ A_{kk}^i$$

on peut employer les résultats donnés par le tracé des cercles de Gershgorine par blocs. En effet, deux matrices qui commutent sont parfois diagonalisables par une même transformation. Ce qui permet après changement de base diagonal par blocs d'utiliser le Théorème 4 (I.2.2) pour obtenir des domaines de localisation de valeurs propres plus fins.

### II.3.2.2. Méthode de conditionnement

L'inconvénient de la méthode précédente réside dans la comparaison entre des cercles dont les centres et les rayons sont fonctions affines des variables d'entrée. Comme nous l'avons déjà mentionné, lorsque m est grand, ce traitement peut s'avérer délicat. De façon à contourner ce problème, nous allons présenter la méthode du conditionnement qui permet de fixer les centres des cercles.

#### a) conditionnement d'une matrice | CIA,82 |

Définitions : 1- Une norme matricielle ||.|| est dite subordonnée si elle est définie à partir d'une norme vectorielle par

$$|| A || = \sup ||Ax||/||x|| \quad x \neq 0$$

2- Le conditionnement d'une matrice A régulière, relativement à une norme matricielle subordonnée ||.||, est la quantité

$$\text{cond}(A) = ||A||\ ||A^{-1}||$$

Dans le cas où la norme choisie est la norme euclidienne, cette quantité est notée $\text{cond}_2(A)$ et possède les propriétés suivantes :

* pour toute matrice A

$$\text{cond}_2(A) = \bar{\mu}(A)/\underline{\mu}(A)$$

où $\bar{\mu}(A)$ et $\underline{\mu}(A)$ désignent respectivement la plus grande et la plus petite des valeurs singulières de A(une valeur singulière est la racine d'une valeur propre de $A^TA$)

* si A est une matrice normale ( $A^TA = AA^T$)

$$(71)\quad \text{cond}_2(A) = \frac{\max_i |\lambda_i(A)|}{\min_i |\lambda_i(A)|}$$

où les nombres $\lambda_i(A)$ sont les valeurs propres de A

b) utilisation

Afin de déterminer les domaines de variations admissibles pour les entrées, nous allons utiliser le théorème de Bauer-Fike | CIA,82 | :

Théorème 6 : Soit A une matrice diagonalisable et P une matrice telles que :

$$P^{-1}AP = \text{diag}\{\lambda_i\}$$

et||.|| une norme matricielle telle que

$$||\text{diag}\{d_i\}|| = \max_i |d_i|$$

Alors pour toute matrice $\Delta A$,

$$Sp(A+\Delta A) \subset \bigcup_{i=1}^{n} D_i \text{ où } D_i = \{ z \in \mathbb{C} \ ; \ |z - \lambda_i| < cond(P) \, \|\Delta A\| \}$$

Dans le cas où A n'est pas diagonalisable, on peut également montrer le résultat plus général suivant :

Théorème 7: Soit A une matrice carrée, J une de ses formes de Jordan associée et P une matrice régulière telles que :

$$P^{-1}AP = J = D + H \text{ où } D = \text{diag } \{\lambda i\} \quad \lambda_i \in Sp(A)$$

Soit $\|.\|$ une norme matricielle subordonnée vérifiant la condition du théorème précédent

Alors pour toute matrice $\Delta A$,

$$Sp(A + \Delta A) \subset \bigcup_{i=1}^{n} Di$$

$$\text{où } Di = \{z \in \mathbb{C} \ ; \ | \ z - \lambda i| \ < \varepsilon + cond(P) \ \|\Delta A\| \}$$

$\varepsilon = 0$ si A diagonalisable, $\varepsilon = 1$ sinon

Démonstration : soit $\lambda$ une valeur propre de $A + \Delta A$, si $\lambda$ est aussi valeur propre de A, le résultat est évident ; sinon $(D - \lambda I)$ est régulière, or $(A + \Delta A - \lambda I)$ étant singulière, l'écriture

$$(72) \qquad P^{-1}(A + \Delta A - \lambda I)P = J - \lambda I + P^{-1}\Delta AP$$
$$= (D - \lambda I) \left\{ I + (D - \lambda I)^{-1} (H + P^{-1}\Delta AP) \right\}$$

indique que $I + (D - I)^{-1}(H + P^{-1} AP)$ est singulière, par conséquent :

$$(73) \qquad 1 \leqslant \|(D-\lambda I)^{-1}(H+P^{-1}\Delta AP)\| < \|(D- I)^{-1}\| \{\|H\| + \| P^{-1}\| \|\Delta A\| \ \|P\| \}$$

or $\|H\| = 1$ si $H \neq 0$

et $\| (D-\lambda I)^{-1} \| = \dfrac{1}{\min\limits_i |\lambda - \lambda_i|}$ #

Remarques : * les normes matricielles eucliciennes , $\max\limits_i \sum\limits_j |a_{ij}|$

et $\max_i \sum_j |a_{ij}|$ vérifient les hypothèses des théorèmes ci-dessus.

* de façon à minimiser les rayons des cercles de localisation des valeurs propres à l'aide de ces deux théorèmes, on peut remplacer cond(P) par la quantité :

$$\Gamma(A) = \inf \{ \text{cond}(P) \ ; \ P^{-1}AP = \text{diag}\{\lambda_i\} + H \}$$

* dans le cas particulier où A est une matrice normale donc diagonalisable par une matrice unitaire, on a :

$$\Gamma_2(A) = 1$$

donc

$$(74) \qquad Sp(A + \Delta A) \subset \bigcup_{i=1}^{n} \{ z \in \mathbb{C} \ ; \ |z - \lambda i| < \|\Delta A\|_2 \}$$

En résumé, la méthode que nous proposons ici, consiste à :

i) sélectionner, dans les matrices constantes déterminant le système bilinéaire (38) une matrice particulière $A_k$,

ii) déterminer l'ensemble des valeurs et vecteurs propres donc $Sp(A_k)$ et $P_k$ telle que $P_k^{-1}A_kP_k = J_k$

iii) appliquer le théorème précédent sur $A_k$ avec la matrice de perturbation :

$$\Delta A_k = u_k^{-1} \sum_{\substack{i=0 \\ i \neq k}}^{m} A_i u_i \quad \text{avec } u_0 = 1$$

On remarquera que l'on a intérêt à sélectionner d'abord des matrices normales et/ou ayant des valeurs propres très différentes.

Par exemple, si dans (38), on sélectionne $A_0$, cela donne l'ensemble de localisation des valeurs propres comme réunion des disques :

$$(75) \qquad i = \{1, \ldots, n\} \quad Di = \{z \in \mathbb{C} \ ; \ |z - \lambda_i^0| < \varepsilon + \text{cond}(P) \left( \sum_{j=1}^{m} |u_i| \, \|A_i\| \right) \}$$

avec $\lambda_i^0 \in S_p(A_0)$ et $\varepsilon = 0$ si $A_0$ diagonalisable, 1 sinon .

Les rayons de ces cercles sont tous égaux ; on a donc un système à deux échelles de temps si l'on a :

$$(76)\quad \exists K,L \quad K \cup L = \{1,\ldots, n\}\ ,\ K \cap L = \emptyset$$
$$\forall k \in K, \forall l \in L \quad \sum_{j=1}^{m} |u_i|\ \|A_i\| < \frac{1}{\text{cond}(P)} \left[ \frac{|\lambda_k^0 - \lambda_l^0|}{2} - \varepsilon \right]$$

Appliquons ceci sur le système défini par (67) ; les matrices $A_0$ et $A_1$ sont symétriques, donc normales et telles que :

$$(77)\quad \begin{aligned} &Sp(A_0) = Sp(A_1) = \{-.89, -10.1098\} \\ &\|A_0\|_2 = \|A_1\|_2 = 10.1098 \end{aligned}$$

l'utilisation de (74) donne les cercles :

* si $A_0$ est privilégiée : de centres -.89, -10.1098 et de rayons $10.1098|u|$ ; il y a donc séparation de dynamiques si $|u| < .456$ la première variable d'état est alors lente.

* si $A_1$ est privilégiée : à partir des mêmes centres, les rayons sont cette fois-ci de $10.1098\ |u|^{-1}$ ; la séparation de dynamiques est effective pour $|u| > 2.193$ et la première variable d'état est dans ce cas rapide.

* pour $|u| \in\ ]\,.456,\ 2.193\,[$, les dynamiques des variables d'état sont équivalentes.

Les résultats donnés par cette méthode semblent moins fins que ceux obtenus par les cercles de Gershgorine, mais le traitement numérique est plus facile et plus rapide (voir, par exemple, le cas $u < 0$ dans (67)).

# CONCLUSION

La mise en évidence de dynamiques différentes dans le cas des systèmes bilinéaires demande des méthodes plus élaborées que dans le cas des systèmes linéaires, mais qui en sont issues. Nous nous sommes restreints dans notre étude car nous n'avons étudié que le comportement "linéarisé" des valeurs propres du système bilinéaire.

Il est, en effet, notoire qu'un système non stationnaire peut être asymptotiquement stable alors qu'à chaque instant, le système linéaire associé est instable ou réciproquement | WIL,70 |. Cependant, pour une telle étude, nous pensons que les outils à utiliser sont encore peu développés pour être utilisables pratiquement. La solution du problème de Floquet, par exemple, n'a toujours pas été trouvée de façon explicite.

C'est en ayant fait cette restriction que nous abordons les chapitres suivants où la mise en évidence des petits paramètres est supposée effective.

# CHAPITRE IV

# PERTURBATIONS SINGULIERES ET SYSTEMES BILINEAIRES

# INTRODUCTION

L'utilisation de la méthode des perturbations singulières pour la réduction de systèmes bilinéaires peut être envisagée lorsqu'on a des variables de dynamiques différentes. Il est à noter que certains systèmes se présentent directement sous forme bilinéaire singulièrement perturbée |MOH,70|. On peut citer, par exemple, le cas typique du système de thermorégulation humaine dont l'équation d'état est de la forme :

$$\begin{bmatrix} c_{m_1} \dot{x}_1 \\ c_{m_2} \dot{x}_2 \end{bmatrix} = \left\{ \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} u \right\} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \theta_a \end{bmatrix} u + \begin{bmatrix} q \\ 0 \end{bmatrix}$$

où $x_1$ est la température moyenne du corps
$x_2$ est la température de la peau
$c_{m_1}$ est la capacité thermique du corps
$c_{m_2}$ est la capacité thermique de la peau
$\theta_a$ est la température ambiante
q est une source indépendante de chaleur (le métabolisme)
u est la conductance thermique qui sert de variable de commande

Comme dans ce schéma, on a $c_{m_2} \gg c_{m_1}$, le système est bien à deux échelles de temps.

Mais, comme nous allons le montrer, d'une part, la méthode des perturbations singulières doit être appliquée avec quelques précautions, hors desquelles les modèles réduits obtenus perdent leur validité ; d'autre part, le modèle réduit lent n'est plus bilinéaire. Cependant, cette méthode a été utilisée pour générer des commandes simplifiées de systèmes bilinéaires |ARM,82 ; ARM,86|.

Sans aborder dans ce chapitre les problèmes de commandes quasioptimales, nous allons étudier tout d'abord l'approximation du système réduit lent par un système bilinéaire. Puis, par la méthode de découplage, on déterminera un modèle réduit rapide également bilinéaire . Enfin, pour améliorer la précision de la modélisation à l'intérieur de la couche limite, on généralisera aux systèmes bilinéaires la notion de transformation réciproque utilisée dans le cas des systèmes linéaires.

## I. SYSTEMES BILINEAIRES SINGULIEREMENT PERTURBES

Considérons un système bilinéaire mis sous la forme singulièrement perturbée (1). Le problème de mise en évidence du petit paramètre étant supposé résolu à l'aide d'une des méthodes présentées dans le chapitre précédent :

$$(1) \qquad \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{m} u_i(t)\, A_i \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ \hline B_2 \end{bmatrix} u(t)$$

$$y = \begin{bmatrix} C_1 \vdots C_2 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \quad ; \; x(0) = x_0 \; ; \; z(0) = z_0$$

$$u_0(t) = 1 \; ; \; u(t) = \left[ u_1(t), \ldots, u_m(t) \right]^T \; , \; y \in \mathbb{R}^n$$

$$B_1 \in \mathbb{R}^{q_1 \times m} \; , \; B_2 \in \mathbb{R}^{q_2 \times m} \; , \; C_1 \in \mathbb{R}^{n \times q_1} \; , \; C_2 \in \mathbb{R}^{n \times q_2}$$

$$\forall\, i \in \{0, \ldots, m\} \qquad A_i = \left[ \begin{array}{c:c} A_{11}^i & A_{12}^i \\ \hline A_{21}^i & A_{22}^i \end{array} \right] ;$$

$$k,l \in \{1,2\}^2 \quad A_{kl}^i \in \mathbb{R}^{q_k \times q_l} \; ; \quad 0 < \varepsilon \ll 1 \; .$$

En supposant que ce système vérifie l'hypothèse suivante :

Hypothèse : $\forall \; t \in \mathbb{R}^+$

- $\sum_{i=0}^{m} u_i(t) A_{22}^i$ est régulière

- les valeurs propres de cette matrice sont constamment plus petites qu'un nombre négatif fixé,

le comportement du système (1), à l'aide du théorème de Tihonov, peut être approché par un modèle réduit lent et un système de couche limite |HOP,71|. Si les hypothèses précédentes ne sont vérifiées qu'à l'intérieur d'un intervalle de temps $[T_1, T_2]$ de $\mathbb{R}^+$, cette approximation ne sera valable que dans cet intervalle de temps.

On a donc soulevé ici un des premiers problèmes de la réduction de dimensionnalité des systèmes bilinéaires par l'application de la méthode des perturbations singulières. En effet, que $\varepsilon$ ait été mis en évidence par une méthode locale ou globale (Cf. chapitre précédent), il faudra ajouter des contraintes supplémentaires sur les variables d'entrée pour que soit vérifiée l'hypothèse du théorème de Tihonov. Ces contraintes supplémentaires pouvant être déterminées par une des méthodes globales (cercles de Gershgorine dynamiques ou conditionnement, Cf. III : II.3.2.).

## I.1. MODELE REDUIT LENT

Si l'on pose $\varepsilon = 0$, dans (1), on obtient l'équation algébrique d'entraînement :

$$(2) \qquad \left[\sum_{i=o}^{m} \bar{u}_i(t)A_{21}^i\right]\bar{x} + \left[\sum_{i=o}^{m} \bar{u}_i(t)A_{22}^i\right]\bar{z} + B_2\bar{u}(t) = o$$

ce qui, si le système vérifie l'hypothèse précédente, permet d'obtenir l'équation d'état du modèle réduit lent :

$$(3) \qquad \dot{\bar{x}} = \left\{\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{11}^i\right] - \left[\sum_{i=o}^{m} \bar{u}_i(t)A_{12}^i\right]\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{22}^i\right]^{-1}\right.$$

$$\left.\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{21}^i\right]\right\}\bar{x} + \left\{B_1 - \left[\sum_{i=o}^{m} \bar{u}_i(t)A_{12}^i\right]\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{22}^i\right]^{-1}\right.$$

$$\left.B_2\right\}\bar{u}(t) \; ; \; \bar{x}(o) = x(o)$$

Si l'on ne s'intéresse qu'au comportement lent du système, on a également l'équation de sortie lente

$$(4) \qquad \bar{y} = \left\{C_1 - C_2\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{22}^i\right]^{-1}\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{21}\right]\right\}\bar{x}_1$$

$$- C_2\left[\sum_{i=o}^{m} \bar{u}_i(t)A_{22}^i\right]^{-1}B_2\bar{u}(t)$$

Bien que ces formules soient analogues à celles du cas des systèmes linéaires, elles impliquent ici une perte de la propriété de bilinéarité pour le modèle réduit lent et ceci à tous les niveaux de la représentation d'état . En effet, le système (3-4) est toujours linéaire par rapport à l'état, mais est non-linéaire par rapport aux variables d'entrée. Ceci est le problème majeur de l'application des perturbations singulières dans le cas des systèmes bilinéaires généraux. Cependant, on peut dégager quelques classes de ces systèmes où ce problème n'existe pas (i.e. le modèle réduit lent est réellement bilinéaire) ; dans les autres cas, nous présenterons quelques procédures d'approximations pour remplacer (3-4) par un système bilinéaire de même ordre.

## I.2. CAS PARTICULIER DE MODELES REDUITS LENTS BILINEAIRES

### I.2.1. Système régulier à entrée scalaire

Ce système a été étudié dans |TZA,84|, qui en a déterminé des retours stabilisants et il se présente sous la forme :

$$(5) \qquad \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = u(t) \left\{ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \right\}$$

Faire $\varepsilon = 0$, dans (5), revient à avoir, si $A_{22}$ régulière :

$$(6) \qquad \bar{z} = -A_{22}^{-1} \{A_{21}\bar{x} + B_2\}$$

Donc, le modèle réduit lent s'écrit sous la forme régulière :

$$(7) \qquad \dot{\bar{x}} = u(t) \left\{ \left[ A_{11} - A_{12}A_{22}^{-1}A_{21} \right] \bar{x} + \left[ B_1 - A_{12}A_{22}^{-1}B_2 \right] \right\}$$

### I.2.2. Systèmes commutatifs

Les systèmes bilinéaires commutatifs sont définis, à partir de (1), par :

$$(8) \qquad \forall\, i,j \in \{0,\ldots,m\}^2 \quad A_i \begin{bmatrix} I_{q_1} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \frac{1}{\varepsilon} I_{q_2} \end{bmatrix} A_j = A_j \begin{bmatrix} I_{q_1} & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \frac{1}{\varepsilon} I_{q_2} \end{bmatrix} A_i$$

Bien que cette propriété de commutativité soit relativement forte pour un système quelconque |KRE,78|, il existe des systèmes physiques décrits par de tels modèles |WEI,78|. Dans certains cas, |GAN,60|, par une même transformation, on peut ramener chacune des matrices $A_i$, i = 0,...,m, à une forme bloc-diagonalisée, soit la transformation :

$$(9) \qquad \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix}$$

telle que (1) devienne :

$$(10) \qquad \begin{bmatrix} v \\ \varepsilon \dot{w} \end{bmatrix} = \left\{ \sum_{i=o}^{m} u_i(t) \begin{bmatrix} J_1^i & 0 \\ 0 & J_2^i \end{bmatrix} \right\} \begin{bmatrix} v \\ w \end{bmatrix} + \begin{bmatrix} B_1' \\ B_2' \end{bmatrix} u$$

Les variables lentes de ce système sont regroupées dans v ; les systèmes lents et rapides étant parfaitement découplés, faire $\varepsilon = 0$ dans (10) n'a aucun sens. La partie lente dans x, et z, sera donc représentée par v et la partie rapide par w. Il vient alors le modèle réduit lent :

$$(11) \qquad \begin{cases} \dot{x}_1 = P_1 \left( \sum_{i=o}^{m} u_i(t) J_1^i \right) P_1^{-1} x_1 + P_1 B_1' u \\ z_1 = P_3 P_1^{-1} x_1 \end{cases}$$

qui est bilinéaire et le modèle réduit rapide :

$$(12) \qquad \begin{cases} \varepsilon \dot{z}_r = P_4 \left( \sum_{i=o}^{m} u_i(t) J_2^i \right) P_4^{-1} z_r + P_4 B_2' u \\ x_r = P_2 P_4^{-1} z_r \end{cases}$$

### I.2.3. Conditions de bilinearité

Ces conditions imposées sur les matrices $A_i$, donc sur la structure du système, ont été étudiées dans |GUI,81 ; ARM,79|. De façon à assurer pour le système réduit lent (3) une matrice d'évolution

linéaire en u et un vecteur de commande constant, on a des groupes de conditions, comme par exemple :

a) $\forall\, i \in \{1,\ldots,m\}.\ A_{22}^{i} = o$

$\quad . \forall\, j \in \{1,\ldots,m\}\ \ A_{12}^{i} A_{22}^{o-1} A_{21}^{j} = o$

$\quad . A_{12}^{i} A_{22}^{o-1} = o$

(3) donne alors :

$$(3\text{-a})\ \dot{\bar{x}} = \left[(A_{11}^{o} - A_{12}^{o}A_{22}^{o-1}A_{21}^{o}) + \sum_{i=1}^{m} u_i(t)(A_{12}^{o}A_{22}^{o-1}A_{21}^{i} + A_{12}^{i}A_{22}^{o-1}A_{21}^{o})\right]\bar{x} + (B_1 - A_{12}^{o}A_{22}^{o-1}B_2)\bar{u}(t)$$

b) $\exists\, M,\ \forall\, i \in \{0,\ldots,m\}\ : A_{12}^{i} = MA_{22}^{i}$

(3) donne dans ce cas :

$$(3\text{-b})\ \dot{\bar{x}} = \left[\sum_{i=o}^{m} u_i(t)\ (A_{11}^{i} - MA_{21}^{i})\right]\bar{x} + (B_1 - MB_2)\bar{u}(t)$$

Il existe d'autres conditions similaires mais, de façon générale, ces conditions sont relativement fortes quant à la structure du système. On peut cependant essayer de s'y ramener par transformation linéaire.

## I.3 SYSTEME COUCHE LIMITE

Dans le domaine de couche limite, le système rapide est défini à partir de (1) écrite dans l'échelle des temps rapides : $\tau = t/\varepsilon$. En posant $\varepsilon = 0$, on obtient (Cf II : II.1.1.)

$$(13)\qquad \tilde{x}(\tau) = \text{cte (à l'échelle des } \tau)$$

Comme d'après le théorème de Tihonov,

$$(14)\qquad \forall\, t \in \mathbb{R}^{+}\ x(t) = \bar{x}(t) + o(\varepsilon)$$

il est naturel de prendre $\tilde{x}(\tau) = 0$, ce qui conduit au système intérieur

$$(15)\qquad \begin{cases} \dfrac{d\tilde{z}}{d\tau} = \left(\sum_{i=o}^{m} \tilde{u}_i(\tau)\ A_{22}^{i}\right)\tilde{z} + B_2\tilde{u}(\tau) \\ \tilde{y} = C_2\tilde{z} \end{cases}$$

Le système rapide conserve donc la propriété de bilinéarité.

Et comme l'on doit avoir :

(16) $\quad z(t) = \bar{z}(t) + \tilde{z}(\tau) + o(\varepsilon)$ (où $\bar{z}(t)$ est défini par (2))

il est nécessaire d'imposer :

$$(17) \quad \begin{cases} u(t) = \bar{u}(t) + \tilde{u}(t/\varepsilon) \\ \tilde{z}(o) = z(o) - \left[\sum_{i=o}^{m} A_{22}^{i}\, \bar{u}_i(o)\right]^{-1} \\ \qquad \left\{\left[\sum_{i=o}^{m} A_{21}^{i} \bar{u}_i(o)\right] x(o) + B_2 \bar{u}(o)\right\} \end{cases}$$

La difficulté d'utilisation de ce schéma, naturel pour résoudre le problème de commande optimale |FOS,83|, réside dans la définition précise de $\bar{u}(t)$ et $\tilde{u}(\tau)$, donc à l'instant initial de $\tilde{z}(o)$.

## I.4. DEVELOPPEMENT ASYMPTOTIQUE DE LA SERIE GENERATRICE

En reprenant l'argument développé dans le cas des systèmes linéaires dans (II : II.2.3), nous allons montrer que l'on peut obtenir un développement analytique en $\varepsilon$ de la série génératrice du système initial, donc des solutions du système.

En utilisant la notation définie en (II : 59-60), on obtient à partir de (1), l'ensemble de sous-systèmes :

$$(18) \quad \begin{cases} (\Sigma_o) \begin{cases} x'(o) = o \\ z'(o) = \sum_{j=o}^{m} u_j \{A_{21}^{j} x(o) + A_{22}^{j} z(o)\} + B_2 u \end{cases} \\ (\Sigma_1) \begin{cases} x'(1) = \sum_{j=o}^{m} u_j \{A_{11}^{j} x(o) + A_{12}^{j} z(o)\} + B_1 u \\ z'(1) = \sum_{j=o}^{m} u_j \{A_{21}^{j} x(1) + A_{22}^{j} z(1)\} \end{cases} \\ (\Sigma_i) \quad (i>2) \begin{cases} x'(i) = \sum_{j=o}^{m} u_j \{A_{11}^{j} x(i-1) + A_{12}^{j} z(i-1)\} \\ z'(i) = \sum_{j=o}^{m} u_j \{A_{21}^{j} x(i) + A_{22}^{j} z(i)\} \end{cases} \end{cases}$$

où ' représente la dérivation par rapport à $\tau(=\frac{t}{\varepsilon})$ et

$$x = \sum_{i=o}^{\infty} \varepsilon^i x_{(i)} \; ; \; z = \sum_{i=o}^{\infty} \varepsilon^i z_{(i)}$$

Ces sous-systèmes sont organisés en cascade suivant le schéma (II : figure 3) et la sortie du système initial s'écrit :

$$(19) \qquad y = \sum_{i=o}^{\infty} \varepsilon^i y_{(i)}$$

où $\forall\, i \in \mathbb{N}$ $y_{(i)} = C_1 x_{(i)} + C_2 z_{(i)}$

Le but de ce paragraphe étant de trouver une écriture analytique de la fonction de transfert codée du système initial, en supposant les conditions initiales nulles, il vient, à l'aide du calcul symbolique et des notations présentées au chapitre I, le codage de chacun des sous-systèmes :

$$(20) \quad \begin{cases} (\Sigma_o) & \begin{cases} X_o = o \\ Z_o = \sum_{j=o}^{m} \alpha_j \{A_{22}^j Z_o\} + \sum_{j=1}^{m} \alpha_j B_2^j \end{cases} \\ (\Sigma_1) & \begin{cases} X_1 = \sum_{j=o}^{m} \alpha_j \{A_{11}^j X_o + A_{12}^j Z_o\} + \sum_{j=1}^{m} \alpha_j B_1^j \\ Z_1 = \sum_{j=o}^{m} \alpha_j \{A_{21}^j X_1 + A_{22}^j X_1 + A_{22}^j Z_1\} \end{cases} \\ (\Sigma_i) \; (i>2) & \begin{cases} X_i = \sum_{j=o}^{m} \alpha_j \{A_{11}^j X_{i-1} + A_{12}^j Z_{i-1}\} \\ Z_i = \sum_{j=o}^{m} \alpha_j \{A_{21}^j X_i + A_{22}^j Z_i\} \end{cases} \end{cases}$$

où Xi (resp.Zi) représente la série génératrice associée à la variable $x_{(i)}$ (resp $z_{(i)}$).

Il vient les expressions analogues à (II : 64) :

$$(21)\quad \begin{cases} \begin{bmatrix} X_o \\ Z_o \end{bmatrix} = \begin{bmatrix} o \\ \Delta_{22}B_2 \end{bmatrix} \alpha = G_o\alpha \\ \begin{bmatrix} X_1 \\ Z_1 \end{bmatrix} = \begin{bmatrix} I_{q_1} \\ \Delta_{22}\Delta_{21} \end{bmatrix} \bar{B}\alpha = G_1\alpha \\ i>2, \quad \begin{bmatrix} X_i \\ Z_i \end{bmatrix} = \underbrace{\begin{bmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{22}\Delta_{21}\Delta_{11} & \Delta_{22}\Delta_{21}\Delta_{12} \end{bmatrix}}_{M} \begin{bmatrix} X_{i-1} \\ Z_{i-1} \end{bmatrix} \end{cases}$$

$$\text{où} \begin{cases} \alpha = [\alpha_1, \ldots, \alpha_m]^T \\ \Delta_{22} = \{I_{q_2} - \sum_{j=o}^{m} \alpha_j A_{22}^j\}^{-1} \\ (k,l) \in \{(1,1), (1,2), (2,1)\} \\ \Delta_{kl} = \sum_{j=o}^{m} \alpha_j A_{kl}^j \\ \bar{B} = B_1 + \Delta_{12}\,\Delta_{22}\,B_2 \end{cases}$$

qui conduisent à la série génératrice donnant les variables $x(i)$, $z(i)$ et $y(i)$ à un ordre quelconque :

$$(22)\qquad \forall\, i \geqslant 1 \quad \begin{bmatrix} X_i \\ Z_i \end{bmatrix} = M^{i-1} G_1\alpha$$

$$Y_i = \begin{bmatrix} C_1 \vdots C_2 \end{bmatrix} \begin{bmatrix} X_i \\ Z_i \end{bmatrix} = C \begin{bmatrix} X_i \\ Z_i \end{bmatrix}$$

$Y_i$ étant la série formelle associée à $y_i$.

Soit Y la série génératrice associée à la sortie du système, celle-ci s'écrit suivant (19) :

$$(23) \qquad Y = \left[ CG_o + \varepsilon C \underbrace{\{ \sum_{i=o}^{\infty} \varepsilon^i M^i \}}_{[I_q - \varepsilon M]^{-1}} G_1 \right] \alpha$$

Or, d'après (21), on a :

$$(24) \qquad \left[ I_q - \varepsilon M \right]^{-1} = \{ I_q + \begin{bmatrix} I_{q_1} \\ \Delta_{22}\Delta_{21} \end{bmatrix} (-\varepsilon) \begin{bmatrix} \Delta_{11} & \Delta_{12} \end{bmatrix} \}^{-1}$$

où $\Delta_1 = \Delta_{11} + \Delta_{12}\Delta_{22}\Delta_{21}$

ce qui, comme dans le cas linéaire, nous conduit à l'expression formelle :

$$(25) \qquad \left[ I_q - \varepsilon M \right]^{-1} = I_q + \varepsilon \begin{bmatrix} I_{q_1} \\ \Delta_{22}\Delta_{21} \end{bmatrix} \left[ I_{q_1} - \varepsilon\Delta_1 \right]^{-1} \begin{bmatrix} \Delta_{11}\Delta_{12} \end{bmatrix}$$

où $\Delta_1 = \Delta_{11} + \Delta_{12}\Delta_{22}\Delta_{21}$

On obtient finalement :

$$(26) \qquad Y = C_2\Delta_{22} B_2 + \varepsilon\bar{C} (I_{q_1} - \varepsilon\Delta_1)^{-1}\bar{B}$$

où $\bar{C} = C_1 + C_2\Delta_{22}\Delta_{21}$

qui est une expression analogue à (II : 69) et qui définit la série génératrice Y comme somme de deux séries génératrices : la série génératrice de la sortie du système bilinéaire rapide et la série associée à la sortie du système lent, qui, comme nous l'avons déjà indiqué, n'est pas bilinéaire. Cependant, cette série est calculable par la formule (26) ; il en vient ainsi les premiers termes :

$$(27) \qquad \bar{C} (I_{q_1} - \varepsilon\Delta_1)^{-1}\bar{B} = y^{(o)}_o + \sum_{j=o}^{m} y^{(1)}_j \alpha_j + \ldots$$

où $y^{(o)}_o = C_1B_1$

$\forall\, j \in \{0,\ldots,m\} \quad y^{(1)}_j = C_1A_{12}^jB_2 + C_2A_{21}^jB_1 + \varepsilon C_1A_{11}^jB_1$

On peut alors, par la méthode proposée par |HES,86|, construire à partir de cette série, un modèle bilinéaire réduit lent d'ordre $q_1$ qui approchera le système défini par la série (27). Nous allons, dans la partie suivante, proposer d'autres méthodes pour déterminer un modèle réduit lent bilinéaire.

## II. BILINEARISATION DU MODELE REDUIT LENT

Nous venons de voir que le problème majeur de l'utilisation des perturbations singulières sur les systèmes bilinéaires concerne la perte de cette propriété sur le modèle réduit lent obtenu. Nous présentons, dans cette partie, quelques procédures permettant de résoudre, de façon approchée, cette question. Chacune de ces approches sera utilisée suivant le contexte d'emploi ou d'analyse du système étudié.

### II.1. DOMINANCE D'ENTREES

Lorsque les domaines de variations des variables d'entrée sont connus, on peut utiliser la notion d'entrée dominante pour construire un modèle approché |ROT,85|.

Définition : La variable d'entrée $u_i(t)$, $i \in \{0,\ldots,m\}$ est dominante sur $[T_1, T_2] \subset \mathbb{R}^+$ si
$\forall\, t \in [T_1, T_2]\ \forall\, k \in \{0,\ldots,m\}$, $k \neq i$, $\dfrac{|u_k(t)|}{|u_i(t)|} \ll 1$

Le principe de l'approximation est, lorsqu'une entrée est dominante sur un intervalle de temps, d'effectuer un développement limité, autour de cette valeur dominante, des quantités non linéaires intervenant dans (3).

Soit $\bar{u}_k$, $k \in \{0,\ldots,m\}$, une entrée dominante sur $[T_1, T_2]$, alors il vient :

$$
\begin{cases}
\text{a) } B_1 - \left[\sum_{i=0}^{m} \overline{u_i}(t) A_{12}^{i}\right]\left[\sum_{i=0}^{m} \overline{u_i}(t) A_{22}^{i}\right]^{-1} B_2 \simeq B_1 - A_{12}^{k} A_{22}^{k\,-1} B_2 = B_1^k \\
\text{b) } \left[\sum_{i=0}^{m} \overline{u_i}(t) A_{12}^{i}\right]\left[\sum_{i=0}^{m} \overline{u_i}(t) A_{22}^{i}\right]^{-1}\left[\sum_{i=0}^{m} \overline{u_i}(t) A_{21}^{i}\right] \simeq A_{12}^{k} A_{22}^{k\,-1} A_{21}^{k} \overline{u_k}(t) \\
\qquad + \sum_{\substack{i=0 \\ i\neq k}}^{m} (A_{12}^{k} A_{22}^{k\,-1} A_{21}^{i} + A_{12}^{i} A_{22}^{k\,-1} A_{21}^{k}) \overline{u_i}(t) = \sum_{i=0}^{m} A_{11}^{k} \overline{u_i}(t)
\end{cases}
\tag{28}
$$

et le modèle obtenu est alors bilinéaire sur $[T_1, T_2]$. Cependant, à l'aide de (28), on dispose d'un ensemble de modèles valables seulement sur un certain intervalle de temps. De façon à éviter les problèmes liés à cette représentation multi-modèle, on construit un modèle approché moyen.

Supposons que l'on connaisse les probabilités de dominance des entrées sur toute l'évolution du système ; soit $p_k$ la probabilité de dominance de l'entrée $u_k(t)$ :

$$(29) \qquad \forall\, k \in \{0,\ldots,m\} \quad p_k \in [0,1] \; ; \; \sum_{k=0}^{m} p_k = 1$$

on construit alors le modèle approché lent défini par :

$$(30) \quad \begin{cases} \bar{\dot{x}} = \{ \sum_{i=0}^{m} u_i(t) \left[ A_{11}^{i} - M_1^i \right] \} \; \bar{x} + B_1 \bar{u}(t) \\ \bar{x}(o) = x\,(o) \\ \text{où } \forall\, i \in \{0,\ldots,m\} \quad M_1^i = \sum_{k=0}^{m} p_k A_{11}^{k} \\ \text{et } B_1 = \sum_{k=0}^{m} p_k B_1^k \end{cases}$$

L'application de cette technique peut conduire à la construction immédiate d'un modèle réduit lent approché bilinéaire dans le cas où les consignes sont constituées de petites variations autour d'une valeur constante. Soit

$$(31) \qquad \forall\, i \in \{1,\ldots,m\} \quad u_i(t) = U_i + \varepsilon_i(t) \text{ où } \left[ \frac{\varepsilon_i(t)}{U_i} \right] \ll 1 \; \forall t \in \mathbb{R}^+$$

La matrice d'évolution de (1) est alors réécrite sous la forme :

$$(32) \qquad A'_o + \sum_{i=1}^{m} \varepsilon_i(t) A_i$$

$$\text{avec } A'_o = A_o + \sum_{i=1}^{m} U_i A_i = \begin{bmatrix} A'_{11} & \vdots & A'_{12} \\ \cdots & \cdots & \cdots \\ A'_{21} & \vdots & A'_{22} \end{bmatrix}$$

l'entrée $u_0$ est alors dominante par rapport aux $\varepsilon_i$ ; le modèle réduit peut être approché par :

$$(33)\qquad \dot{\bar{x}} = \left\{\left[A'_{11} - A'_{12}A'^{-1}_{22}A'_{21}\right] + \sum_{i=1}^{m} \varepsilon_i(t)\left[A^i_{11} - A'_{12}A'^{-1}_{22}A^i_{21} - A^i_{12}A'^{-1}_{22}A'_{21} + A'_{12}A'^{-1}_{22}A^i_{22}A'^{-1}_{22}A'_{21}\right]\right\}\bar{x}$$
$$+\left[B_1 - A'_{12}A'^{-1}_{22}B_2\right]\bar{u}(t)$$

## II.2. METHODE DIRECTE

Dans le cas où l'on ne possède aucun renseignement sur les variables d'entrée, la procédure d'approximation directe consiste à remplacer au sens des moindres carrés les matrices d'évolution et de commande de (3) par des quantités respectivement affines en les entrées et constantes |ROT,86b|. Cette procédure peut être menée en plusieurs étapes suivant le compromis "temps de calcul-précision du modèle" souhaité.

### II.2.1 Approximation de la matrice de commande

Il s'agit de chercher la matrice constante M telle que

$$(34)\qquad \forall\, t \in \mathbb{R}^+,\ \left[\sum_{i=o}^{m} \overline{u_i}(t)A^i_{12}\right]\left[\sum_{i=o}^{m} \overline{u_i}(t)A^i_{22}\right]^{-1} = M,\ M \in \mathbb{R}^{q_1 x q_2}$$

ce qui se met sous la forme du système d'équations matricielles

$$(35)\qquad \left[A^o_{12}, A^1_{12}, \ldots, A^m_{12}\right] = M\left[A^o_{22}, A^1_{22}, \ldots, A^m_{22}\right]$$

La solution, lorsque le système n'est pas compatible, au sens des moindres carrés est donnée par l'utilisation de la pseudo-inverse au sens de Moore-Penrose :

$$(36)\qquad \hat{M} = \left[\sum_{i=o}^{m} A^i_{12}A^{i\,T}_{22}\right]\left[\sum_{i=o}^{m} A^i_{22}A^{i\,T}_{22}\right]^{-1}$$

Lorsque le système est compatible, la solution (36) est encore valable, nous sommes alors dans les conditions de bilinéarité (I.2.3). La méthode présentée résoud donc le système dans tous les cas.

La matrice de commande lentes'exprime par :

$$B_1 = B_1 - \hat{M}B_2 \tag{37}$$

## II.2.2. Approximation de la matrice d'évolution

Alors que l'approximation précédente de la matrice de commande est nécessaire, l'approximation de la matrice système du modèle réduit lent peut se faire de trois façons, chacune nécessitant plus de calcul que la précédente.

a) Utilisation du calcul précédent

Ayant obtenu la relation (34), la matrice intervenant dans (28-b) peut être approchée par :

$$\hat{M} \sum_{i=o}^{m} (\overline{u_i}(t) A_{21}^{i}) \tag{38}$$

où $\hat{M}$ est donné par (36).

La matrice-système du modèle réduit lent (3) s'écrit, sans calculs supplémentaires :

$$A_1 = \sum_{i=o}^{m} \overline{u_i}(t) \, (A_{11}^{i} - \hat{M}A_{21}^{i}) \tag{39}$$

b) Utilisation d'un calcul équivalent

On doit vérifier ici la relation suivante :

$$\forall\, t \in \mathbb{R}^{+} \left[ \sum_{i=o}^{m} \bar{u}_i(t) A_{22}^{i} \right]^{-1} \left[ \sum_{i=o}^{m} \bar{u}_i(t) A_{21}^{i} \right] = L,\ L \in \mathbb{R}^{q_2 x q_1} \tag{40}$$

ce qui est équivalent au système d'équations :

$$(41) \qquad \begin{bmatrix} A_{21}^{o} \\ A_{21}^{1} \\ \vdots \\ A_{21}^{m} \end{bmatrix} = \begin{bmatrix} A_{22}^{o} \\ A_{22}^{1} \\ \vdots \\ A_{22}^{m} \end{bmatrix} L$$

dont une solution au sens des moindres carrés est donnée par :

$$(42) \qquad \hat{L} = \left[ \sum_{i=o}^{m} A_{22}^{iT} A_{22}^{i} \right]^{-1} \cdot \left[ \sum_{i=o}^{m} A_{22}^{iT} A_{21}^{i} \right]$$

Le modèle réduit lent est approché par un système bilinéaire dont la matrice-système est :

$$(43) \qquad A_1 = \sum_{i=o}^{m} \overline{u_i}(t) \, (A_{11}^{i} - A_{12}^{i} \hat{L})$$

où $\hat{L}$ est défini en (42)

c) Utilisation du calcul direct

Le calcul direct consiste à rechercher un ensemble $S_i$, $i=0,\ldots,m$, de matrices constantes de $\mathbb{R}^{q1 \times q1}$ tel que soit vérifié :

$$(44) \qquad \forall t \in \mathbb{R}_{+}, \left[ \sum_{i=o}^{m} A_{12}^{i} \overline{u_i}(t) \right] \left[ \sum_{i=o}^{m} A_{22}^{i} \overline{u_i}(t) \right]^{-1} \left[ \sum_{i=o}^{m} A_{21} \overline{u_i}(t) \right] = \sum_{i=o}^{m} S_i \, \overline{u_i}()$$

Ceci implique des calculs relativement lourds que nous allons détailler dans le cas où m = 1 ; mais, comme le montreront les résultats de simulation, le modèle approché obtenu a un meilleur comportement que ceux déterminés à l'aide des méthodes de calcul précédentes.

Soit m = 1, $u_o(t) = 1$, (44) s'écrit alors :

$$(45)\qquad \left[A_{12}^{o} + A_{12}^{1}\,\bar{u}(t)\right]\left[A_{22}^{o} + A_{22}^{1}\,\bar{u}(t)\right]^{-1}\left[A_{21}^{o} + A_{21}^{1}\bar{u}(t)\right] = S_o + S_1\bar{u}(t)$$

Or, si l'on note :

$$(46)\qquad \left(\mathrm{Com}\ \left[A_{22}^{o} + A_{22}^{1}\,\bar{u}(t)\right]\right)^{T} = \sum_{i=o}^{q_2 - 1} M_i \bar{u}^{i}(t)$$

$$\det\left[A_{22}^{o} + A_{22}^{1}\,\bar{u}(t)\right] = \sum_{i=o}^{q_2} \alpha_i \bar{u}^{i}(t)$$

où Com désigne la matrice des cofacteurs [GAN,60], det le déterminant instantané

$$\forall i \in \{1,\ldots, n_2-1\},\ M_i \in \mathbb{R}^{q_2 x q_2},\ \forall i \in \{1,\ldots, n_2\}\alpha_i \in \mathbb{R}$$

(Remarque : $\alpha_o = \det A_{22}^{o}$ et $\alpha_{q_2} = \det A_{22}^{1}$)

il vient l'ensemble d'équations algébriques matricielles :

$$(47)\qquad \begin{bmatrix} a_o \\ \cdots \\ a_1 \\ \cdots \\ a_2 \\ \cdots \\ \vdots \\ \cdots \\ a_{q_2} \\ \cdots \\ a_{q_2+1} \end{bmatrix} = \begin{bmatrix} \alpha_o I_{q_2} & \vdots & o \\ \cdots & \cdots & \cdots \\ \alpha_1 I_{q_2} & \vdots & \alpha_o I_{q_2} \\ \cdots & \cdots & \cdots \\ \alpha_2 I_{q_2} & \vdots & \alpha_1 I_{q_2} \\ \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots \\ \alpha_{q_2} I_{q_2} & \vdots & \alpha_{q_2-1} I_{q_2} \\ \cdots & \cdots & \cdots \\ o & \vdots & \alpha_{q_2} I_{q_2} \end{bmatrix} \begin{bmatrix} S_o \\ \cdots \\ S_1 \end{bmatrix}$$

avec $a_1 = A_{12}^{o} M_o A_{21}^{o}$ , $a_{q_2} = A_{12}^{1} M_{q_2-1} A_{21}^{o} + A_{12}^{o} M_{q_2-1} A_{21}^{1} + A_{12}^{1} M_{q_2-2} A_{21}^{1}$

$a_2 = A_{12}^{o} M_1 A_{21}^{o} + A_{12}^{o} M_o A_{21}^{1} + A_{12}^{1} M_o A_{21}^{o}$ , $a_{q_2+1} = A_{12}^{1} M_{q_2-1} A_{21}^{1}$

$\forall\, i \in \{2,\ldots, q_2-1\}$

$a_i = A_{12}^{o} M_i A_{21}^{o} + A_{12}^{1} M_{i-1} A_{21}^{o} + A_{12}^{o} M_{i-1} A_{21}^{1} + A_{12}^{1} M_{i-2} A_{21}^{1}$

Une solution de (47) notée :

$$(48)\qquad \left[\beta\right] = \left[\alpha\right]\begin{bmatrix} S_o \\ S_1 \end{bmatrix}$$

au sens des moindres carrés est cette fois ci :

$$(49)\qquad \begin{bmatrix} \hat{S}_o \\ \hat{S}_1 \end{bmatrix} = \left[\alpha^{T}\alpha\right]^{-1}\left[\alpha^{T}\beta\right]$$

Si de façon générale $S_o,\ldots, S_m$ forme une solution au sens des moindres carrés de (44), la matrice-système du modèle réduit approché est donnée par :

$$(50) \qquad A_1 = \sum_{i=o}^{m} (A_{11}^{i} - \hat{S}_i)\, \bar{u}_i(t)$$

## II.3 IDENTIFICATION

Dans le cas où l'on possède une simulation du modèle réduit lent (3), une méthode pour approcher ce comportement entrée-sortie, consiste à utiliser une méthode d'identification propre aux systèmes bilinéaires.

### II.3.1. Principe de la méthode d'identification

De nombreuses méthodes d'identificatiosn ont été développées pour l'identification des systèmes bilinéaires stationnaires ou non stationnaires ; la plupart utilisentun ensemble complet de fonctions orthogonales comme, par exemple, les fonctions de Walsh |KAR,78;CHE,78;DAD,85|, les fonctions block-pulse |CHE,82|, les polynômes de Laguerre, Legendre, Tchebytchev |CHA,82;HWA,82;LIU,84| ou les séries de Fourier |PAR,85|. Ces méthodes sont des extensions de méthodes utilisées dans le cas des systèmes linéaires pour l'identification et la commande optimale |SAN,77;SHI,83;CHE,75a-b; RAO,75;TZA,78;MOU,85;RAO,85|.

Le principe de ces méthodes est d'utiliser une base de fonctions orthogonales complète $\mathfrak{F} = \{f_i(t),\ i \in \mathbb{N}\}$ définies sur un intervalle de temps $\tau = [0,T]$. Dans ce cas, pour toute fonction $x(t)$ de $\tau$ dans $\mathbb{R}^n$, de normeintégrable sur $[0,T]$, on a

$$(51) \qquad \forall\, t \in \tau,\ x(t) = \sum_{n=o}^{\infty} x_n f_n(t)$$

où $x_n = \int_o^T p(t)\, x(t)\, f_n(t)$ où $p(t)$ est une fonction poids normalisant les produits scalaires $\langle f_i(t), f_j(t)\rangle$.

Soit $F_n = [f_o(t),\ldots,f_{n-1}(t)]^T$, une approximation de $x(t)$ à l'ordre n est donnée par :

$$(52) \qquad \forall\, t \in \tau, x(t) \simeq X_n^T F_n \quad \text{où } X_n^T = \left[ x_o, \ldots, x_{n-1} \right]$$

La méthode d'identification utilise plus particulièrement les propriétés de $F_n$ suivantes :

$P^1$) $f_o(t) = 1$

$P^2$) $\exists\, P_n \in \mathbb{R}^{nxn}\ \forall\, t \in \tau \quad \int_o^t F_n d\tau \simeq P_n F_n$

$P^3$) Soit un vecteur constant de $\mathbb{R}^n$

$F_n F_n^T U_n \simeq \tilde{U}_n F_n$

où $\tilde{U}_n$ est une matrice de $\mathbb{R}^{nxn}$ construite à l'aide des composantes de $U_n$.

Lorsque l'on désire identifier un système sous la forme bilinéaire :

$$(53) \qquad \dot{x} = A_o x + \sum_{i=1}^{m} u_i A_i x + BU \quad \text{où } x \in \mathbb{R}^q,\ U = \left[u,\ldots,u_m\right]^T \in \mathbb{R}^m$$

l'intégration de (53) et l'approximation à l'ordre n de x par $X_n^T F_n$ et $u_i$ par $U_{i_n}^T F_n$ fournit la relation :

$$(54) \qquad X_n^T F_n - \left[ x(o), o, \ldots, o \right] F_n = A_o X_n^T \int_o^t F_n d\tau + \sum_{i=1}^{m} A_i X_n^T \int_o^t F_n F_n^T U_{i_n} d\tau + B\, U_n^T \int_o^t F_n d\tau$$

où x(o) est la condition initiale de (53).

D'après les propriétés (P2) et (P3) et après simplification, on obtient le système d'équations :

$$(55) \qquad X_n^T - \left[ x(o), o,\ldots,o \right] = A_o X_n^T P_n + \sum_{i=1}^{m} A_i X_n^T \tilde{U}_{i_n} P_n + B U_n^T P_n$$

dont les inconnues sont les coefficients des matrices $A_0$, $A_i$, B et les données sont les $x_0,\ldots, x_{n-1}$, $U_0,\ldots, U_{n-1}$ issus du système à identifier. Si l'on compte le nombre d'inconnues et le nombre d'équations, le système est résoluble si $n > q + mq + m$. Dans ce cas, le système (55) se met sous la forme vectorielle d'un système linéaire qui peut être résolu au sens des moindres carrés.

### II.3.2. Utilisation des fonctions de Walsh

Comme nous l'avons mentionné dans le paragraphe précédent, les fonctions les plus utilisées sont les fonctions de Walsh [FIN,49] qui possèdent les propriétés supplémentaires d'être constantes par morceaux et d'avoir $p(t) = 1$. Le calcul des $x_n$(51) en est alors facilité.

a) Définition des fonctions de Walsh

Elles sont construites à partir des fonctions de Rademacher. Celles-ci sont définies sur $[0,1]$, mais on peut toujours s'y ramener par dilatation ou contraction du temps :

Fonctions de RADEMACHER :
$$\phi_0(t) = \begin{cases} 1 & t \in [0, 1/2] \\ -1 & t \in [1/2, 1] \end{cases}$$
$$\phi_0(t+1) = \phi_0(t)$$
$$\phi_n(t) = \phi_0(2^n t) \quad \forall\, n > 1$$

Fonctions de WALSH :
$$\psi_0(t) = 1$$
$$\psi_n(t) = \prod_{i=1}^{r} \phi_{n_i}(t) \text{ où les entiers } n_i \text{ sont donnés par : } n = \sum_{i=1}^{r} 2^{n_i},\ n_{i+1} < n_i$$

Ces fonctions de WALSH possèdent les propriétés suivantes :

* dans (51) $x_n = \int_0^1 x(t)\, \psi_n(t)\, dt$
* $\forall m, n \in \mathbb{N}^2$ $\psi_m(t)\, \psi_n(t) = \psi_{m \pm n}(t)$ où $\pm$ représente l'addition modulo-2 sans dépassement.

b) Propriétés pour l'identification

Les propriétés utiles à l'identification sont les constructions de $P_n$ et $\tilde{U}_n$ dans (P2) et (P3). On peut montrer |KAR,78| que l'on a, en notant $\psi_{(m)} = |\psi_o,\ldots,\psi_{m-1}|^T$

P'2] pour $m = 2^n$, $n \in \mathbb{N}$

$\int_o^t \psi_{(m)}\, d\tau = P_m\, \psi_{(m)}$ où $P_{(m)}$ est construit par récurrence à l'aide de l'algorithme :

$$(56)\qquad \begin{cases} P_1 = \frac{1}{2} ; \\ n > 1 \quad P_{2^n} = \left[\begin{array}{c:c} P_{2^{n-1}} & -\frac{1}{2^{n+1}}\, I_{2^{n-1}} \\ \hdashline \frac{1}{2^{n+1}}\, I_{2^{n-1}} & o \end{array}\right] \end{cases}$$

P'3] pour $m = 2^n$, $n \in \mathbb{N}$, $\alpha \in \mathbb{R}^m$

$$\psi_{(m)}\, \psi_{(m)}^T\, \alpha = \Lambda_\alpha^m\, \psi_{(m)}$$

où $\Lambda_\alpha^m = \left[\alpha,\ \Lambda_1^{(m)}\alpha\ ,\ \ldots,\ \Lambda_{m-1}^{(m)}\alpha\right]$

où les $\Lambda_i^{(m)}$, $i \in \{1,\ldots, m-1\}$, sont construits par la récurrence :

$$(57)\qquad \begin{cases} \Lambda_o^{(1)} = 1 \\ \\ n>1 \quad i \in \{o,\ldots, 2^{n-1}-1\} : \\ \qquad \begin{cases} \Lambda_i^{(2n)} = \left[\begin{array}{c:c} \Lambda_i^{(2^{n-1})} & o \\ \hdashline o & \Lambda_i^{(2^{n-1})} \end{array}\right] \\ \\ \Lambda_{2^{n-1}+i}^{(2n)} = \left[\begin{array}{c:c} o & \Lambda_i^{(2^{n-1})} \\ \hdashline \Lambda_i^{(2^{n-1})} & o \end{array}\right] \end{cases} \end{cases}$$

Ainsi, d'après ces propriétés, il est nécessaire d'utiliser des approximations d'état et d'entrée à l'ordre d'une puissance de 2.

## II.4 EXEMPLE

Considérons le système bilinéaire singulièrement perturbé :

$$(58)\qquad \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = \left\{ \begin{bmatrix} -2 & -1 \\ -1 & -1 \end{bmatrix} + u \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} \right\} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$\varepsilon = .1$, $x(o) = z(o) = o$

L'utilisation de la méthode des perturbations singulières conduit au modèle réduit lent :

$$(59)\qquad \begin{cases} \dot{\bar{x}} = \left[ -2 \; -u + \dfrac{(u-1)^2}{1+2u} \right] \bar{x} + \left[ 1 + \dfrac{u-1}{1+2u} \right] u \\ \bar{x}(o) = o \end{cases}$$

avec la condition $\forall\, t \in \mathbb{R}^+$, $1+2u \neq 0$.

On désire approcher ce modèle par un système bilinéaire d'ordre 1 défini par :

$$(60)\qquad \dot{x} = ax + u\tilde{n}x + bu$$

En supposant que l'on ne connaisse pas a priori u, on utilisera donc la méthode directe et la méthode par identification.

### II.4.1. Méthode directe

On a d'abord à résoudre le système incompatible concernant le vecteur de commande

$$(61)\qquad \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} m$$

dont la solution au sens des moindres carrés est $\hat{m} = .2$. Cette valeur fournit un premier modèle approché :

$$(62) \qquad \overset{\circ}{x}_1 = [-2.2 - .8u]\, x_1 + 1.2u$$

L'utilisation du calcul équivalent (II.2.2.b) n'est pas nécessaire puisque l'on est dans le cas scalaire. Pour approcher la matrice-système par (II.2.2.c), on doit résoudre le système incompatible :

$$(63) \qquad \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \end{bmatrix}$$

dont la solution au sens des moindres carrés est $\hat{s}_0 = -.714$, $\hat{s}_1 = .286$. Le système approché est alors :

$$(64) \qquad \overset{\circ}{x}_2 = [-2.714 - .714u]\, x_2 + 1.2u$$

Les résultats de simulation (figure 1-a) indiquent une plus grande précision de ce dernier.

### II.4.2. Identification

Une simulation de (59) donne, pour $u(t) = 1-e^{-t}$, les valeurs suivantes calculées sur $|0,1|$ (après contraction du temps par 10) :

| i | $x_i$ | $u_i$ |
|---|---|---|
| 0 | .2855 | .9 |
| 1 | -.0455 | -.0937 |
| 2 | -.0386 | -.0802 |
| 3 | -.0397 | -.0841 |
| 4 | -.0248 | -.0587 |
| 5 | -.0260 | -.0629 |
| 6 | -.0222 | -.0554 |
| 7 | -.0202 | -.0498 |

TABLE 1

L'équation d'identification (55) fournit le système d'équations incompatibles (obtenu pour n = 8) :

$$(65)\qquad \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} .125 & .1153 & .4129 \\ -.078 & -.0759 & -.2394 \\ -.371 & -.0355 & -.1160 \\ .0044 & .0058 & .0086 \\ -.0178 & -.0171 & -.0562 \\ .0028 & .0035 & .0059 \\ .0024 & .0028 & .0050 \\ .0025 & .0030 & .0053 \end{bmatrix} \begin{bmatrix} a \\ \tilde{n} \\ b \end{bmatrix}$$

La résolution au sens des moindres carrés de (65) donne :

$$(66)\qquad \begin{cases} \hat{a} = -2.5528 \\ \hat{\tilde{n}} = -11.9631 \\ \hat{b} = 4.8078 \end{cases}$$

qui conduit directement au système approché (en divisant chaque terme par 10 pour revenir en temps normal)

$$(67)\qquad \dot{x}_3 = (-.255 - u1.196)x_3 + .481u$$

dont les résultats de simulation (figure 1-b) indiquent une nette amélioration de la précision. Cependant, une étude de la robustesse de ces modélisations par rapport à une variation de l'entrée (figures 2-a et 2-b) montre que les modèles approchés par le calcul direct, donnent une meilleure approximation qualitative des dynamiques que le modèle déterminé par identification.

## III. DECOUPLAGE DES PARTIES LENTES ET RAPIDES |ROT,86b|

Nous allons, dans cette partie, utiliser la méthode de découplage présentée dans le cas des systèmes linéaires (II : II.2.2). On supposera ici que l'on a un modèle réduit lent bilinéaire (obtenu par l'une quelconque des méthodes présentées dans la partie précédente). Soit ce modèle défini par :

$$(68)\qquad \begin{cases} \overset{\circ}{x}_1 = \left(\sum_{i=o}^{m} A_1^i\, u_i(t)\right) x_1 + B_1 u(t) \\ x_1(o) = x_o \end{cases}$$

$x_1$ étant une approximation de x sur toute l'évolution du système, on va donc replacer (68) dans (1), ce qui donne :

$$(69)\qquad \begin{cases} \begin{bmatrix} \overset{\circ}{x}_1 \\ \circ \\ \bar{z} \end{bmatrix} = \sum_{i=o}^{m} u_i(t) \left[\begin{array}{c:c} A_1^i & o \\ \hdashline \dfrac{A_{12}^i}{\varepsilon} & \dfrac{A_{22}^i}{\varepsilon} \end{array}\right] \begin{bmatrix} x_1 \\ \bar{z} \end{bmatrix} + \begin{bmatrix} B_1 \\ \dfrac{B_2}{\varepsilon} \end{bmatrix} u(t) \\ \bar{y} = \left[C_1 \vdots C_2\right] \begin{bmatrix} x_1 \\ \bar{z} \end{bmatrix} \; ; \; x_1(o) = x_o \; ; \; \bar{z}(o) = z_o \end{cases}$$

La transformation

$$(70)\qquad \begin{bmatrix} x_1 \\ z_r \end{bmatrix} = \begin{bmatrix} I_{q_1} & o \\ L & I_{q_2} \end{bmatrix} \begin{bmatrix} x_1 \\ \bar{z} \end{bmatrix}$$

découple les parties lentes et rapides de (69) si sont vérifiées les m+1 équations matricielles.

$$(71)\qquad \forall\, i \in \{0,\ldots,m\} \quad \varepsilon L A_1^i - A_{22}^i L + A_{21}^i = R_i(L) = o$$

Alors que dans le cas linéaire, ce système se résolvait exactement, on a ici un ensemble d'équations incompatibles dans le cas général. L'écriture sous forme vectorielle de(71):

$$(72)\qquad \begin{bmatrix} \mathcal{t}_o \\ \vdots \\ \mathcal{t}_m \end{bmatrix} \operatorname{vec}(L) = \begin{bmatrix} \operatorname{vec} A_{21}^o \\ \vdots \\ \operatorname{vec} A_{21}^m \end{bmatrix}$$

où $\forall\, i \in \{0,\ldots,m\}\ \mathcal{t}_i = -\varepsilon A_i^1 \otimes I_{q_2} + I_{q_1} \otimes A_{22}^i$

(les autres notations étant définies en II : 55)

permet de résoudre (71) au sens des moindres carrés ; la solution est alors donnée par :

$$(73) \qquad \mathrm{vec}(\bar{L}) = \Big(\sum_{i=0}^{m} \mathcal{A}_i^T \mathcal{A}_i\Big)^{-1} \Big(\sum_{i=0}^{m} \mathcal{A}_i^T \, \mathrm{vec}(A_{21}^i)\Big)$$

Après ce changement de variables, en négligeant les matrices d'interconnections $R_i(L)$, $i = 0,\ldots, m$, (69) se met sous la forme découplée :

$$(74) \quad \begin{cases} \text{système lent} \\ a) \begin{cases} \dot{x}_1 = \Big(\sum_{i=0}^{m} A_1^i u_i(t)\Big)x_1 + B_1 u(t) \; ; \; x_1(o) = x_o \\ y_1 = (C_1 - C_2\bar{L})x_1 \end{cases} \\ \text{système rapide} \\ b) \begin{cases} \varepsilon \dot{z}_r = \Big(\sum_{i=0}^{m} A_{22}^i u_i(t)\Big)z_r + (B_2 + \varepsilon \bar{L} B_1)u(t) \\ y_r = C_2 z_r \; ; \; z_r(o) = z_o + Lx_o \end{cases} \\ \bar{y} = y_r + y_1 \end{cases}$$

et la variable rapide z peut être approchée par

$$(75) \qquad \bar{z}(t) = z_r(t) - Lx_1(t)$$

Un résultat de simulation mené sur l'exemple (58) est présenté figure 3. Cette courbe montre que les dynamiques de $\bar{z}$ correspondent bien à celles de z. Cependant, on note en régime statique, une erreur importante qui est due à trois facteurs essentiels :
l'utilisation de la méthode des perturbations singulières qui fournit une bonne approximation sur les variables lentes ; l'annulation des termes $R_i(L)$ qui augmentent l'imprécision ; l'étude en régime statique qui nécessiterait une étude du système linéarisé, donc donnerait de meilleurs résultats, mais ceci ne peut être réalisé que lorsqu'on connait le comportement à l'infini de l'entrée.

De manière à augmenter la précision, on peut agir sur le premier facteur en utilisant la notion de système réciproque développée pour l'étude des systèmes linéaires (II : II.2.4) |DAU,83a|.

## IV. EXTENSION DU SYSTEME RECIPROQUE POUR LES SYSTEMES BILINEAIRES

### IV.1 DEFINITION

Soit le système bilinéaire :

$$(76)\quad \begin{cases} \dot{x} = \sum_{i=0}^{m} A_i u_i x + Bu \\ y = Cx \end{cases}$$

que l'on peut supposer partir de conditions initiales nulles. Le comportement entrée-sortie de ce système est entièrement décrit par la fonction de transfert codée (I : II.3.3.b).

$$(77)\quad F(\alpha_0,\alpha_1,\ldots,\alpha_m) = C(I - \sum_{i=0}^{m} A_i\alpha_i)^{-1}B$$

Or, au cours de notre étude, nous avons relevé les points suivants :

* I : II.2.2.
  La transformée de Laplace-Borel de y est obtenue à partir de la série génératrice du système en remplaçant $\alpha_i$ par l'opérateur en $\alpha_0$ : $\alpha_0[g_{u_i}(\alpha_0)\ \sqcup\!\!\sqcup\ \cdot]$

* II.3.3.b
  La fonction de transfert codée d'un système linéaire s'obtient à partir de sa fonction de transfert en p en faisant la transformation

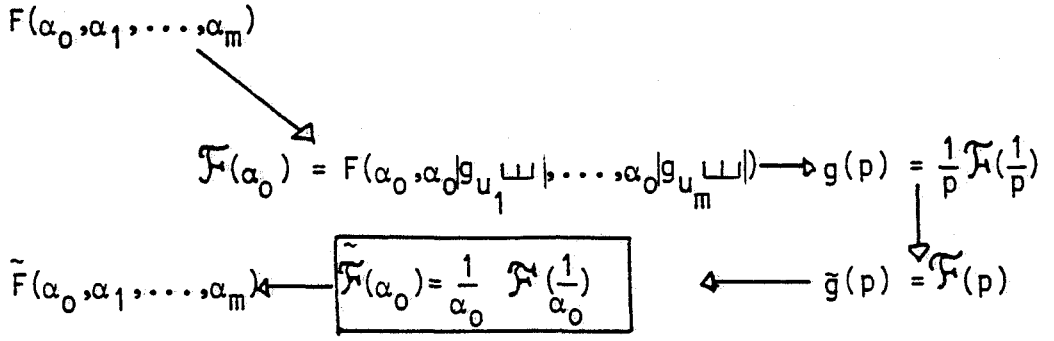
$$F(p) \longrightarrow \frac{1}{\alpha_0} F(\frac{1}{\alpha_0})$$

* II : II.2.4

La fonction de transfert du système réciproque d'un système linéaire s'obtient à partir de sa fonction de transfert par la transformation

$$F(p) \longrightarrow \tilde{F}(p) = \frac{1}{p} F\left(\frac{1}{p}\right)$$

On peut donc définir la fonction de transfert codée du système réciproque de (76) à partir de (77) par le schéma suivant :




et la relation encadrée de ce schéma traduit la définition du système réciproque pour un système bilinéaire.

## IV.2 CONSTRUCTION

A partir de (77), il vient :

$$(78) \qquad \mathcal{F}(\alpha_0) = C(I - \sum_{i=0}^{m} A_i \alpha_0 g_{u_i}(\alpha_0)\sqcup\!\rangle)^{-1}B$$

donc

$$(79) \qquad \tilde{\mathcal{F}}(\alpha_0) = \frac{1}{\alpha_0} C(I - \sum_{i=0}^{m} A_i \frac{1}{\alpha_0} g_{u_i}(\frac{1}{\alpha_0})\sqcup\!\rangle)^{-1}B$$

Or, $\frac{1}{\alpha_0} g_{u_i}(\frac{1}{\alpha_0}) = \tilde{g}_u$ , transformée de Laplace-Borel du signal réciproque de $u_i$ ; (79) s'écrit alors :

$$(80) \qquad \tilde{\mathcal{F}}(\alpha_0) = C(\alpha_0 - A_0 - \sum_{i=1}^{m} A_i \alpha_0 \tilde{g}_{u_i}\sqcup\!\rangle)^{-1}B$$

En codant $\alpha_o \tilde{g}_{u_i} \omega$ par $\tilde{\alpha}_i$, i = 0,..., m, cette écriture est équivalente à

$$(81) \qquad \tilde{F}(\tilde{\alpha}_o, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_m) = C(\tilde{\alpha}_o - A_o - \sum_{i=1}^{m} A_i \tilde{\alpha}_i)^{-1} B$$

Pour arriver à la forme d'une fonction de transfert codée, il suffit, comme dans le cas linéaire, de poser :

$$(82) \qquad A_o = PQ \text{ , P et Q matrices régulières}$$

on obtient alors la fonction de transfert codée du système bilinéaire réciproque de (76) :

$$(83) \qquad \tilde{F}(\tilde{\alpha}_o, \tilde{\alpha}_1, \ldots, \tilde{\alpha}_m) = \tilde{C}(I - \sum_{i=o}^{m} \tilde{A}_i \tilde{\alpha}_i)^{-1} \tilde{B}$$

$$\text{avec} \begin{cases} \tilde{C} = -CQ^{-1} \; ; \; \tilde{B} = P^{-1}B \\ \tilde{A}_o = P^{-1}Q^{-1} \\ \forall i \in \{1,..,m\} \; \tilde{A}_i = -P^{-1}A_iQ^{-1} \end{cases}$$

Le système bilinéaire réciproque de (76) a pour équation d'état :

$$(84) \qquad \begin{cases} \overset{\circ}{\tilde{x}}' = (\sum_{i=o}^{m} \tilde{A}_i \tilde{u}_i')\tilde{x}' + \tilde{\tilde{B}}\tilde{u}' \\ \tilde{y} = \tilde{\tilde{C}}\tilde{x}' \end{cases}$$

Il est à noter que nous aurions obtenu la même chose si, à partir de la fonction de transfert multivariable (I : 30), on définit la transformation :

$$(85) \qquad \forall n \in \mathbb{N} \quad \tilde{F}_n = \frac{1}{p_1 .. p_n} F_n(\frac{1}{p_1}, \ldots, \frac{1}{p_n})$$

qui détermine la fonction de transfert multivariable du système réciproque, ceci souligne bien l'équivalence de ces deux formes de représentation |FLI,79|. Cependant, dans le cas où plusieurs variables d'entrée interviennent dans la matrice système, cette dernière représentation est délicate d'emploi.

## IV.3 UTILISATION AVEC LA METHODE DES PERTURBATIONS SINGULIERES

Reprenons le système (1), un choix naturel de décomposition de la matrice d'évolution en régime libre est $P = I_\varepsilon$ et $Q = A_o$, où $I_\varepsilon$ est définie en (II :78).

Le système réciproque de (1) est régi, dans ces conditions, par l'équation d'état :

$$(86)\quad \begin{bmatrix} \overset{\circ}{\tilde{x}}' \\ \overset{\circ}{\tilde{z}}' \end{bmatrix} = \left[ \begin{bmatrix} \bar{A}_{11}^{o} & \bar{A}_{12}^{o} \\ \varepsilon\bar{A}_{21}^{o} & \varepsilon\bar{A}_{22}^{o} \end{bmatrix} + \sum_{i=1}^{m} \begin{bmatrix} \bar{A}_{11}^{i} & \bar{A}_{12}^{i} \\ \bar{A}_{21}^{i} & \bar{A}_{22}^{i} \end{bmatrix} \tilde{u}_i'(t) \right] \begin{bmatrix} \tilde{x}' \\ \tilde{z}' \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \tilde{u}'(t)$$

$$\tilde{y} = -\left[ C_1\bar{A}_{11}^{o} + C_2\bar{A}_{21}^{o}, C_1\bar{A}_{12}^{o} + C_2\bar{A}_{22}^{o} \right] \begin{bmatrix} \tilde{x} \\ \tilde{z} \end{bmatrix}$$

$$\text{où } * \begin{bmatrix} \bar{A}_{11}^{o} & \vdots & \bar{A}_{12}^{o} \\ \cdots & \vdots & \cdots \\ \bar{A}_{21}^{o} & \vdots & \bar{A}_{22}^{o} \end{bmatrix} = \begin{bmatrix} A_{11}^{o} & \vdots & A_{12}^{o} \\ \cdots & \vdots & \cdots \\ A_{21}^{o} & \vdots & A_{22}^{o} \end{bmatrix}^{-1}$$

(Cf.II : 79 pour la définition des blocs $\bar{A}_{ij}^{o}$)

$$* \ \forall i \in \{1,\ldots,m\} \quad \begin{bmatrix} \bar{A}_{11}^{i} & \bar{A}_{12}^{i} \\ \bar{A}_{21}^{i} & \bar{A}_{22}^{i} \end{bmatrix} = -\begin{bmatrix} A_{11}^{i} & A_{12}^{i} \\ A_{21}^{i} & A_{22}^{i} \end{bmatrix} \begin{bmatrix} \bar{A}_{11}^{o} & \bar{A}_{12}^{o} \\ \bar{A}_{21}^{o} & \bar{A}_{22}^{o} \end{bmatrix}$$

De façon à avoir un système à deux dynamiques, il est nécessaire que l'entrée constante $u_o(t) = 1$ soit dominante par rapport aux autres. Le modèle réduit lent bilinéaire approchant z sera donc obtenu par la méthode des entrées dominantes. Il est donné par :

$$(87)\quad \overset{\circ}{\tilde{z}}_1' = (\tilde{A}_1 + \sum_{i=1}^{m} \tilde{u}_i'(t)\ \tilde{A}_1^i)\tilde{z}_1' + \tilde{B}_1\tilde{u}'(t)$$

$$\text{avec } \tilde{A}_1^o = \varepsilon A_{22}^{o\,-1} \ ; \ \tilde{B}_1 = B_2 + \varepsilon A_{22}^{o\,-1} A_{21}^{o}\, B_2$$

$$\forall i \in \{1,\ldots,m\}$$

$$\tilde{A}_1^i = -A_{22}^{i} A_{22}^{o\,-1} - \varepsilon A_{22}^{o\,-1} A_{21}^{o}\, \bar{A}_{12}^{i}$$

L'utilisation de la transformation réciproque sur ce système conduit à la définition du système bilinéaire rapide à l'intérieur de la couche limite (approximé à l'ordre 0).

$$(88) \qquad \varepsilon z_r = (A_{22}^o + \sum_{i=1}^{m} u_i(t)\, A_{22}^o\, {}^{-1}A_{21}^o\, \bar{A}_{12}^i\, A_{22}^o) z_r + B_2 u$$

avec $z_r(o) = z_o$

L'utilisation de la méthode de découplage à partir du système réciproque, à l'aide de (87) replacée dans (86), permet également de définir une approximation de la variable lente x à l'intérieur de la couche limite.

A titre d'illustration de cette méthode, on peut reprendre l'exemple (58) présenté au II.4. Le système réciproque associé est défini par :

$$(89) \qquad \begin{bmatrix} \overset{\circ}{\tilde{x}}' \\ \overset{\circ}{\tilde{z}}' \end{bmatrix} = \left\{ \begin{bmatrix} -1 & 1 \\ \varepsilon & -2\varepsilon \end{bmatrix} + \tilde{u}' \begin{bmatrix} -2 & 3 \\ 3 & -5 \end{bmatrix} \right\} \begin{bmatrix} \tilde{x}' \\ \tilde{z}' \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tilde{u}'$$

L'application des perturbations singulières par l'introduction sur la première ligne d'un petit paramètre $\mu$ conduit avec $\mu=0$ à l'équation d'entraînement :

$$(90) \qquad \bar{\tilde{x}}' = \frac{1+3\tilde{u}'}{1+2\tilde{u}'} \bar{\tilde{z}}' + \frac{1}{1+2\tilde{u}'} \tilde{u}'$$

donc au modèle réduit lent :

$$(91) \qquad \overset{\circ}{\bar{\tilde{z}}}' = \left[ (-2\varepsilon - 5\tilde{u}') + \frac{(\varepsilon+3\tilde{u}')(1+3\tilde{u}')}{1+2\tilde{u}'} \right] \bar{\tilde{z}}' + \left[ \frac{\varepsilon+3\tilde{u}'}{1+2\tilde{u}'} + 1 \right] \tilde{u}'$$

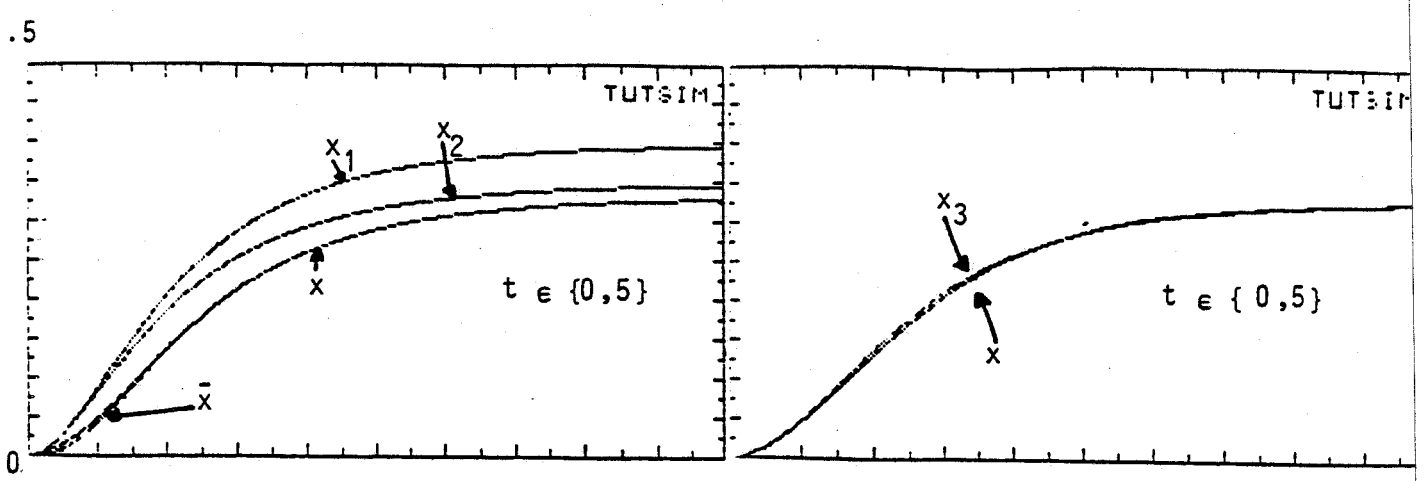
Le modèle bilinéaire approché pour $|\tilde{u}'| \simeq 0$ donne :

$$(92) \qquad \overset{\circ}{\tilde{z}}_1 = \left[ -\varepsilon + \tilde{u}'(\varepsilon - 2) \right] \tilde{z}_1 + (1 + \varepsilon)\, \tilde{u}'$$

L'application de la transformation réciproque sur ce système conduit au modèle bilinéaire valable à l'intérieur de la couche limite, approchant la variable rapide z par :

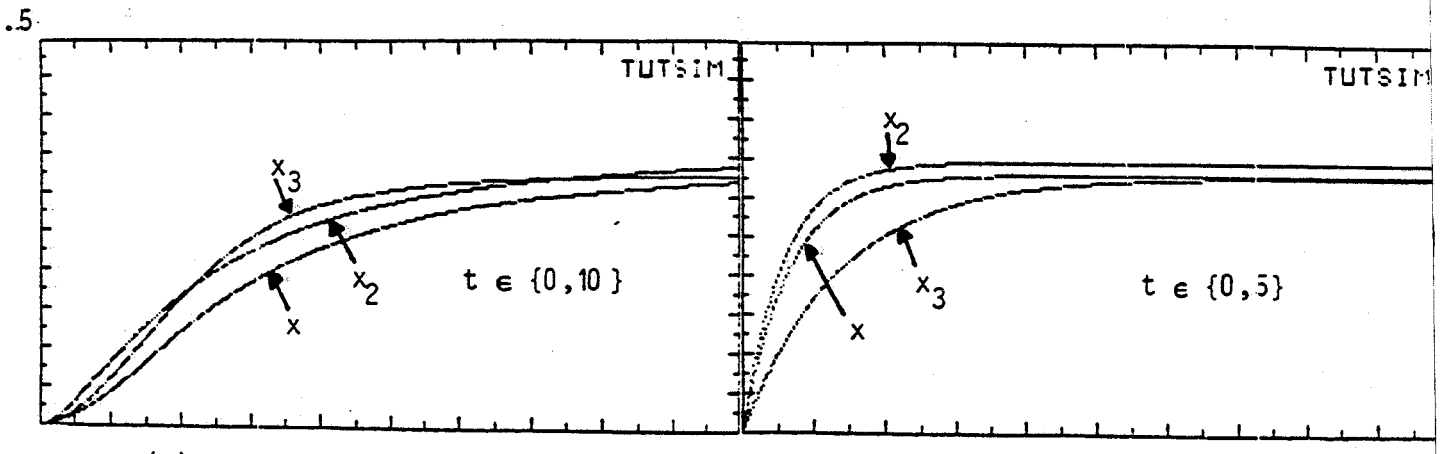
$$(93) \qquad \begin{aligned} \varepsilon \overset{\circ}{z}_r &= \left[ -1 - u(2-\varepsilon) \right] z_r + (1+\varepsilon)u \\ z_r(o) &= z_o \end{aligned}$$

Le résultat d'une simulation est présenté figure 4.




(a) u=1-exp(-t) (b)

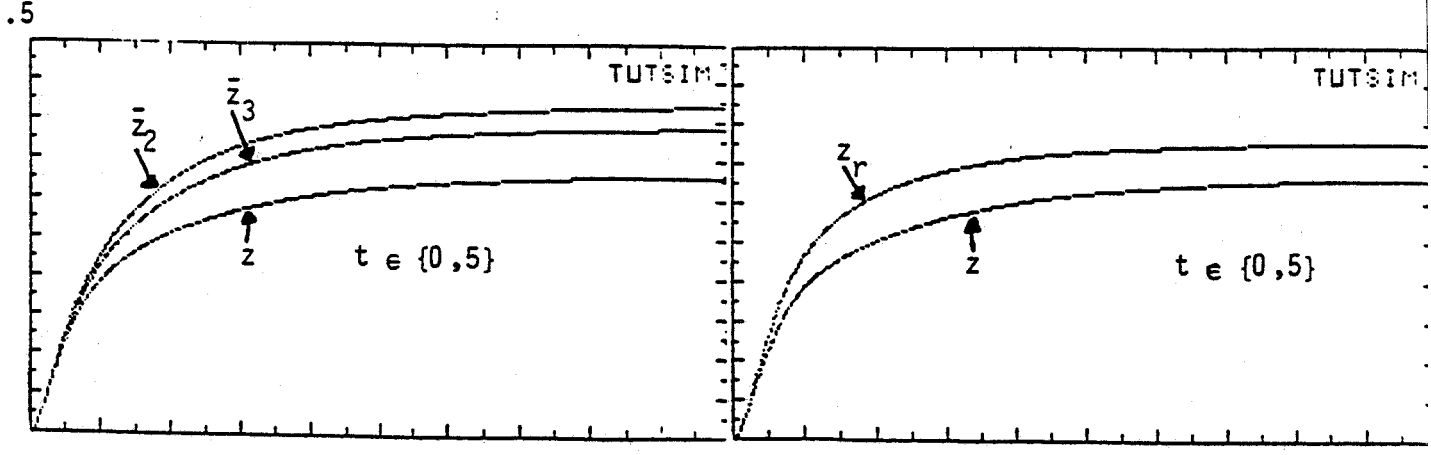
FIGURE 1




(a) u = 1-exp(-t/3) (b) u = 1-exp(-5t)

FIGURE 2

BU LILLE




u = 1-exp(-t) u = 1-exp(-t)

FIGURE 3 FIGURE 4

## CONCLUSION

Nous avons étudié dans ce chapitre, plusieurs méthodes permettant de définir des modèles réduits approchés d'un système bilinéaire singulièrement perturbé.

Ces méthodes ont été adaptées à partir de celles utilisées dans le cas de systèmes linéaires. Elles doivent être utilisées avec précaution, car il faut en permanence vérifier si le système garde sa propriété de double échelle de temps.

Certains problèmes ont été soulevés, notamment celui de la précision des modèles réduits en régime statique par rapport au système initial. En effet, dans le cas des systèmes linéaires, le gain statique est indépendant des variable d'entrée et la méthode des perturbations singulières conserve ce gain. Par contre, en ce qui concerne les systèmes bilinéaires, le gain statique est une fonction non linéaire des variables d'entrée ; ceci explique la mauvaise précision en régime statique des modèles réduits obtenus. Cependant, ces modèles conservent bien les dynamiques des variables d'état initiales.

Disposant, maintenant, de modèles réduits approchant le système initial, nous allons passer dans le chapitre suivant à l'utilisation de ces modèles pour générer une commande quasi-optimale du système initial.

# CHAPITRE V

# COMMANDE

# INTRODUCTION

Les systèmes bilinéaires, comme nous l'avons indiqué dans les chapitres précédents, sont un compromis entre les systèmes linéaires (par leur structure) et les systèmes non-linéaires (par leur comportement). Ce point de vue nous a conduit à commencer ce chapitre par deux parties concernant la commande optimale des systèmes linéaires, pour la première, et des systèmes non linéaires, pour la seconde.

Ces parties ne sont pas constituées essentiellement que de rappels. En effet, dans la première, l'accent est mis sur l'utilisation des techniques de réduction de dimensionnalité pour la construction de commandes sous-optimales et une large description porte sur l'apport de la méthode de découplage des systèmes singulièrement perturbés. Dans la seconde, on développe la construction analytique du retour d'état pour des systèmes non linéaires à commande affine.

La troisième partie est constituée naturellement de l'induction ou de l'application de ces méthodes aux cas des systèmes bilinéaires et, en particulier, lorsqu'ils se présentent sous une forme perturbée.

La commande optimale dépendant du critère à minimiser, nous nous bornerons à étudier le problème de la commande optimale dans le cas d'un critère quadratique non singulier à horizon infini sans contraintes. Dans le cas des problèmes singuliers, l'utilisation de la méthode des perturbations singulières associée à l'emploi du système réciproque permet de contourner la difficulté | DAU,83a | ; nous ne les envisagerons pas ici.

## I. COMMANDE DES SYSTEMES LINEAIRES

Nous ne nous intéressons qu'aux procédures de commande issues des méthodes de réduction de dimensionnalité présentées au chapitre II. Il est bien entendu, que nous n'envisagerons pas les solutions du type commande décentralisée ou commande hiérarchisée | SIN,80 |. Considérons le système linéaire :

$$(1) \qquad \begin{cases} \dot{x} = Ax + Bu \ ; \ x(o) = x_o \ ; \ x \in \mathbb{R}^q, \ u \in \mathbb{R}^m, \ y \in \mathbb{R}^n \\ y = Cx \end{cases}$$

où A, B, C sont des matrices constantes, et le critère à minimiser :

$$(2) \qquad J = \frac{1}{2} \int_0^\infty (y^T Q y + u^T R u) d\tau$$

où Q est non négative et R définie positive.

La solution de ce problème est donnée par le principe du maximum de Pontryagin | ROI,74 | qui fournit :

$$(3) \qquad \begin{aligned} &3a) \quad u = - R^{-1} B^T K x \\ &\quad \text{avec K solution de l'équation algébrique de Riccati :} \\ &3b) \quad KA + A^T K - KBR^{-1}B^T K + C^T Q C = 0 \end{aligned}$$

Introduisons les matrices M et N définies par :

$$(4) \qquad \begin{aligned} BR^{-1}B^T &= MM^T \\ C^T Q C &= N^T N \end{aligned}$$

alors, suivant | KUC,72 |, si le triplet (A, M, N) est stabilisable-détectable, l'équation de Riccati (3b) possède une solution unique définie non négative et la matrice $(A - BR^{-1}B^T K)$ est de Hurwitz (asymptotiquement stable). La construction de la commande optimale implique donc la résolution d'une équation de Riccati d'ordre q; ce qui, dans le cas de grands systèmes, peut s'avérer numériquement délicat. L'utilisation des techniques de réduction de dimensionnalité permet de remplacer cette commande par des commandes sous-optimales dont la détermination demande la résolution d'équations d'ordre inférieur.

## I.1. UTILISATION DE LA SORTIE

La commande (3a) demande l'accession au vecteur état complet ;
lorsque cela n'est pas possible, et si l'on ne veut pas faire appel à un reconstructeur d'état | LAR,77 |, on doit se contenter de la sortie. Nous allons voir que ceci est un moyen de simplifier l'équation de Riccati. En effet, le vecteur de sortie peut être considéré comme un modèle agrégé priviligié de l'état de (1) (Cf. II : I-3). Ce modèle est régi par l'équation d'état, en supposant que C soit de rang plein :

$$(5)\qquad \begin{aligned} &\dot{y} = \tilde{A}y + \tilde{B}u, \text{ avec } \tilde{A} = CAC^* \text{ et } \tilde{B} = CB \\ &y(o) = CX_o \\ &\text{où } C^* \text{ est telle que } CC^* = I_n \end{aligned}$$

Nous allons donc comparer la solution issue du problème direct (1-2) à celle du problème agrégé (5-2). D'après (3), la solution optimale, avec retour de sortie, du problème direct est de la forme :

$$(6)\qquad u = -R^{-1}B^T P y$$

où P est définie par $KC^* = P$
et K est solution de l'équation de Riccati d'ordre q (3).

Considérons maintenant le problème d'optimalité sur le système agrégé (5) ; au sens du critère (2), la solution optimale est donnée par :

$$(7)\qquad \begin{aligned} &7a)\ u = -R^{-1}\tilde{B}^T\tilde{K}y \\ &\text{avec } \tilde{K} \text{ solution de} \\ &7b)\ \tilde{K}\tilde{A} + \tilde{A}^T\tilde{K} - \tilde{K}\tilde{B}R^{-1}\tilde{B}^T\tilde{K} + Q = 0 \end{aligned}$$

Cette dernière équation de Riccati s'écrit, compte tenu de (5), sous la forme :

$$(8)\qquad (\tilde{K}C)(AC^*) + (AC^*)^T(\tilde{K}C)^T - (\tilde{K}C)BR^{-1}B^T(\tilde{K}C)^T + Q = 0$$

Or, en multipliant chacun des termes de l'équation de Riccati (3b), à droite et à gauche respectivement par $C^*$ et $C^{*T}$, il vient :

(9) $$(KC^*)^T (AC^*) + (AC^*)^T(KC^*) - (KC^*)^T BR^{-1}B^T(KC^*) + Q = 0$$

Ce qui entraîne l'identité :

(10) $$\tilde{K}C = (KC^*)^T = P^T$$

soit

(11) $$P = C^T\tilde{K}$$
où $\tilde{K}$ est la solution d'une équation de Riccati d'ordre n.

On a donc, par agrégation, simplifié le problème de commande quasi optimale par retour de sortie en diminuant l'ordre des équations à résoudre. N'importe quel choix d'inverse à droite de C peut être utilisé ; cependant, si l'on regarde les coûts de chaque critère, on a :

(12)
$$\begin{cases} J_{opt} = \frac{1}{2} x_o^T K x_o \text{ avec le retour d'état (3)} \\ \tilde{J} = \frac{1}{2} x_o^T L x_o \text{ avec le retour de sortie } u = -R^{-1}B^T C^T \tilde{K} y \\ \text{où } W = L - K \text{ est solution de l'équation de Lyapounov} \end{cases}$$

$$(A - SKr)^T W + W(A - SKr) + (K - Kr)\, S(Kr - K) = 0$$
avec $S = BR^{-1}B^T$ et $Kr = C^T\tilde{K}C$

De façon à obtenir une commande quasi optimale assurant un coût du critère aussi proche que possible de $J_{opt}$, on doit vérifier, d'après (12) :

(13) $$K = C^T\tilde{K}C = PC$$

Une solution au sens des moindres carrés de cette équation nous est fournie par la pseudo inverse à droite au sens de Moore-Penrose. $C^*$ sera donc donné par :

(14) $$C^* = C^+ = C^T(CC^T)^{-1}$$

La condition d'unicité d'une solution définie positive de (7b) s'écrit dans ce cas : $(CAC^+, CM, NC^+)$ stabilisable-détectable.

Considérons, à titre d'exemple, le système :

$$(15a) \qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

sur lequel on désire minimiser le critère :

$$(15b) \qquad J = \frac{1}{2} \int_0^{\infty} (y^T y + u^T u)\, dt$$

L'équation de Riccati (3b) admet dans ce cas la solution :

$$K = \begin{bmatrix} 1.3865 & 0.6056 \\ 0.6056 & 0.2674 \end{bmatrix}$$

ce qui fournit le système optimal en boucle fermée :

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3.6056 & -4.2674 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

L'agrégation du système initial par le vecteur d'observation $C = [\,2 \quad 1\,]$, conduit à l'équation d'état sur la sortie :

$$\dot{y} = -1.6y + u$$

qui, associé au critère (15b), demande la résolution d'une équation de Riccati scalaire dont la solution est $\tilde{K} = 0.2868$. Le système sous-optimal obtenu est régi par l'équation d'état en boucle fermée :

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3.5736 & -4.2868 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

## I.2 COMMANDE SOUS-OPTIMALE PAR LES METHODES DE PERTURBATION

### I.2.1. UTILISATION DES PERTURBATIONS REGULIERES | BER,82 |

Dans le cas où le système se présente sous la forme régulièrement perturbée :

$$(16)\qquad \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} A_{11} & \varepsilon A_{12} \\ \varepsilon A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

où $\varepsilon$ est un petit paramètre

et $x_1 \in \mathbb{R}^{q1}$, $x \in \mathbb{R}^{q2}$, $x^T = | x_1^T, x_2^T |$, $u_1 \in \mathbb{R}^{m1}$, $u_1 \in \mathbb{R}^{m2}$

la minimisation du critère

$$(17)\qquad J = \frac{1}{2} \int_0^{\infty} (x^T Q x + u^T R u)\, dt, \quad u = | u_1^T, u_2^T |^T$$

où $Q = \left[\begin{array}{c:c} Q_1 & 0 \\ \hdashline 0 & Q_2 \end{array}\right]$ définie non négative

$R = \left[\begin{array}{c:c} R_1 & 0 \\ \hdashline 0 & R_2 \end{array}\right]$ définie positive

conduit à la commande optimale

$$(18)\quad \begin{cases} (18a)\quad \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = - \begin{bmatrix} R_1^{-1} B_1^T & 0 \\ 0 & R_2^{-1} B_2^T \end{bmatrix} K(\varepsilon) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \qquad \text{où } K(\varepsilon) \text{ est solution de l'équation de Riccati :} \\ (18b)\quad K(\varepsilon) \begin{bmatrix} A_{11} & \varepsilon A_{12} \\ \varepsilon A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} A_{11}^T & \varepsilon A_{21}^T \\ \varepsilon A_{12}^T & A_{22}^T \end{bmatrix} K(\varepsilon) \end{cases}$$

$$-K(\varepsilon) \begin{bmatrix} B_1 R_1^{-1} B_1^T & 0 \\ 0 & B_2 R_2^{-1} B_2^T \end{bmatrix} K(\varepsilon) + \begin{bmatrix} Q_{11} & 0 \\ 0 & Q_{22} \end{bmatrix} = 0$$

$K(\varepsilon)$, solution de l'équation (18b), peut être approchée par une série de Taylor en $\varepsilon$ :

$$(19) \qquad K(\varepsilon) = \sum_{i=0}^{\infty} K_i \varepsilon^i$$

et l'utilisation de cette expression dans (18b) permet, par identification des termes en facteur de $\varepsilon^i$ de déterminer de façon itérative les matrices constantes $K_i$. $K_0$ est alors la matrice que l'on obtiendrait si le système était parfaitement découplé ( $\varepsilon = 0$).

## I.2.2 UTILISATION DES PERTURBATIONS SINGULIERES

L'optimisation des systèmes singulièrement perturbés a été largement étudiée (Cf. les articles de synthèse | KOK,76 ; SAK,83 |). L'approche de la commande quasi-optimale de ces systèmes peut se faire par deux voies différentes : la première résulte d'une décomposition de l'équation de Riccati globale ; la seconde est basée sur le principe de décomposition temporelle développée dans les chapitres précédents et conduit à une complète séparation des problèmes d'optimalité.

### I.2.2.1 Décomposition de l'équation de Riccati | KOK,83 |

Considérons le système singulièrement perturbé

$$(20) \quad \begin{cases} \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \; ; \; x(o) = x_o \; ; \; z(o) = z_o \\ y = C_1 x + C_2 z \quad ; \; x \in \mathbb{R}^{q_1} \; ; \; z \in \mathbb{R}^{q_2} \end{cases}$$

Le critère à minimiser étant (2), la solution de ce problème est donnée par (3) où $B^T = | \; B_1^T \quad B_2^T/\varepsilon \; |$ et où l'équation de Riccati se réécrit sous la forme :

$$(21) \quad K \begin{bmatrix} A_{11} & A_{12} \\ A_{21}/\varepsilon & A_{22}/\varepsilon \end{bmatrix} + \begin{bmatrix} A_{11}^T & A_{21}^T/\varepsilon \\ A_{12}^T & A_{22}^T/\varepsilon \end{bmatrix} K - K \begin{bmatrix} B_1 R^{-1} B_1^T & B_1 R^{-1} B_2^T/\varepsilon \\ B_2 R^{-1} B_1^T/\varepsilon & B_2 R^{-1} B_2^T/\varepsilon^2 \end{bmatrix} K$$
$$+ \begin{bmatrix} C_1 & C_2 \end{bmatrix}^T Q \begin{bmatrix} C_1 & C_2 \end{bmatrix} = 0$$

Pour éviter une divergence de la solution de cette équation lorsque $\varepsilon$ tend vers 0, on propose | KOK,72 | la forme suivante pour K :

$$(22) \qquad K = K(\varepsilon) = \begin{bmatrix} K_{11}(\varepsilon) & \varepsilon K_{12}(\varepsilon) \\ \varepsilon K_{12}^{T}(\varepsilon) & \varepsilon K_{22}(\varepsilon) \end{bmatrix}$$

dans laquelle $K_{ij}(\varepsilon)$ sont des matrices de dimensions convenables dépendant de façon analytique de $\varepsilon$ :

$$(23) \qquad \forall (i, j) \in \{1,2\}^2 \qquad K_{ij}(\varepsilon) = \sum_{k=0}^{\infty} K_{ij}^{k} \cdot \varepsilon^{k}$$

L'équation (21) est alors équivalente à un système de trois équations matricielles où, pour simplifier les expressions, l'argument $\varepsilon$ est omis :

$$(24)\quad \begin{aligned} a)\ & K_{11}A_{11} + A_{11}^{T}K_{11} + K_{12}A_{21} + A_{21}^{T}K_{12}^{T} - K_{11}B_1R^{-1}B_1^{T}K_{11} - K_{11}B_1R^{-1} \\ & B_2^{T}K_{12} - K_{12}B_2R^{-1}B_1^{T}K_{11} - K_{12}B_2R^{-1}B_2^{T}K_{12} + C_1^{T}QC_1 = 0 \\ b)\ & K_{11}A_{12} + K_{12}A_{22} + A_{21}^{T}K_{22} - K_{11}B_1R^{-1}B_2^{T}K_{22} - K_{12}B_2R^{-1}B_2^{T}K_{22} + \\ & C_1^{T}QC_2 + \varepsilon\{A_{11}^{T}K_{12} - K_{11}B_1R^{-1}B_1^{T}K_{12} - K_{12}B_2R^{-1}B_1^{T}K_{12}\} = 0 \\ c)\ & K_{22}A_{22} + A_{22}^{T}K_{22} - K_{22}B_2R^{-1}B_2^{T}K_{22} + C_2^{T}QC_2 \\ & + \varepsilon\{A_{12}^{T}K_{12} + K_{12}^{T}A_{12} - K_{12}^{T}B_1R^{-1}B_2^{T}K_{22} - K_{22}B_2R^{-1}B_1^{T}K_{12}\} \\ & + \varepsilon^{2}\{-K_{12}^{T}B_1R^{-1}B_1^{T}K_{12}\} = 0 \end{aligned}$$

De ces expressions, par identification des termes en $\varepsilon^{i}$, $i \in \mathbb{N}$, on déduit les coefficients des expressions asymptotiques (23) donnant $K(\varepsilon)$. On se contente, dans la plupart des cas, de la connaissance des $K_{ij}$ à l'ordre 0 ; ces matrices étant obtenues en faisant $\varepsilon = 0$ dans (24). On est alors ramené à la résolution d'une équation de Riccati d'ordre $q_2$ :

$$(25) \qquad K_{22}^{o}A_{22} + A_{22}^{T}K_{22}^{o} - K_{22}^{o}M_2M_2^{T}K_{22}^{o} + N_2^{T}N_2 = 0$$

où $M_2M_2^{T} = B_2R^{-1}B_2^{T}$ et $N_2^{T}N_2 = C_2^{T}QC_2$,

qui, si $(A_{22}, M_2, N_2)$ est stabilisable-détectable, admet une solution unique définie positive. $K_{12}^{o}$ et $K_{11}^{o}$ peuvent alors être déterminées

par la résolution de l'équation de Riccati :

$$(26) \qquad K_{11}^{o}\,\bar{A} + \bar{A}\,K_{11}^{o} - K_{11}^{o}\bar{B}R^{-1}\bar{B}^{T}K_{11}^{o} + \bar{Q} = 0$$

$$\text{où } \bar{A} = A_{11} + E_1A_{21} + B_1R^{-1}B_2{}^{T}E_2{}^{T} + E_1B_2\,R^{-1}B_2{}^{T}E_2{}^{T}$$

$$\bar{B} = B_1 + E_1B_2$$

$$\bar{Q} = -E_2A_{21} - A_{21}{}^{T}E_2{}^{T} - E_2B_2R^{-1}B_2{}^{T}E_2{}^{T} + C_1{}^{T}QC_1$$

$$E_1 = (B_1R^{-1}B_2{}^{T}K_{22}{}^{o} - A_{12})(A_{22} - B_2R^{-1}B_2{}^{T}K_{22}{}^{o})^{-1}$$

$$E_2 = (A_{21}{}^{T}K_{22}{}^{o} + C_1{}^{T}QC_2)(A_{22} - B_2R^{-1}B_2{}^{T}K_{22}{}^{o})^{-1}$$

et par :

$$(27) \qquad K_{12}{}^{o} = K_{11}{}^{o}E_1 - E_2$$

le système commandé admet pour équation d'état :

$$(28) \qquad \begin{aligned} \dot{x} &= |A_{11} - S_{11}K_{11} - S_{12}K_{12}{}^{T}|x + |A_{12} - S_{12}K_{22} - \varepsilon S_{11}K_{12}\,|z \\ S_{11} &= B_1R^{-1}B_1{}^{T} \qquad S_{12} = B_1R^{-1}B_2{}^{T} \\ \varepsilon\dot{z} &= |A_{21} - S_{12}{}^{T}K_{11} - S_{22}K_{12}{}^{T}|x + |A_{22} - S_{22}K_{22} - \varepsilon S_{12}{}^{T}K_{12}|z \\ S_{22} &= B_2R^{-1}B_2{}^{T} \end{aligned}$$

dont la connaissance à l'ordre 0 substitue à la résolution d'une équation de Riccati d'ordre q la résolution de deux équations de Riccati d'ordres $q_1$ et $q_2$.

### I.2.2.2. Commande quasi-optimale par décomposition temporelle

Le point de vue adopté dans cette partie est l'utilisation de la décomposition temporelle de (20) en sous-systèmes lent et rapide. Les résultats présentés ici diffèrent légèrement de ceux obtenus classiquement par cette méthode | CHO,76 ; MAG,81|, car nous utiliserons la modélisation obtenue par la méthode de découplage (II :II.2.2.). En supposant que l'entrée du système est composée

d'une partie lente $u_l$ et d'une partie rapide $u_r$, le système lent découplé ne sera sensible qu'à $u_l$ alors que la sortie du système rapide sera la superposition de l'effet dû à $u_l$ et de l'effet dû à $u_r$. On obtient donc les équations d'état lentes et rapides et les sorties associées :

$$(29)\quad \begin{cases} l) & \begin{cases} \dot{x}_l = A_l x_l + B_l u_l \\ y_l = C_l x_l + D_l u_l \end{cases} \\ r) & \begin{cases} \varepsilon \dot{z}_r = A_{22} z_r + B_r u_r \\ y_r = C_2 z_r \end{cases} \end{cases}$$

avec $A_l = A_{11} - A_{12}A_{22}^{-1}A_{21}$

$$* \quad \begin{cases} B_l = B_1 - A_{12}A_{22}^{-1}B_2 \\ C_l = C_1 - C_2 L \\ D_l = - C_2 A_{22}^{-1} B_r \\ B_r = (I_{q2} - \varepsilon L A_{12} A_{22}^{-1}) B_2 + \varepsilon L B_1 = B_2 + \varepsilon L B_l \end{cases}$$

où L vérifie l'équation de Lyapounov :

$$A_{21} - A_{22}L + \varepsilon L A_l = 0$$

et la partie lente de la variable d'état rapide, z, est reconstruite à partir de

$$(30)\qquad z_l = -A_{22}^{-1} B_r u_l - L x_l$$

L'utilisation d'une commande globale composée d'un retour d'état lent et d'un retour d'état rapide est justifié à l'aide de la proposition : si les commandes

$$(31)\quad \begin{cases} u_l = G_l x_l \\ u_r = G_r z_r \end{cases}$$

sont appliqués sur les systèmes (29l) et (29r) ; et si la commande composite :

$$(32)\qquad u_c = \left[ (I_m + G_r A_{22}^{-1} B_r) G_l + G_r L \right] x + G_r z$$

est appliquée au système (20) avec l'hypothèse $A_{22} + B_r G_r$ et $A_l + B_l G_l$ stables, alors $\forall\, t \in \left| 0, + \infty \right|$ :

$$
\begin{aligned}
x(t) &= x_1(t) + 0(\varepsilon) \\
z(t) &= -\,|\,A_{22}^{-1}B_r G_1 + L\,|x_1(t) + z_r(t) + 0(\varepsilon) \\
u(t) &= u_1(t) + u_r(t) + 0(\varepsilon) \qquad (33) \\
y(t) &= y_1(t) + y_r(t) + 0(\varepsilon)
\end{aligned}
$$

On peut remarquer que lorsqu'on pose $\varepsilon=0$ dans cette proposition, il vient de (29*), $L = A_{22}^{-1}A_{21}$ et l'on retrouve ainsi le lemme proposé dans | CHO,76 |. La commande pouvant donc être séparée en une partie lente et une partie rapide, le problème d'optimisation du critère (2) va donc être décomposé en deux problèmes :

* problème lent : minimisation de $J_1 = \frac{1}{2}\int_0^\infty (y_1^T Q y_1 + u_1^T R u_1)dt$ pour le système lent (29-1)

* problème rapide : minimisation de $Jr = \frac{1}{2}\int_0^\infty (y_r^T Q y_r + u_r^T R u_r)dt$ pour le système rapide (29-r)

La solution optimale du problème lent est donnée, si $(A_1, B_1, SC_1)$ est un triplet stabilisable-détectable, par | AND,71 | :

$$
\begin{aligned}
u_1 &= -R_1^{-1}(D_1^T Q C_1 + B_1^T K_1)x_1 \qquad (34) \\
&\text{où } R_1 = R + D_1^T Q D_1 \\
&\text{et } K_1 \text{ vérifie l'équation de Riccati :}
\end{aligned}
$$

$$
(A_1 - B_1 R_1^{-1} D_1^T Q C_1)^T K_1 + K_1(A_1 - B_1 R_1^{-1} D_1^T Q C_1) - K_1 B_1 R_1^{-1} B_1^T K_1
$$

$$
+ C_1^T S^T (I - S D_1 R_1^{-1} D_1^T S^T) SC_1 = 0 \qquad \text{où } Q = S^T S .
$$

et celle du problème rapide, si $(A_{22}, B_r, SC_2)$ est stabilisable-détectable, est donnée par :

$$
(35)\begin{cases} u_r = -R^{-1}B_r^T K_r z_r \\ \text{où } K_r \text{ vérifie l'équation de Riccati} \\ K_r A_{22} + A_{22}^T K_r - K_r B_r R^{-1} B_r^T K_r + C_2^T Q C_2 = 0 \end{cases}
$$

La commande composite obtenue alors est d'après (32) :

$$
\begin{aligned}
u_c = -\Big(&\Big[I - R^{-1}B_r^T K_r A_{22}^{-1} B_r\Big] R_1^{-1} \Big[D_1^T Q C_1 + B_1^T K_1\Big] + R^{-1}B_r^T K_r L\Big)x \qquad (36) \\
&- R^{-1}B_r^T K_r z
\end{aligned}
$$

On peut remarquer que si l'on fait $\varepsilon=0$, on retrouve les résultats énoncés dans | CHO,76 |, c'est-à-dire, la commande composite qui y est proposé, est une approximation à l'ordre 0 de (36) ; la commande composite (36) permet, sans augmenter la taille des calculs, de tenir compte de l'influence de $\varepsilon$. De façon à mettre en évidence l'intérêt de la méthode par découplage, on s'intéressera à la quasioptimalité de (36).

D'après (34), on a l'identité

$$(37) \qquad \left[I-R^{-1}B_r{}^T K_r A_{22}^{-1} B_r\right]\left[R + D_1{}^T Q D_1\right]^{-1} = R^{-1}\left[I+B_r{}^T A K_r B_r R^{-1}\right]$$

où $A = \left[A_{22} - B_r R^{-1} B_r{}^T K_r\right]^{T-1}$

ce qui permet de réécrire la commande composite sous la forme :

$$(38) \qquad u_c = R^{-1}\left[B_1{}^T K_1 x + B_r{}^T K_m x + B_r{}^T K_r z\right]$$

avec $K_m = K_r L + A\left[-C_2{}^T Q C_1 + K_r B_r R^{-1} B_1{}^T K_1\right]$

Compte tenu des expressions de $B_1$ et $B_r$ (29), (38) se met sous la forme compacte :

$$(39) \qquad u_c = -R^{-1}\left[B_1{}^T \vdots B_2{}^T/\varepsilon\right]\begin{bmatrix} K_1 & \varepsilon K_3 \\ \varepsilon K_2 & \varepsilon K_4 \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix}$$

où $K_1 = K_1 + \varepsilon L^T K_m$ , $K_3 = L^T K_r$

$K_2 = K_m - A_{22}^{T-1} A_{12}^T K_1 - \varepsilon A_{22}^{-1} {}^T A_{12}$ , $K_4 = (I - \varepsilon A_{22}^{-1T} A_{12}^T L^T) K_r$

$L^T K_m$

Mais, si on fait dans les expressions précédentes une approximation à l'ordre 0, on obtient :

$$(40) \qquad K_1^\circ = K_1^\circ = K_{11}^\circ \; ; \; K_{11}^\circ = K_r^\circ = K_{22}^\circ \; ; \; K_3^\circ = A_{21}^T A_{22}^{-1T} K_{22}^\circ$$

$$K_2^\circ = K_m^\circ - A_{22}^{-1T} A_{12}^T K_1^\circ$$

$$= A\left|K_r^\circ B_2 R^{-1} B_1{}^T K_1^\circ - A_{12}{}^T K_1^\circ - K_2^\circ A_{21} - C_2{}^T Q Q_1\right| = K_{12}^{T\circ}$$

où les $K_{ij}^o$, i, j = 1,2, sont définis par (25 - 26 - 27) et les manipulations effectuées pour obtenir ce résultat sont indiquées dans | CHO,75 |.

Ainsi, d'après | CHO,76 |, la valeur optimale du critère (2) par la commande (36) est approché à un $o(\varepsilon^2)$ près. La considération de modèles découplés conduit donc à obtenir une matrice-gain de retour

d'état (39) de même forme que celle proposée dans (II.2.2.1) et que l'on peut écrire sous la forme d'une série analytique en $\varepsilon$ .
Nous rappelons que la matrice de gain obtenue dans | CHO,76 | était :

$$(41) \qquad \begin{bmatrix} K_{11}^{\circ} & 0 \\ \varepsilon K_{12}^{T\circ} & \varepsilon K_{22}^{\circ} \end{bmatrix}$$

Le retour d'état déterminé par (41), donc la commande composite, est indépendant de $\varepsilon$ ; par contre, la commande composite (36) permet, sans temps de calcul supplémentaire et en conservant l'indépendance des problèmes lents et rapides, de faire intervenir la valeur de $\varepsilon$ . Il serait bien sur possible de ne considérer qu'une commande réduite à partir des variables d'état lentes, la valeur optimale du critère serait alors approchée à un $o(\varepsilon)$ près.

De façon à mettre en évidence l'intérêt du préconditionnement par la méthode de découplage du système à optimiser, considérons l'exemple traité dans | CHO,76 | :

$$(42) \qquad \begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \varepsilon \dot{z}_1 \\ \varepsilon \dot{z}_2 \end{bmatrix} = \left[\begin{array}{cc:cc} 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.345 & 0 \\ \hdashline 0 & -0.524 & -0.465 & 0.262 \\ 0 & 0 & 0 & -1 \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \\ z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u \\ y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ z_1 \\ z_2 \end{bmatrix} \end{cases}$$

où $\varepsilon = 0.1$ ; $[x_1, x_2, z_1, z_2]_{t=0} = [1, 0, 1, 0]$

où le critère à minimiser était :

$$(43) \qquad J = \frac{1}{2} \int_0^{\infty} (y^T y + u^T u) \, dt$$

La comparaison des coûts obtenus donne :

(44)
* retour d'état optimal : $J_{opt}$ : 4.2406
* commande composite | CHO,76 | : $J_c$ = 4.2428
* découplage puis commande composite : $J_{dc}$ = 4.2421

On peut en déduire que, sans nécessiter de calculs supplémentaires au niveau de la détermination de la commande sous-optimale, la méthode du découplage permet d'améliorer la précision dans la valeur du critère. En effet, cette méthode permet de tenir compte du petit paramètre $\varepsilon$ dans les systèmes réduits, ce qui n'est pas possible par la méthode usuelle des perturbations singulières.

### I.2.2.3. Retour de sortie

Nous allons appliquer dans cette partie la méthode de construction d'une commande quasioptimale par retour de sortie (I-1) pour le système singulièrement perturbé (20). En supposant $|C_1, C_2|$ de rang maximal, d'après (5) et (20), il vient une équation d'état approchant la sortie :

$$(45)\qquad \varepsilon\overset{\circ}{\tilde{y}} = (\tilde{A}_o + \varepsilon\tilde{A}_1)\tilde{y} + (\tilde{B}_o + \varepsilon\tilde{B}_1)u$$

$$\text{avec } \tilde{A}_o = C_2\left[A_{21}C_1^T + A_{22}C_2^T\right]\left[C_1C_1^T - C_2C_2^T\right]^{-1}$$

$$\tilde{A}_1 = C_1\left[A_{11}C_1^T + A_{12}C_2^T\right]\left[C_1C_1^T - C_2C_2^T\right]^{-1}$$

$$\tilde{B}_o = C_2B_2$$

$$\tilde{B}_1 = C_1B_1$$

On est donc en présence d'un système singulier singulièrement perturbé (II : II.2.5) et si $A_o$ a un noyau différent de 0 ou de $\mathbb{R}^n$, on peut appliquer les changements de variables mettant en évidence les parties lentes et rapides de y. Dans tous les cas, au problème optimal de minimisation du critère (2) est substitué le problème de minimisation du critère :

$$(46)\qquad \tilde{J} = \frac{1}{2}\int_o^\infty (\tilde{y}^T Q\tilde{y} + u^T Ru)dt$$

dont la solution est

$$(47)\qquad \begin{cases} u = -R^{-1}(\tilde{B}_o + \varepsilon\tilde{B}_1)^T\dfrac{\tilde{K}}{\varepsilon}\tilde{y} \\ \text{où } \tilde{K} \text{ vérifie l'équation de Riccati :} \\ \tilde{K}(\tilde{B}_o + \varepsilon\tilde{B}_1)R^{-1}(\tilde{B}_o^T + \varepsilon\tilde{B}_1^T)\tilde{K} = Q\varepsilon^2 + \varepsilon(\tilde{K}(\tilde{A}_o + \varepsilon\tilde{A}_1) + (\tilde{A}_o^T + \varepsilon\tilde{A}_1^T)\tilde{K}) \end{cases}$$

De façon à construire l'expression analytique de $\tilde{K}$, on peut réécrire cette équation sous la forme :

$$(48)\qquad -\tilde{K}\tilde{B}_o R^{-1}\tilde{B}_o^T\tilde{K} + \varepsilon\left[\tilde{K}\tilde{A}_o + \tilde{A}_o^T\tilde{K} - \tilde{K}(\tilde{B}_1 R^{-1}\tilde{B}_o^T + \tilde{B}_o R^{-1}\tilde{B}_1^T)\tilde{K}\right]$$
$$+ \varepsilon^2\left[\tilde{K}\tilde{A}_1 + \tilde{A}_1^T\tilde{K} + Q - \tilde{K}\tilde{B}_1 R^{-1}\tilde{B}_1^T\tilde{K}\right] = 0$$

en posant

$$(49)\qquad \tilde{K} = \sum_{i=0}^{\infty} \tilde{K}_i\, \varepsilon^i$$

il vient, par identification dans (48), les équations donnant les premiers coefficients matriciels :

$$(50)\qquad \begin{cases} \tilde{K}_o = 0 \\ \tilde{K}_1\tilde{A}_o + \tilde{A}_o^T\tilde{K}_1 - \tilde{K}_1\tilde{B}_o R^{-1}\tilde{B}_o^T\tilde{K}_1 + Q = 0 \\ \tilde{K}_2(\tilde{A}_o - \tilde{B}_o R^{-1}\tilde{B}_o^T\tilde{K}_1) + (\tilde{A}_o - \tilde{B}_o R^{-1}\tilde{B}_o^T\tilde{K}_1)^T\tilde{K}_2 \\ - \tilde{K}_1(\tilde{B}_1 R^{-1}\tilde{B}_o^T + \tilde{B}_o R^{-1}\tilde{B}_1^T)\tilde{K}_1 + \tilde{K}_1\tilde{A}_1 + \tilde{A}_1^T\tilde{K}_1 = 0 \end{cases}$$

et de façon générale, seul $K_1$ demande la résolution d'une équation de Riccati d'ordre n, la détermination de $K_i$ ne nécessitant que la résolution d'une équation de Lyapounov de même ordre. En posant

$$(51)\qquad \begin{cases} S_o = \tilde{B}_o R^{-1}\tilde{B}_o^T \\ S_{o1} = \tilde{B}_1 R^{-1}\tilde{B}_o^T + \tilde{B}_o R^{-1}\tilde{B}_1^T \\ S_1 = \tilde{B}_1 R^{-1}\tilde{B}_1^T \end{cases}$$

cette équation de Lyapounov s'écrit :

$$(52)\qquad \forall\, i > 2 \begin{cases} \tilde{K}_i(\tilde{A}_o - S_o\tilde{K}_1) + (\tilde{A}_o - S_o\tilde{K}_1)^T\tilde{K}_i \\ - \sum_{k=1}^{i-2} \tilde{K}_{i-k}\, S_o\, \tilde{K}_{k+1} \\ - \sum_{k=1}^{i-1} \tilde{K}_{i-k}\, S_{o1}\tilde{K}_k \\ - \sum_{k=2}^{i-1} \tilde{K}_{i-k}\, S_1\, \tilde{K}_{k-1} + \tilde{K}_{i-1}\,\tilde{A}_1 + \tilde{A}_1^T\tilde{K}_{i-1} = 0 \end{cases}$$

Le retour de sortie quasi optimal à appliquer au système (20) sera donc donné par :

$$\tilde{u} = -R^{-1} \left[ \sum_{i=0}^{\infty} G_i \, \varepsilon^i \right] y \qquad (53)$$

$$\text{où } G_i = B_2^T C_2^T \tilde{K}_{i+1} + B_1^T C_1^T \tilde{K}_i, \ \forall i \in \mathbb{N}$$

## II. COMMANDE DES SYSTEMES NON LINEAIRES

Du fait de leur intérêt pour la modélisation des systèmes non linéaires, la commande optimale des systèmes bilinéaires a donné lieu à de nombreux travaux | MOH,73a |. Alors que dans le cas de présences de contraintes sur les variables d'entrée,une commande de type bang-bang permet de conclure | MOH,73b |, dans le cas d'une optimisation sans contraintes, l'approche est plus délicate | SAG,76 |. La difficulté vient du fait qu'il n'y a pas de connection triviale entre l'état du système et l'état adjoint impliqué dans l'Hamiltonien ; les lois de commandes optimales sont donc déterminées en boucle ouverte ou fonction de l'état adjoint | SAG,76 ; NAI,81 ; DIB,82 ; BAN,86b |. Cependant, dans certains cas particuliers, quelques auteurs ont déterminé des retours d'état linéaires localement stabilisantes | DER,80a | ou des retours d'état non linéaires de type commande quadratique (e.g. : $u(x) = -(P + Qx)^T(R + Sx)$) | JAC,77 ; SLE,78 |) ou à commande type diviseur (e.g. : $u(x) = - |F(x)|^{-1} G(x)$ où F ET G sont des polynomes matriciels de même degré). Une approche différente est proposée dans |GIL, 85| : elle consiste à proposer des retours d'état non-linéaires linéarisants. Ils permettent d'optimiser de très larges classes de critères et ont été appliqués sur les systèmes bilinéaires | GIL,83 | ou discrets | GIL,86 |.

Cependant, une procédure d'approximation a été préconisée dans | CEB,84a | qui permet, dans le cas de la minimisation d'un critère quadratrique,de générer une commande quasi-optimale par retour d'état. Cette technique a été employée pour la commande quasi-optimale de cuves chimiques | CEB,84b |.

Nous allons, dans un premier temps, étendre et détailler cette approche dans le cas des systèmes à commande affine. L'expression du retour d'état optimal pour un système bilinéaire sera exposée en tant que cas particulier dans la partie suivante.

## II.1. SYSTEMES A COMMANDE AFFINE ANALYTIQUE

Nous envisageons dans cette partie de résoudre un problème de commande optimale par retour d'état, le système étant défini par l'équation d'état :

$$(54)\quad \begin{cases} 54a) & \dot{x} = f(x) + G(x).u \;,\; x \in \mathbb{R}^q,\; u \in \mathbb{R}^m \\ & y = h(x) \qquad\quad ,\; y \in \mathbb{R}^n \\ & \text{où } f(.),\ G(.),\ h(.) \text{ sont des fonctions non linéaires de l'état supposées analytiques que l'on peut écrire sous la forme} \\ 54b) & f(x) = \sum_{i=1}^{\infty} F_i\, x^{[i]} \;;\; G(x) = [g_1(x), \ldots, g_m(x)] \\ & \forall\, i \in \{1, \ldots, m\}\ g_i(x) = \sum_{j=0}^{\infty} G_{ij}\, x^{[j]} \\ & h(x) = \sum_{i=1}^{\infty} H_i x^{[i]} \end{cases}$$

où les $F_i$, $G_{ij}$, $H_i$ sont des matrices constantes de dimensions adaptées et $x^{[i]}$ représente la $i^{ème}$ puissance de l'état introduite au (I : III-1) (i.e. : $x^{[i]} = \underbrace{x \otimes \ldots\ldots \otimes x}_{i \text{ fois}}$)

Le critère à minimiser, que nous rappelons ici, est :

$$(55)\qquad \tilde{J} = \frac{1}{2} \int_0^{\infty} (y^T Q y + u^T R u)\, dt$$

où Q et R vérifient les hypothèses indiquées en (2).

### II.1.1 CONDITIONS NECESSAIRES D'OPTIMALITE

De façon à résoudre le problème de commande optimale (54) (55), nous utilisons la méthode de programmation dynamique | BOU,69 | pour les systèmes continus. En notant $\hat{\mathcal{R}}(x,t)$ le revenu optimal correspondant à l'instant t à l'état initial x, la condition nécessaire d'optimalité se traduit par l'équation fonctionnelle d'Hamilton-Jacobi :

$$(56)\qquad \frac{\partial}{\partial t}\hat{\mathcal{R}}(x,t) = -\,\mathrm{opt}\left[\frac{1}{2}\left(h^T(x)Qh(x) + u^TRu\right) + \frac{\partial \hat{\mathcal{R}}^T}{\partial x}(x,t)\left(f(x) + G(x)u\right)\right] \qquad u \in \Omega$$

où $\Omega$ est le domaine de variations admissibles des entrées

Comme nous nous plaçons dans le cas particulier stationnaire sans contraintes et à horizon infini, on obtient :

$$(57)\qquad \begin{cases} 57a) & \dfrac{\partial \hat{\mathcal{R}}}{\partial t}(x,t) = 0 \\ 57b) & u^* = -R^{-1}G^T(x)\dfrac{\partial \hat{\mathcal{R}}}{\partial x}(x,t) \end{cases}$$

où $u^*$ représente la commande optimale

ce qui, replacé dans l'équation d'optimalité conduit à :

$$(58)\qquad \frac{1}{2}h^T(x)Qh(x) + \left[\frac{\partial \hat{\mathcal{R}}(x)}{\partial x}\right]^T .f(x) - \frac{1}{2}\left[\frac{\partial \hat{\mathcal{R}}(x)}{\partial x}\right]^T G(x)R^{-1}G^T(x)\left[\frac{\partial \hat{\mathcal{R}}(x)}{\partial x}\right] = 0$$

Le problème à résoudre étant maintenant de déterminer la solution $\partial \hat{\mathcal{R}}/\partial x$ de cette équation.

## II.1.2. EQUATION FONDAMENTALE D'APPROXIMATION

Un procédé naturel est de chercher l'expression analytique de cette solution ; cette démarche a été proposée d'abord dans | MER,64 | puis dans | LUK,69 |. L'approche nouvelle que nous proposons ici consiste dans l'utilisation systématique de la notation tensorielle. L'utilisation de l'algèbre tensorielle simplifiera notablement les calculs. Nous chercherons donc la solution sous la forme :

$$(59)\qquad \frac{\partial \hat{\mathcal{R}}}{\partial x}(x) = \sum_{i=1}^{\infty} P_i\, x^{[i]}$$

où les $P_i$ sont des matrices constantes

et nous aurons besoin de la propriété suivante :

$$(60) \qquad x^T A y = [\mathrm{vec}^T(A^T)]\ (x \otimes y) = [\mathrm{vec}^T(A)]\ (y \otimes x)$$

où vec(.) est la fonction vectorielle de matrice décrite en (II :55)

En utilisant dans (58) les expressions (54b) et (59), on obtient :

$$(61) \quad \Big(\sum_{j=1}^{\infty} x^{[i]} H_j^T\Big) Q \Big(\sum_{j=1}^{\infty} H_j x^{[i]}\Big) + \Big(\sum_{j=1}^{\infty} x^{[i]^T} P_j^T\Big)\Big(\sum_{j=1}^{\infty} F_j x^{[j]}\Big)$$

$$+ \Big(\sum_{j=1}^{\infty} x^{[j]^T} F_j^T\Big)\Big(\sum_{j=1}^{\infty} P_j x^{[j]}\Big)$$

$$- \sum_{k=1}^{m}\sum_{l=1}^{m}\Big\{\Big(\sum_{j=1}^{\infty} x^{[j]^T} P_j^T\Big)\Big(\sum_{j=0}^{\infty} G_{kj} x^{[j]}\Big) r_{kl} \Big(\sum_{j=0}^{\infty} x^{[j]^T} G_{lj}^T\Big)\Big(\sum_{j=1}^{\infty} P_j x^{[j]}\Big)\Big\} = 0$$

où $R^{-1} = [r_{kl}]$ $k \in \{1,\ldots,m\}$ , $l \in \{1,\ldots,m\}$

ce qui s'écrit, compte tenu de (60) :

$$(62) \quad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \mathrm{vec}^T(H_i^T Q H_j + P_i^T F_j + F_i^T P_j)\, x^{[i+j]}$$

$$- \sum_{k=1}^{m}\sum_{l=1}^{m}\Big\{\Big(\sum_{i=1}^{\infty}\sum_{j=0}^{\infty} x^{[i+j]^T} \mathrm{vec}(P_i^T G_{kj})\Big) r_{kl} \Big(\sum_{c=0}^{\infty}\sum_{d=1}^{\infty} \mathrm{vec}^T(P_d^T G_{lc})\, x^{[c+d]}\Big)\Big\} = 0$$

En notant :

$$(63) \qquad \forall i \in \mathbb{N}^*, \forall j \in \mathbb{N}, \forall k \in \{1,\ldots,m\}$$

$$\mathrm{vec}(P_i^T G_{kj}) = V_{ij}^k \ ; \ V_{ij} = [V_{ij}^1 \vdots \ldots \vdots V_{ij}^m]$$

on a la relation :

$$(64) \qquad \forall (i,d) \in \mathbb{N}^{*2}, \forall (j,c) \in \mathbb{N}^2$$

$$\sum_{k=1}^{m}\sum_{l=1}^{m} x^{[i+j]^T} V_{ij}^k r_{kl} V_{dc}^{l\,T} x^{[c+d]} = x^{[i+j]^T} V_{ij} R^{-1} V_{dc}^T x^{[c+d]}$$

$$= \mathrm{vec}^T(V_{ij} R^{-1} V_{dc}^T) x^{[i+j+c+d]}$$

Finalement, l'équation analytique qui détermine les $P_i$, $i \in \mathbb{N}^*$, s'écrit, à partir de (62) :

$$(65)\qquad \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \mathrm{vec}^T(H_i^T Q H_j + P_i^T F_j + F_i^T P_j)x^{[i+j]} - \sum_{k=1}^{\infty}\sum_{d=1}^{\infty}\sum_{l=0}^{\infty}\sum_{c=0}^{\infty} \mathrm{vec}^T(V_{kl}R^{-1}V_{dc}^T)x^{[k+l+c+d]} = 0$$

Les matrices $P_i, i \in \mathbb{N}^*$, seront déterminées en annulant dans (65) les coefficients de chacune des puissances de l'état.

## II.2 DETERMINATION DES Pi

### II.2.1. PREMIER ORDRE

$P_1$ est obtenue en choisissant dans (65), $(i,j,k,d,l,c) = (1,1,1,1,0,0)$, ce qui conduit à l'équation :

$$(66)\qquad \mathrm{vec}^T(H_1^T Q H_1 + P_1^T F_1 + F_1^T P_1) - \mathrm{vec}^T(V_{10}R^{-1}V_{10}^T) = 0$$

Mais l'opérateur vec(.) est linéaire pour des matrices de mêmes dimensions et de plus :

$$(67)\qquad \forall\, k \in \{1,\ldots,m\}\quad V_{10}^k = \mathrm{vec}(P_1^T G_{ko}) = P_1^T G_{ko}$$
$$\text{donc } V_{10} = P_1^T G_o$$
$$\text{où } G_0 = [G_{10} \vdots \cdots \vdots G_{m0}]$$

Ainsi, (66) est équivalente à l'équation de Riccati :

$$(68)\qquad H_1^T Q H_1 + P_1^T F_1 + F_1^T P_1 - P_1^T G_0 R^{-1} G_0^T P_1 = 0$$

Or, $P_1$ correspond à la solution du problème optimal sur le système linéarisé, elle est donc symétrique et si le triplet $(F_1, M_1, N_1)$ où $M_1$ et $N_1$ sont définies par :

$$(69)\qquad \begin{cases} G_0 R^{-1} G_0^T = M_1 M_1^T \\ H_1^T Q H_1 = N_1^T N_1, \end{cases}$$

est stabilisable-détectable, alors $P_1$ est unique, définie non négative et $(F_1 - G_0 R^{-1} G_0^T P_1)$ est de Hurwitz.

### II.2.2. DEUXIEME ORDRE

$P_2$ est obtenue par l'annulation dans (65) du terme en $x[3]$, ce qui donne :

$$(70)\qquad \begin{aligned} &\mathrm{vec}(H_1^T Q H_2 + P_1^T F_2 + F_1^T P_2) + \mathrm{vec}(H_2^T Q H_1 + P_2^T F_1 + P_2^T F_1) \\ &- \{\mathrm{vec}(V_{11} R^{-1} V_{10}^T) + \mathrm{vec}(V_{10} R^{-1} V_{11}^T) + \mathrm{vec}(V_{20} F^{-1} V_{10}^T) \\ &+ \mathrm{vec}\,(V_{10} R^{-1} V_{20}^T)\} = 0 \end{aligned}$$

Pour $P_1$ connue, les termes inconnus de (70) sont ceux regroupés dans $F_1^T P_2$ et $V_{20} R^{-1} V_{20}^T$ car :

$$(71)\qquad \forall\, k \in \{1,\ldots,m\}\quad V_{20}^{\,k} = \mathrm{vec}(P_2^T G_{k0}) \text{ donc } V_{20} = P_2^T G_0$$

Le Théorême 3.2 de | BRE,78 | indique :

$$(72)\qquad \forall\, A \in \mathbb{R}^{pxq},\ \mathrm{vec}(A^T) = E_{pxq}\,\mathrm{vec}(A)$$

où $E_{pxq}$ est une matrice de permutation définie par

$$E_{pxq} = \sum_{i=1}^{p} \sum_{k=1}^{q} (e_i^{(p)} e_k^{(q)T}) \otimes (e_k^{(q)} e_i^{(p)T})$$

où $e_i^{(p)}$ représente le $i^{ème}$ vecteur de base de dimension p. Donc, comme :

$$(73)\qquad \forall A \in \{V_{20}R^{-1}V_{10}^T,\ P_2^TF_1 + H_2^TQH_1 + F_2^TP_1,\ V_{11}R^{-1}V_{10}^T\}$$
$$\mathrm{vec}(A) = E_{qxq}^2\,\mathrm{vec}(A^T)$$

(70) est équivalent à :

$$(74)\qquad U_{qxq}^2\ \{\mathrm{vec}(H_1^TQH_2 + P_1^TF_2 + F_1^TP_2) - \mathrm{vec}(V_{10}R^{-1}V_{11}^T)$$
$$- \mathrm{vec}\,(V_{10}R^{-1}V_{20}^T)\} = 0$$
$$\text{où } U_{qxq}^2 = I_{qxq}^2 + E_{qxq}^2$$

La régularité de $U_{qxq}2$, pour tout q et la linéarité de vec(.) sur des matrices de mêmes dimensions impliquent que $P_2$ soit solution de :

$$(75)\qquad (H_1^TQH_2 + P_1F_2 - V_{10}R^{-1}V_{11}^T) + (F_1^T - P_1G_0R^{-1}G_0^T)P_2 = 0$$

Si le système linéarisé vérifie les conditions de résolution de l'équation de Riccati (68), alors $(F_1^T - P_1G_0R^{-1}G_0^T)$ est inversible en tant que matrice de Hurwitz. On obtient alors directement l'expression de $P_2$ par :

$$(76)\qquad P_2 = -\mathcal{F}^{-1}\mathcal{H}_1$$
$$\text{où } \mathcal{F} = F_1^T - P_1G_0R^{-1}G_0^T$$
$$\mathcal{H}_1 = H_1^TQH_2 + P_1F_2 - P_1G_0R^{-1}V_{11}^T$$
$$\text{avec } V_{11} = \left[V_{11}^1 \vdots \cdots \vdots V_{11}^m\right]$$
$$\forall i \in \{1,\ldots,m\}\quad V_{11}^i = \mathrm{vec}(P_1G_{1i})$$

Ainsi, lorsque $P_1$, la matrice gain du premier ordre du retour d'état est connue, la connaissance de la matrice $P_2$ du second ordre ne demande pas d'hypothèses supplémentaires. L'équation à résoudre étant simplement une équation linéaire.

Cette remarque est également vraie pour les matrices correspondant aux ordres dupérieurs $P_k$, $k > 3$. Nous présentons en annexe la méthode de détermination de ces matrices, qui utilise le

même principe que pour les ordres 1 et 2, mais qui demande des manipulations algébriques plus élaborées. En fait, la conclusion reste la même quel que soit l'ordre : l'équation déterminant la matrice cherchée est linéaire, la solution est alors obtenue par l'inversion de la même matrice $\mathfrak{F}$ .

## II.3 EXPRESSION ANALYTIQUE DU RETOUR D'ETAT

Suivant (57b), le retour d'état optimal ainsi construit s'écrit :

$$(77) \qquad u^* = -R^{-1} \begin{bmatrix} g_1(x)^T \\ \vdots \\ g_m(x)^T \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{\infty} P_i \, x^{[i]} \end{bmatrix}$$

or, d'après (54b) :

$$(78) \qquad \forall\, k \in \{1,\dots,m\} \quad g_k^T(x) \left[ \sum_{i=1}^{\infty} P_i x^{[i]} \right] = \sum_{j=0}^{\infty} x^{[j]^T} G_{kj}^T \left[ \sum_{i=1}^{\infty} P_i x^{[i]} \right] = \sum_{p=1}^{\infty} K_{kp} \, x^{[p]}$$

$$\text{où } K_{kp} = \sum_{i=1}^{p} \mathrm{vec}^T(G_{k,p-i}^T \, P_i)$$

Ainsi, si le système (54) linéarisé, associé au critère (55), est stabilisable détectable, la commande optimale a pour expression :

$$(79) \qquad u^* = \sum_{i=1}^{\infty} U_i x^{[i]}$$

$$\text{où } \forall i \in \mathbb{N}^*, \; U_i = -R^{-1} \begin{bmatrix} K_{1i} \\ \vdots \\ K_{mi} \end{bmatrix}$$

## II.4 CAS DES SYSTEMES SCALAIRES

Lorsque la dimension du système étudié est égale à 1, il n'y a plus lieu de faire appel à la méthode précédente, car l'expression non linéaire du retour d'état peut être directement obtenue. En effet, considérons l'équation d'état :

$$(80) \quad \begin{cases} \dot{x} = f^*(x)x + ug(x) \\ y = h^*(x)x \end{cases}$$

où $u \in \mathbb{R}$, $x \in \mathbb{R}$, $f^*(x)x = f(x)$, $h^*(x)x = h(x)$

et le critère à minimiser :

$$(81) \quad J = \frac{1}{2} \int_0^\infty \left[ qy^2 + ru^2 \right] dt \quad , r > 0, q \geqslant 0$$

les équations d'optimalité deviennent alors en posant, $\partial \hat{\mathcal{R}}/\partial x = p^*(x)x$ :

$$(82) \quad u^* = -\frac{g(x)}{r} p^*(x)x, \text{ où } p^*(x) \text{ est solution de}$$

$$(83) \quad qh^*(x)^2 + 2f^*(x)p^*(x) - \frac{1}{r} g^2(x)\, p^*(x)^2 = 0$$

La seule solution admissible de (83) est :

$$(84) \quad p^*(x) = \frac{r}{g^2(x)} \left[ f^*(x) + \sqrt{f^*(r)^2 + \frac{q}{r} h^*(x)^2 g^2(x)} \right]$$

et le système obtenu en boucle fermée devient :

$$(85) \quad \dot{x} = - \left[ \sqrt{f^*(x)^2 + \frac{q}{r} h^*(x)^2 g^2(x)} \right] x$$

Le retour d'état optimal est, d'après (82-84), donné par :

$$(86) \quad u^*(x) = - \left[ v^*(x) + \text{signe}(g(x)) \sqrt{v^*(x)^2 + \frac{q}{r} h^*(x)^2} \right] x$$

où $v^*(x) = f^*(x)/g(x)$

La remarque que l'on peut tirer de ce paragraphe concerne l'écriture analytique de $u^*(x)$ et son rayon de convergence. L'expression (84) montre que le problème vient ici de l'inversion

de g(x). La commande (86) reste admissible si l'on demeure dans un domaine autour de o où g(x) ne s'annule pas. De façon plus générale, dans le cas où q est quelconque, on peut conjecturer que l'expression analytique (79) converge vers la commande optimale réelle si l'on se trouve pendant tout le mouvement (en particulier à l'instant initial) dans un domaine entourant l'origine où aucun des coefficients de G(x) ne s'annule. Cependant, il est possible d'augmenter le rayon de convergence de la série analytique obtenue pour la commande par un certain nombre de techniques comme, par exemple, l'utilisation des approximants de Padé | GIL,78|.

## II.5 TRANSMISSION DIRECTE

Lorsque le système non linéaire présente une transmission directe de l'information, c'est-à-dire, que l'équation d'état se présente sous la forme (54) où l'équation de sortie est remplacée par

(87) $$y = h(x) + D(x)u$$

où D(x) est une fonction non linéaire analytique de l'état : $D(x) = \left[d_1(x), \ldots, d_m(x)\right]$

$$\forall\, i \in \{1,\ldots,m\} \quad d_i(x) = \sum_{j=0}^{\infty} D_{ij}\, x^{[j]} ,$$

nous allons montrer qu'un simple changement de variable, de la même façon que le cas linéaire, permet de se ramener à un problème sans transmission directe.

D'après (87), le critère (55) à optimiser s'écrit :

(88) $$J = \frac{1}{2} \int_0^{\infty} \left[ h^T(x)\, Qh(x) + 2h^T QD(x)u + u^T Ru \right] dt$$

et si l'on change le vecteur de commande par :

(89) $$u = v - R^{-1}D^T(x)Qh(x)$$

alors J devient :

$$(90) \qquad J = \frac{1}{2} \int_0^{\infty} \left[ h^T(x)S(x)h(x) + v^T R v \right] dt$$

où $S(x) = Q \left[ I + D(x)D^T(x)Q \right]$

L'équation d'état se met sous la forme :

$$(91) \qquad \dot{x} = \phi(x) + G(x)v$$

avec $\phi(x) = f(x) - G(x)R^{-1}P^T(x)Qh(x)$

$$= \sum_{j=1}^{\infty} \phi_j x^{[j]}$$

Ainsi, pour une matrice de transmission directe constante, on peut appliquer directement les résultats présentés dans les paragraphes précédents ; dans le cas le plus général, une légère modification des formules (changement des termes $H_i^T Q H_j$) permet de déterminer une expression analytique du retour d'état optimal $v^*$, donc de $u^*$ à partir de (89).

## III - CAS DES SYSTEMES BILINEAIRES

### III.1 CAS GENERAL : RETOUR D'ETAT ANALYTIQUE

Les systèmes bilinéaires sont un cas particulier des systèmes analytiques étudiés dans la partie II. En effet, leur équation d'état est obtenue en prenant dans (54) :

$$(92) \qquad \begin{aligned} & F_1 = A_o \in \mathbb{R}^{qxq} \\ & \forall i > 1 \; F_i \equiv 0 \\ & H_1 = C^T \in \mathbb{R}^{nxq} \\ & \forall i > 1 \; H_i \equiv 0 \\ & \forall j \in \{1,\ldots,m\} \quad G_{oj} = B_j \in \mathbb{R}^q \\ & \qquad\qquad\qquad\quad G_{1J} = A_j \in \mathbb{R}^{qxq} \\ & \qquad\qquad\qquad\quad \forall i > 1 \; G_{ij} \equiv 0 \end{aligned}$$

ce qui donne

$$(93) \qquad \begin{cases} \dot{x} = A_o x + \sum_{j=1}^{m} u_j A_j x + Bu \\ \qquad \text{où } u = \left[u_1,\ldots,u_m\right]^T \\ \qquad\quad B = \left[B_1 \vdots \ldots \vdots B_m\right] \\ y = C^T x \end{cases}$$

Ainsi, lorsque le critère à minimiser est (55), les équations générales permettant de déterminer les matrices de gain $P_i$, $i > 2$, sont notablement simplifiées. On obtient dans ce cas :

$$(94) \qquad u^* = -R^{-1}\left[B^T + \begin{bmatrix} x^T A_1^T \\ \vdots \\ x^T A_m^T \end{bmatrix}\right]\left[\sum_{i=1}^{\infty} P_i\, x^{[i]}\right]$$

où . $P_1$ est solution de l'équation de Riccati

$$(95) \qquad C^T Q C + P_1 A_o + A_o^T P_1 - P_1 B R^{-1} B^T P_1 = 0$$

. $P_2$ est solution de l'équation linéaire

$$(96) \qquad P_1 B R^{-1}\begin{bmatrix} \mathrm{vec}^T(P_1 A_1) \\ \vdots \\ \mathrm{vec}^T(P_1 A_m) \end{bmatrix} = \left[A_o^T - P_1 B R^{-1} B^T\right] P_2$$

• pour p>2 (les notations sont précisées dans l'annexe)

$$(97) \qquad T_{p+1}^T \Big\{ \mathrm{vec}(A_o^T P_p) + \mathrm{vec}(P_p^T A_o)$$
$$- \sum_{k=1}^{p} \mathrm{vec}(V_{ko} R^{-1} V_{p-k+1,o}^T)$$
$$- \sum_{k=1}^{p-1} \left[\mathrm{vec}(V_{ko} R^{-1} V_{p-k,1}^T) + \mathrm{vec}(V_{k1} R^{-1} V_{p-k,o}^T)\right]$$
$$- \sum_{k=1}^{p-2} \mathrm{vec}(V_{k1} R^{-1} V_{p-k-1,1}^T)\Big\} = 0$$

or, $\forall k \in \mathbb{N}^*$ $\quad V_{ko} = P_k^T B$

$$V_{k1} = \left[\mathrm{vec}(P_k^T A_1) \vdots \cdots\cdots \vdots \mathrm{vec}(P_k^T A_m)\right]$$

Donc, l'équation (97) est équivalente à :

$$
\begin{aligned}
(98)\quad & \mathcal{C}_p (I\alpha_p \otimes (A_o^T - P_1 B R^{-1} B^T))\ \mathrm{vec}\tilde{P}_p \\
& = T_{p+1}^T \left\{ \sum_{k=2}^{p-1} \mathrm{vec}\,(V_{ko} R^{-1} V_{p-k+1,o}^T) \right. \\
& + \sum_{k=1}^{p-1} \left[ \mathrm{vec}(V_{ko} R^{-1} V_{p-k,1}^T) + \mathrm{vec}(V_{k1} R^{-1} V_{p-k,o}^T) \right] \\
& \left. + \sum_{k=1}^{p-2} \mathrm{vec}(V_{k1} R^{-1} V_{p-k-1,1}^T) \right\} \\
& \text{où } P_p = \tilde{P}_p\, T_p^+ \ , \ \mathcal{C}_p = 2\, T_{p+1}^T\, (T_p^{+T} \otimes I_q)
\end{aligned}
$$

## III.2 SYSTEMES BILINEAIRES SINGULIEREMENT PERTURBES

Dans le cas de systèmes singulièrement perturbés, la construction de la commande sous-optimale peut être envisagée en utilisant les deux méthodes rappelées dans le cas des systèmes linéaires : décomposition de l'équation de Riccati et des équations linéaires donnant les $P_i$ ; décomposition temporelle en sous-systemes lent-rapide et construction d'une commande composite. Le système singulièrement perturbé se présente sous la forme (93) avec :

$$
\begin{aligned}
& x^T = \left[ x_1^T,\ x_2^T \right] \quad , \ x_1 \in \mathbb{R}^{q_1},\ x_2 \in \mathbb{R}^{q_2} \\
(99)\quad & \forall i \in \{0,\ldots,m\} \quad A_i = \begin{bmatrix} A_{11}^i & \vdots & A_{12}^i \\ \cdots & \cdots & \cdots \\ A_{21}^i/\varepsilon & \vdots & A_{22}^i/\varepsilon \end{bmatrix} \begin{matrix} \updownarrow q_1 \\ \updownarrow q_2 \end{matrix} \\
& B = \begin{bmatrix} B_1 \\ \cdots \\ B_2/\varepsilon \end{bmatrix} \begin{matrix} \updownarrow q_1 \\ \updownarrow q_2 \end{matrix} \qquad C = \begin{bmatrix} C_1 & \vdots & C_2 \end{bmatrix}
\end{aligned}
$$

### III.2.1 DECOMPOSITION DES EQUATIONS

Suivant le paragraphe II.2.2.1., $P_1$ peut être écrite sous la forme :

$$(100) \qquad P_1(\varepsilon) = \begin{bmatrix} P_{11}(\varepsilon) & \vdots & \varepsilon P_{12}(\varepsilon) \\ \cdots\cdots & \vdots & \cdots\cdots \\ \varepsilon P_{12}^T(\varepsilon) & \vdots & \varepsilon P_{22}(\varepsilon) \end{bmatrix}$$

$$\text{avec } \forall (i,j) \in \{1,2\}^2 \quad P_{ij}(\varepsilon) = \sum_{k=0}^{\infty} P_{ij}^k \varepsilon^k$$

et l'expression des matrices $P_{ij}^k$ est directement calculée par identification des puissances de $\varepsilon$ en remplaçant (100) dans (95). Généralement, on se contentera, comme dans le cas linéaire, de la connaissance à l'ordre 0 des matrices $P_{ij}(\varepsilon)$ déterminées aux étapes précédentes. La méthode de décomposition de l'équation de Riccati est donc applicable directement au cas bilinéaire et même au cas non linéaire singulièrement perturbé.

### III.2.2 DECOMPOSITION TEMPORELLE

Cette méthode, comme nous le verrons dans son application, est beaucoup moins bien adaptée au cas des systèmes bilinéaires (et a fortiori, non linéaires) ; ceci, pour deux raisons principales :

* les modèles lents et rapides obtenus par la méthode des perturbations singulières ne sont pas bilinéaires (ils ne sont même pas à commande affine)

* l'intérêt de la commande composite en linéaire vient du théorème de superposition, qui, ici, n'est plus valable.

Cependant, lorsque cette méthode de commande sera envisagée, afin de réduire la taille des équations à traiter, il conviendra d'élaborer les modèles réduits bilinéaires pour des variables d'entrée proche de 0. Ce choix est guidé par le fait que la commande optimale étudiée ici agira toujours, à la fin du mouvement, autour de l'origine. Malgré les nombreuses hypothèses simplificatrices et bien que cette méthode paraisse peu adaptée aux systèmes non linéaires, l'étude de l'exemple suivant nous montrera que la commande par décomposition temporelle reste intéressante.

## IV. EXEMPLE D'APPLICATION

Considérons le système bilinéaire défini par l'équation d'état :

$$(101)\qquad \begin{bmatrix} \dot{x} \\ \varepsilon \dot{z} \end{bmatrix} = \left\{ \begin{bmatrix} -2 & 1 \\ -1 & -2 \end{bmatrix} + u \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} \right\} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

où $u \in \mathbb{R}$, $\varepsilon = .1$, $x(o) = 1$, $z(o) = 1$

pour lequel on cherche à minimiser le critère :

$$(102)\qquad J = \int_o^\infty (x^2 + z^2 + u^2)dt$$

Cet exemple nous permettra d'utiliser chacune des méthodes de commande évoquées au cours de ce chapitre ; nous commencerons par la commande analytique directe qui nous fournira l'allure des trajectoires optimales, puis nous comparerons les résultats donnés par les méthodes simplificatrices.

### IV.1. COMMANDE ANALYTIQUE GLOBALE

Nous nous sommes donc placés dans le cas étudié dans la partie III.1, avec, de plus :

$$(103)\qquad C^T Q C = I_2 \text{ et } R = 1$$

L'expression analytique de la commande optimale est donc :

$$(104)\qquad u^* = - \left[1-x+z \ , \ 10(1+x-2z)\right] \left[ \sum_{i=1}^{\infty} P_i \begin{bmatrix} x \\ z \end{bmatrix}^{[i]} \right]$$

les $P_i$, $i \in \mathbb{N}^*$, étant déterminés les uns après les autres par les équations (95-96-98).

De façon pratique, nous nous limiterons à $i \in \{1,2,3\}$, la commande alors obtenue sera notée $u_i^a$.

a) calcul de $P_1$

$P_1$ est solution de l'équation de Riccati :

$$(105)\quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + P_1 \begin{bmatrix} -2 & -1 \\ -10 & -20 \end{bmatrix} + \begin{bmatrix} -2 & -10 \\ -1 & -20 \end{bmatrix} P_1 - P_1 \begin{bmatrix} 1 & 10 \\ 10 & 100 \end{bmatrix} P_1 = 0$$

soit

$$(106)\quad P_1 = \begin{bmatrix} .4012 & -.0307 \\ -.0307 & .0253 \end{bmatrix}$$

Cette matrice gain permet d'envisager la commande quasi-optimale de $u_1^a$, ainsi qu'une commande $u_1$ construite à partir du système (101) linéarisé :

$$(107)\quad u_1 = - \begin{bmatrix} 1, 10 \end{bmatrix} P_1 \begin{bmatrix} x \\ z \end{bmatrix}$$

$$(108)\quad u_1^a = - \begin{bmatrix} 1-x+z, & 10(1+x-2z) \end{bmatrix} P_1 \begin{bmatrix} x \\ z \end{bmatrix}$$

b) calcul de $P_2$

Afin de déterminer $P_k$, $k > 2$, il est nécessaire de connaître la matrice $\mathcal{F} = F_1^T - P_1 G_0 G_0^T$, soit ici :

$$(109)\quad \mathcal{F} = \begin{bmatrix} -2.0943 & -10.9432 \\ -1.2223 & -22.223 \end{bmatrix}$$

D'après l'étude présentée dans la partie précédente

$$(110)\quad P_2 = \mathcal{F}^{-1} P_1 G_0 \, \mathrm{vec}^T (P_1 G_1)$$

$$= \begin{bmatrix} -.0072 & .0029 & .0103 & -.0054 \\ .0075 & -.003 & -.0107 & .0057 \end{bmatrix}$$

Donc :

$$(111) \qquad \tilde{P}_2 = \begin{bmatrix} -.0072 & .0132 & -.0054 \\ .0075 & -.0137 & .0057 \end{bmatrix}$$

ce qui fournit la commande :

$$(112) \qquad u_2^a = - \left[1-x+z,\ 10(1+x-2z)\right] \left\{ P_1 \begin{bmatrix} x \\ z \end{bmatrix} + \tilde{P}_2 \begin{bmatrix} x^2 \\ xz \\ z^2 \end{bmatrix} \right\}$$

c) Calcul de $P_3$ (utilisation des notations de l'annexe)

Tous calculs faits, on obtient :

$$(113) \qquad \mathcal{H}_3 = \left[ -.4256,\ 1.5042,\ -1.9194,\ 1.0044,\ -.1829 \right]^T$$

et

$$(114) \qquad \mathcal{C}_3^+ \mathcal{H}_3 = \left[ h_1^T,\ h_2^T,\ h_3^T,\ h_4^T \right]^T$$

avec $h_1^T = \left[ -.2128,\ .37605 \right]$

$h_2^T = \left[ .37605,\ -.47985 \right]$

$h_3^T = \left[ -.47985,\ .25235 \right]$

$h_4^T = \left[ .25235,\ -.09145 \right]$

Ainsi, comme

$$(115) \qquad \tilde{P}_3 = \left[ C_1,\ C_2,\ C_3,\ C_4 \right]$$

avec $i \in \{1 \ldots 4\}$ $C_i = -\mathcal{F}^{-1} h_i$

il vient

$$(116) \qquad \tilde{P}_3 = \begin{bmatrix} -.2667 & .4103 & -.4048 & .1993 \\ .0316 & -.0042 & .0336 & -.0151 \end{bmatrix}$$

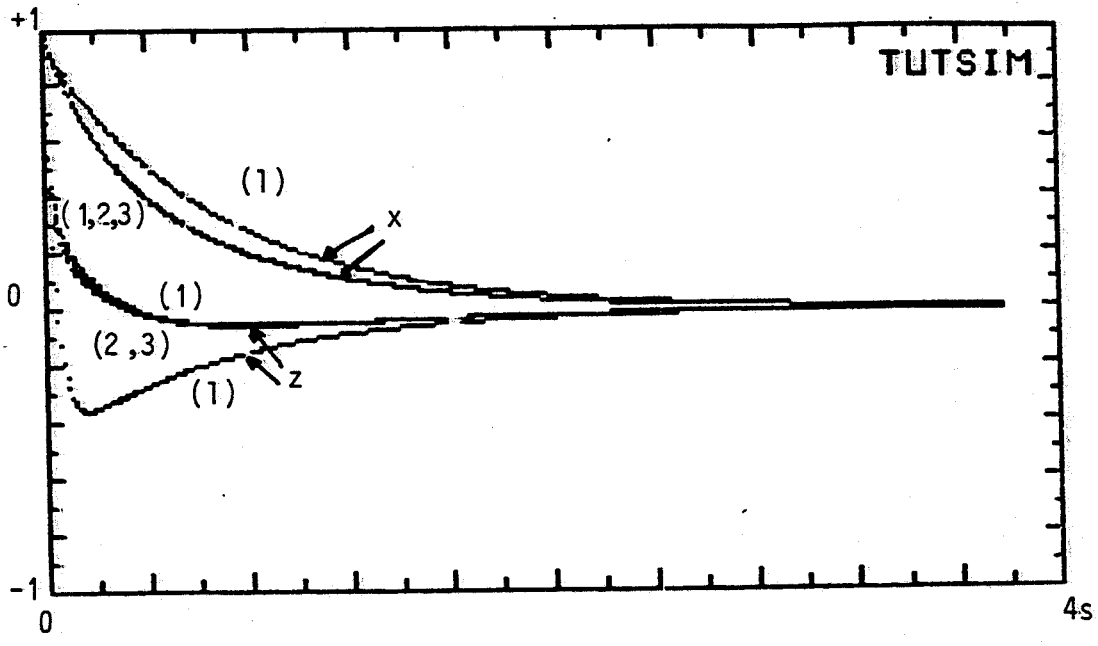
et la commande

$$(117) \qquad u_3^a = - \left[1-x+z,\ 10(1+x-2z)\right] \left\{ P_1 \begin{bmatrix} x \\ z \end{bmatrix} + \tilde{P}_2 \begin{bmatrix} x^2 \\ xz \\ z^2 \end{bmatrix} + \tilde{P}_3 \begin{bmatrix} x^3 \\ x^2 z \\ xz^2 \\ z^3 \end{bmatrix} \right\}$$

d) résultats de simulation

Les réponses du système (101) aux différentes commandes analytiques sont représentées figure 1. La comparaison est faite également avec la commande $u_1$ obtenue à partir du système linéarisé.

FIGURE 1




Réponse du système (101) aux retours d'état analytiques : ((1) : $u_1$, (1) : $u_1^a$, (2) : $u_2^a$, (3) : $u_3^a$)

Les diverses valeurs du critère (102) obtenues au cours de ces simulations sont indiquées ci-dessous :

| COMMANDE | VALEUR DE J |
|---|---|
| $u_1$ | $J_1$ = .37309 |
| $u_1^a$ | $J_1^a$ = .30894 |
| $u_2^a$ | $J_2^a$ = .30266 |
| $u_3^a$ | $J_3^a$ = .29791 |

TABLEAU 1

Ces résultats mettent en évidence les points suivants :

* la prise en compte du caractère non linéaire du vecteur de commande améliore notablement la valeur du critère obtenu.

* l'augmentation de l'ordre de la commande analytique fait converger J vers la valeur optimale.

## IV.2 SIMPLIFICATION DE LA COMMANDE

Le système (101) se présente sous la forme d'un système singulièrement perturbé, on peut donc envisager l'application d'une méthode de simplification. Comme nous l'avons rappelé, on utilisera la décomposition de l'équation de Riccati ou la décomposition temporelle.

### a) décomposition de l'équation de Riccati

La simplification apportée par cette méthode intervient dans le traitement de l'équation de Riccati (105) par la méthode rappelée en I.2.2.1. L'utilisation des formules de ce paragraphe nous conduit à substituer à $P_1$ solution de (105) une matrice $\overline{P_1}$, construite en résolvant des équations de Riccati d'ordre 1,

$$(118) \qquad \overline{P_1} = \begin{bmatrix} P_{11} & \varepsilon P_{12} \\ \varepsilon P_{12} & \varepsilon P_{22} \end{bmatrix} \quad , \varepsilon = .1$$

où $P_{22}$ est la solution positive de :

$$(119) \qquad -4P_{22} - P_{22}^2 + 1 = 0, \text{ soit } P_{22} = .236$$

Pour déterminer les autres coefficients de $\overline{P_1}$, il est nécessaire de calculer les quantités (notations du paragraphe I.2.2.1) :

$$(120) \begin{cases} a) & E_1 = -.5528 \\ & E_2 = .1055 \\ & \\ & \hat{A} = -1.4 \\ b) & \hat{B} = .4472 \\ & \hat{Q} = .7779 \end{cases}$$

$P_{11}$ est alors la solution positive de :

(121) $\quad -2.8\, P_{11} - .2\quad P_{11}^2 + .7779 = 0$, soit $P_{11} = .2725$

et il vient immédiatement

(122) $\quad P_{12} = -.256$

On construit alors une commande analytique $u_1^d$ au premier ordre définie par :

$$(123) \qquad u_1^d = -\left[1-x+z,\ 10(1+x-2z)\right] \bar{P}_1 \begin{bmatrix} x \\ z \end{bmatrix}$$

qui donne une valeur de critère de $J_1^d = .29742$.

Cette méthode n'est à utiliser que quand on envisagera une commande analytique d'ordre 1. En effet, les calculs ne sont simplifiés que dans le calcul de $P_1$, alors que le volume de calculs nécessaires à la détermination de $P_2$, ..., $P_k$, ..., n'est pas diminué.

b) décomposition temporelle

Nous nous contenterons, dans cette partie, d'étudier une commande réduite $u^r$ (i.e. un retour d'état lent) et non une commande composite. Si l'on fait, $\varepsilon = 0$, dans (101), on obtient le modèle réduit lent :

$$(124) \qquad \begin{cases} a)\ \dot{\bar{x}} = -2\bar{x} - \bar{z} + u\left[1-\bar{x}+\bar{z}\right] \\ b)\ \bar{z} = \dfrac{(u-1)}{2(1+u)}\,\bar{x} + \dfrac{u}{2(1+u)} \end{cases}$$

De façon à obtenir un modèle réduit lent bilinéaire et conformément à la présentation générale de cette méthode, on substitue à (124-b) un développement autour de $u = 0$ :

(125) $\quad \bar{z} = (-.5 + u)\,\bar{x} + .5u$

ce qui conduit au système lent bilinéaire :

$$(126) \qquad \dot{\bar{x}} = [-1.5 \ -2.5u\ ]\bar{x} + .5u$$

Le critère à minimiser (102) est remplacé par le critère lent :

$$(127) \qquad \bar{J} = \int_0^\infty (\bar{x}^2 + \bar{z}^2 + u^2)dt$$

ce qui, compte tenu de (125), conduit au problème de commande optimale :

minimisation du critère :

$$(128) \quad \begin{cases} \text{a) } \bar{J} = \int_0^\infty (1.2\ \bar{x}^2 + 1.25v^2)dt \\ \text{pour le système non-linéaire du premier ordre :} \\ \text{b) } \dot{\bar{x}} = -1.4\bar{x} - .5\bar{x}^2 + (.5 - 2.5\bar{x})v \\ \text{où v est relié à u par :} \\ \text{c) } u = v + .2\bar{x} \end{cases}$$

On peut alors construire une solution analytique réduite à l'ordre n sous la forme :

$$(129) \qquad u_n^r = .2x - [.2-x]\left[\sum_{i=1}^{n} p_i x^i\right]$$

où les $p_i$, $i > 1$, sont obtenus par la résolution analytique de

$$(130) \qquad [.25 - 2.5x + 6.25x^2]\ p^2 + [7x + 2.5x^2]\ p - 6x^2 = 0$$

où $p = \sum_{i=1}^{\infty} p_i x^i$

On obtient ainsi pour les premiers termes :

$$(131) \quad \begin{aligned} P_1 &= .8324 \\ P_2 &= -.0488 \\ P_3 &= -.56043 \\ P_4 &= -.06037 \end{aligned}$$

Les simulations pour différentes commandes analytiques

réduites conduisent aux valeurs des critères suivants :

| COMMANDE | VALEUR DE J |
|---|---|
| $u_1^r$ | $J_1^r = .33321$ |
| $u_2^r$ | $J_2^r = .32841$ |
| $u_3^r$ | $J_3^r = .29286$ |
| $u_4^r$ | $J_4^r = .29075$ |

TABLEAU 2

La comparaison entre ce tableau et le tableau 1 montre que la commande réduite permet, pour un ordre donné, d'avoir une bonne approximation par rapport à la commande analytique globale ; d'autre part, l'avantage de l'utilisation de la décomposition temporelle est de diminuer la taille des calculs pour la détermination d'une commande analytique, donc de pouvoir aller plus loin dans l'ordre de la commande envisagé, comme le montre le résultat obtenu par la commande $u_4^r$.

Nous n'avons utilisé qu'une commande réduite, il va de soi que l'utilisation d'une commande composite, tenant compte des parties rapides des variables d'état du système initial permettrait d'affiner les résultats du tableau 2 par rapport aux valeurs optimales.

ANNEXE

## DETERMINATION DE LA MATRICE DE GAIN $P_p$ POUR UN ORDRE QUELCONQUE (LES NOTATIONS EMPLOYEES SONT CELLES DE V : II)

a) Suivant la méthode utilisée pour les deux premiers ordres, l'annulation du coefficient de $x^{[p+1]}$ dans (II : 65) donnera $P_p$. On obtient donc :

$$(A1)\qquad \sum_{i=1}^{p} \mathrm{vec}\,(H_i^T Q H_{p-i+1} + P_i^T F_{p-i+1} + F_i^T P_{p-i+1}) - \underbrace{\sum_{k=1}^{p}\sum_{l=o}^{p-1}\sum_{d=1}^{}\sum_{c=o}^{}}_{k+l+c+d\,=\,p+1} \mathrm{vec}\,(V_{kl} R^{-1} V_{dc}^T) = o$$

En supposant connues les matrices $P_1, \ldots, P_{p-1}$ ; les termes inconnus sont regroupés dans :

$$(A2)\qquad \begin{cases} . \ \mathrm{vec}\,(P_p^T F_1),\ \mathrm{vec}\,(F_1^T P_p) \\ . \ V_{po} R^{-1} V_{10}^T \ (= P_p^T G_o R^{-1} G_o^T P_1),\ V_{10} R^{-1} V_{po}^T \end{cases}$$

Or, comme dans le cas p=2, nous avons

$$(A3)\qquad \forall\, p \geqslant 2 \quad \forall\, A \in \{V_{po} R^{-1} V_{10}^T,\ P_p^T F_1 + H_p^T Q H_1 + F_p^T P_1\}.$$
$$\mathrm{vec}\,(A) = E_{qxqp}\,\mathrm{vec}(A^T)$$

il est à noter ici que, suivant la définition de $E_{qxp}$ (V : 72), $E_{qxqp}$ se construit simplement sous la forme :

$$(A4)\qquad \forall p \geqslant 2 \quad E_{qxqp} = \begin{bmatrix} I_{qp} \otimes e_1^T(q) \\ \vdots \\ I_{qp} \otimes e_k^T(q) \\ \vdots \\ I_{qp} \otimes e_q^T(q) \end{bmatrix}$$

(A1) est donc équivalente à

$$(A5)\qquad \sum_{i=2}^{p-1} \mathrm{vec}(H_i^T Q\, H_{p-i+1} + P_i^T F_{p-i+1} + F_i^T P_{p-i+1})$$

$$- \sum_{\substack{k=1,d=1 \\ k+l+d+c = p+1}}^{p-1} \sum_{c=o,l=o}^{p-1} \mathrm{vec}(V_{kl} R^{-1} V_{dc}^T)$$

$$+ U_{qxq\,p}\left[\mathrm{vec}\,(H_1^T Q H_p + F_1^T P_p + P_1 F_p) - \mathrm{vec}\,(P_1 G_o R^{-1} G_o^T P_p)\right] = o$$

où $U_{qxq\,p} = I_{qxq\,p} + E_{qxq\,p}$

Malheureusement, la matrice $U_{qxqp}$ n'est pas régulière pour tout p (elle l'est seulement pour p pair). De façon à contourner cette difficulté, nous utiliserons la forme non-redondante de la puissance d'un vecteur, notée $\tilde{x}^{[j]}$, et définie par :

$$(A6)\qquad \tilde{x}^{[1]} = x^{[1]} = x = [x_1, \ldots, x_q]^T$$

$\forall\, j \geqslant 2$

$$\tilde{x}^{[j]} = [x_1^j,\ x_1^{j-1}x_2, \ldots, x_1^{j-1}x_q,\ x_1^{j-2}x_2^2,\ x_1^{j-2}x_2x_3, \ldots, x_1^{j-2}x_2x_q, \ldots, x_1^{j-2}x_q^2,\ x_1^{j-3}x_2^3, \ldots, x_q^j]^T$$

La relation entre la forme redondante et la forme non-redondante de la puissance d'un vecteur est donnée par :

$$(A7)\qquad \forall j \in \mathbb{N}\ \ \exists!\ T_j \in \mathbb{R}^{q^j x \alpha_j},\ \text{où}\ \alpha_j = \binom{q+j-1}{j}$$

$$x^{[j]} = T_j \tilde{x}^{[j]}$$

$$\text{donc}\quad \tilde{x}^{[j]} = T_j^+ x^{[j]}\ \text{où}\ T_j^+ = (T_j^T T_j)^{-1} T_j^T$$

BU LILLE

et une fonction analytique de x, écrite en $x^{[j]}$ peut être écrite en $x^{[j]}$ par le changement de coefficients

$$(A8) \qquad A(x) = \sum_{j=o}^{\infty} A_j x^{[j]} = \sum_{j=o}^{\infty} \tilde{A}_j \tilde{x}^{[j]}$$

où $\forall\, j \in \mathbb{N} \quad \tilde{A}_j = A_j T_j$ et $A_j = \tilde{A}_j T_j^{+}$

Ainsi, l'équation (A1) doit être remplacée par, $\forall\, p \geqslant 3$ :

$$(A9) \qquad T_{p+1}^{T} \{K_p\} + T_{p+1}^{T} U_{qxq\,p} \{S_p\} = 0$$

$$\text{où } K_p = \sum_{i=2}^{p-1} \mathrm{vec}(H_i^T Q H_{p-i+1} + P_i^T F_{p-i+1} + F_i^T P_{p-i+1}) - \sum_{\substack{k=1 \\ k+l+d+c=p+1}}^{p-1} \sum_{d=1 \; l=o,c=o}^{p-1} \mathrm{vec}(V_{kl} R^{-1} V_{dc}^T)$$

$$S_p = \mathrm{vec}(H_1^T Q H_p + P_1 F_p) + \mathrm{vec}((F_1^T - P_1 G_o R^{-1} G_o^T) P_p)$$

Or, d'après |BRE,78|, on a également la relation :

$$(A10) \qquad \mathrm{vec}((F_1^T - P_1 G_o R^{-1} G_o^T) P_p)$$
$$= \mathrm{vec}((F_1^T - P_1 G_o R^{-1} G_o^T) \tilde{P}_p T_p^{+})$$
$$= (T_p^{+T} \otimes I_q)(I_{\alpha_p} \otimes (F_1^T - P_1 G_o R^{-1} G_o^T))\, \mathrm{vec}(\tilde{P}_p)$$

ce qui conduit à définir $\mathrm{vec}(\tilde{P}_p)$ comme solution de l'équation linéaire :

$$(A11) \qquad T_{p+1}^{T} \{K_p + U_{qxq\,p} \mathrm{vec}(H_1^T Q H_p + P_1 F_p)\} = \mathcal{H}_p =$$
$$-\mathcal{C}_p (I_{\alpha_p} \otimes (F_1^T - P_1 G_o R^{-1} G_o^T))\, \mathrm{vec}(\tilde{P}_p)$$
$$\text{où } \mathcal{C}_p = T_{p+1}^{T} U_{qxq\,p} (T_p^{+T} \otimes I_q)$$

La matrice $\mathcal{C}_p$ est, pour tout p, une matrice rectangulaire de rang plein de dimension $\alpha_{p+1}$ X $q\alpha_p$. Elle admet donc une pseudo-inverse au sens de Moore-Penrose

$$\text{(A12)} \qquad \mathcal{C}_p^{+} = \mathcal{C}_p^{T}(\mathcal{C}_p \mathcal{C}_p^{T})^{-1}$$

et parmi l'infinité du nombre de solutions de l'équation (A11), celle qui a une norme euclidienne minimale est :

$$\text{(A13)} \qquad \mathrm{vec}(\tilde{P}_p) = -(I_{\alpha_p} \otimes (F_1^{T} - P_1 G_o R^{-1} G_o^{T}))^{-1} \mathcal{C}_p^{+} \mathcal{H}_p$$

mais

$$\text{(A14)} \qquad I_{\alpha_p} \otimes (F_1^{T} - P_1 G_o R^{-1} G_o^{T}) = \begin{bmatrix} F_1^{T}-P_1G_oR^{-1}G_o^{T} & & o \\ & \ddots & \\ o & & F_1^{T}-P_1G_oR^{-1}G_o^{T} \end{bmatrix} \Big\} \alpha_p \text{blocs}$$

et si on note : $\alpha_p$blocs

$$\text{(A15)} \qquad \tilde{P}_p = \begin{bmatrix} \tilde{C}_1 & \vdots & \cdots & \vdots & \tilde{C}_{\alpha_p} \end{bmatrix} \quad , \text{ où } \forall i \in \{1,\ldots,\alpha_p\}\ \tilde{C}_i \in \mathbb{R}^q$$

$$\text{(A16)} \qquad \mathcal{C}_p^{+} \mathcal{H}_p = \begin{bmatrix} h_1 \\ \cdots \\ h_2 \\ \cdots \\ \vdots \\ \cdots \\ h_{\alpha_p} \end{bmatrix} \quad \text{où } \forall i \in \{1,\ldots,\alpha_p\}\ h_i \in \mathbb{R}^q$$

alors la solution (A13) est, en fait, obtenue par l'inversion unique de la matrice $F_1^{T} - P_1 G_o R^{-1} G_o^{T}$ et l'on obtient finalement

$$\text{(A17)} \qquad \forall\, i \in \{1,\ldots,\alpha_p\}\ \tilde{C}_i = -(F_1^{T} - P_1 G_o R^{-1} G_o^{T})^{-1} h_i$$

et

$$\text{(A18)} \qquad P_p = \tilde{P}_p T_p^{+}$$

b) Remarques sur la construction pratique de quelques matrices

b1) Construction itérative de $T_p$

$T_p$ est définie par :

$$(A19) \qquad x^{[p]} = T_p \tilde{x}^{[p]} \quad , \; x \in \mathbb{R}^q$$

on a donc $T_1 = I_q$ et $T_{p+1}$ est construite à partir de $T_p$ par la règle suivante :

$$(A20) \qquad T_{p+1} = \begin{bmatrix} \text{bloc } 1 \\ \vdots \\ \text{bloc } i \\ \vdots \\ \text{bloc } q \end{bmatrix}$$

les blocs 1 à q ayant leurs colonnes indicées par les composantes de $\tilde{x}^{|p+1|}$, si l'on enlève du bloc i(i = 1,...q) les colonnes dont l'indice ne comporte pas de $x_i$, le bloc résiduel est égal à $T_p$. Les autres colonnes du bloc i (celles que l'on a enlevées) sont complètées par des 0.

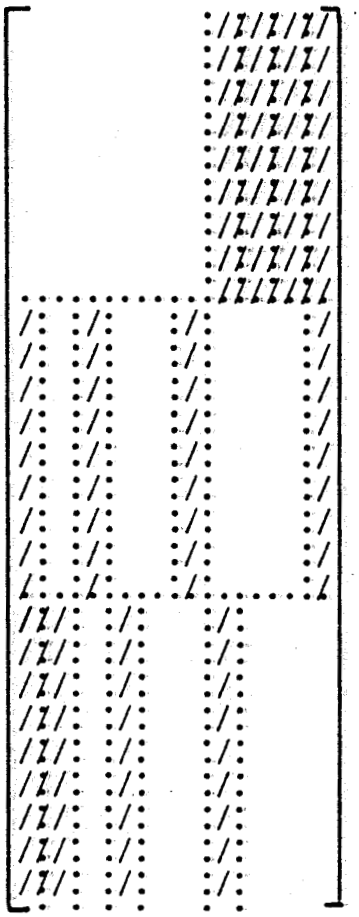
Par exemple :(//// colonnes de 0)

$$* \; q=2 \quad T_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; T_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$T_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ldots\ldots, \quad T_{p+1} = \begin{bmatrix} T_p & 0 \\ 0 & T_p \end{bmatrix}$$

$$* \; q=3 \quad T_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$T_2 = \left[\begin{array}{ccc:cc} 1 & 0 & 0 & & \\ 0 & 1 & 0 & & \\ 0 & 0 & 1 & & \\ \hdashline & 1 & & 0 & 0 \\ & 0 & & 1 & 0 \\ & 0 & & 0 & 1 \\ \hdashline & & 1 & 0 & 0 \\ & & 0 & 1 & 0 \\ & & 0 & 0 & 1 \end{array}\right]$$

$T_3 =$  , etc...

b2) Calcul de $T_p^+$

Par définition : $T_p^+ = (T_p^T T_p)^{-1} T_p^T$, ce qui en posant $T_p^T T_p = D_p$ demande l'inversion de la matrice $D_p$. Or, d'après (A20) :

$$(A21) \qquad D_{p+1} = T_{p+1}^T T_{p+1} = \sum_{i=1}^{q} \left[\text{bloc i}\right]^T \left[\text{bloc i}\right]$$

et d'après la construction du (bloc i), la matrice $|\text{bloci}|^T|\text{bloc i}|$ est une matrice de taille $\alpha_{p+1} x \alpha_{p+1}$, diagonale, et dont la construction est définie par la règle : si l'on enlève de cette matrice les lignes et les colonnes dont l'indice ne présente pas de $x_i$, on obtient $D_p$ ; la matrice entière étant obtenue en remplissant les lignes et les

colonnes enlevées par des 0. Ainsi, comme $T_1 = I_q$, on a montré par récurrence que dans tous les cas, $D_p$ était une matrice régulière diagonale, donc élémentairement inversible. Donnons quelques exemples pour $D_p$ :

* q=2

$$D_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \ \ldots$$

on peut montrer que de façon générale :

(A22) $$\forall\, p \geqslant 1 \quad D_p = \operatorname*{diag}_{i=0}^{p} \{C_p^i\}$$

où les $C_p^i$ sont les coefficients du binôme

* q=3

$$D_2 = \text{diag } \{1,2,2,1,2,1\}$$

$$D_3 = \text{diag } \{1,3,3,3,6,3,1,3,3,1\}$$

b3) Exemples de matrices $\mathcal{C}_p$

Compte tenu des remarques précédentes, le calcul de $\mathcal{C}_p$, dans le cas général, ne pose pas de problèmes particuliers ; nous donnons ici quelques exemples de la forme de $\mathcal{C}_p$ ($= T_{p+1}^T\, U_{qxqp}\, (T_p^{+T} \otimes I_q)$): de façon générale, on peut montrer la propriété suivante :

(A23) $$\forall\, q \in \mathbb{N}^*, \forall p \in \mathbb{N}^* \quad T_{p+1}^T\, U_{qxqp} = 2\, T_{p+1}^T$$

La démonstration vient du fait que $U_{qxqp} = I_{qxqp} + E_{qxqp}$ et que :

(A24) $$x^{[p+1]} = \text{vec}(x.x^{[p]T}) = E_{qp_{xq}}\, \text{vec}(x^{[p]}.x^T) = \text{vec}(x^{[p]}.x^T)$$

et suivant le Théorème 1.5 de |BRE,78|

(A25) $$E_{qp_{xq}}^T = E_{qxqp}$$

ce qui conduit à la relation :

$$x^{[p+1]} = E_{qxqp}^{T}\ T_{p+1}\tilde{x}^{[p+1]} = T_{p+1}\ \tilde{x}^{[p+1]}$$

Ainsi, de façon simplifiée :

$$\mathcal{C}_p = 2T_{p+1}^{T}(T_p + T \otimes I_q), \mathcal{C}_p \in \mathbb{R}^{p+1 x (q.\ p)}$$

ce qui donne pour quelques cas particuliers :

* q=2

$$\mathcal{C}_2 = \left|\begin{array}{cccccc} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right|$$

$$\mathcal{C}_3 = \left|\begin{array}{cccccccc} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right|$$

* q=3

$$\mathcal{C}_2 = \left|\begin{array}{cccccccccccccccccc} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{array}\right|$$

BU LILLE

# CONCLUSION

Dans cette cinquième et dernière partie, nous avons développé deux points de vue différents concernant l'obtention d'une commande sous-optimale pour un système donné. Dans le cas d'un système linéaire singulièrement perturbé, l'utilisation de la méthode de découplage permet d'affiner les résultats pour l'application de la commande composite. Dans le cas des systèmes non-linéaires, l'utilisation du principe de la commande analytique permet de construire l'expression d'un retour d'état polynomial qui approche le retour d'état non-linéaire optimal.

Les systèmes bilinéaires possédant une structure intermédiaire entre les systèmes linéaires et non-linéaires, nous avons utilisé comparativement les deux méthodes précédentes sur un système bilinéaire singulièrement perturbé. Les résultats obtenus montrent que, malgré les hypothèses simplificatrices nécessaires à la construction des modèles réduits lents et rapides, la méthode de décomposition temporelle conduisant à une commande composite, reste très performante.

# CONCLUSION GENERALE

# CONCLUSION GENERALE

Dans notre étude des systèmes bilinéaires singulièrement perturbés, du fait de leur structure intermédiaire entre les modèles linéaires et les modèles non linéaires, nous avons constamment essayé d'induire des méthodes propres aux systèmes linéaires ou d'appliquer des techniques utilisées sur les systèmes non linéaires. Le principe de la commande analytique n'avait été que suggéré dans le passé ; l'utilisation de l'algèbre tensorielle nous a permis d'en proposer une nouvelle formulation et d'en déduire une construction systématique.

Dans un but de clarté de l'exposé, nous avons choisi de ne pas aborder les problèmes théoriques de convergence des algorithmes que nous proposons, en ayant cependant testé leur efficacité sur des exemples. Certains résultats, comme l'utilisation du réciproque pour les systèmes bilinéaires demandent à être affinés pour pouvoir justifier de leur intérêt.

Notre intérêt futur aura donc plusieurs directions : approfondir les parties esquissées dans ce mémoire ; étendre à des systèmes plus généraux (à commandes affines, à états affines) les méthodes d'approximation présentées ; déterminer l'ordre minimal d'une commande analytique pour une erreur fixée a priori. L'application de ces résultats aux systèmes complexes que sont les robots manipulateurs est aussi envisagée.

# REFERENCES BIBLIOGRAPHIQUES

# BIBLIOGRAPHIE


|AIG,49| P.R. AIGRAIN, E.M. WILLIAMS "Synthesys of n-reactance networks for desired transient response". Jour. of App. Phys., vol 20, p. 597, 1949

|AND,78| L. ANDERSON : "Decomposition of two time scale linear systems", Proc. JACC, p. 153,1978

|AND,81| L.R. ANDERSON, W.L. HALLAUER : "A method for order reduction for structural dynamics based on Riccati iteration", AIAA Journal, 19, n°6, p. 796,1981

|AOK,68| M. AOKI : "Control of large scale dynamic systems by aggregation", IEEE-AC-13, juin, p. 246,1968

|APP,78| R.K. APPIAH : "Linear model reduction using Hurwitz polynomial approximation", I.J.C., vol. 28, n° 3, p. 477, 1978

|ARB,80| M. A. ARBIN, E.G. MANES : "Foundations of system theory : the Hankel matrix", J. Computer Syst. Sci., vol. 20, n° 3, p. 330, 1980

|ARD,82| M.D. ARDEMA : "An introduction to singular perturbations in non linear optimal control" in "Singular perturbations in systems and control", Ed. M.D. ARDEMA, Springer Verlag, p. 1, 1982

|ARM,79| M.A. ARMADA : "Simplificacion de systemas bilineales", Thèse de Doctorat, Université de Valladolid, 1979

|ARM,81| M. ARMADA, J. NO : "Systèmes bilinéaires. Application à la modélisation dynamique des robots industriels", Congrès Automatique AFCET, 1981, NANTES, pp. 269-280

|ARM,82| M.A. ARMADA : "Feedback control of singularly perturbed bilinear systems", IFAC workshop on singular perturbation and robustness of control systems, 1982

|ARM,86| M.A. ARMADA : "Simplified feedback stabilization of singularly perturbed bilinear systems", IMACS Symposium on Modelling and Simulation for Control of Lumped and Distributed Parameter Systems, Villeneuve d'Ascq, 1986

|ASH,82| N. ASHOOR, V. SINGH : "A note on low order modelling", IEEE-AC-27, n° 5, p. 1124, 1982

|AVR,79| B. AVRAMOVIC : "Subspace iteration approach to time scale separation", Proc. of IEEE Conf. on Dec. and Cont., p. 684, 1979

|BAL,76| A.V. BALAKRISHNAN : "Are all non-linear systems bilinear?", Joint Automatic Control Conf., SAN FRANCISCO, 1976

|BAN,86b| S.P. BANKS, M.K. YEW : "On the optimal control of bilinear systems and its relations to Lie Algebras", Int. J. Control, vol. 43, n° 3, p. 891, 1986

|BEL,81| D.J. BELL, Q. YE :" A perturbation method for sub-optimal feedback control of bilinear systems", Int. J. Syst. Sci., vol. 12, n° 9, p. 1157, 1981

|BEN,73| T.E. BEN ISRAEL, T.N.E. GREVILLE : "Generalized inverse : theory and applications", Wileg-Intersciences 1973

|BEN,81| A. BENSOUSSAN : "Singular perturbation results for a class of stochastic control problems", IEEE-AC-26, n° 5, p. 1071, 1981

|BEN,82| A. BENSOUSSAN : "Regular perturbations in optimal control" in "Singular perturbations in systems and control", Ed. M.D. ARDEMA, Springer Verlag, p. 169, 1982

|BEN,86| M. BENREJEB, M. GASMI, M.N. ABDELKRIM : "Nouvelle méthode de modélisation des systèmes linéaires singulièrement perturbés. Méthode du cercle", IMACS Symposium on Modelling and Simulation for Control of Lumped and Parameters Systems, p. 569, Lille, 1986

|BER,76| P. BERTRAND, G. MICHAILESCO, J.M. SIRET : "Sur la synthèse de modèles réduits par agrégation", RAIRO, vol. 10, n° 7, p. 105, juillet 1976

|BER,79| P. BERTRAND, J.M. SIRET, G. MICHAILESCO : "On the use of the aggregation techniques" in "Handbook of Large Scale Engineering applications", M. G. SINGH, A. TITLI, North Holland, 1979

|BER,82| J. BERNUSSOU, A. TITLI : "Interconnected dynamical systems : stability decomposition and decentralisation", North Holland, 1982

|BER,84| J. BERSTEL, C. REUTANAUER : "Les séries rationnelles et leurs langages", Masson, 1984

|BER,85| P. BERTRAND, G. DUC, G. MICHAILESCO : "Développements récents sur la réduction de modèles", APII, n° 19, p. 131, 1985

|BIN,78| Z. BINDER, C. COMMAULT : "Une revue des méthodes de simplification de modèles dynamiques linéaires invariants", RAIRO, vol. 12, n° 3, 1978

|BOR,87| P. BORNE, G. DAUPHIN-TANGUY : "Singular perturbations, boundary layer problem", Systems and Control Encyclopedia, M.G. Singh Eds, Pergamon Press, III I-2, 1987

|BOS,72| M.J. BOSLEY, F.P. LEES : "A survey of simple transfer function derivation from high order state variable models", Automatica, p. 765, novembre 1972

|BOS,80| N.K. BOSE, S. BASU : "Two dimensional matrix Pade Approximants : existence, nonuniqueness and recursive computation", IEEE-AC25, n° 3, june 1980

|BOU,69| R. BOUDAREL, J. DELMAS, P. GUICHET : "Commande optimale des processus "T.3, Dunod, 1969

|BRA,64| A. BRAUER : "On the characteristic roots of non-negative matrices" in "Recent advances in matrix theory", Ed. H. SCHNEIDER, Madison and Milwankel, p. 3, 1964

|BRA,81| W. BRANDENBUSCH, W. STEHLE : "A practical closed form solution for systems with variable eigenvalues via bilinear system theory", Circ. Theor. and Appl., vol. 9, p. 103, 1981

|BRE,78| J.W. BREWER : "Kronecker products and matrix calculus in system theory ", IEEE-CAS-25, n° 9, p. 772, Sept. 1978

|BRO,76a| R.W. BROCKETT : "Volterra series and geometric control theory", Automatica, vol. 12, p. 167, 1976

|BRO,76b| R.W. BROCKETT : "Finite and infinite dimensional bilinear realizations", J. Franklin Inst., vol. 301, n° 6, p. 509, 1976

|BRU,71| C. BRUNI, G. DIPILLO, G. KOCH : "On the mathematical models of bilinear systems", Richerche di Automatica, vol. 2, n° 1, 1971

|BRU,74| C. BRUNI, G. DIPILLO, G. KOCH : "Bilinear systems : an appealing class of "nearly linear" systems in theory and applications", IEEE-AC-19, n° 4, p. 334, 1974

|CAR,32| T. CARLEMAN : "Application de la théorie des équations intégrales singulières aux équations différentielles de la dynamique", T. Ark. Mat. Astrom. Fys., 22B, 1, 1932

|CEB,84a| W.A. CEBUHAR, V. COSTANZA : "Approximation procedures for the optimal control of bilinear and non-linear systems", J. of Opt. Th and Appl., vol. 43, n° 4, August 1984

|CEB,84b| W.A. CEBUHAR, V. COSTANZA : "Non linear control of CSTR'S", Chemical Engineering Science, vol. 39, n° 12, p. 1715, 1984

|CHA,71| M. CHAMBAT : "Localisation des valeurs propres : régions de Gudkov", RIRO - n° R-3, p. 82, 1971

|CHA,72| K. W. CHANG : "Singular perturbations of a general boundary value problem", SIAM J. Math. Appl., vol. 3, p. 520, 1972

|CHA,82| R.Y. CHANG, M.L. WANG : "Parameter identification via shifted Legendre polynomials" Int. J. Syst. Sci, vol. 13, p. 1125, 1982

|CHE,61| C.F. CHEN, L.S. SHIEH : "A novel approach to linear model simplification", I.J.C., p. 561, décembre 1961

|CHE,70| C.F. CHEN, L.S. SHIEH : "An algebric method for control sytem design", I.J.C., vol. 11, n° 5, p. 717, 1970

|CHE,74| C.F. CHEN : "Model reduction of multivariable control systems by means of matrix continued functions", I.J.C., vol. 19, n° 2, p. 225, 1974

|CHE,75a| C.F. CHEN, C.H. HSAIO : "Time domain synthesis via Walsh functions", Proc. IEE, vol. 122, p. 575, 1975

|CHE,75b| C.F. CHEN, C.H. HSAIO : "Design of precewise constant gains for optimal control via Walsh functions", IEEE-AC-20, p. 596, 1975

|CHE,75c| A. CHERUY : "Analyse d'une classe importante de biosystèmes : les réactions enzymatiques", RAIRO-Automatique, Février, p. 37, 1975

|CHE,78| W. CHEN, Y. SHIH : "Parameter estimation of bilinear systems via Walsh functions", J. Franklin Inst., vol. 305, n° 5, p.249, 1978

|CHE,82| B. CHENG, N.S. HSU : "Analysis and parameter estimation of bilinear systems via block-pulse functions", I.J.C., vol. 36, p. 53, 1982

|CHI,67| R.M. CHIDAMBARA : "On a method for simplifying linear dynamic system", IEEE-AC-12, p. 119, feb. 1967

|CHI,73| J.S.R. CHISHOLM : "Rational approximants defined from double power series", Math. Comput., vol. 27, p. 841, octobre 1973

|CHI,74| J.S.R. CHISHOLM, J. Mc EWAN : "Rational approximants defined from power series in n variables", Proc. Rog. Soc., vol. A336, p. 421, 1974

|CHO,75| J.H. CHOW : "Two-stage design of singularly perturbed linear regulators", Proc. 13th Annu. Allerton Conf. Circuit and System Theory Univ. Of Illinois, p. 48, 1975

|CHO,76a| J.H. CHOW, P.V. KOKOTOVIC : "Eigenvalue placement in two-time-scale systems", IFAC Symposium on Large Scale Systems, theory and applications, Udine, p. 321, 1976

|CHO,76b| J.H. CHOW, P.V. KOKOTOVIC : "A decomposition of near-optimum regulators for systems with slow and fast modes" IEEE-AC21, p. 701, 1976

|CHO,86| J.H. CHOW : "Time-scale modeling of dynamics networks with applications to power systems", Lecture Notes in Control and Information Sciences, n° 46, Spinger-Verlag, 1986

|CHU,70| S.C. CHUANG : "Application of continued-fraction methods for modelling transfer fuctions to give a More accurate transient response", Electron. Lett., p. 861, décembre 1970

|CIA,82| P.G. CIARLET : "Introduction à l'analyse numérique matricielle et à l'optimisation", Masson, 1982

|CLA,85|

|CLA,85| D. CLAUDE : "Everything you always wanted to know about linearization but were afraid to ask", Méthodes algébriques et géométriques en automatique non linéaire, CNRS, Paris, juin 1985

|COM,77| C. COMMAULT, J.P. GUERIN : "Propriétés et structure des systèmes agrégés, applications au problème de la projection optimale", CRAS, novembre 77

|COM,78| C. COMMAULT, J.P. GUERIN : "Propriétés et structure des systèmes agrégés, application au problème de la projection optimale", RAIRO, vol. 12, n° 4, p. 377, 1978

|DAD,85| DADUGINETO : "Identification de différents modèles discrets de systèmes non linéaires continus : application à deux procédés industriels pilotes", Méthodes algébriques et géométriques en automatique non linéaire, CNRS, Paris, juin 1985

|DAU,82a| G. DAUPHIN-TANGUY, M. NOEL, P. BORNE : "Présentation d'algorithmes de séparation des échelles de temps d'un processus discret", Congrès MECO'82, Tunis, Tunisie, 1982

|DAU,82b| G. DAUPHIN-TANGUY, A. EL MOUDNI, P. BORNE : "Sur une estimation des régimes transitoires des processus à plusieurs échelles de temps", Congrès MECO'82, Tunis, p. 151, 1982

|DAU,83a| G. DAUPHIN-TANGUY : "Sur la représentation multi-modèle des systèmes singulièrement perturbés, application à l'analyse et à la synthèse", Thèse de Docteur es Sciences Physiques", n° 600, Lille, 1983

|DAU,83b| G. DAUPHIN-TANGUY, M. LEBRUN, P. BORNE : "Interprétation de la notion de système réciproque par les bondgraphes pour les processus multi-échelles de temps", Congrès IASTED AI'83, Lille, p. 137, 1983

|DAU,85| G. DAUPHIN-TANGUY, P. BORNE, A. FOSSARD : "Analyse et synthèse des systèmes à plusieurs échelles de temps", APPII, n° 19, p. 169, 1985

|DAU,87| G. DAUPHIN-TANGUY, P. BORNE : "Multi-time scale systems : dynamic location", Systems and Control Encyclopedia, M.G. Singh Eds, Pergamon Press, III I-2, 1987

|DAV,68| E.J. DAVISON : "A new method for simplifying large linear dynamic systems", IEEE-AC-13, p. 214, avril 1968

|DEC,76| M. DECOSTER, E. NOLDUS, A. VAN CAUWENBERGHE : "Réduction des systèmes linéaires", RAIRO, p. 47, avril 1976

|DEI,82| A.S. DEIF : "Advanced matrix theory for scientist and engineers", Abacus Press, 1982

|DEL,86| J. DELFORGE : "on the relation between the identifiability and the conditionning of the identification problem" IMACS Symposium on Modelling and Simulation for Control of Lumped and Distributed Parameter Systems", p. 323, Villeneuve d'Ascq, 1986

|DER,80a| I. DERESE, E. NOLDUS :"Design of linear feedback laws for bilinear systems", Int. J. Control, 31, p. 219, 1980

|DER,80b| I. DERESE, E. NOLDUS : "Non linear control of bilinear systems", Proc. IEE, vol. 127, n° 4, p. 169, 1980

|DIB,82| G. DI BLASIO : "Synthesis of optimal bilinear controls", Journal of Mathematical analysis and applications, n° 88, p. 143, 1982

|ESP,78| M. ESPANA, I.D. LANDAU : "Reduced order bilinear models for distillation columns", Automatica, vol. 14, p. 345, 1978

|FEI,62| D.G. FEINGOLD, R.S. VARGA : "Block diagonally dominant matrices and generalizations of the Gershgorine circle theorem", Pacific J. Math., 12, p. 1241, 1962

|FIN,49| N.J. FINE : "On the Walsh functions", Trans. Amer. Math. Soc., vol. 65, p. 372, 1949

|FLI,74| M. FLIESS : "Matrices de Hankel", J. Math. pures et Appl. n° 53, p. 197, 1974

|FLI,76| M. FLIESS : "Un outil algébrique : les séries formelles non commutatives" in "Mathematical System Theory", Lect. Notes Econom. Math. Syst., Springer-Verlag, vol. 131, p. 122, 1976

|FLI,79| M. FLIESS : "A remark on the transfer function and the realization of homogeneous continuons-time non linear systems", IEEE-AC-24, n° 3, p. 507, 1979

|FLI,80| M. FLIESS : "Une approche algébrique du développement fonctionnel des solutions d'équations différentielles non linéaires forcées", in "Analyse des systèmes" Société Mathématique de France, p. 95, 1980

|FLI,81| M. FLIESS : "Developpements fonctionnels et calcul symbolique non commutatifs" in "Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal", Tome 1, Ed. CNRS, p. 359, 1981

|FLI,83| M. FLIESS, M. LAMNABHI, F. LAMNABHI-LAGARRIGUE : "An algebric approch to nonlinear functionnal expansions", IEEECAS-30, n° 8, p. 554, 1983

|FOS,70| A.J. FOSSARD : "Sur la simplification de modèles complexes" Automatismes, n° 4, p. 141, 1970

|FOS,82| A.J. FOSSARD, J.F. MAGNI : "Modélisation, commande et applications des systèmes à échelles de temps multiples", RAIRO Aut. , vol. 16, n° 1, p. 5, 1983

|FOS,83| A.J. FOSSARD : "Systèmes linéaires à deux échelles de temps" in "Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal", CNRS, T.3, p. 109, 1983

|FOS,86| A.J. FOSSARD, G. DAUPHIN-TANGUY, P. BORNE : "Modelling complexity and control", IMACS Symposium on Modelling and Simulation for Control of Lumped and Distributed Parameter Systems, Villeneuve d'Ascq, p. 5, 1986

|FRA,75| M. FRAYMAN : "On the relationship between bilinear and quadratic systems", IEEE-AC-20, n° 4, p. 567, 1975

|GAN,60| F.R. GANTMACHER : "Théorie des matrices", Dunod, 1960

|GAU,84| J.P. GAUTHIER : "Structure des systèmes non linéaires", Ed CNRS, Paris,1984

|GEN,72| J.C. GENTINA, P. BORNE, F. LAURENT : "Stabilité des systèmes continus non linéaires de grande dimension", RAIROJ-3-Août, p. 69, 1972

|GEN,76| J.C. GENTINA : "Contribution à l'analyse et à la synghèse des systèmes non linéaires de grandes dimensions", Thèse de Doctorat es Sciences Physiques, Lille, 1976

|GIL,78| J. GILEWICZ : "Les approximants de Padé", Springer-Verlag, Lectures Noters in Mathematics, 1978

|GIL,83| G. GILLES : "Une loi de commande optimale d'une classe de systèmes bilinéaires multivariables" Journées de l'Optimisation, Montréal, mai 1983

|GIL,85| G. GILLES, N. LAGGOUNE : "Commande numérique par retour d'état de processus continus non-linéaires" Le point en Automatique, vol. 2, PP. 33-42, Dunod, 1985

|GIL,86| G. GILLES, N. LAGGOUNE, B. NEYRAN, D. THOMASSET : "Non linear discrete time modelling and control of a pilot neutralization chemical reactor" IFAC Symposium on Dynamics and Control of Chemical Reactors and Distillation Columns, Bornemouth (GB), 1986

|GOL,77| M.J. GOLDMAN, C.T. LEONDES : "Truncation of the mixed Cauer form - the approximation in a definite mathematical sense", I.S.C., p. 979, juin 1977

|GRO,63| F.S. GRODINS : "Control Theory and Biological Systems", Columbia University Press, 1963

|GRU,75| Lj. T. GRUJIC : "Novel development of Lyapounov stability of motion", I.J.C., vol. 22, n° 4, p. 525, 1975

|GRU,76| Lj.T. GRUJIC, J.C.GENTINA, P. BORNE : "General aggregation of large scale systems by vector Lyapounov functions and vectors norms", I.J.C., vol. 24, n° 4, p. 529, 1976

|GRU,79a| Lj.T. GRUJIC : "Singular perturbations and large-scale systems", I.J.C., vol. 29, n° 1, p. 159, 1979

|GRU,79b| Lj.T. GRUJIC : "Singular perturbations, large-scale systems and asymptotic stability of invariant sets", Int. J. Systems Sci., vol. 10, n° 12, P. 1323, 1979

|GRU,79c| Lj.T. GRUJIC, P. BORNE, J.C. GENTINA : "Matrix Approaches the absolute stability of time varying Lurre'Postnikov systems", I.J.C., vol. 30, n° 6, p. 967, 1979

|GRU,81| Lj.T. GRUJIC : "Uniform asymptotic stability of non-linear singularly perturbed general and large-scale systems", I.J.C., vol. 33, p. 481, 1981

|GUI,81| J.M. GUILLEN, M.A. ARMADA : "A singular perturbation method for order reduction of large-scale bilinear dynamical systems", in "Large Scale Systems : Theory and Applications", Eds A. TITLI, M.G. SINGH, Pergamon, p. 229, 1981

|GUT,81| S. GUTMAN, E.I. JURY : "A general theory for matrix root-clustering in subregions of the complex plane", IEEE-AC-26, n° 2, p. 859, 1981

|GUT,82| P.O. GUTMAN, C.F. MANNERFELT, P. MOLANDER : "Contributions to the model reduction problem" IEEE-AC-27, n° 2, p. 454, 1982

|HAR,61| J.D. HARDY : "Physiology of temperature regulation", Physiol. Rev., vol. 41, p. 521, 1961

|HER,77| R.H. HERMANN, A.J. KRENER : "Non linear controllability and observability", IEEE-AC-22, n°5, p. 728, 1977

|HES,85| C. HESPEL, G. JACOB : "Approximation de systèmes différentiels par des systèmes bilinéaires", Méthodes algébriques et géométriques en Automatique non linéaire, juin 1985, CNRS, Paris

|HIC,75| J. HICKIN, N.K. SINHA : "Aggregation matrices for a class of low order models for large scale systems", Electronics letters, vol. 11, n° 9, mai 1975

|HOP,71| F. HOPPENSTEADT : "Properties of solutions of ordinary differential equations with smal parameters" Com. Pure Appl. Math, vol. 24, p. 807, 1971

|HOP,74| F. HOPPENSTEADT : "Asymptotic stability in singular perturbations problems : problems having matched asymptotic expansion solutions", J. of differential Equations, n° 15, p. 510, 1974

|HUN,83| L.R. HUNT, R. SU, G. MEYER : "Global transformations of non-linear systems", IEEE-AC-28, n° 1, p. 24, 1983

|HUT,75a| M.F. HUTTON, R. FRIEDLAND " Routh approximation for reducing order of linear, time invariant systems", IEEE-AC-juin, p. 329, 1975

|HUT,75b| M.F. HUTTON, R. FRIEDLAND : "Simplification of High order mechanical systems using the Routh approximant", Trans on ASME, p. 383, décembre 1975

|HWA,82| C. HWANG, Y.P. SHIH : "Parameter identification via Laguerre polynomials", Int. J. Syst. Sci., vol. 13, p. 1125, 1982

|IBE,60| A.S. IBERALL : "Human body as an inconstant heat source and its relation to clothes insulation", II., Journal of Basic Engineering, p. 103, mars 1960

|IMB,75| N. IMBERT, M. CLIQUE, A.J. FOSSARD : "Un critère de gouvernabilité des systèmes bilinéaires", RAIRO, J-3, p.55, nov. 1975

|ISI,73| A. ISIDORI : "Direct construction of bilinear realizations from non linear input-outputs maps", IEEE-AC-18, n° 6, p. 626, 1973

|JAC,77| D.H. JACOBSON : "Extensions of linear-quadratic control, optimization and Matrix theory" Academic Press, 1977

|JAC,81| G. JACOB : "Réalisation des systèmes réguliers et séries génératrices non commutatives" in "Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal", Tome 1, Ed. CNRS, p. 325, 1981

|KAI, 80| T. KAILATH : "Linear systems", Prentice-Hall, 1980

|KAR,78| V.R. KARANAM, P.A. FRICK, R.R. MOHLER : "Bilinear system identification by Walsh functions", IEEE-AC-23, n° 4, p. 709, 1978

|KAT,76| T. KATO : "Perturbation theory of linear operators", Springer Verlag, 1976

|KHA,79| H.K. KHALIL, P.V. KOKOTOVIC : "D. stability and multi-parameter singular perturbation", SIAM J. Cont. and Opt., vol. 17, n° 1, p. 56, 1979

|KHA,80| K.J. KHATWANI, R.K. TIWARI, J.S. BAJWA : "On Chuang's Continued fraction method of model reduction", IEEE-AC-25, N° 4, p. 822, 1980

|KHR,85| N.M. KHRAISHI, H.A. OTHMAN, M.S. MAHMOUND : "Bounds on suboptimality in discrete aggregared models", Large Scale Systems, vol. 8, n° 1, 1985

|KOK,72a| P.V. KOKOTOVIC : "Feedback design of large linear systems", in "Feedback systems" ed. J.B. CRUZ, Mc Graw-Hill, 1972

|KOK,72b| P.V. KOKOTOVIC, R.A. YACKEL : "Singular perturbations of linear regulators : basic theorems", IEEE-AC-17, p. 29, 1972

|KOK,75| P.V. KOKOTOVIC : "A Riccati equation for block-diagonalization of illconditionned systems", IEEE-AC20, p. 812, 1975

|KOK,76| P.V. KOKOTOVIC : "Singular perturbations and order reduction in control theory-an overview", Automatica, vol. 12, p. 123, 1976

|KOK,80| P.V. KOKOTOVIC, J.J. ALLEMONG, J.R. WINKELMAN, J.H. CHOW : "Singular perturbation and iterative separation of time scales", Automatica, vol. 16, n° 23, 1980

|KOK,83| P.V. KOKOTOVIC : "Perturbations singulières en contrôle optimal", in "Outils et modèles mathématiques pour l'automatique, l'analyse des systèmes et le traitement du signal", CNRS, T.3, p. 145, 1983

|KOV,74| Z.V. KOVARIK, D.D. OLESKY : "Sharpness of generalized Gershgorine disks", Linear Albebra and its Applications, 8, p. 477, 1974

|KRE,73| A.J. KRENER : "On the equivalence of control systems and the linearization of non-linear systems", SIAM J. Control, vol.11, n° 4, p. 670, 1973

|KRE,74| A.J. KRENER : "Linearization and bilinearization of control systems", Proc 1974 Allerton Conference on Circuit and System Theory, Urbanca, 1974

|KRE,75| A.J. KRENER : "Bilinear and non-linear realizations of input/output maps", SIAM. J. Control, vol. 13, p. 827, 1975

|KRE,78| A.J. KRENER : "A note on commutative bilinear Optimal control", IEEE-AC-23, n° 6, p. 1111, 1978

|KRE,83| A.J. KRENER, A. ISIDORI : "Linearization by output injection and non-linear observers", Systems and control letters, vol. 3, p. 47, 1983

|KU,83| Y.H. KU : "Theory of non linear systems", J. Franklin Inst., vol. 315, n° 1, p. 1, 1983

|KUC,72| V. KUCERA : "A contribution to Matrix Quadratic equations", IEEE-AC-17, p. 344, 1976

|LAL,74| M. LAL, R. MITRA : "Simplification of large system dynamics using a moment evoluation algorithm", IEEE-AC-19, octobre, p. 602,1974

|LAM,80| M. LAMNABHI : "Séries de Volterra et séries génératrices non commutatives," Thèse de Docteur-Ingénieur, Paris, n° 432, 1980

|LAM,85| F. LAMNABHI-LAGARRIGUE : "Séries de Volterra et Commande optimale singulière", Thèse de Docteur es Sciences physiques, n° 2987, Paris-sud, 1985

|LAR,77| P. LARMINAT, Y. THOMAS : "Automatique des systèmes linéaires", Flammarion Sciences, 1977

|LAS,84| G.J. LASTMAN : "Reduced-order aggregated models for bilinear time-invariant dynamical systems"? IEEE-AC-29, n° 4, p. 359, 1984

|LEJ,85| R. LEJEUNE, W.J. RUGH : "Linearization of non linear systems about constant operating points", IEEE-AC-30, n° 8, p. 804, 1985

|LES,78| C. LESIAK, A.J. KRENER : "The existence and uniqueness of Volterra series for non linear systems", IEEE-AC-23, n° 6, p. 1030, 1978

|LEV,83| J. LEVINE : "Un aperçu élémentaire de la théorie moderne des systèmes non linéaires", RAIRO Automatique, vol. 17, n° 4, p. 277, 1983

|LIO,73| J.L. LIONS : "Perturbations singulières dans les problèmes aux limites et en contrôle optimal", Springer-Verlag, 1973

|LIO,78| J.L. LIONS : "Asymptotic methods in the optimal control of distributed systems", Automatica, vol. 14, p. 199, 1978

|LIT,81| L. LITZ : "Order reduction of linear state-space models via optimal approximation of the nondominant modes", Large Scale Systems, 2, p. 171, 1981

|LIU,84| C.C. LIU, Y.P. SHIH : "Analysis and parameter estimation of bilinear systems via Chebyskhev Polynomials", J. Franklin Inst., vol. 317, n° 6, p. 373, 1984

|LIU,84| C.C. LIU, Y.P. SHIH : "Analysis and parameter estimation of bilinear systems via Chebyskhev Polynomials", J. Franklin Inst., vol. 317, n° 6, p. 373, 1984

|LOB,76| C. LOBRY : "Critère de gouvernabilité des asservissements non linéaires", RAIRO Automatique, vol. 10, n° 11, p. 41, 1976

|LOP,82| K.A. LOPARO, C.S. HSU " Analysis of interconnected systems using bilinear models", Large Scale Systems, vol. 3, p. 237, 1982

|LUE,67| D.G. LUENBERGER "Canonical forms for linear multivariable systems" IEEE-AC-12, n° 3, p. 290,1967

|LUK,69| D.L. LUKES : "Optimal regulation of non-linear dynamical systems", SIAM J. Of Control, vol. 7, p. 75, 1969

|MAG,81| J.F. MAGNI : "Analyse et commande des systèmes à plusieurs échelles de temps", Thèse de Docteur-Ingénieur, ENSAE, Toulouse, 1981

|MAR,66| S.A. MARSHALL : "An approximate method for reducing the order of a linear system", Control, p. 642, décembre 1966

|MAR,85| R. MARINO : "High-gain feedback in non-linear control systems" Int. J. Control, vol. 42, n° 6, p. 1369, 1985 (P.S. Non linéaire et retour à grand gains)

|MEI,67| L. MEIER, D.G. LUENBERGER : "Approximation of linear constant systems", IEEE-AC-12, octobre, p. 585, 1977

|MER,64| C.W. MERRIAH III : "Optimisation theory and the design of feedback control systems" Mac Graw-hill, 1964

|MIC,79| G. MICHAILESCO, J.M. SIRET, p. BERTRAND : "Agrégation de variables d'état et applications" in "Analyse et commande des systèmes complexes", A. TITLI, Cepadues, 1979

|MIC,86| G. MICHAILESCO : "Agrégation des systèmes linéaires : principes et applications", Proc. of IMACS Symp. on "Modelling and Simulation for Control of Lumped and Distributed Parameter Systems" Villeneuve d'Ascq, June 86, p. 593

|MIT,69| D. MITRA : "On the reduction of complexity of linear, time invariant, dynamic systems", Proc 4th IFAC, Warsaw, Technical series, 67, p. 19, 1969

|MIT,79| G.E. MITZEL, S.J. CLANCY, W.J. RUGH : "On the transfer function representations for homogeneous non-linear systems" IEEE-AC-24, n° 2, p. 242, 1979

|MOH,70a| R.R. MOHLER : "Natural bilinear control processes", IEEE-SMC-6, n° 3, p. 192, 1970

|MOH,70b| R.R. MOHLER, C.N. SHEN : "Optimal Control of nuclear reactors", Academic Press, 1970

|MOH,73| R.R. MOHLER "Bilinear control processes", Academic Press,1970

|MOH,76| R.R. MOHLER, C.F. BARTON, V.R. KARANAM " BLS identification by orthogonal functions with application to immunology", Conf. Decis. Control. Symp. Adapt. Processes, Clearwater, p. 143, 1976

|MOH,80| P.R. MOHLER, W.J. KOLODZIEJ : "An overview of bilinear systems : theory and applications", IEEE-SMC-10, n° 10, p. 683, 1980

|MOR,84| O. MOREIGNE : "Contribution à la modélisation et à la synthèse des systèmes à échelles de temps multiples. Application au calcul en temps réel d'une commande quasi-optimale pour les trajectoires d'avions de combat", Thèse de Docteur-Ingénieur, Lille, 1984

|MOU,85| S.G. MOUROUTSOS, P.H. PARASKEVOPOULOS : "Identification of time-varying linear systems using orthogonal functions", J. Franklin Inst., vol. 320, n° 5, p. 249, 1985

|NAI,81| G.G.NAIR : "A note on Optimal control of a class of uncontrained bilinear systems", Int. J. Systems. Sci., vol. 12, n°2, p. 261, 1981

|NAY,73| A.H. NAYFEH : "Perturbation Methods", Wiley, 1973

|NEU,69| H. NEUDECKER : "Some theorems on matrix differentiation with special reference to Kronecker matrix products", J. Amer. Stat. Assoc., vol. 64, p. 953, 1969

|NOU,83| J.P. NOUGIER : "Méthodes de calcul numérique", Masson, 1983

|O'MA,74| R.E. O'MALLEY : "Introduction to singular perturbations", Academic Press, 1974

|O'MA,82| R.O'MALLEY, L.R. ANDERSON : "Time scale decoupling and order reduction for linear time-varying systems", Optimal control applications and methods, vol. 3, p. 139, 1982

|PAR,85| P.N. PARASKEVOPOULOS : "System analysis and synthesis via orthogonal polynomial series and Fourier series", Mathematics and Computer in Simulation, vol. 27, n° 5-6, p. 453, 1985

|PEP,82| G. PEPONIDES, P.V. KOKOTOVIC, J.H. CHOW : "Singular perturbations and time scales in non-linear models of power systems" IEEECAS-29, n° 11, P. 758, 1982

|PIN,81| P. PINGUET : "Sur une classe de méthodes de réduction de l'ordre des systèmes dans l'espace d'état", Thèse de Docteur-Ingénieur, Lille, N° 283, 1981

|POP,73| V.M. POPOV : "L'Hyperstabilité des systèmes automatiques", Dunod, 1973

|POW,66| D.L. POWERS : "A block Gershgorine theorem", Linear Algebra and its applications, 13, p. 45, 1976

|RAO,75| G.P. RAO, L. SIVAKUMAR : "System identification via Walsh functions", Proc. IEE, vol. 122, p. 1160, 1975

|RAO,85| G.P. RAO, S.G. TZAFESTAS : "A decade of piece wise constant orthogonal polynomial series and Fourier series", Mathematics and Computer in Simulation, vol. 27, n° 5-6, p. 453, 1985

|REI,82| H. REINHARD : "Equations différentielles", Gauthiers Villars, 1982

|RIN,68| R.E. RINK, R.R. MOHLER : "Completely controllable bilinear systems", SIAM J. Control, vol. 6, n° 3, p. 477, 1968

|ROI,74| I. ROITENGERG : "Théorie du controle Automatique", Ed. MIR, Moscou, 1974

|ROT,79| E.A. ROTH : "On the higher-order stroboscopic method", J. Appl. Math. and Phys., vol. 30, p. 315, 1979

|ROT,85| F. ROTELLA, G. DAUPHIN-TANGUY, P. BORNE : "Détermination and separation of dynamics for multi-time scale bilinear systems", 1st European Workshop on Parallel Processing, Manchester, 1985

|ROT,86a| F. ROTELLA, I. ZAMBETTAKIS, G. DAUPHIN-TANGUY, P. BORNE : "Bilinéarisation of non-linear analytical Systems : application to the simulation and modelling of dynamic manipulators", 5e IFAC Symposium on Problems in Manufacturing, Suzdal, URSS, 1986

|ROT,86b| F. ROTELLA, I. ZAMBETTAKIS, G. DAUPHIN-TANGUY, P. BORNE : "Modelling and Simulation of non-linear systems based on infinite bilinear realizations", IFAC/IMACS Symposium on Simulation of Control Systems, Septembre 22-26, Vienne, 1986

|ROT,86c| F. ROTELLA, G. DAUPHIN-TANGUY : "Multi-model representation for singularly perturbed bilinear systems", IMACS Symposium on Modelling and Simulation for Control of Lumped and Distributed Parameter Systems, Villeneuve d'Ascq, p. 573, juin 1986

|ROU,73| N. ROUCHE, J. MAWHIN : "Equations différentielles ordinaires", Masson, 1973

|SAK,84| V.R. SAKSENA, J. O'REILLY, P.V. KOKOTOVIC : "Singular perturbations and time-scale methods in control theory : survey 1976-1983", Automatica, vol. 20, n° 3, p. 273, 1984

|SAN,77| P. SANNUTI : "Analysis and synthesis of dynamic system via blockpulse function", Proc. IEE, vol. 124, p. 569, 1977

|SAN,78| N.R. SANDELL, P. VARAIYA, M. ATHANS, M.G. SAFONOV : "Survey of decentralized control methods for large scale systems", IEEE-AC-23, n° 2, p. 108, 1978

|SCH,61| M.P. SCHUTZENBERGER : "on the definition of a familly of Automata", Inf. and Control, vol. 4, p. 245, 1961

|SHA,74a| Y. SHAMASH : "Stable reduced order models using Pade type approximation", IEEE-AC-19, p. 615, 1974

|SHA,74b| Y. SHAMASH : ""Continued fraction for the reduction of discrete-time dynamic systems", I.J.C. vol. 20, n° 2, p. 267, 1974

|SHA,75| Y. SHAMASH : "Model reduction using the Routh stability criterion and the Pade approximation technique", I.J.C., p. 475, mars 1975

|SHI,83| Y.P. SHIH : "Application of Chebyslev polynomials in analysis and identification of linear systems", J. Chin. Inst. Engrs., vol. 6, p. 135, 1983

|SIL,84| R. SILVA-MADRIZ, S.S. SASTRY : "Input-output description of linear systems with multiple time-scales", I.J.C., vol. 40, n° 4, p. 699, 1984

|SIN,80| M.G. SINGH : "Dynamical hierarchical control", North Holland, 1980

|SLE,78| M. SLEMROD : "Stabilization of bilinear control systems with applications to non conservative problems in elasticity", SIAM J. Control and Optimization, vol. 16, p. 131, 1978

|SOL,80| O.A. SOLHEIM : "On the use of a block analog of the

Gershgorine circle theorem in the design of decentralized control of large scale systems", IFAC Symposium on Large Scale Systems, p. 103, Toulouse, 1980

|SUS,76| H.J. SUSSMAN : "Semigroup representations bilinear approximation of input-output maps and generalized inputs", in Mathematical Systems Theory, Springer-Verlag, 1976

|SVO,80| S. SVORONOS, G. STEPHANOPOULOS, R. ARIS : "Bilinear approximation of general non-linear dynamic systems with linear inputs", I.J.C., vol. 31, n° 1, p. 109, 1980

|SYR,82| G. SYRCOS, P. SANNUTI : "Transformation of physical models into singularly perturbed form", IFAC New Delhi Symposium, 1982

|TIH,52| A.N. TIHONOV : "Systems of differential equations containing a small parameter multiplying the derivative", Mat. Sb., vol. 31, n° 73, p. 575, 1952

|TIN,75| J. TING-LO LO : "Global bilinearization of systems with control appearing linearly", SIAM J. Control, vol. 13, n° 4, p. 879, 1975

|TIT,79| A. TITLI, R. HURTEAU : "Partionning and time decomposition for the control of interconnected systems", 2nd IFAC Symp. on Optimization : method, application, aspects, Varna, Bulgarie, 1979

|TZA,78| S.G. TZAFESTAS : "Walsh series approach to lumped and distributed system identification", J. Franklin Inst., vol. 305, p. 199, 1978

|TZA,84| S.G. TZAFESTAS, K.E. ANAGNOSTOU : "Stabilization of singularly perturbed strictly bilinear systems", IEEE-AC-29, n° 10, p. 943, 1984

|VAN,84| A. VAN HARTEN : "Singularly perturbed systems of diffusion type and feedback control", Automatica, vol. 20, n° 1, p. 79, 1984

|VET,70| W.J. VETTER : "Derivative operations on matrices", IEEE-AC-15, n° 2, p. 241, Avril 1970

|WEI,78| K.C. WEI, A.E. PEARSON : "On minimum energy control of commutative bilinear systems", IEEE-AC-23, n° 6, p. 1020, 1978

|WIL,70| J.L. WILLEMS : "Stability theory of dynamical systems", Nelson, 1970

|WIL,72| R.G. WILSON, D.G. FISHER, D.E. SEBORD : "Model reduction of discrete-time dynamic systems", I.J.C., vol. 16, n° 3, p. 459, 1972

|WIL,74| D.A. WILSON : "Model reduction for multivariable systems",

I.J.C., vol. 19, n° 1, p. 57, 1974

|YOU,77| K.D. YOUNG, P.V. KOKOTOVIC, V.I. UTKIN : "A singular perturbation analysis of high gain feedback systems", IEEE-AC-22, n° 6, p. 931, 1977

|ZAG,76| E. ZAGER, D. TABAK : "Optimal control of a class of uncontrained bilinear systems", Int. J. Systems. Sci., vol. 7, n° 12, p. 1349, 1976

|ZAK,73| V. ZAKIAN : "Simplification of linear time-invariant systems by moment-matching", I.J.C. vol. 18, n° 3, p. 455, 1973

ZAK

|ZAK,86| S.H., C.A. MACCALREY : "State feedback control of non-linear systems", I.J.C., vol. 43, n° 5, p. 1497, 1986

# TABLE DES MATIERES

TABLE DES MATIERES


INTRODUCTION GENERALE p. 7

CHAPITRE I : STRUCTURE DES SYSTEMES p. 11

INTRODUCTION p. 13

I. STRUCTURE DES SYSTEMES CONTINUS p. 14

II. CLASSE DE SYSTEMES p. 15

II.1. LES SYSTEMES LINEAIRES p. 15

II.1.1. Forme externe des systèmes linéaires p. 16
II.1.2. Forme interne des systèmes linéaires p. 17

II.2. LES SYSTEMES A COMMANDE AFFINE p. 18

II.2.1. Formule fondamentale p. 18
II.2.2. Calcul symbolique non commutatif p. 20
II.2.3. Estimation du transitoire p. 21
II.2.4. Forme commandable des systèmes non linéaires p. 23
II.2.5. Exemple de systèmes à commande affine p. 24

II.3. LES SYSTEMES BILINEAIRES p. 25

II.3.1. Structure et définitions p. 25
II.3.2. Intérêt des systèmes bilinéaires p. 28
II.3.3. Modèles mathématiques p. 29

a) fonction de transfert multivariable p. 30
b) codage des systèmes bilinéaires p. 31
c) fonction de transfert codée p. 31
d) matrices de Hankel p. 32

III. BILINEARISATION DES SYSTEMES NON LINEAIRES p. 34

III.1. NOTATIONS ET PROPRIETES p. 34

III.2. METHODE D'APPROXIMATION p. 36

III.3. ELIMINATION DE LA REDONDANCE p. 38

III.4. EXEMPLE D'APPLICATION

CONCLUSION p. 42

**CHAPITRE II : SIMPLIFICATION DE MODELES** p. 43

**INTRODUCTION** p. 45

I. METHODES DE SIMPLIFICATION DE MODELES p. 46

I.1. REDUCTION OPTIMALE p. 46

I.2. APPROXIMATION DE FONCTIONS DE TRANSFERT p. 47

I.3. TECHNIQUES D'AGREGATION p. 50

I.4. METHODES DE PERTURBATIONS p. 53

I.4.1. Perturbations régulières p. 54
I.4.2. Perturbations singulières p. 55

II. QUELQUES RESULTATS FONDAMENTAUX CONCERNANT LES SYSTEMES SINGULIEREMENT PERTURBES p. 57

II.1. CAS GENERAL DES SYSTEMES NON LINEAIRES p. 57

II.1.1. Réduction d'ordre et approximation p. 57
II.1.2. Développement asymptotique p. 60
II.1.3. Cas du retour à grand gain p. 62

II.2. CAS DES SYSTEMES LINEAIRES p. 63

II.2.1. Modèles réduits lent et couche limite p. 63

II.2.1.1. Etude de la partie lente p. 64
II.2.1.2. Etude de la partie rapide p. 64

II.2.2. Modélisation par découplage p. 65

II.2.3. Utilisation du développement asymptotique p. 68

II.2.4. Utilisation du système réciproque p. 71

II.2.4.1. Transformation réciproque p. 72
II.2.4.2. Transformation réciproque et perturbations singulières p. 74
II.2.4.3. Utilisation du système réciproque et du découplage p. 76

II.2.5. Systèmes singuliers singulièrement perturbés p. 79

**CONCLUSION** p. 81

**CHAPITRE III : DETECTION DE DYNAMIQUES DIFFERENTES** p. 83

INTRODUCTION

I. MISE EN EVIDENCE DES DYNAMIQUES DANS LE CAS DES SYSTEMES LINEAIRES p. 86

I.1. METHODES ALGEBRIQUES p. 86

I.1.1. Bloc diagonalisation p. 86
I.1.2. Méthodes de détermination des matrices L et M p. 88

I.1.2.1. Utilisation du polynome annulateur P. 89
I.1.2.2. Algorithmes numériques p. 93

I.2. METHODES GEOMETRIQUES p. 96

I.2.1. Utilisation des cercles de Gershgorine p. 97
I.2.2. Utilisation du Théorème de Gershgorine par blocs p. 99

II. APPLICATION AUX SYSTEMES BILINEAIRES p. 104

II.1. SYSTEMES LINEAIRES NON STATIONNAIRES A DEUX ECHELLES DE TEMPS p. 105

II.2. METHODE ALGEBRIQUE DE PSEUDO-DIAGONALISATION p. 107

II.3. UTILISATION DES METHODES GEOMETRIQUES p. 113

II.3.1. Niveau global p. 113
II.3.2. Niveau local p. 114

II.3.2.1. Cercles de Gershgorine p. 115
II.3.2.2. Méthode de conditionnement p. 117

CONCLUSION p. 122

CHAPITRE IV : PERTURBATIONS SINGULIERES ET SYSTEMES BILINEAIRES p. 123

INTRODUCTION p. 125

I. SYSTEMES BILINEAIRES SINGULIEREMENT PERTURBES p. 126

I.1. MODELE REDUIT LENT p. 127

I.2. CAS PARTICULIERS DE MODELES REDUITS LENTS BILINEAIRES p. 128

I.2.1. Système régulier à entrée scalaire p. 128
I.2.2. Systèmes commutatifs p. 128
I.2.3. Conditions de bilinéarité p. 129

I.3. SYSTEMES COUCHE LIMITE p. 130

I.4. DEVELOPPEMENT ASYMPTOTIQUE DE LA SERIE GENERATRICE p. 131

II. BILINEARISATION DU MODELE REDUIT LENT p. 135

II.1. DOMINANCE D'ENTREES p. 135

II.2. METHODE DIRECTE p. 137

II.2.1. Approximation du vecteur de commande p. 137
II.2.2. Approximation de la matrice d'évolution p. 138

a) utilisation du calcul précédent p. 138
b) utilisation d'un calcul équivalent p. 138
c) utilisation du calcul direct p. 139

II.3. IDENTIFICATION p. 141

II.3.1. Principe de la méthode d'identification p. 141
II.3.2. Utilisation des fonctions de Walsh p. 143

a) définition des fonctions de Walsh p. 143
b) propriétés pour l'identification p. 143

II.4. EXEMPLE p. 145

II.4.1. Méthode directe p. 145
II.4.2. Identification p. 146

III. DECOUPLAGE DES PARTIES LENTES ET RAPIDES p. 147

IV. EXTENSION DU SYSTEME RECIPROQUE POUR LES SYSTEMES BILINEAIRES p. 150

IV.1. DEFINITION p. 150

IV.2. CONSTRUCTION p. 151

IV.3. UTILISATION AVEC LA METHODE DES PERTURBATIONS SINGULIERES p. 153

★ Résultats de simulations p. 156

CONCLUSION p. 157

CHAPITRE V : COMMANDE p. 159

INTRODUCTION p. 161

I. COMMANDE DES SYSTEMES LINEAIRES p. 162

I.1. UTILISATION DE LA SORTIE p. 163

I.2. COMMANDE SOUS-OPTIMALE PAR LES METHODES DE PERTURBATION p. 166

I.2.1. Utilisation des perturbations régulières p. 166
I.2.2. Utilisation des perturbations singulières p. 167

I.2.2.1. Décomposition de l'équation de Riccati p. 167
I.2.2.2. Commande quasioptimale par décomposition temporelle p. 167
I.2.2.3. Retour de sortie p. 174

II. COMMANDE DES SYSTEMES NON LINEAIRES p. 176

II.1. SYSTEMES A COMMANDE AFFINE ANALYTIQUES p. 177

II.1.1. Conditions nécessaires d'optimalité p. 177
II.1.2. Equation fondamentale d'approximation p. 178

II.2. DETERMINATION DES $P_i$ p. 180

II.2.1. Premier ordre p. 180
II.2.2. Deuxième ordre p. 181

II.3. EXPRESSION ANALYTIQUE DU RETOUR D'ETAT p. 183

II.4. CAS DES SYSTEMES SCALAIRES p. 184

II.5. TRANSMISSION DIRECTE p. 185

III. CAS DES SYSTEMES BILINEAIRES p. 186

III.1. CAS GENERAL : RETOUR D'ETAT ANALYTIQUE p. 186

III.2 SYSTEMES BILINEAIRES SINGULIEREMENT PERTURBES p. 188

III.2.1. Décomposition des équations p. 188
III.2.2. Décomposition temporelle p. 189

IV. EXEMPLE p. 190

IV.1. COMMANDE ANALYTIQUE GLOBALE p. 190

IV.2. SIMPLIFICATION DE LA COMMANDE p. 194

ANNEXE p. 198

CONCLUSION p. 206

CONCLUSION GENERALE p. 207

BIBLIOGRAPHIE p. 211

BIBLIOTHEQUE INTERUNIVERSITAIRE
SECTION DE SCIENCES
LILLE

**RESUME**

Après la présentation de diverses méthodes permettant de mettre en évidence un modèle bilinéaire à plusieurs dynamiques, l'application de la méthode des perturbations singulières conduit à des modèles réduits découplés. Plusieurs techniques sont proposées pour les rendre bilinéaires.

L'utilisation de ces modèles réduits permet de générer des commandes quasi-optimales simplifiées. Dans ce but, dans le cadre plus général des systèmes non-linéaires analytiques à commande affine, on construit un retour d'état non-linéaire dont l'expression est analytiquement déterminée. Cette méthode, ainsi que les principes de commandes quasi-optimales des systèmes linéaires singulièrement perturbés, sont utilisés dans le cas particulier des systèmes bilinéaires à deux dynamiques.

BIBLIOTHEQUE INTERUNIVERSITAIRE
SECTION DE SCIENCES
LILLE

**MOTS - CLES**

SYSTEMES BILINEAIRES
SYSTEMES NON-LINEAIRES
PERTURBATIONS SINGULIERES
COMMANDE OPTIMALE
MODELISATION
PRODUIT DE KRONECKER
REDUCTION D'ORDRE