



UNIVERSITE de CAEN/BASSE-NORMANDIE

U.F.R. : Sciences

ECOLE DOCTORALE: SIMEM

THESE

présentée par

Mr Youssef ROCHDI

et soutenue

le 21 décembre 2006

en vue de l'obtention du

DOCTORAT de l'UNIVERSITE de CAEN

Spécialité : Automatique-robotique

(Arrêté du 25 avril 2002)

Titre: Identification des systèmes non linéaires blocs

MEMBRES du JURY

Mr José RAGOT , Professeur des universités, INP de Lorraine, Nancy (Rapporteur)

Mr Jean-Claude TRIGEASSOU, Professeur des universités, Université de Poitiers, (Rapporteur)

Mr Luc Dugard, Directeur de recherche CNRS, INP de Grenoble

Mr Mohamed Msaad, Professeur des universités, ENSICAEN

Mr Fouad Giri , Professeur des universités, Université de Caen (Directeur de thèse)

DEDICACE

Je dédie ce travail à tous les membres de ma famille, et spécialement à mes parents. C'est grâce à leur encouragement, leur soutien aux moments critiques, et leur compréhension vis-à-vis de mon indisponibilité pour eux, que ce travail a pu aboutir.

REMERCIEMENTS

J'exprime toute ma gratitude à Monsieur Fouad GIRI, professeur des universités à l'université de Caen, qui a dirigé cette thèse. Je le remercie pour sa contribution scientifique à ce travail, son aide inestimable et les nombreux conseils qu'il n'a cessé de me prodiguer tout au long de ce travail.

Je remercie Monsieur José RAGOT, professeur des universités à l'INP de Lorraine à Nancy, et Monsieur Jean-Claude TRIGEASSOU, professeur des universités à l'université de Poitiers, d'avoir accepté d'examiner ce travail et d'en être les rapporteurs.

Je remercie Monsieur Luc DUGARD, directeur de recherche CNRS à l'INP de Grenoble, et Monsieur Mohamed M'SAAD, professeur des universités à l'ENSICAEN, pour l'intérêt qu'ils ont manifesté pour ce travail en acceptant de faire partie du Jury.

Je remercie également Mme Fatima-zahra CHAOUI, professeur à l'ENSET de Rabat, pour sa contribution directe à l'aboutissement de ce travail.

J'exprime ma gratitude à toute personne qui, de près ou de loin, m'a aidé à réaliser ce travail.

LISTE DES PUBLICATIONS DE L'AUTEUR

F. Giri, F.Z. Chaoui, Y. Rochdi

"Parameter Identification of a class of Hammerstein plants", **Automatica**, 37, pp. 749-756, 2001.

F. Giri, F.Z. Chaoui, Y. Rochdi

"Interval excitation through impulse sequences", **Automatica**, 38, pp. 457-465, 2002.

F.Z. Chaoui, F. Giri, Y. Rochdi, M. Haloua, A. Naitali

"System Identification based on Hammerstein model", **International Journal of Control**, vol. 78, N° 6, 2005, pp. 430-442.

F. Giri, F.Z. Chaoui, Y. Rochdi

"Adaptive identification of plants with nonlinear static gain", **IFAC Symposium on System Identification (SYSID 2000)**, 21-23 June 2000, Santa Barbara, California, USA.

F. Giri, F.Z. Chaoui, Y. Rochdi

"Hammerstein model Identification - The Case of Piecewise-Polynomial Static-Gain", **IEEE-ACC**, Arlington, Virginia, USA, vol.6, pp. 4882 – 4887, 2001.

F. Giri, F.Z. Chaoui, Y. Rochdi

"A unified analysis of self-exciting regulators using internal impulse sequences", **IEEE-ACC**, Arlington, Virginia, USA, vol. 4 , pp. 3299 – 3304, 2001.

F. Giri, F.Z. Chaoui, M. Haloua, Y. Rochdi, A. Naitali,

"Hammerstein model Identification", Proceedings of the **10th Mediterranean Conference on Control and Automation**, Lisbon, Portugal, July 9-12, 2002.

<http://med.ee.nd.edu/MED10/pdf/193.pdf>

A. Naitali, F. Giri, F.Z. Chaoui, M. Haloua, Y. Rochdi

"Parameter Identification Based on Hammerstein Models-A subspace Approach",
IFAC Symposium on System Identification (SYSID), Rotterdam, The Netherlands, pp.
1327-1332, 2003.

F. Giri, F.Z. Chaoui, Y. Rochdi

"Recursive Identification of Systems with Hard Input Nonlinearities of Known Structure".
IEEE American Control Conference, ACC'04, Boston, Massachusetts, USA, 30 June-02
July, 2004, pp. 4764-4769.

F. Giri, F.Z. Chaoui, Y. Rochdi

"System Identification of Hammerstein models including a known structure input
nonlinearity - The Hysterisis case". **IFAC Workshop on Adaptation and Learning in
Control and Signal Processing**, Yokohama, Japan, August 30th- September 1st, 2004.
Session MM3-A-4 (n° pages non disponibles)

F. Giri, Y. Rochdi, F.Z. Chaoui

"Identification of the Nonlinear Element in Wiener Models. A Frequency-Geometric
Solution", **IEEE Conference on Decision and Control, CDC'06**, 2006

SOMMAIRE

I. Première partie : Introduction générale et outils théorique pour l'identification

Chapitre 1: Introduction générale

1.1. Introduction	3
1.2. Modélisation des systèmes non linéaires.....	4
1.3. Identification des modèles de Hammerstein et de Wiener: Aperçu sur les travaux antérieurs.....	7
1.3.1. Architecture des schémas d'identification.....	8
1.3.2. Connaissances a priori dans les schémas d'identification	9
1.3.3. Nature des perturbations et des erreurs de modélisation dans les schémas d'identification.....	11
1.3.4. Analyse de la consistance et techniques d'excitation dans les schémas d'identification.....	11
1.3.5. Complexité des schémas d'identification	12
1.4. Contributions de la thèse	13
1.4.1. Excitation persistante à l'aide de signaux impulsionnels.....	13
1.4.2. Contributions à l'identification du modèle de Hammerstein	14
1.4.3. Contributions à l'identification du modèle de Wiener.....	16
1.5. Organisation du mémoire.....	18
1.5.1. Structure du mémoire	18
1.5.2. Aperçu sur le contenu des chapitres.....	20

Chapitre 2: Outils théoriques pour l'identification

2.1. Introduction.....	23
2.2. Excitation persistante (EP) à l'aide d'un type de séquences impulsionnelles....	23
2.2.1. Intérêt de l'excitation persistante.....	24
2.2.2. Définitions de l'excitation persistante.....	24
2.2.3. Génération de l'excitation persistante.....	25
2.2.4. Lemme technique d'excitation.....	26
2.2.5. Applications du Lemme d'excitation.....	28
2.3. Notions de petitesse en moyenne.....	30

2.4. Algorithmes d'identification paramétrique.....	31
2.4.1. Identification paramétrique en l'absence d'erreur de sortie	31
2.4.2. Identification paramétrique en présence d'une erreur de sortie de modèle inconnu	34
2.4.3. Identification paramétrique en présence d'une erreur de sortie modélisable par un processus stochastique de modèle connu.....	37
2.5. Résolution d'un type de système d'équations algébriques bilinéaires	38

II. Deuxième partie: Identification des systèmes non linéaire de type Hammerstein

Chapitre 3: Identification des systèmes de Hammerstein impliquant un élément non linéaire statique de structure inconnue

3.1. Introduction.....	45
3.2. Position du problème d'identification	46
3.2.1. Classe de procédés à identifier	46
3.2.2. Objectif d'identification	47
3.3. Paramétrisation du modèle du système	48
3.4. Estimation des paramètres du modèle.....	51
3.5. Signal d'entrée procurant une excitation persistante.....	54
3.6. Analyse de convergence des paramètres estimés.....	57
3.7. Approximation de l'élément non linéaire.....	58
3.8. Simulations.....	59
3.9. Conclusions	65

Chapitre 4: Identification des systèmes de Hammerstein impliquant un élément non linéaire statique de structure connue

4.1. Introduction.....	67
4.2. Position du problème d'identification.....	69
4.2.1. Classe de procédés à identifier	69
4.2.2. Objectif d'identification.....	71
4.3. Reparamétrisation du système.....	71
4.4. Estimation des paramètres du système	74
4.5. Excitation persistante et consistance des paramètres estimés	75
4.6. Estimation des paramètres caractéristiques de l'élément non linéaire.....	77
4.7. Simulations	82

4.8. Conclusions	87
------------------------	----

Chapitre 5: Identification des systèmes de Hammerstein impliquant un élément non linéaire de type hystérétique

5.1. Introduction.....	89
5.2. Notations mathématiques préliminaires.....	91
5.2.1. Notions et acronymes.....	91
5.2.2. Une classe de séquences excitantes persistantes pour les systèmes linéaires.....	92
5.3. Position du problème d'identification.....	93
5.3.1. Classe des systèmes à identifier	93
5.3.2. Objectif d'identification	95
5.4. Identification du sous système linéaire.....	95
5.5. Identification du gain non linéaire	104
5.6. Conclusions	113

III. Troisième partie : Identification des systèmes non linéaires de type Wiener

Chapitre 6: Identification de l'élément non linéaire des systèmes de Wiener par la localisation des extremums

6.1. Introduction.....	117
6.2. Position du problème d'identification	119
6.1.1. Classe de procédés à identifier	119
6.1.2. Objectif d'identification	120
6.3. Analyse préliminaire du système	121
6.4. Identification d'éléments non linéaires monotones.....	125
6.5. Identification d'éléments non linéaires non monotones.....	128
6.6. Simulations	134
6.7. Conclusions	136

Chapitre 7: Identification de l'élément non linéaire des systèmes de Wiener à l'aide des courbes de Lissajous

7.1. Introduction.....	139
7.2. Position du problème	140

7.1.1. Classe de procédés à identifier	140
7.1.2. Objectif d'identification	142
7.3. Mise à l'échelle du modèle et reformulation du problème d'identification	142
7.4. Caractérisation géométrique des courbes Sortie/signal d'entrée	145
7.5. Analyse des courbes paramétrées Sortie/ signal interne	151
7.6. Version graphique de la méthode d'identification de l'élément non linéaire.....	153
7.7. Version analytique de la méthode d'identification de l'élément non linéaire.....	157
7.8. Conclusions	159
Chapitre 7: Conclusions générales.....	161
Annexes	
Annexe A	169
Annexe B	177
Annexe C	185
Bibliographie.....	193

PREMIERE PARTIE

INTRODUCTION GENERALE ET OUTILS THEORIQUES POUR L'IDENTIFICATION

CHAPITRE 1

INTRODUCTION GENERALE

1.1. INTRODUCTION

Le problème d'identification des systèmes suscite un grand intérêt et ce depuis plusieurs décennies. L'activité de recherche menée autour de ce problème a fait l'objet de très nombreux colloques et séminaires et a donné lieu à d'innombrables publications. Si l'identification des systèmes linéaires est maintenant arrivée à maturité, celle des systèmes non linéaires reste à ce jour un thème de recherche loin d'être tari. L'intérêt suscité par ce thème s'explique simplement par le fait que le caractère non linéaire des systèmes physiques est plutôt la règle que l'exception. La nature non linéaire peut être d'origine structurelle (par exemple le couple résistant d'un ventilateur varie proportionnellement au carré de sa vitesse) mais elle peut aussi être la conséquence d'imperfections de construction (par exemple la non linéarité d'un potentiomètre de capture de position, les jeux dans les systèmes mécaniques). En outre, ce caractère peut difficilement être ignoré quand le modèle identifié doit servir comme base à la synthèse de prédicteurs ou régulateurs hautement performants. De ce fait, l'intérêt porté à l'identification (comme d'ailleurs à la commande) des systèmes non linéaires n'est pas du tout récent. Certaines publications dans ce domaine remontent en effet aux années quarante! Le nombre de contributions développées jusqu'à ce jour est vraiment impressionnant. La diversité des approches suivies reflète avant tout la complexité du problème de modélisation des systèmes non linéaires. Dans ce chapitre, nous tenterons de dresser un inventaire des principales méthodes d'obtention de modèles par l'approche de l'identification. Nous introduirons les deux structures de modèle qui ont le plus suscité l'intérêt des chercheurs et qui aussi se trouvent au centre du présent mémoire. Nous donnerons un aperçu sur les méthodes d'identification, rapportées dans la littérature, qui s'appuient sur ces structures et situerons nos contributions parmi ces méthodes. Le chapitre sera achevé par une description du plan du rapport de thèse.

1.2.MODELISATION DES SYSTEMES NON LINEAIRES

Toute approche d'identification s'appuie sur une structure de modèle adaptée au système considéré. Cette structure peut être déterminée moyennant une modélisation précise du système. Cependant, il arrive couramment que l'approche par modélisation conduise à une structure de modèle trop complexe pour être utilisée comme base au processus d'identification. Dans certaines situations, la complexité du système et de son interaction avec son environnement est telle qu'il s'avère très difficile, voire impossible, de parvenir à un modèle par cette approche. Dans ces cas, il est normal de se tourner vers l'approche par identification qui consiste à déterminer un modèle dont la structure est choisie à l'avance parmi une variété de structures 'universelles'. Les modèles ainsi obtenus ne rendent pas compte des caractéristiques physiques du système et, de ce fait, ne se prêtent pas à l'élaboration de simulateurs. En revanche, ils s'avèrent très souvent adéquats pour la synthèse de prédicteurs et de régulateurs.

Pour faire face à la diversité des systèmes non linéaires, de très nombreuses méthodes d'identification furent développées. Néanmoins, on peut comme dans le cas linéaire classer ces méthodes en deux grandes catégories : la première étant celle qui regroupe les méthodes qui visent des modèles paramétriques, la seconde regroupe les méthodes qui portent sur les modèles non paramétriques. Les modèles paramétriques sont caractérisés par une structure mathématique connue impliquant un nombre fini de paramètres inconnus ; la détermination de ceux-ci constitue le but de l'identification. Les modèles non paramétriques n'ont aucune structure mathématique connue et de ce fait ils sont caractérisés par un nombre de 'paramètres' infiniment grand. Pour illustrer le propos on peut citer les modèles linéaires :

$$y(t) = \int_0^t g(\tau)u(t-\tau)d\tau \quad (\text{Systèmes linéaires continus}) \quad (1.1)$$

$$y(t) = \sum_{\tau=0}^{t-1} g(\tau)u(t-\tau) \quad (\text{Systèmes linéaires discrets}) \quad (1.2)$$

De tels modèles sont caractérisés par la réponse impulsionnelle du système $g(t)$, donc par l'ensemble des valeurs de celle-ci.

Historiquement, ce sont les modèles non paramétriques qui ont été utilisés en premier dans la modélisation des systèmes non linéaires par l'approche de l'identification. Ces modèles requièrent moins de connaissance préalable sur le système identifié. Plusieurs méthodes utilisées aujourd'hui dans le cadre de l'identification des modèles paramétriques, furent développées à l'origine pour les

modèles non paramétriques. En revanche, ces modèles sont généralement moins commodes pour la synthèse de régulateurs. Parmi les modèles non paramétriques, on peut citer les développements en série de Volterra qui découlent d'une tentative de généralisation immédiate au cas non linéaire des expressions (1.1)-(1.2):

$$y(t) = \sum_{i=0}^n \int_{\tau_1=0}^t \cdots \int_{\tau_i=0}^t g_i(\tau_1, \dots, \tau_i) \prod_{j=1}^i u(t - \tau_j) d\tau_1 \dots d\tau_i \quad (\text{Systèmes continus})(1.3)$$

$$y(t) = \sum_{i=0}^n \dots \sum_{\tau_i=0}^k g_i(\tau_1, \dots, \tau_i) \prod_{j=1}^i u(t - \tau_j) \quad (\text{Systèmes discrets}) \quad (1.4)$$

Ces modèles sont caractérisés par les noyaux de Volterra à savoir les fonctions $g_n(\tau_1, \dots, \tau_n)$; celles-ci ne dépendent que du système et jouent dans (1.3)-(1.4) un rôle comparable à celui de la réponse indicielle $g(t)$ dans les modèles linéaires (1.1)-(1.2). L'identification du système équivaut à l'estimation d'un certain nombre de ces noyaux. Il est largement admis que ce nombre doit être assez élevé si l'on veut atteindre une précision satisfaisante et ce même pour des systèmes non linéaires relativement simples (Haber et Keviczky, 1999a). En outre, les modèles de Volterra s'avèrent inadaptés quand le système comporte des éléments discontinus. Pour toutes ces raisons, ils sont rarement utilisés en pratique. Pour diminuer le nombre de paramètres nécessaires (à l'obtention d'une bonne précision du modèle), Wiener (1958) proposa de faire recours à des développements en série de fonctions orthogonales. A cet effet, les versions multivariées des polynômes de Laguerre et de Hermite sont invoquées donnant lieu à de nouveaux noyaux dits de Wiener. L'intérêt de l'orthogonalité réside aussi dans le fait que l'estimation des différents noyaux peut se faire de façon indépendante. En particulier, si l'on décide d'augmenter d'un cran l'ordre de troncature du développement, il suffit d'estimer les noyaux d'ordre élevé ainsi générés, ceux qui ont été estimés auparavant gardent leurs valeurs.

Le lecteur désireux de creuser cet aspect est invité à consulter les deux volumes de l'ouvrage de Haber et Keviczky (Haber et Keviczky, 1999a-b). Etant relativement récent, cet ouvrage fait aussi l'état de l'art de l'identification des systèmes non linéaires et comprend de ce fait une liste assez complète des références bibliographiques.

Les modèles paramétriques sont beaucoup plus pratiques aussi bien pour l'identification que pour la commande mais ils requièrent généralement plus de connaissance a priori sur le système. Cette connaissance est requise notamment pour sélectionner un type de modèle approprié pour le système et adéquat pour l'usage que l'on veut en faire. A ce propos, il convient de distinguer entre deux

démarches principales. La première, privilégiée en commande, s'efforce à représenter le système par une représentation d'état d'un certain type (par exemple, de type rétroaction stricte ou rétroaction sur la sortie). La seconde démarche fait quant à elle appel aux modèles blocs obtenus en montant en série, en parallèle ou en boucle fermée des blocs dynamiques linéaires et des blocs statiques non linéaires. Les modèles blocs se prêtent généralement mieux à l'identification moyennant une adaptation des méthodes conçues à l'origine pour les systèmes linéaires. Parmi les modèles de ce type qui ont le plus suscité l'intérêt, on peut citer les suivants :

- le modèle de Hammerstein (Fig 1.a) composé d'un élément statique non linéaire suivi d'un bloc dynamique linéaire
- le modèle de Wiener (Fig 1.b) composé d'un bloc dynamique linéaire suivi par un bloc statique non linéaire
- le modèle de Wiener-Hammerstein (Fig 1.c) composé d'un élément statique non linéaire pris en sandwich entre deux blocs dynamiques linéaires
- le modèle de Hammerstein-Wiener (Fig 1.d) composé d'un bloc dynamique linéaire pris en sandwich entre deux blocs statiques non linéaires.
- les modèles se présentant sous la forme d'un feedback comprenant au moins un élément non linéaire ; la Fig 1.e en représente un exemple.

D'un point de vue identification, les configurations citées ci-dessus sont ordonnées dans le sens de complexité croissante. L'on peut donc comprendre pourquoi la plupart des travaux antérieurs ont porté sur les trois premières. Une liste plus fournie en variantes de modèles blocs, contenant des structures série-parallèle et/ou en boucle fermée peut être trouvée dans (Haber et Keviczky, 1999a). Le choix d'une structure adéquate, pour un système donné, est un problème crucial que peu de travaux antérieurs ont abordé. Dans des situations relativement simples, une analyse structurelle du système permet de mettre en évidence ses différents blocs, linéaires et non linéaires, et la manière dont ceux-ci sont interconnectés. Ainsi la présence d'un actionneur électro-mécanique, -hydraulique ou -pneumatique donne-t-elle lieu à une non-linéarité de type saturation placée à l'entrée du modèle. De même, que des phénomènes de type fin-de-course ou trop-plein se traduisent par une non-linéarité placée en sortie. En revanche, dans des situations plus complexes, il s'avère difficile, voire impossible, de justifier d'une configuration de modèle donnée moyennant une analyse structurelle. Dans ces cas, on adopte la démarche dite 'boite noire' et on choisit la configuration de modèle et les paramètres de structure qui réalisent le meilleur compromis précision/complexité (Haber et Unbehauen, 1990).

Les travaux présentés dans ce mémoire, s'inscrivent dans le cadre de l'identification des systèmes non linéaires sur la base des modèles blocs et portent précisément sur les modèles de type Hammerstein et de type Wiener. Il convient de mentionner que l'essentiel des travaux antérieurs s'inscrivant dans ce cadre furent consacrés précisément à ces deux types de modèles. Nous commençons par en faire un aperçu.

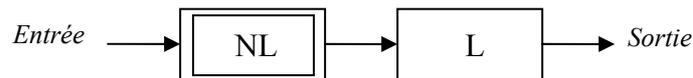


Fig 1.a Modèle de Hammerstein

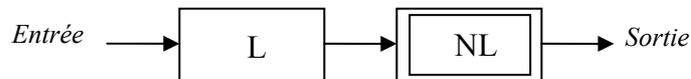


Fig 1.b Modèle de Wiener

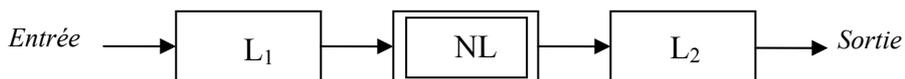


Fig 1.c Modèle de Wiener-Hammerstein

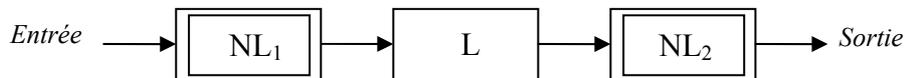


Fig 1.d Modèle de Hammerstein-Wiener

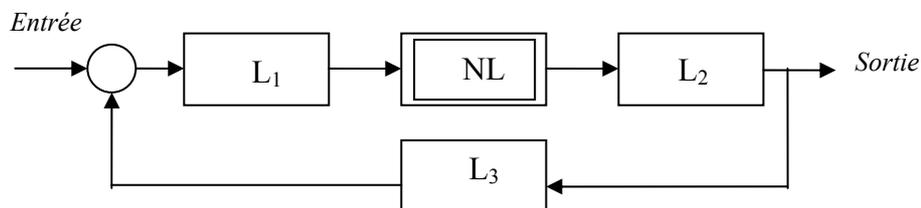


Fig 1.e Modèle de Hammerstein-Wiener bouclé

1.3.IDENTIFICATION DES MODELES DE HAMMERSTEIN ET WIENER : APERÇU SUR LES TRAVAUX ANTERIEURS

Le nombre important de publications dédiées à ce sujet s'explique par la grande diversité des solutions proposées, (voir (Haber et Unbehauen, 1990) et (Haber et Keviczky, 1999a-b) et les références bibliographiques incluses). Ces solutions peuvent être discutées et comparées à partir de différents points de vues, tels que :

- i) l'architecture même du schéma d'identification caractérisée par la nature (paramétrique ou non, fréquentielle ou non, temps réel ou non,) de ce schéma, le type de paramétrisation mis en jeu et des estimateurs impliqués
- ii) les hypothèses utilisées dans la conception de ce schéma, portant généralement sur la connaissance a priori sur le système et sur la nature des bruits
- iii) la complexité de mise en œuvre déterminée notamment par le type d'expériences requises pour la collecte des données et la nature des calculs mis en jeu
- iv) la question de la consistance des estimateurs utilisés et celle sous-jacente d'excitation persistante

Dans ce qui suit, nous tenterons d'embrasser un bon nombre de travaux antérieurs en comparant les schémas proposés par rapport aux critères décrits ci-dessus.

1.3.1. Architecture des schémas d'identification

Nous focalisons l'attention ici sur la nature paramétrique ou non des schémas proposés. En effet, cet aspect revêt une importance plus grande en identification des systèmes non linéaires en blocs dans la mesure où ces derniers sont constitués de plusieurs sous-systèmes ; le choix entre une représentation paramétrique ou non se pose pour chacun de ces sous-systèmes. Concernant le cas précis des systèmes de Hammerstein et de Wiener, on peut distinguer entre trois situations :

- i) système entièrement paramétrique : les deux sous-systèmes (linéaire et non linéaire) sont caractérisés par une représentation paramétrique
- ii) système partiellement paramétrique : un des deux sous-systèmes est non paramétrique
- iii) système entièrement non paramétrique : les deux sous-systèmes sont non paramétriques

Les schémas d'identification proposés se trouvent, à leur tour, classés en trois familles dites paramétriques, partiellement non paramétriques ou (entièrement) non paramétriques. Il arrive souvent que l'on ne fasse pas de distinction entre les deux dernières classes. La plupart des schémas d'identification proposés dans la littérature (pour les systèmes de Hammerstein et de Wiener) sont en fait de type paramétriques (exemples (Boutayeb et al, 1993), (Greblicki, 2000), (Greblicki, 2001), (Krzyzak, 1989), (Kung et Shinh, 1986), (Rangan et al, 1995), (Stoica et Söderström, 1982), (Van Pelt et al), (Zhu, 2002)). Ceci peut s'expliquer par le fait qu'un modèle paramétrique implique un petit nombre de paramètres inconnus, ce qui simplifie le problème d'identification, et se prête mieux à l'élaboration de régulateurs. Ces avantages ont un prix, c'est une connaissance a priori plus importante sur le système. Cette connaissance concerne aussi bien le sous-système dynamique (connaissance de son ordre par exemple) que l'élément non linéaire (polynomial ou non, inversible ou non, si élément polynomial connaissance de son ordre, ...). Moyennant cette connaissance a

priori, il devient possible d'élaborer pour le système des paramétrisations impliquant linéairement ou bilinéairement les paramètres inconnus (du système) pour l'estimation desquels il est fait recours aux estimateurs classiques (moindres carrés et ses variantes) ou moins classiques (méthodes d'optimisation évolutionnaire, ...).

1.3.2. Connaissances a priori dans les schémas d'identification

Le sous-système dynamique linéaire a généralement été supposé paramétrique et représenté par sa fonction de transfert ((Vörös, 2001), (Krzyzak, 1989), (Van Pelt et al), (Stoica et Söderström, 1982)) ou une représentation d'état ; le degré de la dynamique est alors supposé connu ((Kung et Shinh, 1986), (Rangan et al, 1995)). Toutefois, certains travaux ont suggéré des représentations non paramétriques de type réponse impulsionnelle infinie ((Krzyzak, 1989), (Greblicki, 1992), (Greblicki et Pawlak, 1991b)), réponse impulsionnelle finie (FIR) ((Rangan et al, 1995), (Mzyk, 2000), (Mzyk, 2004)) ou réponse fréquentielle ((Bai, 2003)).

Les principales hypothèses ont, en fait, concerné l'élément non linéaire. De très nombreux travaux supposent que cet élément est de type polynomial et son degré est connu ou du moins borné ((Krzyzak, 1989), (Stoica et Söderström, 1982), (Greblicki, 2001), (Greblicki, 2000), (Greblicki, 1992), (Boutayeb et al, 1993)). Il est de ce fait parfaitement caractérisé par un nombre fini de paramètres (inconnus). Dans le même ordre d'idée, certains travaux ont suggéré la représentation de l'élément non linéaire par son développement en série de fonctions orthogonales dans une famille complète choisie a priori ((Rangan et al, 1995), (Mzyk, 2000)). On peut citer à cet égard les fonctions blocs impulsionnelles suggérées dans (Ioannou P., Sun J., 1996), les splines cubiques utilisées dans (Zhu, 2002), les polynômes de Laguerre, de Legendre ou de Hermite proposées respectivement dans (Greblicki et Pawlak, 1991b), (Greblicki, 1994) et (Pawlak, 1991). Le développement en série a également été suggéré. Il va sans dire que le recours à ces développements sous-entend que l'élément non linéaire peut être assimilé à une fonction appartenant à l'espace de Hilbert engendré par la famille des fonctions orthogonales sélectionnée et ce du moins dans le contexte produit pour les besoins de l'identification du système. Cette approche reste néanmoins intéressante d'autant plus qu'elle s'applique à des éléments non linéaires que l'on peut seulement assimiler à une fonction mesurable (au sens de Lebesgue/Borel) ((Krzyzak, 1989), (Mzyk, 2004)). Il va sans dire que lors du processus d'identification les développements mentionnés ci-dessus feront l'objet de troncature. Dans ce cas, l'ordre de troncature est choisi a priori. La qualité du modèle sera d'autant meilleure que l'ordre est plus élevé. Pour des non-linéarités suffisamment lisses ou régulières (continûment dérivables au moins une fois) il est possible

d'atteindre une bonne précision à l'aide de développements limités à un ordre relativement petit, surtout si le choix des fonctions orthogonales est adéquat. En revanche, quand il s'agit d'éléments non lisses, les développements doivent être poussés assez loin pour espérer recouvrer avec une bonne précision les éléments considérés et ce indépendamment du type de fonctions orthogonales. Or, pousser trop loin un développement génère un nombre important de paramètres inconnus ce qui non seulement augmente la complexité de la solution mais, surtout, rend plus ardue la question de synthèse de signaux d'entrée suffisamment excitants.

Il vient de la discussion précédente que le cas des éléments non linéaires et non lisses mérite d'être abordé de façon spécifique. Ceci est d'autant plus vrai que ces éléments sont bien présents dans les systèmes physiques. Nous faisons référence en particulier aux systèmes mécaniques dont le fonctionnement est souvent caractérisé par des phénomènes de 'seuil', de 'zone-morte' ou de 'jeu'. Si les deux premiers aspects peuvent être représentés par des caractéristiques statiques portant les mêmes noms, il n'en est pas ainsi du dernier phénomène. Celui-ci est modélisable par des éléments à mémoire de type hystérésis. Ce phénomène est évidemment rencontré dans d'autres domaines ; l'on sait par exemple que les caractéristiques magnétiques des moteurs électriques et des bobines de puissance sont de type hystérésis. La spécificité des systèmes mécaniques réside dans le fait que les 'jeux mécaniques', responsables du caractère hystérétique, sont introduits exprès, sinon ces systèmes ne peuvent pas fonctionner.

Les travaux dédiés au problème d'identification des systèmes non linéaires comprenant des éléments non lisses sont relativement récents et rares. Le cas de non-linéarités sans-mémoire de type affine par morceaux (incluant en particulier les éléments de type seuils et zone-mortes) a été abordé dans (Vörös, 1999a), (Vörös, 1999b), (Bai, 2002), (Vörös, 2001), (Giri et al, 2001b), (Giri et al, 2004b). Les non-linéarités de type hystérésis ont été considérées dans (Bai, 2002), (Giri et al, 2004a). D'un point de vue identification, la difficulté de tous ces éléments réside, entre autres, dans le fait que leur description mathématique fait intervenir plusieurs expressions analytiques, chacune d'entre elles étant valable sur un intervalle spécifique de l'entrée, sachant que les extrémités de ces intervalles font partie des paramètres inconnus. Cette difficulté est en partie contournée moyennant des hypothèses ad hoc (symétrie de l'élément, connaissance de majorants des paramètres, etc). Cet aspect sera rediscuté au paragraphe suivant et dans la suite de ce mémoire.

La connaissance a priori concernant l'élément non linéaire peut porter sur d'autres caractéristiques de cet élément. Ainsi, a-t-il été supposé inversible dans (Greblicki, 2001), (Greblicki, 1992),

(Greblicki, 1994), continu et lisse dans (Bai, 2003), dérivable de dérivée bornée dans (Greblicki, 2001), (Greblicki, 1994), Lipschitzien dans (Greblicki, 1992), borné par un polynôme dans (Greblicki, 2001).

1.3.3. Nature des perturbations et des erreurs de modélisation dans les schémas d'identification

Les perturbations sont un fait réel que l'on peut difficilement ignorer dans le processus de modélisation ou d'identification des systèmes physiques. Elles représentent en fait l'influence non contrôlée de l'environnement du système sur celui-ci. Quand on essaie d'en tenir compte, on est confronté à deux types d'incertitudes : la nature (stochastique ou déterministe) de la source génératrice de ces perturbations et la manière dont celles-ci affectent le système (leurs points d'action). En identification des systèmes non linéaires, l'effet des perturbations est très souvent pris en compte dans le modèle à travers une erreur de sortie de type bruit blanc (avec des distributions plus ou moins clairement spécifiées) éventuellement filtré par un filtre stable. Toutefois, certains travaux ont suggéré une prise en compte des perturbations à travers une erreur d'équation ou d'entrée assimilée à un bruit blanc filtré ou non ((Greblicki, 1994), (Greblicki, 2001), (Kreisselmeier, 1989), (Greblicki et Pawlak, 1991b)). Les cas extrêmes sont ceux où aucune hypothèse n'est émise sur le processus générateur des erreurs ((Chaoui et al, 2005), (Giri et al, 2001a-c), (Giri et al, 2002a-b), (Giri et al, 2004a-b)) et le cas opposé où ces erreurs sont supposées nulles ((Vörös, 2001), (Stoica et Söderström, 1982)). Il va sans dire que les hypothèses considérées dans tel ou tel travail sont avant tout motivées par le souci d'analyse de consistance des estimateurs considérés. Il convient de remarquer aussi que les erreurs en question (de sortie, d'équation et d'entrée) sont aussi le résultat d'erreurs de modélisation inhérente à l'approche 'boite noire'.

1.3.4. Analyse de la consistance et techniques d'excitation dans les schémas d'identification

Toute méthode d'identification doit prouver son efficacité théorique par le biais d'une analyse de consistance des estimateurs mis en jeu. Par consistance on entend la convergence des estimés vers leurs vraies valeurs. Selon le contexte stochastique ou déterministe de l'identification, la convergence est à prendre au sens probabiliste ou déterministe. La consistance ou non d'un estimateur dépend, au moins en partie, du degré d'excitation des signaux d'entrée appliqués au système lors du processus d'identification. On convient de qualifier de 'persistante' une excitation qui suffit pour garantir la consistance d'un estimateur donné. Cette question a été bien élucidée dans le cas des systèmes linéaires, généralement abordés par des estimateurs de type moindres carrés, et on dispose maintenant de formulations précises de ce que signifie, dans ce contexte, l'exigence d'excitation persistante (voir les ouvrages (Ljung, 1997); (Soderstrom, et Stoica, 1988), (Goodwin

et Sin 1983), (Lozano et Taoutaou, 2001)). Spécifiquement, cette exigence stipule que le vecteur de régression sous-jacent doit effectuer un balayage périodique de l'espace. Pour s'y conformer il suffit que la puissance du signal d'entrée soit répartie sur un nombre de fréquences au moins égal à la dimension du vecteur de régression.

Il est évident que la question devient beaucoup plus complexe quand l'identification porte sur les systèmes non linéaires. On peut même penser, a priori, que le caractère non linéaire du système favorise la génération d'harmoniques supplémentaires rendant de ce fait plus excitants les signaux mis en jeu. Sauf que cette intuition n'a jamais été rigoureusement analysée, et aucun résultat de ce type n'a été formellement établi. Dans le cas des systèmes non linéaires blocs, il est cependant possible d'établir, pour une bonne partie des schémas d'identification proposés, une formulation précise d'une condition ad hoc d'excitation persistante. En revanche, il est rarement établi comment choisir le signal d'entrée en vue de se conformer avec cette exigence ((Zhu, 2002), (Rangan et al, 1995), (Van Pelt et al)). Parmi les rares travaux ayant abordé cette question, on peut citer (Stoica et Söderström, 1982). L'exigence d'excitation persistante peut être satisfaite moyennant un signal d'entrée de type bruit blanc (qui doit en outre être non corrélé avec les perturbations). L'intérêt pratique de ce résultat est limité par la difficulté de réalisation du signal en question. Dans le cadre de cette thèse, nous avons démontré que l'excitation persistante peut être assurée moyennant des signaux déterministes, ((Chaoui et al, 2005), (Giri et al, 2000), (Giri et al, 2001a-c), (Giri et al, 2002a-b), (Giri et al, 2004a-b)). Cet aspect sera repris au chapitre 3.

1.3.5. Complexité des schémas d'identification

Le degré de complexité de mise en œuvre des méthodes proposées est très variable. Il se situe notamment au niveau de la stratégie d'estimation et au niveau de la collecte des données. La synthèse d'estimateurs implique la résolution de problèmes d'optimisation plus ou moins complexes. Selon l'architecture du schéma d'identification, trois stratégies sont envisageables pour appréhender l'estimation des deux sous-systèmes (linéaire et non linéaire). La première, suivie au chapitre 3, est celle qui consiste estimer simultanément les paramètres des deux sous-systèmes ((Chaoui et al, 2005), (Giri et al, 2001b), (Giri et al, 2002a), (Giri et al, 2004b)). L'intérêt de cette stratégie réside, entre autres, dans le fait que la collecte des données nécessite une seule expérience. La seconde stratégie dite séquentielle est celle où on procède en deux temps : on commence par estimer un des deux sous-systèmes et on profite des estimés ainsi obtenus pour identifier, dans un second temps, l'autre sous-système ((Bai, 2003), (Bai, 2002), (Giri et al, 2004a-b)). Cette stratégie peut requérir deux expériences ou plus. Elle sera illustrée au chapitre 5. La troisième approche dite

itérative par alternance consiste à estimer de manière itérative et à tour de rôle les deux sous-systèmes : si le modèle du sous-système linéaire fait l'objet d'une mise à jour lors de l'étape n , alors c'est celui de l'élément non linéaire qui fera l'objet d'une mise à jour lors de l'étape $n+1$. Une telle approche nécessite en général une seule expérience ((Vörös, 1999b)).

1.4.CONTRIBUTIONS DE LA THESE

Cette thèse est centrée sur le problème d'identification des systèmes non linéaires sur la base du modèle de Hammerstein et du modèle de Wiener. Nous avons noté au paragraphe précédent que la majeure partie des travaux antérieurs a porté sur le modèle de Hammerstein et que la plupart des schémas d'identification proposés ont esquivé la question d'excitation persistante. Notre but est de concevoir de nouveaux schémas d'identification avec les caractéristiques suivantes:

- les signaux d'entrée utilisés lors du processus d'identification doivent être facilement réalisables et assez riches, pour apporter l'excitation persistante nécessaire à la consistance des estimateurs mis en jeu
- les hypothèses faites sur le système, notamment en termes d'informations a priori sur le modèle, doivent être réalistes afin de garantir un champ d'application le plus large possible pour les schémas développés. En outre, ceux-ci doivent être les plus simples possibles pour en faciliter la mise en oeuvre.

1.4.1. Excitation persistante à l'aide de signaux impulsionnels

Les contributions élaborées dans le cadre de cette thèse portent essentiellement sur l'identification du modèle de Hammerstein et celui de Wiener. Mais, étant donné que la question d'excitation persistante est centrale dans ces problèmes, nous lui avons accordée une attention particulière. Nous nous sommes plus particulièrement intéressés à la création des conditions d'excitation persistante moyennant des signaux déterministes de nature impulsionnelle. A l'origine, cette façon de parvenir à une excitation persistante a été conçue dans un contexte de commande adaptative des systèmes linéaires. Les boucles de commande mises en évidence dans ce contexte sont caractérisées par leur caractère non linéaire et par rapport à cette difficulté que les signaux impulsionnels déterministes se sont avérés efficaces.

Dans le cadre de cette thèse, nous avons réussi à formuler sous forme d'un lemme technique ((Giri et al, 2002b), (Giri et al, 2001c)) comment l'exigence d'excitation persistante est-elle satisfaite, moyennant ce type de signaux, que ce soit en identification ou en commande adaptative des

systèmes linéaires. Un aperçu sur ce résultat sera fait au Chapitre 2. En outre, nous verrons au Chapitre 3 que ce même lemme technique est profitable à l'identification des systèmes non linéaires blocs de type Hammerstein.

1.4.2. Contributions à l'identification du modèle de Hammerstein

Dans le cadre de thèse, nous avons apporté trois contributions à l'identification des systèmes blocs de type Hammerstein. La première, faisant l'objet du Chapitre 3, a consisté à développer un schéma d'identification qui, contrairement à la plupart des travaux antérieurs, ne fait aucune hypothèse sur le sous-système non linéaire à l'exception du fait que celui-ci soit statique et l_∞ -stable (stable au sens 'entrée bornée sortie bornée'). Spécifiquement, cet élément peut être caractérisé par une fonction quelconque $F(\cdot)$ qui vérifie seulement : $|F(v)| < \infty, \forall |v| < \infty$. En particulier, il n'est pas du tout exigé que $F(\cdot)$ soit lisse et encore moins qu'elle soit de type polynomial. La dynamique linéaire est quant à elle représentée par sa fonction de transfert. S'agissant des perturbations, on sait seulement qu'elles sont bornées, plaçant ainsi le problème dans un contexte déterministe. Le schéma d'identification développé permet d'estimer, simultanément, les paramètres de la dynamique linéaire et un certain nombre (disons N) de points de la caractéristique non linéaire $F(\cdot)$. Le choix du nombre de points et de leur répartition est totalement laissé à l'utilisateur. Il s'ensuit que notre approche est de nature partiellement non paramétrique et nécessite une seule expérience pour collecter les données. Pour l'estimation des paramètres inconnus, il est fait usage d'un algorithme de type moindres carrés et d'une décomposition en valeurs singulières. Pour que l'estimation se fasse dans un contexte d'excitation favorable, le système est excité par un signal périodique impulsionnel. Il convient de noter que les amplitudes et les signes des impulsions peuvent être choisis arbitrairement, ce dont on peut profiter pour faire en sorte que le signal soit, par exemple, de moyenne nulle. En invoquant le lemme technique mentionné au paragraphe précédent, on démontre que le niveau d'excitation générée par le signal d'entrée impulsionnel est proportionnel à la taille moyenne de l'erreur de modélisation (celle-ci étant due, entre autres, aux perturbations). Il s'ensuit que la qualité du modèle estimé est d'autant meilleure que l'erreur est plus petite en moyenne. Dans le cas idéal où cette erreur est inexistante, l'excitation apportée par le signal d'entrée devient persistant ce qui assure la consistance du modèle estimé. Nous avons déjà observé (au paragraphe 1.3.4) que les travaux dédiés à l'identification des systèmes non linéaires blocs et qui ont réellement établi la capacité de leurs schémas à produire une excitation persistante sont à ce jour très rares. En effet, dans la grande majorité des travaux il est simplement 'espéré' que l'excitation procurée par le signal d'entrée soit persistante.

Notre seconde contribution, développée au Chapitre 4, a consisté à concevoir un schéma d'identification adapté aux systèmes de Hammerstein dont l'élément non linéaire a une structure connue a priori. Spécifiquement, le schéma intéresse les systèmes dont on sait que l'élément non linéaire est un polynôme de degré connu et ceux dont on sait que l'élément non linéaire est un seuil avec zone-morte. Si le premier type d'élément a largement été pris en compte dans les travaux antérieurs, il n'en est pas ainsi du second type. Cette observation suffit pour se rendre compte du fait que le problème d'identification est plus ardu quand il s'agit de systèmes impliquant des non-linéarités non lisses. C'est en particulier le cas des éléments de type seuil, avec ou sans zone-morte. La difficulté se situe, entre autres, dans le fait que la description analytique de ces éléments (et d'une façon générale de tous ceux de type affine par morceaux) nécessite plusieurs expressions mathématiques, chacune d'elles étant valable sur un intervalle donné et les bornes de ces intervalles ne sont pas connues a priori. Le schéma d'identification proposé s'appuie, en partie, sur le schéma développé au Chapitre 3. Ce dernier est intéressant dans la mesure où il ne fait aucune hypothèse sur la nature de l'élément non linéaire. Son utilisation ici permettra d'identifier le sous-système dynamique linéaire et N points de l'élément non linéaire. A partir du nuage de points ainsi obtenu il devient possible de déterminer les paramètres caractérisant l'élément non linéaire. Dans le cas d'un élément polynomial il s'agit de déterminer ses coefficients ce qui ne pose aucun problème de principe. Dans le cas d'un élément de type seuil et zone-morte, il s'agit de déterminer les extrémités du seuil et celles de la zone-morte, ce que le schéma proposé réussira à faire effectivement, mais à un prix : il faut connaître un majorant des paramètres inconnus.

Nous avons déjà mentionné (paragraphe 1.3.2) que (Bai, 2002) est parmi les rares auteurs à avoir abordé dans (Bai, 2002) le problème d'identification des systèmes de Hammerstein en présence d'éléments non linéaires affines par morceaux. Dans le cas où l'élément est parfaitement symétrique, de telle sorte qu'il puisse être caractérisé par un seul paramètre inconnu positif (noté a), l'auteur propose un schéma d'identification impliquant une recherche graphique du minimum d'une fonction monovariante dont le minimum se produit précisément en a . Pour qu'une telle recherche puisse aboutir au minimum global, il faut qu'elle s'effectue dans un intervalle contenant la vraie valeur de a ce qui, de fait, revient à supposer connu un majorant de ce paramètre. L'intérêt de notre solution est qu'elle n'exige pas nécessairement une symétrie parfaite de l'élément non linéaire. Dans (Vörös, 1999a), on a proposé un schéma d'identification itérative en alternance mais aucune analyse de consistance n'a été faite. Cet algorithme peut en fait diverger dans certaines situations.

Notre troisième contribution, développé au Chapitre 5, a consisté à concevoir un schéma d'identification approprié pour les systèmes non linéaires blocs de type Hammerstein dont l'élément non linéaire est à mémoire. Spécifiquement, le schéma porte sur les systèmes dont on sait que l'élément non linéaire est de type hystérésis ou hystérésis-relais. Il va sans dire, que l'effet 'hystérésis' conjugué avec le caractère 'non-lisse' de ces non-linéarités rend le problème d'identification plus complexe que dans les cas discutés ci-dessus. Le schéma d'identification proposé est de type séquentiel et de ce fait opère en deux temps. Dans un premier temps, il identifie le sous-système dynamique linéaire. A cet effet, un signal d'entrée périodique de type impulsionnel est conçu de telle sorte que :

- i) le signal interne commandant le sous-système linéaire devienne 'mesurable'
- ii) l'excitation produite s'avère excitante

La première propriété permet d'envisager l'identification du sous-système linéaire en utilisant à cet effet un estimateur de type moindres carrés. La seconde propriété garantit que le modèle estimé est consistant en présence de perturbations dont l'effet sur la sortie peut être assimilé à un bruit blanc ergodique. Dans une deuxième phase, le schéma d'identification œuvre pour l'identification de l'élément non linéaire en exploitant le modèle estimé du sous-système linéaire. Cette phase nécessite d'effectuer deux autres expériences pendant lesquelles le système est soumis à deux signaux périodiques de forme triangulaire. Les données ainsi collectées sont judicieusement exploitées en vue d'obtenir des estimations consistantes des paramètres de l'élément non linéaire. A notre connaissance, c'est la première fois que de tels résultats sont atteints en présence d'une non-linéarité de type hystérésis asymétrique.

1.4.3. Contributions à l'identification du modèle de Wiener

Dans le cadre de cette thèse, nous nous sommes également intéressés au problème d'identification des systèmes non linéaires de type Wiener. Le problème peut très bien être abordé par un schéma d'identification opérant de façon simultanée (comparable à ce qui est fait au Chapitre 3 pour les systèmes de Hammerstein) l'estimation des deux sous-systèmes. Nous avons abandonné cette voie car il n'a pas été possible de trouver un signal d'entrée assurant une excitation persistante. Nous nous sommes donc tournés vers une démarche séquentielle consistant à estimer d'abord un des deux sous-systèmes. Ainsi, avons-nous élaboré deux solutions au problème d'identification de l'élément non linéaire, en supposant que le sous-système linéaire est encore indéterminé. Il est évident que si l'élément non linéaire était identifié, il deviendrait plus facile de déterminer le modèle de la partie

linéaire, notamment si l'élément non linéaire est inversible. Les solutions proposées permettent d'estimer un nombre de points, aussi grand que l'on veut, de la caractéristique graphique de cet élément. Elles sont de nature fréquentielle et n'exigent aucune connaissance a priori sur l'élément non linéaire à part le fait qu'il est statique. En particulier, ce dernier peut être inversible ou non, lisse ou non.

La première solution, développée au Chapitre 6, repose sur deux idées essentielles. La première utilise le résultat fondamental qui stipule que si l'on excite le système par un signal sinusoïdal, alors le signal interne du modèle (généré par la partie linéaire) est également sinusoïdal (en régime permanent) mais son amplitude et sa phase sont inconnues. Or, il est bien connu que le modèle recherché n'est pas unique; il est en fait défini, à un facteur constant près réparti, entre la dynamique linéaire et la caractéristique non linéaire. Cette observation nous autorise à supposer, sans porter atteinte à la généralité du problème, que l'amplitude du signal interne est égale à 1. Ceci revient en fait à focaliser la recherche sur un modèle particulier du système (par les innombrables modèles de celui-ci), Par conséquent, la seule incertitude qui reste encore sur le signal interne concerne la phase. La seconde idée consiste à exploiter le fait que le signal de sortie est périodique de même période que le signal interne et est en phase avec ce dernier. L'analyse des extremums de ces deux signaux montre qu'à chaque instant où le signal interne (que l'on sait être sinusoïdal et en phase avec la sortie) passe par un extremum, il produit un extremum au niveau de la sortie. En outre, cette catégorie d'extremums de la sortie se produit aux mêmes instants même quand on modifie l'amplitude du signal interne. En revanche, les éventuels extremums de la sortie générée par la non-linéarité elle-même changent eux d'instant quand le signal interne change d'amplitude. Le schéma d'identification proposé exploite judicieusement cette observation et permet de ce fait d'estimer un nuage de point de l'élément non linéaire. Cette estimation est consistante en présence de perturbations dont l'effet est assimilable à un bruit blanc ergodique affectant la sortie du système.

La seconde solution, développée au Chapitre 7, est conçue en exploitant la théorie des courbes de Lissajous. Comme dans la solution précédente, le point de départ est le fait que, si l'on soumet le système à une entrée sinusoïdale, il en résulte (en régime permanent) un signal interne lui aussi sinusoïdal et en plus son amplitude peut être supposée égale à 1, sans que cela ne remette en cause la généralité du problème. Il en résultera aussi, au niveau de la sortie du système, un signal périodique de même période que le signal interne et en phase avec celui-ci. Alors, si on trace la sortie du système en fonction d'un signal sinusoïdal d'amplitude 1 et de phase quelconque (notée φ), on obtiendra forcément une courbe de Lissajous constituée d'une ou plusieurs boucles. Nous

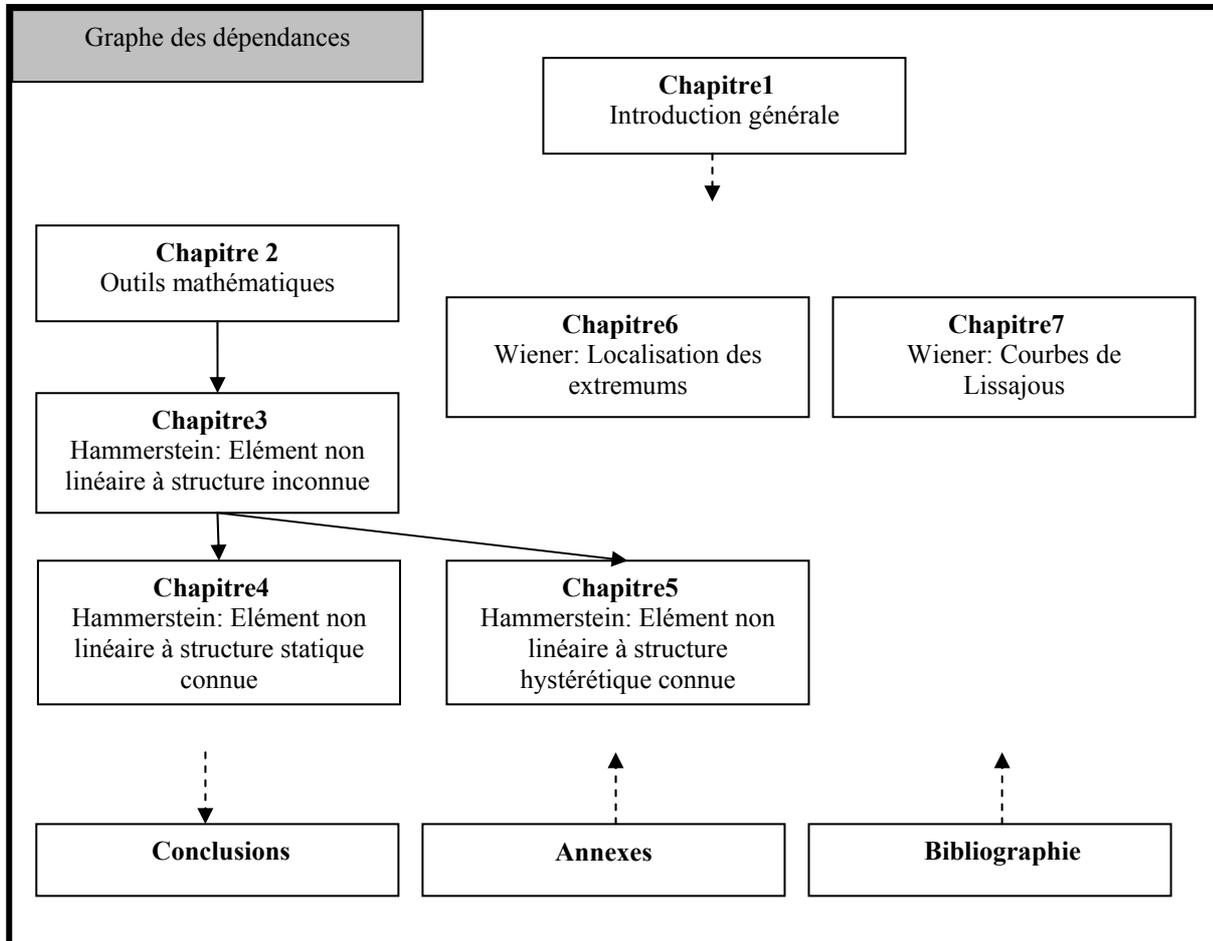
avons réussi à démontrer que si le paramètre φ coïncide (à π près) avec la vraie phase du signal interne, alors la courbe de Lissajous obtenue aura l'allure d'une ligne courbée (aucune boucle n'est visible) et, en outre, l'allure de cette ligne représente la caractéristique de l'élément non linéaire. Ces observations constituent le fondement théorique du schéma d'identification proposé. On démontre que l'estimation d'un nuage de points de la non-linéarité par ce schéma est consistante en présence d'un bruit blanc ergodique affectant directement la sortie du système. Le schéma nécessite une seule expérience consistant à soumettre le système à un signal d'entrée sinusoïdal.

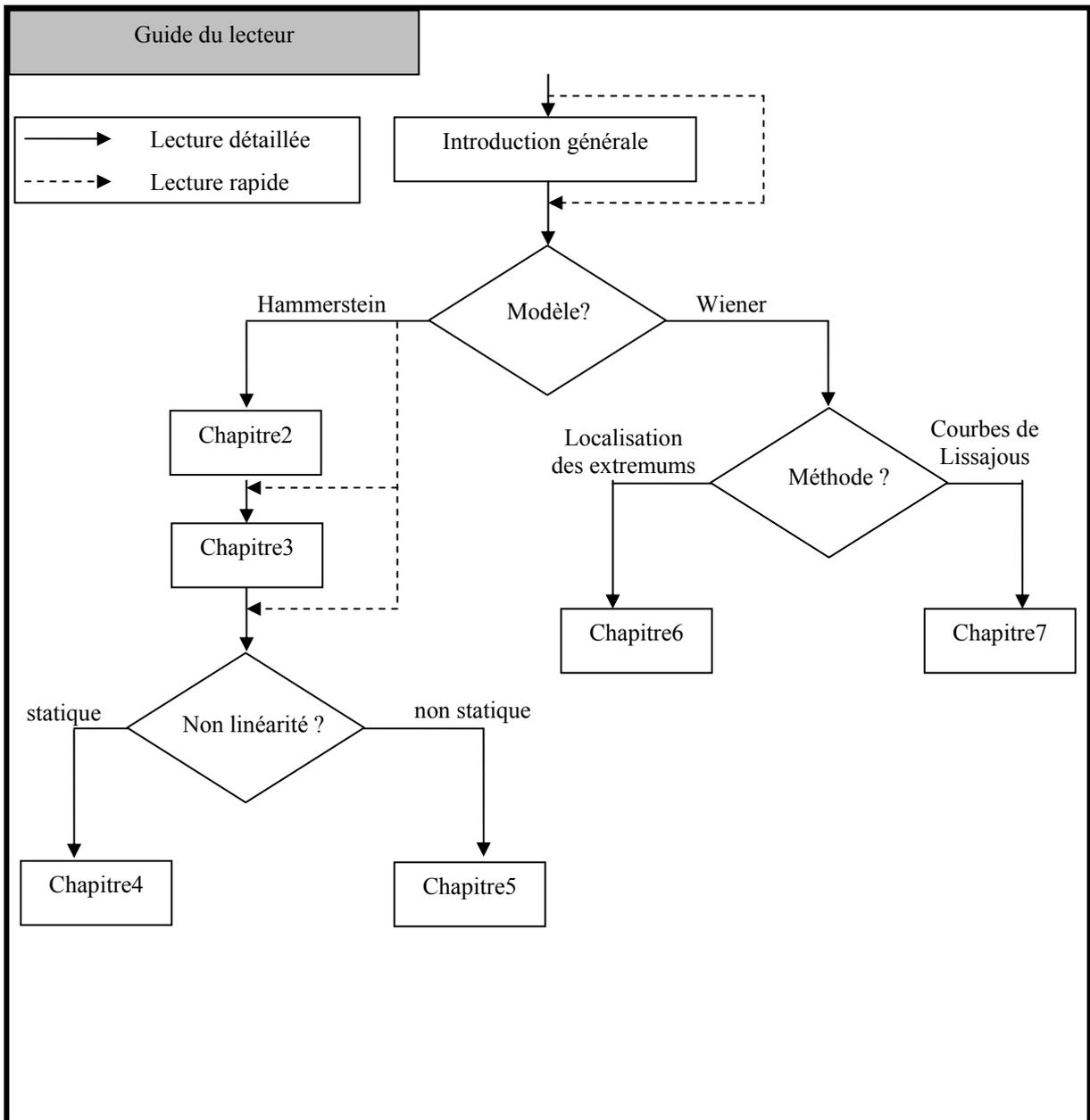
Les solutions présentées précédemment sont plus performantes que des solutions de même type (fréquentielles) présentées antérieurement ((Mzyk, 2004), (Bai, 2003)). En effet, dans (Mzyk, 2004), la dynamique linéaire a été représentée par une réponse impulsionnelle finie (FIR) ce qui pose le problème de choix de l'ordre de troncature. Dans (Bai, 2003), l'auteur propose une méthode fréquentielle, impliquant une décomposition en série de Fourier discrète. L'algorithme proposé pour estimer la phase du signal interne présente un sérieux problème de consistance, ce qui rend la méthode invalide dans la plupart des cas; en plus, la non linéarité est soumise à des conditions de lissage.

1.5.ORGANISATION DU MEMOIRE

1.5.1. Structure du mémoire

Le mémoire est structuré en trois parties : la première comprend l'introduction générale et le chapitre 2 qui rassemble des outils mathématiques utiles. La seconde, constituée des chapitres 3, 4 et 5, rassemble les contributions concernant l'identification des systèmes non linéaires sur la base du modèle de Hammerstein. La troisième partie, constituée des chapitres 6 et 7, regroupe les contributions intéressant l'identification des systèmes non linéaires sur la base du modèle de Wiener. L'organigramme ci-dessous donne une idée sur la dépendance qui existe entre les différents éléments et aide le lecteur à appréhender la lecture du mémoire. Le mémoire se termine par une conclusion générale, résumant les principaux résultats de ce travail et traçant de nouvelles perspectives pour la poursuite de cette activité de recherche. Pour alléger le corps du mémoire la plupart des démonstrations ont été regroupées en annexe. La liste bibliographique est placée en fin du mémoire. Les résultats présentés dans ce rapport ont fait l'objet de publications, dont la liste est placée au début de ce mémoire.





1.5.2. Aperçu sur le contenu des chapitres

Dans le chapitre 2, nous avons regroupé les outils mathématiques qui seront utilisés par la suite, notamment dans les chapitres 3, 4 et 5. Certains de ces outils sont bien connus (algorithmes d'estimation, de décomposition en valeurs singulières) et sont donc présentés juste par commodité. En revanche, d'autres aspects le sont beaucoup moins, c'est le cas de l'excitation persistante par les signaux impulsionnels et de la petitesse en moyenne.

Les chapitres 3, 4 et 5 portent tous sur l'identification des modèles de Hammerstein. La lecture du chapitre 3 est indispensable si l'on veut vraiment cerner le contenu des deux autres. Le schéma

d'identification du chapitre 3 n'exige aucune connaissance a priori sur la non-linéarité. En revanche, les chapitres 4 et 5 portent sur des systèmes impliquant des non-linéarités dont on connaît la structure. Les non-linéarités statiques, de type polynomial ou affines par morceaux, font l'objet du chapitre 4. Ceux de type hystérésis sont abordés au chapitre 6.

Les chapitres 6 et 7 portent sur l'identification des modèles de Wiener. L'accent est en fait mis sur l'identification de l'élément non linéaire. Ces chapitres présentent deux solutions fréquentielles élaborées à l'aide d'outils analytiques, au chapitre 6, et d'outils géométriques, au chapitre 7.

CHAPITRE 2

OUTILS THEORIQUES POUR L'IDENTIFICATION

2.1. INTRODUCTION

Dans ce chapitre, nous allons présenter quelques outils qui vont se révéler, dans les prochains chapitres, très utiles aussi bien dans la formulation des problèmes d'identifications abordés que dans l'élaboration des solutions. Au premier paragraphe, nous introduirons un lemme technique qui permettra de faire face à l'exigence de l'excitation persistante dans certains problèmes d'identification impliquant des systèmes linéaires. Au second paragraphe, nous présenterons un recueil d'algorithmes d'estimation auxquels nous ferons couramment appel. Le troisième paragraphe sera consacré à une notion de petitesse en moyenne dont on rappellera les principales propriétés. Dans le dernier paragraphe, nous aborderons un problème algébrique qui consiste à résoudre un système d'équations bilinéaires d'un certain type, en faisant appel à la technique de décomposition en valeurs singulières.

2.2. EXCITATION PERSISTANTE (EP) A L'AIDE D'UN TYPE DE SEQUENCES IMPULSIONNELLES

Dans ce paragraphe, on rappelle la définition de la notion d'excitation persistante, on montre l'intérêt de garantir l'excitation persistante dans un contexte d'identification paramétrique, puis on présente un lemme technique qui définit une classe de séquences d'entrée particulières possédant cette propriété d'EP.

2.2.1. Intérêt de l'excitation persistante

On se place dans un contexte d'identification paramétrique et on suppose, pour simplifier, que le système à identifier puisse être décrit par une régression linéaire de la forme :

$$y(t) = \varphi(t)^T \theta^* + v(t) \quad (2.1)$$

où $y(t)$ représente la sortie du système, $\varphi(t)$ est le vecteur de régression faisant intervenir les valeurs précédentes de l'entrée et la sortie, $v(t)$ est une erreur de modélisation qui rend compte de l'effet de perturbations externes et de dynamiques négligées, θ^* désigne le vecteur des paramètres à estimer. Dans le cas d'un système linéaire, $\varphi(t)$ dépend linéairement des signaux d'entrée/sortie.

Typiquement,

$$\varphi(t) = [-y(t-1) \quad \dots \quad -y(t-n) \quad u(t-1) \quad \dots \quad u(t-n)]^T \quad (2.2)$$

Soit $\hat{\theta}(t)$ une estimation de θ^* obtenue à par des mesures disponibles à l'instant t à l'aide de l'estimateur des moindres carrés. On a alors le résultat suivant :

Lemme 2.1. Si $\{\varphi(t)\}$ possède la propriété d'excitation, alors l'estimateur des moindres carrés est consistant dans les cas suivants :

- a) $\{v(t)\} = \{0\}$: dans ce cas $\{\hat{\theta}(t)\}$ converge vers θ^* , au sens déterministe.
- b) $\{v(t)\}$ un bruit blanc : $\{\hat{\theta}(t)\}$ converge vers θ^* avec probabilité 1 \square

La condition d'excitation persistante est de ce fait très importante et représente un aspect capital dans tout schéma d'identification paramétrique. En son absence et même dans le cas idéal, la convergence des estimés des paramètres vers leurs vraies valeurs n'est pas garantie. En effet, la plupart des algorithmes d'identification paramétrique, cherchent une solution qui minimise un certain critère. Sans une excitation persistante du vecteur de régression, ces algorithmes ne garantissent qu'une convergence locale. Les estimés obtenus, permettent de minimiser le critère considéré localement et ne correspondent pas forcément aux vrais paramètres recherchés.

2.2.2. Définitions de l'excitation persistante

Définition 2.1: Soit $\{\varphi(t)\}$ une suite vectorielle dans R^p ($p \in N$). $\{\varphi(t)\}$ possède la propriété d'excitation persistante (EP), si : $\exists \tau > 0, \exists \delta > 0, \exists T > 0, \forall W \in R^p - \{0\}, \forall t > T :$

$$\max_{0 \leq i \leq \tau} |w^T \varphi(t+i)| \geq \delta \quad (\text{avec } w = \frac{W}{\|W\|}) \quad (2.3) \quad \square$$

On peut donner à cette définition la formulation suivante qui peut parfois s'avérer plus commode.

Définition 2.2 : $\{\varphi(t)\}$ possède la propriété EP si : $\exists \delta > 0, \exists T > 0, \forall t > T :$

$$\sum_{i=0}^M \varphi(t+i)\varphi(t+i)^T \geq \delta.I \quad (I \text{ matrice identité de dimension } p) \quad (2.4) \quad \square$$

Les deux définitions étant équivalentes, on utilisera indifféremment l'une ou l'autre, dans la suite de ce rapport. La première montre que la propriété d'EP signifie la capacité du vecteur $\varphi(t)$ à balayer périodiquement l'espace avec une période égale à τ . On parle parfois de balayage persistant de l'espace.

2.2.3. Génération de l'excitation persistante

Dans un contexte d'identification de systèmes, ce sont les entrées du système qui déterminent la variation de la séquence de régression $\{\varphi(t)\}$. Pour conférer la propriété d'EP à cette séquence, il faut que les entrées soient suffisamment riches en fréquences. Cette exigence est parfaitement maîtrisée dans le cas où l'identification se fait en boucle ouverte et porte sur un système linéaire, (Ljung, 1997)- (Goodwin et Sin, 1983). En effet, il suffit alors que le spectre d'énergie de l'entrée soit réparti sur au moins p fréquences différentes (p représente l'ordre du système) ; cette propriété peut facilement être réalisée moyennant une somme de $p/2$ sinusoides (ou une séquence pseudo aléatoire). En revanche, la situation se complique très vite dès que l'identification porte sur des systèmes non linéaires qui font l'objet de cette thèse. L'analyse fréquentielle de tels systèmes est généralement très complexe et de, toute façon, se fait au cas par cas. Cette difficulté nous a incité à envisager d'autres types de signaux afin de procurer la propriété d'EP à la séquence de régression. Ainsi, nous sommes-nous tournés vers les séquences de type impulsionnel qui furent utilisés au cours des années 90, par (Kreisselmeier et Smith, 1986), (Kreisselmeier, 1987), (Giri et al, 1988a-c), (Giri et al, 1989a), (Giri et al, 1990a-c), (Giri et al, 1993), pour doter les schémas de commande adaptative de capacité d'auto-excitation. Dans le cadre de cette thèse, ce type de signaux s'est avéré très commode pour créer de l'excitation persistante notamment quand il s'agit d'identifier des systèmes de type Hammerstein. Des résultats intéressants à cet égard vont être présentés aux chapitres 3, 4 et 5. Pour en simplifier l'exposé nous donnons au paragraphe suivant un lemme technique décrivant la nature des signaux en question et l'excitation qu'ils procurent à un système linéaire.

2.2.4. Lemme technique d'excitation

On s'intéresse ici à l'identification des paramètres de systèmes linéaires dont le comportement d'entrée-sortie peut être décrit par un modèle de la forme suivante :

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}u(t) + \eta_p(t) \quad (2.5a)$$

avec

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_nq^{-n} \quad \text{et} \quad B(q^{-1}) = b_1q^{-1} + \dots + b_nq^{-n} \quad (2.5b)$$

où $u(t)$ et $y(t)$ dénotent l'entrée et la sortie du système, respectivement; $\eta_p(t)$ est l'erreur de sortie qui rend compte de l'effet de perturbations externes et de dynamiques négligées; les a_i et b_i sont des paramètres inconnus réels. En outre, on suppose que :

H2.1: L'entier n est connu

H2.2: $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux

Remarques 2.1.

1. L'entier n n'est pas exactement égal à l'ordre du système, parce que a_n, b_n, a_{n-1}, \dots peuvent être nuls.
2. L'hypothèse H2.2 assure que la dynamique du système est commandable. Si le système est seulement stabilisable, les pôles non commandables seront simplifiés, donnant lieu à une erreur, prise en compte par $\eta_p(t)$, qui disparaît exponentiellement dans le temps.
3. La stabilité du système (2.5a) n'est pas requise.

Pour élargir le champ d'application du résultat développé dans ce paragraphe, nous supposons que l'entrée du système (2.5a) est générée par une loi de la forme :

$$R(t, q^{-1})u(t) + S(t, q^{-1})y(t) = f(t) + \eta_c(t) \quad (2.6a)$$

$R(t, q^{-1})$ et $S(t, q^{-1})$ désignent des polynômes variant dans le temps, de degré $n-1$, qui vérifient pour tout k et tout $t \in [4nk, 4n(k+1)[$:

$$R(t, q^{-1}) = R(4nk, q^{-1}) \quad \text{et} \quad S(t, q^{-1}) = S(4nk, q^{-1}) \quad (2.6b)$$

ce qui autorise les expressions suivantes:

$$\begin{aligned} R(4nk, q^{-1}) &= 1 + r_1(k)q^{-1} + \dots + r_{n-1}(k)q^{-n+1} \\ S(4nk, q^{-1}) &= s_0(k) + s_1(k)q^{-1} + \dots + s_{n-1}(k)q^{-n+1} \end{aligned} \quad (2.6c)$$

$\{f(t)\}$ est une séquence (impulsionnelle) définie ci-après et dont le rôle est de produire l'excitation recherchée ; $\{\eta_c(t)\}$ rend compte de l'effet de gel des paramètres (2.6b) et de la présence d'une éventuelle séquence de référence.

L'expression (2.6a) recouvre en particulier les situations suivantes :

- i) identification en boucle ouverte, dans ce cas $R(t, q^{-1}) = 1$, $S(t, q^{-1}) = 0$, $\eta_c(t) = 0$; $\forall t$
- ii) commande à l'aide d'un régulateur fixe, dans ce cas : $R(t, q^{-1}) = R(0, q^{-1})$ et $S(t, q^{-1}) = S(0, q^{-1})$; $\forall t$
- iii) commande adaptative, dans ce cas : $R(4nk, q^{-1})$ et $S(4nk, q^{-1})$ sont des fonctions de k ;

Les notations suivantes vont s'avérer utile dans la suite de l'exposé:

$$\varphi(t) = [-y(t-1) \cdots -y(t-n) \quad u(t-1) \cdots u(t-n)]^T \quad (2.7a)$$

$$\eta(k) = [\eta_p(4nk-n) \quad \dots \quad \eta_p(4nk+3n) \quad \eta_c(4nk+1) \quad \dots \quad \eta_c(4nk+2n+1)]^T \quad (2.7b)$$

$$\theta(k) = [r_1(k) \quad \dots \quad r_{n-1}(k) \quad s_0(k) \quad \dots \quad s_{n-1}(k)] \quad (2.7c)$$

Lemme 2.2. (Giri et al, 2002b) Considérons le système décrit par (2.5a-b) et soumis aux hypothèses H2.1-H2.2. Supposons que l'entrée du système soit engendrée par une loi de la forme (2.6a) avec :

$$f(t) = \begin{cases} g(k) & \text{si } t = 4nk + 2n; \quad k = 0, 1, 2, \dots \\ 0 & \text{sinon} \end{cases} \quad (2.8)$$

Alors, il existe une constante réelle α et une fonction continue $g(\cdot)$ telles que, pour tout k :

$$\sum_{i=0}^{4n-1} \varphi(4nk+i) \varphi(4nk+i)^T \geq \alpha (g(k))^2 I_{2n} - \Delta(\eta(k), \theta(k)) \quad (2.9a)$$

où la fonction $\Delta(\eta, \theta)$ est continue en fonction de η (donc de η_p et η_c) et vérifie:

$$\Delta(0, \theta) = 0, \quad \text{pour tout } \theta \quad (2.9b) \quad \square$$

Pour une raison de commodité, la preuve est rappelée en Annexe A.1.

Remarques 2.2.

1. La séquence $f(t)$ est une séquence impulsions telle que l'intervalle séparant deux impulsions successives ne peut pas être plus petit que $4n-1$. La $k^{\text{ème}}$ impulsion est appliquée au centre de l'intervalle $[4kn, 4kn+4n-1]$ et elle est d'amplitude $g(k)$. Selon (2.6a), elle produit une excitation

sur le même intervalle. De ce fait, la propriété (2.9a) est parfois appelée 'excitation par intervalle'.

2. Le gel des paramètres (2.6c) est souvent pratiqué en commande adaptative notamment quand l'EP est un objectif recherché, ((Goodwin et Sin,1983), (Anderson, 1985b), (Helliott et al, 1985), (Cristi,)). Dans ce cas, le gel empêche que la variation des paramètres ne vienne, éventuellement, neutraliser, sur l'intervalle $[4kn, 4kn+4n-1]$, l'effet excitant de la séquence $\{f(t)\}$.
3. Dans le cas particulier $\eta(t)=0$, on obtient à partir de (2.9a-b) pour tout k:

$$\sum_{i=0}^{4n-1} \varphi(4nk+i)\varphi(4nk+i)^T \geq \alpha(g(k))^2 I_{2n}$$

Il est alors clair que, si $g(k) \neq 0$, la $k^{\text{ième}}$ impulsion procure de l'excitation sur l'intervalle correspondant $[4nk, 4nk+4n-1]$. Ainsi, si $|g(k)|$ est strictement supérieure à zéro de façon uniforme (c'est-à-dire $\min_k |g(k)| \geq \varepsilon > 0$), alors l'excitation devient persistante.

2.2.5. Applications du Lemme d'excitation

2.2.5.1. Identification en boucle ouverte

Dans ce cas, on a :

$$R(t, q^{-1})=I, S(t, q^{-1})=0, \eta_c(t)=0$$

Il s'ensuit que $u(t)=f(t)$ et, par conséquent, une condition suffisante pour que $\{\varphi(t)\}$ possède la propriété EP est:

$$\min_k |g(k)| \geq \varepsilon > 0 \text{ et } \sup_t |\eta(t)| \text{ suffisamment petit} \quad (2.10)$$

En effet, comme $\Delta(\eta, \theta)$ est continue par rapport à η , il découle de (2.9b) que si $\sup_t |\eta(t)|$ est suffisamment petit alors $\sup_k \|\Delta(\eta(k), \theta(k))\|$ l'est aussi. Ce résultat sera appliqué au chapitre 3, dans un contexte d'identification de systèmes de type Hammerstein.

Dans le cas particulier où $\eta_p(t)=0$, on a $\Delta(\eta(k), \theta(k))=0$. L'EP est alors garantie pourvu que: $|g(k)| \geq \varepsilon > 0, \forall k$. On peut prendre par exemple:

$$g(k) = g_0 \quad \forall k; \quad \text{avec } g_0 > 0.$$

Si l'on souhaite que la séquence $f(t)$ soit de moyenne nulle, il faut prendre plutôt:

$$g(2k) = g_0 \text{ et } g(2k + 1) = -g_0, \forall k; \text{ avec } g_0 > 0 .$$

Ces remarques montrent que la qualité de l'excitation de la séquence construite dépend de la taille du terme $\eta_p(t)$. Il stipule que si cette taille est suffisamment petite, l'EP peut être facilement obtenue par une séquence d'entrée très simple à réaliser.

2.2.5.2. Commande adaptative

Dans ce cas, la paire $(R(t, q^{-1}), S(t, q^{-1}))$ est estimée à partir des entrées/sorties du système soit directement (schémas directs) soit indirectement en passant par l'estimation de la paire $(A(q^{-1}), B(q^{-1}))$ (schémas indirects), voir par exemple (Goodwin et Sin, 1983), (Ioannou P., Sun J., 1996), (Narendra et Annaswamy 1989), (Sastry et Bodson, 1989.). Bien que (2.10) constitue toujours une condition suffisante pour procurer à $\{\varphi(t)\}$ la propriété d'EP, elle n'a qu'un faible intérêt dans un contexte de commande adaptative. En effet, dans ce cas, la séquence $\{\eta(t)\}$ est la conséquence, entre autres, des dynamiques négligées du système et de l'effet du gel paramétrique ce qui en fait une fonction (généralement inconnue) des signaux d'entrée/sortie. De ce fait, elle n'est pas a priori bornée, en tout cas pas avant que la stabilité du système de commande ne soit établie. Or, la stabilité est elle-même tributaire de l'excitation. Pour contourner cette difficulté, on opte pour des signaux d'excitation internes (dépendant des signaux d'entrée/sortie). Typiquement, on prend :

$$g(k) = \lambda.m(k) \quad (\lambda > 0) \quad (2.11a)$$

$$m(k) = \sigma m(k-1) + \|\varphi(k-1)\| + m_0 \quad (\text{avec } 0 < \sigma < 1, m_0 > 0 \text{ et } m(0) > 0) \quad (2.11b)$$

$m(t)$ est une séquence interne générée à partir des signaux internes de telle façon que, si ceux-ci n'étaient pas bornés, elle ne peut pas diverger plus rapidement qu'exponentiellement. Alors, une condition suffisante pour réaliser la propriété EP sur $\varphi(t)$ est que l'erreur normalisée $\eta(t)/m(t)$ soit suffisamment petite. Une telle idée est illustrée dans (Cristi.), où la propriété EP et la régulation asymptotique sont simultanément réalisées moyennant un régulateur adaptatif direct. Plus généralement, $g(k)$ peut être choisi comme suit :

$$g(k) = \lambda(k).m(k) \quad (2.12a)$$

Alors, la propriété EP peut être atteinte étant donné que :

$$\min_k |\lambda(k)| \geq \varepsilon > 0 \quad (2.12b)$$

En commande adaptative indirecte, il est possible de formuler des conditions suffisantes de stabilité impliquant notamment le modèle estimé du système. Typiquement, le modèle estimé doit être

uniformément stabilisable et lentement variant dans le temps, (Giri et al, 1988b). Pour forcer le modèle estimé à être admissible, la séquence $\{\lambda(k)\}$ peut être choisie comme suit :

$$\lambda(k) = \begin{cases} \lambda_0 & \text{si le modèle estimé est stabilisable} \\ 0 & \text{sinon} \end{cases} \quad (2.13)$$

Ainsi, si la stabilisabilité se perdait infiniment longtemps alors la condition (2.12b) serait vérifiée garantissant l'EP et impliquant, de ce fait, la convergence du modèle estimé vers le vrai modèle, qui, lui, est admissible par hypothèse. Moyennant un raisonnement de ce type on montre ((Anderson, 1985b), (Giri et al, 1993), (Giri et al, 1988c), (Giri et al, 1990a)) que le modèle estimé devient forcément admissible après une période transitoire de durée finie, au terme de laquelle $\lambda(k)$ s'annule et demeure nulle. Des régulateurs adaptatifs opérant selon ce principe ont été développés dans plusieurs endroits (Anderson, 1985a-b), (Giri et al, 1993), (Giri et al, 1988c), (Giri et al, 1990), (Cristi,), (Polderman, 1989). La solution (2.13) est parfois qualifiée de mécanisme d'excitation commutée, (Ioannou P., Sun J., 1996).

2.3. NOTION DE PETITESSE EN MOYENNE

La taille moyenne d'une suite réelle peut être évaluée en utilisant le concept de petitesse en moyenne.

Définition 2.3. Une suite réelle $\{s(t)\}$ est dite α -petite en moyenne ($\alpha \in R^+$) si:

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \sum_{t=h+1}^{h+k} |s(t)| \leq \alpha \quad (2.14)$$

Pour une valeur donnée de α , l'ensemble des suites vérifiant (2.14) est noté $S(\alpha)$ □

Le lemme suivant rassemble quelques propriétés de la notion de petitesse définie précédemment.

Lemme 2.3. (Cristi,)

1) Si $\{s(t)\} \in S(\alpha)$, alors $\{|s(t)|\}$ est bornée.

2) Si $\{s_1(t)\} \in S(\alpha_1)$ et $\{s_2(t)\} \in S(\alpha_2)$ alors, pour toutes constantes réelles $\beta_1, \beta_2 \geq 0$:

$$\{\beta_1 s_1(t) + \beta_2 s_2(t)\} \in S(|\beta_1| \alpha_1 + |\beta_2| \alpha_2)$$

3) Soient $\{s_1(t)\} \in S(\alpha_1)$ et $\{s_2(t)\} \in S(\alpha_2)$. Si $\{x_1(t)\}$ et $\{x_2(t)\}$ sont des suites réelles bornées, respectivement par X_1 et X_2 (c'est-à-dire $|x_i(t)| \leq X_i, \forall t$). Alors, on a:

$$\{x_1(t)s_1(t) + x_2(t)s_2(t)\} \in S(X_1\alpha_1 + X_2\alpha_2)$$

4) Soit $\{s(t)\} \in S(\alpha)$ et $\varepsilon > 0$ un réel quelconque. Soit $\{s(t_i)\}$ la sous-suite comprenant tous les termes de $\{s(t)\}$ vérifiant $s(t_i) \leq \alpha + \varepsilon$. Alors, la suite incrémentale $\{t_i - t_{i-1}\}$ est bornée.

5) Soit $\{s(t)\} \in S(\alpha)$ et $\{s(t_i)\}$ une sous suite telle que $\{t_i - t_{i-1}\}$ est bornée. Alors, $\{s(t_i)\} \in S(\alpha\beta)$ avec $\beta = \sup_i (t_i - t_{i-1})$ \square

2.4. ALGORITHMES D'IDENTIFICATION PARAMETRIQUE

Dans ce paragraphe on rappelle les caractéristiques des algorithmes d'identification paramétrique les plus utilisés en identification (et en commande adaptative) des systèmes linéaires. Ces derniers sont décrits ici par le modèle (2.1) que l'on soumet aux hypothèses H2.1 et H2.2. Le choix de l'algorithme adéquat est alors influencé par la nature (déterministe ou stochastique, bornée a priori ou non) de l'erreur de sortie. Rappelons que cette erreur rend compte de l'effet de perturbations externes et de dynamiques négligées. La situation la plus simple à aborder, d'un point de vue théorique, est bien entendu celle où l'erreur est nulle. C'est par là que nous commençons pour une raison de pédagogie.

2.4.1. Identification paramétrique en l'absence d'erreur de sortie

Dans ce cas, l'équation (2.5a-b) peut être mise sous la forme de régression suivante :

$$y(t) = \varphi(t)^T \theta^* \quad (2.15a)$$

avec:

$$\theta = [a_1 \dots a_n \ b_1 \dots b_n]^T \quad (2.15b)$$

$$\varphi(t) = [-y(t-1) \dots -y(t-n) \ u(t-1) \dots u(t-n)]^T \quad (2.15c)$$

Le vecteur des paramètres inconnus, θ^* , peut être estimé par l'algorithme des moindres carrés ou l'une de ses nombreuses variantes. Un bon nombre d'estimateurs de ce type peuvent être décrits par la forme récursive suivante:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{P(t-1)\varphi(t)e(t)}{1 + \varphi^T(t)P(t-1)\varphi(t)} \quad (2.16a)$$

$$e(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) \quad (2.16b)$$

où $e(t)$ est l'erreur de prédiction a priori et $P(t)$ le gain d'adaptation. Les différentes variantes de cet algorithme correspondent en fait à différentes façons de réajuster le gain d'adaptation. Quelques unes de ces variantes sont présentées dans ce qui suit.

2.4.1.1. Algorithme du gradient

Dans ce cas, le gain est maintenu constant :

$$P(t) = \frac{1}{\lambda} I \quad (2.17)$$

où λ est une constante positive non nulle et I la matrice identité. L'estimateur (2.16a) s'écrit alors simplement :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)e(t)}{\lambda + \varphi^T(t)\varphi(t)} \quad (2.18)$$

On peut aisément montrer que l'estimateur $\hat{\theta}(t)$ minimise globalement la fonction coût :

$$J(x) = (y(t) - \varphi(t)^T x)^2 + \lambda \|x - \hat{\theta}(t-1)\|^2 \quad (2.19)$$

2.4.1.2. Algorithme des moindres carrés à gain décroissant

Dans ce cas, le gain est réajusté selon la loi :

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1 + \varphi^T(t)P(t-1)\varphi(t)} \quad (2.20)$$

L'estimateur, constitué de (2.16a) et (2.20), minimise globalement la fonction coût :

$$J(x) = \frac{1}{2} \sum_{i=1}^t (y(i) - \varphi(i)^T x)^2 + \frac{1}{2} (x - \hat{\theta}(0))^T P_0^{-1} (x - \hat{\theta}(0)) \quad (2.21)$$

Le premier terme du membre droit de (2.21) représente la somme des carrés de l'erreur de prédiction. Le second tient compte des conditions initiales ; P_0 mesure l'inverse de la confiance mise dans la condition initiale $\hat{\theta}(0)$. En général, on adopte comme valeurs initiales pour cet algorithme:

$$\hat{\theta}(0) = 0 \text{ et } P(0) = \alpha I \text{ (avec } \alpha \gg 1)$$

La nature décroissante du gain d'adaptation implique l'usure du pouvoir d'adaptation de l'estimateur. La vitesse avec laquelle le gain décroît devient exponentielle en présence d'excitation persistante. De ce fait, il est déconseillé d'utiliser cet estimateur quand le vecteur θ^* est susceptible de changer dans le temps. Dans ce dernier cas, on peut avoir recours aux variantes décrites dans les sous paragraphes suivants.

2.4.1.3. Algorithme des moindres carrés avec réinitialisation du gain

Dans ce cas, le gain est réinitialisé périodiquement :

$$P(t) = \begin{cases} C.I & \text{si } t = kT, k \in N \\ P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{1 + \varphi^T(t)P(t-1)\varphi(t)} & \text{sinon} \end{cases} \quad (2.22)$$

où C est une constante réelle positive et T est la période de réinitialisation du gain.

2.4.1.4. Algorithme des moindres carrés avec facteur d'oubli

Dans ce cas, le gain est mis à jour comme suit :

$$P(t) = \frac{1}{\lambda} \left(P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{\lambda + \varphi^T(t)P(t-1)\varphi(t)} \right) \text{ où } 0 < \lambda < 1 \quad (2.23)$$

L'estimateur constitué de (2.16a-b) et (2.23) minimise globalement la fonction coût :

$$J(x) = \frac{1}{2} \sum_{i=1}^t \lambda^{t-i} (y(i) - \varphi(i)^T x)^2 + \frac{1}{2} (x - \hat{\theta}(0))^T P_0^{-1} (x - \hat{\theta}(0)) \quad (2.24)$$

Ainsi, le poids accordé à chaque observation décroît exponentiellement au fur et à mesure que celle-ci vieillit. L'utilisation pratique de cet estimateur requiert de ce fait certaines précautions. En particulier, il faut s'assurer que $\varphi(t)$ possède bien la propriété d'excitation persistante ; autrement il y a un risque que le gain d'adaptation diverge.

La liste des lois d'adaptation du gain présentée jusque là est loin d'être exhaustive. La littérature spécialisée est abondante en variantes d'ajustement du gain (voir par exemple (Lozano et Taoutaou, 2001)).

2.4.1.5. Propriétés des algorithmes d'identification en l'absence d'erreur de sortie

Les principales caractéristiques des estimateurs présentés précédemment sont résumées dans le lemme suivant.

Lemme 2.4. Soit $\{\hat{\theta}(t)\}$ la suite générée par l'estimateur, constitué de (2.16a-b), et l'une des lois d'ajustement (2.17), (2.20), (2.22) ou (2.23). Si les données utilisées par l'estimateur sont générées par la régression linéaires (2.15a-c), alors $\{\hat{\theta}(t)\}$ a les propriétés suivantes :

- 1) $\{\|\hat{\theta}(t)\|\}$ est uniformément bornée.
- 2) $\left\{ \frac{|\varphi^T(t)\tilde{\theta}(t)|}{\sqrt{I + \varphi^T(t)\varphi(t)}} \right\} \in L_2 \Rightarrow \lim_{t \rightarrow \infty} \frac{|\varphi^T(t)\tilde{\theta}(t)|}{\sqrt{I + \varphi^T(t)\varphi(t)}} = 0$ avec $\tilde{\theta}(t) = \theta - \hat{\theta}(t)$
- 3) $\{\|\hat{\theta}(t) - \hat{\theta}(t-1)\|\} \in L_2 \Rightarrow \lim_{t \rightarrow \infty} \|\hat{\theta}(t) - \hat{\theta}(t-1)\| = 0$
- 4) a) $\{\|\tilde{\theta}(t)\|\}$ est non croissante pour l'algorithme du gradient.

- b) $\{\|\hat{\theta}(t)\|\}$ converge pour l'algorithme des moindres carrés.
- c) $\forall k \in N, \forall t \in N \cap [kT, (k+1)T[: \|\tilde{\theta}(t)\| \leq \|\tilde{\theta}(kT)\|$
- 5) Si, en outre, $\{\varphi(t)\}$ est bornée et possède la propriété d'EP, alors $\{\hat{\theta}(t)\}$ converge vers sa vraie valeur θ^* . La convergence est exponentielle pour les algorithmes du gradient et des moindres carrés avec réinitialisation du gain \square

La preuve de ce lemme se trouve dans plusieurs endroits, voir par exemple (Lozano et Taoutaou, 2001).

Remarque 2.3.

Les points 1 à 4 du lemme précédent montrent que tous les estimateurs produisent asymptotiquement un modèle (induit par $\hat{\theta}(t)$) présentant une bonne qualité de prédiction de la sortie. Cette propriété est d'autant plus appréciable qu'elle est obtenue pour n'importe quelle entrée du système (et sans que celui-ci soit nécessairement stable). L'excitation persistante et la bornitude de $\{\varphi(t)\}$ (donc de la stabilité du système) n'est en fait requise que pour garantir la consistance des estimateurs, ce qui dans le cas déterministe considéré, signifie la convergence de $\hat{\theta}(t)$ vers sa vraie valeur.

2.4.2. Identification paramétrique en présence d'une erreur de sortie de modèle inconnu

On se place ici dans le cas où aucune connaissance n'est disponible sur la suite $\{\eta_p(t)\}$ à l'exception du fait qu'elle est bornée. Dans ce cas, toute identification du système (2.5a-b) s'appuiera forcément sur la régression linéaire :

$$y(t) = \varphi(t)^T \theta^* + v(t) \tag{2.25}$$

où $v(t) = A(q^{-1})\eta_p(t)$ et θ^* , $\varphi(t)$ sont définis par (2.15b-c). L'usage de l'estimateur (2.16a-b) peut alors s'avérer dangereux et ce notamment en l'absence d'excitation persistante. En effet, l'estimateur est, dans ce cas, susceptible d'exhiber un comportement d'intégrateur ce qui, en présence d'une erreur $v(t)$ de moyenne non nulle, peut entraîner la divergence des estimés. Pour faire face au cas où l'excitation persistante n'est pas garantie, il faut modifier l'estimateur de manière à ce qu'il ne puisse pas diverger, tout en préservant l'essentiel de ses qualités décrites dans le lemme 2.4. A ce propos, on peut citer la modification qui consiste à projeter les estimés sur un

convexe ainsi que celle qui consiste à introduire une zone morte. Ces deux modifications sont illustrées ci-après sur l'algorithme du gradient et le seront plus tard sur l'algorithme des moindres carrés. Pour une présentation exhaustive de cet aspect le lecteur peut consulter [Lozano et Taoutaou, 2001, chapitres 1 à 3].

2.4.2.1. Algorithme du gradient avec projection

Cet algorithme se prémunit contre toute divergence des estimés moyennant une projection de ces derniers sur un convexe $C \subset \mathbb{R}^{2n}$, déterminé à l'avance. Pour s'assurer que cette projection préserve les qualités de l'estimateur initial, notamment celles décrites dans le Lemme 2.4, il faut s'assurer que le convexe contient le vrai vecteur des paramètres. A cet effet, on émet l'hypothèse suivante :

H2.3: $\theta^* \in C$

L'algorithme du gradient avec projection est alors défini comme suit:

$$\theta^p(t) = \hat{\theta}(t-1) + \frac{\varphi(t)e(t)}{1 + \varphi^T(t)\varphi(t)} \quad (2.26a)$$

$$\hat{\theta}(t) = \text{proj}_C(\theta^p(t)) \quad (2.26b)$$

$$e(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) \quad (2.256c)$$

où $\text{proj}_C(x)$ désigne l'opérateur de projection orthogonale sur C . Les propriétés de l'estimateur ainsi obtenu, sont décrites dans le lemme suivant, où μ désigne le plus petit réel tel que:

$$\{v(t)\} \in S(\mu) \quad (2.27)$$

Notons que $0 \leq \mu < \infty$ puisque $\{v(t)\}$ est bornée.

Lemme 2.5. Soit le système décrit par (2.5a-b), ainsi que par la régression linéaire (2.25), et soumis aux hypothèses H2.1 et H2.3. Soit $\{\hat{\theta}(t)\}$ la suite engendrée par l'algorithme (2.26a-c), à partir des données générées par la régression (2.25). Alors, on a les propriétés suivantes :

1) $\hat{\theta}(t) \in C, \quad \forall t$

2) Il existe un réel $K > 0$, indépendant de μ , tel que :

$$\left\{ \frac{e(t)}{\sqrt{1 + \varphi^T(t)\varphi(t)}} \right\}, \left\{ \|\hat{\theta}(t) - \hat{\theta}(t-1)\| \right\} \in S(K\mu)$$

- 3) Si $\{\eta_p(t)\}$ est nulle, alors les propriétés de l'algorithme de gradient sans projection décrites dans le lemme 2.4, sont préservées □

La preuve de ce lemme peut être trouvée, entre autres, dans [Lozano et Taoutaou, 2001, chapitre 3].

2.4.2.2. Algorithme du gradient avec zone morte

Dans un estimateur avec zone-morte, la recherche des paramètres inconnus est active uniquement quand l'erreur de prédiction a priori est supérieure à un seuil réglable, noté dz . En deçà de ce seuil, la recherche est gelée. La formulation de cette idée, dans le cadre d'une estimation par l'algorithme du gradient, donne lieu à l'estimateur suivant :

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \frac{\varphi(t)D(e(t))}{1 + \varphi^T(t)\varphi(t)} \quad (2.28a)$$

$$e(t) = y(t) - \varphi^T(t)\hat{\theta}(t-1) \quad (2.28b)$$

$$D(x) = \begin{cases} x - dz & \text{si } x > dz \\ 0 & \text{si } |x| < dz \\ x + dz & \text{si } x < -dz \end{cases} \quad (2.28c)$$

Les propriétés essentielles de cet estimateur sont rassemblées dans le lemme suivant, où l'on utilise la notation :

$$\mu = \sup_t |v(t)| \quad (2.29)$$

Notons que $0 \leq \mu < \infty$ puisque $\{v(t)\}$ est borné.

Lemme 2.6. Soit le système décrit par (2.5a-b), ainsi que par la régression linéaire (2.25), et soumis à l'hypothèse H2.1. Soit $\{\hat{\theta}(t)\}$ la suite engendrée par l'algorithme (2.28a-c), à partir des données générées par la régression (2.25). Si $\mu \leq dz$ alors on a les propriétés suivantes :

- 1) $\{\hat{\theta}(t)\}$ est bornée
- 2) $\left\{ \frac{D(e(t))}{\sqrt{1 + \varphi^T(t)\varphi(t)}} \right\} \in L_1 \cap L_2 \Rightarrow \lim_{t \rightarrow \infty} \frac{D(e(t))}{\sqrt{1 + \varphi^T(t)\varphi(t)}} = 0$
- 3) $\|\hat{\theta}(t) - \hat{\theta}(t-1)\| \in L_2$
- 4) si $\{\varphi(t)\}$ est bornée, alors $\{\hat{\theta}(t)\}$ converge.

La preuve de ce lemme peut être trouvée, entre autres, dans [Lozano et Taoutaou, 2001, ch. 3].

A la lumière du lemme précédent (notamment le point 2), il vient que la qualité du modèle est d'autant meilleure que la taille de la zone-morte (dz) est plus proche de la taille de l'erreur d'équation (μ). L'idéal serait $dz = \mu$ ce qui suppose une bonne connaissance a priori sur le système objet de l'identification, ce qui n'est pas toujours le cas en pratique.

2.4.3. Identification paramétrique en présence d'une erreur de sortie modélisable par un processus stochastique de modèle connu

On suppose ici que l'erreur de sortie $\{\eta_p(t)\}$ est quelconque mais l'on sait qu'elle est générée par un modèle de la forme :

$$\eta_p(t) = \frac{C(q^{-1})}{D(q^{-1})} \xi(t) \quad (2.30)$$

où $\{\xi(t)\}$ est une suite de variables aléatoires indépendantes (régies par la même loi) de moyenne nulle (bruit blanc) et $C(q^{-1})/D(q^{-1})$ est un filtre stable. Cette situation est traitée de façon exhaustive dans la littérature spécialisée, voir notamment (Ljung, 1997), (Soderstrom, et Stoica, 1988), (Goodwin et Sin, 1983). Nous nous contentons de rappeler ici que la majorité des méthodes d'identification paramétriques visent, dans ce cas, à estimer les deux fonctions de transfert $B(q^{-1})/A(q^{-1})$ et $C(q^{-1})/D(q^{-1})$, en s'appuyant généralement sur une régression pseudo-linéaire. Le seul cas de base où l'on peut associer au système une forme de régression linéaire est quand :

$$D(q^{-1}) = A(q^{-1}) \text{ et } C(q^{-1}) = I \quad (2.31)$$

La régression en question est simplement :

$$y(t) = \varphi^T(t) \theta^* + \xi(t) \quad (2.32)$$

Pour estimer θ^* , il peut être fait appel à l'algorithme des moindres carrés standard dont la forme récursive a été présentée dans le sous paragraphe 2.3.1.2. La forme non récursive est elle définie par l'expression :

$$\hat{\theta}(N) = \left[\frac{1}{N} \sum_{t=1}^N \varphi(t) \varphi^T(t) \right]^{-1} \left[\frac{1}{N} \sum_{t=1}^N y(t) \varphi(t) \right] \quad (2.33)$$

Bien que les propriétés de cet estimateur sont très connues, nous les rappelons ici pour la clarté de l'exposé.

Lemme 2.7. Soit le système décrit par les équations (2.5a-b), (2.30), (2.31) et soumis à l'hypothèse H2.1. Soit $\hat{\theta}(N)$ le vecteur estimé par l'algorithme (2.28a-c) à partir des données générées par la régression (2.32). Si $\{\varphi(t)\}$ possède la propriété d'EP, alors on a les propriétés suivantes de $\hat{\theta}(N)$:

- 1) $\hat{\theta}(N)$ est un estimé consistant de θ^* , ce qui signifie que :

$$\hat{\theta}(N) \rightarrow \theta^*, \text{ avec probabilité 1, quand } N \rightarrow \infty$$
- 2) $\hat{\theta}(N)$ est un estimé non biaisé, ce qui veut dire que $E(\hat{\theta}(N)) = \theta^*$

La preuve de ce lemme peut être trouvée dans (Ljung, 1997), (Soderstrom, et Stoica, 1988), (Goodwin et Sin, 1983) (entre autres).

2.5. RESOLUTION D'UN TYPE DE SYSTEME D'EQUATIONS ALGEBRIQUES BILINEAIRES

Les schémas d'identification paramétrique élaborés dans le cadre de cette thèse et ailleurs, notamment pour les systèmes non linéaires de type Hammerstein et Wiener, mettent en évidence des systèmes d'équations algébriques de la forme :

$$b_i d_j = \mu_{ij} ; i = 1, \dots, n ; j = 1, \dots, N \quad (2.34)$$

où les entiers n et N et les réels μ_{ij} sont connus ; les inconnues étant les b_i et les d_j . En vue de résoudre ce système, on commence par lui donner la forme matricielle suivante :

$$\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [d_1 \quad \dots \quad d_N] = \begin{bmatrix} \mu_{11} & \dots & \mu_{1N} \\ \vdots & & \vdots \\ \mu_{n1} & \dots & \mu_{nN} \end{bmatrix} \quad (2.35)$$

Par commodité, on introduit les notations suivantes :

$$M = \begin{bmatrix} \mu_{11} & \dots & \mu_{1N} \\ \vdots & & \vdots \\ \mu_{n1} & \dots & \mu_{nN} \end{bmatrix} \quad (2.36)$$

$$B = [b_1 \quad \dots \quad b_n]^T ; D = [d_1 \quad \dots \quad d_N]^T \quad (2.37)$$

Moyennant ces notations, l'équation (2.35) se réécrit sous forme plus compacte comme suit :

$$BD^T = M \quad (2.38)$$

Remarquons que cette équation n'a pas une solution unique. En effet, si (B^*, D^*) est une solution alors, pour tout réel α non nul, le couple $\left(\alpha B^*, \frac{D^*}{\alpha}\right)$ est aussi une solution. Si l'on s'intéresse à une solution particulière, il faut en préciser les caractéristiques. On peut par exemple s'intéresser à l'unique solution qui vérifie les conditions suivantes:

$$\sum_{i=1}^n b_i^2 = 1 \text{ et } \rho([b_1 \dots b_n]) > 0 \quad (2.39)$$

où $\rho([b_1 \dots b_n])$ désigne la première composante du vecteur B ayant la plus grande valeur absolue.

Cette composante vérifie :

$$|\rho([b_1 \dots b_n])| = \sup_{1 \leq j \leq n} |b_j| \quad (2.40a)$$

En outre, en posant $\rho([b_1 \dots b_n]) = b_i$, il vient que, pour tout k :

$$|b_k| = \sup_{1 \leq j \leq n} |b_j| \Rightarrow i \leq k \quad (2.40b)$$

A la lumière des observations précédentes, il vient que le problème formulé par l'équation (2.38) et les contraintes (2.39) a une solution unique. Le lemme qui suit décrit comment cette solution peut être déterminée.

Lemme 2.8. Soit $M \in \mathbb{R}^{n \times N}$ une matrice réelle de rang 1. Sa décomposition en valeurs singulières a la forme suivante:

$$M = \Gamma \begin{bmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & 0 & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \Sigma \quad \text{où } \Gamma \in \mathbb{R}^{n \times n}, \Sigma \in \mathbb{R}^{N \times N}, \quad (2.41)$$

et σ_1 est l'unique valeur non singulière de M . Il s'ensuit que :

$$M = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [d_1 \quad \dots \quad d_N] \quad (2.42)$$

avec:

$$B = \frac{\Gamma [\sigma_1 \quad 0 \quad \dots \quad 0]^T}{\|\Gamma [\sigma_1 \quad 0 \quad \dots \quad 0]^T\|} \text{sign}(\gamma) \quad (2.43)$$

$$D^T = \|\Gamma [\sigma_1 \quad 0 \quad \dots \quad 0]^T\| \cdot [I \quad 0 \quad \dots \quad 0] \Sigma \text{sign}(\gamma) \quad (2.44)$$

où $\gamma = \rho(\Gamma[\sigma_1 \ 0 \ \dots \ 0]^T)$. Le couple (B,D) ainsi défini est l'unique solution du problème défini par (2.38)-(2.39) \square

La preuve de ce lemme est placée en Annexe A.2.

DEUXIEME PARTIE

IDENTIFICATION PARAMETRIQUE DES SYSTEMES NON LINEAIRES DE TYPE HAMMERSTEIN

Cette partie est dédiée à l'identification paramétrique des systèmes non linéaires sur la base d'un modèle de type Hammerstein (Fig3.1). Ce modèle est constitué d'un gain statique non linéaire, caractérisé par une fonction F , et d'un sous-système dynamique linéaire de fonction de transfert $G(p)$ ou $G(z)$ (selon le contexte temporel considéré). Les deux éléments sont montés en série, l'élément non linéaire étant placé à l'entrée. L'intérêt pratique de ce modèle est double :

- (i) il convient bien aux procédés industriels du fait que ceux-ci comprennent un actionneur (électrique, hydraulique, pneumatique) placé à l'entrée et dont le comportement est souvent assimilé à une non-linéarité statique.
- (ii) il peut représenter, pour de nombreux autres types de systèmes non linéaires, un bon compromis qualité des estimés/simplicité.

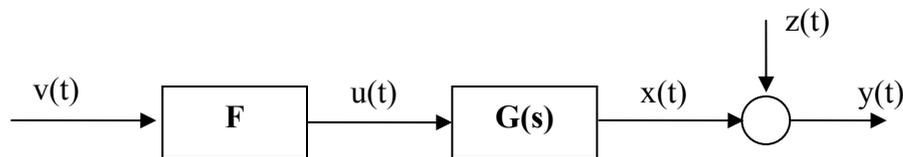


Fig.3.1 : Structure d'un modèle de Hammerstein

L'identification des procédés non linéaires sur la base de ce modèle a suscité un grand intérêt, au sein de la communauté scientifique, et ce notamment au cours des vingt dernières années. On peut citer à cet égard comme travaux (Bai, 2002), (Billings et Fakhouri, 1979), (Giri et al, 2000)-(Giri et al, 2001a-c), (Greblicki, 1989), (Greblicki, et Pawlak, 1991a), (Krzyzak, 1989), (Lang, 1993a), (Lang, 1997), (Pawlak, 1991), (Stoica et Söderström, 1982), (Vörös, 1999a). Globalement, le problème fut abordé en suivant deux approches principales:

- Approche paramétrique : dans ce cas, le gain non linéaire est représenté par une fonction polynomiale ou, plus généralement, par un développement (limité) en série de fonctions orthogonales. De ce fait, il est caractérisé par un nombre fini de paramètres inconnus. Le problème d'identification consiste alors à estimer l'ensemble des paramètres inconnus : ceux de l'élément non linéaire et ceux de la dynamique linéaire. L'estimation des paramètres inconnus est généralement effectuée à l'aide d'algorithmes récursifs de type moindres carrés, ((Billings et Fakhouri, 1979), (Stoica et Söderström, 1982), (Lang, 1993a), (Giri et al, 2001b), (Vörös, 1999), (Bai, 2002)).

- Approche non paramétrique : l'élément non linéaire n'est alors pas caractérisé par une fonction analytique de type connu. Il peut être n'importe quelle fonction continue comme dans (Lang, 1997) ou juste une fonction mesurable (au sens de Lebesgue/Borel) comme dans ((Greblicki, et Pawlak, 1991a), (Greblicki, 1989), (Pawlak, 1991), (Krzyzak, 1989), (Kung et Shinh, 1986)). Le but de l'identification, concernant cet élément, est l'estimation d'un nombre fini de points de sa caractéristique graphique. Cet ensemble de points peut par la suite faire l'objet d'une interpolation de type polynomiale ou autre.

L'analyse de convergence des paramètres estimés par les différentes approches a généralement été faite à l'aide d'outils stochastiques. Ainsi, a-t-on établi dans (Stoica et Söderström, 1982) que la consistance d'une méthode d'identification paramétrique (impliquant la variable instrumentale) peut être garantie à condition d'exciter le système à l'aide d'un bruit blanc. Des entrées aléatoires spécifiques ont également été développées pour assurer la consistance de méthodes non paramétriques, ce fut par exemple le cas dans (Pawlak, 1991), (Lang, 1993a).

Dans cette partie nous présentons trois méthodes d'identification dont la consistance sera assurée à l'aide de signaux d'entrée déterministes. Ces méthodes intéressent des systèmes de Hammerstein impliquant différents types d'éléments non linéaires. Au chapitre 3, aucune hypothèse n'est émise sur l'élément non linéaire à part le fait qu'il soit statique (sans mémoire) ; le but de l'identification étant l'estimation des paramètres de la dynamique linéaire du système et d'un ensemble de points de la caractéristique de l'élément non linéaire. Dans le chapitre 4, nous nous intéresserons aux systèmes de Hammerstein impliquant des éléments non linéaires de structure connue. Pour une raison de pédagogie, nous aborderons d'abord le cas trivial où la non-linéarité est de type polynomial. L'attention sera ensuite portée sur les non-linéarités non lisses de type zone morte et/ou seuil. Le chapitre 5 portera sur les systèmes de Hammerstein impliquant des éléments non linéaires à mémoire, l'attention sera précisément focalisée sur les éléments de type hystérésis et hystérésis-relais. Il convient de souligner que l'originalité des schémas présentés dans les chapitres 3, 4 et 5 réside, d'une part, dans la simplicité de mise en oeuvre et, d'autre part, dans le fait que des propriétés fondamentales telles que l'excitation persistante et la consistance des estimateurs, sont assurées moyennant des signaux d'entrée déterministes.

CHAPITRE 3

IDENTIFICATION DES SYSTEMES DE HAMMERSTEIN IMPLIQUANT UN ELEMENT NON LINEAIRE STATIQUE DE STRUCTURE INCONNUE

3.1. INTRODUCTION

Dans ce chapitre, le problème d'identification des systèmes de type Hammerstein est abordé sous des hypothèses très générales concernant l'élément non linéaire. Etant seulement supposé statique, cet élément est analytiquement caractérisé à l'aide d'une fonction $F(.)$ dont la nature est totalement inconnue. Le but de l'identification est l'estimation du modèle de la dynamique linéaire du système et la détermination d'un ensemble de points $(V_i, F(V_i))$ ($i=1, \dots, N$) de la caractéristique de l'élément non linéaire. L'entier N et les réels V_i sont arbitrairement choisis par l'utilisateur. Il va sans dire que la connaissance d'un nuage de points $(V_i, F(V_i))$ rend possible l'identification de la nature et la structure de l'élément non linéaire ce qui, à son tour, permet de mieux appréhender la question d'interpolation de ce nuage de points.

Le schéma d'identification exposé dans ce chapitre est conçu en combinant un certain nombre d'idées brièvement présentées dans ce qui suit. L'ensemble des points $(V_i, F(V_i))$ déterminent de manière unique un polynôme $P(v)$ de degré N qui coïncide avec $F(v)$ sur l'ensemble $\{V_i; i=1, \dots, N\}$. Ainsi, la détermination de l'ensemble des points $\{(V_i, F(V_i)); i=1, \dots, N\}$ équivaut à l'identification du polynôme $P(v)$. En outre, le polynôme $P(v)$ peut très bien se substituer à $F(v)$ dans le modèle du procédé à condition d'astreindre l'entrée du système à prendre ses valeurs dans l'ensemble $\{V_i, i=1, \dots, N\}$. Dans ces conditions, le problème d'identification initial se trouve converti en un nouveau problème impliquant comme élément non linéaire le polynôme $P(v)$. Le nouveau problème d'identification porte sur un système dont le modèle est caractérisé par un nombre fini de paramètres inconnus ; ceux-ci incluent les paramètres de la partie dynamique linéaire et les coefficients du polynôme $P(v)$. Ces paramètres interviennent de façon bilinéaire dans le nouveau modèle qui, de ce fait, se présente comme étant linéaire par rapport aux paramètres

composés (produits des paramètres recherchés). Il est donc logique de faire porter l'identification paramétrique sur les paramètres composés et faire appel à cet effet à un estimateur de type moindres carrés. La consistance des estimés, du moins dans le cas idéal, sera assurée en faisant recours à un signal d'entrée de type impulsionnel. La dernière étape consiste à recouvrer les paramètres recherchés à partir des estimés des paramètres composés. A cet effet, une procédure algébrique, impliquant une décomposition en valeurs singulières, est élaborée. Notons enfin que le schéma d'identification décrit dans ce chapitre s'appuie étroitement sur les résultats présentés dans ((Giri et al, 2001b), (Chaoui et al, 2005)).

Le chapitre est organisé comme suit : le système à identifier est décrit dans la section 2 ; le schéma d'identification est présenté dans les sections 3 à 5 et il est analysé dans la section 6. Finalement, une simulation illustre l'efficacité de la méthode d'identification proposée.

3.2. POSITION DU PROBLEME D'IDENTIFICATION

3.2.1. Classe de procédés à identifier

On s'intéresse aux systèmes dont le comportement entrée/sortie, en temps discret, peut être décrit par le modèle suivant:

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}u(t) + z(t) \quad (3.1a)$$

$$u(t) = F(v(t))$$

avec

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_naq^{-na} \quad (3.2a)$$

$$B(q^{-1}) = b_1q^{-1} + \dots + b_n bq^{-nb} \quad (3.2b)$$

où q^{-1} représente l'opérateur retard, $v(t)$ et $y(t)$ désignent respectivement l'entrée et la sortie; $u(t)$ est un signal interne *non mesurable* ; $z(t)$ est une erreur qui rend compte de l'effet des perturbations externes (et éventuellement d'erreurs de modélisation), il est seulement supposé que la suite $\{z(t)\}$ est bornée. $F(.)$ est une fonction non identiquement nulle représentant le gain statique et vérifiant les propriétés suivantes :

$$F(0) = 0 \quad (3.3a)$$

$$|F(v)| < \infty \text{ pour tout } |v| < \infty. \quad (3.3b)$$

L'égalité (3.3a) garantit que l'origine est bien un point d'équilibre du système ; la condition (3.3b) assure que la fonction F est définie partout sur \mathbb{R} , ce qui exclut des points d'échappement fini. La dynamique du système est quant à elle soumise aux hypothèses suivantes :

H3.1. Il existe un majorant connu n tel que: $n \geq \max \{n_a, n_b\}$.

H3.2. $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux.

H3.3. Tous les zéros de $q^{na}A(q^{-1})$ sont à l'intérieur du cercle unité.

Remarques 3.1.

- a) L'hypothèse H3.1 porte sur la structure du sous-système dynamique, elle est habituelle en identification. L'hypothèse H3.2 assure la commandabilité de la fonction de transfert $B(q^{-1})/A(q^{-1})$. L'hypothèse H3.3 est cohérente avec le fait que l'on se place ici dans un contexte d'identification en boucle ouverte ; la stabilité du système est dans ce cas une hypothèse indispensable.
- b) Mis à part les hypothèses précédentes, le système est quelconque. En particulier, les paramètres dynamiques (a_i, b_i) sont inconnus et les coefficients principaux (b_1, b_2, \dots) peuvent être nuls ; cette dernière remarque implique que le vrai retard est inconnu (néanmoins, on sait qu'il n'est pas plus grand que n).
- c) Notons que dans (Giri et al, 2001b), l'élément non linéaire $F(\cdot)$ est supposé affine par morceaux avec une dérivée non nulle à l'origine. Dans ce chapitre, $F(\cdot)$ est supposée seulement non identiquement nulle.

3.2.2. Objectif d'identification

Notre but est de concevoir un schéma d'identification en ligne qui, dans le cas idéal, détermine exactement le vrai modèle de la dynamique du système $B(q^{-1})/A(q^{-1})$ ainsi qu'un ensemble de points de la caractéristique $F(\cdot)$ de l'élément non linéaire. Dans le cas non idéal, il est attendu de ce schéma qu'il génère un modèle dont la qualité est d'autant plus élevée que la taille de l'erreur $\{z(t)\}$ est plus petite.

Remarques 3.2.

- a) Il est clair que la difficulté du problème d'identification ainsi posé réside, entre autres, dans le fait que le signal interne $u(t)$ n'est pas mesurable. Seules les séquences d'entrée $\{v(t)\}$ et de sortie $\{y(t)\}$ peuvent être utilisées dans le schéma d'identification (Fig.3.1).
- b) Le problème d'identification tel qu'il est formulé ci-dessus considère les deux situations extrêmes concernant la séquence d'erreur $\{z(t)\}$ à savoir :

i) $\{z(t)\}=\{0\}$

ii) $\{z(t)\}$ quelconque mais bornée

Pour traiter ces deux cas, il va être fait appel qu'à des outils déterministes. Le cas intermédiaire où $\{z(t)\}$ serait un processus stochastique obtenu par filtrage linéaire d'un bruit blanc fut traité dans (Chaoui et al , 2005).

3.3. PARAMETRISATION DU MODELE DU SYSTEME

Le schéma d'identification présenté dans ce chapitre sera développé en trois étapes. La première, exposée dans ce paragraphe, consiste à caractériser la non-linéarité par un nombre fini de paramètres et à élaborer pour le système une paramétrisation impliquant de façon bilinéaire les paramètres inconnus de ce dernier. A cet effet, soit $(V_j, F(V_j))$ ($j=1, \dots, N$) l'ensemble des points de $F(.)$ que l'on cherche à déterminer. Le choix de l'entier N et des réels V_j est entièrement laissé à l'utilisateur. Il est seulement supposé que :

$$0 < V_1 < V_2 < \dots < V_N < \infty \tag{3.4}$$

Moyennant une interpolation polynomiale, on associe aux couples $(V_j, F(V_j))$ un polynôme unique de degré N tel que:

$$P(0)=0 \tag{3.5}$$

$$P(V_j) = F(V_j) \quad \text{pour} \quad j=1, \dots, N \tag{3.6}$$

Ce polynôme peut être écrit sous la forme:

$$P(v) = v \sum_{j=1}^N d_j P_j(v) \quad \text{avec} \quad P_j(v) = \prod_{\substack{i=1 \\ i \neq j}}^N \frac{v - V_i}{V_j - V_i} \tag{3.7}$$

Il est évident que:

$$P_j(V_i) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad \text{et} \quad d_j = \frac{P(V_j)}{V_j} \tag{3.8}$$

L'identification des N couples $(V_j, F(V_j))$ équivaut à identifier le polynôme $P(v)$, c'est-à-dire les coefficients d_j ($j=1, \dots, N$). En plus, si l'entrée du système est astreinte à prendre ses valeurs dans

l'ensemble $\{0, V_1, \dots, V_N\}$ alors, il découle de (3.6) que le système (3.1a-b) peut être représenté par le modèle suivant:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + A(q^{-1})z(t) \quad (3.9a)$$

$$u(t) = P(v(t)) \quad (3.9b)$$

Il convient de souligner que le remplacement de $F(\cdot)$ par $P(\cdot)$ dans (3.9b) ne génère aucune erreur et ce tant que $v(t)$ appartienne à l'ensemble $\{0, V_1, \dots, V_N\}$. Dans ces conditions, les deux modèles (3.1a-b) et (3.9a-b) sont parfaitement équivalents. Ainsi, le problème d'identification initial (qui consistait à déterminer les N couples $(V_j, F(V_j))$ et les paramètres (a_i, b_i)) a-t-il été converti en un nouveau problème portant sur la détermination des paramètres a_i, b_i et d_j . Une fois ce nouveau problème résolu, les N couples $(V_j, F(V_j))$ pourront être recouverts à partir des (estimés des) paramètres d_j . En effet, il vient de (3.8) que :

$$F(V_j) = P(V_j) = d_j \cdot V_j \quad \text{pour } j=1, \dots, N \quad (3.9c)$$

Contrairement à la représentation initiale (3.1a-b), la nouvelle représentation (3.9a-b) permet d'élaborer une forme de régression pour le système considéré. En effet, en substituant (3.7) et (3.9b) dans (3.9a) on obtient successivement :

$$\begin{aligned} A(q^{-1})y(t) &= B(q^{-1})P(v(t)) + A(q^{-1})z(t) \\ &= \sum_{i=1}^n b_i q^{-i} \sum_{j=1}^N d_j P_j(v(t))v(t) + \eta(t) \\ &= \sum_{i=1}^n \sum_{j=1}^N \mu_{ij} v(t-i)P_j(v(t-i)) + \eta(t) \end{aligned} \quad (3.10a)$$

avec $\eta(t) = A(q^{-1})z(t)$ et :

$$\mu_{ij} = b_i d_j \quad (i=1, \dots, n; \quad j=1, \dots, N) \quad (3.10b)$$

L'équation (3.10a) donne lieu à la forme de régression suivante:

$$y(t) = \Phi(t)^T \Theta^* + \eta(t) \quad (3.11a)$$

$$\begin{aligned} \Phi(t)^T &= [-y(t-1) \quad -y(t-n) \quad v(t-1)P_1(v(t-1)) \quad v(t-n)P_1(v(t-n)) \quad \dots \\ &\dots \quad v(t-1)P_N(v(t-1)) \quad \dots \quad v(t-n)P_N(v(t-n))] \end{aligned} \quad (3.11b)$$

$$\Theta^* = [a_1 \dots a_n \mu_{11} \dots \mu_{n1} \dots \mu_{1N} \dots \mu_{nN}]^T \quad (3.11c)$$

Le vecteur Θ^* intervient linéairement dans l'équation (3.11a) qui, de ce fait, peut servir comme base pour l'estimation de ce vecteur, en utilisant à cet effet un estimateur de type moindres carrés. Cet aspect sera développé au paragraphe 3.

A la lumière de ce qui précède, il est devenu clair que nous allons pouvoir disposer d'estimés des paramètres a_i et μ_{ij} . Ceci nous conduit naturellement à nous poser la question suivante : comment peut-on obtenir des estimés de (b_i, d_j) à partir des estimés des μ_{ij} ?

En vue d'élaborer une réponse adéquate à cette question, il convient de lui donner une reformulation mathématique précise. A cet effet, remarquons que l'ensemble des égalités (3.10b) peut être regroupés sous la forme compacte suivante :

$$M = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [d_1 \quad \cdots \quad d_n] \quad \text{avec} \quad M = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1N} \\ \vdots & & \vdots \\ \mu_{1n} & \cdots & \mu_{nN} \end{bmatrix} \quad (3.12)$$

Dans cette équation, la matrice M est supposée connue ; le but étant de déterminer les vecteurs figurant à droite. Ce problème fut abordé au chapitre 2 (§2.4) où l'on a remarqué qu'il n'a pas une solution unique. L'unicité peut néanmoins être garantie en soumettant la solution aux contraintes suivantes :

$$\sum_{i=1}^n b_i^2 = 1 \quad \text{et} \quad \rho([b_1 \dots b_n]) > 0 \quad (3.13)$$

où $\rho([b_1 \dots b_n]) > 0$ représente la première composante du vecteur $[b_1 \dots b_n]^T$ possédant la plus grande valeur absolue. Autrement dit, la solution recherchée est telle que la norme du vecteur $[b_1 \dots b_n]^T$ est égale à 1 et sa première composante avec la plus grande valeur absolue est positive.

En appliquant le Lemme 2.8, il vient que la solution du problème (3.12)-(3.13) est la suivante

$$[b_1 \quad \cdots \quad b_n]^T = \frac{\Gamma[\sigma_1 \quad \cdots \quad 0]^T}{\|\Gamma[\sigma_1 \quad \cdots \quad 0]^T\|} \text{sign}(\gamma) \quad (3.14a)$$

$$[d_1 \quad \cdots \quad d_n]^T = \|\Gamma[\sigma_1 \quad \cdots \quad 0]^T\| \cdot [I \quad 0 \quad \cdots \quad 0] \Sigma \text{sign}(\gamma) \quad (3.14b)$$

où les notations sont celles du Lemme 2.8.

3.4. ESTIMATION DES PARAMETRES DU MODELE

3.4.1. Estimation des paramètres a_i et μ_{ij}

2.4.1.1. Algorithme d'estimation

Des estimés $\hat{\Theta}(t)$ de Θ^* peuvent être obtenus de manière récursive en utilisant, par exemple, l'algorithme suivant :

$$\bar{\Theta}(t) = \hat{\Theta}(t-1) + \frac{\Phi(t)e(t)}{1 + \Phi(t)^T \Phi(t)} \quad (3.15a)$$

$$\hat{\Theta}(t) = \min(\rho^*, \|\bar{\Theta}(t)\|) \cdot \frac{\bar{\Theta}(t)}{\|\bar{\Theta}(t)\|}$$

$$e(t) = y(t) - \Phi(t)^T \hat{\Theta}(t-1) \quad (3.15b)$$

où $\hat{\Theta}(0)$ est arbitrairement choisi et la constante réelle ρ^* est un majorant quelconque de $\|\Theta^*\|$.

Les équations (3.15a-b) définissent en fait un algorithme de type gradient avec projection sur une sphère centrée sur l'origine et de rayon ρ^* . Une telle projection est utilisée pour éviter toute divergence de $\|\hat{\Theta}(t)\|$, en dépit de l'erreur de modélisation $\{\eta(t)\}$.

La qualité des estimés $\hat{\Theta}(t)$ dépend, en partie, de la taille moyenne de $\{\eta(t)\}$. Pour établir ceci, notons μ le plus petit réel tel que :

$$\{\eta(t)\} \in S(\mu) \quad (3.16)$$

Notons que μ existe puisque $\{\eta(t)\}$ est bornée. Les propriétés de l'algorithme d'estimation (3.15a-b) peuvent alors être exprimées en terme de μ et de l'erreur de prédiction à posteriori définie comme suit :

$$e_p(t) = y(t) - \Phi(t)^T \hat{\Theta}(t) \quad (3.17)$$

Proposition 3.1. (Propriétés générales de $\hat{\Theta}(t)$)

1) Il existe une constante réelle K_I , indépendante de μ , telle que:

$$\left\{ \left\| \hat{\Theta}(t) - \hat{\Theta}(t-1) \right\| \right\} \text{ et } \{e_p(t)\} \in S(K_I \mu)$$

2) Dans le cas idéal (où $\{z(t)\} = \{0\}$) on a :

$$\left\{ \left\| \hat{\Theta}(t) - \hat{\Theta}(t-1) \right\| \right\} \in L_2, \quad \{e_p(t)\} \in L_2. \quad \square$$

Preuve: Cette proposition fut démontrée pour la première fois dans (Edgard, 1979). Cette preuve peut être trouvée dans des références plus récentes, voir par exemple ('Panorama des algorithmes d'identification' dans (Lozano et Taoutaou, 2001)).

Cette proposition montre que la qualité du modèle, induite asymptotiquement par le vecteur $\hat{\Theta}(t)$, dépend de la taille moyenne, mesurée par μ , de l'erreur de modélisation $\eta(t)$. La qualité prédictive du modèle est d'autant meilleure que μ est plus petite. La puissance de ce résultat réside dans le fait qu'il est obtenu sans émettre aucune hypothèse sur la séquence d'entrée $\{v(t)\}$. En revanche, la Proposition 3.1 ne stipule pas que $\hat{\Theta}(t)$ converge vers sa vraie valeur Θ^* et ce même dans le cas idéal (absence de perturbations). Pour garantir la convergence des estimés vers leur vraie valeur il faut utiliser une **séquence d'entrée $\{v(t)\}$ suffisamment riche**, afin de procurer au **vecteur de régression sous jacent** la propriété **d'excitation persistante**. Un exemple d'une telle séquence sera développé dans le paragraphe 5 de ce chapitre.

3.4.2. Reconstruction de $A(q^{-1})$, $B(q^{-1})$ et des N points de la non linéarité $F(.)$

A partir du vecteur des paramètres estimés :

$$\hat{\Theta}(t) = [\hat{a}_1(t) \quad \dots \quad \hat{a}_{n-1}(t) \quad \hat{\mu}_{11}(t) \quad \dots \quad \hat{\mu}_{n1}(t) \quad \hat{\mu}_{12}(t) \quad \dots \quad \hat{\mu}_{n2}(t) \quad \dots \quad \hat{\mu}_{1N}(t) \quad \dots \quad \hat{\mu}_{nN}(t)]^T \quad (3.16a)$$

on peut construire immédiatement l'estimé de $A(q^{-1})$:

$$\hat{A}_t(q^{-1}) = 1 + \hat{a}_1(t)q^{-1} + \dots + \hat{a}_n(t)q^{-n} \quad (3.16b)$$

Les estimés (\hat{b}_i, \hat{d}_j) de (b_i, d_j) sont obtenus à partir des $\hat{\mu}_{ij}$ en suivant la démarche présentée au paragraphe 3.3, fondée sur le Lemme 2.8. A cet effet, on construit la matrice suivante:

$$\hat{M}(t) = \begin{bmatrix} \hat{\mu}_{11}(t) & \cdots & \hat{\mu}_{1N}(t) \\ \vdots & & \vdots \\ \hat{\mu}_{1n}(t) & \cdots & \hat{\mu}_{nN}(t) \end{bmatrix} \quad (3.17)$$

Tant que la convergence des $\hat{\mu}_{ij}$ vers leurs vraies valeurs n'est pas prouvée, la matrice $\hat{M}(t)$ ne peut être supposée de rang 1, ce qui est le cas de sa vraie valeur M définie par (3.12). Sa décomposition en valeur singulière prend donc la forme générale suivante :

$$\hat{M}(t) = \Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2(t) & \ddots & & \vdots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & \sigma_n(t) & 0 & \cdots & 0 \end{bmatrix} \Sigma(t) \quad (3.18)$$

où l'on a supposé que $n < N$. Il est clair que si $\hat{M}(t)$ converge vers M alors $\sigma_1(t)$ converge vers σ_1 et $\sigma_i(t)$ ($i=2, \dots, n$) converge vers zéro. Compte tenu de cette remarque, la démarche exposée au paragraphe 3.3 suggère pour les (b_i, d_i) les estimés suivants:

$$\begin{bmatrix} \hat{b}_1(t) & \cdots & \hat{b}_n(t) \end{bmatrix}^T = \frac{\Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 \end{bmatrix}^T}{\left\| \Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 \end{bmatrix}^T \right\|} \text{sign}(\gamma(t)) \quad (3.19a)$$

$$\begin{bmatrix} \hat{d}_1(t) & \cdots & \hat{d}_n(t) \end{bmatrix}^T = \left\| \Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 \end{bmatrix}^T \right\| \cdot [I \ 0 \ \cdots \ 0] \Sigma(t) \text{sign}(\gamma(t)) \quad (3.19b)$$

où $\gamma(t) = \rho\left(\Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 \end{bmatrix}^T\right)$ et $\rho(X)$ désigne la première composante de X ayant la plus grande valeur absolue. Les estimés ainsi obtenus sont les seuls qui vérifient $\sum_{i=1}^n \hat{b}_i^2 = I$ et $\rho\left(\begin{bmatrix} \hat{b}_1(t) & \cdots & \hat{b}_n(t) \end{bmatrix}\right) > 0$.

Proposition 3.2. Soit $\{\hat{M}(t)\}$ la suite matricielle définie par (3.17). Soit $\begin{bmatrix} \hat{b}_1(t) & \cdots & \hat{b}_n(t) \end{bmatrix}^T$ et $\begin{bmatrix} \hat{d}_1(t) & \cdots & \hat{d}_n(t) \end{bmatrix}^T$ les vecteurs obtenus à partir de $\hat{M}(t)$ selon les expressions (3.19a-b). Si les estimés $\hat{\mu}_{ij}$ convergent vers leurs vraies valeurs μ_{ij} , alors $(\hat{b}_i(t), \hat{d}_j(t))$ convergent à leur tour vers leur vraies valeurs (b_i, d_j) . □

Preuve: Annexe B.1.

Etant donné les estimés $(\hat{b}_i(t), \hat{d}_j(t))$, les équations (3.2b), (3.7) et (3.9b) suggèrent, pour $B(q^{-1})$ et les $F(V_i)$ ($i=1, \dots, N$), les estimés suivants:

$$\hat{B}(q^{-1}) = \hat{b}_1(t)q^{-1} + \dots + \hat{b}_n(t)q^{-n} \quad (3.20a)$$

$$\hat{F}(V_i) = V_i \hat{d}_i(t) \quad \text{pour } i=1, \dots, N \quad (3.20b)$$

Remarque 3.3

- La décomposition (3.18) est l'une des particularités de ce schéma d'identification par rapport à celui présenté dans (Giri et al, 2001b).
- Il va sans dire, que la caractérisation de l'élément non linéaire $F(\cdot)$ par N points sera d'autant plus précise que le nombre N est plus élevé. Le choix de ce paramètre découlera donc d'un compromis entre, d'une part, la précision avec laquelle le nuage de points décrit l'élément $F(\cdot)$ et, d'autre part, la complexité de la méthode qui croît avec la valeur du paramètre.
- En l'absence de toute information sur la structure de l'élément $F(\cdot)$, on procédera à une approximation de ce dernier par un polynôme ou une fonction affine par morceaux. Cet aspect sera rediscuté dans la Section 7 où il sera établi pourquoi notre démarche, qui consiste à déterminer N points de $F(\cdot)$ puis opérer dessus une interpolation d'ordre N , est plus cohérente que celle, suivie dans la plupart des travaux antérieurs, qui consiste à substituer à $F(\cdot)$ un polynôme sans se préoccuper de la cohérence entre le choix du degré de ce polynôme et le type du signal excitant.

3.5. SIGNAL D'ENTREE PROCURANT UNE EXCITATION PERSISTANTE

Le choix de la séquence d'entrée $\{v(t)\}$ sera considéré comme étant approprié, si les trois conditions suivantes sont remplies :

- (i) $\{v(t)\}$ doit être facilement réalisable,
- (ii) elle doit prendre ses valeurs seulement dans l'ensemble $\{0, V_1, \dots, V_N\}$,
- (iii) elle doit procurer au vecteur de régression $\Phi(t)$ résultant la propriété d'excitation persistante au sens précisé au chapitre 2.

Le signal d'entrée, noté $\{v(t)\}$, proposé dans ce paragraphe s'inspire en partie du résultat général présenté dans le Lemme 2.2. Il est périodique de période $T=n(N+3)$ et est défini comme suit: pour tout entier k et tout instant t dans l'intervalle $[t_k, t_{k+1}[$ avec $t_k=kT$:

$$v(t) = \begin{cases} V_1 & \text{pour } t = t_k + 2n \\ V_j & \text{pour } t = t_k + (2 + j)n \quad ; j = 2, \dots, N \\ 0 & \text{ailleurs} \end{cases} \quad (3.21)$$

Le caractère périodique de ce signal le rend facilement réalisable. En plus, il prend ses valeurs dans l'ensemble $\{0, V_1, \dots, V_N\}$. Enfin, il sera établi dans la suite de ce paragraphe que ce signal confère bien la propriété d'excitation persistante au vecteur de régression $\Phi(t)$. A cet effet, nous suivrons de très près l'analyse développée dans (Chaoui et al , 2005).

Le point de départ est d'opérer la partition suivante sur le vecteur $\Phi(t)$:

$$\Phi(t) = [\phi_1^T(t) \quad \phi_2^T(t) \quad \dots \quad \phi_N^T(t)]^T \quad (3.22a)$$

où:

$$\phi_1^T(t) = [-y(t-1) \quad \dots \quad -y(t-n) \quad v(t-1)P_1(v(t-1)) \quad \dots \quad v(t-n)P_1(v(t-n))] \quad (3.22b)$$

$$\phi_j^T(t) = [v(t-1)P_j(v(t-1)) \quad \dots \quad v(t-n)P_j(v(t-n))] \quad \dots \text{pour } j = 2, \dots, N \quad (3.22c)$$

Introduisons également le vecteur, noté $\Phi'(t)$, défini comme étant la version non perturbée de $\Phi(t)$:

$$\Phi'(t) = [\phi_1'^T(t) \quad \phi_2'^T(t) \quad \dots \quad \phi_N'^T(t)]^T \quad (3.22d)$$

avec :

$$\phi_1'^T(t) = [-x(t-1) \quad \dots \quad -x(t-n) \quad v(t-1)P_1(v(t-1)) \quad \dots \quad v(t-n)P_1(v(t-n))] \quad (3.22e)$$

et

$$x(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t) = y(t) - z(t) \quad (3.22f)$$

$x(t)$ étant la sortie non perturbée du système. Il découle immédiatement de (3.22a-f) que :

$$\phi_j(t) = \phi'_j(t) + \tilde{\phi}(t) \quad \text{avec} \quad \tilde{\phi}^T(t) = [-z(t-1) \quad \dots \quad -z(t-n) \quad 0 \quad \dots \quad 0] \quad (3.23a)$$

et, de ce fait :

$$\Phi(t) = \Phi'(t) + \tilde{\Phi}(t) \quad \text{avec} \quad \tilde{\Phi}^T(t) = [\phi^T(t) \quad 0 \quad \dots \quad 0] \quad (3.23b)$$

Les deux propositions suivantes stipulent respectivement que les vecteurs $\phi'_j(t)$ et $\phi_j(t)$ ($j=2, \dots, N$) possèdent effectivement la propriété d'EP.

Proposition 3.3. (L'excitation persistante de $\phi'_j(t)$).

Il existe un réel positif ε_0 , tel que pour tout k :

$$\sum_{\tau=1}^{4n} \phi'_j(t_k + \tau) \phi'_j(t_k + \tau)^T > \varepsilon_0 I_{2n} \quad (3.24)$$

où I_{2n} désigne la matrice identité de dimension $2n \times 2n$ □

Preuve: Ce résultat est en fait un cas particulier du Lemme 2.2 (Chapitre 2) ; la preuve détaillée est développée dans l'Annexe B.2.

Proposition 3.4. (L'excitation persistante des $\phi_j(t)$, $j=2, \dots, N$). Pour tout k , on a :

$$\sum_{\tau=1}^n \phi_j(t_k + n(j+2) + \tau) \phi_j(t_k + n(j+2) + \tau)^T \geq \varepsilon_j \cdot I_n \quad \text{pour } j = 2, \dots, N \quad (3.25)$$

où $\varepsilon_j = \inf_{j=2, \dots, N} (V_j)^2$ et I_n est la matrice identité de dimension $n \times n$ □

Preuve: Voir Annexe B.3.

Le fait que le vecteur de régression global, $\Phi'(t)$, soit la concaténation des sous-vecteurs $\phi'_j(t)$ et $\phi_j(t)$ ($j=2, \dots, N$), tous dotés de l'EP, et que ces derniers n'opèrent jamais tous en même temps, va faire que $\Phi'(t)$ est à son tour doté de l'EP. C'est ce qui fait l'objet de la proposition suivante:

Proposition 3.5 (L'excitation persistante de $\Phi(t)$).

- 1) Le vecteur de régression $\Phi'(t)$ est excitant persistant ; ainsi il existe une constante réelle $\varepsilon > 0$ telle que, pour tout k et tout vecteur unitaire W :

$$\max_{1 \leq \tau \leq (N+3)_n} \left| \Phi'(t_k + \tau)^T W \right| > \varepsilon \quad (3.26a)$$

- 2) Par conséquent, il existe une constante réelle $K_1 > 0$, indépendante de μ , et une séquence réelle positive $\{\delta_0(k)\} \in S(K_1, \mu)$ telle que, pour tout k et tout vecteur unitaire W :

$$\max_{1 \leq \tau \leq (N+3)_n} \left| \Phi(t_k + \tau)^T W \right| > \varepsilon - \delta_0(k) \quad (3.26b)$$

où le réel $\varepsilon > 0$ est proportionnel à $\min\{V_1^2, \dots, V_N^2\}$ □

- 3) Dans le cas idéal ($\{z(t)\} = \{0\}$), la séquence $\{\delta_0(k)\}$ dans l'expression précédente devient identiquement nulle □

Preuve: se reporter à l'Annexe B.4.

Remarque 3.4.

- 1) La seconde partie de la proposition précédente montre que le vecteur de régression $\Phi(t)$ possède bien la propriété d'EP à une erreur près de magnitude moyenne proportionnelle à μ (la taille moyenne de l'erreur de modèle). La qualité d'excitation procurée à $\Phi(t)$ est donc d'autant meilleure que μ est plus faible.
- 2) Dans le cas idéal ($\{z(t)\} = \{0\}$), $\Phi(t)$ se confond avec $\Phi'(t)$ et se trouve, de ce fait, doté à son tour de l'EP au vu de la Proposition 3.5 (partie 1).

Au paragraphe suivant, nous allons voir quelles conséquences les propriétés d'excitation ainsi établies vont avoir sur la consistance de l'estimé $\hat{\Theta}(t)$ généré par l'algorithme (3.15).

3.6. ANALYSE DE CONVERGENCE DES PARAMETRES ESTIMES

Proposition 3.6. (Convergence du vecteur $\hat{\Theta}(t)$).

- 1) Dans le cas général ($\{z(t)\}$ quelconque mais bornée), il existe une constante réelle K_2 , indépendante de μ , telle que:

$$\left\{ \left\| \hat{\Theta}(t) - \Theta^* \right\| \right\} \in \mathcal{S}(K_2, \mu) \quad (3.27)$$

2) Dans le cas idéal ($\{z(t)\} = \{0\}$):

$$\left\{ \left\| \hat{\Theta}(t) - \Theta^* \right\| \right\} \text{ converge vers zéro, exponentiellement.} \quad (3.28)$$

Dans ce cas, la convergence est exponentielle et est d'autant plus rapide que $\min\{V_1^2, \dots, V_N^2\}$ est plus grand □

Preuve: La preuve de la partie 2 est standard et peut être trouvée dans plusieurs références ((Goodwin et Sin, 1983), (Lozano et Taoutaou, 2001)). La première partie de la proposition est elle donnée en Annexe B.5.

Cette proposition montre que la qualité des estimés dépend à son tour de la taille moyenne, mesurée par μ , de l'erreur de modélisation $\{\eta(t)\}$. Dans le cas idéal, cette erreur est nulle ce qui, compte tenu du fait que $\Phi(t)$ est alors dotée de l'EP, implique la consistance de $\hat{\Theta}(t)$. Ces résultats se transposent immédiatement sur les estimés $(\hat{b}_i(t), \hat{d}_i(t))$ du fait de la continuité de l'estimateur (3.19a-b).

3.7. APPROXIMATION DE L'ÉLÉMENT NON LINÉAIRE

Le schéma d'identification présenté précédemment tente de déterminer N points de l'élément non linéaire $F(\cdot)$. Cette détermination est exacte en l'absence d'erreur de modélisation ; autrement elle est approximative. Moyennant une interpolation appropriée de ce nuage de points, il est possible d'arriver à une description analytique ou graphique complète de l'élément non linéaire. Il est évident que la détermination d'une telle description serait facilitée si des informations sur la structure de l'élément non linéaire étaient disponibles.

Intuitivement, la qualité de l'interpolation dépendra, d'une part, de la taille N du nuage de points et de sa dispersion et, d'autre part, de la technique d'interpolation appliquée. Le paramètre N peut à priori être choisi aussi grand que l'on veut ; seule la complexité du schéma d'identification en limite la valeur. Si aucune connaissance n'est disponible à priori sur le type de la non-linéarité, l'emplacement des points, déterminé par les V_i ($i=, \dots, N$), est quelconque. Quant à l'interpolation,

elle peut être abordée de plusieurs façons possibles. Dans le présent contexte, on peut envisager les approches suivantes :

- a) Interpolation affine par morceaux : convient à une fonction $F(.)$ qui est non lisse et/ou discontinue (zone-morte, seuil, ...).
- b) Interpolation par des splines: convient à une fonction $F(.)$ quelconque lisse et continue.
- c) Interpolation polynomiale: convient si $F(.)$ est de type polynomial.

La question est quel type d'interpolation choisir quand aucune connaissance n'est a priori disponible sur $F(.)$?

La question peut être abordée de la manière suivante: pour chaque type d'interpolation, on procède à la validation du modèle obtenu $\hat{F}(.), \hat{A}(q^{-1}), \hat{B}(q^{-1})$ en excitant le système par un signal (différent de celui utilisé pour l'identification) balayant toute la plage utile de l'entrée. Le choix se portera alors sur le type d'interpolation qui donnera lieu à la meilleure validation du modèle au sens prédictif par exemple.

3.8. SIMULATION

Dans ce paragraphe, on rapporte les résultats de simulations effectuées dans un but d'illustrer l'efficacité du schéma d'identification présenté dans ce chapitre.

On considère le système décrit par le sous-système linéaire suivant:

$$A(q^{-1}) = 1 - 1.5 q^{-1} + 0.56 q^{-2};$$

$$B(q^{-1}) = q^{-1} - 2 q^{-2};$$

et l'élément statique non linéaire de type non polynomial

$$F(v) = \frac{2v}{1 + v^2}$$

En vue de la condition (3.13), les paramètres à identifier sont, pour la partie linéaire sont:

$$a_1 = 0.56, \quad a_2 = -1.5, \quad b_1 = -1/\sqrt{5} = -0.447, \quad b_2 = 2/\sqrt{5} = 0.894$$

Pour le gain non linéaire, et pour ne pas trop encombrer les figures, on prend $N=8$; Donc seulement 8 points seront identifiés sur la caractéristique $F(v)$. En réalité rien ne nous empêche de prendre N plus grand. Les points choisis sont:

V_j	-4	-3	-2	-1	1	2	3	4
$F(V_j)$	-0.47	-0.6	-0.8	-1	1	0.8	0.6	0.47

En vue de (3.6), (3.8) et (3.13) les d_j à identifier sont alors :

d_j	-0.263	-0.447	-0.894	-2.236	-2.236	-0.894	-0.447	-0.263
-------	--------	--------	--------	--------	--------	--------	--------	--------

Remarquons ici que les N points sont identifiés sur la caractéristique $-\sqrt{5} * F(v)$ et non pas sur la caractéristique $F(v)$.

La séquence d'entrée est générée selon la loi (3.21) en utilisant, bien sûr, les valeurs V_j qu'on a choisies.

Dans un premier temps nous allons supposer que la perturbation extérieure $z(t)$ est nulle, pour montrer la convergence des paramètres vers leurs vraies valeurs, dans le cas idéal.

La fig.3.2 montre la convergence des paramètres du sous système linéaire. Alors que la fig.3.3 montre le tracé de la caractéristique $-\sqrt{5} * F(v)$ à estimer et les N points estimés (en signe +). Comme le montre ces deux figures, la convergence est rapide et parfaite vers les vraies valeurs, et les N points se trouvent exactement sur la caractéristique à identifier. Ce ci confirme l'efficacité de l'algorithme (3.14a-b) proposé, en l'absence de perturbation.

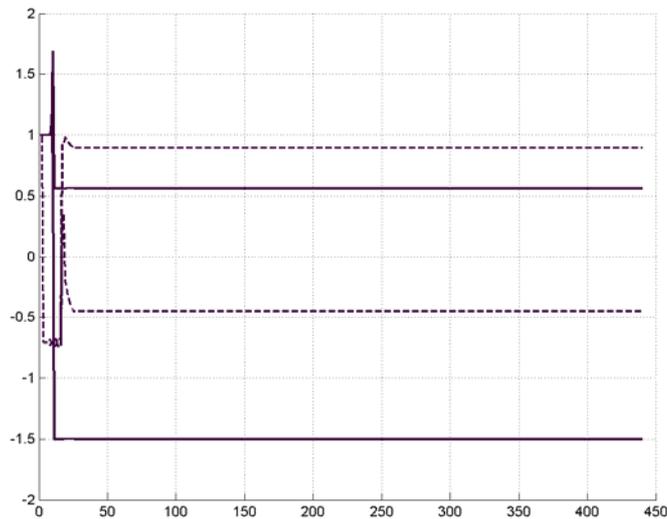
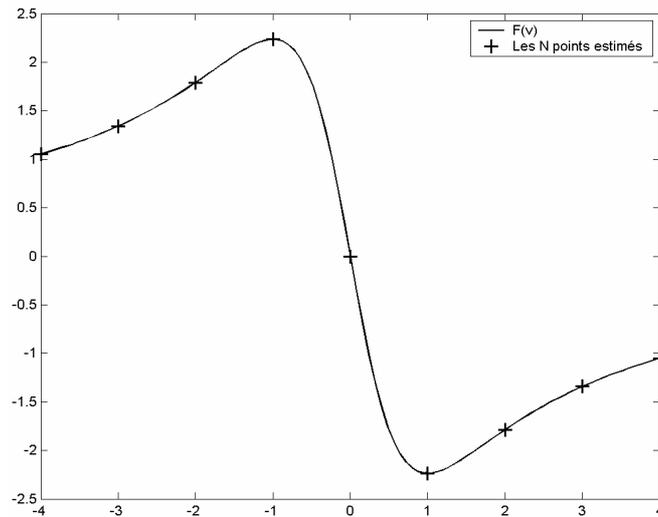


Figure 3.2 Convergence des paramètres du sous système linéaire (a_i et b_i).

Figure 3.3 Les N points estimés sur la caractéristique $F(v)$

Après avoir obtenu les N points, on peut procéder à une caractérisation globale du gain non linéaire, par interpolation. Comme on l'a mentionné dans le paragraphe 7, plusieurs types d'interpolations sont possibles. La fig. 3.4 montre les caractéristiques obtenues par deux types d'interpolations (polynomiales et par splines cubiques).

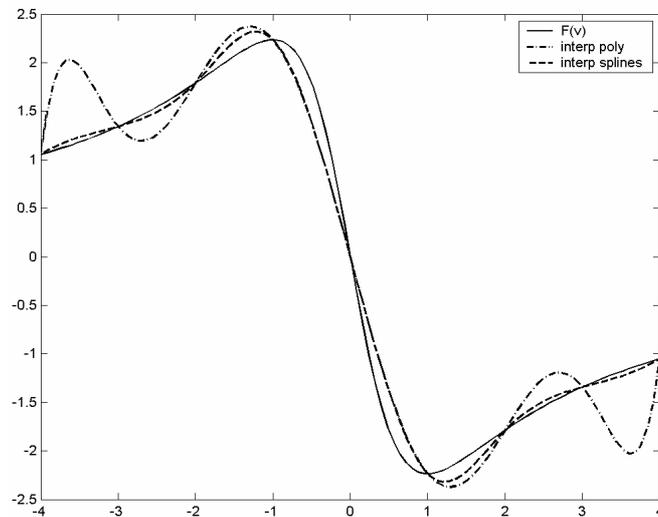


Figure 3.4 Caractéristiques obtenues par deux types d'interpolation

Si on procède à une validation des deux modèles du gain non linéaire, obtenus avec les deux types d'interpolations, on observe que les deux modèles donnent la même erreur de modélisation entre $[-1, 1]$, alors qu'au-delà de cette intervalle, l'interpolation par splines cubiques donne, pour la non linéarité considérée, une erreur plus petite que l'interpolation polynomiale. Ce ci rend l'interpolation

par splines cubiques préférable à l'interpolation polynomiale, pour ce cas. Ce ci illustre, de manière assez simple, les propos de la section 7.

Nous allons maintenant étudier l'influence d'un bruit externe sur la qualité des estimés obtenus par l'algorithme proposé. On ajoute à la sortie un bruit externe de type:

$$z(t) = \frac{b(t)}{H(q^{-1})}$$

où $b(t)$ est une séquence de nombres aléatoires dans l'intervalle $[-\sigma, +\sigma]$ et $H(q^{-1}) = 1 - 1.7q^{-1} + 0.72q^{-2}$.

1^{er} cas: $\sigma=0.1$

La fig.3.5 montre la sortie non perturbée du système et le bruit externe ajouté, et illustre le rapport signal/bruit.

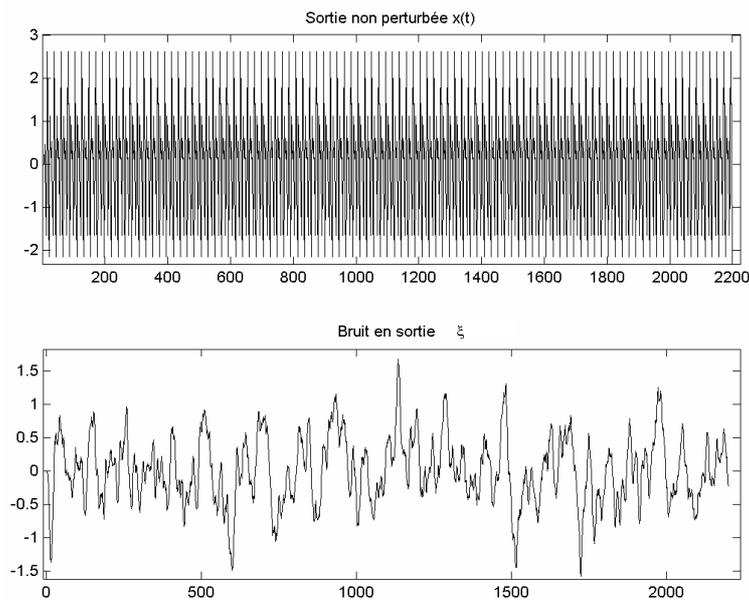


Figure 3.5 Sortie non perturbée du système et bruit en sortie

La fig.3.6 montre la convergence des paramètres du sous système linéaire. Les estimés obtenus sont $\hat{a}_1 = -1.545$, $\hat{a}_2 = 0.574$, $\hat{b}_1 = -0.418$, $\hat{b}_2 = 0.908$. Ils sont légèrement différents des vraies valeurs. La fig.3.7 montre les N points estimés par rapport à la caractéristique du gain non linéaire à identifier. On voit aussi que les points estimés sont légèrement décalés para rapport aux vraies valeurs.

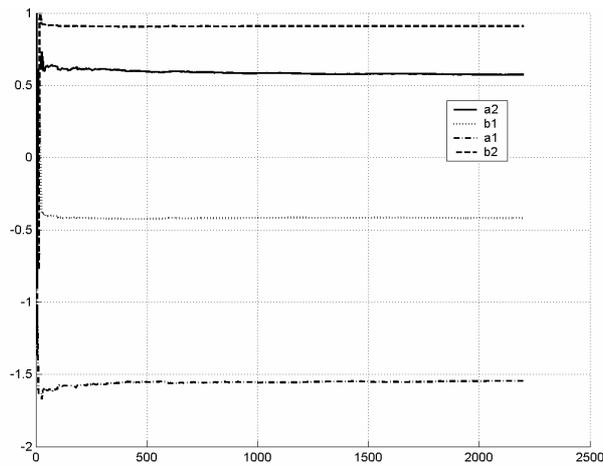


Figure 3.6 Convergence des estimés des paramètres du sous système linéaire

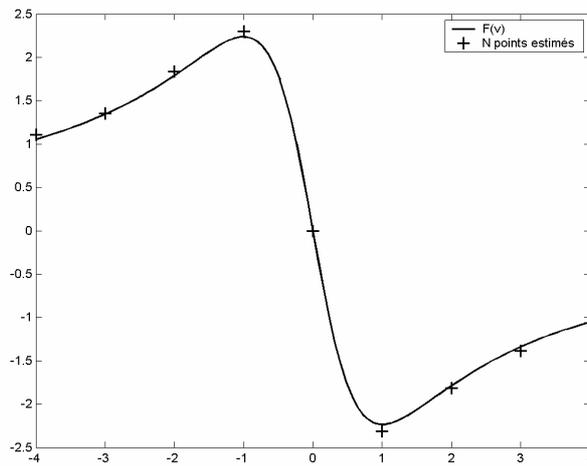


Figure 3.7 les N points estimés sur la caractéristique $F(v)$

2^{ème} cas: $\sigma=0.3$

Les estimés obtenus pour les paramètres du sous-système linéaire sont $\hat{a}_1 = -1.6$, $\hat{a}_2 = 0.623$, $\hat{b}_1 = -0.417$, $\hat{b}_2 = 0.908$. Alors que les N points estimés sont représentés sur la fig.3.8

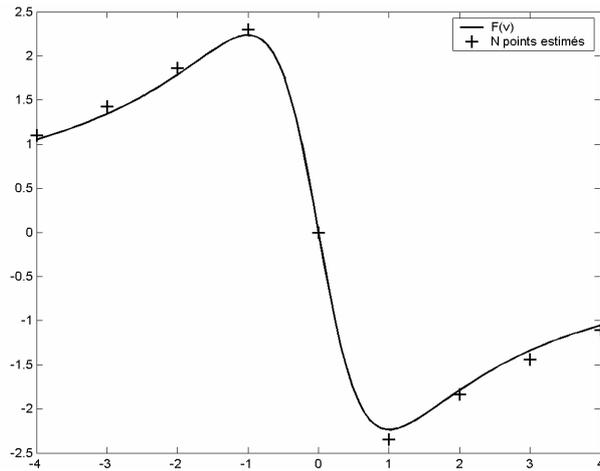


Figure 3.8 les N points estimés sur la caractéristique $F(v)$

A partir des résultats de simulation obtenus on remarque que plus la perturbation externe est grande en moyenne plus la qualité des estimés obtenus est médiocre. Ce ci confirme les résultats théoriques déjà démontrés. Toutefois, pour des rapports signal/bruit acceptables, l'algorithme d'identification donne des résultats satisfaisants, en garantissant la convergence des estimés vers un voisinage proche de leurs vraies valeurs.

3.9. CONCLUSIONS

Ce chapitre a porté sur l'identification des systèmes non linéaires sur la base du modèle de Hammerstein. Les seules hypothèses faites sur le gain non linéaire $F(\cdot)$ sont les suivantes:

- i) $F(\cdot)$ n'est pas identiquement nulle.
- ii) pour toute valeur finie de v , $F(v)$ existe, est finie et, en particulier, $F(0)=0$.

Le schéma d'identification élaboré permet d'estimer les paramètres a_i et b_i ($i=1, \dots, n$) du sous-système dynamique linéaire ($B(q^{-1})/A(q^{-1})$) ainsi que N points $(V_j, F(V_j))$ de l'élément non linéaire, où N et les V_j 's sont arbitrairement choisis par l'utilisateur. Les principaux atouts de ce schéma, par rapport aux travaux antérieurs de même type, sont rappelés dans ce qui suit.

Les travaux antérieurs font l'hypothèse que l'élément non linéaire est un polynôme de degré connu, ce qui n'est pas le cas dans ce chapitre. On pourrait reprocher au présent travail d'avoir recouru à la substitution d'un polynôme $P(v)$, en l'occurrence celui défini par (3.8), à la non-linéarité $F(v)$ afin d'obtenir la régression linéaire (3.11a) à partir du modèle initial (3.1a-b). Nous rétorquons à cette critique en remarquant que cette substitution est toujours possible à condition d'astreindre l'argument v à prendre au plus $N+1$ valeurs différentes. Cette condition a été respectée dans ce chapitre en prenant un signal d'entrée prenant ses valeurs dans l'ensemble $\{0, V_1, \dots, V_N\}$. En revanche, les travaux antérieurs ne se privaient pas de appel recours à des signaux prenant un très grand nombre, voire un nombre infini, de valeurs différentes. Dans ces conditions, la substitution d'un polynôme $P(v) = \sum_{j=1}^m c_j v^j$ à la non-linéarité initiale $F(v)$ conduira forcément à une erreur de modélisation qui empêchera les paramètres estimés de converger vers leurs vraies valeurs et ce même en l'absence de perturbations. Ainsi, les résultats antérieurs ne sont vraiment valides que si l'élément non linéaire est un polynôme de degré fini.

La propriété d'excitation persistante a rarement été garantie dans les travaux antérieurs et quand elle l'a été c'est à l'aide de signaux difficilement réalisables pratiquement. C'est le cas par exemple dans (Stoica et Söderström, 1982) où l'on a fait appel, pour ce faire, à un signal d'entrée aléatoire de type bruit blanc. Dans le présent travail, cet objectif a été atteint moyennant un signal d'entrée périodique déterministe et facilement réalisable.

Le schéma présenté dans ce chapitre ne suppose aucune connaissance structurelle sur l'élément non linéaire $F(v)$ et aucune hypothèse sur la nature de la perturbation. Quand on connaît la structure de la non-linéarité et la nature de la perturbation, il sera possible, en approchant le problème différemment, de déterminer exactement l'élément non linéaire. Les deux prochains chapitres vont illustrer cette possibilité pour deux types de non-linéarités dont justement le cas $F(v) = \sum_{j=1}^m c_j v^j$ mentionné ci-dessus.

CHAPITRE 4

IDENTIFICATION DES SYSTEMES DE HAMMERSTEIN IMPLIQUANT UN ELEMENT NON LINEAIRE STATIQUE DE STRUCTURE CONNUE

4.1. INTRODUCTION

Dans ce chapitre, le problème d'identification des systèmes de Hammerstein est abordé en supposant que l'élément non linéaire sous-jacent possède une structure connue à l'avance. Cette hypothèse est effectivement vérifiée dans certaines situations pratiques. Il en est ainsi quand le système physique considéré comprend des composants placés à l'entrée, en l'occurrence des actionneurs, dont on connaît la forme de la caractéristique statique. C'est également le cas dans certains problèmes de commande où il s'avère utile, voire nécessaire, de placer une non-linéarité (logicielle) à l'entrée du système commandé afin de se conformer à certaines clauses du cahier des charges (par exemple, limiter en amplitude ou en vitesse le signal de commande envoyé à l'entrée du système physique).

Deux familles d'éléments non linéaires statiques (sans-mémoire) seront considérées dans ce chapitre. La première comprend tout élément continu dont le comportement peut être décrit à l'aide d'un polynôme de degré connu. De très nombreux travaux antérieurs furent précisément consacrés à l'identification de systèmes faisant intervenir ce type de non-linéarités. Ce choix s'appuie sur le postulat que l'élément non linéaire $F(.)$ puisse, dans le domaine limité de variation de l'entrée, faire l'objet d'un développement en série d'ordre fini. Les développements polynomiaux ont souvent été privilégiés mais d'autres types de développements ont été utilisés, par exemple le développement en série de Fourier ((Krzyzak, 1989)). La représentation ainsi obtenue de l'élément non linéaire s'avère continue par rapport à son argument (l'entrée de commande) et dépend linéairement des paramètres inconnus (coefficients du développement). Ces caractéristiques simplifient relativement le problème d'identification.

La seconde classe de non-linéarités considérées dans ce chapitre est constituée d'éléments de type seuil avec zone-morte. Les travaux d'identification consacrés aux systèmes caractérisés par des non-linéarités affines par morceaux et discontinues restent à ce jour relativement rares, ((Bai, 2002), (Vörös, 1999a), (Taware et Tao, 2003)). La difficulté du problème d'identification réside essentiellement dans le caractère discontinu de ces non-linéarités qui fait que celles-ci dépendent de leurs paramètres inconnus de façon non linéaire. Paradoxalement, le problème de commande de systèmes impliquant des non linéarités affines par morceaux est devenu depuis quelques années un champ de recherche très actif (voir par exemple (Taware et Tao, 2003) et la liste des références qui y figure). Concernant le problème d'identification, (Vörös, 1999a) est l'un des premiers travaux où l'on considère un élément non linéaire de type seuil constitué de deux demi-droites. Le processus d'identification proposé s'appuie sur une régression pseudo linéaire en les paramètres inconnus du système. Pour faire face à cette nature pseudo linéaire, l'auteur propose deux estimateurs opérant en alternance, chaque estimateur porte sur une partie des paramètres en supposant que les autres paramètres sont connus. Comme ceux-ci ne le sont pas, ils sont remplacés par leurs estimés. L'efficacité du schéma d'identification n'a été illustrée que par voie de simulation ; aucune analyse formelle de consistance n'a été faite. Dans (Bai, 2002), l'accent a été mis sur une large classe de non linéarités (dont ceux de type seuil avec zone-morte). Le problème d'identification a été abordé sous une hypothèse de taille à savoir que l'élément non linéaire est parfaitement symétrique et, de ce fait, se trouve caractérisé par un seul paramètre inconnu. L'auteur fait alors appel à l'approche non linéaire des moindres carrés séparés pour recouvrer les paramètres du modèle. Seulement cette approche implique la minimisation d'une fonction non linéaire et non convexe, ce qui a amené l'auteur à suggérer une recherche graphique du minimum global (qui n'est autre que la vraie valeur du paramètre caractérisant la non-linéarité). Pour qu'un tel procédé conduise au minimum global il faut focaliser la recherche sur un intervalle comprenant ce minimum ce qui revient, de ce fait, à supposer qu'un voisinage de ce minimum est connu a priori. Quand l'élément non linéaire est asymétrique, il se trouve caractérisé par plusieurs paramètres. En théorie, la solution précédente peut être envisagée, mais il conduit à la minimisation d'une fonction non linéaire, non convexe et multivariable. L'auteur de la solution reconnaît lui-même que, dans ce cas, la méthode graphique proposée trouve bien sa limite et ce même en supposant connu un voisinage du minimum. Il propose alors une solution alternative exigeant une multitude d'expériences avec différents signaux de commande ; ces derniers étant supposés de type bruit blanc. Le procédé d'estimation des paramètres inconnus implique l'usage multiple d'algorithmes de type moindres carrés mais aussi des techniques de corrélation de signaux. L'auteur lui-même admet que la qualité de l'estimation des fonctions de corrélation est tributaire de la longueur des données recueillies. Il s'ensuit que le

problème de compromis entre la qualité des résultats et la complexité de la méthode se pose de façon cruciale dans cette solution. Enfin, il n'est pas vraiment établi que les estimateurs de paramètres proposés opèrent dans un contexte d'excitation persistante, ni que l'estimation est consistante.

Le schéma d'identification présenté dans ce chapitre est relativement plus simple dans la mesure où il exige une seule expérience et fait vraiment l'objet d'une analyse rigoureuse de consistance. Il s'inspire du schéma développé au Chapitre 3 et exploite judicieusement les résultats de consistance élaborés dans ce dernier. A cet effet, le problème d'identification est reformulé de telle sorte qu'il porte sur les paramètres du sous-système dynamique linéaire et sur un nuage de points de la caractéristique graphique de l'élément non linéaire. La taille de ce nuage et sa répartition dépendront de la nature de cet élément. L'ensemble de ces inconnues est estimé simultanément à l'aide d'un algorithme de type moindres carrés opérant dans un contexte d'excitation persistante. Les paramètres caractérisant la non-linéarité sont déduits, en temps réel, du nuage de points ainsi obtenu.

Le chapitre est organisé comme suit : la classe des systèmes à identifier est décrite à la section 2, le schéma d'identification est présenté et analysé dans les sections 3 à 6 et est illustré par voie de simulation dans la section 7.

4.2. POSITION DU PROBLEME D'IDENTIFICATION

4.2.1. Classe de systèmes à identifier

On s'intéresse aux systèmes dont le comportement entrée/sortie peut être décrit par un modèle de la forme :

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})} u(t) + z(t) \quad (4.1)$$

$$u(t) = F(v(t)) \quad (4.2)$$

avec

$$A(q^{-1}) = 1 + a_1 q^{-1} + \dots + a_{na} q^{-na} \quad (4.3a)$$

$$B(q^{-1}) = b_1 q^{-1} + \dots + b_{nb} q^{-nb} \quad (4.3b)$$

où $v(t)$ et $y(t)$ désignent respectivement l'entrée et la sortie; $u(t)$ est un signal interne *non mesurable* ; $z(t)$ est une erreur de modélisation due aux perturbations externes (et éventuellement

aux dynamiques négligées). On suppose que $z(t)$ est bornée. $F(.)$ est une fonction à structure connue représentant le gain statique. Deux cas de figures seront considérés :

1^{er} cas: $F(.)$ est une fonction polynomiale de la forme :

$$F(v) = \sum_{i=1}^m c_i v^i \tag{4.4}$$

Le degré m est supposé connu mais les paramètres c_i ne le sont pas. Pour une raison de commodité, la non-linéarité (4.4) sera notée $\mathcal{P}(c_1, \dots, c_m)$. Il est possible que $c_1 = c_2 = \dots = c_p = 0$ et $c_m = c_{m-1} = \dots = c_q = 0$ pour certains entiers p et q ($p < q$). Il s'ensuit que la connaissance précise de la structure du polynôme n'est pas exigée. La connaissance d'un majorant du degré de ce polynôme est suffisante.

2^{ème} cas: $F(.)$ est une fonction affine par morceaux dont la forme est définie graphiquement par la Figure 4.1. Il s'agit donc d'une non-linéarité de type zone-morte et/ou seuil et caractérisée par six paramètres, en l'occurrence $D_1, D_2, S_1, S_2, L_1, L_2$ supposés tous inconnus. On convient de désigner un tel élément par la notation $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$. Parmi les cas particuliers importants recouverts par cette description, on peut citer les éléments de type zone-morte ($L_1=L_2=0$) et les seuils ($D_1=D_2=0$).

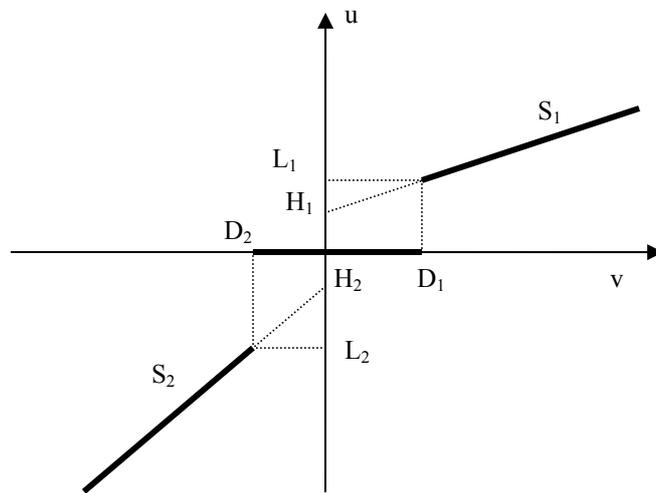


Fig 4.1. Non-linéarité $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$

La dynamique du procédé est elle soumise aux hypothèses suivantes :

- H4.1.** Il existe n tel que: $n \geq \max \{na, nb\}$.
- H4.2.** $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux.
- H4.3.** Tous les zéros de $q^{na}A(q^{-1})$ sont à l'intérieur du cercle unité.

Ces hypothèses sont semblables aux hypothèses H3.1-H3.3 du Chapitre 3. Par conséquent, les remarques 3.1 de ce dernier sont valables ici.

4.2.2. Objectif d'identification

Le but est de concevoir un schéma d'identification qui soit capable, du moins dans le cas idéal, d'estimer exactement le modèle du sous-système dynamique $(B(q^{-1})/A(q^{-1}))$ et les paramètres caractérisant le gain non linéaire, à savoir les réels c_i ($i=1, \dots, m$) dans le cas où $F=\mathcal{P}(c_1, \dots, c_m)$ et les réels $D_1, D_2, S1, S2, L_1, L_2$ dans le cas où $F=\mathcal{ZMS}(D_1, D_2, S1, S2, L_1, L_2)$.

Le schéma d'identification proposé dans ce qui suit sera élaboré en deux étapes. Dans la première étape, on se fixera comme objectif l'estimation des paramètres de la dynamique du système $B(q^{-1})/A(q^{-1})$ et un ensemble de points $(V_i, F(V_i))$ ($i=1, \dots, N$) de la caractéristique $F(\cdot)$ du gain non linéaire ; le choix de l'entier N sera discuté ultérieurement. A cet effet, on suivra la démarche développée au paragraphe 3.3 du chapitre précédent. Ainsi, nous commencerons par la mise en évidence d'un polynôme auxiliaire, noté $P(\cdot)$, tel que $P(V_i)=F(V_i)$ ($i=1, \dots, N$). Ce polynôme pourra alors se substituer à $F(\cdot)$ dans le modèle du système sans que cela donne lieu à une quelconque erreur à condition d'astreindre l'entrée du système à prendre ses valeurs dans l'ensemble $\{V_1, \dots, V_N\}$. Cette substitution conduira à une reparamétrisation du modèle impliquant bi-linéairement les paramètres du sous-système dynamique $B(q^{-1})/A(q^{-1})$, d'une part, et ceux du polynôme $P(\cdot)$, d'autre part. L'estimation portera alors sur les paramètres composés, qui eux interviennent linéairement dans la représentation précédente, en utilisant à cet effet un algorithme de type moindres carrés.

La seconde étape dans la synthèse du schéma d'identification proposé consiste à retrouver les paramètres qui font réellement l'objet du problème d'identification à savoir ceux du sous-système dynamique et ceux des non-linéarités considérées $\mathcal{P}(c_1, \dots, c_m)$ et $\mathcal{ZMS}(D_1, D_2, S1, S2, L_1, L_2)$

4.3. REPARAMETRISATION DU SYSTEME

Dans un premier temps notre objectif sera de déterminer les paramètres du sous-système dynamique $B(q^{-1})/A(q^{-1})$ et N couples $(V_j, F(V_j))$ ($j=1, \dots, N$). Le choix de l'entier N et des réels V_j dépendra du type de l'élément non linéaire, il sera discuté plus tard au paragraphe 6. Pour l'instant, on soumet ce choix aux seules conditions :

$$V_1 < V_2 < \dots < V_N \quad (4.5a)$$

$$V_j \neq 0 \quad (j = 1, \dots, N) \quad (4.5b)$$

Les N couples $(V_j, F(V_j))$ déterminent un polynôme unique P de degré N tel que:

$$P(0) = 0 \quad (4.6a)$$

$$P(V_j) = F(V_j) \quad \text{pour } j = 1, \dots, N \quad (4.6b)$$

Ce polynôme peut toujours être écrit sous la forme :

$$P(v) = v \sum_{j=1}^N d_j P_j(v) \quad \text{avec} \quad P_j(v) = \prod_{\substack{i=1 \\ i \neq j}}^N \frac{v - V_i}{V_j - V_i} \quad (4.7)$$

Il s'ensuit que :

$$P_j(V_i) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad \text{et} \quad d_j = \frac{P(V_j)}{V_j} \quad (4.8)$$

En outre, si l'on astreint (lors du processus d'identification) l'entrée du système à prendre ses valeurs seulement dans l'ensemble $\{0, V_1, \dots, V_N\}$ alors, au vu de (4.6b), P peut se substituer à F dans le modèle (4.1)-(4.2), sans que cela ne génère la moindre erreur. Il s'ensuit que:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + A(q^{-1})z(t) \quad (4.9a)$$

$$u(t) = P(v(t)) \quad (4.9b)$$

Le problème d'identification porte désormais sur les paramètres a_i, b_i et d_j et, de ce fait, s'avère identique à celui déjà étudié au chapitre précédent. Il peut donc être appréhendé en suivant rigoureusement la même démarche (voir en particulier paragraphe 3.3 à partir des équations (3.10a-b) et les paragraphes 3.4 et 3.5). Par commodité nous reproduisons l'essentiel de cette solution dans le présent chapitre. Cette façon de faire a le mérite de permettre au présent chapitre de se lire, en grande partie, indépendamment du précédent.

Ainsi, les équations (4.9a-b) et (4.7) impliquent-elles successivement que :

$$\begin{aligned} A(q^{-1})y(t) &= B(q^{-1})P(v(t)) + A(q^{-1})z(t) \\ &= \sum_{i=1}^n b_i q^{-i} \sum_{j=1}^N d_j P_j(v(t)) v(t) + \eta(t) \\ &= \sum_{i=1}^n \sum_{j=1}^N \mu_{ij} v(t-i) P_j(v(t-i)) + \eta(t) \end{aligned} \quad (4.10a)$$

avec :

$$\eta(t) = A(q^{-1})z(t) \quad \text{et} \quad \mu_{ij} = b_i d_j \quad (i = 1, \dots, n; \quad j = 1, \dots, N) \quad (4.10b)$$

L'équation (4.10a) met en évidence des paramètres composés, en l'occurrence les μ_{ij} . Son intérêt réside dans le fait qu'elle implique linéairement les paramètres inconnus (a_i et μ_{ij}) et peut de ce fait servir de base à l'estimation de ceux-ci. A cet effet, on commence par la mettre sous la forme de régression suivante:

$$y(t) = \Phi(t)^T \Theta^* + \eta(t) \quad (4.11a)$$

avec :

$$\Phi(t)^T = \begin{bmatrix} -y(t-1) & -y(t-n) & v(t-1)P_1(t-1) & v(t-n)P_1(t-1) & \cdots \\ \cdots & v(t-1)P_N(v(t-1)) & \cdots & v(t-1)P_N(v(t-1)) \end{bmatrix} \quad (4.11b)$$

$$\Theta^* = [a_1 \cdots a_n \quad \mu_{11} \cdots \mu_{n1} \cdots \cdots \mu_{1N} \cdots \mu_{nN}]^T \quad (4.11c)$$

L'estimation de Θ^* fera l'objet du paragraphe suivant. Auparavant, arrêtons-nous sur le problème de passage des paramètres composés μ_{ij} vers les paramètres qui nous intéressent réellement, à savoir les (b_i, d_j) . La procédure élaborée au Chapitre 2 (Lemme 2.8) et utilisée déjà au Chapitre 3, exploite la relation suivante :

$$M \stackrel{def}{=} \begin{bmatrix} \mu_{11} & \cdots & \mu_{1N} \\ \vdots & & \vdots \\ \mu_{n1} & \cdots & \mu_{nN} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [d_1 \quad \cdots \quad d_n] \quad (4.12)$$

Cette équation ne détermine pas de façon unique les deux vecteurs contenant les b_i et les d_j .

L'unicité peut être forcée si l'on impose à ceux-ci des conditions supplémentaires. Par exemple :

$$\sum_{i=1}^n b_i^2 = 1 \quad \text{et} \quad \rho([b_1 \dots b_n]) > 0 \quad (4.13)$$

où $\rho([b_1 \dots b_n]) > 0$ représente la première composante du vecteur $[b_1 \dots b_n]^T$ ayant la plus grande valeur absolue. Dans ce cas, l'unique solution de (4.13) est complètement décrite par le Lemme 2.8 du chapitre 2. Avec les notations de ce lemme, la solution est donnée par :

$$[b_1 \quad \cdots \quad b_n]^T = \frac{\Gamma[\sigma_1 \quad 0 \quad \cdots \quad 0]^T}{\|\Gamma[\sigma_1 \quad 0 \quad \cdots \quad 0]^T\|} \text{sign}(\gamma) \quad (4.14a)$$

$$[d_1 \quad \cdots \quad d_n]^T = \|\Gamma[\sigma_1 \quad 0 \quad \cdots \quad 0]^T\| \cdot [1 \quad 0 \quad \cdots \quad 0] \Sigma \text{sign}(\gamma) \quad (4.14b)$$

4.4. ESTIMATION DES PARAMETRES DU SYSTEME

4.4.1. Estimation des paramètres composés

L'estimation du vecteur de paramètres Θ^* peut se faire, de manière récursive ou non, en utilisant des variantes adéquates de l'estimateur des moindres carrés. Un exemple en est l'algorithme suivant :

$$\begin{aligned}\bar{\Theta}(t) &= \hat{\Theta}(t-1) + \frac{P_1(t)\Phi(t)e(t)}{1 + \Phi^T(t)P_1(t-1)\Phi(t)} \\ \hat{\Theta}(t) &= \min(\rho^*, \|\bar{\Theta}(t)\|) \cdot \frac{\bar{\Theta}(t)}{\|\bar{\Theta}(t)\|}\end{aligned}\quad (4.15a)$$

$$\begin{aligned}P_1(t) &= P_1(t-1) - \frac{P_1(t-1)\Phi(t)\Phi^T(t)P_1(t-1)}{1 + \Phi^T(t)P_1(t-1)\Phi(t)} \\ e(t) &= y(t) - \Phi(t)^T \hat{\Theta}(t-1)\end{aligned}\quad (4.15b)$$

où le vecteur initial $\hat{\Theta}(0)$ est arbitrairement choisi et la constante réelle ρ^* est un majorant quelconque de $\|\Theta^*\|$, $P_1(0) = \alpha I$ et α un réel assez grand. Ces équations définissent un algorithme de type moindres carrés avec projection sur une sphère centrée sur l'origine et de rayon ρ^* . Cette projection est introduite pour pallier la nature intégrale de l'algorithme des moindres carrés standard ; elle empêchera les estimés de diverger en dépit de l'erreur $\eta(t)$. Les autres propriétés générales de cet estimateur sont celles décrites par la Proposition 3.1 du chapitre précédent. En outre, les commentaires qui ont immédiatement suivi cette proposition restent valables ici. En particulier, il n'est pas du tout garanti à ce stade que $\hat{\Theta}(t)$ converge vers sa vraie valeur Θ^* , et ce même dans le cas idéal ($\eta(t) = 0$). La consistance de l'estimateur dépend en fait de la capacité du signal d'entrée $\{v(t)\}$ à procurer ou non la propriété d'excitation persistante au vecteur de régression sous-jacent. Cet aspect sera abordé plus tard au paragraphe 5 de ce chapitre.

4.4.2. Reconstruction de $A(q^{-1})$, $B(q^{-1})$ et des N points de la non linéarité $F(.)$

Etant donné le vecteur des estimés suivant :

$$\hat{\Theta}(t) = [\hat{a}_1(t) \dots \hat{a}_n(t) \quad \hat{\mu}_{11}(t) \dots \hat{\mu}_{n1}(t) \quad \hat{\mu}_{12}(t) \dots \hat{\mu}_{n2}(t) \dots \dots \hat{\mu}_{1N}(t) \dots \hat{\mu}_{nN}(t)]^T \quad (4.16a)$$

On peut en déduire immédiatement une estimation de $A(q^{-1})$:

$$\hat{A}_t(q^{-1}) = 1 + \hat{a}_1(t)q^{-1} + \dots + \hat{a}_n(t)q^{-n} \quad (4.16b)$$

Des estimations (\hat{b}_i, \hat{d}_i) de (b_i, d_i) sont alors obtenues à partir des $\hat{\mu}_{ij}$ à travers une décomposition en valeur singulières de la matrice suivante:

$$\hat{M}(t) = \begin{bmatrix} \hat{\mu}_{11}(t) & \cdots & \hat{\mu}_{1N}(t) \\ \vdots & & \vdots \\ \hat{\mu}_{n1}(t) & \cdots & \hat{\mu}_{nN}(t) \end{bmatrix} \quad (4.17)$$

Tant que les $\hat{\mu}_{ij}$ n'ont pas encore convergé vers leurs vraies valeurs, la matrice $\hat{M}(t)$ n'est pas forcément de rang 1. Ainsi, la décomposition de cette matrice prend-elle la forme générale suivante (en supposant que $n < N$) :

$$\hat{M}(t) = \Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2(t) & \ddots & & & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & & \vdots & \vdots \\ 0 & \cdots & 0 & \sigma_n(t) & 0 & \cdots & 0 \end{bmatrix} \Sigma(t) \quad (4.18)$$

La Lemme 2.8 du Chapitre 2 suggère alors d'estimer les coefficients (b_i, d_i) de la façon suivante :

$$\begin{bmatrix} \hat{b}_1 & \cdots & \hat{b}_n \end{bmatrix}^T = \frac{\Gamma(t) [\sigma_1(t) \ 0 \ \cdots \ 0]^T}{\| \Gamma(t) [\sigma_1(t) \ 0 \ \cdots \ 0]^T \|} \text{sign}(\gamma(t)) \quad (4.19a)$$

$$\begin{bmatrix} \hat{d}_1 & \cdots & \hat{d}_n \end{bmatrix}^T = \| \Gamma(t) [\sigma_1(t) \ 0 \ \cdots \ 0]^T \| \cdot [1 \ 0 \ \cdots \ 0] \Sigma(t) \text{sign}(\gamma(t)) \quad (4.19b)$$

où $\gamma(t) = \rho(\Gamma(t) [\sigma_1(t) \ 0 \ \cdots \ 0]^T)$ et $\rho(X)$ désigne la première composante de X ayant la plus grande valeur absolue. Les estimés ainsi obtenus sont en fait les seuls qui vérifient $\sum_{i=1}^n \hat{b}_i^2 = 1$ et $\rho(\begin{bmatrix} \hat{b}_1(t) & \cdots & \hat{b}_n(t) \end{bmatrix}) > 0$. En outre, il découle de la Proposition 3.2 du précédent chapitre que, si $\hat{\Theta}(t)$ converge vers sa vraie valeur Θ^* , il en sera de même pour les $(\hat{b}_i(t), \hat{d}_i(t))$.

Etant donné les estimés $(\hat{b}_i(t), \hat{d}_i(t))$, les équations (4.2b), (4.7) et (4.9b) suggèrent pour $B(q^{-1})$ et les $F(V_i)$ ($i=1, \dots, N$) les estimés suivants :

$$\hat{B}_i(q^{-1}) = \hat{b}_1(t)q^{-1} + \cdots + \hat{b}_n(t)q^{-n} \quad (4.20a)$$

$$\hat{F}(V_i) = V_i \hat{d}_i(t) \quad (i=1, \dots, N) \quad (4.20b)$$

4.5. EXCITATION PERSISTANTE ET CONSISTANCE DES PARAMETRES ESTIMES

On recours au signal impulsionnel déjà utilisé au chapitre précédent (paragraphe 3.5). Ce signal est périodique de période $T=n(N+3)$ et est analytiquement défini, sur une période $[t_k, t_{k+1}[$ (avec $t_k=kT$), par l'expression suivante :

$$v(t) = \begin{cases} V_1 & \text{pour } t = t_k + 2n \\ V_j & \text{pour } t = t_k + (2 + j)n \\ 0 & \text{ailleurs} \end{cases} ; j = 2, \dots, N \quad (4.21)$$

Ce signal est conçu de façon à prendre ses valeurs dans l'ensemble $\{0, V_1, \dots, V_N\}$, permettant ainsi de substituer le modèle (4.9a-b), impliquant le polynôme $P(\cdot)$, au modèle initial (4.1)-(4.2), impliquant la non-linéarité $F(\cdot)$. Le caractère périodique de ce signal en fait un signal déterministe et facilement réalisable en pratique. En outre, les amplitudes et les signes des impulsions (V_j) peuvent théoriquement être choisis arbitrairement. En pratique, les signes peuvent par exemple être choisis de telle sorte que le signal $\{v(t)\}$ soit de moyenne nulle. Quant aux amplitudes, elles ne doivent être :

- i) ni trop grandes, afin de garantir le maintien du système dans le domaine de validité du modèle,
- ii) ni trop petites, afin d'assurer une vitesse de convergence raisonnable pour les paramètres estimés.

Le dernier aspect est, en partie, ressorti par la Proposition 3.5 du précédent chapitre. Compte tenu de son importance, nous la reproduisons ici:

Proposition 4.1 (L'excitation persistante de $\Phi(t)$).

- 1) Il existe une constante réelle $K_1 > 0$, indépendante de μ , et une séquence réelle positive $\{\delta_0(k)\} \in S(K_1, \mu)$ telle que, pour tout k et tout vecteur unitaire W :

$$\max_{1 \leq \tau \leq (N+3)n} |\Phi(t_k + \tau)^T W| > \varepsilon - \delta_0(k) \quad (4.22)$$

où le réel $\varepsilon > 0$ est proportionnel à $\min\{V_1^2, \dots, V_N^2\}$.

- 2) Dans le cas idéal ($\{z(t)\} = \{0\}$), la séquence $\{\delta_0(k)\}$ du précédent point est nulle \square

Les commentaires ayant fait l'objet de la Remarque 3.1 du précédent chapitre sont tous applicables à la Proposition 4.1 ci-dessus. En particulier, celle dernière montre bien que le degré d'excitation procuré à $\Phi(t)$, et mesuré par ε , est proportionnel à la taille des impulsions. Plus grand est $\min\{V_1^2, \dots, V_N^2\}$, plus élevé sera le degré d'excitation et, par conséquent, plus rapide sera la convergence des paramètres estimés. Cette convergence est parfaitement décrite par la Proposition 3.6 du chapitre précédent, elle est reproduite ici par commodité pédagogique.

Proposition 4.2. L'algorithme (4.15a-b), appliqué au système (4.1)-(4.2) excité par le signal (4.21), génère une suite d'estimés $\{\hat{\Theta}(t)\}$ possédant les propriétés de convergence suivantes:

- 1) Cas général ($\{\eta(t)\}$ quelconque mais bornée): il existe une constante réelle K_2 , indépendante de μ , telle que:

$$\left\{ \left\| \hat{\Theta}(t) - \Theta^* \right\| \right\} \in S(K_2 \mu) \quad (4.23)$$

- 2) Cas idéal ($\{\eta(t)\} = \{0\}$):

$$\left\{ \left\| \hat{\Theta}(t) - \Theta^* \right\| \right\} \text{ converge vers zéro} \quad (4.24)$$

Dans ce cas, la convergence est exponentielle et est d'autant plus rapide que $\min\{V_1^2, \dots, V_N^2\}$ est plus grand \square

Cette proposition montre que la qualité moyenne des estimés est proportionnelle à la taille moyenne (μ) de l'erreur $\eta(t)$ de modélisation. En l'absence d'erreur ($\eta(t)=0$), l'excitation persistante est garantie pour $\Phi(t)$ ce qui implique la consistance de $\hat{\Theta}(t)$. Ces résultats se transposent immédiatement sur les estimés $(\hat{b}_i(t), \hat{d}_i(t))$ du fait de la continuité de l'estimateur (4.19a-b).

Proposition 4.3. Sous les mêmes conditions que dans la Proposition 3.2, les estimés $(\hat{b}_i(t), \hat{d}_i(t))$ obtenus par l'estimateur (4.19a-b) possèdent les propriétés suivantes :

- 1) Dans le cas général ($\{\eta(t)\}$ est quelconque mais bornée), il existe une constante réelle K_3 , indépendante de μ , telle que:

$$\left\{ \left| \hat{b}_i(t) - b_i \right| \right\}, \left\{ \left| \hat{d}_j(t) - d_j \right| \right\} \in S(K_3 \mu) \quad (4.23)$$

où $i=1, \dots, n; j=1, \dots, N$.

- 2) Dans le cas idéal ($\{\eta(t)\} = \{0\}$):

$$\left\{ \left| \hat{b}_i(t) - b_i \right| \right\}, \left\{ \left| \hat{d}_j(t) - d_j \right| \right\} \text{ convergent vers zéro} \quad (4.24) \quad \square$$

4.6. ESTIMATION DES PARAMETRES CARACTERISTIQUES DE L'ELEMENT NON LINEAIRE F(.)

Le schéma d'identification présenté précédemment permet, dans le cas idéal, d'identifier parfaitement N points de la caractéristique $F(.)$. Dans le cas non idéal, les N points $(V_j, \hat{F}_t(V_j))$ fournis par ce schéma sont situés (pour $t \sim \infty$) à proximité de la caractéristique graphique de $F(.)$. Ce

nuage de points sera d'autant plus proche de $F(\cdot)$ que la taille moyenne de l'erreur $\eta(t)$ est plus petite. Ce résultat a en fait déjà été établi au chapitre précédent et ce pour tout élément $F(\cdot)$ possédant les deux propriétés élémentaires :

$$F(0)=0 \text{ et } (|v|<\infty \Rightarrow |F(v)|<\infty)$$

En revanche, le précédent chapitre n'a pas soulevé la question suivante :

Comment les paramètres caractéristiques de l'élément non linéaire se déduisent-ils des N points $(V_j, \hat{F}_t(V_j))$ identifiés précédemment ?

Il va sans dire que la réponse à cette question dépend fortement de la nature de l'élément $F(\cdot)$. En d'autres termes, la détermination des paramètres de $F(\cdot)$ ne peut être envisagée qu'au cas par cas. Dans la suite de ce paragraphe, nous traitons cet aspect pour les deux non-linéarités considérées en début de chapitre, à savoir $\mathcal{P}(c_1, \dots, c_m)$ et $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$.

4.6.1. Estimation des paramètres d'éléments de type $\mathcal{P}(c_1, \dots, c_m)$

On se focalise ici sur le cas où l'élément non linéaire $F(\cdot)$ est un polynôme de la forme (4.4) dont le degré m est connu. Il s'agit d'estimer ses coefficients (c_1, \dots, c_m) sachant qu'on dispose d'un nuage de points $(V_j, \hat{F}_t(V_j))$ ($j=1, \dots, N$). A cet effet, il suffit de prendre $N=m$, d'écrire (4.4) pour $v=V_j$ ($j=1, \dots, m$) et d'utiliser (4.6) et (4.8). Ce faisant, on obtient le système d'équations :

$$F(V_j) = \sum_{i=1}^m c_i V_j^i = P(V_j) = V_j d_j \quad (j=1, \dots, m) \tag{4.25}$$

ce que l'on peut mettre sous la forme matricielle suivante :

$$\begin{bmatrix} V_1 d_1 \\ \vdots \\ V_m d_m \end{bmatrix} = \begin{bmatrix} V_1 & \dots & V_1^m \\ \vdots & \ddots & \vdots \\ V_m & \dots & V_m^m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = A \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \tag{4.26}$$

Il découle de (4.5) que la matrice A est inversible, par conséquent:

$$\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} = A^{-1} \begin{bmatrix} V_1 d_1 \\ \vdots \\ V_m d_m \end{bmatrix} \tag{4.27}$$

Comme les V_j sont choisis a priori, le calcul de l'inverse de A se fait une fois pour toute au départ. L'expression (4.27) suggère l'estimateur suivant pour les coefficients c_i :

$$\begin{bmatrix} \hat{c}_1(t) \\ \vdots \\ \hat{c}_m(t) \end{bmatrix} = A^{-1} \begin{bmatrix} V_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & V_m \end{bmatrix} \begin{bmatrix} \hat{d}_1(t) \\ \vdots \\ \hat{d}_m(t) \end{bmatrix} \quad (4.28)$$

Il découle de (4.28) que les propriétés de convergence des $\hat{c}_j(t)$ sont identiques à celles des $\hat{d}_j(t)$ données par la Proposition 4.3.

4.6.2. Estimation des paramètres d'éléments de type $ZMS(D_1, D_2, S_1, S_2, L_1, L_2)$

On s'intéresse aux éléments de type seuil avec zone-morte dont il s'agit de déterminer les paramètres $(D_1, D_2, S_1, S_2, L_1, L_2)$ à partir d'un nuage de points $(V_j, \hat{F}_i(V_j))$ ($j=1, \dots, N$). Au vu de la caractéristique graphique (fig. 4.1) de ces éléments, il apparaît qu'il suffit de disposer de quatre points tous situés en dehors de la zone-morte. En outre, une paire de points doit être située à gauche de cette zone et l'autre à droite. Il s'ensuit que $N=4$ et, compte-tenu (4.5), les V_j ($j=1, \dots, 4$) doivent être choisis de la façon suivante :

$$V_1 < V_2 < D_2 < 0 < D_1 < V_3 < V_4$$

Or, les paramètres D_1 et D_2 sont inconnus. C'est pour contourner cette difficulté que nous faisons recours à l'hypothèse suivante :

H4.4: il existe un réel connu D_M tel que $D_M > \max(D_1, -D_2)$.

En pratique il est assez facile de déterminer un majorant D_M des deux zones mortes. Moyennant cette hypothèse, il devient possible de choisir les V_j ($j=1, \dots, 4$) de la façon suivante :

$$V_1 < V_2 < -D_M < 0 < D_M < V_3 < V_4 \quad (4.29)$$

Ce qui assure que la paire de points $\{(V_1, F(V_1)), (V_2, F(V_2))\}$ est située sur la demi-droite à gauche de la zone-morte et que la paire $\{(V_3, F(V_3)), (V_4, F(V_4))\}$ est située à droite. La donnée de ces deux paires de points (ou plutôt de leurs estimations) va permettre de caractériser parfaitement les deux demi-droites situées de part et d'autre de la zone-morte de l'élément $F(\cdot)$. Mais, auparavant il convient de faire une remarque d'ordre bibliographique concernant l'hypothèse H4.4. (Bai, 2002) propose une méthode d'identification applicable au cas d'une non-linéarité du $ZMS(D_1, D_2, S_1, S_2, L_1, L_2)$ présentant une symétrie parfaite, c'est-à-dire $D_1=D_2=S_1=S_2=L_1=L_2=a$. La méthode implique la minimisation d'une fonction non linéaire et non convexe du seul paramètre inconnu a . L'auteur suggère alors de procéder graphiquement pour localiser le minimum global. Or, une méthode graphique porte nécessairement sur un intervalle de recherche fini choisi a priori. Elle

ne peut donc localiser le minimum global que si ce dernier appartient à l'intervalle de recherche choisi ce qui revient de fait à connaître a priori un majorant du module de ce minimum global.

Revenons à présent au problème de détermination des paramètres inconnus $(D_1, D_2, S_1, S_2, L_1, L_2)$ en supposant disponibles quatre points $(V_j, \hat{F}_t(V_j))$ ($j=1, \dots, 4$) vérifiant la condition (4.29). Il vient de la Fig. 3.1 que :

$$F(v) = \begin{cases} S_1 v + H_1 & \text{si } v > D_M \\ S_2 v + H_2 & \text{si } v < -D_M \end{cases} \quad (4.30)$$

Il découle alors de (4.29) que :

$$S_1 = \frac{F(V_3) - F(V_4)}{V_3 - V_4} \text{ et } S_2 = \frac{F(V_1) - F(V_2)}{V_1 - V_2} \quad (4.31)$$

$$H_1 = \frac{V_3 F(V_4) - V_4 F(V_4)}{V_3 - V_4} \text{ et } H_2 = \frac{V_1 F(V_2) - V_2 F(V_1)}{V_1 - V_2} \quad (4.32)$$

Par ailleurs, il vient de (4.6)-(4.8) que :

$$F(V_j) = P(V_j) = d_j \cdot V_j \quad (j=1, 2, 3, 4) \quad (4.33)$$

Les expressions (4.31)-(4.29) suggèrent les estimateurs suivants :

$$\hat{S}_1(t) = \frac{\hat{d}_3(t)V_3 - \hat{d}_4(t)V_4}{V_3 - V_4} \text{ et } \hat{S}_2(t) = \frac{\hat{d}_1(t)V_1 - \hat{d}_2(t)V_2}{V_1 - V_2} \quad (4.34a)$$

$$\hat{H}_1(t) = V_3 V_4 \frac{\hat{d}_3(t) - \hat{d}_4(t)}{V_3 - V_4} \text{ et } \hat{H}_2(t) = V_1 V_2 \frac{\hat{d}_2(t) - \hat{d}_1(t)}{V_1 - V_2} \quad (4.34b)$$

Les paramètres qu'il faut encore estimer sont L_1, L_2, D_1, D_2 . Au regard de la Fig. 4.1, il apparaît que :

$$L_1 = S_1 D_1 + H_1 \text{ et } L_2 = S_2 D_2 + H_2 \quad (4.35)$$

Ces deux équations sont insuffisantes pour déterminer les quatre paramètres inconnus. Cette détermination sera seulement possible si on soumet ces paramètres à des conditions supplémentaires conduisant à des cas particuliers importants de la non-linéarité. Ces cas sont décrits dans ce qui suit.

Eléments de type zone-morte

Dans ce cas $L_1=L_2=0$. Les expressions (4.35) suggèrent alors les estimateurs suivants pour D_1 et D_2 :

$$\hat{D}_1(t) = -\frac{\hat{H}_1(t)}{\hat{S}_1(t)} \quad \text{et} \quad \hat{D}_2(t) = -\frac{\hat{H}_2(t)}{\hat{S}_2(t)} \quad (4.36)$$

Eléments de type seuil

Dans ce cas $D_1=D_2=0$. Les expressions (4.35) suggèrent alors les estimateurs suivants pour L_1 et L_2 :

$$\hat{L}_i(t) = \hat{H}_i(t) \quad (i=1, 2) \quad (4.37)$$

Eléments de type $ZMS(D_1, D_2, S_1, S_2, L_1, L_2)$ partiellement symétriques

La symétrie peut concerner les pentes (S_1, S_2) ou les seuils (L_1, L_2) et les zones mortes (D_1, D_2).

On suppose que $L_1=-L_2 \stackrel{\text{def}}{=} L$ et $D_1=-D_2 \stackrel{\text{def}}{=} D$. Il est alors possible de déterminer les inconnues L et D mais à condition que $S_1 \neq S_2$. En effet, il découle des expressions (4.35) que :

$$D = \frac{H_1 + H_2}{S_2 - S_1} \quad \text{et} \quad L = S_1 D + H_1 \quad (4.38)$$

ce qui suggère les estimateurs suivants :

$$\hat{D}(t) = \frac{\hat{H}_1(t) + \hat{H}_2(t)}{\hat{S}_2(t) - \hat{S}_1(t)} \quad \text{et} \quad \hat{L}(t) = \hat{S}_1(t)\hat{D}(t) + \hat{H}_1(t) \quad (4.39)$$

Eléments de type $ZMS(D_1, D_2, S_1, S_2, L_1, L_2)$ totalement symétriques

On suppose que $S_1 = S_2 \neq 1$, $H_1 = -H_2$ et $L_1 = -L_2 = D_1 = -D_2 = X$ alors on peut estimer l'inconnu X par :

$$\hat{X}(t) = \frac{\hat{H}_1(t)}{1 - \hat{S}_1(t)} \quad (4.40)$$

Remarque 4.1. Il est important de noter ici que les cas de figures considérés ci-dessus ne vont pas jusqu'à supposer une symétrie parfaite de l'élément non linéaire comme cela fut le cas dans (Bai, 2002). Mieux encore les pentes S_1 et S_2 ne sont pas forcément égales à un, comme cela a été supposé dans (Bai, 2002). Le schéma d'identification proposé embrasse de ce fait une plus grande variété d'éléments non linéaires en comparaison avec les travaux antérieurs.

4.7. SIMULATION

Dans ce paragraphe, on rapporte les résultats de simulations effectuées dans un but d'illustrer l'efficacité du schéma d'identification présenté dans ce chapitre. Nous considérons un système dont la dynamique linéaire est caractérisée par:

$$A(q^{-1}) = 1 - 1.5q^{-1} + 0.56q^{-2}; B(q^{-1}) = q^{-1} - 2q^{-2};$$

Pour l'élément statique non linéaire on considère deux cas; Dans le premier cas $F(v)$ est polynomiale et dans le deuxième cas $F(v)$ est une caractéristique asymétrique de type zone morte avec un seuil.

Dans un premier temps on considère que la perturbation extérieure $z(t)$ est nulle, pour vérifier la convergence des paramètres vers leurs vraies valeurs.

En vue de la condition (4.13) et de (4.14), les paramètres à identifier sont, pour le sous-système linéaire sont:

$$a_2 = 0.56, a_1 = -1.5, b_1 = -1/\sqrt{5} = -0.447, b_2 = 2/\sqrt{5} = 0.894$$

1^{er} cas: Gain non linéaire polynomial

$$F(v) = \sum_{j=1}^m c_j v^j \quad \text{avec } m=5 \text{ et } c_1=1, c_2=0.0, c_3=0.2, c_4=0.0, c_5=0.009$$

Etant donné la condition (4.13) et les expressions (4.14a-b), les paramètres à identifier seront, pour le gain non linéaire:

$$c_1 = \sqrt{5} = 2.236, c_2 = 0.0, c_3 = -0.447, c_5 = -0.0201$$

On suppose qu'on ne connaît pas d'avance m , mais plutôt un majorant de m , soit $m'=7$. La séquence d'entrée est générée selon la loi (4.21), en utilisant les sept valeurs suivantes :

$$V_1 = -3; V_2 = -2; V_3 = -1; V_4 = 1; V_5 = 2; V_6 = 3; V_7 = 4.$$

La fig.4.2 montre la convergence des paramètres du sous-système linéaire. Les estimés des paramètres qu'on obtient pour le gain non linéaire sont parfaitement égaux à leurs vraies valeurs; et en particulier on obtient :

$$\hat{c}_7 = \hat{c}_6 = \hat{c}_4 = \hat{c}_2 = 0$$

Ce qui nous permet de déduire que le degré du polynôme est égal à 5.

Les estimés des paramètres donnés par l'algorithme (4.14a-b) convergent rapidement vers leurs vraies valeurs en l'absence de perturbations externes. Ceci prouve la consistance de l'algorithme en absence de perturbations externes.

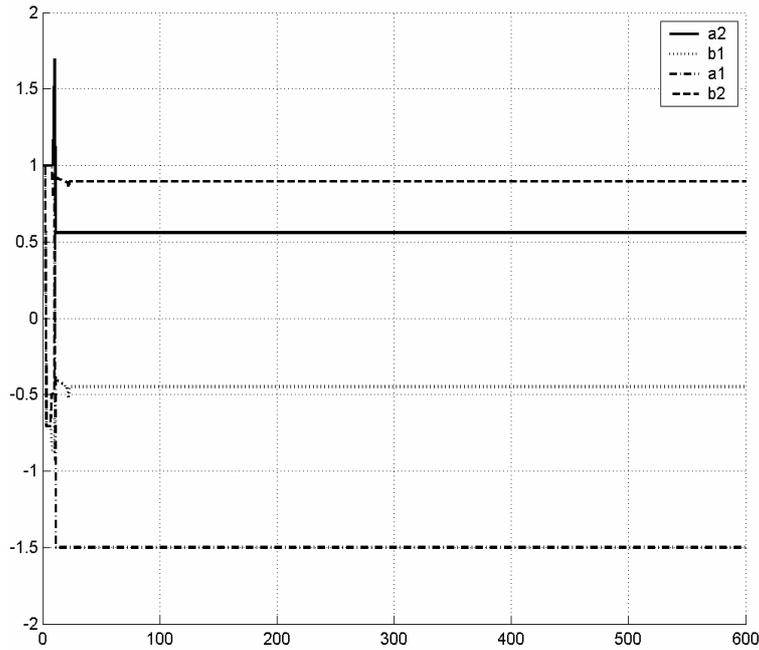


Figure 4.2 Convergence des estimés des paramètres du sous-système linéaire

Nous considérons maintenant le même système mais en présence d'un bruit externe en sortie de type:

$$z(t) = \frac{b(t)}{H(q^{-1})} \text{ où } b(t) \text{ est une séquence de nombres aléatoires dans l'intervalle}$$

$[-2.5 \ +2.5]$ et $H(q^{-1}) = 1 - 1.7q^{-1} + 0.72q^{-2}$ (voir fig.4.4).

On obtient les estimés suivants: $\hat{a}_1 = -1.51$, $\hat{a}_2 = 0.5625$, $\hat{b}_1 = -0.442$, $\hat{b}_2 = 0.897$

$\hat{c}_1 = -2.551$, $\hat{c}_2 = -0.111$, $\hat{c}_3 = -0.3765$, $\hat{c}_4 = 0.0142$, $\hat{c}_5 = -0.0252$, $\hat{c}_6 = 0.0001$, $\hat{c}_7 = -0.0000$.

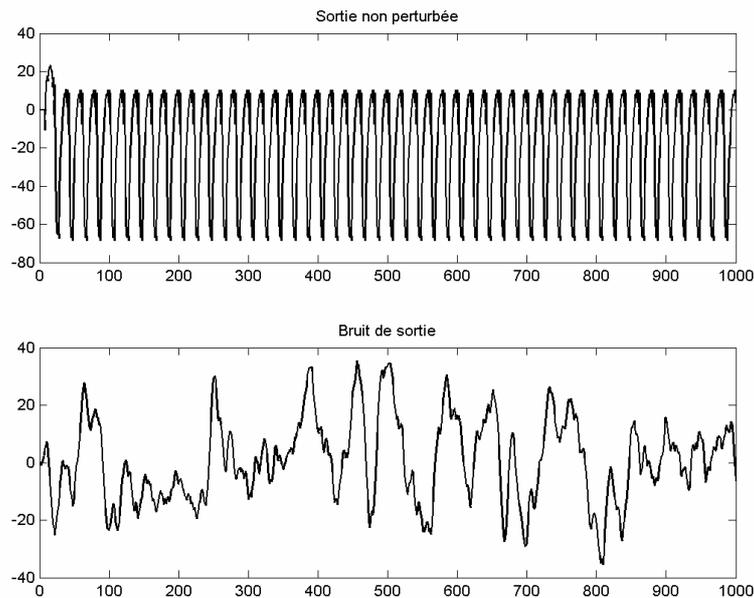


Figure 4.4 Sortie non perturbée et le bruit externe

La figure 4.5 montre l'erreur relative d'estimation de $F(v)$; on voit qu'elle ne dépasse pas 10%.

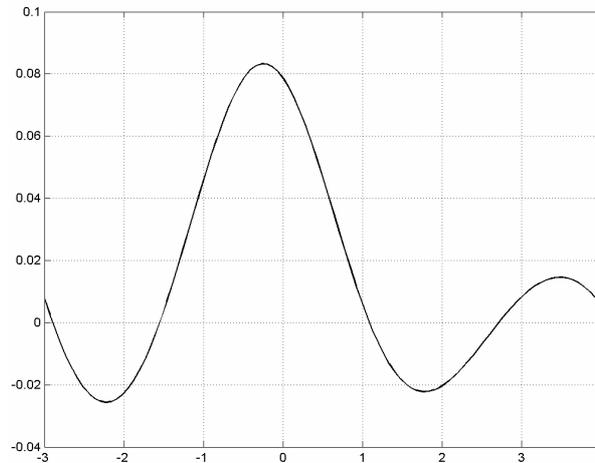


Figure 3.5 Erreur relative d'estimation de $F(v)$

D'après les valeurs obtenus pour les \hat{c}_j , on peut conclure que le degré du polynôme m est de 6, mais $\hat{c}_6 = 0.0001$ est tellement faible que dans la plage d'excitation considérée sa contribution peut être négligée. Ce problème d'estimation du degré du polynôme n'est pas spécifique à notre approche mais il est commun à toutes les approches présentées dans les précédents travaux.

D'après les résultats obtenus, on observe que les paramètres convergent vers un voisinage de leurs vraies valeurs, en présence d'une perturbation. D'autres résultats de simulation montrent que plus la perturbation est faible en moyenne, plus ce voisinage des vraies valeurs est très réduit.

2^{ème} cas: Zone morte et seuil asymétrique avec comme paramètres :

$$D_1 = -D_2 = 2; L_1 = -L_2 = 3; S_1 = H_1 = 1; S_2 = 2; H_2 = 1; \text{ (voir fig.4.1)}$$

Etant donné la condition (4.13) et les expressions (4.14), les paramètres à identifier seront, pour le gain non linéaire:

$$D = D_1 = -D_2 = 2; L = L_1 = -L_2 = -6.708; S_1 = H_1 = -\sqrt{5} = -2.236;$$

$$S_2 = 2\sqrt{5} = 4.47; H_2 = -2.236$$

La séquence d'entrée est générée selon la loi (4.21), en utilisant les quatre valeurs suivantes :

$$V_1 = -6; V_2 = -3; V_3 = 4; V_4 = 6.$$

La fig.4.6 montre la convergence des paramètres du sous-système linéaire, alors que la fig.4.7 montre la convergence de ceux du gain non linéaire. Les estimés des paramètres donnés par

l'algorithme (4.14a-b) convergent rapidement vers leurs vraies valeurs en l'absence de perturbations externes.

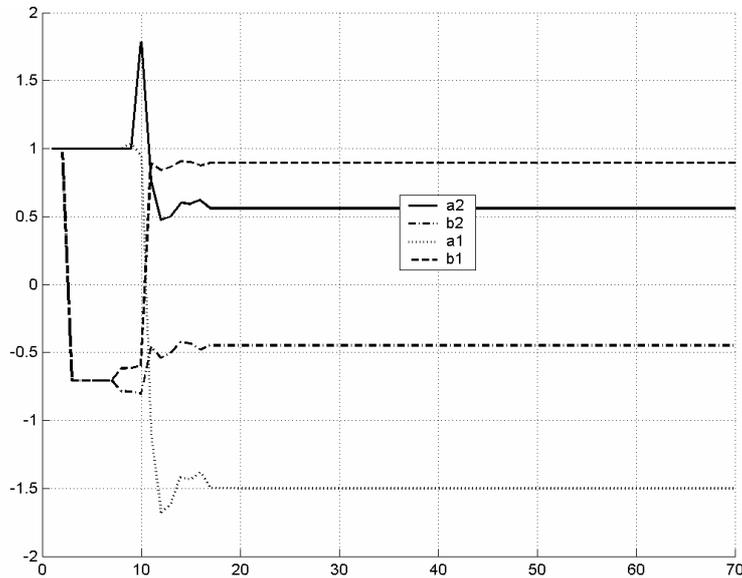


Figure 4.6 Convergence des estimés des paramètres du sous-système linéaire

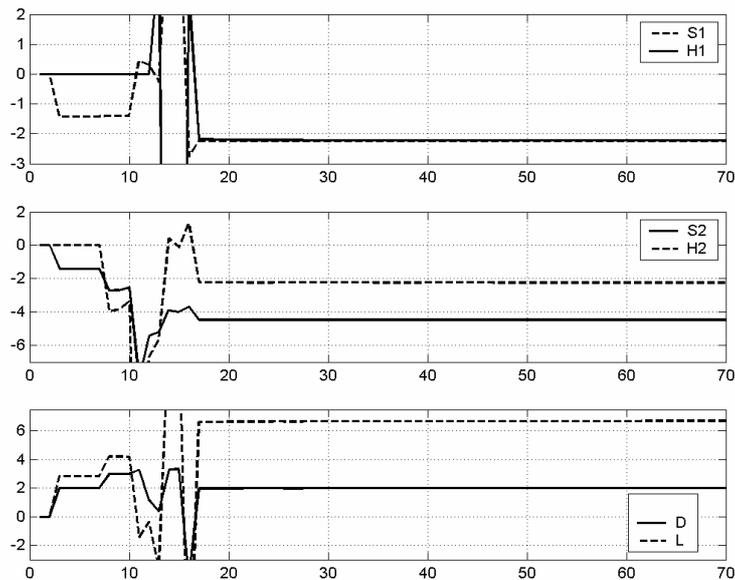


Figure 4.7 Convergence des estimés des paramètres du gain non linéaire

Nous considérons maintenant le même système mais en présence d'un bruit externe en sortie de type:

$$z(t) = \frac{b(t)}{H(q^{-1})} \text{ où } b(t) \text{ est une séquence de nombres aléatoires dans l'intervalle}$$

$$[-2.5 \ +2.5] \text{ et } H(q^{-1}) = 1 - 1.7q^{-1} + 0.72q^{-2}.$$

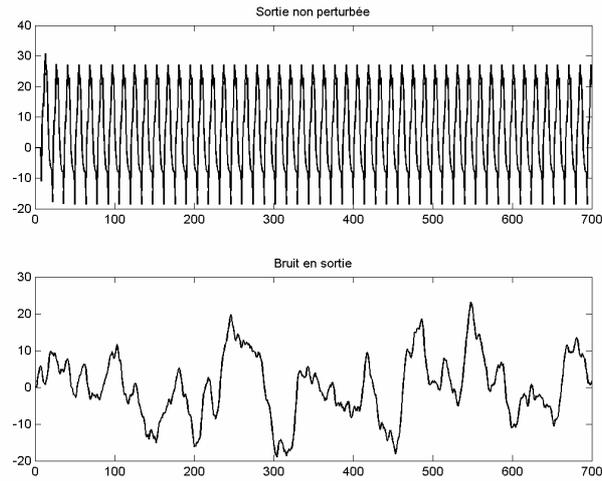


Figure 4.8 Sortie non perturbée et bruit externe

Les estimés obtenus convergent vers un voisinage de leur vraies valeurs. L'existence de ce biais est due aux perturbations externes comme on l'a démontré théoriquement. En effet les valeurs obtenues des estimés sont:

$$\hat{a}_1 = -1.543 \quad , \quad \hat{a}_2 = 0.569 \quad , \quad \hat{b}_1 = -0.429 \quad , \quad \hat{b}_2 = 0.903$$

$$\hat{D} = \hat{D}_1 = -\hat{D}_2 = 2.31 \quad , \quad S_1 = -2.27 \quad , \quad H_1 = -2.67 \quad , \quad S_2 = -4.683 \quad , \quad H_2 = -2.95$$

L'écart entre $\hat{D}_1 = 2.31$ et sa vraie valeur, se justifie par la taille assez importante des perturbations.

4.8. CONCLUSIONS

Dans ce chapitre nous avons abordé le problème d'identification des systèmes non linéaires sur la base du modèle de Hammerstein, en supposant que la structure du gain non linéaire mis en jeu est connue a priori. L'incertitude porte de ce fait sur les paramètres du sous-système dynamique et ceux du gain statique. Deux classes de non-linéarités ont été considérées : celle de type polynomiale $\mathcal{P}(c_1, \dots, c_m)$ et celle de type seuil avec zone-morte $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$. La seconde classe est caractérisée par sa nature non lisse ce qui explique le peu de travaux antérieurs consacrés à ces éléments. Le schéma d'identification proposé dans ce chapitre exploite de façon judicieuse, notamment en ce que concerne les éléments de type $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$, les résultats du chapitre précédent. Il permet ainsi d'estimer parfaitement, dans le cas idéal, la partie dynamique système et un nuage de points de la caractéristique du gain non linéaire. A partir de ce nuage de points il a été possible de déterminer les paramètres de tout élément de type $\mathcal{P}(c_1, \dots, c_m)$ et ceux de certaines catégories d'éléments de type $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$. En présence d'une erreur de modélisation, la qualité de l'estimation est proportionnelle à la taille moyenne de l'erreur. Par rapport aux travaux antérieurs, notre solution présente les caractéristiques suivantes :

- la nature déterministe de la solution contribue à sa simplicité, aussi bien au niveau de l'analyse/synthèse théorique qu'au niveau de la mise en œuvre,
- la solution prend en compte des éléments de type $\mathcal{ZMS}(D_1, D_2, S_1, S_2, L_1, L_2)$ asymétriques,
- le schéma proposé opère une estimation simultanée (et non en deux ou trois temps) du sous-système dynamique linéaire et du gain statique non linéaire ; il nécessite ainsi une seule expérience,
- la solution ne se contente pas de supposer que la propriété d'excitation persistante est vérifiée mais démontre qu'il en est réellement ainsi.

CHAPITRE 5

IDENTIFICATION SYSTEMES DE HAMMERSTEIN IMPLIQUANT UN ELEMENT NON LINEAIRE DE TYPE HYSTERETIQUE

5.1. INTRODUCTION

En général, les non linéarités non lisses présentent toutes une difficulté principale: elles nécessitent en effet plusieurs formulations mathématiques pour les décrire. Chacune de ces formulations est valide sur une plage (ou intervalle) d'entrée bien spécifique de l'entrée. En plus, ces intervalles sont définis par des paramètres inconnus; sur la Fig 5.1, par exemple, ces intervalles sont $] -\infty, h_2]$, $]h_2, h_1[$ et $[h_1, +\infty[$, où h_1 et h_2 sont des paramètres inconnus. Une étape clé dans la conception des méthodes d'identification sera le développement d'une formulation mathématique unique pour l'élément non linéaire qui est valide sur tous les intervalles. Ceci n'est pas toujours possible, et les schémas d'identification doivent alors s'accommoder de cette contrainte. En plus dans la pratique, certains systèmes présentent des non linéarités qui ne sont pas statiques ; Pour ce genre de non linéarités qu'on qualifie de non linéarités avec effet mémoire (à la différence des non linéarités statiques), la sortie n'est pas toujours la même pour la même entrée. La sortie de l'élément non linéaire est fonction de son entrée mais aussi de l'évolution antérieure de cette entrée. Cet effet mémoire ajoute une difficulté supplémentaire au problème d'identification. Ceci explique un peu le nombre assez réduit de contributions pour l'identification de tels systèmes.

Dans ce chapitre nous aborderons l'identification des systèmes non linéaires sur la base du modèle de Hammerstein à gain non linéaire avec effet mémoire. Deux schémas d'identification seront présentés. Le premier s'applique au cas où le gain non linéaire est de type hystérésis-relais (Fig 5.1) alors que l'autre s'applique au cas où le gain non linéaire est de type hystérésis (Fig 5.2). Les types de non linéarités considérées dans ce chapitre sont loin d'être simplistes et sont très présentes dans certains systèmes réels. Leur présence peut être volontaire, comme c'est le cas de l'hystérésis-relais, largement utilisé dans le contrôle des procédés thermiques (thermostats), ou électriques (commandes en fourchette de courant des machines électriques), ou complètement involontaire,

comme c'est le cas de l'hystérésis (aussi appelé backlash ou jeu), qui est presque inévitable dans les systèmes de transmission mécanique (jeu dans les engrenages par exemple).

Pour surmonter toutes les difficultés liées à l'effet mémoire, on identifie séparément les paramètres du sous système linéaire et ceux du gain non linéaire. Dans une première étape, on procède à une identification des paramètres du sous système linéaire moyennant une paramétrisation adéquate du système et une séquence d'entrée qui masque l'effet mémoire du gain non linéaire. Ces paramètres, sont alors utilisés, dans une autre étape, pour estimer les paramètres du gain non linéaire en utilisant une nouvelle paramétrisation et des séquences d'entrée bien spécifiques.

A l'inverse de (Bai, 2002), notre approche est valable pour les non linéarités symétriques ou asymétriques. Ces dernières sont caractérisées par plusieurs paramètres et non pas par un seul. Ceci rend la tâche d'identification un peu plus difficile. Plus précisément, l'hystérésis sera définie par trois paramètres (S, h_1, h_2) , alors que le hystérésis-relais sera défini par quatre paramètres (M_1, M_2, h_1, h_2) . En plus, la méthode d'identification proposée utilise, pour les deux étapes, des séquences d'entrée déterministes, excitantes persistantes et périodiques. Ces séquences peuvent donc être facilement générées. L'approche, dans sa totalité, présente l'avantage de ne pas nécessiter ni calculs d'optimisation complexes, ni approximations simplificatrices.

Ce chapitre est organisé comme suit : la section 2, rappelle certains concepts mathématiques et introduit des notations qu'on utilisera par la suite; La section 3 introduit le problème d'identification considéré; la section 4 présente le schéma d'identification du sous système linéaire ; la section 5 celui du gain non linéaire, pour les deux types de non linéarités considérés (relais-hystérésis et l'hystérésis). Les sections 4 et 5 se terminent chacune par un paragraphe qui donne quelques résultats de simulation pour concrétiser l'efficacité de l'approche.

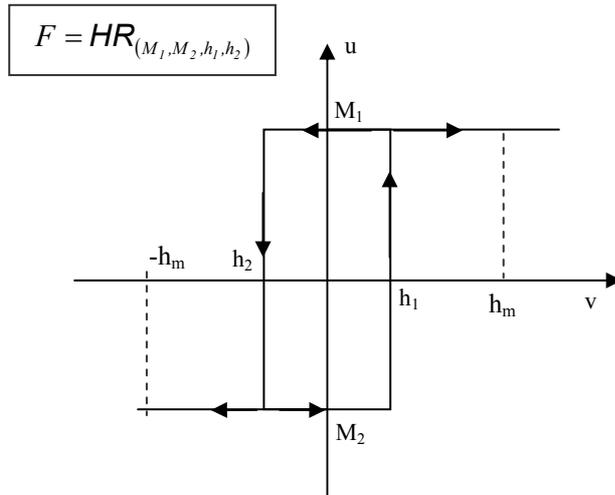


Figure 5.1 Non linéarité de type Hystérésis -Relais

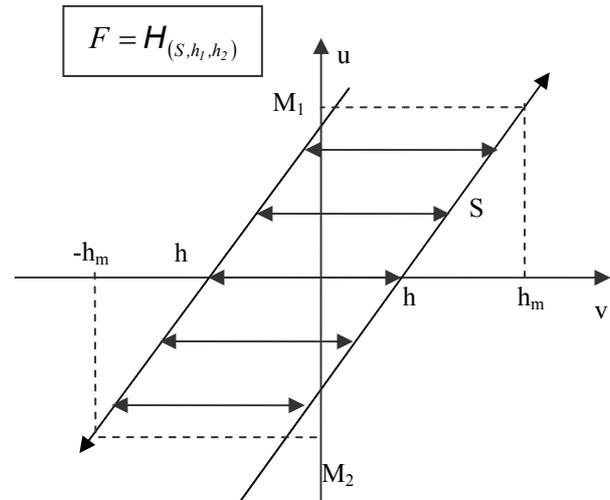


Figure 5.2 Non linéarité de type Hystérésis

5.2. NOTATIONS MATHÉMATIQUES PRELIMINAIRES

Nous allons, d'abord, introduire certaines notations et concepts mathématiques utilisés le long de ce chapitre. En plus, nous rappelons, en version simplifiée, le lemme technique d'excitation du chapitre 2.

5.2.1. Notations et acronymes

$HS-H$: Système de Hammerstein avec hystérésis

$HS-HR$: Système de Hammerstein avec Hystérésis-Relais

$H_{(S, h_1, h_2)}$: Non linéarité de type Hystérésis définie par les paramètres (S, h_1, h_2) , (fig. 5.2).

$HR_{(M_1, M_2, h_1, h_2)}$: Non linéarité de type Hystérésis définie par les paramètres (M_1, M_2, h_1, h_2) , (fig. 5.1).

t : temps discret ($t=0, 1, 2, \dots$)

$E(x(t))$: Espérance mathématique (pour une séquence stochastique $\{x(t)\}$)

$\bar{x}(N)$: Moyenne temporelle de la séquence $\{x(t)\}$ c'est-à-dire $\bar{x}(N) = \frac{1}{N} \sum_{i=1}^N x(i)$

a.p.1: avec probabilité 1

Si $\{x(t)\}$, $\{y(t)\}$ sont des processus stochastiques stationnaires et ergodiques, alors on a,

a.p.1:

$$\lim_{N \rightarrow \infty} \bar{x}(N) = E(x(t)) ,$$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N x(t)y(t+\tau) = E(x(t)y(t+\tau))$$

5.2.2. Une classe de séquences excitantes persistantes pour les systèmes linéaires

Considérons un système linéaire contrôlable décrit par:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + \xi(t) \quad (5.1a)$$

où $\xi(t)$ est un terme borné qui représente les perturbations externes sur le système; $A(q^{-1})$, $B(q^{-1})$ sont deux opérateurs polynomiaux en q^{-1} définis par:

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_naq^{-na} , \quad B(q^{-1}) = b_1q^{-1} + \dots + b_nbq^{-nb} \quad (5.1b)$$

Puisque $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux, grâce à la contrôlabilité du système , il existe une seule paire de polynômes $(P(q^{-1}), Q(q^{-1}))$ tels que:

$$A(q^{-1})P(q^{-1}) + B(q^{-1})Q(q^{-1}) = 1 \quad (5.2a)$$

Introduisons la variable d'état interne $z(t)$ définie par:

$$z(t) = P(q^{-1})u(t) + Q(q^{-1})x(t) \quad (5.2b)$$

où $x(t)$ représente la sortie non perturbée:

$$A(q^{-1})x(t) = B(q^{-1})u(t) \quad (5.2c)$$

Ainsi, en opérant $A(q^{-1})$ sur (5.2b) et en utilisant (5.2c) et (5.2a) on obtient:

$$A(q^{-1})z(t) = u(t) \quad (5.3)$$

Soit le vecteur d'état suivant :

$$Z(t) = [z(t) \dots z(t-2n+1)]^T \in \mathfrak{R}^{2n} \quad (5.4)$$

On peut alors formuler le lemme technique suivant:

Lemme 5.1. Soit le système (5.1a-b) excité par le signal d'entrée qui vérifie:

$$u(t) = \begin{cases} \text{une valeur quelconque (bornée)} & \text{si } t = t_k + 2n; \quad k = 0, 1, 2, \dots \\ 0 & \text{ailleurs} \end{cases} \quad (5.5)$$

où $\{t_k\}$ une séquence quelconque d'entiers telle que $t_k \geq t_{k-1} + 4n$. Donc, il existe un réel $\lambda > 0$ tel que, pour tout entier all k :

$$\sum_{i=0}^{4n-1} Z(t_k + i)Z(t_k + i)^T \geq \lambda [u(t_k + 2n)]^2 I_{2n} \quad (5.6)$$

□

Remarques 5.1.

1. Ce lemme est une conséquence directe du résultat plus général établi dans le lemme technique du chapitre 2.
2. La séquence d'entrée $\{u(t)\}$ est un train d'impulsions. L'intervalle séparant deux impulsions successives ne peut pas être plus petit que $4n-1$. La $k^{ième}$ impulsion est appliquée au centre de l'intervalle $[t_k, t_k+4n-1]$ et a, comme amplitude, $u(t_k)$. En vue de (5.6), elle produit l'effet d'excitation dans le même intervalle. Ceci explique l'appellation de séquences à excitation par intervalles, qu'on donne souvent à ce type d'excitation.
3. Dans le cas où $t_k = t_{k-1}+4n$, et $|u(t_k)| \geq \gamma > 0$ (pour tout k) l'inégalité (5.6) donne au vecteur d'état $Z(t)$ la propriété d'excitation persistante. Dans le cas où $t_k = t_{k-1}+4n$, et $|u(t_k)| = \gamma > 0$ (pour tout k), en plus du fait que $Z(t)$ sera excitant persistant, la séquence $\{u(t)\}$ devient alors périodique. Cette dernière propriété sera intéressante dans la suite de ce chapitre.

5.3. POSITION DU PROBLEME D'IDENTIFICATION

5.3.1. Classe de systèmes à identifier

On considère les systèmes non linéaires dont le comportement entrée/sortie peut être décrit par le modèle de Hammerstein suivant:

$$A(q^{-1})y(t) = B(q^{-1})u(t) + \zeta(t) \text{ et } u(t) = F(v(t)) \tag{5.7a}$$

avec

$$A(q^{-1}) = 1 + a_1q^{-1} + \dots + a_{na}q^{-na}, \quad B(q^{-1}) = b_1q^{-1} + \dots + b_{nb}q^{-nb} \tag{5.7b}$$

où $v(t)$ et $y(t)$ désignent respectivement l'entrée et la sortie du système; $u(t)$ est un signal interne non mesurable ; $\zeta(t)$ représente une perturbation externe; $\{\zeta(t)\}$ est une séquence de variables aléatoires indépendantes, stationnaire, ergodique et de moyenne nulle. La caractéristique $F(.)$ est, soit de type relais-hystérésis (Fig 5.1) ou hystérésis (Fig 5.2).

Par ailleurs, le modèle du système respecte les hypothèses suivantes:

H5.1. Il existe un majorant connu n tel que: $n \geq \max \{n_a, n_b\}$.

H5.2. $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux.

H5.3. Tous les zéros de $q^{na}A(q^{-1})$ sont à l'intérieur du cercle unité.

H5.4. Il existe un réel connu h_m tel que $h_m > \max(|h_1|, |h_2|)$ et soit $M_1 \stackrel{\text{def}}{=} F(h_m)$ et

$$M_2 \stackrel{\text{def}}{=} F(-h_m).$$

H5.5. $B(1) \neq 0$, c'est à dire que le système est à gain statique non nul.

Remarque 5.2

1. L'hypothèse H5.1, concernant la structure du système, est assez courante pour l'identification des systèmes linéaires.
2. L'hypothèse H5.2 assure la contrôlabilité de la fonction de transfert $B(q^{-1})/A(q^{-1})$. Elle est nécessaire pour appliquer le lemme d'excitation du chapitre 1.
3. L'hypothèse H5.4 n'est pas restrictive puisque h_m peut être de manière arbitraire assez grand. Dans (Bai, 2002), les non linéarités sont supposées symétriques, c'est-à-dire $h_1 = -h_2 = a$ et $M_1 = -M_2 = I$. La méthode proposée implique alors un problème de minimisation en fonction du seul paramètre inconnu a . Pour avoir un estimé du paramètre a , une recherche graphique du minimum d'une fonction de a , a été utilisée. Pour que cette recherche soit fructueuse, cela suppose la connaissance d'un majorant du paramètre a (ce ci n'a pas été mentionnée explicitement dans (Bai, 2002) !). Cette supposition n'est pas moins restrictive que l'hypothèse H5.4. L'une des contributions majeures de notre travail est le fait que h_1 peut être différent de $-h_2$, et M_1 peut être aussi différent de $-M_2$. L'approche présentée dans (Bai, 2002) est très difficile, voire impossible, à appliquer dans ce cas.
4. Les systèmes tels que $B(1) = 0$ incluent un opérateur de dérivation $(1 - q^{-1})$. Cette propriété est particulièrement rare dans les applications pratiques et quand elle se présente elle est facilement détectable. Dans ce cas l'hypothèse H5.5 peut être assurée en prenant comme signal d'entrée $(1 - q^{-1})u(t)$.
5. A l'exception des hypothèses H5.1-H5.4, le système peut être quelconque. Ainsi, les paramètres de la dynamique linéaire (a_i, b_i) sont inconnus, et les premiers coefficients (b_1, b_2, \dots) peuvent être nuls, ce qui implique que le vrai retard pur du système est inconnu (mais il n'est pas plus grand que n).

5.3.2. Objectif d'identification

Notre objectif est de concevoir un schéma d'identification qui permet d'estimer, de manière consistante, les paramètres du sous système linéaire $B(q^{-1})/A(q^{-1})$ et ceux de la caractéristique du gain non linéaire $F(\cdot)$. Deux difficultés majeures doivent être surmontées:

- L'effet mémoire engendré par les non linéarités considérées;
- Le signal $u(t)$ est non mesurable et le schéma d'identification doit utiliser seulement les séquences $\{v(t)\}$ et $\{y(t)\}$ (Fig. 5.3).

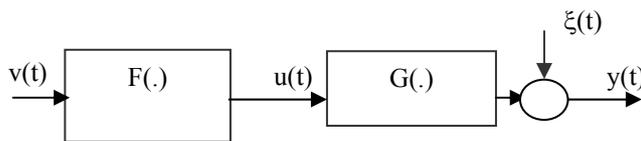


Fig. 5.3 Modèle d' Hammerstein

5.4. IDENTIFICATION DU SOUS SYSTÈME LINÉAIRE

Dans cette section on traite le problème d'identification du sous système linéaire. L'approche proposée opère en trois étapes : d'abord, dans le paragraphe 5.1, une reparamétrisation est utilisée pour réduire le nombre de paramètres à estimer; la représentation obtenue est transformée, dans le paragraphe 5.2, pour surmonter le fait que le signal interne $u(t)$ n'est pas mesurable. Cette nouvelle représentation fait apparaître, de manière linéaire, les paramètres du sous système linéaire et sera donc utilisée pour les estimer de manière classique. Finalement, dans la sous-section 5.3, on conçoit une séquence d'entrée, et on montre qu'elle est capable d'assurer une meilleure convergence des estimés des paramètres vers leurs vraies valeurs.

5.4.1. Reparamétrisation du modèle

Remarquons d'abord que le problème d'identification considéré n'a pas qu'une seule solution. En effet, si le couple $(B(q^{-1})/A(q^{-1}), F(\cdot))$ est une solution, alors pour tout réel $\alpha \neq 0$, le couple $(\alpha B(q^{-1})/A(q^{-1}), F(\cdot)/\alpha)$ est aussi une solution. Soit $(B^*(q^{-1})/A(q^{-1}), F^*(\cdot))$ la solution correspondante à $\alpha = M_2 - M_1$, c'est-à-dire

$$B^*(q^{-1}) \stackrel{\text{def}}{=} (M_2 - M_1)b_1 q^{-1} + \dots + (M_2 - M_1)b_n q^{-n}$$

$$\stackrel{def}{=} b_1^* q^{-1} + b_2^* q^{-2} + \dots + b_n^* q^{-n} \quad (5.8)$$

$$F^*(.) \stackrel{def}{=} \frac{F(.)}{M_2 - M_1} \quad (5.9)$$

A partir des figures 5.1 et 5.2, on voit clairement que :

$$F = HR_{(M_1, M_2, h_1, h_2)} \Rightarrow F^* = HR_{(m_1, m_2, h_1, h_2)} \quad (5.10)$$

$$F = H_{(s, h_1, h_2)} \Rightarrow F^* = H_{(s, h_1, h_2)} \quad (5.11)$$

avec:

$$m_1 = \frac{M_1}{M_2 - M_1}, m_2 = \frac{M_2}{M_2 - M_1}, s = \frac{S}{M_2 - M_1} \quad (5.12)$$

où les deux premières relations sont valables pour les deux types de systèmes de type *HS-HR* et *HS-H* alors que la dernière est valable pour les systèmes de type *HS-H* seulement. Remarquons qu'en plus on a les relations suivantes :

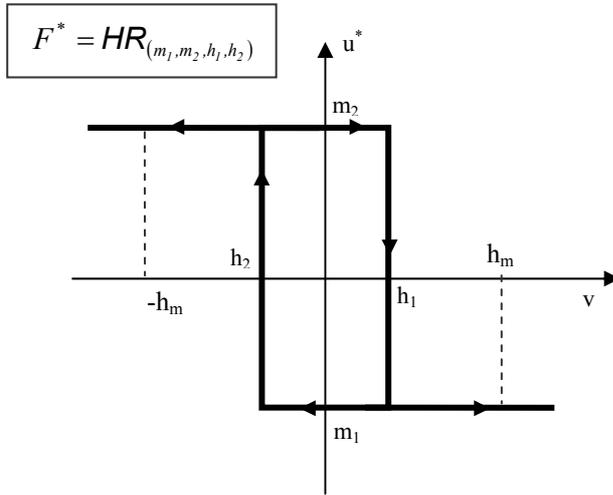
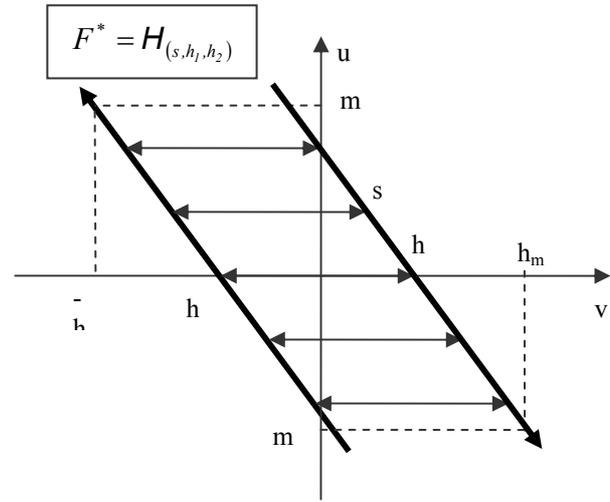
$$-m_1 + m_2 = 1 \quad (5.13)$$

$$S = \frac{M_1}{h_m - h_1} = \frac{M_2}{h_m - h_2} \quad (5.14)$$

$$s = \frac{m_1}{(h_m - h_1)} = \frac{m_2}{(h_m - h_2)} \quad (5.15)$$

Les deux dernières concernent *HS-H* seulement.

Les figures suivantes illustrent les allures des nouvelles fonctions non linéaires obtenues:


 Figure 5.4 Non linéarité F^* de type Hystérésis - Relais

 Figure 5.5 Non linéarité F^* de type Hystérésis - Transformation

En utilisant (5.8)-(5.9), le système (5.7a-b) peut être écrit comme suit :

$$A(q^{-1})y(t) = B^*(q^{-1})u^*(t) + \xi(t) \quad (5.16a)$$

$$u^*(t) \stackrel{def}{=} F^*(v(t)) \quad (5.16b)$$

La nouvelle représentation du système (5.16a-b) invoque pour le cas d'une non linéarité relais-hystérésis une caractéristique $HR_{(m_1, m_2, h_1, h_2)}$ qui, grâce à (5.13), fait appel à trois paramètres seulement (grâce à 5.13), au lieu de quatre qu'on avait dans la représentation initiale (5.7a-b) invoquant la caractéristique $HR_{(M_1, M_2, h_1, h_2)}$. Par contre, le signal interne, $u^*(t)$, est toujours non mesurable. Pour contourner ce problème, la représentation du système (5.16a-b) sera donc transformée dans le paragraphe 5.4.2.

5.4.2. Conception d'une forme de régression pour le sous système linéaire:

Soit $\{y_1(t)\}$ la solution de (5.16a) correspondant à la séquence d'entrée suivante:

$$v(0) = 0 \text{ et } v(t) = v_1(t) \stackrel{def}{=} h_m, \text{ pour tout } t > 0 \quad (5.17)$$

Alors, à partir de (5.10)-(5.11) et les fig. 5.1 et 5.2 on déduit que, pour tout $t \geq n$:

$$A(q^{-1})y_1(t) = B^*(q^{-1})u_1^*(t) + \xi_1(t) \quad (5.18a)$$

$$u_1^*(t) = F^*(h_m) \stackrel{def}{=} m_1 \quad (5.19)$$

où $\xi_l(t)$ représente la réalisation de $\xi(t)$ pour cette expérience. Après le régime transitoire, la sortie non perturbée se stabilise et garde une valeur constante, et en moyennant, sur un intervalle $1 \leq t \leq L$, les deux termes de l'équation (5.18a), on obtient:

$$A(q^{-1})\bar{y}_1(L) = B^*(q^{-1})m_l + \bar{\xi}_1(L) \quad (5.20a)$$

L'ergodicité de $\{\xi_l(t)\}$ implique que $\bar{\xi}_1(L) \rightarrow E(\xi(t)) = 0$ lorsque $L \rightarrow \infty$ (a.p. 1). Soit \bar{y}_1 la limite de $\bar{y}_1(L)$ lorsque $L \rightarrow \infty$. On peut déduire à partir de (5.20a) qu'une telle limite existe et satisfait l'équation :

$$A(1)\bar{y}_1 = B^*(1)m_l \quad (5.20b)$$

En pratique, \bar{y}_1 est déterminée à partir d'un nombre assez grand d'échantillons $\{y(t); t=1, \dots, L\}$. la soustraction de (5.20) de (5.16a) donne:

$$A(q^{-1})(y(t) - \bar{y}_1) = B^*(q^{-1})(u^*(t) - m_l) + \xi(t) \quad (5.21)$$

Pour alléger les équations, on utilisera les notations suivantes:

$$\tilde{y}(t) \stackrel{def}{=} y(t) - \bar{y}_1 \quad (5.22a)$$

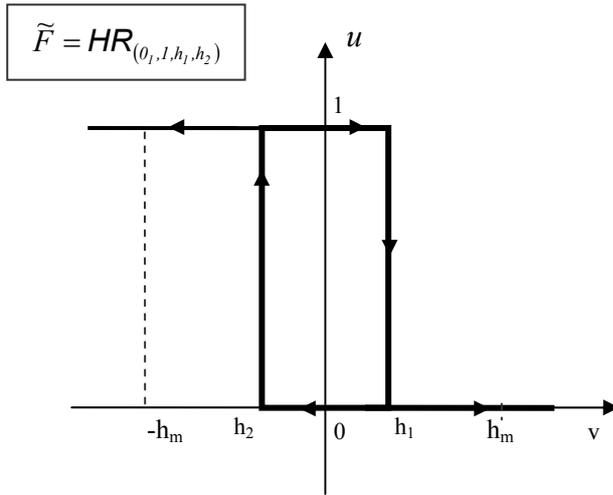
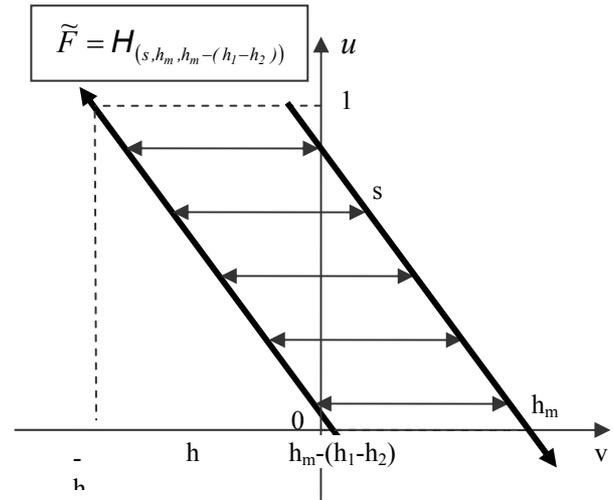
$$\tilde{u}(t) \stackrel{def}{=} u^*(t) - m_l \quad (5.22b)$$

Ainsi le système à identifier (5.16a-b) peut être écrit maintenant sous la forme compacte suivante en utilisant (5.22a-b) et (5.21):

$$A(q^{-1})\tilde{y}(t) = B^*(q^{-1})\tilde{u}(t) + \xi(t) \quad (5.23)$$

$$\tilde{u}(t) = \tilde{F}(v(t))$$

Les figures suivantes illustrent l'allure de la caractéristique $\tilde{F}(v)$ pour les deux types de non linéarités considérées:


 Figure 5.6 Non linéarité \tilde{F} de type Hystérésis -Relais

 Figure 5.7 Non linéarité \tilde{F} de type Hystérésis

Cette nouvelle représentation du système possède les propriétés clés suivantes :

$$v(t) = h_m \Rightarrow \tilde{u}(t) = 0 \quad (5.24a)$$

$$v(t) = -h_m \Rightarrow \tilde{u}(t) = 1 \quad (5.24b)$$

En quelque sorte, la séquence interne $\{\tilde{u}(t)\}$ dans (5.24a-b) devient parfaitement mesurable, si la séquence $\{v(t)\}$ prend ses valeurs dans l'ensemble $\{-h_m, h_m\}$. L'identification paramétrique des polynômes $A(q^{-1})$ et $B^*(q^{-1})$ peut être alors effectuée en se basant sur l'équation d'erreur (5.23). Cette dernière possède la forme de régression suivante:

$$\tilde{y}(t) = \tilde{\Phi}(t)^T \theta^* + \xi(t) \quad (5.25)$$

avec:

$$\tilde{\Phi}(t)^T = [-\tilde{y}(t-1) \quad \dots \quad -\tilde{y}(t-n) \quad \tilde{u}(t-1) \quad \dots \quad \tilde{u}(t-n)] \quad (5.26a)$$

$$\theta^* = [a_1 \quad \dots \quad a_n \quad b_1^* \quad \dots \quad b_n^*] \quad (5.26b)$$

où b_i^* ($i=1, \dots, n$) représentent les coefficients de $B^*(q^{-1})$.

5.4.3. Estimation des paramètres du sous système linéaire

Dans ce paragraphe la forme de régression (5.25) sera utilisée pour estimer le vecteur des paramètres inconnus, θ^* . On a déjà indiqué que la séquence d'entrée $\{v(t)\}$ doit prendre ses valeurs dans $\{-h_m, h_m\}$, pour que la séquence interne $\tilde{u}(t)$ soit mesurable (c'est-à-dire que ses valeurs

appartiennent à $\{0, 1\}$ d'après (5.24a-b)). Par acquisition d'un nombre suffisamment grand de données $\{v(t), y(t); 1 \leq t \leq N\}$, les paramètres peuvent être estimés en utilisant par exemple l'algorithme classique des moindres carrés (en version batch) :

$$\hat{\theta}(N) = \left[\frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{\Phi}(i)^T \right]^{-1} \left[\frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{y}(i)^T \right] \quad (5.27)$$

Ceci n'est possible que si l'inverse de la matrice $\left[\frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{\Phi}(i)^T \right]$ existe. Pour cela le vecteur de régression $\tilde{\Phi}(t)$ doit posséder la propriété d'excitation persistante. Ce ci est intimement lié à la nature de l'excitation qu'on utilisera pour l'estimation.

5.4.4. Développement d'une séquence d'entrée excitante persistante

Le choix de la séquence d'entrée $\{v(t)\}$ sera approprié si les trois conditions suivantes sont satisfaites :

- (i) $\{v(t)\}$ doit être facilement réalisable.
- (ii) Elle doit prendre ses valeurs dans l'ensemble $\{-h_m, h_m\}$, ainsi la séquence interne $\{\tilde{u}(t)\}$ est connue.
- (iii) Le vecteur de régression résultant $\tilde{\Phi}(t)$ doit posséder la propriété d'excitation persistante, dans un certain sens bien défini (Propositions 5.1).

En s'inspirant du Lemme 5.1, cité en début de ce chapitre, on propose comme séquence d'entrée, la séquence périodique suivante (de période $T = 4n$) où k est un entier, $t_k = kT$, $t_k \leq t < t_{k+1}$:

$$v_2(t) = \begin{cases} -h_m & \text{pour } t = t_k + 2n \\ h_m & \text{ailleurs} \end{cases} \quad (5.28a)$$

Grâce à (5.24a-b), le signal interne $\{\tilde{u}(t)\}$ prend l'expression suivante :

$$\tilde{u}(t) = \begin{cases} 1 & \text{pour } t = t_k + 2n \\ 0 & \text{ailleurs} \end{cases} \quad (5.28b)$$

5.4.5. Analyse de la convergence des estimés des paramètres du sous système linéaire

Soit $\{\tilde{z}(t)\}$ la séquence interne définie comme suit :

$$A(q^{-1})\tilde{z}(t) = \tilde{u}(t) \quad (5.29a)$$

Introduisons le vecteur d'état suivant:

$$\tilde{Z}(t) = [\tilde{z}(t) \dots \tilde{z}(t-2n+1)]^T \in \mathfrak{R}^{2n} \quad (5.29b)$$

Ce vecteur possède alors la propriété d'excitation persistante comme le stipule la proposition suivante:

Proposition 5.1 (Excitation persistante de $\{\tilde{Z}(t)\}$) Considérons le système (5.7a-b), avec les hypothèses H5.1-H5.5, la non linéarité $F(\cdot) \in \{HR_{(M_1, M_2, h_1, h_2)}, H_{(S, h_1, h_2)}\}$. Le système est excité par la séquence (5.28a). Ce système peut avoir la représentation d'équation d'erreur (5.23) et la forme de régression (5.25). Alors, la séquence de vecteurs $\{\tilde{Z}(t)\}$ possède la propriété d'excitation persistante, c'est-à-dire, qu'il existe une constante réelle $\lambda > 0$ telle que, pour tout entier k:

$$\sum_{i=0}^{4n-1} \tilde{Z}(t_k+i) \tilde{Z}(t_k+i)^T \geq \lambda I_{2n} \quad \square$$

Preuve. Ce résultat est un cas particulier du Lemme 5.1 en utilisant les équations (5.23)-(5.25).

Proposition 5.2 (Consistance) Considérons le schéma d'identification décrit dans proposition 5.1. En appliquant l'algorithme d'estimation (5.27) à la régression (5.25)-(5.26), on obtient un estimé consistant $\hat{\theta}(t)$ c'est-à-dire

$$\hat{\theta}(N) \rightarrow \theta^* \text{ quand } N \rightarrow \infty \text{ (a.p. 1)} \quad \square$$

Preuve. L'étape clé sera de montrer que la propriété d'excitation persistante de $\tilde{Z}(t)$ s'applique aussi au vecteur de régression $\tilde{\Phi}(t)$. En opérant $B^*(q^{-1})$ sur (5.29a) et en utilisant (5.23) on obtient:

$$\tilde{y}(t) = B^*(q^{-1})\tilde{z}(t) + \eta(t) \quad (5.30a)$$

où $\eta(t)$ est définie par:

$$A(q^{-1})\eta(t) = \xi(t) \quad (5.30b)$$

Notons tout simplement que les équations (5.29a) et (5.30a) définissent la forme différentielle contrôlable du système (Giri et al, 2001c). Ceci nous permet d'écrire la relation suivante entre $\tilde{\Phi}(t)$ et $\tilde{Z}(t)$:

$$\begin{aligned}
 \tilde{\Phi}(t) &= \begin{bmatrix} 0 & -b_1^* & -b_2^* & & & -b_n^* & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -b_1^* & -b_2^* & & & -b_n^* & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 0 & -b_1^* & -b_2^* & \cdots & \cdots & -b_n^* \\ 1 & a_1 & a_2 & a_3 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & 1 & a_1 & a_2 & \ddots & & \ddots & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & 1 & a_1 & a_2 & a_3 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 & a_1 & a_2 & \cdots & \cdots & a_n \end{bmatrix} \tilde{Z}(t) + \begin{bmatrix} -\eta(t) \\ -\eta(t-1) \\ \vdots \\ -\eta(t-n+2) \\ -\eta(t-n+1) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \\
 &\stackrel{\text{def}}{=} \Omega \tilde{Z}(t) + \Xi(t)
 \end{aligned} \tag{5.31}$$

où Ω est en fait la matrice de Sylvester correspondant au couple $-B^*(q^{-1})$ et $A(q^{-1})$ (voir (Giri et al, 2001c)). Puisque ces deux derniers sont premiers entre eux, Ω est alors non singulière. On déduit à partir de (5.31) que:

$$\begin{aligned}
 \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{\Phi}(i)^T &= \Omega \left[\frac{1}{N} \sum_{i=1}^N \tilde{Z}(i) \tilde{Z}(i)^T \right] \Omega^T + \Omega \left[\frac{1}{N} \sum_{i=1}^N \tilde{Z}(i) \Xi(i)^T \right] \\
 &+ \left[\frac{1}{N} \sum_{i=1}^N \Xi(i) \tilde{Z}(i)^T \right] \Omega^T + \left[\frac{1}{N} \sum_{i=1}^N \Xi(i) \Xi(i)^T \right]
 \end{aligned} \tag{5.32a}$$

En utilisant la proposition 5.1 et l'ergodicité, il découle à partir de (5.32a) que (avec une probabilité 1):

$$\begin{aligned}
 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{\Phi}(i)^T &\geq \lambda \sigma I_{2n} + \Omega E(\tilde{Z}(i) \Xi(i)^T) \\
 &+ E(\Xi(i) \tilde{Z}(i)^T) \Omega^T + E(\Xi(i) \Xi(i)^T)
 \end{aligned} \tag{5.32b}$$

où $\sigma > 0$ représente la plus petite valeur singulière de la matrice qui est définie positive $\Omega \Omega^T$, et $\lambda > 0$. Rappelons que le vecteur $\tilde{Z}(t)$ est totalement déterminé par l'entrée $\tilde{u}(t)$, selon l'équation (5.29a), alors que le vecteur $\Xi(t)$ est totalement déterminé par $\xi(t)$ (voir (5.30b)). Comme les séquences $\tilde{u}(t)$ et $\xi(t)$ sont indépendantes, d'un point de vue stochastique, le deuxième et troisième terme du membre à droite de l'équation (5.32b) sont alors nuls. Et puisque le dernier terme est une matrice définie non négative, on déduit à partir de (5.32b) que:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \tilde{\Phi}(i) \tilde{\Phi}(i)^T \geq \lambda \sigma I_{2n} \tag{a.p.1}$$

Il est bien connu que cette excitation persistante avec le fait que $\{\xi(t)\}$ est un processus stochastique impliquent que l'algorithme des moindres carrés (5.27) est consistant (voir (Giri et al, 1993)).□

5.4.6. Simulation

Pour illustrer les performances de l'approche proposée, nous avons procédé à des simulations en considérant un modèle de Hammerstein avec les caractéristiques suivantes :

$$A(q^{-1})=1 - 1,5 q^{-1} + 0,56 q^{-2} ; B(q^{-1})= q^{-1} - 0.5q^{-2} \quad (5.33)$$

où $\xi(t)$ est une séquence de variables aléatoires identiquement distribuées et indépendantes dans $[-0.5, 0.5]$. Les non linéarités considérées sont définies par :

$$\text{Pour } F=H_{(S,h_1,h_2)} : S = 0.5, h_1=2, h_2 = -3, \text{ et on a choisi } h_m = 4, \quad (5.34a)$$

$$\text{Pour } F=HR_{(M_1,M_2,h_1,h_2)} : h_1 = 1, h_2 = -1.5, M_1 = 2, M_2 = -3, \text{ et on a choisi } h_m = 2 \quad (5.34b)$$

Pour (5.34a) on a $(M_1=F(h_m)=1 \text{ et } M_2=F(-h_m)= -0.5)$ qui sont implicitement utilisés (voir Fig. 5.2).

On obtient alors, à partir de (5.8) :

$$B^*(q^{-1})= - 1.5q^{-1} - 0.75q^{-2} \text{ pour } F=H_{(S,h_1,h_2)} \quad (5.35a)$$

$$B^*(q^{-1})= - 5q^{-1} + 2.5q^{-2} \text{ pour } F=HR_{(M_1,M_2,h_1,h_2)} \quad (5.35b)$$

D'autre part, à partir de (5.9)-(5.12) on a :

$$\text{Pour } F=H_{(S,h_1,h_2)} : F^*=H_{(s,h_1,h_2)} \text{ avec } m_1=-2/3, s=-1/3, h_1=2, h_2= -3, \quad (5.36a)$$

$$\text{Pour } F=H_{(S,h_1,h_2)} : F^*=HR_{(m_1,m_2,h_1,h_2)} \text{ avec } m_1=-0.4, h_1=1, h_2=-1.5, \quad (5.36b)$$

avec $m_2=1+m_1$. Les vecteurs des paramètres résultant à estimer sont :

$$\theta^* = [-1.5 \ 0.56 \ -1.5 \ 0.75]^T \text{ pour } F=H_{(S,h_1,h_2)} \quad (5.37a)$$

$$\theta^* = [-1.5 \ 0.56 \ -5 \ 2.5]^T \text{ pour } F=HR_{(M_1,M_2,h_1,h_2)} \quad (5.37b)$$

Notez que les non linéarités sont volontairement choisies asymétriques, à l'inverse de ce qui a été supposé dans (Bai, 2002). L'algorithme (5.28a-b) est appliqué pour trouver un estimé $\hat{\theta}(N)$, pour les deux cas de non linéarités considérés. Les estimés obtenus pour deux valeurs de N sont rapportés dans les tables 1 et 2.

Table 1: $F^* = H_{(s, h_1, h_2)}$	
Vecteur des vrais paramètres θ^*	$[-1.5 \ 0.56 \ -1.5 \ 0.75]^T$
Estimé $\hat{\theta}(N)$ pour $N=400$	$[-1.472 \ 0.541 \ -1.528 \ 0.756]^T$
Estimé $\hat{\theta}(N)$ pour $N=10000$	$[-1.506 \ 0.566 \ -1.501 \ 0.766]^T$

Table 2: $F^* = HR_{(M_1, M_2, h_1, h_2)}$	
Vecteur des vrais paramètres θ^*	$[-1.5 \ 0.56 \ -5 \ 2.5]^T$
Estimé $\hat{\theta}(N)$ pour $N=400$	$[-1.478 \ 0.533 \ -4.939 \ 2.334]^T$
Estimé $\hat{\theta}(N)$ pour $N=10000$	$[-1.500 \ 0.559 \ -4.964 \ 2.534]^T$

Les estimés obtenus sont très proches de leur vraies valeurs surtout lorsque N est suffisamment grand, malgré la présence d'une perturbation en sortie. Ceci conclut la partie concernant l'identification des paramètres du sous système linéaire.

5.5. IDENTIFICATION DU GAIN NON LINEAIRE

Soit $\hat{\theta}(N)$ l'estimé de θ^* obtenu à partir de (5.27), pour un N assez grand. Soient $\hat{A}_N(q^{-1})$ et $\hat{B}_N(q^{-1})$ les estimés de $A(q^{-1})$ et $B^*(q^{-1})$ déduis à partir de $\hat{\theta}(N)$, à travers (5.26b). Ces estimés seront maintenant utilisés pour déterminer ceux de l'élément non linéaire $F^*(.) \in \{HR_{(m_1, m_2, h_1, h_2)}, H_{(s, h_1, h_2)}\}$.

5.5.1. Estimation du paramètre m_1

Nous avons déjà mentionné que si $v(t)=v_1(t)=h_m$ (pour tout t), on obtient à partir de (5.19), $u_1^*(t) = F^*(h_m) = m_1$. En plus en moyennant les deux termes de (5.18a) sur un intervalle $[1, L]$, on obtient l'équation (5.20) que l'on écrit, pour plus de clarté, sous la forme:

$$A(1)\bar{y}_1(L) = B^*(1)m_1 + \bar{\xi}(L) \quad (5.38)$$

Ceci suggère comme estimé pour m_1 , $\hat{m}_1(L, N)$ défini comme suit:

$$\hat{m}_1(L, N) = \frac{\hat{A}_N(1)}{\hat{B}_N(1)} \bar{y}_1(L), \quad (\text{pour } L \text{ suffisamment grand}) \quad (5.39)$$

Le processus d'estimation, ainsi construit est décrit et analysé dans la proposition suivante:

Proposition 5.3. Considérons le système (5.7a-b) soumis aux hypothèses H5.1-H5.5, où $F(.) \in \{HR_{(m_1, m_2, h_1, h_2)}, H_{(s, h_1, h_2)}\}$. Le système peut être représenté également par l'équation d'erreur (5.16a-b) où $F^* \in \{HR_{(m_1, m_2, h_1, h_2)}, HR_{(s, h_1, h_2)}\}$, et (m_1, m_2, s) sont définis par (5.12). Supposons que le système est excité par une entrée constante $v(t) = h_m$ (pour tout t). Son comportement moyen est décrit par l'équation (5.38), qui permet d'estimer, à partir de la règle (5.39), le paramètre inconnu m_1 . L'estimé $\hat{m}_1(L, N)$ ainsi obtenu converge (a.p 1) vers sa vraie valeur m_1 quand L et $N \rightarrow \infty$.

□

Preuve. A partir de (5.38) on a successivement:

$$\begin{aligned} B^*(1)(\hat{m}_1(L, N) - m_1) &= B^*(1)\hat{m}_1(L, N) - A(1)\bar{y}_1(L) + \bar{\xi}(L) \\ &= \hat{B}_N(1)\hat{m}_1(L, N) - (\hat{B}_N(1) - B^*(1))\hat{m}_1(L, N) - A(1)\bar{y}_1(L) + \bar{\xi}(L) \end{aligned} \quad (5.40)$$

En utilisant (5.39), on déduit à partir de (5.40) que:

$$B^*(1)(\hat{m}_1(L, N) - m_1) = -(\hat{B}_N(1) - B^*(1))\hat{m}_1(L, N) + (\hat{A}_N(1) - A(1))\bar{y}_1(L) + \bar{\xi}(L) \quad (5.41)$$

D'après la proposition 5.2, les deux premiers termes à droite de (5.41) convergent vers zéro (avec probabilité 1) quand $N \rightarrow \infty$. Le dernier disparaît quand $L \rightarrow \infty$ puisque $\xi(t)$ est ergodique stationnaire et de moyenne nulle. Ceci implique alors que le terme à gauche de (5.41) converge vers zéro (a.p. 1), ceci termine la preuve de la proposition 5.3 □.

5.5.2. Estimation des paramètres du gain non linéaire

Dans ce paragraphe on propose une approche pour estimer les paramètres h_1, h_2 pour les systèmes de type HS-HR, et les paramètres s, h_1, h_2 pour les systèmes de type HS-H. Notons d'abord que $u^*(t) = F^*(v(t))$, avec $F^* = HR_{(m_1, m_2, h_1, h_2)}$ ou $F^* = H_{(s, h_1, h_2)}$, et qu'à partir de (5.22b) on déduit que:

$$\tilde{u}(t) = \tilde{F}(v(t)) \quad (5.42a)$$

avec $\tilde{F} \stackrel{def}{=} HR_{(0,1, h_1, h_2)}$ pour HS-HR (5.42b)

ou $\tilde{F} \stackrel{def}{=} H_{(s, h_m, h_m - (h_1 - h_2))}$ pour HS-H (5.42c)

Soit la séquence d'entrée périodique (de période T) $v(t) = v_3(\alpha, t)$ définie par:

$$v_3(k) = h_m - k\Delta_1 \quad \text{pour } k = 0, 1, \dots, \alpha T \quad (5.43a)$$

$$\text{et } v_3(k) = -h_m + k\Delta_2 \quad \text{pour } k = \alpha T + 1, \dots, T \quad (5.43b)$$

où $0 < \alpha < 1$ est tel que, αT soit un entier, et $\Delta_1 \stackrel{\text{def}}{=} \frac{2h_m}{\alpha T}$, $\Delta_2 \stackrel{\text{def}}{=} \frac{2h_m}{(T - \alpha T)}$; la période $T > 0$ et le rapport cyclique α sont quelconques et sont choisis par l'utilisateur. La Fig. 5.8 illustre la forme de la séquence $v(t)$ ainsi définie.

Il existe des entiers $t_1, t_2, t'_1, t'_2, \in]0, T[$ tels que:

$$h_m - t_1 \Delta_1 = h_2 + \varepsilon_1(\alpha, T) \quad \text{avec } 0 \leq \varepsilon_1 < \Delta_1 \quad (5.44a)$$

$$-h_m + (t_2 - \alpha T) \Delta_2 = h_1 - \varepsilon_2(\alpha, T) \quad \text{avec } 0 \leq \varepsilon_2 < \Delta_2 \quad (5.44b)$$

$$h_m - t'_1 \Delta_1 = h_m - (h_1 - h_2) + \varepsilon_1(\alpha, T) \quad \text{avec } 0 \leq \varepsilon_1 < \Delta_1 \quad (5.44c)$$

$$-h_m + (t'_2 - \alpha T) \Delta_2 = -h_m + (h_1 - h_2) - \varepsilon_2(\alpha, T) \quad \text{avec } 0 \leq \varepsilon_2 < \Delta_2 \quad (5.44d)$$

Les instants t_1, t_2 (respect t'_1, t'_2) approximent ceux auxquels la séquence $v(t) = v_3(\alpha, t)$ approche les valeurs particulières h_1 et h_2 quand $F = HR_{(M_1, M_2, h_1, h_2)}$ (respect $(h_m - (h_1 - h_2), -h_m + (h_1 - h_2))$ quand $F = H_{(S, h_1, h_2)}$). Les équations (5.44a-d) montrent que, pour un α donné tel que $0 < \alpha < 1$, la période T doit être assez grande pour que les erreurs d'estimations $\varepsilon_1(\alpha, T)$ et $\varepsilon_2(\alpha, T)$ soient suffisamment négligeables.

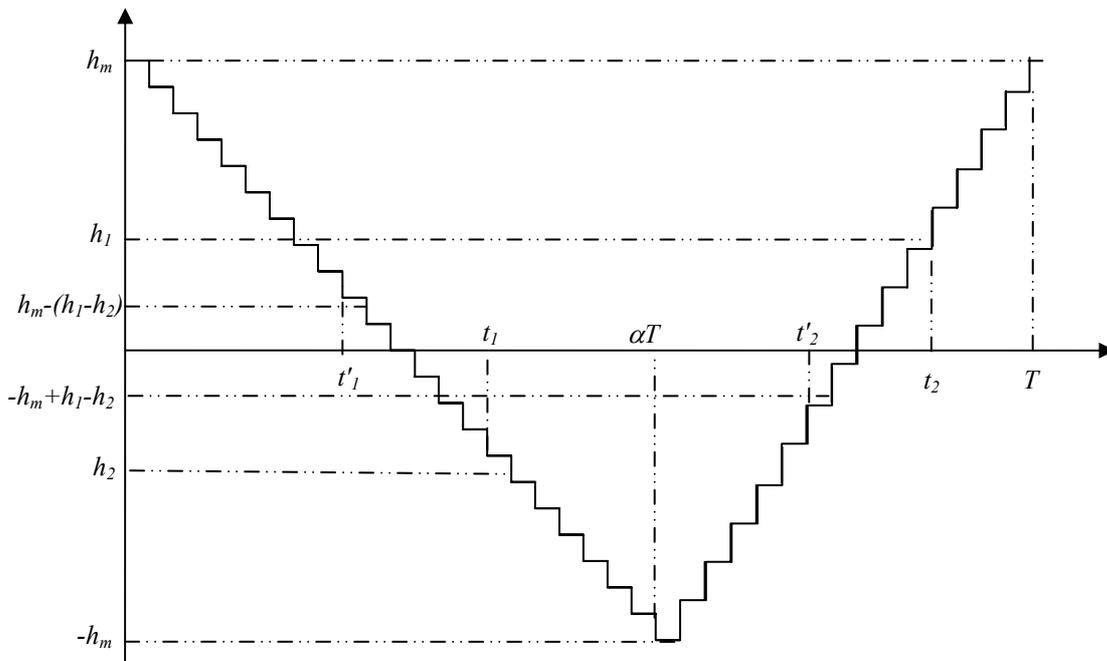


Figure 5.8. Forme de la séquence d'entrée définie par (5.43a-b)

A partir de (5.22b) et des Fig. 5.2 et Fig. 5.3, on déduit que, pour $t \geq T$, $\tilde{u}(t)$ devient périodique de période T . Le régime transitoire durant la première période $[0, T[$ peut être ignoré en prenant comme origine des temps $t=T$. Ainsi, $\tilde{u}(t)$ sera définie sur $t \in [0, T[$ comme suit :

$$\tilde{u}(t) = \gamma(\alpha, t) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{si } t_1 < t \leq t_2 \\ 0 & \text{ailleurs} \end{cases} \quad \text{pour } F^* = HR_{(m_1, m_2, h_1, h_2)} \quad (5.45a)$$

$$\tilde{u}(t) = \delta(\alpha, t) \stackrel{\text{def}}{=} \begin{cases} \frac{t-t'_1}{\alpha T - t'_1} & \text{si } t'_1 < t \leq \alpha T \\ 1 & \text{si } \alpha T < t \leq t'_2 \\ -\frac{t-t'_2}{T-t'_2} + 1 & \text{si } t'_2 < t \leq T \\ 0 & \text{ailleurs} \end{cases} \quad \text{pour } F^* = H_{(s, h_1, h_2)} \quad (5.45b)$$

Ceci est illustré sur la fig. 5.9.

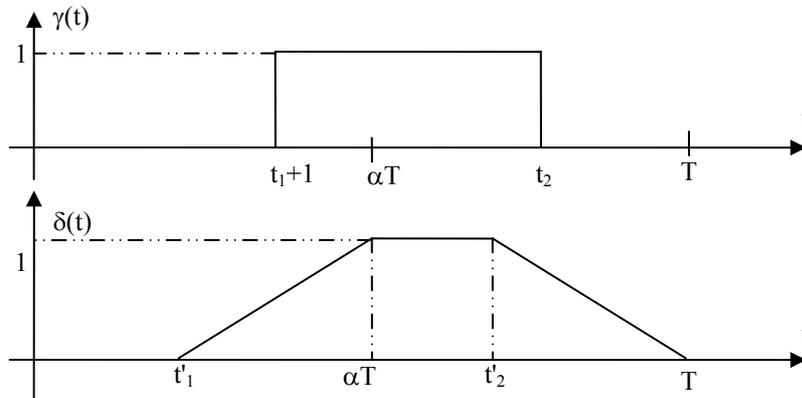


Figure 5.9. Forme des séquences $\gamma(\alpha, t)$ et $\delta(\alpha, t)$

Soient $\bar{\gamma}(\alpha), \bar{\delta}(\alpha)$ les moyennes temporelles de $\gamma(\alpha, t), \delta(\alpha, t)$ respectivement, sur l'intervalle $[0, T[$. Il est alors facile de démontrer à partir des équations (5.45a-b) que:

$$\bar{\gamma}(\alpha) = \frac{t_2 - t_1}{T} + \varepsilon(\alpha, T) \quad \text{pour } F^* = HR_{(m_1, m_2, h_1, h_2)} \quad (5.46a)$$

$$\bar{\delta}(\alpha) = \frac{\alpha T - t'_1}{2T} + \frac{t'_2 - \alpha T}{T} + \frac{T - t'_2}{2T} + \varepsilon(\alpha, T) \quad \text{pour } F^* = H_{(s, h_1, h_2)} \quad (5.46b)$$

Dans (5.46a-b), et toutes les équations sous jacentes, le terme $\varepsilon(\alpha, T)$ représente une erreur générique, résultante de $\varepsilon_1(\alpha, T)$ et $\varepsilon_2(\alpha, T)$, telle que pour un $0 < \alpha < 1$ fixé:

$$\lim_{T \rightarrow \infty} \varepsilon(\alpha, T) = 0 \quad (5.47)$$

En utilisant (5.44a-d) et les expressions de Δ_1 et Δ_2 , il s'ensuit à partir de (5.46a-b) que:

$$\bar{\gamma}(\alpha) = \frac{1}{2} + \frac{h_1}{2h_m}(1-\alpha) + \frac{h_2}{2h_m}\alpha + \varepsilon(\alpha, T) \quad \text{pour } F^* = HR_{(m_1, m_2, h_1, h_2)} \quad (5.48a)$$

$$\bar{\delta}(\alpha) = \frac{1}{2} + \frac{(h_1 - h_2)}{4h_m}(1 - 2\alpha) + \varepsilon(\alpha, T) \quad \text{pour } F^* = H_{(s, h_1, h_2)} \quad (5.48b)$$

Et en utilisant (5.13) et (5.15), (5.48b) peut être réécrite comme suit :

$$\bar{\delta}(\alpha) = \frac{1}{2} + \frac{(2h_m + 1/s)}{4h_m}(1 - 2\alpha) + \varepsilon(\alpha, T) \quad (5.48c)$$

Notez qu'on a exprimé la moyenne du signal interne, non mesurable, en fonction des paramètres inconnus qu'on veut estimer. Nous avons besoin maintenant d'estimer ces valeurs moyennes $\bar{\gamma}(\alpha), \bar{\delta}(\alpha)$ pour pouvoir estimer les paramètres inconnus des non linéarités. Pour cela, on procède au calcul de la moyenne des deux parties de (5.23) sur $[0, MT]$, quand $v(t) = v_3(\alpha, t)$, et après une période transitoire due à la dynamique du sous-système linéaire. De cette manière, on obtient :

$$A(q^{-1})\bar{y}(M) = B^*(q^{-1})\bar{u}(M) + \bar{\xi}(M) \quad (5.49)$$

Puisque $\tilde{u}(t)$ est périodique de période T , $\bar{u}(M)$ coïncide avec la moyenne de $\tilde{u}(t)$ sur $[0, T]$. Ainsi,

$$\bar{u}(M) = \begin{cases} \bar{\gamma}(\alpha) & \text{pour } F^* = HR_{(m_1, m_2, h_1, h_2)} \\ \bar{\delta}(\alpha) & \text{pour } F^* = H_{(s, h_1, h_2)} \end{cases} \quad (5.50)$$

En utilisant (5.49), (5.50) et l'érgodicité de $\{\xi(t)\}$, on obtient alors:

$$\lim_{M \rightarrow \infty} [B^*(1)\bar{\gamma}(\alpha) - A(1)\bar{y}(M)] = 0 \quad \text{pour } F^* = HR_{(m_1, m_2, h_1, h_2)} \quad (\text{a.p. } 1) \quad (5.50a)$$

$$\lim_{M \rightarrow \infty} [B^*(1)\bar{\delta}(\alpha) - A(1)\bar{y}(M)] = 0 \quad \text{pour } F^* = HR_{(s_2, h_1, h_2)} \quad (\text{a.p. } 1) \quad (5.50b)$$

Maintenant on dispose d'une manière pour estimer $\bar{\gamma}(\alpha), \bar{\delta}(\alpha)$, l'équation (5.48) peut nous servir pour estimer les paramètres inconnus des non linéarités.

5.5.3. Estimation de (h_1, h_2) pour HS-HR

L'équation (5.48a) toute seule n'est pas suffisante pour déterminer les deux paramètres inconnus (h_1, h_2) . On procède alors par deux expériences qui utilisent deux séquences d'entrée de même forme que (5.43a-b) avec la même période T mais deux rapports cycliques différents α_1 et α_2 . Les séquences internes résultantes seront notées $\gamma_1(t), \gamma_2(t)$ et auront pour moyenne:

$$\bar{\gamma}_1(\alpha) = \frac{1}{2} + \frac{h_1}{2h_m}(1 - \alpha_1) + \frac{h_2}{2h_m}\alpha_1 + \varepsilon(\alpha_1, T) \quad (5.52a)$$

$$\bar{\gamma}_2(\alpha) = \frac{1}{2} + \frac{h_1}{2h_m}(1 - \alpha_2) + \frac{h_2}{2h_m}\alpha_2 + \varepsilon(\alpha_2, T) \quad (5.52b)$$

La résolution du système d'équations linéaires (5.52a-b) donne (h_1, h_2) :

$$\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \frac{(1-\alpha_1)}{2h_m} & \frac{\alpha_1}{2h_m} \\ \frac{(1-\alpha_2)}{2h_m} & \frac{\alpha_2}{2h_m} \end{bmatrix}^{-1} \begin{bmatrix} \bar{\gamma}_1(\alpha) - \frac{1}{2} - \varepsilon(\alpha_1, T) \\ \bar{\gamma}_2(\alpha) - \frac{1}{2} - \varepsilon(\alpha_2, T) \end{bmatrix} \quad (5.53)$$

Mais l'utilisation de (5.53), est d'abord tributaire de la connaissance de $\bar{\gamma}_1(\alpha), \bar{\gamma}_2(\alpha)$. En utilisant (5.50) et (5.50a), on obtient les estimés suivants:

$$\hat{\gamma}_1(M, N) = \frac{\hat{A}_N(1)}{\hat{B}_N(1)} \tilde{y}_1(M), \quad \hat{\gamma}_2(M, N) = \frac{\hat{A}_N(1)}{\hat{B}_N(1)} \tilde{y}_2(M) \quad (5.54)$$

où $\tilde{y}_1(t), \tilde{y}_2(t)$ sont respectivement les sorties de (5.22)-(5.23) pour les entrées $v(t) = v_3(t, \alpha_1)$ et $v(t) = v_3(t, \alpha_2)$. Ce ci, avec (5.53), définit une expression pour estimer (h_1, h_2) :

$$\begin{bmatrix} \hat{h}_1(M, N) \\ \hat{h}_2(M, N) \end{bmatrix} = \begin{bmatrix} \frac{(1-\alpha_1)}{2h_m} & \frac{\alpha_1}{2h_m} \\ \frac{(1-\alpha_2)}{2h_m} & \frac{\alpha_2}{2h_m} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\gamma}_1(M, N) - \frac{1}{2} \\ \hat{\gamma}_2(M, N) - \frac{1}{2} \end{bmatrix} \quad (5.55)$$

Proposition 5.4. (Consistance de (5.54)-(5.55)) Considérons le système (5.7a-b), soumis aux hypothèses H5.1-H5.5, où $F(\cdot) = HR_{(M_1, M_2, h_1, h_2)}$. Le système peut être représenté par (5.23) et (5.42a) où $\tilde{F} = HR_{(0, 1, h_1, h_2)}$. Si deux séquences triangulaires et périodiques $v(t) = v_3(\alpha, t)$ (définies par (5.43a-b) avec deux valeurs α différentes) sont appliquées lors de deux expériences distinctes à l'entrée du système, alors l'algorithme (5.54)-(5.55) génère les estimés $\hat{h}_1(M, N), \hat{h}_2(M, N)$ qui convergent a.p.1 vers leur vraies valeurs h_1, h_2 quand $M, N \rightarrow \infty$ □

Preuve. En soustrayant (5.55) de (5.53) on obtient:

$$\begin{bmatrix} \hat{h}_1(M, N) - h_1 \\ \hat{h}_2(M, N) - h_2 \end{bmatrix} = \begin{bmatrix} \frac{(1-\alpha_1)}{2h_m} & \frac{\alpha_1}{2h_m} \\ \frac{(1-\alpha_2)}{2h_m} & \frac{\alpha_2}{2h_m} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\gamma}_1(M, N) - \bar{\gamma}_1(M, N) + \varepsilon(\alpha, T) \\ \hat{\gamma}_2(M, N) - \bar{\gamma}_2(M, N) + \varepsilon(\alpha, T) \end{bmatrix} \quad (5.56)$$

D'autre part, l'application de (5.50a) aux couples $(\bar{\gamma}_1(\alpha), \bar{y}_1(M))$ et $(\bar{\gamma}_2(\alpha), \bar{y}_2(M))$ donne (a.p. 1):

$$\lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} [B^*(1)\bar{\gamma}_1(\alpha) - A(1)\bar{y}_1(M)] = \lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} [B^*(1)\bar{\gamma}_2(\alpha) - A(1)\bar{y}_2(M)] = 0 \quad (5.57)$$

Utilisons maintenant la proposition 5.2, (5.57) s'écrit alors:

$$\lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} [\hat{B}_N(1)\bar{\gamma}_1(\alpha) - \hat{A}_N(1)\bar{y}_1(M)] = \lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} [\hat{B}_N(1)\bar{\gamma}_2(\alpha) - \hat{A}_N(1)\bar{y}_2(M)] = 0 \quad (5.58)$$

Mais, de (5.54) on aussi:

$$\hat{B}_N(1)\hat{\gamma}_1(M, N) - \hat{A}_N(1)\bar{y}_1(M) = \hat{B}_N(1)\hat{\gamma}_2(M, N) - \hat{A}_N(1)\bar{y}_2(M) = 0 \quad (5.59)$$

En comparant (5.58) et (5.59), il découle (en utilisant la Proposition 5.2) que:

$$\lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} (\hat{\gamma}_1(M, N) - \bar{\gamma}_1(\alpha)) = \lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} (\hat{\gamma}_2(M, N) - \bar{\gamma}_2(\alpha)) = 0$$

laquelle, combinée avec (5.56) et (5.47), établit la Proposition 5.4 \square

5.5.4. Estimation de (s, h_1, h_2) pour HS-H

On s'intéresse maintenant au cas où $F^* = H_{(s, h_1, h_2)}$. Les équations (5.48c), (5.49) et (5.50) suggèrent directement pour le paramètre s , l'estimé suivant $\hat{s}(M, N)$:

$$\hat{s}(M, N) = \left[4h_m \left(\hat{\delta}(M, N) - \frac{1}{2} - \varepsilon(\alpha, T) \right) \left(\frac{1}{1-2\alpha} \right) - 2h_m \right]^{-1} \quad (5.60)$$

où
$$\hat{\delta}(M, N) \stackrel{def}{=} \frac{\hat{A}_N(1)}{\hat{B}_N(1)} \bar{y}(M) \quad (5.61)$$

et $\bar{y}(t)$ est la réponse de (5.22)-(5.23) à l'entrée $v(t) = v_3(t, \alpha)$ avec $0 < \alpha < 1$ mais $\alpha \neq 0.5$. Par conséquence, une seule expérience est nécessaire pour estimer le paramètre s . D'autre part, à partir de (5.15) et Fig. 5.2 on tire les relations suivantes:

$$h_1 = h_m - \frac{m_1}{s} \quad \text{et} \quad h_2 = -h_m - \frac{1 + m_1}{s}$$

Celles ci suggèrent pour (h_1, h_2) les estimés $(\hat{h}_1(L, M, N), \hat{h}_2(L, M, N))$ définis comme suit:

$$\hat{h}_1(L, M, N) = h_m - \frac{\hat{m}_1(L, N)}{\hat{s}(M, N)} \quad \text{et} \quad \hat{h}_2(L, M, N) = -h_m - \frac{1 + \hat{m}_1(L, N)}{\hat{s}(M, N)} \quad (5.62)$$

Proposition 5.5. (Consistance de (5.60)-(5.61). Considérons le système (5.7a-b), soumis aux hypothèses H3.1-H3.5, où $F(\cdot) = H_{(s, h_1, h_2)}$. Le système peut aussi être représenté par (5.23) et (5.42a) où $\tilde{F} = H_{(s, h_1, h_2)}$. Si la séquence $v(t) = v_3(\alpha, t)$, triangulaire et périodique, définie par

(5.43a-b) avec $\alpha \neq 0.5$, est appliquée à l'entrée du système alors l'algorithme (5.60)-(5.62) donne les estimés $\hat{s}(M, N)$, $\hat{h}_1(L, M, N)$, $\hat{h}_2(L, M, N)$ qui convergent a.p.1 vers les vraies valeurs s , h_1 , h_2 quand $L, M, N \rightarrow \infty$ \square

Preuve. A partir de (5.60) et (5.48c) obtient, respectivement:

$$s^{-1} = 4h_m \left(\bar{\delta}(\alpha) - \frac{1}{2} - \varepsilon(\alpha, T) \right) \left(\frac{1}{1-2\alpha} \right) - 2h_m \quad (5.63)$$

$$(\hat{s}(M, N))^{-1} = 4h_m \left(\hat{\delta}(M, N) - \frac{1}{2} \right) \left(\frac{1}{1-2\alpha} \right) - 2h_m \quad (5.64)$$

En soustrayant (5.64) à partir de (5.63), on obtient:

$$s^{-1} - (\hat{s}(M, N))^{-1} = 4h_m \left(\bar{\delta}(\alpha) - \hat{\delta}(M, N) - \varepsilon(\alpha, T) \right) \left(\frac{1}{1-2\alpha} \right) \quad (5.65)$$

D'autre part, en appliquant (5.50b) au cas où $v(t) = v_3(\alpha, t)$ et $\tilde{u}(t) = \delta(\alpha, t)$, on a:

$$\lim_{M \rightarrow \infty} \left[B^*(1) \bar{\delta}(\alpha) - A(1) \bar{y}(M) \right] = 0 \quad (\text{a.p.1}) \quad (5.66)$$

où $\bar{y}(t)$ représente la réponse de (5.22)-(5.23) à l'entrée $v(t) = v_3(t, \alpha)$. En utilisant la proposition 5.2, on déduit à partir de (5.66) que:

$$\lim_{M, N \rightarrow \infty} \left[\hat{B}_N(1) \bar{\delta}(\alpha) - \hat{A}_N(1) \bar{y}(MT) \right] = 0 \quad (\text{a.p.1}) \quad (5.67)$$

Mais à partir de (5.61) on a aussi, pour tous les entiers (M, N) :

$$\hat{B}_N(1) \hat{\delta}(M, N) - \hat{A}_N(1) \bar{y}(MT) = 0 \quad (5.68)$$

La comparaison de (5.67) et (5.68), permet d'écrire (en utilisant la proposition 5.2) que:

$$\lim_{\substack{M \rightarrow \infty \\ N \rightarrow \infty}} \left[\bar{\delta}(\alpha) - \hat{\delta}(M, N) \right] = 0 \quad (5.69)$$

Cette dernière, avec (5.65) et (5.47), établit la proposition 5.4 pour s . Le même résultat de convergence peut être obtenu, de manière similaire, pour h_1 , h_2 \square

5.5.5. Simulation

Considérons le système de type *HS-HR* caractérisé par (5.33), (5.34b), (5.35b), (5.36b) et (5.37b).

L'algorithme (5.39) est appliqué en premier pour obtenir un estimé de $m_I = -0.4$.

On a obtenu les résultats suivants:

$$\hat{m}_1 = -0.3621 \text{ pour } N=400 \text{ et } \hat{m}_1 = -0.403 \text{ pour } N=10000$$

On constate que l'estimé \hat{m}_1 est très proche de sa vraie valeur quand N est suffisamment grand, comme le stipule la proposition 5.3. Puis on applique l'algorithme (5.54)-(5.55), avec les paramètres de la table 3:

	T	M	α
Expérience 1	600	100	$\alpha_1 = 1/3$
Expérience 2	600	100	$\alpha_2 = 1/2$

On obtient les estimés suivants:

$$\hat{h}_1 = 1.041, \hat{h}_2 = -1.507,$$

qui sont très proches des vraies valeurs pour $h_1=1$ et $h_2=-1.5$.

Considérons maintenant le système de type $HS-H$ caractérisé par (5.33), (5.34a), (5.35a), (5.36a) et (5.37a). L'algorithme (5.39) est appliqué en premier, pour avoir un estimé de $m_1 = -2/3$. Ce paramètre nous sera utile pour estimer les paramètres de la non linéarité. On a obtenu les résultats suivants:

$$\hat{m}_1 = -0.751 \text{ pour } N=400 \text{ et } \hat{m}_1 = -0.6708 \text{ pour } N=10000$$

Là encore, on voit que l'estimé est proche de sa vraie valeur quand N est suffisamment grand. On applique l'algorithme (5.60)-(5.61), avec les paramètres de la table 4:

	T	L	α
Expérience 1	600	100	$\alpha=1/3$

On obtient comme estimés suivants:

$$\hat{s} = -0.3348 \text{ et } \hat{h}_1 = 1.996, \hat{h}_2 = -3.016.$$

Les estimés obtenus pour les paramètres des non linéarités sont très proches de leur vraies valeurs.

5.6. CONCLUSIONS

Dans ce chapitre nous avons abordé l'identification des systèmes non linéaires sur la base du modèle de Hammerstein ayant une non linéarité non douce et présentant un effet hystérétique. Plus précisément, on a considéré les cas des non linéarités de type Hystérésis –relais et Hystérésis. A la différence de ce qui a été supposé dans les précédents travaux (très rares), ces non linéarités peuvent être symétriques ou asymétriques; en général elles sont caractérisées par plusieurs paramètres inconnus. L'approche adoptée pour l'identification consiste à identifier d'abord le sous-système linéaire, puis l'élément non linéaire. Pour cela, la première étape consiste en une paramétrisation (5.25) du système, qui présente deux propriétés clés : (i) Les paramètres linéaires inconnus apparaissent de manière linéaire dans cette nouvelle paramétrisation; (ii) le signal interne correspondant $\{\tilde{u}(t)\}$ est alors connu tant que l'entrée $v(t)$ appartient à l'intervalle $\{-h_m, +h_m\}$ (Fig. 5.2). Cette nouvelle paramétrisation permet aussi de réduire le nombre de paramètres à estimer pour le cas d'une non linéarité de type hystérésis-relais $HR_{(m_1, m_2, h_1, h_2)}$. En effet le nombre de paramètres est réduit à trois au lieu de quatre. Les paramètres du sous-système linéaire sont alors estimés en utilisant un algorithme classique de type moindres carrés (5.27). D'autre part, quand $v(t)$ est générée selon (5.24a-b) le vecteur de régression $\tilde{\Phi}(t)$ possède la propriété d'excitation persistante (voir preuve de la proposition 5.2). Ceci est nécessaire pour assurer la consistance des estimés (proposition 5.2). Dans deuxième étape, on choisit des séquences d'entrées bien particulières qui vont permettre d'estimer les paramètres de l'élément non linéaire. Ainsi, le paramètre $m_1 = M_1 / (M_2 - M_1)$ est estimé en utilisant une entrée constante $v_I(t) = h_m$ (pour tout t), et l'algorithme (5.39) qui est consistant (Proposition 5.3). Ensuite les paramètres h_1, h_2 pour $HR_{(m_1, m_2, h_1, h_2)}$, ou les paramètres s, h_1, h_2 pour $H_{(s, h_1, h_2)}$, sont estimés en utilisant des séquences d'entrée périodiques et de forme triangulaire (5.43). Avec de telles séquences on montre que la moyenne de la séquence interne $\{\tilde{u}(t)\}$ ne dépend que des paramètres inconnus h_1, h_2 ou s ((5.50) et (5.51)). Cette moyenne peut être calculée à partir de celle de la sortie $\{\tilde{y}(t)\}$, étant donné que la perturbation est ergodique et de moyenne nulle; Ceci est l'idée clé de l'estimation des paramètres des non linéarités. L'analyse montre que les paramètres sont consistants. C'est la première fois que de tels résultats de convergence sont obtenus pour un système de type Hammerstein avec des non linéarités asymétriques et à effet mémoire. Les outils mathématiques utilisés dans les différentes étapes de ce schéma d'identification sont très simples, bien connus et ne nécessitent pas de calculs trop complexes.

L'hypothèse d'ergodicité est le plus souvent admise, mais si tel n'est pas le cas, on peut alors garder la même approche mais au lieu de faire une expérience ou deux on fera une multitude d'expériences et on utilisera la moyenne (non plus temporelle) statistique pour faire disparaître le terme perturbateur.

Notons que l'algorithme proposé ne se prête pas à une utilisation en ligne; en effet l'estimation des paramètres du sous-système linéaire et celle des paramètres du gain non linéaire ne se font pas simultanément mais séquentiellement, et en utilisant des séquences d'excitation différentes. Remarquons en fin que cette approche d'identification séparée du sous-système linéaire et du gain non linéaire peut être appliquée à des modèles de Hammerstein avec d'autres types de non linéarités non douces mais à structures connues (telles que celles étudiées en chapitre 4 par exemple).

TROISIEME PARTIE

IDENTIFICATION NON PARAMETRIQUE DES SYSTEMES NON LINEAIRES DE TYPE WIENER

CHAPITRE 6

IDENTIFICATION FREQUENTIELLE DE L'ELEMENT NON LINEAIRE DES SYSTEMES DE WIENER – APPROCHE PAR ANALYSE DES EXTREMUMS DE LA SORTIE

6.1. INTRODUCTION

Dans ce chapitre, nous nous intéresserons au problème d'identification des systèmes non linéaires blocs de type Wiener (Fig 6.1). La plupart des travaux consacrés à ce problème ont proposé des solutions qui reposent sur l'hypothèse que l'élément non linéaire est un polynôme de degré connu et le sous-système linéaire est une fonction de transfert d'ordre connu, il en est ainsi dans (Hasiewicz, 1987), (Hunter and Korenberg, 1986), (Pajunen, 1992), (Wigren, 1993), (Voros, 1997), (Chou et al, 1999), (Nordsjö and Zetterberg, 2001), (Gardiner, 1993). La difficulté du problème vient en grande partie de ce que le signal interne n'est pas accessible à la mesure. Par conséquent, toute représentation entrée-sortie du système fait intervenir les paramètres inconnus de façon bilinéaire. Pour surmonter cette bilinéarité, rencontrée déjà dans le cas des systèmes de Hammerstein, certains auteurs ont en recours aux mêmes approches. Ainsi, Voros (1997) procède-t-il par optimisation itérative alternée paramètres linéaires/paramètres non linéaires. La consistance de l'estimateur ainsi obtenu, n'est généralement pas assurée. La technique des moindres carrés séparés a été suivie par Chou et al (1999) alors que Gardiner (1993) a suggéré une approche fréquentielle.

Certains travaux, plus rares, se sont attaqués au cas où l'élément non linéaire n'est caractérisé par aucun modèle paramétrique, c'est par exemple la cas dans (Greblicki, 1992, 1997), (Bai, 2003), (Hu and Chen, 2005). La solution proposée dans (Greblicki, 1992, 1997) nécessite l'excitation du système par des bruits blancs et exige que l'élément non linéaire soit inversible. Ces hypothèses limitent bien entendu la portée pratique de la solution. Le schéma d'identification proposé récemment dans (Hu and Chen, 2005) opère également dans un contexte stochastique. Il génère une estimation pour chaque paramètre du sous-système linéaire et de $F(x)$, pour une valeur fixée de x .

La consistance des estimateurs mis en jeu est bien établie mais la solution suppose que le système linéaire est de type MA (moyenne mobile) et l'élément non linéaire est continu et ne peut pas croître plus rapidement qu'un polynôme. Jusqu'à très récemment, le travail de référence, s'agissant de l'identification des systèmes de Wiener, fut celui de Bai (2003). En effet, la méthode d'identification fréquentielle proposée se voulait générale et à cet effet l'élément non linéaire n'est pas supposé inversible. En outre, la contribution se voulait théoriquement irréprochable et à cet effet une analyse de consistance est développée. Toutefois, nous avons soulevé dans (Giri et al, 2006b) un grave problème théorique mettant en cause la consistance de la méthode d'identification proposée par Bai et limitant pratiquement son champ d'application au cas où le sous-système linéaire possède une phase variant entre 0 et 180°. L'annexe C décrit en détail la défaillance soulevée et comprend une amélioration ad-hoc du schéma d'identification de telle manière qu'il soit consistant au moins quand la phase du sous-système linéaire varie entre 0 et 180°.

Dans ce chapitre, nous présentons une nouvelle solution fréquentielle pour le problème d'identification des systèmes de Wiener. Spécifiquement, l'accent sera mis sur l'identification de l'élément non linéaire placé en sortie et supposé sans mémoire. L'identification de cet élément ouvrira la voie à celle du sous-système linéaire en appliquant à cet effet des solutions disponibles rappelées dans (Bai, 2003). L'approche d'identification fréquentielle consiste à exciter le système par des signaux sinusoïdaux et à comparer, à chaque fois, l'entrée et la sortie du système en profitant du fait que, en régime permanent, la sortie est périodique et de même période que l'entrée. Cette propriété est très connue et largement exploitée dans le cas des systèmes linéaires (dans ce cas la sortie est même sinusoïdale). S'agissant des systèmes non linéaires blocs, elle est également vérifiée bien que plus rarement exploitée. Le schéma d'identification développé ici est théoriquement fondé sur l'analyse des extremums de la sortie périodique du système. Plus explicitement, le système est soumis à un signal sinusoïdal $v(t) = V \cos(\omega t)$. Il s'ensuit en régime permanent un signal interne de la forme $u(t) = U \cos(\omega t - \varphi)$. L'incertitude concernant l'amplitude de ce signal peut être levée moyennant un rééquilibrage des deux parties, linéaire et non linéaire, du modèle en exploitant à cet effet la non unicité de ce dernier. Ce faisant, on obtient $u(t) = \cos(\omega t - \varphi)$ et par conséquent la seule incertitude qui subsiste au niveau du signal interne est le déphasage φ . L'idée clé est précisément de déterminer ce déphasage en explorant la relation entre les extremums du signal interne et ceux du signal de sortie. Une technique a ainsi été mise au point permettant la reconnaissance des instants d'occurrence des extremums du signal interne à partir d'un enregistrement de la sortie. Une fois levée l'incertitude sur le déphasage φ , le signal

interne devient disponible, permettant ainsi de déterminer la caractéristique de l'élément non linéaire $F(\cdot)$. Dans le cas où l'élément non linéaire est strictement monotone, le schéma d'identification s'avère relativement simple et nécessite une seule expérience, impliquant un signal sinusoïdal de fréquence et amplitude fixes. Dans la cas d'éléments non linéaires non monotones, l'approche proposée exige au moins deux expériences impliquant des signaux sinusoïdaux ayant la même fréquence mais des amplitudes différentes. La consistance de la méthode est assurée, en présence d'un bruit blanc ergodique, moyennant un filtrage approprié de la sortie, exploitant à cet effet sa nature périodique.

Le chapitre est organisé de la manière suivante: le problème d'identification étudié sera formulé au paragraphe 2. Au paragraphe 3, le système fera l'objet d'une analyse mathématique qui permettra de mettre en évidence un certain nombre de propriétés qui s'avéreront très utiles dans l'élaboration de la solution. Au paragraphe 4, nous présenterons un schéma d'identification simple adapté au cas où l'élément non linéaire est strictement monotone. Au paragraphe 5, le schéma précédent est étendu aux non-linéarités non nécessairement monotones. Les performances du schéma d'identification sont illustrées par voie de simulation au paragraphe 6. Une conclusion situant le présent travail par rapport aux contributions antérieures achèvera le chapitre.

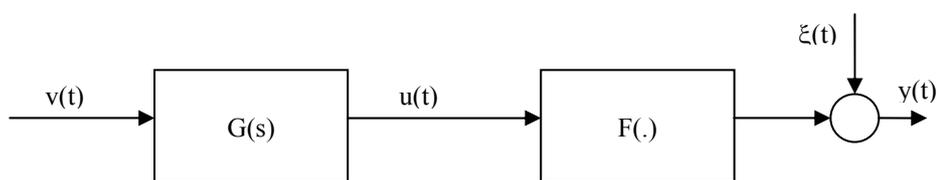


Fig. 6.1. Modèle de Wiener

6.2. POSITION DU PROBLÈME D'IDENTIFICATION

6.2.1. Classe des systèmes à identifier

On s'intéresse aux systèmes dont le comportement entrée-sortie peut être parfaitement décrit par le modèle de Wiener de la Fig. 6.1, où $G(s)$ représente la fonction de transfert du sous-système linéaire et $F(\cdot)$ représente la caractéristique statique du gain non linéaire. Ce modèle est analytiquement décrit par les équations suivantes :

$$u(t) = g(t) * v(t) \quad (6.1)$$

$$y(t) = F(u(t)) + \xi(t) \quad (6.2)$$

où :

- . $g(t)$ désigne la transformée inverse de Laplace de $G(s)$;
- . le symbole ' * ' dénote l'opérateur de convolution;
- . $v(t)$ et $y(t)$ sont respectivement les signaux d'entrée et de sortie du système ;
- . $u(t)$ est un signal interne non mesurable ;
- . $\xi(t)$ est une erreur rendant compte de l'effet des perturbations externes, voire d'erreurs de modélisation.

Le système est en outre soumis aux hypothèses suivantes:

H6.1. $G(s)$ est asymptotiquement stable.

H6.2. $\{\xi(t)\}$ est un processus stochastique, stationnaire, indépendant, ergodique et de moyenne nulle.

H6.2. La fonction $F(\cdot)$ est dérivable presque partout.

Remarques 6.1

- 1) L'hypothèse H6.1 est réaliste du moment que l'identification se fait ici en boucle ouverte.
- 2) L'hypothèse d'ergodicité s'avérera utile en permettant par la suite de remplacer la moyenne mathématique par la moyenne arithmétique.
- 3) A l'exception des hypothèses H6.1-H6.2, le sous système linéaire est quelconque. En particulier aucune hypothèse n'est faite sur la nature et la structure de $G(s)$. En plus, $F(\cdot)$ n'est pas supposé inversible ni même paramétrique.

6.2.2. Objectif d'identification

Notre but est d'élaborer un schéma d'identification capable de fournir un nuage de points de la caractéristique $F(\cdot)$ sur un intervalle de la forme : $-V|G(j\omega)| \leq u(t) \leq V|G(j\omega)|$, où V est l'amplitude du signal d'entrée et ω sa pulsation, ceux-ci étant choisis arbitrairement. Notons que cet objectif est conforme au fait que la structure de $F(\cdot)$ n'est pas connue a priori.

La principale difficulté vient du fait que le signal interne $u(t)$ est non mesurable. Le processus d'identification devra donc s'appuyer uniquement sur les mesures des signaux d'entrée et de sortie, $v(t)$ et $y(t)$. Le fait que $u(t)$ soit non mesurable entraîne que le problème d'identification n'a pas une solution unique. En effet, si le couple $(G(s), F(u))$ est représentatif du système, d'un point de vue entrée-sortie, tout couple $(G(s)/\alpha, F(\alpha u))$ l'est aussi et ce pour tout réel α non nul. Il s'ensuit que

le signal interne est un objet théorique qui n'a généralement aucune existence physique. Le manque d'unicité du modèle sera exploité ultérieurement pour lever l'incertitude qui pèse sur l'amplitude du signal interne $u(t)$ en régime harmonique.

6.3. ANALYSE PRÉLIMINAIRE DU SYSTÈME

Dans ce paragraphe, nous allons mettre en évidence un certain nombre de propriétés qui vont s'avérer être des ingrédients essentiels à la synthèse du schéma d'identification. Dans le paragraphe 6.3.1, on montrera qu'en présence d'un signal d'entrée $v(t)$ sinusoïdal, le signal interne $u(t)$, lui aussi sinusoïdal en régime permanent, peut être considéré comme à amplitude unitaire. Dans le paragraphe 6.3.2, nous explorerons la correspondance dans le temps entre les extremums de la sortie et ceux du signal interne. Cette analyse va permettre la mise au point, au paragraphes 6.4 et 6.5, d'une technique d'estimation du déphasage du signal interne.

6.3.1. Rééquilibrage du modèle de Wiener

Le long de ce paragraphe, le système à identifier est excité par une sinusoïde de la forme :

$$v(t) = V \cos(\omega t) \quad (t \geq 0) \quad (6.3)$$

où l'amplitude $V > 0$ et la pulsation ω sont maintenues constantes., Il vient de (6.1) que le signal interne $u(t)$ prend la forme suivante en régime permanent :

$$u(t) = u_v(t) \stackrel{def}{=} U_V \cos(\omega t - \varphi) \quad (6.4a)$$

avec :

$$U_V = V |G(j\omega)| \quad \text{et} \quad \varphi = -\arg(G(j\omega))$$

Le signal de sortie résultant est :

$$y(t) = F(u_v(t)) + \xi(t)$$

Par commodité, nous réécrivons $u_v(t)$ sous la forme suivante :

$$u_v(t) = U_V \cos(\omega(t - t_\varphi)) \quad (6.4b)$$

Avec :

$$t_\varphi \stackrel{def}{=} \frac{\varphi}{\omega} \quad (6.4c)$$

Si le signal $u_V(t)$ était mesurable, il serait possible d'identifier un ensemble de points de $F(\cdot)$ à partir d'un ensemble de couples $(u_V(t), y_f(t))$ où $y_f(t)$ désigne une version filtrée de la sortie (un tel filtrage permettrait de réduire l'effet du bruit). Mais, $u_V(t)$ est non mesurable puisque ni son amplitude U_V ni sa phase φ ne sont connues. L'incertitude affectant l'amplitude peut être contournée en substituant au modèle (6.1), caractérisé par le couple $(G(s), F(u))$, un modèle de la forme $(G(s)/\alpha, F(\alpha u))$ pour une valeur appropriée de α . A cet égard, deux valeurs de α sont particulièrement indiquées, à savoir $\alpha=U_V$ et $\alpha=-U_V$. Il en résulte respectivement les deux modèles suivants :

$$M^+(V) \stackrel{\text{def}}{=} (F_V^+, G_V^+(s)) \quad (6.5a)$$

$$M^-(V) \stackrel{\text{def}}{=} (F_V^-, G_V^-(s)) \quad (6.5b)$$

où:

$$F_V^+(u) \stackrel{\text{def}}{=} F(U_V u), \quad F_V^-(u) \stackrel{\text{def}}{=} F(-U_V u) \quad (6.6)$$

$$G_V^+(s) \stackrel{\text{def}}{=} \frac{1}{U_V} G(s), \quad G_V^-(s) \stackrel{\text{def}}{=} \frac{1}{-U_V} G(s) \quad (6.7)$$

Les signaux internes correspondant à ces modèles sont respectivement:

$$u^+(t) = \cos(\omega(t-t^+)), \quad u^-(t) = \cos(\omega(t-t^-)) \quad (6.8)$$

où

$$t_\varphi^+ = t_\varphi$$

et $t_\varphi^- = t_\varphi + \frac{\pi}{\omega}$.

Ainsi, le signal interne (dans les deux modèles) ne dépend plus de l'amplitude V , contrairement au signal interne $u_V(t)$ du modèle initial (6.1). En revanche, les deux nouveaux modèles et le modèle initial génèrent tous le même signal de sortie, à chaque fois qu'ils sont soumis au signal sinusoïdal (6.3). Le signal de sortie en question est caractérisé par les expressions suivantes :

$$y(t) = y_V(t) + \xi(t) \quad (6.9)$$

avec :

$$y_V(t) = F_V^+(u^+(t)) = F_V^-(u^-(t)) = F(u_V(t)) \quad (6.10)$$

Cette propriété sera très utile pour identifier ou la phase t_ϕ^+ ou la phase t_ϕ^- dans (6.8). Comme il est impossible de distinguer entre les deux modèles: l'élément non linéaire qui sera estimé pourra correspondre aussi bien à $F_V^+(\cdot)$ qu'à $F_V^-(\cdot)$. La méthode d'identification que l'on va concevoir à cet effet consistera à reconnaître les extremums du signal de sortie qui ont un correspondant au niveau du signal interne (non mesurable). Cette correspondance entre extremums des deux signaux va maintenant être clarifiée dans le sous-paragraphe suivant.

6.3.2. Correspondance entre les extremums du signal interne et ceux de la sortie non bruitée

Tout au long de ce sous-paragraphe, le système est excité par un signal sinusoïdal de la forme $v(t) = V \cos(\omega t)$ où l'amplitude $V > 0$ et la fréquence $\omega > 0$ sont maintenues inchangées. Le but est de clarifier la correspondance entre les extremums de $u^+(t)$ et $u^-(t)$, d'une part, et ceux de la sortie $y_V(t)$, d'autre part.

Il découle de (6.8) que les extremums des signaux internes $u^+(t)$ et $u^-(t)$ surviennent aux instants :

$$t_i = t_\phi + \frac{i\pi}{\omega} \quad (i=0, 1, 2, \dots) \quad (6.11)$$

Ceux-ci découlent des équations $\frac{du^+(t)}{dt} = 0$ et $\frac{du^-(t)}{dt} = 0$. Par ailleurs, il vient de (6.10) que :

$$\frac{dy_V(t)}{dt} = \frac{dF_V^+}{du}(u^+(t)) \cdot \frac{du^+(t)}{dt} = \frac{dF_V^-}{du}(u^-(t)) \cdot \frac{du^-(t)}{dt} \quad (6.12)$$

ce qui implique que :

$$\frac{dy_V(t_i)}{dt} = 0 \quad (i=0, 1, 2, \dots) \quad (6.13)$$

Ainsi, chaque extremum de $u^+(t)$ et $u^-(t)$ donne lieu à un extremum de $y_V(t)$ se produisant au même instant. En plus, d'après (6.12), il est clair que $y_V(t)$ peut présenter d'autres extremums à d'autres instants. Ces derniers sont les solutions des équations :

$$\frac{dF_V^+}{du}(u^+(t)) = 0, \quad \frac{dF_V^-}{du}(u^-(t)) = 0 \quad (6.14)$$

Mais dans le reste de ce chapitre nous allons nous intéresser aux extremums de $(u^+(t), u^-(t), y_V(t))$, qui surviennent aux instants t_i définis par (6.11). En effet, si l'un d'eux est localisé il sera possible d'en déduire les quantités $(t_\varphi^+, t_\varphi^-)$ ce qui rendra disponible les signaux internes $u^+(t)$ et $u^-(t)$. La question cruciale est : comment reconnaître les instants t_i sur un enregistrement de la sortie perturbée? Le paragraphe suivant constitue une première étape dans l'élaboration de la réponse.

6.3.3. Caractérisation des extremums du signal interne en utilisant la sortie mesurée

Pour s'affranchir de l'effet du bruit et recouvrer la sortie non bruitée $y_V(t)$ à partir de la sortie mesurée, il suffit d'opérer sur cette dernière un filtrage approprié. Compte-tenu de la périodicité de $y_V(t)$, l'Hypothèse H6.2 suggère un filtrage par moyennage :

$$y(t, N) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{k=0}^{N-1} y(t+kT), \quad \text{pour } t \in [0, T[\quad (6.15a)$$

$$\text{et } y(t+kT, N) = y(t, N), \quad \text{pour tout entier } k \text{ et tout réel } t \in [0, T[\quad (6.15b)$$

où $T = 2\pi/\omega$ et N un entier quelconque mais suffisamment grand. Par ailleurs, on suppose que la période T est un multiple de la période d'échantillonnage. Pour tout t fixé, la quantité $y(t, N)$ représente la moyenne des valeurs de la sortie mesurées aux instants $t+kT$ ($k=0, 1, 2, \dots$). Le lien entre la sortie filtrée ainsi définie et la sortie non bruitée $y_V(t)$ est précisé dans la proposition suivante.

Proposition 6.1. Considérons le système (6.1), soumis aux hypothèses H6.1-H6.3, excité par le signal sinusoïdal (6.3) où $V > 0$ et $\omega > 0$ sont quelconques mais constants. Alors, pour tout t :

$$\lim_{N \rightarrow \infty} y(t, N) = y_V(t) \quad (\text{a.p. } 1)$$

Par conséquent, on a pour tout $t \in \{t_0, t_1, t_2, \dots\}$:

$$\frac{d}{dt} \left(\lim_{N \rightarrow \infty} y(t, N) \right) = 0 \quad (\text{a.p. } 1)$$

où les t_i sont définis par (6.11). \square

Preuve. En utilisant le fait que $y_V(t)$ est périodique de période T , il découle de (6.9) que :

$$y(t+kT) = y_V(t) + \xi(t+kT) \quad (\text{pour tout réel } t \text{ et tout entier } k) \quad (6.16)$$

D'autre part, à partir de l'hypothèse H6.2 on déduit que pour tout réel t :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \xi(t+kN) = E(\xi(t+kN)) = 0 \quad (\text{a.p. } 1) \quad (6.17)$$

En combinant (6.16) et (6.17), on obtient pour tout réel t :

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N y(t+kT) = y_V(t) \quad (\text{a.p. } 1)$$

Ceci achève la preuve de la proposition 6.1. \square

On a ainsi établi que, pour N suffisamment grand, la sortie filtrée $y(t, N)$ possède des extremums aux instants t_i définis par (6.11). Toutefois, il se peut que $y(t, N)$ possède d'autres extremums en d'autres instants (qui sont solutions de (6.14)). Mais les extremums qui nous intéressent dans cette étude sont ceux qui surviennent aux instants t_i . Une procédure qui permet de reconnaître ces instants sur un enregistrement de la sortie du système va maintenant être présentée dans le paragraphe suivant, pour le cas où la non-linéarité $F(\cdot)$ est strictement monotone. Le cas général sera traité dans le paragraphe 5.

6.4. IDENTIFICATION DES NON-LINEARITES MONOTONES

Tout au long de ce paragraphe, le système est excité par le signal sinusoïdal $v(t) = V \cos(\omega t)$ où l'amplitude $V > 0$ et la fréquence $\omega > 0$ sont maintenues constantes. La fonction $F(\cdot)$ est supposée strictement monotone, mais le sens de monotonie (croissant ou décroissant) n'est pas connu. Il en va de même pour les deux fonctions $F_V^+(\cdot)$ et $F_V^-(\cdot)$. Dans ce cas, on a pour tout u :

$$\frac{dF_V^+}{du}(u) \neq 0 \quad \text{et} \quad \frac{dF_V^-}{du}(u) \neq 0$$

Il vient alors de (6.12) que :

$$\frac{dy_V(t)}{dt} = 0 \Leftrightarrow t = t_i \quad (i = 1, 2, \dots) \Leftrightarrow \frac{du^+(t)}{dt} = \frac{du^-(t)}{dt} = 0 \quad (6.18)$$

Il s'ensuit que $y_V(t)$ n'a pas d'autres extremums que ceux se produisant aux instants t_i . En plus, la monotonie stricte de $F_V^+(\cdot)$ et $F_V^-(\cdot)$ implique que, sur chaque période, ces extremums sont globaux tout comme ceux de $u^+(t)$ et $u^-(t)$. Plus précisément, si les maximums de $y_V(t)$ se produisent aux instants t_{2j} , ses minimums auront lieu aux instants t_{2j+1} et inversement.

Notons τ_i ($i = 1, 2, \dots$) les instants où $y_V(t)$ atteint réellement ses maximums. Alors, comme $y_V(t)$ est de même période que $u(t)$, on a:

$$\tau_i - \tau_{i-1} = \frac{2\pi}{\omega} \quad (i = 1, 2, \dots) \quad (6.19)$$

La donnée des instants τ_i ($i = 1, 2, \dots$) est suffisante pour la détermination de la non-linéarité $F(\cdot)$. Ceci est précisé dans la Proposition suivante.

Proposition 6.2. Considérons le système (6.1) excité par le signal (6.3).

- 1) Si $F(\cdot)$ est (strictement) monotone, il existe un modèle de Wiener M^* appartenant à l'ensemble $\{M^+(V), M^-(V)\}$ tel que le signal interne correspondant, noté $u^*(t)$, atteint ses maximums aux instants τ_i ($i = 1, 2, \dots$).
- 2) Plus précisément, si $F(\cdot)$ est croissant (resp. décroissant) alors $M^* = M^+(V)$ et $u^*(t) = u^+(t)$ (resp. $M^* = M^-(V)$ et $u^*(t) = u^-(t)$).
- 3) Soit $F^* \in \{F_V^+(\cdot), F_V^-(\cdot)\}$ l'élément non linéaire associée au modèle M^* . On a, pour tout t :

$$\lim_{N \rightarrow \infty} y(t, N) = F^*(u^*(t)) \quad (\text{a.p. 1}) \quad \square$$

Preuve. Supposons que $F(\cdot)$ soit strictement croissante. Alors, il découle de (6.7) que $F_V^+(\cdot)$ l'est aussi. De ce fait, la relation $F_V^+(u^+(t)) = y_V(t)$ implique que les maximums de $u^+(t)$ et $y_V(t)$ sont atteints par ces signaux aux mêmes instants, en l'occurrence aux instants τ_i ($i = 1, 2, \dots$). Dans ce cas, $M^* = M^+(V)$ et $u^*(t) = u^+(t)$. Le cas où la fonction $F(\cdot)$ est décroissante peut être

traité de la même façon. Ceci établit les parties 1 et 2 de la proposition. La partie 3 est une conséquence immédiate de la partie 1 de la Proposition 6.2 et de la Proposition 6.1 \square

Puisque y_V est non mesurable, les instants τ_i doivent être estimés à partir de la sortie filtrée $y(t, N)$. A cet effet, nous proposons la procédure ci-dessous, inspiré par les Propositions 6.1 et 6.2. La détermination des instants τ_i sera mise à profit pour estimer la non-linéarité du modèle.

6.4.1. Procédure d'identification des non-linéarités monotones (INM) :

INM-1. Appliquer au système un signal sinusoïdal $v(t) = V \cos \omega t$, où $V > 0$ et $\omega > 0$ sont choisies par l'utilisateur en tenant compte domaine de validité du modèle. Attendre le régime permanent et enregistrer la sortie $y(t)$ pendant un intervalle de temps suffisamment grand (sur un nombre de période N assez élevé).

INM-2. Calculer la version filtrée de la sortie $y(t, N)$. Noter τ^* un instant quelconque où $y(t, N)$ atteint son maximum (sur une période T , il doit exister un seul instant de ce type).

INM-3. Soit $u^*(t)$ le signal interne sinusoïdal (de période $2\pi/\omega$ et d'amplitude 1) qui atteint son maximum à l'instant τ^* . (En fait, $u^*(t) = \cos(\omega(t - \tau^*))$ représente le signal interne associé au modèle M^* , défini dans la Proposition 6.2).

INM-4. Tracer la courbe $(u^*(t), y(t, N))$ pour $\tau^* \leq t \leq \tau^* + 2\pi/\omega$; cette courbe représente l'élément non linéaire $F^*(.)$ (défini dans la Proposition 6.2) avec une précision qui dépend de N . Une estimation $\hat{F}_N(.)$ de $F^*(.)$ est obtenue à partir d'un échantillon de points $(u^*(t_j), y(t_j, N))$ ($j=1, \dots, P$), moyennant une interpolation appropriée. La qualité d'estimation de $\hat{F}_N(.)$ est d'autant meilleure que N et P sont plus grands.

Remarques 6.2

1. On peut se demander comment savoir si la non-linéarité est strictement monotone ou non ? la réponse à cette question est assez simple ; en effet soit on dispose d'une connaissance préalable sur la structure de la non linéarité système; sinon après avoir généré la sortie filtrée (dans l'étape INM-2), il suffit de vérifier que sur une période T , elle ne présente que deux extremums (un maximum et un minimum) ; si c'est le cas on est en présence d'une non linéarité strictement monotone et on continue alors la procédure INM. Si ce n'est pas le cas il faut utiliser la procédure générale décrite ci-dessous dans le paragraphe 6.5.
2. La non-linéarité est supposée strictement monotone mais il est permis qu'elle soit discontinue d'être discontinue, ce qui constitue un progrès par rapport à (Bai, 2003) où la non-linéarité est supposée lisse et inversible.

6.5. IDENTIFICATION DES NON-LINEARITES NON MONOTONES

La méthode proposée dans le paragraphe 4 va maintenant être étendue pour prendre en considération le cas des fonctions $F(\cdot)$, non strictement monotones. Dans ce cas, la difficulté principale est liée au fait que la sortie non bruitée $y_V(t)$ (résultant en régime permanent d'une excitation sinusoïdale $v(t) = V \cos(\omega t)$) peut présenter d'autres extremums que ceux correspondant au signal interne ($u^+(t)$ ou $u^-(t)$) et survenant aux instants t_i définis par (6.11). En effet, de nouveaux extremums peuvent apparaître si les équations (6.14) ont des solutions en t . En outre, l'ensemble de telles solutions peut être non dénombrable ; il en sera ainsi si la non-linéarité présente des paliers horizontaux (zones-mortes, saturations). Dans la suite de ce paragraphe, l'intérêt sera porté uniquement sur les extremums de $y_V(t)$ survenant aux instants t_i , car la connaissance de ceux-ci permet de lever l'incertitude qui pèse sur le signal interne rendant ainsi possible l'identification de l'élément non linéaire. De ce fait, ces extremums seront qualifiés d'utiles et tous les autres d'inutiles. Il convient de noter que les extremums utiles ne sont plus forcément globaux comme cela fut le cas au Paragraphe 6.4. Il s'agit maintenant de concevoir une technique permettant de reconnaître ces extremums quand un enregistrement de la sortie bruitée $y(t)$ est disponible.

Pour répondre à cette question, commençons par analyser l'effet d'un changement d'amplitude de l'entrée sur la sortie du système. A cet effet, considérons deux signaux d'entrée sinusoïdaux ne différant que par leurs amplitudes:

$$v_1(t) = V_1 \cos \omega t, \quad v_2(t) = V_2 \cos \omega t \quad (6.20)$$

où $V_1 \neq V_2$. Soient $(y_V^1(t), y_V^2(t))$ et $(y_1(t), y_2(t))$ les sorties, non bruitées et réelles, résultant des entrées précédentes. En se référant au modèle initial $(F(\cdot), G(s))$, défini par (6.1), les signaux internes correspondants sont:

$$u_1(t) = U_1 \cos(\omega t - \varphi), \quad u_2(t) = U_2 \cos(\omega t - \varphi) \quad (6.21)$$

où $\varphi = -\arg(G(j\omega))$ est indépendant de l'amplitude. Il s'ensuit que :

$$\frac{du_1(t)}{dt} = 0 \Leftrightarrow t = k_i = \frac{\varphi + \pi.i}{\omega} \quad (i = 0, 1, 2, \dots) \quad (6.22a)$$

$$\frac{du_2(t)}{dt} = 0 \Leftrightarrow t = k_i = \frac{\varphi + \pi.i}{\omega} \quad (i = 0, 1, 2, \dots) \quad (6.22b)$$

Les extremums de $u_1(t)$ et $u_2(t)$ surviennent donc aux mêmes instants. Par conséquent, un changement d'amplitude du signal l'entrée n'affecte pas l'emplacement, sur l'axe du temps, des extremums du signal interne $u(t)$. Par ailleurs, comme :

$$\frac{dy_V^1(t)}{dt} = \frac{dF}{dx}(u_1(t)) \cdot \frac{du_1(t)}{dt} \quad (6.23a)$$

$$\frac{dy_V^2(t)}{dt} = \frac{dF}{dx}(u_2(t)) \cdot \frac{du_2(t)}{dt} \quad (6.23b)$$

il vient que :

$$\left. \frac{dy_V^1(t)}{dt} \right|_{t=k_i} = \left. \frac{dy_V^2(t)}{dt} \right|_{t=k_i} = 0 \quad (6.24)$$

et

$$y_V^1(k_i) = F((-1)^i V_1) \quad ; \quad y_V^2(k_i) = F((-1)^i V_2) \quad (6.25)$$

où les k_i désignent les instants définis par (6.22a-b). Par ailleurs, il vient de (6.21) que, en général :

$$\frac{dF}{dx}(u_1(t)) = 0 \quad \not\Rightarrow \quad \frac{dF}{dx}(u_2(t)) = 0 \quad (6.26)$$

Les observations précédentes sont récapitulées dans la proposition suivante.

Proposition 6.3. Soit $y_V(t)$ la sortie non bruitée générée par le système (6.1) en réponse à l'entrée sinusoïdale $v(t)$.

- 1) Un changement d'amplitude au niveau de l'entrée du système n'affecte pas les instants où surviennent les extremums utiles de $y_V(t)$. En revanche, la taille de ces extremums s'en trouve affectée. On exprime ceci en disant qu'un changement d'amplitude à l'entrée provoque en sortie une variation verticale des extremums utiles.
- 2) Un changement d'amplitude au niveau de l'entrée affecte généralement les instants où surviennent les extremums inutiles de $y_V(t)$. En revanche, la taille de ces extremums reste inchangée. On exprime ceci en disant qu'un changement d'amplitude à l'entrée provoque en sortie un déplacement horizontal des extremums inutiles \square

Remarques 6.3

1. Le résultat précédent n'est pas directement utilisable, pour déterminer les instants des extremums utiles, puisqu'il invoque la sortie non mesurable $y_V(t)$. Ce problème fut contourné dans le Paragraphe 4 en substituant à $y_V(t)$ la version filtrée $y(t, N)$ de la sortie mesurée $y(t)$.
2. Le cas particulier où le point 2 de la proposition précédente n'est pas vérifié est celui où la fonction $F(\cdot)$ présente un extremum local en zéro. Des extremums inutiles vont alors apparaître sur $y_V(t)$ lors du passage par zéro de $u(t)$. Ces extremums ne vont pas glisser horizontalement, comme le feraient les autres extremums de même nature, suite à un changement d'amplitude du signal d'entrée. Ils ne vont pas non plus glisser verticalement comme le feraient les extremums utiles. Dans ce cas, on va observer sur chaque période T deux extremums distants de $T/2$, se situant toujours aux mêmes instants et ayant la même valeur et ce quelle que soit l'amplitude de l'entrée. Ces extremums ne suivent donc ni l'évolution des extremums utiles ni celle des extremums inutiles, suite à un changement d'amplitude à l'entrée. Grâce à cette propriété il sera facile de les reconnaître si un enregistrement de $y_V(t)$ était disponible. Pour simplifier l'exposé, ce cas particulier sera écarté dans la suite de ce chapitre.

Le schéma d'identification développé dans ce chapitre s'appuie en partie sur l'analyse faite jusqu'ici. La partie de ce schéma qui repose sur ce qui vient d'être dit est présentée dans le sous paragraphe suivant.

6.5.1. Identification des non-linéarités non monotones (INNM) : première partie

INNM-1. Appliquer à tour de rôle au système quelques signaux sinusoïdaux $(v_1(t), v_2(t), \dots)$ ayant une même fréquence mais des amplitudes différentes et enregistrer les sorties résultantes en régime permanent $(y_1(t), y_2(t), \dots)$. Elaborer les versions filtrées $(y_1(t, N), y_2(t, N), \dots)$ en appliquant (6.15).

INNM-2. Comparer les extremums des signaux de sortie filtrés et sélectionner tout ceux qui se produisent (dans les différents enregistrements filtrés) aux mêmes instants en notant qu'ils sont distants de $T/2$ et qu'ils n'ont pas forcément la même valeur (glissement vertical). Les extremums ainsi retenus sont ceux qu'on qualifie d'utiles. Durant cette sélection on utilise le fait que, en passant d'un enregistrement à l'autre, les extremums utiles ne font que varier verticalement là où tous les autres ne font que se déplacer horizontalement.

INNM-3. Si c'est nécessaire augmenter la valeur de N au niveau du filtrage et/ou multiplier les expériences en utilisant des amplitudes différentes tout en maintenant inchangé la fréquence et ce jusqu'à ce qu'on distingue clairement les extremums utiles des autres.

Remarque 6.4

1. L'étape INNM-1 suppose que des entrées avec différentes amplitudes provoquent des sorties différentes avec différentes amplitudes. Ceci est toujours vérifié sauf si $F(.)$ est vraiment une fonction constante ; ce cas est écarté puisqu'il ne présente aucun intérêt pratique.
2. La méthode précédente peut aussi s'avérer laborieuse dans le cas extrême où $F(.)$ comporte des paliers horizontaux dominant la plage utile. C'est le cas par exemple des éléments non linéaires de type relais (fonction signe), de type saturation (avec une zone linéaire trop étroite) ou de type zone-morte et saturation.

Maintenant que les extremums utiles sont repérés, il s'agit d'exploiter cette information en déterminant la non-linéarité. A cet effet, on utilisera un seul enregistrement parmi ceux relevés lors des expériences effectuées dans l'étape INNM-1. On peut a priori sélectionner celui obtenu lors de la première expérience. Par commodité, on convient d'adopter ici les notations du paragraphe 4, c'est-à-dire :

- . signal d'entrée : $v(t) = V \cos \omega t$ avec $V=V_1$
- . signal de sortie non bruitée : $y_v(t)$,
- . signal de sortie bruité : $y(t)$,
- . signal de sortie filtré : $y(t, N)$,
- . modèles rééquilibrés : $M^+(V), M^-(V)$,
- . éléments non linéaires : $F_V^+(\cdot), F_V^-(\cdot)$
- . signaux internes : $u^+(t) = \cos(\omega(t-t_\varphi^+))$, $u^-(t) = \cos(\omega(t-t_\varphi^-))$.

Comme $u^+(t), u^-(t)$ sont périodiques de période $2\pi/\omega$, la sortie non bruitée $y_v(t)$ est forcément périodique et sa période est un sous-multiple de la précédente, soit $2\pi/n\omega$ (pour un certain entier $n \geq 1$). Il s'ensuit que, pour N suffisamment grand, la sortie filtrée $y(t, N)$ est périodique elle aussi de période $2\pi/n\omega$. Ainsi, en examinant un enregistrement de $y(t, N)$, on verra sur tout intervalle de durée $2\pi/\omega$ deux (et seulement deux) extremums utiles (correspondant aux extremums du signal interne). Selon la nature de $F(\cdot)$, ces deux extremums peuvent être soit deux maximums, soit deux minimums soit un maximum et un minimum. Nous allons focaliser notre attention sur une période $[t_0, t_0 + T]$ de $y(t, N)$ où t_0 est arbitrairement choisi. Notons T_1 et T_2 les deux instants correspondant aux deux extremums utiles de $y(t, N)$ sur la période considéré. Ces deux instants vérifient :

$$T_2 - T_1 = \pi / \omega = T / 2$$

Ces deux instants peuvent être obtenus à l'issue de l'étape INN-2. Il est évident que si T_1 correspond à l'instant de passage du signal interne $u(t)$ par un maximum, alors T_2 sera l'instant de passage de $u(t)$ par son minimum et vice versa. Nous ne pouvons pas savoir lequel des deux cas de figure correspond à la réalité ! Mais ceci n'a aucune importance. On peut supposer l'un ou l'autre cas, sans que cela n'ait d'incidence négative sur le résultat d'identification. En effet, ces deux cas de figures renvoient simplement aux deux modèles du système $\{M^+(V), M^-(V)\}$. Cette observation nous permet d'énoncer le résultat suivant.

Proposition 6.4. Considérons le système (6.1) soumis à l'entrée sinusoïdale $v(t)$. Soit T_1 un instant quelconque où survient un extremum utile de $y(t, N)$. Alors, il existe un modèle de Wiener M^* appartenant $\{M^+(V), M^-(V)\}$ tel que le signal interne correspondant, noté u^* , atteigne son

maximum (égal à 1) à l'instant T_l □

Preuve. Au vu de (6.8), (6.9) et (6.11), il vient que $u^+(t)$ atteint ses maximums aux instants t_{2j} ($j=0, 1, 2, \dots$) alors que $u^-(t)$ atteint les siens aux instants t_{2j+1} ($j=0, 1, 2, \dots$). D'autre part, il est évident que :

$$T_l \in \{t_{2j}; j=0, 1, 2, \dots\} \text{ ou } T_l \in \{t_{2j+1}; j=0, 1, 2, \dots\}$$

Si c'est la première égalité qui est vraie, alors $u^*(t) = u^+(t)$ et le modèle qui sera identifié est $M^* = M^+(V)$. Dans l'autre cas, on a $u^*(t) = u^-(t)$ et $M^* = M^-(V)$ □

Remarque 6.5

1. En suivant les notations utilisées dans la Proposition 6.4, l'élément non linéaire associé à M^* est noté $F^*(.)$. Il s'ensuit que $F^* \in \{F_V^+(.), F_V^-(.)\}$ et

$$u^*(t) = F^*(v(t))$$

2. Le résultat que l'on vient d'établir est au centre de la deuxième partie du schéma d'identification de la fonction $F^*(.)$.

6.5.2. Identification des non-linéarités non monotones (INNMM): deuxième partie

INNMM-4. Soit $y(t, N)$ l'un des signaux filtrés obtenus dans l'étape INNMM-1. Repérer dans l'enregistrement correspondant un instant de type T_l (instant d'occurrence d'un extremum utile). Poser $u^*(t) = \cos(\omega(t - T_l))$.

INNMM-5. Tracer la courbe $(u^*(t), y(t, N))$ pour $T_l \leq t \leq T_l + 2\pi/\omega$ puis en extraire un échantillon de points $(u^*(t_j), y(t_j, N))$ ($j=1, \dots, P$). Elaborer une estimation $\hat{F}_N(.)$ de la non-linéarité de la fonction $F^*(.)$ en opérant une interpolation sur l'échantillon des points acquis. Il va de soi que la qualité de l'estimation sera d'autant meilleure que P et N sont grands.

6.6. SIMULATION

Pour illustrer la procédure d'identification proposée, on considère le système de Wiener caractérisé par:

$$G(s) = \frac{12(s-1)}{(s+1)(s+2)}, \quad F(u) = \frac{10u}{1+u^2}$$

Le bruit $\xi(t)$ est simulé par une séquence de nombres aléatoires de variance $\sigma^2 = 0.5$ et de moyenne nulle. Comme $F(\cdot)$ n'est pas strictement monotone, on fait recours à la procédure INNMM. L'étape 1 consiste à appliquer successivement au système au moins deux signaux sinusoïdaux de même fréquence. Prenons par exemple :

$$v_1(t) = V_1 \cos \omega t, \quad v_2(t) = V_2 \cos \omega t \quad (6.27)$$

avec:

$$V_1 = 1.1, \quad V_2 = 0.7, \quad \omega = \pi \text{ rad/s} \quad (6.28)$$

Comme suggéré plus haut, on adoptera la notation :

$$v(t) = V \cos \omega t \quad \text{avec } V=V_1 \quad (6.29)$$

Le but de l'identification est alors de déterminer la non-linéarité $F^* \in \{F_V^+, F_V^-\}$ du modèle M^* défini dans la Proposition 6.4. En appliquant (6.6), il vient de (6.28)-(6.29) que les non-linéarités mises en jeu ici sont les suivantes :

$$F_V^+(u) = \frac{35.44u}{1+(3.544u)^2} \quad (6.30a)$$

$$F_V^-(u) = -\frac{35.44u}{1+(3.544u)^2} \quad (6.30b)$$

Comme le suggère l'étape 1, on commence par enregistrer les sorties $(y_1(t), y_2(t))$ générées par le système en réponse aux deux entrées (v_1, v_2) . La Fig.6.2 montre ces réponses sur une période $T=2\pi/\omega=2$. Les sorties filtrées $y_1(t, N)$ et $y_2(t, N)$, obtenues en appliquant (6.15) pour $N = 95$, sont représentées sur une période par la Fig. 6.3.

En suivant l'étape 2 de la procédure INNMM, il apparaît clairement que, en passant de $y_1(t, N)$ à $y_2(t, N)$, deux extremums glissent verticalement, il s'agit des extremums utiles dus à $u(t)$. Nous pouvons également observer que d'autres extremums, qui sont ici au nombre de quatre, glissent horizontalement. Ce sont des extremums inutiles, dus à $F(u)$. L'étape 3 est donc sans objet dans cet

exemple.

Sur la Fig 6.3, on lit que l'un des deux extremums utiles (en l'occurrence le premier) survient à l'instant $T_1 = \frac{63 * T}{200}$. L'étape 4 de la procédure suggère alors de prendre comme signal interne $u^*(t) = \cos(\omega(t - T_1))$.

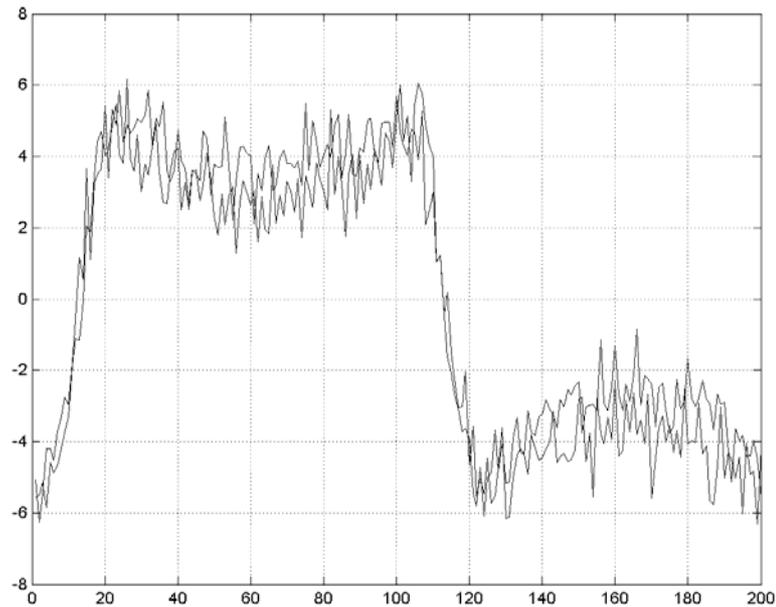


Fig. 6.2. Les sorties $(y_1(t), y_2(t))$ sur une période.

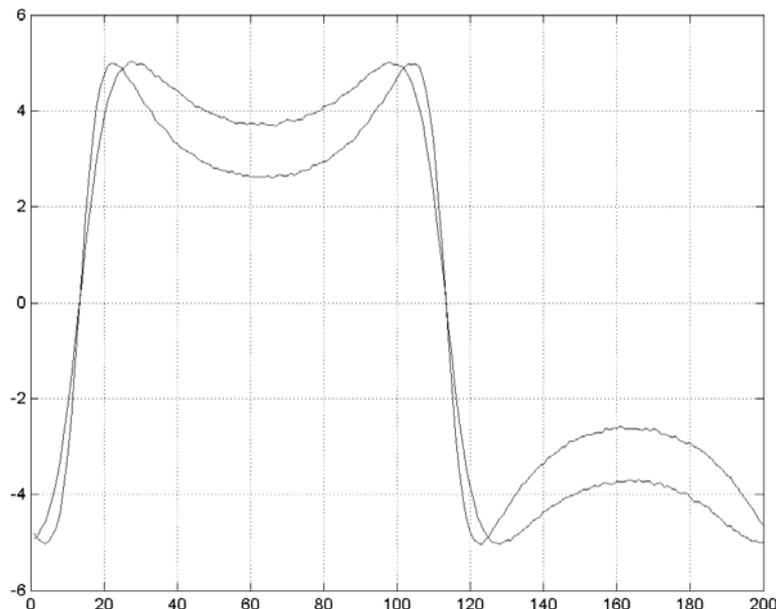


Fig 6.3. Les sorties filtrées $(y_1(t, N), y_2(t, N))$

Conformément à l'étape, on a tracé la courbe $(u^*(t), y(t, N))$ (Fig. 6.4). L'allure de cette courbe

donne déjà une idée sur la nature de l'élément non linéaire et aide dans le choix d'un type d'interpolation approprié (voir la discussion Paragraphe 7 du Chapitre 3). Dans le cas présent, le choix s'est porté sur une interpolation polynomiale, soit $\hat{F}(u) = \sum_{i=1}^p c_i u^i$. Pour un degré p donné, les meilleurs coefficients au sens des moindres sont déterminés à partir d'un échantillon de points $(u^*(t_j), y(t_j, N))$. La fig. 6.4 illustre la qualité de l'estimation de $\hat{F}(u) = \sum_{i=1}^p c_i u^i$ pour $p=14$. Cette valeur élevée de p est justifiée par le fait que la non linéarité n'est pas de nature polynomiale.

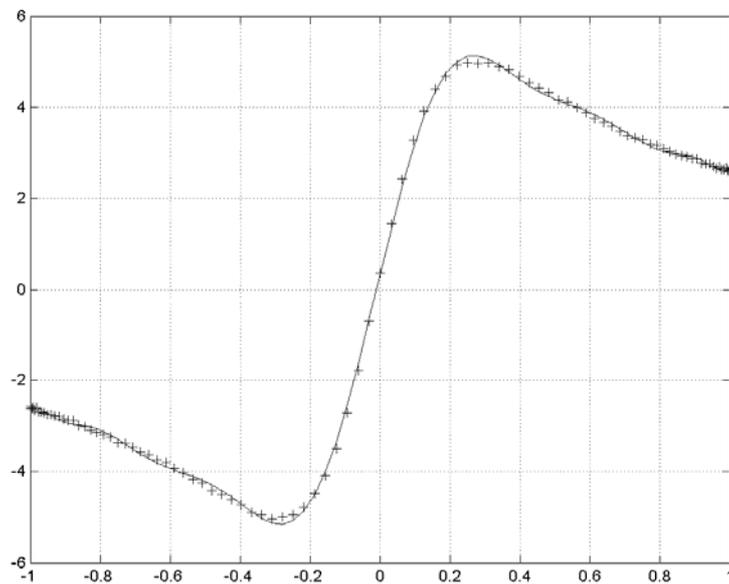


Fig.6.4. Echantillon de points $(u^*(t_j), y(t_j, N))$ (++) et caractéristique de $\hat{F}(u)$ pour $p=14$ (—)

6.7. CONCLUSION

Dans ce chapitre, nous avons développé une solution au problème d'identification de l'élément non linéaire, noté $F(\cdot)$, d'un modèle de Wiener. La solution proposée est de nature fréquentielle dans la mesure où elle ne fait intervenir que des signaux d'excitation sinusoïdaux de la forme $v(t)=V\cos(\omega t)$. Elle est aussi non paramétrique puisque, d'une part, aucune hypothèse n'est émise sur la structure de l'élément non linéaire et, d'autre part, le résultat de l'identification est la caractéristique graphique de la fonction $F(\cdot)$. En outre, et contrairement à plusieurs travaux antérieurs, cette dernière n'est pas supposée inversible ni même lisse. La détermination de $F(\cdot)$ revient à identifier le signal interne $u(t)=F(v(t))$ qui, dans le contexte fréquentiel considéré, est à son tour sinusoïdal et oscille à la même fréquence que l'entrée. L'amplitude de ce signal peut être fixé a priori du fait de la non unicité du modèle. L'intérêt s'est donc porté sur les deux modèles, notés $M^+(U), M^-(U)$, pour lesquels l'amplitude du signal interne vaut 1. A partir de là, la seule incertitude qui continua à peser sur ce signal limita à son déphasage (par rapport à l'entrée), noté φ .

Par commodité pédagogique, nous avons traité séparément le cas de non-linéarités strictement monotones (paragraphe 6.4). La procédure d'identification proposée dans ce cas est très simple et n'exige qu'une seule expérience. Le cas général fut ensuite appréhendé au paragraphe 6.5. La procédure d'identification nécessite alors au moins deux expériences impliquant des signaux d'entrée sinusoïdaux, ayant la même fréquence mais avec des amplitudes différentes. Dans tous les cas, la procédure d'identification consiste à déterminer le déphasage du signal interne du système par rapport à l'entrée. Cette détermination passe essentiellement par la reconnaissance des extremums utiles de la sortie du système, en mettant à profit l'analyse théorique du Paragraphe 6.3. L'effet du bruit, supposé blanc et ergodique, fut contourné moyennant un filtrage de la sortie.

L'originalité de la solution élaborée dans ce chapitre réside, d'une part, dans sa simplicité de mise en œuvre et, d'autre part, dans son large champ d'applications puisque l'élément non linéaire peut être inversible ou non, lisse ou non, polynomial ou non, etc. En plus, la détermination de la caractéristique graphique de l'élément non linéaire (même si le tracé est entaché d'une certaine erreur due à un filtrage insuffisant de la sortie) permet d'avoir une idée sur la structure de la non-linéarité et aider dans le choix d'un type d'interpolation adéquat : polynomial, série de fonctions orthogonales, fonctions de spline, etc. Le type d'interpolation est important puisqu'il détermine le nombre de coefficients nécessaires pour arriver à une bonne estimation de la non-linéarité. Il

convient de souligner ce nombre peut s'avérer trop grand en présence de non-linéarités non lisses, notamment avec une interpolation polynomiale.

L'approche développée dans ce chapitre s'appuie sur une démarche graphique en vue de localiser les extremums utiles. Ce procédé a l'avantage d'être simple mais, comme dans toute démarche graphique, la localisation est forcément approximative. Il va sans dire que la qualité de l'identification peut être améliorée en ne se limitant pas à deux expériences. Si cela s'avérait insuffisant, on peut toujours recourir à des procédés analytiques de recherche d'extremums.

CHAPITRE 7

IDENTIFICATION FREQUENTIELLE DE L'ELEMENT NON LINEAIRE DES SYSTEMES DE WIENER

APPROCHE PAR TRACE DES COURBES DE LISSAJOUS

7.1. INTRODUCTION

Dans ce chapitre, un nouveau schéma d'identification, basé sur des outils géométriques et fréquentiels, est élaboré. L'accent est mis sur l'identification du gain non linéaire, car si ce dernier est estimé, l'identification du sous-système linéaire ne présente pas un grand problème surtout si le gain non linéaire est inversible. L'intérêt de cette méthode d'identification consiste en l'utilisation d'un signal sinusoïdal de fréquence unique à l'entrée du système $v(t) = V \cos(\omega t)$ pour une seule fréquence. En régime permanent, le signal interne est alors sinusoïdal de la forme $u(t) = U \cos(\omega t - \varphi)$. Il est toujours possible, en utilisant une mise à l'échelle adéquate, de supposer que le signal interne est d'amplitude unitaire (c'est-à-dire $U=1$). Si on trace le lieu fermé $(\cos(\omega t - \psi), y(t, N))$ pour différentes valeurs de $\psi \in [0, \pi]$ on obtient des courbes de Lissajous (Weisstein, 2006); $y(t, N)$ est une version filtrée du signal de sortie $y(t)$. On démontre, sous certaines conditions que, la surface bordée par cette courbe est non nulle sauf pour $\psi = \varphi + k\pi$ (pour tout entier k). De plus lorsque $\psi = \varphi + k\pi$, le lieu précédent décrit $F(x)$ ou $F(-x)$. Cette notion mathématique constitue l'ingrédient principal de la méthode d'identification présentée dans ce chapitre.

Ce chapitre est organisé comme suit le problème d'identification est défini dans le paragraphe 2, puis il est reformulé dans le paragraphe 3; dans les paragraphes 4 et 5 une analyse mathématique des lieux $(\cos(\omega t - \psi), y(t, N))$ précédemment décrit est établie, les paragraphes 6 et 7 présentent deux versions de l'approche proposée: une graphique et l'autre analytique.

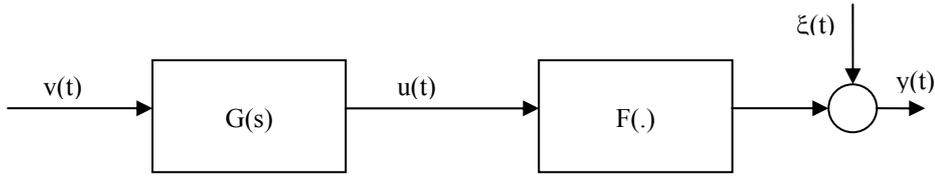


Fig. 7.1. Modèle de Wiener

7.2. POSITION DU PROBLEME D'IDENTIFICATION

7.2.1. Classe de systèmes à identifier

On considère les systèmes décrits par le modèle de Wiener (Fig. 7.1) où $G(s)$ représente la fonction de transfert du sous système linéaire et $F(.)$ représente la caractéristique statique du gain non linéaire. Ce modèle est analytiquement décrit par les équations suivantes :

$$u(t) = g(t) * v(t) \quad (7.1)$$

$$y(t) = F(u(t)) + \xi(t) \quad (7.2)$$

où $g(t)$ représente la transformée inverse de Laplace de $G(s)$; le symbole ' * ' représente l'opérateur de convolution; $v(t)$ et $y(t)$ représentent le signal appliqué à l'entrée et le signal mesuré à la sortie du système, respectivement. $u(t)$ est le signal interne non mesurable et $\xi(t)$ est un bruit externe.

Hypothèses pour le sous-système linéaire:

H7.1. $G(s)$ est asymptotiquement stable.

H7.2. $\{\xi(t)\}$ est un processus stochastique, stationnaire et ergodique, de moyenne nulle.

Remarques 7.1

- 1) L'hypothèse H7.1 est nécessaire car l'identification se fait en boucle ouverte.
- 2) L'hypothèse d'ergodicité est suffisante pour que l'analyse qui sera présentée soit valide. Elle sera utile par la suite pour remplacer la moyenne probabiliste par la moyenne arithmétique, lors du calcul de la version filtrée de la sortie.
- 3) A l'exception des hypothèses H7.1-H7.2, le sous système linéaire peut être arbitraire. En particulier aucune hypothèse n'est faite sur la nature et la structure de $G(s)$. En plus, les variables aléatoires du processus $\{\xi(t)\}$ sont nécessairement indépendantes.

Hypothèses pour le gain non linéaire $F(\cdot)$

L'identification fréquentielle de $F(\cdot)$ nécessite l'application d'une sinusoïde en entrée, $v(t) = V \cos(\omega t)$. Le choix de V et ω doit être tel que la fonction $F(\cdot)$ soit définie dans l'intervalle $-V|G(j\omega)| \leq u \leq V|G(j\omega)|$ et satisfasse les hypothèses suivantes:

H7.3:

a) Si $F(\cdot)$ est paire alors, $F(0)$ est connu et $F^{-1}(F(0)) = \{0\}$

b) Si $F(\cdot)$ n'est pas paire il doit exister alors σ_0 et σ_1 avec $-1 \leq \sigma_0 < \sigma_1 \leq 1$ tels que $F(\cdot)$ est bijective sur le sous intervalle $]\sigma_0 V|G(j\omega)|, \sigma_1 V|G(j\omega)|[$. i.e $\forall u \in]\sigma_0 V|G(j\omega)|, \sigma_1 V|G(j\omega)|[$:
 $F^{-1}(F(u)) = \{u\}$.

Remarques 7.2

1) L'hypothèse H7.3 (partie b) n'est pas très restrictive puisque la taille du l'intervalle $]\sigma_0 V|G(j\omega)|, \sigma_1 V|G(j\omega)|[$ peut être arbitrairement petit et totalement inconnu. En particulier la fonction $F(\cdot)$ peut être non inversible, à l'exception d'un petit intervalle ouvert $]\sigma_0 V|G(j\omega)|, \sigma_1 V|G(j\omega)|[$.

2) A l'exception de ces hypothèses, $F(\cdot)$ peut être quelconque. En particulier, elle peut être discontinue, non inversible et non lisse. De ce fait, notre contribution se distingue des précédents travaux, puisque l'approche proposée est valable pour un large domaine de non linéarité, à l'inverse des autres approches qui restent valables pour des classes de non linéarités bien spécifiques (polynomiales, ou continues, ou lisses, ou inversibles ou tout simplement à structure connue). L'hypothèse H7.3 est vérifiée par la plus part des non linéarités rencontrées en pratique (gain avec saturation, zone morte, seuil, valeur absolue, arc tangente, et tout polynôme qui vérifie H7.3 ...). Si on ne dispose d'aucune information sur la nature de $F(v)$ par rapport aux exigences de H7.3, l'approche est toujours applicable mais l'admissibilité de la fonction estimée dépendra de son conformité avec H7.3. Ce point sera détaillé par la suite.

7.2.2. Objectif d'identification

Notre objectif est d'élaborer un schéma d'identification capable de fournir, de manière non paramétrique, la caractéristique $F(.)$ du gain non linéaire sur un intervalle choisi $-V|G(j\omega)| \leq u \leq V|G(j\omega)|$.

La principale difficulté vient du fait que le signal interne $u(t)$ est non mesurable. Le processus d'identification doit donc utiliser seulement les mesures des signaux d'entrée et de sortie $v(t)$ et $y(t)$ respectivement. De ce fait le problème d'identification n'a pas de solution unique, i.e si le couple $(G(s), F(u))$ représente une solution au problème alors tout couple $(G(s)/\alpha, F(\alpha u))$ est aussi solution, pour tout réel α non nul. Ceci implique aussi que le signal interne peut être complètement théorique et n'a aucune existence physique.

Donc, l'objectif d'identification se ramène à l'identification d'un couple $(G(s), F(u))$ qui satisfait les équations (7.1)-(7.2). Ce manque d'unicité de la solution sera exploité dans notre approche (paragraphe 3) pour supposer que l'amplitude de $u(t)$ est connue et unitaire.

7.3. MISE A L'ECHELLE DU MODELE ET REFORMULATION DU PROBLEME D'IDENTIFICATION

Le schéma d'identification que nous allons élaborer se base sur deux idées principales. La première consiste en une mise à l'échelle du modèle, ce qui va permettre de diminuer l'incertitude qu'on a sur le signal interne non mesurable. La seconde consiste en la caractérisation de la relation entre le signal interne du nouveau modèle mis à l'échelle, et le signal de sortie et ainsi tracer la caractéristique du gain non linéaire.

7.3.1. Mise à l'échelle

Le long de cette section, le système à identifier est excité par une sinusoïde de la forme:

$$v(t) = V \cos(\omega t) \quad (t \geq 0) \tag{7.3}$$

où l'amplitude $V > 0$ et la pulsation ω sont maintenues constantes. En régime permanent, et d'après (7.1), le signal interne $u(t)$ prend la forme suivante:

$$u_v(t) = U_v \cos(\omega t - \varphi) \tag{7.4}$$

avec $U_V = V|G(j\omega)|$, $\varphi = -\arg(G(j\omega))$ et le signal de sortie résultant sera $y(t) = F(u_V(t)) + \xi(t)$. D'autre part si le signal $u_V(t)$ était mesurable, il serait possible de déterminer un estimé de $F(\cdot)$ en utilisant l'ensemble de points $(u_V(t), y(t))$ et un processus de filtrage de la sortie. Mais, $u_V(t)$ est non mesurable puisque ni son amplitude U_V ni sa phase φ ne sont connues. La première incertitude peut être résolue avec une mise à l'échelle du modèle (7.1), c'est-à-dire en exploitant le fait que le modèle n'est pas unique. Pour cela, nous allons considérer les deux modèles mis à l'échelle suivants:

$$M^+(U) \stackrel{\text{def}}{=} (F_V^+, G_V^+(s)) \quad (7.5a)$$

$$M^-(U) \stackrel{\text{def}}{=} (F_V^-, G_V^-(s)) \quad (7.5b)$$

où:

$$F_V^+(u) \stackrel{\text{def}}{=} F(U_V u), \quad F_V^-(u) \stackrel{\text{def}}{=} F(-U_V u) \quad (7.6)$$

$$G_V^+(s) \stackrel{\text{def}}{=} \frac{1}{U_V} G(s), \quad G_V^-(s) \stackrel{\text{def}}{=} \frac{1}{-U_V} G(s) \quad (7.7)$$

$$u(t) = \cos(\omega t - \varphi)$$

Les signaux internes correspondant aux deux modèles précédents sont respectivement:

$$u^+(t) = \cos(\omega t - \varphi), \quad u^-(t) = -\cos(\omega t - \varphi) \quad (7.8)$$

De ce fait, le signal interne (pour les deux nouveaux modèles) ne dépend plus de l'amplitude V (contrairement au signal $u_V(t)$ correspondant au modèle initial (7.1)). Toutefois, les deux modèles génèrent le même signal de sortie que le modèle initial (7.1), quand le signal d'entrée $v(t)$ est un signal sinusoïdal (7.3). C'est-à-dire,

$$y(t) = y_V(t) + \xi(t) \quad (7.9)$$

où:

$$y_V(t) = F_V^+(u^+(t)) = F_V^-(u^-(t)) = F(u_V(t)) \quad (7.10)$$

Cette propriété sera très utile pour résoudre le problème de la non connaissance de la phase φ dans (7.8). Nous n'avons aucun moyen pour l'instant de distinguer les deux modèles: l'estimé qu'on peut

obtenir pour $F(\cdot)$ peut bien correspondre à l'un des deux modèles (soit $F_V^+(u^+(t))$ ou $F_V^-(u^-(t))$).

7.3.2. Reformulation du problème d'identification

Dans la suite de ce chapitre nous allons chercher à estimer l'une ou l'autre des deux fonctions $F^+(\cdot)$ ou $F^-(\cdot)$, puisque ces deux fonctions ne peuvent pas être distinguées, en utilisant seulement l'entrée et la sortie du système. Toutefois, il n'est pas très important de savoir laquelle des deux fonctions sera estimée. La méthode d'identification que nous allons présenter est basée sur deux propriétés clés qui sont communes aux deux fonctions $F^+(\cdot)$ et $F^-(\cdot)$. La première est une conséquence directe de l'hypothèse H7.3. Elle est le sujet de la proposition 7.1. La seconde propriété, de nature géométrique, sera établie dans le paragraphe 4.

Proposition 7.1. Considérons le système (7.1), sous les hypothèses H7.1-H7.3, excité par l'entrée sinusoïdale (7.3) où $V > 0$ et $\omega > 0$ sont arbitraires mais constantes. Alors, pour tout $H \in \{F^+, F^-\}$, l'une des propositions suivantes est vérifiée:

- 1) H est paire, $H(0) = F(0)$ est connue et $H^{-1}(H(0)) = \{0\}$
- 2) H est non paire mais il existe $-l \leq \sigma_0 < \sigma_1 \leq l$ tels que pour tout $u \in (\sigma_0, \sigma_1)$: $H^{-1}(H(u)) = \{u\}$

□

Cette proposition stipule tout simplement que si $F^+(\cdot)$ ou $F^-(\cdot)$ est une solution du problème d'identification tel qu'il est reformulé, alors elle vérifie l'une ou l'autre des deux propositions de l'hypothèse H7.3. On définit alors la notion d'admissibilité de toute fonction qui est solution du problème, puisque on s'intéresse à l'identification de $F^+(\cdot)$ ou $F^-(\cdot)$.

Définition 7.1. Une fonction $H : [-l, +l] \rightarrow IR$ est dite admissible, comme solution du problème d'identification considéré, si elle satisfait l'une des deux propriétés de la proposition 7.1 □

Ceci nous donne alors un moyen de vérifier après estimation si la fonction obtenue peut être admissible ou non. Si aucune des solutions obtenues n'est admissible, on peut conclure alors que le

système ne vérifie pas les hypothèses H7.3 et ne peut pas être donc identifié par cette approche.

7.4. CARACTERISATION GEOMETRIQUE DES COURBES $(u^-(t), y_V(t))$ ET $(u^+(t), y_V(t))$

Dans cette section, nous allons chercher les propriétés géométriques des courbes paramétrées $(u^-(t), y_V(t))$ et $(u^+(t), y_V(t))$. Le but est de pouvoir reconstituer $(u^-(t))$ et $(u^+(t))$ à partir de ces propriétés.

Notons d'abord que tout signal sinusoïdal de même fréquence que $u^-(t)$ et $u^+(t)$ et d'amplitude unitaire aura la forme suivante :

$$u_\psi(t) = \cos(\omega t - \psi) \quad (7.11)$$

où $\psi \in \mathbf{IR}$ est quelconque. On peut alors écrire:

$$u^+(t) = u_\psi(t), \quad u^-(t) = u_{(\psi+\pi)}(t) \quad (7.12)$$

Soit C_ψ le lieu des points de coordonnées $(u_\psi(t), y_V(t)); (t \geq 0)$. $u_\psi(t)$ est périodique de période $T = 2\pi / \omega$, et $y_V(t)$ est périodique de période $T/n = 2\pi / n\omega$; n un entier ($n \geq 1$) qui dépend de la nature de $F(\cdot)$. La courbe C_ψ devient alors un lieu orienté et fermé. En général, C_ψ est constitué d'une ou plusieurs boucles (orientées) (Fig. 7.2). Dans le cas particulier où, $u_\psi(t)$ et $y_V(t)$ sont sinusoïdaux, C_ψ est une courbe de Lissajous et peut avoir plusieurs formes selon le déphasage entre les deux signaux, par exemple: une ellipse, un cercle ou une droite ((Weisstein, 2006)).

La seule caractéristique de C_ψ qui nous intéresse dans cette étude sera son aire géométrique qu'on notera $A(\psi)$. Précisons, d'abord, que l'aire géométrique d'une simple boucle (orientée) ne tient pas compte de son orientation, et elle est toujours positive ou nulle. Il est nécessaire de la distinguer de l'aire algébrique qui peut être positive, nulle ou négative selon le sens d'orientation de la boucle. Dans le cas d'un lieu avec plusieurs boucles orientées, l'aire géométrique globale est la somme des aires géométriques des différentes boucles. Sur la Fig. 7.2 on a représenté un lieu constitué de deux cercles de même rayon r , mais orientés en sens inverse, l'aire géométrique globale est égale alors à $2\pi r^2$ alors que l'aire algébrique est nulle.

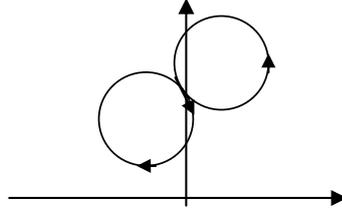


Fig. 7.2. Exemple de lieu orienté, fermé

Cette introduction nous permet de donner alors la définition suivante:

Définition 7.2. Le lieu fermé C_ψ est dit statique si son aire géométrique est nulle ($A(\psi) = 0$). Dans ce cas, le lieu C_ψ a la forme d'une ligne courbée (sans aucune boucle). Inversement, C_ψ est dit une courbe non statique si $A(\psi) \neq 0$.

Proposition 7.2. Considérons le problème défini dans la proposition 7.1. $C_{\varphi+k\pi}$ est alors statique pour tout entier k ($k = 0, \pm 1, \pm 2, \dots$) \square

Preuve: A partir (7.10) et (7.12) on déduit que:

$$y_V(t) = F_V^+(u_\varphi(t)) = F_V^-(u_{\varphi+\pi}(t))$$

ce qui implique, grâce à (7.11), que:

$$y_V(t) = H_k(u_{\varphi+k\pi}(t)) \quad (7.13)$$

où

$$H_k = \begin{cases} F_V^+ & \text{si } k \text{ est pair} \\ F_V^- & \text{si } k \text{ est impair} \end{cases} \quad (7.14)$$

La proposition est alors établie à partir de (7.13) en utilisant le fait que $H_k(\cdot)$ est une fonction. \square

Proposition 7.3. Considérons le problème défini dans la proposition 7.1 et supposons que C_ψ est un lieu statique. Alors, pour tout θ , on a:

$$F_V^+(cos(\theta)) = F_V^+(cos(\theta - 2(\psi - \varphi)))$$

$$F_V^-(cos(\theta)) = F_V^-(cos(\theta - 2(\psi - \varphi))) .$$

Preuve. A partir de (7.10) on a, pour tout t:

$$y_V(t) = F_V^+(cos(\omega t - \varphi)) = F_V^-(-cos(\omega t - \varphi))$$

D'autre part, si C_ψ est statique alors il existe une fonction $H(.)$ telle que:

$$y_V(t) = H(cos(\omega t - \psi)) \quad (7.15)$$

Donc, on a, pour tout t:

$$F_V^+(cos(\omega t - \varphi)) = H(cos(\omega t - \psi)) \quad (7.16a)$$

$$F_V^-(-cos(\omega t - \varphi)) = H(cos(\omega t - \psi)) \quad (7.16b)$$

D'un autre côté, on peut vérifier aisément que pour tout t:

$$cos\left(\omega\left(t + \frac{\psi}{\omega}\right) - \psi\right) = cos\left(\omega\left(T - t + \frac{\psi}{\omega}\right) - \psi\right) \quad (7.17)$$

où $T = 2\pi / \omega$. Donc, on déduit que pour tout t:

$$H\left(cos\left(\omega\left(t + \frac{\psi}{\omega}\right) - \psi\right)\right) = H\left(cos\left(\omega\left(T - t + \frac{\psi}{\omega}\right) - \psi\right)\right)$$

cette dernière équation, combinée avec (7.16a-b), implique, pour tout t:

$$F_V^+\left(cos\left(\omega\left(t + \frac{\psi}{\omega}\right) - \varphi\right)\right) = F_V^+\left(cos\left(\omega\left(T - t + \frac{\psi}{\omega}\right) - \varphi\right)\right)$$

$$F_V^-\left(-cos\left(\omega\left(t + \frac{\psi}{\omega}\right) - \varphi\right)\right) = F_V^-\left(-cos\left(\omega\left(T - t + \frac{\psi}{\omega}\right) - \varphi\right)\right)$$

Ce qui donne en développant d'avantage, respectivement, pour tout t:

$$F_V^+(cos(\omega t + \psi - \varphi)) = F_V^+(cos(\omega T - \omega t + \psi - \varphi))$$

$$= F_V^+(cos(\omega t - \psi + \varphi))$$

De même, on a

$$F_V^-(-cos(\omega t + \psi - \varphi)) = F_V^-(-cos(\omega T - \omega t + \psi - \varphi))$$

$$= F_v^-(-\cos(\omega t - \psi + \varphi))$$

En introduisant la variable $\theta = \omega t + \psi - \varphi$ dans les équations précédentes, on établit alors la proposition 7.3 \square

Proposition 7.4. Considérons le problème défini dans la proposition 7.1 et supposons que C_ψ est statique. Alors, on a les propriétés suivantes:

1) si $F(\cdot)$ est non paire, alors:

$$\psi - \varphi = h\pi, \text{ pour } h=0, \pm 1, \pm 2, \dots$$

2) si la fonction $F(\cdot)$ est paire, alors:

$$\psi - \varphi = h\pi \text{ ou } \frac{\pi}{2} + h\pi, \text{ pour } h=0, \pm 1, \pm 2, \dots$$

Preuve. Introduisons les notations suivantes:

$$2(\psi - \varphi) = \delta + 2h\pi \tag{7.18}$$

où:

$$0 \leq \delta < 2\pi \text{ et } h=0, \pm 1, \pm 2, \dots \tag{7.19}$$

Alors en utilisant la proposition 7.3, on obtient, pour tout $\theta \in [0, 2\pi[$

$$F_v^+(\cos(\theta)) = F_v^+(\cos(\theta - \delta)) \tag{7.20}$$

Nous allons traiter les deux cas séparément, d'abord le cas où $F(\cdot)$ est non paire, puis le cas où elle est paire.

Cas 1: La fonction $F(\cdot)$ est non paire

A partir de (7.6), en utilisant l'hypothèse H7.3.b, on déduit que la fonction F_v^+ est une bijection sur un intervalle $] \sigma_0, \sigma_1 [$. Alors, (7.20) implique que, pour tout $\theta \in]0, 2\pi[$ tel que,

$$\begin{aligned} \cos(\theta) &\in] \sigma_0, \sigma_1 [\\ \cos(\theta) &= \cos(\theta - \delta) \end{aligned} \tag{7.21}$$

Mais, on peut vérifier aisément que, si $\delta \neq 0$ alors, pour tout $\theta \in]0, 2\pi[- \left\{ \frac{\delta}{2}, \pi + \frac{\delta}{2} \right\}$:

$$\cos(\theta) \neq \cos(\theta - \delta) \tag{7.22}$$

Ceci contredit alors (7.21) sur $] \sigma_0, \sigma_1 [$. Par contre, si $\delta = 0$, c'est-à-dire $(\psi - \varphi) = h\pi$, d'après (7.18), (7.21) sera vérifié pour tout $\theta \in [0, 2\pi[$, et $F(\cdot)$ sera statique.

Case 2: la fonction $F(.)$ est paire

A partir de (7.6) on voit clairement que F_V^+ est aussi paire. Posons $\theta = \pi/2$ dans (7.20), on obtient

$F_V^+(0) = F_V^+\left(\cos\left(\frac{\pi}{2} - \delta\right)\right)$. Donc l'hypothèse H7.3.a) implique que $\cos\left(\frac{\pi}{2} - \delta\right) = 0$ ou, de manière

équivalente:

$$\delta = 2i\pi \quad \text{ou} \quad \delta = \pi - 2i\pi \quad \text{pour un entier } i=0, \pm 1, \pm 2 \dots \quad (7.23)$$

D'après (7.18), Ceci implique que:

$$\psi - \varphi = (h-i)\pi \quad \text{ou} \quad \psi - \varphi = \frac{\pi}{2} + (h-i)\pi \quad \square$$

Proposition 7.5. Considérons le problème défini dans Proposition 7.1. Si C_ψ est statique alors il existe une seule fonction admissible $H(.)$, telle que:

$$H(\cos(\omega t - \psi)) = y_V(t), \quad \forall t$$

Plus précisément,

$$H(.) = \begin{cases} F_V^+(.) & \text{si } \psi - \varphi = 2h\pi \\ F_V^-(.) & \text{si } \psi - \varphi = \pi + 2h\pi \quad (\text{pour tout entier } h) \end{cases}$$

Preuve. Le fait que C_ψ est statique garantit l'existence d'une fonction $H(.)$ telle que:

$$H(\cos(\omega t - \psi)) = y_V(t), \quad \forall t \quad (7.24)$$

Pour prouver l'unicité de $H(.)$, les deux cas de l'hypothèse H7.3 sont considérés séparément.

Cas 1: La fonction $F(.)$ est non paire

En utilisant la Proposition 7.4 (partie 1) on a $\psi - \varphi = h\pi$, pour $h=0, \pm 1, \pm 2, \dots$. Donc (7.24), implique alors, que:

$$H((-1)^h \cos(\omega t - \varphi)) = y_V(t), \quad \forall t \quad (7.25)$$

En comparant (7.25) et (7.10), on a:

$$H((-1)^h \cos(\omega t - \varphi)) = F_V^+(u^+(t)) = F_V^-(u^-(t))$$

cette équation donne, grâce à (7.8),

$\forall z \in [-1, +1]$:

$$H(z) = F_V^+(z) \text{ si } h \text{ est pair} \quad (7.26)$$

$$H(z) = F_V^-(z) \text{ si } h \text{ est impair} \quad (7.27)$$

Donc, la proposition 7.5 est vérifiée pour le cas 1.

Cas 2: la fonction $F(\cdot)$ est paire

En utilisant la Proposition 7.4 (Partie 2), il s'ensuit que:

$$\psi - \varphi = h\pi \text{ ou } \psi - \varphi = \frac{\pi}{2} + h\pi \text{ pour } h=0, \pm 1, \pm 2, \dots \quad (7.28)$$

Nous allons montrer par contradiction que la deuxième solution de (7.28) ne peut pas être vérifiée.

Pour cela supposons que pour un certain h :

$$\psi - \varphi = \frac{\pi}{2} + h\pi \quad (7.29)$$

A partir de (7.24) on a:

$$H((-1)^h \sin(\omega t - \varphi)) = y_V(t), \quad \forall t \quad (7.30)$$

En comparant (7.30) et (7.10), on obtient:

$$H((-1)^h \sin(\omega t - \varphi)) = F_V^+(\cos(\omega t - \varphi))$$

Cette équation peut s'écrire dans un format plus compact, en posant $\theta = \omega t - \varphi$:

$$H((-1)^h \sin(\theta)) = F_V^+(\cos(\theta)), \quad \forall \theta \quad (7.31)$$

La substitution de $\theta + \pi$ à θ dans (7.31) implique que, $\forall \theta$:

$$H(-(-1)^h \sin(\theta)) = F_V^+(-\cos(\theta)) \quad (7.32)$$

Puisque F_V^+ est paire, on déduit alors, en comparant (7.31) et (7.32), que $H(\cdot)$ est aussi paire. Donc, on obtient à partir de (7.31) que, $\forall \theta$:

$$F_V^+(\cos(\theta)) = H\left(\sqrt{1 - \cos^2(\theta)}\right) \quad (7.33)$$

En utilisant le changement de variable $z = \sqrt{1 - \cos^2(\theta)}$, on déduit à partir de (7.33) que, $\forall z \in [0, 1]$:

$$H(z) = F_V^+\left(\sqrt{1 - z^2}\right) \quad (7.34)$$

On peut montrer maintenant que cette solution $H(\cdot)$, est non admissible (voir Définition 7.1). En effet, on peut écrire directement à partir de (7.34) que:

$$H(0) = F_V^+(1) \quad (7.35)$$

D'autre part, à partir de la Proposition 7.1 (Partie 1) on a:

$$F_V^+(u) = F_V^+(0) \Rightarrow u = 0 \quad (7.36)$$

En utilisant (7.36), on obtient à partir de (7.35) que $H(0) \neq F_V^+(0)$ c'est-à-dire $H(0) \neq F(0)$. Ceci prouve clairement que $H(\cdot)$ est non admissible (voir Proposition 7.1 (Partie 1)). Ainsi, la solution (7.29) ne doit pas être retenue. En vue de (7.28), nous avons alors nécessairement: $\psi - \varphi = k\pi$. Le reste de la preuve est similaire au premier cas. Ceci complète la preuve de la Proposition 7.5 \square

7.5. ANALYSE DES COURBES PARAMETRÉES $(u^-(t), y(t)), (u^+(t), y(t))$

Tout ce qu'on vient de démontrer jusqu'à maintenant fait référence à la sortie non perturbée du système, c'est-à-dire $y_V(t)$, qui n'est pas mesurable; Or seule la sortie $y(t)$ est mesurable. Comment peut-on retrouver $y_V(t)$ à partir de $y(t)$?

Introduisons la notation de sortie filtrée suivante:

$$y(t, N) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{k=1}^N y(t + kT) \text{ pour } t \in [0, T[\quad (7.37a)$$

$$\text{Et } y(t + kT, N) = y(t, N) \quad \forall k \in \mathbb{N} \quad (7.37b)$$

N est un entier quelconque mais suffisamment grand. Pour un instant t donné, dans le régime permanent, la quantité $y(t, N)$ représente la moyenne des valeurs de la sortie du système aux instants $t + kT$.

Soit $C_{\psi, N}$ le lieu paramétré des points de coordonnées $(u_{\psi}(t), y(t, N))$; ($t \geq 0$). Alors, nous pouvons énoncer la proposition suivante:

Proposition 7.7. Considérons le problème défini dans la Proposition 7.1. On a alors:

$$1^\circ) \quad \lim_{N \rightarrow \infty} C_{\psi, N} = C_{\psi} \text{ dans le sens où, pour tout } t \geq 0 :$$

$$\lim_{N \rightarrow \infty} (u_{\psi}(t), y(t, N)) = (u_{\psi}(t), y_V(t)) \quad (7.38)$$

2°) Par conséquent, si $C_{\psi, N}$ devient (presque) statique pour une certaine valeur de ψ et un certain N suffisamment grand, alors:

a) $\psi = \varphi + 2h\pi$ (pour $h=0, \pm 1, \pm 2, \pm 3, \dots$) et le mappage $\cos(\omega t - \psi) \rightarrow y(t, N)$ coïncide (approximativement) avec $F_V^+(\cdot)$

ou

b) $\psi = \varphi + \pi + 2h\pi$ (pour $h=0, \pm 1, \pm 2, \pm 3, \dots$) et le mappage $\cos(\omega t - \psi) \rightarrow y(t, N)$ coïncide (approximativement) avec $F_V^-(\cdot)$.

Il est évident que, plus N est grand, meilleure est la qualité des approximations précédentes. \square

Preuve. A partir de (7.9) on a, pour tout $t \in [0, T[$:

$$y(t) = y_V(t) + \xi(t) \quad (7.39)$$

Donc, en utilisant le fait que $y_V(t) = y_V(t+T)$, on déduit à partir de (7.39) que, pour $t \in [0, T[$ et tout entier k :

$$y(t+kT) = y_V(t) + \xi(t+kT)$$

Ceci implique, pour tout $t \in [0, T[$ et tout entier k :

$$(u_\psi(t+kT), y(t+kT)) = (u_\psi(t), y_V(t)) + (0, \xi(t+kT)) \quad (7.40)$$

où on a utilisé le fait que $u_\psi(t) = \cos(\omega t - \psi)$ est périodique de période T . D'autre part, à partir de l'hypothèse H7.2 on a, pour tout $t \in [0, T[$:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \xi(t+kT) = E(\xi(t+kT)) = 0 \quad (7.41)$$

En combinant (7.40) et (7.41), on obtient, $\forall t \in [0, T[$:

$$\left(u_\psi(t), \lim_{N \rightarrow \infty} \sum_{k=1}^N y(t+kT) \right) = (u_\psi(t), y_V(t))$$

Ceci prouve la partie 1 de la proposition. La partie 2 est une conséquence directe de la première partie en utilisant la proposition 7.5. \square

Après avoir donné tous les éléments nécessaires pour prouver la validité de l'approche d'identification proposée dans ce chapitre, nous allons maintenant décrire son mode opératoire en deux versions. La première version dite graphique se base sur une simple observation graphique de l'évolution du lieu $C_{\psi, N}$ en fonction ψ pour trouver la valeur qui correspond à φ (à $k\pi$ près). La deuxième dite analytique, permet de trouver la valeur du déphasage φ (à $k\pi$ près), en cherchant, par

calcul, la valeur de ψ qui minimise (voire annule) l'aire géométrique du lieu $C_{\psi,N}$.

7.6. VERSION GRAPHIQUE DE LA METHODE D'IDENTIFICATION DE L'ELEMENT NON LINEAIRE

7.6.1. Méthode d'identification graphique (MIG)

La proposition 7.6 suggère la méthode graphique suivante pour obtenir un estimé $\hat{F}(\cdot)$ pour $F^+(\cdot)$ ou $F^-(\cdot)$. On rappelle, ici, qu'il n'est pas important de savoir la quelle est estimée, puisque l'une ou l'autre est une solution au problème d'identification initial.

MIG1. On applique un signal sinusoïdal, $v(t) = V \sin \omega t$ à l'entrée du système non linéaire considéré.

MIG2. On observe la sortie $y(t)$ jusqu'à ce qu'on atteigne le régime permanent, (c'est à dire le signal de sortie devient presque périodique, en faisant abstraction du bruit). Puis, on enregistre le signal de sortie mesuré $y(t)$ pour un nombre de périodes N assez grand. On note cet intervalle de temps $[l, NT]$. On génère la sortie filtrée $y(t, N)$ selon (7.37a-b).

MIG3. On trace la courbe fermée paramétrique:

$$C_{\psi,N} = \{(\cos(\omega t - \psi), y(t, N)), 1 \leq t \leq NT\}$$

pour différentes valeurs de ψ (dans l'intervalle $[0, \pi[$) jusqu'à ce qu'une courbe presque statique soit obtenue (Ceci peut être fait par programme pour accélérer la procédure). Notons ψ^* la valeur de ψ , pour laquelle $C_{\psi^*,N}$ est (presque) statique.

MIG4. Prélever un nombre suffisant de points (u_i, y_i) à partir de $C_{\psi^*,N}$ et déterminer par interpolation adéquate la meilleure fonction $H(\cdot)$ (par exemple au sens des moindres carrés) telle que $y_i = H(u_i)$, pour tout i . Le type d'interpolation à utiliser dépend de la forme $C_{\psi^*,N}$ obtenu.

MIG5. Si $H(\cdot)$ obtenue, n'est pas paire, alors prendre $\hat{F} = H$.

MIG6. Si $H(\cdot)$ obtenue, est paire alors:

$$\hat{F}(z) = \begin{cases} H(z) & \text{si } H(0) = F(0) \\ H(\sqrt{1-z^2}) & \text{sinon} \end{cases} \quad (7.42)$$

Remarques 7.3.

- 1) La procédure graphique proposée est assez simple; elle ne nécessite qu'une seule expérience fréquentielle avec une seule fréquence. Les méthodes fréquentielles utilisées en identification nécessitent, en général, plusieurs expériences avec, comme signaux d'entrée du système, des signaux sinusoïdaux avec des fréquences différentes ((Soderstrom, et Stoica, 1988), (Bai, 2002)).
- 2) Il est évident que $C_{w,N} = \{(\cos(\omega t - \psi), y(t, N)), 1 \leq t \leq NT\}$ sera tracé sur un nombre fini de points; on procède en effet à un échantillonnage de $y(t)$ avec une période T_s . Dans ces conditions, et pour que le filtrage (7.37) soit correcte il faut que T/T_s soit entier (bien sûr assez élevé).

7.6.2. Evaluation de MIG par simulation

Pour illustrer la MIG présentée ci dessus, on considère le système de Wiener suivant caractérisé par:

$$G(s) = \frac{s-1}{(s+2)(s+1)}, \quad F(u) = \text{sign}(u)(\text{abs}(u) + 0.5)$$

Notons que $G(s)$ est à phase non minimale. La non linéarité $F(\cdot)$ est de type seuil avec une discontinuité en $u=0$. Ce choix est volontaire pour prouver que notre approche fonctionne bien pour des non linéarités discontinues; ce qui n'est pas le cas de la plupart des autres approches d'identification.

Le système est excité par $v(t) = V \cos(\omega t)$ avec $V = 2$, $\omega = \pi$ (rad / s) et une perturbation de moyenne nulle $\xi(t)$ avec une variance $\sigma^2 = 0.05$. La Fig. 7.3 montre la sortie observée $y(t)$ sur quelques périodes.

L'objectif est de déterminer F_V^+ ou F_V^- lesquelles, dans notre cas, s'écrivent:

$$F_V^+(u) = \text{sign}(u)(\text{abs}(0.5455u) + 0.5),$$

$$F_V^-(u) = -\text{sign}(u)(\text{abs}(0.5455u) + 0.5)$$

Pour y arriver la sortie moyennée $y(t, N)$ est générée selon (7.37), avec $N = 100$, et utilisée pour

construire les courbes $(\cos(\omega t - \psi), y(t, N))$ pour différentes valeurs de ψ . Les courbes obtenues, pour deux valeurs de ψ , sont représentées par la Fig. 7.4. On voit qu'une courbe statique est obtenue pour $\psi^* = 1.965 \text{ rad}$. La forme de la courbe statique montre qu'elle correspond à F_V^+ .

Si on opte pour une interpolation polynomiale, alors l'estimé qu'on peut proposer pour $F(\cdot)$, s'écrit

$$\hat{F}(u) = \sum_{i=1}^p c_i u^i, \text{ pour un entier } p. \text{ La discontinuité autour de zéro nous impose de prendre } p$$

suffisamment grand. On détermine alors les paramètres c_i qui minimise l'erreur entre $\hat{F}(u)$ et la caractéristique $(\cos(\omega t - \psi), y(t, N))$ au sens des moindres carrés sur p points choisis (en utilisant l'expression de Vandermonde). L'estimé obtenu, pour $p=50$ est représenté par la Fig. 7.5. Vu la forme de l'estimé obtenu une interpolation linéaire aura été un meilleur choix qu'une interpolation polynomiale. Toutefois notons la qualité de l'estimé obtenu malgré la présence de perturbations.

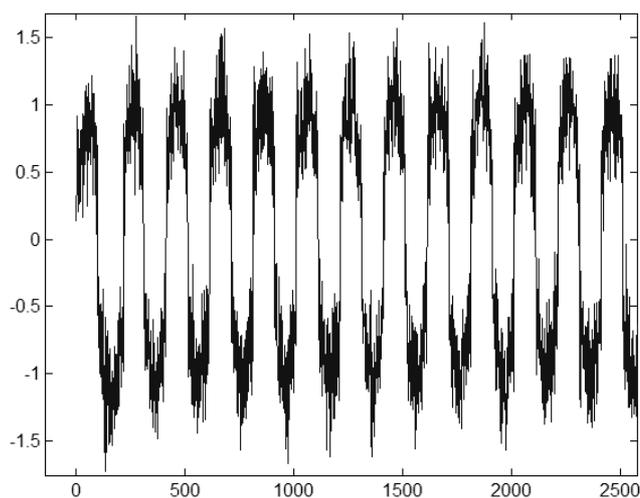


Fig. 7.3 La sortie observée sur quelques périodes

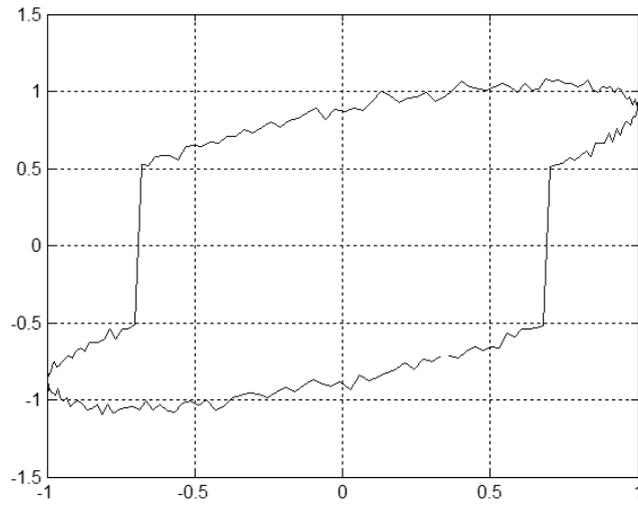


Fig. 7.4.a Courbe $C_{\psi,N}$ pour $\psi = 1,2 \text{ rad}$

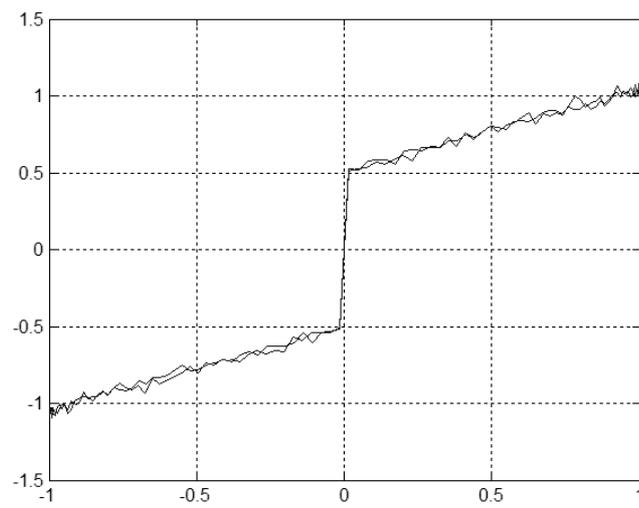


Fig. 7.4.b Courbe $C_{\psi,N}$ pour $\psi^* = 1.965 \text{ rad}$

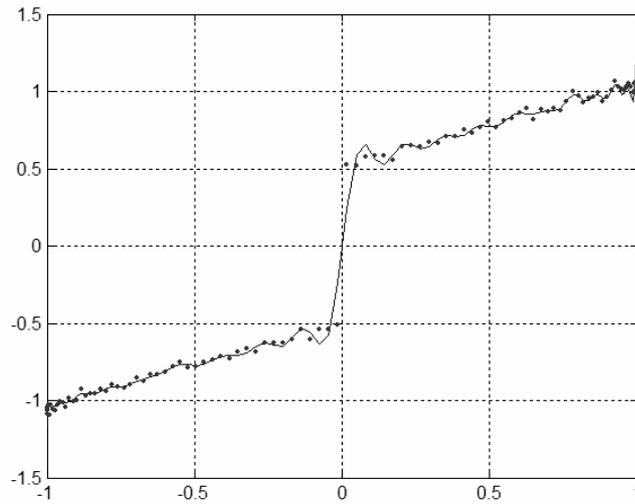


Fig. 7.5. les points du lieu $C_{\psi^*, N}$ et l'estimée polynomiale \hat{F} pour $p=50$

7.7. VERSION ANALYTIQUE DE LA METHODE D'IDENTIFICATION DE L'ELEMENT NON LINEAIRE

7.7.1. Reformulation analytique de la MIG

Dans ce paragraphe nous allons donner une reformulation plus rigoureuse de la méthode graphique présentée dans le paragraphe 6. Le point de départ sera l'évaluation analytique de l'aire géométrique $A(\psi, N)$ de $C_{\psi, N}$. En utilisant la Proposition 7.3 et l'équation (7.17), on peut alors écrire:

$$\begin{aligned} A(\psi, N) &= \int_0^{NT} \left| y\left(T - t + \frac{\psi}{\omega}, N\right) - y\left(t + \frac{\psi}{\omega}, N\right) \right| \left| \frac{du_\psi}{dt} \right| dt \\ &= \int_0^{NT} \left| y\left(T - t + \frac{\psi}{\omega}, N\right) - y\left(t + \frac{\psi}{\omega}, N\right) \right| |\sin(\omega t - \psi)| dt \end{aligned} \quad (7.43)$$

La proposition 7.6 stipule que, si pour un entier N assez grand, $A(\psi, N) \approx 0$ alors le mappage $\cos(\omega t - \psi) \rightarrow y(t, N)$ constitue un estimé de F^+ ou de F^- . Plus $A(\psi, N)$ est faible, plus la qualité de l'estimé est meilleure. Le problème devient alors un problème de minimisation du terme non linéaire $A(\psi, N)$, fortement non linéaire, en fonction de ψ . Toutefois, comme c'est un

problème à une seule dimension et le domaine de recherche de l'optimum est bien connu et limité, c'est-à-dire $[0, \pi]$, la résolution de ce problème peut être faite aisément, de manière graphique par exemple en traçant par programme l'évolution de la quantité $A(\psi, N)$ en fonction de ψ , ou par une recherche dichotomique plus rapide. Ceci nous amène à la méthode d'identification analytique suivante (MIA):

MIA1-MIA2. Identiques aux MIG1- MIG2

MIA3. Minimiser $A(\psi, N)$ en fonction de ψ sur l'intervalle $[0, \pi]$. Ceci peut se faire de manière graphique en traçant la courbe $\psi \rightarrow A(\psi, N)$. Soit ψ^* le minimum recherché.

MIA4-MIA6. Identiques aux MIG4-MIG6

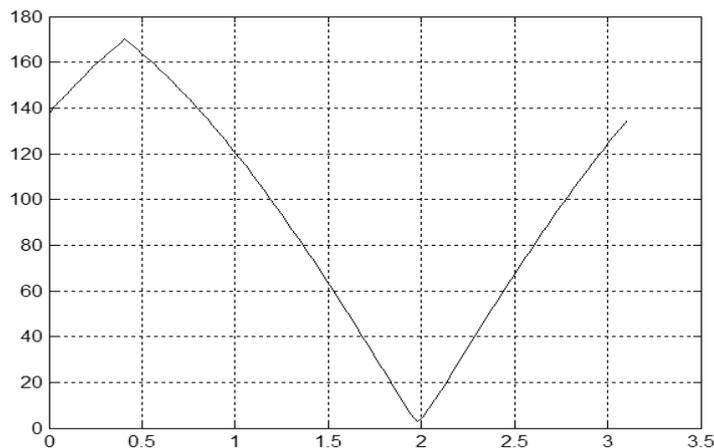
7.7.2. Evaluation de la MIA par simulation

Les conditions de simulation sont identiques à celles du paragraphe 6.2. La représentation obtenue de $A(\psi, N)$ est indiquée par la Fig. 7.6; à vrai dire pour annuler il suffit d'annuler $A(\psi, N)$ la quantité

$$Q(\psi, N) = \sum_{k=0}^M \left| y\left(T - kT_s + \frac{\psi}{\omega}, N\right) - y\left(kT_s + \frac{\psi}{\omega}, N\right) \right|$$

Avec $T_s / T = 1 / M$

On y voit clairement que $A(\psi, N)$ a un minimum global $\psi^* = 1.98rad$ dans l'intervalle $[0, \pi]$, très proche de la valeur obtenue avec la méthode MIG (paragraphe 6.2).

Figure 7.6. Tracé de $A(\psi, N)$ pour $0 \leq \psi < \pi$

Ces résultats de simulation confirment alors l'efficacité de la méthode, déjà prouvée théoriquement.

7.8. CONCLUSIONS

Dans ce chapitre nous avons développé une nouvelle approche, déterministe, non paramétrique permettant d'identifier la caractéristique du gain non linéaire d'un système de Wiener. Cette approche n'exige pas que cette caractéristique non linéaire soit inversible, continue, lisse ou polynomiale comme on l'a fait dans plusieurs travaux précédents. Elle est donc valable pour une large classe de non linéarité que l'on rencontre en pratique. L'approche permet de déterminer l'estimé de $F(x)$ pour tout x donné. Si $F(\cdot)$ est de structure connue et caractérisée par un certain nombre de paramètres, il est toujours possible de retrouver ces paramètres à partir du mappage $(x, F(x))$ estimé. Sinon, en procédant par interpolation (linéaire, par des splines, ou polynomiale) on peut trouver une expression mathématique pour estimer $F(\cdot)$.

Dans sa mise en œuvre, l'approche proposée ne nécessite qu'une seule expérience avec une excitation sinusoïdale (déterministe donc) avec une fréquence et une amplitude fixes. Ceci n'est pas le cas pour les autres approches fréquentielles qui nécessitent plusieurs expériences avec des fréquences différentes et/ ou des amplitudes différentes. L'analyse théorique faite dans la section 4 montre que les courbes de Lissajous C_ψ sont non statiques sauf lorsque $\psi = \varphi + k\pi$, c'est-à-dire lorsque ψ recouvre le déphasage inconnu φ (à $k\pi$ près). Ainsi la caractéristique non linéaire, ($F^+(\cdot)$ ou $F^-(\cdot)$) du système mis à l'échelle peut être retrouvée à partir de la courbe statique C_ψ avec $\psi = \varphi + k\pi$, pour tout k . L'identification du sous-système linéaire, peut être assez aisée si $F(\cdot)$

est inversible, mais un peu plus délicate si $F(.)$ n'est pas inversible (certains travaux précédents ont déjà abordé la question avec des hypothèses plus ou moins simplificatrices). Nous sommes entrain de développer une nouvelle approche pour identifier le sous système linéaire.

CHAPITRE 8

CONCLUSIONS GENERALES

L'identification des systèmes non linéaires est l'un des sujets de recherche d'actualité. En effet, d'une part l'identification joue un rôle capital dans la conduite des procédés industriels, et d'autre part les non linéarités des systèmes constituent une réalité à laquelle on s'intéresse de plus en plus afin de garantir une commande ayant des performances meilleures. Ce travail de recherche s'inscrit dans cette perspective. Nous nous sommes intéressés à l'identification des systèmes modélisables par un modèle d'Hammerstein ou par le modèle de Wiener. Certes le thème n'est pas nouveau, car dans la littérature, on trouve beaucoup de travaux dédiés à ce sujet. La plupart de ces travaux proposent des méthodes d'identification en se fixant des hypothèses simplificatrices, ou font appel à des calculs très complexes pour pouvoir être facilement mise en œuvre, ou mettent en jeu des signaux difficilement réalisables. Et en général, ces méthodes ne s'appliquent pas à une large gamme de systèmes. Notre objectif principal était de trouver de nouvelles méthodes d'identification qui:

- n'imposent pas des hypothèses simplificatrices, afin d'élargir leur champ d'application.
- ne nécessitent pas de calculs complexes, afin de faciliter leur mise en œuvre.
- opèrent, de préférence, dans un contexte déterministe avec des signaux facilement réalisables.
- ne supposent pas de connaissances a priori sur la nature des perturbations.

L'un des problèmes, le plus rencontré dans l'identification des systèmes à l'aide des méthodes paramétriques, est l'excitation persistante du vecteur de régression. Ce type d'excitation

permet, idéalement, une convergence des estimés vers leurs vraies valeurs. Ainsi, dans le chapitre 2 nous avons montré qu'à l'aide de séquences d'entrée impulsionnelles, on peut atteindre pour des systèmes linéaires des objectifs d'excitation persistante pour l'identification, et des objectifs relatifs aux performances de la commande des systèmes linéaires. Par rapport aux excitations classiques type bruit blanc qui ont été utilisées dans beaucoup de travaux, les séquences impulsionnelles présentent l'avantage d'être déterministes, facilement réalisables et peuvent même être périodiques. Elles utilisent le concept d'excitation par intervalles de temps. Dans la première partie de ce mémoire, on a traité l'identification sur la base du modèle de Hammerstein dans un contexte le plus général possible (pour la classe de non linéarité considérée et pour la nature des perturbations). Dans le chapitre 3 nous n'avons imposé aucune condition sur la non linéarité, sinon le fait qu'elle soit mesurable (au sens Borel/Lebesgue). Les travaux précédents qui ont fait de même sont rares. Le schéma d'identification proposé, permet d'estimer en ligne, les paramètres de la dynamique linéaire et N points sur la caractéristique non linéaire. Ces N points, sont arbitrairement choisis par l'utilisateur. On peut qualifier cette approche de semi paramétrique. A l'inverse des travaux précédents nous n'approximons pas la non linéarité par un polynôme; nous la modélisons plutôt par un polynôme qui passe par les N points choisis. Ceci n'engendre aucune erreur de modélisation que si les valeurs de l'entrée se limitent aux abscisses des N points choisis. Tout l'enjeu alors est de trouver une séquence qui soit excitante persistante et qui ne soit composée que des N valeurs choisies. Pour cela nous avons utilisé le principe d'excitation par intervalles et nous sommes appuyés sur les résultats du chapitre 2. Ces résultats ont été adaptés au cas des systèmes non linéaires modélisables par le modèle de Hammerstein. La consistance a été prouvée dans le cas d'absence de perturbations. Par ailleurs, la seule hypothèse faite à propos des perturbations est qu'elles soient petites en moyenne. Aucune hypothèse sur leur nature n'a été formulée (bruit blanc, distribution gaussienne ...). Ceci nous rapproche un peu plus de la réalité. Dans ce contexte, nous avons prouvé la convergence des estimés vers un voisinage de leurs vraies valeurs. La taille de ce voisinage est d'autant plus réduite que les perturbations sont petites en moyenne.

Le chapitre 4, étend le résultat obtenu dans le chapitre 3, pour l'appliquer aux deux cas particuliers: non linéarités de type polynomial (cas trivial) et non linéarités affines par morceaux (zone morte avec seuil). L'idée de base réside dans le bon choix des N points à utiliser lors du processus d'excitation. Les non linéarités de type zone morte avec seuil n'ont pas reçu beaucoup d'intérêt dans les travaux précédents. Paradoxalement, ce type de non

linéarité est très présent dans les procédés industriels, et il est de plus en plus utilisé dans le contrôle des systèmes. Certains contrôleurs intègrent des zones mortes. Les schémas d'identification présentés dans le chapitre 4, s'appuient, en partie, sur la même analyse de la consistance développée dans le chapitre 3. Ils permettent d'estimer, en ligne, les paramètres de la dynamique linéaire et les paramètres caractérisant la non linéarité. A l'inverse des travaux précédents, ils considèrent le cas d'une non linéarité symétrique ou asymétrique. Le concept qui a été utilisé dans ce chapitre est facilement extensible à d'autres types de non linéarité statique (saturation, zone morte et saturation ...), avec un minimum d'effort d'adaptation.

Dans le chapitre 5, nous traitons un cas qui a rarement été considéré dans les travaux précédents. En effet, presque tous ces travaux supposent que la non linéarité est statique. Mais, il se trouve que certains systèmes présentent des non linéarités avec un effet d'hystérésis. Dans ce chapitre, nous avons considéré le cas d'une non linéarité de type hystérésis ou hystérésis-relais. Là encore, on considère le cas général d'une non linéarité symétrique ou asymétrique. Les schémas d'identification proposés opèrent en deux étapes, et utilisent deux types de séquence d'excitation. Une séquence impulsionnelle périodique pour estimer les paramètres de la dynamique linéaire. Une séquence triangulaire pour estimer les paramètres de la non linéarité, en exploitant ceux déjà estimés précédemment. Ces deux séquences sont déterministes et facilement réalisables. A notre connaissance, c'est la première fois qu'on arrive à de tels résultats pour ce type de systèmes. Des travaux sont en cours pour étendre cette approche à d'autres types de non linéarité de nature hystérétique.

Dans les chapitres 6 et 7 on s'intéresse au problème de l'identification des systèmes non linéaires sur la base du modèle de Wiener. Plus exactement on s'intéresse seulement à l'identification de la non linéarité du modèle. En effet, comme dans certains travaux précédents, l'identification de la non linéarité du modèle peut être découplée de celle de la dynamique linéaire. L'identification de cette dernière est un problème assez aisé à résoudre, si la non linéarité est estimée et elle est inversible, sinon le problème devient un peu plus compliqué. Les schémas d'identification que nous avons proposés, permettent de retracer, point par point, un estimé de la non linéarité. Les approches sont donc de type non paramétrique.

Dans le chapitre 6, aucune hypothèse n'est formulée concernant le type de la non linéarité statique. Celle-ci peut être quelconque (mais mesurable au sens Borel/Lebesgue). L'idée de base est d'utiliser deux (ou plusieurs selon le besoin) excitations sinusoïdales avec différente

amplitudes et de reconstruire le signal interne (non mesurable), et ainsi tracer un estimé de la non linéarité en utilisant une version filtrée de la sortie mesurable. Cette reconstruction est basée sur une idée assez simple qui consiste en la localisation, sur une période, d'un extremum du signal interne à partir des extremums du signal de sortie. Pour présenter l'idée de manière simple nous avons traité d'abord le cas d'une non linéarité strictement monotone. Puis nous avons généralisé le principe à une non linéarité quelconque. Ainsi, et à l'inverse de beaucoup de travaux précédents, notre approche reste applicable pour tout type de non linéarité. Ceci, en plus de la simplicité de sa mise en œuvre, constitue un avantage de taille, par rapport à ce qui a été fait jusqu'à maintenant.

Dans le chapitre 7, nous avons utilisé le principe des courbes de Lissajous pour pouvoir estimer la non linéarité du modèle de Wiener. Ces courbes sont bien connues dans le cadre d'un système linéaire : la forme d'une courbe de Lissajous dépend du déphasage du signal sinusoïdal d'entrée par rapport au signal sinusoïdal de sortie.

L'approche nécessite une seule excitation de type sinusoïdal, afin de garantir que le signal d'entrée de l'élément non linéaire est sinusoïdal. Le tracé de la sortie filtrée en fonction d'un signal sinusoïdal de phase variable donne des courbes de Lissajous. Sous certaines conditions, on a démontré que lorsque la courbe de Lissajous est statique, elle représente un estimé de la non linéarité du modèle de Wiener. Cette approche est simple à mettre en œuvre, puisqu'elle ne nécessite qu'une seule expérience et une excitation de type sinusoïdal. Elle est valable pour une large gamme de non linéarités, rencontrées en pratique.

Dans les chapitres 6 et 7, et comme dans plusieurs travaux précédents, les perturbations ont été supposées générées par un processus stochastique ergodique. Cette hypothèse a permis d'obtenir une version filtrée de la sortie. Si cette hypothèse n'est pas valable, les approches du chapitre 6 et 7 peuvent donner une esquisse de l'allure de la non linéarité, sous l'hypothèse moins forte de perturbations relativement faibles. Ceci peut alors servir pour obtenir des connaissances préalables sur la forme de la non linéarité. Partant de ces connaissances, il faut chercher des techniques qui permettent d'estimer la non linéarité avec plus de précision. Ceci est un problème de traitement de l'effet des perturbations, valables pour toutes les approches et pas seulement à la notre.

En résumé, dans cette thèse, nous avons développé un certain nombre de schémas d'identification basé sur le modèle de Hammerstein ou de Wiener. Dans un contexte idéal (perturbations nulles), les algorithmes proposés sont toujours consistants et s'appliquent à une

large classes de systèmes non linéaires. En général, dans la littérature, la présence de perturbations rend le problème d'identification plus compliqué, surtout si on ne fait pas d'hypothèses simplificatrices sur la nature de ces perturbations. En présence de telles perturbations, les algorithmes deviennent moins consistants. Ceci reste un problème assez général en identification et commande. Pour le modèle de Hammerstein les approches proposées assurent une convergence des paramètres vers un voisinage de leurs vraies valeurs en présence de perturbations faibles en moyenne. Pour le modèle de Wiener les approches supposent l'érgodicité des perturbations et garantissent une bonne qualité des estimés obtenus.

Les systèmes en blocs ne se limitent pas au modèle de Hammerstein et Wiener. Comme extension à ce travail, nous sommes en train d'étudier la possibilité d'application des résultats obtenus pour d'autres modèles en blocs tels le modèle de Hammerstein-Wiener (N-L-N) ou Wiener-Hammerstein (L-N-L). La tâche n'est pas facile, mais le défi est tentant ...

ANNEXES

ANNEXE A

COMPLEMENTS DU CHAPITRE 2

ANNEXE A.1: PREUVE DU LEMME 2.2

La preuve contient quatre étapes. D'abord, une forme contrôlable du système est établie. Puis, une représentation dans l'espace d'état en boucle fermée est créée. Dans la troisième étape on montre que le vecteur d'état impliqué vérifie la propriété d'excitation. Ce ci est utilisé pour établir la propriété (4a) pour le vecteur d'observation $\phi(t)$.

1. Forme différentielle contrôlable du système.

Puisque $A(q^{-1})$ et $B(q^{-1})$ sont premiers entre eux, par l'hypothèse H2.2, il existe une paire unique de polynômes $P(q^{-1})$, $Q(q^{-1})$ d'ordre $n-1$, telle que:

$$A(q^{-1})P(q^{-1}) + B(q^{-1})Q(q^{-1}) = 1 \quad (\text{A1.1a})$$

Introduisons l'état interne:

$$\zeta(t) = P(q^{-1})u(t) + Q(q^{-1})x(t) \quad (\text{A1.1b})$$

où $x(t)$ représente la sortie non perturbée:

$$x(t) = \frac{B(q^{-1})}{A(q^{-1})}u(t) \quad (\text{A1.2})$$

Donc, en opérant $A(q^{-1})$ sur (A1.1b) et en utilisant (A1.1a) et (A1.2) on obtient:

$$A(q^{-1})\zeta(t) = u(t) \quad (\text{A1.3a})$$

De la même façon, en opérant $B(q^{-1})$ sur (A1.1b) et en utilisant (A1.1a) et (A1.2) on obtient:

$$x(t) = B(q^{-1})\zeta(t)$$

Laquelle avec (A1.2) et (2.3a), donne:

$$y(t) = B(q^{-1})\zeta(t) + \eta_p(t) \quad (\text{A1.3b})$$

Les équations (A1.3a-b) définissent la forme différentielle contrôlable du système.

2. Représentation de la boucle fermée dans l'espace d'état.

En substituant (A1.3a-b) dans (2.6a) on obtient pour tout k et tout $t \in [4nk, 4n(k+1)[$:

$$\left[R(4nk, q^{-1})A(q^{-1}) + S(4nk, q^{-1})B(q^{-1}) \right] \zeta(t) = f(t) + \delta_I(t) \quad (\text{A1.4a})$$

$$\delta_I(t) = \eta_c(t) - S(4nk, q^{-1})\eta_p(t) \quad (\text{A1.4b})$$

Définissons le vecteur d'état suivant:

$$X(t) = [\zeta(t) \cdots \zeta(t-2n+1)]^T \in \mathbb{R}^{2n} \quad (\text{A1.5})$$

L'équation (A1.4a) implique alors que, pour tout k et tout $t \in [4nk, 4n(k+1)[$:

$$X(t) = \mathbf{A}(k)X(t-1) + \mathbf{B}(f(t) + \delta_I(t)) \quad (\text{A1.6})$$

Avec:

$$\mathbf{A}(k) = \begin{bmatrix} -\alpha_1(k) & -\alpha_2(k) & \cdots & \cdots & -\alpha_{2n-1}(k) & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots & \vdots \\ \vdots & 0 & \ddots & 0 & 0 & 0 \\ \vdots & \cdots & \ddots & 1 & 0 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 0 \end{bmatrix}; \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad (\text{A1.7})$$

Où les $\alpha_i(k)$'s sont définis comme suit:

$$R(4nk, q^{-1})A(q^{-1}) + S(4nk, q^{-1})B(q^{-1}) = I + \alpha_1(k)q^{-1} + \cdots + \alpha_{2n-1}(k)q^{-2n+1} \quad (\text{A1.8})$$

3. Propriété d'excitation du vecteur d'état $X(t)$.

Il découle aisément de (A1.7) que le polynôme caractéristique de $\mathbf{A}(k)$ est donné par:

$$P(z) = z^{2n} + \alpha_1(k)z^{2n-1} + \cdots + \alpha_{2n-1}(k)z \quad (\text{A1.9})$$

Maintenant, définissons le vecteur auxiliaire suivant, pour n'importe quel k et tout $t \in [4nk, 4nk+2n [$:

$$V(t) = X(t+2n) + \alpha_1(k)X(t+2n-1) + \cdots + \alpha_{2n-1}(k)X(t+1) \quad (\text{A1.10})$$

A partir de (A1.6) on obtient, pour tout (t, j) tels que $t+j \in [4nk, 4n(k+1)[$:

$$X(t+j) = \mathbf{A}(k)^j X(t) + \sum_{i=1}^j \mathbf{A}(k)^{j-i} \mathbf{B}(f(t+i) + \delta_I(t+i)) \quad (\text{A1.11})$$

En utilisant (A1.11) il s'en suit à partir de (A1.10) que, pour tout $t \in [4nk, 4nk+2n[$:

$$\begin{aligned} V(t) &= \mathbf{A}(k)^{2n} X(t) + \sum_{i=1}^{2n} \mathbf{A}(k)^{2n-i} \mathbf{B}(f(t+i) + \delta_I(t+i)) \\ &+ \alpha_1(k) \left[\mathbf{A}(k)^{2n-1} X(t) + \sum_{i=1}^{2n-1} \mathbf{A}(k)^{2n-i-1} \mathbf{B}(f(t+i) + \delta_I(t+i)) \right] \\ &\vdots \\ &+ \alpha_{2n-1}(k) [\mathbf{A}(k)X(t) + \mathbf{B}(f(t+1) + \delta_I(t+1))] \end{aligned} \quad (\text{A1.12})$$

En utilisant le Théorème de Cayley-Hamilton (une matrice est une racine de son polynôme caractéristique), il s'en suit à partir de (A1.12) que la contribution de la condition initiale (c'est à dire de $X(t)$) dans $V(t)$ est égale à zéro. Donc, pour tout $t \in [4nk, 4nk+2n[$:

$$\begin{aligned}
V(t) = & \left[A(k)^{2n-1} \mathbf{B} + \alpha_1(k) A(k)^{2n-2} \mathbf{B} + \dots + \alpha_{2n-1}(k) \mathbf{B}, \right. \\
& A(k)^{2n-2} \mathbf{B} + \alpha_1(k) A(k)^{2n-3} \mathbf{B} + \dots + \alpha_{2n-2}(k) \mathbf{B}, \\
& \vdots \\
& \left. A(k) \mathbf{B} + \alpha_1(k) \mathbf{B}, \quad \mathbf{B} \right] \left(\begin{array}{c} f(t+1) \\ f(t+2) \\ \vdots \\ f(t+n) \end{array} + \begin{array}{c} \delta_1(t+1) \\ \delta_1(t+2) \\ \vdots \\ \delta_1(t+n) \end{array} \right) \quad (A1.13a)
\end{aligned}$$

Qui peut être représenté sous forme condensée:

$$V(t) = G(k)(F(t+1) + \Delta_1(\eta(t+1), \theta(k))) \quad (A1.13b)$$

Avec

$$\eta(t) = [\eta_p(t), \eta_c(t)]^T \quad (A1.13c)$$

et les définitions évidentes de $G(k)$, $F(t+1)$ et $\Delta_1(\eta(t+1), \theta(k))$. Il apparaît clairement à partir de (A1.4b) que $\Delta_1(\cdot, \cdot)$ est une fonction continue, par rapport à $\eta(t)$, qui vérifie:

$$\Delta_1(0, \theta) = 0, \text{ for all } \theta \quad (A1.13d)$$

D'autre part, à partir de (A1.7) on obtient les égalités suivantes:

$$A(k) \mathbf{B} + \alpha_1(k) \mathbf{B} = \begin{bmatrix} -\alpha_1 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \alpha_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (A1.14a)$$

$$\begin{aligned}
A(k)^2 \mathbf{B} + \alpha_1(k) A(k) \mathbf{B} + \alpha_2(k) \mathbf{B} &= A(k)(A(k) \mathbf{B} + \alpha_1(k) \mathbf{B}) + \alpha_2(k) \mathbf{B} \\
&= A(k) \begin{bmatrix} 0 \\ I \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \alpha_2(k) \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I \\ \vdots \\ 0 \end{bmatrix} \quad (A1.14b)
\end{aligned}$$

$$A(k)^{2n-1} \mathbf{B} + \alpha_1(k) A(k)^{2n-2} \mathbf{B} + \dots + \alpha_{2n-1}(k) \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ I \end{bmatrix} \quad (A1.14c)$$

A partir des $2n$ égalités précédentes on obtient en particulier:

$$G(k) = G = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ \vdots & & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & & \vdots \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (\text{A1.14d})$$

et $\det G = 1$. Maintenant, à partir de (A1.13b), il s'en suit:

$$\begin{aligned} G \sum_{i=0}^{2n-1} F(4nk+i+1)F(4nk+i+1)^T G^T \\ = \sum_{i=0}^{2n-1} (V(4nk+i) - G\Delta_1(\eta(4nk+i+1), \theta(k)))(V(4nk+i) - G\Delta_1(\eta(4nk+i+1), \theta(k)))^T \\ \leq 2 \sum_{i=0}^{2n-1} V(4nk+i)V(4nk+i)^T + G\Delta_1(\eta(4nk+i+1), \theta(k))\Delta_1(\eta(4nk+i+1), \theta(k))^T G^T \end{aligned} \quad (\text{A1.15a})$$

en utilisant l'inégalité $[X+Y][X+Y]^T \leq 2[XX^T + YY^T]$. Soit:

$$\Delta_2(\eta(k), \theta(k)) = 2 \sum_{i=0}^{2n-1} G\Delta_1(\eta(4nk+i+1), \theta(k))\Delta_1(\eta(4nk+i+1), \theta(k))^T G \quad (\text{A1.15b})$$

En utilisant (A1.13c), il s'en suit à partir de (A1.15b) que la fonction $\Delta_2(\eta, \theta)$ est continue par rapport à η et vérifie:

$$\Delta_2(0, \theta, g) = 0, \quad \text{pour tout } (\theta, g) \quad (\text{A1.15c})$$

En plus, à partir de (2.4b) on a:

$$\begin{aligned} \sum_{i=0}^{2n-1} F(4nk+i+1)F(4nk+i+1)^T &= \begin{bmatrix} 0 \\ 0 \\ \vdots \\ g(k) \end{bmatrix} [0 \ 0 \ \cdots \ g(k)] + \\ &+ \cdots + \begin{bmatrix} g(k) \\ 0 \\ \vdots \\ 0 \end{bmatrix} [g(k) \ 0 \ \cdots \ 0] \\ &= (g(k))^2 I_{2n} \end{aligned} \quad (\text{A1.16})$$

où I_{2n} représente la matrice identité de dimension $2n \times 2n$.

En combinant (A1.14a), (A1.15a) et (A1.16) on obtient:

$$\sum_{i=0}^{2n-1} V(4nk+i)V(4nk+i)^T \geq \frac{g(k)^2}{2} GG^T - \Delta_2(\eta(k), \theta(k)) = \frac{g(k)^2}{2} I_{2n} - \Delta_2(\eta(k), \theta(k))$$

(A1.17)

Maintenant, soit

$$Y(t) = [V(t)V(t+1) \cdots V(t+2n-1)] \quad (A1.18)$$

Alors, à partir de (A1.17) et pour tout k on a:

$$Y(4nk)Y(4nk)^T \geq \frac{g(k)^2}{2} I_{2n} - \Delta_2(\eta(k), \theta(k)) \quad (A1.19)$$

D'autre part, on obtient à partir de (A1.10) et (A1.18) que:

$$\begin{aligned} Y(4nk) &= [X(4nk+1) \cdots X(4nk+4n-1)]M(k) \\ &= W(k)M(k) \end{aligned} \quad (A1.20)$$

où $W(\cdot)$ est évident et $M(\cdot)$ est la matrice de dimension $(4n-1) \times (2n)$ définie comme suit :

$$M(k) = \begin{bmatrix} \alpha_{2n-1} & 0 & \cdots & 0 \\ \alpha_{2n-2} & \alpha_{2n-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha_1 & \cdots & \cdots & \alpha_{2n-1} \\ 1 & \alpha_1 & \cdots & \alpha_{2n-2} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \alpha_1 \\ 0 & \cdots & 0 & 1 \end{bmatrix} \quad (A1.21)$$

À partir de (A1.20) on a pour tout k :

$$Y(4nk)Y(4nk)^T = W(k)M(k)M(k)^T W(k)^T \leq \beta_1 W(k)W(k)^T \quad (A1.22)$$

avec:

$$\beta_1 = \sup_k \lambda_{\max}(M(k)M(k)^T)$$

où $\lambda_{\max}(\cdot)$ représente la plus grande valeur propre. Notez que β_1 est un nombre réel fini car $M(k)$ est bornée. En combinant (A1.19), (A1.20) et (A1.22) il s'en suit, pour tout k :

$$\sum_{i=0}^{4n-1} X(4nk+i)X(4nk+i)^T \geq \frac{(g(k))^2 I_{2n} - 2\Delta_2(\eta(k), \theta(k), g(k))}{2\beta_1} \quad (A1.23)$$

Propriété d'excitation du vecteur d'état $\varphi(t)$.

En substituant (A1.3a-b) dans (2.5a) et en utilisant (A1.5) on obtient:

$$\begin{aligned}
\varphi(t) &= \begin{bmatrix} 0 & -b_1 & -b_2 & & & -b_n & 0 & \cdots & \cdots & 0 \\ 0 & 0 & -b_1 & -b_2 & & & -b_n & 0 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & & \ddots & 0 \\ 0 & & \cdots & 0 & 0 & -b_1 & -b_2 & \cdots & \cdots & -b_n \\ 1 & a_1 & a_2 & a_3 & \cdots & a_n & 0 & 0 & \cdots & 0 \\ 0 & 1 & a_1 & a_2 & \ddots & & \ddots & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & & \ddots & \ddots & 0 \\ \vdots & \cdots & \ddots & 1 & a_1 & a_2 & a_3 & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 & a_1 & a_2 & \cdots & \cdots & a_n \end{bmatrix} X(t) + \begin{bmatrix} -\eta_p(t-1) \\ -\eta_p(t-2) \\ \vdots \\ -\eta_p(t-n+1) \\ -\eta_p(t-n) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \\
&= FX(t) + \delta_2(t) \tag{A1.24}
\end{aligned}$$

où F n'est autre que le transposé de la matrice Sylvester associé à $-B(q^{-1})$ et $A(q^{-1})$. Puisque ces derniers sont premiers entre eux alors F est non singulière.

A partir de (A1.23) et (A1.24) on a pour tout k :

$$\begin{aligned}
\sum_{i=0}^{4n-1} (\varphi(4nk+i) - \delta_2(4nk+i)) (\varphi(4nk+i) - \delta_2(4nk+i))^T &= F \sum_{i=0}^{4n-1} X(4nk+i) X(4nk+i)^T F^T \\
&\geq \beta_2 \sum_{i=0}^{4n-1} X(4nk+i) X(4nk+i)^T \\
&\geq \frac{\beta_2}{2\beta_1} [g(k)^2 I_{2n} - 2\Delta_2(\eta(k), \theta(k))] \tag{A1.25}
\end{aligned}$$

où $\beta_2 = \lambda_{\min}(FF^T)$ est un réel positif puisque F est non singulière. En appliquant une autre fois l'inégalité: $(X+Y)(X+Y)^T \leq 2XX^T + 2YY^T$ à la partie gauche de (A1.25) on obtient:

$$\begin{aligned}
\sum_{i=0}^{4n-1} \varphi(4nk+i) \varphi(4nk+i)^T + \sum_{i=0}^{4n-1} \delta_2(4nk+i) \delta_2(4nk+i)^T \\
\geq \frac{1}{2} \sum_{i=0}^{4n-1} (\varphi(4nk+i) - \delta_2(4nk+i)) (\varphi(4nk+i) - \delta_2(4nk+i))^T \\
\geq \frac{\beta_2}{4\beta_1} [g(k)^2 I_{2n} - 2\Delta_2(\eta(k), \theta(k))] \tag{A1.26}
\end{aligned}$$

Qui peut s'écrire sous la forme suivante:

$$\sum_{i=0}^{4n-1} \varphi(4nk+i) \varphi(4nk+i)^T \geq \frac{\beta_2}{4\beta_1} (g(k))^2 I_{2n} + \Delta(\eta(k), \theta(k)) \tag{A1.27}$$

avec

$$\Delta(\eta(k), \theta(k)) = \frac{\beta_2}{2\beta_1} \Delta_2(\eta(k), \theta(k)) - \sum_{i=0}^{4n-1} \delta_2(4nk+i) \delta_2(4nk+i)^T \quad (\text{A1.28})$$

Du point de vue de (A1.24) et (A1.15c), il s'en suit à partir de (A1.28) que la fonction $\Delta(\eta, \theta)$ est continue par rapport à η et vérifie $\Delta(0, \theta) = 0$ pour tout θ . Ce ci combine avec (A1.26) établit (2.6a) avec $\alpha = \beta_2/2\beta_1$ et complète la preuve du lemme technique \square

ANNEXE A.2: PREUVE DU LEMME 2.8

La matrice à droite de l'expression (2.41) peut être décomposée comme suit:

$$\begin{bmatrix} \sigma_l & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \sigma_l \\ 0 \\ \vdots \\ 0 \end{bmatrix} [I \ 0 \ \cdots \ 0] \quad (\text{A2.1})$$

Donc (2.23) devient:

$$M = \Gamma \begin{bmatrix} \sigma_l \\ 0 \\ \vdots \\ 0 \end{bmatrix} \cdot [I \ 0 \ \cdots \ 0] \Sigma \quad (\text{A2.2})$$

Ce ci suggère la décomposition suivante:

$$M = B'D'^T \quad \text{avec} \quad B' = \Gamma \begin{bmatrix} \sigma_l \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{et} \quad D'^T = [I \ 0 \ \cdots \ 0] \Sigma \quad (\text{A2.3})$$

Une telle décomposition ne satisfait la condition (2.22). Cette dernière suggère la solution modifiée suivante :

$$B = \frac{B'}{\|B'\|} \cdot \text{sign}(\rho(B')) \quad \text{et} \quad D = \|B'\| \cdot D' \cdot \text{sign}(\rho(B')) \quad (\text{A2.4})$$

□

ANNEXE B

COMPLEMENTS DU CHAPITRE 3

ANNEXE B.1 : PREUVE DE LA PROPOSITION 3.2

A partir de (3.18) la matrice $\hat{M}(t)$ peut être décomposée comme suit :

$$\hat{M}(t) = \Gamma(t) \cdot \begin{bmatrix} \sigma_1(t) & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \cdot \Sigma(t) + \Delta_0(t) \quad (\text{B1.1})$$

où :

$$\Delta_0(t) = \Gamma(t) \cdot \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \sigma_2(t) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n(t) \end{bmatrix} \cdot \Sigma(t) \quad (\text{B1.2})$$

Supposons que les $\hat{\mu}_{ij}(t)$ dans (3.17) convergent vers μ_{ij} de (3.10b). Ce ci assure qu'à partir de (3.12), que :

$$\lim_{t \rightarrow \infty} \hat{M}(t) = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} [d_1 \quad \cdots \quad d_N] \quad (\text{B1.3})$$

Ce ci implique que le rang de $\hat{M}(t)$ tend vers un. Ainsi les valeurs singulières $\sigma_2(t), \dots, \sigma_n(t)$ convergent vers zéro et par conséquent:

$$\lim_{t \rightarrow \infty} \Delta_0(t) = 0 \quad (\text{B1.4})$$

En plus à partir de (3.19a-b) on a :

$$\begin{bmatrix} \hat{b}_1(t) \\ \vdots \\ \hat{b}_n(t) \end{bmatrix} \begin{bmatrix} \hat{d}_1(t) & \dots & \hat{d}_N(t) \end{bmatrix} = \Gamma(t) \begin{bmatrix} \sigma_1(t) & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} \Sigma(t) \quad (\text{B1.5})$$

en combinant les deux relations précédentes on a:

$$\lim_{t \rightarrow \infty} \left(M(t) - \begin{bmatrix} \hat{b}_1(t) \\ \vdots \\ \hat{b}_n(t) \end{bmatrix} \begin{bmatrix} \hat{d}_1(t) & \dots & \hat{d}_N(t) \end{bmatrix} \right) = 0 \quad (\text{B1.6})$$

la proposition 2.1 stipule que la décomposition (2.29) est l'unique relation qui respecte la condition:

$$\sum_{i=1}^n \hat{b}_i(t)^2 = 1 \quad (\text{B1.7})$$

Ce ci implique :

$$\lim_{t \rightarrow \infty} \begin{bmatrix} \hat{b}_1(t) \\ \vdots \\ \hat{b}_n(t) \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \quad \text{et} \quad \lim_{t \rightarrow \infty} \begin{bmatrix} \hat{d}_1(t) & \dots & \hat{d}_N(t) \end{bmatrix} = \begin{bmatrix} d_1 & \dots & d_N \end{bmatrix} \quad \square$$

ANNEXE B.2 : PREUVE DE LA PROPOSITION 3.3

(3.22f) peut être écrite pour tout k et tout $t \in [t_k, t_k+4n-1]$ comme suit:

$$A(q^{-1})x(t) = B(q^{-1})u(t) \quad (\text{B2.1})$$

En plus, à partir (3.21) on a pour tout $t \in [t_k, t_k+4n-1]$:

$$v(t) = \begin{cases} V_1 & \text{pour } t = t_k + 2n \\ 0 & \text{par ailleurs} \end{cases} \quad (\text{B2.2})$$

Comme $v(t) \in \{0, V_1\}$ pour tout $t \in [t_k, t_k+4n-1]$, on a successivement:

$$u(t) = F(v(t)) = P(v(t)) \quad (\text{à partir de (3.6a-b)})$$

$$= v(t) \sum_{j=1}^N d_j P_j(v(t)) \quad (\text{à partir de (3.7)})$$

$$= v(t) d_1 P_1(v(t)) \quad (\text{à partir de (3.8)})$$

Donc, (B2.1) peut être réécrite de la manière suivante:

$$A(q^{-1})x(t) = B'(q^{-1})v(t)P_1(v(t)); \quad (\text{avec } B'(q^{-1}) = d_1 B'(q^{-1})) \quad (\text{B2.3a})$$

Aussi, à partir de (3.8) et (B2.2), on a:

$$v(t)P_1(v(t)) = \begin{cases} V_1 & \text{pour } t = t_k + 2n \\ 0 & \text{ailleurs} \end{cases} \quad (\text{B2.3b})$$

(B2.3a-b) nous permet d'appliquer le Lemme technique 2.1 du chapitre 2, et d'en déduire que le vecteur de régression associé c'est-à-dire

$$\phi_1^T(t) = [-x(t-1) \quad \dots \quad -x(t-n) \quad v(t-1)P_1(v(t-1)) \quad \dots \quad v(t-n)P_1(v(t-n))]$$

est excitant persistant sur l'intervalle $[t_k, t_k+4n[$. □

ANNEXE B.3 : PREUVE DE LA PROPOSITION 3.4

A partir de (3.21) que, pour tout k et $j=2, \dots, N$:

$$\phi_j(t_k + n(j+2) + \tau) \phi_j(t_k + n(j+2) + \tau)^T = \begin{bmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \dots & 0 & x_j^2 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \leftarrow \text{ligne } \tau$$

\uparrow colonne τ

(B3.1)

où $\tau=1, \dots, n$ et :

$$x_j = V_j P_j(V_j) = V_j \tag{B3.2}$$

Alors, (3.25) s'ensuit directement à partir de (B3.1).

ANNEXE B.4 : PREUVE DE LA PROPOSITION 3.5

Supposons que cette proposition n'est pas vraie. Donc pour toute séquence de nombre réels positifs $\{\varepsilon(i)\}$ telle que :

$$\lim_{i \rightarrow \infty} \varepsilon(i) = 0 \quad (\text{B4.1})$$

il existe une séquence de vecteurs unitaires $\{W(i)\}$ (c'est à dire $\|W(i)\| = 1$ pour tout i) et une sous séquence t_{k_i} telle que pour tout i ,

$$\max_{0 \leq \tau \leq n(3+N)} \left| \Phi'(t_{k_i} + \tau)^T W(i) \right| \leq \varepsilon(i) \quad (\text{B4.2})$$

En correspondance avec (3.22a) , $W(i)$ peut être décomposé comme suit:

$$W(i) = [w_1(i)^T \ w_2(i)^T \ \dots \ w_N(i)^T]^T \quad (\text{B4.3})$$

D'autre part, à partir de (3.21) et (3.22a) on obtient successivement, pour tout k et tout i :

$$\Phi'(t_k + \tau)^T W(i) = \phi_1'(t_k + \tau)^T w_1(i) \quad \text{pour } \tau=1, \dots, 4n \quad (\text{B4.4})$$

$$\Phi'(t_k + n(j+2) + \tau)^T W(i) = \phi_j'(t_k + n(j+2) + \tau)^T w_j(i) + \phi_j(t_k + n(j+3) + \tau)^T w_j(i)$$

$$\text{pour } j=2, \dots, N ; \tau=1, \dots, n \quad (\text{B4.5})$$

En utilisant la Proposition 3.3, on obtient à partir de A2.9 que, pour tout k et tout i ,

$$\|w_1(i)\|^2 < \frac{4n \left(\max_{1 \leq \tau \leq 4n} \left| \Phi'(t_k + \tau)^T W(i) \right| \right)^2}{\varepsilon_0}$$

laquelle, avec (B4.1) et (B4.2) impliquent que:

$$\lim_{i \rightarrow \infty} \|w_1(i)\|^2 = 0 \quad (\text{B4.6})$$

De manière similaire, en utilisant la Proposition 3.4, on obtient à partir de (B4.5) que pour tout k , et tout i et $j=2 \dots N$:

$$\|w_j(i)\|^2 < n \frac{\left(\max_{1 \leq \tau \leq n} \left\| \Phi'(t_k + n(j+2) + \tau)^T W(i) \right\| \right)^2 + \left(\max_{1 \leq \tau \leq n} \left\| \phi_j'(t_k + n(j+2) + \tau)^T w_j(i) \right\| \right)^2}{\varepsilon_j}$$

$$(\text{B4.7})$$

En utilisant (B4.2), (B4.6) et le fait que $\phi_j'(t)$ est borné, il s'ensuit à partir de (B4.7):

$$\lim_{i \rightarrow \infty} \|w_j(i)\| = 0 \quad \text{pour } j = 2, \dots, N \quad (\text{B4.8})$$

En combinant (B4.8) et (B4.6) on obtient à partir de (B4.3) que:

$$\lim_{i \rightarrow \infty} \|W(i)\| = 0$$

Ce ci contredit le fait que $W(i)$ est un vecteur unitaire. Donc, la Proposition 3.5 est vraie. \square

ANNEXE B.5 : PREUVE DE LA PROPOSITION 3.6

Pour établir la première partie de cette proposition, appliquons (3.26a) avec

$$W = \frac{\tilde{\Theta}(t_k)}{\|\tilde{\Theta}(t_k)\|} \quad \text{où} \quad \tilde{\Theta}(t_k) = \hat{\Theta}(t_k) - \Theta^* \quad (\text{B5.1})$$

c'est à dire:

$$\tilde{\Theta}(t_k)^T \left[\sum_{\tau=0}^{nN+3n} \Phi'(t_k + \tau) \Phi'(t_k + \tau)^T \right] \tilde{\Theta}(t_k) \geq \varepsilon^2 \|\tilde{\Theta}(t_k)\|^2, \quad \text{pour tout } k \quad (\text{B5.2})$$

qui peut s'écrire comme suit :

$$\varepsilon^2 \|\tilde{\Theta}(t_k)\|^2 \leq \sum_{\tau=0}^{nN+3n} (\Phi'(t_k + \tau)^T \tilde{\Theta}(t_k))^2, \quad \text{pour tout } k \quad (\text{B5.3})$$

En plus, en utilisant (3.11a) on a :

$$e_p(t) = y(t) - \Phi(t)^T \hat{\Theta}(t) = -\Phi(t)^T \tilde{\Theta}(t) + \eta(t) \quad (\text{B5.4})$$

D'autre part, on a successivement pour tout k :

$$\begin{aligned} \tilde{\Theta}(t_k)^T \Phi'(t_k + \tau) &= \tilde{\Theta}(t_k)^T \Phi(t_k + \tau) - \tilde{\Theta}(t_k)^T \tilde{\Phi}(t_k + \tau) && \text{(utilisant (2.23b))} \\ &= -e_p(t_k + \tau) + \eta(t_k + \tau) - (\tilde{\Theta}(t_k + \tau)^T - \tilde{\Theta}(t_k)^T) \Phi(t_k + \tau) - \tilde{\Theta}(t_k)^T \tilde{\Phi}(t_k + \tau) \end{aligned} \quad (\text{B5.5})$$

avec:

$$\eta(t) = A(q^{-1}) z(t)$$

où (B5.4) , (3.11a) ont été utilisées dans la dernière égalité. On tire à partir de la Proposition 3.1 (partie 1) que, pour tout entier $l \leq \tau \leq (N+3)n$:

$$\{ e_p(t_k + \tau)^2 \} \in S(K_4 \mu) \quad \text{et} \quad \left\{ \|\tilde{\Theta}(t_k + \tau) - \tilde{\Theta}(t_k)\|^2 \right\} \in S(K_4 \mu) \quad (\text{B5.6})$$

avec $K_4 = K_l(nN+3n)$. Puisque $\{\Phi(t)\}$ est bornée, il s'ensuit à partir de (B5.6) qu'il existe un réel K_5 , indépendant de μ , tel que pour tout entier $l \leq \tau \leq nN+3n$:

$$\left\{ \left| (\tilde{\Theta}(t_k + \tau) - \tilde{\Theta}(t_k))^T \Phi(t_k + \tau) \right|^2 \right\} \in S(K_5 \mu) \quad (\text{B5.7})$$

et puisque $\left\{ \|\tilde{\Theta}(t)\| \right\}$ est bornée (par $2\rho^*$), il s'ensuit directement à partir de (3.23a-b) et (3.15c) que, pour tout entier $\tau (l \leq \tau \leq nN+3n)$:

$$\left\{ \left| (\tilde{\Theta}(t_k))^T \tilde{\Phi}(t_k + \tau) \right|^2 \right\} \in S(K_6 \mu) \quad (\text{B5.8})$$

pour une constante réelle K_6 , indépendante de μ . Maintenant, en utilisant (B5.6) et (B5.8), on obtient à partir de (B5.5) que:

$$\left\{ \left| \left(\tilde{\Theta}(t_k) \right)^T \Phi'(t_k + \tau) \right|^2 \right\} \in S(K_7\mu), \quad \text{pour tout } \tau \ (1 \leq \tau \leq nN+3n) \quad (\text{B5.9})$$

pour une certaine constante réelle K_7 indépendante de μ . Aussi, il s'ensuit à partir de (B5.2) que :

$$\varepsilon^2 \left\| \tilde{\Theta}(t_k) \right\|^2 \leq \tilde{\Theta}(t_k)^T \left[\sum_{\tau=1}^{nN+3n} \Phi'(t_k + \tau) \Phi'(t_k + \tau)^T \right] \tilde{\Theta}(t_k) \leq \sum_{\tau=1}^{nN+3n} (\Phi'(t_k + \tau)^T \tilde{\Theta}(t_k))^2 \quad (\text{B5.10})$$

qui, grâce à (B5.9), implique que $\left\{ \left\| \tilde{\Theta}(t_k) \right\|^2 \right\} \in S(K_8\mu)$ avec $K_8 = (nN+3n)K_7$. Ainsi, en utilisant le fait que $\left\{ \left\| \hat{\Theta}(t) - \hat{\Theta}(t-1) \right\|^2 \right\} \in S(K_1\mu)$, on obtient que $\left\{ \left\| \tilde{\Theta}(t) \right\|^2 \right\} \in S(K_2\mu)$, pour une certaine constante K_2 qui dépend seulement de K_8 , n et N . Ce ci établit la première partie de la proposition 3.6. La preuve de la partie 2 est un cas particulier de la partie 1. \square

Annexe C

Comment on ‘Frequency domain identification of Wiener models’, by E.W. Bai, *Automatica* 39 (2003), 1521-1530

Abstract. A frequency-domain identification method is proposed by Bai (2003) for Wiener systems. The key component is a phase estimator that gives estimates of the linear subsystem phase for different pulsations. These estimates are then used to identify the whole system. In this note, it is shown that the phase estimator is *generally* inconsistent, making inconsistent the whole identification method.

1. INTRODUCTION AND PROBLEM STATEMENT⁽¹⁾

The paper commented-on considers the problem of identifying Wiener systems (Fig. 1). The objective is to get estimates $(\hat{G}(\cdot), \hat{f}(\cdot))$ of $(G(\cdot), f(\cdot))$, using only measurable signals $(u(t), y(t))$. In the method of Bai (2003), the inputs are sinusoids:

$$u(t) = A \cos(\omega_k t); \quad t \in \left[0, L \frac{2\pi}{\omega_k}\right]; \quad k = 1, 2, \dots, q \quad (1)$$

for some real numbers $A, L, \omega_k > 0$. The system is submitted to the following assumptions:

Assumptions 1 (Assumptions 2.1 in Bai (2003))

1. The linear part $G(s)$ is stable and $\sup_{\omega} |G(j\omega)| < \infty$.
2. The nonlinearity $f(x)$ is static and piecewise smooth and $v(t)$ is an additive white noise.

The system of Fig. 1 generates in response to the input (1) the following (steady-state) signals:

$$x(t) = A |G(j\omega_k)| \cos(\omega_k t + \angle G(j\omega_k)) \quad (2)$$

$$y(t) = f\left(A |G(j\omega_k)| \cos(\omega_k t + \angle G(j\omega_k))\right) + v(t) \quad (3)$$

System rescaling. Let $\omega_{\bar{k}}$ be any fixed pulsation (with $1 \leq \bar{k} \leq q$). Then, it can be assumed that

$$A |G(j\omega_{\bar{k}})| = 1, \text{ which in turn implies that:}$$

¹ The notations used throughout this note are generally those of Bai (2003). The rare exceptions are clearly pointed out.

$$x(t) = \cos(\omega_k^- t + \angle G(j\omega_k^-)) \quad (4)$$

$$y(t) = f\left(\cos(\omega_k^- t + \angle G(j\omega_k^-))\right) + v(t) \quad (5)$$

Indeed, if $A|G(j\omega_k^-)| \neq 1$ then one just has to introduce the following rescaling $G^*(s) = \frac{G(s)}{A|G(j\omega_k^-)|}$, $f^*(x) = f\left(A|G(j\omega_k^-)|x\right)$. The new couple $(G^*(.), f^*(.))$ also represents the system, just as $(G(.), f(.))$ does. At this point, note that the comments following Assumption 3.1 in (Bai, 2003) should be corrected substituting $f(\alpha x)$ for $\alpha f(x)$.

Coexistence of two system representations. Equations (4)-(5) do not define the system uniquely.

Indeed, the couple $(G^-(.), f^-(.))$ with:

$$G^-(s) \stackrel{def}{=} -G(s) \quad \text{and} \quad f^-(x) \stackrel{def}{=} f(-x) \quad (6)$$

also represents the system. Actually, when excited by a sinusoid $u(t) = A \cos(\omega_k t)$, both representations generate the same output, i.e.:

$$y(t) = f(\cos(\omega_k t + \angle G(j\omega_k))) + v(t) = f^-(\cos(\omega_k t + \angle G^-(j\omega_k))) + v(t) \quad (7)$$

the corresponding internal signals are respectively $x(t)$ and $x^-(t)$ with:

$$x^-(t) = \cos(\omega_k^- t + \angle G^-(j\omega_k^-)) = -\cos(\omega_k^- t + \angle G(j\omega_k^-)) = -x(t) \quad (8)$$

The identification problem consists in getting an estimate $(\hat{G}(.), \hat{f}(.))$ of $(G(.), f(.))$. It will not be well posed if $(G(.), f(.))$ is not clearly characterized and well distinguished from $(G^-(.), f^-(.))$. Such a discrimination issue has been largely ignored in Bai (2003). Though this issue is fixed hereafter (see Subsection 3.3), the Bai's phase estimator will still be inconsistent.

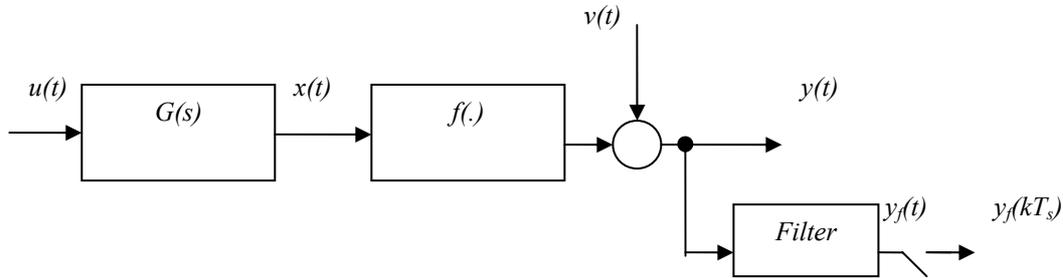


Figure 1. Wiener system structure

2. THE IDENTIFICATION METHOD PROPOSED BY BAI (2003)

2.1. Bai's identification method (an outline)

In Bai (2003), it is shown (see Lemma 2.1 and Section 3) that if $v_f(t)$ (the filtered noise) is a continuous time random signal that is the output of an unknown stable and proper linear system driven by a white noise of zero mean and finite variance, then there exists an integer \bar{i} such that for $1 \leq i \leq \bar{i}$:

$$Y_{f,DFT}(i\omega_k) \rightarrow \frac{r_i}{2} \cos(i\angle G(j\omega_k)) + j \frac{r_i}{2} \sin(i\angle G(j\omega_k)), \text{ in probability, as } L \rightarrow \infty \quad (9a)$$

The notations are those of Bai (2003). At this point, $(G(\cdot), f(\cdot))$ and $(G^-(\cdot), f^-(\cdot))$ play a symmetric role. Consequently, equation (9a) also applies to $(G^-(\cdot), f^-(\cdot))$, i.e.:

$$Y_{f,DFT}(i\omega_k) \rightarrow \frac{r_i^-}{2} \cos(i\angle G^-(j\omega_k)) + j \frac{r_i^-}{2} \sin(i\angle G^-(j\omega_k)), \text{ in probability, as } L \rightarrow \infty \quad (9b)$$

It can be easily checked that $r_i^- = (-1)^i r_i$. Motivated by (9a), Bai (2003) proposed the estimator:

$$\angle \hat{G}(j\omega_k) = \frac{1}{i} \tan^{-1} \left(\frac{\text{Im}(Y_{f,DFT}(i\omega_k))}{\text{Re}(Y_{f,DFT}(i\omega_k))} \right); \quad k = 1, 2, \dots, q \quad (10)$$

where i is any integer such that $Y_{f,DFT}(i\omega_k) \neq 0$. It readily follows from (9a-b)-(10) and (6) that:

$$\tan(i\angle \hat{G}(j\omega_k)) \rightarrow \tan(i\angle G(j\omega_k)), \text{ in probability, as } L \rightarrow \infty \quad (11a)$$

$$\angle G^-(j\omega) - \angle G(j\omega) = \pm\pi \quad \text{and} \quad \tan(i\angle G(j\omega)) = \tan(i\angle G^-(j\omega)) \quad (\text{for all } \omega) \quad (11b)$$

2.2. Definitions

Inverse tangent function. The estimator (10) involves the inverse tangent function $\tan^{-1}(\cdot)$. In the sequel, this refers to its *principal* value. Consequently, it follows from (10) that, for all k :

$$-\frac{\pi}{2} \leq i\angle \hat{G}(j\omega_k) < \frac{\pi}{2} \quad (12)$$

Consistency of the phase estimator. The estimator (10) is said to be consistent if, for some integer m : $\angle \hat{G}(j\omega_k) \rightarrow \angle G(j\omega_k) + 2m\pi$, in probability, as $L \rightarrow \infty$. Otherwise, it is inconsistent.

3. CONSISTENCY ANALYSIS OF BAI'S METHOD WHEN $i = 1$

3.1. Inconsistency of phase estimator (10) ⁽²⁾

In the case where $i=1$, (11a) implies that, for any ω_k , there is an integer p such:

$\angle \hat{G}(j\omega_k) \rightarrow \angle G(j\omega_k) + p\pi$ (in probability) as $L \rightarrow \infty$. Then, the phase estimator is consistent if:

² In the rest of the paper, all convergence results are intended in probability, as $L \rightarrow \infty$.

for any ω_k , the corresponding integer p is even ($p=2m$). As $\angle\hat{G}(j\omega_k) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, it follows that

phase estimator consistency is ensured if: for any ω_k , there is an integer m such that:

$\angle G(j\omega_k) \in \left]-\frac{\pi}{2} + 2m\pi, \frac{\pi}{2} + 2m\pi\right[$. Therefore, it is not surprising that for the simulation example

in Bai (2003), the author has chosen $G(s) = \frac{2s+1}{s^2+2s+1}$ that satisfies $\angle G(j\omega) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ (for

all ω) and a nonlinearity $f(\cdot)$ for which $i=1$.

Inversely, if $\angle G(j\omega_k) \in \bigcup_{m=-\infty}^{+\infty} \left(\left]-\pi + 2m\pi, -\frac{\pi}{2} + 2m\pi\right[\cup \left]\frac{\pi}{2} + 2m\pi, \pi + 2m\pi\right[\right)$ (for an ω_k) then

the phase estimator is inconsistent for that ω_k . In effect, since $\tan(\angle\hat{G}(j\omega_k)) \rightarrow \tan(\angle G(j\omega_k))$ (in

probability) as $L \rightarrow \infty$ and $\angle\hat{G}(j\omega_k) \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, then one has $\angle\hat{G}(j\omega_k) \rightarrow \angle G(j\omega_k) + \pi + 2m\pi$

(for some integer m). In fact, this means that $\angle\hat{G}(j\omega_k) \rightarrow \angle G^-(j\omega_k) + 2m\pi$ (for some integer m).

Now, let us show that the phase estimator inconsistency does not prevent from achieving a consistent estimation of the system nonlinear element $f(\cdot)$. To this end, let the system be excited by $u(t) = A \cos(\omega_{\bar{k}}t)$. As observed above, the estimate $\angle\hat{G}(j\omega_{\bar{k}})$ provided by (10) then converges (in probability) to either $\angle G(j\omega_{\bar{k}}) + 2m\pi$ or $\angle G^-(j\omega_{\bar{k}}) + 2m\pi$, as $L \rightarrow \infty$ (for some integer m).

Substituting $\angle\hat{G}(j\omega_{\bar{k}})$ to $\angle G(j\omega_{\bar{k}})$ in (4) one gets the internal signal

$\hat{x}(t) \stackrel{def}{=} \cos(\omega_{\bar{k}}t + \angle\hat{G}(j\omega_{\bar{k}}))$. It is clear that \hat{x} will uniformly converge either to x or to x^- . Letting

\hat{f} denote the nonlinearity determined by the couples $(\hat{x}(t), y(t))$, it follows that \hat{f} will in turn uniformly converge either to $f(\cdot)$ or to $f^-(\cdot)$. Both of $f(\cdot)$ and $f^-(\cdot)$ are convenient representations of the system nonlinearity. Therefore, inconsistency of the estimator (10) is not an issue as long as the nonlinearity is concerned. The main reason is that the nonlinearity estimation necessitates phase estimation at a single pulsation $\omega_{\bar{k}}$.

After \hat{f} has been obtained, the next step in Bai's identification method is to estimate the phase $\angle G(j\omega_k)$, for a number of ω_k 's, using (10). The point is that the estimate $\angle\hat{G}(j\omega_{\bar{k}})$ converges (in

probability) to $\angle G(j\omega_k)$ (modulo 2π) as $L \rightarrow \infty$, only when

$$\angle G(j\omega_k) \in \bigcup_{m=0}^{\infty} \left] -\frac{\pi}{2} + 2m\pi, \frac{\pi}{2} + 2m\pi \right[; \text{ otherwise it converges to } \angle G^-(j\omega_k) \text{ (modulo } 2\pi).$$

Consequently, if $G(s)$ is such that $\angle G(j\omega_k) \in \bigcup_{m=0}^{\infty} \left] -\frac{\pi}{2} + 2m\pi, \frac{\pi}{2} + 2m\pi \right[$ for a set of pulsations,

denoted Ω_1 , and $\angle G(j\omega_k) \notin \bigcup_{m=0}^{\infty} \left] -\frac{\pi}{2} + 2m\pi, \frac{\pi}{2} + 2m\pi \right[$ for an other set, say Ω_2 , then $\angle \hat{G}(j\omega_k)$

will converge (in probability) to $\angle G(j\omega_k)$ for all $\omega_k \in \Omega_1$ and to $\angle G^-(j\omega_k)$ for all $\omega_k \in \Omega_2$.

This is, for instance, the case when $G(s)$ has a relative degree higher than 2 (e.g. $G(s) = 1/(s+1)^n$

with $n > 2$) and when it includes a delay (i.e. $G(s) = e^{-\tau s} B(s)/A(s)$). In such situations, the

estimates $\angle \hat{G}(j\omega_k)$ will not all be compatible with $\angle \hat{G}(j\omega_{\bar{k}})$ which was used to get \hat{f} . Then,

the model (\hat{G}, \hat{f}) will not converge (in probability) to (G, f) or to (G^-, f^-) .

3.2. Improvement of Bai's phase estimator when $i = 1$

The inconsistency of estimator (10) can be removed using the following improved rule:

$$\angle \hat{G}(j\omega_k) = \begin{cases} \tan^{-1} \left(\frac{\text{Im}(Y_{f,DFT}(i\omega_k))}{\text{Re}(Y_{f,DFT}(i\omega_k))} \right) & \text{if } \text{Re}(Y_{f,DFT}(i\omega_k)) > 0 \\ \tan^{-1} \left(\frac{\text{Im}(Y_{f,DFT}(i\omega_k))}{\text{Re}(Y_{f,DFT}(i\omega_k))} \right) + \text{sign}[\text{Im}(Y_{f,DFT}(i\omega_k))] \pi & \text{if } \text{Re}(Y_{f,DFT}(i\omega_k)) < 0 \end{cases} \quad (13)$$

This estimator is motivated by relation (9a) which (for $i=1$) reduces to:

$$Y_{f,DFT}(\omega_k) \rightarrow \frac{r_1}{2} \cos(\angle G(j\omega_k)) + j \frac{r_1}{2} \sin(\angle G(j\omega_k)) \quad (\text{in probability}), \text{ as } L \rightarrow \infty \quad (14)$$

Recall that $Y_{f,DFT}(\omega_k) \neq 0$ (otherwise $i > 1$). This in turn implies that $r_1 \neq 0$. Since $r_1^- = -r_1$, one

has either ($r_1 > 0$ and $r_1^- < 0$) or ($r_1 < 0$ and $r_1^- > 0$). Until now, $(G(\cdot), f(\cdot))$ and $(G^-(\cdot), f^-(\cdot))$ play

a symmetric role. From now on, these will be distinguished from each other considering that $(G(\cdot),$

$f(\cdot))$ is such that $r_1 > 0$. Let us show that the estimator (13) is actually consistent. First, suppose that:

$$\angle G(j\omega_k) \in \bigcup_{m=0}^{\infty} \left] -\frac{\pi}{2} + 2m\pi, \frac{\pi}{2} + 2m\pi \right[. \quad \text{Then,} \quad (14) \quad \text{yields:}$$

$$\tan^{-1} \left[\frac{\text{Im}(Y_{f,DFT}(\omega_k))}{\text{Re}(Y_{f,DFT}(\omega_k))} \right] \rightarrow \angle G(j\omega_k) + 2m\pi \quad (\text{for some integer } m). \text{ On the other hand, since}$$

$Real(G(j\omega_k)) > 0$, the first line of (13) yields $\angle \hat{G}(j\omega_k) = \tan^{-1} \left[\frac{Im(Y_{f,DFT}(\omega_k))}{Re(Y_{f,DFT}(\omega_k))} \right]$. Combining the

above results, one gets that $\angle \hat{G}(\omega_k) \rightarrow \angle G(j\omega_k) + 2m\pi$. Now, let

$\angle G(j\omega_k) \in \bigcup_{m=-\infty}^{+\infty} \left[-\pi + 2m\pi, -\frac{\pi}{2} + 2m\pi \right] \cup \left[\frac{\pi}{2} + 2m\pi, \pi + 2m\pi \right]$. Then, (14) implies that

$\tan^{-1} \left[\frac{Im(Y_{f,DFT}(\omega_k))}{Re(Y_{f,DFT}(\omega_k))} \right] \rightarrow \angle G(j\omega_k) + 2m\pi + \pi$ (for some integer m). On the other hand,

since $Real(G(j\omega_k)) < 0$, the second line of (13) yields $\angle \hat{G}(j\omega_k) = \tan^{-1} \left[\frac{Im(Y_{f,DFT}(\omega_k))}{Re(Y_{f,DFT}(\omega_k))} \right] \pm \pi$.

Combining the above results, one gets $\angle \hat{G}(j\omega_k) \rightarrow \angle G(j\omega_k) + 2m'\pi$ (with $m'=m$ or $m+1$).

Hence, the new phase estimator (13) is actually consistent.

4. INCONSISTENCY OF BAI'S METHOD WHEN $i > 1$

When $i > 1$, (11a) implies that: $i\angle \hat{G}(j\omega_k) \rightarrow i\angle G(j\omega_k) + p\pi$ (for some integer p) which shows that the phase estimator (10) is consistent if, for any pulsation ω_k , the corresponding integer p is such that $p = 2im$ (for some integer m). This result, together with (12), implies that the consistency

of (10) is ensured only for those pulsations such that: $\angle G(j\omega_k) \in \bigcup_{m=-\infty}^{m=+\infty} \left[-\frac{\pi}{2i} + 2m\pi, \frac{\pi}{2i} + 2m\pi \right]$.

Otherwise, it is inconsistent and then one has:

$$\angle \hat{G}(j\omega_k) \rightarrow \angle G(j\omega_k) + \frac{p}{i}\pi ; \text{ for some integer } p \neq 2im \quad (15)$$

In this situation, \hat{f} will not (uniformly) converge to f or f^- . Indeed, when $u(t) = A \cos(\omega_k t)$, the internal signal is estimated by $\hat{x}(t) \stackrel{def}{=} \cos(\omega_k t + \angle \hat{G}(j\omega_k))$. Then, it follows from (15) that, for all t :

$$\hat{x}(t) \rightarrow \cos\left(\omega_k t + \angle G(j\omega_k) + \frac{p}{i}\pi\right), \text{ in probability, as } L \rightarrow \infty \quad (16)$$

Since $p \neq 2im$, it readily follows from (16) and (8) that \hat{x} does not (uniformly) converge to x or to x^- . Consequently, the estimate $\hat{f}(\cdot)$ obtained from (\hat{x}, y) does not (uniformly) converge to f or to f^- . To illustrate this, let us consider the following example.

Example. Let the identified system be defined by:

$$G(s) = 10/(1+s)^4, f(x) = x^2, v(t) = 0 \quad (\text{no noise}) \quad (17)$$

Notice that, since f is even, one has $f = f^-$. Letting $\omega_{\bar{k}} = 4 \text{ (rd/s)}$, one gets $|G(j\omega_{\bar{k}})| \cong 0.034$ and

$\angle G(j\omega_{\bar{k}}) \cong -5.3 \text{ rd} \cong -1.69 \pi \text{ (rd)}$. Taking $u(t) = A \cos(\omega_{\bar{k}}t)$ with $A=1/0.034$, it follows that:

$$x(t) = \cos(\omega_{\bar{k}}t + \angle G(j\omega_{\bar{k}})) = \cos(\omega_{\bar{k}}t - 1.69\pi) \quad (18)$$

$$y(t) = (x(t))^2 = \frac{1}{2} + \frac{1}{2} \cos(2\omega_{\bar{k}}t - 3.38\pi) = \frac{1}{2} + \frac{1}{4} e^{-j3.38\pi} e^{+j2\omega_{\bar{k}}t} + \frac{1}{4} e^{+j3.38\pi} e^{-j2\omega_{\bar{k}}t} \quad (19)$$

Here, filtering is not needed as the output spectrum is limited (to $2\omega_{\bar{k}}$); thus $y_f = y$. Then, (19) yields:

$$Y_{f,DFT}(\omega_{\bar{k}}) = 0 \text{ and } Y_{f,DFT}(2\omega_{\bar{k}}) = \frac{\pi}{2} e^{-j3.38\pi} = -0.578 + j 1.46 \quad (20)$$

Equation (20) shows that $i = 2$. Then, (10) and (18) yield:

$$\angle \hat{G}(j\omega_{\bar{k}}) = \frac{1}{2} \tan^{-1} \left\{ \frac{1.46}{-0.578} \right\} = \frac{1}{2} \times (-1.1938) = -0.5969 \approx -0.19\pi$$

$$\hat{x}(t) = \cos(\omega_{\bar{k}}t + \angle \hat{G}(j\omega_{\bar{k}})) = \cos(\omega_{\bar{k}}t - 0.19\pi) = -\sin(\omega_{\bar{k}}t - 1.69\pi)$$

It is clearly seen that \hat{x} is quite different from both x and x^- . The functions $\hat{f}(\cdot)$ (represented by a number of couples $(\hat{x}(kT_s), y(kT_s))$) and $f(\cdot)$ are plotted in Fig 2 which clearly shows that $\hat{f}(\cdot)$ and $f(\cdot)$ are quite different from each other.

5. CONCLUSION

The identification method proposed in Bai (2003) is generally inconsistent. In the case where $i=1$, the method can be made consistent using the improved phase estimator (13).

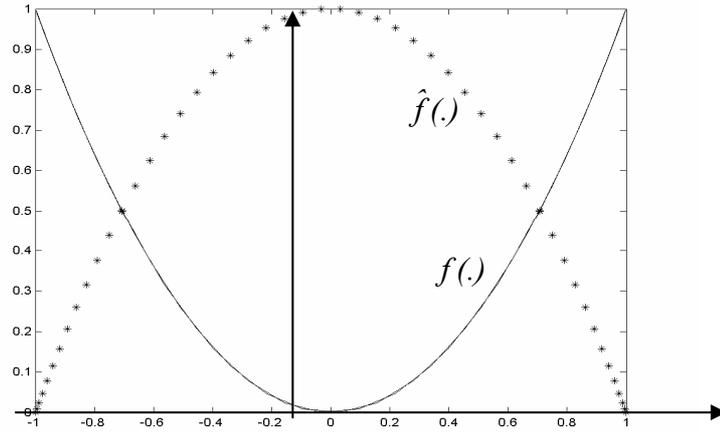


Figure 2. True nonlinearity $f(\cdot)$ (solid) and its estimate $\hat{f}(\cdot)$ (***)

BIBLIOGRAPHIE

Anderson B.D.O., "Adaptive systems, lack of persistency of excitation and bursting phenomena", *Automatica*, Vol. 21, pp. 247-258, 1985a.

Anderson B.D.O., Johnstone R.M., "Global adaptive pole positioning", *IEEE TAC*, AC-30, pp. 11-22, 1985b.

Anderson B.D.O., "Exponential stability of linear equations arising in adaptive identification", *IEEE TAC*, Vol. 22, pp. 83-88, 1977.

Bai E.W., "Frequency domain identification of Wiener models", *Automatica*, 39, 1521-1530, 2003.

Bai E.W., "Identification of systems with hard input nonlinearities of unknown structure", *Automatica*, vol 38, pp. 853-860, 2002.

Billings S.A., Fakhouri S.Y., "Nonlinear system identification using the Hammerstein model", *Int. J. Syst. Sci.*, 10, pp. 567-578, 1979.

Bitmead R.R., "Persistence of excitation conditions and the convergence of adaptive schemes", *IEEE-TAC*, Vol. 30, pp. 183-191, 1984.

Boutayeb M., Darouach M., Rafaralahy H., Krzakala G., "A new technique for identification of MISO Hammerstein model ", *Proceedings ACC San Francisco, California Juin 1993*.

Boyd S. and Sastry S., "Necessary and sufficient conditions for parameter convergence in adaptive control", *Automatica*, Vol. 22, pp. 629-639, 1986.

Chaoui F.Z., Giri F. Rochdi Y., Haloua M. , and Naitali A., "System Identification based on Hammerstein model", *International Journal of Control*, vol. 78, N° 6, 2005, pp. 430-442, 2005.

Chou, C.T., Heverkamp B., and Verhaegen M., "Linear and nonlinear system identification using separable least squares", *European Journal of Control*. 5, 116-128, 1999.

Cristi R., "Internal persistency of excitation in indirect adaptive control", *IEEE TAC*, Vol AC-32, pp. 1101-1103.

Egardt, E., "Stability of adaptive controllers", Springer Verlag, 1979.

Gardiner, A., "Frequency domain identification of nonlinear systems". *IFAC Symposium of System Identification and Estimation*, Haghe, Netherlands, 831-834, 1993.

Giri F., Dugard L., Dion J. M., M. M'Saad, "Stable pole placement direct adaptive control with arbitrary zeros", *Int. J. Control*, Vol. 47, pp. 1443-1454, 1988a.

Giri F., M. M'Saad, Dion J. M., L. Dugard, "Cautious approach for robust adaptive regulation", *Int. J. Adaptive Contr. Signal Processing*, Vol. 2, pp.273-290, 1988b.

Giri F., M'Saad M., Dugard L., Dion J. M., "Robust pole placement indirect adaptive control", *Int. J. Adaptive Control Signal Processing*, vol. 2, pp. 33-47, 1988c.

Giri F., M'Saad M., Dugard L., Dion J. M., "Robust pole placement direct adaptive control", *IEEE TAC*, Vol. 34, pp. 356-359, 1989a.

Giri F., Dion J. M., M'Saad M., Dugard L., "A globally convergent pole placement indirect adaptive controller", *IEEE-TAC*, Vol. 34, N° 3, pp. 353-356, 1989b.

Giri F., M'Saad M., Dion J. M., and Dugard L., "On the robustness of discrete-time indirect linear adaptive controllers", *Automatica*, vol. 27, pp. 153-160, 1990a.

Giri F., Dugard L., M'Saad M., and Dion J. M., "Parameter estimation aspects in adaptive control", *Automatica*, Vol.27, pp. 399-402, 1990b.

Giri F., M'Saad M., Dion J. M., Dugard L., "Pole placement direct adaptive control for time-varying ill-modeled plants", *IEEE TAC*, Vol.35, N°6, pp.723-726, 1990c.

Giri F., Ioannou P., Ahmed-Zaid F., "A stable indirect adaptive control scheme for 1st order plants with no prior knowledge on the parameters", IEEE-TAC, Vol. 38, pp.766-771, 1993.

Giri F., Chaoui F.Z., Rochdi Y., "Identification of nonlinear static gain plants", 12th IFAC Symposium on System Identification, Santa Barbara, California, USA, SYSID 2000.

Giri F., Chaoui F.Z., Rochdi Y., "Parameter identification of a class of Hammerstein plants", Automatica, vol. 37, pp 749-756, 2001a.

Giri F., Chaoui F.Z., Rochdi Y., "Hammerstein model Identification - The Case of Piecewise-Polynomial Static-Gain", IEEE-ACC, Arlington, Virginia, USA, Vol. 6, pp. 4882 – 4887, 2001b

Giri F., Chaoui F.Z., Rochdi Y., "A unified analysis of self-exciting regulators using internal impulse sequences", IEEE-ACC, Arlington, Virginia, USA, Vol. 4 , pp. 3299 – 3304, 2001c

Giri F., Chaoui F.Z., Rochdi Y., Haloua M., Naitali A., "Hammerstein model Identification", Proceedings of the 10TH Mediterranean Conference on Control and Automation, Lisbon, Portugal, 2002a. <http://med.ee.nd.edu/MED10/pdf/193.pdf>

Giri F., Chaoui F.Z. Rochdi Y., "Interval excitation through impulse sequences. A technical lemma", 38, pp. 457-465, Automatica, 2002b.

Giri F., Chaoui F.Z. and Rochdi Y., "System Identification of Hammerstein models including a known structure input nonlinearity - The Hysteresis case". IFAC Workshop on Adaptation and Learning in Control and Signal Processing, Yokohama, Japan , 2004a. Session MM3-A-4

Giri F., Chaoui F.Z., Rochdi Y., "Recursive Identification of Systems with Hard Input Nonlinearities of Known Structure", IEEE American Control Conference, ACC'04, Boston, Massachusetts, USA, 2004b, pp. 4764-4769.

Giri F., Rochdi Y., Chaoui F.Z. "Identification of the Nonlinear Element in Wiener Models. A Frequency-Geometric Solution", IEEE Conference on Decision and Control, CDC'06, 2006a.

Giri F., Rochdi Y., Chaoui F.Z. "Comment on 'Frequency domain identification of Wiener models, by E.W. Bai, Automatica 39 (2003), 1521-1530", En révision pour Automatica, 2006b.

Goodwin G.C. and Sin S., "Adaptive filtering prediction and control", Prentice Hall, 1983.

Goodwin G.C., Teoh E.K., "Persistency of excitation in the presence of possibly unbounded signals", IEEE TAC, AC-30, pp. 595-597.

Greblicki, W. "Non parametric orthogonal series identification of Hammerstein systems", Int. J. Systems Sciences, vol. 20, pp. 2355-2367, 1989.

Greblicki, W. and M. Pawlak, "Nonparametric identification of a cascade nonlinear time series system", Signal Processing, vol. 22, pp. 61-75, 1991a.

Greblicki W. and Pawlak M. , "Nonparametric recovering nonlinearities in block oriented systems with the help of Laguerre polynomials", Control-Theory and Advanced Technology, vol. 10, part 1, pp. 771-791, 1991b.

Greblicki, W., Nonparametric identification of Wiener systems. IEEE Transactions on Information Theory, 38, 1487-1493, 1992.

Greblicki W. "Nonparametric identification of Wiener systems by orthogonal series", IEEE Transactions on Automatic Control, vol. 39, No. 10, pp. 2077-2086, October 1994.

Greblicki, W., Pawlak M., "Nonparametric recovering nonlinearities in block oriented systems with the help of Laguerre polynomials", Control-Theory and Advanced Technology, vol. 10, part 1, pp. 771-791, 1994.

Greblicki, W., "Nonparametric approach to Wiener system identification". IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications, 38, 538-545, 1997.

Greblicki, W. "Continuous-Time Hammerstein System Identification", IEEE Transactions on Automatic Control, vol. 45, No. 6, pp. 1232-1236, June 2000.

Greblicki, W., "Recursive identification of Wiener system". Applied Mathematics and Computer Science, 11, 977-991, 2001.

Haber R. and Unbehauen H. "Structure identification of nonlinear dynamic Systems – A Survey on Input/Output Approaches", Automatica, Vol. 26, No. 4, pp. 651-677, 1990.

Haber R. and Keviczky L., "Nonlinear system identification. Input-Output Modeling Approach", Volume 1: Nonlinear system parameter identification, Kluwer Academic Publishers, Boston, 1999a, ISBN 0-7923-5857-0.

Haber R. and Keviczky L. "Nonlinear System Identification. Input/Output Modelling approach" Volume 2: Nonlinear system structure identification, Kluwer Academic Publishers, , 1999b, ISBN 0-7923-5857-0.

Hasiewicz, Z., "Identification of a linear system observed through zero-memory nonlinearity". Int. J. of Systems Science, 18, 1595-1607, 1987.

Helliot H. , Cristi R. , Das M. , "Global stability of adaptive pole placement algorithms", IEEE TAC, AC-30, pp. 348-356, 1985.

Hu, X.L. and H.F. Chen, "Strong consistence of recursive identification for Wiener systems". Automatica, 41, 1905-1916, 2005.

Hunter, I.W. and M.J. Korenberg "The identification of nonlinear biological systems: Wiener and Hammerstein cascade models". Biological Cybernetics, 55, 135-144, 1986.

Ioannou P., SunJ., "Robust adaptive control", Prentice Hall, 1996.

Johnstone R.M., Anderson B.D.O., "Exponential convergence of recursive least squares with exponential forgetting factor– adaptive control", Syst. Contr. Letters, Vol 2, pp. 69-76, 1982.

Kreisselmeier G., Smith M., "Stable adaptive regulation of arbitrary nth order plants", IEEE-TAC, Vol. AC-31, pp. 299-305, 1986.

Kreisselmeier G. "An indirect adaptive controller with self-excitation capability", IEEE-TAC, Vol. AC-34, pp. 524-528, 1989.

Krzyzak A., "Identification of discrete Hammerstein systems by the Fourier series regression estimate", Int. J. Systems Sciences, vol. 20, pp. 1729-1744 , 1989.

Kung F.C. and Shinh D.H., "Analysis and identification of Hammerstein model nonlinear systems using block-pulse function expansion", Int. J. Control, vol. 43, pp. 139-147, 1986.

Lang, Z.Q., "Controller design oriented model identification method for Hammerstein systems", Automatica, vol. 29, pp. 767-771, 1993a.

Lang Z.Q. "A nonparametric polynomial identification algorithm for the Hammerstein system", IEEE-TAC, vol. 42, pp. 1435-1441, 1997.

Lang Z.Q., "Controller design oriented model identification method for Hammerstein systems", Automatica, vol. 29, pp. 767-771, 1993b.

Ljung, L., "System identification. Theory for the user", Prentice-Hall, ISBN: 0-13-656695-2, 1997.

Lozano R., Taoutaou D., "Identification and adaptive control", Hermes-Science, ISBN: 2-7462-0217-4, 2001.

Mareels I.M.Y., Bitmead R.R. , "Nonlinear dynamics in adaptive control: chaos and periodic stabilization", Automatica, Vol. 22, pp. 641-655, 1985.

Mzyk, G. "Hammerstein System Identification by a Semi-Parametric Method", Proceedings of the ICSES, 2000.

Mzyk, G., " Nonparametric identification of Wiener systems - A median based approach", Proceedings of the 10th IEEE International Conference MMAR, 2004.

Naitali A., Giri F., F.Z. Chaoui, M. Haloua, Y. Rochdi, "Parameter Identification Based on Hammerstein Models-A subspace Approach", IFAC Symposium on System Identification (SYSID), Rotterdam, The Netherlands, pp. 1327-1332, 2003.

Narendra K.S., Annaswamy A.M., "Persistent excitation of adaptive systems", Int. J. Control, Vol. 31, pp. 306-315, 1986.

Narendra K.S., Annaswamy A.M., "Stable adaptive systems", Prentice Hall, 1989.

Nordsjö, A.E. and Zetterberg L.H., "Identification of certain time-varying nonlinear Wiener and Hammerstein systems". IEEE Trans on Signal Processing, 49, 577-592, 2001.

Pajunen, G.A., "Adaptive control of Wiener type nonlinear systems". Automatica, 28, 781-785, 1992.

Pawlak, M., "On the series expansion approach to the identification of Hammerstein systems", IEEE-TAC, vol. 36, pp. 736-767, 1991.

Polderman J.W., "A state space approach to the problem of adaptive pole assignment", Math. Control Signals Systems, Vol. 2, pp. 71-94, 1989.

Rangan, S., Wolodkin, G., Poolla K. "New Results for Hammerstein System Identification". Proceedings 34th CDC News Orleans, LA - December 1995.

Sastry S., Bodson M., "Adaptive control: stability, convergence and robustness", Prentice Hall, 1989.

Soderstrom, T. and Stoica P. "System identification", Prentice-Hall, ISBN:0-138-81236-5, 1988.

Stoica P. and Söderström T., "Instrumental variable methods for identification of Hammerstein systems", Int. J. control, vol. 35, pp. 459-476, 1982.

Tao G. and Kokotovic P., "Adaptive Control of Systems with Actuator and Sensor Nonlinearities". John Wiley & Sons, 1996. ISBN 0-471-15654-X .

Tao G., and Lewis F. F. (Eds.). "Adaptive Control of Non smooth Dynamic Systems", Springer, 2001, ISBN 1-85233-384-7.

Taware A., Tao G., "Control of Sandwich Nonlinear Systems", Springer, ISBN 3-540-44115-8, 2003.

Van Pelt, T.H Roup, A. V. and Bernstein, D. S. "Nonlinear system identification using dual-input Hammerstein models with application to thrust vectoring " .

Voros J., "Parameter identification of discontinuous Hammerstein systems", Automatica, vol.33, pp.1141-1145. , 1999a .

Vörös J., "Iterative algorithm for identification of Hammerstein systems with two-segment nonlinearities", IEEE Transactions on Automatic Control, vol. 44, pp. 2145-2149, 1999b.

Vörös, J., "Parameter identification of Wiener systems with discontinuous nonlinearities" .Systems and Control Letters, 44, 363-372, 1997.

Vörös J. , "Parameter identification of Wiener systems with discontinuous nonlinearities", Systems & Control Letters, vol. 44, pp. 363-372, 2001.

Weisstein E.W. "Lissajous Curve". From MathWorld, a Wolfram Web Resource. <http://mathworld.wolfram.com/LissajousCurve.html>, 2006.

Westwick, D.T. and R.E. Kearney, "A new algorithm for the identification of multiple-input Wiener systems". Biological Cybernetics, 68, 75-85, (1992).

Wiener N., "Nonlinear problems in random theory", MIT Press, Cambridge, MA, USA, 1958

Wigren, T., "Recursive prediction error identification using the nonlinear Wiener model", Automatica, 29, 1011-1025, 1993.

Zhu, Y., "Estimation of an N-L-N Hammerstein-Wiener model". Automatica Vol. 38, pp. 1607-1614, 2002.

RESUME. Cette thèse porte sur le problème d'identification des systèmes non linéaires sur la base des modèles blocs. Deux types de modèles sont considérés ici, ceux de type Hammerstein et ceux de type Wiener. La plupart des solutions antérieures ont été élaborées sous des hypothèses assez contraignantes concernant le sous-système non linéaire du modèle. Celui-ci a souvent été supposé lisse (voire polynomial), inversible et sans mémoire. Les travaux présentés dans cette thèse tentent de repousser ces limites. Dans le cas des systèmes de type Hammerstein, nous élaborons un schéma d'identification qui ne fait aucune hypothèse sur l'élément non linéaire, à l'exception du fait qu'il est sans mémoire et L_∞ -stable ; ce schéma estime parfaitement (du moins dans la cas idéal) les paramètres du sous-système dynamique linéaire et un nuage de points de la caractéristique de l'élément non linéaire. Ce schéma est ensuite adapté au cas où l'on connaît la structure de l'élément non linéaire ; à ce propos, les non-linéarités statiques affines par morceaux et discontinues jouissent d'une attention particulière. Nous terminons la partie consacrée à l'identification des systèmes de Hammerstein, en abordant le problème des systèmes impliquant un élément non linéaire à mémoire. Deux familles d'éléments de cette nature sont considérées : celle comprenant des éléments de type hystérisis sans saturation et celle comprenant des hystérisis-relais. Le problème est appréhendé à l'aide d'un schéma d'identification dont on établit la consistance en présence de perturbations assimilables à un bruit blanc appliqué en sortie du système. La dernière partie du mémoire est centrée sur l'identification des systèmes de Wiener, dont l'élément non linéaire n'est pas supposé inversible. A cet effet, nous présentons deux schémas d'identification de type fréquentiel et établissons leur consistance dans les mêmes conditions que précédemment concernant les perturbations. L'exigence d'excitation persistante occupe une place centrale dans cette thèse. Pour procurer cette propriété aux différents schémas d'identifications proposés, il a été fait appel à une famille de signaux d'excitation de type impulsionnel. Dans ce cadre, nous avons élaboré un lemme technique précisant, pour les systèmes linéaires, le lien entre cette famille de signaux et la propriété d'excitation persistante. L'adaptation de ce lemme au cas des systèmes non linéaires est illustrée dans les différents schémas d'identification.

ABSTRACT. This thesis deals with the problem of identifying nonlinear systems based on bloc models. Two types of models are considered, those of Hammerstein and those of Wiener. Most of previous solutions have been designed under restrictive constraints regarding the nonlinear element of the model. This has been usually supposed to be smooth (or even polynomial), invertible and memoryless. In the case of Hammerstein systems, an identification scheme is designed that involves no assumptions on the nonlinear element except that it is memoryless and L_∞ -stable. In ideal situations (undisturbed system), the proposed scheme determines exactly the parameters of the linear subsystem as well as a set of points of the nonlinear subsystem characteristic. This scheme is then adapted to the case where the structure of the nonlinear element is known. Then, the focus is particularly made on piecewise affine discontinuous nonlinearities. The part of this thesis that deals with Hammerstein systems identification is completed considering the problem of identifying systems that includes memory nonlinear elements. Two families of this type are focused on: the first one includes (unsaturated) hysteresis elements; the second one involves hysteresis-relay elements. The problem is coped with building up an identification scheme for which consistence results are achieved in presence of disturbances that can be assimilated to white noise that affects the output. The last part of this report is centred on the identification of Wiener systems whose nonlinear element is not necessarily invertible. Two identification schemes are constructed, to deal with this problem, and shown to be consistent in the same conditions as previously concerning the disturbances. The persistent excitation requirement plays a central role in the in this thesis. The different identification schemes are given this property through impulse type input signals. To this end, a technical lemma is developed that describes, in a general framework, how persistent excitation can be provided to linear systems. The exploitation of this lemma in the nonlinear context is illustrated through the analysis and design of the different identification schemes.