

Titre: Etude des régulateurs à structure variable pour systèmes
Title: stochastiques

Auteur: Pierre Gauthier
Author:

Date: 1990

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Gauthier, P. (1990). Etude des régulateurs à structure variable pour systèmes
Citation: stochastiques [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
<https://publications.polymtl.ca/59269/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/59269/>
PolyPublie URL:

**Directeurs de
recherche:**
Advisors:

Programme: Non spécifié
Program:

UNIVERSITÉ DE MONTRÉAL

**ÉTUDE DES RÉGULATEURS À STRUCTURE VARIABLE POUR
SYSTÈMES STOCHASTIQUES**

par

**Pierre GAUTHIER
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE**

**MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU GRADE DE MAÎTRE ES SCIENCES APPLIQUÉES (M. Sc. A.)**

Janvier 1990

© Pierre Gauthier 1990



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-58117-4

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

**ÉTUDE DES RÉGULATEURS À STRUCTURE VARIABLE POUR
SYSTÈMES STOCHASTIQUES**

Présenté par: Pierre GAUTHIER

en vue de l'obtention du grade de: M.Sc.A.

a été dûment accepté par le jury d'examen constitué de:

M. El Kébir Boukas , Ph.D. , président

M. Roland Malhamé , Ph.D.

M. Mario Lefebvre , Ph.D.

Sommaire

Une première tentative de généralisation de la théorie de la commande à structure variable aux systèmes dynamiques linéaires stochastiques est amorcée ici. Dans ce but nous proposons une caractérisation de tels régulateurs, comme modèles stochastiques hybrides. À partir de cette représentation stochastique hybride, on montre que l'évolution de la fonction de densité de probabilité des états du système dans l'espace de phase peut être décrite au moyen d'un système d'équations aux dérivées partielles couplées; les équations de Fokker-Planck du modèle hybride.

Les équations de Fokker-Planck étant en général très difficiles à résoudre analytiquement, un schéma numérique s'inspirant de la méthode des directions alternées implicites est formulé pour la simulation du système. Cet outil de simulation est utilisé afin d'étudier les effets du mode de glissement au niveau des deux premiers moments de la densité de probabilité conjointes des états.

Les résultats obtenus permettent d'apprécier les propriétés d'invariance par rapport au bruit lorsque le système est en mode de glissement et confirment la validité potentielle de l'approche du modèle équivalent pour l'étude du mouvement moyen du processus contrôlé.

Une conjecture sur l'existence d'une équation de Riccati équivalente, dont les paramètres sont les matrices obtenues en utilisant le modèle équivalent en mode de glissement, est proposée. La plausibilité de cette conjecture est démontrée sur une

base analytique et par simulation. Finalement on utilise l'équation de Riccati équivalente pour montrer que l'invariance par rapport au bruit peut s'étendre à l'évolution des covariances.

Abstract

This work presents a first attempt to generalize the variable structure control theory to linear stochastic systems. With this objective in mind, a characterization of these controllers as hybrid stochastic models is made. Based on that hybrid stochastic representation we show that the evolution of the probability density function of the states in phase space can be described by a system of coupled partial differential equations; the Fokker-Planck equations of the hybrid model.

Since analytical solutions to such equations are generally rarely known, a numerical scheme for the simulation of the Fokker-Planck equations based on the alternated directions implicit method is formulated. This simulation tool is then used to study the effects of sliding control on the first two moments of the joint probability density of states.

Simulation results when the system is in sliding mode motion allow us to appraise the invariance property with respect to noise of these controllers and to confirm the potential adequacy of the equivalent model approach for the study of the controlled process mean motion.

A conjecture about the existence of an equivalent Riccati equation, the parameters of which are the matrices derived from the equivalent model of the controller in the sliding mode, is then introduced.

The plausibility of the conjecture is assessed both by analytical and numerical means. Finally we use the equivalent Riccati equation to show that the noise/parameters invariance property can be extended to the evolution of the covariances of the process.

Remerciements

Mes remerciements s'adressent d'abord à mon directeur de recherches , le professeur Roland Malhamé, qui par son intelligence et son humanité a rendu possible ce travail. Ce mémoire est d'abord et avant tout le résultat de sa patience, de son sens profond de la communication et de sa compétence scientifique. Nos discussions auront été un terrain d'apprentissage inestimable!

Je remercie aussi les membres du jury, mes professeurs Mario Lefebvre et El Kébir Boukas pour avoir accepté d'évaluer ce mémoire.

J'exprime aussi mes remerciements à mes amis Maher Hammami et Sofiène Kamoun pour leur support moral et technique.

Enfin ma plus grande gratitude va à Carole pour sa patience, son affection, ses conseils et son soutien moral , et à ma famille pour leur encouragements et leur compréhension.

Table des matières

Sommaire	iv
Abstract	vi
Remerciements	viii
Liste des figures	xiii
Introduction	1
1 Préliminaires théoriques	4
1.1 Systèmes à structure variable	4
1.1.1 Définition	4
1.1.2 Propriétés générales des systèmes à structure variable	5
1.1.3 Méthode du prolongement de Filippov	6
1.1.4 Relais non-idéaux et mouvement moyen	9
1.1.5 Stabilisation des systèmes à structure variable (SSV)	10
1.1.6 Existence du mode de glissement	12
1.1.7 Équations du mode de glissement	14
1.2 Équations différentielles stochastiques	15
1.2.1 Équation d'Itô	15
1.2.2 Équation de Fokker-Planck	16
1.2.3 Courant de probabilité	18

1.2.4	Bruit blanc gaussien et équations d'Itô	19
1.3	Conditions frontières	20
1.3.1	Paroi réfléchissante	20
1.3.2	Paroi absorbante	21
1.3.3	Paroi de discontinuité	21
2	Modèle d'un régulateur à structure variable en présence de bruit	22
2.1	Origine du modèle	22
2.2	Relation avec les régulateurs à structure variable	27
2.3	Modèle dans le plan de phases	28
2.4	Modèle dans un repère non-orthogonal	31
2.5	Nature des équations à la paroi de réinsertion de probabilité	34
2.5.1	Schéma de réinsertion des probabilités	35
2.5.2	Perte de probabilité à la paroi absorbante	36
2.5.3	Réinsertion de probabilités à la paroi de réinsertion	38
2.5.4	Terme d'absorption/réinsertion	46
2.6	Discontinuité paramétrique	46
2.7	Parois réfléchissantes	50
3	Schéma numérique	53
3.1	Décomposition d'opérateur	53
3.2	Décomposition de l'équation de Fokker-Planck	55
3.3	Discretisation de l'équation de Fokker-Planck excluant discontinuités et conditions frontières	57
3.4	Conditions frontières	62
3.4.1	Parois réfléchissantes	62
3.4.2	Paroi de réinsertion de probabilité	74
3.4.3	Paroi absorbante	76
3.4.4	Discontinuité Paramétrique	76

4	Implantation du schéma numérique	77
4.1	Constitution du maillage	77
4.2	Calcul de la densité initiale	77
4.3	Choix de l'origine du système pour l'écriture des matrices	78
4.4	Calcul de l'évolution des densités des états	79
4.4.1	Etude du premier demi-pas	80
4.4.2	Etude du second demi-pas	81
4.4.3	Quelques sous-matrices particulières pour le premier demi-pas	83
4.4.4	Quelques sous-matrices particulières pour le second demi-pas .	83
4.4.5	Calcul des matrices tridiagonales générales G	84
4.4.6	Première distribution	84
4.4.7	Deuxième distribution	86
4.4.8	Cycle de calcul	88
4.5	Calcul des moments du système	88
4.6	Équation de transformation des moments	89
5	Vérification de la précision et stabilité du modèle	91
5.1	Vérification du modèle sans discontinuités paramétriques et avec structure unique	91
5.2	Équation de Riccati	92
5.3	Conservation de la probabilité totale	94
5.4	Précision du schéma au niveau des moyennes	96
5.5	Précision du schéma obtenu au niveau des covariances	98
5.6	Hypothèse de linéarité	102
6	Conditions d'invariances et modes glissants	104
6.1	Théorie classique de l'invariance	104
6.2	Invariance du mode de glissement en présence de bruit	106
6.3	Application au mouvement dans l'espace de phase	109

6.4	Invariance par rapport aux variations paramétriques	110
6.5	Résultats de simulation	112
6.6	Domaines de validité de l'équation équivalente en présence de bruit .	117
	Conclusion	128
	Bibliographie	132
	Annexes	135

Liste des figures

1.1	Comportements possibles de la trajectoire dans l'espace de phase au voisinage de $S(X) = 0$	7
1.2	Construction géométrique du champ équivalent	8
1.3	Trajectoires dans l'espace de phase pour $\dot{X} = f^1(X)$	11
1.4	Trajectoires dans l'espace de phase pour $\dot{X} = f^2(X)$	11
1.5	Trajectoires dans l'espace de phase pour la combinaison des systèmes	12
2.1	Fonctions de densité du modèle hybride	25
2.2	Représentation du modèle de transition	29
2.3	Régions correspondant aux deux sous-ensembles de structures $m(t)=1,-1$	30
2.4	Systèmes de coordonnées utilisés.	32
2.5	Réinsertion et absorption de probabilités	35
2.6	Mécanisme de pertes de probabilité	36
2.7	Mécanisme de réinsertion de probabilité	38
2.8	Volume élémentaire pour le traitement de la discontinuité paramétrique	47
2.9	Volume élémentaire pour le traitement des parois réfléchissantes . . .	52
3.1	Tampon imaginaire utilisé pour les parois réfléchissantes	63
4.1	Système de coordonnées pour chacune des distributions	79
4.2	Formation de la matrice G_{\mp}^1 : premier demi-pas	81
4.3	Formation de la matrice G_{\mp}^2 : second demi-pas	82

5.1 Evolution de la probabilité dans chaque distribution. Paramètres: $\theta = 0, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$. . . 95

5.2 Evolution de la probabilité dans chaque distribution. Paramètres: $\theta = 35, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$. . . 95

5.3 Evolution de $E(x_1)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 96

5.4 Evolution de $E(x_1)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 97

5.5 Evolution de $E(x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 97

5.6 Evolution de $E(x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 98

5.7 Evolution de $E(x_1^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 99

5.8 Evolution de $E(x_1^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 99

5.9 Evolution de $E(x_2^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 100

5.10 Evolution de $E(x_2^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35, a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$ 100

5.11	Evolution de $E(x_1x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$	101
5.12	Evolution de $E(x_1x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$	101
5.13	Evolution dans l'espace de phases par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$	102
5.14	Evolution dans l'espace de phases par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$	103
6.1	Validité du modèle équivalent dans le plan de phase	114
6.2	Equation de Riccati équivalente pour Ex_1^2	115
6.3	Equation de Riccati équivalente pour Ex_2^2	115
6.4	Equation de Riccati équivalente pour Ex_1x_2	116
6.5	Conservation de probabilité	116
6.6	Dynamique du transfert de probabilité	117
6.7	Domaines de stabilité en présence du bruit $K_1 = .1$, $\delta = 5, 10, 15$	127
6.8	Domaines de stabilité en présence du bruit $K_1 = .05$, $\delta = 5, 10, 15$	127

Introduction

Le premier chapitre présente un certain nombre de préliminaires théoriques utiles à une bonne compréhension du mémoire. Il est essentiellement constitué de deux parties. Une première partie qui introduit des définitions et présente des résultats issus de la théorie des régulateurs à structure variable déterministes. Et une seconde partie qui aborde la théorie des équations différentielles stochastiques en introduisant les équations d'Îto et l'équation de Fokker-Planck, qui sera le point d'appui de toutes nos simulations. On donne une interprétation "physique" de cette équation comme divergence d'un courant de probabilité. Cette interprétation est ensuite utilisée pour traiter des trois classes de conditions frontières qu'on utilisera par la suite.

Le deuxième chapitre introduit notre modèle des régulateurs à structure variable pour systèmes stochastiques. On présente son origine, le modèle stochastique hybride des charges électriques de Malhamé et Chong. À partir de cette représentation stochastique hybride, on montre que l'évolution de la fonction de densité de probabilité des états du système dans l'espace de phase peut être décrite au moyen d'un système d'équations aux dérivées partielles couplées; les équations de Fokker-Planck du modèle hybride. On dérive ensuite les équations de Fokker-Planck associées au modèle hybride dans un repère non-orthogonal adapté à la géométrie du problème. Une attention particulière est apportée à la dérivation des équations de Fokker-Planck aux parois de réinsertion et d'absorption de probabilités. On dérive finalement les équations de Fokker-Planck sur les parois réfléchissantes et de discontinuité paramétrique.

Dans le troisième chapitre on formule un schéma numérique pour la simulation du système couplé d'équations aux dérivées partielles du modèle hybride. En s'inspirant de la méthode des directions alternées implicites et de la notion plus générale de fractionnement d'opérateur différentiel, on introduit une décomposition de l'équation de Fokker-Planck qui conduit à une méthode de résolution numérique stable et précise en deux demi-pas. On établit ensuite la discrétisation, que suggère notre méthode de résolution numérique, des expressions analytiques de l'équations de Fokker-Planck dérivées au second chapitre.

Le quatrième chapitre traite de l'implantation du schéma numérique. On montre comment sont constituées les matrices des systèmes tridiagonaux associés à l'application de notre méthode numérique, et ce pour chaque demi-pas. Une fois ces matrices constituées, on montre comment s'opère la solution complète de notre système en exposant les différentes étapes d'un cycle complet de calcul. On introduit les approximations utilisées pour calculer les différents moments du système, les équations de transformation des moments du système de coordonnées non-orthogonal au repère cartésien, les contraintes sur le maillage utilisé et la procédure de calcul de la densité initiale.

Le cinquième chapitre a pour objectif la vérification de la précision et de la stabilité du modèle. Au moyen de simulations, dont les paramètres sont choisis conformément à certaines hypothèses de linéarité près de la paroi de commutation, on caractérise la précision des solutions obtenues pour les deux premiers moments de la densité de probabilité conjointes des états. Cette caractérisation est faite au moyen d'une comparaison des résultats obtenus par simulation du modèle hybride et de l'équation de Riccati pour des conditions initiales identiques. On vérifie ensuite la validité du modèle de réinsertion/absorption de probabilité par l'analyse des propriétés de conservation de la probabilité totale et au moyen d'un examen de la convergence des solutions du modèle hybride vers le modèle linéaire loin de la paroi de commutation.

Après s'être assuré de la validité du schéma numérique on passe au sixième chapitre dont l'objet est l'étude des effets du mode de glissement au niveau des deux premiers moments de la densité de probabilité conjointe des états. On expose la théorie de l'invariance classique pour des systèmes déterministes, puis on réexamine cette théorie d'un point de vue stochastique. On construit un système qui assure l'invariance par rapport au bruit et aux paramètres. On simule ce système afin de montrer la validité du modèle équivalent pour l'étude du mouvement moyen et pour évaluer la plausibilité d'une équation de Riccati équivalente, dont les paramètres sont obtenus en utilisant le modèle équivalent en mode de glissement. Finalement on étudie les domaines de validité de l'équation équivalente en présence de bruit à l'aide d'une fonction de Lyapounov. Une analyse détaillée de la différentielle de cette fonction près du sous-espace de commutation conduit à une forme hermitienne. Cette forme est alors utilisée pour montrer que le taux moyen d'accroissement de la fonction de Lyapounov est strictement négatif à l'extérieur, et strictement positif à l'intérieur, d'une certaine zone près de l'origine.

Chapitre 1

Préliminaires théoriques

Ce chapitre a pour but d'introduire le lecteur à certains concepts utiles à une bonne compréhension des régulateurs à structure variable pour systèmes stochastiques. Dans un premier temps on présentera les régulateurs à structure variable déterministes, puis on examinera la nature des processus solutions d'une classe d'équations différentielles stochastiques qui servira de référence pour l'ensemble du travail.

1.1 Systèmes à structure variable

1.1.1 Définition

Un système à structure variable est un système dynamique, Σ , continu par rapport à ses arguments, (X, u) , et dont la commande, u , est une fonction continue par morceaux des variables d'états, X . A chaque instant la commande, u , appartient à un ensemble fini, \mathcal{U} de fonctions continues, l'élément particulier choisi est déterminé par une fonction de commutation, $S(X)$, qui dépend de l'état du système. Un système à structure variable peut donc être décrit par la concaténation d'un ensemble de sous-systèmes continus, $\Sigma_j(u_j, X, S(X))$ liés par une fonction de commutation $S(X)$.

1.1.2 Propriétés générales des systèmes à structure variable

La compréhension du comportement des solutions résultant de l'application d'une commande non-linéaire est fondamentale dans la théorie des systèmes à structure variable. Les techniques habituelles de résolution d'équations différentielles ne s'appliquent pas. On utilise la théorie des équations différentielles discontinues à droite. La description du comportement de la solution près du sous-espace de commutation $S(X)$ est particulièrement importante. Soit le système à structure variable invariant dans le temps suivant:

$$\dot{X} = f(X, u) \quad (1.1)$$

avec $X \in \mathfrak{R}^n$, $u \in \mathfrak{R}^m$ et $f = (f^1, \dots, f^n)^T$ continue par rapport à X et u . Avec u une fonction scalaire continue par morceaux de X :

$$u(X) = \begin{cases} u^+(X) & \text{si } S(X) > 0 \\ u^-(X) & \text{si } S(X) < 0 \end{cases} \quad (1.2)$$

Le sous-espace de commutation, $S(X)$, dans l'espace de phase \mathfrak{R}^n est continu par rapport à X . Dans le cas du contrôle scalaire $S(X) \in \mathfrak{R}^{n-1}$.

On peut réécrire le système (1.1) comme suit:

$$\dot{X} = \begin{cases} f^+(X) & \text{si } S(X) > 0 \\ f^-(X) & \text{si } S(X) < 0 \end{cases} \quad (1.3)$$

On suppose que les fonctions $f^+(X)$ et $f^-(X)$ satisfont chacune à la condition de Lipschitz, i.e. $\|f^+(X)\| \leq K^+ \|X^+\|$ et $\|f^-(X)\| \leq K^- \|X^-\|$ avec $K^+ > 0$ et

$K^- > 0$. De sorte que la solution $X(t)$ est bien définie (la solution existe et elle est unique (Coddington 1951)) lorsque $X(t)$ ne croise pas le sous-espace de commutation $S(X) = 0$. Cependant lorsque $X(t)$ croise le sous-espace de commutation, le théorème d'existence et d'unicité de la solution ne s'applique plus (Utkin 1974). Les méthodes habituelles de prolongement des solutions d'équations différentielles ne s'y appliquent plus.

Une théorie qui traite rigoureusement les équations différentielles discontinues à droite, permettant le prolongement des solutions lorsque la trajectoire croise $S(X)$, a été développé par Filippov (Filippov 1964), (Paden et Sastry 1986).

1.1.3 Méthode du prolongement de Filippov

En se limitant au cas où $X \in \mathfrak{R}^2$, examinons les comportements possibles de la trajectoire dans l'espace de phases selon les directions du champ f^+ et f^- au voisinage de $S(X) = 0$. [Figure 1.1]

Trois cas sont possibles:

a) f^+ pointe vers $S(X) = 0$ et f^- pointe hors de $S(X) = 0$. La trajectoire est donc dirigée vers $S(X) = 0$ à partir de $S(X) > 0$ et se poursuit dans l'autre région subissant un changement de vitesse à $S(X) = 0$.

b) Les deux champs vectoriels s'éloignent de $S(X) = 0$. La trajectoire ne croise jamais $S(X) = 0$.

c) Les deux champs pointent vers $S(X) = 0$. La trajectoire partant d'une région ou l'autre est dirigée vers $S(X) = 0$ et y est coincée.

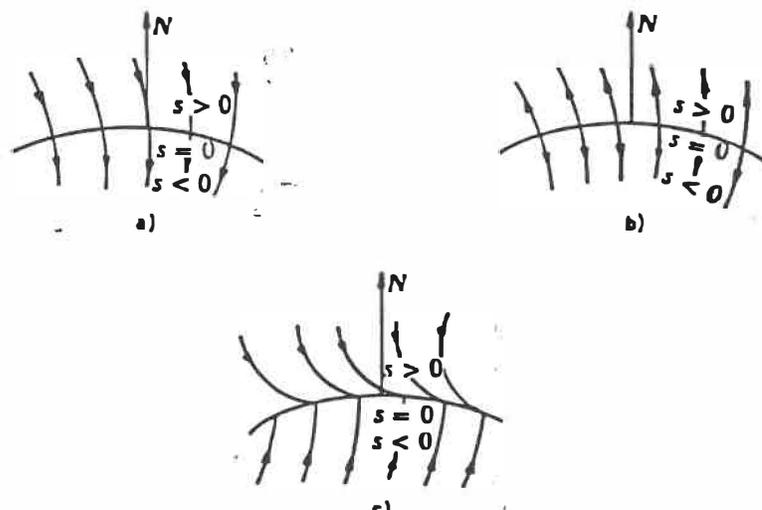


Figure 1.1: Comportements possibles de la trajectoire dans l'espace de phase au voisinage de $S(X) = 0$

Le premier cas se produit dans de nombreux systèmes à relais simples. Ce type de comportements est analysé en raccordant les solutions associées à la trajectoire dans $S(X) > 0$ et $S(X) < 0$. Le deuxième cas est analysé par les techniques habituelles des équations différentielles puisque la trajectoire ne croise jamais $S(X) = 0$.

Le comportement illustré en (c) est le mode glissant et il est analysé en utilisant la méthode du prolongement de Filippov. Une trajectoire qui atteint le sous-espace de commutation en X_0 doit par la suite demeurer sur la surface $S(X) = 0$. Un champ vectoriel f^0 équivalent est donné comme suit (Filippov 1964):

$$f^0 = \frac{(\text{grad}S) \cdot f^-}{(\text{grad}S) \cdot (f^- - f^+)} f^+ - \frac{(\text{grad}S) \cdot f^+}{(\text{grad}S) \cdot (f^- - f^+)} f^- \quad (1.4)$$

Le champ f^0 est une moyenne pondérée des champs vectoriels f^+ et f^- . On peut le construire graphiquement en prolongeant f^+ et f^- dans les régions $S(X) < 0$ et

$S(X) > 0$ respectivement. On trace une ligne connectant les extrémités des deux vecteurs et f^0 est alors le vecteur tangent à $S(X) = 0$ en X_0 dont l'extrémité est sur la ligne de connexion. L'équation différentielle résultante est alors $\dot{X} = f^0(X)$. Notons qu'en général on obtiendra le comportement limite en considérant que $S(X) = 0$ est une contrainte sur l'équation différentielle (1.1) permettant ainsi sa réduction à une équation différentielle d'ordre $(n-1)$.

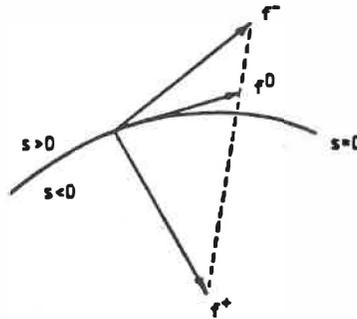


Figure 1.2: Construction géométrique du champ équivalent

Une forme explicite pour f^0 est garantie (Utkin 1977) si le système est linéaire par rapport à la commande, i.e si (1.1) est de la forme:

$$\dot{X} = f(X) + B(X)u \quad (1.5)$$

où $f(X)$ et X sont des vecteurs colonnes de dimension n , $B(X)$ est une matrice $n \times m$ et u un vecteur colonne de dimension m .

1.1.4 Relais non-idéaux et mouvement moyen

Dans les relais non-idéaux, des propriétés physiques comme l'hystérèse ou des délais peuvent modifier les comportements qu'on vient de voir. Les cas (a) et (b) peuvent être analysés en ajoutant des perturbations à la solution. Pour le cas (c) on doit vérifier que le mode de glissement existe encore.

En général avec des relais non-idéaux la trajectoire d'une région se poursuit un peu dans l'autre avant que la commutation n'ait lieu. Puisque la trajectoire est alors contenue dans une région où le champ vectoriel est continu, le mouvement est bien défini. Elle est ensuite redirigée vers la surface de commutation où la commutation se produit effectivement après un temps suffisamment long pour que la trajectoire soit entièrement contenue dans l'autre région. Le comportement résultant en est un de haute fréquence et de faible amplitude appelé "chattering".

Généralement le "chattering" est négligé dans la modélisation du comportement sur la surface de commutation et on ne retient que le mouvement moyen. Ce dernier étant le mouvement limite obtenu lorsque les paramètres non-idéaux tendent vers zéro.

Pour le cas des anomalies de type hystérèse et délais en commande scalaire discontinue, on peut montrer que le mouvement moyen en mode glissant est égal au mouvement obtenu en utilisant la méthode du prolongement de Filippov. On peut aussi montrer que le même système sous d'autres anomalies peut converger vers un autre modèle de prolongement: le modèle de la commande équivalente. Un résultat important pour notre étude est que ces différences disparaissent si le système est linéaire par rapport à la commande (Utkin 1974).

1.1.5 Stabilisation des systèmes à structure variable (SSV)

L'étude de la stabilité d'un système à retour d'état discontinu est complexe comme pour la majorité des systèmes non-linéaires. Pour prouver la stabilité à l'origine on utilise souvent des fonctions de Lyapounov. Une modification de cette méthode visant son utilisation sur des systèmes à commande discontinue est que la fonction V doit être continûment différentiable presque partout, i.e. V ou \dot{V} n'ont pas besoin d'être définies sur la surface de commutation $S(X) = 0$.

L'étude de la stabilité des systèmes SSV est décomposée en deux problèmes distincts (DeCarlo 1988). En premier lieu on doit trouver des conditions d'atteinte de la surface de commutation. Soit que l'on démontre l'attractivité globale, de sorte qu'un point à un lieu arbitraire de l'espace d'état est dirigé asymptotiquement vers le sous-espace glissant. Soit que l'on démontre l'attractivité locale, de sorte qu'un point dans un certain voisinage du sous-espace glissant y soit dirigé asymptotiquement. On utilise généralement une fonction de Lyapounov pour atteindre cet objectif.

En second lieu et ceci à des fins de stabilisation, on doit s'assurer qu'une fois en mode glissant, l'origine ou le point d'équilibre désiré est asymptotiquement stable. On peut là aussi utiliser une fonction de Lyapounov, cependant notons que si le système à contrôler est linéaire, en mode glissant la dynamique résultante le sera également de sorte qu'on pourra utiliser les techniques des systèmes linéaires pour déterminer la stabilité asymptotique, en étudiant par exemple les valeurs propres de l'équation caractéristique pendant le glissement.

La méthodologie de design utilisée en SSV se décompose en deux parties. D'abord choisir une surface de commutation qui engendre un mode glissant stable. Puis choisir une commande qui force le système à atteindre le mode glissant à partir d'une condition initiale arbitraire dans l'espace de phases.

Examinons le comportement possible d'un système de second ordre SSV. Les figures [1.3 et 1.4] représentent les trajectoires dans l'espace de phase des systèmes décrits par $\dot{X} = f^1(X)$ et $\dot{X} = f^2(X)$ respectivement. On combine les systèmes [fig 1.5] en commutant entre les deux lorsque la trajectoire croise les droites de commutation $S_1(X) = X_1 = 0$ et $S_2(X) = X_2 + CX_1 = 0$. La commande est choisie de sorte que lorsque $X(t)$ appartient aux régions I et III le système conjoint est $f^1(X)$ et lorsque $X(t)$ appartient aux régions II et IV le système conjoint est $f^2(X)$. Le système résultant est illustré à la figure [1.5] et possède un mode glissant sur $S_2(X) = 0$. Une remarque souvent faite dans le contexte des systèmes à structure variable, est que même si aucun système n'est stable en-soi, leur combinaison par une loi de commutation appropriée peut donner un système stable.

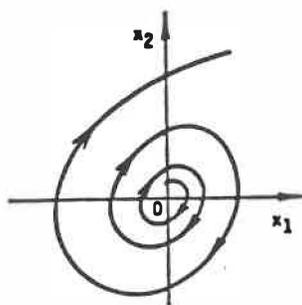


Figure 1.3: Trajectoires dans l'espace de phase pour $\dot{X} = f^1(X)$

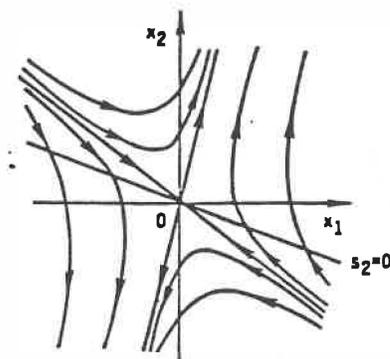


Figure 1.4: Trajectoires dans l'espace de phase pour $\dot{X} = f^2(X)$

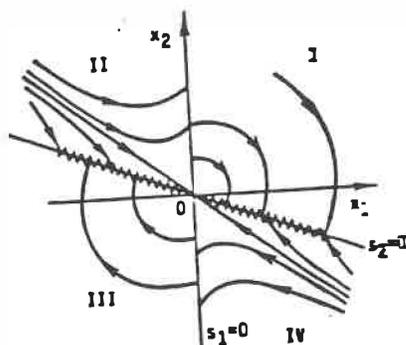


Figure 1.5: Trajectories dans l'espace de phase pour la combinaison des systèmes

1.1.6 Existence du mode de glissement

Considérons le système linéaire par rapport à la commande suivant:

$$\dot{X} = f(X) + B(X)u \quad (1.6)$$

avec $u \in \mathfrak{R}^m$ défini par:

$$u_i(X) = \begin{cases} u_i^+(X) & \text{si } S_i(X) > 0 \\ u_i^-(X) & \text{si } S_i(X) < 0 \end{cases} \quad (1.7)$$

$i = 1, \dots, m$

Chaque S_i représente une hypersurface de commutation dans l'espace des états du système. A l'aide de m entrées, on pourra espérer imposer m hypersurfaces de commutation différentes. Le mode de glissement peut se produire sur chacune des hypersurfaces $S_i = 0$ et à l'intersection de celles-ci, c'est-à-dire sur l'hypersurface d'ordre $n-m$, $S = (S_1, \dots, S_m)^T = 0$. Pour qu'un mode glissant existe sur l'hypersurface $S_i = 0$, les trajectoires des deux côtés de celle-ci doivent être dirigées vers cette hypersurface. Une condition suffisante pour que cela se produise (localement) est:

$$\lim_{S_i \rightarrow 0_+} \dot{S}_i \leq 0 \quad \text{et} \quad \lim_{S_i \rightarrow 0_-} \dot{S}_i \geq 0 \quad (1.8)$$

ou de manière équivalente:

$$\lim_{S_i \rightarrow 0} \dot{S}_i S_i \leq 0 \quad (1.9)$$

Cette condition équivaut à analyser la stabilité du sous-espace $S_i = 0$ en considérant une fonction de Lyapounov de la forme $V(S_i) = \frac{S_i^2}{2}$. $V(S_i)$ est semi-définie positive avec $V(S_i) = 0$ uniquement pour $S_i = 0$. De plus l'équation (1.9) implique que $\dot{V}(S_i) = \dot{S}_i S_i \leq 0$ dans un petit voisinage autour de $S_i = 0$. Ce qui implique que le système est localement stable par rapport au sous-espace $S_i = 0$. Afin de garantir la stabilité asymptotique autour de S_i il suffira d'imposer l'inégalité stricte suivante:

$$\lim_{S_i \rightarrow 0} \dot{S}_i S_i < 0 \quad (1.10)$$

Montrons comment utiliser la fonction de Lyapounov $V(S) = \frac{S^2}{2}$ pour imposer la stabilité asymptotique sur l'hypersurface $S(X) = 0$ dans le cas d'un système complètement linéaire à commande scalaire, c'est-à-dire pour lequel $f(X) = A$ et $B(X) = B$ dans l'équation (1.6) où B est un vecteur colonne de dimension $n \times 1$ (Astrom et Wittenmark 1989).

Soit $S(X) = P^T X = 0$, avec P un vecteur $1 \times n$, alors $V(S) = \frac{(P^T X)^2}{2}$ qui est évidemment semi-définie positive.

La condition de stabilité asymptotique est:

$$\dot{V}(S) = S \dot{S} = X^T P P^T (A + BL) X < 0 \quad (1.11)$$

où L représente la matrice $1 \times n$ de gains du système.

On tire de la dernière équation que:

$$\dot{V}(S) = X^T P (P^T A + P^T B L) X$$

$$\begin{aligned}
&= X^T P (P^T B) \left(\frac{P^T A}{P^T B} + L \right) X \\
&= S(X) (P^T B) (\alpha + L) X
\end{aligned} \tag{1.12}$$

où

$$\alpha = \frac{P^T A}{P^T B} \tag{1.13}$$

Alors les conditions suivantes sur les gains L sont suffisantes pour assurer l'existence du mode de glissement:

$$l_i < -\alpha_i \quad \text{si} \quad S(X)(P^T B)X_i > 0 \quad \text{et} \quad l_i > -\alpha_i \quad \text{si} \quad S(X)(P^T B)X_i < 0 \tag{1.14}$$

1.1.7 Équations du mode de glissement

La méthode de prolongement du contrôle équivalent est la plus directe pour trouver la solution du système pendant le glissement. Pour que le mode de glissement existe, la trajectoire doit demeurer sur l'hypersurface $S(X) = 0$, de sorte que la dérivée de S par rapport au temps est nulle pendant le glissement:

$$\dot{S}(X) = G(X)\dot{X} = G(X)f(X) + G(X)B(X)u = 0 \quad \text{avec} \quad G(X) = \text{grad}(S) \tag{1.15}$$

La solution de cette équation est le contrôle équivalent:

$$u_{eq}(X) = -(G(X)B(X))^{-1}G(X)f(X) \tag{1.16}$$

L'équation du mode de glissement est alors obtenue par substitution de cette dernière expression dans (1.6):

$$\dot{X} = f(X) - B(X)(G(X)B(X))^{-1}G(X)f(X) \quad (1.17)$$

Rappelons que cette équation est la même que celle qu'on aurait obtenue avec le prolongement de Filippov si on suppose que le système est linéaire par rapport à la commande.

1.2 Équations différentielles stochastiques

1.2.1 Équation d'Itô

Une classe importante d'équations différentielles stochastiques est la classe des équations d'Itô:

$$dX(t) = F(X(t), t)dt + G(X(t), t)dB(t) \quad (1.18)$$

où $X(t)$ est le processus vectoriel solution à n dimensions, $F(X(t), t)$ une fonction matricielle $[n \times 1]$, $G(X(t), t)$ une fonction matricielle $[n \times m]$ et $B(t)$ un processus de Wiener à m dimensions dont chacune des composantes $B_j(t)$, $j=1, 2, \dots, m$, a les propriétés suivantes:

$$\begin{aligned} E(dB_j(t)) &= E(B_j(t+dt) - B_j(t)) = 0 \\ E(dB_i(t)dB_j(t)) &= 2D_{ij}dt, t > t_0, i, j = 1, 2, \dots, m \end{aligned}$$

(1.19)

On peut montrer, moyennant certaines hypothèses de continuité, que la solution, $X(t)$, de cette équation existe, est unique, et constitue un processus stochastique markovien (Soong 1973). Si la densité de probabilité initiale $f(x_0, t_0)$ est donnée alors, les densités de probabilité futures $f(x, t)$ sont entièrement déterminées, si toutefois le processus n'est pas observé.

De plus on peut montrer que la densité de probabilité de transition satisfait à l'équation de Chapman-Kolmogorov suivante:

$$f(X_3, t_3 | X_1, t_1) = \int_{-\infty}^{\infty} f(X_3, t_3 | X_2, t_2) f(X_2, t_2 | X_1, t_1) dX_2 \quad (1.20)$$

où $f(X_{t'}, t' | X_t, t)$ est la probabilité conditionnelle que le processus ait la valeur $X_{t'}$ à l'instant t' étant donné X_t au temps t .

1.2.2 Équation de Fokker-Planck

La densité de probabilité de transition associée à ce processus markovien peut être déterminée au moyen de l'équation de Fokker-Planck ("Forward") suivante, lorsque les conditions frontières appropriées sont spécifiées.

Soit $X(t)$ le processus markovien solution de l'équation (1.18) alors l'équation de Fokker-Planck associée est donnée par (Gardiner 1980):

$$\frac{\partial f(X, t | X_0, t_0)}{\partial t} = - \sum_{j=1}^m \frac{\partial \alpha_j(X, t) f(X, t)}{\partial X_j} + 1/2 \sum_{i,j=1}^m \frac{\partial^2 \alpha_{ij} f(X, t)}{\partial X_i \partial X_j} \quad (1.21)$$

avec :

$$\begin{aligned}
\alpha_j(X, t) &= \lim_{dt \rightarrow 0} \frac{1}{dt} E([X_j(t+dt) - X_j(t)] | X(t) = x) \\
\alpha_{ij}(X, t) &= \lim_{dt \rightarrow 0} \frac{1}{dt} E([X_i(t+dt) - X_i(t)][X_j(t+dt) - X_j(t)] | X(t) = x)
\end{aligned}
\tag{1.22}$$

où les coefficients α_j et α_{ij} sont obtenus en écrivant l'équation d'Itô sous forme incrémentale:

$$\begin{aligned}
dX_j(t) &= F_j(X(t), t)dt + \sum_{k=1}^m G_{jk}(X(t), t)dB_k(t) + o(dt) \\
dX_i(t)dX_j(t) &= \sum_{k,l=1}^m G_{ik}(X(t), t)G_{jl}(X(t), t)dB_k(t)dB_l(t) + \\
&\quad F_i(X(t), t) \sum_{k=1}^m G_{jk}(X(t), t)dB_k(t) + \\
&\quad F_j(X(t), t) \sum_{k=1}^m G_{ik}(X(t), t)dB_k(t) + \\
&\quad F_i(X(t), t)F_j(X(t), t)(dt)^2 + o(dt)
\end{aligned}
\tag{1.23}$$

où

$$\lim_{dt \rightarrow 0} \frac{o(dt)}{dt} = 0$$

En utilisant l'équation (1.19) et en notant que $dB(t)$ est un processus indépendant de $X(t)$ on obtient:

$$\begin{aligned}
\alpha_j(X, t) &= \lim_{dt \rightarrow 0} \frac{1}{dt} E([X_j(t+dt) - X_j(t)] | X(t) = x) \\
&= F_j(X(t), t)
\end{aligned}
\tag{1.24}$$

$$\begin{aligned}
\alpha_{ij}(X, t) &= \lim_{dt \rightarrow 0} \frac{1}{dt} E([X_i(t+dt) - X_i(t)][X_j(t+dt) - X_j(t)] | X(t) = x) \\
&= 2 \sum_{k,l=1}^m D_{kl} G_{ik}(X(t), t) G_{jl}(X(t), t) \\
&= 2(GDG^T)_{ij} \quad i, j = 1, 2, \dots, n
\end{aligned} \tag{1.25}$$

avec D la matrice de covariance du bruit.

L'équation de Fokker-Planck résultante est alors:

$$\frac{\partial f}{\partial t} = - \sum_{j=1}^m \frac{\partial F_j(X, t) f}{\partial X_j} + \sum_{i,j=1}^m \frac{\partial^2 [GDG^T]_{ij} f}{\partial X_i \partial X_j} \tag{1.26}$$

1.2.3 Courant de probabilité

Remarquons que l'équation de Fokker-Planck peut s'écrire comme suit (Fuller 1970):

$$\frac{\partial f}{\partial t} = - \sum_{j=1}^m \frac{\partial J_j(X, t)}{\partial X_j} = -div(\vec{J}) \tag{1.27}$$

où \vec{J} est la densité de courant de probabilité dont les composantes sont:

$$J_j = F_j(X, t) f - \sum_{i=1}^m \frac{\partial [GDG^T]_{ji} f}{\partial X_i} \tag{1.28}$$

Envisagée sous cet angle l'équation de Fokker-Planck est l'expression d'un principe de conservation: celui du courant (flux) de probabilité. Ce courant permettra de dériver de manière naturelle un certain nombre de conditions frontières. On montrera plus loin [Annexe 1] que cette formulation en terme de divergence d'un flux de probabilité est indépendante de l'orthogonalité du système de coordonnées choisi.

1.2.4 Bruit blanc gaussien et équations d'Itô

Par la suite on sera intéressé par la forme suivante de l'équation (1.18):

$$\dot{X} = F(X(t), t) + G(X(t), t)W(t) \quad (1.29)$$

avec W un vecteur de bruits blancs gaussiens.

Le lecteur averti notera que cette équation n'a pas de sens par rapport aux règles ordinaires du calcul différentiel. La dérivée \dot{X} n'existe pas. Cela est dû au fait que le bruit blanc gaussien n'est pas un processus stochastique (Kozin 1969). La représentation formelle d'un bruit blanc gaussien peut se faire de la façon suivante:

$$W(t) = \frac{dB(t)}{dt} \quad (1.30)$$

où $B(t)$ est un processus de Wiener, i.e. un processus gaussien à moyenne nulle avec incréments indépendants et stationnaires pour lequel:

$$E(B(t)B(s)) = \min(s, t) \quad \text{et} \quad P(B(0) = 0) = 1 \quad (1.31)$$

Wiener a montré que presque toutes les réalisations de ce processus sont des fonctions continues différentiables en aucun point.

C'est Itô qui a donné une interprétation rigoureuse à l'équation (1.29) en utilisant l'intégrale stochastique généralisée. Depuis les travaux d'Itô l'équation (1.29) s'écrit en utilisant la relation formelle (1.30) sous forme de différentielle comme en (1.18):

$$dX(t) = F(X(t), t)dt + G(X(t), t)dB(t) \quad (1.32)$$

On donne un sens rigoureux à la dernière équation en utilisant l'équation intégrale stochastique suivante:

$$x(t) - x(t_0) = \int_{t_0}^t F(X(\tau), \tau) d\tau + \int_{t_0}^t G(X(\tau), \tau) dB(\tau) \quad (1.33)$$

Sous certaines conditions de Lipschitz sur F et G et en utilisant des bornes sur les coefficients de F et G , $X(t)$ devient un processus markovien avec des réalisations qui sont des fonctions continues (Itô, 1961). Dans ce travail lorsqu'on écrit un modèle d'état sous la forme (1.29) on fera implicitement référence à son expression sous la forme de (1.32).

1.3 Conditions frontières

L'équation de Fokker-Planck est une équation aux dérivées partielles de second ordre parabolique. Pour solutionner cette équation on a besoin de conditions initiales et des conditions à l'extrémité de l'intervalle où X est contraint de se propager. D'autre part on admet qu'il existe des discontinuités paramétriques dans l'équation d'Itô à l'interface de certains domaines connexes. Examinons ici dans toute leur généralité la nature des conditions frontières qu'on rencontrera par la suite.

1.3.1 Paroi réfléchissante

Soit la situation où le mouvement étudié est contraint à rester dans une région R . Le flux net de probabilité à travers S , la frontière de R , doit être nul. La condition est alors que:

$$\vec{N} \cdot \vec{J}(X, t) = 0, \quad X \in S \quad (1.34)$$

où \vec{N} est la normale à S .

1.3.2 Paroi absorbante

Soit maintenant la situation où lorsque le mouvement atteint S il ne peut revenir en arrière alors, pour un processus de diffusion on peut montrer que (Karlin et Taylor 1981):

$$f(X, t) = 0 \quad X \in S \quad (1.35)$$

Dans ce cas on dit que S est une paroi absorbante.

1.3.3 Paroi de discontinuité

Ici on admet que les coefficients F_j et G_j peuvent être discontinus sur la paroi S . Puisque la probabilité est conservée, la composante du courant normale à cette paroi doit être continue à travers S :

$$\lim_{X \rightarrow S^+} \vec{N} \cdot \vec{J}(X, t) = \lim_{X \rightarrow S^-} \vec{N} \cdot \vec{J}(X, t) \quad (1.36)$$

Nous admettons également la continuité de la densité à travers la discontinuité:

$$\lim_{X \rightarrow S^+} f(X, t) = \lim_{X \rightarrow S^-} f(X, t) \quad (1.37)$$

On note que la définition du courant de probabilité (1.28) n'implique pas que les dérivées de $f(X, t)$ soient nécessairement continues en S .

Après s'être munis de ces notions préliminaires nous sommes maintenant en mesure de présenter notre modèle des régulateurs à structure variable en présence de bruit, ce qui fait l'objet du prochain chapitre.

Chapitre 2

Modèle d'un régulateur à structure variable en présence de bruit

2.1 Origine du modèle

Notre modélisation des régulateurs à structures variables en présence de bruits repose essentiellement sur l'utilisation des modèles stochastiques hybrides tels que développés dans le cadre de la modélisation stochastique des charges électriques par Malhamé et Chong, (Malhamé et Chong 1983, 1985, 1988). Rappelons ici quelques éléments de ces travaux.

Un modèle stochastique des charges électriques associées à des éléments de chauffage dont les états d'opération sont contrôlés par un thermostat est donné comme suit:

$$Cdx(t) = -\alpha(x(t) - X_a(t))dt + p(t)m(t)b(t)dt + dv'(t) \quad (2.1)$$

avec:

C : la capacité thermique moyenne de la maison

α : le taux de perte de chaleur moyen

$x(t)$: la température moyenne de la maison

$X_a(t)$: la température ambiante

$p(t)$: le taux de chaleur fourni par l'élément chauffant

$m(t)$: l'état discret d'opération de l'élément (0 pour "OFF" et 1 pour "ON")

$b(t)$: l'état du système de puissance (0 pour "OFF" et 1 pour "ON")

$v'(t)$: un processus de bruit de Wiener

En divisant cette équation par C on obtient:

$$dx(t) = -a(x(t) - X_a(t))dt + Rm(t)b(t)dt + dv(t) \quad (2.2)$$

avec $v(t)$ un processus de Wiener de variance σ .

L'évolution de la variable $m(t)$ est gouvernée par un thermostat qui a un point de bascule à x_- et une bande morte ($x_+ - x_-$). $m(t)$ passe de 1 à 0 quand $x(t)$ atteint x_+ et de 0 à 1 quand $x(t)$ atteint x_- .

Le modèle des transitions de $m(t)$ est $m(t + dt) = \pi + m(t)$ avec:

$$\pi(x, m, x_+, x_-) = \begin{cases} 0 & \text{si } x_- < x < x_+ \\ -m & \text{si } x > x_+ \\ (1 - m) & \text{si } x < x_- \end{cases} \quad (2.3)$$

On peut alors considérer que le système oscille entre les deux structures suivantes ($b(t)=1$):

$$dx(t) = -a(x(t) - X_a(t))dt + Rdt + dv(t) \quad (2.4)$$

$$dx(t) = -a(x(t) - X_a(t))dt + dv(t) \quad (2.5)$$

selon que $m(t)=1$ ou que $m(t)=0$ respectivement.

Ce type de système peut être considéré comme composé de deux sous-systèmes interconnectés. Une partie linéaire caractérisée par l'état stochastique continu $x(t)$ dont l'évolution dépend de $m(t)$. Chaque valeur de $m(t)$ induisant une structure particulière. Une partie non-linéaire $m(t)$ qui est un processus de sauts dépendant de l'état $x(t)$ de façon déterministe. Le processus vectoriel $[x(t), m(t)]$ constitue alors le processus de Markov hybride.

On est naturellement conduit à associer à chacune des structures une fonction de densité comme suit:

$$\begin{aligned} f_1(\lambda, t)d\lambda &= Pr[\lambda < x(t) \leq \lambda + d\lambda \cap m(t) = 1] \\ f_0(\lambda, t)d\lambda &= Pr[\lambda < x(t) \leq \lambda + d\lambda \cap m(t) = 0] \end{aligned} \quad (2.6)$$

On peut alors utiliser les équations de Fokker-Planck pour modéliser l'évolution de la fonction de densité associée à chacune des structures. Une question particulièrement importante est la nature du couplage de ces équations aux dérivées partielles. Ce couplage est produit par le cycle d'échange de probabilité entre les deux structures. On parlera alors du mécanisme de réinsertion de probabilité d'une structure à l'autre.

A titre d'exemple d'utilisation des notions introduites dans le chapitre 1, dérivons les équations de Fokker-Planck de ce système couplé.

En utilisant la version scalaire de l'équation (1.27) on obtient:

$$\frac{\partial f_1}{\partial t}(\lambda, t) = \frac{\partial}{\partial \lambda} [(a(\lambda - x_a(t)) - b(t)R)f_1(\lambda, t)] + \frac{\sigma^2}{2} \frac{\partial^2}{\partial \lambda^2} f_1(\lambda, t) \quad (2.7)$$

pour la structure associée à $m(t)=1$.

Et

$$\frac{\partial f_0}{\partial t}(\lambda, t) = \frac{\partial}{\partial \lambda} [(a(\lambda - x_a(t)))f_0(\lambda, t)] + \frac{\sigma^2}{2} \frac{\partial^2}{\partial \lambda^2} f_0(\lambda, t) \quad (2.8)$$

pour la structure associée à $m(t)=0$.

Puis en utilisant l'équation (1.28) on définit les courants de probabilité associés à chacune des distributions comme suit:

$$J_1 = -[(a(\lambda - x_a(t)) - b(t)R)f_1(\lambda, t)] - \frac{\sigma^2}{2} \frac{\partial}{\partial \lambda} f_1(\lambda, t) \quad (2.9)$$

$$J_0 = -[(a(\lambda - x_a(t)))f_0(\lambda, t)] - \frac{\sigma^2}{2} \frac{\partial}{\partial \lambda} f_0(\lambda, t) \quad (2.10)$$

En se référant à la figure 2.1, x_+ constitue une paroi absorbante pour la distribution f_1 et x_- une paroi de réinsertion pour f_1 (au sens où l'évolution locale de la densité de probabilité f_1 est perturbée par la réinsertion de probabilité venant de la distribution f_0). De même x_+ constitue une paroi de réinsertion pour la distribution f_0 et x_- une paroi absorbante pour f_0 .

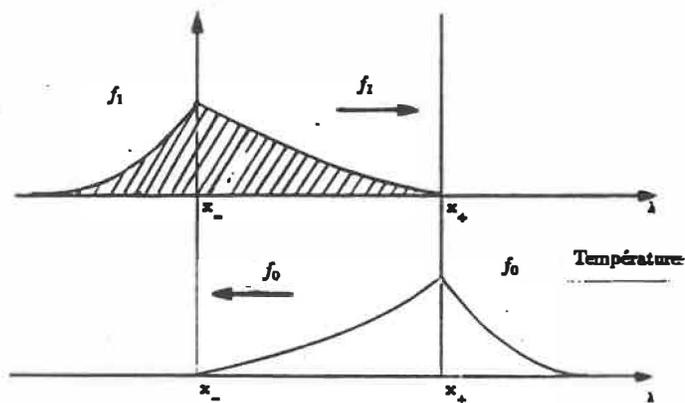


Figure 2.1: Fonctions de densité du modèle hybride

Les conditions frontières suivantes sont alors utilisées:

Paroi absorbante:

$$f_{1b}(x_+, t) = f_{0b}(x_-, t) = 0$$

Conditions à l'infini

$$f_{1a}(-\infty, t) = f_{0c}(+\infty, t) = 0$$

Conditions de continuité sur les parois de réinsertion

$$f_{1a}(x_-, t) = f_{1b}(x_-, t)$$

$$f_{0b}(x_+, t) = f_{0c}(x_+, t)$$

Conditions de conservation du courant sur les parois de réinsertion

En utilisant le principe de conservation des courants on obtient:

$$J_{0c}(x_+, t) = J_{0b}(x_+, t) + J_{1b}(x_+, t)$$

$$J_{1b}(x_-, t) = J_{1a}(x_-, t) - J_{0b}(x_-, t)$$

Ce qui en utilisant le fait que $f_{1b}(x_+, t) = 0$ et $f_{0b}(x_-, t) = 0$ donne:

$$-\frac{\partial}{\partial \lambda} f_{0c}(x_+, t) = -\frac{\partial}{\partial \lambda} f_{0b}(x_+, t) - \frac{\partial}{\partial \lambda} f_{1b}(x_+, t) \quad (2.11)$$

et

$$-\frac{\partial}{\partial \lambda} f_{1b}(x_-, t) = +\frac{\partial}{\partial \lambda} f_{0b}(x_-, t) - \frac{\partial}{\partial \lambda} f_{1a}(x_-, t) \quad (2.12)$$

Ces équations permettent de caractériser complètement la nature du couplage entre les deux distributions.

2.2 Relation avec les régulateurs à structure variable

Soit maintenant le système à structure variable univarié suivant:

$$dx(t) = ax(t)dt + bu(t)dt + dv(t) \quad (2.13)$$

avec

$$u = \psi x \quad (2.14)$$

où

$$\psi = \alpha(1 + m(t))/2 + \beta(1 - m(t))/2 \quad (2.15)$$

et

$$S(x) = x = 0 \quad (2.16)$$

On définit un modèle de transition $m(t + dt) = \pi + m(t)$ avec:

$$\pi(x, m, \varepsilon_+, \varepsilon_-) = \begin{cases} 0 & \text{si } S(x) - \varepsilon_- < x < S(x) + \varepsilon_+ \\ -(1 + m) & \text{si } x > S(x) + \varepsilon_+ \\ (1 - m) & \text{si } x < S(x) - \varepsilon_- \end{cases} \quad (2.17)$$

où $S(x)$ est la paroi de commutation et $m(t) \in [-1, 1]$ les états discrets caractérisant les structures entre lesquelles le système oscille. Les paramètres ε_+ et ε_- représentent les non-idéalités de la commutation.

Par analogie avec la modélisation stochastique des éléments chauffant on associe une fonction de densité à chacune des structures entre lesquelles le système oscille soit:

$$\begin{aligned} f_1(\lambda, t)d\lambda &= Pr[\lambda < x(t) \leq \lambda + d\lambda \cap m(t) = 1] \\ f_{-1}(\lambda, t)d\lambda &= Pr[\lambda < x(t) \leq \lambda + d\lambda \cap m(t) = -1] \end{aligned} \quad (2.18)$$

De la même façon qu'auparavant on pourrait en utilisant les conditions frontières appropriées dériver un système d'équations de Fokker-Planck couplées. On remarque l'analogie avec le modèle stochastique des éléments chauffants.

2.3 Modèle dans le plan de phases

Par la suite on s'intéressera au système scalaire dans l'espace de phases suivant:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \omega(t) \quad (2.19)$$

avec $u(t) = \psi_1(x_1)x_1 + \psi_2(x_2)x_2$

et $\psi_1(x_1) = \alpha_1\{1 + m(t)\text{sgn}(x_1)\}/2 + \beta_1\{1 - m(t)\text{sgn}(x_1)\}/2$

et $\psi_2(x_2) = \alpha_2\{1 + m(t)\text{sgn}(x_2)\}/2 + \beta_2\{1 - m(t)\text{sgn}(x_2)\}/2$

où

$$\text{sgn}(x) = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x < 0 \end{cases} \quad (2.20)$$

Le modèle peut alors s'écrire:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a + B_1\psi_1 & b + B_1\psi_2 \\ c + B_2\psi_1 & d + B_2\psi_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \omega(t) \quad (2.21)$$

La loi de transition de $m(t)$ est $m(t + dt) = \pi(z, \varepsilon) + m(t)$ avec:

$$\pi(z, \varepsilon) = \begin{cases} 0 & \text{si } -\varepsilon < z < \varepsilon \\ -(1 + m) & \text{si } z < -\varepsilon \\ (1 - m) & \text{si } z > \varepsilon \end{cases} \quad (2.22)$$

où $z = S(X) = P^T X$.

Remarquons que ce modèle de transition implique une non-idéalité de type hystérese au niveau de la commutation, ce qu'illustre la figure [2.2].

Les valeurs de $m(t) \in [1, -1]$ caractérisent alors deux ensembles de structures:

Un ensemble où $m(t) = 1$ défini par la loi de retour d'état:

$$u(t) = \psi_1(x_1)x_1 + \psi_2(x_2)x_2$$

$$\begin{aligned}
 &\text{avec } \psi_1(x_1) = \alpha_1 \{1 + \text{sgn}(x_1)\} / 2 + \beta_1 \{1 - \text{sgn}(x_1)\} / 2 \\
 &\text{et } \psi_2(x_2) = \alpha_2 \{1 + \text{sgn}(x_2)\} / 2 + \beta_2 \{1 - \text{sgn}(x_2)\} / 2
 \end{aligned}
 \tag{2.23}$$

Et un autre ensemble où $m(t) = -1$ défini par la loi de retour d'état:

$$\begin{aligned}
 u(t) &= \psi_1(x_1)x_1 + \psi_2(x_2)x_2 \\
 &\text{avec } \psi_1(x_1) = \alpha_1 \{1 - \text{sgn}(x_1)\} / 2 + \beta_1 \{1 + \text{sgn}(x_1)\} / 2 \\
 &\text{et } \psi_2(x_2) = \alpha_2 \{1 - \text{sgn}(x_2)\} / 2 + \beta_2 \{1 + \text{sgn}(x_2)\} / 2
 \end{aligned}
 \tag{2.24}$$

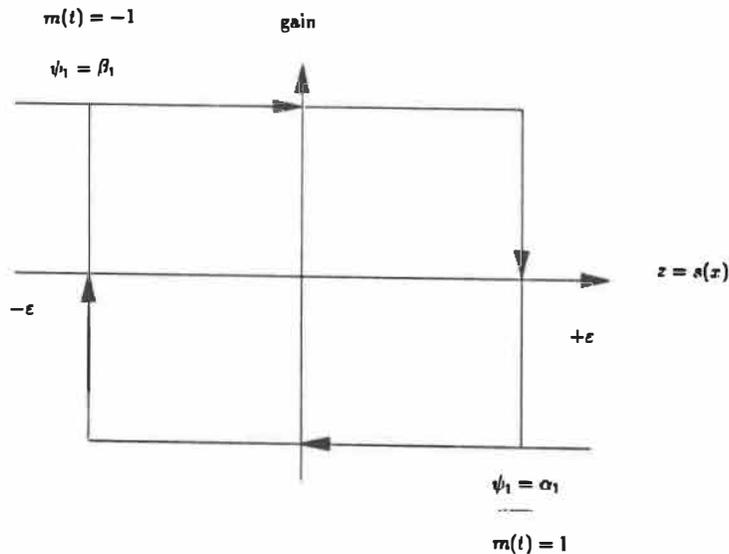


Figure 2.2: Représentation du modèle de transition

On associe à chacun de ces ensembles de structures une fonction de densité comme suit:

$$\begin{aligned}
 f_1(\lambda_1, \lambda_2, t) d\lambda_1 d\lambda_2 &= Pr[\lambda_1 < x_1(t) \leq \lambda_1 + d\lambda_1, \lambda_2 < x_2(t) \leq \lambda_2 + d\lambda_2, \cap m(t) = 1] \\
 f_{-1}(\lambda_1, \lambda_2, t) d\lambda_1 d\lambda_2 &= Pr[\lambda_1 < x_1(t) \leq \lambda_1 + d\lambda_1, \lambda_2 < x_2(t) \leq \lambda_2 + d\lambda_2, \cap m(t) = -1]
 \end{aligned}
 \tag{2.25}$$

L'évolution des densités de probabilité dans chacun des ensembles de structures est alors décrite au moyen d'un système d'équations de Fokker-Planck couplées comme dans le cas du modèle univarié que l'on vient de présenter. Remarquons qu'on parle d'ensemble de structures puisque pour $m(t)$ donné la valeur du retour d'état peut varier selon $sgn(x_1)$ et $sgn(x_2)$ définissant des structures différentes au sein d'un même ensemble de structures qu'on associe à une même densité de probabilité. On distinguera ce type de phénomène avec le changement d'ensemble de structures défini par une transition de $m(t)$ en parlant de discontinuité paramétrique dans le premier cas et de discontinuité de structure dans le second cas.

La figure 2.3 montre quelles sont les régions correspondant au deux sous-ensembles de structures ($m(t) \in [1, -1]$) de notre modèle et parmi elles les différentes structures associées selon les valeurs de $sgn(x_1)$ et $sgn(x_2)$.

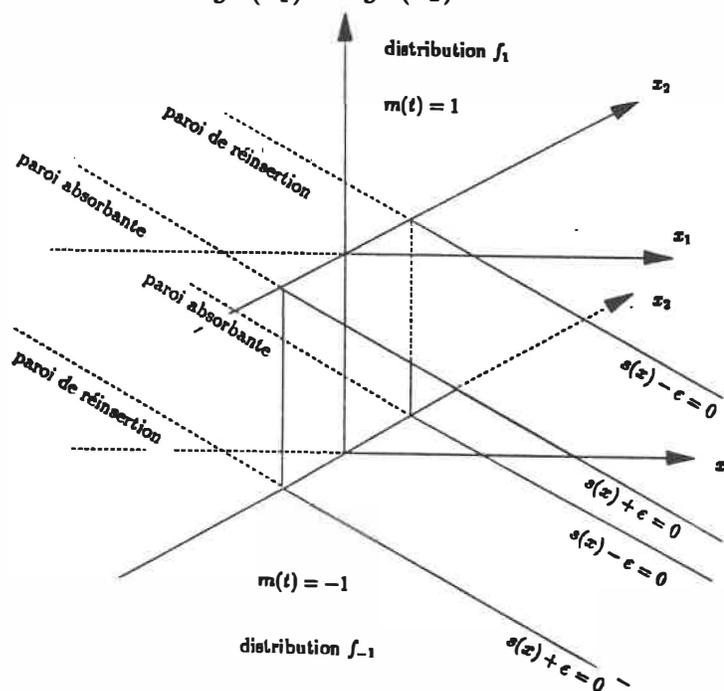


Figure 2.3: Régions correspondant aux deux sous-ensembles de structures $m(t)=1,-1$

Soit pour l'ensemble $m(t)=1$, les régions 1,2 et 6 définies par:

Région 1: $sgn(x_1) > 0, sgn(x_2) > 0, \psi_1 = \alpha_1, \psi_2 = \alpha_2$.

Région 2: $sgn(x_1) > 0, sgn(x_2) < 0, \psi_1 = \alpha_1, \psi_2 = \beta_2$.

Région 6: $sgn(x_1) < 0, sgn(x_2) > 0, \psi_1 = \beta_1, \psi_2 = \alpha_2$.

Et pour l'ensemble $m(t)=-1$, les régions 3,4 et 5 définies par:

Région 3: $sgn(x_1) > 0, sgn(x_2) < 0, \psi_1 = \beta_1, \psi_2 = \alpha_2$.

Région 4: $sgn(x_1) < 0, sgn(x_2) < 0, \psi_1 = \alpha_1, \psi_2 = \alpha_2$.

Région 5: $sgn(x_1) < 0, sgn(x_2) > 0, \psi_1 = \alpha_1, \psi_2 = \beta_2$.

Par la suite en vue de simplifier l'algorithme de simulation on écrira notre modèle dans un repère non-orthogonal adapté à la géométrie de notre problème. Puis on dérivera les conditions frontières et les mécanismes de couplage appropriés à la simulation du système dans ce nouveau repère.

2.4 Modèle dans un repère non-orthogonal

Afin de simplifier l'algorithme de simulation écrivons le modèle dans le repère non-orthogonal (θ_1, θ_2) représenté à la figure 2.4.

Soit $\theta = \tan^{-1}(c)$ où c est la pente de la droite de commutation.

La matrice de transformation de (x_1, x_2) à (θ_1, θ_2) est:

$$M(\theta) = \begin{bmatrix} 1/\cos \theta & 0 \\ \tan \theta & 1 \end{bmatrix} \quad (2.26)$$

et

$$M^{-1}(\theta) = \begin{bmatrix} \cos \theta & 0 \\ -\sin \theta & 1 \end{bmatrix} \quad (2.27)$$

De sorte que:

$$\Theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = M(\theta) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad (2.28)$$

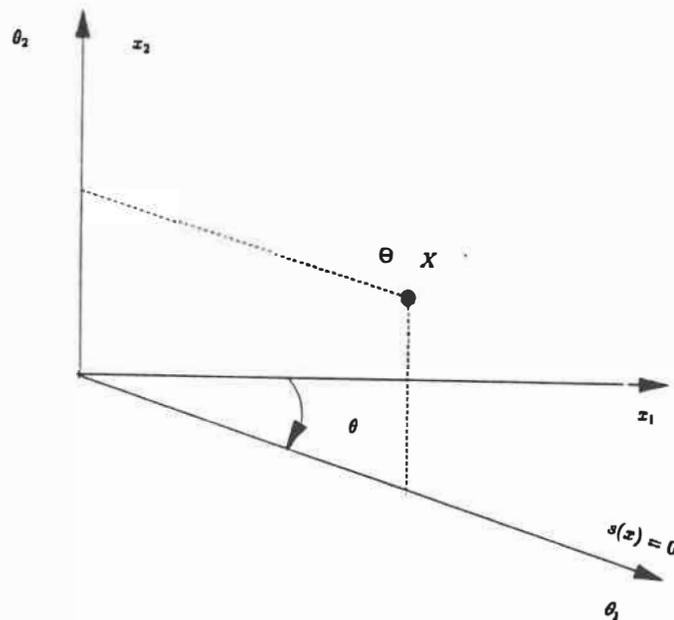


Figure 2.4: Systèmes de coordonnées utilisés.

De $\Theta = MX$ et $\dot{X} = AX + K\omega$ on tire que $\dot{\Theta} = MAM^{-1}\Theta + MK\omega$.

En définissant les variables intermédiaires suivantes:

$$\begin{aligned} \Delta_1 &= a + B_1\psi_1 & \Delta_2 &= b + B_1\psi_2 \\ \Delta_3 &= c + B_2\psi_1 & \Delta_4 &= d + B_2\psi_2 \end{aligned}$$

Alors en définissant $T = MAM^{-1}$ et $G = MK$ on obtient:

$$T = \begin{bmatrix} \Delta_1 - \tan \theta \Delta_2 & 1 / \cos \theta \Delta_2 \\ (\Delta_1 - \Delta_4) \sin \theta - \Delta_2 \tan \theta + \Delta_3 \cos \theta & \Delta_4 + \Delta_2 \tan \theta \end{bmatrix} \quad (2.29)$$

et

$$G = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} K_1 / \cos \theta \\ K_1 \tan \theta + K_2 \end{bmatrix} \quad (2.30)$$

L'équation d'état dans (θ_1, θ_2) s'écrit donc:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} + G\omega(t) \quad (2.31)$$

avec $E(\omega(t)) = 0$ et $E(\omega(t)^2) = 1\delta(t)$

où:

$$\Gamma_1 = T_{11}\theta_1 + T_{12}\theta_2$$

$$\Gamma_2 = T_{21}\theta_1 + T_{22}\theta_2$$

En utilisant l'expression (1.27), l'équation de Fokker-Planck dans le système de coordonnées (θ_1, θ_2) est:

$$\frac{\partial f}{\partial t} = - \sum_{j=1}^2 \frac{\partial \Gamma_j f}{\partial \theta_j} + \sum_{i,j=1}^2 \frac{\partial^2 [GDG^T]_{ij} f}{\partial \theta_i \partial \theta_j} \quad (2.32)$$

où D est la matrice de covariance du bruit.

Soit:

$$(GDG^T) = \begin{bmatrix} Dg_1^2 & Dg_1g_2 \\ Dg_1g_2 & Dg_2^2 \end{bmatrix} = \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix} \quad (2.33)$$

Notons que le système (2.31) est équivalent à :

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} = \begin{bmatrix} \Gamma_1 \\ \Gamma_2 \end{bmatrix} + \begin{bmatrix} \omega_1(t) \\ \omega_2(t) \end{bmatrix} \quad (2.34)$$

avec:

$$E(\omega_1(t)) = 0 \quad \text{et} \quad E(\omega_1(t)^2) = \xi_{11}\delta(t) = g_1^2\delta(t)$$

$$E(\omega_2(t)) = 0 \quad \text{et} \quad E(\omega_2(t)^2) = \xi_{22}\delta(t) = g_2^2\delta(t)$$

$$E(\omega_1(t)\omega_2(t)) = E(\omega_2(t)\omega_1(t)) = \xi_{12}\delta(t) = \xi_{21}\delta(t) = g_1g_2\delta(t)$$

En développant l'expression (2.32) l'équation de Fokker-Planck s'écrit comme suit:

$$\frac{\partial f}{\partial t} = -\frac{\partial \Gamma_1 f}{\partial \theta_1} - \frac{\partial \Gamma_2 f}{\partial \theta_2} + 2\xi_{12} \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} + \xi_{11} \frac{\partial^2 f}{\partial \theta_1^2} + \xi_{22} \frac{\partial^2 f}{\partial \theta_2^2} \quad (2.35)$$

où l'on reconnaît la forme parabolique.

Comme on l'a vu dans le premier chapitre, un résultat capital est que l'équation de Fokker-Planck peut s'exprimer à l'aide de la divergence du courant de probabilité \vec{J} comme suit :

$$\frac{\partial f}{\partial t} + \sum_{j=1}^2 \frac{\partial J_j}{\partial \theta_j} = 0 \quad (2.36)$$

Remarquons qu'on peut montrer que l'expression de l'équation de Fokker-Planck comme divergence d'un courant de probabilité est toujours valable même si le système de coordonnées n'est pas orthogonal [voir Annexe 1].

Les composantes de la densité de courant probabilité sont alors fournies par l'équation (1.28) comme suit:

$$\begin{aligned} J_1 &= T_{11}\theta_1 + T_{12}\theta_2 - \xi_{11} \frac{\partial f}{\partial \theta_1} - \xi_{12} \frac{\partial f}{\partial \theta_2} \\ J_2 &= T_{21}\theta_1 + T_{22}\theta_2 - \xi_{21} \frac{\partial f}{\partial \theta_1} - \xi_{22} \frac{\partial f}{\partial \theta_2} \end{aligned} \quad (2.37)$$

2.5 Nature des équations à la paroi de réinsertion de probabilité

On s'intéresse ici à la forme de l'équation de Fokker-Planck sur la paroi de réinsertion. Rappelons que sur cette paroi l'évolution de la densité de probabilité $f_1, (f_{-1})$ est perturbée localement par une réinsertion de la probabilité provenant de $f_{-1}, (f_1)$ et résultant d'une transition de $m(t)$.

En effet lorsqu'une réalisation du processus stochastique croise la paroi d'absorption elle se trouve réinsérée dans l'autre ensemble de structures au point

correspondant sur la paroi de réinsertion, ce phénomène engendre une variation locale de la densité de probabilité associée à la structure d'accueil, ce qu'on représente à la figure 2.5.

On montrera que l'équation de Fokker-Planck résultante à la paroi de réinsertion pourra être représentée comme suit:

$$\frac{\partial f}{\partial t} = -\text{div}(\vec{J}) + \vartheta(x, t) \quad (2.38)$$

Les prochaines sections permettront de déterminer $\vartheta(x, t)$.

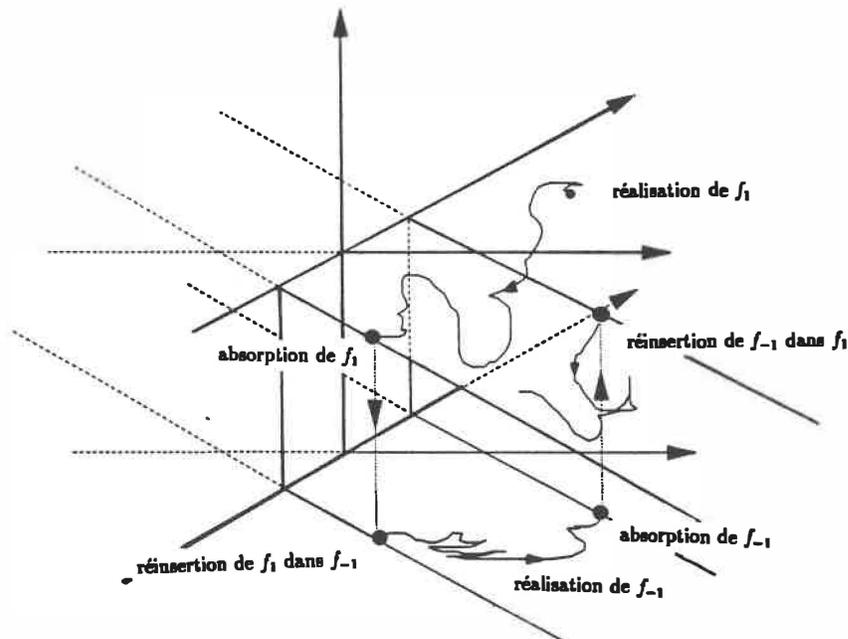


Figure 2.5: Réinsertion et absorption de probabilités

2.5.1 Schéma de réinsertion des probabilités

On décompose ici le problème de la réinsertion des probabilités en deux parties distinctes. On examine d'abord la nature des équations qui régissent la perte totale de probabilité d'une région via le mécanisme d'absorption. Puis on formule un schéma de

réinsertion de cette probabilité dans l'autre région. Finalement on pourra constater l'identité des quantités absorbées (perdues) et réinsérées.

2.5.2 Perte de probabilité à la paroi absorbante

La première question consiste à déterminer une expression pour la perte de probabilité à la paroi absorbante qui est représentée à la figure 2.6.

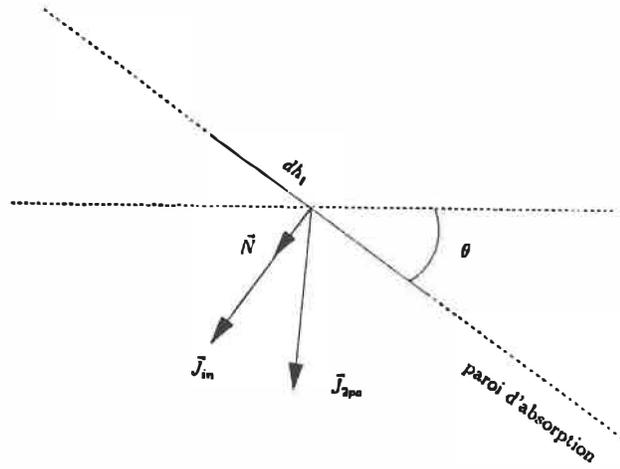


Figure 2.6: Mécanisme de pertes de probabilité

Soit \vec{N} la normale unitaire à la droite d'absorption dans le plan (x_1, x_2) :

$$\vec{N} = (1, 1/c) = (1, 1/\tan \theta) = (\sin \theta, \cos \theta)$$

Soit $\vec{J}_{pa}(\theta)$ le vecteur de densité de flux de probabilité dans le plan à la paroi absorbante (θ_1, θ_2) .

Appliquons la transformée $M^{-1}(\theta)$ à $\vec{J}_{pa}(\theta)$ de sorte à obtenir les composantes du courant dans le plan (x_1, x_2) soit:

$$\vec{J}_{pa}(X) = \begin{bmatrix} J_{1x1} \\ J_{2x2} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ -\sin \theta & 1 \end{bmatrix} \begin{bmatrix} J_{1pa} \\ J_{2pa} \end{bmatrix} = \begin{bmatrix} J_{1pa} \cos \theta \\ -J_{1pa} \sin \theta + J_{2pa} \end{bmatrix} \quad (2.39)$$

Maintenant évaluons le produit scalaire $\vec{N} \cdot \vec{J}_{pa}(X)$:

$$(J_{1pa} \cos \theta, -J_{1pa} \sin \theta + J_{2pa}) \cdot (\sin \theta, \cos \theta) = J_{2pa} \cos \theta$$

Dans (x_1, x_2) le vecteur résultant est:

$$J_{2pa} \cos \theta (\sin \theta, \cos \theta) = (J_{2pa} \cos \theta \sin \theta, J_{2pa} \cos^2 \theta) \quad (2.40)$$

Maintenant revenons au plan (θ_1, θ_2) par la transformée $M(\theta)$:

$$\vec{J}_{in} = \begin{bmatrix} J_{1in}(\theta) \\ J_{2in}(\theta) \end{bmatrix} = \begin{bmatrix} 1/\cos \theta & 0 \\ \tan \theta & 1 \end{bmatrix} \begin{bmatrix} J_{2pa} \cos \theta \sin \theta \\ J_{2pa} \cos^2 \theta \end{bmatrix} = \begin{bmatrix} J_{2pa} \sin \theta \\ J_{2pa} \end{bmatrix} \quad (2.41)$$

Ce dernier vecteur représente le courant réinséré dans l'autre distribution. Il servira à écrire les équations de couplage du système d'équations de Fokker-Planck.

Mais d'abord calculons la probabilité perdue à la paroi absorbante. Pour ce faire remarquons que :

$$\vec{N} \cdot \vec{J}_{pa} = J_{2pa} \cos \theta = \frac{Prob}{(\text{unite de longueur} \times \text{unite de temps})} \quad (2.42)$$

La probabilité qui s'échappe par longueur dh_1 pendant l'intervalle de temps dt est donc $J_{2pa} \cos \theta dt dh_1$.

Et puisque la densité est :

$$\frac{Prob}{dh_1 dh_2 \cos \theta} \quad (2.43)$$

on obtient que la densité qui s'échappe par longueur dh_1 sur l'intervalle dt est:

$$\Delta f_{elem} = \frac{J_{2pa} \cos \theta dt dh_1}{dh_1 dh_2 \cos \theta} = \frac{J_{2pa} dt}{dh_2} \quad (2.44)$$

La perte totale de densité (en considérant le flux constant par rapport au temps durant l'intervalle dt) est:

$$\Delta f_{tot} = dt \int_0^l \frac{J_{2pa}}{dh_2} dh_1 \quad (2.45)$$

où l représente la longueur de la paroi absorbante, théoriquement infinie, mais qui dans notre cas est bornée par des parois réfléchissantes.

2.5.3 Réinsertion de probabilités à la paroi de réinsertion

La deuxième question consiste à dégager le mécanisme par lequel la probabilité absorbée est réinsérée dans l'autre distribution.

On considère que localement le flux conserve à la réinsertion la direction qu'il avait au moment de son absorption. La paroi absorbante permet donc de déterminer de manière univoque l'orientation du flux lorsqu'il est réinséré dans la distribution d'accueil. La figure 2.7 représente le phénomène à l'étude.

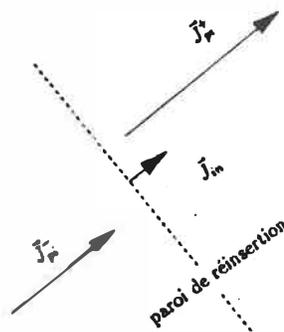


Figure 2.7: Mécanisme de réinsertion de probabilité

En appliquant la conservation des courants on peut écrire:

$$\vec{J}_{pi}^+ - \vec{J}_{pi}^- = \vec{J}_{in} \quad (2.46)$$

où p_i représente la paroi de réinsertion.

Et

$$\vec{J}_{pi}^+ = \lim_{\varepsilon \rightarrow 0} \vec{J}_{pi}(x + \varepsilon)$$

$$\vec{J}_{pi}^- = \lim_{\varepsilon \rightarrow 0} \vec{J}_{pi}(x - \varepsilon)$$

avec x la paroi de réinsertion, ε un réel strictement positif.

Ce qui en reprenant les expressions des courants obtenues précédemment i.e:

$$J_{2pi}^+ = (T_{21}\theta_1 + T_{22}\theta_2)f^{pi} \Big|_+ - \xi_{21} \frac{\partial f^{pi}}{\partial \theta_1} \Big|_+ - \xi_{22} \frac{\partial f^{pi}}{\partial \theta_2} \Big|_+ \quad (2.47)$$

$$J_{2pi}^- = (T_{21}\theta_1 + T_{22}\theta_2)f^{pi} \Big|_- - \xi_{21} \frac{\partial f^{pi}}{\partial \theta_1} \Big|_- - \xi_{22} \frac{\partial f^{pi}}{\partial \theta_2} \Big|_- \quad (2.48)$$

$$J_{1pi}^+ = (T_{11}\theta_1 + T_{12}\theta_2)f^{pi} \Big|_+ - \xi_{11} \frac{\partial f^{pi}}{\partial \theta_1} \Big|_+ - \xi_{12} \frac{\partial f^{pi}}{\partial \theta_2} \Big|_+ \quad (2.49)$$

$$J_{1pi}^- = (T_{11}\theta_1 + T_{12}\theta_2)f^{pi} \Big|_- - \xi_{11} \frac{\partial f^{pi}}{\partial \theta_1} \Big|_- - \xi_{12} \frac{\partial f^{pi}}{\partial \theta_2} \Big|_- \quad (2.50)$$

et en utilisant la continuité des densités

$$f^{pi} \Big|_+ = f^{pi} \Big|_- \quad (2.51)$$

permet d'écrire la matrice de réinsertion suivante:

$$\begin{bmatrix} J_{1pi}^+ - J_{1pi}^- \\ J_{2pi}^+ - J_{2pi}^- \end{bmatrix} = \begin{bmatrix} J_{2pa} \sin \theta \\ J_{2pa} \end{bmatrix} = - \begin{bmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial \theta_1} \Big|_+ - \frac{\partial f}{\partial \theta_1} \Big|_- \\ \frac{\partial f}{\partial \theta_2} \Big|_+ - \frac{\partial f}{\partial \theta_2} \Big|_- \end{bmatrix} \quad (2.52)$$

Par la suite on conviendra que: $f \triangleq f^{pi}$ et

$$\left\{ \frac{\partial f}{\partial \theta_1} \Big|_+ - \frac{\partial f}{\partial \theta_1} \Big|_- \right\} = \Phi_{\theta_1} \quad (2.53)$$

$$\left\{ \frac{\partial f}{\partial \theta_2} \Big|_+ - \frac{\partial f}{\partial \theta_2} \Big|_- \right\} = \Phi_{\theta_2} \quad (2.54)$$

Considérons de plus les développements de Taylor suivants (où les termes d'ordre trois ou plus sont négligés):

$$f(\theta_1, \theta_2 + h_2) = f(\theta_1, \theta_2) + \frac{\partial f}{\partial \theta_2} \Big|_+ h_2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ h_2^2 \quad (2.55)$$

$$f(\theta_1, \theta_2 - h_2) = f(\theta_1, \theta_2) - \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} h_2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} h_2^2 \quad (2.56)$$

$$\begin{aligned} f(\theta_1 + h_1, \theta_2 + h_2) &= f(\theta_1, \theta_2) + \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} h_1 + \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} h_2 + \\ &\quad \frac{1}{2} \frac{\partial^2 f}{\partial \theta_1^2} \Big|_{\downarrow} h_1^2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} h_2^2 + \\ &\quad \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} \Big|_{\downarrow} h_1 h_2 \end{aligned} \quad (2.57)$$

$$\begin{aligned} f(\theta_1 - h_1, \theta_2 - h_2) &= f(\theta_1, \theta_2) - \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} h_1 - \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} h_2 + \\ &\quad \frac{1}{2} \frac{\partial^2 f}{\partial \theta_1^2} \Big|_{\downarrow} h_1^2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} h_2^2 + \\ &\quad \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} \Big|_{\downarrow} h_1 h_2 \end{aligned} \quad (2.58)$$

$$\begin{aligned} f(\theta_1 - h_1, \theta_2 + h_2) &= f(\theta_1, \theta_2) - \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} h_1 + \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} h_2 + \\ &\quad \frac{1}{2} \frac{\partial^2 f}{\partial \theta_1^2} \Big|_{\downarrow} h_1^2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} h_2^2 - \\ &\quad \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} \Big|_{\downarrow} h_1 h_2 \end{aligned} \quad (2.59)$$

$$\begin{aligned} f(\theta_1 + h_1, \theta_2 - h_2) &= f(\theta_1, \theta_2) + \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} h_1 - \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} h_2 + \\ &\quad \frac{1}{2} \frac{\partial^2 f}{\partial \theta_1^2} \Big|_{\downarrow} h_1^2 + \frac{1}{2} \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} h_2^2 - \\ &\quad \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} \Big|_{\downarrow} h_1 h_2 \end{aligned} \quad (2.60)$$

L'approximation retenue consiste à considérer l'expression suivante:

$$\frac{\partial f}{\partial t} = \frac{1}{2} \left\{ \frac{\partial f}{\partial t} \Big|_{\downarrow} + \frac{\partial f}{\partial t} \Big|_{\uparrow} \right\} \quad (2.61)$$

Ce qui devient en substituant dans l'équation de Fokker-Planck (2.35):

$$\begin{aligned} \frac{\partial f}{\partial t} = & \frac{1}{2} \left[-\frac{\partial \Gamma_1 f}{\partial \theta_1} \Big|_+ - \frac{\partial \Gamma_2 f}{\partial \theta_2} \Big|_+ + 2 \frac{\partial^2 \xi_{12} f}{\partial \theta_1 \partial \theta_2} \Big|_+ \right. \\ & \left. + \frac{\partial^2 \xi_{22} f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 \xi_{11} f}{\partial \theta_1^2} \Big|_+ \right. \\ & \left. - \frac{\partial \Gamma_1 f}{\partial \theta_1} \Big|_- - \frac{\partial \Gamma_2 f}{\partial \theta_2} \Big|_- + 2 \frac{\partial^2 \xi_{12} f}{\partial \theta_1 \partial \theta_2} \Big|_- \right. \\ & \left. + \frac{\partial^2 \xi_{22} f}{\partial \theta_2^2} \Big|_- + \frac{\partial^2 \xi_{11} f}{\partial \theta_1^2} \Big|_- \right] \end{aligned} \quad (2.62)$$

$$\begin{aligned} \frac{\partial f}{\partial t} = & - \left\{ -\frac{\partial \Gamma_1}{\partial \theta_1} - \frac{\partial \Gamma_2}{\partial \theta_2} \right\} f - \frac{1}{2} \left\{ \frac{\partial f}{\partial \theta_1} \Big|_+ + \frac{\partial f}{\partial \theta_1} \Big|_- \right\} \Gamma_1 \\ & - \frac{1}{2} \left\{ \frac{\partial f}{\partial \theta_2} \Big|_+ + \frac{\partial f}{\partial \theta_2} \Big|_- \right\} \Gamma_2 + \frac{\xi_{11}}{2} \left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} \\ & + \frac{\xi_{22}}{2} \left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2^2} \Big|_- \right\} + \xi_{12} \left\{ \frac{\partial^2 f}{\partial \theta_1 \theta_2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1 \theta_2} \Big|_- \right\} \end{aligned} \quad (2.63)$$

Cherchons maintenant à exprimer:

$$\left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2^2} \Big|_- \right\} \quad (2.64)$$

En rajoutant (2.55) à (2.56) on obtient:

$$\begin{aligned} f(\theta_1, \theta_2 + h_2) - 2f(\theta_1, \theta_2) + f(\theta_1, \theta_2 - h_2) &= \frac{1}{2} \left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2^2} \Big|_- \right\} h_2^2 \\ &+ \left\{ \frac{\partial f}{\partial \theta_2} \Big|_+ - \frac{\partial f}{\partial \theta_2} \Big|_- \right\} h_2 \end{aligned} \quad (2.65)$$

De sorte que:

$$\left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2^2} \Big|_- \right\} = 2 \frac{[f(\theta_1, \theta_2 + h_2) - 2f(\theta_1, \theta_2) + f(\theta_1, \theta_2 - h_2) - \Phi_{\theta_2} h_2]}{h_2^2} \quad (2.66)$$

Où l'on voit apparaître le terme de réinsertion Φ_{θ_2} .

Cherchons maintenant le terme:

$$\left\{ \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_2} \Big|_{\uparrow} \right\} \quad (2.67)$$

De (2.55)-(2.56) on tire que:

$$\begin{aligned} f(\theta_1, \theta_2 + h_2) - f(\theta_1, \theta_2 - h_2) &= \left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\downarrow} - \frac{\partial^2 f}{\partial \theta_2^2} \Big|_{\uparrow} \right\} h_2^2 \\ &+ \left\{ \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_2} \Big|_{\uparrow} \right\} h_2 \end{aligned} \quad (2.68)$$

Puisque le premier terme est de l'ordre de h_2^2 on tire que:

$$\lim_{h_2 \rightarrow 0} \left\{ \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_2} \Big|_{\uparrow} \right\} = \frac{[f(\theta_1, \theta_2 + h_2) - f(\theta_1, \theta_2 - h_2)]}{h_2} \quad (2.69)$$

Puis cherchons :

$$\left\{ \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_1} \Big|_{\uparrow} \right\} \quad (2.70)$$

De (2.57)-(2.58) on tire que:

$$\begin{aligned} f(\theta_1 + h_1, \theta_2 + h_2) - f(\theta_1 - h_1, \theta_2 - h_2) &= \left\{ \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_1} \Big|_{\uparrow} \right\} h_1 \\ &+ \left\{ \frac{\partial f}{\partial \theta_2} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_2} \Big|_{\uparrow} \right\} h_2 \end{aligned} \quad (2.71)$$

Ce qui en utilisant le résultat (2.69) nous permet d'écrire:

$$\begin{aligned} \left\{ \frac{\partial f}{\partial \theta_1} \Big|_{\downarrow} + \frac{\partial f}{\partial \theta_1} \Big|_{\uparrow} \right\} &= \\ \frac{[f(\theta_1 + h_1, \theta_2 + h_2) - f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1, \theta_2 + h_2) + f(\theta_1, \theta_2 - h_2)]}{h_1} \end{aligned} \quad (2.72)$$

Le terme contenant les dérivées mixtes:

$$\left\{ \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_- \right\} \quad (2.73)$$

est obtenu comme suit:

De (2.58)-(2.60) on obtient:

$$f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1 + h_1, \theta_2 - h_2) = -2h_1 \frac{\partial f}{\partial \theta_1} \Big|_- + 2h_1 h_2 \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_- \quad (2.74)$$

Et de (2.57)-(2.59) on obtient que:

$$f(\theta_1 + h_1, \theta_2 + h_2) - f(\theta_1 - h_1, \theta_2 + h_2) = 2h_1 \frac{\partial f}{\partial \theta_1} \Big|_+ + 2h_1 h_2 \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_+ \quad (2.75)$$

La somme de ces deux dernières équations donne:

$$\begin{aligned} & f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 + h_1, \theta_2 + h_2) - \\ & f(\theta_1 - h_1, \theta_2 + h_2) = \left\{ \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_+ - \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_- \right\} 2h_1 h_2 + \left\{ \frac{\partial f}{\partial \theta_1} \Big|_+ - \frac{\partial f}{\partial \theta_1} \Big|_- \right\} 2h_1 \end{aligned} \quad (2.76)$$

D'où :

$$\left\{ \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_- \right\} = \frac{1}{2h_1 h_2} [f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 + h_1, \theta_2 + h_2) - f(\theta_1 - h_1, \theta_2 + h_2) - \Phi_{\theta_1 h_1}] 2h_1 h_2 \quad (2.77)$$

avec Φ_{θ_1} le terme de réinsertion.

Reste maintenant à évaluer le terme:

$$\left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} \quad (2.78)$$

À partir de (2.57)+(2.58) on obtient:

$$\begin{aligned}
& f(\theta_1 + h_1, \theta_2 + h_2) + f(\theta_1 - h_1, \theta_2 - h_2) - 2f(\theta_1, \theta_2) = \\
& \left\{ \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} \Big|_- \right\} h_1 h_2 + \left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} h_1^2 / 2 \\
& \left\{ \frac{\partial^2 f}{\partial \theta_2^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_2^2} \Big|_- \right\} h_2^2 / 2 + \left\{ \frac{\partial f}{\partial \theta_1} \Big|_+ - \frac{\partial f}{\partial \theta_1} \Big|_- \right\} h_1 + \left\{ \frac{\partial f}{\partial \theta_2} \Big|_+ - \frac{\partial f}{\partial \theta_2} \Big|_- \right\} h_2
\end{aligned} \tag{2.79}$$

Ce qui devient en substituant les expressions trouvées précédemment:

$$\begin{aligned}
& f(\theta_1 + h_1, \theta_2 + h_2) + f(\theta_1 - h_1, \theta_2 + h_2) = 2f(\theta_1, \theta_2) + h_1 \Phi_{\theta_1} + h_2 \Phi_{\theta_2} + \\
& + [2\{f(\theta_1, \theta_2 + h_2) - 2f(\theta_1, \theta_2) + f(\theta_1, \theta_2 - h_2) - \Phi_{\theta_2} h_2\} / h_2^2] h_2^2 / 2 + \\
& [[f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 + h_1, \theta_2 + h_2) \\
& - f(\theta_1 - h_1, \theta_2 + h_2) - \Phi_{\theta_1} h_1] / 2h_1 h_2] h_1 h_2 + \left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} h_1^2 / 2 \tag{2.80}
\end{aligned}$$

On trouve finalement:

$$\begin{aligned}
\left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} &= \frac{1}{h_1^2} [f(\theta_1 + h_1, \theta_2 + h_2) + f(\theta_1 - h_1, \theta_2 - h_2) \\
& - 2f(\theta_1, \theta_2 + h_2) - 2f(\theta_1, \theta_2 - h_2) \\
& - f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 - h_1, \theta_2 + h_2)] \tag{2.81}
\end{aligned}$$

Substituant ces résultats dans l'équation de Fokker-Planck (2.63) on trouve:

$$\begin{aligned}
\frac{\partial f}{\partial t} &= -\frac{\partial \Gamma_1}{\partial \theta_1} f - \frac{\partial \Gamma_2}{\partial \theta_2} f \\
& \frac{\Gamma_1 [f(\theta_1 + h_1, \theta_2 + h_2) - f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1, \theta_2 + h_2) + f(\theta_1, \theta_2 - h_2)]}{2h_1} \\
& \frac{\Gamma_2 [f(\theta_1, \theta_2 + h_2) - f(\theta_1, \theta_2 - h_2)]}{2h_2} \\
& + \frac{\xi_{11}}{2h_1^2} [f(\theta_1 + h_1, \theta_2 + h_2) + f(\theta_1 - h_1, \theta_2 + h_2) - 2f(\theta_1, \theta_2 + h_2)]
\end{aligned}$$

$$\begin{aligned}
& -2f(\theta_1, \theta_2 - h_2) + f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 - h_1, \theta_2 - h_2)] \\
& + \frac{\xi_{22}}{h_2^2} [f(\theta_1, \theta_2 + h_2) + f(\theta_1, \theta_2 - h_2) - 2f(\theta_1, \theta_2) - \Phi_{\theta_2} h_2] \\
& + \frac{\xi_{12}}{2h_1 h_2} [f(\theta_1 - h_1, \theta_2 - h_2) - f(\theta_1 + h_1, \theta_2 - h_2) + f(\theta_1 + h_1, \theta_2 + h_2) \\
& + f(\theta_1 - h_1, \theta_2 + h_2) - 2h_1 \Phi_{\theta_1}] \tag{2.82}
\end{aligned}$$

De plus si on considère les approximations suivantes:

$$\frac{\partial f}{\partial \theta_1} \approx \frac{1}{2} \left\{ \frac{\partial f}{\partial \theta_1} \Big|_+ + \frac{\partial f}{\partial \theta_1} \Big|_- \right\} \tag{2.83}$$

$$\frac{\partial^2 f}{\partial \theta_1^2} \approx \frac{1}{2} \left\{ \frac{\partial^2 f}{\partial \theta_1^2} \Big|_+ + \frac{\partial^2 f}{\partial \theta_1^2} \Big|_- \right\} \tag{2.84}$$

Et en utilisant la matrice de réinsertion définie en (2.52) on obtient:

$$- \xi_{22} \Phi_{\theta_2} - \xi_{12} \Phi_{\theta_1} = J_{2pa} \tag{2.85}$$

Finalement l'équation de Fokker-Planck obtenue est l'équation habituelle à laquelle est ajouté le terme de réinsertion qui est:

$$\vartheta(x, t) = \frac{J_{2pa}}{h_2} \tag{2.86}$$

Lorsque $h_2 \rightarrow 0$ ce terme de réinsertion produit localement une variation Δf_{elem} sur un intervalle de temps dt qui est:

$$\Delta f_{elem} = \frac{J_{2pa} dt}{dh_2} \tag{2.87}$$

On remarque qu'il s'agit de la même expression que celle obtenue pour la perte locale de densité dans la première partie. Ce qui démontre que le schéma obtenu est conservatif au niveau des échanges de probabilité aux parois d'absorption et de réinsertion, condition évidemment essentielle à la validité du schéma.

2.5.4 Terme d'absorption/réinsertion

Rappelons que le courant absorbé/réinséré J_{2pa} est:

$$J_{2pa} = (P_{21}\theta_1 + P_{22}\theta_2)f^{pa} - \xi_{21}\frac{\partial f^{pa}}{\partial\theta_1} - \xi_{22}\frac{\partial f^{pa}}{\partial\theta_2} \quad (2.88)$$

où f^{pa} représente la densité à la paroi absorbante. Or par définition cette dernière est nulle, i.e. $f^{pa} = 0$.

De plus la dérivée selon θ_1 est nulle puisque f^{pa} est constant le long de la paroi absorbante i.e:

$$\frac{\partial f^{pa}}{\partial\theta_1} = 0 \quad (2.89)$$

De sorte qu'on peut exprimer J_{2pa} comme suit:

$$J_{2pa} = -\xi_{22}\frac{\partial f^{pa}}{\partial\theta_2} \quad (2.90)$$

ce qui permet finalement d'écrire l'expression désirée.

2.6 Discontinuité paramétrique

Notons qu'au sein d'un même ensemble de structures l'état discret associé, $m(t)$, peut subir une transition. On écrit ici les expressions sur les parois où ces transitions ont lieu. On étudie d'abord la discontinuité selon l'axe θ_2 . Les paramètres du retour d'état étant différents de part et d'autre de cet axe, on cherche à exprimer l'équation de Fokker-Planck sur cet axe.

On a vu au chapitre 1 que la composante du flux normale à l'axe de discontinuité doit être continue et de plus que la densité f doit y être continue.

On a donc:

$$\vec{N} \cdot \vec{J}|_{s+} = \vec{N} \cdot \vec{J}|_{s-} \quad (2.91)$$

et

$$f|_{s+} = f|_{s-} \quad (2.92)$$

Puisque $J_2|_{s+}$ et $J_2|_{s-}$ sont tangentielles à l'axe de discontinuité (et ne contribuent pas au produit scalaire dans (2.91)) on obtient directement que:

$$J_1|_{s+} = J_1|_{s-} \quad (2.93)$$

Considérons le volume élémentaire centré sur la droite de discontinuité représenté à la figure 2.8.

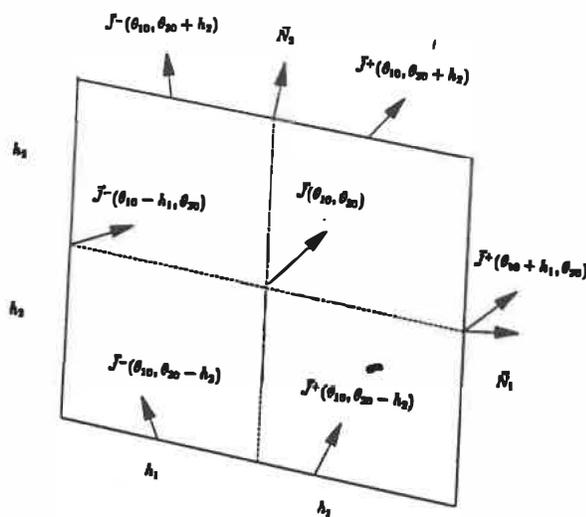


Figure 2.8: Volume élémentaire pour le traitement de la discontinuité paramétrique

Puisque l'équation de Fokker-Planck peut s'exprimer au moyen de la divergence du courant de probabilité, étudions la divergence dans ce volume:

Par définition de la divergence:

$$\begin{aligned}
 \operatorname{div} \vec{J} &= \lim_{h_1, h_2 \rightarrow 0} \frac{\text{Flux s'échappant du volume élémentaire}}{4h_1 h_2 \cos \theta} \\
 &= \lim_{h_1, h_2 \rightarrow 0} \vec{N}_1 \cdot \frac{[\vec{J}^+(\theta_{10} + h_1, \theta_{20}) - \vec{J}^-(\theta_{10} - h_1, \theta_{20})] 2h_2 + \vec{N}_2}{4h_1 h_2 \cos \theta} \\
 &\quad \cdot \frac{[\vec{J}^+(\theta_{10}, \theta_{20} + h_2) - \vec{J}^+(\theta_{10}, \theta_{20} - h_2) + \vec{J}^-(\theta_{10}, \theta_{20} + h_2) - \vec{J}^-(\theta_{10}, \theta_{20} - h_2)] 2h_1}{4h_1 h_2 \cos \theta}
 \end{aligned} \tag{2.94}$$

De l'évaluation des produits scalaires et de la simplification des cosinus on tire que:

$$\begin{aligned}
 \operatorname{div} \vec{J} &= \lim_{h_1, h_2 \rightarrow 0} \frac{[J_1^+(\theta_{10} + h_1, \theta_{20}) - J_1^+(\theta_{10} - h_1, \theta_{20})]}{2h_1} + \\
 &\quad \frac{[J_2^+(\theta_{10}, \theta_{20} + h_2) - J_2^+(\theta_{10}, \theta_{20} - h_2) + J_2^-(\theta_{10}, \theta_{20} + h_2) - J_2^-(\theta_{10}, \theta_{20} - h_2)]}{4h_2}
 \end{aligned} \tag{2.95}$$

En utilisant la condition de continuité des courants (2.93):

$$J_1^+(\theta_{10}, \theta_{20}) = J_1^-(\theta_{10}, \theta_{20}) \tag{2.96}$$

En utilisant les expressions des courants:

$$J_1^+ = (T_{11}^+ \theta_1 + T_{12}^+ \theta_2) f^+ - \xi_{11} \frac{\partial f^+}{\partial \theta_1} - \xi_{12} \frac{\partial f^+}{\partial \theta_2} \tag{2.97}$$

$$J_1^- = (T_{11}^- \theta_1 + T_{12}^- \theta_2) f^- - \xi_{11} \frac{\partial f^-}{\partial \theta_1} - \xi_{12} \frac{\partial f^-}{\partial \theta_2} \tag{2.98}$$

$$J_2^+ = (T_{21}^+ \theta_1 + T_{22}^+ \theta_2) f^+ - \xi_{21} \frac{\partial f^+}{\partial \theta_1} - \xi_{22} \frac{\partial f^+}{\partial \theta_2} \tag{2.99}$$

$$J_2^- = (T_{21}^- \theta_1 + T_{22}^- \theta_2) f^- - \xi_{21} \frac{\partial f^-}{\partial \theta_1} - \xi_{22} \frac{\partial f^-}{\partial \theta_2} \tag{2.100}$$

On obtient une expression ne faisant intervenir que des dérivées parfaitement définies de chaque côté de la discontinuité:

$$\begin{aligned} \operatorname{div} \vec{J} &= \lim_{h_1, h_2 \rightarrow 0} \frac{[J_1^+(\theta_{10} + h_1, \theta_{20}) - J_1^+(\theta_{10}, \theta_{20}) + J_1^-(\theta_{10}, \theta_{20}) - J_1^-(\theta_{10} - h_1, \theta_{20})]}{2h_1} \\ &+ \frac{[J_2^+(\theta_{10}, \theta_{20} + h_2) - J_2^+(\theta_{10}, \theta_{20} - h_2) + J_2^-(\theta_{10}, \theta_{20} + h_2) - J_2^-(\theta_{10}, \theta_{20} - h_2)]}{4h_2} \\ &= 1/2 \left[\frac{\partial(J_1^+ + J_1^-)}{\partial \theta_1} \right] + 1/2 \left[\frac{\partial(J_2^+ + J_2^-)}{\partial \theta_2} \right] \end{aligned} \quad (2.101)$$

En utilisant la condition (2.92) de continuité des densités on obtient:

$$J_1^+ + J_1^- = (T_{11}^+ + T_{11}^-)\theta_1 + (T_{12}^+ + T_{12}^-)\theta_2 f - \xi_{11} \frac{\partial f^+}{\partial \theta_1} - \xi_{11} \frac{\partial f^-}{\partial \theta_1} - \xi_{12} \frac{\partial f^+}{\partial \theta_2} - \xi_{12} \frac{\partial f^-}{\partial \theta_2} \quad (2.102)$$

et

$$J_2^+ + J_2^- = (T_{21}^+ + T_{21}^-)\theta_1 + (T_{22}^+ + T_{22}^-)\theta_2 f - \xi_{21} \frac{\partial f^+}{\partial \theta_1} - \xi_{21} \frac{\partial f^-}{\partial \theta_1} - \xi_{22} \frac{\partial f^+}{\partial \theta_2} - \xi_{22} \frac{\partial f^-}{\partial \theta_2} \quad (2.103)$$

De sorte que finalement l'équation de Fokker-Planck sur la paroi de discontinuité paramétrique s'écrit:

$$\begin{aligned} \frac{\partial f}{\partial t} &= - \frac{\partial [1/2(T_{11}^+ + T_{11}^-)\theta_1 + 1/2(T_{12}^+ + T_{12}^-)\theta_2] f}{\partial \theta_1} \\ &- \frac{\partial [1/2(T_{21}^+ + T_{21}^-)\theta_1 + 1/2(T_{22}^+ + T_{22}^-)\theta_2] f}{\partial \theta_2} \\ &+ \xi_{12} \frac{\partial^2 f^+}{2\partial \theta_1 \partial \theta_2} + \xi_{12} \frac{\partial^2 f^-}{2\partial \theta_1 \partial \theta_2} + \xi_{11} \frac{\partial^2 f^+}{2\partial \theta_1^2} + \xi_{11} \frac{\partial^2 f^-}{2\partial \theta_1^2} \\ &+ \xi_{21} \frac{\partial^2 f^+}{2\partial \theta_1 \partial \theta_2} + \xi_{21} \frac{\partial^2 f^-}{2\partial \theta_1 \partial \theta_2} + \xi_{22} \frac{\partial^2 f^+}{2\partial \theta_2^2} + \xi_{22} \frac{\partial^2 f^-}{2\partial \theta_2^2} \end{aligned} \quad (2.104)$$

L'approximation retenue par la suite consistera à écrire localement l'équation de Fokker-Planck avec la moitié des paramètres de "dérive" de chaque côté de la

discontinuité:

$$\begin{aligned} \frac{\partial f}{\partial t} = & - \frac{\partial [1/2(T_{11}^+ + T_{11}^-)\theta_1 + 1/2(T_{12}^+ + T_{12}^-)\theta_2] f}{\partial \theta_1} \\ & - \frac{\partial [1/2(T_{21}^+ + T_{21}^-)\theta_1 + 1/2(T_{22}^+ + T_{22}^-)\theta_2] f}{\partial \theta_1} \\ & + \xi_{12} \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} + \xi_{11} \frac{\partial^2 f}{\partial \theta_1^2} + \xi_{21} \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} + \xi_{22} \frac{\partial^2 f}{\partial \theta_2^2} \end{aligned} \quad (2.105)$$

avec $T_{12}^+ = T_{12}^-$ et $T_{22}^+ = T_{22}^-$.

On peut montrer qu'en appliquant les mêmes arguments pour une discontinuité selon l'axe θ_1 on obtient une expression du même type mais avec cette fois-ci, $T_{11}^+ = T_{11}^-$ et $T_{21}^+ = T_{21}^-$.

2.7 Parois réfléchissantes

On a vu au chapitre 1 que sur ces parois la composante du flux normale à cette surface doit être nulle, i.e:

$$\vec{N} \cdot \vec{J} = 0 \quad (2.106)$$

En se référant à la figure 2.9 on obtient les résultats qui suivent.

Puisque J_2 est tangentielle à la surface A on a que $J_1 \cos \theta = 0$ i.e:

$$J_1 = (T_{11}\theta_1 + T_{12}\theta_2)f - \xi_{11} \frac{\partial f}{\partial \theta_1} - \xi_{12} \frac{\partial f}{\partial \theta_2} = 0 \quad (2.107)$$

De plus puisque J_1 est tangentielle à la surface B on a que $J_2 \cos \theta = 0$ i.e:

$$J_2 = (T_{21}\theta_1 + T_{22}\theta_2)f - \xi_{21} \frac{\partial f}{\partial \theta_1} - \xi_{22} \frac{\partial f}{\partial \theta_2} = 0 \quad (2.108)$$

Ce qui complète l'étude des conditions frontières et des mécanismes de couplage de notre modèle.

Maintenant que notre modèle est bien spécifié sur la plan analytique passons à la formulation d'un schéma numérique approprié qui permettra d'en évaluer les propriétés par simulation.

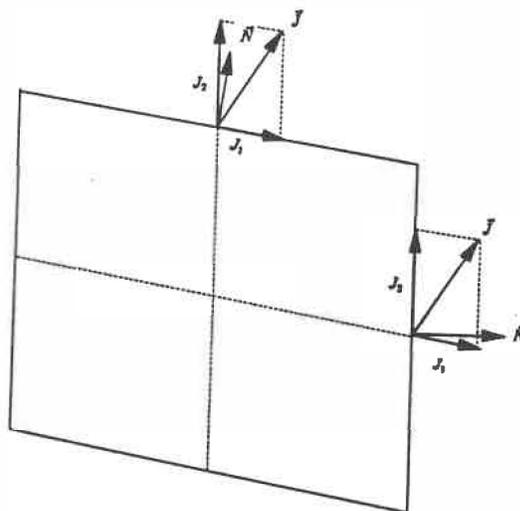


Figure 2.9: Volume élémentaire pour le traitement des parois réfléchissantes

Chapitre 3

Schéma numérique

3.1 Décomposition d'opérateur

La simulation du système d'équations de Fokker-Planck couplées paraboliques se fera à l'aide d'un schéma aux différences finies. On peut montrer (Gerard et Wheatley 1988) que le schéma numérique le plus précis est purement explicite, cependant on est confronté à un problème de stabilité. D'autre part le schéma numérique le plus stable est implicite mais il manque de précision.

Une alternative est d'utiliser le schéma de Crank-Nicholson (Ames 1965) qui est à la fois implicite et explicite ce qui assure une bonne stabilité et précision de la solution. Son application implique cependant l'inversion de systèmes couplés d'équations linéaires ce qui entraîne des problèmes numériques considérables.

La méthode des directions alternées implicites permet de poser le problème en termes de résolution de systèmes tridiagonaux simples et conserve les propriétés de stabilité et de précision du schéma de Crank-Nicholson (Press 1986). Elle repose sur la division de chaque pas temporel en deux demi-pas. A chaque demi-pas l'une des directions est traitée implicitement tandis que l'autre est traitée explicitement. Chaque demi-pas ne nécessite que la résolution d'un système tridiagonal.

La méthode ADI (“alternating-direction implicit method”) peut être envisagée comme une méthode de décomposition d’un opérateur différentiel \mathcal{L} (Press 1986),(Yamenko 1971).

Chaque équation de Fokker-Planck est alors conçue comme suit:

$$\frac{\partial f}{\partial t} = \mathcal{L}f \quad (3.1)$$

où l’on suppose que \mathcal{L} peut se décomposer comme une somme d’opérateurs linéaires $\mathcal{L}_1, \dots, \mathcal{L}_m$. Dans la méthode ADI, m , est la dimension du système puisqu’on suppose que \mathcal{L} peut se décomposer comme la somme de m opérateurs qui peuvent être traités alternativement de façon explicite et implicite. Par exemple, dans le cas bidimensionnel, une décomposition naturelle est $\mathcal{L} = \mathcal{L}_x + \mathcal{L}_y$

Convenons d’appeler \mathcal{U}_1 une méthode de résolution qui inclut tous les termes de l’opérateur global \mathcal{L} mais qui ne soit stable que par rapport à la partie de l’opérateur global \mathcal{L}_1 et \mathcal{U}_2 une méthode stable pour \mathcal{L}_2 .

Alors une façon de passer de f^t à $f^{t+\Delta t}$ est la suivante:

$$\begin{aligned} f^{t+\frac{\Delta t}{2}} &= \mathcal{U}_1\left(f^t, \frac{\Delta t}{2}\right) \\ f^{t+\Delta t} &= \mathcal{U}_2\left(f^{t+\frac{\Delta t}{2}}, \frac{\Delta t}{2}\right) \end{aligned} \quad (3.2)$$

3.2 Décomposition de l'équation de Fokker-Planck

Reprenons l'expression de l'équation de Fokker-Planck et suggérons une décomposition qui approxime le mieux un schéma ADI pur.

$$\frac{\partial f}{\partial t} = \frac{\partial \Gamma_1 f}{\partial \theta_1} - \frac{\partial \Gamma_2 f}{\partial \theta_2} + 2\xi_{12} \frac{\partial^2 f}{\partial \theta_1 \partial \theta_2} + \xi_{11} \frac{\partial^2 f}{\partial \theta_1^2} + \xi_{22} \frac{\partial^2 f}{\partial \theta_2^2} \quad (3.3)$$

avec

$$\begin{aligned} \Gamma_1 &= T_{11}\theta_1 + T_{12}\theta_2 \\ \Gamma_2 &= T_{21}\theta_1 + T_{22}\theta_2 \end{aligned} \quad (3.4)$$

L'opérateur (qui est le générateur différentiel) \mathcal{L} est :

$$\mathcal{L} = -\frac{\partial \Gamma_1}{\partial \theta_1} - \frac{\partial \Gamma_2}{\partial \theta_2} + 2\xi_{12} \frac{\partial^2}{\partial \theta_1 \partial \theta_2} + \xi_{11} \frac{\partial^2}{\partial \theta_1^2} + \xi_{22} \frac{\partial^2}{\partial \theta_2^2} \quad (3.5)$$

Décomposons maintenant cet opérateur comme suit:

$$\mathcal{L}_1 = -\frac{\partial \Gamma_1}{\partial \theta_1} + \xi_{11} \frac{\partial^2}{\partial \theta_1^2} \quad (3.6)$$

et

$$\mathcal{L}_2 = -\frac{\partial \Gamma_2}{\partial \theta_2} + \xi_{22} \frac{\partial^2}{\partial \theta_2^2} \quad (3.7)$$

et

$$\mathcal{L}_3 = 2\xi_{12} \frac{\partial^2}{\partial\theta_1 \partial\theta_2} \quad (3.8)$$

De sorte que

$$\frac{\partial f}{\partial t} = (\mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3)(f) \quad (3.9)$$

On remarque que seule la partie \mathcal{L}_3 de l'opérateur global \mathcal{L} ne se prête pas à un schéma purement explicite ou implicite. La stratégie utilisée consistera à identifier les parties de \mathcal{L}_3 qui peuvent être envisagées de façon purement implicite et les autres qui doivent être prises de façon purement explicite.

La solution proposée est alors la suivante:

$$\begin{aligned} f^{t+\frac{\Delta t}{2}} &= \mathcal{U}_1\left(f^t, \frac{\Delta t}{2}\right) \\ f^{t+\Delta t} &= \mathcal{U}_2\left(f^{t+\frac{\Delta t}{2}}, \frac{\Delta t}{2}\right) \end{aligned} \quad (3.10)$$

avec \mathcal{U}_1 une méthode spécifiée par l'application des opérateurs suivants:

$$\mathcal{U}_1 = \mathcal{L}_1^{exp} + \mathcal{L}_2^{imp} + \mathcal{L}_3^2$$

où

\mathcal{L}_2^{imp} est prise implicitement et \mathcal{L}_1^{exp} est prise explicitement.

et \mathcal{U}_2 une méthode spécifiée par l'application des opérateurs :

$$\mathcal{U}_2 = \mathcal{L}_1^{imp} + \mathcal{L}_2^{exp} + \mathcal{L}_3^1$$

\mathcal{L}_1^{imp} est prise explicitement et \mathcal{L}_2^{exp} est prise implicitement.

Finalement \mathcal{L}_3^i , $i=1,2$, est pris avec le maximum de termes implicites possibles selon la direction i .

Clarifions notre propos en étudiant la discrétisation de l'équation de Fokker-Planck et sa décomposition.

3.3 Discrétisation de l'équation de Fokker-Planck excluant discontinuités et conditions frontières

Introduisons les notations suivantes:

$$\begin{aligned} & f(\pm h_1, \pm h_2) \Delta f(\theta_{1n}, \theta_{2n}) \\ & \lambda_1(\pm h_1, \pm h_2) \Delta T_{11}(\theta_{1n} \pm h_1) + T_{12}(\theta_{2n} \pm h_2) \\ & \lambda_2(\pm h_1, \pm h_2) \Delta T_{21}(\theta_{1n} \pm h_1) + T_{22}(\theta_{2n} \pm h_2) \\ & f(\pm h_1, \pm h_2) \Delta f(\theta_{1n} \pm h_1, \theta_{2n} \pm h_2) \end{aligned}$$

Voyons d'abord ce que donnerait notre schéma sur la dérivée mixte intervenant dans l'opérateur \mathcal{L}_3 .

L'approximation retenue pour cette dérivée est:

$$\begin{aligned} \frac{\partial^2 f(\theta_{1n}, \theta_{2n})}{\partial \theta_1 \partial \theta_2} &= \frac{1}{2h_1 h_2} [f(h_1, h_2) + 2f(\theta_{1n}, \theta_{2n}) + f(-h_1, -h_2) \\ & - f(\theta_{1n} + h_1, \theta_{2n}) - f(\theta_{1n} - h_1, \theta_{2n}) \\ & - f(\theta_{1n}, \theta_{2n} + h_2) - f(\theta_{1n}, \theta_{2n} - h_2)] \end{aligned} \quad (3.11)$$

On constate que l'approximation résultante ne peut s'intégrer directement dans un schéma ADI. Certains des termes de l'approximation ne se situant uniquement ni sur une colonne (i constant) ou sur une ligne (j constant). La stratégie que l'on utilisera consiste à intégrer le maximum de termes dans un schéma ADI puis d'approximer les autres termes de façon purement explicite.

Les approximations pour les dérivées premières et secondes à l'ordre de $\mathcal{O}(h^2)$ sont données par les formules qui suivent:

$$\frac{\partial^2 f(\theta_{1n}, \theta_{2n})}{\partial \theta_1^2} = \frac{f(\theta_{1n} + h_1, \theta_{2n}) - 2f(\theta_{1n}, \theta_{2n}) + f(\theta_{1n} - h_1, \theta_{2n})}{h_1^2} \quad (3.12)$$

$$\frac{\partial^2 f(\theta_{1n}, \theta_{2n})}{\partial \theta_2^2} = \frac{f(\theta_{1n}, \theta_{2n} + h_2) - 2f(\theta_{1n}, \theta_{2n}) + f(\theta_{1n}, \theta_{2n} - h_2)}{h_2^2} \quad (3.13)$$

$$\frac{\partial f(\theta_{1n}, \theta_{2n})}{\partial \theta_1} = \frac{f(\theta_{1n} + h_1, \theta_{2n}) - f(\theta_{1n} - h_1, \theta_{2n})}{2h_1} \quad (3.14)$$

$$\frac{\partial f(\theta_{1n}, \theta_{2n})}{\partial \theta_2} = \frac{f(\theta_{1n}, \theta_{2n} + h_2) - f(\theta_{1n}, \theta_{2n} - h_2)}{2h_2} \quad (3.15)$$

Reprenant la décomposition suggérée plus haut on obtient que:

$$\begin{aligned} \mathcal{L}_1^{imp}(f) &= -\frac{\lambda_1(h_1, 0)f(h_1, 0)^{k+1} - \lambda_1(-h_1, 0)f(-h_1, 0)^{k+1}}{2h_1} + \\ &\quad \xi_{11} \frac{f(h_1, 0)^{k+1} - 2f(\cdot)^{k+1} + f(-h_1, 0)^{k+1}}{h_1^2} \end{aligned} \quad (3.16)$$

$$\begin{aligned} \mathcal{L}_1^{exp}(f) &= -\frac{\lambda_1(h_1, 0)f^k(h_1, 0) - \lambda_1(-h_1, 0)f^k(-h_1, 0)}{2h_1} + \\ &\xi_{11} \frac{f^k(h_1, 0) - 2f^k() + f^k(-h_1, 0)}{h_1^2} \end{aligned} \quad (3.17)$$

$$\begin{aligned} \mathcal{L}_2^{imp}(f) &= -\frac{\lambda_2(0, h_2)f^{k+1}(0, h_2) - \lambda_2(0, -h_2)f^{k+1}(0, -h_2)}{2h_2} + \\ &\xi_{22} \frac{f^{k+1}(0, h_2) - 2f^{k+1}() + f^{k+1}(0, -h_2)}{h_2^2} \end{aligned} \quad (3.18)$$

$$\begin{aligned} \mathcal{L}_2^{exp}(f) &= -\frac{\lambda_2(0, h_2)f^k(0, h_2) - \lambda_2(0, -h_2)f^k(0, -h_2)}{2h_2} + \\ &\xi_{22} \frac{f^k(0, h_2) - 2f^k() + f^k(0, -h_2)}{h_2^2} \end{aligned} \quad (3.19)$$

$$\begin{aligned} \mathcal{L}_3^1(f) &= \frac{\xi_{12}}{h_1 h_2} [f^k(h_1, h_2) + 2f^{k+1}() + f^k(-h_1, -h_2) \\ &\quad - f^{k+1}(h_1, 0) - f^{k+1}(-h_1, 0) \\ &\quad - f^k(0, h_2) - f^k(0, -h_2)] \end{aligned} \quad (3.20)$$

et

$$\begin{aligned} \mathcal{L}_3^2(f) &= \frac{\xi_{12}}{h_1 h_2} [f^k(h_1, h_2) + 2f^{k+1}() + f^k(-h_1, -h_2) \\ &\quad - f^k(h_1, 0) - f^k(-h_1, 0) \\ &\quad - f^{k+1}(0, h_2) - f^{k+1}(0, -h_2)] \end{aligned} \quad (3.21)$$

Comme on l'a vu la méthode de décomposition repose sur la division du pas temporel Δt en deux demi-pas de $\tau = \frac{\Delta t}{2}$ tels que:

$$\begin{aligned}
f^{t+\frac{\Delta t}{2}} &= \mathcal{U}_1(f^t, \frac{\Delta t}{2}) \\
f^{t+\Delta t} &= \mathcal{U}_2(f^{t+\frac{\Delta t}{2}}, \frac{\Delta t}{2})
\end{aligned}
\tag{3.22}$$

Le premier demi-pas s'opère sur l'intervalle temporel $[k\Delta t, k\Delta t + \tau]$ et l'équation de Fokker-Planck sur ce premier demi-pas s'écrit en introduisant la notation suivante:

$$(k+a) \triangleq [k\Delta t, k\Delta t + a\tau] \tag{3.23}$$

$$\begin{aligned}
\frac{f^{k\Delta t+\tau}() - f^{k\Delta t}()}{\tau} &= \frac{\lambda_1(h_1, 0)f^k(h_1, 0) - \lambda_1(-h_1, 0)f^k(-h_1, 0)}{2h_1} + \\
&\xi_{11} \frac{f^k(h_1, 0) - 2f^k() + f^k(-h_1, 0)}{h_1^2} + \\
&\frac{\lambda_2(0, h_2)f^{k+1}(0, h_2) - \lambda_2(0, -h_2)f^{k+1}(0, -h_2)}{2h_2} + \\
&\xi_{22} \frac{f^{k+1}(0, h_2) - 2f^{k+1}() + f^{k+1}(0, -h_2)}{h_2^2} + \\
&\frac{\xi_{12}}{h_1 h_2} [f^k(h_1, h_2) + 2f^{k+1}() + f^k(-h_1, -h_2) \\
&- f^k(h_1, 0) - f^k(-h_1, 0) \\
&- f^{k+1}(0, h_2) - f^{k+1}(0, -h_2)]
\end{aligned}
\tag{3.24}$$

Et pour le second demi-pas, i.e dans l'intervalle $[k\Delta t + \tau, k\Delta t + 2\tau]$:

$$\begin{aligned}
\frac{f^{k\Delta t+2\tau}() - f^{k\Delta t+\tau}()}{\tau} &= \frac{\lambda_1(h_1, 0)f^{k+2}(h_1, 0) - \lambda_1(-h_1, 0)f^{k+2}(-h_1, 0)}{2h_1} + \\
&\xi_{11} \frac{f^{k+2}(h_1, 0) - 2f^{k+2}() + f^{k+2}(-h_1, 0)}{h_1^2} +
\end{aligned}$$

$$\begin{aligned}
& \frac{\lambda_2(0, h_2)f^{k+1}(0, h_2) - \lambda_2(0, -h_2)f^{k+1}(0, -h_2)}{2h_2} + \\
& \xi_{22} \frac{f^{k+1}(0, h_2) - 2f^{k+1}() + f^{k+1}(0, h_2)}{h_2^2} + \\
& \frac{\xi_{12}}{h_1 h_2} [f^{k+1}(h_1, h_2) + 2f^{k+2}() + f^{k+1}(-h_1, -h_2) \\
& - f^{k+2}(h_1, 0) - f^{k+2}(-h_1, 0) \\
& - f^{k+1}(0, h_2) - f^{k+1}(0, -h_2)] \tag{3.25}
\end{aligned}$$

En définissant les variables intermédiaires suivantes:

$$r_1 = \frac{\xi_{12}\tau}{h_1 h_2}, r_2 = \frac{\xi_{11}\tau}{h_1^2}, r_3 = \frac{\xi_{22}\tau}{h_2^2}, r_4 = \frac{\tau}{2h_1}, r_5 = \frac{\tau}{2h_2} \tag{3.26}$$

et en groupant les termes explicites et implicites on obtient,

pour le premier demi-pas:

$$\begin{aligned}
& f^{k+1}()(1 - 2r_1 + 2r_3) + f^{k+1}(0, h_2)(r_1 - r_3 + \lambda_2(0, h_2)r_5 + \\
& f^{k+1}(0, h_2)(r_1 - r_3 + \lambda_2(0, -h_2)r_5 = \\
& f^k(h_1, h_2)r_1 + f^k(-h_1, -h_2)r_1 + f^k()(1 - 2r_2) + \\
& f^k(h_1, 0)(-r_1 + r_2 - \lambda_1(h_1, 0)r_4) + f^k(-h_1, 0)(-r_1 + r_2 + \lambda_1(-h_1, 0)r_4) \tag{3.27}
\end{aligned}$$

et pour le second demi-pas:

$$\begin{aligned}
& f^{k+2}()(1 - 2r_1 + 2r_2) + f^{k+2}(h_1, 0)(r_1 - r_2 + \lambda_1(h_1, 0)r_4) + \\
& f^{k+2}(h_1, 0)(r_1 - r_2 + \lambda_1(-h_1, 0)r_4) =
\end{aligned}$$

$$\begin{aligned}
& f^{k+1}(h_1, h_2)r_1 + f^{k+1}(-h_1, -h_2)r_1 + f^{k+1}()(1 - 2r_3) + \\
& f^{k+1}(0, h_2)(-r_1 + r_3 - \lambda_2(0, h_2)r_5) + f^{k+1}(0, -h_2)(-r_1 + r_3 + \lambda_2(0, -h_2)r_5)
\end{aligned}
\tag{3.28}$$

3.4 Conditions frontières

Les expressions (3.27) et (3.28) représentent le système en absence de toute condition frontière et de discontinuité . Voyons maintenant ce que deviennent ces expressions lorsqu'on introduit certaines conditions frontières.

3.4.1 Parois réfléchissantes

On utilise un développement de Taylor pour trouver les composantes explicites du type $f^k(h_1, h_2)$ et $f^k(-h_1, -h_2)$. Entourons les parois réfléchissantes d'un tampon imaginaire de profondeur unitaire, qu'on représente à la figure 3.1. Les valeurs du tampon seront utilisées pour respecter la condition frontière.

Premier demi-pas: paroi L^A

Examinons les équations sur la paroi L^A :

En reprenant l'expression du premier demi-pas,

$$\begin{aligned}
& f^{k+1}()(1 - 2r_1 + 2r_3) + f^{k+1}(0, h_2)(r_1 - r_3 + \lambda_2(0, h_2)r_5) + \\
& f^{k+1}(0, h_2)(r_1 - r_3 + \lambda_2(0, -h_2)r_5) = \\
& f^k(h_1, h_2)r_1 + f^k(-h_1, -h_2)r_1 + f^k()(1 - 2r_2) \\
& + f^k(h_1, 0)(-r_1 + r_2 - \lambda_1(h_1, 0)r_4) + f^k(-h_1, 0)(-r_1 + r_2 + \lambda_1(-h_1, 0)r_4)
\end{aligned}
\tag{3.29}$$

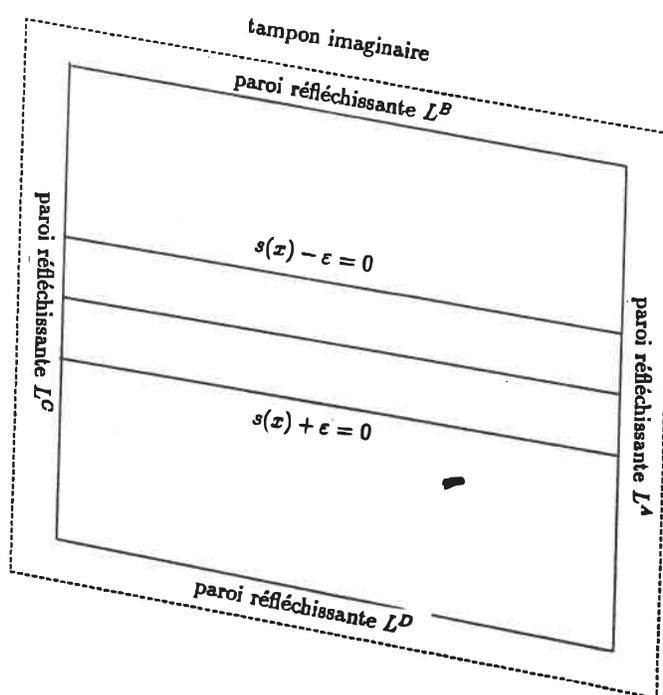


Figure 3.1: Tampon imaginaire utilisé pour les parois réfléchissantes

On voit que les termes $f^k(h_1, h_2)$ et $f^k(h_1, 0)$ appartiennent au tampon et qu'ils apparaissent explicitement dans l'équation. Or si au début du cycle de calcul la condition frontière est respectée alors $f^k()$ doit satisfaire à la condition frontière $J_1 = 0$ i.e:

$$J_1 = (T_{11}\theta_1 + T_{12}\theta_2)f - \xi_{11} \frac{\partial f}{\partial \theta_1} - \xi_{12} \frac{\partial f}{\partial \theta_2} = 0 \quad (3.30)$$

En utilisant les approximations habituelles on obtient que:

$$(T_{11}\theta_1 + T_{12}\theta_2)f^k() - \xi_{11} \frac{f^k(h_1, 0) - f^k(-h_1, 0)}{2h_1} - \xi_{12} \frac{f^k(0, h_2) - f^k(0, -h_2)}{2h_2} = 0 \quad (3.31)$$

i.e

$$f^k(h_1, 0) = \frac{2h_1}{\xi_{11}} \lambda_1(0, 0) f^k() + f^k(-h_1, 0) - \frac{\xi_{12} h_1}{\xi_{11} h_2} f^k(0, h_2) - f^k(0, -h_2) \quad (3.32)$$

Ce qui permet de déterminer la valeur de $f^k(h_1, 0)$.

Maintenant pour déterminer la valeur de $f^k(h_1, h_2)$ utilisons le développement de Taylor suivant autour de $f^k(0, 0)$:

$$f(h_1, h_2) = f() + h_1 \frac{\partial f()}{\partial \theta_1} + h_2 \frac{\partial f()}{\partial \theta_2} + \frac{1}{2} \frac{\partial^2 f()}{\partial \theta_1^2} h_1^2 + \frac{1}{2} \frac{\partial^2 f()}{\partial \theta_2^2} h_2^2 + \frac{\partial^2 f()}{\partial \theta_1 \partial \theta_2} h_1 h_2 =$$

$$f() + h_1 \frac{f(h_1, 0) - f(-h_1, 0)}{2h_1} + h_2 \frac{f(0, h_2) - f(0, -h_2)}{2h_2}$$

$$\begin{aligned}
& + \frac{h_1^2 f(h_1, 0) - f() + f(-h_1, 0)}{2h_1^2} + \frac{h_2^2 f(0, h_2) - f() + f(0, -h_2)}{2h_2^2} \\
& + 2h_1 h_2 \frac{f(h_1, h_2) + 2f() + f(-h_1, -h_2) - f(h_1, 0) - f(-h_1, 0) - f(0, h_2) - f(0, -h_2)}{2h_1 h_2}
\end{aligned} \tag{3.33}$$

Cette expression devient après simplifications:

$$f^k(h_1, h_2) \approx f^k(h_1, 0) - f^k(-h_1, 0) + f^k(0, h_2) + f^k(0, -h_2) + f^k(-h_1, -h_2) \tag{3.34}$$

Ce qui permet de déterminer $f^k(h_1, h_2)$ à partir de quantités connues.

En substituant ces expressions dans la formule aux différences pour le premier demi-pas on trouve que:

$$\begin{aligned}
& f^{k+1}() (1 - 2r_1 + 2r_3) + f^{k+1}(0, h_2) (r_1 - r_3 + \lambda_2(0, h_2)r_5) + \\
& f^{k+1}(0, h_2) (r_1 - r_3 + \lambda_2(0, -h_2)r_5) = f^k(-h_1, -h_2) 2r_1 \\
& + f^k() [1 - 2r_2 - (\lambda_1(0, 0) 2h_1 / \xi_{11}) (r_2 - \lambda_1(h_1, 0))] \\
& + f^k(-h_1, 0) (-2r_1 + 2r_2 + r_4 [\lambda_1(-h_1, 0) - \lambda_1(-h_1, 0)]) \\
& + f^k(0, h_2) (r_1 - r_2 r_6 + \lambda_1(h_1, 0) r_4 r_6) \\
& + f^k(0, -h_2) (-r_1 + r_2 r_6 - \lambda_1(h_1, 0) r_4 r_6)
\end{aligned} \tag{3.35}$$

$$\text{avec } r_6 = \frac{\xi_{12} h_1}{\xi_{11} h_2}$$

Second demi-pas: paroi L^A

Maintenant examinons l'équation du second demi-pas en supposant une fois de plus que la condition frontière $J_1 = 0$ est respectée par les composantes explicites, i.e:

$$\frac{2h_1}{\xi_{11}} \lambda_1(0,0) f^{k+1}() + f^{k+1}(-h_1,0) - \frac{\xi_{12} h_1}{\xi_{11} h_2} [f^{k+1}(0, h_2) - f^{k+1}(0, -h_2)] = f^{k+1}(h_1, 0) \quad (3.36)$$

On constate à l'examen de l'équation du second demi-pas:

$$\begin{aligned} & f^{k+2}()(1 - 2r_1 + 2r_2) + f^{k+2}(h_1, 0)(r_1 - r_2 + \lambda_1(h_1, 0)r_4) + \\ & f^{k+2}(h_1, 0)(r_1 - r_2 + \lambda_1(-h_1, 0)r_4) = f^{k+1}(h_1, h_2)r_1 + f^{k+1}(-h_1, -h_2)r_1 + \\ & f^{k+1}()(1 - 2r_3) + f^{k+1}(0, h_2)(-r_1 + r_3 - \lambda_2(0, h_2)r_5) + \\ & f^{k+1}(0, -h_2)(-r_1 + r_3 + \lambda_2(0, -h_2)r_5) \end{aligned} \quad (3.37)$$

que les points tampons apparaissent de manière explicite et implicite, i.e. $f^{k+1}(h_1, h_2)$ et $f^{k+2}(h_1, 0)$.

On utilise encore une fois le développement de Taylor pour trouver $f^{k+1}(h_1, h_2)$ ce qui donne:

$$\begin{aligned} f^{k+1}(h_1, h_2) & \approx f^{k+1}(h_1, 0) - f^{k+1}(-h_1, 0) + f^{k+1}(0, h_2) + f^{k+1}(0, -h_2) \\ & + f^{k+1}(-h_1, -h_2) \end{aligned} \quad (3.38)$$

Reste maintenant à déterminer $f^{k+2}(h_1, 0)$. On constate que cette valeur appa-

raraît implicitement de sorte qu'on doit simultanément résoudre l'équation de Fokker-Planck et satisfaire à la condition frontière pour trouver sa valeur.

Puisqu'il s'agit du second demi-pas la condition frontière $J_1 = 0$ peut s'écrire:

$$\lambda_1(0,0)f^{k+2}() - \xi_{11} \frac{f^{k+2}(h_1,0) - f^{k+2}(-h_1,0)}{2h_1} - \xi_{12} \frac{f^{k+1}(0,h_2) - f^{k+1}(0,-h_2)}{2h_2} = 0 \quad (3.39)$$

De sorte qu'on peut isoler $f^{k+2}(h_1,0)$ comme suit:

$$\frac{2h_1}{\xi_{11}} \lambda_1(0,0)f^{k+2}() + f^{k+2}(-h_1,0) - \frac{\xi_{12}h_1}{\xi_{11}h_2} f^{k+1}(0,h_2) - f^{k+1}(0,-h_2) = f^{k+1}(h_1,0) \quad (3.40)$$

En substituant ces expressions dans la formule pour le second demi-pas on trouve que:

$$\begin{aligned} & f^{k+2}()(1 - 2r_1 + 2r_2 + \frac{2h_1}{\xi_{11}} \lambda_1(0,0)[r_1 - r_2 + \lambda_1(h_1,0)r_4]) + \\ & f^{k+2}(-h_1,0)(2r_1 - 2r_2 + r_4[\lambda_1(h_1,0) - \lambda_1(-h_1,0)]) = \\ & + f^{k+1}(0,h_2)(-r_2r_6 + \lambda_1(h_1,0)r_4r_6 + r_3 - \lambda_2(0,h_2)r_5) \\ & + f^{k+1}(0,-h_2)(r_2r_6 - \lambda_1(h_1,0)r_4r_6 + r_3 + \lambda_2(0,-h_2)r_5 - 2r_1) \\ & + f^{k+1}()(1 - 2r_3 + \frac{r_1 2h_1}{\xi_{11}} \lambda_1(0,0)) + f^{k+1}(-h_1,-h_2)2r_1 \end{aligned} \quad (3.41)$$

Premier demi-pas: paroi L^B

Maintenant examinons le comportement du schéma sur la frontière L^B :

Toujours en examinant l'équation du premier demi-pas on observe que les composantes appartenant au tampon sont $f^k(h_1, h_2)$ et $f^{k+1}(0, h_2)$. Rappelons que sur cette paroi la condition frontière est:

$$J_2 = (T_{21}\theta_1 + T_{22}\theta_2)f - \xi_{21} \frac{\partial f}{\partial \theta_1} - \xi_{22} \frac{\partial f}{\partial \theta_2} = 0 \quad (3.42)$$

En utilisant les mêmes arguments que ceux du deuxième demi-pas pour la paroi A on obtient que:

$$f^{k+1}(0, h_2) = \frac{2h_2}{\xi_{22}} \lambda_2(0, 0) f^{k+1}() + f^{k+1}(0, -h_2) - \frac{\xi_{21} h_2}{\xi_{22} h_1} f^k(h_1, 0) - f^k(-h_1, 0) \quad (3.43)$$

$$f^k(h_1, h_2) \approx f^k(h_1, 0) - f^k(-h_1, 0) + f^k(0, h_2) - f^k(0, -h_2) + f^k(-h_1, -h_2) \quad (3.44)$$

Ces dernières expressions substituées dans l'équation du premier demi-pas donnent :

$$\begin{aligned} & f^{k+1}() \left(1 - 2r_1 + 2r_3 + \frac{2h_2}{\xi_{22}} \lambda_2(0, 0) [r_1 - r_3 + \lambda_2(0, h_2) r_5] \right) + \\ & f^{k+1}(0, -h_2) (2r_1 - 2r_3 + r_5 [\lambda_2(0, h_2) - \lambda_2(0, -h_2)]) = \\ & + f^k(h_1, 0) (-r_3 r_7 + \lambda_2(0, h_2) r_5 r_7 + r_2 - \lambda_1(h_1, 0) r_4) \\ & + f^k(h_1, 0) (r_3 r_7 - \lambda_2(0, h_2) r_5 r_7 + r_2 + \lambda_1(-h_1, 0) r_4 - 2r_1) \end{aligned}$$

$$+f^k()\left(\frac{r_{12}h_2}{\xi_{22}}\lambda_2(0,0)+1-2r_2\right)+f^k(-h_1,-h_2)2r_1 \quad (3.45)$$

$$\text{avec } r_7 = \frac{\xi_{21}h_2}{\xi_{22}h_1}$$

Second demi-pas: paroi L^B

Pour le second demi-pas sur la paroi L^B on réutilise les arguments du premier demi-pas pour la paroi L^A . On obtient alors que:

$$f^{k+1}(0,h_2) = \frac{2h_2}{\xi_{22}}\lambda_2(0,0)f^{k+1}() + f^{k+1}(0,-h_2) - \frac{\xi_{21}h_2}{\xi_{22}h_1}f^{k+1}(h_1,0) - f^{k+1}(-h_1,0) \quad (3.46)$$

$$\begin{aligned} f^{k+1}(h_1,h_2) &\approx f^{k+1}(h_1,0) - f^{k+1}(-h_1,0) + f^{k+1}(0,h_2) - f^{k+1}(0,-h_2) \\ &\quad + f^{k+1}(-h_1,-h_2) \end{aligned} \quad (3.47)$$

ce qui substitué dans l'expression du second demi-pas donne:

$$\begin{aligned} &f^{k+2}()(1-2r_1+2r_2) + f^{k+1}(h_1,0)(r_1-r_2+\lambda_1(h_1,0)r_4) + \\ &f^{k+2}(-h_1,0)(r_1-r_2-\lambda_1(-h_1,0)r_4) = f^{k+1}(-h_1,-h_2)2r_1 \\ &+ f^{k+1}()[1-2r_3 + [\lambda_2(0,0)\frac{2h_2}{\xi_{22}}][-\lambda_2(0,h_2)r_5 + r_3] \\ &+ f^{k+1}(-h_1,0)[-r_1 + r_7r_3 - r_5r_7\lambda_2(0,h_2)] \\ &+ f^{k+1}(h_1,0)[r_1 - r_7r_3 + r_5r_7\lambda_2(0,h_2)] \end{aligned}$$

$$+f^{k+1}(0, -h_2)(-2r_1 + 2r_3 - r_5[\lambda_2(0, h_2) - \lambda_2(0, -h_2)]) \quad (3.48)$$

Premier demi-pas: paroi L^C

En examinant l'équation du premier demi-pas sur la paroi L^C on voit que les termes $f^k(-h_1, -h_2)$ et $f^k(-h_1, 0)$ apparaissent explicitement dans l'équation. En utilisant la condition frontière $J_1 = 0$ au début du cycle on obtient que:

$$f^k(-h_1, 0) = -\frac{2h_1}{\xi_{11}}\lambda_1(0, 0)f^k() + f^k(h_1, 0) + \frac{\xi_{12}h_1}{\xi_{11}h_2}f^k(0, h_2) - f^k(0, -h_2) \quad (3.49)$$

En utilisant un développement de Taylor autour de $f(0, 0)$ on obtient que:

$$f^k(-h_1, -h_2) \approx -f^k(h_1, 0) + f^k(-h_1, 0) - f^k(0, h_2) + f^k(0, -h_2) + f^k(h_1, h_2) \quad (3.50)$$

Ce qui permet d'écrire l'expression du premier demi-pas:

$$\begin{aligned} & f^{k+1}()(1 - 2r_1 + 2r_3) + f^{k+1}(0, h_2)(r_1 - r_3 + \lambda_2(0, h_2)r_5 + \\ & f^{k+1}(-0, h_2)(r_1 - r_3 - \lambda_3(0, h_2)r_5 = \\ & f^k(h_1, h_2)2r_1 + f^k()[1 - 2r_2 + [\lambda_1(0, 0)\frac{2h_1}{\xi_{11}}][-\lambda_1(-h_1, 0)r_4 - r_2] \\ & + f^k(h_1, 0)[-2r_1 + 2r_2 + r_4[\lambda_1(-h_1, 0) - \lambda_1(h_1, 0)] \\ & + f^k(0, h_2)[-r_1 + r_2r_6 + r_4r_6\lambda_1(-h_1, 0)] \end{aligned}$$

$$+f^k(0, -h_2)[-r_1 - r_2r_6 - r_4r_6\lambda_1(-h_1, 0)] \quad (3.51)$$

Second demi-pas: paroi L^C

En examinant l'équation du second demi-pas sur cette paroi, on constate la présence des termes $f^{k+1}(-h_1, -h_2)$ qui est explicite et de $f^{k+2}(-h_1, 0)$ qui y est implicite.

En utilisant un développement de Taylor autour de $f(0,0)$ on obtient que:

$$\begin{aligned} f^{k+1}(-h_1, -h_2) &\approx -f^{k+1}(h_1, 0) + f^{k+1}(-h_1, 0) - f^{k+1}(0, h_2) + f^{k+1}(0, -h_2) \\ &\quad + f^{k+1}(h_1, h_2) \end{aligned} \quad (3.52)$$

De plus assumons que les composantes explicites satisfont à la condition frontière $J_1 = 0$ de sorte que:

$$f^{k+1}(-h_1, 0) = -\frac{2h_1}{\xi_{11}}\lambda_1(0, 0)f^{k+1}() + f^{k+1}(h_1, 0) + \frac{\xi_{12}h_1}{\xi_{11}h_2}f^{k+1}(0, h_2) - f^{k+1}(0, -h_2) \quad (3.53)$$

Pour déterminer la valeur de $f^{k+2}(-h_1, 0)$ on utilise le fait qu'au second demi-pas on peut utiliser la condition frontière $J_1 = 0$ pour écrire:

$$f^{k+2}(-h_1, 0) = -\frac{2h_1}{\xi_{11}}\lambda_1(0, 0)f^{k+2}() + f^{k+2}(h_1, 0) + \frac{\xi_{12}h_1}{\xi_{11}h_2}f^{k+1}(0, h_2) - f^{k+1}(0, -h_2) \quad (3.54)$$

La substitution de ces expressions dans l'équation du second demi-pas fournit après simplifications:

$$\begin{aligned}
& f^{k+2}() (1 - 2r_1 + 2r_2 - \frac{2h_1}{\xi_{11}} \lambda_1(0,0) [r_1 - r_2 - \lambda_1(-h_1,0)r_4]) + \\
& f^{k+2}(h_1,0) (2r_1 - 2r_2 + r_4 [\lambda_1(h_1,0) - \lambda_1(-h_1,0)]) = \\
& + f^{k+1}(0, h_2) (r_2 r_6 + \lambda_1(-h_1,0)r_4 r_6 + r_3 - \lambda_2(0, h_2)r_5 - 2r_1) \\
& + f^{k+1}(0, -h_2) (-r_2 r_6 - \lambda_1(h_1,0)r_4 r_6 + r_3 + \lambda_2(0, -h_2)r_5) \\
& + f^{k+1}() (1 - 2r_3 - \frac{r_1 2h_1}{\xi_{11}} \lambda_1(0,0)) + f^{k+1}(h_1, h_2) 2r_1 \tag{3.55}
\end{aligned}$$

Premier demi-pas: paroi L^D

En examinant l'équation du premier demi-pas on observe que les composantes appartenant au tampon sont $f^k(-h_1, -h_2)$ et $f^k(0, -h_2)$. De plus sur cette paroi la condition frontière est $J_2 = 0$ de sorte qu'on peut utiliser des arguments similaires à ceux utilisés pour le second demi-pas de la paroi L^C .

En utilisant un développement de Taylor autour de $f(0,0)$ on obtient que:

$$f^k(-h_1, -h_2) \approx -f^k(h_1, 0) + f^k(-h_1, 0) - f^k(0, h_2) + f^k(0, -h_2) + f^k(h_1, h_2) \tag{3.56}$$

Avec la condition frontière $J_2 = 0$ on obtient:

$$f^k(0, -h_2) = -\frac{2h_2}{\xi_{22}} \lambda_2(0,0) f^k() + f^k(0, h_2) + \frac{\xi_{21} h_2}{\xi_{22} h_1} f^k(h_1, 0) - f^k(-h_1, 0) \tag{3.57}$$

et

$$f^{k+1}(0, -h_2) = -\frac{2h_2}{\xi_{22}} \lambda_2(0,0) f^{k+1}() + f^{k+1}(0, h_2) + \frac{\xi_{21} h_2}{\xi_{22} h_1} f^k(h_1, 0) - f^k(-h_1, 0) \tag{3.58}$$

Ce qui substitué dans l'équation du premier demi-pas donne:

$$\begin{aligned}
& f^{k+1}() (1 - 2r_1 + 2r_3 - \frac{2h_3}{\xi_{22}} \lambda_2(0,0) [r_1 - r_3 - \lambda_2(0, -h_2)r_5]) + \\
& f^{k+1}(0, h_2) (2r_1 - 2r_3 + r_5 [\lambda_2(0, h_2) - \lambda_2(0, -h_2)]) = \\
& + f^k(h_1, 0) (r_3 r_7 + \lambda_2(0, -h_2) r_5 r_7 + r_2 - \lambda_1(h_1, 0) r_4 - 2r_1) \\
& + f^k(-h_1, 0) (-r_3 r_7 - \lambda_2(0, -h_2) r_5 r_7 + r_2 + \lambda_1(-h_1, 0) r_4) \\
& + f^k() (1 - 2r_2 - \frac{r_1 2h_2}{\xi_{22}} \lambda_2(0, 0)) + f^k(h_1, h_2) 2r_1
\end{aligned} \tag{3.59}$$

Second demi-pas: paroi L^D

En réutilisant les arguments lors du premier demi-pas pour la paroi L^D on obtient que:

$$\begin{aligned}
f^{k+1}(-h_1, -h_2) & \approx -f^{k+1}(h_1, 0) + f^{k+1}(-h_1, 0) - f^{k+1}(0, h_2) + f^{k+1}(0, -h_2) \\
& + f^{k+1}(h_1, h_2)
\end{aligned} \tag{3.60}$$

Avec la condition frontière $J_2 = 0$ on obtient:

$$\begin{aligned}
f^{k+1}(0, -h_2) & = -\frac{2h_2}{\xi_{22}} \lambda_2(0,0) f^{k+1}() + f^{k+1}(0, h_2) + \frac{\xi_{21} h_2}{\xi_{22} h_1} f^{k+1}(h_1, 0) \\
& - f^{k+1}(-h_1, 0)
\end{aligned} \tag{3.61}$$

Ce qui finalement substitué dans l'équation du second demi-pas conduit à:

$$f^{k+2}() (1 - 2r_1 + 2r_2) + f^{k+2}(h_1, 0) (r_1 - r_2 + \lambda_1(h_1, 0) r_4 +$$

$$\begin{aligned}
& f^{k+2}(-h_1, 0)(r_1 - r_2 - \lambda_1(-h_1, 0)r_4 = \\
& f^k(h_1, h_2)2r_1 + f^k() [1 - 2r_3 + [\lambda_2(0, 0) \frac{2h_2}{\xi_{22}}] [-\lambda_2(0, -h_2)r_5 - r_3] \\
& + f^k(0, h_2)[-2r_1 + 2r_3 + r_5[\lambda_2(0, -h_2) - \lambda_2(0, h_2)]] \\
& + f^k(h_1, 0)[-r_1 + r_3r_7 + r_5r_7\lambda_2(0, -h_2)] \\
& + f^k(-h_1, 0)[r_1 - r_3r_7 - r_5r_7\lambda_2(0, -h_2)]
\end{aligned}
\tag{3.62}$$

Ce qui complète le traitement des parois réfléchissantes.

3.4.2 Paroi de réinsertion de probabilité

Le traitement de la paroi de réinsertion de probabilité se fait en écrivant l'expression (2.38) pour chaque demi-pas où $\vartheta(x, t)$ est calculé explicitement. Ce calcul se fait sur chaque demi-pas, i.e. la probabilité réinsérée est calculée à partir des valeurs les plus récentes de la distribution qui est absorbée.

On remarque que l'expression (2.86) pour $\vartheta(x, t)$, ne fait intervenir que la composante diffusive et est indépendante de la composante de "dérive", cela résulte du fait que pour un mouvement brownien la composante de bruit agit beaucoup plus rapidement que la composante de "dérive". L'information sur les composantes de "dérive" est essentiellement contenue dans les pentes des densités au niveau de la paroi absorbante. Cette circonstance découle des conditions frontières sur la paroi absorbante qui résultent des caractéristiques de vitesse infinie de la composante de diffusion du mouvement brownien. Cette caractéristique est normalement compensée par la propriété de vitesse moyenne nulle de la même composante. Cependant ce mécanisme de compensation ne tient plus à la paroi absorbante où il n'existe plus de flux de probabilité de retour. De sorte qu'immédiatement après un comportement possiblement impulsif à la paroi la densité tombe à zéro sur cette paroi. (Malhamé

et Chong 1985)

L'approximation retenue consiste à évaluer le flux de probabilité J_{2pa} non pas directement sur la paroi absorbante via l'expression théorique développée en (2.90) mais à partir d'un développement de Taylor sur le flux un pas avant la paroi absorbante.

Rappelons l'expression théorique de J_{2pa} :

$$J_{2pa} = -\xi_{22} \frac{\partial f^{pa}}{\partial \theta_2} \quad (3.63)$$

On note, en premier lieu, que le calcul de cette quantité implique l'approximation d'une dérivée première dans un voisinage où la densité a un caractère discontinu, puisqu'on impose brusquement sa nullité. De plus les seules approximations directes possibles pour cette dérivée sont nécessairement de l'ordre de $\mathcal{O}h$, ce qui laisse supposer qu'il faudrait utiliser dans cette région un pas spatial d'au moins un ordre inférieur à celui de toutes nos approximations qui sont d'ordre de $\mathcal{O}h^2$. Notons simplement, qu'il existe des techniques dite de "maillage variable" qui permettent d'avoir des pas spatiaux différents selon les régions où se concentre l'activité du phénomène qu'on étudie. L'implantation d'un tel maillage qui, près de la paroi absorbante, serait plus fin que partout ailleurs est possiblement une solution au problème de l'approximation directe de cette dérivée.

Pour notre part, et en acceptant que "localement" le courant est continu, on a choisi une approximation de l'ordre de $\mathcal{O}h^2$ qui utilise les courants près de la paroi de commutation plutôt que l'approximation directe de la dérivée dans J_{2pa} . De plus cette approximation par les courants est tout à fait conforme à l'idée que l'on se fait du mécanisme "physique" de l'absorption de probabilité. Pour nous l'absorption de probabilité est d'abord une absorption d'un courant qui préexiste à la nullité de la densité. Présentons notre approximation:

Soit J_{2pa} le flux à la paroi absorbante, J_{2pa+} le flux un pas avant la paroi absorbante

et J_{2pa++} le flux deux pas avant la paroi absorbante. Maintenant évaluons $J_{2pa-\epsilon}$ lorsque $\epsilon \rightarrow 0$ par un développement de Taylor sur J_{2pa+} , i.e:

$$J_0 = \lim_{\epsilon \rightarrow 0} J_{2pa-\epsilon} = J_{2pa+} - \frac{\partial J_{2pa+}}{\partial \theta_2} h_2 + \frac{\partial^2 J_{2pa+}}{\partial \theta_2^2} h_2^2 \quad (3.64)$$

En utilisant les approximations habituelles on obtient que:

$$J_0 = \lim_{\epsilon \rightarrow 0} J_{2pa-\epsilon} = J_{2pa+} - J_{2pa++}/2 + J_{2pa}/2 \quad (3.65)$$

On utilise alors J_0 au lieu de J_{2pa} dans l'expression de la probabilité réinsérée (2.86).

3.4.3 Paroi absorbante

Le traitement de la paroi absorbante se fait simplement en écrivant les expressions du premier et du second demi-pas un pas avant la paroi absorbante avec toutes les composantes concernant la fonction de densité sur la paroi à zéro. De plus on écrit une équation supplémentaire pour chaque demi-pas directement à la paroi absorbante assurant la nullité de la fonction de densité en ce point.

3.4.4 Discontinuité Paramétrique

Le traitement des discontinuités paramétrique se fait en écrivant les expressions qu'on vient de dériver pour le premier et le second demi-pas mais en utilisant la moitié des paramètres de "dérive" comme on l'a vu au chapitre 2, section 2.3.6.

La discrétisation de notre modèle étant complétée, voyons maintenant comment planter notre schéma numérique en exploitant au maximum les propriétés de tridiagonalité des matrices résultantes.

Chapitre 4

Implantation du schéma numérique

On décrit ici quels sont les différents modules essentiels à l'implantation du schéma numérique et la méthode de solution proposée.

4.1 Constitution du maillage

La contrainte sur le maillage est: $h_1 = h_2 / \sin \theta$ où rappelons-le, θ est lié à la pente, c , de la droite de commutation par $\theta = \tan^{-1}(c)$. La procédure de calcul du maillage génère h_1 , h_2 , n_1 et n_2 satisfaisant aux contraintes sur le maillage et les données initiales qui sont la pente de la droite de commutation c , la largeur de la bande morte ε et la valeur maximum de x_1 et x_2 .

4.2 Calcul de la densité initiale

Cette procédure génère la densité initiale à partir des entrées sur la moyenne et la covariance du système dans (x_1, x_2) . La distribution choisie est une binormale

possédant les caractéristiques suivantes dans (x_1, x_2) :

$$E[x_1] = \mu_1, E[x_2] = \mu_2 \quad (4.1)$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix} \quad (4.2)$$

où $\sigma_1^2 = E[x_1^2] - E[x_1]^2$ et $\sigma_2^2 = E[x_2^2] - E[x_2]^2$

De sorte que la densité conjointe dans (x_1, x_2) [voir Annexe 2] est:

$$f_{x_1, x_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp \left[-\frac{(x_1 - \mu_1)^2}{2\sigma_1^2} - \frac{(x_2 - \mu_2)^2}{2\sigma_2^2} \right] \quad (4.3)$$

qui devient dans (θ_1, θ_2) :

$$\begin{aligned} f_{\theta_1, \theta_2}(\theta_1, \theta_2) &= \cos \theta f_{x_1, x_2}(\theta_1 \cos \theta, -\sin \theta \theta_1 + \theta_2) \\ &= \frac{\cos \theta}{2\pi\sigma_1\sigma_2} \exp \left[-\frac{(\theta_1 \cos \theta - \mu_1)^2}{2\sigma_1^2} - \frac{((-\sin \theta \theta_1 + \theta_2) - \mu_2)^2}{2\sigma_2^2} \right] \end{aligned} \quad (4.4)$$

4.3 Choix de l'origine du système pour l'écriture des matrices

On s'assure du maximum de simplicité dans l'écriture du schéma. La figure qui suit représente le système de coordonnées choisi pour chacune des distributions.

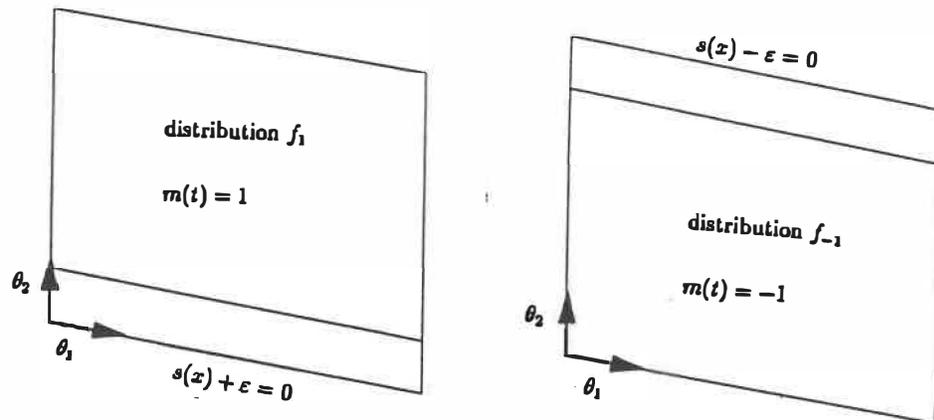


Figure 4.1: Système de coordonnées pour chacune des distributions

4.4 Calcul de l'évolution des densités des états

Soit n_1 le nombre de pas selon θ_1 et n_2 le nombre de pas selon θ_2 , i.e. n_1 le nombre de colonnes et n_2 le nombre de lignes. La résolution du problème via la méthode de fractionnement d'opérateur suppose que l'on balaie le plan selon θ_2 implicite d'abord puis selon θ_1 implicite ensuite. On résout donc le problème tridiagonal du premier demi-pas selon θ_2 en balayant toutes les colonnes du plan (θ_1, θ_2) , puis le système tridiagonal du second demi-pas en balayant toutes les lignes du plan, et ceci en appliquant les méthodes \mathcal{U}_1 et \mathcal{U}_2 respectivement.

On a donc $n_2 \times n_1$ lignes et colonnes à balayer.

où chacune des sous-matrices $G(i)$, $F(i)$ et $R(i)$ représente le système tridiagonal associé à la résolution d'une colonne. Ce système général conserve évidemment la propriété de tridiagonalité de ses sous-matrices constituantes. La solution complète issue du système (4.6) par l'application de la méthode \mathcal{U}_1 constitue la solution du premier demi-pas, i.e. $f^{t+\Delta t/2}$. On adopte la convention d'appeler G_+^1 et G_-^1 ces matrices pour la première et la seconde distribution respectivement.

La figure 4.2 représente la nature du balayage effectué et l'ordre de remplissage de la matrice tridiagonale générale G_+^1 pour le premier demi-pas de calcul de la distribution f_1 .

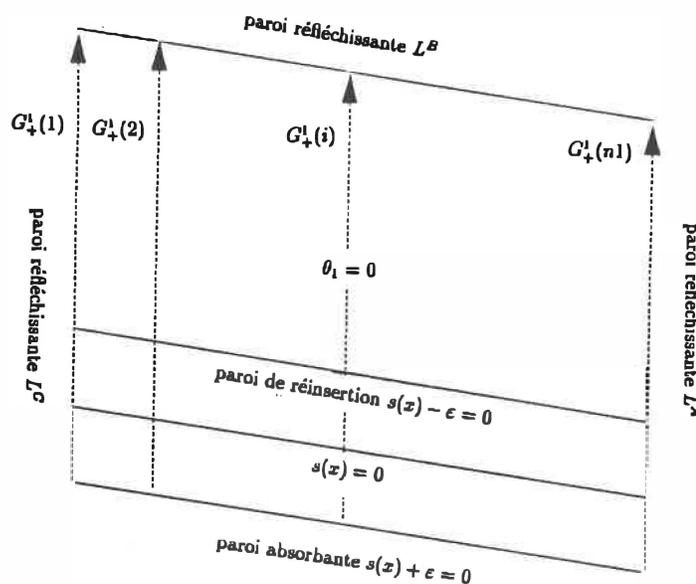


Figure 4.2: Formation de la matrice G_+^1 : premier demi-pas

4.4.2 Etude du second demi-pas

L'application de la méthode \mathcal{U}_2 associée au second demi-pas implique le balayage du système selon θ_1 puisqu'on utilise l'opérateur \mathcal{L}_1^{imp} . On balaie donc le plan selon ses n_2 lignes contenant chacune n_1 éléments. Comme auparavant on obtient pour la

4.4.3 Quelques sous-matrices particulières pour le premier demi-pas

D'abord observons que chaque sous-matrice $G(i)$ contient au moins une ligne spécifiant:

1-le comportement du système sur la paroi de réinsertion de probabilité. On utilise alors le schéma de réinsertion. Notons que puisque la probabilité réinsérée est toujours calculée explicitement, cette ligne est la même que celle qu'on utilise pour le schéma normal, i.e. l'équation (3.27).

2-le comportement du système sur la paroi absorbante. On applique alors la méthode proposée à la section 3.4.3 pour la paroi absorbante, en notant que la densité f_y est nulle.

3-le comportement du système sur les parois réfléchissantes (L^B ou L^D). En ces points on utilise le fait que $J_2 = 0$ et les expressions associées, i.e. (3.46) ou (3.60).

D'autre part une de ces sous-matrices, $G(i)$, spécifie le comportement du système sur la discontinuité paramétrique selon $\theta_1 = 0$, i.e. l'axe θ_2 , où l'on applique la méthode vue en 3.4.4.

Les matrices $G(1)$ et $G(n_1)$ sont uniquement associées à une paroi réfléchissante verticale où l'on applique la condition $J_1 = 0$ et les expressions qui en découlent, i.e. (3.36) ou (3.52).

4.4.4 Quelques sous-matrices particulières pour le second demi-pas

Comme auparavant observons que chaque sous-matrice $G(j)$ contient nécessairement une ligne spécifiant:

1-le comportement du système pour les discontinuités paramétriques selon θ_2 , i.e. $\theta_1 = 0$. On utilise alors le schéma avec la moitié des paramètres de "dérive", toujours selon la méthode vue en 3.4.4.

2-le comportement du système sur chaque paroi réfléchissante (L^A et L^C). En ces points on utilise le fait que $J_1 = 0$ et les expressions qui en découlent, i.e. (3.56) ou (3.42).

Par ailleurs il existe une sous-matrice $G(j)$ spécifiant uniquement le comportement du système sur la droite de réinsertion où l'on utilise le schéma normal (3.28).

D'autre part la matrice, $G(n_2)$, spécifie le comportement du système sur les parois réfléchissantes, L^B ou L^D , et alors on utilise (3.63) ou (3.49).

La matrice $G(1)$ est uniquement associée à la paroi absorbante.

4.4.5 Calcul des matrices tridiagonales générales G

Cette procédure génère automatiquement les entrées des matrices G_+^1 , G_-^1 , G_+^2 et G_-^2 à partir des expressions symboliques associées aux points d'intérêts pour chacun des demi-pas.

4.4.6 Première distribution

Pour la première distribution, $(S + \varepsilon) > 0$, de sorte que les gains sont donnés par l'expression suivante:

$$\psi_i = \begin{cases} \alpha_i & \text{si } x_i(S + \varepsilon) > 0 \\ \beta_i & \text{si } x_i(S + \varepsilon) < 0 \end{cases} \quad (4.8)$$

i.e.

$$\psi_i = \begin{cases} \alpha_i & \text{si } x_i > 0 \\ \beta_i & \text{si } x_i < 0 \end{cases} \quad (4.9)$$

Ajoutons à cette condition $\psi_i = \frac{\alpha_i + \beta_i}{2}$ si $x_i = 0$, ce qui traduit la condition de discontinuité paramétrique telle que présentée dans la section 2.3.6.

De sorte que l'expression générale des gains devient:

$$\psi_i = \begin{cases} \frac{\alpha_i + \beta_i}{2} & \text{si } x_i = 0 \\ \alpha_i & \text{si } x_i(S + \varepsilon) > 0 \\ \beta_i & \text{si } x_i(S + \varepsilon) < 0 \end{cases} \quad (4.10)$$

On sait maintenant que les lignes de chacune des sous-matrices $G(i)$ et $G(j)$ de la matrice tridiagonale générale G peuvent prendre les formes suivantes pour le premier et le second demi-pas de la première distribution.

1-la forme normale (en dehors de toute paroi de réinsertion, d'absorption, de réflexion mais incluant la discontinuité paramétrique via le calcul des gains.)

2-la forme avec condition frontière sur L^B .

3-la forme avec condition frontière sur L^A .

4-la forme avec condition frontière sur L^C .

5-la forme avec condition frontière à l'intersection de L^C et L^B .(coin)

6-la forme avec condition frontière à l'intersection de L^A et L^B .(coin)

7-la forme sur la paroi de réinsertion.

8-la forme sur la paroi absorbante.

4.4.7 Deuxième distribution

Pour la seconde distribution, $(S - \varepsilon) < 0$, les gains sont donnés par l'expression suivante:

$$\psi_i = \begin{cases} \alpha_i & si & x_i(S - \varepsilon) > 0 \\ \beta_i & si & x_i(S - \varepsilon) < 0 \end{cases} \quad (4.11)$$

i.e.

$$\psi_i = \begin{cases} \alpha_i & si & x_i < 0 \\ \beta_i & si & x_i > 0 \end{cases} \quad (4.12)$$

De sorte que l'expression générale des gains devient:

$$\psi_i = \begin{cases} \frac{\alpha_i + \beta_i}{2} & si & x_i = 0 \\ \alpha_i & si & x_i < 0 \\ \beta_i & si & x_i > 0 \end{cases} \quad (4.13)$$

Comme auparavant on associe aux lignes des sous-matrices de G les formes suivantes pour les deux demi-pas:

1-la forme normale

2-la forme avec condition frontière sur L^D .

3-la forme avec condition frontière sur L^A .

4-la forme avec condition frontière sur L^C .

5-la forme avec condition frontière à l'intersection de L^C et L^D .(coin)

6-la forme avec condition frontière à l'intersection de L^A et L^D .(coin).

7-la forme sur la paroi de réinsertion.

8-la forme sur la paroi absorbante.

Puisque que les équations de $x_2 = 0$ et $x_1 = 0$ sont connues dans le plan (θ_1, θ_2) et qu'on s'est assuré que le maillage a une intersection non-vide avec ces droites, le programme peut générer les matrices tridiagonales générales G_+^1 , G_-^1 , G_+^2 et G_-^2 .

Ce faisant le programme doit balayer le plan (θ_1, θ_2) selon les directions associées aux demi-pas et aux deux distributions et:

1- tester la condition (4.10 ou 4.13) et produire les gains ψ_1 et ψ_2 selon la distribution.

2- opérer la transformation du plan (x_1, x_2) au plan (θ_1, θ_2) en calculant la matrice P des gains du système dans ce plan.

3- vérifier la forme de l'équation de Fokker-Planck qui s'applique (1-8).

4-générer l'élément d'un vecteur pointeur de fonction (1-8).

5-générer les éléments de la ligne de G concernée.

Puisque les matrices G sont constantes elles ne sont calculées qu' à l'initialisation. Reste le problème des vecteurs R . Rappelons que ces derniers contiennent des expressions qui font référence aux valeurs les plus récentes de la fonction de densité, i.e les valeurs qui interviennent de façon explicite dans notre schéma. C'est ici que le vecteur pointeur de fonctions intervient. Celui-ci indique comment chaque élément de la matrice R doit être calculé, i.e. quelle forme d'équation de Fokker-Planck

s'applique (1-8) et avec quels gains P . Ce vecteur pointe sur la fonction qui est évaluée et dont la valeur de retour est inscrite dans le vecteur R . La constitution de ce vecteur pointeur s'est faite à l'initialisation.

4.4.8 Cycle de calcul

Supposons les matrices G_+^1, G_-^1, G_+^2 et G_-^2 déjà calculées. Soit une condition initiale sur la densité f_1^t et f_{-1}^t alors le calcul de $f_1^{t+\Delta t}$ et $f_{-1}^{t+\Delta t}$ se fait comme suit:

1-Calculer la matrice R_+^1 en utilisant les valeurs de f_1^t et f_{-1}^t et résoudre le système $G_+^1 F = R_+^1$. Les éléments de la matrice F obtenue sont alors affectés à $f_1^{t+\Delta t/2}$.

2-Calculer la matrice R_-^1 en utilisant les valeurs de $f_1^{t+\Delta t/2}$ et f_{-1}^t et résoudre le système $G_-^1 F = R_-^1$. Les éléments de la matrice F obtenue sont alors affectés à $f_{-1}^{t+\Delta t/2}$.

3-Calculer la matrice R_+^2 en utilisant les valeurs de $f_1^{t+\Delta t/2}$ et $f_{-1}^{t+\Delta t/2}$ et résoudre le système $G_+^2 F = R_+^2$. Les éléments de la matrice F obtenue sont alors affectés à $f_1^{t+\Delta t}$.

4-Calculer la matrice R_-^2 en utilisant les valeurs de $f_1^{t+\Delta t}$ et $f_{-1}^{t+\Delta t/2}$ et résoudre le système $G_-^2 F = R_-^2$. Les éléments de la matrice F obtenue sont alors affectés à $f_{-1}^{t+\Delta t}$. Ce qui complète le cycle de calcul.

4.5 Calcul des moments du système

Cette procédure calcule les moments $E[\theta_1], E[\theta_2], E[\theta_1\theta_2], E[\theta_1^2]$ et $E[\theta_2^2]$ en utilisant les approximations suivantes:

$$E(\theta_1) = \int \int \theta_1 f(\theta_1, \theta_2) d\theta_1 d\theta_2 \approx \sum_i \sum_j \theta_1(i) f(\theta_1(i), \theta_2(j)) h_1 h_2 \cos \theta$$

$$\begin{aligned}
E(\theta_2) &= \int \int \theta_2 f(\theta_1, \theta_2) d\theta_1 d\theta_2 \approx \sum_i \sum_j \theta_2(j) f(\theta_1(i), \theta_2(j)) h_1 h_2 \cos \theta \\
E(\theta_1^2) &= \int \int \theta_1^2 f(\theta_1, \theta_2) d\theta_1 d\theta_2 \approx \sum_i \sum_j \theta_1^2(i) f(\theta_1(i), \theta_2(j)) h_1 h_2 \cos \theta \\
E(\theta_2^2) &= \int \int \theta_2^2 f(\theta_1, \theta_2) d\theta_1 d\theta_2 \approx \sum_i \sum_j \theta_2^2(j) f(\theta_1(i), \theta_2(j)) h_1 h_2 \cos \theta \\
E(\theta_1 \theta_2) &= \int \int \theta_1 \theta_2 f(\theta_1, \theta_2) d\theta_1 d\theta_2 \approx \sum_i \sum_j \theta_1(i) \theta_2(j) f(\theta_1(i), \theta_2(j)) h_1 h_2 \cos \theta
\end{aligned} \tag{4.14}$$

4.6 Équation de transformation des moments

Après avoir simulé le système dans le plan plan (θ_1, θ_2) et en avoir calculé les moments on veut connaître les moments associés dans le plan (x_1, x_2) .

Puisque $X = M^{-1}\Theta$ on a $E(X) = M^{-1}E(\Theta)$ où $\Theta = (\theta_1, \theta_2)^T$.

Maintenant:

$$XX^T = \begin{bmatrix} X_1^2 & X_1 X_2 \\ X_2 X_1 & X_2^2 \end{bmatrix} \tag{4.15}$$

$$\text{Or } XX^T = M^{-1}\Theta[M^{-1}\Theta]^{-1} = M^{-1}\Theta\Theta^T M^{-T}$$

Avec:

$$M^{-1} = \begin{bmatrix} \cos \theta & 0 \\ -\sin \theta & 1 \end{bmatrix} \tag{4.16}$$

et

$$M^{-T} = \begin{bmatrix} \cos \theta & -\sin \theta \\ 0 & 1 \end{bmatrix} \tag{4.17}$$

et

$$\Theta\Theta^T = \begin{bmatrix} \theta_1^2 & \theta_1 \theta_2 \\ \theta_2 \theta_1 & \theta_2^2 \end{bmatrix} \tag{4.18}$$

on obtient que:

$$\begin{aligned}
 XX^T &= \begin{bmatrix} \cos \theta & 0 \\ -\sin \theta & 1 \end{bmatrix} \begin{bmatrix} \theta_1^2 & \theta_1 \theta_2 \\ \theta_2 \theta_1 & \theta_2^2 \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ 0 & 1 \end{bmatrix} \\
 &= \begin{bmatrix} \cos^2 \theta \theta_1^2 & -\sin \theta \cos \theta \theta_1^2 + \cos \theta \theta_1 \theta_2 \\ -\sin \theta \cos \theta \theta_1^2 + \cos \theta \theta_1 \theta_2 & \sin^2 \theta \theta_1^2 - 2 \sin \theta \theta_1 \theta_2 + \theta_2^2 \end{bmatrix} \quad (4.19)
 \end{aligned}$$

De sorte que finalement la matrice de covariance dans (x_1, x_2) est:

$$\begin{aligned}
 E[XX^T] &= \begin{bmatrix} E[X_1^2] & E[X_1 X_2] \\ E[X_2 X_1] & E[X_2^2] \end{bmatrix} \\
 &= \begin{bmatrix} \cos^2 \theta E[\theta_1^2] & -\sin \theta \cos \theta E[\theta_1^2] + \cos \theta E[\theta_1 \theta_2] \\ -\sin \theta \cos \theta E[\theta_1^2] + \cos \theta E[\theta_1 \theta_2] & \sin^2 \theta E[\theta_1^2] - 2 \sin \theta E[\theta_1 \theta_2] + E[\theta_2^2] \end{bmatrix} \quad (4.20)
 \end{aligned}$$

Après avoir montré comment on implante notre schéma numérique nous sommes en mesure de passer à la phase de vérification de la précision et de la stabilité de notre modèle, ce qui fait l'objet du prochain chapitre.

Chapitre 5

Vérification de la précision et stabilité du modèle

Ce chapitre a pour but de vérifier la précision et la stabilité du schéma de simulation. Pour ce faire on se limite au cas où il n'existe pas de discontinuités paramétriques et où les deux ensembles de structures ont les mêmes gains. On montrera que dans ces conditions l'utilisation de l'équation de Riccati constitue un bon moyen de vérification de la précision du modèle.

De plus on montrera que le mécanisme proposé d'échange de probabilité entre les deux ensembles de structures conduit à un schéma précis et stable au sens où la probabilité totale du système est conservée.

5.1 Vérification du modèle sans discontinuités paramétriques et avec structure unique

Notre technique de vérification de la précision du schéma numérique repose sur l'utilisation de l'équation de Riccati. Montrons comment cette dernière constitue une bonne approximation du comportement des moments du système hybride. Il

est clair que lorsque le système existe dans une seule structure et que la densité est concentrée dans une seule distribution, loin de la paroi absorbante, son comportement est essentiellement linéaire. De sorte que les moments solutions de l'équation de Riccati nous permettent de vérifier la validité du schéma numérique dans cette région.

Cette technique de vérification est cependant d'application délicate lorsque la densité se rapproche de la paroi absorbante, i.e. de la région où se manifeste la non-linéarité associée au phénomène d'absorption de probabilité. Le fait qu'il n'existe pas de flux de retour à la paroi absorbante (voir discussion section 3.4.2) implique que l'équation de Riccati est une approximation de moins en moins valable du comportement réel du système. Notons cependant que l'intensité du flux de retour est liée au niveau de bruit, de sorte que si le bruit est faible cette non-linéarité devient négligeable. Dans ces conditions les moments calculés par la simulation de l'équation de Riccati devraient converger vers les moments calculés par la simulation directe du système d'équation de Fokker-Planck couplé. Cependant à mesure que le bruit augmente cette approximation est de moins en moins justifiée. On ne peut plus négliger la perturbation introduite par la présence de la paroi absorbante. On conclut donc que notre technique de vérification est limitée par l'intensité du bruit qu'on utilise. Cela explique pourquoi dans toutes les simulations qui suivent on utilise un bruit relativement faible.

Un aspect crucial est évidemment la vérification de la conservation de la probabilité totale du système. La conservation de cette probabilité est indicative de la pertinence du mécanisme proposé pour le transfert de probabilité d'une distribution à l'autre.

5.2 Équation de Riccati

Soit le système dans une seule structure dont l'équation d'état est:

$$\dot{X} = FX + Gw(t) \quad (5.1)$$

Alors on a clairement que:

$$E(\dot{X}) = FE(X) \quad (5.2)$$

De plus on peut montrer [Maybeck,1970] que l'équation de Riccati associée à (5.1) est:

$$\dot{R}^T(t) = FR(t) + R(t)F^T + GQ(t)G^T \quad (5.3)$$

où R est la matrice d'entrées $R_{ij}(t) = E(x_i(t)x_j(t))$ et $Q(t) = \sigma^2\delta(t)I$

Dans notre cas on a vu que:

$$GQ(t)G^T = \begin{bmatrix} K_1^2 & K_1K_2 \\ K_1K_2 & K_2^2 \end{bmatrix} \quad (5.4)$$

Il est facile de montrer que le modèle d'état suivant est équivalent à l'équation de Riccati (Fowler 1987):

$$\begin{bmatrix} \dot{R}_{11}(t) \\ \dot{R}_{12}(t) \\ \dot{R}_{22}(t) \end{bmatrix} = \begin{bmatrix} 2F_{11} & 2F_{12} & 0 \\ F_{21} & F_{11} + F_{22} & F_{12} \\ 0 & 2F_{21} & 2F_{22} \end{bmatrix} \begin{bmatrix} R_{11}(t) \\ R_{12}(t) \\ R_{22}(t) \end{bmatrix} + \begin{bmatrix} K_1^2 \\ K_1K_2 \\ K_2^2 \end{bmatrix} \quad (5.5)$$

avec $R_{12}(t) = R_{21}(t)$.

Les figures qui suivent présentent les résultats de simulation sur un même système avec des intensités de bruits relativement faibles et des pentes différentes.

Rappelons que la dérivation du terme de réinsertion de probabilité, $v(X, t)$, était intimement liée à la valeur de la pente de la droite de commutation. C'est pourquoi on cherche à vérifier la justesse du schéma pour différentes valeurs de pentes.

5.3 Conservation de la probabilité totale

Un aspect important dans l'étude de la stabilité de notre modèle est la vérification de la conservation de la probabilité totale du système.

Les simulations présentées aux figures 5.1 et 5.2 montrent que le schéma est conservatif et que la validité des approximations pour le mécanisme d'injection ne dépendent pas de la pente de la droite de commutation. On voit que toute la densité réussit à passer d'un ensemble de structures à un autre sans que la probabilité totale du système ne soit modifiée.

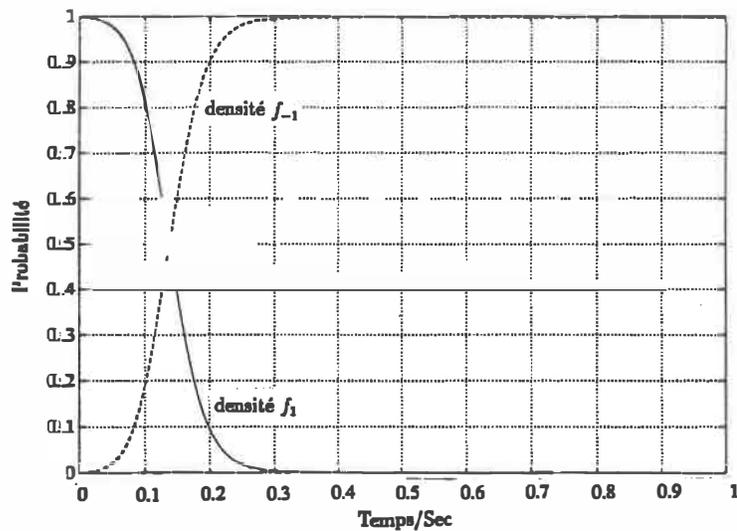


Figure 5.1: Evolution de la probabilité dans chaque distribution. Paramètres: $\theta = 0$, $a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$

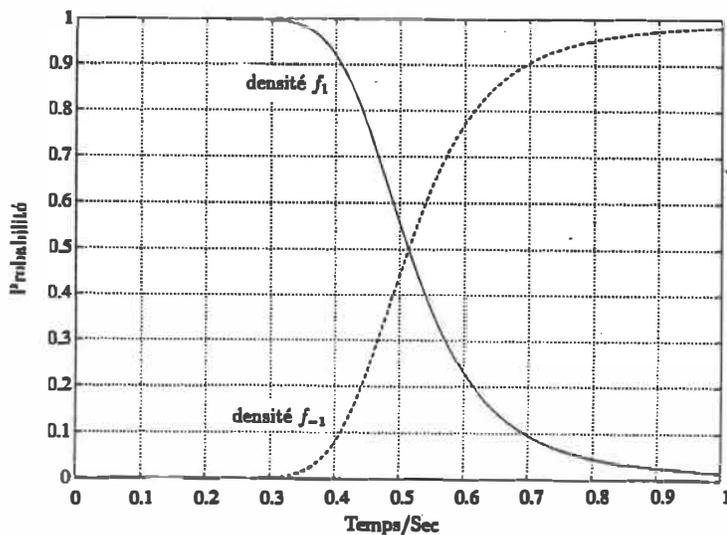


Figure 5.2: Evolution de la probabilité dans chaque distribution. Paramètres: $\theta = 35$, $a=0, b=1, c=-4, d=-4, \psi_1 = 0, \psi_2 = 0, k_1=.04, k_2=.04, dt=.001$

5.4 Précision du schéma au niveau des moyennes

Les figures qui suivent 5.3, 5.4, 5.5 et 5.6 présentent la superposition des résultats de simulation pour $E(x_1(t))$ et $E(x_2(t))$ tels qu'obtenus par la simulation directe des équations de Fokker-Planck et du modèle d'état (5.2) pour des valeurs initiales, $E(x_1(0))$ et $E(x_2(0))$, identiques. Pour ce faire on simule les fonctions de densité du modèle hybride dans le repère (θ_1, θ_2) puis on applique les équations de transformation (4.21) pour obtenir $E(x_1(t))$ et $E(x_2(t))$. On utilise alors $E(x_1(0))$ et $E(x_2(0))$ comme condition initiale du modèle d'état (5.2). Remarquons qu'on utilise un même pas temporel dans le simulateur du modèle hybride et dans le simulateur de l'équation d'état. Cette dernière simulation est effectuée sur MATLAB en utilisant la fonction LSIM qui est essentiellement une méthode d'intégration de Runge-Kutta. On remarque la grande précision du schéma obtenu.

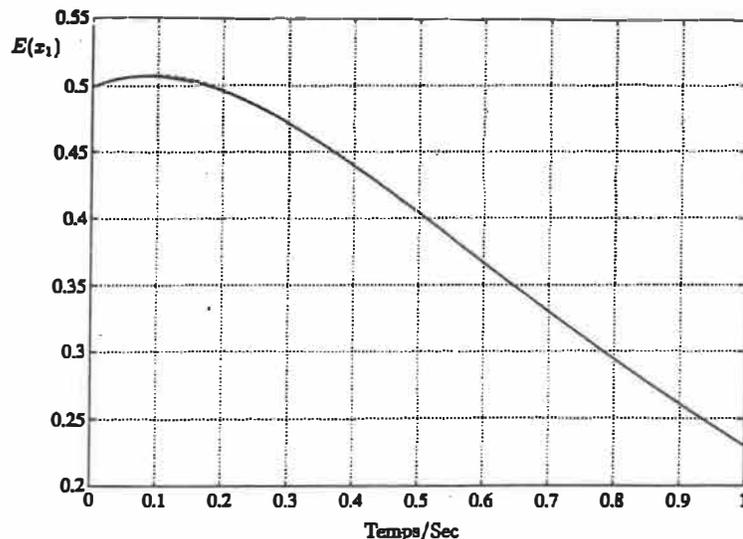


Figure 5.3: Evolution de $E(x_1)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

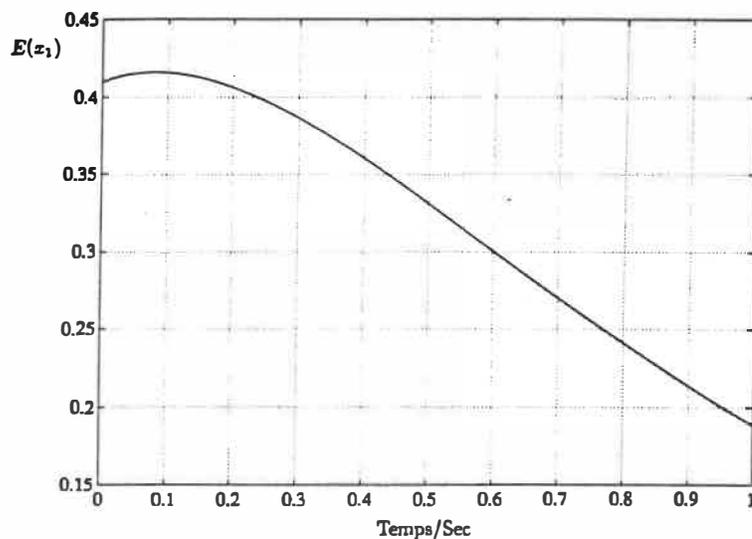


Figure 5.4: Evolution de $E(x_1)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

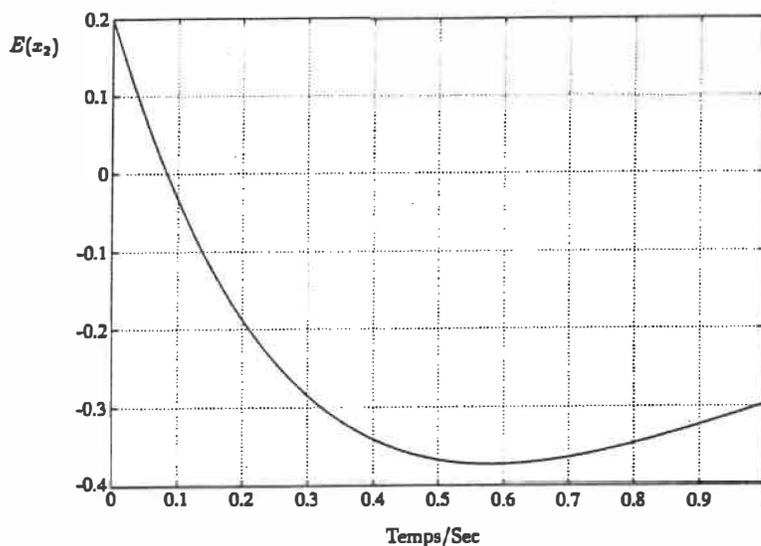


Figure 5.5: Evolution de $E(x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

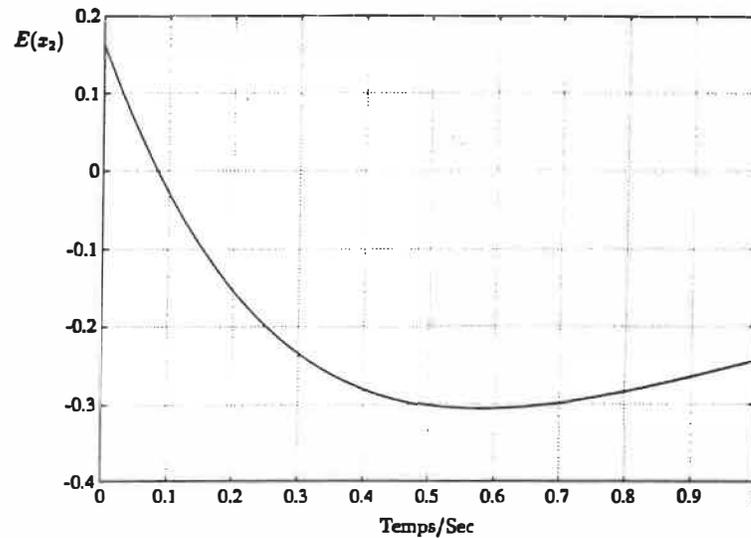


Figure 5.6: Evolution de $E(x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

5.5 Précision du schéma obtenu au niveau des covariances

Les figures 5.7, 5.8, 5.9, 5.10, 5.11 et 5.12 représentent les résultats obtenus pour les moments $E(x_1^2(t))$, $E(x_2^2(t))$ et $E(x_1(t)x_2(t))$ tels que calculés par la simulation directe des équations de Fokker-Planck et de l'équation de Riccati dont les conditions initiales $E(x_1^2(0))$, $E(x_2^2(0))$ et $E(x_1(0)x_2(0))$ sont calculées dans le simulateur du modèle hybride. La simulation de l'équation de Riccati est faite à partir du modèle d'état (5.4) toujours en utilisant la fonction LSIM de MATLAB.

On remarque encore une fois la grande précision du schéma obtenu et la validité des approximations (4.15) utilisées pour approximer les moments du modèle hybride.

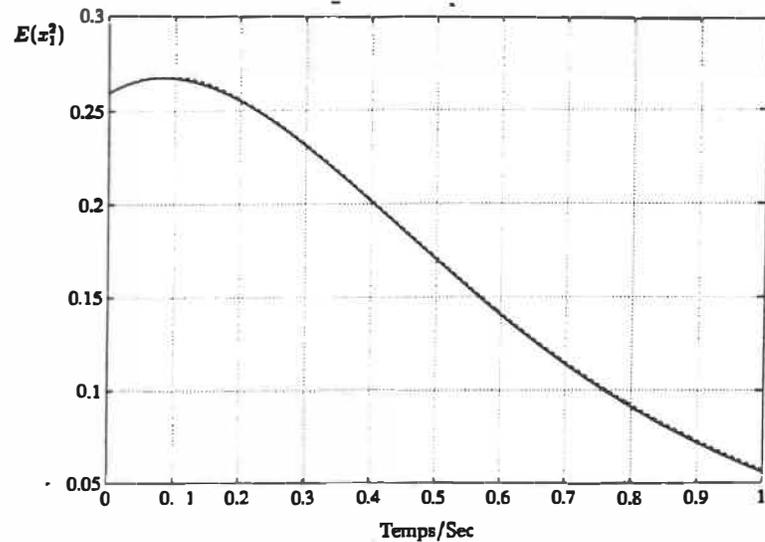


Figure 5.7: Evolution de $E(x_1^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

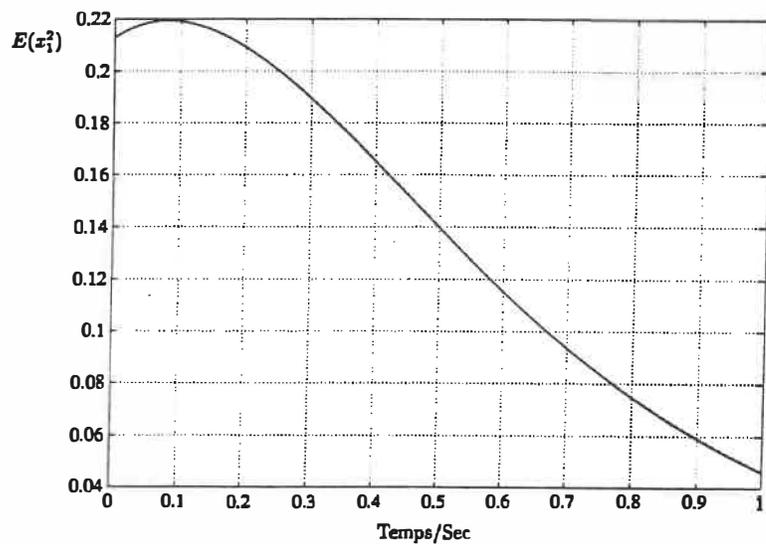


Figure 5.8: Evolution de $E(x_1^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

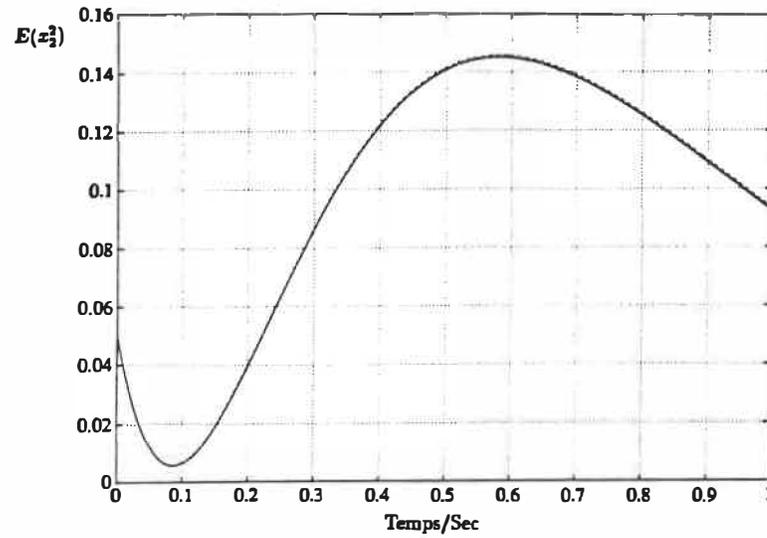


Figure 5.9: Evolution de $E(x_2^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

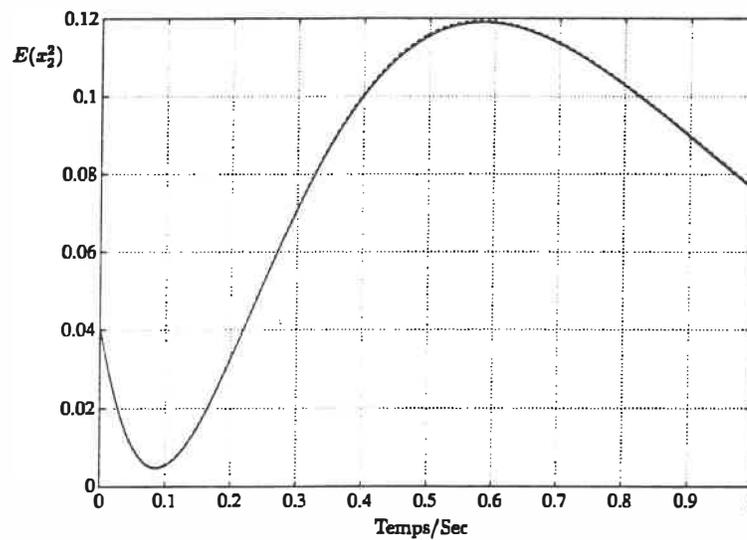


Figure 5.10: Evolution de $E(x_2^2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

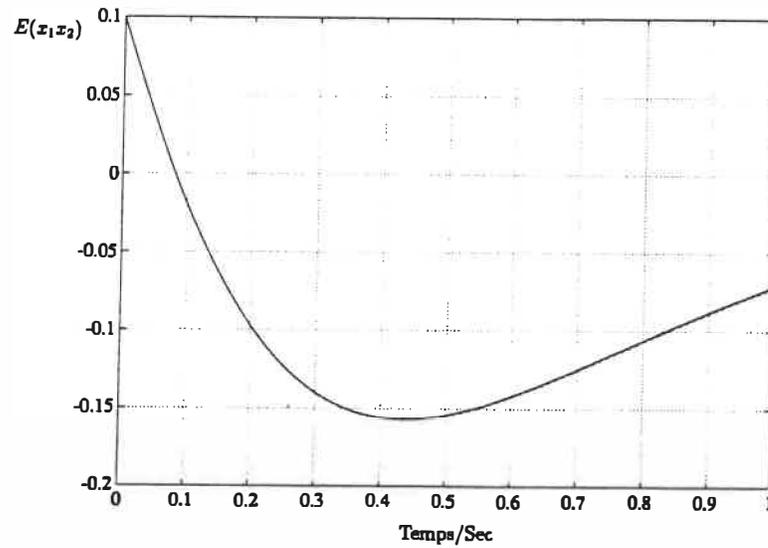


Figure 5.11: Evolution de $E(x_1, x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$; $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

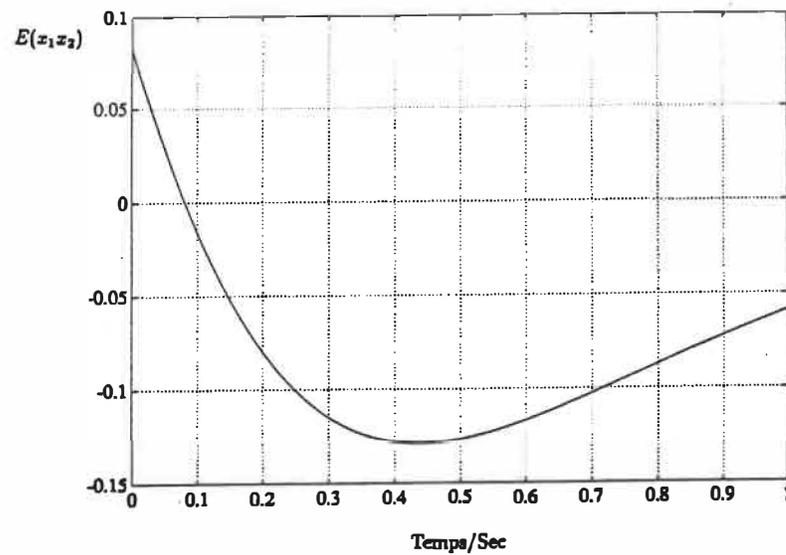


Figure 5.12: Evolution de $E(x_1, x_2)$ par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$; $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

5.6 Hypothèse de linéarité

Afin de bien montrer la validité de notre technique de vérification on présente les simulations dans l'espace de phases. Les figures 5.13, 5.14 permettent d'évaluer l'hypothèse de linéarité en présence de faibles bruits. Notons la très faible déviation entre les trajectoires issues de la simulation modèle hybride et de l'équation de Riccati au voisinage de la paroi de commutation. Remarquons aussi que cette différence disparaît à mesure que la probabilité s'éloigne de la paroi et se reforme dans l'autre distribution.

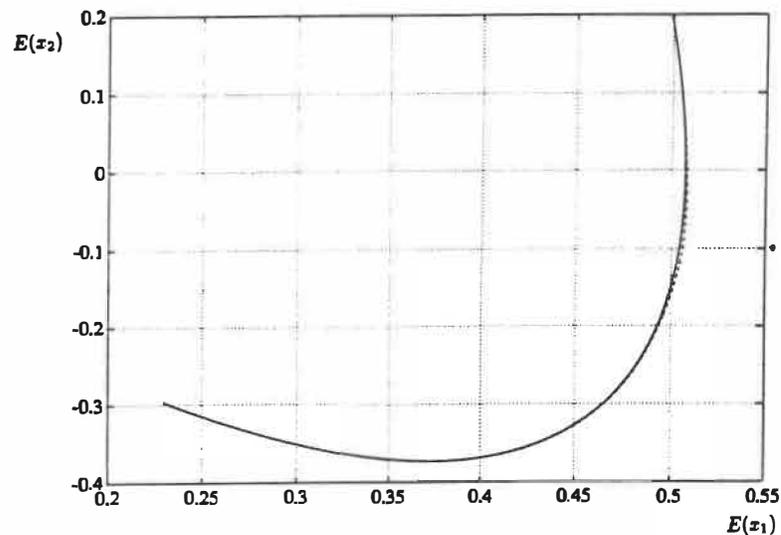


Figure 5.13: Evolution dans l'espace de phases par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 0$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k1=.04$, $k2=.04$, $dt=.001$

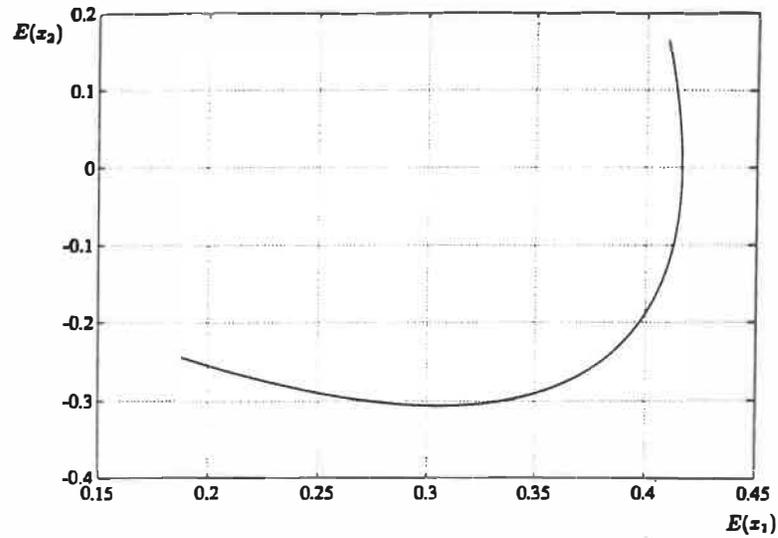


Figure 5.14: Evolution dans l'espace de phases par simulation du système couplé de Fokker-Planck et de l'équation de Riccati. Paramètres: $\theta = 35$, $a=0$, $b=1$, $c=-4$, $d=-4$, $\psi_1 = 0$, $\psi_2 = 0$, $k_1=.04$, $k_2=.04$, $dt=.001$

Après s'être assurés de la validité du schéma numérique proposé pour la simulation du système d'équations de Fokker-Planck couplées, passons à la simulation du système en mode de glissement et à l'étude des propriétés de ses moments solutions.

Chapitre 6

Conditions d'invariances et modes glissants

Une propriété essentielle des systèmes à structure variable est que sous certaines conditions leur comportement en mode glissant est indépendant des valeurs des perturbations et des variations possibles de ses paramètres dynamiques. On parle alors d'invariance du mode de glissement par rapport au bruit ou aux variations paramétriques. Ce chapitre a pour but de présenter certains résultats de simulation confirmant ces propriétés et leurs implications en ce qui concerne les deux premiers moments de la distribution de probabilité dans l'espace de phases d'un système en mode de glissement.

6.1 Théorie classique de l'invariance

Un problème central de la théorie du contrôle automatique est d'assurer l'invariance du système contrôlé sous la présence de perturbations et de variations de paramètres.

Soit l'équation du système contrôlé:

$$\dot{X} = f(X, t) + B(X, t)u + h(X, t) \quad (6.1)$$

où le vecteur $h(X, t)$ caractérise l'ensemble des facteurs dont l'influence sur le processus doit être éliminé si l'invariance est réalisée.

Une condition générale pour que cela soit possible est que:

$$h(X, t) \in \mathcal{B} \quad \forall X, \forall t \quad (6.2)$$

où \mathcal{B} est l'espace R^n engendré par la base des vecteurs colonnes de B . Si cette condition est satisfaite alors il existe une commande u_k , pour laquelle $Bu_k = -h(X, t)$ et qui assure l'invariance. Notons que généralement $h(X, t)$ n'est pas mesurable de sorte que l'invariance est difficilement réalisable en pratique.

Les systèmes à commande discontinue ont cette propriété importante que si:

$$h(X, t) \in \mathcal{B} \quad \forall X, \forall t \quad (6.3)$$

alors l'équation de glissement sur $S(X) = 0$ est invariante par rapport au vecteur $h(X, t)$ (Drazenovic 1969), (Zinober, El-Ghezawi et Billings 1982). On peut donc assurer l'invariance en mode de glissement même si on ignore la valeur de $h(X, t)$ pourvu que $h(X, t)$ satisfasse cette dernière condition.

Dans la théorie classique de l'invariance on considère généralement qu'il est suffisant d'assurer l'invariance d'un sous-ensemble du vecteur d'état, le vecteur d'observation $Y = CX$ où $Y \in R^k$ et $k < n$. Deux voies ont été empruntées pour régler ce problème. La première approche repose sur le fait que pour assurer l'invariance du système, ce dernier doit être non-commandable par rapport à $h(X, t)$. On doit de plus s'assurer que le vecteur Y est dans le sous-espace non-commandable. Dans

(Wonham 1979) on peut trouver une étude de la classe de systèmes pour lesquels cette condition est vérifiée. La seconde approche vise l'étude des systèmes invariants sous l'action de perturbations engendrées par un système dynamique linéaire à paramètres constants connus mais à conditions initiales inconnues. On cherche alors à faire le design d'un compensateur dynamique qui engendre la dite perturbation, (Davison 1972).

Une contribution importante de la théorie des régulateurs à structures variables est qu'elle permet d'étendre la classe des systèmes étudiés dans ces deux approches, (Young 1982).

Comme on vient de le voir les équations du mode glissant sont invariantes pour la projection du vecteur de perturbations sur le sous-espace \mathcal{B} . La classe de systèmes étudiés par Wonham peut donc s'étendre puisque la condition de non-commandabilité par rapport à $h(X, t)$ est réduite à la condition de non-commandabilité par rapport aux composantes de $h(X, t)$ qui sont orthogonales au sous-espace \mathcal{B} .

D'autre part la classe de perturbations incluse dans la seconde approche peut s'étendre aux modèles de perturbations à paramètres variables dans le temps. En effet on peut montrer qu'en mode glissant l'invariance est réalisée pour des variations de ces paramètres dans le temps. La seule information nécessaire est alors l'intervalle de variations de ces paramètres, (Utkin 1974).

6.2 Invariance du mode de glissement en présence de bruit

Considérons le système linéaire à commande et perturbation scalaires suivant:

$$\dot{X}(t) = AX(t) + Bu(t) + G\omega(t) \quad (6.4)$$

Considérons l'hypersurface $S(X) = P^T X = 0$. En mode de glissement les trajectoires dans l'espace de phases sont toutes dans le sous-espace défini par l'équation $P^T X = 0$. De plus on considère que \dot{X} appartient au même sous-espace de sorte que $P^T \dot{X} = 0$.

Ces deux hypothèses permettent alors de déterminer la valeur d'une commande moyenne, continue, u malgré le caractère discontinu de la commande sur la surface de commutation. Cette commande est appelée la commande équivalente u_{eq} .

En substituant (6.4) dans $P^T \dot{X} = 0$ on obtient que:

$$P^T \dot{X} = P^T AX + P^T B u_{eq} + P^T G\omega = 0 \quad (6.5)$$

De sorte que si $P^T B$ est de plein rang on peut déterminer u_{eq} comme suit:

$$u_{eq} = -(P^T B)^{-1} P^T (AX + G\omega) \quad (6.6)$$

Ce qui substitué dans (6.4) donne l'équation du mode de glissement:

$$\dot{X} = [I - B(P^T B)^{-1} P^T](AX + G\omega) \quad (6.7)$$

Qu'on peut encore écrire sous la forme suivante:

$$\dot{X} = A_{eq}X + G_{eq}\omega \quad (6.8)$$

avec $A_{eq} = [I - B(P^T B)^{-1}P^T]A$ et $G_{eq} = [I - B(P^T B)^{-1}P^T]G$.

Il découle de cette équation que si la condition suivante est satisfaite alors l'équation du mode de glissement est invariante par rapport à la perturbation:

$$[I - B(P^T B)^{-1}P^T]G = 0 \quad (6.9)$$

Si cette condition est satisfaite alors la perturbation ω disparaît de l'équation du mode glissant, la trajectoire sur le mode glissant ne dépend des valeurs du bruit que par le biais des conditions initiales du mode de glissement. Le système devient donc insensible au bruit et ce jusqu'à un certain voisinage de l'origine, comme on le verra par la suite.

On peut montrer (Drazenovic 1969) que la condition (6.9) est satisfaite si le vecteur colonne G est linéairement dépendant du vecteur colonne B , i.e.:

$$\text{rang}[B:G] = \text{rang}B \quad (6.10)$$

On peut remarquer que cette condition d'invariance par rapport au bruit ne dépend d'aucune façon de la matrice A mais que des vecteurs colonnes qui définissent les points d'entrées de la commande et de la perturbation dans le système soit B et G .

Si on étudie maintenant le premier moment de l'équation équivalente on obtient que:

$$\begin{aligned}
E[\dot{X}] &= [I - B(P^T B)^{-1} P^T] E[(AX + G\omega(t))] \\
E[\dot{X}] &= [I - B(P^T B)^{-1} P^T] (AE[X] + GE[\omega(t)]) \\
E[\dot{X}] &= [[I - B(P^T B)^{-1} P^T] A] E[X] \\
E[\dot{X}] &= A_{eq} E[X]
\end{aligned}$$

(6.11)

On voit donc que le mouvement moyen en présence de bruit blanc est indépendant de la nature de la matrice G , on parlera ici d'invariance du mouvement moyen par rapport au bruit.

On voit que si l'objectif du contrôleur est de minimiser un critère de trajectoire moyenne sur le plan glissant, alors en mode glissant cette condition est satisfaite pour des perturbations à large bande ayant des caractéristiques de bruits blancs. Les simulations qui suivent montrent de manière convaincante ce résultat et confirment la validité du modèle équivalent pour le cas d'une hystérèse (à faible largeur de zone morte) au niveau de la commutation.

6.3 Application au mouvement dans l'espace de phase

En général $S(X) = P^T X = px_1 + x_2 = 0$ de sorte que $P = [c, 1]$ avec p la pente de la droite de commutation. Remarquons que P^T n'est rien d'autre que le gradient de $S(X)$. De plus soit $B = [b_1, b_2]^T$ et $G = [k_1, k_2]^T$ alors:

$$A_{eq} = [I - B(P^T B)^{-1} P^T] A = \begin{bmatrix} a_{11}(1 - \alpha b_1 p) - \alpha b_1 a_{21} & a_{12}(1 - \alpha b_1 p) - \alpha b_1 a_{22} \\ a_{21}(1 - \alpha b_2) - \alpha b_2 p a_{11} & a_{22}(1 - \alpha b_2) - \alpha b_2 p a_{12} \end{bmatrix} \quad (6.12)$$

et

$$G_{eq} = [I - B(P^T B)^{-1} P^T] G = \begin{bmatrix} k_1(1 - \alpha b_1 p) - \alpha b_1 k_2 \\ k_2(1 - \alpha b_2) - \alpha b_2 p k_1 \end{bmatrix} \quad (6.13)$$

où $\alpha = 1/pb_1 + b_2$

Notre conjecture est que l'évolution des covariances $R(t)$ en mode glissant est décrite par l'équation différentielle suivante:

$$\dot{R}(t) = A_{eq} R(t) + R(t) A_{eq}^T + G_{eq} Q G_{eq}^T \quad (6.14)$$

De plus si la condition de rang (6.10) est satisfaite alors $G_{eq} = 0$ et l'évolution des covariances n'est déterminée que par les composantes déterministes du système. Notons que dans ce cas la stabilité de A_{eq} conditionne la stabilité de l'équation de Riccati. Cette indépendance des trajectoires des covariances en mode glissant par rapport au bruit constitue ce que l'on conviendra d'appeler l'invariance des covariances par rapport au bruit.

6.4 Invariance par rapport aux variations paramétriques

Considérons la matrice $A = A_v + A_c$ où A_v représente tous les éléments de A qui pourraient subir des variations et A_c le reste des éléments de A .

L'équation correspondante en mode de glissement est alors:

$$\dot{X} = [I - B(P^T B)^{-1} P^T] (A_c X + A_v X + G\omega(t)) \quad (6.15)$$

Il découle de cette équation que si la condition suivante est satisfaite alors les paramètres associés à la matrice A_v disparaissent de l'équation du mode de glissement:

$$[I - B(P^T B)^{-1} P^T](A_v X) = 0 \quad (6.16)$$

On peut montrer (Drazenovic 1969) que cette condition est satisfaite si et seulement si tous les vecteurs de base T_i du sous-espace de commutation, $P^T X = 0$, satisfont à l'équation:

$$\text{rang}[B:A_v T_i] = \text{rang} B \quad (6.17)$$

uniformément pour tout l'hypercube de variation de la matrice A_v .

De sorte que les paramètres A_v n'affectent le système que jusqu'au moment où le système entre en mode de glissement. De même façon qu'auparavant l'utilisation d'une commande qui réduit la durée de cette transitoire implique une faible sensibilité du système face aux variations paramétriques.

Notons que si le mode glissant est réalisé en utilisant uniquement les dérivées alors les conditions d'invariances par rapport aux perturbations et aux paramètres sont simultanément satisfaites. En effet puisque ces systèmes peuvent toujours s'écrire sous la forme canonique suivante:

$$\begin{aligned} x_i &= x_{i+1} \quad i = 1, 2, \dots, n-1 \\ x_n &= \sum_{i=1}^n a_i x_i + u + g \end{aligned} \quad (6.18)$$

alors toutes les lignes des matrices A , B , G sont nulles sauf la dernière.

Une conséquence intéressante du résultat (6.17) est la caractérisation de la nature du mouvement moyen en mode de glissement pour les systèmes à paramètres variables dans le temps. En effet si on suppose que cette partie est décrite par A_v et que la condition de rang (6.17) est satisfaite alors l'équation du mouvement moyen est:

$$\begin{aligned}
 E[\dot{X}] &= E([I - B(P^T B)^{-1} P^T][(A_c X + A_v X + G\omega(t))]) \\
 E[\dot{X}] &= [I - B(P^T B)^{-1} P^T](A_c E[X] + G E[\omega(t)]) + E[[I - B(P^T B)^{-1} P^T] A_v X] \\
 E[\dot{X}] &= [[I - B(P^T B)^{-1} P^T] A_c] E[X] \\
 E[\dot{X}] &= A_{eq} E[X]
 \end{aligned}
 \tag{6.19}$$

où A_{eq} ne dépendra pas des composantes de A_v . Une conclusion similaire s'applique à l'équation de propagation équivalente des covariances si la conjecture est valable.

Illustrons par un exemple particulier ces propriétés pour un système dynamique dans l'espace de phases.

6.5 Résultats de simulation

On a choisi d'illustrer simultanément les propriétés d'invariance par rapport au bruit et aux variations paramétriques.

En posant $b_1 = 0$ dans (6.12) on obtient :

$$A_{eq} = \begin{bmatrix} a_{11} & a_{12} \\ 0 & -p a_{12} \end{bmatrix}
 \tag{6.20}$$

alors l'équation du mode glissant est totalement insensible aux paramètres a_{21} et a_{22} et donc à toute variation de ces paramètres dans le temps.

De plus de (6.13) on tire que:

$$G_{eq} = \begin{bmatrix} k_1 \\ -pk_1 \end{bmatrix} \quad (6.21)$$

de sorte que si $k_1 = 0$ alors $G_{eq} = 0$, le système est insensible au bruit ω pour toute valeur de k_2 . Clairement $\text{rang}[B:G] = \text{rang}[B]$ puisque $B = [0 \ 1]^T$ et $G = [0 \ k_2]$.

Si la conjecture est valable alors l'équation de propagation des covariances en mode de glissement devrait être:

$$\dot{R}(t) = A_{eq}R(t) + R(t)A_{eq}^T \quad (6.22)$$

qui ne dépend que des paramètres a_{11} et a_{22} .

On considère que le système est en mode de glissement lorsque la probabilité $P(X \in [S(X) - \varepsilon, S(X) + \varepsilon]) \geq .85$. La procédure de simulation est la suivante:

- 1-Simuler le système d'équations de Fokker-Planck.
- 2-Calculer la probabilité $P(X \in [S(X) - \varepsilon, S(X) + \varepsilon])$.
- 3-Si $P(X \in [S(X) - \varepsilon, S(X) + \varepsilon]) \geq .85$ conserver les valeurs de $E(x_1)$, $E(x_2)$, $E(x_1x_2)$, $E(x_1x_1)$ et $E(x_2x_2)$.
- 4-Insérer les valeurs de $E(x_1)$, $E(x_2)$ comme conditions initiales dans le simulateur de l'équation équivalente pour la propagation des moyennes; $E[\dot{X}] = A_{eq}E[X]$.
- 5-Insérer les valeurs de $E(x_1x_2)$, $E(x_1x_1)$ et $E(x_2x_2)$ comme conditions initiales

dans le simulateur de l'équation de Riccati équivalente pour la propagation des covariances; $\dot{R}(t) = A_{eq}R(t) + R(t)A_{eq}^T + G_{eq}QG_{eq}^T$.

L'observation du système dans le plan de phases permet de confirmer la validité du modèle équivalent pour la propagation des moyennes. Après une transitoire qui dépend essentiellement des conditions initiales, le système suit effectivement la trajectoire spécifiée par l'équation équivalente $E(\dot{X}) = A_{eq}E(X)$. La figure 6.1. présente la superposition des espérances obtenues via la simulation du modèle hybride et de celles obtenues par la simulation de l'équation équivalente pour la propagation des moyennes.

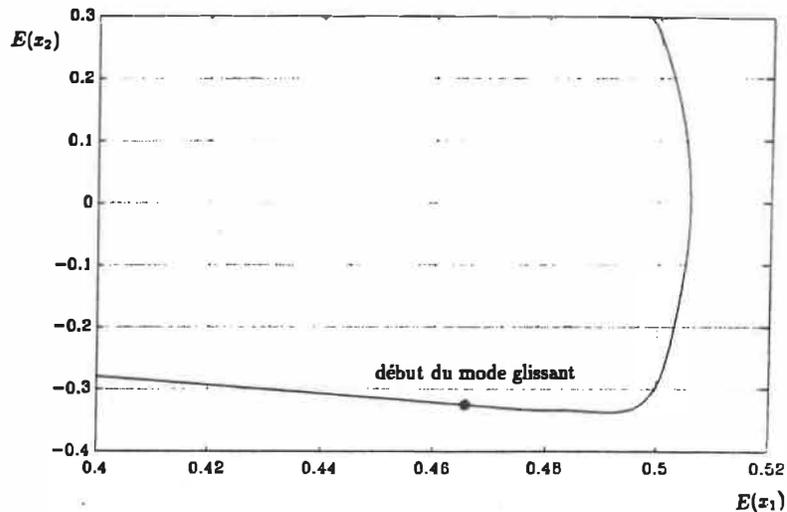


Figure 6.1: Validité du modèle équivalent dans le plan de phase

On présente ici les résultats associés au mode de propagation des covariances. Les figure 6.2, 6.3 et 6.4 présentent la superposition des covariances obtenues via la simulation du modèle hybride et de celles obtenues par la simulation de l'équation équivalente de Riccati pour la propagation des covariances. Ces résultats suggèrent fortement que notre conjecture est plausible !

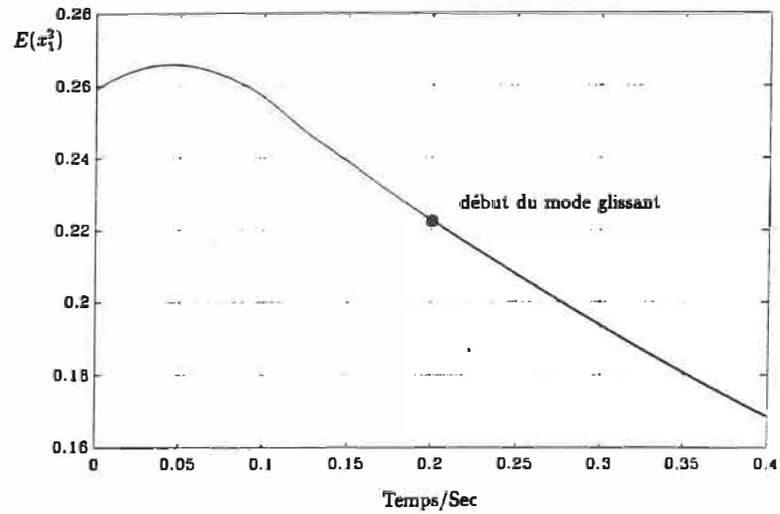


Figure 6.2: Equation de Riccati équivalente pour $\bar{E}x_1^2$

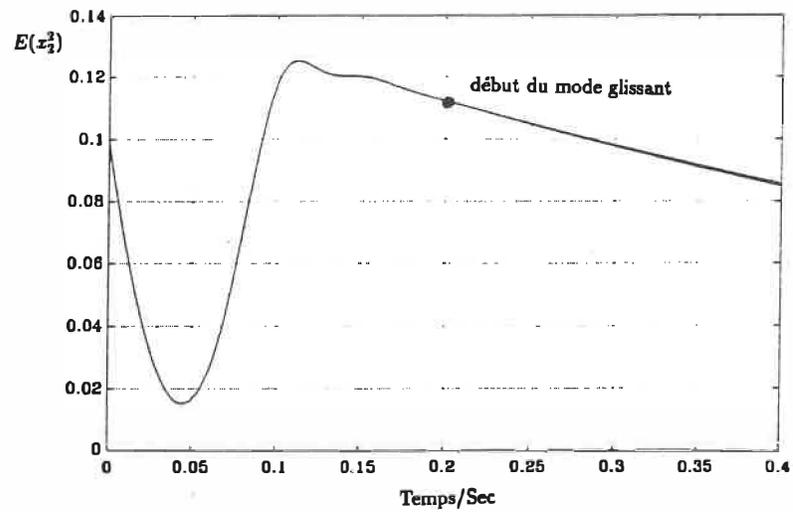


Figure 6.3: Equation de Riccati équivalente pour $\bar{E}x_2^2$

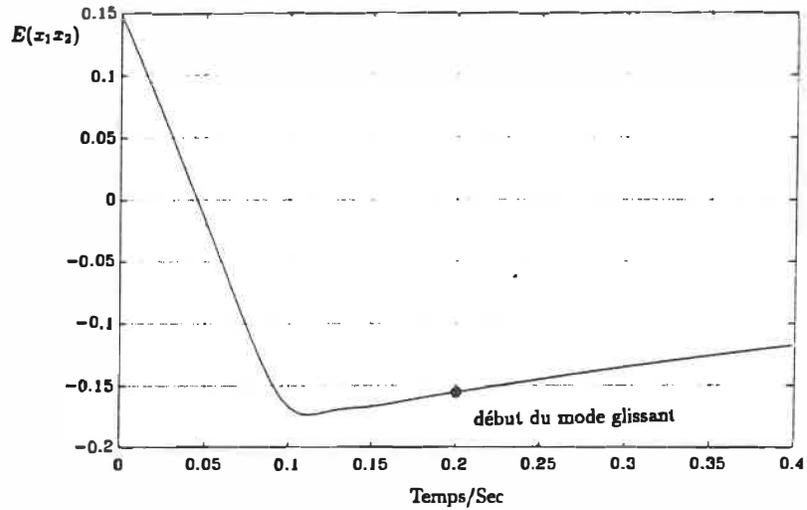


Figure 6.4: Equation de Riccati équivalente pour $E x_1 x_2$

La figure 6.5 permet de vérifier la stabilité (conservation de la probabilité totale du système) ce qui confirme la validité du schéma d'absorption/réinsertion de probabilité.

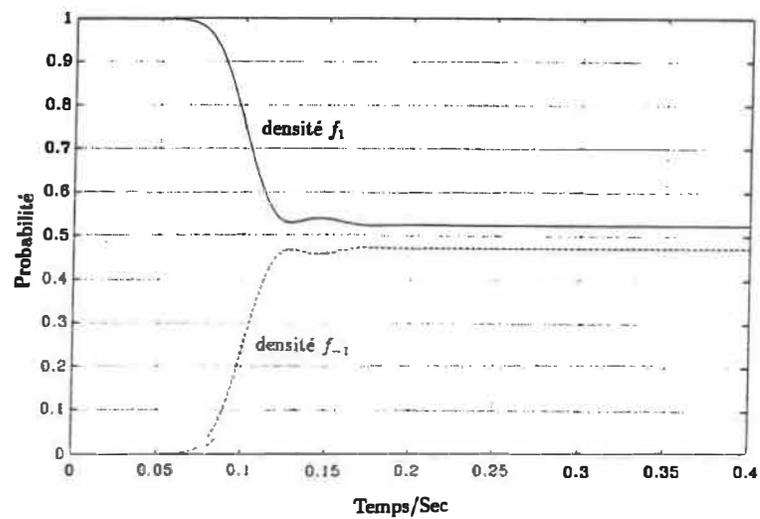


Figure 6.5: Conservation de probabilité

La figure 6.6 montre la dynamique d'échange de probabilité entre les deux struc-

tures. Elle suggère l'existence d'un régime permanent d'échange de probabilités entre les deux structures.

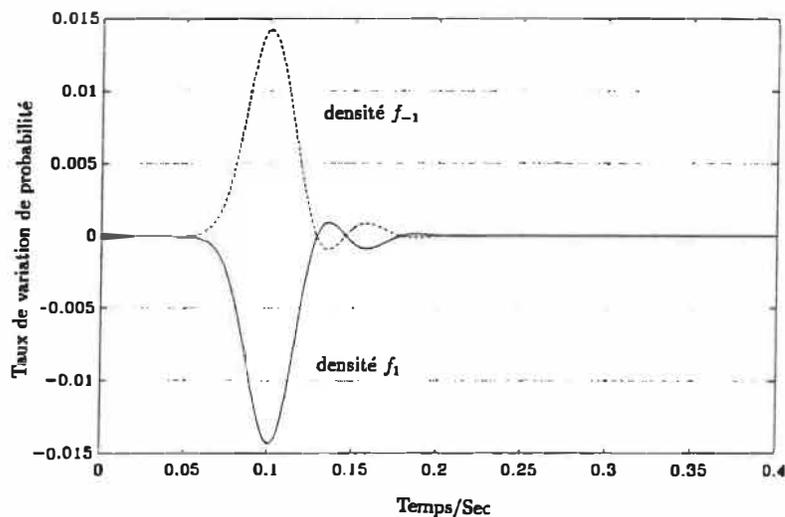


Figure 6.6: Dynamique du transfert de probabilité

6.6 Domaines de validité de l'équation équivalente en présence de bruit

Une question importante est de définir le domaine de validité des équations équivalentes.

Supposons que le design du contrôleur se soit fait en négligeant le bruit. On veut connaître le domaine d'application possible de l'équation de Riccati équivalente en faisant l'hypothèse que cette équation est valable si toutes les réalisations appartenant à ce domaine sont attirées vers le plan glissant avec une probabilité P_{max} jugée satisfaisante.

Soit $F = A + BL$ où L est la matrice de gains calculés à l'aide du design déterministe, alors l'équation d'état du système en présence de bruit est:

$$\dot{X} = FX + G\omega(t) \quad (6.23)$$

où $F = F^+$ si $S(X) > 0$ et $F = F^-$ si $S(X) < 0$

Soit $S(X) = cx_1 + x_2$ avec, $c > 0$, la pente de la droite de commutation.

Évidemment $dS(X) = cdx_1 + dx_2$.

En conformité avec le choix de la fonction de Lyapounov qu'on a fait au chapitre 1 pour étudier les systèmes scalaires, $V(S) = \frac{S^2}{2}$, on veut étudier le signe de la différentielle suivante près du sous-espace de commutation $S(X) = 0$ en présence de bruit:

$$\begin{aligned} d\left(\frac{S^2}{2}\right) &= SdS + \frac{(dS)^2}{2} \\ &= (cx_1 + x_2)(cdx_1 + dx_2) + \frac{c^2(dx_1)^2 + 2cdx_1dx_2 + (dx_2)^2}{2} \\ &= c^2x_1dx_1 + cx_1dx_2 + cx_2dx_1 + x_2dx_2 \\ &\quad + \frac{c^2(dx_1)^2 + 2cdx_1dx_2 + (dx_2)^2}{2} \end{aligned} \quad (6.24)$$

En considérant le modèle d'état suivant:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \omega(t) \quad (6.25)$$

dont l'interprétation formelle est (voir section 1.2.4):

$$\begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} dt + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} dB(t) \quad (6.26)$$

avec $B(t)$ un processus de Wiener standard.

On trouve que:

$$\begin{aligned}
d\left(\frac{S^2}{2}\right) &= c_2 x_1 (F_{11} x_1 + F_{12} x_2) dt + C^2 x_1 K_1 dB(t) \\
&+ c x_1 (F_{21} x_1 + F_{22} x_2) dt + c x_1 K_2 dB(t) \\
&+ c x_2 (F_{11} x_1 + F_{12} x_2) dt + c x_2 K_1 dB(t) \\
&+ x_2 (F_{21} x_1 + F_{22} x_2) dt + x_2 K_2 dB(t) \\
&+ c^2/2((F_{11} x_1 + F_{12} x_2) dt + K_1 dB(t))^2 \\
&+ c((F_{11} x_1 + F_{12} x_2) dt + K_1 dB(t))((F_{21} x_1 + F_{22} x_2) dt + K_2 dB(t)) \\
&+ 1/2((F_{11} x_1 + F_{12} x_2) dt + K_1 dB(t))^2
\end{aligned} \tag{6.27}$$

qui devient en développant:

$$\begin{aligned}
d\left(\frac{S^2}{2}\right) &= c^2 F_{11} x_1^2 dt + c^2 F_{12} x_1 x_2 dt + c^2 x_1 K_1 dB(t) \\
&+ c F_{21} x_1^2 dt + c F_{22} x_1 x_2 dt + c x_1 K_2 dB(t) \\
&+ c F_{11} x_1 x_2 dt + c F_{12} x_2^2 dt + c x_2 K_1 dB(t) \\
&+ F_{21} x_1 x_2 dt + F_{22} x_2^2 dt + x_2 K_2 dB(t) \\
&+ \frac{c^2}{2} [F_{11}^2 x_1^2 dt^2 + F_{11} F_{12} x_1 x_2 dt^2 + K_1 F_{11} x_1 dB(t) dt \\
&+ F_{12} F_{11} x_1 x_2 dt^2 + F_{12}^2 x_2^2 dt^2 + K_1 F_{12} x_2 dB(t) dt \\
&+ F_{11} x_1 K_1 dB(t) dt + K_1 F_{12} x_2 dB(t) dt + K_1^2 dB^2(t)] \\
&+ c [F_{11} F_{21} x_1^2 dt^2 + F_{11} F_{22} x_1 x_2 dt^2 + K_2 F_{11} x_1 dB(t) dt \\
&+ F_{12} F_{21} x_1 x_2 dt^2 + F_{12} F_{22} x_2^2 dt^2 + K_2 F_{12} x_2 dB(t) dt \\
&+ F_{21} K_1 x_1 dB(t) dt + K_1 F_{22} x_2 dB(t) dt + K_1 K_2 dB^2(t)] \\
&+ \frac{1}{2} [F_{21}^2 x_1^2 dt^2 + F_{21} F_{22} x_1 x_2 dt^2 + K_2 F_{21} x_1 dB(t) dt \\
&+ F_{21} F_{22} x_1 x_2 dt^2 + F_{22}^2 x_2^2 dt^2 + K_2 F_{22} x_2 dB(t) dt \\
&+ F_{21} x_1 K_2 dB(t) dt + K_2 F_{22} x_2 dB(t) dt + K_2^2 dB^2(t)]
\end{aligned} \tag{6.28}$$

A présent en négligeant les termes de l'ordre de $(dt)^2$ et en utilisant la propriété centrale du processus de Wiener standard qui veut que $dB(t)$ soit de l'ordre de \sqrt{dt} pour un processus de Wiener standard, alors $dB(t)dt$ est d'ordre $\sqrt[3]{dt}$, terme qui est aussi négligeable.

On obtient:

$$\begin{aligned}
d\left(\frac{S^2}{2}\right) &\approx (c^2 F_{11} + cF_{21})x_1^2 dt \\
&+ (c^2 F_{12} + cF_{22} + cF_{11} + F_{21})x_1 x_2 dt \\
&+ (cF_{12} + F_{22})x_2^2 dt \\
&+ (c^2 K_1 + cK_2)x_1 dB(t) + (cK_1 + K_2)x_2 dB(t) \\
&+ \left(\frac{c^2}{2}K_1^2 + cK_1 K_2 + \frac{1}{2}K_2^2\right)dB^2(t)
\end{aligned} \tag{6.29}$$

où l'on remarque que:

$$(c^2 K_1 + cK_2)x_1 dB(t) + (cK_1 + K_2)x_2 dB(t) = (cK_1 + K_2)S(X)dB(t) \tag{6.30}$$

$$\begin{aligned}
d\left(\frac{S^2}{2}\right) &= (c^2 F_{11} + cF_{21})x_1^2 dt \\
&+ (c^2 F_{12} + cF_{22} + cF_{11} + F_{21})x_1 x_2 dt \\
&+ (cF_{12} + F_{22})x_2^2 dt \\
&+ \left(\frac{c^2}{2}K_1^2 + cK_1 K_2 + \frac{1}{2}K_2^2\right)dB^2(t)
\end{aligned} \tag{6.31}$$

Ce terme doit être analysé avec précaution. D'une part on doit admettre que tout processus de commutation réel est nécessairement non-idéal, doit contenir

une certaine bande morte. Or le processus $dB(t)$ agit sur une échelle de temps qui est beaucoup plus rapide que la réponse physiquement réalisable du commutateur, de sorte qu'on exclut la possibilité d'un changement de signe de $S(X)$ qui soit causé par le terme $dB(t)$. Notons que c'est le seul terme restant qui peut osciller entre des valeurs positives et négatives puisque $dB^2(t)$ est nécessairement positif. D'autre part, et parce que le processus $dB(t)$ est de l'ordre de \sqrt{dt} , le contrôleur va intégrer d'une manière ou d'une autre les fluctuations de $dS(X)$ causées par $dB(t)$, de sorte que si l'on étudie l'effet global sur un intervalle de temps t' , il faudra se pencher sur le signe de $U(t')$ où $U(t')$ est donné par:

$$U(t') = \int_0^{t'} dB(\tau) \quad (6.32)$$

Si on étudie maintenant la différentielle de cette intégrale de Wiener: $dU(t) = U(t + dt) - U(t)$. On peut montrer en utilisant la propriété de martingale que:

$$E[dU(t)|U(t'), 0 \leq t' \leq t] = 0 \quad (6.33)$$

ce qui confirme la propriété de moyenne nulle sur tout intervalle $t - t'$. De sorte que le contrôleur lui-même a cette propriété de moyenne nulle. Il s'ensuit, que sous l'hypothèse de non-idéalité de la commutation, on pourra négliger le terme $S(X)dB(t)$, quant à son action sur le contrôleur.

Remarquons en passant que:

$$E[dU^2(t)|U(t'), 0 \leq t' \leq t] = dt \quad (6.34)$$

de sorte que le système contrôlé n'a pas cette propriété de moyenne nulle par rapport à $dB^2(t)$.

Convenons d'appeler:

$$\begin{aligned}
\gamma_{11} &= (c^2 F_{11} + cF_{21}) \\
\gamma_{12} &= (c^2 F_{12} + cF_{22} + cF_{11} + F_{21}) \\
\gamma_{22} &= (cF_{12} + F_{22}) \\
\phi &= \left(\frac{c^2}{2} K_1^2 + cK_1 K_2 + \frac{1}{2} K_2^2\right)
\end{aligned}
\tag{6.35}$$

On peut montrer que le processus de Wiener standard, $B(t)$, conduit à la fonction de densité gaussienne suivante pour $dB(t) = B(t + dt) - B(t)$ (Larson et Shubert 1984):

$$f_{dB(t)}(u) = \frac{1}{\sqrt{2\pi dt}} \exp^{-u^2/2dt} \tag{6.36}$$

L'écart-type cette distribution est $\sigma = \sqrt{dt}$, de sorte que la probabilité à $3\sigma(t)$ est:

$$P(-3\sqrt{dt} \leq dB(t) \leq 3\sqrt{dt}) = .99903 = P_{max} \tag{6.37}$$

Il suit que:

$$P(dB^2(t) \leq 9dt) = .99903 = P_{max} \tag{6.38}$$

Si on substitue cette borne pour $dB^2(t)$ dans l'équation (6.30), on obtient:

$$d\left(\frac{S^2}{2}\right) \leq (\gamma_{11}x_1^2 + \gamma_{12}x_1x_2 + \gamma_{22}x_2^2 + 9\phi)dt \tag{6.39}$$

avec $dt > 0$, et une probabilité P_{max} .

Pour garantir la validité de la condition de stabilité asymptotique $d\left(\frac{S^2}{2}\right) < 0$, il devient donc suffisant d'assurer que:

$$(\gamma_{11}x_1^2 + \gamma_{12}x_1x_2 + \gamma_{22}x_2^2 + 9\phi) < 0 \tag{6.40}$$

Soit le système suivant:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ K_2 \end{bmatrix} \omega(t) \quad (6.41)$$

avec $u(t) = l_1 x_1 + l_2 x_2$ la commande scalaire et $\omega(t)$ le bruit blanc scalaire. On cherche la matrice de gains L qui fait en sorte que la surface de commutation $S(X) = cx_1 + x_2$ soit attractive. En utilisant les résultats du chapitre 1 (1.13, 1.14) pour un design ne tenant pas compte du bruit, on doit avoir:

$$\alpha = \frac{P^T A}{P^T B} = (2, 1 + c) \quad (6.42)$$

et

$$\begin{aligned} l_2 &< -2 & \text{si } S(X)x_2 < 0 \\ l_2 &> -2 & \text{si } S(X)x_2 > 0 \\ l_1 &> -(1+c) & \text{si } S(X)x_1 < 0 \\ l_1 &< -(1+c) & \text{si } S(X)x_1 > 0 \end{aligned} \quad (6.43)$$

En se limitant à la région où $x_1 > 0$, on obtient les modèles d'états suivant selon que $S(X) > 0$ ou que $S(X) < 0$.

Si $S(X) > 0$ on a:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 - (2 + \delta) & 1 - (1 + c + \delta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ K_2 \end{bmatrix} \omega(t) \quad (6.44)$$

et

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2 - (2 - \delta) & 1 - (1 + c + \delta) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ K_2 \end{bmatrix} \omega(t) \quad (6.45)$$

si $S(X) < 0$ avec $\delta > 0$ un paramètre de design qui caractérise l'ampleur des gains dans chacune des structures.

En utilisant ces modèles et les développements qui précèdent on obtient:

$$\begin{aligned}
 \gamma_{11} &= c\delta \\
 \gamma_{12} &= -c\delta + \delta \\
 \gamma_{22} &= -\delta \\
 \phi &= \frac{1}{2}K_2^2
 \end{aligned}
 \tag{6.46}$$

si $S(X) < 0$.

et

$$\begin{aligned}
 \gamma_{11} &= -c\delta \\
 \gamma_{12} &= c\delta - \delta \\
 \gamma_{22} &= \delta \\
 \phi &= \frac{1}{2}K_2^2
 \end{aligned}
 \tag{6.47}$$

si $S(X) > 0$.

Rappelons que notre condition de stabilité asymptotique est vérifiée (pour le niveau de probabilité qu'on s'est donné!) si les conditions suivantes sont simultanément satisfaites.

$$(\gamma_{11}^- x_1^2 + \gamma_{12}^- x_1 x_2 + \gamma_{22}^- x_2^2 + 9\phi) < 0
 \tag{6.48}$$

si $S(X) < 0$.

et

$$(\gamma_{11}^+ x_1^2 + \gamma_{12}^+ x_1 x_2 + \gamma_{22}^+ x_2^2 + 9\phi) < 0
 \tag{6.49}$$

si $S(X) > 0$.

En réécrivant les deux dernières équations sous une forme hermitienne on obtient:

$$X^T M^\pm X + 9\phi < 0 \quad (6.50)$$

où $M^\pm = M^{\pm T}$ est la matrice hermitienne définie par:

$$M^\pm = \begin{bmatrix} \gamma_{11}^\pm & \frac{\gamma_{12}^\pm}{2} \\ \frac{\gamma_{12}^\pm}{2} & \gamma_{22}^\pm \end{bmatrix} \quad (6.51)$$

Or sait que (Chen 1984):

$$\lambda_{min}^\pm \| X \| \leq X^T M^\pm X \leq \lambda_{max}^\pm \| X \| \quad (6.52)$$

où λ_{min}^\pm et λ_{max}^\pm sont les valeurs propres minimale et maximale de M^\pm .

De sorte qu'une condition suffisante pour que la condition de stabilité ne soit plus satisfaite est que:

$$\lambda_{min}^\pm \| X \| + 9\phi \geq 0 \quad (6.53)$$

et où l'on écrit que la limite de stabilité du plan glissant est:

$$\lambda_{min}^\pm \| X \| + 9\phi = 0 \quad (6.54)$$

Notons que ce résultat est tout a fait général puisque par théorème (Chen 1984), toutes les valeurs propres d'une matrice hermitienne sont strictement réelles.

En appliquant ce résultat à nos données on obtient les graphiques (6.7) et (6.8). Le premier graphique est produit en faisant varier le paramètre de design, δ , de 5 à 15 pour un niveau de bruit de $K_2 = .1$. On constate l'accroissement de la zone de

stabilité à mesure que δ croît. La seconde figure montre l'effet d'une réduction du bruit sur les zones de stabilité pour la même séquence de paramètres de design δ .

On constate donc qu'il existe un voisinage près de l'origine dont la grandeur est proportionnelle au bruit où les réalisations ne sont pas dirigées vers le plan glissant et où par conséquent l'équation de Riccati équivalente n'est pas valable.

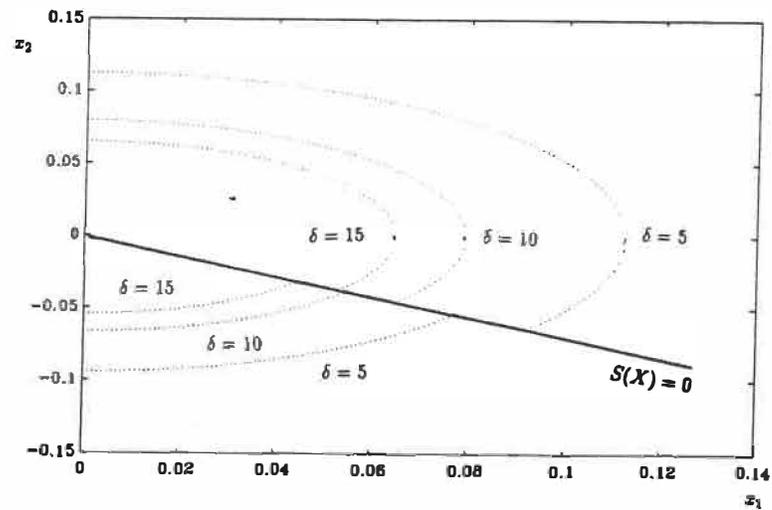


Figure 6.7: Domaines de stabilité en présence du bruit $K_1 = .1$, $\delta = 5, 10, 15$

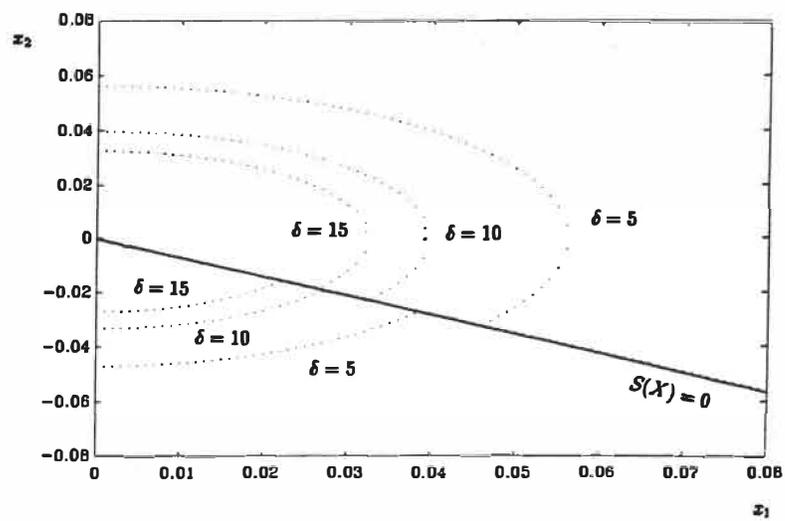


Figure 6.8: Domaines de stabilité en présence du bruit $K_1 = .05$, $\delta = 5, 10, 15$

Conclusion

L'utilisation du concept de modèle stochastique hybride et son application à la modélisation des régulateurs à structure variable en présence de bruit nous a permis de caractériser d'un point de vue probabiliste le comportement de ces systèmes. Ces systèmes étant forcément non linéaires, la théorie actuelle nous fournit peu de résultats analytiques directement utilisables.

Au meilleur de nos connaissances, le modèle stochastique hybride ici proposé est le seul mécanisme actuellement existant qui permet de commencer à saisir le comportement stochastique des systèmes de commande à structure variable. Un des objectifs de ce travail était précisément de fournir un outil de simulation qui puisse, après expérimentation, conduire à certaines conjectures qui plus tard pourraient être converties en théorèmes. Durant ce travail, un effort particulier a été consacré à l'évaluation de la possibilité d'approximer le comportement en mode glissant à l'aide d'équations différentielles linéaires stochastiques, dont les propriétés sont bien connues et qui sont sûrement moins lourdes à simuler que le système d'équations de Fokker-Planck couplées, seul instrument analytique complet. C'est ce qui nous a par exemple amené à examiner l'existence d'une équation de Riccati équivalente pour la propagation des covariances.

L'aspect sûrement le plus difficile du travail accompli a été la formulation d'un modèle d'échange inter-structure de probabilité qui évite l'ambiguïté sur le sens à

donner au courant de probabilité absorbé/réinséré relevée dans les travaux de Kazakov (Kazakov 1977) par Khadr et Loparo (Khadr et Loparo 1983), ces deux derniers étant à notre connaissance les seuls à avoir essayé de modéliser d'un point de vue probabiliste des systèmes hybrides. Notre modèle à l'avantage d'être relativement simple et est globalement axé sur l'utilisation d'un principe unique: la conservation des courants de probabilité. On a pu montrer que les équations de couplage des densités du modèle de Malhamé-Chong pour les éléments chauffants pouvaient être dérivées facilement à l'aide de ce principe.

Pour utiliser efficacement le schéma numérique de fractionnement d'opérateur on a dû travailler dans un système de coordonnées non-orthogonal ce qui nous a permis de constater la généralité de l'expression des équations de Fokker-Planck comme divergence d'un courant de probabilité. Le schéma numérique lui-même présente à notre avis un certain nombre d'innovations. Rappelons qu'il est une variante des schémas à directions alternées implicites. Nous ne connaissons aucun autre travail qui utilise cette variante dans le cadre d'une modélisation stochastique. Les propriétés de stabilité et de précision obtenues sont remarquables compte tenu de la relative simplicité de la méthode retenue pour le fractionnement de l'opérateur infinitésimal de l'équation de Fokker-Planck. De plus le fait d'utiliser le schéma à directions alternées nous a permis d'obtenir un schéma conservatif (stable) même si les réinsertions de probabilité sont faites explicitement, ce qui entraîne nécessairement un délai entre le moment où la probabilité est absorbée et le moment où elle est réinsérée. Une solution potentielle à ce problème serait la parallélisation du schéma numérique. Enfin le traitement des différentes conditions frontières s'est fait de façon naturelle en utilisant des principes élémentaires de conservation de courants et de continuité des fonctions de densité.

On a pu clairement vérifier la validité du modèle équivalent proposé par Utkin pour les systèmes stochastiques avec une non-idéalité faible de commutation. D'autre

part on a pu réexaminer d'un point de vue stochastique les propriétés d'invariance en mode de glissement. Nos simulations nous permettent de croire qu'il existe une équation de type Riccati qui étend ces propriétés d'invariance aux seconds moments de la distribution des états en mode de glissement.

De plus une caractérisation des domaines de stabilité par rapport au plan glissant à l'aide d'une fonction de Lyapunov nous a permis de comprendre et de restreindre l'utilisation de cette équation équivalente. On a pu définir une borne sur la région où le mode glissant peut continuer à exister en présence de bruit. La matrice hermitienne qui sert à définir ce domaine est dérivée à partir de l'analyse de la différentielle stochastique d'une fonction de Lyapounov. L'étude des valeurs propres de cette forme hermitienne est extrêmement intéressante dans la mesure où l'on pourrait les lier aux pôles du contrôleur issus du modèle équivalent. Cela pourrait conduire à une méthodologie de design où l'on pourrait à la fois spécifier les pôles du contrôleur et optimiser la région où l'on est assuré d'avoir une grande robustesse. On a vu que la présence d'hystérèse nous a permis de négliger les termes de la forme $S(X)dB(t)$ dans la différentielle de la fonction de Lyapounov. Une question intéressante serait d'étudier le cas limite où la commutation devient idéale.

Une autre question intéressante concerne la possibilité, via l'équation équivalente de Riccati, d'utiliser la matrice du modèle équivalent pour atteindre un objectif de variance minimale. On a vu que si la condition d'invariance par rapport au bruit est satisfaite, alors cette équation ne dépend que des paramètres déterministes associés à la matrice du modèle équivalent. Cette matrice équivalente est elle-même, sous certaines conditions, invariante par rapport aux variations paramétriques. De sorte qu'il est possible de penser à une forme de contrôleur qui assure un objectif de variance minimale jusqu'à la région spécifiée par la forme hermitienne, et qui soit robuste. L'idée serait ici d'utiliser le plan glissant non seulement pour profiler les transitoires des premiers moments, c'est ce qu'on fait généralement avec le plan glissant par un

placement de pôles, mais de profiler aussi les transitoires des covariances en placant les pôles de l'équation de Riccati équivalente dans un contexte d'optimisation dynamique en temps fini. Cette valeur du temps étant spécifiée par le temps moyen d'atteinte de la région définie par la forme hermitienne. Lorsque le système entre dans la région où le plan glissant est détruit on pourrait, en utilisant les fonctions de Lyapounov de Malhamé (Malhamé 1988), s'assurer d'une forme de stabilité en probabilité qu'on tente d'optimiser.

Une conclusion très claire se dégage de ce qui précède, c'est qu'il est possible et rentable d'élaborer une théorie essentiellement stochastique des systèmes à structure variable. Ce travail en représentait un jalon...!

BIBLIOGRAPHIE

- [1] Ames,W.F. "*Nonlinear Partial Differential Equations in Engineering*" Academic Press (1965)
- [2] Astrom,K.J., Wittenmark, B. "*Adaptive Control*" Addison Wesley (1989)
- [3] Coddington,E.A. "*Theory of Differential Equations*" Mc-Graw Hill (1955)
- [4] DeCarlo,R.A, Zak,S.H., Matthews,G.P "*Variable Structure Control of Nonlinear Multivariable Systems*" IEEE Proc.76,No.3 (1988)
- [5] Drazenovic,B. "*The Invariance Conditions in Variable Structure Systems*" Automatica,vol.5 (1969)
- [6] Chen,C.T. "*Linear System Theory and Design*" Holt,Rinerhart and Winston, Orlando (1984)
- [7] Davison,E.J. "*The Output Control of Linear Time-Invariant Systems with Unmeasurable Arbitrary Disturbances*" IEEE Trans., AC-17,No.5 (1972)
- [8] Filippov,A.F. "*Differential Equations with Discontinuous Right-Hand Side*" American Math. Soc. Translations, vol.42 (1964)
- [9] Fowler, T. "*A Numerical Method for the Propagation of Uncertainty in Nonlinear Systems*" Int.J.General Syst., vol 13 (1987)
- [10] Fuller A.T. "*Nonlinear Stochastic Control Systems*" Taylor and Francis Ltd. (1970)
- [11] Gardiner,G.W. "*Handbook of Stochastic Methods*" Springer-Verlag, New York (1983)
- [12] Itô,K. "*Lectures on Stochastic Processes*" Tata Institute, Bombay (1961)
- [13] Karlin.S, Taylor.H "*A Second Course in Stochastic Processes*" Academic Press (1981)

- [14] Kazakov, I.E. "*Probability Analysis of a Class of Systems with Variable Structure*" Automatic and Remote Control, 38(3), (1977)
- [15] Kazakov, I.E. "*Probabilistic Analysis of Change of Operating Mode of Automatic System*" Automatic and Remote Control, 38(1), (1977)
- [16] Khadr, A.S., Loparo, K. "*Variable Structure Stochastic Systems, Conservation Laws and the Method of Characteristics*" Proc. of the Berkeley-Ames Conf. on Nonlinear Problems in Control and Fluid Dynamics (1983)
- [17] Kozin, F. "*A Survey of Stability of Stochastic Systems*" Automatica, Vol.5 (1969)
- [18] Larson, H.J., Schubert, B.O. "*Probabilistic Models in Engineering Sciences*" Holt, Rinehart and Winston, Orlando (1984)
- [19] Malhamé, R., Chong, C.Y. "*Statistical Approach to Physically Based Load Modeling*" U.S. Department of Energy (1983) Contract No. DE-AS01-77ET29116
- [20] Malhamé, R., Chong, C.Y. "*Electric Load Model Synthesis by Diffusion Approximation of a High-Order Hybrid-State Stochastic System*" IEEE Trans. Aut. Control, vol AC-30, No.9 (1985)
- [21] Malhamé, R., "*Sliding Mode Control of Stochastic Systems in the Plane: Some Preliminary Results*" Proc. of 1987 Symposium on the Mathematical Theory of Networks and Systems, Phoenix, Arizona (1988)
- [22] Malhamé, R., Chong, C.Y. "*On the Statistical Properties of a Cyclic Diffusion Process Arising in the Modeling of Thermostat-Controlled Electric Power Systems Loads*" SIAM J. APPL. MATH. Vol.48, No.2 (1988)
- [23] Maybeck, P.S. "*Stochastic Models, Estimation and Control*" Academic Press, New York (1970)
- [24] Paden, B.E., Sastry, S.S. "*A Calculus for Computing Filippov's Differential Inclusion with Application to the Variable Structure Control of Robot Manipulators*" Proc. of 25th Conference on Decision and Control (1986)
- [25] Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. "*Numerical Recipes*" Cambridge University Press (1986)
- [26] Soong, T.T., "*Random Differential Equations in Science and Engineering*" Academic Press (1973)
- [27] Utkin, V.I. "*Sliding Mode in their Applications in Variable Structure Systems Applications*" Mir Publ., Moscow (1974)

- [28] Utkin, V.I. "*Variable Structure Systems with Sliding Modes*" IEEE Trans. Aut. Control, vol AC-22, No.2 (1977)
- [29] Gerald, F.G., Wheatley, P.O. "*Applied Numerical Analysis*" Addison Wesley (1988)
- [30] Wonham, W.M. "*Linear Multivariable Control: A Geometric Approach*" Springer-Verlag (1979)
- [31] Yamenko, N.N. "*The Method of Fractional Steps*" Springer-Verlag (1971)
- [32] Young, K.-K.D. "*Disturbance Decoupling by High Gain Feedback*" IEEE Trans., AC-27, No.4 (1982)
- [33] Zinober, A.S.I, El-Ghezawi, O.M.E, Billing, S.A "*Multivariable Variable-Structure Adaptive Model-Following Control Systems*" IEEE Proc.129, No.1 (1982)

Annexe 1: Divergence du courant de probabilité dans le repère non-orthogonal et équation de Fokker-Planck

Démonstration de la validité de la divergence de $\vec{J}(\theta)$ comme expression de l'équation de Fokker-Planck.

Soit J_{1x} et J_{2x} les composantes du flux $\vec{J}(X)$ dans le repère cartésien.

Les composantes de ce flux dans le repère non-orthogonal (θ_1, θ_2) sont:

$$J_1(\theta) = 1/\cos\theta J_{1x}$$

$$J_2(\theta) = \tan\theta J_{1x} + J_{2y}$$

En utilisant la dérivation en chaîne:

$$\frac{\partial J_1(\theta)}{\partial \theta_1} = \frac{\partial J_1(\theta)}{\partial x_1} \frac{\partial x_1}{\partial \theta_1} + \frac{\partial J_1(\theta)}{\partial x_2} \frac{\partial x_2}{\partial \theta_1}$$

$$\frac{\partial J_2(\theta)}{\partial \theta_2} = \frac{\partial J_2(\theta)}{\partial x_1} \frac{\partial x_1}{\partial \theta_2} + \frac{\partial J_2(\theta)}{\partial x_2} \frac{\partial x_2}{\partial \theta_2}$$

On trouve:

$$\begin{aligned} \frac{\partial J_1(\theta)}{\partial x_1} &= \frac{1}{\cos \theta} \frac{\partial J_{1x}}{\partial x_1} & \frac{\partial J_1(\theta)}{\partial x_2} &= \frac{1}{\cos \theta} \frac{\partial J_{1x}}{\partial x_2} \\ \frac{\partial J_2(\theta)}{\partial x_1} &= \tan \theta \frac{\partial J_{1x}}{\partial x_1} + \frac{\partial J_{2x}}{\partial x_1} & \frac{\partial J_2(\theta)}{\partial x_2} &= \tan \theta \frac{\partial J_{1x}}{\partial x_2} + \frac{\partial J_{2x}}{\partial x_2} \end{aligned}$$

Et puisque :

$$\begin{aligned} x_1 &= \theta_1 \cos \theta \\ x_2 &= -\theta_1 \sin \theta + \theta_2 \end{aligned}$$

On tire que:

$$\begin{aligned} \frac{\partial x_1}{\partial \theta_1} &= \cos \theta & \frac{\partial x_1}{\partial \theta_2} &= 0 \\ \frac{\partial x_2}{\partial \theta_1} &= -\sin \theta & \frac{\partial x_2}{\partial \theta_2} &= 1 \end{aligned}$$

De sorte que finalement:

$$\frac{\partial J_1(\theta)}{\partial \theta_1} + \frac{\partial J_2(\theta)}{\partial \theta_2} = \frac{\partial J_{1x}}{\partial x_1} + \frac{\partial J_{2x}}{\partial x_2} = -\frac{\partial f}{\partial t}$$

Annexe 2: Équations de transformation des densités

Ici on cherche à exprimer $f_x(x_1, x_2)$ étant donné $f_\theta(\theta_1, \theta_2)$ connu.

On peut montrer que:

$$f_x(x_1, x_2) = f_\theta(\theta_1, \theta_2) |J(\theta_1, \theta_2)|^{-1}$$

où J est le Jacobien défini par:

$$J = \begin{bmatrix} \frac{\partial g_1}{\partial \theta_1} & \frac{\partial g_1}{\partial \theta_2} \\ \frac{\partial g_2}{\partial \theta_1} & \frac{\partial g_2}{\partial \theta_2} \end{bmatrix}$$

et

$$x_1 = g_1(\theta_1, \theta_2) = \cos \theta \theta_1$$

$$x_2 = g_2(\theta_1, \theta_2) = -\sin \theta \theta_1 + \theta_2$$

On obtient:

$|J(\theta_1, \theta_2)|^{-1} = \frac{1}{\cos \theta}$ de sorte qu'on peut finalement écrire l'équation de transformation suivante:

$$f_x(\cos \theta \theta_1, -\sin \theta \theta_1 + \theta_2) = \frac{f_\theta(\theta_1, \theta_2)}{\cos \theta}$$

ÉCOLE POLYTECHNIQUE DE MONTRÉAL



3 9334 00290916 4

C.
U.
19
G: