
TABLE DES MATIERES

TABLE DES MATIÈRES	1
NOTATION	6
CONVENTIONS	7
VARIABLES DU VEHICULE AUTOMOBILE	7
PARAMÈTRES DU VÉHICULE AUTOMOBILE	8
INTRODUCTION	9
ELÉMENTS D'IDENTIFICATION DE MODÈLES PARAMÉTRIQUES CONTINUS	13
1 INTRODUCTION	13
2 PROPRIÉTÉS DES MODÈLES	14
2.1 LINÉARITÉ DES MODÈLES	14
2.2 PROPRIÉTÉS STRUCTURELLES DES MODÈLES	15
2.2.1 SENSIBILITÉ	15
2.2.2 EXEMPLE	17
2.2.3 IDENTIFIABILITÉ	17
2.2.3.1 Méthode par transformée de Laplace	19
2.2.3.2 Exemple	19
2.2.3.3 Représentation géométrique.....	20
2.2.3.4 Méthode numérique	21
2.2.4 DISCERNABILITÉ.....	21
3 CRITÈRES DE COÛT	23
3.1 PROPRIÉTÉS DES ESTIMATEURS	23
3.2 CRITÈRES QUADRATIQUES DES MOINDRES CARRÉS	25
3.2.1 FORMULATION DU CRITÈRE	25

3.2.2 MINIMISATION DU CRITÈRE DES MOINDRES CARRÉS	26
3.2.3 EFFET DU BRUIT DE MESURE	26
3.3 ESTIMATEUR DU MAXIMUM DE VRAISEMBLANCE	26
3.3.1 BRUIT GAUSSIEN VECTORIEL STATIONNAIRE DE COVARIANCE INCONNUE	28
3.3.2 PROPRIÉTÉS DES ESTIMATEURS AU SENS DU MAXIMUM DE VRAISEMBLANCE	29
4 ROBUSTIFICATION	29
5 IDENTIFICATION DES SYSTÈMES À TEMPS CONTINU	34
5.1 SIMULATION DE SYSTÈME À TEMPS CONTINU	34
5.1.1 CAS GÉNÉRAL (MODÈLE LE OU NON)	35
5.1.2 MODÈLE LE	35
5.2 ESTIMATION DES PARAMÈTRES D'UNE ÉQUATION DIFFÉRENTIELLE	36
5.2.1 MÉTHODES INDIRECTES	36
5.2.1.1 Identification de modèle paramétrique à temps continu à partir de modèle non-paramétrique	37
5.2.1.2 Identification de modèle à temps continu à partir de modèle paramétrique discret	37
5.2.2 MÉTHODES DIRECTES	37
5.3 ESTIMATION DES PARAMÈTRES PHYSIQUES DE SYSTÈMES CONTINUS	38
5.3.1 MÉTHODES INDIRECTES	38
5.3.2 MÉTHODES DIRECTES	38
6 CONCLUSION	39
TECHNIQUES D'OPTIMISATION	41
1 INTRODUCTION	41
1.1 POSITION DU PROBLÈME	42
1.2 APPLICATION À L'IDENTIFICATION	43
2 CRITÈRES D'ARRÊT ET DE PERFORMANCE	43
2.1 CRITÈRES DE PRÉCISION ET DE FIABILITÉ	44
2.2 CRITÈRE D'ARRÊT	45
2.3 ASPECTS NUMÉRIQUES	46
2.4 FONCTIONS DE TEST	47
3 ALGORITHMES	48
3.1 ALGORITHMES DE TYPE GRIMPEURS	48
3.1.1 MÉTHODE DES POLYÈDRES FLEXIBLES	49
3.1.2 MÉTHODE DE NEWTON	51
3.1.3 MÉTHODE DE QUASI-NEWTON	51
3.1.4 CONCLUSION	52
3.2 ALGORITHMES D'OPTIMISATION GLOBALE	52
4 ALGORITHMES GÉNÉTIQUES	53
4.1 PRINCIPE	54
4.2 CODAGE	54
4.3 INITIALISATION	56
4.4 FONCTION D'ADÉQUATION	56
4.4.1 ADÉQUATION PROPORTIONNELLE	56
4.4.2 ADÉQUATION BASÉE SUR LE RANG	57
4.5 SÉLECTION	58
4.6 RECOMBINAISON	59
4.7 MUTATION	61
4.8 RÉINSERTION	61
4.9 CONVERGENCE	63
5 ALGORITHMES GÉNÉTIQUES HYBRIDES	63
5.1 HYBRIDATION SÉQUENTIELLE	64
5.2 PRINCIPES DARWINIENS	64
5.3 PRINCIPES LAMARCKIENS	65

5.4 DURÉE DES “ VIES GRIMPEURS ”	65
5.5 ALGORITHMES GÉNÉTIQUES À RECOMBINAISON SIMPLEX.....	66
5.6 ALGORITHMES GÉNÉTIQUES “ 2 HYBRIDES ”	67
5.7 EVALUATIONS EXPÉRIMENTALES.....	67
5.8 DES ALGORITHMES POLYVALENTS.....	70
6 CONCLUSION.....	72
VALIDATION	73
1 INTRODUCTION.....	73
2 INVALIDATION DES RÉSIDUS	74
3 INCERTITUDES SUR LES PARAMÈTRES	75
3.1 ERREUR DE BIAIS D'ESTIMATION	75
3.2 ERREUR DE VARIANCE	75
3.2.1 ESTIMATION STATISTIQUE THÉORIQUE : INÉGALITÉ DE CRAMER-RAO.....	76
3.2.1.1 Inégalité de Cramer-Rao.....	77
3.2.1.2 Evaluation stochastique.....	78
3.2.2 REPRÉSENTATION GÉOMÉTRIQUE DE L'ERREUR DE VARIANCE : ISOCRITÈRES.....	79
3.2.2.1 Approche statistique	79
3.2.2.2 Approche heuristique.....	80
3.2.3 RELATIONS GÉOMÉTRIQUES	82
3.2.3.1 Sensibilité/insensibilité.....	82
3.2.3.2 Corrélation	83
3.2.3.3 Bornes de Cramer-Rao.....	83
3.2.3.4 Axes principaux.....	84
3.2.3.5 Exploitation des isocritères.....	84
3.2.4 ESTIMATION EN PRÉSENCE D'ERREUR DE MODÉLISATION.....	85
3.2.4.1 Constat.....	86
3.2.4.2 Causes.....	87
3.2.4.3 Solution.....	88
3.2.4.4 Validation stochastique par Monte-Carlo.....	89
4 POUVOIR PRÉDICTIF	91
4.1 DOMAINE FRÉQUENTIEL	91
4.1.1 ERREUR DE BIAIS.....	91
4.1.2 ERREUR DE VARIANCE.....	91
4.2 DOMAINE TEMPOREL	93
4.2.1 QUANTIFICATION.....	93
4.2.2 ERREUR DE BIAIS ET VARIANCE.....	94
5 CHOIX DU MODÈLE.....	95
5.1 POUVOIR PRÉDICTIF/SIMPLICITÉ DU MODÈLE	95
5.2 INCERTITUDES SUR LES PARAMÈTRES	96
6 CONCLUSION.....	97
APPLICATION AU VÉHICULE AUTOMOBILE	98
1 INTRODUCTION.....	98
2 OBJECTIF.....	99
3 MODÈLES DE CONNAISSANCE	100
3.1 NOTIONS DE DYNAMIQUE VÉHICULE	100
3.1.1 ELÉMENT DE DYNAMIQUE LATÉRALE DU VÉHICULE.....	100
3.1.2 RÉFÉRENTIEL	101

3.1.3 ROULIS.....	102
3.1.4 PNEUMATIQUES	102
3.1.4.1 Angle de dérive.....	103
3.1.4.2 Caractéristique d'adhérence transversale	103
3.1.4.3 Ballant et longueur de relaxation	105
3.2 EQUATIONS DE LA DYNAMIQUE DU CORPS SOLIDE : PREMIERS MEMBRES DES ÉQUATIONS DU MOUVEMENT.....	106
3.3 DOMAINE DE VALIDITÉ, HYPOTHÈSES DES MODÈLES	107
3.4 MODÈLE LAROUDÉBAL.....	108
3.4.1 FORCES ET MOMENTS APPLIQUÉS SUR LE VÉHICULE.....	108
3.4.1.1 Accélération transversale.....	109
3.4.1.2 Moments dynamiques.....	109
3.4.1.3 Effort transversal à la roue.....	109
3.4.1.4 Egalisation, matrices d'états	110
3.4.2 DYNAMIQUES MISES EN JEU	111
3.5 MODÈLE LAROUDÉ	112
3.6 MODÈLE LADÉBAL.....	112
3.7 MODÈLE LADÉ	113
4 PROPRIÉTÉS STRUCTURELLES DES MODÈLES.....	113
4.1 SENSIBILITÉ	114
4.2 IDENTIFIABILITÉ	117
4.2.1 MODÈLE LAROUDÉBAL.....	117
4.2.2 MODÈLE LAROUDÉ.....	119
4.2.3 MODÈLE LADÉBAL.....	119
4.2.4 MODÈLE LADÉ.....	119
4.3 DISCERNABILITÉ	120
4.4 CONCLUSION	120
5 PROTOCOLE EXPÉRIMENTAL	121
5.1 EXCITATION.....	121
5.2 FRÉQUENCE D'ÉCHANTILLONNAGE	122
5.3 PRÉTRAITEMENT DES SIGNAUX	123
5.3.1 DÉRIVE.....	123
5.3.2 BRUIT DE MESURE, OUTLIERS.....	123
5.4 VÉRIFICATION DE LA LINÉARITÉ	123
6 IDENTIFICATION DES PARAMÈTRES PHYSIQUES DU VÉHICULE AUTOMOBILE	125
6.1 ETAT DE L'ART	125
6.2 RÉSULTATS STOCHASTIQUES.....	126
6.2.1 PARAMÈTRES IDENTIFIÉS.....	127
6.2.2 GLOBALITÉ ET PRÉCISION DES SOLUTIONS.....	128
6.2.3 INVALIDATION DES RÉSIDUS.....	129
6.2.4 POUVOIR PRÉDICTIF.....	130
6.2.5 RÉSULTATS EN PRÉSENCE DE BRUIT COLORÉ SUR LES DONNÉES	130
6.3 RÉSULTATS EXPÉRIMENTAUX.....	131
6.3.1 MODÈLE LADÉ.....	132
6.3.1.1 Paramètres identifiés	133
6.3.1.2 Incertitudes sur les paramètres.....	133
6.3.1.3 Globalité et précision des solutions	134
6.3.1.4 Invalidation des résidus	134
6.3.1.5 Pouvoir prédictif	136
6.3.2 MODÈLE LADÉBAL.....	137
6.3.2.1 Paramètres identifiés	137
6.3.2.2 Incertitude sur les paramètres	139
6.3.2.3 Globalité et précision des solutions	140
6.3.2.4 Invalidation des résidus	140
6.3.2.5 Pouvoir prédictif	142
6.3.3 MODÈLE LAROUDÉ.....	143
6.3.3.1 Paramètres identifiés	144
6.3.3.2 Globalité et précision des solutions	145

6.3.4 MODÈLE LAROUDÉBAL.....	145
6.3.4.1 Paramètres identifiés.....	145
6.3.4.2 Incertitudes sur les paramètres.....	147
6.3.4.3 Globalité et précision des solutions.....	147
6.3.4.4 Invalidation des résidus.....	149
6.3.4.5 Pouvoir prédictif.....	150
6.3.5 IDENTIFICATION ROBUSTE.....	151
7 CONCLUSION.....	152
CONCLUSION	154
BIBLIOGRAPHIE	157
ANNEXES	162
1 PARAMÉTRISATION DES ALGORITHMES D'OPTIMISATION.....	162
1.1 FONCTION DE DE JONG.....	162
1.2 FONCTION DE DE GRIEWANK.....	163
2 INVALIDATION DES RÉSIDUS.....	164
2.1 MOYENNE NULLE.....	164
2.2 NORMALITÉ DES RÉSIDUS.....	164
2.2.1 TEST GRAPHIQUE.....	164
2.2.2 TEST STATISTIQUE.....	165
2.3 STATIONNARITÉ.....	166
2.3.1 TEST GRAPHIQUE.....	166
2.3.2 TEST STATISTIQUE.....	166
2.4 INDÉPENDANCE.....	167
2.4.1 TEST GRAPHIQUE.....	167
2.4.2 TEST STATISTIQUE.....	167
3 EXPRESSION DES FORCES ET MOMENTS APPLIQUÉS AU CENTRE DE GRAVITÉ DU VÉHICULE.....	168
3.1 PRINCIPE FONDAMENTAL DE LA DYNAMIQUE.....	168
3.1.1 CALCUL DE L'ACCÉLÉRATION D'ENTRAÎNEMENT :.....	168
3.1.2 CALCUL DE L'ACCÉLÉRATION RELATIVE.....	168
3.1.3 CALCUL DE L'ACCÉLÉRATION DE CORIOLIS.....	169
3.1.4 CONCLUSION.....	169
3.2 THÉORÈME DU MOMENT DYNAMIQUE.....	170
4 VITESSE DU POINT DE CONTACT PNEU/SOL.....	171
5 FONCTIONS DE SENSIBILITÉ DE LA DERIVE ET DE LA VITESSE DE ROULIS.....	172

NOTATION

$\arg \min(J(\underline{\theta}))$	valeur (ou ensemble de valeurs) minimisant $J(\underline{\theta})$
$\arg \max(J(\underline{\theta}))$	valeur (ou ensemble de valeurs) maximisant $J(\underline{\theta})$
$B_x^{\hat{x}}$	bassin d'attraction au voisinage $V_y^{\hat{y}}$.
$\det(M)$	déterminant de la matrice M
$\text{Grad}(\underline{\theta})$	gradient d'un critère $J(\cdot)$ (ou d'une fonction $f(\cdot)$) par rapport aux paramètres, évalué en $\underline{\theta}$
$\text{Hess}(\underline{\theta})$	Hessien d'un critère $J(\cdot)$ par rapport aux paramètres, évalué en $\underline{\theta}$
$\text{Im}(z)$	partie imaginaire du nombre complexe z
$J(\cdot)$	critère de coût à minimiser
LE	modèle dont la sortie est linéaire par rapport à l'entrée
LP	modèle dont la sortie est linéaire par rapport aux paramètres
$\text{med}(x)$	valeur médiane de x
M^T	transposée de la matrice M
$M(\underline{\theta})$	modèle de connaissance de structure $M(\cdot)$ et de paramètres $\underline{\theta}$
$M_F(\underline{\theta})$	matrice d'information de Fisher associée aux paramètres $\underline{\theta}$
na	degré du polynôme $A(z)$ d'un modèle à temps discret
nb	degré du polynôme $B(z)$ d'un modèle à temps discret
ny	nombre de sortie (système MO)
np	nombre de paramètres identifiés (estimation la matrice de covariance du bruit de mesure)
$N(m, v)$	distribution d'une variable aléatoire normale de moyenne m et de variance v
N	nombre de point de mesure
$P_{\underline{\theta}}$	matrice de covariance asymptotique des paramètres $\underline{\theta}$
$P(Y)$	probabilité de l'événement Y
$P(y, \underline{\theta})$	densité de probabilité conjointe de y et de $\underline{\theta}$
$P(y \underline{\theta})$	densité de probabilité de y si les données sont générées à partir d'un modèle de paramètres $\underline{\theta}$ (vraisemblance de y)
$\text{Re}(z)$	partie réelle du nombre complexe z

s.d.	structurellement discernable
s.n.d.	structurellement non discernable
s.g.i.	structurellement globalement identifiable
s.l.i.	structurellement localement identifiable
s.n.i.	structurellement non identifiable
$\text{sgn}(x)$	fonction signe de la valeur x valant $\begin{cases} 1 \text{ si } x > 0 \\ 0 \text{ si } x = 0 \\ -1 \text{ si } x < 0 \end{cases}$
$V_X^{\underline{x}_0}$	voisinage du point \underline{x}_0 dans l'espace paramétrique
$V_Y^{y_0}$	voisinage du point y_0 dans l'espace d'état
\underline{x}^* :	point stationnaire vers lequel converge l'algorithme d'optimisation
$\hat{\underline{x}}$:	solution finale estimée
$\hat{\underline{x}}(k)$:	solution estimée à l'itération k
\underline{x}_0 :	solution supposée connue du problème d'optimisation
$\underline{x}_0(i)$:	$i^{\text{ème}}$ composante du vecteur \underline{x}_0
$\tilde{\underline{x}}$:	erreur d'estimation $\tilde{\underline{x}} = \underline{x}_0 - \hat{\underline{x}}$
y	vecteur des sorties observées sur le système
$y(t)$	vecteur des sorties observées à l'instant t
$\hat{y}(t, \underline{\theta})$	vecteur des sorties du modèle correspondant à $y(t)$
\tilde{y}	erreur de prédiction du vecteur de sortie $\tilde{y} = y - \hat{y}$
v.a.i	variables (ou vecteurs) aléatoires indépendant(e)s
$\varepsilon(t)$	variable aléatoire à moyenne nulle appartenant à une suite de v.a.i.
Λ	matrice de covariance du bruit de mesure
λ	variance du bruit de mesure sur une variable de sortie
$\underline{\theta}_0$	vecteur des paramètres supposé connu d'un système
$\hat{\underline{\theta}}$	vecteur des paramètres, estimé d'un modèle donné
$\hat{\underline{\theta}}_{\text{mv}}$	vecteur des paramètres, par la méthode du maximum de vraisemblance (Maximum Likelihood)
$\tilde{\underline{\theta}}$	erreur d'estimation $\tilde{\underline{\theta}} = \underline{\theta}_0 - \hat{\underline{\theta}}$
θ_i	$i^{\text{ème}}$ composante du vecteur de paramètres $\underline{\theta}$

CONVENTIONS

$i = \{n(k) m\}$	i variant de n à m par pas de k
\dot{x}	dérivée temporelle de x

VARIABLES DU VEHICULE AUTOMOBILE

α_v	angle au volant	rad
δ_1	angle de braquage de la roue avant = α_v / dem	rad
$\dot{\psi}$	vitesse de lacet	rad/s
β	angle de dérive	rad

β_{t1}	angle de dérive du train avant	rad
β_{t2}	angle de dérive du train arrière	rad
β_{p1}	angle de dérive du pneu avant	rad
β_{p2}	angle de dérive du pneu arrière	rad
$\dot{\theta}$	vitesse de roulis	rad/s
γ_t	accélération transversale au centre de gravité	m/s ²
V	vitesse longitudinale moyenne sur la durée d'un essai du véhicule	m/s

PARAMETRES DU VEHICULE AUTOMOBILE

aro	amortissement antiroulis de la voiture	Nm/rad/s
b1	longueur de relaxation du train avant	m
b2	longueur de relaxation du train arrière	m
dem	rapport de démultiplication de la direction	---
D1	rigidité de dérive train avant	N/rad
D2	rigidité de dérive train arrière	N/rad
hg	hauteur du centre de gravité/sol	m
ho	hauteur de l'axe de roulis/centre de gravité	m
h1	hauteur de l'axe de roulis de l'essieu avant/sol	m
h2	hauteur de l'axe de roulis de l'essieu arrière/sol	m
lr	inertie de la masse suspendue autour de l'axe de roulis	kg.m ²
lxx	inertie de roulis de la masse suspendue	kg.m ²
lxz	produits d'inertie roulis – lacet	kg.m ²
lzz	inertie de lacet de l'ensemble du véhicule	kg.m ²
Kp1	rigidité de ballant du train avant	N/m
Kp2	rigidité de ballant du train arrière	N/m
kro	raideur en roulis de la voiture	N/rad
l	distance de l'essieu avant à l'essieu arrière	m
l1	distance de l'essieu avant au centre de gravité	m
l2	distance de l'essieu arrière au centre de gravité	m
M	masse totale du véhicule	kg
ms	masse suspendue	kg
mns	masse non suspendue, mns = M - ms	kg
ε_1	braquage induit par le roulis pour le train avant	---
ε_2	braquage induit par le roulis pour le train arrière	---

INTRODUCTION

Les fantastiques progrès réalisés ces trois dernières décennies dans le domaine de l'informatique et l'avènement de puissants logiciels ont bouleversé les tâches de conception et de régulation de système. Ainsi l'automatique, avec l'utilisation de techniques de commande prédictive par modèle interne, ou plus récemment le diagnostic, sont devenus un gros « consommateurs de modèles » [RIC91]

La tâche de modélisation est donc un des points sensibles de l'automatique. Elle se situe à la frontière entre le monde réel perturbé et bruité et sa représentation mathématique idéale. Dans la mesure du possible, la tâche de modélisation vise l'obtention de modèles robustes vis à vis des différents phénomènes perturbateurs. Ces phénomènes sont, soit de nature extrinsèque au modèle (tel le bruit de mesure, les variations de l'environnement ...), soit de nature intrinsèque au modèle (erreur de caractérisation, variation de paramètres ...). **Les travaux présentés dans ce mémoire traitent de l'identification des paramètres physiques de modèles à temps continu.**

□ Pourquoi identifier en continu ?

Généralement, deux grandes classes de modèles se distinguent en automatique [RIC91], [WAL97] :

- **Les modèles de connaissance** (ou phénoménologiques) qui sont établis à l'aide des lois de la physique.
- **Les modèles de comportement** (ou de représentation) qui se limitent à la reproduction du comportement de systèmes en se basant sur des données mesurées. Ce type de modèle requiert en général aucune connaissance physique sur le processus modélisé.

Si les modèles de comportement sont d'obtention relativement aisée, ils possèdent quelques inconvénients :

- Ils possèdent généralement un **domaine de validité restreint**.
- Leur structure doit être appropriée au processus décrit
- Le processus modélisé peut varier dans le temps. Sa description par le biais de modèles invariants dans le temps est alors problématique.

Ces inconvénients ne sauraient masquer leurs avantages. Généralement, ces modèles sont représentés sous forme de fonctions de transfert discrètes ou sous forme d'équations aux différences. Leur rapidité et simplicité de simulation sont appréciables, notamment pour la commande de processus en temps réel, les nombreuses limitations des modèles comportementaux

ont motivé nos travaux. En effet, le travail présenté dans ce mémoire traite de l'identification de systèmes continus par le biais de modèles de connaissance. En effet, généralement les processus physiques sont à temps continu, il paraît alors naturel de les décrire par le biais de modèles à temps continu. La motivation principale liée à l'utilisation de ce type de modèle dépend de l'usage de celui-ci.

Le but de nos travaux est, par le biais de l'identification de modèles de connaissance, d'améliorer de la connaissance physique d'un système et de permettre une interprétation physique de son comportement.

Malgré tout, par rapport au très nombreux travaux réalisés ces trente dernières années dans le domaine de l'identification de modèles comportementaux, l'identification des paramètres physiques de systèmes à temps continu a fait l'objet d'un nombre de publications plus restreint. Des travaux ont été menés sur plusieurs systèmes : des générateurs électriques [LAI96], des haut-parleurs [KNU96], des robots manipulateurs ([GAU98], [ISE92]) et des véhicules automobiles ([HOR96], [BOT82], [HUA93], [LET95]).

¶ **Objectifs**

Ces travaux s'inscrivent dans une des directions de recherche du groupe de recherche MIAM. Ils visent à plus long terme l'amélioration de la connaissance du comportement de systèmes complexes, et du véhicule automobile en particulier. A ces fins, l'identification de modèles de connaissance est recherchée.

L'élaboration de modèle de connaissance à temps continu entraîne l'identification de coefficients ou de paramètres physiques inconnus. En diagnostic, de nombreuses applications ont récemment été réalisées ([BRU90], [ISE97]). Isermann souligne dans son article [ISE93] que la connaissance des paramètres physiques facilite les tâches de détection et de localisation de défauts.

Les travaux présentés dans ce mémoire ont donc pour but de contribuer à la formalisation des différents points constitutifs d'une stratégie visant l'identification des paramètres physiques de systèmes complexes : l'étude des propriétés structurelles du modèle, la formalisation d'un critère d'erreur, la minimisation de ce critère, la validation du modèle identifié.... En effet, ces différents aspects sont souvent traités de manière distincte et/ou sur des cas d'école.

A des fins de validation, cette démarche sera appliquée à un système complexe par essence même : le **véhicule automobile**. Le véhicule est un système bruité, difficilement modélisable dont les paramètres varient dans le temps (usure, adhérence ...). Le processus de modélisation doit donc être robuste par rapport aux différentes perturbations agissant sur le véhicule.

¶ **Organisation du mémoire**

Le travail décrit dans ce mémoire résulte d'une réflexion qui constitue son fil conducteur. L'organisation de ce document s'organise logiquement autour de 4 chapitres.

Le **premier chapitre** présente les différents **aspects spécifiques à l'identification de modèles de connaissance à temps continu**. Notamment l'accent est mis sur les propriétés structurelles des modèles. Afin que la démarche de modélisation entreprise ait des chances d'aboutir, certaines des propriétés doivent être satisfaites (en particulier les propriétés **d'identifiabilité** des modèles). Il est également montré que l'étude des propriétés de **discernabilité** et de **sensibilité** se révèlent de précieuses sources d'informations complémentaires lorsque des problèmes d'identifiabilité des paramètres apparaissent.

Ensuite les avantages et inconvénients des différentes formulations de critère de coût sont brièvement abordés et le choix d'un estimateur au sens du **maximum de vraisemblance** est justifié.

Afin d'éviter dans la mesure du possible une estimation biaisée des paramètres, toute idée de termes de pondération en faveur de valeur a priori des paramètres est rejetée.

Afin de minimiser l'influence de l'éventuelle présence de données aberrantes, différents schémas de robustification du critère de coût sont ensuite présentés, discutés et évalués. Enfin, les différentes stratégies permettant l'identification des paramètres physiques de modèles à temps continu sont discutées. Une fois la stratégie établie et le critère de coût formulé, les paramètres identifiables sont alors estimés par minimisation du critère d'erreur défini.

Le **second chapitre** présente les différents **algorithmes d'optimisation** à la disposition du modélisateur pour minimiser le critère de coût précédemment défini. En effet, la tâche de minimisation du critère d'erreur, posée par une stratégie d'identification directe des paramètres physiques de système continu, n'est pas triviale. **Le critère de coût n'est pas linéaire en les paramètres** et des algorithmes itératifs d'optimisation non linéaire doivent être employés. Ces algorithmes d'optimisation doivent être en mesure de localiser fidèlement l'optimum global du problème. Devant les limitations inhérentes aux algorithmes classiques d'optimisation de type grimpeur, l'utilisation des algorithmes d'optimisation globale est préconisée. Il est montré que ces derniers parviennent à localiser fidèlement l'optimum global, avec une bonne précision, tout en requérant un temps de calcul raisonnable. Généralement, l'obtention des paramètres au sens du critère d'erreur minimisé n'est pas suffisant, un certain degré de confiance sur les résultats est souhaitable.

Le **troisième chapitre** traite de la **validation de modèles**. Il convient notamment d'associer un certain degré de confiance aux paramètres estimés. Dans le cas d'une estimation non biaisée des paramètres, ces degrés de confiance se quantifient par association d'intervalles d'incertitude aux valeurs des paramètres identifiés. Les propriétés intrinsèques à l'estimateur du maximum de vraisemblance permettent une quantification des intervalles d'incertitude associés aux paramètres estimés, à partir d'outils de la théorie stochastique (inégalité de Cramer-Rao). La validité de ces outils repose essentiellement sur le respect des hypothèses formulées sur la nature des résidus. Une importance toute particulière est donc consacrée au développement d'outils permettant de vérifier ces hypothèses.

En raison de la quasi-inévitable présence d'erreur de caractérisation, l'hypothèse d'indépendance des résidus n'est que partiellement remplie. Il est montré que cela induit une sous-estimation des intervalles d'incertitude prédits par l'inégalité de Cramer-Rao. Une méthode de correction utilisée lors de l'estimation des coefficients aérodynamiques d'aéronef sera présentée. Son efficacité est vérifiée en simulation. Les intervalles d'incertitude des paramètres sont ensuite mis à profit pour associer des intervalles de confiance aux prédictions temporelles et fréquentielles des modèles identifiés.

Au sein du **quatrième chapitre**, les points précédents **sont appliqués à l'identification des paramètres physiques d'un véhicule automobile**. Ces paramètres interviennent dans la modélisation du comportement transversal du véhicule automobile. Après une brève présentation des modèles utilisés, leurs propriétés structurelles sont étudiées. Les paramètres identifiables d'un modèle sont ensuite identifiés dans un contexte théorique idéalisé, afin de vérifier les différents outils développés. Enfin, les paramètres physiques d'un véhicule automobile sont ensuite identifiés à partir d'essais réels effectués sur piste avec un des véhicules d'essais du laboratoire.

Huit paramètres physiques d'un des véhicules du laboratoire sont estimés par minimisation d'un critère d'erreur multivariable d'erreur de sortie, formulé au sens du maximum de vraisemblance. Les résultats expérimentaux soulignent notamment que les erreurs de variance associées aux paramètres estimés sont, dans la majorité des cas, bien inférieures aux intervalles de confiance associés aux données nominales du constructeur automobile.

Finalemant, dans la conclusion, un regard critique et constructif est porté sur l'ensemble des travaux de ce mémoire. En particulier, les limites en terme de véracité des valeurs des paramètres identifiés sont soulevées. Les perspectives d'amélioration et d'application de ces résultats sont avancées et les contributions originales de ce travail sont résumées.

ELEMENTS D'IDENTIFICATION DE MODELES PARAMETRIQUES CONTINUS

1 INTRODUCTION

Les travaux présentés dans ce mémoire présentent des résultats obtenus dans le domaine de l'identification des paramètres physiques de systèmes continus. L'essentiel des travaux repose donc sur l'emploi de modèles de connaissances de complexité plus ou moins élevée suivant le degré de description du système. Les spécificités de l'identification des paramètres physiques seront présentées. Les paramètres physiques (dit réifiables [RIC91]) sont des coefficients intervenant dans les modèles à temps continu. Avant d'entreprendre leur identification, il convient de s'assurer que la démarche a, sur le plan théorique, une possibilité d'aboutir.

Dans ce cadre, aboutir à un résultat probant est conditionné par le respect de certaines propriétés des modèles utilisés. Ces propriétés - la sensibilité, l'identifiabilité et la discernabilité - font l'objet de la première partie de ce chapitre. Une fois ces propriétés satisfaites la phase d'estimation des paramètres peut débuter. Généralement, cette étape est basée sur la minimisation d'une fonctionnelle (appelée aussi fonction ou critère de coût) représentative de l'écart de comportement entre le modèle et le processus.

De très nombreuses formulations de cette fonctionnelle ont été développées à ce jour. Chaque formulation de critère de coût est adaptée à la connaissance de l'expérimentateur disponible sur le système. Ces connaissances portent typiquement sur les caractéristiques et la nature du bruit présent sur les variables, des valeurs des paramètres, de la distribution de ces valeurs... Notre choix s'est porté sur un critère offrant un bon compromis entre les connaissances disponibles sur le système et les caractéristiques de l'estimateur.

Un estimateur des paramètres est alors obtenu par minimisation du critère de coût. Cette minimisation s'opère par le biais de puissantes techniques d'optimisation, présentées dans le second chapitre.

Or les variables enregistrées sur le système peuvent être corrompues par des données aberrantes. Ces données conduisent généralement à une estimation biaisée. Différentes méthodes de robustification du critère de coût vis à vis de données aberrantes seront étudiées.

La partie suivante traite du problème bien spécifique de l'identification des paramètres physiques de système complexes à temps continu. Les paramètres physiques sont des coefficients intervenant dans les modèles à temps continu. Afin de disposer des nombreux outils théoriques propres aux systèmes linéaires, ces modèles sont fréquemment décrits sous forme d'équations différentielles linéarisées. Deux voies sont alors possibles pour l'identification des paramètres physiques de systèmes à temps continu.

- La première dite **indirecte** consiste à estimer les paramètres physiques à partir des coefficients de la fonction de transfert à temps continu liant les sorties du système à ses entrées. Différentes techniques sont disponibles pour l'identification des coefficients de cette fonction de transfert :
 - Les méthodes **indirectes** utilisant un modèle intermédiaire. Ce modèle intermédiaire peut être un modèle non paramétrique ou un modèle paramétrique à temps discret.
 - Les méthodes **directes** estimant directement le modèle continu : les filtres de variables d'états, les moments partiels réinitialisés, minimisation d'un critère d'erreur de sortie...
- La seconde approche dite **directe** pour l'identification des paramètres physiques de systèmes à temps continu exploite des algorithmes itératifs relevant des techniques de programmation non linéaire. Ces algorithmes jouent alors directement sur les paramètres physiques du modèle continu.

Les avantages et inconvénients respectifs de ces méthodes seront discutés. La méthode retenue conservera au maximum les relations intrinsèques au modèle, sans pour autant perdre de vue les aspects d'efficacité, de robustesse et de simplicité de mise en œuvre.

2 PROPRIETES DES MODELES

2.1 Linéarité des modèles

Deux types de linéarité de modèle sont généralement distingués :

- **La linéarité par rapport aux entrées.**
- **La linéarité par rapport aux paramètres.**

Une structure de modèle est dite linéaire si sa sortie vérifie le principe de superposition. Un modèle de paramètre $\underline{\theta}$ excité par une excitation U sera dit linéaire par rapport à l'entrée (noté **LE**) si sa sortie $y(t, \underline{\theta}, U)$ vérifie le principe de superposition par rapport à l'entrée :

$$y(t, \underline{\theta}, k_1 U_1 + k_2 U_2) = k_1 y(t, \underline{\theta}, U_1) + k_2 y(t, \underline{\theta}, U_2) \quad \forall k_1, k_2 \in \mathbb{R}$$

Ce modèle sera dit linéaire par rapport aux paramètres (noté **LP**) si sa sortie vérifie le principe de superposition par rapport aux paramètres :

$$y(t, k_1 \underline{\theta}_1 + k_2 \underline{\theta}_2, U) = k_1 y(t, \underline{\theta}_1, U) + k_2 y(t, \underline{\theta}_2, U) \quad \forall k_1, k_2 \in \mathbb{R}$$

Nous verrons qu'il est avantageux de disposer de modèles LE et LP. Le problème d'estimation des paramètres des modèles LP permet une formulation analytique de la solution. L'estimation directe des paramètres de modèles non-LP nécessite l'emploi de techniques d'optimisation itératives (développées dans le second chapitre).

Une description des systèmes réels par modèle de connaissance conduit dans la majorité des cas à l'écriture de modèles continus non-LP. De plus, les systèmes réels sont fréquemment non-LE et

possèdent des caractéristiques non linéaires telles que des phénomènes de saturation. Dans le cas de non-linéarités statiques, le comportement d'un modèle non-LE peut être approximé par plusieurs modèles LE valables autour des différents points de fonctionnement.

2.2 Propriétés structurelles des modèles

Les propriétés de sensibilité par rapport aux paramètres, d'identifiabilité et de discernabilité sont des propriétés souhaitables voire indispensables pour mener à terme toute procédure d'identification des paramètres physiques de systèmes continus.

L'analyse de la sensibilité caractérise l'influence des paramètres d'un modèle sur sa réponse. L'étude de l'identifiabilité garantit, au moins sur le plan théorique, que la démarche d'identification entreprise des paramètres recherchés est possible. La discernabilité est une propriété liant plusieurs modèles. En effet, différents modèles peuvent être pressentis pour décrire le comportement d'un système et identifier ses paramètres physiques. L'étude de la discernabilité permet avant même de débiter le processus de collecte de données d'affirmer, en considérant uniquement les pouvoirs prédictifs des différents modèles, s'il sera possible d'effectuer un choix parmi les différents modèles mis en concurrence.

Le paragraphe suivant présente les différents éléments constitutifs de l'analyse de la sensibilité.

2.2.1 Sensibilité

L'analyse de la sensibilité permet une quantification relative de l'influence de la variation de chaque paramètre du modèle sur sa réponse. Elle est particulièrement utile pour l'analyse globale et la synthèse de systèmes mécaniques. Elle permet par exemple d'isoler les modes mécaniques les plus affectés par la variation d'un paramètre et/ou de déterminer les paramètres prédominants dans la réponse d'un système. Les domaines d'application sont très larges, de la mise au point de système de contrôle adaptatif à la conception de systèmes complexes.

Dans le contexte de l'identification, l'analyse de la sensibilité révèle les paramètres influant peu le comportement des différents modèles de connaissances pressentis pour décrire le système. L'identification précise de ces paramètres est alors improbable. Dans le cas où des problèmes d'identifiabilité (cf. § 2.2.3) surviennent du fait de l'estimation simultanée de plusieurs paramètres, l'analyse de la sensibilité permet d'isoler les paramètres les moins influents : leurs valeurs peuvent alors être fixées à des valeurs nominales connues ou définies a priori.

Une interprétation de ces fonctions de sensibilité est possible soit dans le domaine temporel soit dans le domaine fréquentiel. Toutefois l'analyse dans le domaine temporel a comme principal inconvénient de dépendre des caractéristiques du signal d'entrée. Pour cette raison l'analyse fréquentielle est préférée.

En effet l'identification avec précision d'un paramètre requiert une précision élevée de la fonction de réponse en fréquence du modèle identifié. Plus précisément, la fonction de réponse en fréquence du modèle doit être représentative du comportement du système aux fréquences où l'influence du paramètre en question est élevée [LJU87]. Or, une grande sensibilité de la fonction de réponse en fréquence vis à vis des paramètres identifiés est nécessaire, mais non suffisante, pour assurer l'identifiabilité du modèle.

Le calcul de la sensibilité dans le domaine fréquentiel nécessite que le modèle soit formalisable sous forme de fonction de transfert. Considérons la fonction de transfert complexe :

$$G(\omega) = P(\omega) + jQ(\omega) \text{ ayant pour phase } \Phi(\omega) = \tan^{-1} \left(\frac{Q(\omega)}{P(\omega)} \right) \quad (1.2-1)$$

En supposant que $G(\underline{\theta})$ soit continue en $\underline{\theta}$, son développement autour de $\underline{\theta}_0$ conduit à :

$$G(\underline{\theta}) = G(\underline{\theta}_0) + (\underline{\theta} - \underline{\theta}_0) \frac{\partial G(\underline{\theta})}{\partial \underline{\theta}} + \frac{1}{2} (\underline{\theta} - \underline{\theta}_0)^T \frac{\partial^2 G(\underline{\theta})}{\partial \underline{\theta}^2} (\underline{\theta} - \underline{\theta}_0) + o(\underline{\theta}) \quad (1.2-2)$$

Où $o(\underline{\theta})$ est un terme tel que $o(x)/x \rightarrow 0$ lorsque $x \rightarrow 0$

Cette expression s'exprime également sous la forme

$$\Delta G = \sum_{i=1}^m W_i \Delta \theta_i + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m W_{i,j} \Delta \theta_i \Delta \theta_j + o(\underline{\theta}) \quad (1.2-3)$$

où $\Delta G = G(\underline{\theta}) - G(\underline{\theta}_0)$, W_i et $W_{i,j}$ sont les matrices de sensibilité du premier et du second ordre [NAL88] :

$$W_i = \left. \frac{\partial G}{\partial \theta_i} \right|_{\underline{\theta}=\underline{\theta}_0} \quad \text{et} \quad W_{i,j} = \left. \frac{\partial^2 G}{\partial \theta_i \partial \theta_j} \right|_{\underline{\theta}=\underline{\theta}_0} \quad i = \{1(1)m\} \quad (1.2-4)$$

Les calculs formels induits par l'analyse de la sensibilité deviennent rapidement lourds. Les calculs sont considérablement simplifiés en fixant une valeur $\underline{\theta} = \underline{\theta}_0$ pour le vecteur des paramètres et en calculant numériquement (en posant $s=j\omega$) les dérivées de l'expression (1.2-4). De plus, l'évaluation de cette expression nécessite une connaissance a priori de la valeur des paramètres recherchés. Cependant, les valeurs des paramètres influence peu l'amplitude des fonctions de sensibilité du système. Ainsi l'approximation réalisée en étudiant la sensibilité du modèle en considérant les valeurs des paramètres à priori demeure raisonnable.

Ces matrices de sensibilité sont évaluées pour une valeur nominale des paramètres. Celles-ci diffèrent suivant les valeurs de $\underline{\theta}_0$. **Les propriétés de sensibilité de la réponse des modèles ne sont donc pas à proprement parler des propriétés structurelles.**

Les matrices de sensibilités sont des fonctions complexes de ω . Ces matrices décrivent l'influence d'une variation d'un paramètre sur la réponse fréquentielle du système. Il est admis que les fonctions de sensibilité du deuxième ordre sont plus sensibles à une variation de valeur des paramètres que les fonctions de sensibilité du premier ordre. Toutefois, l'information recherchée (essentiellement qualitative) se satisfait souvent des seules magnitudes des fonctions de sensibilités du premier ordre [NAL88].

Les ordres de grandeur des différentes fonctions de sensibilité par rapport aux paramètres peuvent différer. L'introduction de coefficients de normalisation par rapport aux paramètres et à par rapport à l'amplitude de la fonction de transfert complexe définit les fonctions de sensibilité logarithmiques du premier ordre :

$$S_i = \frac{\partial \ln G}{\partial \ln \theta_i} = \frac{\partial G}{\partial \theta_i} \frac{\theta_i}{G} = W_i \frac{\theta_i}{G} \quad (1.2-5)$$

En séparant les parties réelle et imaginaire de la fonction de sensibilité du premier ordre, la fonction de sensibilité logarithmique de la magnitude de la fonction de transfert $G(\underline{\theta})$ s'exprime :

$$S_i^{|G|} = \frac{\partial \ln |G|}{\partial \ln \theta_i} = \frac{\partial |G|}{\partial \theta_i} \frac{\theta_i}{|G|} = \text{Re}(S_i) \quad (1.2-6)$$

Les résultats de l'analyse de la sensibilité peuvent être à la base d'une première approche de planification d'expérience optimale. Cependant, une augmentation de la précision des paramètres identifiés n'est pas garantie. En effet, les corrélations entre les différents paramètres sont ignorées. Une planification d'expérience tenant compte de ces couplages sera avantageusement réalisée à

l'aide des très nombreuses méthodes disponibles dans la littérature (planification d'expérience A-optimale, D-optimale, ...).

2.2.2 Exemple

Afin d'illustrer nos propos, la fonction de transfert du second ordre suivante est utilisée :

$$\frac{u(s)}{i(s)} = \frac{K \omega_n^2}{s^2 + 2 \xi \omega_n s + \omega_n^2} \quad (1.2-7)$$

La sensibilité de la sortie de cette fonction de transfert vis à vis des différents paramètres (pulsation naturelle ω_n , amortissement ξ et gain statique K) est évaluée. Dans ce but les valeurs suivantes sont considérées :

$$\xi = 0.2; \omega_n = 6.28 \text{ rad/s}; K = 1 \quad (1.2-8)$$

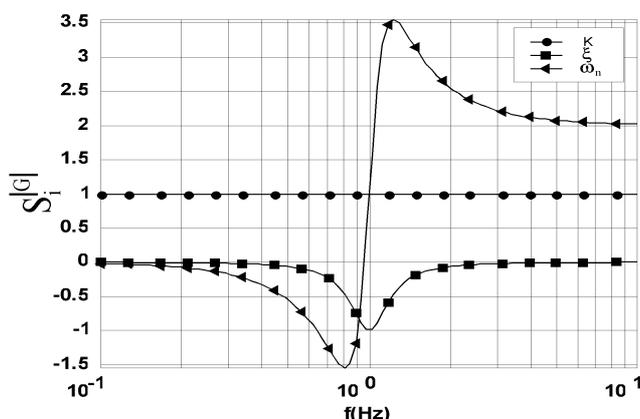


figure 1 - 1 : Fonctions de sensibilité du système du second ordre

Les tracés des magnitudes des fonctions de sensibilité (cf. figure 1 - 1) de cette fonction de transfert confirme les faits connus suivants :

- La valeur positive de la magnitude de la fonction de sensibilité vis à vis du gain statique signifie qu'une augmentation de celui-ci provoque une augmentation de la magnitude de la fonction de transfert sur tout le domaine fréquentiel.
- Les valeurs négatives de la magnitude de la fonction de sensibilité vis à vis de l'amortissement confirme la diminution de la magnitude de la fonction de transfert aux alentours de la fréquence naturelle lorsque l'amortissement croit.
- Le tracé de la sensibilité vis à vis de la pulsation naturelle montre qu'une augmentation de celle-ci provoque une diminution de la magnitude de la fonction de transfert aux fréquences amont à la fréquence naturelle. Naturellement, aux fréquences supérieures une augmentation de la fréquence naturelle entraîne une augmentation de la magnitude de la fonction de transfert.

Dans le cas présent, pour les valeurs définies, les ordres de grandeur des fonctions de sensibilité des différents paramètres sont sensiblement identiques.

2.2.3 Identifiabilité

La propriété d'identifiabilité assure, avant même de procéder au recueil des données, que la démarche d'identification est susceptible d'aboutir¹.

¹ L'étude de l'identifiabilité est indépendante du protocole expérimental, qui est considéré ici comme satisfaisant.

Pour cela, le cadre idéalisé décrit figure 1 - 2 est employé.

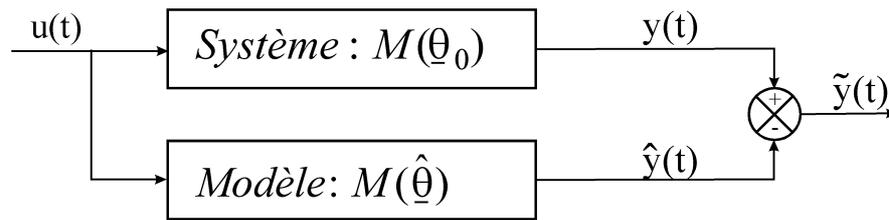


figure 1 - 2 : contexte idéalisé de l'identifiabilité structurelle

Le système est remplacé par un modèle $M(\underline{\theta}_0)$ avec :

- $M(\cdot)$ la structure du modèle du système supposée connue (absence d'erreur de caractérisation).
- $\underline{\theta}_0$ les valeurs des paramètres supposés connues également.

L'excitation est supposée être générée indépendamment de la sortie. La présence d'un bouclage de la sortie sur l'entrée entraîne généralement une perte d'identifiabilité. De plus les variables sont sans bruit. Dans ce contexte, il est toujours possible d'ajuster $\hat{\theta}$ de manière à obtenir une identité de comportement des deux modèles :

$$M(\hat{\theta}) = M(\underline{\theta}_0) \quad (1.2-9)$$

Les définitions adoptées sont celles de Walter [WAL97] dont les travaux sur l'identifiabilité font référence. Le paramètre θ_i est **structurellement globalement identifiable** (noté **sgi**) si pour presque² toute valeur de $\underline{\theta}_0$:

$$M(\hat{\theta}) = M(\underline{\theta}_0) \Rightarrow \hat{\theta}_i = \theta_{0_i} \quad (1.2-10)$$

Le paramètre θ_i est **structurellement localement identifiable** (noté **sli**) si pour presque toute valeur de $\underline{\theta}_0$ il existe un voisinage³ tel que si $\hat{\theta} \in V_x^{\underline{\theta}_0}$ et $M(\hat{\theta}) = M(\underline{\theta}_0) \Rightarrow \hat{\theta}_i = \theta_{0_i}$. L'identifiabilité structurelle locale est une condition nécessaire à l'identifiabilité structurelle globale. Un paramètre θ_i non-sli est **structurellement non identifiable** (**sni**). Il existe alors une infinité de valeurs de θ_i entraînant une identité de comportement de $M(\hat{\theta})$ et de $M(\underline{\theta}_0)$.

Un modèle est **structurellement globalement identifiable** (noté **sgi**) si tous ces paramètres le sont. Un modèle est **structurellement non identifiable** (noté **sni**) si un de ces paramètres l'est. Un modèle **sni** peut donc posséder des paramètres **structurellement localement identifiables** et **structurellement globalement identifiables**.

Différentes méthodes pour tester l'identifiabilité de modèles existent. Ces méthodes dépendent essentiellement des caractéristiques de linéarité des modèles par rapport aux paramètres et par rapport à l'entrée. Pour les modèles de connaissance LE et non-LP, les plus connues sont celles opérant transformée de Laplace, par changement de base dans l'espace d'état et par développement en série de Taylor (applicable également au modèle non-LE). La plus simple à appréhender est sans aucun doute l'approche par transformée de Laplace.

² La restriction concerne les valeurs atypiques des points de l'espace paramétrique.

³ Voir définition chapitre deuxième de ce mémoire.

2.2.3.1 Méthode par transformée de Laplace

Cette approche consiste à exprimer le modèle sous forme de fonction de transfert canonique. Soit $H(\hat{\theta})$ et $H(\underline{\theta}_0)$ respectivement les fonctions de transfert canoniques du modèle et du système. Il s'agit maintenant d'exprimer l'égalité de comportement $M(\hat{\theta}) = M(\underline{\theta}_0)$, c'est à dire $H(\hat{\theta}) = H(\underline{\theta}_0)$. Cette égalité se traduit par un ensemble d'équations en $\hat{\theta}$ et $\underline{\theta}_0$. Le nombre de solutions de ce système d'équations permet de conclure quant à l'identifiabilité du modèle considéré.

2.2.3.2 Exemple

Un circuit passe-bas du premier ordre (circuit RC série, cf. figure 1 - 3) sera utilisé dans ce mémoire pour illustrer sur un exemple académique les différentes notions de sensibilité, d'identifiabilité et de discernabilité. Ce même circuit sera mis à profit pour valider en simulation (Monte-Carlo) différents points présentés dans ce mémoire (intervalles d'incertitude sur les paramètres estimés par exemple).

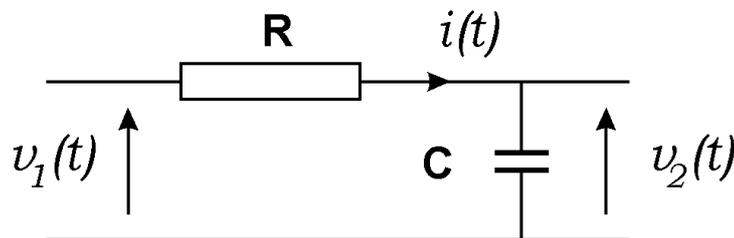


figure 1 - 3 : circuit RC passe-bas

Ce circuit illustrera nos propos sans aucune perte de généralité. De surcroît, le nombre de paramètre réduit de ce circuit permet une visualisation graphique de l'espace de recherche. Cette représentation graphique, autorise une interprétation graphique particulièrement explicite du concept d'identifiabilité. Ce système évite une dispersion de l'attention sur la complexité intrinsèque du système aux dépends de la problématique posée. Toutes les études illustrées par le biais de ce système s'étendent sans aucune perte de généralités à tout système non-LP de complexité supérieure. Le modèle décrivant le système possède plusieurs avantages :

- Il est simple (1° ordre).
- Il permet l'identification de paramètres physiques (R et C).
- Suivant les variables d'entrée/sortie choisies, il est identifiable ou non.

Considérons deux descriptions différentes de ce circuit. Dans un premier temps considérons la fonction de transfert liant la tension $v_2(t)$ à la tension $v_1(t)$, puis dans un second temps la fonction de transfert liant le courant $i(t)$ à la tension $v_1(t)$.

Description 1

$$G_{v_2/v_1}(s) = \frac{v_2(s)}{v_1(s)} = \frac{1}{1 + R C s}$$

Identité de comportement

$$M(\hat{\theta}) = M(\underline{\theta}_0) \Rightarrow \hat{R} \hat{C} = R C \quad (1.2-11)$$

Description 2

$$G_{i/v_1}(s) = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$$

Identité de comportement

$$M(\hat{\theta}) = M(\underline{\theta}_0) \Rightarrow \begin{cases} \hat{C} = C_0 \\ \hat{R} \hat{C} = R_0 C_0 \end{cases} \quad (1.2-12)$$

Conclusion

Il existe une infinité de solutions satisfaisant l'équation (1.2-11), les deux paramètres sont s_{ni}
 \Rightarrow **le modèle est structurellement non identifiable.**

Conclusion

La solution du système d'équation (1.2-12) précédent est unique. Les deux paramètres sont s_{gi}
 \Rightarrow **le modèle est structurellement globalement identifiable.**

Cet exemple montre l'importance des variables d'état choisies pour la description d'un système, sur les propriétés d'identifiabilité des modèles. Ces défauts d'identifiabilité trouvent une représentation géométrique particulièrement explicite.

2.2.3.3 Représentation géométrique

Les surfaces et contours de coût engendrés par les deux descriptions précédentes sont reproduite sur la figure 1 - 4. Ces surfaces représentent l'évolution de la valeur d'un critère quadratique d'erreur de sortie en fonction des valeurs des composants R et C (application numérique avec $\underline{\theta}_0 = (R, C) = (470 \Omega, 100 \mu F)$, sans bruit sur les mesures et une excitation de type séquences binaires pseudo-aléatoires). La formulation du critère d'erreur est la suivante :

$$J(\underline{\theta}) = \log \left(\frac{1}{N} \sum_{t=1}^N (y(t, \underline{\theta}_0) - \hat{y}(t, \underline{\theta}))^2 \right) \tag{1.2-13}$$

Le logarithmique en favorisant les faibles variations de coût rend ces surfaces plus « lisibles ».

Description 1

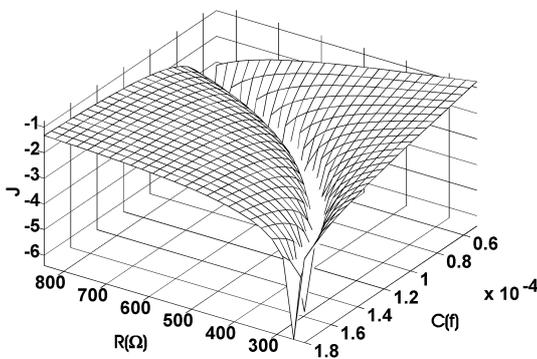


figure 1 - 4a : description 1, $J = f(R, C)$

Description 2

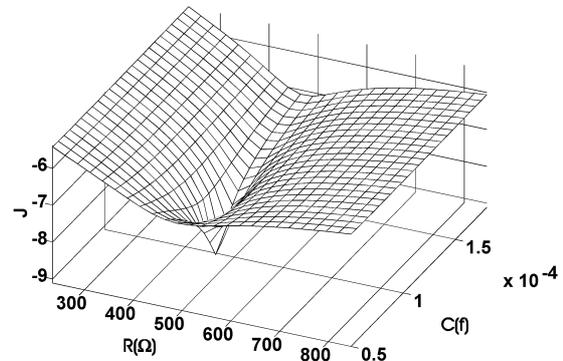


figure 1 - 4b : description 2, $J = f(R, C)$

figure 1 - 4 : surfaces de coût, circuit RC

Pour la première description, où le modèle est structurellement non-identifiable, l'infinité de solutions possibles est flagrante. Tandis qu'avec la seconde description (modèle structurellement globalement identifiable) la solution est unique. Nous verrons dans le troisième chapitre les répercussions de la forme de ces surfaces sur des intervalles d'incertitude associés aux paramètres identifiés.

2.2.3.4 Méthode numérique

Lorsque l'ordre du modèle s'élève et que les différents paramètres sont fortement couplés, les calculs formels induits par l'étude de l'identifiabilité structurelle deviennent très vite lourds⁴. Pour cette raison, l'approche numérique locale a été préférée. Cette approche garantit au moins l'identifiabilité structurelle locale d'un modèle. Une condition nécessaire à l'identifiabilité structurelle locale d'un modèle est que le déterminant de la matrice d'information de Fisher soit non-nul⁵ [WAL90].

$$\det \sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \bigg|_{\underline{\theta}=\underline{\theta}_0} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \bigg|_{\underline{\theta}=\underline{\theta}_0} \neq 0 \quad (1.2-14)$$

La méthode numérique détaillée dans l'ouvrage de Walter [WAL97] consiste à :

- Prendre dans un premier temps pour $\underline{\theta}$ une valeur nominale tirée au hasard dans l'espace de recherche et de simuler le modèle $M(\underline{\theta})$ pour obtenir des données fictives $y(t, \underline{\theta})$.
- Estimer dans un second temps $\hat{\underline{\theta}}$ à partir des données précédentes en minimisant un critère de coût quadratique par une méthode d'optimisation de type Newton initialisée à $\hat{\underline{\theta}}(0) = \underline{\theta}_0$. Si à la fin du processus d'optimisation $\hat{\underline{\theta}} = \underline{\theta}_0$ alors $M(\cdot)$ est structurellement localement identifiable dans le voisinage $V_{\underline{x}}^{\underline{\theta}_0}$. En revanche, un problème numérique ou une solution $\hat{\underline{\theta}} \neq \underline{\theta}_0$ peut être significatif d'un modèle structurellement non identifiable ou du fait que $\underline{\theta}_0$ est proche d'un point atypique (par exemple, avec notre seconde description, $C \approx 0 \Rightarrow R \approx 0$).

D'après Walter, les algorithmes de type Levenberg-Marquardt et quasi-Newton ne conviennent pas, ils possèdent en effet des procédures de régulation en cas de problème mal conditionné. Si la démarche précédente est appliquée à un système structurellement non identifiable, le processus d'identification échouera systématiquement sur un problème numérique (dû au fait que l'algorithme de Newton inverse une matrice Hessienne mal conditionnée, conséquence d'une direction désensibilisée de l'espace de recherche).

Notre expérience laisse penser que, non seulement l'utilisation des algorithmes de type Levenberg-Marquardt, quasi-Newton et méthodes d'ordre 0 (simplex par exemple) n'est pas à éviter mais est bien au contraire souhaitable. Non seulement le problème d'inversion de matrice singulière est évité et un défaut d'identifiabilité se traduira par l'inégalité $\hat{\underline{\theta}} \neq \underline{\theta}_0$. De plus, en cas de modèle structurellement localement identifiable, l'utilisation d'algorithme d'optimisation globale (voir deuxième chapitre) permet la localisation des multiples minima globaux.

Alors que l'identifiabilité est une propriété recherchée pour une structure de modèle $M(\cdot)$, la discernabilité est une propriété recherchée liant plusieurs modèles entre eux.

2.2.4 Discernabilité

Pour décrire le comportement de systèmes complexes imparfaitement connus, plusieurs structure de modèle sont souvent en compétition. Le problème est alors de déterminer le modèle adéquat à partir de relevés expérimentaux. Généralement, le modèle décrivant le plus précisément le comportement du système est retenu. Ce choix peut s'opérer suivant plusieurs critères (cf. troisième chapitre).

⁴ Voir même irréalisable par le biais de logiciel de calcul formel tel que Maple[®] et Mathematica[®].

⁵ Une planification d'expérience optimale maximisant l'expression (1.2-15), maximisant donc l'identifiabilité, transforme un problème initialement qualitatif en un problème quantitatif.

L'étude de la discernabilité permet d'affirmer, lorsque plusieurs structures de modèle différentes sont en compétition, s'il est possible d'effectuer un choix parmi celles-ci. Le pouvoir prédictif des modèles est le seul paramètre pris en compte par l'étude de la discernabilité. L'étude de la discernabilité se base sur le même contexte idéalisé que pour l'identifiabilité, à l'exception du fait que les structures des deux modèles diffèrent.

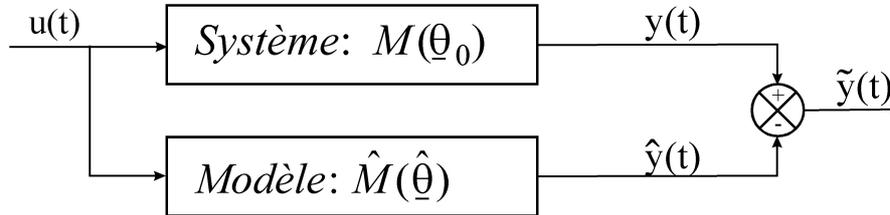


figure 1 - 5 : contexte idéalisé de la discernabilité structurelle

Le vecteur de paramètres $\hat{\theta}$ du modèle $\hat{M}(\cdot)$ et le vecteur θ_0 de $M(\cdot)$ peuvent différer et être de dimensions différentes. Etant donné que les structures du modèle et du processus diffèrent, il peut être impossible d'ajuster les paramètres $\hat{\theta}$ de manière à obtenir une identité de comportement $\hat{M}(\hat{\theta}) = M(\theta_0)$. Cette impossibilité permettra de choisir la structure $M(\cdot)$ aux dépens de la structure $\hat{M}(\cdot)$. Cette façon de procéder dépend uniquement des pouvoirs prédictifs des différents modèles en compétition. D'aucune manière les valeurs prises par les paramètres n'entrent en ligne de compte (test de consistance de la valeur des paramètres).

Le modèle $\hat{M}(\cdot)$ est dit **structurellement discernable** (noté **sd**) de $M(\cdot)$ si pour presque⁶ toute valeur de θ_0 de $M(\cdot)$ il n'existe aucune valeur $\hat{\theta}$ telle que $\hat{M}(\hat{\theta}) = M(\theta_0)$. Le fait que $\hat{M}(\cdot)$ soit structurellement discernable de $M(\cdot)$ n'implique pas que la réciproque soit vraie : cas typique où un modèle est un sur-ensemble de l'autre. Lorsque que $\hat{M}(\cdot)$ est sd de $M(\cdot)$ et que $M(\cdot)$ est structurellement discernable de $\hat{M}(\cdot)$, $M(\cdot)$ et $\hat{M}(\cdot)$ sont dits **structurellement discernables** (**structurellement indiscernables** (noté **si**) dans le cas contraire). Contrairement à l'étude de l'identifiabilité ou l'unicité d'une solution est recherchée, l'étude de la discernabilité cherche au contraire à montrer la non-existence de solutions. Les algorithmes à mettre en œuvre pour tester la discernabilité entre modèles sont parfaitement identiques à ceux utilisés pour les tests d'identifiabilité : approche par transformée de Laplace et approche numérique.

¶ Exemple

Supposons maintenant que le circuit RC précédent (seconde description : $M(\theta_0) = G_{i/u}(s) = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$) soit décrit par le modèle faux suivant :

$\hat{M}(\hat{\theta}) = \hat{G}_{i/u}(s) = \frac{i(s)}{v_1(s)} = \frac{\hat{R}\hat{C}s}{1 + \hat{C}s}$. Vérifions si le test de discernabilité permet d'éliminer $\hat{M}(\cdot)$ au

profit de $M(\cdot)$. L'identité de comportement des deux modèles implique :

$$M(\hat{\theta}) = M(\theta_0) \Rightarrow \begin{cases} \hat{R}\hat{C} = C \\ \hat{C} = R C \end{cases} \quad (1.2-15)$$

⁶ La restriction concerne les valeurs atypiques des points de l'espace paramétrique.

Quelles que soient les valeurs de θ_0 , il est toujours possible de trouver des valeurs pour $\hat{\theta}$ telles que (1.2-15) soit vérifiée. La réciproque étant également vraie, $\hat{M}(\cdot)$ et $M(\cdot)$ sont donc structurellement indiscernables. En se basant donc uniquement sur les pouvoirs prédictifs des modèles, il est impossible d'effectuer un choix entre $\hat{M}(\cdot)$ et $M(\cdot)$. Nous verrons (cf. troisième chapitre), comment lever l'ambiguïté relative au choix de structures de modèle. Les propriétés de discernabilité sont totalement découplées des valeurs prises par les paramètres. La connaissance disponible par le modéliste sur le système permet dans certains cas la résolution du problème de discernabilité entre deux modèles. Cette connaissance, sous forme de valeurs a priori admissible sur les paramètres permet le cas échéant d'éliminer une structure de modèle au profit d'une seconde.

3 CRITERES DE COUT

De nombreuses formulations de critères de coût pour l'estimation de modèle paramétrique sont répertoriées dans la littérature. Chacune de ces formulations est adaptée aux connaissances disponibles sur le système et aux caractéristiques du bruit présent sur les données [LJU87, SOD89]. Le problème est alors de choisir en toute connaissance de cause le critère le plus adéquat. Ces critères conditionnent les propriétés des estimateurs obtenus. Généralement, plus la quantité d'information a priori présente dans la formulation du critère est élevée, plus les propriétés de l'estimateur sont intéressantes.

Dans ce contexte, le champ de l'étude est volontairement restreint aux formulations les plus courantes pour l'identification des modèles paramétriques à temps continu.

Dans une première partie les différentes formulations possibles sont brièvement abordées. Les formulations des critères au sens des moindres carrés et du maximum de vraisemblance seront traitées. La simplicité de la formulation de la méthode de moindres carrés est particulièrement séduisante. Puis, nous verrons que moyennant certaines hypothèses sur la nature du bruit de mesure, l'estimateur du maximum de vraisemblance des propriétés particulièrement attractives. Enfin, une robustification des différents critères de coût permettra de se prémunir de l'influence néfaste d'éventuelles données aberrantes.

3.1 Propriétés des estimateurs

Classiquement on distingue les estimateurs :

- Des moindres carrés (MC).
- De Markov.
- Du maximum de vraisemblance (MV).
- De Bayes (risques minimum, maximum a posteriori : MAP).

Ces différents estimateurs offrent des propriétés plus ou moins intéressantes qui sont fonction des hypothèses nécessaires à leurs formulations. Ainsi, les propriétés des estimateurs (absence de biais, distribution des estimées ...) sont fonctions de la quantité et de la nature des informations présentes dans le critère d'erreur. Les estimateurs de Bayes requièrent le maximum d'information a priori ; en revanche l'estimateur des moindres carrés est utilisable si aucune connaissance a priori n'est disponible.

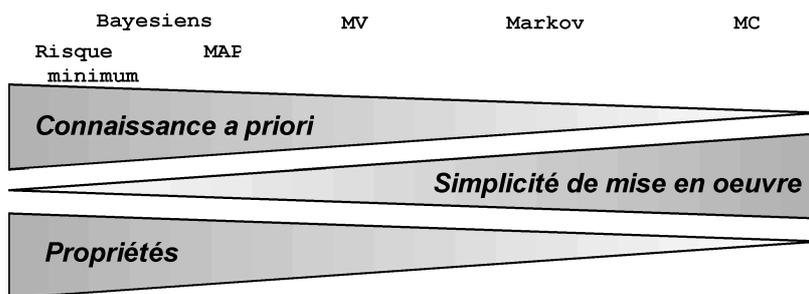


figure 1 - 6 : propriétés des estimateurs

Les connaissances disponibles sur le système se distinguent suivant leur nature statistique ou déterministe. Les connaissances de nature déterministes sont généralement prises en compte dans le modèle. Sa structure sera dictée par l'objectif de la démarche de modélisation entreprise. Les connaissances de nature statistique concernent essentiellement

- Le bruit sur les données.
- Les valeurs des paramètres.

Soient plus explicitement

- Λ : la matrice de covariance du bruit de mesure.
- $P(\varepsilon)$: la fonction de distribution du bruit.
- P_θ : la matrice de covariance a priori des paramètres.
- $P(\underline{\theta})$: la fonction de distribution des paramètres.
- $C(\hat{\underline{\theta}}, \underline{\theta}_0)$: l'évaluation supposée possible du coût de croire que les paramètres valent $\hat{\underline{\theta}}$ alors qu'en réalité ils valent $\underline{\theta}_0$.

Le tableau 1 - 1 résume succinctement les différentes connaissances a priori indispensables à la formulation des principaux critères.

Estimateur	Connaissances a priori					
	Λ	$P(\varepsilon)$	P_θ	$P(\underline{\theta})$	$C(\hat{\underline{\theta}}, \underline{\theta}_0)$	
Moindres carrés	0	0	0	0	0	
Markov	1	0	0	0	0	
Maximum vraisemblance	1	1	0	0	0	
Bayesiens	MAP	1	1	1	1	0
	Risque minimum	1	1	1	1	1

tableau 1 - 1 : connaissances a priori et estimateurs

Les **critères Bayesiens** nécessitent la connaissance des fonctions de distribution des paramètres et de la fonction de distribution du bruit de mesure. Ces estimateurs combinent donc l'information a priori disponible sur les paramètres avec l'information issue des mesures. Asymptotiquement, l'estimateur de Bayes (MAP) possède les mêmes propriétés que l'estimateur du maximum de vraisemblance ; en effet, le poids de l'information a priori sur les paramètres diminue lorsque le nombre de mesures augmente.

Les estimateurs de Bayes favorisent les valeurs a priori des paramètres. Si ceux-ci sont asymptotiquement non-biaisés ; un biais peut donc persister lorsque les données sont peu nombreuses. Ils sont donc réservés aux situations où les informations disponibles sur les paramètres sont **fiables** (valeur des paramètres a priori et intervalle de confiance associés). Leur

utilisation permet malgré tout de résoudre certains problèmes d'identifiabilité ; voir par exemple les travaux de Le Letty [LET95] pour l'identification des paramètres physiques de véhicules automobiles.

L'estimateur du **maximum de vraisemblance** est équivalent à l'estimateur de Bayes si la fonction de distribution des paramètres est uniforme (hypothèse raisonnable quand la distribution des paramètres est inconnue). L'estimateur du maximum de vraisemblance se formule pour toutes les distributions de bruit de mesure. Il possède sous certaines hypothèses de nombreuses propriétés attrayantes (cf. 3.3).

L'estimateur de **Markov** nécessite la seule connaissance de la matrice de covariance du bruit de mesure. Il peut être utilisé si le bruit ne suit pas une loi normale. Si la matrice de covariance du bruit de mesure est remplacée par toute autre matrice de pondération positive définie, l'estimateur de Markov se réduit à celui des moindres carrés pondérés. Si le bruit de mesure suit une distribution Gaussienne, de moyenne nulle, et que la matrice de covariance du bruit de mesure est connue, l'estimateur de Markov équivaut à l'estimateur du MV. Si de surcroît le modèle est LP, l'estimateur est le **meilleur estimateur linéaire non biaisé** (BLUE : Best Linear Unbiased Estimator).

L'estimateur des **moindres carrés** ne requiert quant à lui aucune connaissance relative à la distribution du bruit de mesure. Sans aucun doute, l'estimateur des moindres carrés est le plus répandu.

3.2 Critères quadratiques des moindres carrés

Introduisons le critère quadratique suivant :

$$J(\underline{\theta}) = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t, \underline{\theta}))^T (y(t) - \hat{y}(t, \underline{\theta})) \tag{1.3-1}$$

$$J(\underline{\theta}) = \frac{1}{N} \sum_{t=1}^N (\tilde{y}(t, \underline{\theta}))^T (\tilde{y}(t, \underline{\theta}))$$

où $\tilde{y}(t, \theta)$ est appelé résidu ou erreur de prédiction. Les algorithmes à mettre en œuvre pour minimiser ce critère de coût dépendent essentiellement du caractère LP du modèle utilisé. Nos travaux sont consacrés à l'identification des paramètres physiques de modèle phénoménologiques. Ces paramètres physiques seront identifiés par le biais d'une approche directe ajustant directement les paramètres recherchés (cf. 5.2.1.1). Ces modèles étant exceptionnellement LP, seul l'estimateur des moindres carrés dans le cas non-LP sera développé.

Dans ce cas, une linéarité locale de la sortie vis à vis des paramètres peut être définie autour d'une valeur nominale $\underline{\theta}_n$. Considérons le développement de $\hat{y}(t, \underline{\theta}_n)$ autour de $\underline{\theta}_n$:

$$\hat{y}(t, \underline{\theta}_n + \Delta \underline{\theta}) = \hat{y}(t, \underline{\theta}_n) + \underline{\sigma}(t)^T \Delta \underline{\theta} + o(\Delta \underline{\theta}) \tag{1.3-2}$$

Avec

$$\underline{\sigma}(t)^T = (\sigma_1(t) \ \sigma_2(t) \ \dots \ \sigma_m(t)) \tag{1.3-3}$$

$$\sigma_i(t) = \left. \frac{\partial \hat{y}(t, \underline{\theta})}{\partial \theta_i} \right|_{\underline{\theta}=\underline{\theta}_n} \quad \text{la fonction de sensibilité de la sortie vis à vis du paramètre } i$$

3.2.1 Formulation du critère

Exprimons maintenant l'erreur de prédiction en fonction des fonctions de sensibilité et les surfaces de coût résultantes. Dans un premier temps, plaçons-nous dans un cadre déterministe en l'absence

d'erreur de caractérisation. Soit $\underline{\theta}_0$ le vecteur des paramètres exacts supposé connu et $\hat{\underline{\theta}} = \underline{\theta}_0 + \Delta \underline{\theta}$ une estimation de $\underline{\theta}_0$. Pour de faibles variations $\Delta \underline{\theta}$, nous avons :

$$\hat{y}(t, \hat{\underline{\theta}}) = \hat{y}(t, \underline{\theta}_0 + \Delta \underline{\theta}) \approx \hat{y}(t, \underline{\theta}_0) + \underline{\sigma}(t)^T \Delta \underline{\theta} \quad (1.3-4)$$

D'où l'expression de l'erreur de prédiction

$$\tilde{y} \approx -\underline{\sigma}(t)^T \Delta \underline{\theta} \quad (1.3-5)$$

En injectant (1.3-5) dans (1.3-1) il vient

$$J(\underline{\theta}) = \Delta \underline{\theta}^T \mathbf{W} \Delta \underline{\theta} \quad (1.3-6)$$

$$\mathbf{W} = \frac{1}{N} \sum_{t=1}^N \underline{\sigma}(t) \underline{\sigma}(t)^T \quad (1.3-7)$$

\mathbf{W} est une matrice définie positive. L'hypersurface de coût décrite par $J(\underline{\theta})$ dans l'espace paramétrique est donc un paraboloïde seulement dans le voisinage de $\underline{\theta}_0$. Il peut exister des optima secondaires locaux.

3.2.2 Minimisation du critère des moindres carrés

Le critère précédent $J(\underline{\theta})$ n'étant pas linéaire par rapport aux paramètres, une formulation analytique de la solution comme pour le cas LP est impossible. Nous devons donc faire appel à des méthodes d'optimisation itératives relevant des techniques de **programmation non-linéaire** (cf. deuxième chapitre).

L'estimation des paramètres de modèles non-LP repose sur la simulation du modèle de sortie $\hat{y}(t, \hat{\underline{\theta}})$ en réponse à une excitation $u(t)$. L'erreur de prédiction est du type erreur de sortie.

3.2.3 Effet du bruit de mesure

Quelle que soit la structure du modèle (LP ou non-LP), dans un contexte stochastique où la sortie est bruitée par une variable aléatoire, la surface décrite par $J(\underline{\theta})$ conserve sa forme de paraboloïde. Toutefois, l'optimum global n'est plus $\underline{\theta}_0$ mais $\underline{\theta}_0 + \Delta \underline{\theta}$ ou $\Delta \underline{\theta}$ est un vecteur aléatoire nommé erreur d'estimation. L'amplitude de cette erreur d'estimation dépend des caractéristiques du bruit présent sur les données. Nous verrons au troisième chapitre, que l'amplitude de la variable aléatoire $\Delta \underline{\theta}$ est proportionnelle à la racine carrée de la variance du bruit de mesure.

Après avoir brièvement détaillé l'estimateur des moindres carrés, intéressons-nous à une autre classe d'estimateur est très largement répandue, l'estimateur au sens du maximum de vraisemblance. Celui-ci est particulièrement intéressant ; car il possède des propriétés statistiques attractives et offre sur le plan théorique toutes les qualités d'un estimateur idéal (estimateur non-biaisé, à erreur de variance minimale...).

3.3 Estimateur du maximum de vraisemblance

La méthode du maximum de vraisemblance consiste à choisir parmi les différentes valeurs des paramètres à estimer celles qui maximisent la vraisemblance des données observées (Y) sur le processus [FIS12].

$$\hat{\underline{\theta}}_{mv} = \arg \max(P_Y(Y | \underline{\theta})) \quad (1.3-8)$$

avec $P_Y(Y | \underline{\theta})$ la vraisemblance des observations (Y).

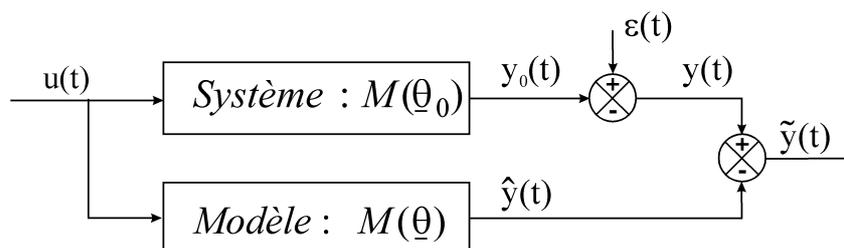


figure 1 - 7 : erreur de sortie

Supposons ensuite que les sorties observées satisfont (cf. figure 1 - 7).

$$y(t) = y_0(t, \underline{\theta}_0) + \varepsilon(t) \quad (1.3-9)$$

$\underline{\theta}_0$ est la valeur supposée connue du vecteur des paramètres.

$y_0(t, \underline{\theta}_0)$ est la sortie d'un modèle déterministe, sans erreur de caractérisation, du processus étudié.

$\varepsilon(t)$ est une séquence de variable aléatoire indépendante (vai) de moyenne nulle de densité de probabilité $P(\varepsilon)$.

Comme le bruit de mesure est une suite de vai, sa densité de probabilité s'exprime

$$P_\varepsilon(\varepsilon) = \prod_{t=1}^N P_\varepsilon(\varepsilon(t)) \quad (1.3-10)$$

Considérons alors que l'erreur de prédiction est une erreur de sortie (cf. figure 1 - 7).

$$\tilde{y}(t, \underline{\theta}) = y(t) - \hat{y}(t, \underline{\theta}) \quad (1.3-11)$$

Comme la sortie du modèle est de nature déterministe, la vraisemblance d'une observation de la sortie $y(t)$ est égale à la fonction de densité de probabilité du bruit de mesure

$$P_{y(t)}(y(t) | \underline{\theta}) = P_{\varepsilon(t)}(\tilde{y}(t, \underline{\theta})) = P_{\varepsilon(t)}(\varepsilon(t)) \quad (1.312)$$

Ainsi la fonction de vraisemblance des N observations de la sortie s'exprime donc

$$P_Y(Y | \underline{\theta}) = \prod_{t=1}^N P_{\varepsilon(t)}(\tilde{y}(t, \underline{\theta})) = \prod_{t=1}^N P_{\varepsilon(t)}(y(t) - \hat{y}(t, \underline{\theta})) \quad (1.3-13)$$

L'expression de la fonction de log-vraisemblance transforme l'expression précédente en une somme de termes, fonction de l'erreur de sortie à chaque instant

$$\ln P_Y(Y | \underline{\theta}) = \sum_{t=1}^N \ln P_\varepsilon(y(t) - \hat{y}(t, \underline{\theta})) \quad (1.3-14)$$

La forme de la distribution des résidus détermine la formulation finale du critère à optimiser. Parmi les distributions possibles, la distribution normale est la plus fréquemment utilisée. Considérons maintenant l'expression (1.3-14) en considérant que $\varepsilon(t)$ est une variable vectorielle aléatoire indépendante de distribution **normale** et de **variance inconnue**.

3.3.1 Bruit Gaussien vectoriel stationnaire de covariance inconnue

Plusieurs sorties sont observées simultanément (cas multivariable) sur le système et satisfont

$$y(t) = \hat{y}(t, \underline{\theta}_0) + \varepsilon(t) \quad t = \{1(1)N\}$$

Où les variables $y(t)$, $\hat{y}(t, \underline{\theta}_0)$ et $\varepsilon(t)$ forment une suite de variables de dimensions (n_y, N) et où $\varepsilon(t)$ est une variable aléatoire indépendante (vai) de moyenne nulle et de covariance Λ inconnue, $\varepsilon(t) \xrightarrow{\text{dist}} N(0, \Lambda)$. La quantité n_y désigne le nombre de sorties considérées dans la formulation du critère d'erreur au sens du maximum de vraisemblance. Comme $\varepsilon(t)$ suit une loi $N(0, \Lambda)$, la vraisemblance de Y s'exprime :

$$P(Y | \underline{\theta}, \Lambda) = \prod_{t=1}^N \frac{1}{((2\pi)^{n_y} \det \Lambda)^{1/2}} \exp\left(-\frac{1}{2} \tilde{y}(t, \underline{\theta})^T \Lambda^{-1} \tilde{y}(t, \underline{\theta})\right) \quad (1.3-15)$$

D'où l'expression de la fonction de la log-vraisemblance

$$\ln(P(Y | \underline{\theta}, \Lambda)) = -\frac{n_y N}{2} \ln(2\pi) - \frac{N}{2} \ln(\det(\Lambda)) - \frac{1}{2} \sum_{t=1}^N \tilde{y}(t, \underline{\theta})^T \Lambda^{-1} \tilde{y}(t, \underline{\theta}) \quad (1.3-16)$$

Si la matrice de covariance du bruit de mesure était connue, maximiser la fonction de vraisemblance reviendrait à minimiser le critère de coût suivant

$$J_{mv}(\underline{\theta}) = \sum_{t=1}^N \tilde{y}(t, \underline{\theta})^T \Lambda^{-1} \tilde{y}(t, \underline{\theta}) \quad (1.3-17)$$

Si la matrice Λ est connue l'estimateur au sens du maximum de vraisemblance revient à la formulation des moindres carrés pondérés. Comme le vecteur des paramètres $\underline{\theta}$ et la matrice de covariance du bruit de mesure Λ sont inconnues, l'expression (1.3-16) doit être maximisée par rapport à chacun de ses deux derniers termes. Une minimisation séquentielle en deux temps est possible ; dans un premier temps, l'expression est minimisée par rapport à $\underline{\theta}$ et dans un second temps par rapport à Λ . Cette méthodologie est implémentée dans la « mml3 toolbox » [MIL88]. Goodwin [GOO77] a montré qu'une estimation $\hat{\underline{\theta}}_{mv}$ pouvait être obtenue par minimisation de l'expression

$$J_{mv}(\underline{\theta}) = \det \sum_{t=1}^N \tilde{y}(t, \underline{\theta})^T \tilde{y}(t, \underline{\theta}) \quad (1.3-18)$$

Les paramètres sont estimés par minimisation de l'expression (1.3-18), soit encore

$$\hat{\underline{\theta}}_{mv} = \arg \min(J_{mv}(\underline{\theta})) \quad (1.3-19)$$

Une estimation de la matrice de covariance Λ du bruit de mesure s'obtient alors par l'expression

$$\hat{\Lambda}_{mv}(\hat{\underline{\theta}}) = \frac{1}{N + np} \sum_{t=1}^N (y(t) - \hat{y}(t, \hat{\underline{\theta}}_{mv})) (y(t) - \hat{y}(t, \hat{\underline{\theta}}_{mv}))^T \quad (1.3-20)$$

où np désigne le nombre de paramètres identifiés $N + np \approx N$ si $N \gg np$.

Ainsi une estimation des paramètres au sens du maximum de vraisemblance est possible par minimisation du coût (1.3-18) sans aucune connaissance de la valeur de la matrice de covariance du bruit de mesure. L'expression (1.3-18) a également l'avantage de résoudre le problème de la

pondération des différentes variables dans la formulation de critère de coût multivariable sans introduire de coefficient de pondération plus ou moins arbitraire.

Notons que dans l'expression (1.3-17) le poids relatif des différentes variables entrant dans la formulation du critère de coût est fixé par l'inverse de la matrice de covariance du bruit de mesure. L'estimateur au sens du maximum de vraisemblance pondère donc ses différentes variables par l'inverse de la variance associée à celles-ci. Plus une variable est bruitée, plus sa pondération dans le critère de coût est faible.

3.3.2 Propriétés des estimateurs au sens du maximum de vraisemblance

Dans le contexte de l'identification et sous réserve que les hypothèses suivantes soient satisfaites :

- ☞ **Absence d'erreur de caractérisation.**
- ☞ **Le modèle est identifiable.**
- ☞ **Le bruit sur les données est assimilable à une suite de vai.**

Les estimateurs du maximum de vraisemblance possèdent les propriétés suivantes. Ils sont :

- ☞ **asymptotiquement convergents**
 $\forall \delta > 0, P(|\hat{\underline{\theta}} - \underline{\theta}_0| > \delta) \rightarrow 0 \text{ quand } N \rightarrow \infty$ (1.3-21)
- ☞ **asymptotiquement efficaces**

Il n'existe pas d'estimateur convergent offrant une covariance $P_{\hat{\underline{\theta}}}$ plus faible sur les paramètres estimés :

$$P_{\hat{\underline{\theta}}} = \text{var}(\hat{\underline{\theta}}) = E\{(\hat{\underline{\theta}} - \underline{\theta}_0)(\hat{\underline{\theta}} - \underline{\theta}_0)^T\} \quad (1.3-22)$$

- ☞ **asymptotiquement normaux et non-biaisés**
 $\hat{\underline{\theta}} \xrightarrow{\text{dist}} N(0, P_{\hat{\underline{\theta}}}) \text{ lorsque } N \rightarrow \infty$ (1.3-23)

Si la fonction de vraisemblance est exprimée pour une distribution normale du bruit de mesure, la théorie nous affirme que l'estimateur est à minimum de variance si le bruit de mesure suit une distribution normale. Néanmoins l'estimateur demeure applicable pour une fonction de distribution du bruit différente. Sibille [SIB90] a montré dans ses travaux (maximum de vraisemblance formulé en considérant une erreur d'équation) que les résultats demeuraient satisfaisants lorsque la distribution du bruit était uniforme. Cependant, comme il le fait remarquer, sur le plan théorique rien ne garantit alors que l'estimateur est à minimum de variance.

Généralement, les variables recueillies sont perturbées par la présence de données aberrantes. Le paragraphe suivant présente quelques techniques visant à rendre les différents estimateurs robustes vis à vis de violations des hypothèses relatives à la distribution des résidus.

4 ROBUSTIFICATION

Précédemment il a été montré que les estimateurs des moindres carrés et du maximum de vraisemblance supposent que les données recueillies sont entachées d'un bruit de mesures de moyenne nulle. Lors du recueil des ces données expérimentales, certains échantillons peuvent être corrompus par une erreur de mesure particulièrement importante. Les causes possibles de ces erreurs sont multiples : perturbation, présence d'erreur de conversion ...

En l'absence de prise de mesures spécifiques, un biais d'estimation sur les paramètres demeure possible. Afin de minimiser l'influence de ces données marginales sur les paramètres estimés,

diverses techniques inspirées de la statistique ont été développées. Parmi celles-ci les M-estimateurs sont les plus répandus ([SOD89], [LJU87], [GNA97]).

Le principe des M-estimateurs est de minimiser l'influence des données aberrantes en remplaçant la pondération quadratique des résidus aberrants dans les critères par une autre fonctionnelle alternative des résidus :

$$J_{\text{rob}}(\underline{\theta}) = \sum_{t=1}^N \rho(y(t) - \hat{y}(t, \underline{\theta})) \quad (1.4-1)$$

où $\rho(\cdot)$ est une fonction symétrique positive définie présentant un unique minimum en 0.

La fonction $\rho(\cdot)$ est choisie de croissance inférieure à la fonction quadratique. Généralement, une fonction d'influence à partir de la fonction $\rho(\cdot)$ se définit telle que⁷ :

$$\Psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+ \quad (1.4-2)$$

$$\Psi(x) = \frac{\partial \rho}{\partial x}$$

La fonction d'influence $\Psi(\cdot)$ quantifie l'influence d'une donnée sur le critère d'erreur. Pour le critère des moindres carrés, $\rho(x) = \frac{x^2}{2}$ et la fonction d'influence $\Psi(x) = x$ signifie que l'influence d'une donnée sur le critère croît linéairement avec la valeur de x ; ceci confirme la non-robustesse de l'estimateur des moindres carrés vis à vis de données aberrantes. Pour minimiser l'influence de données aberrantes de nombreuses fonctions $\rho(\cdot)$ ont été développées pour l'estimation robuste. Citons pour mémoire les fonctions de Cauchy, German-Mc Lure, Welsh, Tukey, Huber... Parmi ces fonctions, trois sont retenues pour leur efficacité et leur représentativité des autres fonctions existantes.

¶ Fonction de Huber

Les résultats obtenus avec cette fonction sont satisfaisants dans de nombreux cas de figure, et son utilisation est préconisée dans la majorité des cas [ZEH95]. La fonction de Huber s'exprime

$$\rho(x) = \begin{cases} \frac{x^2}{2} & \text{si } |x| \leq \delta \\ \delta x - \frac{1}{2} \delta^2 & \text{sinon} \end{cases} \quad (1.4-3)$$

Et sa fonction d'influence

$$\Psi(x) = \begin{cases} x & \text{si } |x| \leq \delta \\ \delta \operatorname{sgn}(x) & \text{sinon} \end{cases} \quad (1.4-4)$$

Où $\delta = \lambda \hat{\sigma}$ avec $1 \leq \lambda \leq 1.8$ et $\hat{\sigma}$ est un estimateur robuste de l'écart type des erreurs vis à vis de données aberrantes [SOD89], [LJU87] :

$$\hat{\sigma} = \frac{\operatorname{med}(|x - \operatorname{med}(x)|)}{0.7} \quad (1.4-5)$$

où $\operatorname{med}(x)$ désigne la valeur médiane du vecteur x .

⁷ Afin d'alléger la notation et de conserver une formulation générale, l'erreur de prédiction $\tilde{y}(t, \underline{\theta}) = y(t) - \hat{y}(t, \underline{\theta})$ est remplacée par la variable x pour la description des différentes fonctions de robustification.

L'estimateur de Huber se comporte comme un estimateur des moindres carrés pour les faibles valeurs de x et comme un estimateur des moindres modules pour les valeurs de x plus importantes. L'utilisation de la fonction de Huber peut poser problème aux méthodes d'optimisation du second ordre. Les problèmes numériques induits sont la conséquence de la discontinuité de la dérivée seconde de la fonction $\rho(\cdot)$.

$$\frac{\partial^2 \rho}{\partial x^2} = \begin{cases} 1 & \text{si } |x| \leq \delta \\ 0 & \text{sinon} \end{cases} \quad (1.4-6)$$

Note : la fonction d'influence de la fonction de Huber ne s'annule jamais. Ainsi l'influence de données particulièrement aberrantes ne s'annule jamais (cf. figure 1 - 8, pour $\delta = 1$). D'autres fonctions $\rho(\cdot)$ limitent l'influence de ce type de données. La fonction de Tukey est une de ces fonctions.

¶ **Fonction de Tukey**

Avec une formulation de la quantité δ identique à la précédente, la fonction $\rho(\cdot)$ de Tukey s'exprime

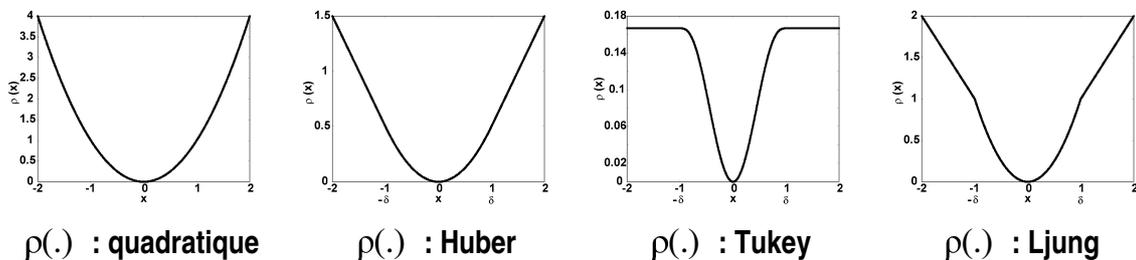
$$\rho(x) = \begin{cases} \frac{1}{2} \left(x^2 - \frac{x^4}{\delta^2} + \frac{x^6}{3\delta^4} \right) & \text{si } |x| \leq \delta \\ \frac{\delta^2}{6} & \text{sinon} \end{cases} \quad (1.4-7)$$

Et sa fonction d'influence

$$\psi(x) = \begin{cases} \frac{1}{2} \left(2x - \frac{4x^3}{\delta^2} + \frac{6x^5}{3\delta^4} \right) & \text{si } |x| \leq \delta \\ 0 & \text{sinon} \end{cases} \quad (1.4-8)$$

La forme de la fonction d'influence $\Psi(\cdot)$ définit le type de M-estimateur. La fonction de Huber définit un estimateur robuste non-redescendant (cf. figure 1 - 8) ; dans le cas de la fonction de Tukey il est fait référence à un estimateur redescendant. Cependant ce dernier type d'estimateur peut entraîner l'apparition de minima locaux secondaires.

Ljung [LJU91], propose une robustification différente et originale et qui est implémentée dans le code de la fonction « pem » de la « toolbox identification » de Matlab[®]. Reprenons rapidement cette dernière formulation afin de la comparer aux deux précédentes.



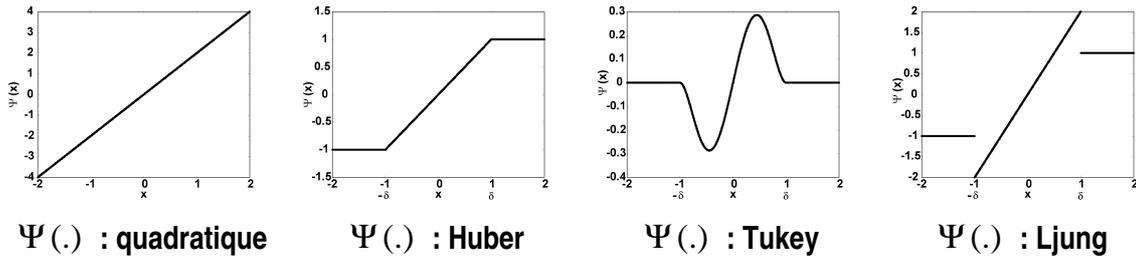


figure 1 - 8 : fonctions de robustification

¶ **Fonction de « Ljung »**

Avec une formulation de δ identique à la précédente, l'étude du code de « pem » permet de déduire la fonction $\rho(.)$ de Ljung :

$$\rho(x) = \begin{cases} x^2 & \text{si } |x| \leq \delta \\ \frac{\delta x^2}{|x|} & \text{sinon} \end{cases} \quad (1.4-9)$$

Et sa fonction d'influence

$$\psi(x) = \begin{cases} 2x & \text{si } |x| \leq \delta \\ 2\delta \operatorname{sgn}(x) & \text{sinon} \end{cases} \quad (1.4-10)$$

Aucune trace de cette fonction de robustification n'a été trouvée dans la littérature. La discontinuité présente sur la fonction d'influence est surprenante et est une source problèmes pour les méthodes d'optimisation du premier et second ordre. La figure 1 - 8 récapitule les principales formes de fonction de robustification reprises dans ce mémoire.

¶ **Exemple illustratif**

L'étude comparative expérimentale décrite souligne les avantages et inconvénients liés à l'utilisation de ces différents schémas de robustification. Pour cela, reprenons le circuit RC série et la fonction de transfert tension/courant , $G_{i/u}(s) = \frac{i(s)}{v_1(s)} = \frac{Cs}{1+R Cs}$. Identifions ensuite sur des données simulées le vecteur de paramètres $\underline{\theta} = (R, C)$, connaissant $\underline{\theta}_0 = (R_0, C_0) = (470 \Omega, 100 \mu F)$. Le critère d'erreur minimisé est du type maximum de vraisemblance

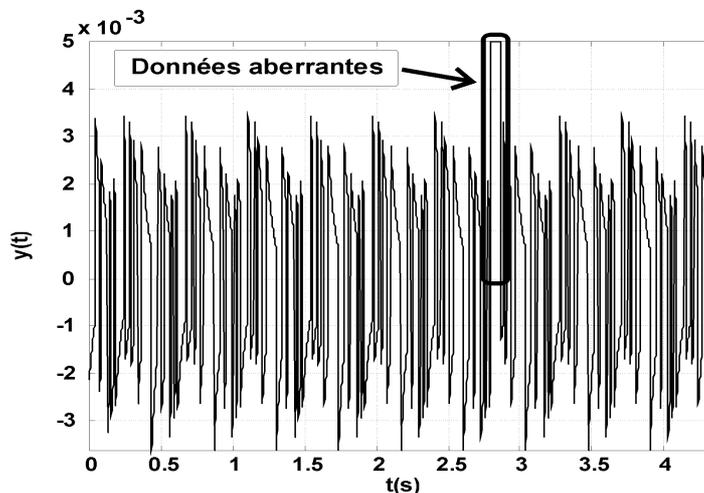


figure 1 - 9 : données corrompues

Le système est excité par un signal de type SBPA : période d'échantillonnage = 2ms, temps élémentaire (durée d'un bit) = 14 ms, nombre de bit = 5, amplitude = $\pm 1V$, nombre de périodes = 10. Des données aberrantes sont introduites sur les données simulées (sans bruit), telles que :

$$y(t_{\text{aberrantes}}) = 5 \cdot 10^{-3}, \quad t_{\text{aberrantes}} = 2.998 (0.002) 3.070 \quad (1.4-11)$$

Ces données aberrantes peuvent être représentatives de la présence d'un faux contact ou de parasites durant 7/100 de seconde, cf. figure 1 - 9.

$\rho(\cdot)$	--	Quadratique		Huber		Tukey		Ljung	
$\underline{\theta}$	$\underline{\theta}_0$	$\hat{\underline{\theta}}_{mv}$	$\Delta \underline{\theta}$						
R(Ω)	470	477.71	-7.71	470.000	-8.52 10^{-6}	470.000	1.83 10^{-6}	470.000	-8.52 10^{-6}
C (μ F)	100	82.89	17.11	99.999	1.22 10^{-5}	100.000	-5.31 10^{-6}	99.999	1.22 10^{-5}

tableau 1 - 2 : comparaison des différents schémas de robustification

Dans le tableau 1 - 2 sont reportés les résultats obtenus par minimisation d'un critère de coût formulé au sens du maximum de vraisemblance. Les deux premières colonnes du tableau comportent respectivement les valeurs nominales des paramètres $\underline{\theta}_0$ et les résultats sans robustification du critère $\hat{\underline{\theta}}_{mv}$. Puis les résultats obtenus selon par robustification selon Huber, Tukey et Ljung ($\Delta \underline{\theta} = \underline{\theta}_0 - \hat{\underline{\theta}}_{mv}$) sont présentés dans les colonnes suivantes.

Les différentes formes de robustification du critère de coût offre des résultats très proches des valeurs nominales. Les résultats obtenus à l'aide de fonction de robustification de type Huber et Ljung sont quasiment identiques. Néanmoins, les discontinuités de la dérivée première de ce dernier peuvent poser des problèmes numériques aux méthodes d'optimisation du premier et second ordre évaluant explicitement le gradient et Hessien de la fonction de coût. Comme mentionné auparavant, la robustification par fonction de Tukey est la plus efficace. Le but de ce mémoire n'étant pas de réaliser une étude exhaustive des différents schémas de robustification existant, le lecteur désireux d'approfondir ce point consultera avantageusement les ouvrages suivants [LJU87], [SOD89], [GNA97], [ZHE95].

Voyons maintenant les problèmes d'optimisation posés par ces différentes formulations du critère d'erreur. Traçons les surfaces et contours de coût (figure 1 - 10) induites ces critères de coût robustifiés en présence de bruit de mesure. Le bruit de mesure simulé ajouté aux données suit une loi normale telle que $\varepsilon(t) \xrightarrow{\text{dist}} N(0, \lambda)$, $\lambda = \sigma^2 = (1.7 \cdot 10^{-3})^2 = 2.9 \cdot 10^{-6}$. Une robustification par fonction de Tukey entraîne la création de nombreux minima locaux. Il faudra donc veiller à ce que l'algorithme d'optimisation utilisé soit capable de localiser l'optimum global parmi tous les optima secondaires. La présence de ces optima locaux, motive les travaux consacrés à la mise en œuvre de puissantes techniques d'optimisation globale, dans le chapitre suivant de ce mémoire.

Les résultats obtenus en simulation, montrent l'efficacité de ces différentes fonctions de robustification. Cependant, il faut noter que ces différents schémas de robustification offrent de remarquables résultats en présences de données aberrantes d'amplitude sensiblement différentes de la moyenne des données (exemple typique de faux contacts ou de dysfonctionnement de capteurs à $\pm V_{cc}$). Pour minimiser l'influence de données aberrantes d'amplitudes plus restreintes, le modéliste s'orientera plutôt :

- Soit vers des estimateurs minimisant des critères de coût plus robuste, tels que les critères des moindres module ou de la médiane des carrés des résidus par exemple.
- Soit vers un filtrage « intelligent » préalable des données [BAS99] par le biais de filtres de Kalman par exemple ; la mise en œuvre de ce type de filtrage n'est pas triviale et requiert une bonne connaissance des caractéristiques des bruits présent sur les données.

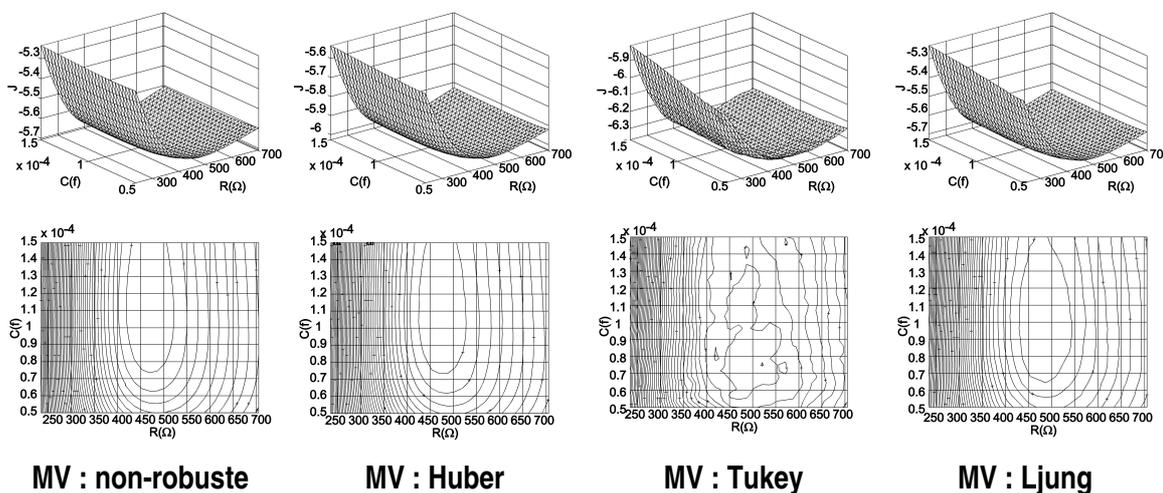


figure 1 - 10: surfaces de coût, critère du maximum de vraisemblance robustifié

Dans ce paragraphe nous avons montré qu'une robustification efficace des solutions vis à vis de données aberrantes était possible. Cette robustification est atteinte par l'introduction d'une pondération relative des différentes mesures dans la formulation du critère de coût.

Le paragraphe suivant traite de l'identification des paramètres de systèmes à temps continu. Les différentes formulations de critères sont applicables aux différentes méthodes d'identification des paramètres des systèmes à temps continu.

5 IDENTIFICATION DES SYSTEMES A TEMPS CONTINU

Les processus physiques sont généralement à temps continu. Il semble donc naturel de vouloir les décrire par des modèles à temps continu. Typiquement, les paramètres physiques interviennent dans la description de ce type de modèle. Après avoir brièvement décrit quelques éléments nécessaires à la simulation de systèmes à temps continu, nous aborderons différentes techniques possibles pour l'identification des paramètres physiques de modèle à temps continu. Parmi ces techniques, deux voies se distinguent :

- L'approche indirecte, basée sur les résultats de l'identification des paramètres d'une équation différentielle régissant le système étudié.
- L'approche directe conduisant à l'obtention directe des paramètres recherchés.

Ces deux approches sont applicables pour une représentation temporelle ou fréquentielle des signaux relevés sur le système.

5.1 Simulation de système à temps continu

Une démarche d'identification passe généralement par une phase de recherche d'identité de comportement entrée/sortie du processus avec le modèle. Suivant les cas, pour l'identification de systèmes à temps continu, une étape de simulation est-elle indispensable, permettant ainsi de juger

de l'adéquation avec le modèle au système réel. Le simulateur doit permettre de simuler la réponse d'un modèle aux instants souhaités. Plaçons-nous dans le cadre d'une période d'échantillonnage contante dans le temps. La simulation de la réponse du système continu aux instants d'échantillonnage est alors recherchée : $\{y(k), k = 1(1)N\}$ en réponse à l'entrée $\{u(k), k = 1(1)N\}$. Du caractère LE ou non-LE des modèles continus à simuler dépendent les techniques à mettre en œuvre pour leurs simulations.

5.1.1 Cas général (modèle LE ou non)

Pour les modèles non-LE, il est possible de fournir une réponse approximée du système via diverses méthodes d'intégration itératives. Citons par exemple les méthodes de Runge-Kutta, d'Adams-Bashforth, de Gear...[PRE86]. Ces algorithmes opèrent en un processus itératif d'intégration pour chaque point calculé et entraînent rapidement des charges de calcul élevées. Pour des modèles non-LE seules les méthodes d'intégration sont applicables. Ainsi, lorsque cela est possible, il est avantageux de linéariser le modèle et d'adopter un mode de représentation sous forme de matrice d'état ou de fonction de transfert.

5.1.2 Modèle LE

Considérons le modèle à temps continu linéarisé (LE) suivant :

$$\begin{aligned} \dot{x}(t) &= A x(t) + B u(t) \\ y(t) &= C x(t) + D u(t) \end{aligned} \quad (1.5-1)$$

Supposons maintenant que le système est dans l'état $x(k T_e)$ à l'instant $k T_e$ et que l'on souhaite connaître sa réponse pour tout temps $t \geq k T_e$. La solution exacte à (1.5-1) à un instant t quelconque est donné par :

$$x(t) = e^{A(t-kT_e)} x(k T_e) + \int_{kT_e}^t e^{A(t-\tau)} B u(\tau) d\tau \quad (1.5-2)$$

Sous l'hypothèse que l'excitation demeure constante entre deux instants d'échantillonnage :

$$u(t) = u(k T_e) \quad (k T_e < t < (k+1) T_e) \quad (1.5-3)$$

L'expression (1.5-2) s'exprime alors en posant $\Phi(t) = e^{A t}$

$$x((k+1)T_e) = \Phi(t - k T_e) x(k T_e) + \int_{k T_e}^{(k+1)T_e} \Phi(t - \tau) B d\tau u(k T_e) \quad (1.5-4)$$

Soit encore dans le cas d'une période d'échantillonnage T_e constante

$$x((k+1)T_e) = \Phi(T_e) x(k T_e) + \int_0^{T_e} \Phi(T_e - \tau) B d\tau u(k T_e) \quad (1.5-5)$$

Sous forme abrégée

$$x((k+1)T_e) = \Phi(T_e) x(k T_e) + \Gamma(T_e) u(k T_e) \quad (1.5-6)$$

Le mode de calcul de la matrice exponentielle par développement en série de Taylor transforme le système différentiel continu en un système d'équations aux différences. Cette formulation permet de simuler la réponse du système continu avec une bonne précision et un temps de calcul raisonnable. La méthode de l'exponentielle de la matrice calcule la matrice $\Phi(T_e)$ par approximations successives, soit :

$$\Phi(t) = e^{A T_e} \approx \sum_{i=0}^{\infty} \frac{(A T_e)^i}{i!} \quad (1.5-7)$$

$$\Gamma = \int_0^{T_e} e^{A(T_e-\tau)} B d\tau = A^{-1}(e^{A T_e} - I) B = A^{-1}(\Phi(T_e) - I) B \quad (1.5-8)$$

Note : lorsque l'excitation varie continûment la méthode de la matrice exponentielle demeure valide à condition :

- Soit d'appliquer un bloqueur d'ordre supérieur à 0 sur l'excitation $u(t)$.
- Soit de décrire analytiquement les variations de $u(t)$ entre les instants d'échantillonnage.

De nombreux logiciels de conception de loi de commande permettent la simulation de systèmes continus (Matlab[®], MatrixX[®], Scilab[®], Octave[®]...), en offrant une large gamme de méthodes d'intégration numérique. L'écriture de ces routines de simulation présente donc un intérêt limité. Cependant, une étude quantitative comparative de l'influence des différents algorithmes de simulation a été conduite sur des signaux réels et les différents modèles présentés dans le dernier chapitre de ce mémoire. Ces algorithmes sont :

- Méthode de la matrice exponentielle. Le système est discrétisé avec un bloqueur d'ordre 0
- Runge-Kutta(4,5)
- Runge-Kutta(2,3)
- d'Adams-Bashforth

Les solutions obtenues $\hat{\theta}_{mv}$ (paramètres physiques du véhicule automobile), estimées au sens du maximum de vraisemblance, sont très robustes vis à vis de l'algorithme de simulation employé. Les différences observées (outre le temps de calcul) d'un algorithme à l'autre sont **au maximum** de 10^{-4} . Cette quantité exprimée en pourcentage des paramètres recherchés peut raisonnablement être considérée comme négligeable. En effet, lors de l'identification des paramètres physique d'un véhicule automobile, l'amplitude du paramètre identifié la plus faible est de l'ordre de 0.5. Une différence d'estimation de 10^{-4} représente alors dans le cas le plus défavorable une différence de **0.002 %**.

Une fois le problème de la simulation de systèmes continus traité, nous pouvons aborder la problématique de l'identification des modèles à temps continu. Le paragraphe suivant est consacré à l'identification des paramètres d'une équation différentielle ou d'une fonction de transfert. Le problème est alors d'estimer les paramètres d'un modèle à temps **continu** à partir de données **discrétisées** recueillies sur le processus.

5.2 Estimation des paramètres d'une équation différentielle linéaire

Considérons un modèle paramétrique à temps continu, exprimé sous forme de fonction de transfert

$$H(s) = \frac{Y(s)}{U(s)} = \frac{a_0 + a_1 s + \dots + a_{na} s^{na}}{1 + b_1 s + \dots + b_{nb} s^{nb}} \quad nb \geq na \quad (1.5-9)$$

Deux approches sont possibles pour l'identification des paramètres de systèmes à temps continu à partir de données discrétisées [SIN82]. La première dite **indirecte** passe par l'exploitation de modèles non-paramétriques ou par l'utilisation de modèles paramétriques discrets. La seconde appelée **directe** conduit directement à l'obtention des paramètres intervenant dans l'expression (1.5-9).

5.2.1 Méthodes indirectes

Les méthodes indirectes opèrent généralement en deux temps. La première étape consiste à élaborer un modèle intermédiaire. Ce modèle est ensuite converti dans une représentation à temps

continu. Les caractéristiques de ce modèle intermédiaire différencient les différentes méthodes indirectes existantes. Essentiellement deux types de modèles intermédiaires sont employés : les modèles non-paramétriques et les modèles paramétriques discrets.

5.2.1.1 Identification de modèle paramétrique à temps continu à partir de modèle non-paramétrique

Le processus d'identification opère en deux temps. Dans un premier temps le comportement du système est décrit sous forme de réponse impulsionnelle, réponse indicielle ou sous forme de fonction de réponse en fréquence (analyse spectrale). Puis dans un second temps, le modèle continu est calculé par ajustement au modèle non-paramétrique. Les techniques à mettre en œuvre pour réaliser cet ajustement dépendent du domaine de représentation.

Dans le domaine temporel, citons les méthodes de l'analyse indicielle, de Strejc, de Hudzovic. De nombreux travaux ont été consacrés dans le domaine fréquentiel dans le passé et plus récemment ([PIN94], [STO93], [KOL95]...) au problème d'estimation non-linéaire et au problème de synthèse de la fonction de transfert à partir de la fonction de réponse en fréquence.

5.2.1.2 Identification de modèle à temps continu à partir de modèle paramétrique discret

Dans cette approche, le problème est subdivisé en deux sous problèmes :

- L'estimation des paramètres d'un **modèle paramétrique à temps discret** à partir des données recueillies.
- Le **calcul du modèle continu** correspondant.

Les nombreux outils théoriques et informatiques développés ces 20 dernières années, dans le domaine de l'identification de modèles linéaires discrets, permettent une résolution théorique du premier sous problème. Pour un exposé détaillé des différents outils disponibles consulter les références de Ljung [LJU87] et de Södeström [SOD89].

La seconde étape du problème consistant à obtenir un modèle continu à partir du modèle discret précédent n'est pas triviale. Sur le plan théorique il convient de distinguer les cas où l'entrée $u(t)$ varie entre les instants d'échantillonnage. Le passage d'un modèle à temps discret à un modèle à temps continu est possible par le biais d'opérateurs de transformation (transformation bilinéaire ou de méthode Tustin par exemple). Du fait de l'existence de différents opérateurs de transformations possibles, différents modèles continus peuvent être obtenus à partir du même modèle discret. Le passage d'un modèle discret à un modèle continu est en général un problème délicat et hasardeux [COI97].

5.2.2 Méthodes directes

Les méthodes directes agissent directement sur les coefficients du modèle continu. Formulons l'expression (1.5-9) sous forme d'équation différentielle où les coefficients a_i et b_i sont constants :

$$y(t) + b_1 y'(t) + \dots + b_{nb} y^{nb}(t) = a_0 u(t) + a_1 u'(t) + \dots + a_{nb} u^{nb}(t) \quad (1.5-10)$$

L'équation (1.5-10) se formule sous forme de régression linéaire

$$y(t) = -b_1 y'(t) - \dots - b_{nb} y^{nb}(t) + a_0 u(t) + a_1 u'(t) + \dots + a_{nb} u^{nb}(t) \quad (1.5-11)$$

avec

$$y(t, \underline{\theta}) = \underline{\varphi}(t)^T \underline{p} \quad (1.5-12)$$

$$\underline{\varphi}(t) = (-y'(t) \dots - y^{nb}(t) \quad u(t) \dots u^{na}(t)) \quad (1.5-13)$$

$$\underline{p} = (b_1 \dots b_{na} \ a_0 \dots a_{nb})^T \quad (1.5-14)$$

Le modèle obtenu est LP mais impossible à identifier directement car les dérivées successives de $y(t)$ et de $u(t)$ du régresseur $\underline{\varphi}(t)$ sont non-mesurables. L'idée est alors de conditionner les signaux d'entrée et de sortie avant de leur appliquer un estimateur. Parmi les conditionnements possibles le filtrage linéaire est le plus simple, on aboutit au filtre de variable d'états. Une autre possibilité réside dans l'utilisation d'algorithme d'optimisation non-linéaire. Cette méthode agit directement sur les coefficients de la fonction de transfert (1.5-9). L'algorithme d'optimisation non-linéaire ajuste les coefficients de la fonction de transfert de manière à minimiser l'erreur de sortie entre le modèle (1.5-9) et les mesures. Cette méthode requiert donc à chaque itération la simulation du modèle en réponse à l'entrée $u(t)$.

Les résultats issus des méthodes directes et indirectes pour l'estimation des paramètres d'une équation différentielle peuvent être mis à profit pour l'estimation des paramètres physiques des systèmes continus.

5.3 Estimation des paramètres physiques de systèmes continus

Comme pour l'identification des paramètres de systèmes à temps continu, deux approches sont envisageables. Les méthodes indirectes et directes.

5.3.1 Méthodes indirectes

Les premières consistent en l'exploitation des résultats de l'identification des paramètres de systèmes à temps continu, puis en l'estimation les paramètres physiques ($\underline{\theta} = (\theta_1 \dots \theta_m)^T$) à partir des coefficients du modèle continu identifié ($\underline{p} = (b_1 \dots b_{na} \ a_0 \dots a_{nb})^T$). Les paramètres continus \underline{p} dépendent des paramètres physiques $\underline{\theta}$:

$$\underline{p} = f(\underline{\theta}) \quad (1.5-19)$$

La méthode consiste donc dans un premier temps à estimer les paramètres \underline{p} du système continu puis d'identifier les paramètres physiques $\underline{\theta}$ par inversion de la relation (1.5-19) :

$$\underline{\theta} = f^{-1}(\underline{p}) \quad (1.5-20)$$

Généralement les paramètres physiques $\underline{\theta}$ dépendent des paramètres continus \underline{p} par le biais de relations algébriques non-linéaires. Sur des exemples académiques, l'inversion (1.5-20) est souvent extrêmement simple ; circuit RC par exemple, $G_{i/u} = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$. Toutefois dès que des

systèmes complexes sont abordés les calculs formels induits deviennent rapidement trop lourds et complexes pour être réalisés manuellement. L'utilisation de logiciels de calcul formel devient alors indispensable. Les méthodes directes permettent d'éviter ces calculs fastidieux.

5.3.2 Méthodes directes

Cette approche a été préférée aux précédentes. L'algorithme d'optimisation ajuste directement les paramètres physiques recherchés. Une résolution analytique du problème étant impossible, cette méthode nécessite également l'usage d'algorithme de programmation non-linéaire. Pour éviter un

biais sur les paramètres, le critère de coût est basé sur la formulation d'un critère d'erreur de sortie dans le domaine temporel.

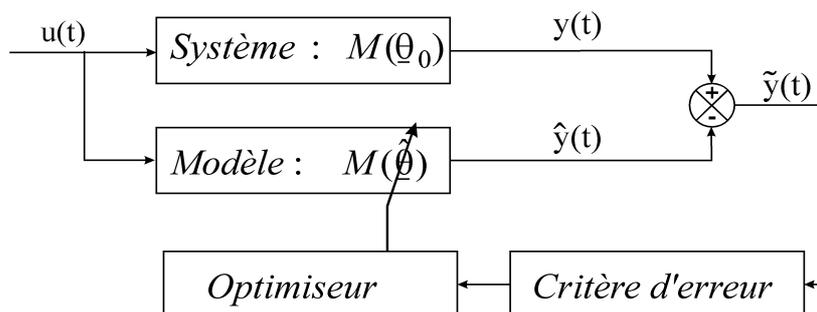


figure 1 - 11 : identification des paramètres physiques, méthode directe

L'algorithme d'optimisation non-linéaire étant d'ores et déjà incontournable, il apparaît logique de le faire ajuster directement les paramètres physiques, évitant ainsi le problème d'inversion (1.5-20). Aucun calcul formel n'est donc nécessaire et la seule possibilité de simuler le système à temps continu est requise. Dans ces travaux, la méthode de la matrice exponentielle (avec un bloqueur d'ordre zéro) a été retenue. Le bloc diagramme suivant schématise cette approche dans le domaine temporel :

Comme pour les méthodes précédentes, cette méthode pose l'hypothèse d'absence d'erreur de caractérisation. Cette manière de procéder est aisément transposable au domaine fréquentiel. Dans un premier temps, notre étude se restreint au domaine temporel. Dans le domaine fréquentiel le critère d'erreur repose sur une fonctionnelle de l'écart entre le diagramme de Bode du modèle et le résultat d'une analyse spectrale réalisée sur les signaux relevés sur le processus.

Elle permet donc de profiter des avantages propres au domaine fréquentiel, à savoir

- La possibilité d'augmentation du rapport signal/bruit ; les fréquences non excitées peuvent être éliminées (d'où réduction des intervalles d'incertitude associés aux paramètres).
- La compression de données ; les nombreuses données temporelles sont avantageusement remplacées par quelques raies fréquentielles.
- L'absence de problèmes liés aux conditions initiales du système.
- La combinaison aisée de données issues de plusieurs expériences.

Cette approche fréquentielle a déjà été partiellement exploitée et implémentée dans la toolbox « Frequency Domain System Identification Toolbox » de Matlab[®] [KOL95]. Notons ici que les algorithmes d'optimisation utilisés dans ces programmes sont des algorithmes de recherche locale : rien ne garantit donc la globalité de la solution obtenue. De plus, ces modèles continus recherchés doivent être exprimés sous la forme de **fonctions de transfert continues**. L'utilisation de cette boîte à outils de Matlab[®] pour réaliser une approche **directe** pour l'identification des paramètres physiques est donc impossible.

6 CONCLUSION

Dans ce chapitre les **éléments d'identification** nécessaires pour justifier les choix et converger vers une solution au projet décrit au quatrième chapitre de ce mémoire ont été exposés. Les principes et les algorithmes permettant de garantir les propriétés d'identifiabilité et discernabilité sont décrits. Ces propriétés doivent être tout ou en partie remplies pour mener à terme un processus d'identification des paramètres physiques de système à temps continu.

Différentes formulations de critères d'erreur permettant de juger de l'adéquation du modèle au processus réel sont présentées. Ces différentes formulations tiennent compte d'une éventuelle information disponible a priori sur le système. Parmi celles-ci, la formulation au sens du **maximum de vraisemblance** offre toutes les qualités propres à un estimateur idéal tout en requérant un minimum d'informations sur le système et sur les perturbations agissant sur celui-ci.

Les hypothèses nécessaires aux différentes formulations de critères d'erreur peuvent être violées par la présence de données aberrantes sur le jeu de données prélevées sur le système réel. Afin de se prémunir de l'influence de telles données, différents schémas particulièrement efficaces de robustification (M-estimateur) du critère de coût sont présentés. Parmi ces schémas, l'utilisation des plus efficaces théoriquement (M-estimateurs redescendant) peut être source de problème pour les méthodes d'optimisation non-linéaires : création de minima locaux et de discontinuités. Une **robustification par fonction de Huber semble offrir le meilleur compromis efficacité/inconvénient**.

Puis l'identification des systèmes à temps continu est brièvement abordée. Les résultats issus de l'estimation des paramètres de systèmes continus peuvent être le point de départ d'une démarche d'identification des paramètres physiques de systèmes à temps continu. Généralement ces approches nécessitent rapidement de lourds calculs formels. Pour cette raison, nous retenons un processus d'identification des paramètres physiques reposant sur **la minimisation d'un critère d'erreur au sens du maximum de vraisemblance de l'erreur de sortie**. L'algorithme de programmation non-linéaire utilisé ajustant directement les paramètres recherchés, cette approche ne requiert aucun calcul formel. De plus, sous réserve que le modèle soit parfaitement caractérisé, cette démarche offre sur le plan théorique une estimation non-biaisée des paramètres en présence de bruit de mesure sur les données.

La formulation du critère de coût au sens du maximum de vraisemblance est réputée être source de difficultés pour les algorithmes d'optimisation non-linéaires. Le chapitre suivant a pour but de mettre en œuvre et d'évaluer les performances de ces algorithmes afin de retenir le plus approprié à l'identification des paramètres physiques de systèmes continus complexes.

TECHNIQUES D'OPTIMISATION

1 INTRODUCTION

Dans le chapitre précédent différents éléments nécessaires à l'identification de paramètres physiques de systèmes complexes à temps continu sont abordés. Généralement les paramètres physiques sont identifiés par recherche d'une identité de comportement entre un modèle et le processus réel. Différentes formulations de critère de coût représentatif d'un écart de comportement entre le processus et le modèle ont été abordées. Lorsque le critère de coût est non-linéaire par rapport aux paramètres l'expression analytique de la solution $\hat{\theta}$ minimisant ce critère $J(\cdot)$ est impossible. La minimisation de ce critère doit alors s'opérer à l'aide d'algorithmes relevant des techniques de programmation non-linéaires. Afin de garantir la globalité de la solution obtenue de puissantes méthodes d'optimisation globales ont été présentées dans ce chapitre.

D'autres disciplines faisant intervenir un processus d'optimisation, où d'une part les problèmes ne sont pas formalisables analytiquement et d'autre part font intervenir un nombre important d'inconnues, gagnent à utiliser ces méthodes. Citons pour mémoire :

- ☞ Le monde **l'économie** où elles permettent des simulations de l'évolution des marchés.
- ☞ Le domaine de la **biologie moléculaire** pour la détermination de la géométrie d'une molécule.
- ☞ Le domaine de **l'automatique** avec la majorité des problèmes de **commande optimale** et certains problèmes **d'identification**.

Ces algorithmes d'optimisation sont itératifs et requièrent donc un critère d'arrêt. Après avoir présenté la problématique de l'optimisation, un critère d'arrêt s'appuyant sur plusieurs considérations sera développé, ce critère d'arrêt des itérations exploite la convergence de l'algorithme vers une solution stable, tant dans l'espace paramétrique que dans l'espace d'état. Puis, afin de comparer les différents algorithmes d'optimisation, des critères de performances en terme de précision et de fiabilité seront présentés. Les limitations intrinsèques aux algorithmes d'optimisation classiquement utilisés en identification seront soulevées. L'apport potentiel lié à l'utilisation de **techniques d'optimisation globales** sera montré. Parmi celles-ci les **algorithmes génétiques** (dans une de

leurs variantes) offrent d'intéressantes perspectives. Leurs avantages et inconvénients seront soulignés. En particulier, **différentes hybridations avec des algorithmes plus classiques** permettront une significative augmentation de la précision de la solution et une nette réduction du temps de calcul. Les performances de ces algorithmes seront évaluées sur des fonctions de test classiques dans le domaine de l'optimisation globale.

1.1 Position du problème

Afin de conférer un caractère général à ce chapitre, la problématique de l'optimisation est posée de manière générale (indépendamment de l'application). Cette généralité est atteinte en remplaçant les différentes expressions de critère de coût $J(\cdot)$ du premier chapitre par une fonction $f(\cdot)$ à extrémiser. Définissons maintenant quelques termes indispensables à la poursuite de ce chapitre.

- **Fonction objectif** : f est une fonction définie de $X \rightarrow Y$ à optimiser. La valeur $y = f(\underline{x})$ sera appelée valeur objectif.
- **Optimisation** : Suivant le problème à résoudre, l'optimisation consiste à minimiser un critère de coût ou à maximiser une performance. Un extremum de f est recherché $f(\underline{x}) = f(x_1, x_2, \dots, x_n) \Rightarrow \text{extr}$.
- **Espace de recherche** : l'ensemble X est appelé espace de recherche ou espace paramétrique. $\underline{x} \in X$ définit un point de l'espace paramétrique.
- **Espace de d'état** : l'ensemble Y est appelé espace d'état, espace objectif ou espace des coûts dans le cas particulier d'un problème de minimisation. Un point dans l'espace d'état est défini par $y = f(\underline{x})$.
- **Distance** : une fonction distance (norme euclidienne par exemple) peut être définie dans l'espace paramétrique et dans l'espace d'état.
 - Espace paramétrique : $d_X(\underline{x}_1, \underline{x}_2), X^2 \rightarrow \mathbb{R}^+, \underline{x}_1$ et \underline{x}_2 désignent deux points dans l'espace paramétrique.
 - Espace d'état : $d_Y(\underline{y}_1, \underline{y}_2), Y^2 \rightarrow \mathbb{R}^+, \underline{y}_1$ et \underline{y}_2 désignent deux points dans l'espace d'états.
- **Voisinage** : la notion de voisinage ou de région s'étend aux 2 espaces paramétrique et d'états :
 - Voisinage du point \underline{x}_0 dans l'espace paramétrique :

$$V_X^{\underline{x}_0} = \{ \underline{x} \in X \mid d_X(\underline{x}, \underline{x}_0) < \varepsilon_{d_X} \}$$
 - Voisinage du point \underline{y}_0 dans l'espace d'état : $V_Y^{\underline{y}_0} = \{ \underline{y} \in Y \mid d_Y(\underline{y}, \underline{y}_0) < \varepsilon_{d_Y} \}$
- **Conditions d'optimalité** : comment reconnaître un optimum ? Les conditions nécessaires et les conditions suffisantes sont bien connues. Le calcul de ces conditions nécessite que la fonction objectif soit différentiable en $\hat{\underline{x}}$.
 - **Conditions nécessaires (C.N.)** : si $\hat{\underline{x}}$ est optimal alors
 - § C.N du premier ordre : le gradient est nul.
 - § C.N du second ordre : le Hessien est défini positif.
 - **Condition suffisante** : si le gradient et le Hessien est défini positif alors $\hat{\underline{x}}$ est un optimum local (dans le voisinage de $\hat{\underline{x}}$), $\forall \underline{x} \in V_X^{\hat{\underline{x}}} f(\underline{x}) > f(\hat{\underline{x}})$.

- **Bassin d'attraction** : au voisinage $V_Y^{\hat{y}}$, l'ensemble $B_X^{\hat{x}}$ des points dans X correspondant à ce voisinage constitue un bassin d'attraction centré en \hat{x} , si $\forall \underline{x} \in B_X^{\hat{x}}, f(\hat{x}) < f(\underline{x})$. Dans ce cas \hat{x} est nommé **optimum local**.
- **Solution** : les solutions globales (points optimaux) $\underline{x}_0 \in X$ sont définis par $\forall \underline{x} \in X, \underline{x} \neq \underline{x}_0, f(\underline{x}_0) < f(\underline{x})$ où « < » explicite l'extrémisation recherchée (minimisation ou maximisation). Dans la suite du mémoire, lorsque l'extrémisation n'est pas explicitée, l'extrémisation considérée sera une minimisation de coût.

1.2 Application à l'identification

Dans le cadre de l'identification, les algorithmes de programmation non-linéaire permettent le passage d'une représentation non-paramétrique (temporelle ou fréquentielle) à une représentation paramétrique.

La fonction f (fonction de coût) permet le **passage de l'espace d'état à l'espace paramétrique**. En général f n'est pas bijective, il peut donc exister plusieurs optima globaux : cas typique d'un modèle non structurellement globalement identifiable mais sli (cf. premier chapitre). Le critère de coût formulé au sens du maximum de vraisemblance est réputé pour présenter des optima locaux [WAL97]. La présence d'erreurs de modélisation est également source de minima locaux.

Dans le chapitre précédent, une robustification efficace des paramètres estimés vis à vis de données aberrantes a été rendue possible par l'usage de M estimateurs. L'utilisation des M estimateurs n'est pas sans inconvénient. En effet, les surfaces de coût engendrées par les critères de coût robustifiés présentent de très nombreux minima locaux. En particulier, lors d'une robustification par fonction de Tukey.

Ces surfaces montrent l'intérêt de disposer d'un algorithme d'optimisation globale pour l'identification des paramètres de modèle phénoménologique.

Dans son livre [LJU87], Ljung déplore le manque d'adéquation des algorithmes classiques d'optimisation pour l'identification. En particulier, il constate que dans leur grande majorité, ils sont sensibles aux différents minima locaux entraînés par les critères de type erreur de prédiction. D'après lui, ces difficultés motivent la recherche et l'utilisation d'autres méthodes d'optimisation.

Avant d'aborder ces algorithmes et pour pouvoir les comparer objectivement, nous allons définir des critères de performance. De plus, ces méthodes d'optimisation sont itératives et requièrent donc un critère d'arrêt. Ces deux aspects font l'objet du paragraphe suivant.

2 CRITERES D'ARRET ET DE PERFORMANCE

Le principe des algorithmes de type PNL est de progresser itérativement dans l'espace paramétrique à partir d'un ou de plusieurs points initiaux. Afin d'opérer un choix parmi les nombreuses techniques existantes, des critères de comparaison doivent être utilisés. Avant d'aborder ces critères définissons les points suivants dans l'espace paramétrique.

- \underline{x}^* : point stationnaire vers lequel converge l'algorithme.
- \hat{x} : solution finale estimée.
- $\hat{x}(k)$: solution estimée intermédiaire à l'itération k .
- \underline{x}_0 : solution théorique, supposée connue du problème d'optimisation.
- \tilde{x} : erreur d'estimation $\tilde{x} = \underline{x}_0 - \hat{x}$.

Principalement 6 critères se dégagent de la littérature consacrée à ce sujet. Ceux-ci quantifient respectivement la convergence de l'algorithme vers un point stationnaire, la robustesse de la méthode, la simplicité d'utilisation, le temps de calcul, la précision et la fiabilité.

- **La convergence** : montrer la convergence d'un algorithme itératif vers une solution revient à montrer que la séquence des points $\hat{\underline{x}}(k)$ tend vers une valeur \underline{x}^* pour un nombre croissant d'itérations, soit $\lim_{k \rightarrow \infty} d_X(\hat{\underline{x}}(k), \underline{x}^*) \rightarrow 0$.
- **La robustesse** : la robustesse illustre les capacités de l'algorithme à résoudre une large gamme de problèmes, c'est à dire des problèmes de complexités variables de l'espace de recherche. La robustesse n'est pas indispensable si le type de problème est connu et fixe dans le temps.
- **La simplicité d'utilisation** : elle est liée au nombre de paramètres internes à l'algorithme à ajuster pour assurer un bon déroulement du processus d'optimisation.
- **La précision** : la précision estime la distance de la solution finale à la solution réelle du problème, soit $d_X(\hat{\underline{x}}, \underline{x}_0)$.
- **La fiabilité** : un algorithme d'optimisation est fiable si la solution obtenue est globale. La solution estimée appartient alors au bassin d'attraction de \underline{x}_0 , $\hat{\underline{x}} \in B_X^{\underline{x}_0}$.

Dans le cadre de l'identification

Dans le contexte de l'identification des paramètres physiques, il est avantageux de disposer d'un algorithme général d'optimisation convergent, fiable, précis et requérant un temps de calcul raisonnable. Sa robustesse permet l'identification des paramètres physiques de systèmes variés de complexités diverses.

Parmi les critères précités, la convergence, la fiabilité, la précision et le temps de calcul sont les plus critiques.

2.1 Critères de précision et de fiabilité

Toutes les méthodes numériques itératives d'optimisation approximent une solution avec une précision donnée. Dans le cas idéal où la solution \underline{x}_0 au problème d'optimisation est connue, la précision de la solution finale $\hat{\underline{x}}$ peut alors être quantifiée. A l'exception de fonctions des test connues, la solution \underline{x}_0 n'est généralement pas connue. Une quantification de l'erreur commise par l'approximation de \underline{x}_0 par $\hat{\underline{x}}$ est donc impossible. Tout au plus, via l'algorithme présenté ci-dessous, une estimation de l'erreur commise par l'approximation de \underline{x}^* par $\hat{\underline{x}}$ est possible. Sous réserve que l'algorithme converge vers $\underline{x}^* \rightarrow \underline{x}_0$, la précision calculée est représentative de l'erreur d'estimation normalisée commise sur chacune des composantes du vecteur solution :

$$\tilde{x}_{\text{Nor } i} = \frac{x_{0i} - \hat{x}_i}{\hat{x}_i} \quad \forall i = \{1(1)m\} \quad (2.2-1)$$

La structure de la fonction objectif, le critère d'arrêt, la précision de la machine sont alors des facteurs déterminants de la précision maximale possible. Compte tenu de ces considérations, la méthode présentée ci-dessous inspirée des travaux de Hyslop [HYS72] permet de borner la précision de la solution obtenue.

Considérant la meilleure solution estimée $\hat{\underline{x}} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m)^T$, ses différentes composantes sont perturbées (toutes combinaisons possibles) par les valeurs :

$$\Delta x_i = \begin{cases} \pm \delta_j & \text{si } x_i = 0 \\ \pm \delta_j x_i & \text{si } x_i \neq 0 \end{cases} \quad (2.2-2)$$

δ_j permet de borner la précision de la solution obtenue.

$$\delta = (\delta_1, \delta_2, \dots, \delta_j, \dots, \delta_{\max})^T$$

Par exemple:

$$\delta = (10^{-8}, 10^{-6}, 10^{-3}, 10^{-1})^T$$

(2.2-3)

La solution \hat{x} est alors qualifiée de précise à au moins δ_j près si pour toutes les combinaisons possibles du vecteur solution

$$f(\hat{x}) - f(\hat{x}, \delta_j) > 0 \quad (2.2-4)$$

L'exemple suivant à deux dimensions ($m=2$) conduit à tester 8 ($3^m - 1$), [HYS72], valeurs du vecteur solution perturbé pour chaque valeur de δ_j , voir figure 2 - 1.

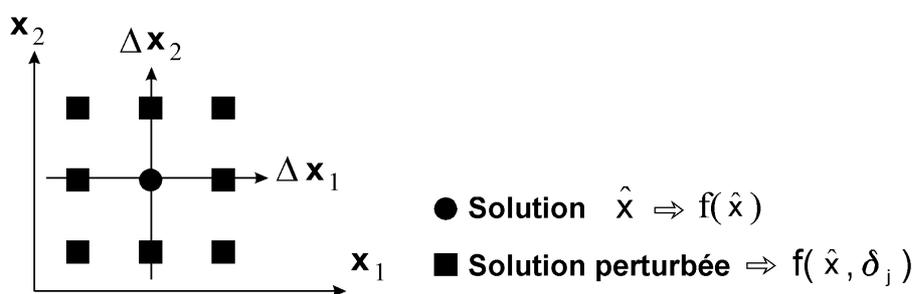


figure 2 - 1 : précision de la solution

Les solutions \underline{x}_0 des fonctions de tests sur lesquelles les performances des algorithmes seront évaluées sont connues. Il est donc possible de donner une estimation de la précision de la solution en calculant l'expression suivante en lieu et place de (2.2-4)

$$f(\hat{x}) - f(\underline{x}_0, \delta_j) > 0 \quad (2.2-5)$$

L'algorithme d'optimisation est alors qualifié de fiable si toutes les composantes de solution \hat{x} sont localisées avec une précision supérieure ou égale à la plus grande valeur de δ . La valeur maximale de δ est fixée de manière à rejeter les solutions rendues inexploitable par leur faible précision.

2.2 Critère d'arrêt

L'arrêt du processus itératif de recherche numérique, en assurant une réelle convergence vers un optimum est un point délicat. Généralement, les critères d'arrêt diffèrent d'un algorithme d'optimisation à l'autre. Classiquement ces critères d'arrêt (ca) se définissent soit dans l'espace paramétrique ou dans l'espace d'état.

□ Espace paramétrique

Un terme est mis aux itérations lorsque :

$$\bullet \quad d_x(\hat{x}(k - \Delta k), \hat{x}(k)) \leq \varepsilon_{d_x} \quad (\text{ca1})$$

$$\bullet \quad |\hat{x}(k - \Delta k) - \hat{x}(k)| \leq \varepsilon_x \quad (\text{ca2})$$

Avec Δk un nombre d'itération prédéfini. La valeur absolue représente dans le cas présent la valeur absolue composante par composante du vecteur et non son module.

□ Espace d'état

Un terme est mis aux itérations lorsque :

$$\bullet \quad d_Y(f(\hat{x}(k - \Delta k)), f(\hat{x}(k))) \leq \varepsilon_{d_Y} \quad (\text{ca3})$$

$$\bullet \quad \left| \frac{\partial f(\hat{x}(k))}{\partial \hat{x}(k)} \right| \leq \varepsilon_{\text{gr}} \quad (\text{ca4})$$

Une possibilité alternative est de fixer un nombre prédéfini d'itération maximum (ca5) ou un nombre maximum d'évaluation de la fonction objectif (ca6).

Isolément, les critères d'arrêt (ca1) à (ca6) ne peuvent être considérés comme satisfaisant. Les critères (ca5) et (ca6) relèvent de choix plus ou moins arbitraires et conditionnent l'arrêt en cas de problème de convergence. Pourtant, ces critères sont très fréquemment utilisés pour stopper le processus itératif des méthodes de type algorithmes génétiques.

Les conditions nécessaires aux deux premiers critères (ca1) et (ca2) peuvent être remplies même loin de l'optimum global, voir cas (1) figure 2 - 2. Une fonction objectif présentant un relief modéré peut satisfaire les critères (ca3) et (ca4), voir cas (2) figure 2 - 2.

Le critère (ca4) souvent utilisé pour les méthodes basées sur une descente de gradient n'est pas toujours approprié en présence de contraintes sur les paramètres ou de discontinuités de la surface de coût.

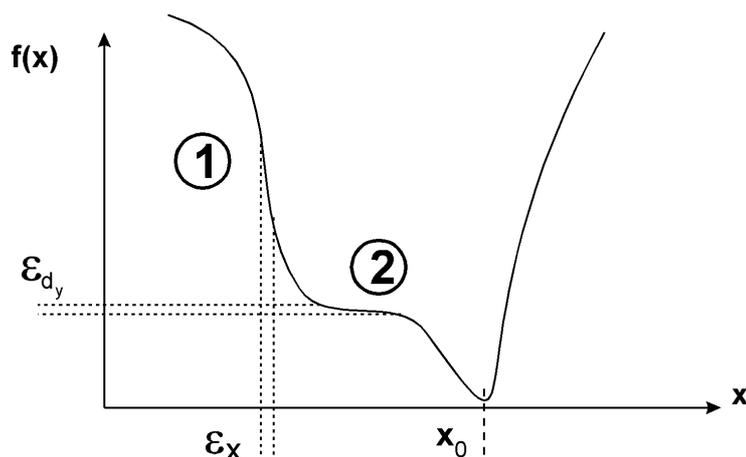


figure 2 - 2 : critère d'arrêt

Pour diverses raisons, le critère d'arrêt final retenu s'inspire des travaux d'Himmelblau [HIM72] et de Press & al [PRE86] en associant (ca2), (ca3) et (ca5) d'où le critère final $ca = ca2 \& ca3 \mid ca5$. Ce critère garantit l'arrêt de l'algorithme d'optimisation dans le voisinage de \underline{x}^* sous réserve que le nombre d'itération maximum ne soit pas dépassé.

2.3 Aspects numériques

Les critères d'arrêt et de précision sont indissociables de la plate forme utilisée, à savoir un ordinateur de type PC Pentium II 233 MHz, 64 Mo de RAM, logiciel Matlab[®] v5.2.

La taille finie du format de représentation des nombres borne les valeurs des nombres flottants, soit :

$$\bullet \quad \text{Plus petit nombre réel absolu : } 2.225 * 10^{-308} \quad (2.2-6)$$

$$\bullet \quad \text{Plus grand nombre réel absolu : } 1.7976 * 10^{308} \quad (2.2-7)$$

La taille de la mantisse est également limitée. Elle impose une limite de précision relative exprimée par les égalités suivantes :

$$\bullet \quad \varepsilon_1 = 0 \text{ si } |\varepsilon_1| < 1.1125 \cdot 10^{-308} \quad (2^{-1023}) \quad (2.2-8)$$

$$\bullet \quad \varepsilon_2 + 1 = 1 \text{ si } |\varepsilon_2| < 1.1102 \cdot 10^{-16} \quad (2^{-53}) \quad (2.2-9)$$

Ces quantités sont importantes lors du test de certaines inégalités (par ex. (2.2-4) et (2.2-5)). La quantité ε_2 est généralement appelée précision de la machine. Les valeurs minimales de δ (2.2-2) doivent tenir compte de ces limites. Notamment la valeur de ε_2 est importante lors du test du critère d'arrêt. Press & al [PRE86] conseillent de fixer ε_{dx} (ca1) à une valeur de l'ordre de la racine carrée de ε_2 , et chacune des composantes de $\underline{\varepsilon}_x$ (ca2) à des valeurs de l'ordre de grandeur de ε_2 . Selon les mêmes auteurs une valeur de ε_{dy} (ca3) de l'ordre de ε_2 est appropriée (préférer une valeur légèrement supérieure afin de ne pas être affecté par des erreurs d'arrondi). L'efficacité et la pertinence des critères d'arrêt et de performance précédemment décrits ne peuvent être aisément démontrées que sur des fonctions de test dont les solutions sont connues. Le paragraphe suivant présente quelques fonctions de tests classiquement utilisées dans le domaine de l'optimisation, sur lesquelles les performances des algorithmes seront évaluées.

2.4 Fonctions de test

Les performances des algorithmes présentés dans ce mémoire seront évaluées sur les fonctions de test suivantes. Afin de représenter ces fonctions, la dimension de l'espace de recherche est limitée à 2.

□ Fonction de De Jong n°1

Cette fonction est continue convexe et unimodale. La minimisation de cette fonction ne pose aucun problème pour toutes les méthodes d'optimisation.

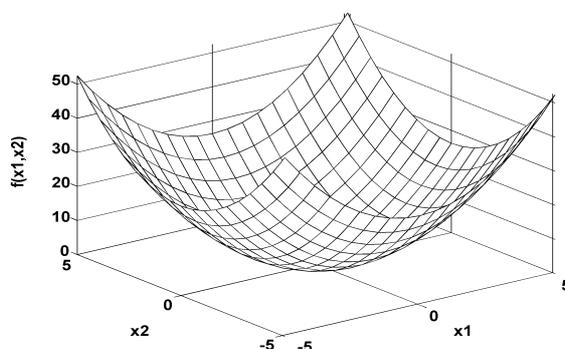


figure 2 - 3 : fonction de De Jong n°1

$$f_1(\underline{x}) = \sum_{i=1}^m x_i^2, \quad -5.12 \leq x_i \leq 5.12 \quad (2.2-10)$$

$$\text{A l'optimum } f_1(\underline{x}_0) = 0, \quad \underline{x}_0(i) = 0, \quad i = 1(1)m \quad (2.2-11)$$

□ Fonction de Griewank n°8

Cette fonction s'inspire de la précédente avec l'addition d'un terme cosinus entraînant de nombreux minima locaux. Ainsi cette fonction est hautement multimodale. Les locaux sont régulièrement distribués sur tout l'espace de recherche (cf. figure 2 - 4). Sur cette figure, les valeurs de la fonction de Griewank sont reproduites de deux points de vue différents (zoom).

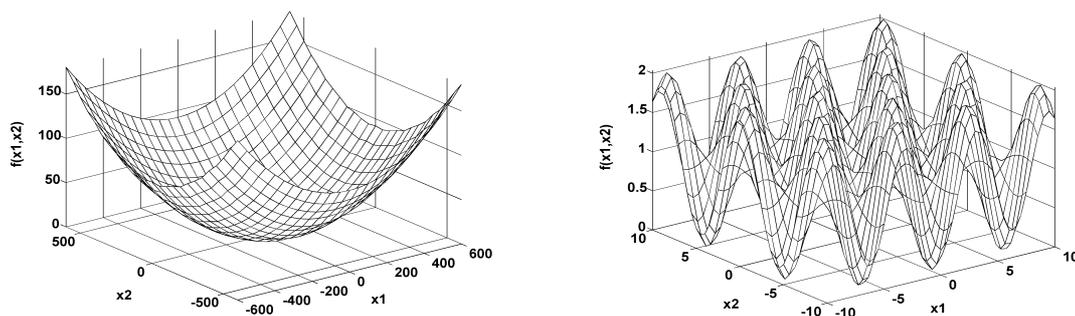


figure 2 - 4 : fonction de Griewangk n°8

$$f_g(\underline{x}) = \sum_{i=1}^m \frac{x_i^2}{4000} - \prod_{i=1}^m \cos(x_i / \sqrt{i}) + 1, \quad -600 \leq x_i \leq 600 \quad (2.2-12)$$

$$\text{A l'optimum } f_g(\underline{x}_0) = 0, \quad \underline{x}_0(i) = 0, \quad i = 1(1)m \quad (2.2-13)$$

3 ALGORITHMES

De très nombreux algorithmes relevant des techniques de programmation non-linéaire ont été développés dans les années 60 et 70. Les algorithmes se distinguent suivant leurs stratégies d'exploration et d'exploitation de l'espace de recherche. Une première stratégie est d'explorer l'espace de recherche à partir d'un unique point. Ces méthodes basées sur ce type de stratégie sont qualifiées de grimpeurs et n'exploitent que les informations locales au point courant pour converger vers une solution. Lorsque ces méthodes convergent vers un optimum, il n'y a aucune garantie que ce point soit un optimum global. Les extrema atteints sont optimaux dans le voisinage du point de départ.

Une seconde stratégie est de travailler sur un ensemble de solutions candidates distribuées sur tout l'espace de recherche. Les extrema alors obtenus sont alors globaux avec une probabilité plus ou moins grande selon les algorithmes. La majorité de ces algorithmes travaillent sur un ensemble de solutions.

3.1 Algorithmes de type grimpeurs

Le principe de ces algorithmes est de progresser itérativement dans l'espace paramétrique, dans des directions choisies de manière à converger vers l'optimum local. Une formulation usuelle de ces algorithmes est :

$$\hat{\underline{x}}(k+1) = \hat{\underline{x}}(k) + \alpha(k) \underline{d}(k) \quad (2.3-1)$$

Où $\hat{\underline{x}}(k)$ est le vecteur des paramètres estimés à l'itération k , $\underline{d}(k)$ est la direction de recherche à l'itération k et $\alpha(k)$ est un scalaire fixe ou calculé tel qu'une extrémisation de $f(\underline{x})$ soit obtenue. Les stratégies de choix des directions de recherche et du pas de calcul $\alpha(k)$ différencient les différents algorithmes de type grimpeurs. Généralement ces méthodes se distinguent par l'ordre de la dérivée de la fonction objectif calculée pour la mise à jour du vecteur des paramètres (2.3-1).

□ Les méthodes d'ordre 0 (ou méthodes directes) ne requièrent aucun calcul de dérivée. Les plus connues sont :

- La méthode de Gauss.
- La méthode de Powell.
- La méthode des polyèdres flexibles de Nelder & al [NEL65].

- Les méthodes du premier ordre utilisent les dérivées partielles premières de $f(x)$, typiquement la méthode du gradient.
- Les méthodes du second ordre (ou de Newton) exploitent les dérivées secondes partielles de $f(x)$.

Dans un premier temps, nous allons nous attarder sur les méthodes d'ordre 0 et notamment particulièrement sur celle des polyèdres flexibles, réputée particulièrement robuste. Les méthodes du premier ordre de base (algorithme de la plus grande pente par exemple) sont généralement très lentes et fortement déconseillées.

Les algorithmes de type Newton sont en général rapides mais peuvent poser des problèmes de stabilité. Des modifications ont été apportées aux algorithmes de base afin d'augmenter leur stabilité tout en conservant leur rapidité de convergence. L'algorithme de quasi-Newton est un de ces algorithmes. Le choix de l'algorithme quasi-Newton est justifié et son principe brièvement présenté.

3.1.1 Méthode des polyèdres flexibles

Bien que relativement lente lorsque la dimension de l'espace de recherche dépasse 8-10, elle se révèle dans certains cas extrêmement robuste [PRE86]. En effet, aucune hypothèse n'est faite sur la topologie posée par la fonction à optimiser. Son pseudo-code est brièvement présenté⁸.

Cette méthode progresse dans l'espace paramétrique par construction successive d'un polyèdre (simplex) à $m+1$ sommets (m : dimension de l'espace de recherche). Plaçons-nous dans le cas où l'extrémisation souhaitée est une minimisation. A chaque itération, dans la majorité des cas, le sommet du polyèdre présentant la valeur objective la plus élevée est remplacé par un point de valeur objective plus faible.

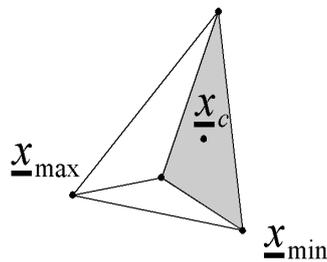


figure 2 - 5a : initialisation

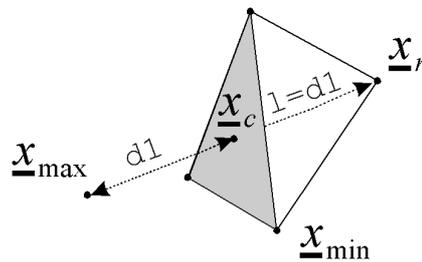


figure 2 - 5b : réflexion

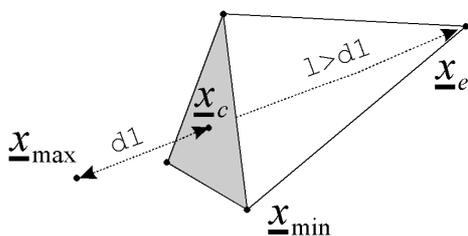


figure 2 - 5c : expansion

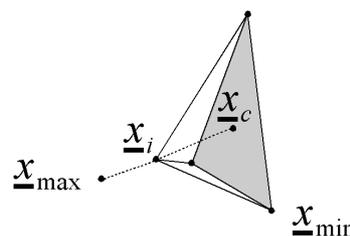


figure 2 - 5d : contraction partielle

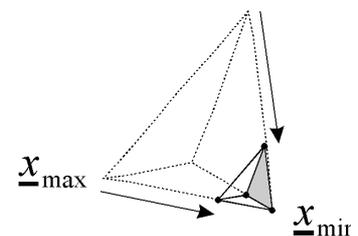


figure 2 - 5e : contraction totale

figure 2 - 5 : algorithme des polyèdres flexibles

⁸ L'heuristique de cet algorithme sera reprise ultérieurement dans ce mémoire pour créer un opérateur original de recombinaison des algorithmes génétiques.

Afin de pouvoir visualiser son principe de fonctionnement sur une figure, fixons, sans perte de généralité, la dimension de l'espace de recherche à 3 ($m=3$). Cette figure illustre les cinq transformations géométriques élémentaires réalisées sur le simplex. Le pseudo-code suivant explicite le fonctionnement de l'algorithme du simplex.

(1) Initialisation

Les différents sommets initiaux du simplex sont obtenus par $\underline{x}_j = \underline{x}_{ini} + \kappa \underline{e}_j, \forall 1 \leq j \leq m+1$ (cf. figure 2 - 5a) avec

- \underline{x}_{ini} : une première estimation de \underline{x}_0
- \underline{e}_j : vecteur unitaire base dans la direction j de l'espace de recherche
- κ : scalaire typiquement de valeur 1
- Valeurs de constantes classiquement ([NEL65], [PRE86]) utilisées par l'algorithme $\nu = 1, \xi = 2, \eta = 0.5$

(2) Détermination de \underline{x}_{min} , \underline{x}_{max} et \underline{x}_c , (cf. figure 2 - 3a)

- $\underline{x}_{max} \mid f(\underline{x}_{max}) = \max(f(\underline{x}_j)) \forall 1 \leq j \leq m+1$
- $\underline{x}_{min} \mid f(\underline{x}_{min}) = \min(f(\underline{x}_j)) \forall 1 \leq j \leq m+1$
- \underline{x}_c : centre de gravité des sommets **autres** que \underline{x}_{max} , $\underline{x}_c = \frac{1}{m} \sum_{j=1, j \neq j_{max}}^{m+1} \underline{x}_j$

(3) Point réfléchi

Le point réfléchi \underline{x}_r est obtenu par déplacement de \underline{x}_{max} dans une direction perpendiculaire au plan défini par les autres points $\underline{x}_r = \underline{x}_c + \nu(\underline{x}_c - \underline{x}_{max})$.

(4) Expansion/réflexion

Si $f(\underline{x}_r) < f(\underline{x}_{min})$,

% l'algorithme tente une expansion en obtenant le point \underline{x}_e

$$\underline{x}_e = \underline{x}_c + \xi(\underline{x}_r - \underline{x}_c)$$

Si $f(\underline{x}_e) > f(\underline{x}_{max})$, $\underline{x}_{max} = \underline{x}_r$ % Réflexion (cf. figure 2 - 5b)

sinon $\underline{x}_{max} = \underline{x}_e$ % Expansion (cf. figure 2 - 5c)

Fin_Si

(5) Contraction partielle/totale

Sinon Si $f(\underline{x}_r) > f(\underline{x}_j)$ pour tout $j = \{1(1)m+1\}, j \neq j_{max}$

$$\underline{x}_i = \underline{x}_c + \eta(\underline{x}_{max} - \underline{x}_c)$$

Si $f(\underline{x}_i) < f(\underline{x}_{max})$, $\underline{x}_{max} = \underline{x}_i$

% Contraction partielle (cf. figure 2 - 5d)

sinon

$$\underline{x}_j = \frac{1}{2}(\underline{x}_{min} + \underline{x}_j), \forall j = 1(1)m+1$$

% Contraction totale (cf. figure 2 - 5e)

Fin_Si

Sinon $\underline{x}_{max} = \underline{x}_r$

% Réflexion (cf. figure 2 - 5b)

Fin_Si

En répétant le processus (2)-(5) un certain nombre de fois, il est possible de localiser le minimum d'une fonction dans un espace de recherche multidimensionnel. La robustesse de cet algorithme est reconnue [PRE86]. Cependant, par rapport aux algorithmes de type Newton, cet algorithme requiert un nombre élevé d'évaluations de la fonction objectif. En particulier lorsque la dimension de l'espace de recherche croît.

3.1.2 Méthode de Newton

Cette méthode est basée sur un développement de $f(\underline{x})$ au second ordre. Dans le voisinage du point $\hat{\underline{x}}(k)$ nous avons

$$f(\underline{x}) = f(\hat{\underline{x}}(k)) + (\underline{x} - \hat{\underline{x}}(k))^T \left. \frac{\partial f(\underline{x})}{\partial \underline{x}} \right|_{\underline{x}=\hat{\underline{x}}} + \frac{1}{2} (\underline{x} - \hat{\underline{x}}(k))^T \left. \frac{\partial^2 f(\underline{x})}{\partial \underline{x}^2} \right|_{\underline{x}=\hat{\underline{x}}} (\underline{x} - \hat{\underline{x}}(k)) + o(\underline{x}) \quad (2.3-2)$$

Soit

$$f(\underline{x}) = f(\hat{\underline{x}}(k)) + (\underline{x} - \hat{\underline{x}}(k))^T \text{Grad}(\hat{\underline{x}}(k)) + \frac{1}{2} (\underline{x} - \hat{\underline{x}}(k))^T \text{Hess}(\hat{\underline{x}}(k)) (\underline{x} - \hat{\underline{x}}(k)) + o(\underline{x}) \quad (2.3-3)$$

Dans cet algorithme, la mise à jour du vecteur des paramètres $\hat{\underline{x}}(k)$ est calculée en annulant les dérivées premières partielles de $f(\underline{x})$ (C.N. du 1^o ordre). Au voisinage du minimum l'expression $\frac{\partial f(\underline{x})}{\partial \underline{x}}$ s'annule. L'incrément à apporter à $\hat{\underline{x}}(k)$ pour annuler $\frac{\partial f(\underline{x})}{\partial \underline{x}}$ s'obtient en dérivant (2.4) et en remplaçant \underline{x} par $\hat{\underline{x}}(k+1)$:

$$\frac{\partial f(\underline{x})}{\partial \underline{x}} = \text{Grad}(\hat{\underline{x}}(k)) + \text{Hess}(\hat{\underline{x}}(k)) (\hat{\underline{x}}(k+1) - \hat{\underline{x}}(k)) = 0 \quad (2.3-4)$$

D'où :

$$\hat{\underline{x}}(k+1) = \hat{\underline{x}}(k) - \text{Hess}(\hat{\underline{x}}(k))^{-1} \text{Grad}(\hat{\underline{x}}(k)) \quad (2.3-5)$$

L'expression (2.6) constitue l'algorithme de Newton de base. Le calcul du Hessien requiert de nombreux calculs. De plus une inversion de cette matrice est nécessaire. Les calculs nécessaires à chaque itérations sont plus lourds qu'avec la méthode du gradient. Il est possible montrer que si $f(\underline{x})$ est quadratique, cette méthode converge en une seule itération vers l'optimum quel que soit le point initial. Même lorsque le critère n'est pas quadratique la convergence est en général très rapide. Cette rapidité se paie par un risque d'instabilité. En effet, pour une initialisation donnée $\hat{\underline{x}}(0)$, rien ne garantit que la valeur de $\hat{\underline{x}}(k+1) - \hat{\underline{x}}(k)$ dans (2.3-4) est suffisamment petite pour que le développement au second ordre de $f(\underline{x})$ soit valable.

3.1.3 Méthode de quasi-Newton

L'algorithme du gradient assure une certaine stabilité au détriment de la rapidité. En revanche, l'algorithme de Newton assure une bonne vitesse de convergence au détriment de la stabilité. Différentes méthodes cherchent à combiner les avantages des deux méthodes. Citons les algorithmes de Gauss-Newton, Newton-Raphson, Levenberg-Marquardt (LM) et quasi-Newton (QN). Parmi ceux-ci effectuer un choix est délicat. Dans les faits LM et QN sont les plus courants. Bard [BAR70] affirme que QN est particulièrement adapté aux problèmes où la fonction à extrémiser s'exprime sous forme de somme de carrés (surface quadratique). Afin de permettre à un non expert d'opérer un choix parmi ces méthodes Lemaréchal [LEM89] suggère «d'utiliser une méthode QN chaque fois que cela est possible». Renders [REN95] quant à lui préconise plutôt LM pour l'optimisation de fonction s'exprimant sous forme de somme de carrés.

Dans le cadre de l'identification, un critère de coût de type maximum de vraisemblance conduit à une surface de coût quadratique. Une robustification du critère peut entraîner de petites modifications de cette surface (cf. premier chapitre). La polyvalence de QN et ses performances comparables à celles de LM sur des surfaces quadratiques ont motivé son choix. L'algorithme QN reprend le développement au second ordre de $f(x)$, et ne calcule pas explicitement la matrice Hessienne $Hess(\hat{x}(k))$ b utilise une approximation de celle-ci notée $H(k)$.

De nombreuses stratégies de mise à jour de la matrice Hessienne ont été développées. Citons les plus connues, les formules de Davidon-Fletcher-Powell (DFP) et celle de Broyden-Fletcher-Goldfarb-Shanno (BFGS). Il est généralement admis que la stratégie BFGS est supérieure à DFP en terme d'erreur d'arrondi et de convergence [PRE86, LEM89, JAN97]. La mise à jour de la matrice Hessienne est obtenue par :

$$H(k+1) = H(k) + \frac{q(k) q(k)^T}{q(k)^T s(k)} - \frac{H(k)^T s(k) s(k)^T H(k)}{s(k)^T H(k)^T s(k)} \quad (2.3-6)$$

où

$$s(k) = \hat{x}(k+1) - \hat{x}(k) \quad (2.3-7)$$

$$q(k) = \left. \frac{\partial f(x)}{\partial x} \right|_{x=\hat{x}(k+1)} - \left. \frac{\partial f(x)}{\partial x} \right|_{x=\hat{x}(k)} \quad (2.3-8)$$

$$H(0) = I \quad (2.3-9)$$

3.1.4 Conclusion

Dans ce paragraphe, les méthodes les plus représentatives des différents algorithmes de type grimpeurs ont été abordées. Leurs avantages et inconvénients sont soulevés. Leurs points forts seront mis à profit pour développer des algorithmes d'optimisation globale. En particulier les algorithmes de type grimpeurs sont incapables de résoudre des problèmes multimodaux. Par exemple, trouver le minimum de la fonction de Griewank avec ce type d'algorithmes est illusoire. La résolution de ce type de problème est possible via l'utilisation d'algorithmes d'optimisation globale.

3.2 Algorithmes d'optimisation globale

L'utilisation d'algorithmes d'optimisation globale a pour but de trouver les optima globaux de $f(x)$. Ces méthodes permettent de contourner les problèmes d'initialisation posés par les méthodes locales (grimpeurs). Rappelons que la recherche de solution dans un espace complexe implique dans la majorité des cas de réaliser un compromis entre 2 objectifs apparemment contradictoires :

- ☐ Une **exploitation** maximale des meilleures solutions disponibles à chaque itération.
- ☐ Une **exploration** robuste de l'espace de recherche.

Les algorithmes précédents de type grimpeurs répondent principalement au seul premier objectif et ne garantissent pas la globalité de l'optimum trouvé. Les méthodes de type recherche aléatoire ou de type énumération systématique conduisent généralement à **des charges de calcul inacceptables en raison de leur lenteur** et de leur ignorance aveugle. D'autres méthodes moins aveugles mettant en œuvre une heuristique plus ou moins poussée existent. Elles tirent profit de la connaissance acquise sur l'espace de recherche lors des itérations précédentes pour guider et accélérer la recherche des solutions futures.

De très nombreuses méthodes d'optimisation globales existent (pour un aperçu plus complet consulter les références [HAN92], [ZHI91]). Parmi ces méthodes, deux grandes approches se dessinent :

Les approches déterministes (analyse par intervalles, le tunneling, le multistart, le branch and bound ...). Les performances des techniques déterministes sont limitées : notamment la complexité des problèmes d'optimisation qu'elles permettent de résoudre est restreinte.

Les approches stochastiques explorent plus ou moins aléatoirement l'espace de recherche pour localiser les solutions optimales. Ces méthodes sont en général associées à des fonctions de mémorisation de l'espace exploré qui dirigent et accélèrent la recherche. Les caractéristiques suivantes sont communes aux algorithmes stochastiques.

- Les points initiaux sont choisis plus ou moins aléatoirement dans l'espace de recherche. Il est clair que cette multiplication du nombre de points de recherche augmente la probabilité d'obtenir plusieurs optima (donc l'optimum global), moyennant une augmentation du temps de calcul.
- Ils se meuvent dans l'espace de recherche par transformations déterministes (opérateurs particuliers, algorithmes de type grimpeurs) et aléatoires (pouvant conduire temporairement à une augmentation de $f(\hat{x}(k))$).

Les méthodes stochastiques sont aussi appelées calculs évolutifs (evolutionary computation). Les plus connues sont le **branch and bound stochastique**, le **recuit simulé**, la **recherche tabou**, la « **threshold accepting strategy** » et le « **great deluge algorithm** » [Han92]. Ces méthodes sont d'utilisation délicate ; les résultats sont très dépendants des paramètres internes de l'algorithme à ajuster. Elles sont donc peu robustes et rarement efficaces sur des problèmes de grande dimension.

D'autres méthodes stochastiques d'inspiration similaire ont été développées ces 30 dernières années et offrent un compromis exploitation/exploration raisonnable. Citons les **algorithmes génétiques** (genetic algorithms) principalement développés aux Etats Unis par **Holland** [HOL75], les **stratégies d'évolution** (evolution strategie) en Allemagne par **Rechenberg** [REC73] et la **programmation évolutive** (evolutionary programming) par **Fogel** [FOG66]. Chacune de ces voies constitue une approche différente mais elles s'inspirent toutes des principes de l'évolution naturelle.

Ces méthodes ont été évaluées lors de benchmarks. Ces travaux, réalisés par des adeptes d'une méthode particulière ne semblent pas aboutir à des conclusions parfaitement objectives.

Les algorithmes génétiques (dans leur version de base) sont largement dépassés en terme de précision, simplicité d'utilisation et temps de calcul par les stratégies d'évolution et la programmation évolutive. Mais le développement relativement récent de nouveaux opérateurs génétiques et leurs hybridations avec des algorithmes de type grimpeurs conduit à l'obtention de méthodes très certainement plus robustes, plus fiables, plus précises et plus rapides. Il est à noter qu'à ce niveau, les AGs hybrides n'ont d'algorithmes génétiques que le nom.

Dans le paragraphe suivant, les principes constitutifs des AGs sont présentés et leurs performances seront évaluées sur les fonction de tests présentées au début de ce chapitre.

4 ALGORITHMES GENETIQUES

Les algorithmes génétiques se distinguent des méthodes d'optimisation classiques sur plusieurs points :

- Ils travaillent sur un **ensemble de points** et non un point unique.
- Les AGs ne nécessitent ni dérivée partielle de la fonction à optimiser ni aucune connaissance auxiliaire (comme les algorithmes de type grimpeur d'ordre 0). Seule la valeur de $f(x)$ est nécessaire. L'utilisateur ne se préoccupe ni de la continuité ni de la différentiabilité de la fonction à optimiser.

- Leurs **règles de transition probabilistes** leur confèrent d'intéressantes propriétés d'exploration.
- Malgré un nombre de paramètres internes relativement élevé, **leur robustesse élevée** simplifie leur utilisation par rapport à d'autres méthodes d'optimisation globale.
- A un problème donné peut correspondre un ensemble de solutions. Les AGs sont particulièrement efficaces pour identifier cet **ensemble de solutions** potentielles⁹.

4.1 Principe

Les AGs sont des algorithmes de recherche stochastiques mimant les lois de l'évolution naturelle. Ils travaillent sur un ensemble de solutions (individus). Les individus les mieux adaptés à leur environnement survivent d'une génération à l'autre. A chaque génération les meilleurs individus se reproduisent par le biais d'opérateurs directement inspirés de la génétique : sélection, recombinaison, mutation et réinsertion. Ce processus répété sur un certain nombre de générations entraîne une meilleure adaptation des individus à leur environnement (cf. figure 2 - 6). Le rôle de l'environnement est joué par la fonction objectif.

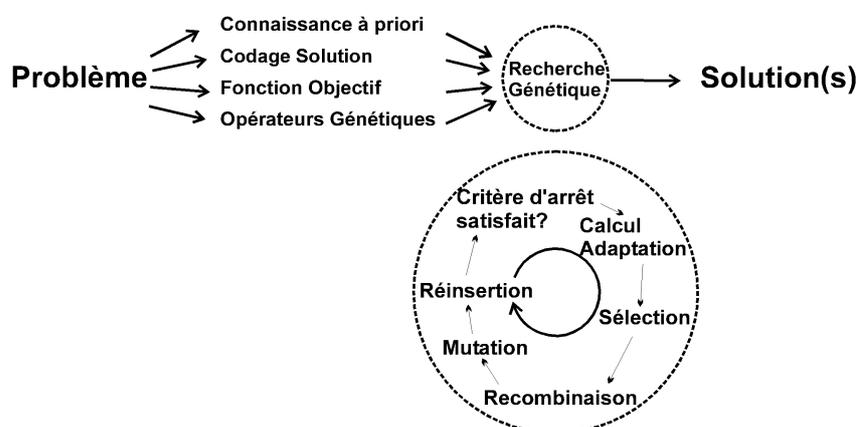


figure 2 - 6 : principe des algorithmes génétiques

La puissance d'un tel algorithme est souvent suffisante pour la résolution d'un grand nombre de problèmes. Cependant, de meilleurs résultats sont parfois obtenus (dans le domaine de l'optimisation multicritère notamment) par l'introduction de plusieurs populations appelées sous-populations. Chacune évolue isolément durant quelques générations (tel un algorithme mono-population), puis un certain nombre d'individus sont échangés entre les sous-populations. Cette diversité de populations isolées conduit plus aisément à la formation d'espèces (satisfaisant plus un critère qu'un autre) qu'un algorithme mono-population.

4.2 Codage

Les algorithmes génétiques travaillent sur un ensemble de solutions potentielles. Le principe des AGs consiste à opérer un codage de ces solutions, puis à travailler sur ces solutions codées nommée chromosomes ou individus. L'ensemble de ces solutions potentielles (individus) constitue une population. Dans la très grande majorité des cas, le génome d'un individu est constitué d'un seul chromosome¹⁰. Donc chaque individu (solution) est codé sur un chromosome constitué de

⁹ Par exemple lors d'une optimisation multicritère, où les solutions se situent sur la frontière de Pareto.

¹⁰ Des individus hétérozygotes (dont les génomes sont codés sur plusieurs chromosomes), avec des notions de dominance sont utilisés lorsque des propriétés d'adaptation à un environnement variable sont recherchées.

gènes (composantes de solution) pouvant prendre plusieurs allèles (valeurs de gène). Le transcoding d'une solution en un individu est défini par le génotype. Le génotype crée donc un espace génotypique associé à l'espace de recherche. L'AG crée donc à partir du génotype un espace génotypique dans lequel les individus sont représentés et manipulés. Des fonctions de codage/décodage permettent le passage de l'espace de recherche à l'espace phénotypique.

Traditionnellement, les gènes sont représentés sous forme de code binaire (Gray). Toutefois des études [JAN91, DAV91] ont montré qu'un codage en nombres réels conduisait à une diminution du temps de calcul, une précision accrue et une utilisation plus aisée. De plus, les résultats sont plus cohérents d'une génération à l'autre du fait de la manipulation directe par les opérateurs génétiques du phénotype de chaque individu et non de sa représentation.

Concernant la notation, les différents éléments constitutifs d'un algorithme génétique, la convention suivante est adoptée :

- λ : nombre d'enfants
- μ : nombre de parents
- $I \in \mathbb{R}^m$: espace paramétrique
- $x \in I$: un individu
- $x_i \in \mathbb{R}$: le $i^{\text{ème}}$ gène de l'individu x
- $P \in I^\mu$: population constituée de μ individus
- $x(j) \in I$: le $j^{\text{ème}}$ individu de la population P (cf. figure 2 - 7).

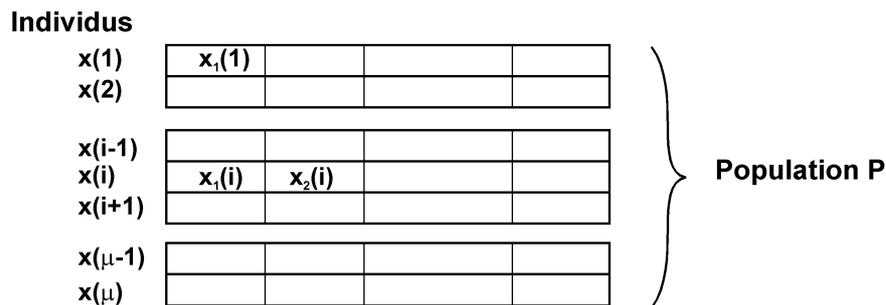


figure 2 - 7 : organisation d'une population d'algorithme génétique

Que les gènes soient codés en nombres binaires ou réels, chaque chromosome est constitué de la concaténation sous forme d'un vecteur ligne de l'ensemble des composantes de solution. La structure du chromosome peut différer de celle de la solution :

<u>Solution</u>	<u>Individus (génotypes possibles)</u>
$x = (x_1 \ x_2 \ x_3 \ x_4)$	$\Rightarrow \text{ind} = (x_4 \ x_2 \ x_1 \ x_3)$
	$\Rightarrow \text{ind} = (x_2 \ x_1 \ x_4 \ x_3)$

Par souci de simplicité, le chromosome sera codé identiquement à la solution :

$$\text{ind} = x = (x_1 \ x_2 \ x_3 \ x_4)$$

Dans la suite de ce mémoire, à moins de le spécifier explicitement, l'espace de recherche et l'espace génotype seront considérés comme équivalents ; un phénotype représentant directement une solution.

4.3 Initialisation

Si aucune connaissance sur la position de l'optimum dans l'espace de recherche n'est disponible, la population initiale est générée aléatoirement par tirage uniforme sur l'espace de recherche défini [MIC91]. Disposant maintenant d'une population initiale, les individus sont évalués par une fonction d'adéquation.

4.4 Fonction d'adéquation

L'opérateur de sélection choisit les parents de la nouvelle génération et détermine le nombre moyen d'enfant de chaque parent. La première étape consiste à calculer l'adéquation de chaque individu à l'environnement. La valeur de la fonction objectif est déjà une indication de l'adéquation de l'individu à son environnement. Dans le cas d'une minimisation, plus cette valeur sera faible, meilleur sera l'individu. La valeur de la fonction objectif n'est qu'une valeur intermédiaire lors du calcul de l'adéquation (aussi nommé fitness dans la littérature anglo-saxonne).

Généralement une fonction $\varphi(\cdot)$ transforme la valeur objectif (cf. chapitre précédent) en une mesure d'adéquation relative à l'ensemble de la population.

$$\varphi : I \rightarrow R$$

$$\Phi(x) = \varphi(f(x))$$

La fonction $\varphi(\cdot)$ transforme la valeur de la fonction objectif $f(x)$ en un nombre réel non-négatif. $\Phi(x)$ est l'adéquation relative de x . Généralement $\Phi(x)$ représente le nombre probable d'enfants engendré par l'individu $x(i)$ dans la génération suivante. La fonction $\varphi(\cdot)$ la plus simple est la fonction d'adéquation proportionnelle

4.4.1 Adéquation proportionnelle

L'adéquation de l'individu i ($\Phi(x(i))$) de la population se calcule à partir de la valeur $f(x(i))$ et de la moyenne de celles de tous les individus de la population :

$$\Phi(x(i)) = \frac{f(x(i))}{\frac{1}{\mu} \sum_{j=1}^{\mu} f(x(j))} \quad (2.4-1)$$

où μ est la taille de la population et $x(i)$ le phénotype de l'individu i .

Cette fonction d'adéquation est uniquement applicable en cas de valeur objectif positive. Goldberg [GOL89] propose alors d'opérer une transformation linéaire de $f(x(i))$ avant de calculer l'adéquation des individus :

$$\Phi(x(i)) = \frac{a f(x(i)) + b}{\sum_{j=1}^{\mu} a f(x(j)) + b} \quad (2.4-2)$$

où a est un scalaire positif ou négatif suivant l'extrémisation recherchée.

L'offset b assure le non-négativité de l'expression $a f(x(i)) + b$. Ce mécanisme possède beaucoup d'inconvénients. En effet, si un individu est très supérieur à la moyenne de la population, la génération suivante sera presque exclusivement constituée de la duplication de cet individu. La population aura alors perdu toute sa diversité génétique (avec un risque de convergence prématurée de l'algorithme). Une adéquation calculée sur le rang pallie ce problème.

4.4.2 Adéquation basée sur le rang

Les individus sont classés par rapport à leur valeur objectif. L'adéquation résultant ne dépend uniquement de sa position dans la population [BAC85]. Considérant μ le nombre d'individu, k sa position dans la population classée ($k=1$ pour le moins bon, $k=\mu$ pour le meilleur) et sp la pression de sélection¹¹, deux possibilités de calcul d'adéquation sont possibles.

□ **Affectation linéaire**

$$\Phi(x(i)) = 2 - sp + \frac{2(sp - 1)(i - 1)}{\mu - 1} \quad (2.4-3)$$

Cette affectation linéaire permet d'imposer une pression de sélection comprise dans l'intervalle [1.0 ... 2.0]. Une affectation non-linéaire permet d'imposer des valeurs de pression de sélection plus élevées.

□ **Affectation non-linéaire**

La fonction d'affectation d'adéquation non-linéaire s'exprime :

$$\Phi(x(i)) = \frac{\mu \gamma^{(i-1)}}{\sum_{j=1}^{\mu} \gamma^{(j-1)}}$$

où γ est racine du polynôme

$$(sp - \mu) \gamma^{(\mu-1)} + sp \gamma^{(\mu-2)} + sp \gamma^{(\mu-3)} + \dots + sp \gamma + sp = 0 \quad (2.4-4)$$

Dans ce cas, la pression de sélection peut se situer dans l'intervalle [1.0 ... $\mu - 2.0$]. La figure 2 - 8 compare graphiquement ces deux affectations linéaire et non-linéaire basées sur le rang.

Dans tous les processus d'optimisation présentés dans ce mémoire, une adéquation basée sur le rang non-linéaire est effectuée. Cette stratégie permet de converger plus rapidement vers une solution au problème d'optimisation posé [BAK85, BAC91, BLI91]. En effet, les individus les plus adaptés à l'environnement possèdent une plus forte probabilité d'être sélectionné pour la phase de reproduction.

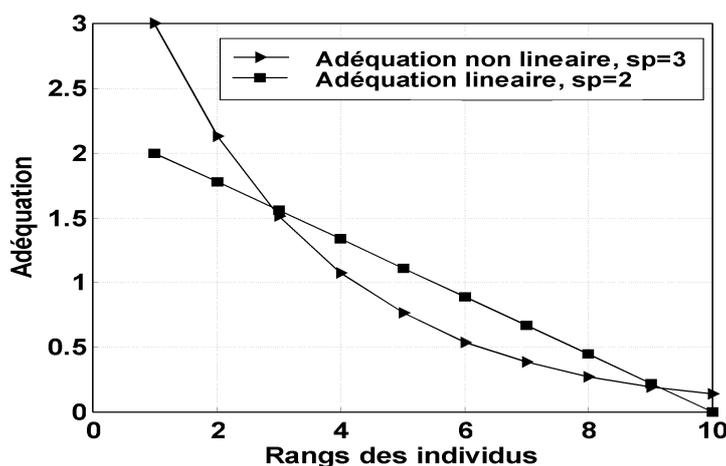


figure 2 - 8 : comparatif de l'adéquation linéaire et non-linéaire basées sur le rang

¹¹ La pression de sélection (selective pressure : sp) définit le nombre probable d'enfants générés par le meilleur individu de la population.

La valeur de l'adéquation de chaque individu représente statistiquement le nombre moyen d'enfant de chaque individu dans la nouvelle génération. L'opérateur de sélection converti ce nombre théorique rarement entier en une nouvelle population d'individus dénombrable.

4.5 Sélection

L'opérateur de sélection sélectionne les individus (relativement à leur adéquation) et les duplique un certain nombre de fois pour créer un nouveau groupe d'individus (enfants). L'algorithme de sélection (parfois appelé reproduction) idéal dupliquerait chaque parent un nombre de fois défini par sa valeur d'adéquation. Ses valeurs sont exceptionnellement entières : différents algorithmes convertissent les valeurs d'adéquation des parents en une population intermédiaire d'individus dénombrable.

□ Echantillonnage stochastique avec remplacement.

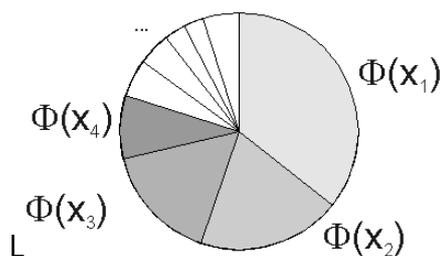


figure 2 - 9 : échantillonnage stochastique avec remplacement

L'algorithme le plus répandu, mais le plus critiquable est celui dit de la roue de loterie ou échantillonnage stochastique avec remplacement [HOL75]. Imaginons une roue de loterie biaisée pour laquelle chaque individu occupe un secteur angulaire proportionnel à son adéquation, cf. figure 2 - 9.

Pour réaliser la sélection, la roue de loterie est tournée λ (nombre d'enfant) fois et sa position d'équilibre définit le parent à dupliquer. Notons qu'avec cette méthode la probabilité de sélectionner λ fois le même individu existe. Ce problème est partiellement résolu avec l'échantillonnage stochastique avec remplacement partiel.

□ Echantillonnage stochastique avec remplacement partiel

Pour résoudre le problème de la duplication un trop grand nombre de fois le même individu, chaque fois qu'un individu est sélectionné la taille de son secteur angulaire est réduite de la taille de celle d'un individu. Si cette taille devient négative, le secteur est effacé de la roue. Cette deuxième technique limite le nombre maximum de parents identiques mais une différence notable est possible entre le nombre de copies d'un individu et sa valeur d'adéquation (présence d'un biais) [BAK87]. La technique suivante permet de réduire ce biais.

□ Echantillonnage stochastique du reste avec (et sans) remplacement

Les méthodes basées sur un échantillonnage du reste opèrent en deux temps. La première étape génère un nombre de copies de chaque individu de manière déterministe. Le nombre de copies est égal à la partie entière de la valeur d'adéquation de chaque individu. Les copies manquantes sont obtenues par échantillonnage stochastique avec remplacement (roue de loterie). Les tailles des secteurs angulaires sont alors proportionnelles à la partie non-entière de l'adéquation de chaque parent. Cette technique souffre du même inconvénient que le tirage par roue de loterie, à savoir une probabilité non-nulle de dupliquer plusieurs fois le même individu lors de la phase d'échantillonnage stochastique. Cet inconvénient peut être contourné en n'effectuant pas de remplacement dans la seconde phase, pour obtenir un échantillonnage stochastique du reste sans remplacement. Devant la complexité de ce dernier algorithme et les inconvénients liés à l'échantillonnage stochastique avec

remplacement partiel Baker [BAK87] propose une stratégie de sélection particulièrement séduisante par sa simplicité et ses propriétés ; l'échantillonnage universel stochastique.

¶ Echantillonnage stochastique universel

Cet algorithme opère en une étape. Dans une optique similaire à la roue de loterie, les individus occupent un segment linéaire sur une droite, proportionnelle à leur adéquation. La somme cumulée de toutes les adéquations est normée à 1. Puis sur la longueur cumulée de ces segments, λ pointeurs (équidistants $\text{distance} = 1/\lambda$) sont disposés. La position du premier pointeur est alors obtenue par tirage aléatoire uniforme entre 0 et 1. La figure 2 - 10 donne un exemple avec un nombre d'enfant $\lambda = 8$.

Cet algorithme assure que chaque individu est sélectionné :

- Au moins un nombre de fois égal à la partie entière de son adéquation.
- Au plus un nombre de fois égal à l'arrondi supérieur de sa valeur d'adéquation.

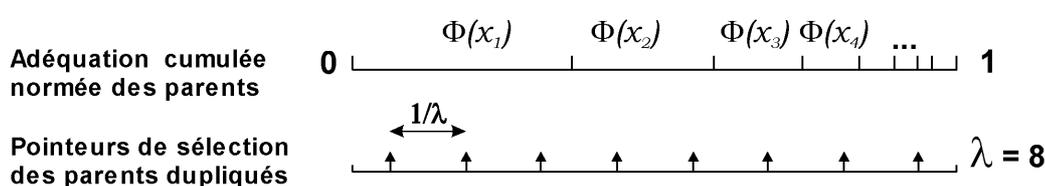


figure 2 - 10 : échantillonnage stochastique universel (sp=3)

L'opération de sélection a permis de dupliquer certains parents et de constituer un ensemble de λ individus enfants. Pour l'instant, ces enfants possèdent les mêmes caractéristiques que leurs parents. L'opérateur de recombinaison va opérer un travail sur le matériel génétique de ces enfants de façon à créer des individus distincts de leurs parents.

4.6 Recombinaison

L'opérateur de recombinaison¹² produit de nouveaux individus en combinant les individus sélectionnés. Généralement, ces individus sont combinés 2 à 2 par tirage aléatoire dans la population. Les individus sont classiquement sélectionnés avec une forte probabilité [0.6 ... 0.95], [GOL89]. Les gènes de ces individus sont échangés et parfois modifiés pour créer de nouveaux individus.

L'opérateur de croisement est indissociable du **théorème fondamental des algorithmes génétiques** [GOL89] : **l'hypothèse de reconstructibilité par blocs** («building blocks hypothesis»). L'opérateur de sélection sélectionne les meilleurs individus de la population. L'hypothèse est faite qu'une combinaison du matériel génétique des meilleurs individus conduise à la formation d'individus aux performances supérieures aux parents initiaux. D'une génération à l'autre, les opérateurs de sélection et de recombinaison cherche à combiner les gènes des meilleurs individus pour aboutir à une combinaison optimale. Si la formulation du problème ne satisfait pas cette hypothèse les AGs échoueront. La validité de cette hypothèse explique l'efficacité des AGs pour la résolution de problèmes générant des surfaces de coût possédant une structure combinatoire sous-jacente (analogie avec le théorème de superposition en électricité) ; ces surfaces engendrent un relief «périodique» où les positions des optima sont interdépendantes.

¹² Le terme de recombinaison s'applique à une représentation réelle des chromosomes. Lorsque qu'un codage binaire est utilisé le terme de croisement est approprié.

Plusieurs implémentations sont possibles. Contrairement à l'opérateur génétique de sélection où les différentes stratégies sont concurrentes, les différents schémas de recombinaison sont complémentaires. Les travaux présentés dans ce mémoire utilisent les recombinaisons :

- **Discrète**
- **Intermédiaire**
- **Continue**

☞ **Recombinaison discrète**

La recombinaison discrète opère un échange entre les gènes des deux individus. Les individus générés (enfants) se situent à l'intérieur d'un hypercube défini par les deux parents (cf. figure 2 - 11 pour une illustration dans un espace de dimension 2).

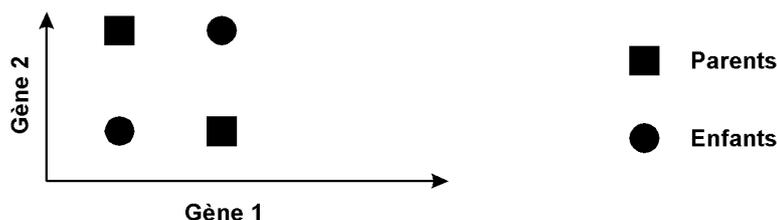


figure 2 - 11 : recombinaison discrète (m=2)

☞ **Recombinaison intermédiaire**

Cet opérateur peut générer des individus à l'extérieur de l'hypercube précédent. Les enfants sont produits par application de la loi

$$\text{rejetons} = \text{parent1} + \alpha(\text{parent2} - \text{parent1}) \quad (2.4-5)$$

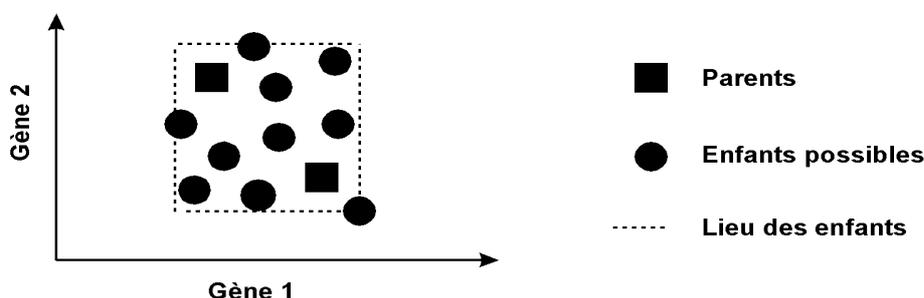


figure 2 - 12 : recombinaison intermédiaire (m=2)

Où α est un facteur d'échelle obtenu par tirage aléatoire uniforme dans l'intervalle $[-d \dots 1+d]$. Une valeur de $d = 0.25$ est recommandée par Mühlenheim [MUH93]. Pour chaque gène une nouvelle valeur de α est tirée (cf. figure 2 - 12 pour une illustration dans un espace de dimension 2).

☞ **Recombinaison continue**

La recombinaison est obtenue par moyennage arithmétique des gènes des parents. Les enfants sont donc le barycentre des parents dans l'espace paramétrique.

L'opérateur de recombinaison est le moteur principal des AGs mais n'est pas suffisant. En particulier, cet opérateur peut conduire à un appauvrissement de la diversité du matériel génétique de la population. L'opérateur de mutation permet de maintenir une certaine diversité génétique au sein de la population. La mutation est un opérateur de second plan dans les AGs.

4.7 Mutation

L'opérateur de mutation agissant sur les enfants consiste à perturber, avec une amplitude restreinte et avec une faible probabilité, la valeur d'un ou plusieurs gènes. Des articles [BAC93, MUH93] sont consacrés à la recherche d'un taux de mutation optimal. Mülhenheim a constaté empiriquement qu'un taux de mutation inversement proportionnel à la dimension de l'espace de recherche offre de bons résultats sur de nombreuses fonctions de test. Notons que ce taux de mutation est indépendant de la taille de la population. Bäck [BAC93] a montré qu'un taux de $1/m$ (m : dimension de l'espace de recherche) était optimal pour l'optimisation de fonction unimodale (où l'utilisation des AGs est sans intérêt !). Ses travaux soulignent par contre l'intérêt d'adopter un taux de mutation auto adaptatif lors de l'optimisation de problèmes générant des surfaces multimodales. Dans nos travaux le taux de mutation est choisi de manière à muter en moyenne un gène sur deux chromosomes (individus).

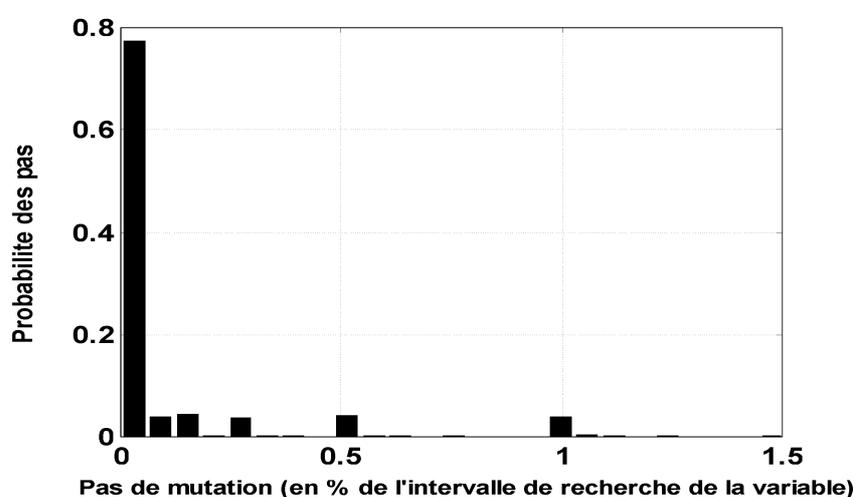


figure 2 - 13 : taille des pas de l'échantillonnage stochastique universel

Le choix de l'amplitude (pas de mutation) de la perturbation apportée au gène muté est un problème réputé délicat. Le pas optimal dépend du problème d'optimisation. Généralement, une faible amplitude est suffisante mais un pas de mutation d'amplitude plus élevée accélèrent souvent la convergence de l'algorithme [MUH93]. Mülhenheim, dans ce même article, propose un opérateur de mutation pour un codage réel des individus offrant ces amplitudes de pas de mutation (amplitudes de pas faibles et élevées). Sa description algorithmique est la suivante :

$$\text{allèle_mutée} = \text{allèle_non_mutée} \pm v \delta \quad (\pm \text{ sont équiprobables}) \quad (2.4-6)$$

$$v = 0.5 * \text{intervalle de recherche de la variable} \quad (2.4-7)$$

$$\delta = \sum_{k=0}^{19} \alpha_k 2^{-k}, \quad \alpha_k = 1 \text{ avec une probabilité } 1/m, \text{ sinon } \alpha_k = 0; \quad m = 20; \quad (2.4-8)$$

Les quantités apportées aux gènes mutés sont statistiquement faibles (cf. figure 2 - 13, les statistiques sont réalisées sur 10000 occurrences). Une fois ces enfants mutés, leurs adéquations sont calculées. Puis tout ou une partie de ces enfants sont réinsérés dans la population initiale.

4.8 Réinsertion

Généralement, le taille de la population est maintenue constante d'une génération à l'autre. Différents schémas de réinsertion des enfants dans la population des parents sont possibles. Les

parents peuvent être remplacés en totalité ou partiellement, de façon déterministe ou aléatoire. Dans notre cas, une stratégie élitiste (déterministe) est adoptée : les parents les plus faibles sont remplacés par les meilleurs enfants. Un taux de réinsertion est ainsi défini. Un taux de 0.95 signifie que 95% des plus mauvais parents sont remplacés par les meilleurs enfants. Notre taux de réinsertion est calculé en fonction du nombre d'individus de la population de manière à conserver le meilleur individu d'une génération à l'autre. Cette stratégie garantit au moins une non-diminution de l'adéquation du meilleur individu de la population. Dans le cas le plus défavorable, cette adéquation stagne (effet de cliquet).

La convention de notation suivante (cf. tableau 2 - 1) est adoptée.

Notation	Signification
λ	nombre d'enfants
μ	nombre de parents
$x \in I$	un individu
$P \in I^\mu$	population constituée de μ individus
$I \in R^m$	espace paramétrique
$\Phi : I \rightarrow R$	fonction de d'adéquation
$f : R^m \rightarrow R$	fonction objectif
$ca : I^\mu \rightarrow \{\text{oui, non}\}$	critère d'arrêt
$s_{\{Ps, \mu\}}$	opérateur de sélection de μ individus, avec une probabilité $\{Ps\}$ par individu
$r_{\{Pc\}} : I^\mu \rightarrow I$	opérateur de recombinaison appliqué avec une probabilité $\{Pc\}$
$m_{\{Pm\}} : I \rightarrow I$	opérateur de mutation appliqué avec une probabilité $\{Pm\}$

tableau 2 - 1 : algorithmes génétiques, convention de notation.

Le pseudo-code ci-dessous (dont la notation s'inspire en partie de [BAC93]) résume succinctement l'algorithmique mise en œuvre pour la programmation d'un AG codé réel intégrant une stratégie élitiste de réinsertion

```

génération := 0;
Initialisation P(0) := {x1(0), ..., xμ(0)} ∈ Iμ
Tant_que ca(P(t) ≠ vrai) faire,
  Evaluer P(t) : {Φ(x1(t)), ..., Φ(xμ(t))}
  pour {x1(t), ..., xμ(t)} ∈ Iμ
  Sélectionner P'(t) := s{Ps, λ}(P(t))
    avec Ps(xk(t)) := Φ(xk(t)) / (1/μ ∑j=1μ Φ(xj(t)));
  Recombiner xk''(t) := r{Pc}(P'(t)) ∀ k ∈ {1, ..., λ};
  Muter xk'''(t) := m{Pm}(xk''(t)) ∀ k ∈ {1, ..., λ};
  Evaluer P''(t) : {Φ(x1'''(t)), ..., Φ(xλ'''(t))}
  Pour {x1'''(t), ..., xλ'''(t)} ∈ Iλ
  Sélectionner P(t+1) := s{Ps, μ}(P(t) ∪ P''(t)) % Réinsertion
    avec Ps(xk(t)) := Φ(xk(t)) / (1/(μ+λ) ∑j=1μ+λ Φ(xj(t)));
génération := génération + 1;
Fin_tant_que

```

Il sera montré (§ 5.7 de ce chapitre) que les AGs offrent de réels avantages par rapport aux méthodes grimpeurs par leurs capacités d'exploration et permettent d'isoler fiablement un optimum global même en présence de nombreux minima locaux. En contrepartie, **la solution est peu précise et le temps de calcul demeure important**. Les AGs hybrides tentent de conjuguer les algorithmes génétiques aux méthodes de type grimpeurs pour aboutir à un algorithme offrant des solutions à la fois précises et fiables en un temps de calcul raisonnable.

4.9 Convergence

L'arrêt du processus d'optimisation par AG est conditionné par la vérification du critère développé au § 2.2 de ce chapitre. Une visualisation temps réelle de la convergence de l'algorithme vers une solution stable tant dans l'espace paramétrique que dans l'espace des coûts est possible via des techniques de visualisation appropriées. Dans l'espace des coûts, une idée précise de l'algorithme est obtenue en traçant (en fonction des générations) :

- La valeur objectif du meilleur individu
- La valeur objectif moyenne de la population
- L'écart type autour de cette moyenne

Et dans l'espace paramétrique en observant le tracé (en fonction des générations) de la valeur de chaque composante de solution :

- Du meilleur individu
- De la moyenne des individus
- De l'écart type autour de la moyenne

Ces graphiques donnent une réelle indication de la convergence de l'algorithme génétique vers une solution stable tant dans l'espace paramétrique que dans l'espace des coûts (cf. § 6.3.4.3 du quatrième chapitre).

5 ALGORITHMES GENETIQUES HYBRIDES

Dans le paragraphe précédent, les éléments constitutifs des AGs ont été abordés et leurs limitations en terme de temps de calcul et de précision ont été soulevées. L'idée largement répandue que les nouveaux algorithmes d'optimisation de type calculs évolutifs rendent obsolètes les méthodes plus classiques est fautive. Les méthodes d'optimisation de type calculs évolutifs possèdent d'indéniables qualités d'exploration mais sont médiocres du point de vue de l'exploitation. Les algorithmes locaux (de type grimpeurs) ont intrinsèquement une puissance d'exploitation élevée mais un pouvoir exploratoire quasi nul. L'idée est donc d'associer ces algorithmes pour combiner leurs avantages en réduisant leurs inconvénients. Pour simplifier, l'AG localise les bassins d'attraction de la fonction à optimiser et l'algorithme local escalade les montagnes associées.

Un algorithme génétique hybride désigne un AG utilisant des algorithmes de recherche autres que génétiques ; comme par exemple des techniques locales de type grimpeur. Plusieurs auteurs [MIC91, POW91] ont proposé de telles hybridations mais peu de travaux décrivent des résultats d'étude sur des AGs hybrides. Un des meilleurs articles à ce sujet est certainement celui de Renders [REN96].

Pour le développement d'algorithmes hybrides, le parallèle entre les disciplines de la génétique et de l'optimisation est poursuivi, en introduisant la notion de connaissance acquise par l'individu durant son existence. Ainsi le concept d'évolution globale de la population est associé aux algorithmes génétiques et celui d'apprentissage de la vie à une phase d'adaptation à l'environnement (par le biais d'opérateur de recherche locale par exemple). Dans ce paragraphe, quatre schémas

d'hybridation sont proposés avec l'utilisation d'un opérateur de croisement original dérivé de l'heuristique de la méthode des polyèdres flexibles.

Parmi ces stratégies d'hybridation, l'hybridation séquentielle est sans aucun doute celle d'inspiration la plus simple. Les performances en terme de temps de calcul sont toutefois modestes. Pour augmenter la rapidité de tel algorithmes, deux autres stratégies d'hybridation s'inspirant des principes de l'évolution selon Darwin et Lamarck sont proposées. Enfin, une nette amélioration des performances de ces algorithmes est possible par le développement d'un nouvel opérateur de recombinaison inspiré de l'heuristique de l'algorithme du simplexe. Ces différentes stratégies d'hybridation sont successivement décrites.

5.1 Hybridation séquentielle

Dans une première phase les algorithmes génétiques explorent l'espace de recherche dans le but de localiser des solutions grossières prometteuses (avec une précision limitée) ; c'est à dire situer ces solutions à l'intérieur du bassin d'attraction de l'optimum global. Dans une seconde phase la méthode d'optimisation locale prend le relais en prenant pour valeurs initiales les 5-10% des meilleurs individus de la dernière génération [GOL89] (cf. figure 2 - 14).



figure 2 - 14 : hybridation séquentielle

Afin de garantir que les AGs fournissent des solutions bien situées un certain nombre de générations est indispensable. Par conséquent, le temps de calcul demeure important. Pour pallier ce problème, deux voies s'inspirent respectivement des thèses de l'évolution selon Darwin et Lamarck. Consulter les travaux de Dobzhansky [DOB70] pour une description génétique plus détaillée de ces deux interprétations de la théorie de l'évolution des espèces. Renders [REN96] exploite ces deux visions de l'évolution pour développer des algorithmes génétiques s'inspirant des théories de Lamarck et Darwin.

5.2 Principes Darwiniens

□ Description génétique

La sélection des individus dépend de leur adéquation à l'environnement. Le matériel génétique utilisé lors de la reproduction est le matériel donné à la naissance (l'inné). Ce matériel est invariant durant la vie de l'individu. Au fil des générations, le matériel génétique composant la population produit des adultes de plus en plus performants.

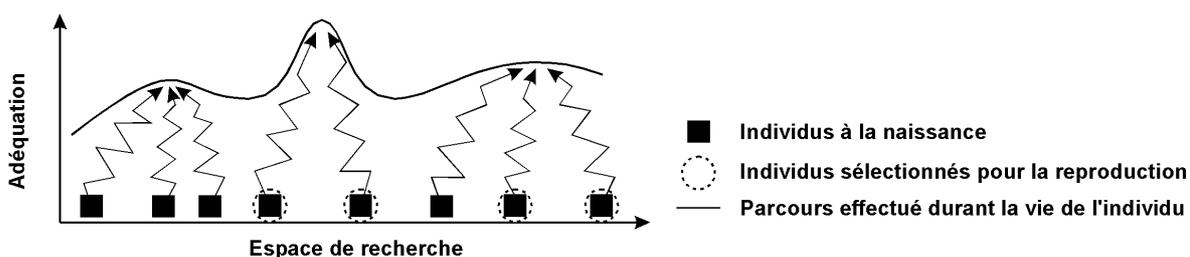


figure 2 - 15 : adaptation de la population à l'algorithme AGD

□ Description algorithmique

Les AGs produisent itérativement de meilleures estimées initiales pour les algorithmes de type grimpeur (vies). Ainsi d'une génération à l'autre les points initiaux des grimpeurs sont ceux situés à l'intérieur du bassin d'attraction de l'optimum global. Lors de l'évaluation de la qualité d'adéquation des parents, les AGs utilisent les résultats issus d'algorithmes grimpeurs initialisés par le phénotype de chaque parent. Puis, pendant les phases de sélection, recombinaison et mutation les algorithmes génétiques travaillent sur les points initiaux donnés aux grimpeurs et non pas le résultat de ceux-ci (cf. figure 2 - 15).

5.3 Principes Lamarckiens

□ Description génétique

Dans ce cas également, la sélection des individus dépend de leurs adéquations après leurs vies. Mais ici, toutes les connaissances acquises par l'individu durant son existence (l'acquis) est transmis lors de la reproduction. Le code génétique de l'individu change donc au cours de sa vie (voici un bel exemple de manipulation génétique !).

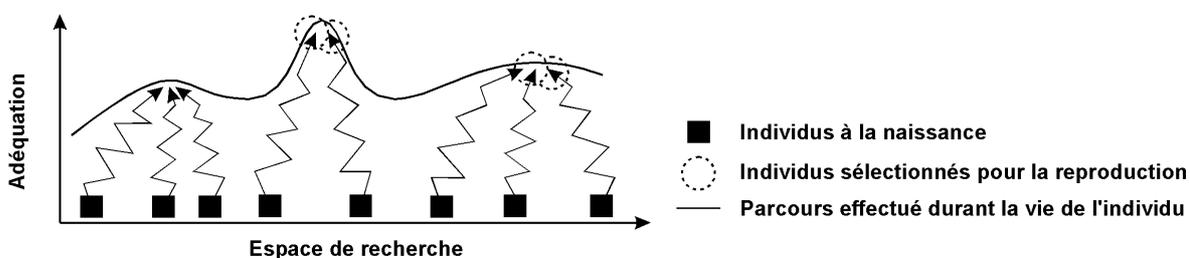


figure 2 - 16 : adaptation de la population l'algorithme AGL

□ Description algorithmique

Lors de l'évaluation de l'adéquation de chaque individu, les AGs utilisent les résultats issus de grimpeurs dont les points initiaux correspondent aux phénotypes des individus (cf. figure 2 - 16). Contrairement aux principes Darwiniens, durant les phases de reproduction, recombinaison et mutation, les AGs travaillent sur les points finaux de l'algorithme grimpeur.

5.4 Durée des “ vies grimpeurs ”

Avec les deux stratégies d'hybridation précédentes (selon Darwin et Lamarck), une itération longue des algorithmes de type grimpeur jusqu'à satisfaction complète des critères d'arrêt (développés paragraphe 2.2) n'est pas indispensable. En effet, la majorité de l'information (les plus grands déplacements vers l'optimum) est acquise lors des premières itérations. Ce fait se rencontre essentiellement avec une hybridation selon le schéma de Lamarck où la recherche est poursuivie les générations suivantes sur l'information acquise lors des phases grimpeurs.

L'hybridation selon Darwin nécessite quant à elle une durée de vie plus longue des individus durant leur vie « grimpeur », car les résultats des vies « grimpeurs » conditionnent significativement la sélection des individus potentiellement performants. Une erreur de sélection due à une imprécision de valeur d'adéquation peut avoir de fâcheuses conséquences (convergence prématurée). Par contre fait que les résultats des vies « grimpeurs » ne soient pas directement utilisés d'une génération à l'autre, offre des capacités d'exploration plus élevées qu'avec le schéma d'hybridation selon Lamarck.

Pour les deux hybridations précédentes, après satisfaction d'un critère d'arrêt, les meilleurs individus sont retenus comme points de départ d'un algorithme de recherche local (grimpeur), dans le but d'obtenir une solution finale plus précise.

Les algorithmes génétiques présentés jusqu'ici associent les AGs et les méthodes d'optimisation locales pour combiner leurs avantages. L'idée est maintenant d'opérer une hybridation selon une seconde voie en introduisant un opérateur génétique de recombinaison offrant une heuristique de type grimpeur : l'algorithme génétique à recombinaison simplex.

5.5 Algorithmes génétiques à recombinaison simplex

L'idée de recombinaison simplex¹³ a été employé avec succès entre autres par Lerman [LER96] et Jang [JAN97]. L'opérateur de recombinaison couplé à celui de sélection sont le moteur principal des algorithmes génétiques. Les AGs sont donc efficaces uniquement si le problème possède une structure sous-jacente (hypothèses de reconstructibilité par blocks). En l'absence de structure sous-jacente, l'opérateur de recombinaison est inefficace et inutile. En présence de structure sous-jacente, il faut que l'opérateur recombiner les éléments adéquats. Ce constat motive la création de la recombinaison simplex. Cet opérateur se base sur une heuristique proche de l'algorithme des polyèdres flexibles (ou simplex, présenté en 3.1.1).

Description sommaire

La population initiale est constituée de $\gamma(m+1)$ individus (m dimension de l'espace de recherche et γ nombre de sous-groupes de population). A chaque génération seuls γ enfants (un par sous-groupe de population) sont engendrés et réinsérés dans leur sous-groupe respectif (stratégie élitiste de réinsertion).

- **Sélection**

Parmi la population totale, $\gamma(m+1)$ individus sont sélectionnés avec une probabilité définie par une fonction d'adéquation basée sur le rang (non-linéaire). Les individus sélectionnés sont divisés en γ sous-groupes par tirage aléatoire uniforme.

- **Recombinaison**

Trois types de recombinaison sont utilisés et définis sur chaque sous-groupe ($(m+1)$ individus) : recombinaison simplex, continue et intermédiaire. Soient P_{rs} la probabilité de recombinaison simplex, P_{rc} la probabilité de recombinaison continue, et P_{ri} la probabilité de recombinaison intermédiaire.

§ La recombinaison **discrète** est obtenue en choisissant aléatoirement chaque gène de l'enfant parmi les gènes d'un des $(m+1)$ individus du sous-groupe.

§ La recombinaison **continue** consiste à générer l'enfant tel qu'il soit égal au barycentre des $(m+1)$ individus du sous-groupe dans l'espace paramétrique.

§ La recombinaison **simplex** est directement inspirée du pseudo-code (voir 3.1.1) : l'enfant de chaque sous-groupe est calculé à partir des formules de réflexion, expansion et contraction partielle. L'unique différence concerne la contraction totale où l'enfant (unique !) est engendré par :

$$\text{rejeton} = \frac{1}{2}(\underline{x}_{\min} + \underline{x}_{\max})$$

¹³ Ou croisement simplex pour un codage binaire des chromosomes

Quatre hybridations, séquentielles, selon les principes de l'évolution décrits par Darwin (AGD), par Lamarck (AGL) et un opérateur de recombinaison «d'heuristique grimpeur» (AGX) ont été implémentés sur la base d'article et d'actes de congrès. La question est maintenant de savoir dans quelles mesures ces méthodes sont concurrentes ou complémentaires. Aussi pourquoi ne pas combiner AGD et AGL avec ce dernier opérateur de recombinaison simplex pour aboutir à des algorithmes génétiques «2 hybrides» (AGDX et AGLX) ?

5.6 Algorithmes génétiques “ 2 hybrides ”

L'objectif de cet algorithme est de dépasser les performances des algorithmes AGD et AGL. En effet, AGL et AGD possèdent d'excellentes propriétés d'exploitation apportées par le grimpeur et d'exploration inhérentes aux AGs. L'ajout de la recombinaison simplex ne peut qu'augmenter les capacités d'exploration de ces algorithmes au niveau global [REN95] ; dans le principe, cela revient à initialiser une méthode des polyèdres flexibles sur les optima locaux de la fonction à extrémiser (combinaison de AGL et AGX).

Jusqu'à présent, dans ce chapitre, de nombreux algorithmes d'optimisation ont été développés. L'association judicieuse de ces algorithmes permet la résolution d'une large gamme de problèmes d'optimisation. Le paragraphe suivant résume succinctement les domaines d'application des différents algorithmes développés en soulignant leurs points forts et leurs points faibles.

5.7 Evaluations expérimentales

Les algorithmes d'optimisation globale développés ont tous pour base les AGs. Les différentes possibilités d'algorithmes génétiques sont :

- Codés binaire (sans intérêt).
- Code réel.
- Mono ou multipopulation avec différents schémas de migration.
- Avec des opérateurs de recombinaison classiques ou simplex.
- Hybridation séquentielle ou selon les théories de Darwin ou Lamarck.

Ces algorithmes génétiques peuvent être associés à toute une panoplie d'algorithmes de type grimpeur afin de résoudre une très large gamme de problème d'optimisation. Les algorithmes grimpeurs ayant retenu notre attention sont :

- Méthode des **polyèdres flexibles** (du simplex).
- Algorithme de **quasi-Newton**.
- Instruction «**pem**¹⁴» de « l'identification toolbox » de Matlab[®] [LJU87] : algorithme de type Gauss-Newton.
- Encapsulation logicielle de la « **mml3**¹⁵ toolbox » de Matlab[®] [MIL88] : algorithme de type **Levenberg-Marquardt**.

Afin de se découpler d'un choix arbitraire de structure de modèle, ces algorithmes sont évalués sur les fonctions de test présentées paragraphe 2.4. L'optimisation de la fonction de Griewank est réputée difficile. Le relief présenté par cette fonction est sans aucun doute plus ardu du point de vue de l'optimiseur que ceux posés par la majorité des problèmes pour l'identification. Ainsi l'algorithme résolvant le problème de l'optimisation de cette fonction sera à même de résoudre un grand nombre de problèmes d'optimisation posés par l'identification des paramètres physiques de systèmes. Les

¹⁴ Pem : **P**redict **E**rror **M**ethod configurable en erreur de sortie et erreur d'équation

¹⁵ Mml3 : **M**odified **M**aximum **L**ikelihood **E**stimator version **3**, configurable en erreur de sortie et erreur d'équation

fonctions de test supportant cette étude comparative sont les fonctions de De Jong et de Griewank en deux dimensions ($m=2$).

Les différents algorithmes considérés sont les grimpeurs de type quasi-Newton (**QN**) et simplex, les algorithmes génétiques codés réels (**AGR**), algorithmes génétiques codés réel à recombinaison simplex (**AGX**), les AGs hybrides selon Darwin (**AGD**) et Lamarck (**AGL**) et une combinaison double hybrides AGL et AGX (**AGLX**)¹⁶.

Les résultats sont reportés dans le tableau 2 - 2 et le tableau 2 - 3. Les algorithmes génétiques étant basés sur des règles de transition probabilistes, les résultats portés dans ces tableaux résultent de statistiques opérées sur 100 processus d'optimisation.

La dernière colonne de ces tableaux comporte le temps de calcul moyen nécessaire à l'obtention d'une solution. La colonne précédente «fiabilité » indique le pourcentage de solutions fiables obtenues. Une solution est jugée fiable si sa précision est inférieure ou égale à 10^{-1} . La précision des solutions fiables est évaluée pour les valeurs de $\delta = (10^{-8}, 10^{-6}, 10^{-3}, 10^{-1})^T$. Ces colonnes indiquent le pourcentage du nombre de solutions fiables pour les valeurs de δ correspondantes. Les deux colonnes précédentes affichent les statistiques (valeurs moyennes et écart types des solutions fiables). De manière similaire les deux premières colonnes indiquent les mêmes statistiques sur toutes les 100 solutions.

La paramétrisation des algorithmes est reportée en annexe 1. Le matériel informatique employé est brièvement décrit paragraphe 2.3 du présent chapitre.

☐ **Fonction de De Jong n°1** (figure 2 - 3)

L'intérêt d'une étude approfondie des performances des différents algorithmes sur cette fonction est relativement limité. Quelques points sont remarquables (cf. tableau 2 - 2).

- Les algorithmes de type grimpeur résolvent ce problème rapidement en moins de $4/10^0$ de seconde avec une précision de 10^{-8} et une fiabilité de 100%.
- AGR et AGX localisent la solution avec une fiabilité de 100% également. Cependant, dans 90% des cas la précision des solutions obtenues par AGR est limitée à 10^{-3} et ceci au bout d'un temps de calcul de près d'une minute. Notons également la nette réduction du temps de calculs et la très significative augmentation de la précision des solutions de AGX par rapport à AGR.

FONCTION DE DE JONG N°1														
Algorithmes	Toutes solutions				Solutions fiables, précision								Fiab (%)	Tps calcul (s)
	Moyenne		Ecart type		Moyenne		Ecart type		10^{-8}	10^{-6}	10^{-3}	10^{-1}		
	x_1	x_2	x_1	x_2	x_1	x_2	x_1	x_2						
QN	(1)	(1)	(0)	(0)	(1)	(1)	(0)	(0)	100	0	0	0	100	0.54
Simplex	1.10^{-8}	4.10^{-9}	(0)	(0)	1.10^{-8}	4.10^{-9}	(0)	(0)	100	0	0	0	100	0.83
AGR	-2.10^{-5}	2.10^{-5}	5.10^{-4}	5.10^{-4}	-2.10^{-5}	2.10^{-5}	5.10^{-4}	5.10^{-4}	0	0	90	10	100	60
AGX	4.10^{-10}	3.10^{-10}	-4.10^{-9}	3.10^{-9}	4.10^{-10}	3.10^{-10}	-4.10^{-9}	3.10^{-9}	99	1	0	0	100	35

tableau 2 - 2 : évaluation expérimentale sur la fonction de De Jong n°1

Cependant l'objet de ces quelques lignes est de montrer que le choix des algorithmes génétiques (relativement en vogue dans certains milieux ces dernières années) n'est pas toujours judicieux. Bien des problèmes peuvent être résolus plus efficacement via des méthodes classiques éprouvées.

¹⁶ AGD étant largement surpassé par AGL, une combinaison AGX-AGD est de ce fait dépassée par AGLX. Cette combinaison n'est donc pas reportée dans le tableau

(0) : les règles d'évolution des algorithmes QN et simplex étant déterministes, l'écart type autour des valeurs obtenues est nul.

(1) : la précision machine entraîne un arrondi du résultat à 0

□ **Fonction de Griewank n°8.**

Comparer des algorithmes de type grimpeur sur cette fonction est illusoire. Les solutions de ces algorithmes sont optimales localement, dans le voisinage du point initial. Les algorithmes d'optimisation globale résolvent ce problème de manière plus ou moins efficace. Les capacités de ces algorithmes à résoudre ce problème diffèrent grandement d'un algorithme à l'autre (cf. tableau 2 - 3).

FONCTION DE GRIEWANK N°8															
Algorithmes	Toutes solutions				Solutions fiables, précision								Fiab (%)	Tps calcul (s)	
	Moyenne		Ecart type		Moyenne		Ecart type		10 ⁻⁸	10 ⁻⁶	10 ⁻³	10 ⁻¹			
	x ₁	x ₂	x ₁	x ₂	x ₁	x ₂	x ₁	x ₂							
AGR	2.10 ⁻¹	9.10 ⁻²	3	2	8.10 ⁻⁵	6.10 ⁻³	3.10 ⁻²	4.10 ⁻²	0	0	0	100	36	277	
AGX	9.10 ⁻²	4.10 ⁻²	1	2	2.10 ⁻⁹	-10 ⁻¹¹	9.10 ⁻⁹	10 ⁻⁸	97.6	2.4	0	0	85	63	
AGD	QN	6.10 ⁻²	8.10 ⁻¹	4	4	-3.10 ⁻⁶	-2.10 ⁻⁵	10 ⁻⁵	2.10 ⁻⁵	22.2	51.8	25.9	0	27	346
	Splx	5.10 ⁻¹	9.10 ⁻¹	5	5	-8.10 ⁻⁶	6.10 ⁻⁶	7.10 ⁻⁵	10 ⁻⁴	18.7	56.3	25	0	32	260
AGL	QN	10 ⁻¹	3.10 ⁻¹	1	2	-7.10 ⁻⁷	-2.10 ⁻⁷	3.10 ⁻⁶	5.10 ⁻⁶	79.7	17.2	2.5	0	79	170
	Splx	3.10 ⁻⁵	8.10 ⁻⁷	8.10 ⁻⁵	10 ⁻⁴	3.10 ⁻⁶	8.10 ⁻⁷	8.10 ⁻⁵	10 ⁻⁴	93	4	3	0	100	116
AGLX	QN	2.10 ⁻¹	-4.10 ⁻¹	8.10 ⁻¹	1	-3.10 ⁻⁸	-4.10 ⁻⁸	2.10 ⁻⁷	4.10 ⁻⁸	97.8	0	2.2	0	93	97
	Splx	2.10 ⁻⁸	4.10 ⁻⁹	4.10 ⁻⁷	4.10 ⁻⁷	2.10 ⁻⁸	4.10 ⁻⁹	4.10 ⁻⁷	4.10 ⁻⁷	94	4	2	0	100	107

tableau 2 - 3 : évaluation expérimentale sur la fonction de Griewank n°8

Malgré leurs propriétés d'exploration, les algorithmes génétiques (AGR) se révèlent peu fiables (36% de solutions fiables) et requièrent un temps de calcul important. De plus, lorsque les solutions sont fiables, celles-ci sont peu précises (100% des solutions fiables sont précises à 10⁻¹).

L'opérateur de recombinaison simplex associé aux algorithmes génétiques (AGX) conduit à une sensible amélioration des résultats. En particulier, le taux de fiabilité passe à 85% et les solutions fiables sont dans la grande majorité des cas (97.6%) précises à 10⁻⁸. Le temps de calcul est également inférieur d'un facteur 4 par rapport à l'algorithme AGR. La précision accrue de AGX vis à vis de AGR se répercute par une nette amélioration des statistiques des solutions fiables (moyenne et écart type).

Les algorithmes résultant de l'hybridation des algorithmes génétiques avec un grimpeur (AGD) offrent théoriquement de meilleurs résultats en terme de fiabilité, mais requiert intrinsèquement un temps de calcul important pour converger. Le temps de calcul devenant très rapidement excessif, le nombre maximum de générations a été fixé à 30. Les taux de fiabilité demeurent faibles, ils valent respectivement 27 et 32%, avec les algorithmes grimpeurs de type quasi-Newton et simplex.

L'hybridation selon les principes Lamarckiens (AGL) améliore sensiblement ces résultats. Le fait d'appliquer les opérateurs de sélection, recombinaison et mutation sur les résultats d'algorithmes de type grimpeurs augmente la vitesse de convergence. Parallèlement à cette diminution du temps de calcul, le taux de fiabilité croît significativement (79% pour une hybridation avec quasi-Newton 100% avec simplex). Le taux de fiabilité supérieur de AGL associé au simplex (par rapport à quasi-Newton) est dû aux capacités d'exploration supérieures de simplex par rapport à quasi-Newton.

Les résultats prometteurs de AGX et AGL laissent supposer qu'une combinaison 2 hybrides (AGLX) conduise à un algorithme encore plus performant. Les taux de fiabilité d'AGLX sont particulièrement remarquables. L'algorithme AGLX hybridé avec le simplex conserve le taux de fiabilité de 100% de AGL hybridé avec ce même grimpeur tout en diminuant le temps de calcul. Lors d'une hybridation

avec le grimpeur quasi-Newton l'apport de l'opérateur de recombinaison simple est encore plus remarquable. En effet l'ajout de l'opérateur de recombinaison simple augmente le taux de fiabilité de 79% pour AGL à 93% pour AGLX, en diminuant le temps de calcul d'un facteur 1.75 (97s contre 170s). En terme de précision des solutions, AGLX surpasse les performances de AGX et AGL. Quel que soit l'algorithme grimpeur utilisé, les solutions fiables sont localisées avec une précision supérieure ou égale 10^{-8} dans au moins 94% des cas.

Cette brève étude comparative ne pourrait être représentative de tous les problèmes d'optimisation possibles. Son objectif est de montrer par le biais d'exemples qu'une judicieuse combinaison des différents algorithmes répandus dans la littérature permet l'élaboration d'algorithmes d'optimisation très efficaces et de complexité mise en œuvre modérée.

5.8 Des algorithmes polyvalents

Le tableau 2 - 4 est établi sur des considérations quantitatives et qualitatives, basées sur l'expérience acquise durant ces travaux. En effet, les performances des différents algorithmes furent évaluées sur diverses fonctions de test classiquement utilisées dans le domaine de l'optimisation : les fonctions de De Jong n°1, Schwefel n°7, Griewank n°8, Branins, Easom, Michalewicz n°12, Rosenbrock et Rastrigin n°6. Les performances des algorithmes spécifiquement dédiés à l'identification furent quant à elles évaluées sur de réels problèmes d'identification :

- Identification des paramètres physiques des différents modèles de véhicules présentés dans le chapitre quatre sur des données réelles et issus de simulation.
- Identification des paramètres physiques du circuit RC présenté dans le premier chapitre, à partir de données réelles et simulées.

TABLEAU COMPARATIF																
Fonction de coût Unimodale/Multimodale		Unimodale							Multimodale							
Fonction Différentiable/ Propriétés Solution(s)		Différentiab.		Solutions					Différentiab.		Solutions					
		OUI	NON	Précis	Fiabil	Robus	Simpli	Tps	OUI	NON	Précis	Fiabil	Robus	Simpli	Tps	
Grimpeurs	QN	1	0	+++	+++	+++	+++	+++	1	0	+++	--	--	+++	+++	
	Simplex	1	1	+++	+++	+++	+++	++	1	1	+++	--	--	+++	+++	
	Pem	1	SO	+++	+++	+++	+++	+++	1	SO	+++	--	--	+++	+++	
	Mmle3	1	SO	+++	+++	+++	--	+++	1	SO	+++	--	--	--	+++	
Algorithmes d'optimisation. globale	AGR	1	1	--	+++	+++	++	--	1	1	--	+	++	++	--	
	AGX	1	1	+	+++	+++	++	+	1	1	+	++	+++	++	+	
	AGD	QN	1	0	+++	+++	+++	++	--	1	0	+++	+	++	++	--
		Splx	1	1	+++	+++	+++	++	--	1	1	+++	+	++	++	--
		Pem	1	SO	+++	+++	+++	++	--	1	SO	+++	+	++	++	--
		Mmle3	1	SO	+++	+++	+++	--	--	1	SO	+++	+	++	--	--
	AGL	QN	1	0	+++	+++	+++	++	+	1	0	+++	++	++	++	+
		Splx	1	1	+++	+++	+++	++	+	1	1	+++	++	++	++	+
		Pem	1	SO	+++	+++	+++	++	+	1	SO	+++	++	++	++	+
		Mmle3	1	SO	+++	+++	+++	--	+	1	SO	+++	++	++	--	+
	AGLX	QN	1	0	+++	+++	+++	++	+	1	0	+++	+++	+++	++	+
		Splx	1	1	+++	+++	+++	++	+	1	1	+++	+++	+++	++	+
		Pem	1	SO	+++	+++	+++	++	+	1	SO	+++	+++	+++	++	+
		Mmle3	1	SO	+++	+++	+++	--	+	1	SO	+++	+++	+++	--	+

tableau 2 - 4 : tableau comparatif des algorithmes d'optimisation

Le système d'appréciation suivant est adopté :

□ Solutions

- Précision : précision de la solution.

- Fiabilité : fiabilité de la solution.
- Robustesse : robustesse de l'algorithme.
- Simplicité : simplicité d'utilisation.
- Temps : temps de calcul.
- -- : très mauvais.
- - : mauvais.
- + : bon.
- ++ : très bon.
- +++ : excellent.

☞ Adéquation surface de coût/algorithme d'optimisation

- SO : sans objet (impossibilité logicielle).
- 1 : numériquement possible, sans garantie de résultat.
- 0 : l'utilisation de cet algorithme posera un problème numérique.
(calcul de dérivées de fonctions discontinues par exemple).

☞ Algorithmes

- AGR : AG codé réel.
- AGX : AG codé réel à recombinaison simplex.
- AGD : AG codé réel, hybridation selon Darwin.
- AGL : AG codé réel, hybridation selon Lamarck.
- AGLX : AG codé réel, hybridation selon Lamarck, à croisement simplex.

☞ Remarques

- L'instruction « **pem** » de « l'identification toolbox » de Matlab[®] et la « **mmle3 toolbox** » permettent uniquement la résolution de problèmes d'identification (modèle de représentation ou modèle de connaissance sous forme de représentation d'état). De plus l'identification d'un modèle paramétrique selon un critère de type moindre module est impossible. Les autres algorithmes d'optimisation globale (AGL avec simplex ou AGLX avec simplex) présentés dans ce tableau sont capables de résoudre la grande majorité des critères de coût existant dans la littérature.
- L'algorithme **simplex** devient particulièrement lent lorsque la dimension de l'espace de recherche dépasse 8-10. Pour une dimension de l'espace de recherche supérieure il conviendra de le remplacer par un autre algorithme plus adapté (du type gradients conjugués, quasi-Newton avec une stratégie de type programmation séquentielle quadratique pour la mise à jour de la matrice Hessienne).
- La « **mmle3 toolbox** » est d'utilisation particulièrement délicate, conséquence d'un grand nombre de paramètres internes à régler et d'une implémentation logicielle qui gagnerait à s'inspirer des règles du génie logiciel. Toutes les variantes d'algorithmes basées sur l'utilisation de ce programme se voient affectées d'une note -- pour la simplicité d'utilisation.

☞ Discussion

Empiriquement, pour l'optimisation de fonctions multimodales différentiables ou non, l'identification des paramètres physiques de systèmes complexes (dimension de l'espace de recherche inférieure à 10, présentant des minima locaux) les algorithmes suivants sont particulièrement robustes, précis, fiables, simples d'utilisation et requièrent un temps de calcul raisonnable. Ils sont donc fortement conseillés (cellules grisées dans le tableau 2 - 4) :

- Algorithme génétique codé réel mono-population « 2 hybrides » (hybridation selon Lamarck : grimpeur = simplex, recombinaison simplex : $P_{rs} = 1/3$, $P_{rc} = 1/3$, $P_{ri} = 1/3$) avec stratégie élitiste de réinsertion.
- Si la fonction objectif est différentiable, le même algorithme est conseillé avec pour variante l'algorithme grimpeur quasi-Newton ; celui-ci entraîne une diminution du temps de calcul.

A la lecture de ce tableau, il est tentant d'utiliser systématiquement un algorithme de type AGLX. Avant d'utiliser cet algorithme, il est souhaitable de s'assurer qu'un simple algorithme grimpeur ne résous pas le problème posé en un temps de calcul nettement inférieur.

6 CONCLUSION

Avec l'objectif de se prémunir de la convergence de l'algorithme vers un optimum local, **une étude comparative des performances de différents algorithmes d'optimisation a été menée**. Il est montré que l'idée largement répandue que les algorithmes d'optimisation de type calcul évolutif rendent les algorithmes classiques de type grimpeurs obsolète est fausse.

En effet, pour l'optimisation de problèmes unimodaux les algorithmes classiques de type grimpeur se révèlent particulièrement performant en terme de précision et de temps de calcul. En revanche, pour l'optimisation de problèmes multimodaux, ces algorithmes n'offrent aucune garantie quant à la globalité de la solution.

Pour résoudre ce type de problème les algorithmes génétiques possèdent d'indiscutables avantages (capacités d'exploration) mais pèchent par un temps de convergence élevé une précision relativement médiocre. Une première hybridation des algorithmes génétiques par l'introduction d'un opérateur de recombinaison inspiré de la méthode des polyèdres flexibles améliore sensiblement les performances de l'algorithme génétique de base. Plus précisément, la fiabilité et la précision des solutions sont accrues.

Parmi les algorithmes hybrides développés, **les algorithmes génétiques hybridés selon Lamarck associé à l'opérateur génétique de recombinaison inspiré du simplex offrent des performances remarquables**. L'algorithme résultant allie les propriétés d'exploitation des algorithmes de type grimpeur aux capacités d'exploration des algorithmes génétiques, tout en requérant un temps de calcul raisonnable. Cet algorithme offre des prestations vraisemblablement surdimensionnées par rapport aux problèmes d'identification des paramètres physiques relatés dans ce mémoire. Cependant, la polyvalence et la puissance de cet algorithme permettent d'envisager avec sérénité l'aspect optimisation lors de l'identification des paramètres physiques de systèmes complexes. Par ailleurs, la fiabilité de cet algorithme permet dans une certaine mesure de garantir les résultats des études numériques de l'identifiabilité et de la discernabilité : la probabilité d'obtention d'un minimum local est faible.

Les valeurs des paramètres identifiés sont obtenues par minimisation des critères d'erreurs définis dans le premier chapitre de ce mémoire via les algorithmes présentés dans ce chapitre. Une phase de validation, faisant l'objet du chapitre suivant, permettra d'accepter ou de rejeter modèle et les valeurs des différents paramètres identifiés. Cette validation permettra d'effectuer un choix parmi les différentes structures de modèle en concurrence pour modéliser le système.

VALIDATION

1 INTRODUCTION

La minimisation des critères d'erreurs décrits au premier chapitre par le biais des algorithmes d'optimisation non-linéaire (cf. deuxième chapitre) vise l'obtention de la meilleure valeur possible des paramètres au sens du pouvoir prédictif du modèle choisi. Se pose alors la question de la suffisance du modèle ainsi obtenu vis à vis du but poursuivi de la modélisation. Ce problème revêt plusieurs aspects.

- La capacité du modèle à décrire les dynamiques mises en jeu dans le système : la **flexibilité** du modèle [LJU87], [SOD89].
- La **véracité** du modèle et des paramètres estimés par rapport au processus réel.
- Les **incertitudes** sur les paramètres estimés et sur les variables de sortie prédites.

Les réponses à ces questions doivent être modulées en fonction du but poursuivi par la démarche de modélisation. Dans une optique de conception de loi de commande, l'attention sera portée sur le pouvoir prédictif du modèle identifié et sur sa robustesse vis à vis de perturbations extérieures. Si l'objectif assigné est l'augmentation de la connaissance disponible sur un système, les valeurs des paramètres identifiés, les intervalles d'incertitude associés à ces paramètres et les phénomènes physiques mis en jeu sont également considérés dans la phase de validation. Dans ce contexte le problème de la validation de modèle est loin d'être résolu de manière satisfaisante. Notamment la vérification de la véracité des valeurs des paramètres estimés est particulièrement délicate.

Afin de prédire aussi précisément que possible les incertitudes sur les paramètres estimés, certaines hypothèses relatives aux caractéristiques du bruit de mesure doivent être vérifiées. Dans un premier temps, quelques éléments dédiés à la **validation des résidus seront donc présentés**.

Puis, dans un second temps, nous nous intéresserons aux notions de biais et de variance sur les paramètres estimés et aborderons le problème de l'estimation de ces erreurs de biais et de variance. L'accent sera essentiellement porté sur les différentes techniques disponibles pour l'estimation et l'interprétation de l'erreur de variance associée aux paramètres estimés. En premier lieu, nous donnerons une estimation statistique de celles-ci à partir des propriétés asymptotiques de l'estimateur du maximum de vraisemblance. Puis différentes interprétations géométriques possibles de cette estimation statistique seront abordées. Cette approche géométrique sera le support d'une

recherche visant l'isolation des paramètres entraînant une erreur de variance élevée sur les paramètres estimés.

Une estimation des intervalles d'incertitude (erreur de variance) sur les paramètres estimés nécessite une bonne connaissance de la nature des erreurs de prédiction. Or, lors de l'identification des paramètres physiques de systèmes complexes, la présence d'erreur de modélisation est quasi inévitable. Le processus d'estimation des incertitudes associées aux paramètres estimés doit donc tenir compte de la présence d'erreur de modélisation. Classiquement, la quantification de l'erreur de variance des paramètres estimés s'opère par le biais d'outils classiques de la théorie stochastique ; les résidus sont assimilés à une suite de variables aléatoires indépendantes. **La présence d'erreur de modélisation (erreur déterministe) fausse donc l'estimation des intervalles d'incertitude des paramètres estimés.**

Une méthode de quantification des intervalles d'incertitude en présence d'erreur de modélisation, utilisée pour l'identification des coefficients aérodynamiques d'aéronefs, sera utilisée puis validée par Monte-Carlo.

Enfin, pour conclure ce chapitre, la problématique du choix du « bon modèle » sera abordée. La question du choix du modèle, reçoit à ce jour que des réponses partielles. Cette problématique dépend de plusieurs facteurs lorsque plusieurs modèles entrent en concurrence pour décrire le même processus. Ces facteurs sont principalement :

- Le pouvoir prédictif des modèles.
- La véracité et la précision des paramètres identifiés des différents modèles.
- Les propriétés de discernabilité, d'identifiabilité et sensibilité des différents modèles.

En fonction des priorités du modéliste, ces différents points guident alors son choix.

2 INVALIDATION DES RESIDUS

Les différents tests sur les résidus visent, d'une part à vérifier la flexibilité ([LJU87], [SOD89]) du modèle identifié, et d'autre part, à valider les hypothèses formulées sur la nature du bruit de mesure présent sur les données. La flexibilité du modèle définit sa capacité à décrire les différentes dynamiques du système. La vérification des hypothèses relatives à la nature du bruit est importante. En effet, la satisfaction de ces hypothèses conditionne une bonne partie des propriétés de l'estimateur utilisé. Si le modèle identifié par minimisation d'un critère de coût du maximum de vraisemblance de l'erreur de sortie, est flexible, alors :

$$\tilde{y}(t, \hat{\theta}) = y(t) - \hat{y}(t, \hat{\theta}) \rightarrow \varepsilon(t) \quad (3.2-1)$$

L'invalidation des résidus a donc pour objet de vérifier les différentes hypothèses formulées sur $\varepsilon(t)$. Il s'agit de vérifier que la variable $\varepsilon(t)$ est bien assimilable à une variable aléatoire indépendante stationnaire, de moyenne nulle et de distribution normale. Pour ce faire, **l'espérance mathématique, la normalité, la stationnarité et l'indépendance** des résidus sont successivement vérifiées. Dans ce but, différents outils statistiques et graphiques ont été développés dans la littérature. Les ouvrages suivants traitent de manière approfondie ces différents aspects : [WAL97], [PRE86], [SOD89], [LJU87], [GOO77]. Les principaux algorithmes utilisés dans ce mémoire sont rappelés (sans démonstration) en annexe 2.

3 INCERTITUDES SUR LES PARAMETRES

L'obtention de la meilleure valeur des paramètres au sens du critère de coût minimisé est en général insuffisante. Il est aussi souhaitable d'évaluer les incertitudes associées au résultat. Les paramètres estimés sont affectés par deux types d'erreur : **l'erreur de biais et l'erreur de variance**.

3.1 Erreur de biais d'estimation

Le biais d'estimation se définit statistiquement par la présence d'une erreur systématique répétitive. Si $\hat{\theta}$ est un estimateur de θ_0 alors $\hat{\theta} = \operatorname{argmin} J(\theta_0, \varepsilon(t), w(t))$, avec $\varepsilon(t)$ le bruit de mesure, et $w(t)$ le bruit d'état dû à la présence d'erreur de modélisation :

$$\begin{aligned} \dot{x}(t) &= A x(t) + B u(t) + w(t) \\ y(t) &= C x(t) + D u(t) + \varepsilon(t) \end{aligned} \quad (3.3-1)$$

La présence du bruit d'état $w(t)$ est la conséquence du fait que le modèle est une approximation du système réel. Pour une entrée définie, le critère de coût J est fonction de la vraie valeur des paramètres, du bruit de mesure et du bruit d'état. Il est alors possible de définir une fonction de biais d'estimation $b(\cdot)$ telle que :

$$b(\theta_0) = E\{\hat{\theta} - \theta_0\} = E\{\operatorname{argmin} J(\theta_0, \varepsilon(t), w(t))\} - \theta_0 \quad (3.3-2)$$

Un estimateur non-biaisé est un estimateur pour lequel $b(\theta_0) = 0 \quad \forall \theta_0, \varepsilon(t)$ et $w(t)$.

Sous certaines hypothèses, l'estimateur du maximum de vraisemblance sur l'erreur de sortie est asymptotiquement non-biaisé. L'absence d'erreur de modélisation est une de ces hypothèses. Lorsque les paramètres estimés ont une signification physique, ou que des décisions dépendent de leurs valeurs (diagnostic par exemple), il convient d'évaluer, dans la mesure du possible, les erreurs induites par la présence d'erreurs structurelles.

A ce jour, l'évaluation du biais d'estimation est encore un problème non résolu. D'une part l'expression (3.3-2) est extrêmement complexe et nécessite la connaissance de la vraie valeur des paramètres. De plus pour des systèmes réels complexes, la présence d'un biais d'estimation est quasiment inévitable. Afin de le minimiser, il faut donc veiller à valider les différentes hypothèses relatives à la nature des résidus (paragraphe précédent) et vérifier les capacités du modèle à décrire le processus.

Pour toutes ces raisons, la présence d'une erreur de biais est difficilement détectable. De plus, une estimation biaisée des paramètres avec une faible variance sur les paramètres estimés est également possible. Ce biais peut aussi être la conséquence de la présence d'une erreur d'instrumentation, de calibration de capteur.... Dans les travaux relatés dans ce mémoire, le biais sera réduit au minimum en soignant les étapes de modélisation, d'instrumentation et de traitement du signal. Une fois les résidus validés, la présence d'un biais d'estimation est essentiellement évaluée à partir de la connaissance disponible a priori sur le processus.

3.2 Erreur de variance

La quantification de la variance vise à associer aux valeurs des paramètres identifiées (valeurs biaisées ou non) un intervalle de confiance. En répétant donc un certain nombre de fois la même expérience, il est possible de juger de la dispersion des valeurs autour de la moyenne des estimées.

Par exemple reprenons le circuit RC (de fonction de transfert $G_{i/u}(s) = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$) décrit

dans le premier chapitre où l'identification de la capacité est menée. La valeur de cette capacité est identifiée par minimisation d'un critère d'erreur de sortie au sens du maximum de vraisemblance, pour différents jeux de données. Sur la figure 3 - 1 sont représentés les résultats du processus d'identification sur 35 jeux de données différents (jeux de données simulés sur 35 réalisations différentes de bruit de mesure). Sous forme de points sont présentées les valeurs de la capacité identifiée, sous forme de carrés la moyenne des 35 estimations, et sous forme de triangle l'intervalle de confiance à 95 % calculé sur ces 35 points.

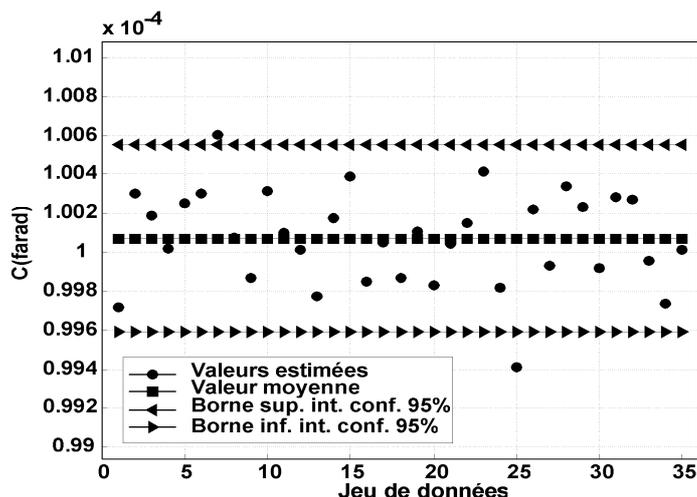


figure 3 - 1 : intervalles de confiance expérimentaux

Cette manière de procéder a l'avantage de tenir compte de tous les facteurs possibles (erreur et condition de mesure, perturbations de l'environnement ...). Les techniques reposant sur l'exploitation de l'inégalité de Cramer-Rao tentent de prédire cette variance dans un contexte théorique stochastique et ne peuvent généralement rendre compte de tous les facteurs influents. Le principal inconvénient de cette méthode est de nécessiter un très grand nombre de jeu de données pour être significative. Un second inconvénient est de ne pas disposer des couplages existant entre les différents paramètres. Cette méthode sera donc avantageusement complétée en évaluant la matrice de variance/covariance des paramètres, qui dans le cas d'un estimateur non-biaisé s'exprime

$$P_{\hat{\theta}} = E \{ (\hat{\theta} - \theta_0)(\hat{\theta} - \theta_0)^T \} \quad (3.3-3)$$

Une interprétation géométrique de ces intervalles de confiance sous forme d'ellipse d'incertitude dans l'espace paramétrique offre également d'intéressantes perspectives. Par ailleurs lorsque le nombre de jeux de données est limité, une évaluation statistique expérimentale par Monte-Carlo est possible en simulant des données fictives par les techniques du bootstrap ou du jackknife par exemple [WAL97].

3.2.1 Estimation statistique théorique : inégalité de Cramer-Rao

Lors de l'estimation des paramètres, il semble naturel de vouloir tendre vers un estimateur offrant la plus faible variance sur les estimées. Or la plus faible valeur possible pour la matrice de covariance sur les paramètres est bornée par l'inégalité de Cramer-Rao [FIS12]. L'inégalité de Cramer-Rao, indépendante de l'estimateur utilisé, fixe une limite théorique sur la variance des paramètres estimés.

3.2.1.1 Inégalité de Cramer-Rao

Pour tout estimateur non-biaisé, la matrice de covariance est bornée par l'inégalité de Cramer-Rao :

$$P_{\hat{\theta}} = \text{var}(\hat{\theta}) \geq M_F^{-1}(\theta_0) \quad (3.3-4)$$

$$P_{\hat{\theta}} = \begin{bmatrix} \sigma_{(1,1)} & \sigma_{(1,2)} & \cdots & \sigma_{(1,m)} \\ \sigma_{(1,2)} & \sigma_{(2,2)} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{(m,1)} & \sigma_{(m,2)} & \cdots & \sigma_{(m,m)} \end{bmatrix}$$

où $M_F^{-1}(\theta_0)$ est l'inverse de la matrice d'information de Fisher.

L'inégalité de Cramer-Rao permet donc pour des conditions expérimentales définies d'évaluer la variance minimale que l'on est en droit d'attendre. Notons que cette estimation est possible à partir d'un **unique** jeu de données. Si les intervalles de confiance désirés sont inférieurs à cette limite, le processus d'identification doit être partiellement ou en totalité remis en cause¹⁷.

La matrice d'information de Fisher [FIS12] s'exprime

$$M_F(\theta_0) = E \left\{ \left(\frac{\partial \ln P(y | \theta, \Lambda)}{\partial \theta} \right) \left(\frac{\partial \ln P(y | \theta, \Lambda)}{\partial \theta} \right)^T \right\} \quad (3.3-5)$$

Un estimateur dont la variance égale l'inverse de la matrice d'information de Fisher est dit efficace $P_{\hat{\theta}} = M_F^{-1}(\theta_0)$. Or, l'estimateur du maximum de vraisemblance est asymptotiquement normal, non-biaisé et efficace (cf. premier chapitre). Le vecteur $\hat{\theta}_{mv}$ tend donc à être asymptotiquement distribué suivant une loi $N(\theta_0, M_F^{-1}(\theta_0))$. Détaillons maintenant l'expression (3.3-5) dans le cas d'une distribution des résidus $N(0, \Lambda)$, d'un modèle non-LP et pour la vraie valeur des paramètres θ_0 . La fonction de la log-vraisemblance s'exprime alors

$$\ln(P(Y | \theta_0, \Lambda)) = -\frac{n_y N}{2} \ln(2\pi) - \frac{N}{2} \ln(\det(\Lambda)) - \frac{1}{2} \sum_{t=1}^N \tilde{y}(t, \theta_0)^T \Lambda^{-1} \tilde{y}(t, \theta_0) \quad (3.3-6)$$

où $\tilde{y}(t, \theta_0)$ désigne l'erreur de prédiction définie par l'expression (1.3-1)

En injectant (3.3-6) dans (3.3-5), il vient

$$M_F(\theta_0) = \sum_{t=1}^N \left. \frac{\partial y(t, \theta)}{\partial \theta} \right|_{\theta=\theta_0} \Lambda^{-1} \left. \frac{\partial y(t, \theta)}{\partial \theta^T} \right|_{\theta=\theta_0} \quad (3.3-7)$$

Or, la valeur de θ_0 n'est pas connue. On suppose alors que la valeur $\hat{\theta}$ est suffisamment proche de θ_0 pour pouvoir assimiler $M_F(\hat{\theta})$ à $M_F(\theta_0)$. Comme le calcul (3.3-7) est peu sensible aux variations de θ_0 [LJU87], l'approximation $M_F(\theta_0) = M_F(\hat{\theta}_{mv})$ demeure tout à fait acceptable. De plus la matrice de covariance du bruit de mesure est généralement inconnue. Cependant, une fois les paramètres identifiés, elle est estimée par

$$\hat{\Lambda}_{mv} = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t, \hat{\theta}_{mv})) (y(t) - \hat{y}(t, \hat{\theta}_{mv}))^T \quad (3.3-8)$$

¹⁷ Par exemple par réduction du nombre de paramètres identifiés ou par planification d'expérience optimale afin de réduire les intervalles d'incertitude associés aux paramètres.

La valeur obtenue pour la matrice de covariance des paramètres est d'autant plus exacte que le nombre de données est élevé, que le modèle est LP et que les résidus sont assimilables à une suite de variables aléatoires indépendantes de distribution normale. La matrice $P_{\underline{\theta}}$ estime la variance des paramètres estimés. Classiquement les intervalles de confiance à 95 % sont considérés.

3.2.1.2 Evaluation stochastique

Reprenons le circuit RC décrit chapitre un, (de fonction de transfert $G_{i/u}(s) = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$).

Identifions alors le vecteur de paramètres $\underline{\theta}_0 = (R, C)$, au sens du maximum de vraisemblance, par minimisation d'un critère d'erreur de sortie. L'estimateur obtenu est théoriquement non-biaisé et efficace. A partir d'un unique jeu de données, il est donc possible de prédire la variance sur les paramètres estimés grâce à l'expression (3.3-4). Vérifions maintenant expérimentalement (en utilisant la méthode de Monte-Carlo) que cette variance estimée est bien représentative de la variance expérimentale. Cette dernière est calculée sur plusieurs estimations issues de jeux de données différents (100 jeux de données avec 100 réalisations du bruit de mesure différentes¹⁸).

$\underline{\theta}$	$\underline{\theta}_0$	$\hat{\underline{\theta}}_{mv}$	Intervalles de confiance à 95 % estimés par Cramer-Rao (%)	Intervalles de confiance à 95 % estimés expérimentalement (%)
R (Ω)	470	469.96	0.31	0.32
C (μ F)	100	100.06	0.52	0.54

tableau 3 - 1 : paramètres estimés et intervalles d'incertitude en présence de bruit blanc

Le tableau 3 - 1 présente les résultats obtenus : dans la quatrième colonne sont portés les intervalles de confiance à 95 % estimés à partir des bornes de Cramer-Rao d'un unique jeu de données ; la cinquième colonne contient les intervalles de confiance à 95 %, estimés à partir de l'écart type calculé sur les 100 estimations. Ces intervalles de confiance sont exprimés en % de la valeur nominale des paramètres ($\underline{\theta}_0$).

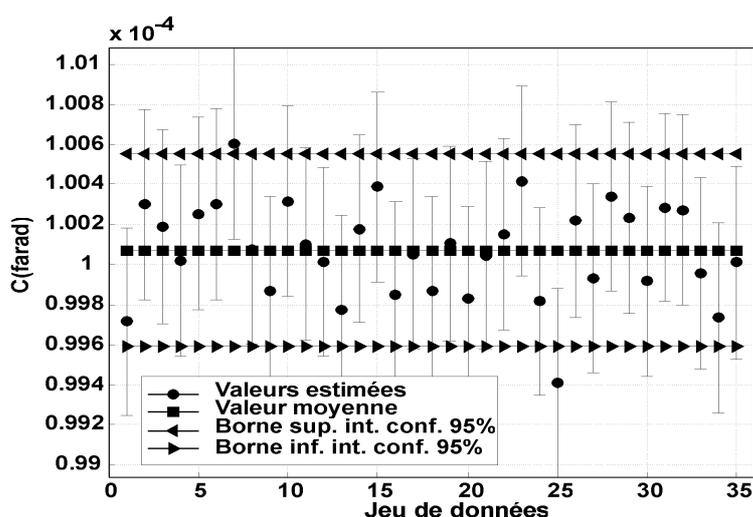


figure 3 - 2 : intervalles de confiance et bornes de Cramer-Rao

¹⁸ Conditions de simulation : le système est excité par un signal de type sbpa : période d'échantillonnage = 2ms, temps élémentaire (durée d'un bit) = 14 ms, nombre de bit = 5, amplitude = $\pm 1V$, nombre de périodes = 10. Un bruit de mesure $N(0, 10^{-8})$ est ajouté sur la variable de sortie.

Ces valeurs montrent l'excellente prédiction des intervalles d'incertitude sur les données estimées à partir de la seule inégalité de Cramer-Rao lorsque toutes les hypothèses relatives à la nature du bruit de mesure sont remplies. Ces résultats peuvent être présentés sous une forme graphique.

Ainsi sur la figure 3 - 2 sont tracés les 35 premières occurrences de la valeur de la capacité identifiée, la valeur moyenne sur ces 35 valeurs et les intervalles de confiance à 95% calculés sur ces mêmes valeurs. Pour chaque valeur estimée (donc chaque jeu de données), sont représentés sous forme de barres verticales, les intervalles de confiance estimés à partir de l'inégalité de Cramer-Rao. La figure 3 - 2 montre très clairement que l'amplitude des intervalles de confiance estimés à partir de chaque jeu de donnée est représentative des intervalles de confiance réellement constatés sur l'ensemble des valeurs estimées.

Dans le paragraphe suivant, l'interprétation géométrique des intervalles d'incertitude sous forme d'isocritères est étudiée. Nous verrons comment ces isocritères définissent la topologie de l'espace de recherche auquel les algorithmes d'optimisation sont confrontés.

3.2.2 Représentation géométrique de l'erreur de variance : isocritères

Les isocritères, ou ellipses d'incertitude, sont des approximations des intervalles de confiance calculés à partir de l'inégalité de Cramer-Rao. Cette représentation géométrique des intervalles de confiance, sous forme d'ellipses d'incertitude, s'explique soit par une approche statistique, soit par une approche heuristique.

3.2.2.1 Approche statistique

Les intervalles de confiance se définissent à partir de la fonction de densité de probabilité des paramètres estimés. Comme nous l'avons déjà mentionné, le vecteur des paramètres estimés au sens du maximum de vraisemblance, suit une distribution $N(\underline{\theta}_0, P_{\underline{\theta}})$. Sa fonction de vraisemblance s'exprime donc :

$$P(\underline{\theta}) = \prod_{t=1}^N \frac{1}{((2\pi)^m \det P_{\underline{\theta}})^{1/2}} \exp\left(-\frac{1}{2}(\underline{\theta} - \underline{\theta}_0)^T P_{\underline{\theta}}^{-1} (\underline{\theta} - \underline{\theta}_0)\right) \quad (3.3-9)$$

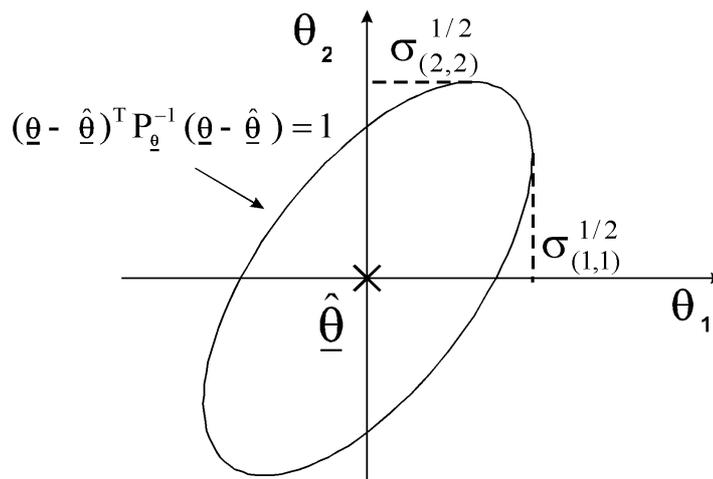


figure 3 - 3 : représentation des ellipses d'incertitudes

Les paramètres estimés d'iso-vraisemblance se situeront donc, dans l'espace paramétrique, sur des courbes isocritères $P(\underline{\theta})$:

$$P(\underline{\theta}) = \text{constante} \quad (3.3-10)$$

Pour une distribution normale des estimées, ces surfaces se concrétisent sous la forme d'ellipses d'incertitude. En exprimant la fonction de log-vraisemblance de (3.3-9), il vient :

$$(\underline{\theta} - \underline{\theta}_0)^T \mathbf{P}_{\underline{\theta}}^{-1} (\underline{\theta} - \underline{\theta}_0) = \text{constante} \quad (3.3-11)$$

La constante dans l'expression précédente est communément choisie égale à 1. Elle définit ainsi une région dans l'espace paramétrique, autour de $\underline{\theta}_0$, où la probabilité de trouver $\hat{\underline{\theta}}$ est égale à 68.3%. Or pour l'estimation des paramètres, $\underline{\theta}_0$ n'est en général pas connu. Par conséquent, ces ellipses ne peuvent pas être représentées telles quelles. En substituant $(\underline{\theta} - \hat{\underline{\theta}})$ à la quantité $(\underline{\theta} - \underline{\theta}_0)$, il est possible de centrer les ellipses d'incertitudes sur $\hat{\underline{\theta}}$. On fait alors la supposition que $\hat{\underline{\theta}}$ est suffisamment proche de $\underline{\theta}_0$ pour pouvoir assimiler $\underline{\theta}_0$ à $\hat{\underline{\theta}}$. La faible sensibilité de l'expression (3.3-9) vis à vis de la valeur $\underline{\theta}_0$ permet cette approximation [LJU87]. Il est alors possible de définir les intervalles d'incertitude autour des paramètres estimés $\hat{\underline{\theta}}$.

Ce jeu d'écriture permet de définir la variance des paramètres estimés autour de $\hat{\underline{\theta}}$ que l'on obtiendrait autour de $\underline{\theta}_0$. Précédemment (cf. paragraphe 3.2.1.2) il a été montré, en simulation, que la variance estimée à partir d'un unique jeu de données est bien représentative de celle estimée sur un grand nombre de jeux de données. L'équation des ellipses d'incertitude est donnée par l'expression (3.3-11). Sur la figure 3 - 3 est tracée une représentation graphique possible (2 paramètres) de ces ellipses d'incertitude.

3.2.2.2 Approche heuristique

Les paramètres estimés au sens du maximum de vraisemblance, s'obtiennent par minimisation d'un critère de coût quadratique multivariable. Dans le cas de modèle non-LP, au voisinage de l'optimum global, ce minimum est quadratique. Au point $\hat{\underline{\theta}}$, la valeur minimale du critère de coût vaut :

$$\mathbf{J}_{\text{mv}}(\hat{\underline{\theta}}_{\text{mv}}) = \sum_{t=1}^N \tilde{\mathbf{y}}(t, \hat{\underline{\theta}}_{\text{mv}})^T \hat{\Lambda}_{\text{mv}}^{-1} \tilde{\mathbf{y}}(t, \hat{\underline{\theta}}_{\text{mv}}) \quad (3.3-12)$$

Pour chaque réalisation différente du bruit de mesure, une estimation différente de $\hat{\underline{\theta}}$ est obtenue. Or $\hat{\underline{\theta}}$ tend à suivre une distribution $\mathbf{N}(\underline{\theta}_0, \mathbf{P}_{\underline{\theta}})$. La répétition un grand nombre de fois du processus d'identification conduit donc à un nuage elliptique de points centré sur $\underline{\theta}_0$.

Développons l'expression du critère de coût autour du point $\underline{\theta}_0$.

$$\mathbf{J}(\underline{\theta}) = \mathbf{J}(\underline{\theta}_0) + \frac{1}{2} (\underline{\theta} - \underline{\theta}_0)^T \left. \frac{\partial^2 \mathbf{J}}{\partial^2 \underline{\theta}} \right|_{\underline{\theta}=\hat{\underline{\theta}}} (\underline{\theta} - \underline{\theta}_0) \quad (3.3-13)$$

Comme autour de l'optimum global, le minimum est quadratique, les autres termes d'ordre supérieurs sont théoriquement négligeables. Il est maintenant possible de définir une amplitude pour $\mathbf{J}_{\text{mv}}(\underline{\theta})$ telle que $\mathbf{J}_{\text{mv}}(\underline{\theta}) > \mathbf{J}_{\text{mv}}(\hat{\underline{\theta}}_{\text{mv}})$, telle qu'un certain pourcentage des points estimés ($\hat{\underline{\theta}}$) se situe sous la surface décrite par l'ellipse d'équation (3.3-13). En choisissant $\Delta \mathbf{J}$ tel que $\mathbf{J}(\underline{\theta}) = \mathbf{J}_{\text{mv}}(\underline{\theta}_0) + \Delta \mathbf{J}$ et $\Delta \mathbf{J} = 0.5$, une ellipse, centrée sur $\underline{\theta}_0$ dans l'espace paramétrique contenant 68.3% des points estimés (cf. figure 3 - 4) est définie.

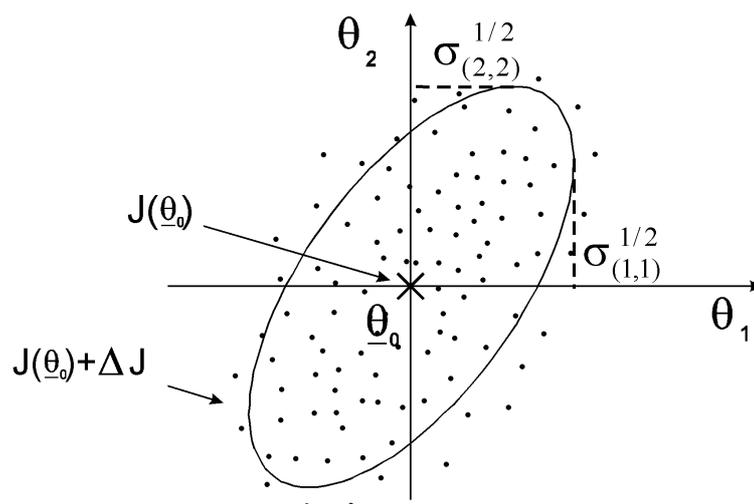


figure 3 - 4 : approche heuristique des ellipses d'incertitude

Les deux approches statistique et géométrique fournissent une image précise de la variance sur les paramètres estimés. L'exploitation de toute l'information contenue dans ces ellipses, à partir de leur seule représentation géométrique est délicate. En effet, tant que le nombre de paramètres recherchés est inférieur à 3, il est possible de donner une interprétation géométrique de ces ellipses d'incertitude. Pour un nombre supérieur, plusieurs solutions sont envisageables :

- La première consiste à projeter le nuage de points, sur des sous-espaces paramétriques de dimension 2.
- La seconde est de déterminer l'équation du nuage de points assimilé à une ellipse par le biais des expressions (3.3-11) ou (3.3-13).
- La troisième méthode consiste à analyser le nuage de points par le biais de méthodes d'analyse en composante spectrale (PCA) permettant d'obtenir différents indicateurs, relatifs aux corrélations entre les différents paramètres.
- Une autre manière de procéder [RIC91], est de parcourir itérativement les sous-espaces de dimension 2 par pas d'angle $\Delta\alpha$ (cf. figure 3 - 5) autour de l'optimum $\hat{\theta}$. La méthode consiste alors à rechercher les points autour du minimum entraînant une certaine augmentation du critère de coût.

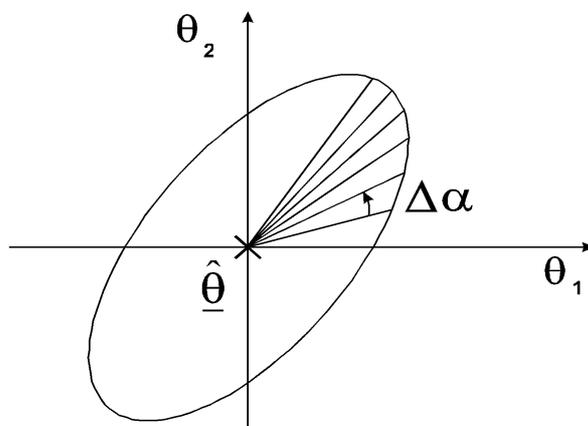


figure 3 - 5 : incertitude, méthode de « Richalet »

- Enfin la dernière, est basée sur l'étude de l'équation des ellipses d'incertitude définies par l'expression (3.3-11), [MIL88] et notamment sur l'exploitation des informations contenues dans la matrice P_{θ} .

Dans les quatre premiers cas, une approche globale du problème est impossible et l'analyse demeure limitée à deux dimensions. Une extrapolation des analyses réalisées en deux dimensions aux espaces de dimension supérieure est trompeuse et peut conduire à des conclusions hâtives et erronées. Pour être significatives, les méthodes reposant sur l'exploitation d'un nuage de points requièrent un grand nombre de jeux de données. Cette contrainte est parfois difficile à satisfaire, lorsque le protocole expérimental est lourd de mise en œuvre.

Pour ces raisons, la dernière méthode a été préférée. Elle exploite les propriétés géométriques des ellipses d'incertitude et s'applique à une dimension quelconque de l'espace de recherche.

3.2.3 Relations géométriques

Afin de pouvoir illustrer graphiquement nos propos, nous limiterons **sans aucune perte de généralité**, la dimension de l'espace de recherche à 2.

3.2.3.1 Sensibilité/insensibilité

La sensibilité est à rapprocher de l'analyse de la sensibilité décrite dans le premier chapitre. La sensibilité quantifie l'influence d'un paramètre sur la sortie du modèle, ou sur le critère d'erreur $J(\cdot)$.

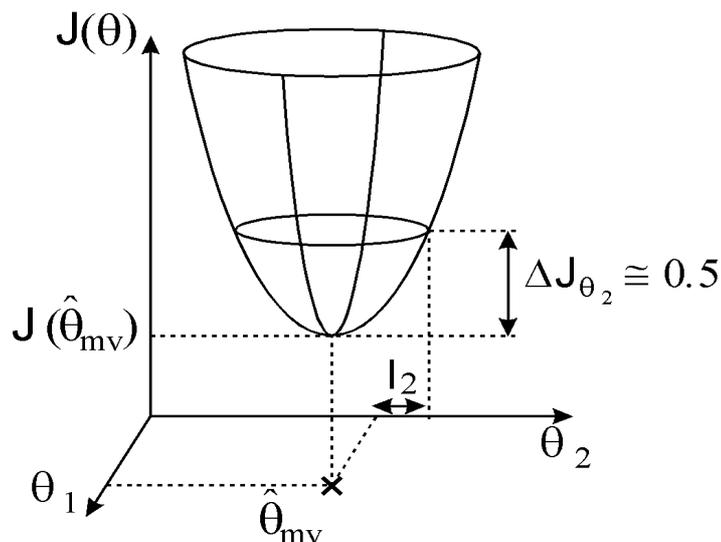


figure 3 - 6 : vérification de la globalité de la solution obtenue

En adoptant le terme d'insensibilité défini par Iliff [IL84], on définit l'amplitude d'une perturbation à apporter à une composante du vecteur des paramètres estimé $\hat{\theta}$, provoquant une augmentation du critère de coût $J(\theta) = J_{mv}(\hat{\theta}_{mv}) + \Delta J$ de $\Delta J = 0.5$. La figure 3 - 6 montre un exemple illustratif avec une dimension de l'espace de recherche limitée à 2). L'amplitude de cette perturbation se déduit aisément de (3.3-11)

$$I_i = P_{\hat{\theta}(i,i)}^{-1/2} = M_{F(i,i)}^{1/2} \quad (3.3-14)$$

En modifiant donc la $i^{\text{ème}}$ composante du vecteur des paramètres estimés d'une quantité égale à I_i , il est possible de vérifier que la solution estimée se situe bien au fond d'un minimum quadratique (minimum global). Dans le cas contraire, il est probable que l'algorithme n'ait pas convergé vers

l'optimum global. Notons ici, que l'algorithme d'optimisation implémenté dans le code de l'instruction « pem » de la « toolbox identification » du logiciel Matlab[®] v4.2 ne satisfait jamais ce test. Il semble que ce problème a été en partie résolu à partir de la version 5.1 de Matlab[®].

Du point de vue statistique, l'insensibilité représente l'écart type sur l'estimation d'un paramètre, en supposant les valeurs des autres paramètres connues. Lorsque plusieurs paramètres sont simultanément estimés, l'insensibilité borne la variance minimale sur les différents paramètres. En effet, les corrélations, entre les différents paramètres induisent des écarts types d'amplitudes aux moins égales à l'insensibilité de chacun des paramètres.

Une valeur d'insensibilité élevée révèle une direction de l'espace paramétrique insensibilisée le long d'un axe de l'espace paramétrique : conséquence possible d'un problème d'identifiabilité. L'inaptitude de l'insensibilité à quantifier les intervalles d'incertitude lors de l'estimation simultanée de différents paramètres conduit à considérer les corrélations existantes entre les différents paramètres.

3.2.3.2 Corrélation

La corrélation entre les paramètres i et j du vecteur de paramètres estimés $\hat{\theta}$ se définit classiquement à partir des éléments non-diagonaux de la matrice de variance/covariance des paramètres.

$$\rho_{(i,j)} = \frac{\sigma_{(i,j)}}{\sigma_{(i,i)}^{1/2} \sigma_{(j,j)}^{1/2}} \quad -1 < \rho_{(i,j)} < 1 \quad (3.3-15)$$

Ce coefficient n'apporte aucune information quant à la taille des ellipses d'incertitude. En effet, des paramètres fortement corrélés peuvent être estimés précisément (petite ellipse d'incertitude) et inversement, deux paramètres parfaitement non corrélés pourront présenter une précision médiocre (cercle de grande dimension).

Par ailleurs, lors de l'identification d'un nombre de paramètres supérieur à deux, ce coefficient de corrélation est incapable de détecter le fait que plus de 2 paramètres sont corrélés.

3.2.3.3 Bornes de Cramer-Rao

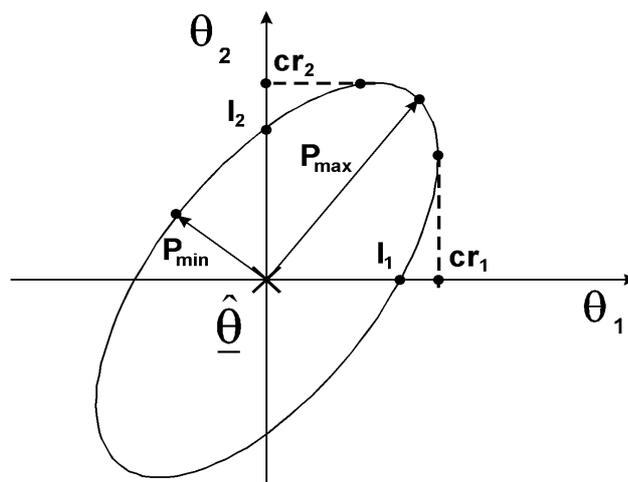


figure 3 - 7 : incertitude et bornes de Cramer-Rao

Par définition, les intervalles d'incertitude, bornés par l'inégalité de Cramer-Rao, s'expriment à partir de la racine carrée des éléments non-diagonaux de la matrice de variance/covariance des paramètres.

$$cr_i = \sigma_{(i,i)}^{1/2} = M_{(i,i)}^{-1/2} \quad (3.3-16)$$

Contrairement à la mesure de l'insensibilité, les bornes de Cramer-Rao tiennent compte du fait de l'estimation simultanée de plusieurs paramètres.

3.2.3.4 Axes principaux

Afin que l'algorithme d'optimisation puisse localiser l'optimum du critère de coût, il est important que la surface engendrée par ce coût dans l'espace paramétrique soit bien conditionnée [RIC91]. Les caractéristiques de cette surface au voisinage du minimum dépendent des fonctions de sensibilité. Si les différentes fonctions de sensibilités sont d'amplitudes voisines, les variations de la surface de coût dans le voisinage du minimum sont du même ordre de grandeur dans les différentes directions de l'espace de recherche : les ellipses d'incertitudes seront alors approximativement des sphères. Le conditionnement est alors qualifié de sphérique et idéal pour la recherche de l'optimum. Une surface mal conditionnée (longue vallée encaissée) peut traduire un défaut d'identifiabilité, ou un protocole expérimental peu sensibilisant pour un paramètre (cf. figure 3 - 8).

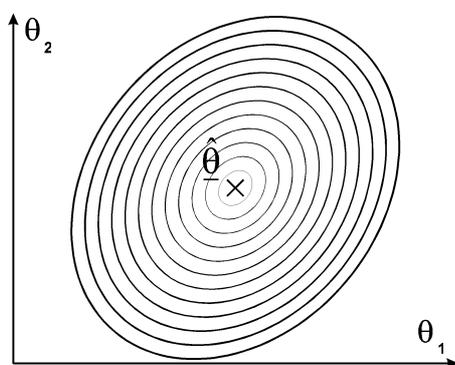


figure 3 - 8: conditionnement sphérique

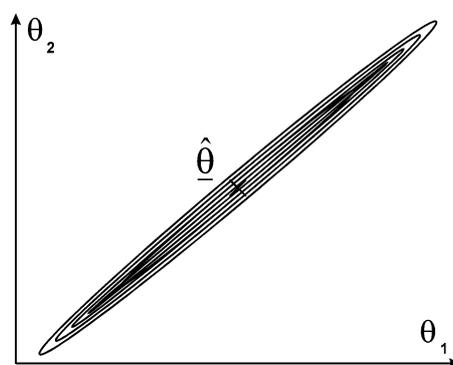


figure 3 - 8 : mauvais conditionnement

figure 3 - 8 : conditionnement du problème d'optimisation

Le calcul des axes principaux donne une indication sur le conditionnement du problème. Pour cela, les valeurs propres et les vecteurs propres de la matrice d'information de Fisher sont calculés. Soit \underline{V}_{\max} le vecteur propre associé à la valeur propre la plus élevée (λ_{\max}), et \underline{V}_{\min} le vecteur propre associé à la valeur propre la plus faible (λ_{\min}). Les axes principaux de l'ellipse d'incertitudes \underline{P}_{\min} et \underline{P}_{\max} (cf. figure 3 - 7) s'expriment alors :

$$\underline{P}_{\max} = \frac{\underline{V}_{\min}}{\lambda_{\min}^{1/2}} \quad \text{et} \quad \underline{P}_{\min} = \frac{\underline{V}_{\max}}{\lambda_{\max}^{1/2}} \quad (3.3-17)$$

Le rapport des normes des axes principaux donne alors une indication du conditionnement du problème.

$$\frac{\|\underline{P}_{\max}\|}{\|\underline{P}_{\min}\|} = \frac{\lambda_{\max}^{1/2}}{\lambda_{\min}^{1/2}} \quad (3.3-18)$$

Plus ce rapport est proche de un, meilleur est le conditionnement du problème.

3.2.3.5 Exploitation des isocritères

Lorsque les intervalles d'incertitude des paramètres estimés présentent des amplitudes élevées, il est intéressant d'isoler les causes de ces erreurs de variances élevées. Dans un premier temps, il

convient de vérifier si l'amplitude des insensibilités sur les paramètres n'est pas supérieure aux intervalles d'incertitude désirés sur les paramètres.

□ Une valeur d'insensibilité trop élevée sur un paramètre signifie que la réponse du modèle n'est pas sensible à la valeur prise par ce paramètre. Dans ce cas deux possibilités s'offrent au modélisateur :

- **Le paramètre a une influence négligeable** et ne peut être identifié.
- **Le paramètre n'a pas été sensibilisé** par le protocole expérimental initialement défini. Une modification du protocole expérimental par le biais de techniques de planification d'expériences optimales permet dans certains cas de sensibiliser la réponse du système vis à vis du paramètre recherché.

□ Si la valeur de l'insensibilité est raisonnable, un intervalle de confiance d'amplitude élevée peut être la conséquence de la corrélation de deux paramètres. Ces corrélations sont aisément mises en évidence par le biais du coefficient de corrélation (3.3-15). Sous forme matricielle, ces coefficients se calculent aisément en normalisant la matrice de variance/covariance des paramètres par les insensibilités des paramètres [MIL88].

$$P_{\theta \text{ nor}} = T P_{\theta} T \quad (3.3-19)$$

$$\text{avec } T = \text{diag}(I_1, \dots, I_m) \quad (3.3-20)$$

Sous cette forme, tous les paramètres présentent une variance unité. Deux paramètres corrélés induiront donc des coefficients non-diagonaux de la matrice $P_{\theta \text{ nor}}$ de valeurs non négligeables devant 1.

Il convient également de vérifier le conditionnement du problème par le calcul du rapport des axes principaux des ellipses d'incertitudes. Notons qu'une ellipse d'incertitude, dont le rapport des normes des axes principaux est élevé (ellipse fortement excentrique) n'induit pas systématiquement des valeurs de $gdop$ élevées. En effet les axes principaux de l'ellipse peuvent être parallèles aux axes du repère paramétrique.

□ Si les points précédents ne permettent pas l'isolation des causes des intervalles d'incertitude élevés identifiées, le modélisateur doit effectuer un choix parmi les différentes solutions qui s'offrent à lui :

- Accepter les intervalles d'incertitude d'amplitudes élevées.
- Estimer la valeur des paramètres responsables des incertitudes élevées par le biais d'une source d'information supplémentaire (mesure, information a priori ...) et réitérer le processus d'estimation.
- Inclure une information a priori sur les paramètres, en les estimant par minimisation d'un critère de coût formulé au sens du maximum a priori.
- Reconsidérer le problème de l'excitation par le biais de techniques de planification d'expériences optimales.
- Adopter un modèle mettant en jeu un nombre plus restreint de paramètres.

Ce paragraphe a principalement été consacré à la quantification des intervalles d'incertitude associés aux paramètres estimés. Cette quantification est réalisée à partir des propriétés statistiques de l'estimateur utilisé, sous réserve que les différentes hypothèses relatives à la formulation du critère de coût soit remplies. Le paragraphe suivant traite de l'estimation des intervalles d'incertitude associés aux paramètres en présence d'erreur de modélisation.

3.2.4 Estimation en présence d'erreur de modélisation

Lors de l'estimation des paramètres de processus réels, la présence de deux types d'erreur est inévitable. Le premier type est composé d'erreurs de mesure ($\varepsilon(t)$). Ces erreurs de mesure (ou

encore bruit de mesure) induisent des incertitudes sur les paramètres estimés. Le second type d'erreur est constitué des erreurs de modélisation (ou erreurs structurelles : $w(t)$). La description de ces deux types d'erreurs conduit à la formulation suivante sous forme de représentation d'état :

$$\begin{aligned}\dot{x}(t) &= A x(t) + B u(t) + w(t) \\ y(t) &= C x(t) + D u(t) + \varepsilon(t)\end{aligned}\tag{3.3-22}$$

Jusqu'à présent, il a été supposé que les données étaient corrompues uniquement par des erreurs de mesures décrites de manière stochastique. Dans les travaux relatant l'estimation des paramètres physiques de systèmes complexes, peu d'études fournissent une quantification réaliste des intervalles d'incertitude associés aux paramètres. Dans le domaine de l'identification des paramètres physiques de véhicules automobiles, lorsque cet aspect est abordé, les techniques mises en œuvre pour leurs quantifications sont passées sous silence et leur représentativité des erreurs de variance réelle ignorée. La majorité des travaux [TNO97], [HOR96], [LET95], [HUA93], n'aborde même pas le problème de la quantification des erreurs de variance associées aux paramètres. Dans la majorité des travaux consacrés à l'identification des paramètres physiques de systèmes, autres que le véhicule automobile (voir par exemple les travaux de Knudsen [KNU96]), les différentes stratégies de quantification des intervalles d'incertitude reposent alors sur l'exploitation de la seule matrice d'information de Fisher. La représentativité des erreurs de variance ainsi estimées n'est alors pas traitée. Seul Iliff [ILI84] semble s'être penché sur ce problème lors de l'estimation des paramètres aérodynamiques d'aéronefs. Il a étudié et quantifié partiellement l'influence de la présence d'erreur de modélisation sur l'estimation de la variance des paramètres estimés.

3.2.4.1 Constat

Dans ses travaux relatant l'identification des coefficients aérodynamiques d'aéronefs, Iliff [ILI84] constate que les intervalles d'incertitude estimés à partir de la seule matrice d'information de Fisher sont sous-évalués par rapport aux intervalles d'incertitude constatés expérimentalement ; l'estimateur n'est pas efficace.

Nous observons ce même phénomène sur l'identification des paramètres physiques du véhicule automobile. Pour éviter ce problème plusieurs solutions sont alors envisageables. Parmi celles-ci, trois philosophies se dégagent.

- Une approche consiste à modifier (restriction) l'ensemble des paramètres admissible sous certaines contraintes : principe de l'estimation ensembliste à erreur bornée [WAL97] par exemple. Le principe de l'estimation ensembliste à erreur bornée consiste à obtenir à partir des données mesurées, d'un modèle paramétrique et d'une borne sur l'amplitude des résidus, une région dans l'espace paramétrique compatible avec ces trois informations. La mise en œuvre de ce schéma d'estimation n'est pas triviale. La principale difficulté est la caractérisation de la région de l'espace paramétrique adéquate à partir des différentes informations. De surcroît, la région des paramètres admissibles de l'espace paramétrique peut se constituer de différentes sous régions non-connexes de l'espace paramétrique.
- Une seconde voie consiste à modéliser le système par un modèle nominal (représentant le comportement du système aux basses fréquences) et d'un modèle d'erreur de modélisation, représentant les dynamiques rapides négligées par le modèle nominal [OUV97]. Puis, l'erreur d'estimation paramétrique est alors quantifiée par les outils classiques de la théorie stochastique (inégalité de Cramer-Rao). Deux critiques sont alors à émettre et s'opposent à cette démarche. En effet, il est difficile de garantir la véracité de la quantification des erreurs de modélisation par rapport au comportement du modèle nominal. La seconde réserve concerne la quantification de l'erreur de modélisation qui ne peut être que la conséquence de dynamiques rapides par rapport aux dynamiques du modèle nominal. L'erreur de modélisation doit alors se cantonner aux

hautes fréquences. Or généralement, l'erreur de modélisation se manifeste aux fréquences sollicitées par l'excitation (aux basses fréquences dans notre cas).

- Pour ces raisons, une troisième approche plus classique basée sur la minimisation d'un critère d'erreur de sortie au sens du maximum de vraisemblance a été préférée. La région des paramètres admissibles de l'espace de recherche reste inchangée, tandis que le processus de quantification des intervalles d'incertitude, basé sur la matrice d'information de Fisher, est adapté selon les travaux d'Illiff. Cette adaptation permet, dans une certaine mesure, de tenir compte de la présence d'erreur de modélisation dans les basses fréquences.

3.2.4.2 Causes

Dans son étude, Illiff a partiellement isolé la cause de ce phénomène de sous-estimation des intervalles d'incertitude des paramètres estimés. La cause de ce phénomène est essentiellement due à la **violation de l'hypothèse d'indépendance des résidus** ; la puissance spectrale des résidus est essentiellement concentrée sur les basses fréquences (cf. figure 3 - 9).

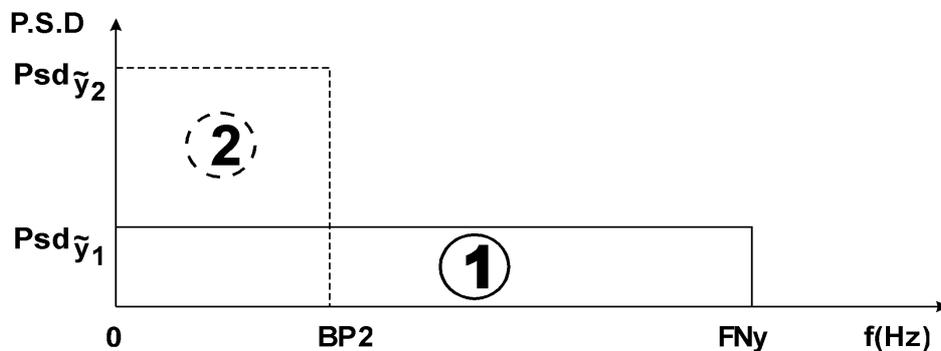


figure 3 - 9 : distribution de la puissance spectrale des résidus

La formulation du critère de coût au sens du maximum de vraisemblance, suppose une distribution uniforme des résidus (qui sont censés être constitués uniquement du bruit de mesure) entre 0 Hz et la fréquence de Nyquist.

Etudions maintenant quantitativement les répercussions de ces différentes distributions des résidus. Pour cela, exprimons l'inégalité de Cramer-Rao, non plus en terme de variance des résidus, mais en terme de leur puissance spectrale. Considérons dans le cas monovarié, les deux distributions (1) et (2) présentant une puissance totale identique (cf. figure 3 - 9).

Reprenons l'expression de l'inégalité de Cramer-Rao :

$$P_{\underline{\theta}} \geq M_F^{-1}(\hat{\underline{\theta}}) = \left(\sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\underline{\theta}}} \Lambda^{-1} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \Big|_{\underline{\theta}=\hat{\underline{\theta}}} \right)^{-1} \quad (3.3-23)$$

qui s'exprime en terme de puissance spectrale des résidus

$$P_{\underline{\theta}} \geq \left(\sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\underline{\theta}}} \left(\frac{\text{Psd}_{\tilde{y}}}{2 \text{Te}} \right)^{-1} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \Big|_{\underline{\theta}=\hat{\underline{\theta}}} \right)^{-1} \quad (3.3-24)$$

avec Te : période d'échantillonnage

Dans les deux cas (1) et (2) précédents nous avons respectivement pour l'amplitude de la puissance spectrale $\text{Psd}_{\tilde{y}}$

$$(1) : \text{Psd}_{\tilde{y}} = 2 \text{Te} \Lambda \quad (3.3-25)$$

$$(2) : \text{Psd}_{\tilde{y}} = \Lambda \text{BP}_2^{-1} \quad (3.3-26)$$

En exprimant (3.3-24) dans ces deux cas, il vient :

$$(1) : \mathbf{P}_{\hat{\theta}} \geq \left(\sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\theta}} \Lambda^{-1} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \Big|_{\underline{\theta}=\hat{\theta}} \right)^{-1} \quad (3.3-27)$$

$$(2) : \mathbf{P}_{\hat{\theta}} \geq \frac{1}{2 \text{BP}_2 \text{Te}} \left(\sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\theta}} \Lambda^{-1} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \Big|_{\underline{\theta}=\hat{\theta}} \right)^{-1} \quad (3.3-28)$$

Dans le premier cas, on retrouve via l'expression (3.3-27) la formulation initiale de l'inégalité de Cramer-Rao (3.3-23). En revanche, dans le second cas, la variance sur les paramètres est sous-estimée d'un facteur $\frac{1}{2 \text{BP}_2 \text{Te}}$. En présence d'erreur structurelle, si malgré tous les efforts

réalisés pour « blanchir les résidus », le spectre des résidus est du type (2), il convient d'appliquer un terme correctif à l'inégalité de Cramer-Rao.

3.2.4.3 Solution

Pour contrer ce phénomène de sous-estimation plusieurs voies sont envisageables. La première méthode consiste à utiliser un facteur multiplicatif (empirique) sur les intervalles d'incertitude. L'avantage indéniable de cette méthode réside dans sa simplicité de mise en œuvre. En revanche, cette méthode nécessite un nombre d'essais élevé et le manque de lisibilité intrinsèque n'est pas souhaitable. En effet, puisque les causes de ce phénomène sont partiellement identifiées et expliquées, il semble naturel et plus judicieux de les exploiter.

La seconde se base sur le calcul d'un facteur correctif tenant compte des remarques formulées ci-dessus. Jusqu'à présent, les puissances spectrales totales dans les deux cas précédents ont été supposées identiques. Dans le cas (2), l'absence totale d'erreur de structure au dessus de la fréquence BP_2 a été supposée. Une approche développée par Maine [MAI81] adapte un terme correctif en fonction de l'amplitude (supposée constante des résidus) entre 0 et BP_2 . Cette méthode consiste dans un premier temps à filtrer les résidus via un filtre passe-bas (dans notre cas filtre de Butterworth d'ordre 10) de fréquence de coupure égale à BP_2 .

L'énergie des résidus est alors calculée

$$\mathbf{J}_{\text{Fil}} = \frac{1}{2} \sum_{t=1}^N \tilde{y}_{\text{Fil}}(t, \hat{\theta})^T \hat{\Lambda}^{-1} \tilde{y}_{\text{Fil}}(t, \hat{\theta}) \quad (3.3-29)$$

Cette énergie convertie en puissance spectrale est alors utilisée pour corriger l'inégalité de Cramer-Rao telle que

$$\mathbf{P}_{\hat{\theta}} \geq \mathbf{M}_{\text{FCor}}^{-1}(\hat{\theta}) \quad (3.3-30)$$

avec

$$\mathbf{M}_{\text{FCor}}^{-1} = \mathbf{K}^2 \frac{\mathbf{J}_{\text{Fil}}}{2 \text{BP}_2 \text{Te} N \text{ny}} \left(\sum_{t=1}^N \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}} \Big|_{\underline{\theta}=\hat{\theta}} \hat{\Lambda}^{-1} \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \Big|_{\underline{\theta}=\hat{\theta}} \right)^{-1} \quad (3.3-31)$$

Bien qu'efficace, cette méthode nécessite l'emploi d'un terme empirique \mathbf{K} . Cependant, lors d'applications réelles (cf. chapitre suivant), sa valeur est bien inférieure au facteur de sous-estimation des intervalles d'incertitude par l'inégalité de Cramer-Rao. La présence de ce facteur se justifie pour au moins deux raisons. La première découle du fait que dans l'expression (3.3-31) la

matrice de covariance du bruit de mesure $\hat{\Lambda}$ comprend également les erreurs structurelles¹⁹. La seconde découle du fait que le spectre des erreurs de prédiction ne présente pas trait pour trait (même après filtrage) les caractéristiques décrites figure 3 - 9.

La valeur de la fréquence BP_2 dépend essentiellement de l'excitation utilisée et définit sa largeur de bande. La détermination empirique du facteur K est moins triviale ; sa valeur est obtenue par approximations successives et comparaison des intervalles de confiance prédits par l'expression (3.3-31) avec les intervalles de confiance obtenus expérimentalement.

3.2.4.4 Validation stochastique par Monte-Carlo

Validons cette démarche en simulation par la méthode de Monte-Carlo, sur notre exemple, le circuit RC de fonction de transfert $G_{i_u}(s) = \frac{i(s)}{v_1(s)} = \frac{C s}{1 + R C s}$. L'identification du vecteur des paramètres $\underline{\theta}_0 = (R, C)$ est réalisée sur 100 jeux de données²⁰. Sur ces données, un bruit blanc de mesure et un bruit rose simulant la présence d'erreur de modélisation sont ajoutés. Généralement, les résidus induits par une erreur de modélisation sont corrélés au signal d'excitation. Dans notre cas, ce fait importe peu, le but étant de montrer que :

- La seule distribution spectrale des résidus est à l'origine de la sous-estimation des intervalles d'incertitude estimés à partir de la seule matrice d'information de Fisher.
- La méthode de Maine pour la quantification des intervalles d'incertitude en présence d'erreurs structurelle est efficace.

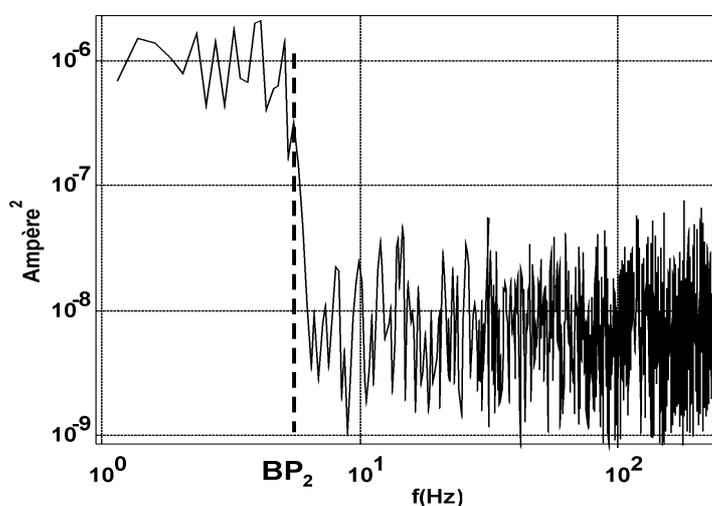


figure 3 - 10 : distribution de la puissance spectrale du bruit simulé

Un bruit de mesure $N(0, 10^{-8})$ est ajouté sur la variable de sortie. L'erreur de modélisation est simulée par filtrage d'une variable aléatoire indépendante $N(0, 10^{-6})$ par un filtre de Butterworth d'ordre 10 de fréquence de coupure égale à 5 Hz. Le spectre résultant du bruit de mesure et de l'erreur de modélisation est présenté figure 3 - 10.

¹⁹ Pour cette raison, l'appellation de « matrice de covariance du bruit de mesure » peut paraître abusive.

²⁰ Conditions de simulation : le système est excité par un signal de type sbpa : période d'échantillonnage = 2ms, temps élémentaire (durée d'un bit) = 14 ms, nombre de bit = 5, amplitude = $\pm 1V$, nombre de périodes = 10.

$\underline{\theta}$	$\underline{\theta}_0$	$\bar{\hat{\theta}}_{mv}$	Int. Conf. à 95 % estimé par C-R (%) (3.3-23)	Int. Conf. à 95 % estimé par C-R (%) corrigées (3.3-31)	Int. Conf. à 95 % estimé expérimentalement (%)
R(Ω)	470	469.86	1.084	2.818	2.816
C(μ F)	100	99.70	1.532	5.014	5.048

tableau 3 - 2 : paramètres estimés et intervalles d'incertitude en présence de bruit blanc et de bruit coloré

Les statistiques résultant des 100 estimations sont portés dans le tableau 3 - 2. Dans ce tableau, les intervalles de confiance expérimentaux en présence de bruit coloré sur les mesures (6^o colonne) sont largement supérieurs à ceux prédits par l'inégalité de Cramer-Rao non corrigée (facteur de l'ordre de 3). La figure 3 - 11a montre les intervalles d'incertitude expérimentaux et estimés (sous forme de barres verticales) pour chaque jeu de données sur les 35 premières estimations²¹ ($P_{\underline{\theta}} \geq M_F^{-1}(\hat{\underline{\theta}})$).

En corrigeant les intervalles d'incertitude par le biais de l'expression (3.3-31), les intervalles d'incertitude estimés pour chaque jeu de données sont nettement plus représentatifs des intervalles expérimentaux (cf. tableau 3 - 2). Dans ce cas de figure la valeur du facteur K est de 1.2 ; cette valeur est bien inférieure au facteur de sous-estimation (approximativement 3) des intervalles d'incertitude par l'inégalité de Cramer-Rao. La figure 3 - 11b montre les intervalles d'incertitude expérimentaux et estimés (sous forme de barres verticales) pour chaque jeu de données sur les 35 premières estimations ($P_{\underline{\theta}} \geq M_{FCor}^{-1}(\hat{\underline{\theta}})$).

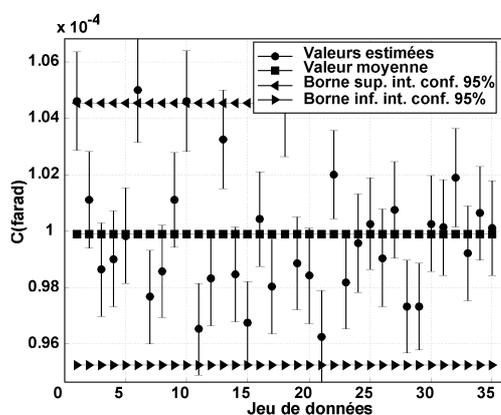


figure 3 - 11a : bornes de Cramer-Rao non-corrigées

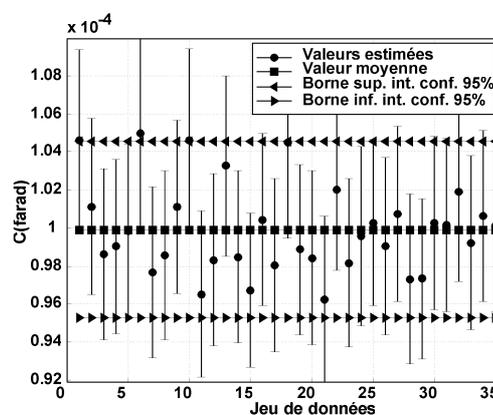


figure 3 - 11b : bornes de Cramer-Rao corrigées

figure 3 - 11: intervalles de confiance expérimentaux et bornes de Cramer-Rao

Malgré l'extrême simplicité du modèle utilisé, le nombre de paramètres limité à 2 et le cadre idéalisé de la simulation, cette méthode de quantification des intervalles en présence d'erreur structurale conserve son caractère général. Des résultats similaires, obtenus à partir de modèles de complexité plus élevée et sur des données expérimentales, seront présentés dans le chapitre suivant.

²¹ Nombre limité à 35 sur les figures pour des raisons de lisibilité. Le tableau 3 - 2 contient les statistiques réalisées sur 100 estimations.

La vérification du pouvoir prédictif du modèle identifié est une des étapes du processus de validation. Dans le paragraphe suivant consacré à la vérification du pouvoir prédictif du modèle identifié, les intervalles d'incertitude associés aux paramètres seront mis à profit pour élaborer des intervalles de confiance sur les prédictions temporelles et sur la fonction de réponse en fréquence du modèle.

4 POUVOIR PREDICTIF

La vérification des hypothèses quant à la nature du bruit présent sur les données, et les intervalles d'incertitude associés aux paramètres ne sont généralement pas suffisants. Il convient aussi de vérifier que le modèle est bien capable de prédire le comportement du processus.

Par ailleurs, pour être significatif, le pouvoir prédictif du modèle doit être évalué en validation croisée. La validation croisée valide le modèle sur des données n'ayant pas servi au processus d'identification. Un apprentissage « par cœur » des données est ainsi évité ; par apprentissage, il est entendu qu'il s'agit du comportement du système et d'une réalisation du bruit de mesure présent sur les données. Un modèle dont le pouvoir prédictif est validé de cette manière est donc théoriquement robuste par rapport au bruit présent sur les données. La démarche de validation croisée garantit, dans une certaine mesure, l'invariance du modèle dans le temps.

Par le biais de critère d'erreur, il est alors possible de juger de l'adéquation du modèle à décrire le comportement du processus réel. Une interprétation du pouvoir prédictif est possible dans le domaine temporel et fréquentiel. Plutôt que concurrentes ces deux approches sont complémentaires. Les notions de biais et de variance vues auparavant pour l'estimation des paramètres, s'étendent aux comportements fréquentiel et temporel du modèle.

4.1 Domaine fréquentiel

Afin d'évaluer le pouvoir prédictif du modèle identifié dans le domaine fréquentiel, il est avantageux de comparer le diagramme de Bode du modèle identifié avec le résultat de l'analyse spectrale réalisée sur les données d'estimation (ou de validation). Avant de poursuivre, distinguons les deux composantes de l'erreur fréquentielle : l'erreur de biais et l'erreur de variance. L'erreur de biais résulte du fait qu'il peut subsister une erreur de modélisation entre le modèle et le processus. L'erreur de variance, quant à elle, est la conséquence de la présence de bruit de mesure.

4.1.1 Erreur de biais

L'erreur fréquentielle de biais d'un modèle LE peut s'évaluer par les outils d'analyse spectrale. Le biais s'estime alors par comparaison de la fonction de réponse en fréquence expérimentale du système avec le diagramme de Bode du modèle identifié. La fonction de réponse en fréquence s'évalue en calculant le rapport de l'interspectre des signaux d'entrée et de sortie sur l'autospectre du signal d'entrée [LJU87], [SOD89] :

$$T_{xy}(\omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} \quad (3.4-1)$$

4.1.2 Erreur de variance

Lorsque le modèle est d'ordre suffisamment élevé pour décrire les dynamiques du système, et que l'estimateur est non biaisé, l'erreur fréquentielle se compose essentiellement de l'erreur de variance. Sous ces conditions, il est possible de juger de la précision de l'estimateur dans le domaine fréquentiel.

Pour cela, il convient de distinguer les cas où la réponse fréquentielle du modèle est LP ou non-LP, [NIN95]. Les travaux présentés dans ce mémoire relatent essentiellement l'identification des paramètres physiques de systèmes continus, la réponse fréquentielle du modèle $G(j\omega, \hat{\underline{\theta}})$ est non-LP. Cependant, cette réponse peut être linéarisée autour d'une valeur du vecteur de paramètres $\hat{\underline{\theta}}$ via un développement en série de Taylor.

$$G(j\omega, \underline{\theta}) \approx G(j\omega, \hat{\underline{\theta}}) + \left. \frac{\partial G(j\omega, \underline{\theta})}{\partial \underline{\theta}^T} \right|_{\underline{\theta}=\hat{\underline{\theta}}} (\underline{\theta} - \hat{\underline{\theta}}) + o(\underline{\theta}) \quad (3.4-2)$$

En définissant l'erreur fréquentielle $g(f)$ telle que [OUV97]

$$g(f) = \begin{bmatrix} \text{Re} \{ G(j\omega, \underline{\theta}_0) - G(j\omega, \hat{\underline{\theta}}) \} \\ \text{Im} \{ G(j\omega, \underline{\theta}_0) - G(j\omega, \hat{\underline{\theta}}) \} \end{bmatrix} \quad (3.4-3)$$

Et en posant

$$\underline{\Omega}^T(f) = \begin{bmatrix} \text{Re} \left\{ \left. \frac{\partial G(j\omega, \underline{\theta})}{\partial \underline{\theta}} \right|_{\underline{\theta}=\hat{\underline{\theta}}} \right\} \\ \text{Im} \left\{ \left. \frac{\partial G(j\omega, \underline{\theta})}{\partial \underline{\theta}} \right|_{\underline{\theta}=\hat{\underline{\theta}}} \right\} \end{bmatrix} \quad (3.4-4)$$

On obtient l'expression suivante de l'erreur fréquentielle

$$g(f) \approx \underline{\Omega}(f)(\underline{\theta}_0 - \hat{\underline{\theta}})^T \quad (3.4-5)$$

qui admet pour matrice de covariance

$$P_f = \text{cov}(g(f)) = E\{g(f) g^T(f)\} \quad (3.4-6)$$

$$P_f = \underline{\Omega}(f) P_{\underline{\theta}} \underline{\Omega}^T(f) \quad (3.4-7)$$

Compte tenu de l'expression asymptotique (3.3-28) l'expression (3.4-7) permet de quantifier les intervalles d'incertitude associés à la réponse fréquentielle du modèle identifié. Sous forme de diagramme de Bode ces intervalles d'incertitude se concrétisent par des intervalles de confiance autour du diagramme de Bode du modèle identifié. Dans le plan de Nyquist, ces intervalles de confiance se concrétisent sous forme d'ellipses d'incertitude.

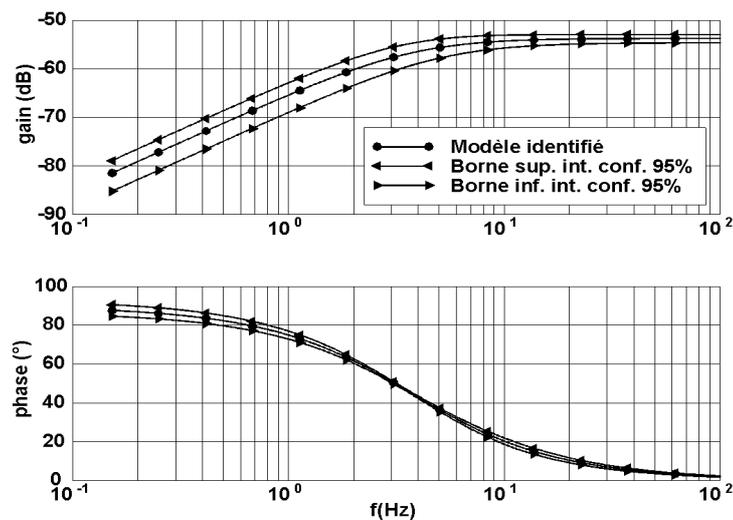


figure 3 - 12 : réponse fréquentielle et intervalles d'incertitude

Ces intervalles de confiance dans le domaine fréquentiel sont liés à la sensibilité du modèle vis à vis des paramètres identifiés et aux intervalles d'incertitude associés aux paramètres identifiés. Un paramètre à sensibilité élevée, identifié avec une précision médiocre, entraînera des intervalles d'incertitude d'amplitude importante dans le domaine fréquentiel. Plus précisément ces intervalles d'amplitude d'amplitudes élevées seront essentiellement concentrés aux fréquences où la sensibilité de la réponse fréquentielle est élevée. Contrairement à l'analyse de la sensibilité (cf. premier chapitre), les couplages éventuels entre les paramètres sont ici pris en compte par le biais de la matrice de covariance $P_{\hat{\theta}}$ des paramètres.

L'obtention d'un modèle nominal et d'intervalles d'incertitude autour de celui-ci est séduisante. Sur la figure 3 - 12 sont représentés les intervalles de confiance associés au diagramme de Bode du circuit

RC ($G_{u/i}(s) = \frac{i}{v_1} = \frac{C s}{1 + R C s}$), après identification de la valeur de la résistance et de la capacité. Il

faut cependant remarquer que les intervalles d'incertitude sont calculés dans un contexte stochastique et que l'hypothèse consistant à considérer l'erreur de modélisation nulle est difficile à satisfaire. Malgré tout, lorsque :

- Le nombre de données est important.
- L'estimateur est non-biaisé (erreur de sortie et maximum de vraisemblance).
- L'ordre du modèle est suffisamment élevé.

Alors l'erreur de biais et l'erreur de variance sont alors d'amplitude voisines. Dans ce cas, Ljung [LJU91] suggère de calculer les intervalles d'incertitude fréquentiels à l'aide de l'expression (3.4-6), puis de les doubler. Moyennant un modèle d'ordre élevé, Ninness [NIN95] remarque que cette approche aboutit à des résultats tout à fait acceptables du point de vue de la quantification des intervalles d'incertitude fréquentiels. Il convient cependant de rester vigilant aux implications de cette méthode lorsque que l'amplitude de l'erreur de modélisation est élevée. Sa caractérisation en tant que réalisation d'une variable aléatoire est discutable.

En tout état de cause, l'approche adoptée consistant à tenir compte de l'éventuelle présence d'erreur de modélisation par le biais de l'expression corrigée de la matrice de covariance des paramètres (3.3-30) procède de la même philosophie. Elle permet de plus une correction plus en aval et est directement réutilisable pour le calcul des intervalles de confiance dans le domaine temporel.

4.2 Domaine temporel

4.2.1 Quantification

Afin de pouvoir comparer le pouvoir prédictif de différents modèles, il est souhaitable d'établir un critère de comparaison. Le calcul de l'erreur de prédiction « efficace » normalisée remplit ce rôle :

$$\tilde{y}_{\text{RMS Nor}}(\hat{\theta}) = 100 \times \sqrt{\frac{\sum_{t=1}^N \tilde{y}^2(t, \hat{\theta})}{\sum_{t=1}^N y^2(t)}} \quad (3.4-8)$$

Ce critère est particulièrement bien adapté à des réponses de systèmes $y(t)$ présentant une composante continue. Dans le cas où une partie de l'essai se déroule à sollicitation quasi nulle, la seule présence d'un bruit de mesure de très faible amplitude dégradera significativement la valeur de cet indicateur ; il sera alors préférable de calculer la quantité non-normalisée suivante.

$$\tilde{y}_{\text{RMS}}(\hat{\underline{\theta}}) = \sqrt{\frac{1}{N} \sum_{t=1}^N \tilde{y}^2(t, \hat{\underline{\theta}})} \quad (3.4-9)$$

4.2.2 Erreur de biais et variance

Les considérations précédentes relatives au biais et à la variance dans le domaine fréquentiel s'applique au domaine temporel. Comme dans le domaine fréquentiel, il convient ici de distinguer le cas où la réponse temporelle du modèle est LP ou non-LP. Dans le cadre de l'identification des paramètres physiques de modèle phénoménologique, certains modèles extrêmement simples peuvent présenter une réponse temporelle LP.

☞ Dans ce cas la réponse temporelle s'exprime sous forme de régression linéaire

$$\hat{y}(t, \hat{\underline{\theta}}) = \underline{\varphi}(t)^T \hat{\underline{\theta}} \quad (3.4-10)$$

Afin de conserver la notation adoptée pour le domaine fréquentiel exprimons l'erreur temporelle par $g(t) = \tilde{y}(t, \hat{\underline{\theta}})$ et posons $\underline{\Omega}(t) = \underline{\varphi}(t)^T$. On obtient alors l'expression de l'erreur temporelle

$$g(t) = \underline{\varphi}^T(t)(\underline{\theta}_0 - \hat{\underline{\theta}}) = \underline{\Omega}(t)(\underline{\theta}_0 - \hat{\underline{\theta}}) \quad (3.4-11)$$

☞ Dans le cas où la réponse du modèle est non-LP il faut à nouveau définir une linéarité locale en utilisant le développement limité de la sortie autour des paramètres estimés $\hat{\underline{\theta}}$

$$y(t, \underline{\theta}) \approx y(t, \hat{\underline{\theta}}) + \left. \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \right|_{\underline{\theta}=\hat{\underline{\theta}}} (\underline{\theta} - \hat{\underline{\theta}}) + o(\underline{\theta}) \quad (3.4-12)$$

En posant $\underline{\Omega}(t) = \left. \frac{\partial y(t, \underline{\theta})}{\partial \underline{\theta}^T} \right|_{\underline{\theta}=\hat{\underline{\theta}}}$ l'expression approchée de l'erreur temporelle devient

$$g(t) = \underline{\Omega}(t)(\underline{\theta}_0 - \hat{\underline{\theta}}) \quad (3.4-12)$$

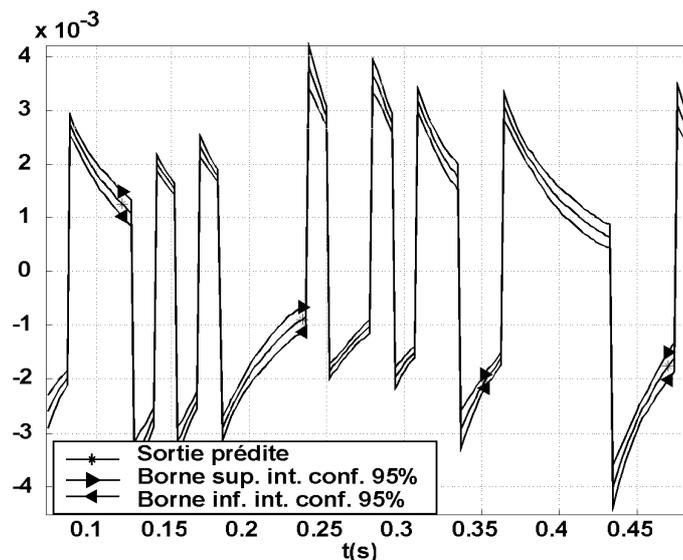


figure 3 - 13 : réponse temporelle et intervalles de confiance

Dans les deux cas précédents (réponse temporelle LP ou non-LP), l'erreur de prédiction temporelle admet pour matrice de covariance

$$P_t = \text{cov}(g(t)) = E\{g(t)g^T(t)\} \quad (3.4-13)$$

$$P_t = \Omega(t) P_{\theta} \Omega^T(t) \quad (3.4-14)$$

Comme pour le domaine fréquentiel, on obtient la réponse nominale du modèle et des intervalles d'incertitude autour de cette réponse nominale (figure 3 - 13).

5 CHOIX DU MODELE

Effectuer un choix parmi les différents modèles présentés pour la description d'un système est délicat. Ce choix est conditionné par de nombreux facteurs. La priorité de ces différents éléments dépend principalement du but poursuivi par la démarche de modélisation entreprise. Il est généralement conseillé d'utiliser différents tests visant à vérifier les caractéristiques des modèles identifiés [LJU87], [SOD89], [WAL97]. Classiquement ces tests concernent (pour les modèles comportementaux et phénoménologiques) :

- Le pouvoir prédictif des modèles.
- Les propriétés des résidus (blancheur, stationnarité, indépendance, distribution, inter corrélation avec l'entrée, ...).
- Le domaine de validité des modèles.
- La compensation pôles/zéros (lié au problème de l'identifiabilité).
- La simplicité des modèles.

A ces points, s'ajoutent les aspects suivants, plus spécifiques à l'identification de paramètres physiques :

- Les propriétés de discernabilité.
- Les propriétés d'identifiabilité.
- La véracité de la valeur des paramètres identifiés.

Les tests relatifs aux propriétés des résidus et au pouvoir prédictif des modèles sont respectivement décrits aux paragraphes 2 et 4 de ce chapitre. Le choix de l'ordre et de la complexité du modèle vise à satisfaire un compromis pouvoir prédictif/simplicité du modèle satisfaisant.

5.1 Pouvoir prédictif/simplicité du modèle

Le compromis pouvoir prédictif/simplicité du modèle se gère généralement par l'emploi de critère de coût $C(\cdot)$ incluant un terme de pénalisation, fonction de la complexité du modèle. Classiquement ces critères sont de la forme

$$C(m_i) = J(\hat{\theta}) + \mu(\dim(\hat{\theta})) \quad (3.5-1)$$

Où m_i est le modèle candidat, $\hat{\theta}$ le vecteur de paramètre identifié associé et $J(\hat{\theta})$ l'erreur quadratique moyenne

$$J_{\text{EQM}}(\hat{\theta}) = \frac{1}{N} \sum_{t=1}^N \tilde{y}(t, \hat{\theta})^T \tilde{y}(t, \hat{\theta}) \quad (3.5-2)$$

Généralement le terme $J(\hat{\theta})$ décroît lorsque l'ordre de la fonction de transfert augmente. Ce terme permet d'éviter une sous-paramétrisation du modèle. Le second terme de complexité, $\dim(\hat{\theta})$, fonction du nombre de paramètres estimés, augmente lorsque le nombre de paramètre croît, pénalise une surparamétrisation potentielle du modèle. Plusieurs critères s'expriment suivant la

forme (3.5-1). Les plus courants sont le critère AIC (Aikake Information Criterion) et le critère FPE (Final Prediction Error) [LJU87] :

$$C_{AIC}(m_i) = N \ln(J_{EQM}(\hat{\theta})) + 2 \dim(\hat{\theta}) \quad (3.5-3)$$

$$C_{FPE}(m_i) = \frac{1 + \dim(\hat{\theta})}{1 - \dim(\hat{\theta})} \frac{N}{N} J_{EQM}(\hat{\theta}) \quad (3.5-4)$$

Ces deux critères permettent d'isoler une structure optimale au sens pouvoir prédictif/ordre du modèle. Dans le cadre de l'identification de paramètres physiques, plusieurs objections s'opposent à l'utilisation de tels critères. Si dans le cadre de l'identification de modèles paramétriques discrets ces critères permettent d'isoler la structure optimale, il n'en est pas de même pour l'identification des paramètres physiques de modèle de connaissance.

En effet, dans ce cas, l'identification d'un paramètre supplémentaire ne traduit pas systématiquement l'augmentation de l'ordre de la fonction de transfert du modèle. De plus, la formulation du critère FPE se base sur une estimation de la variance des résidus [LJU87]. Cette estimation est donc valable si la suite des erreurs de prédiction est assimilable à une suite de variables aléatoires indépendantes. Or, lors de la détermination de la structure du modèle, il est fortement probable que le modéliste soit amené à considérer des modèles sous-paramétrés (où la suite des erreurs de prédiction n'est plus assimilable à une suite de variables aléatoires indépendantes).

Ces critères associent le pouvoir prédictif du modèle à la dimension de l'espace de recherche pour dégager des indicateurs sur la structure optimale du modèle. Malgré les inconvénients liés à leurs formalisations, ces critères peuvent être employés pour dégager certaines tendances. L'utilisation de ces seuls critères n'est pas suffisante. Pour cela, considérons maintenant une approche complémentaire, exploitant essentiellement les intervalles d'incertitude associés aux paramètres estimés.

5.2 Incertitudes sur les paramètres

Dans leur ouvrage consacré à l'identification des modèles paramétriques²² à temps continu, Schoukens et Pintelon [SCH91] proposent d'augmenter le nombre de paramètres estimé jusqu'à ce que l'amplitude d'un des paramètres estimés soit du même ordre de grandeur que l'intervalle d'incertitude associé. Celui-ci conseille alors de supprimer le(s) paramètre(s) en question. Dans le cadre bien précis de l'identification des paramètres de fonction de transfert à temps continu, la présence d'une compensation pôles/zéros se traduit généralement par des intervalles d'incertitude d'amplitude élevée.

Ce raisonnement n'est pas directement transposable à l'identification des paramètres physiques de systèmes continus. En effet, comme pour les critères AIC et FPE, l'ajout ou le retrait d'un paramètre de la fonction estimée ne se traduit pas systématiquement par une augmentation (ou une diminution) de l'ordre du modèle considéré. Toutefois ce raisonnement peut être tenu si l'amplitude d'un paramètre estimé est de l'ordre de grandeur des intervalles d'incertitude associés ; rien ne garantit alors le caractère déterministe de la valeur obtenue.

²² Modèles de la forme $H(s) = \frac{Y(s)}{U(s)} = \frac{a_0 + a_1 s + \dots + a_{na} s^{na}}{1 + b_1 s + \dots + b_{nb} s^{nb}}$ $nb \geq na$

La satisfaction des différents points cités au début du paragraphe 5 vise l'obtention d'un modèle décrivant au mieux le processus. Or, le choix d'un modèle, parmi plusieurs modèles concurrents non-discernables, où que les paramètres estimés présentent des valeurs « raisonnables » et où les résidus n'ont pas pu être invalidés est très difficile. En effet, comme l'affirme Ljung [LJU87], il est philosophiquement impossible de savoir si le modèle théorique décrit de manière rigoureusement exacte le système modélisé.

La question de la véracité des modèles utilisés se pose alors, et l'ultime choix revient au modéliste. Ce choix se base sur certains indicateurs objectifs (incertitude sur les paramètres, pouvoir prédictif, invalidation des résidus) et d'autres plus au moins subjectifs (valeurs a priori des paramètres, connaissances et expériences acquises sur le processus). Finalement, le but ultime de la démarche de modélisation entreprise dictera les différentes options possibles dans les phases d'écriture et de validation du modèle.

Pour notre application (cf. chapitre suivant), ces différents aspects seront successivement abordés lors de l'identification des paramètres des différents modèles de connaissances pressentis pour la description du comportement du véhicule automobile.

6 CONCLUSION

Dans ce chapitre, les principaux aspects relatifs à la **validation de l'identification des paramètres physiques de systèmes à représentation continue ont été abordés**. Après un bref aperçu des outils dédiés à la **validation des résidus**, l'accent est mis sur la quantification et l'étude des intervalles de confiance sur les paramètres estimés.

Lors de l'identification des paramètres physiques de systèmes complexes à temps continu, la présence d'erreur de modélisation est quasi inévitable. Pour ces raisons, la quantification des intervalles d'incertitude associés aux paramètres, par le biais des outils classiques de la théorie stochastique demeure imparfaite.

Cette présence entraîne une sensible différence entre les erreurs de variance expérimentale des paramètres estimés, et celles prévues par les outils classiques de la théorie stochastique. Il a été montré qu'une méthode de quantification des intervalles d'incertitude (inspirée de travaux consacrés à l'identification des paramètres aérodynamiques d'aéronefs) tenant compte de la présence d'erreur structurelle, pallie partiellement ce problème. Cette technique a été validée par Monte-Carlo. Le principal inconvénient de cette méthode est de négliger les dynamiques élevées du système.

Afin de quantifier une erreur de variance de la prédiction tant sur le plan temporel que fréquentiel, les intervalles de confiance ainsi obtenus sont exploités. Ces erreurs de variance de prédiction sont alors utiles pour la conception de loi de commande robuste par exemple.

Enfin, différentes considérations relatives au choix du modèle final sont abordées. Le modèle final satisfera, dans la mesure du possible, un certain nombre de critères. La question du choix du modèle final parmi les différents modèles pressentis ne reçoit à ce jour que des réponses partielles. Ce chapitre apporte au modéliste un maximum d'éléments pour guider son choix ultime. Cependant, la validation d'un modèle ne sera réalisée que dans la mesure où son utilisation aura permis de résoudre le problème motivant sa recherche.

Le chapitre suivant se propose d'appliquer les techniques et les outils présentés dans les trois premiers chapitres de ce mémoire, à l'identification des paramètres physiques d'un véhicule automobile.

APPLICATION AU VEHICULE AUTOMOBILE

1 INTRODUCTION

Lors de l'identification des paramètres physiques de systèmes complexes, les différentes hypothèses relatives aux différentes méthodes d'identification sont de plus en plus difficiles à satisfaire. Dans la mesure où ces hypothèses sont plus ou moins remplies, la phase de validation du modèle revêt un aspect primordial. Dans ce contexte, l'objectif de ce chapitre est de mener une étude de faisabilité de l'identification des paramètres physiques du véhicule automobile.

Le véhicule automobile est un système complexe, difficilement modélisable où les paramètres sont sujets à variation (usure, adhérence, vitesse ...). Dans un premier temps, les facteurs motivant le choix de l'identification des paramètres physiques du véhicule automobile seront présentés. Puis, après avoir brièvement présenté les quelques éléments essentiels de la dynamique du véhicule automobile, différents modèles du véhicule étudié supportant cette étude seront développés. Dans cette optique, 4 modèles de complexité plus ou moins élevée seront décrits. Leurs domaines de validité seront précisés.

L'écriture de ces modèles est la première étape d'une démarche de modélisation visant à décrire le **comportement dynamique latéral du véhicule**. Les différentes propriétés structurelles des modèles seront alors étudiées. Les propriétés d'identifiabilité et de discernabilité souligneront les limitations inhérentes à l'utilisation de ces modèles pour l'identification des paramètres physiques d'un véhicule automobile.

Dans un second temps, compte tenu de ces propriétés, les paramètres physiques des différents modèles de connaissance seront estimés. La véracité des valeurs des paramètres identifiés sera discutée. Une quantification réaliste des erreurs de variance associées aux paramètres estimés sera recherchée et obtenue. Les pouvoirs prédictifs des modèles obtenus confirmeront leurs capacités à décrire précisément les régimes statiques et dynamiques régissant le véhicule sur le plan latéral, tant dans le domaine temporel que fréquentiel.

2 OBJECTIF

L'objectif de cette étude est une augmentation de la connaissance du comportement du véhicule automobile. Les motivations de ces travaux sont multiples, elles peuvent en effet contribuer à la phase de **conception** du véhicule automobile, et à l'amélioration de la **compréhension des phénomènes dynamiques** régissant le comportement du véhicule automobile. Différentes approches sont alors possibles. Une des plus naturelle consiste certainement à décrire le comportement du processus à partir des lois de la physique le gouvernant. Le système est alors étudié par le biais de simulations, études temporelles, études fréquentielles... Une seconde approche consiste à établir des modèles de représentation afin de prédire le comportement du véhicule en réponse à des sollicitations spécifiques.

Or l'identification du comportement transversal du véhicule par le biais de l'identification de modèles paramétriques discrets se heurte à de nombreuses limitations. Les premiers travaux entamés en début de thèse illustrent ces limitations, à savoir :

- Le **comportement transversal** du véhicule possède théoriquement **une infinité de points de fonctionnement**. En effet, les gains statiques et les dynamiques des différentes fonctions de transferts du véhicule automobile dépendent directement de la vitesse ou du carré de celle-ci. Les modèles comportementaux identifiés sont donc peu robustes vis à vis de variations de vitesse. **L'identification du comportement latéral du véhicule** sur une large gamme de vitesse est donc **impossible** par le biais d'un unique modèle paramétrique discret.
- En fonction des sorties variables considérées, la structure des modèles paramétriques discrets dépend de la vitesse longitudinale du véhicule. A titre d'exemple, l'ordre et le retard pur de la fonction de transfert liant la vitesse de lacet à l'angle volant demeurent identiques pour des vitesses comprises entre 40 et 110 km/h. En revanche, la structure de la fonction de transfert entre l'accélération transversale et l'angle volant dépend significativement de la vitesse.

Les **modèles paramétriques discrets** du comportement transversal du véhicule possèdent un **domaine de validité restreint** autour d'une vitesse donnée. Un modèle global peut alors être obtenu par unification de ces différents modèles locaux. La vitesse, représentée sous formes de fonctions d'appartenance floues, ou encore par un réseau neuronal à fonctions de base radiale, permettrait alors la synthèse des différents modèles locaux. Les travaux de M. Farges [FAR95] consacrés à la mise en œuvre de réseaux neuronaux pour l'identification et la stabilisation du porte-avions Charles De Gaulle mettent en œuvre une stratégie analogue.

Cependant, le manque de lisibilité induit par ce type d'approche ne favorise pas une interprétation physique des paramètres identifiés et, dans le cas présent, du comportement du véhicule. Il semble alors logique d'inclure la connaissance disponible a priori sur le système. Une description du comportement du véhicule par le biais de **modèles phénoménologiques** paraît alors logique.

Les domaines de validité des modèles phénoménologiques sont généralement plus larges que ceux des modèles de comportement. Cependant, leur obtention est souvent plus lourde. Un des problèmes majeurs consiste à obtenir les valeurs des paramètres nécessaires à leur écriture. Les valeurs de ces paramètres se déterminent classiquement par :

- Mesure directe sur le processus lorsque la mesure est accessible et financièrement acceptable.
- Prise en compte des valeurs nominales fournies par le constructeur.

Toutefois, cette approche est limitée par plusieurs facteurs :

- L'indisponibilité des données constructeur.

- La confidentialité de certaines valeurs constructeur et/ou de leurs imprécisions (exemple typique des pneumatiques dont les caractéristiques sont souvent très difficiles à obtenir et dont les valeurs sont peu précises : ces valeurs sont typiquement associées à des **intervalles d'incertitude avec un degré de confiance de 95% de l'ordre de 60%**).
- La non-stationnarité à plus ou moins long terme des différentes données. A titre d'exemple, les **masses**, les **inerties** et plus généralement les caractéristiques de la dynamique du véhicule sont **sujettes au chargement** du véhicule.

Le processus d'estimation des paramètres physiques retenu repose sur la simulation de modèle(s) de connaissance du véhicule automobile.

3 MODELES DE CONNAISSANCE

L'écriture des différents modèles de connaissance utilisés est une étape cruciale. De la véracité des modèles à décrire le comportement du véhicule, dépend la qualité des résultats finaux. Cette étape doit donc être particulièrement soignée. Dans le but de sélectionner un modèle décrivant le comportement latéral réel du véhicule de manière la plus exacte que possible, tout en veillant à ne pas aboutir à un modèle non-identifiable, plusieurs modèles de différentes complexités sont décrits et étudiés. L'exactitude de ces différents modèles à rendre compte des différents phénomènes régissant le comportement réel d'un véhicule automobile sera discutée. Dans la mesure du possible leurs domaines de validité seront précisés. Avant d'entamer la procédure de modélisation du comportement latéral du véhicule, détaillons quelques notions de la dynamique véhicule indispensables.

3.1 Notions de dynamique véhicule

Dans ce paragraphe quelques éléments de base indispensable à la compréhension du comportement des véhicules sur le plan latéral seront présentés. Ce paragraphe ne se voulant pas exhaustif, la lecture de l'excellent travail de synthèse réalisé par Chamailard dans son mémoire [CHA96] et des références [MIL95], [GIL92], [REN93], [DOE80], [DEU70], [ELL69] est conseillée.

3.1.1 Élément de dynamique latérale du véhicule

Généralement, le comportement global du véhicule se décrit par le biais de l'expressions de ses mouvements de rotation (lacet, roulis, tangage) et de translation (x , y , z). Sur la figure 4 - 1 sont représentés ces mouvements dans un repère lié à la caisse du véhicule et centré sur le centre de gravité du véhicule.

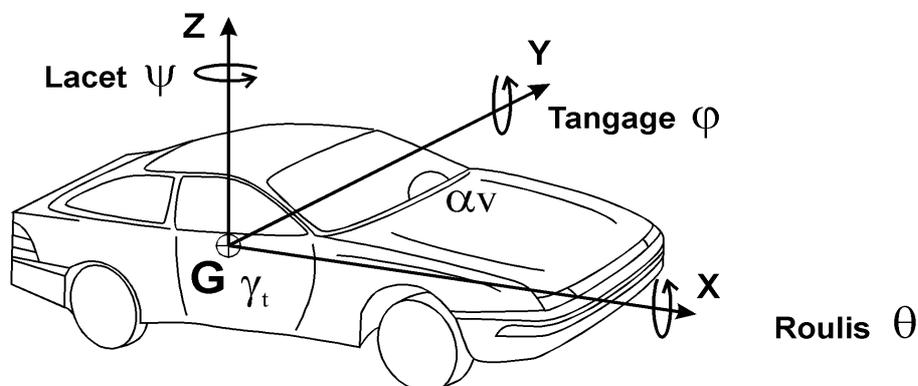


figure 4 - 1 : dynamique du véhicule

Le conducteur perçoit le comportement du véhicule dans le plan transversal et le plan longitudinal. Les comportements du véhicule dans ces deux plans sont étroitement couplés. A long terme, notre étude s'inscrit dans une démarche d'augmentation de la connaissance du comportement global du véhicule. Cependant, **dans un premier temps, l'étude est restreinte à son comportement transversal**. Pour ces raisons, seuls les mouvements mettant en jeu la dynamique latérale du véhicule sont décrits et pris en compte. Plus précisément, l'attention est concentrée sur la description des mouvements de lacet, de roulis et d'accélération transversale. Les modèles employés décrivent le comportement transversale du véhicule automobile du véhicule à vitesse constante du véhicule, ce qui permet dans une certaine mesure de découpler les comportements longitudinaux et transversaux du véhicule.

Afin d'identifier les paramètres physiques régissant le comportement transversal du véhicule, diverses variables sont prélevées, via un système d'acquisition, sur le véhicule d'essais du laboratoire : ces variables sont celles de **l'angle donné au volant** (α_v), la **vitesse de lacet** ($\dot{\psi}$), **l'accélération transversale** (γ_t) et la **vitesse longitudinale du véhicule** (V). L'angle volant est mesuré par le biais d'un capteur de type **potentiomètre rotatif**. La vitesse de lacet est mesurée par le biais d'un **gyroscope**, tandis que l'accélération transversale est mesurée par l'intermédiaire d'un **accéléromètre piezo-résistif**. Les **capteur inductifs** (utilisés par le système ABS d'origine) disposés sur les 4 roues du véhicule permettent quant à eux, la mesure la vitesse longitudinale du véhicule.

3.1.2 Référentiel

La signification des variables utilisées dans ce chapitre et les paramètres propres au véhicule sont explicités dans le paragraphe notation situé au début de ce mémoire. Les référentiels employés au nombre de 3, sont donnés par la figure 4 - 2 et figure 4 - 3.

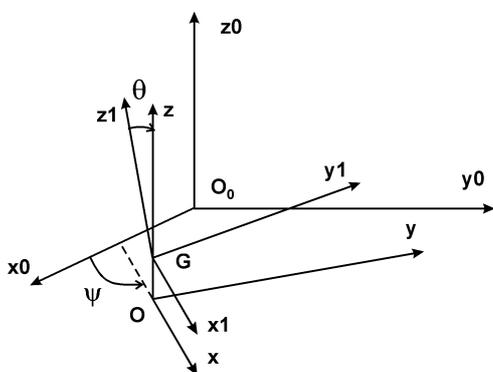


figure 4 - 2a : trièdres

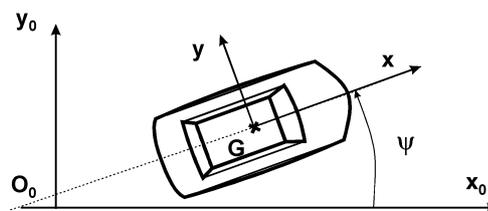


figure 4 - 2b : angle de lacet

figure 4 - 2 : référentiels

Les trois trièdres sont respectivement :

- Le trièdre $R_0 = (O_0, x_0, y_0, z_0)$ galiléen, par rapport auquel le principe fondamental de la dynamique sera appliqué.
- Le trièdre intermédiaire $R = (O, x, y, z)$ où $(\overrightarrow{O_0 x_0}, \overrightarrow{O x}) = \psi$.
- Le trièdre lié à la caisse du véhicule (masse suspendue) $R_1 = (G, x_1, y_1, z_1)$ où $(\overrightarrow{O y}, \overrightarrow{G y_1}) = \theta$.

Ces repères définissent les angles d'Euler de lacet ψ et de roulis θ . Détaillons brièvement l'angle de roulis et certains des paramètres intervenant dans la modélisation de ce phénomène.

3.1.3 Roulis

Pour limiter les mouvements du roulis qui vont à l'encontre de la stabilité du véhicule, les constructeurs introduisent des barres antiroulis. Chaque train du véhicule présente une raideur et un amortissement antiroulis.

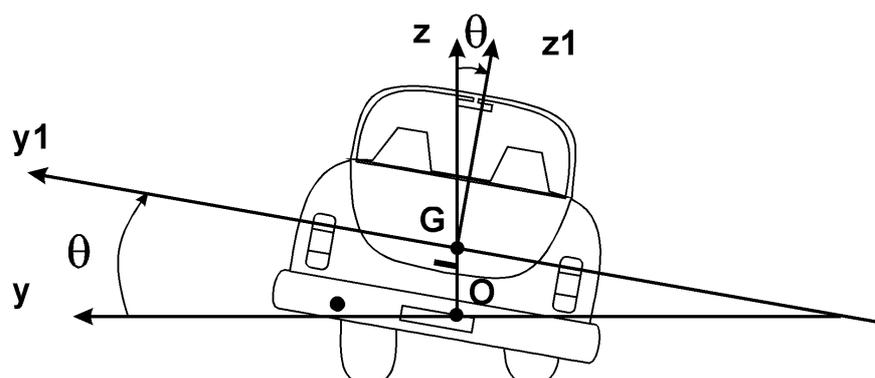


figure 4 - 3 : angle de roulis

La présence d'un écart angulaire (de roulis) entre la caisse (masse suspendue) et le train est limité par ces raideurs et ces amortissements via un moment d'antiroulis. Ce moment est appliqué en un point appelé centre de roulis autour d'un axe (axe de roulis) dans le plan de symétrie longitudinal du véhicule. La direction de l'axe de roulis est définie usuellement par rapport au centre de gravité du véhicule et par les hauteurs des centres de roulis avant et arrière (respectivement h_1 et h_2 sur la figure 4 - 4).

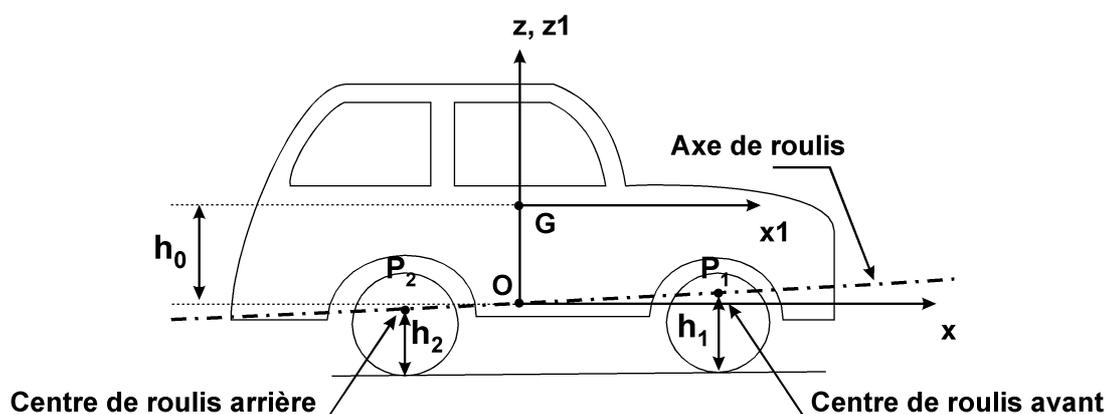


figure 4 - 4 : axe de roulis

3.1.4 Pneumatiques

Tous les efforts (efforts aérodynamiques exceptés) appliqués sur le véhicule sont générés à l'interface pneumatique/sol. L'importance prépondérante des pneumatiques sur le comportement routier d'un véhicule est alors aisément compréhensible.

Les pneumatiques remplissent essentiellement trois fonctions [GIL92] :

- Supporter la charge verticale du véhicule et amortir les déformations du sol.
- Développer les efforts longitudinaux qui accélèrent et freinent le véhicule.

- Développer les efforts transversaux qui lui permettent de tourner.

Dans le cadre de ces travaux, l'étude du véhicule se limite à son comportement transversal à vitesse quasi-constante : le régime établi²³ du véhicule est alors considéré. Ce comportement est conditionné par les efforts transversaux générés par les pneumatiques. Ces efforts s'expriment généralement par une fonction de l'angle de dérive des pneumatiques.

3.1.4.1 Angle de dérive

Lorsqu'une combinaison roue plus pneumatique est soumise à un effort transversal, la surface de contact du pneumatique sur le sol glisse selon une direction opposée à cet effort. La déformation de la surface de contact crée un angle entre l'axe longitudinal de la roue et la direction de son mouvement. Cet angle est appelé angle de dérive du pneumatique. La figure 4 - 5 illustre le phénomène dérive du pneumatique.

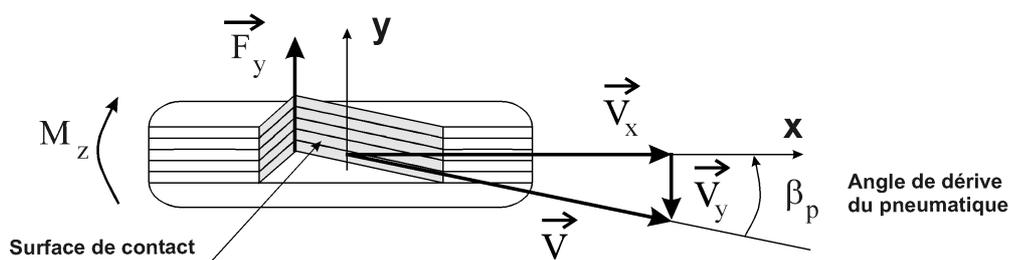


figure 4 - 5 : dérive du pneumatique

La surface de contact avec le sol se déplace vers l'arrière (respectivement l'avant) si le véhicule avance (recule). En réaction à ces déformations, le pneu développe un effort transversal F_y en fonction de l'angle de dérive et un moment d'auto alignement M_z proportionnel à l'effort transversal F_y . L'influence mineure du moment d'auto alignement justifie le fait qu'il soit négligé dans notre étude. Cette simplification réduira les problèmes posés par les propriétés d'identifiabilité des paramètres. L'angle de dérive du pneumatique se définit classiquement²⁴ par :

$$\beta_p = \text{atan}(V_y / V_x) \quad (4.3-1)$$

3.1.4.2 Caractéristique d'adhérence transversale

La caractéristique d'adhérence transversale du pneumatique sur le sol est fortement non-linéaire en fonction de la dérive. Elle suit une allure donnée figure 4 - 6 pour différentes conditions d'adhérence et différentes charges verticales supportées par le pneumatique.

La figure 4 - 6a représente l'évolution de l'effort transversal généré par le pneumatique pour un angle de dérive variant de 0 à 10 degrés et différentes conditions d'adhérence et pour une charge verticale de 7000 N. La figure 4 - 6b traduit l'évolution de la caractéristique d'adhérence transversale en fonction de l'angle de dérive, pour une adhérence très forte (sol de type macadam sec) et différentes charges verticales exercées sur le pneumatique. La figure 4 - 6a montre que pour un angle de dérive défini, l'effort latéral transmis au véhicule est d'autant plus élevé que l'adhérence est forte, ou pour une adhérence donnée, l'effort latéral transmis au véhicule est d'autant plus élevé que la force verticale exercée sur le pneumatique est élevée. Un des éléments conditionnant la stabilité

²³ Le régime établi est aussi nommé régime statico-dynamique par les constructeurs automobile.

²⁴ De manière similaire, l'angle de dérive du véhicule se définit comme l'angle présent entre son axe longitudinal et le vecteur vitesse en son centre de gravité.

du véhicule est donné par la pente à l'origine de la caractéristique d'adhérence. Cette pente est appelée **rigidité de dérive** et est définie par.

$$D = \frac{\partial F_y}{\partial \beta_p} \quad (4.3-2)$$

Plus la valeur de la rigidité de dérive d'un pneumatique est élevée, plus ce dernier est apte à répondre à une sollicitation transversale importante et meilleur sera le comportement du véhicule. La valeur de la rigidité de dérive d'un pneumatique est généralement connue avec une précision limitée et dépend d'un grand nombre de facteurs tels que :

- Le mode de fabrication des pneumatiques : sa hauteur, sa largeur, ses sculptures.
- Son conditionnement : la charge verticale supportée par celui-ci, sa pression de gonflage, sa température...

Par ailleurs, le pneumatique admet une limite au-delà de laquelle il ne peut plus supporter d'effort transversal supplémentaire. Le phénomène de saturation de la caractéristique d'adhérence transversal illustre clairement cette limite. Au-delà de cette limite de saturation, le pneumatique dérape transversalement. Cette limite est d'autant plus élevée que l'adhérence est forte et qu'une charge verticale est exercée sur le pneumatique.

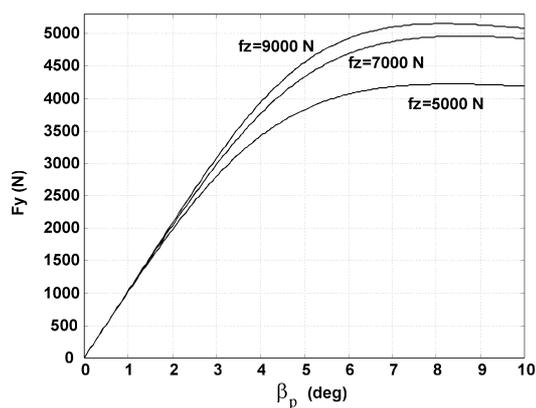
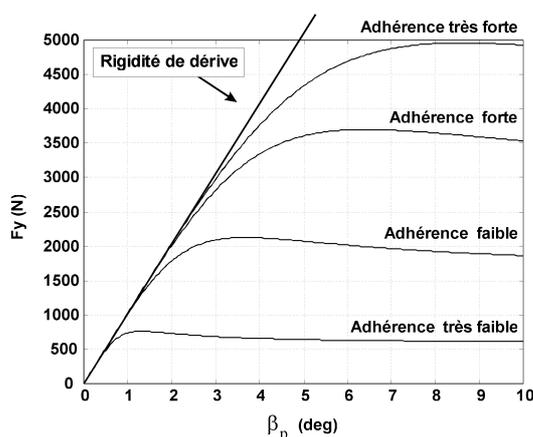


figure 4 - 6a : en fonction de l'adhérence

figure 4 - 6b : en fonction de la charge

figure 4 - 6 : caractéristique d'adhérence transversale

La compréhension du comportement non-linéaire du pneumatique est essentielle à l'étude du comportement du véhicule. Vu l'extrême complexité des phénomènes mis en jeu au niveau de l'interface pneumatique/sol, le comportement dynamique du pneumatique est encore de nos jours, mal expliqué. Une description analytique des phénomènes mis en jeu est de ce fait impossible. Parmi les travaux consacrés à l'étude du comportement des pneumatiques, ceux de Pacejka font référence. Pacejka [PAC91] propose une description du comportement du pneumatique par le biais de modèles de représentation. Ces modèles de représentation macroscopiques complètent alors les modèles de connaissance de véhicule.

Ainsi, lors de la modélisation d'un véhicule, si l'essentiel des composantes mécaniques (telles que les inerties, raideurs, amortissement...) peuvent être décrites par le biais de modèle de connaissance, le modèle complet du véhicule dotés de ses pneumatiques aboutira à un modèle mixte de type boîte grise.

3.1.4.3 Ballant et longueur de relaxation

Tout effet latéral généré par le pneumatique provoque un **déplacement de la surface de contact** roue/sol par rapport au plan de la jante. Ce déplacement est le ballant du pneumatique (cf. figure 4 - 7). Intuitivement, il est aisé de comprendre que pour un effort latéral donné et une pression de gonflage nominale, le ballant d'un pneu taille basse sera plus faible que celui d'un pneu conventionnel.

Pour de faibles déformations le ballant est sensiblement proportionnel à l'effort latéral

$$F_y \cong K_p d \quad \text{avec } K_p \text{ la rigidité de ballant} \quad (4.3-3)$$

La valeur du ballant d est rarement explicitement usitée par les constructeurs automobiles. Le concept de longueur de relaxation est alors préféré. Lors d'une prise de virage, le pneumatique se déforme et doit rouler une certaine distance pour pouvoir se mettre en appui afin de générer un effort latéral. La distance que le pneumatique doit parcourir est la **longueur de relaxation du pneumatique** (b). Aussi curieux que cela puisse paraître, quels que soient les articles consacrés à ce sujet, la valeur de la longueur de relaxation s'exprime indépendamment de la vitesse et sa valeur est toujours de l'ordre de 0.5 m ([REN93], [LET95]).

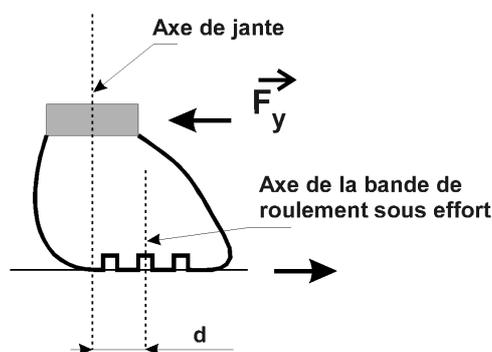


figure 4 - 7 : ballant d'un pneumatique

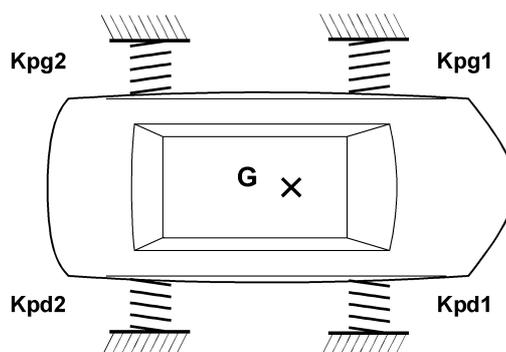


figure 4 - 8 : effet du ballant

Sur le plan latéral, le ballant offre au véhicule la liberté d'osciller autour de son axe de lacet et intervient donc sur son comportement routier. Son influence est particulièrement sensible sur l'accélération transversale du véhicule. Pour une vitesse longitudinale donnée, une excitation volant de type sinus vobulé, il apparaît une fréquence d'excitation pour laquelle le véhicule ne génère plus d'accélération transversale ni d'angle de roulis. Un phénomène d'antirésonance sur la fonction de transfert liant l'accélération transversale à l'angle volant apparaît. A cette fréquence (approximativement 2.2 Hz sur un des véhicules d'essai du laboratoire), l'angle donné au volant est en opposition de phase avec le mouvement de lacet du véhicule. Ainsi, l'effet du ballant ou encore de la longueur de relaxation sur le comportement du véhicule est loin d'être négligeable.

L'évolution temporelle des signaux d'angle volant et d'accélération transversale pour une excitation volant de type harmonique, à vitesse constante est représentée figure 4 - 9. Les données sont issues d'essais expérimentaux réalisés sur un des véhicules du laboratoire.

Etablissons maintenant les différents modèles utilisés dans cette étude. Dans un premier temps, les premiers membres des équations différentielles (équations de la dynamique du corps solide) sont établis.

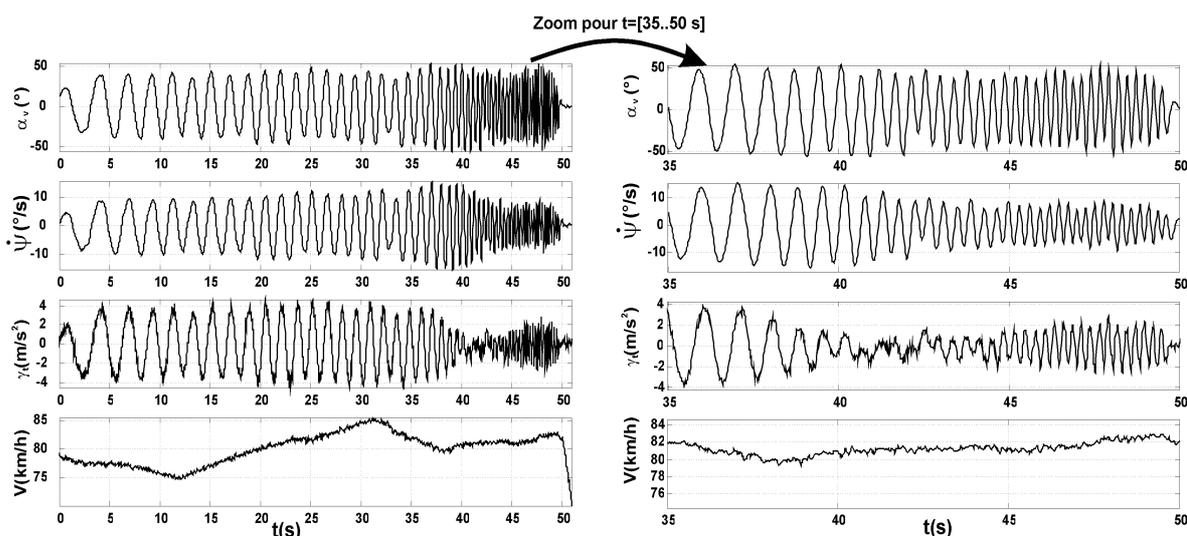


figure 4 - 9 : effet du ballant sur la réponse temporelle en accélération transversale

3.2 Equations de la dynamique du corps solide : premiers membres des équations du mouvement

En appliquant le principe fondamental de la dynamique, le calcul de l'accélération transversale absolue du centre de gravité du véhicule dans le trièdre R_0 conduit à l'expression des forces :

$$\begin{aligned} F_x &= -M V (\dot{\beta} + \dot{\psi}) \sin \beta - m_s h_0 \dot{\theta} \dot{\psi} \\ F_y &= M V (\dot{\beta} + \dot{\psi}) \cos \beta - m_s h_0 \ddot{\theta} \\ F_z &= 0 \end{aligned} \quad (4.3-4)$$

Les variables suivantes décrivant le comportement du véhicule, dépendent du temps : $\beta, \dot{\beta}, \theta, \dot{\theta}, \ddot{\theta}, \psi, \dot{\psi}, \alpha_v$. L'expression du moment dynamique dans le repère R_0 s'obtient alors en appliquant le théorème du moment dynamique.

$$\begin{aligned} L_x &= (I_{xx} + m_s h_0^2) \ddot{\theta} - I_{xz} \dot{\psi} - m_s h_0 V \cos \beta (\dot{\beta} + \dot{\psi}) \\ M_y &= -m_s h_0 V \sin \beta (\dot{\beta} + \dot{\psi}) + (I_{xx} - I_{zz} + m_s h_0^2) \dot{\theta} \dot{\psi} + I_{xz} (\dot{\theta}^2 - \dot{\psi}^2) \\ N_z &= m_s h_0 V \cos \beta \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} \end{aligned} \quad (4.3-5)$$

Les calculs intermédiaires des expressions précédentes sont consignés en annexe 3. Ces équations seront utilisées partiellement ou intégralement selon les degrés de liberté inclus dans les différents modèles de connaissance décrits. Ces modèles de connaissance du comportement transversal du véhicule tiennent compte des principaux phénomènes régissant la dynamique transversale du véhicule : les mouvements de lacet/dérive, de roulis et de ballant. De nombreux autres phénomènes (tels que le carrossage dynamique, déformations élastocinématiques des trains, l'influence de la pince sur la dérive...) sont très difficiles à modéliser. De plus, les valeurs des paramètres nécessaires à leur description sont peu ou pas connus : dans le meilleur des cas seuls leurs ordres de grandeur sont connus. De plus, à l'heure actuelle, l'influence respective de ces différents phénomènes sur le comportement global du véhicule est encore mal expliquée par les constructeurs automobiles. Dans ce mémoire, les modèles supportant notre étude sont au nombre de quatre :

q Modèle LaRouDéBal

Ce modèle décrit les mouvements de **LA**cet, **ROU**lis, **DE**rive et **BAL**lant du véhicule. Il est établi à partir des connaissances disponibles sur le véhicule, des travaux de Le Letty [LET95], Pham [PHA85], Song [SON92] et des références [MIL95], [REN93] et [DEU70]. C'est le modèle le plus complet de notre étude. D'un point de vue théorique, ce modèle possède le meilleur pouvoir prédictif. En revanche, vu le nombre élevé de paramètres physiques nécessaires à son écriture et des couplages existant entre eux, nous montrerons que le processus d'estimation des paramètres physiques se heurte à des problèmes d'identifiabilité.

☞ **Modèle LaRouDé**

Ce modèle décrit les mouvements de **LA**cet, **ROU**lis et **DE**rive. Ce modèle, comme les suivants, est obtenu par simplification du modèle LaRouDéBal. Ce type de modèle incluant les mouvements de lacet/dérive et de roulis est fréquemment utilisé dans la littérature (voir [SON92] par exemple) lorsque le simple modèle bicyclette (LaDé, voir ci-dessous) se révèle insuffisant.

☞ **Modèle LaDéBal**

Les mouvements de **LA**cet, **DE**rive et **BAL**lant sont décrits par ce modèle. Peu utilisé dans la littérature (pas à la connaissance de l'auteur), il offre cependant de très intéressantes perspectives dans le cadre de l'identification des paramètres physiques.

☞ **Modèle LaDé**

Ce modèle est fréquemment utilisé pour l'étude du comportement de la dynamique transversale de véhicule automobile [LAU98], [MIL95] ou le contrôle de trajectoire [ACK96]. Rickert est à l'origine de ce modèle [RIC40]. Il permet de décrire les principales composantes (mouvement de **LA**cet et **DE**rive) transversales du mouvement du véhicule à vitesse constante. Toutefois son domaine de validité est restreint aux faibles fréquences d'excitation (approximativement inférieures à 1.5 Hz ; au-delà l'influence du ballant n'est plus négligeable et doit être prise en compte). Malgré tout, son utilisation permet de se familiariser aux nombreux problèmes posés par l'étude du comportement transversal d'un véhicule automobile.

3.3 Domaine de validité, hypothèses des modèles

Les formulations des modèles de connaissance décrits dans ce mémoire se basent sur les hypothèses suivantes :

- **L'accélération transversale est inférieure à 0.4g** afin d'éviter tout phénomène de saturation des pneumatiques (rigidité de dérive des pneumatiques assimilable à la pente à l'origine de la caractéristique d'adhérence transversale).
- **Aucun effort aérodynamique** n'intervient sur le véhicule.
- **Les angles de dérive sont de faibles amplitudes** : $\sin \beta \approx \beta$, $\cos \beta \approx 1$.
- **La vitesse longitudinale du véhicule est constante** : absence de transferts de charge longitudinaux (tangage).
- **Les masses sont concentrées sur les trains**, le véhicule est symétrique par rapport à son axe longitudinal.

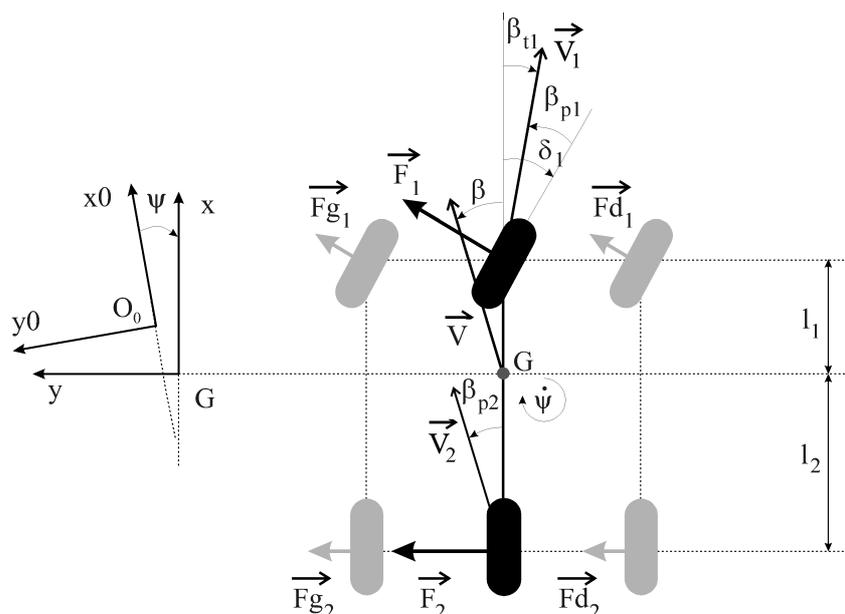


figure 4 - 10 : modèle simplifié du véhicule

La fréquence d'excitation maximale applicable au volant est limitée approximativement à 2.5 Hz. Au-delà, de nombreux phénomènes « hautes fréquences » encore mal définis interviennent. Pour l'ensemble des modèles décrits dans ce mémoire, une simplification consistant à **remplacer les deux roues d'un essieu par une seule** est réalisée. Pour le véhicule complet, seules deux roues (une par essieu) sont représentées, cf. figure 4 - 10. Une identité des efforts sur les deux roues d'un même essieu est supposée, soit pour l'avant :

$$F_1 = F_{G1} + F_{D1} = -D_1 \beta_{p1} \quad (4.3-6)$$

L'effort équivalent généré sur l'essieu s'approche alors par la somme des deux forces transversales générées par chaque pneu. Une rigidité de dérive équivalente ($D1$ pour le train avant et $D2$ pour le train arrière) au train, résultant de la somme des rigidités de dérive des deux pneus d'un même essieu est alors définie.

3.4 Modèle LaRouDéBal

3.4.1 Forces et moments appliqués sur le véhicule

Les seconds membres des équations du mouvement représentent les forces et moments exprimés dans le trièdre R_0 extérieurs appliqués soit sur l'ensemble du véhicule, soit sur la caisse (masse suspendue). Ces forces et moments proviennent :

- Des actions de guidage des pneumatiques.
- Des différentes raideurs et amortissements (essentiellement la raideur et l'amortissement antiroulis).

Pour exprimer ces différentes forces et moments, considérons successivement l'expression de l'accélération transversale du véhicule, les moments dynamiques et les efforts transversaux générés au niveau des roues (ramenées au train).

3.4.1.1 Accélération transversale

L'accélération transversale du véhicule résulte de la somme des forces latérales générées au niveau des pneumatiques. L'hypothèse de petit angle pour l'angle de dérive permet l'expression suivante de l'effort transversal au centre de gravité :

$$F_y = M V (\dot{\beta} + \dot{\psi}) - m s h_0 \ddot{\theta} = F_{y1} + F_{y2} = M \gamma_t \quad (4.3-7)$$

3.4.1.2 Moments dynamiques

Le moment des forces autour de (G, x) est noté L_x et le moment des forces autour de (G, z) est noté N_z . En posant $I_r = I_{xx} + m s h_0^2$ pour l'inertie de la masse suspendue autour de l'axe de roulis et en faisant l'hypothèse de petits angles il vient :

$$\begin{aligned} L_x &= I_r \ddot{\theta} - I_{xz} \ddot{\psi} - m s h_0 V (\dot{\beta} + \dot{\psi}) = -k_{ro} \theta - a_{ro} \dot{\theta} \\ N_z &= m s h_0 V \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} = I_1 F_{y1} - I_2 F_{y2} \end{aligned} \quad (4.3-8)$$

Avec k_{ro} et a_{ro} respectivement la raideur et l'amortissement de roulis global du véhicule. Les expressions (4.3-7) et (4.3-8) nécessitent l'expression de l'effort transversal au niveau de chaque roue du véhicule.

3.4.1.3 Effort transversal à la roue

L'effort transversal est égal au produit de la rigidité de dérive avec la dérive au pneumatique, soit pour le train avant

$$F_1 = -D_1 \beta_{p1} \quad (4.3-9)$$

Et pour le train arrière

$$F_2 = -D_2 \beta_{p2} \quad (4.3-10)$$

Ces efforts s'expriment en fonction de la dérive au pneu de chaque train. Il faut donc estimer la vitesse transversale des points de contact pneumatique/sol par rapport au plan de jante. Dans un premier temps, calculons la vitesse transversale au train avant dans le repère caisse R_1 (cf. annexe 4) :

$$\vec{V}_1 = \begin{pmatrix} V_{1X} \\ V_{1Y} \\ V_{1Z} \end{pmatrix} = \begin{pmatrix} V_x \\ V_y + h_1 \dot{\theta} + l_1 \dot{\psi} \\ 0 \end{pmatrix} \quad (4.3-11)$$

Alors l'expression de la dérive au train avant

$$\beta_{t1} = \text{atan}(V_{1Y} / V_{1X}) \quad (4.3-12)$$

Soit avec l'hypothèse de petits angles

$$\beta_{t1} = \frac{V_y}{V_x} + \frac{h_1 \dot{\theta}}{V_x} + \frac{l_1 \dot{\psi}}{V_x} \quad (4.3-13)$$

Par souci de clarté, la vitesse longitudinale du véhicule sera dorénavant notée V . La dérive au train avant s'exprime alors en fonction de la dérive β au centre de gravité du véhicule

$$\beta_{t1} = \beta + \frac{h_1 \dot{\theta}}{V} + \frac{l_1 \dot{\psi}}{V} \quad (4.3-14)$$

En tenant compte de l'angle de braquage des roues avant ($\delta_1 = \alpha_v / \text{dem}$), la dérive au pneumatique avant (cf. figure 4 - 11) s'exprime :

$$\beta_{p1} = \beta_{t1} - \delta_1 \quad (4.3-15)$$

$$\beta_{p1} = \beta + \frac{h_1 \dot{\theta}}{V} + \frac{l_1 \dot{\psi}}{V} - \delta_1 \quad (4.3-16)$$

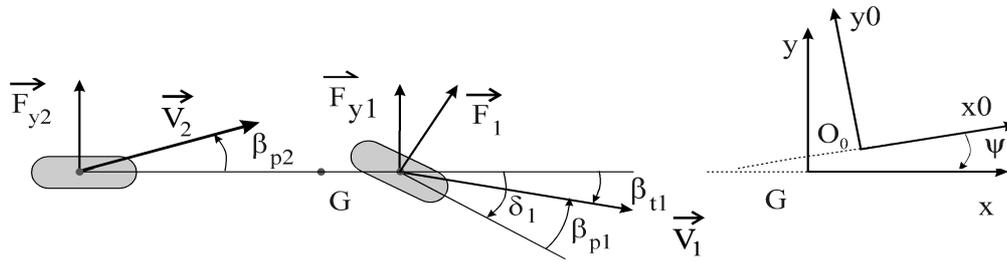


figure 4 - 11 : expression de la dérive au pneumatique

Les caractéristiques géométriques et la technologie de construction des trains provoquent la naissance de braquages induits en présence de roulis. Ces braquages induits s'ajoutent à l'angle de braquage des roues. Au train arrière (roues non-directrices) il provoque un léger braquage des roues. Généralement, l'angle ajouté est supposé proportionnel à l'angle de roulis au centre de gravité du véhicule. Compte tenu de ces braquages induits (décrit par le biais des coefficients ε_1 et ε_2 respectivement pour le train avant et le train arrière), la dérive au pneu avant s'exprime :

$$\beta_{p1} = \beta + \frac{h_1 \dot{\theta}}{V} + \frac{l_1 \dot{\psi}}{V} - (\delta_1 + \varepsilon_1 \theta) \quad (4.3-17)$$

La longueur de relaxation des pneus doit maintenant être prise en compte. Le point de contact pneumatique/sol se déplace latéralement avec une certaine vitesse par rapport à la caisse. Ce déplacement est essentiellement causé par le phénomène de ballant. En supposant la rigidité de dérive indépendante du temps, cette vitesse transversale de déplacement vaut :

$$\dot{d} = b_1 \dot{\beta}_{p1} \quad (4.3-18)$$

En tenant compte de cette vitesse de déplacement transversal, la dérive au pneu avant s'exprime :

$$\beta_{p1} = \beta + \frac{h_1 \dot{\theta}}{V} + \frac{l_1 \dot{\psi}}{V} + \frac{b_1 \dot{\beta}_{p1}}{V} - (\delta_1 + \varepsilon_1 \theta) \quad (4.3-19)$$

D'où l'expression de la dérivée de la dérive au pneu avant :

$$b_1 \dot{\beta}_{p1} = V \beta - V \beta_{p1} + h_1 \dot{\theta} + l_1 \dot{\psi} - V \delta_1 - V \varepsilon_1 \theta \quad (4.3-20)$$

De manière similaire, l'expression de dérivée de la dérive au pneu arrière conduit à :

$$b_2 \dot{\beta}_{p2} = V \beta - V \beta_{p2} + h_2 \dot{\theta} - l_2 \dot{\psi} - V \varepsilon_2 \theta \quad (4.3-21)$$

3.4.1.4 Egalisation, matrices d'états

L'égalisation des premiers et second membres des équations du mouvement du véhicule (eq (4.3-7), (4.3-8), (4.3-9), (4.3-10), (4.3-20) et (4.3-21)), permet de décrire les mouvements **de lacet, de dérive, de roulis et de ballant** du véhicule par le biais d'un unique modèle linéaire sous forme de matrice d'état. Récapitulons les différentes équations nécessaires :

$$\dot{\beta}_{p1} : b_1 \dot{\beta}_{p1} = V \beta - V \beta_{p1} + h_1 \dot{\theta} + l_1 \dot{\psi} - V \delta_1 - V \varepsilon_1 \theta \quad (4.3-22)$$

$$\dot{\beta}_{p2} : b_2 \dot{\beta}_{p2} = V \beta - V \beta_{p2} + h_2 \dot{\theta} - l_2 \dot{\psi} - V \varepsilon_2 \theta \quad (4.3-23)$$

$$M \gamma_t : M V (\dot{\beta} + \dot{\psi}) - m s h_0 \ddot{\theta} = -D_1 \beta_{p1} - D_2 \beta_{p2} \quad (4.3-24)$$

$$L_x : I_r \ddot{\theta} - I_{xz} \ddot{\psi} - m s h_0 V (\dot{\beta} + \dot{\psi}) = -k_r \theta - a_r \dot{\theta} \quad (4.3-25)$$

$$N_z : ms h_0 V \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} = -I_1 D_1 \beta_{p1} + I_2 D_2 \beta_{p2} \quad (4.3-26)$$

La mise sous forme de matrices d'état s'obtient en passant par la forme intermédiaire :

$$\begin{aligned} G \dot{X}(t) &= H X(t) + N U(t) \\ Y(t) &= C X(t) + D U(t) \end{aligned} \quad (4.3-27)$$

La formulation (4.3-27) est ensuite mise sous la forme plus classique :

$$\begin{aligned} \dot{X}(t) &= A X(t) + B U(t) \\ Y(t) &= C X(t) + D U(t) \end{aligned} \quad (4.3-28)$$

Le calcul des matrices A et B requiert l'inversion de la matrice G. Pour des raisons évidentes de simplicité de calcul (inversion symbolique d'une matrice 6x6), la matrice G est inversée numériquement, une fois que des valeurs sont associées aux paramètres. Les différents vecteurs et matrices de l'expression (4.3-27) s'expriment donc :

$$X(t) = (\beta_{p1} \ \beta_{p2} \ \beta \ \dot{\psi} \ \theta \ \dot{\theta})^T \quad (4.3-29)$$

$$U(t) = \alpha_v \quad (4.3-30)$$

$$N = (0 \ 0 \ -V/dem \ 0 \ 0 \ 0)^T \quad (4.3-31)$$

$$G = \begin{bmatrix} 0 & 0 & 0 & I_{zz} & ms h_0 V & -I_{xz} \\ 0 & 0 & MV & 0 & 0 & -ms h_0 \\ b_1 & 0 & 0 & 0 & -h_1 & 0 \\ 0 & b_2 & 0 & 0 & -h_2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -ms h_0 V & -I_{xz} & 0 & I_{xx} + ms h_0^2 \end{bmatrix} \quad (4.3-32)$$

$$H = \begin{bmatrix} -I_1 D_1 & I_2 D_2 & 0 & 0 & 0 & 0 \\ D_1 & D_2 & 0 & -MV & 0 & 0 \\ -V & 0 & V & I_1 & -V \varepsilon_1 & 0 \\ 0 & -V & V & -I_2 & -V \varepsilon_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & ms h_0 V & -kro & -aro \end{bmatrix} \quad (4.3-33)$$

Le modèle linéaire ainsi obtenu décrit les mouvements de lacet/dérive, de roulis, et de ballant d'un véhicule automobile à vitesse constante. La description de nombreux autres phénomènes aurait pu être ajoutée au modèle. Toutefois, les propriétés d'identifiabilité des paramètres sont le principal frein à l'utilisation de modèles de complexité plus élevée. De plus, dans le paragraphe 6.3 il sera montré que l'identification des paramètres inconnus ou difficilement mesurables de ce modèle à partir des variables enregistrées sur le véhicule pose déjà de sérieux problème d'identifiabilité.

3.4.2 Dynamiques mises en jeu

Avant de décrire brièvement les autres modèles de connaissance étudiés, penchons-nous sur les principales dynamiques mises en jeu dans le modèle LaRouDéBal. Le tableau suivant résume succinctement les fréquences de coupures des différents phénomènes régissant le comportement dynamique transversal d'un véhicule automobile de type berline [DIX96], [MIL95].

Le tableau 4 - 1, établi à partir de [REN93], montre que les dynamiques mises en jeu par les phénomènes de lacet/dérive et de roulis sont proches. Les mouvements de lacet et de dérive sont associés et mettent en jeu les mêmes dynamiques. Ainsi, en fonction des phénomènes modélisés

sur les différents modèles considérés dans notre étude, il est probable que la non-modélisation du mouvement de roulis (dans le cas des modèles LaDéBal et LaDé) entraîne une estimation biaisée des paramètres physiques en relation avec les mouvements de lacet et de dérive.

Phénomènes	Fréquences de coupures associées
Lacet/Dérive	0.3 – 1.3 Hz
Roulis	0.3 – 3 Hz
Ballant	2.0 – 4.0 Hz

tableau 4 - 1 : dynamiques du comportement transversal d'un véhicule automobile de type berline

Le phénomène de ballant présente quant à lui, une fréquence de coupure sensiblement supérieure au mouvement de lacet/dérive. Ainsi, lors de l'identification des paramètres physiques en relation avec le ballant, un phénomène de dominance des dynamiques de ballant par celles des mouvements de lacet et de dérive est possible. Cette dominance entraînerait une perte d'identifiabilité des paramètres physiques associés au mouvement ballant (i.e. les longueurs de relaxation des pneumatiques).

Avant d'aborder les propriétés structurelles de ce modèle, décrivons brièvement les modèles simplifiés LaRouDé, LaDéBal et LaDé qui ont également été étudiés.

3.5 Modèle LaRouDé

Les équations d'états régissant les phénomènes de lacet/dérive et de roulis se déduisent du modèle précédent. Les angles de braquages induits par le roulis sont ici négligés, dans la suite du mémoire (analyse de la sensibilité des modèles § 4.1) leurs influences mineures sur le comportement du véhicule seront montrées.

$$\dot{\beta}_{p1} : V \dot{\beta}_{p1} = V \beta + h_1 \dot{\theta} + I_1 \dot{\psi} - V \delta_1 \quad (4.3-34)$$

$$\dot{\beta}_{p2} : V \dot{\beta}_{p2} = V \beta - h_2 \dot{\theta} - I_2 \dot{\psi} \quad (4.3-35)$$

$$M \gamma_t : M V (\dot{\beta} + \dot{\psi}) - ms h_0 \ddot{\theta} = -D_1 \dot{\beta}_{p1} - D_2 \dot{\beta}_{p2} \quad (4.3-36)$$

$$L_x : I_r \ddot{\theta} - I_{xz} \ddot{\psi} - ms h_0 V (\dot{\beta} + \dot{\psi}) = -k_r \theta - a_r \dot{\theta} \quad (4.3-37)$$

$$N_z : ms h_0 V \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} = -I_1 D_1 \dot{\beta}_{p1} + I_2 D_2 \dot{\beta}_{p2} \quad (4.3-38)$$

La non-description du phénomène de ballant le rend inapte à décrire les dynamiques transversales rapides des pneumatiques. L'étude précédente du lieu des pôles du modèle LaRouDéBal a montré que ces dynamiques se situent théoriquement entre 2.9 Hz et 4.1 Hz pour une vitesse longitudinale du véhicule entre 20 km/h et 120 km/h.

Afin d'éviter dans une certaine mesure une estimation biaisée des paramètres, l'identification des paramètres physiques s'effectuera sur des données expérimentales ou le phénomène de ballant possède une influence mineure (fréquence de l'excitation maintenue inférieure à 1.5 Hz).

3.6 Modèle LaDéBal

Les équations d'états, régissant les phénomènes de lacet/dérive et de ballant se déduisent du modèle LaRouDéBal.

$$\dot{\beta}_{p1} : b_1 \dot{\beta}_{p1} = V \beta + I_1 \dot{\psi} - V \dot{\beta}_{p1} - V \delta_1 \quad (4.3-39)$$

$$\dot{\beta}_{p2} : b_2 \dot{\beta}_{p2} = V \beta - I_2 \dot{\psi} - V \dot{\beta}_{p2} \quad (4.3-40)$$

$$M \gamma_t : M V (\dot{\beta} + \dot{\psi}) = -D_1 \beta_{p1} - D_2 \beta_{p2} \quad (4.3-41)$$

$$N_z : I_{zz} \ddot{\psi} = -I_1 D_1 \beta_{p1} + I_2 D_2 \beta_{p2} \quad (4.3-42)$$

La modélisation du phénomène de ballant permet de décrire les dynamiques de fréquences élevées, tandis que la modélisation du phénomène de lacet/dérive décrit les dynamiques plus lentes. Or les dynamiques du mouvement de roulis sont proches de celles de lacet/dérive et de ballant. Afin d'obtenir un comportement du modèle identique à celui du système, il est probable que le processus d'identification assimilera les dynamiques propres au roulis aux mouvements de lacet/dérive et de ballant. Suivant l'importance du mouvement de roulis non-modélisé sur le comportement global du véhicule, un biais d'estimation sur les paramètres identifiés sera donc présent.

Dans la majorité des travaux consacrés à l'étude du comportement du véhicule faiblement sollicité, les phénomènes de roulis et de ballant sont négligés. Ces travaux reposent essentiellement sur le modèle LaDé. Outre la sensibilité du modèle vis à vis des paramètres et les propriétés d'identifiabilité du modèle, la comparaison relative des valeurs des paramètres identifiés par l'intermédiaire des modèles LaDéBal et LaRouDéBal mettra en évidence l'influence de la prise en compte du mouvement de roulis et/ou de ballant sur l'identification des paramètres physiques du véhicule.

Les données expérimentales employées pour l'identification des paramètres physiques de ce modèle possèdent les mêmes caractéristiques que celles employées pour l'identification du modèle LaRouDéBal (le signal d'excitation appliqué au volant est un signal sinusoïdal vobulé en fréquence entre les fréquences de 0.2 et 2.2 Hz)

3.7 Modèle LaDé

Les équations d'états régissant les phénomènes de lacet/dérive et de ballant se déduisent du modèle LaRouDéBal.

$$M \gamma_t : M V (\dot{\beta} + \dot{\psi}) = -D_1 (\beta + (I_1 / V) \dot{\psi} - \delta_1) - D_2 (\beta - (I_2 / V) \dot{\psi}) \quad (4.3-43)$$

$$N_z : I_{zz} \ddot{\psi} = (I_2 D_2 - I_1 D_1) \beta + (I_2^2 D_2 - I_1^2 D_1) \dot{\psi} / V + I_1 D_1 \delta_1 \quad (4.3-44)$$

Ce modèle très répandu dans la littérature décrit essentiellement les dynamiques basses fréquences du comportement latéral du véhicule. L'inaptitude de ce modèle à décrire les dynamiques rapides (effet de ballant) diminue significativement son pouvoir prédictif. L'extrême simplicité de ce modèle ainsi que l'absence de description des mouvements de roulis et de ballant sont autant de sources potentielles de présence de biais d'estimation sur les paramètres identifiés.

Afin d'éviter dans la mesure du possible une estimation biaisée des paramètres, l'identification des paramètres physiques s'effectuera sur des données expérimentales ou le phénomène de ballant possède une influence mineure (fréquence de l'excitation maintenue inférieure à 1.5 Hz).

L'étude qualitative par l'examen de la position des pôles, dans le plan de Laplace du modèle LaRouDéBal, en fonction de la vitesse, laisse entrevoir des difficultés pour l'identification des paramètres physiques d'un véhicule automobile. Les études de sensibilité, d'identifiabilité et de discernabilité apportent des éléments d'information supplémentaires.

4 PROPRIETES STRUCTURELLES DES MODELES

Avant de débiter l'identification des différents paramètres, l'étude des propriétés structurelles des modèles définit les paramètres physiques que le modélisateur est légitimement en droit d'estimer. Plutôt que concurrentes les études de sensibilité et d'identifiabilité sont complémentaires.

Par.	Désignation	Valeurs	Unité	Ecart type (%)
D1	Rigidité de dérive train avant	84085	N/rad	30
D2	Rigidité de dérive train arrière	87342	N/rad	30
M	Masse totale du véhicule	1400	kg	0.1
mns	Masse non-suspendue	138	kg	1
ms	Masse suspendue, $ms = M - mns$	1262	kg	0.1
lzz	Inertie de lacet de l'ensemble du véhicule	1571	kg.m ²	20
lxx	Inertie de roulis de la masse suspendue	252	kg.m ²	50
lxz	Produits d'inertie roulis – lacet	50	kg.m ²	100
aro	Amortissement antiroulis de la voiture	2750	Nm/rad/s	40
kro	Raideur en roulis de la voiture	150000	N/rad	15
b1	Longueur de relaxation du train avant	0.5	m	60
b2	Longueur de relaxation du train arrière	0.5	m	60
hg	Hauteur du centre de gravité/sol	0.535	m	10
h1	Hauteur axe de roulis de l'essieu avant/sol	0.0903	m	20
h2	Hauteur axe de roulis de l'essieu arrière/sol	0.00	m	non précisé
ε_1	Braquage induit par le roulis au train avant	-0.0269	---	non précisé
ε_2	Braquage induit par le roulis au train arrière	0.02	---	80

tableau 4 - 2 : paramètres étudiés et intervalles d'incertitude constructeur associés

Dans les cas où des choix sont à effectuer parmi plusieurs paramètres non-identifiables, l'étude de la sensibilité permet d'isoler les paramètres les moins influents sur le comportement du modèle de véhicule. Sans entraîner de dramatiques conséquences sur les paramètres estimés, les valeurs de ces paramètres d'influence mineure sur les sorties du modèle, peuvent alors être fixées à des valeurs typiques (valeurs constructeur ou a priori).

Les paramètres du véhicule étudiés sont ceux dont les valeurs sont inconnues, imprécisément connues, difficilement mesurables ou sujets à d'importantes variations. Ces paramètres, leurs valeurs nominales et les écarts types fournis par le constructeur automobile sont listés dans le tableau 4 - 2. A ces paramètres, les valeurs des masses aisément mesurables par pesées, ont été ajoutées. Cependant, à des fins de contrôle adaptatif, la détermination des masses, variables en fonction du chargement, pourrait être recherchée.

4.1 Sensibilité

L'étude de la sensibilité des différents modèles est réalisée à partir des modèles théoriques paramétrés avec les valeurs nominales des paramètres (données du constructeur) du véhicule d'essais du laboratoire et une **vitesse du véhicule de 80 km/h**.

Le tracé des fonctions de sensibilité logarithmique induit un nombre considérable de courbes. A titre d'exemple, seules les fonctions de sensibilité logarithmique du premier ordre des paramètres du modèle LaRouDéBal sont reproduites dans ce mémoire. L'examen des fonctions de sensibilité des autres modèles conduit à des conclusions identiques. Seules les fonctions de sensibilité des variables entrant dans la formulation du critère de coût (vitesse de lacet et accélération transversale) sont reproduites sur la figure 4 - 12. Les tracés des fonctions de sensibilité des variables de dérive, et de vitesse de roulis sont reproduits en annexe 5. En effet, pour le protocole défini (cf. § 5 de ce chapitre), l'amplitude de l'angle de dérive est faible (inférieure à 2°) ; le signal d'angle de dérive est noyé dans le bruit et est, par conséquent, inexploitable. Par ailleurs, la variable de vitesse de roulis n'est pas relevée sur le véhicule.

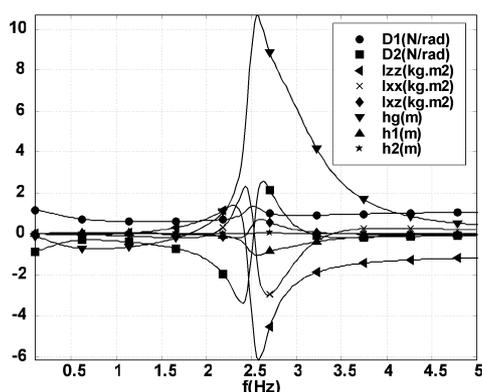


figure 4 - 12a : vitesse de lacet

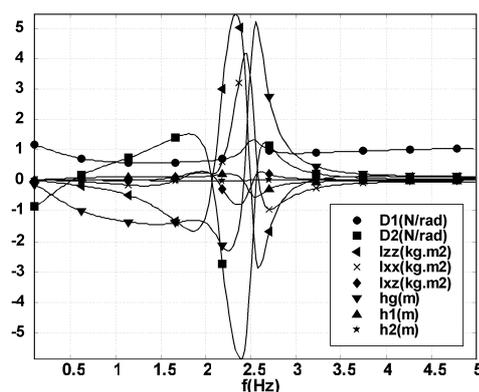


figure 4 - 12b : accélération transversale

figure 4 - 12 : magnitude des fonctions de sensibilité logarithmique du premier ordre de la vitesse de lacet et de l'accélération transversale

L'examen des fonctions de sensibilité est réalisé sur l'intervalle fréquentiel [0.1 .. 5 Hz]. La fréquence de 5 Hz est approximativement la fréquence maximale d'excitation applicable au volant par le biais d'un opérateur humain. L'étude des différentes fonctions de sensibilité suscite les réflexions suivantes :

☐ **Angle de dérive**

Les hauteurs des centres de roulis avant et arrière (h_1 et h_2), le produit d'inertie l_{xz} , les longueurs de relaxation des trains avant et arrière (b_1 et b_2) ainsi que les coefficients de braquage induit par le roulis (ε_1 et ε_2) ont une faible influence sur l'angle de dérive du véhicule.

☐ **Vitesse de lacet**

La hauteur du centre de roulis arrière (h_2) et les coefficients de braquage induit par le roulis (ε_1 et ε_2) possèdent une influence négligeable sur la vitesse de lacet. Dans une moindre mesure, la hauteur du centre de roulis avant (h_1) et le produit d'inertie (l_{xz}) exercent une faible influence sur la vitesse de lacet. La vitesse de lacet se révèle peu sensible à la longueur de relaxation du train avant pour les fréquences inférieures à 2.5 Hz. La sensibilité de la vitesse de lacet vis à vis de la longueur de relaxation avant croit pour les fréquences supérieures à 2.5 Hz, mais demeure toutefois plus faible par rapport aux autres paramètres.

☐ **Accélération transversale**

L'influence des coefficients de braquages induits par le roulis (ε_1 et ε_2) et la hauteur du centre de roulis arrière (h_2) sont négligeables. Dans une moindre mesure, la hauteur du centre de roulis avant (h_1) est faible. De manière similaire à la vitesse de lacet, l'influence de la longueur de relaxation avant (b_1) est négligeable aux fréquences inférieures à 2.5 Hz. Bien que d'influence plus élevée aux fréquences plus hautes, relativement aux autres paramètres, son influence demeure faible. Relativement aux autres paramètres, le produit d'inertie lacet/roulis (I_{xz}) influence peu la variable d'accélération transversale.

□ Vitesse de roulis

Les conclusions relatives à la vitesse de roulis sont en tout point identiques à celles formulées au sujet de l'accélération transversale. Ceci s'explique par le fait qu'en première approximation, il est communément admis dans le monde automobile, que pour de faibles sollicitations, l'angle de roulis est proportionnel à l'accélération transversale²⁵.

□ Influence de la vitesse

La valeur de la vitesse longitudinale du véhicule modifie significativement les fonctions de sensibilité vis à vis des paramètres. La figure 4 - 13 illustre ce phénomène. Sur cette figure sont représentées les fonctions de sensibilité logarithmique de l'accélération transversale vis à vis de certains paramètres, pour des vitesses de 20 km/h et 80 km/h. Malgré la forte dépendance des tracés des fonctions de sensibilité vis à vis de la vitesse, les constatations précédentes quant à l'influence mineure de certains paramètres demeurent valides pour des vitesses comprises entre 20 km/h et 120 km/h. Une diminution de la vitesse entraîne essentiellement :

- Un étalement des variations des fonctions de sensibilité à vitesse plus élevée, sur une gamme de fréquences plus large.
- Une diminution de l'amplitude des fonctions de sensibilité pour certains paramètres intervenant sous forme de produit avec la vitesse du véhicule, dans les différentes fonctions de transfert.

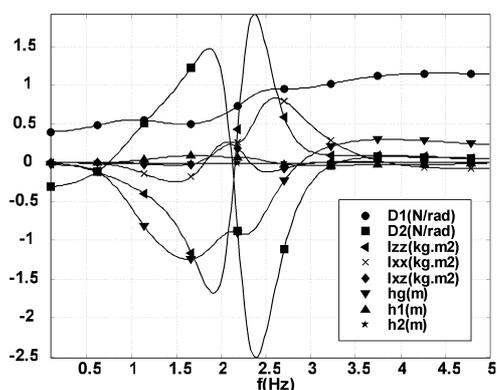


figure 4 - 13a : vitesse = 20 km/h

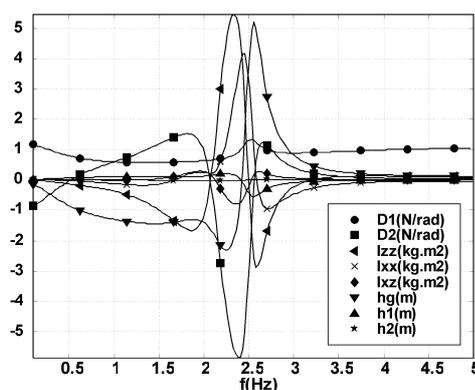


figure 4 - 13b : vitesse = 80 km/h

figure 4 - 13 : influence de la vitesse sur les magnitudes des fonctions de sensibilité logarithmique du premier ordre de l'accélération transversale

Des points communs se dégagent des fonctions de sensibilité des variables de lacet et d'accélération transversale. Les coefficients de braquages induits par le roulis (ε_1 et ε_2) et la hauteur du centre de roulis arrière n'affecte pas significativement le comportement du véhicule en vitesse de

²⁵ Le coefficient liant l'angle de roulis à l'accélération transversale est nommé roulis spécifique par les constructeurs automobiles.

lacet et en accélération transversale. Dans une moindre mesure, l'influence de la valeur du produit d'inertie (I_{xz}), de la hauteur du centre de roulis avant (h_1) et de la longueur de relaxation arrière (b_1) est faible relativement aux autres paramètres étudiés.

Les conclusions relatives à l'influence respective des différents paramètres sur les variables de sortie motiveront un éventuel arbitrage quant aux paramètres identifiés, dans le cas où l'identification simultanée de tous les paramètres rendrait le modèle non-identifiable.

4.2 Identifiabilité

L'identifiabilité des différents modèles de connaissance est abordée. Les propriétés d'identifiabilité des différents modèles est vérifiée par l'approche numérique décrite dans le premier chapitre. L'approche par transformée de Laplace nécessite l'emploi de logiciel de calcul formel (Maple[©] par exemple). La puissance de résolution de tels logiciels est insuffisante pour exprimer le modèle LaRouDéBal sous une forme littérale de la forme

$$\begin{aligned}\dot{X}(t) &= A X(t) + B U(t) \\ Y(t) &= C X(t) + D U(t)\end{aligned}\tag{4.4-1}$$

L'approche symbolique nécessite le calcul de la fonction de transfert formelle $Y(s)/U(s)$ qui est alors irréalisable par le biais de tels logiciels. L'approche numérique consiste à [WAL97] :

- Choisir une valeur nominale des paramètres θ_0 et à simuler un grand nombre de données fictives y .
- Estimer $\hat{\theta}$ à partir des données précédentes par minimisation d'un critère de coût quadratique.

Dans le cas présent, le critère de coût multivariable inclut les variables de vitesse de lacet et d'accélération transversale (variables fiables relevées sur le véhicule). La pondération relative de ces variables dans le critère de coût est définie par l'inverse de la matrice de covariance du bruit de mesure expérimental (maximum de vraisemblance). L'excitation (l'angle volant) appliquée aux modèles est de type SBPA²⁶. La vitesse du véhicule est maintenue constante à 80 km/h

4.2.1 Modèle LaRouDéBal

Les résultats de l'étude de l'identifiabilité structurelle locale du modèle en fonction des paramètres recherchés sont reproduits dans le tableau 4 - 3. En premier lieu, l'identifiabilité des paramètres influents dont les valeurs sont connues de manière imprécise est étudiée. Chaque colonne du tableau comporte les paramètres identifiés marqués d'une croix. Les dernières lignes contiennent les conclusions sur l'identifiabilité du modèle considéré. Le modèle est identifiable oui (O)/non (N).

Des appréciations supplémentaires sous forme de notes sont ajoutées lorsque les intervalles d'incertitude associés de certains paramètres identifiés les rendent inexploitable : sur les principaux paramètres, lorsque les amplitudes des écarts types²⁷ sont supérieures à 10 % de la valeur nominale des paramètres, les intervalles de confiance à 95 % sur les sorties prédites deviennent inacceptables (les intervalles de confiance sur les sorties sont de l'ordre de l'amplitude du signal de sortie). Dans ce cas de figure, le processus d'identification n'est pas jugé viable. De tels résultats, dans un contexte théorique idéalisé, en l'absence d'erreur structurelle et de bruit de mesure et où

²⁶ Période d'échantillonnage = 20 ms, temps élémentaire (durée d'un bit) = 140 ms, nombre de bit = 5, amplitude = $\pm 25^\circ$, nombre de périodes = 10.

²⁷ Les écarts types sont estimés à partir de l'inégalité de Cramer-Rao.

l'excitation est parfaitement maîtrisée, offrent peu de perspectives dans des conditions expérimentales réelles.

Le tableau 4 - 3 montre que les **7 principaux paramètres** (D1, D2, lzz, lxx, kro, aro et b2) **sont identifiables simultanément** à l'exception de la hauteur h_g présente entre l'axe de roulis et le centre de gravité du véhicule. L'identification de ce paramètre entraîne une dégradation des intervalles de confiance sur certains paramètres (de l'ordre de 20 % sur la raideur et l'amortissement de roulis, kro et aro) et sur les sorties prédites par le modèle identifié.

Para.	Paramètres identifiés																			
D1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
D2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
lzz	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
lxx	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
lxz	X	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X		
h_1					X		X		X					X		X	X			
h_2						X		X		X					X	X		X	X	
M		X																		
ms				X																
h_g			X		X	X	X	X	X										X	
kro	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
aro	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
ε_1											X		X	X	X	X			X	
ε_2												X	X	X	X	X				
b1	X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X	X	
b2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Iden.	O	N	O	N	O	O	O	N	O	O	O	O	O	O	O	O	O	O	O	
Note	-	-	1	-	1	1	1	-	2	2	-	-	-	2	2	1	2	-	-	1

tableau 4 - 3 : étude de l'identifiabilité des paramètres du modèle LaRouDéBal

Notes :

1 : les intervalles de confiance associés aux paramètres estimés (avec un degré de confiance de 68%), sont d'amplitudes très élevées (plus de 100%). Les résultats de l'identification sont alors jugés inexploitable.

2 : les intervalles de confiance sur les sorties prédites par le modèle identifié sont d'amplitudes élevées (de l'ordre de l'amplitude des variables de sortie). Le pouvoir prédictif du modèle identifié est jugé inacceptable.

Ces paramètres ont été retenus en raison de leurs influences non-négligeables sur le comportement latéral du véhicule. De plus les valeurs de ces paramètres sont généralement connues avec une précision limitée (cf. tableau 4 - 2).

Aux **7 paramètres majeurs** précédents, **4 paramètres mineurs** (le produit d'inertie lxz , la longueur de relaxation avant $b1$, et les coefficients de braquage induit par le roulis ε_1 , ε_2) supplémentaires sont identifiables. L'identification simultanée de la hauteur du centre de roulis arrière avec les 7 paramètres précédents induit des intervalles d'incertitude inacceptables sur les paramètres tels que la raideur et l'amortissement de roulis (100 à 200 % de la valeur nominale).

Théoriquement, **11 paramètres** ($D1$, $D2$, I_{zz} , I_{xx} , I_{xz} , kro , aro , ϵ_1 , ϵ_2 , $b1$ et $b2$) **peuvent être identifiés**. Un nombre de paramètres plus élevé (13 au total) est cependant identifiable ; les intervalles de confiance (de l'ordre de 100% à 200%) associés à certains paramètres rendent cependant ces résultats inexploitable en pratique.

Pour résumer, dans le cadre théorique idéalisé et le protocole idéal défini (signal de type SPBA), du point de vue du modéliste et de l'importance relative des différents paramètres, le cas de figure le plus intéressant (au sens de leurs influences sur le comportement du véhicule et au sens des incertitudes associées aux données constructeurs) concerne l'identification des 9 paramètres suivants : **$D1$, $D2$, I_{zz} , I_{xx} , I_{xz} , kro , aro , $b1$ et $b2$** . Dans la mesure du possible, l'identification à partir de données expérimentales de ces paramètres sera recherchée.

Cependant, rien ne garantit la faisabilité de l'identification de ces paramètres dans le cadre expérimental réel où :

- L'excitation du véhicule n'est pas maîtrisable à 100 % (vitesse constante approximative maintenue par le pilote, excitation harmonique approximative ...).
- Les données expérimentales sont entachées de bruit de mesure.
- La présence d'erreurs structurelles même minimales est inévitable.

Avant d'aborder la partie expérimentale de ce travail, l'identifiabilité des modèles LaRouDé, LaDéBal et LaDé est traitée.

4.2.2 Modèle LaRouDé

Mis à part les conclusions relatives aux longueurs de relaxation avant et arrière ($b1$ et $b2$), les propriétés d'identifiabilité du modèle LaRouDé sont identiques au modèle LaRouDéBal.

Ce modèle permet théoriquement l'identification simultanée des deux rigidités de dérive avant et arrière ($D1$ et $D2$), des inerties de lacet et de roulis (I_{zz} et I_{xx}), du produit d'inertie lacet/roulis (I_{xz}), de la distance entre le centre de gravité et l'axe de roulis (hg), de la raideur et de l'amortissement de roulis (kro et aro).

Cependant, l'identification simultanée de la distance entre le centre de gravité et l'axe de roulis (hg) avec les paramètres précédents induit des intervalles d'incertitude inacceptables sur certains paramètres tels que la raideur et l'amortissement de roulis (kro et aro), l'inertie de roulis (I_{xx}), le produit d'inertie lacet/dérive (I_{xz}), et la hauteur hg . Pour cette raison l'identification des paramètres suivants est souhaitable **$D1$, $D2$, I_{zz} , I_{xx} , I_{xz} , kro et aro** .

4.2.3 Modèle LaDéBal

Les principaux paramètres difficilement mesurables (les deux rigidités de dérive ($D1$ et $D2$), l'inertie de lacet (I_{zz}) ainsi que les deux longueurs de relaxation avant et arrière ($b1$ et $b2$)) nécessaire à la paramétrisation de ce modèle sont identifiables simultanément. L'identification des paramètres suivant sera donc recherchée : **$D1$, $D2$, I_{zz} , $b1$ et $b2$** . A des fins de contrôle adaptatif, une identification de la masse (M) peut être utile. Son identification simultanée avec les paramètres précédents rend le modèle non-identifiable.

4.2.4 Modèle LaDé

Les deux rigidités de dérives (**$D1$ et $D2$**) et l'inertie de lacet (**I_{zz}**) sont identifiables simultanément. L'identification de la masse (M) rend le modèle non-identifiable. Après avoir étudié l'identifiabilité structurelle locale des différents modèles, considérons leurs propriétés de discernabilité.

4.3 Discernabilité

Les modèles LaDé, LaDéBal et LaRouDé sont issus de simplifications du modèle LaRouDéBal. En d'autres termes, **le modèle LaRouDéBal est un sur-ensemble des trois autres modèles**. Pour ces raisons, le modèle LaRouDéBal est non-discernable des modèles LaDé, LaDéBal et LaRouDé. Un raisonnement identique peut être tenu pour les autres modèles. Ainsi, le modèle LaRouDé et LaDéBal sont structurellement non-discernables du modèle LaDé.

Donc, à ce stade, il est donc déjà possible d'affirmer que les modèles

- LaDé et LaRouDé sont structurellement indiscernables.
- LaDé et LaDéBal sont structurellement indiscernables.
- LaRouDé et LaRouDéBal sont structurellement indiscernables.
- LaDéBal et LaRouDéBal sont structurellement indiscernables.

Il reste donc à déterminer les propriétés de discernabilité structurelle respective des modèles LaRouDé et LaDéBal. Pour cela, l'approche numérique est alors à nouveau employée. Des données fictives sont simulées à l'aide des deux modèles. Puis les paramètres du modèle LaRouDé sont identifiés à partir des données fictives simulées par l'intermédiaire du modèle LaDéBal (et réciproquement). S'il est impossible d'ajuster les paramètres du modèle LaDéBal pour obtenir une identité de comportement avec le modèle LaRouDé, alors le modèle LaDéBal sera structurellement discernable du modèle LaRouDé. Si la réciproque se vérifie, les modèles seront dits structurellement discernables.

Quels que soient les paramètres identifiés, les modèles LaRouDé et LaDéBal sont structurellement discernables. Les propriétés de discernabilité des différents modèles permettront d'orienter le choix final du modélisateur par rapport au pouvoir prédictif des modèles. Toutefois, dans le cas présent, le choix du modélisateur sera essentiellement motivé par :

- Les valeurs des paramètres estimés.
- Les intervalles d'incertitude associés aux paramètres estimés.
- Le pouvoir prédictif du modèle.
- Les différentes caractéristiques des résidus par rapport aux hypothèses formulées sur leur nature.

La discernabilité constitue donc un élément supplémentaire à la disposition du modélisateur pour effectuer un choix du modèle final parmi tous les modèles pressentis.

4.4 Conclusion

Dans cette partie consacrée à l'étude des propriétés structurelles des modèles de véhicule automobile établi il a été montré que l'identification des paramètres physiques de véhicule automobile n'est pas triviale et pose de réels problèmes d'identifiabilité et de discernabilité.

Tous les paramètres inconnus ou difficilement mesurables ne sont pas identifiables simultanément : un choix doit donc être effectué. L'analyse de la sensibilité permet dans une certaine mesure d'éliminer les paramètres d'influences « mineures » et les paramètres dominants le comportement du véhicule, peuvent alors être retenus en vue de leur identification.

Compte tenu de ces limitations, un compromis sur les paramètres identifiés doit alors être réalisé. De plus il est probable que dans le cadre expérimental réel (présence de bruit de mesure, présence d'erreur de modélisation, limitations inhérentes au protocole expérimental ...) que d'autres difficultés apparaissent. Aussi avant d'entreprendre l'identification des paramètres à partir de données expérimentales, les différents aspects relatifs au protocole expérimental sont abordés.

5 PROTOCOLE EXPERIMENTAL

Les aspects relatifs au signal d'excitation appliqué au véhicule, au choix de la fréquence d'échantillonnage des variables relevées sur le véhicule et au filtrage des données sont successivement abordés. Enfin, en utilisant la fonction de cohérence entre les variables d'entrée et de sortie, le caractère linéaire en l'entrée (LE) du véhicule sera évalué.

5.1 Excitation

Les variables d'entrée du véhicule sont au nombre de deux : la vitesse longitudinale et l'angle appliqué au volant. Au paragraphe 4.1 de ce chapitre, il a été montré qu'une diminution de la vitesse induisait un étalement des courbes de sensibilité dans le domaine fréquentiel. Or la génération d'un signal d'excitation (au volant) de fréquence supérieure à 4 Hz, par le biais d'un opérateur humain est difficilement concevable. Ainsi, pour concentrer l'influence des paramètres recherchés dans une gamme de fréquence inférieure à 4 Hz, la vitesse du véhicule est fixée à 80 km/h.

Tout en maintenant, dans la mesure du possible, la vitesse du véhicule constante, le pilote exerce une excitation au volant, définie par le protocole expérimental. Idéalement, l'excitation appliquée au volant doit minimiser les intervalles d'incertitude associés aux paramètres estimés. Ce type d'excitation se définit par l'intermédiaire des techniques de planification d'expérience optimale. Or, ces techniques nécessitent généralement une connaissance exacte du modèle et des paramètres recherchés. De plus, elles ne permettent pas de mettre en évidence la présence de phénomènes non modélisés.

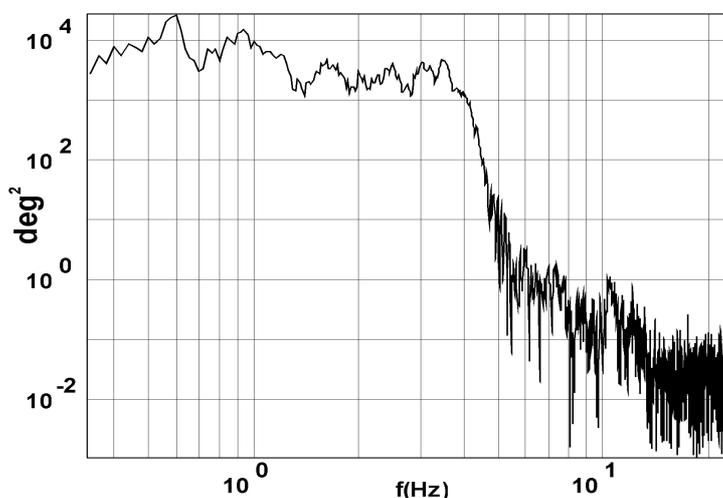


figure 4 - 14 : spectre de l'angle appliqué au volant (excitation harmonique)

Lorsque le modèle n'est pas précisément défini, l'excitation appliquée au volant est idéalement persistante, parfaitement reproductible et possède un spectre en accord avec les dynamiques du système identifié. Les signaux de type séquences binaires pseudo-aléatoires (SBPA) possèdent ces propriétés. Ces signaux ne sont pas utilisables dans le cas présent. Ils sont en effet irréalisables manuellement et pour des raisons évidentes de sécurité leurs générations par le biais d'effecteurs couplés au volant n'est pas souhaitable.

Pour l'identification de modèles comportementaux de véhicules routiers, Mimuro [MIM90] et Xia [XIA95] préconisent l'emploi de sollicitation impulsionnelle (« dirac » d'angle volant d'amplitude de

l'ordre de 70° et de 0.4 s de durée). Les spectres résultants s'étendent des basses fréquences jusqu'à environ 2 Hz. De plus, ce type de sollicitation est discutable du point de vue de la sécurité.

D'autres auteurs ([HEY90], [WAD86]) utilisent une excitation de type harmonique. Ce type de sollicitation possède d'intéressantes propriétés. Par nature, elles sensibilisent les modes de résonance et d'antirésonance du véhicule. Une identification précise des paramètres en rapport avec ces modes est alors possible (par exemple les longueurs de relaxation des pneumatiques en rapport avec l'antirésonance présente sur la fonction de transfert liant l'accélération transversale à l'angle volant). Ce type d'excitation réclame cependant une durée d'expérimentation plus longue et une dextérité certaine de la part du pilote. Le spectre d'excitation obtenu est quasiment plat entre les fréquences comprises entre 0.3 et 4 Hz (cf. figure 4 - 14).

5.2 Fréquence d'échantillonnage

Dans le cadre de l'identification de modèles comportementaux de processus, plusieurs règles empiriques sont énoncées dans la littérature. Ces règles se basent essentiellement sur la connaissance du temps de réponse de 5% à 95 % (t_r) du processus. La fréquence d'échantillonnage ($f_e = 1/T_e$) est alors définie en fonction de la valeur du rapport t_r/T_e .

Les règles énoncées dans la littérature se résument à 3 valeurs de rapport t_r/T_e

- $t_r / T_e \approx 10$ [LJU87], [SOD89].
- $t_r / T_e \in [30 \dots 50]$ [RIC91], [BOR92].
- $t_r / T_e > 30$ [FLA94].

Compte tenu de ces indications, un choix judicieux de la période d'échantillonnage n'est pas toujours évident. Richalet [RIC91] propose d'adapter la valeur du rapport t_r/T_e en fonction du système. Si la réponse du système ne présente pas d'oscillation, il suggère l'utilisation d'un rapport t_r/T_e de 30. En revanche, si la réponse du système présente des oscillations, il préconise plutôt un rapport de 50.

Du point de vue de l'identification des paramètres physiques, il convient d'adapter la période d'échantillonnage aux différentes dynamiques du système. Dans le cas présent, le véhicule possède des dynamiques lentes et rapides. La période d'échantillonnage adoptée ne doit pas masquer les dynamiques les plus rapides. En adoptant une valeur du rapport $t_r/T_e=10$ et en se basant sur des simulations (réponse indiciale) du modèle LaRouDéBal, une période d'échantillonnage de $T_e = 0.025s$ est appropriée à la description du mode de lacet/dérive (pôles dominants). Toutefois, la description des dynamiques mises en jeu par les dynamiques rapides de roulis et de ballant nécessiterait une période d'échantillonnage inférieure de l'ordre de $T_e=0.005 s$ ($f_e=200$ Hz). La fréquence d'échantillonnage est donc fixée à 200 Hz. Selon les règles énoncées par Richalet [RIC91], Borne [BOR92] et Flaus [FLA94] cette valeur est également appropriée à la description de la seule dynamique de lacet/dérive. Cette fréquence semble donc offrir un bon compromis.

Plus rigoureusement, et dans la mesure où les résultats sont applicables au véhicule automobile (augmentation de la durée d'expérimentation), il serait intéressant d'évaluer l'apport du choix d'une période d'échantillonnage optimale, basée sur l'expression de la matrice d'information de Fisher moyenne par échantillons [WAL97].

Les signaux obtenus par l'échantillonnage des variables continues sur le véhicule ne sont pas directement utilisables par les algorithmes d'identification de modèle paramétrique à temps continu. Outre le filtrage anti-repliement, les signaux nécessitent souvent une opération de prétraitement.

5.3 Prétraitement des signaux

Les signaux prélevés sur le véhicule peuvent être entachés de défauts. Ces défauts sont souvent la conséquence de

- Défaillance d'instrumentation (dérive, étalonnage...).
- Défaillance du système d'acquisition (parasite, bruit de mesure...).
- L'apparition de perturbations sur le processus (modifications de l'environnement, ...).

5.3.1 Offset et dérive des signaux

Fréquemment les signaux issus d'une chaîne d'acquisition subissent des dérives. Leurs éliminations sont possibles via l'utilisation de filtre passe-haut ou de régression linéaire. L'usage de filtre passe-haut de fréquence de coupure très basse se heurte au problème des conditions initiales du filtre. L'élimination de la dérive par l'intermédiaire de l'usage de régression linéaire [SOD89] s'affranchit de ces conditions initiales. Cependant l'usage de la régression linéaire se limite à l'élimination de dérive linéaire en fonction du temps.

Les dérives et offsets supposées linéaires en fonction du temps s'éliminent en calculant la régression linéaire du signal en fonction du temps ($\hat{s}(t) = s_0 + s_1 t$) et en soustrayant cette valeur au signal originel. Cette manière de procéder s'applique dans le cas où les signaux utilisés sont théoriquement centrés sur 0.

La phase d'élimination des éventuelles dérives est primordiale lorsque les modèles sont identifiés par minimisation d'un critère d'erreur de sortie. En effet, en présence d'offset et de dérive, ce type de critère est principalement constitué de ces quantités. L'influence de la valeur des paramètres du modèle sur la valeur d'un critère de l'erreur de sortie est alors secondaire [LJU87].

5.3.2 Bruit de mesure, points aberrants

L'influence dans le critère de coût de la présence de données aberrantes se minimise soit par l'intermédiaire de critère de coût robustifié, soit en filtrant les données par l'intermédiaire d'un filtre passe-bas. Comme pour le choix de la fréquence d'échantillonnage, de nombreuses règles empiriques ont été énoncées dans la littérature pour définir la fréquence de coupure de ces filtres.

Les données aberrantes doivent être éliminées pour éviter une estimation biaisée des paramètres. Or, la suppression ou la réduction du bruit de mesure par l'intermédiaire de filtre passe-bas n'est pas toujours souhaitable. Le filtrage des données irait à l'encontre des hypothèses formulées sur la nature du bruit de mesure présent sur les données. Dans le cas du maximum de vraisemblance par exemple, le bruit sur les données est assimilé à une suite de variables aléatoires indépendante. Expérimentalement (cf. § 0), la présence ou l'absence de filtrage des variables exerce une influence imperceptible sur les valeurs des paramètres estimés. Pour ces raisons, les données ne sont pas filtrées.

Afin de s'assurer que le modèle puisse être décrit par un modèle LE, il convient de vérifier la linéarité en l'entrée du processus.

5.4 Vérification de la linéarité vis à vis de l'entrée

La linéarité du modèle vis à vis de l'entrée se vérifie par la fonction de cohérence :

$$C_{xy}(\omega) = \frac{|S_{xy}(\omega)|^2}{S_{xx}(\omega) S_{yy}(\omega)} \quad (4.5-1)$$

où

$S_{xy}(\omega)$ désigne l'interspectre entre les signaux d'entrée et de sortie

$S_{xx}(\omega)$ désigne l'autospectre du signal d'entrée

La fonction de cohérence est assimilable à un coefficient de corrélation, dépendant de la fréquence, entre l'entrée et la sortie du processus étudié [LJU87]. Une valeur de fonction de cohérence proche de 1 à une fréquence donnée signifie que la relation liant la sortie du système à son entrée est parfaitement linéaire et que les signaux sont peu bruités. En revanche, une valeur proche de 0 est synonyme d'absence de relation de cause à effet entre l'entrée et la sortie du processus. Pratiquement, un système est considéré linéaire lorsque sa fonction de cohérence est supérieure à 0.8 [TIS95]. Aux fréquences où un phénomène d'antirésonance apparaît, la fonction de cohérence présente des valeurs généralement faibles ; un mauvais rapport signal à bruit est normal.

Vérifions le caractère linéaire par rapport à l'entrée des variables utilisées pour l'identification des paramètres physiques du véhicule automobile : la vitesse de lacet et l'accélération transversale. Des résultats typiques sont reproduits sur la figure 4 - 15.

Pour le protocole défini, évitant la saturation des pneumatiques, le caractère linéaire en l'entrée des variables de vitesse de lacet et d'accélération transversale est évident. Sur la variable de vitesse de lacet, quelles que soient les fréquences excitées, la fonction de cohérence est proche de la valeur l'unité. La variable d'accélération transversale présente à certaines fréquences, une fonction de cohérence d'amplitude plus faible. Ces valeurs moins élevées s'expliquent par un rapport signal à bruit moindre sur la variable d'accélération transversale que sur la variable de vitesse de lacet et aussi par la présence du phénomène d'antirésonance (approximativement situé à une fréquence de 2.2 Hz) déjà mentionné.

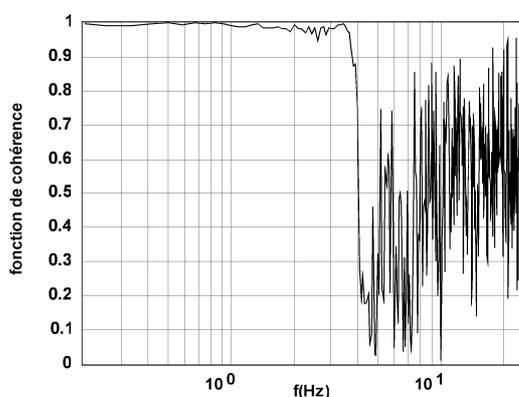


figure 4 - 15a : vitesse de lacet

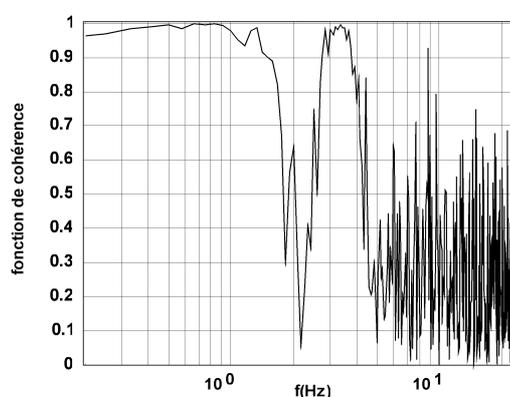


figure 4 - 15b : accélération transversale

figure 4 - 15 : fonctions de cohérence entre les sorties et l'angle volant

Au vu de ces courbes, l'adoption d'un modèle de véhicule linéaire (LE) à coefficients constants semble être un choix raisonnable pour la description du comportement dynamique latéral du véhicule dans les conditions expérimentales définies (cf. § 5.1). Après avoir conditionné les signaux et s'être assuré le caractère linéaire du véhicule, l'identification des paramètres physique à proprement parler peut débuter.

6 IDENTIFICATION DES PARAMETRES PHYSIQUES DU VEHICULE AUTOMOBILE

Dans ce paragraphe, un bref exposé de l'état de l'art de l'identification des paramètres physiques de véhicules automobiles sera dressé. Puis, compte tenu des propriétés structurelles des différents modèles, les paramètres physiques inconnus ou difficilement mesurables sont estimés par minimisation d'un critère d'erreur de sortie formulé au sens du maximum de vraisemblance. Certains paramètres identifiés sont communs à tous les modèles de connaissance : les deux rigidités de dérive ($D1$ et $D2$) et l'inertie de lacet (I_{zz}). Il sera alors montré que dans certains cas de figure, suivant les modèles utilisés, des différences notables apparaissent sur les valeurs des paramètres identifiés.

L'adéquation de la quantification théorique des intervalles d'incertitude par rapport aux écarts types expérimentaux sera recherchée et en partie atteinte. La véracité de la valeur des paramètres identifiés sera discutée. Contre toute attente, l'augmentation du nombre de paramètres identifiés ne se traduira pas systématiquement par une augmentation des intervalles d'incertitude associés aux paramètres estimés.

Une étude comparative des résultats issu de la minimisation de critères non-robustifiés et robustifié démontrera la robustesse des résultats de l'identification vis à vis de la présence de données aberrantes.

6.1 Etat de l'art

Vu les nombreuses applications potentielles, le nombre extrêmement réduit d'articles consacrés à l'identification des paramètres physiques de véhicules automobiles est surprenant. La recherche bibliographique effectuée recense 4 articles consacrés plus ou moins directement à l'identification des paramètres physiques de véhicules automobile ([HOR96], [BOT82], [HUA93], [LET95]).

Les travaux d'Horiuchi [HOR96] exploitent le résultat d'une procédure d'identification de certains paramètres du modèle lacet/dérive (les rigidités de dérive avant et arrière ($D1$ et $D2$) et l'inertie de lacet (I_{zz}) à des fins de contrôle de trajectoire. Peu de détails sur la procédure d'identification sont précisés et aucun intervalle de confiance n'est associé aux paramètres estimés. Les paramètres sont identifiés par minimisation d'un critère d'erreur de sortie formulée au sens du maximum de vraisemblance. Les variables entrant dans la composition du critère ne sont pas spécifiées. L'algorithme d'optimisation et le protocole expérimental ne sont pas reproduits dans l'article.

Bötiger [BOT82] identifie les paramètres d'inertie de lacet et de roulis (I_{zz} et I_{xx}), l'amortissement de roulis (α) et les rigidités de dérive avant et arrière ($D1$ et $D2$). A ces fins, un modèle de véhicule lacet/dérive/roulis est employé. Les autres paramètres intervenant dans ce modèle sont supposés connus. Les paramètres sont identifiés par minimisation d'un critère d'erreur quadratique fréquentiel. Les variables entrant dans la composition de ce critère ne sont pas mentionnées. L'algorithme d'optimisation, les problèmes d'identifiabilité éventuellement rencontrés et le protocole expérimental employé ne sont pas non plus reproduits. Les résultats sont exprimés par inspection visuelle du pouvoir prédictif du modèle dans les domaines temporel et fréquentiel. De plus, les valeurs des paramètres identifiés ne sont pas reproduites dans l'article et aucun intervalle d'incertitude n'est associé aux estimations.

Les travaux de Huang [HUA93], plus récents, semblent refléter un travail plus approfondi. Celui-ci identifie les deux rigidités de dérives ($D1$ et $D2$), l'inertie de roulis (I_{xx}), le produit masse suspendue/hauteur du centre de gravité ($m_s h_g$), les coefficients de braquages induits par le roulis (ε_1 et ε_2) et deux paramètres supplémentaires relatifs à l'actionneur de la direction assistée du

véhicule. Ces paramètres sont estimés par minimisation d'un critère d'erreur fréquentiel multivariable des moindres carrés. Ce critère se compose des variables de vitesse de lacet et de l'angle de roulis. Cependant, la pondération relative de ces variables et l'algorithme d'optimisation employé ne sont pas précisés. Huang soulève des problèmes numériques parfaitement symptomatiques de défauts d'identifiabilité. Toutefois, l'identifiabilité des paramètres n'est pas traitée, et le protocole expérimental non précisé. Contrairement aux travaux de Bötiger, les valeurs des paramètres identifiés sont ici portées à la connaissance du lecteur.

Enfin, les travaux relatés dans l'article de Le Letty [LET95] sont particulièrement intéressants. L'auteur parvient à identifier un nombre de paramètres remarquable par minimisation d'un critère de coût multivariable. Les variables entrant dans la composition de ce critère ne sont pas précisées. Les paramètres identifiés sont : les deux rigidités de dérivées (D1 et D2), l'amortissement et la raideur de roulis (k_{ro} et a_{ro}), les coefficients de braquages induits par le roulis (ϵ_1 et ϵ_2), les longueurs de relaxation des pneus avant et arrière (b_1 et b_2), les hauteurs des centres de roulis avant et arrière (h_1 et h_2) et le demi-empattement avant (l_1). L'identification simultanée de l'ensemble de ces paramètres est rendue possible par la formation quelque peu « originale » du critère d'erreur minimisé. Ce critère inclut entre autres, un terme de pondération en faveur de valeurs connues a priori des paramètres. Les paramètres identifiés sont donc biaisés en faveur des valeurs a priori des paramètres. Cette formulation, inspirée du maximum a priori, du critère de coût parvient donc à lever bon nombre de difficultés posées par l'identifiabilité des paramètres. Cependant, l'identifiabilité des paramètres n'est pas étudiée, et le protocole expérimental n'est pas précisé. Les résultats obtenus sont validés par inspection visuelle des sorties prédites par le modèle avec les données expérimentales et par comparaison des valeurs de paramètres aux données constructeur. Les valeurs des paramètres identifiés sont reproduites sans intervalles d'incertitude associés. Un algorithme d'optimisation de type quasi-Newton est utilisé pour minimiser le critère de coût. Malgré la dimension élevée de l'espace de recherche engendrée par le grand nombre de paramètres identifiés et la formulation complexe du critère de coût, aucun problème de minima locaux n'est mentionné. Les résultats sont obtenus à partir de données simulées et sont donc à considérer avec circonspection. La transposition de la démarche adoptée sur des données expérimentales réelles n'est pas évidente et n'est pas abordée.

6.2 Résultats stochastiques (simulation)

Avant d'entamer l'identification des paramètres à partir de données expérimentales réelles, il semble naturel de valider les fondements théoriques, les résultats asymptotiques et les différents algorithmes implémentés. **Ce cadre théorique permet également dans la mesure du possible de vérifier que le protocole expérimental adopté est adapté à l'identification des paramètres physiques recherchés.**

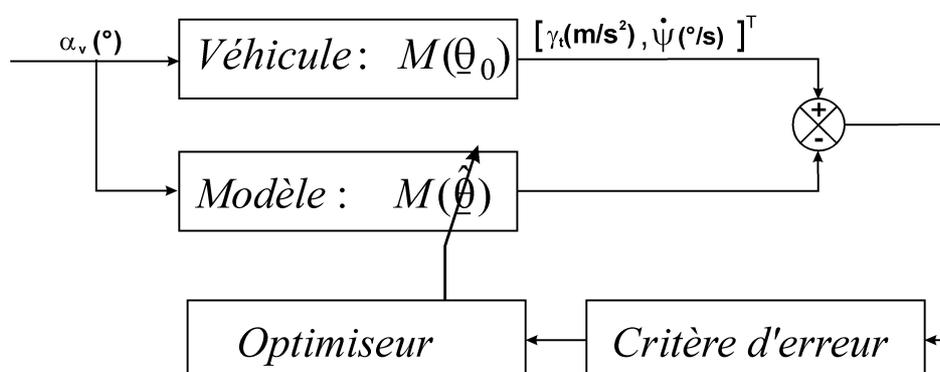


figure 4 - 16 : cadre de l'identification des paramètres physiques du véhicule automobile

La figure 4 - 16) récapitule sous forme de blocs diagrammes le cadre théorique et expérimental d'identification des paramètres physiques du véhicule automobile.

Pour cela, dans un premier temps, les paramètres physiques du véhicule automobile sont identifiés à partir de données simulées, par minimisation d'un critère d'erreur de sortie multivariable (vitesse de lacet et accélération transversale) formulé au sens du maximum de vraisemblance. Un bruit de mesure constitué d'une suite de variables aléatoires indépendante, de distribution normale est ajouté sur les variables simulées. La variance de ces bruits de mesure est calée sur l'amplitude du bruit de mesure expérimentalement constaté sur ces mêmes variables (soit une variance de 0.02 sur la vitesse de lacet, et une variance de 0.03 sur l'accélération transversale).

Un grand nombre de jeux de données (100) est simulé. Les propriétés asymptotiques de l'estimateur au sens du maximum de vraisemblance peuvent alors être vérifiées expérimentalement (Monte-Carlo). Afin de rendre compte du protocole expérimental réel, l'excitation appliquée au volant est de type harmonique modulée linéairement en fréquence de 0.2 à 5 Hz dans un laps de temps de 40 secondes.

Dans un second temps, un bruit rose supplémentaire est ajouté au bruit blanc précédent. Du point de vue de sa distribution spectrale, ce bruit rose simule la présence de résidus non blancs. Dans ce second cas de figure, l'aptitude de l'algorithme de quantification des intervalles d'incertitude en présence de résidus colorés sera vérifiée.

6.2.1 Paramètres identifiés

Les résultats reproduits concernent l'identification des paramètres physiques du modèle LaRouDéBal. Sans perte de généralité ces résultats sont transposables aux autres modèles de véhicule décrits dans ce mémoire. En effet, une démarche identique opérée sur les autres modèles conduit à des résultats similaires en terme de propriétés de l'estimateur du maximum de vraisemblance.

θ	Unité	θ_0	$\hat{\theta}_{mv}$	$\sigma_{95\% \text{ c.r}} (\%)$	$\sigma_{95\% \text{ exp.}} (\%)$
D1	N/rad	84085	83870	1.34	1.28
D2	N/rad	87342	87028	1.80	1.60
Izz	kg.m ²	1571.5	1570.5	1.18	1.02
Ixx	kg.m ²	252.0	252.4	0.92	0.96
Ixz	kg.m ²	50.0	50.9	8.72	8.48
kro	N/rad	150000	150476	1.20	1.28
aro	Nm/rad/s	2750	2751	0.92	0.82
b1	m	0.500	0.497	2.44	2.50
b2	m	0.500	0.499	0.78	0.74

tableau 4 - 4 : paramètres estimés en présence de bruit blanc

Les statistiques résultant de l'identification sur 100 jeux de données sont reproduites dans le tableau 4 - 4, en présence de bruit blanc Gaussien ($N(0, 0.07)$ sur la vitesse de lacet, et $N(0, 0.08)$ sur l'accélération transversale). Ce tableau contient :

- 1° colonne, θ_0 : les valeurs nominales des paramètres, utilisées pour générer les données fictives.
- 2° colonne, $\hat{\theta}_{mv}$: la valeur moyenne des paramètres sur les 100 estimations.

- 3° colonne, $\sigma_{95\% \text{ c.r.}} (\%)$: les intervalles de confiance à 95% prédits par l'inégalité de Cramer-Rao, exprimés en pour-cent de $\hat{\theta}_{mv}$.
- 4° colonne, $\sigma_{95\% \text{ exp.}} (\%)$: les intervalles de confiance à 95 % calculés à partir des 100 estimations, exprimés en pour-cent de $\hat{\theta}_{mv}$.

Les valeurs reportées dans ce tableau soulignent la proximité de la moyenne des estimées $\hat{\theta}_{mv}$ avec les valeurs nominales θ_0 , synonymes d'une estimation non-biaisée des paramètres. De plus, l'histogramme reproduit figure 4 - 17 tend à montrer que l'estimateur du maximum de vraisemblance est asymptotiquement normal.

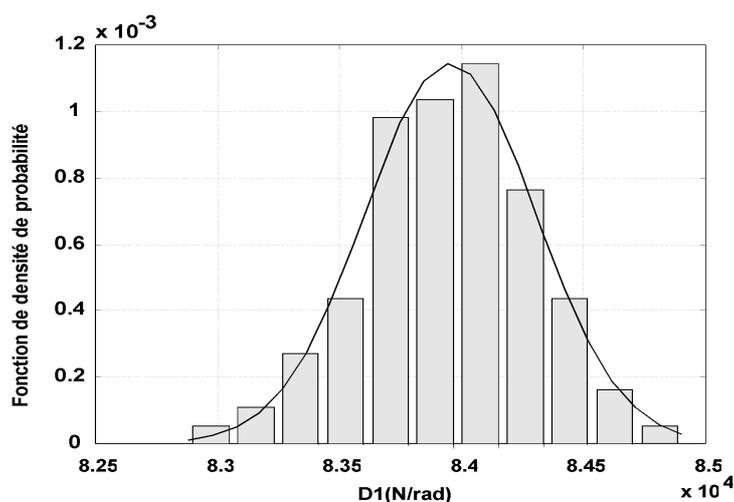


figure 4 - 17 : histogramme des estimées de la rigidité de dérive avant D1

6.2.2 Globalité et précision des solutions

Il a été montré dans le premier chapitre, que le minimum du critère de coût formulé au sens du maximum de vraisemblance était quadratique. La globalité des solutions obtenues se vérifie alors aisément en perturbant (sous forme de somme ou de soustraction) successivement chacune des composantes du vecteur solution par l'insensibilité de chaque paramètre. Si l'optimum localisé par l'algorithme d'optimisation est global (quadratique), chaque évaluation de la fonction de coût à partir du vecteur solution perturbé, se traduira par une augmentation de coût de 0.5.

θ	$\Delta J (-)$	$\Delta J (+)$
D1 (N/rad)	0.4998	0.4998
D2 (N/rad)	0.5007	0.5011
lzz (kg.m ²)	0.5000	0.5014
lxx (kg.m ²)	0.5006	0.5002
lxz (kg.m ²)	0.5003	0.5011
kro (N/rad)	0.5004	0.5011
aro (Nm/rad/s)	0.4985	0.5007
b1 (m)	0.4997	0.5005
b2 (m)	0.5002	0.4985

tableau 4 - 5 : vérification de la globalité des solutions

Dans le tableau 4 - 5 sont reproduites les augmentations induites par le vecteur solution perturbé positivement (+ : addition) et négativement (- : soustraction) par les insensibilités respectives de chaque paramètre. Cet exemple est prélevé sur une des 100 identifications réalisées. L'optimum global a systématiquement été localisé :

Ces valeurs typiques (pour les 100 solutions) montre clairement que les optima localisés sont quadratiques et globaux. L'algorithme d'optimisation du simplex utilisé dans le cas présent se révèle donc adapté pour la minimisation du critère de coût précédent. L'algorithme d'optimisation est initialisé dans un proche voisinage de l'optimum global ($\underline{x}(0) = 1.2 \underline{x}_0$). Les solutions sont localisées avec une précision $\delta = 10^{-4}$. Le critère d'arrêt employé est de la forme $ca = ca2 \& ca3 \mid ca5$ (cf. chapitre 2), avec $\varepsilon_{d_y} = 10^{-4}$, $\varepsilon_x = 10^{-4}$, et le nombre maximum d'itérations fixé à 100 fois la dimension de l'espace de recherche [GRA92], soit 900 dans le cas présent. L'identification des paramètres du modèle à partir d'un jeu de données nécessite un temps de calcul moyen de 100 s.

6.2.3 Invalidation des résidus

Les conditions de simulations étant idéales, les différents tests (cf. chapitre précédent et annexe 2) visant à vérifier avec un degré de confiance de 95 % l'espérance mathématique nulle des résidus, leur normalité, leur stationnarité et leur indépendances se révèlent positifs. A titre d'exemple, la fonction d'autocorrélation normalisée des résidus et la fonction d'intercorrélation normalisée des résidus avec l'entrée sont reproduites sur la figure 4 - 18.

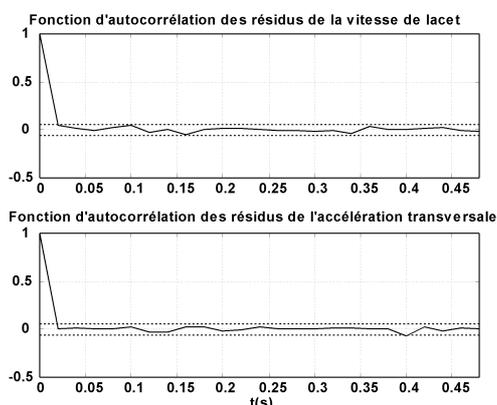


figure 4 - 18a : fonction d'autocorrélation

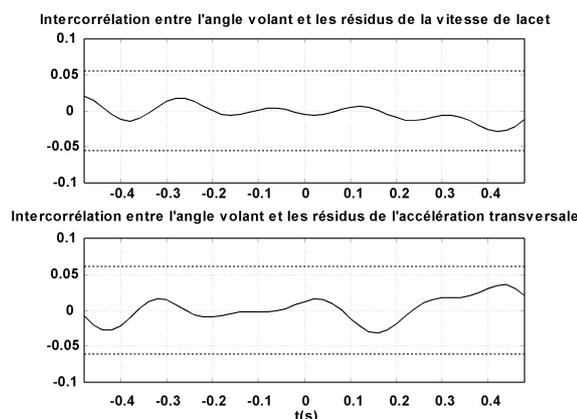


figure 4 - 18b : fonction d'intercorrélation

figure 4 - 18 : blancheur des résidus et corrélation des résidus avec l'entrée

Les résidus normalisés et le tracé de leurs régressions polynomiales (figure 4 - 19a) à l'ordre 2 met clairement en évidence la stationnarité des résidus. L'histogramme des résidus reproduit figure 4 - 19b, comparé à une distribution normale, illustre le caractère Gaussien de la distribution des résidus.

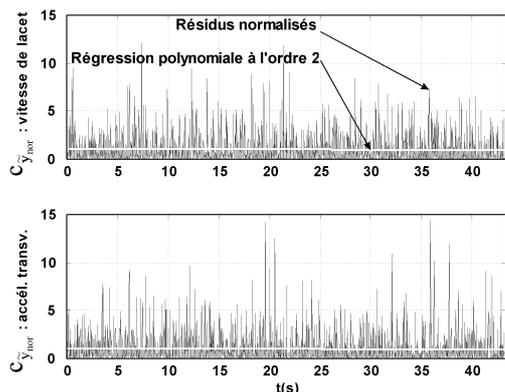


figure 4 - 19a : résidus normalisés

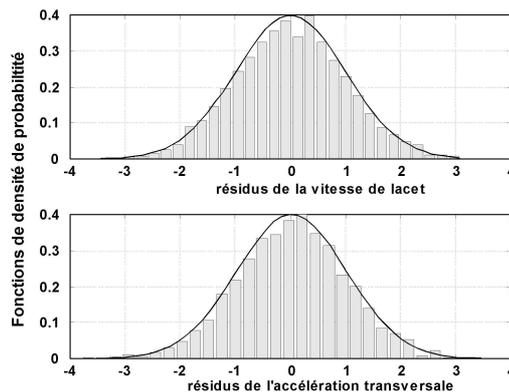


figure 4 - 19b : histogramme des résidus

figure 4 - 19 : stationnarité et distribution des résidus

6.2.4 Pouvoir prédictif

Les erreurs efficaces moyennes sur les variables de sorties sont respectivement de $0.26^\circ/s$ sur la vitesse de lacet et de 0.29 m/s^2 sur l'accélération transversale. Ces valeurs traduisent un excellent ajustement temporel et fréquentiel. A titre d'exemple, les prédictions temporelles du modèle LaRouDéBal sur les variables de vitesse de lacet et d'accélération transversale sont reproduites sur la figure 4 - 20. Les sorties prédites sont comparées aux données simulées (donnée d'identification). Pour des raisons de lisibilité de la figure, un zoom est réalisé sur un intervalle temporel restreint, de l'ordre de 6 s, soit une durée légèrement supérieure à une période du signal de type sbpa durant 4.3 s.

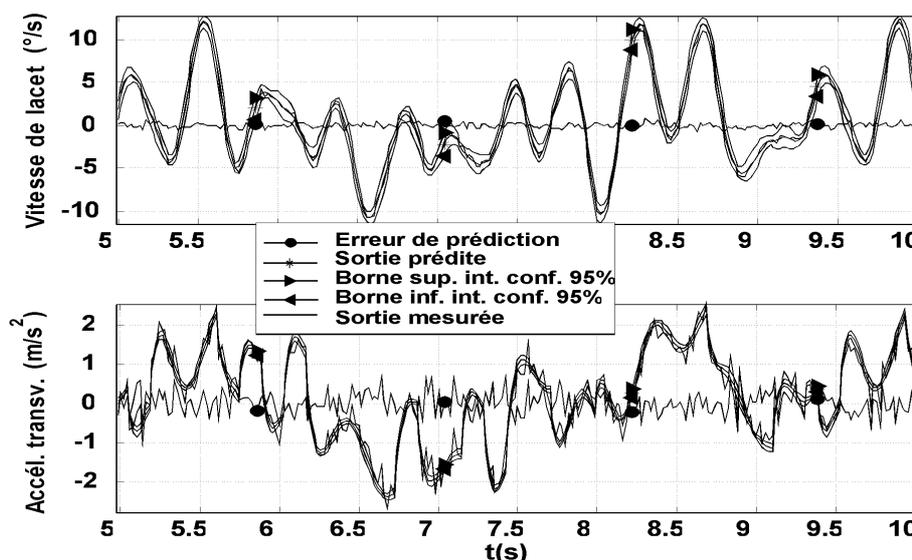


figure 4 - 20 : pouvoir prédictif temporel du modèle identifié

6.2.5 Résultats en présence de bruit coloré sur les données

Dans ce cas, avant d'opérer l'estimation des paramètres, un bruit coloré est ajouté aux variables et au bruit de mesure précédent. Ce bruit coloré est obtenu par filtrage de bruit blanc ($N(0, 0.8)$ pour la vitesse de lacet et $N(0, 0.6)$ pour l'accélération transversale) via un filtre de Butterworth d'ordre 10 de fréquence de coupure 3.5 Hz. Les variances des bruits blancs filtrés sont adaptées en fonction de l'amplitude du spectre des résidus résultant de l'identification des paramètres du modèle LaRouDéBal.

Le tableau 4 - 6 comporte les résultats de l'identification des paramètres physiques à partir des données corrompues par la présence de bruit blanc de mesure et la présence de bruit coloré. En présence de bruit coloré, les intervalles d'incertitude prédits par les bornes de Cramer-Rao sont sensiblement sous-évalués. Leurs corrections (6^o colonne du tableau) par l'intermédiaire du facteur correctif développé par Maine [MAI81] détaillé dans le troisième chapitre, rend celles-ci nettement plus représentatives des intervalles d'incertitude expérimentaux.

Dans le cadre idéalisé de la simulation, la viabilité du processus d'identification de certains paramètres physiques du véhicule automobile a été montrée. Validons cette méthodologie dans un contexte plus réaliste, en exploitant des données réelles prélevées sur le véhicule lors d'essais sur piste.

θ	θ_0	$\hat{\theta}_{-mv}$	$\sigma_{95\% \text{ c.r.}} (\%)$	$\sigma_{95\% \text{ exp.}} (\%)$	$\sigma_{95\% \text{ cr}_{\text{corrigées}}} (\%)$
D1 (N/rad)	84085	83618	4.16	5.24	5.20
D2 (N/rad)	87342	86764	5.47	6.98	6.83
lzz (kg.m ²)	1571.5	1572.7	3.66	4.58	4.57
lxx (kg.m ²)	252.0	253.9	2.97	4.66	3.71
lxz (kg.m ²)	50.0	51.2	27.15	32.06	33.88
kro (N/rad)	150000	1509867	4.10	7.06	5.12
aro (Nm/rad/s)	2750	2743	2.96	4.24	3.69
b1 (m)	0.500	0.493	7.56	8.70	9.44
b2 (m)	0.500	0.498	2.52	3.26	3.15

tableau 4 - 6 : paramètres estimés en présence de bruit coloré

Dans ce cadre théorique en présence de bruit coloré sur les données il apparaît que le produit d'inertie croisée lacet/roulis lxz n'est pas identifiable et que par conséquent le nombre de paramètres identifiables décroît de 9 à 8. Il est donc probable que ce phénomène se rencontre également dans le cadre expérimental.

6.3 Résultats expérimentaux

Les résultats obtenus successivement avec les modèles LaDé, LaDéBal et LaRouDéBal sont tour à tour présentés. Pour ces modèles, différents aspects sont traités. En premier lieu la véracité des valeurs des paramètres identifiés sera essentiellement jugée à partir de la connaissance a priori disponible sur le véhicule. Dans une certaine mesure, les intervalles d'incertitude associés aux paramètres identifiés, permettent de caractériser le caractère aléatoire ou déterministe des valeurs obtenues. Dans un second temps, une importance toute particulière sera consacrée à l'étude des pouvoirs prédictifs des modèles identifiés et à la nature des résidus. Enfin, en fonction des modèles,

en fonction des conditions expérimentales (notamment la vitesse du véhicule), la cohérence et la véracité des paramètres seront discutées.

Les résultats expérimentaux sont le fruit de l'exploitation de données expérimentales prélevées sur le véhicule du laboratoire, lors d'essais sur piste. Etant donné la lourdeur des moyens mis en œuvre, un nombre limité de jeux de données supporte cette partie applicative. Dans la mesure du possible, une identité de protocole expérimental sera recherchée. Les conditions expérimentales sont les suivantes :

- Une excitation au volant de type sinus vobulé en fréquence de 0.2 à 5 Hz.
- Une fréquence d'échantillonnage de 200 Hz.
- Une vitesse longitudinale du véhicule approximativement constante de 80km/h.

Ces différentes conditions expérimentales, permettent de définir diverses séries de jeux de données. Ces différents groupes sont principalement définis par la vitesse du véhicule et son chargement (nombre et répartitions des personnes à bord).

☐ **Vitesse de 80 km/h :**

- **Série 1** : 2 personnes (p1 et p2), 5 fichiers de données.
- **Série 2** : 2 personnes (p2 et p3), 4 fichiers de données.
- **Série 3** : 3 personnes (p1, p2 et p4), 5 fichiers de données.
- **Série 4** : 2 personnes (p1 et p2, ou p2 et p3), 9 fichiers de données.

☐ **Vitesse de 100 km/h :**

- **Série 5** : 2 personnes (p1 et p2), 4 fichiers de données.
- **Série 6** : 3 personnes (p1, p2 et p4), 5 fichiers de données.

Afin d'obtenir un nombre de fichiers plus important au sein d'une même série (**série 4**), les séries 1 et 2 peuvent être additionnées. En effet le poids cumulé des personnes (p1+p2) est approximativement identique au poids cumulé des personnes (p2+p3). Cependant la répartition légèrement différente des masses au sein du véhicule peut entraîner des différences sur les paramètres estimés, notamment les inerties.

Les résultats sont présentés pour les différents modèles de véhicule développés. Cette présentation débute avec les résultats issus du modèle LaDé. En effet, l'augmentation de la complexité des modèles conjuguée à l'augmentation du nombre de paramètres estimés pose un nombre croissant de difficultés avec la complexité des modèles.

6.3.1 Modèle LaDé

Le tableau 4 - 8 résume les résultats obtenus pour l'identification des deux rigidités de dérive avant et arrières (D1 et D2) et de l'inertie de lacet (Izz) pour les différentes séries de jeux de données. La première colonne de ce tableau contient les valeurs nominales fournies par le constructeur, puis dans les colonnes suivantes sont reproduits les résultats pour les différentes séries de données. A chaque paramètre estimé sont associés deux ou trois (suivant les cas) intervalles d'incertitude avec un degré de confiance de 95 %, exprimés en pour-cent des valeurs estimées (cf. tableau 4 - 8).

74141	3.36	▶ intervalles confiance à 95 % (estimé expérimentalement)
	7.12	▶ intervalles de confiance à 95 % (inégalité de Cramer-Rao corrigée)
	4.70	▶ intervalles de confiance à 95 % (inégalité de Cramer-Rao non-corrigée)

tableau 4 - 7 : présentation des résultats

Uniquement pour la série 4, le premier (en partant du haut) noté dans la colonne est calculé sur l'ensemble des estimations réalisées sur une série de jeux de données. Le second représente l'intervalle de confiance à 95 % estimé à partir de l'inégalité de Cramer-Rao corrigée selon Maine [MAI81]. Enfin, le troisième est l'intervalle de confiance à 95 % estimé à partir de l'inégalité de Cramer-Rao non-corrigée, cf. tableau 4 - 7.

θ	$\hat{\theta}_{mv}$ série 1 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 2 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 3 V = 80 km/h M = 1475 kg		$\hat{\theta}_{mv}$ série 4 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 5 V = 100 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 6 V = 100 km/h M = 1475 kg	
	D1 N/rad	74991	-- 7.84 4.36	73078	-- 6.22 3.34	74794	-- 8.80 4.64	74141	3.36 7.12 4.70	74643	-- 10.2 5.68	75960
D2 N/rad	89434	-- 9.14 4.08	88249	-- 7.66 4.18	73659	-- 8.74 4.60	88909	8.08 8.48 4.15	89536	-- 10.8 5.98	87568	-- 9.22 5.12
Izz kg.m ²	1690	-- 7.88 4.38	1714	-- 6.12 3.28	1601	-- 8.16 4.30	1700	8.16 7.08 4.18	1651	-- 10.2 5.68	1736	-- 9.12 5.06

tableau 4 - 8 : paramètres estimés avec le modèle LaDé

6.3.1.1 Paramètres identifiés

Pour les séries de jeux de données (**séries 1, 2, 4 et 5**), les valeurs des paramètres identifiés sont toutes voisines. Compte tenu du nombre restreint de jeux de données, les différences observées peuvent raisonnablement être attribuées à la présence de bruit sur les mesures.

Du point de vue du modéliste, les valeurs obtenues des deux rigidités de dérive identifiées sont proches des valeurs constructeur avec 2 personnes de 75 kg à bord (D1 = 84085 N/rad, D2 = 87342 N/rad). L'obtention de ces valeurs est rassurante et tends à confirmer la véracité des valeurs obtenues. Cependant, le poids non négligeable du système d'acquisition présent dans le coffre exerçant un effort de poussée verticale supplémentaire tendrait du point de vue théorique à augmenter la valeur de la rigidité de dérive arrière. La valeur de l'inertie de lacet (Izz) de l'ordre de 1700 kg.m² est supérieure à la valeur constructeur (Izz = 1575 kg.m²) donnée pour un véhicule chargé de 2 personnes à l'avant. Ce supplément d'inertie est attribuable à la masse du système d'acquisition présent dans le coffre.

□ Influence du chargement

L'interprétation des résultats issus de l'identification des paramètres à partir des séries 3 et 6 pose certaines questions. Dans les cas de figure où les paramètres physiques sont identifiés à partir des séries 3 et 6, la masse du véhicule a été modifiée de manière à tenir compte du supplément de poids (75 kg) engendré par la présence de la troisième personne.

A 80 km/h (**série 4 et 5**), la présence d'un troisième passager entraîne une diminution de l'inertie de lacet (Izz : - 5.9%) et de la rigidité de dérive arrière (D2 : - 17.7%) lorsque ces paramètres sont identifiés à partir de la série 3. Ce fait ne reçoit pas d'explication satisfaisante. En effet, la présence d'un passager supplémentaire entraîne logiquement une augmentation de la valeur de la rigidité de dérive arrière et une augmentation de l'inertie de lacet.

L'hypothèse suivante peut être avancée : la personne supplémentaire est présente sur le coté arrière droit du véhicule. Non seulement sa présence modifie l'inertie de lacet, mais aussi l'inertie de

roulis (I_{xx}). Par voie de conséquence, le comportement en roulis du véhicule se trouve affecté. Or le phénomène de roulis n'est pas modélisé au sein du modèle LaDé, et l'algorithme d'optimisation ajuste les paramètres du modèle dans le but d'obtenir un comportement du modèle identique à celui du système réel. Les variations inexpliquées des paramètres identifiés sont alors peut être la conséquence de la présence d'erreur de modélisation et notamment de l'absence de description du phénomène de roulis.

¶ **Influence de la vitesse**

En dépit de la vitesse moyenne de 100 km/h de la série 5, comparativement aux résultats obtenus à partir de la série 4 (80 km/h), les résultats sont peu affectés par la variation de la vitesse du véhicule. Relativement à leurs valeurs, l'inertie de lacet est le paramètre le plus affectée par ce changement de vitesse.

Par ailleurs, l'identification des paramètres du véhicule avec un troisième passager à bord pour les vitesses de 80 km/h (**série 3**) et 100 km/h (**série 6**) entraîne des résultats sensiblement différents. L'augmentation de la vitesse du véhicule de 80 km/h à 100 km/h entraîne une nette augmentation de la rigidité de dérive arrière ($D2 : + 18.9\%$) et de l'inertie de lacet ($I_{zz} : +8.5\%$)

6.3.1.2 Incertitudes sur les paramètres

Les statistiques réalisées sur les séries constituées d'un nombre de jeux de données restreint sont peu significatives. Considérons cependant la série 4 constituée de 9 jeux de données et les intervalles d'incertitude (avec un degré de confiance de 95 %) associés aux paramètres identifiés. Abstraction faite de la rigidité de dérive avant ($D1$), les intervalles de confiance calculés expérimentalement à partir des 9 identifications réalisées sont bien supérieurs à ceux prédits par l'inégalité de Cramer-Rao.

En corrigeant l'inégalité de Cramer-Rao ($K=1.2$), selon la méthode de Maine, les bornes de Cramer-Rao corrigées sont plus représentatives des intervalles d'incertitude expérimentaux. Malgré l'extrême simplicité du modèle utilisé, les intervalles d'incertitude associés aux paramètres identifiés sont d'ores et déjà bien inférieurs à ceux associés aux données constructeur (voir tableau 4 - 2)²⁸.

6.3.1.3 Globalité et précision des solutions

L'espace de recherche associé au problème d'optimisation posé par l'identification des paramètres du modèle LaDé est de dimension modeste (dimension 3). Les algorithmes de type grimpeurs résolvent ce problème aisément. Dans le cas présent, l'algorithme du simplexe a été employé avec succès. La globalité de toutes les solutions a été vérifiée avec succès. Le critère d'arrêt employé est identique à celui utilisé pour l'identification des paramètres à partir des jeux de données simulés. Une précision $\delta = 10^{-4}$ est atteinte dans 100% des cas. Ce critère d'arrêt sera conservé par la suite. L'identification des paramètres du modèle à partir d'un jeu de données nécessite un temps de calcul moyen de 15 s.

6.3.1.4 Invalidation des résidus

L'étude des propriétés des résidus révèle que le nombre réduit de degré de liberté du modèle LaDé n'est pas suffisant pour décrire l'intégralité des phénomènes régissant la dynamique latérale du véhicule entre 0.2 et 1.5 Hz.

²⁸ Dans ce tableau sont reportés les écarts types associés aux valeurs constructeur et non pas les intervalles de confiance à 95 %.

Par le biais des indicateurs présentés dans le chapitre précédent, les résultats propres à l'invalidation des résidus sont consignés dans le tableau 4 - 9. Dans ce tableau sont reportés les cas exprimés en pour-cent ou l'invalidation des résidus a échoué.

Propriétés des résidus	Pourcentage de jeux de données non-invalidé (%)	
	Vitesse de lacet	Accélération transversale
Moyenne nulle	77.77	100
Skewness	22.22	55.55
Kurtosis	22.22	44.44
Stationnarité	0	0
Indépendance	0	22.22

tableau 4 - 9 : invalidation des résidus

La série 4 étant constituée de 9 jeux de donnée, un jeu de données représente donc 11.11 % des cas. L'invalidation de la moyenne nulle des résidus, de la comparaison de leur skewness et de leur kurtosis avec ceux d'une distribution normale, de leur stationnarité et de leur indépendance est successivement recherchée et les résultats sont reportés dans ce tableau.

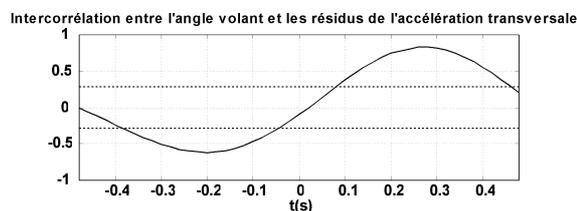
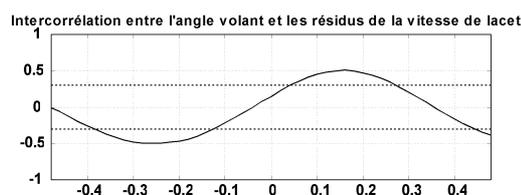
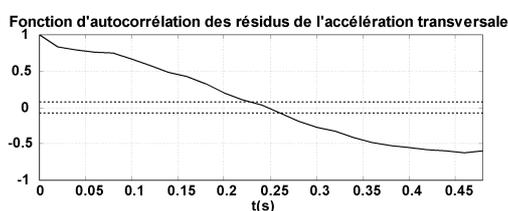
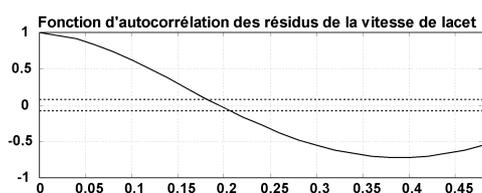


figure 4 - 21a : fonction d'autocorrélation

figure 4 - 21b : fonction d'intercorrélation

figure 4 - 21 : blancheur des résidus et corrélation des résidus avec l'entrée

L'espérance mathématique non-nulle des résidus est relativement difficile à invalider (notamment sur l'accélération transversale, cf. tableau 4 - 8). La non-normalité des résidus est quant à elle plus facile à montrer (cf. tableau 4 - 8 et figure 4 - 22b). La stationnarité des résidus a été invalidée dans 100% des cas et leur indépendance n'a pas pu être invalidée dans seulement 2 cas sur l'accélération transversale. Les tracés (figure 4 - 21) des fonctions d'intercorrélation des résidus avec l'entrée et de leur fonction d'autocorrélation confirme la non-blancheur des résidus des différentes variables.

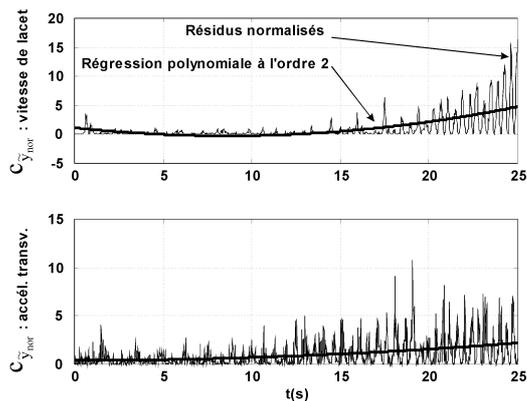


figure 4 - 22a : résidus normalisés

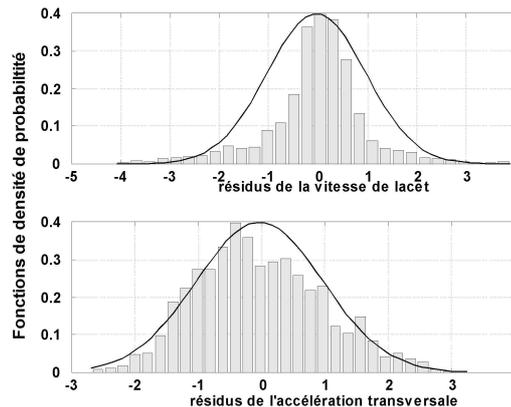


figure 4 - 22b : histogramme des résidus

figure 4 - 22 : stationnarité et distribution des résidus

Ces figures montrent notamment la forte dépendance des résidus de l'accélération transversale et de la vitesse de lacet vis à vis de l'angle volant. Les résidus normalisés et le tracé de leurs régressions polynomiales (figure 4 - 22a) à l'ordre 2 met clairement en évidence la non-stationnarité des résidus. L'histogramme des résidus reproduit figure 4 - 22b, comparé à une distribution normale (paramétrée par la valeur moyenne et la variance expérimentale), illustre le caractère non-Gaussien de la distribution des résidus

6.3.1.5 Pouvoir prédictif

□ Domaine temporel

Les conclusions précédentes sur la nature des résidus se reflètent sur le pouvoir prédictif du modèle dans le domaine temporel. Les prédictions en validation croisée et les intervalles de confiance à 95 % corrigés sont reproduits sur la figure 4 - 23. Ces intervalles de confiance sur les sorties prédites sont calculés à partir des intervalles de confiance corrigés sur les paramètres estimés.

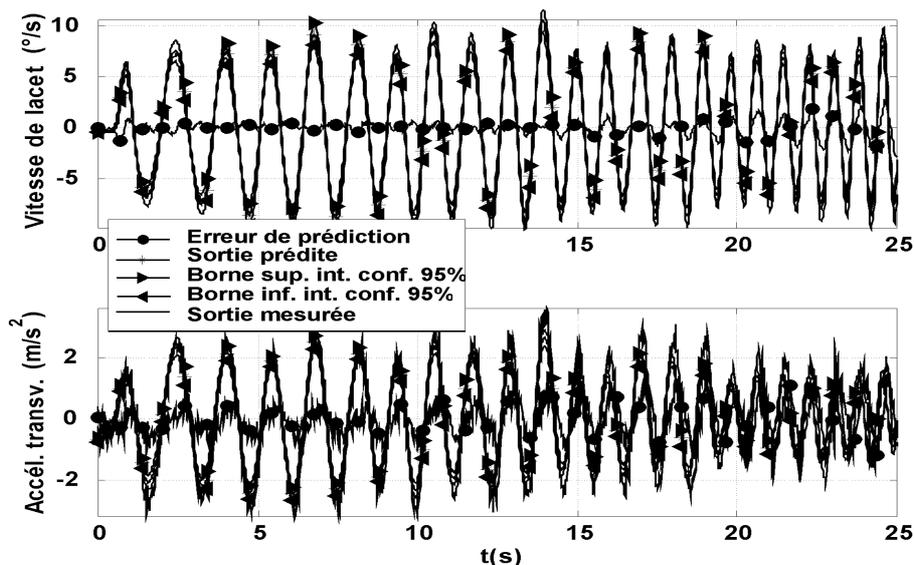


figure 4 - 23 : pouvoir prédictif temporel du modèle identifié

La figure 4 - 22a et la figure 4 - 23 montrent clairement que le pouvoir prédictif du modèle diminue d'autant plus que la fréquence du signal d'excitation (angle volant) croît. Ce fait motive partiellement le choix d'un modèle de complexité plus élevée, intégrant l'effet de ballant (dynamiques plus rapides). Sur l'exemple présenté sur la figure 4 - 23, les erreurs RMS commises sur la vitesse de lacet et sur l'accélération transversale sont respectivement de $0.71^\circ/s$ et de 0.58 m/s^2 .

□ Domaine fréquentiel

Sur le diagramme de Bode reproduit sur la figure 4 - 24, il convient de considérer uniquement les fréquences inférieures à 1.5 Hz, les fréquences supérieures n'étant pas significatives car non prises en compte dans le processus d'estimation : le jeu de données est tronqué aux fréquences supérieures à 1.5 Hz (cf. § 3.7 de ce chapitre). L'algorithme d'optimisation ajuste les paramètres du modèle de façon à obtenir un bon ajustement des variables entrant dans la composition du critère de coût. Le modèle LaDé étant manifestement sous caractérisé, le pouvoir prédictif fréquentiel du modèle final est relativement médiocre. Ces figures confirment l'inaptitude du modèle à décrire le comportement du véhicule pour les fréquences approximativement supérieures à 1.0 Hz.

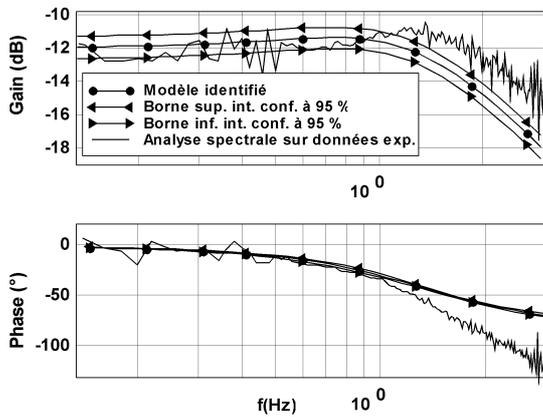


figure 4 - 24a : vitesse de lacet

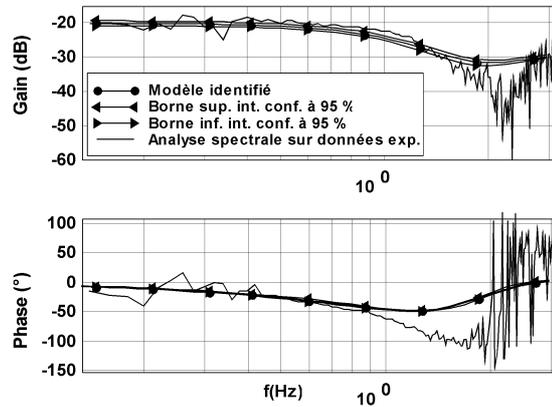


figure 4 - 24b : accélération transversale

figure 4 - 24 : pouvoir prédictif fréquentiel du modèle LaDé identifié

6.3.2 Modèle LaDéBal

Les deux rigidités de dérive ($D1$ et $D2$), l'inertie de lacet (I_{zz}) et les deux longueurs de relaxation des pneumatiques ($b1$ et $b2$) sont identifiées à partir des différentes séries de jeux de données. Les résultats obtenus sont consignés dans le tableau 4 - 10.

6.3.2.1 Paramètres identifiés

Comme pour le modèle précédent, les paramètres identifiés à partir des données expérimentales des séries 1, 2, 4 et 5 sont d'amplitudes voisines. Toutefois, une plus grande disparité entre les résultats issus des différentes séries de fichiers est notable. Comparons maintenant les valeurs des paramètres identifiés communs aux modèles LaDé et LaDéBal. En considérant les résultats obtenus à partir de la série 4 (le nombre de jeux de données le plus élevé), aucune différence significative n'est perceptible sur la valeur de l'inertie de lacet (I_{zz}). Une légère différence s'observe sur la valeur de la rigidité de dérive avant ($D1$) : une augmentation de 2.2% de sa valeur estimée avec le modèle LaDéBal par rapport à la valeur estimée avec le modèle LaDé. La valeur de ces deux paramètres semble donc relativement insensible à la description du phénomène de ballant.

θ	$\hat{\theta}_{mv}$ série 1 V = 80 km/h M = 1400 kg	$\hat{\theta}_{mv}$ série 2 V = 80 km/h M = 1400 kg	$\hat{\theta}_{mv}$ série 3 V = 80 km/h M = 1475 kg	$\hat{\theta}_{mv}$ série 4 V = 80 km/h M = 1400 kg	$\hat{\theta}_{mv}$ série 5 V = 100 km/h M = 1400 kg	$\hat{\theta}_{mv}$ série 6 V = 100 km/h M = 1475 kg
D1 N/rad	76311	--	--	--	3.66	--
	4.68	75176	3.62	70778	3.84	75806
	1.72	--	--	--	4.20	75686
D2 N/rad	97225	--	--	--	5.22	--
	6.68	95481	5.58	71726	4.56	100864
	2.48	--	--	--	6.19	97467
Izz kg.m ²	1647	--	--	--	8.78	--
	6.90	1747	9.68	1451	4.48	1780
	2.56	--	--	--	3.44	1909
b1 m	0.51	--	29.2	--	15.5	--
	8.96	0.44	20.0	0.52	4.26	0.48
	3.10	--	6.94	--	3.26	0.48
b2 m	0.44	--	--	--	11.9	--
	6.26	0.48	11.3	0.41	3.68	0.48
	2.32	--	3.68	--	2.82	0.48
					8.50	9.10
					3.05	3.38
						5.42
						3.62

tableau 4 - 10 : paramètres estimés avec le modèle LaDéBal

En revanche, une variation sensible de la valeur de la rigidité de dérive arrière (D2) est flagrante. L'amplitude de la valeur de la rigidité de dérive identifiée par le biais du modèle LaDéBal est supérieure de 8.5 % à celle identifiée par le biais du modèle LaDé. Cette différence semble être attribuable à la modélisation du phénomène de ballant. En effet, le phénomène de ballant se manifeste par l'apparition d'un retard entre la sollicitation donnée au volant et la génération d'un effort transversal au niveau du pneumatique. Le modèle LaDé ne tenant pas compte de ce phénomène, l'algorithme d'optimisation ajuste les paramètres du modèle LaDé de manière à obtenir une identité de comportement modèle/système réel.

Or, pour tenir compte de la présence de ce retard dû au phénomène de ballant lors de l'identification des paramètres du modèle LaDé, l'algorithme d'optimisation dispose uniquement de trois degrés de liberté que sont les trois paramètres identifiés (D1, D2 et Izz). Afin de tenir compte de la présence de ce retard, cet algorithme tâchera donc de « ralentir » le modèle en jouant sur ces trois degrés de liberté. Ce ralentissement ne peut alors s'obtenir qu'en diminuant la valeur des deux rigidités de dérive avant et arrière : cette diminution entraînant une diminution de l'effort transversal généré au niveau des pneumatiques durant le retard dû au phénomène de ballant.

Pourtant la valeur de la rigidité de dérive avant (D1) semble peu affectée par ce phénomène. Si l'hypothèse précédente se révélait exacte, les deux rigidités de dérive avant et arrière seraient affectées. Or, l'étude de la sensibilité des variables de vitesse de lacet et d'accélération transversale a montré la relative insensibilité de ces deux variables vis à vis de la longueur de relaxation avant (b1). Cette faible sensibilité semble expliquer la faible augmentation de la valeur de la rigidité de dérive avant identifiée avec le modèle LaDéBal par rapport à celle identifiée avec le modèle LaDé.

□ Influence du chargement

L'identification des paramètres physiques du véhicule avec trois personnes à bord du véhicule conduit également avec ce modèle à des résultats a priori illogiques (**série 3 et 6**). Avec l'utilisation du modèle LaDéBal, comme pour le modèle LaDé, la présence d'une troisième personne entraîne une diminution des rigidités de dérive avant et arrière (D1 et D2).

A 80 km/h (**série 3 et 4**) la présence d'une troisième personne entraîne une forte diminution (de l'ordre de 25.6 %) de la valeur de la rigidité de dérive arrière.

En revanche à 100 km/h (**série 5 et 6**) les résultats semblent plus cohérents. La masse du passager supplémentaire entraîne une augmentation de la rigidité de dérive avant (1.7%) et une augmentation de la valeur de l'inertie de lacet (6.75%). La diminution de la rigidité de dérive arrière demeure ici inexplicable.

□ **Influence de la vitesse**

L'attribution des variations des paramètres identifiés en fonction de la vitesse (comparaison des valeurs obtenues avec les **séries 4 et 5**) à un phénomène précis est nettement plus délicate. Dans le cas présent, la rigidité de dérive arrière (D2) et l'inertie de lacet (Izz) sont les paramètres les plus affectés. Les valeurs de ces paramètres croissent en fonction de la vitesse : augmentation de 4.6% de la rigidité de dérive arrière (D2) et de 5.2% de l'inertie de lacet.

6.3.2.2 Incertitude sur les paramètres

A des fins de comparaison statistique, et en raison du faible nombre de jeux de données présents au sein des séries 1, 2, 3, 5 et 6 de jeux de données, seuls les résultats issus de la série 4 seront exploités. Sur l'identification des deux rigidités de dérive (D1 et D2) une bonne correspondance des intervalles d'incertitude expérimentaux et de ceux prédits par l'inégalité de Cramer-Rao corrigée (K=2) est obtenue. Par contre, les intervalles d'incertitude expérimentaux sont plus élevés que ceux prédits pour les longueurs de relaxation avant et arrière (b1 et b2).

Les amplitudes des intervalles de confiance relatifs à l'inertie de lacet (Izz) sont approximativement d'amplitudes voisines à ceux associés à l'inertie de lacet identifiée avec le modèle LaDé. En revanche, les intervalles de confiance associés aux deux rigidités de dérive (D1 et D2) sont remarquables. L'augmentation de la complexité du modèle et du nombre de paramètres identifiés ne se traduit pas par une augmentation de l'amplitude des intervalles d'incertitude associés aux paramètres. Bien au contraire, tous les intervalles d'incertitude (expérimentaux et bornes de Cramer-Rao corrigés ou non) des deux rigidités de dérive sont bien inférieurs à ceux associés aux paramètres identifiés avec le modèle LaDé.

Dans la littérature consacrée à l'identification, un unique auteur, Walter [WAL90] aborde ce point. Dans cet article, différents modèles de description du métabolisme d'une drogue administrée par voie orale sont employés. Les propriétés d'identifiabilité des paramètres propres à ces modèles sont étudiées. Il apparaît que l'augmentation de la complexité du modèle et du nombre de paramètres identifiés n'entraîne pas systématiquement une perte d'identifiabilité (ou une augmentation des intervalles d'incertitude associés aux paramètres identifiés).

Dans notre cas, ce phénomène s'observe uniquement sur des paramètres identifiés à partir des données expérimentales. Des intervalles d'incertitude d'amplitudes plus faibles traduisent un gain en identifiabilité des paramètres. Ce gain d'identifiabilité semble être la conséquence du fait que le modèle LaDéBal est plus apte à décrire le comportement du véhicule automobile réel. Le modèle LaDé plus simple, opère un compromis de manière à décrire un comportement moyen du système. Manifestement le modèle LaDé est sous caractérisé et présente d'importantes erreurs de caractérisation.

Ainsi, dans le cas présent, le modèle LaDé ne semble pas posséder un degré de description suffisant pour décrire les différentes dynamiques régissant le comportement latéral du véhicule automobile. En augmentant artificiellement la valeur de la rigidité de dérive arrière, l'algorithme d'optimisation abouti finalement à un modèle moyen reflétant le comportement moyen du véhicule, où les dynamiques de lacet/dérive et de ballant sont assimilées au sein d'une dynamique moyenne

unique de « lacet/dérive/ballant ». Voyons comment se traduisent ces phénomènes sur le plan de Laplace.

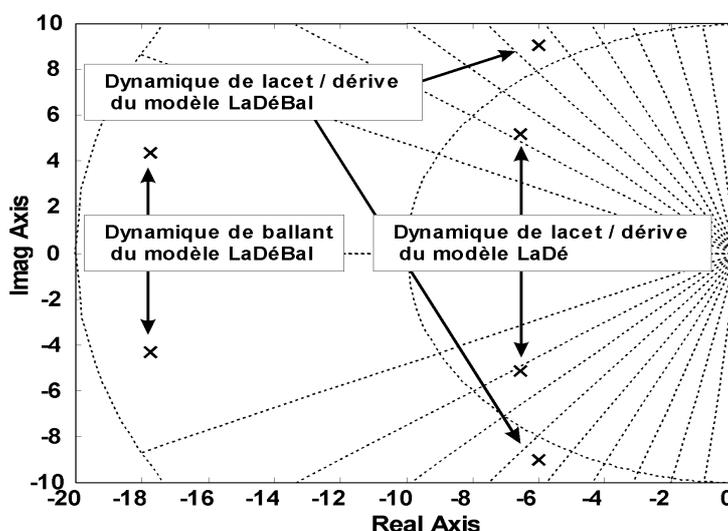


figure 4 - 25 : comparaison de la position des pôles (plan de Laplace) des modèles LaDé et LaDéBal identifiés

Sur la figure 4 - 25 sont représentés respectivement uniquement les pôles du modèle LaDé identifié et les pôles du modèle LaDéBal identifié. Les pôles du modèle LaDé (identifié avec les jeux de données constituant la série 4, à 80 km/h) décrivent un comportement dynamique moyen par le biais de 2 pôles. Ces pôles se situent entre les pôles lents dominants de lacet/dérive et les pôles rapides de ballant du modèle LaDéBal. Les intervalles de confiance plus faibles sur les paramètres identifiés par le biais du modèle LaDéBal démontrent la nécessité de modéliser le comportement latéral du véhicule avec au minimum deux dynamiques différentes, une lente et une rapide. Rien ne garantit a priori la véracité de la description des dynamiques rapides par le phénomène de ballant. Toutefois, les connaissances acquises sur le véhicule et les valeurs des paramètres identifiés accréditent cette hypothèse.

Sous réserve que les autres phénomènes non modélisés (notamment le roulis) exercent une influence moindre sur les valeurs des paramètres identifiés, l'approche multi-modèle adoptée, par augmentation successive de la complexité du modèle LaDé, a l'avantage de permettre dans une certaine mesure la quantification de l'amplitude du biais d'estimation résultant de la non-modélisation du phénomène de ballant au sein du modèle LaDé.

Ainsi dans le cas présent, en supposant une juste caractérisation du véhicule par le modèle LaDéBal, le biais d'estimation sur la rigidité de dérive arrière s'élève à 7541 N/rad soit 7.1 % de la valeur identifiée avec le modèle LaDéBal. Les différences observées sur les autres paramètres sont moins significatives. L'utilisation des modèles de complexité encore plus élevée complétera alors cette première analyse.

6.3.2.3 Globalité et précision des solutions

Le problème d'optimisation posé par l'identification des paramètres physiques du modèle LaDéBal pose quelques difficultés aux algorithmes classiques de type grimpeurs : en effet, suivant l'initialisation de l'algorithme, celui-ci peut tomber dans un minimum local. Expérimentalement, il a été constaté que ces minima locaux sont peu nombreux. Une stratégie de multistart permet aisément de s'affranchir de ces minima. Une solution plus élégante est d'employer l'algorithme d'optimisation AGL, paramétré par un petit nombre d'individus et une ou deux générations. Ainsi,

des individus sont générés aléatoirement sur tout l'espace de recherche. Cette stratégie extrêmement simple de mise en œuvre évite au modélisateur le délicat choix des conditions initiales de l'algorithme d'optimisation. Les résultats présentés dans le tableau 4 - 10 sont calculés au moyen d'un algorithme GQL hybridé avec la méthode du simplexe. L'algorithme génétique est paramétré par 10 individus et une génération. Pour tous les jeux de données, la solution globale a été localisée dans 100% des cas au terme de la première génération avec une précision $\delta = 10^{-4}$. L'identification des paramètres du modèle à partir d'un jeu de données nécessite un temps de calcul moyen de 7 minutes et 30 s.

6.3.2.4 Invalidation des résidus

Par son degré de complexité plus élevé, ce modèle permet une description plus précise des différentes dynamiques du véhicule. Les résultats (cf. tableau 4 - 11) concernent l'invalidation des résidus des 9 jeux de données composant le série 4 (un jeu de données représente donc 11.11%).

Propriétés des résidus	Pourcentage de jeux de données non-invalidé (%)	
	Vitesse de lacet	Accélération transversale
Moyenne nulle	44.44	100
Skewness	22.22	77.77
Kurtosis	22.22	33.33
Stationnarité	0	22.22
Indépendance	0	33.33

tableau 4 - 11 : invalidation des résidus

Avec ce modèle une espérance mathématique nulle des résidus de la vitesse de lacet est plus difficile à atteindre qu'avec le modèle LaDé (4 invalidations ont échoué contre 7 avec le modèle LaDé). En revanche, les skewness et kurtosis des résidus d'accélération transversale se rapprochent plus de ceux d'une distribution normale lorsque les paramètres sont identifiés avec le modèle LaDéBal. De plus, sur deux jeux de données, la stationnarité des résidus de l'accélération transversale n'a pas pu être invalidée (avec le modèle précédent, la stationnarité des résidus d'accélération transversale a toujours été invalidée). Les différences observées, concernant le nombre de jeux de données non-invalidés, diffèrent sur les deux variables de vitesse de lacet et d'accélération transversale s'expliquent par un rapport signal à bruit plus faible sur l'accélération transversale.

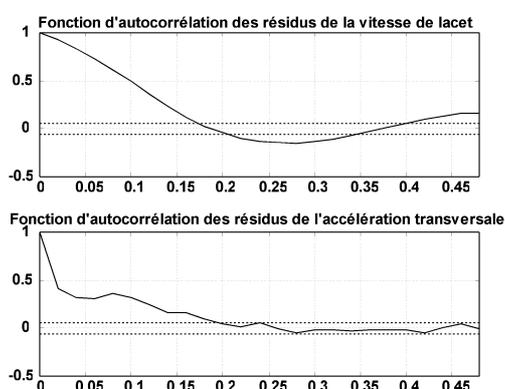


figure 4 - 26a : fonction d'autocorrélation

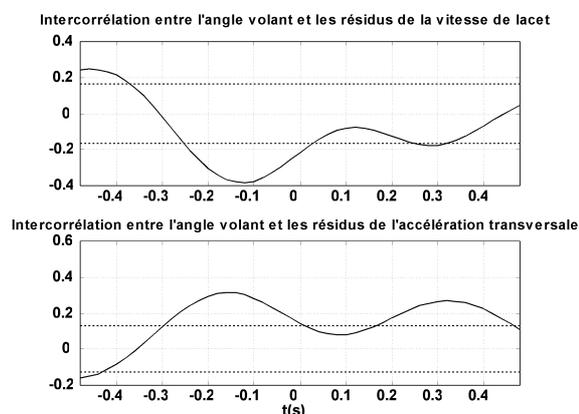


figure 4 - 26b : fonction d'intercorrélation

figure 4 - 26 : blancheur des résidus et corrélation des résidus avec l'entrée

La figure 4 - 26a et figure 4 - 26b illustrent les résultats reproduits tableau 4 - 11. Le tracé de la fonction d'autocorrélation des résidus et de la fonction d'intercorrélation des résidus avec l'entrée illustrent ces propos. Notamment, comparativement avec le modèle LaDé, une plus faible corrélation entre les résidus et l'entrée s'observe. Par rapport aux résidus générés avec le modèle LaDé, les résidus issus de la simulation du modèle LaDéBal identifié s'apparente plus à une suite de variables aléatoires normales. Bien qu'imparfaitement stationnaires, une très nette amélioration de la stationnarité des résidus issus du modèle LaDéBal s'observe par rapport à ceux issus de la simulation du modèle LaDé (cf. figure 4 - 27a). Sans être parfaitement normales leurs distributions s'apparentent plus à une distribution normale (cf. figure 4 - 27b).

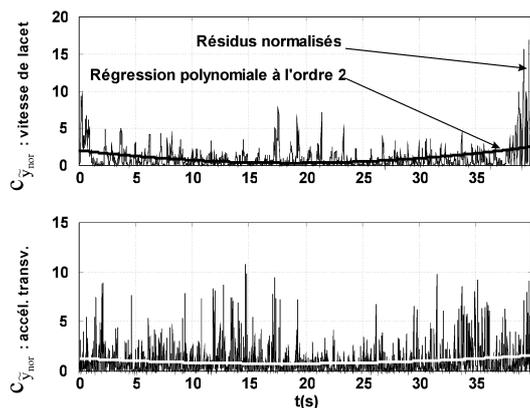


figure 4 - 27a : résidus normalisés

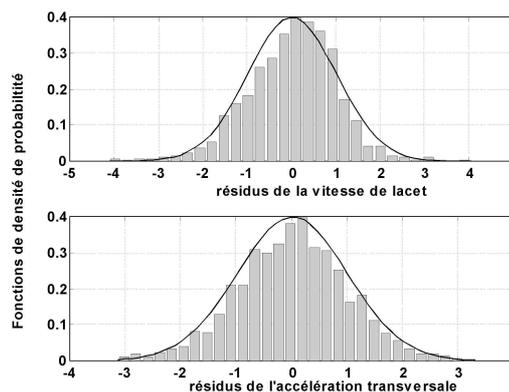


figure 4 - 27b : histogramme des résidus

figure 4 - 27 : stationnarité et distribution des résidus

6.3.2.5 Pouvoir prédictif

q Domaine temporel

Les propriétés statistiques des résidus traduisent sur le plan temporel un pouvoir prédictif accru (cf. figure 4 - 28) par rapport au modèle LaDé. Sans être parfaitement stationnaire, l'amplitude des résidus est visuellement plus stable qu'auparavant.

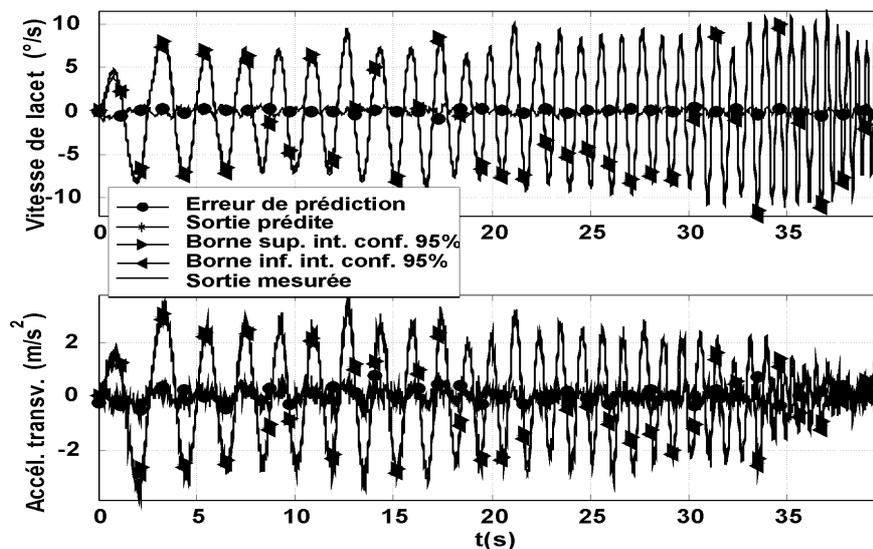


figure 4 - 28 : pouvoir prédictif temporel du modèle identifié

Sur l'exemple présenté sur la figure 4 - 28, les erreurs RMS commises sur la vitesse de lacet et sur l'accélération transversale sont respectivement de $0.31^\circ/s$ et de 0.29 m/s^2 . Les amplitudes de ces erreurs RMS sont approximativement égales à la moitié à celles commises avec le modèle LaDé sur le même jeu de données de validation. Seule différence notable, le jeu de données n'est plus tronqué à la fréquence de 1.5 Hz. La fréquence maximale considérée a été augmentée à 2.2 Hz. La fréquence de 2.2 Hz correspond à la fréquence où un phénomène d'antirésonance (attribué au ballant du pneumatique) est observable sur la variable d'accélération transversale.

□ Domaine fréquentiel

Les résultats précédents dans le domaine temporel traduisent également un bon ajustement fréquentiel, cf. figure 4 - 29. Sur ces courbes, il convient de considérer uniquement les fréquences inférieures à 2.2 Hz, les fréquences supérieures n'étant pas significatives car non prises en compte dans le processus d'estimation : le jeu de données est tronqué aux fréquences supérieures à 2.2 Hz (cf. § 3.6). L'amélioration en terme de pouvoir prédictif fréquentiel par rapport au modèle LaDé est flagrante.

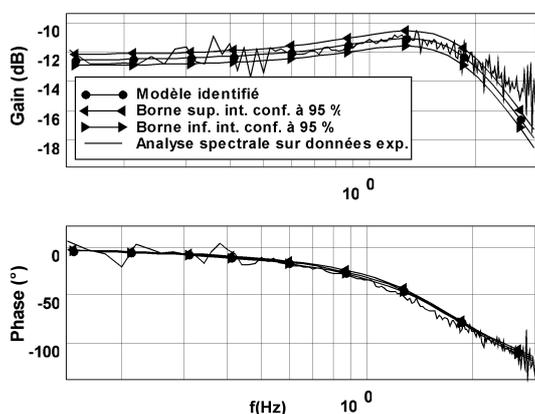


figure 4 - 29a : vitesse de lacet

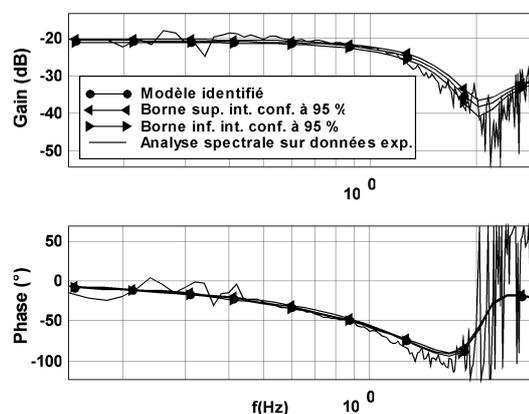


figure 4 - 29b : accélération transversale

figure 4 - 29 : pouvoir prédictif fréquentiel du modèle LaDéBal identifié

Considérons maintenant les apports potentiels liés à l'utilisation des modèles intégrant la description du phénomène de roulis.

6.3.3 Modèle LaRouDé

La description du phénomène de roulis ajouté au modèle LaDé permet du point vu théorique, l'identification de 4 paramètres supplémentaires. Aux paramètres identifiés avec le modèle LaDé, s'ajoute l'inertie de roulis (I_{xx}), la raideur et l'amortissement de roulis (k_{ro} et a_{ro}).

Cependant, en pratique, l'identification simultanée de ces paramètres se révèle numériquement impossible. Les symptômes reflètent vraisemblablement une perte d'identifiabilité du modèle. Le problème d'optimisation est mal conditionné ; il en résulte des intervalles d'incertitude quasi infinis. Cette perte d'identifiabilité peut s'expliquer par :

- La présence de bruit de mesure sur les données.
- Le protocole expérimental différent du cadre idéalisé théorique : la génération du signal d'excitation par le biais d'un opérateur humain peut entraîner l'absence de sollicitation de certaines fréquences.

- La présence d'erreur de caractérisation (absence de description du phénomène de ballant).

En poursuivant la métaphore consistant à comparer la réponse du véhicule à celle d'un moteur à courant continu, il est théoriquement possible dans le cadre idéalisé sans bruit, de distinguer la réponse du modèle simplifié (en négligeant l'inductance dans la partie électromécanique) du moteur, de la réponse du modèle complet (intégrant l'inductance). En présence de bruit de mesure, cette distinction n'est plus possible. En considérant le véhicule, l'identification du produit d'inertie lacet/roulis (I_{xz}) d'influence mineure (cf. analyse de la sensibilité des modèles) simultanément avec les autres paramètres, est expérimentalement et numériquement impossible pour le protocole expérimental défini.

Les valeurs des paramètres identifiables et identifiés par l'intermédiaire de ce modèle sont reproduits dans le tableau 4 - 12.

6.3.3.1 Paramètres identifiés

Les valeurs obtenues pour les 3 paramètres communs aux modèles LaDé et LaDéBal (D1, D2 et I_{zz}) semblent en accord avec les données constructeur et les valeurs précédemment obtenues.

θ	$\hat{\theta}_{mv}$ série 1 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 2 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 3 V = 80 km/h M = 1475 kg		$\hat{\theta}_{mv}$ série 4 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 5 V = 100 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 6 V = 100 km/h M = 1475 kg	
	D1 N/rad	76116	-- 6.00 3.52	74492	-- 5.38 3.16	76087	-- 7.64 4.49	75039	5.51 5.72 3.36	75323	-- 7.12 4.18	76384
D2 N/rad	89041	-- 7.68 4.51	89124	-- 7.27 4.28	74036	-- 7.70 4.52	89077	10.9 7.49 4.41	89183	-- 5.72 3.36	87221	-- 8.54 5.02
I_{zz} kg.m ²	1468	-- 16.9 9.95	1538	-- 19.5 11.5	1525	-- 11.5 6.76	1496	9.26 18.1 10.6	1573	-- 14.8 8.70	1684	-- 14.1 8.30
I_{xx} kg.m ²	111	-- 96.1 56.5	202	-- 19.4 11.4	167	-- 27.0 15.9	151	92.0 62.0 36.5	32.5	-- 123 72.5	70.8	-- 338 199
kro N/rad	520626	-- 61.1 35.9	689684	-- 58.8 34.6	780348	-- 63.0 15.9	489543	69.4 60.1 35.3	699307	-- 130 76.4	989756	-- 216 127
aro Nm/rad/s	-913	-- 71.0 41.8	-957	-- 67.3 39.6	-617	-- 88.6 52.2	-932	28.0 69.4 40.8	-487	-- 129 76.0	-299	-- 251 147

tableau 4 - 12 : paramètres estimés avec le modèle LaRouDé

En revanche, les valeurs de l'inertie de roulis (I_{xx}) estimées à partir des essais réalisés à 80 km/h sont de l'ordre du triple de la valeur constructeur ($I_{xx} = 50 \text{ kg.m}^2$) et sont surprenantes. Toutefois les intervalles d'incertitude associés à ce paramètre (de l'ordre de $\pm 90 \%$ avec un degré de confiance de 95%) justifient les valeurs obtenues. Malheureusement aucun élément supplémentaire ne permet d'évaluer la véracité des valeurs obtenues.

De plus, les valeurs de la raideur de roulis (kro) et de l'amortissement de roulis (aro) posent un certain nombre d'interrogations. En premier lieu, l'amplitude de la raideur de roulis est anormalement élevée. Les valeurs obtenues sont au moins trois fois supérieures aux valeurs nominales

constructeur. Les intervalles de confiance constructeur associés à la valeur nominale de ce paramètre ($\pm 30\%$) confirment le caractère excessif de ce résultat. En second lieu la valeur de l'amortissement de roulis est également surprenante. Sa valeur négative et sa faible amplitude comparée à sa valeur nominale ($\text{aro} = 2750 \text{ Nm/rad/s}$) sont a priori inexplicables. Ceci traduit vraisemblablement le fait que pour le protocole expérimental employé, ce paramètre n'est pas suffisamment sollicité et est par conséquent non identifiable.

Ces résultats quelques peu surprenants sont vraisemblablement liés à la présence d'erreur de caractérisation. Pour les sollicitations employées (excitation harmonique), l'expérience acquise confirme la prépondérance du phénomène de ballant sur le comportement latéral du véhicule. Ainsi, à la fréquence d'antirésonance (2.2 Hz) présent sur la variable d'accélération transversale, le véhicule n'est plus animé d'aucune accélération transversale ni de mouvement de roulis. L'algorithme d'optimisation ajuste les paramètres du modèle LaRouDé de manière à obtenir une identité de comportement modèle avec le système réel. Or, pour tenir compte de la présence de ce retard propre au phénomène de ballant lors de l'identification des paramètres du modèle LaRouDé, l'algorithme d'optimisation dispose des six degrés de liberté que sont les six paramètres identifiés (D1, D2, lzz, lxx, kro et aro). Il semble que l'absence de modélisation du phénomène de ballant soit la cause des estimations quelques peu excentriques de la raideur et de l'amortissement de roulis. L'identification des paramètres de roulis et de ballant par le biais du modèle LaRouDéBal validera ou invalidera cette hypothèse.

Les mauvais résultats obtenus en terme de véricité des valeurs des paramètres ne justifient pas l'analyse détaillée des intervalles d'incertitude associés aux paramètres, des propriétés statistiques des résidus et des pouvoirs prédictifs du modèle ainsi identifié.

6.3.3.2 Globalité et précision des solutions

Le problème d'optimisation posé par l'identification des paramètres physiques du modèle LaRouDé pose de sérieuses difficultés en terme de minima locaux. Pour résoudre ce problème d'optimisation par le biais d'algorithmes de type grimpeur, une stratégie de multistart de type « bracketting » particulièrement gourmande en ressources informatiques pourrait être employée. Les algorithmes d'optimisation globale localisent aisément l'optimum global avec une précision $\delta = 10^{-4}$ tout en réclamant une puissance de calcul plus raisonnable.

A puissance machine consommée quasi identique, l'algorithme AGLX offre des performances bien supérieures à l'algorithme AGL (cf. deuxième chapitre). L'algorithme AGLX²⁹ a donc été retenu. L'utilisation d'algorithme d'optimisation globale accroît significativement le temps de calcul nécessaire à l'identification des paramètres du modèle à partir d'un jeu de données : ce temps est en moyenne égal à 2 heures et 10 minutes.

Voyons maintenant si la prise en compte des phénomènes de lacet/dérive, de roulis et de ballant conduit à des résultats plus probants en terme de valeurs des paramètres identifiés.

6.3.4 Modèle LaRouDéBal

Ce modèle permet d'identifier les deux rigidités de dérive (D1 et D2), l'inertie de lacet (lzz), la raideur et l'amortissement de roulis (kro et aro) et les deux longueurs de relaxation avant et arrières (b1 et b2). Les résultats obtenus à partir des différentes séries de jeux de données sont reproduits dans le tableau 4 - 13.

²⁹ AGLX paramétré comme suit : pression de sélection de 1.2, nombre de génération égal à 9, une population est constituée de 28 individus, les probabilités de recombinaison [1/3 discrète, 1/3 continue, 1/3 simplex], absence de mutation, stratégie élitiste de réinsertion.

6.3.4.1 Paramètres identifiés

Contrairement au modèle LaRouDé précédent, les valeurs des paramètres identifiés sont ici réalistes. Toutefois, l'amplitude des paramètres identifiés communs aux modèles LaDé, LaDéBal et LaRouDéBal diffèrent sensiblement (paramètres identifiés à partir de la série 4 de jeux de données). Comparons maintenant les valeurs des paramètres (communs à tous les modèles) identifiés par le biais du modèle LaRouDéBal à celles obtenues par l'intermédiaire des autres modèles.

Les valeurs des rigidités de dérive et de l'inertie de lacet sont plus élevées que celles obtenues à partir du modèle LaDé (+9.3% sur la rigidité de dérive avant D1, +18.3% sur la rigidité de dérive arrière D2 et + 3.1 % sur l'inertie de lacet lzz).

Dans une mesure moindre, ce phénomène s'observe également sur les paramètres identifiés à partir du modèle LaDéBal (+7.3% sur la rigidité de dérive avant D1, +11.6% sur la rigidité de dérive arrière D2, +8.5 % sur l'inertie de lacet lzz et +37.8% sur la longueur de relaxation arrière b2). Face à l'augmentation de l'amplitude de certains paramètres, une nette diminution s'observe sur la longueur de relaxation avant (b1 : -23.1%). D'un point de vue phénoménologique, l'attribution de ces différences à un phénomène particulier est délicate. D'un point de vue comportemental, la proximité des dynamiques (cf. tableau 4 - 1) mises en jeu par les pôles de roulis, de ballant et de lacet/dérive semble bien être la cause de ces différences.

θ	$\hat{\theta}_{mv}$ série 1 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 2 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 3 V = 80 km/h M = 1475 kg		$\hat{\theta}_{mv}$ série 4 V = 80 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 5 V = 100 km/h M = 1400 kg		$\hat{\theta}_{mv}$ série 6 V = 100 km/h M = 1475 kg	
D1 N/rad	82501	--	80892	--	83063	3.68	81786	2.69	78468	--	83485	--
		2.54		2.32				3.68		2.44		3.34
		1.69		1.55				2.45		1.63		2.22
D2 N/rad	110884	--	106808	--	90123	4.38	109072	3.21	108291	--	105994	--
		3.68		3.29				4.38		3.51		4.50
		2.45		2.19				2.92		2.34		3.00
lzz kg.m ²	1831	--	1872	--	1875	4.85	1849	3.71	1881	--	1947	--
		3.84		3.68				4.85		3.77		5.12
		2.56		2.45				3.23		2.51		3.41
lxx kg.m ²	92.1	--	75.5	--	70.7	26.1	84.9	68.5	121	--	60	--
		23.6		25.7				26.1		24.5		47.6
		15.8		17.11				17.4		16.4		31.7
kro N/rad	182902	--	155648	--	134991	14.8	170789	9.22	351169	--	210271	--
		12.8		12.2				14.8		12.5		27.3
		8.53		8.13				9.88		8.35		18.2
aro Nm/rad/s	329	--	434	--	1900	141	376	119	1089	--	337	--
		155		171				141		162		206
		103		114				93.8		108		137
b1 m	0.38	--	0.40	--	0.54	80.8	0.39	71.1	0.52	--	0.44	--
		49.7		54.1				80.8		51.7		79.4
		33.1		36.1				53.9		34.5		52.9
b2 m	0.71	--	0.77	--	0.79	7.94	0.74	7.74	0.69	--	0.77	--
		5.74		5.65				7.94		5.70		11.8
		3.82		3.77				5.29		3.80		7.87

tableau 4 - 13 : paramètres estimés avec le modèle LaRouDéBal

Cependant, la comparaison avec les données constructeurs, des valeurs de la raideur et de l'amortissement de roulis (positif dans le cas présent) plaide en faveur de la véracité des paramètres identifiés. Plus généralement, tous les paramètres identifiés présentent des valeurs réalistes et proches des valeurs nominales du constructeur.

En supposant une parfaite caractérisation du véhicule par le biais du modèle LaRouDéBal, les différences constatées sur l'identification des paramètres communs aux trois modèles LaDé, LaDéBal et LaRouDéBal permettent une quantification de l'erreur de biais présente sur les paramètres identifiés (l'erreur de biais due à la présence d'erreur de caractérisation).

☞ **Influence du chargement**

De manière similaire aux modèles précédents, l'identification des paramètres physiques avec une troisième personne à bord du véhicule conduit à des résultats a priori illogiques (**série 3 et 6**). Avec ce modèle, la présence d'une troisième personne à bord entraîne une diminution des rigidités de dérive avant et arrière (D1 et D2).

A 80 km/h (**série 3 et 4**), la présence d'une troisième personne entraîne une diminution de la rigidité de dérive arrière de 17.4%. En ce qui concerne les paramètres en relation avec le phénomène de roulis, la valeur de la raideur de roulis chute de 21% tandis que la valeur de l'amortissement croît de 405%. L'inertie de roulis est quant à elle peu affectée par la présence d'une personne supplémentaire.

A 100 km/h (**série 5 et 6**), la variation de la valeur de l'inertie de lacet semble plus logique (augmentation 3.5%). En revanche, les diminutions de 100% des valeurs de l'inertie de roulis I_{xx} , et de 223% de l'amortissement de roulis, sont difficilement explicables. Quelles que soient les séries de jeux de données, les valeurs identifiées de l'inertie de roulis paraissent faibles au regard de sa valeur nominale et de l'écart type associé.

☞ **Influence de la vitesse**

Contrairement aux résultats issus de l'utilisation du modèle LaDéBal, certains paramètres identifiés (rigidités de dérive avant et arrière, inertie de lacet et longueur de relaxation arrière) avec le modèle LaRouDéBal ne semblent plus être affectés par un changement de vitesse du véhicule (**série 4 et 5**). En revanche, les valeurs de la raideur et de l'amortissement de roulis varient dans de grandes proportions (+105% sur la raideur de roulis, +189% sur l'amortissement de roulis). Dans une proportion moindre, la longueur de relaxation avant subit une augmentation de 33.3%.

6.3.4.2 Incertitudes sur les paramètres

A des fins de comparaisons statistiques, seuls les résultats issus de la série 4 seront exploités. Une bonne correspondance des intervalles de confiance à 95% expérimentaux et de ceux prédits par l'inégalité de Cramer-Rao corrigée ($K=1.2$), est obtenue sur la majorité des paramètres estimés (rigidités de dérive, inerties de lacet et de roulis, longueur de relaxation arrière). Une adéquation dégradée est obtenue sur les intervalles d'incertitude associés à la longueur de relaxation avant ainsi que ceux associés à la raideur et l'amortissement de roulis.

Notons également que les intervalles d'incertitude expérimentaux et de ceux prédits par l'inégalité de Cramer-Rao corrigées sur les paramètres communs aux modèles LaDé, LaDéBal et LaRouDéBal sont d'amplitude plus faible lorsque les paramètres sont identifiés par le biais du modèle LaRouDéBal. Ce fait tend à démontrer l'importance de la modélisation du phénomène de roulis. Il apparaîtrait donc que le modèle LaRouDéBal identifié n'est pas surparamétré. Ainsi, dans la mesure où l'identifiabilité des paramètres est conservée, une augmentation supplémentaire de son degré de description est souhaitable.

6.3.4.3 Globalité et précision des solutions

Comme pour l'identification du modèle LaRouDé, l'identification des paramètres physiques du modèle LaRouDéBal nécessite l'utilisation d'algorithmes d'optimisation globale particulièrement puissants. L'algorithme d'optimisation AGLX³⁰ a été retenu pour sa polyvalence et ses performances. Les optima globaux sont localisés dans 100% des cas avec une précision $\delta = 10^{-4}$.

	Algorithme du simplex			Algorithme AGLX		
	$\hat{\theta}$	$\Delta J (-)$	$\Delta J (+)$	$\hat{\theta}$	$\Delta J (-)$	$\Delta J (+)$
D1 (N/rad)	78917	0.01	0.98	82626	0.50	0.49
D2 (N/rad)	103597	4.49	-3.47	110949	0.49	0.50
lzz (kg.m ²)	1815	-7.42	8.41	1869	0.50	0.49
lxx(kg.m ²)	126	-2.18	2.52	146	0.38	0.32
kro (N/rad)	169856	6.61	-5.57	189443	0.49	0.48
aro (Nm/rad/s)	452	1.58	-0.38	428	0.73	0.55
b1 (m)	0.33	-7.27	8.27	0.34	0.50	0.50
b2 (m)	0.75	4.58	-3.58	0.73	0.49	0.49

tableau 4 - 14 : vérification de la globalité des solutions

A titre de comparaison, dans le tableau 4 - 14 sont reproduits les résultats obtenus sur un jeu de données par l'intermédiaire de deux algorithmes d'optimisation différents (algorithme du simplex et AGLX). Cette comparaison démontre l'indéniable efficacité de l'algorithme AGLX face à celui du simplex. Ces performances accrues se paient par un temps de calcul plus élevé pour l'algorithme AGLX (2 heures et 50 minutes) que pour l'algorithme du simplex (3 minutes).

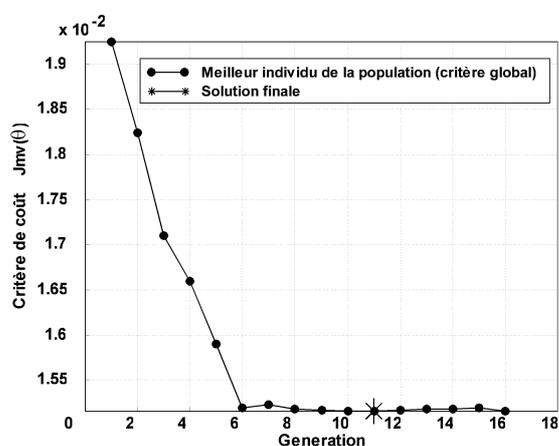


figure 4 - 30a : espace d'état

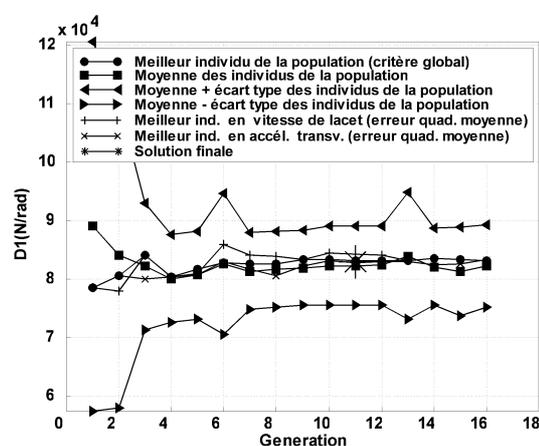


figure 4 - 30b : espace paramétrique

figure 4 - 30 : convergence de l'algorithme AGLX vers l'optimum global

La figure 4 - 30 illustre la convergence de l'algorithme d'optimisation AGLX vers l'optimum global lors de l'identification des paramètres de ce modèle. Dans l'espace d'état, la figure 4 - 30a montre

³⁰ AGLX paramétré comme suit : pression de sélection de 1.2, nombre de génération égal à 9, une population est constituée de 36 individus, les probabilités de recombinaison [1/3 discrète, 1/3 continue, 1/3 simplex], absence de mutation, stratégie élitiste de réinsertion.

l'évolution de la valeur de la fonction objectif du meilleur individu de la population. L'algorithme a convergé à l'optimum global au terme de la 13^o génération.

Dans l'espace paramétrique, l'évolution au fil des générations de la valeur de la rigidité de dérive avant D1 est reproduite sur la figure 4 - 30b. Les valeurs de ce paramètre (cf. figure 4 - 30a) :

- du meilleur individu (au sens du critère d'erreur multivariable du maximum de vraisemblance) de la population.
- de la moyenne de l'ensemble des individus composant la population.
- des écarts types autour de cette moyenne.
- des meilleurs individus minimisant chaque critère d'erreur sur les variables de vitesse de lacet et d'accélération transversale (au sens de l'erreur quadratique moyenne sur chacune de ces variables).

Notons que la variable de vitesse de lacet a tendance à augmenter la valeur de la rigidité de dérive avant tandis que la variable d'accélération transversale tend à diminuer sa valeur. La convergence de l'algorithme d'optimisation s'observe également au terme de la 13^o génération. La réduction des écarts types autour de la moyenne traduit la convergence de l'ensemble de la population vers une solution stable dans l'espace paramétrique.

6.3.4.4 Invalidation des résidus

Propriétés des résidus	Pourcentage de jeux de données non-invalidé (%)	
	Vitesse de lacet	Accélération transversale
Moyenne nulle	44.44	100
Skewness	44.44	66.66
Kurtosis	33.33	22.22
Stationnarité	33.33	44.44
Indépendance	0	33.33

tableau 4 - 15 : invalidation des résidus

Le modèle LaRouDéBal possédant le degré de description le plus élevé, il semble naturel que les résultats en terme d'invalidation des résidus soient meilleurs que ceux obtenus à partir des modèles précédents. Les résultats (cf. tableau 4 - 15) concernent l'invalidation des résidus des 9 jeux de données composant le série 4 (un jeu de données représente 11.11%).

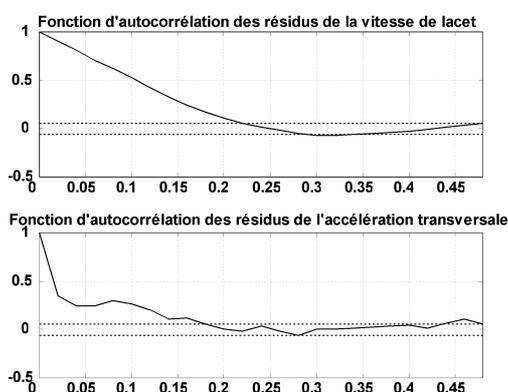


figure 4 - 31a : fonction d'autocorrélation

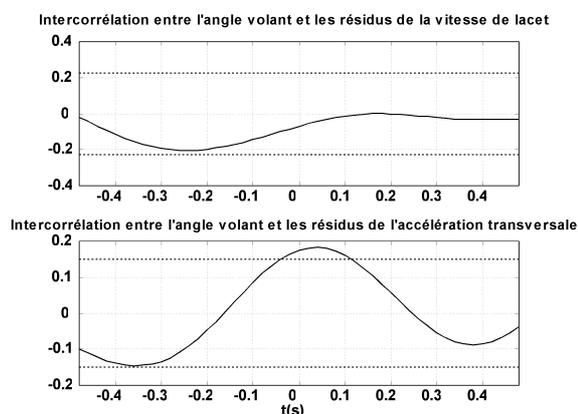


figure 4 - 31b : fonction d'intercorrélation

figure 4 - 31 : blancheur des résidus et corrélation des résidus avec l'entrée

Globalement, le pourcentage de jeux de données non-invalidés croit significativement. Par rapport au modèle LaDéBal, les principales améliorations concernent la stationnarité et la normalité des résidus de la vitesse de lacet. Malgré le fait que l'indépendance des résidus ait été invalidée sur la variable de vitesse de lacet, les tracés de la fonction d'autocorrélation des résidus et de la fonction d'intercorrrelation des résidus avec l'entrée montrent une nette amélioration. Notamment, une corrélation moindre des résidus de la vitesse de lacet et de l'accélération transversale avec l'entrée (angle volant) est observable. Ces améliorations sensibles se visualisent également sur la figure 4 - 31.

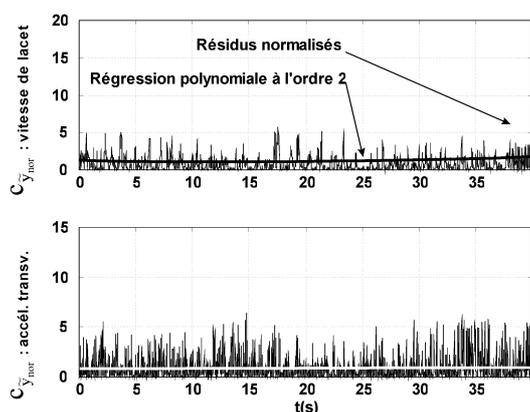


figure 4 - 32a : résidus normalisés

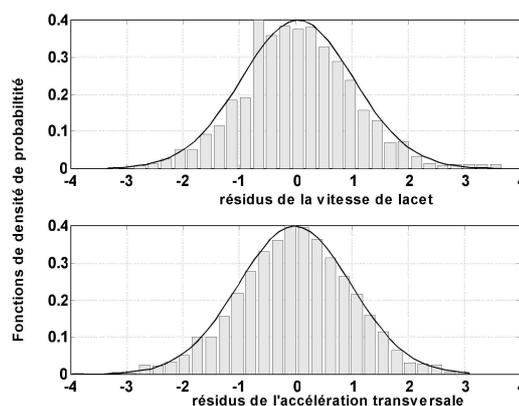


figure 4 - 32b : histogramme des résidus

figure 4 - 32 : stationnarité et distribution des résidus

La figure 4 - 32 montre clairement les améliorations réalisées en terme de stationnarité et de distribution des résidus comparativement au modèle LaDéBal.

6.3.4.5 Pouvoir prédictif

□ Domaine temporel

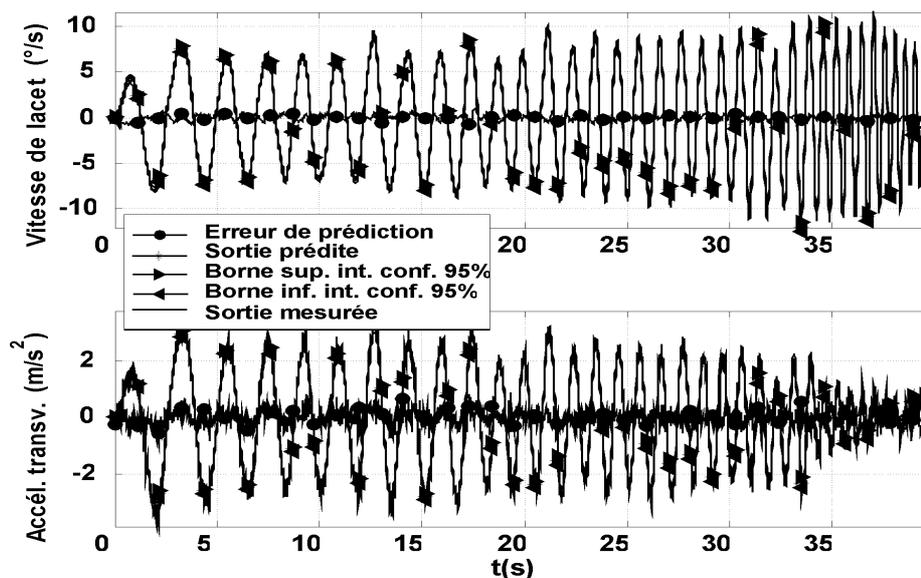


figure 4 - 33 : pouvoir prédictif temporel du modèle identifié

Du point de vue du pouvoir prédictif, dans le domaine temporel, l'amélioration est difficilement appréciable par inspection visuelle. Sur l'exemple reproduit sur la figure 4 - 33, les erreurs RMS commises sur les variables de vitesse de lacet et d'accélération transversale sont respectivement de $0.30^\circ/s$ et de 0.27 m/s^2 . Ainsi, l'amélioration en terme de pouvoir prédictif du modèle LaRouDéBal par rapport au modèle LaDéBal est faible (erreurs RMS sur les variables de vitesse de lacet et d'accélération transversale respectivement de $0.31^\circ/s$ et de 0.29 m/s^2).

¶ Domaine fréquentiel

Comme pour le domaine fréquentiel, l'amélioration en termes de pouvoir prédictif entre le modèle LaDéBal et LaRouDéBal est ténue (cf. figure 4 - 34)

Si la différence observée entre les pouvoirs prédictifs des modèles LaDéBal et LaRouDéBal est ténue dans le domaine temporel, une amélioration sensible s'observe dans le domaine fréquentiel. Notamment sur la variable de vitesse de lacet aux fréquences supérieures à 1,5 Hz (cf. figure 4 - 34a).

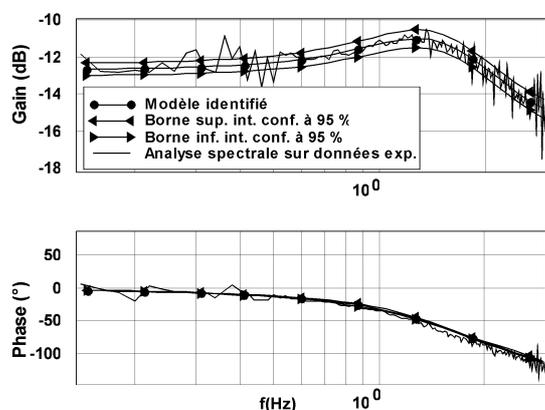


figure 4 - 34a : vitesse de lacet

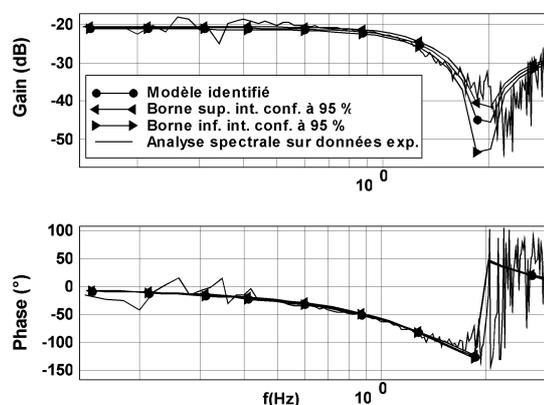


figure 4 - 34b : accélération transversale

figure 4 - 34 : pouvoir prédictif fréquentiel du modèle LaRouDéBal identifié

6.3.5 Identification robuste

Identifions maintenant les paramètres physiques du véhicule automobile par l'intermédiaire des techniques d'identification robustes. Voyons si les résultats sont robustes vis à vis d'éventuelles données aberrantes présentes dans les différents jeux de données expérimentaux. Dans ce but, nous allons uniquement considérer la série 4 de jeux de données (constituée du nombre de jeux de données le plus élevé ; la vitesse moyenne du véhicule est de 80 km/h, 2 personnes sont présentes à bord, le véhicule possède alors une masse totale de 1400 kg). Les résultats sont alors successivement comparés en terme de valeurs des paramètres estimés par minimisation des critères de coûts suivants :

- Maximum de vraisemblance non-robustifié, données non filtrées.
- Maximum de vraisemblance non-robustifié, données filtrées par l'intermédiaire d'un filtre de Butterworth d'ordre 8 de fréquence de coupure de 4 Hz, à déphasage nul (filtre aller-retour).
- Maximum de vraisemblance robustifié par fonction de Huber, données non filtrées.
- Maximum de vraisemblance robustifié par fonction de Tukey, données non filtrées.
- Maximum de vraisemblance robustifié par fonction de Ljung, données non filtrées.

Les conclusions relatives à ces résultats demeurent identiques quel que soit le modèle employé. Dans un souci de concision et sans perte de généralité, les résultats reproduits dans le tableau 4 -

16 résultent de l'identification des paramètres du modèle LaDéBal par l'intermédiaire des différentes stratégies de robustification précitées.

Les résultats retranscrits dans ce tableau reflètent la robustesse des valeurs des paramètres identifiés vis à vis de l'éventuelle présence de données aberrantes. Les valeurs des paramètres identifiés sont d'amplitudes sensiblement identiques.

$\rho(\cdot)$	m.v données non filtrées.	m.v données filtrées.	Huber	Tukey	Ljung
D1 (N/rad)	75806	76208	75753	75772	75565
D2 (N/rad)	96450	96825	97132	97142	97162
lzz (kg.m ²)	1692	1720	1699	1703	1713
b1 (m)	0.48	0.46	0.48	0.48	0.48
b2 (m)	0.46	0.47	0.47	0.47	0.48

tableau 4 - 16 : comparaison des différents schémas de robustification

Les résultats issus de l'identification des paramètres physiques à partir des données filtrées sont quasiment égaux à ceux issus de l'identification à partir de données non filtrées. Le filtrage des données ne se justifie donc pas. De plus le filtrage tend à détruire l'hypothèse d'indépendance du bruit de mesure, donc des résidus (hypothèse posée lors de la formulation du critère de coût au sens du maximum de vraisemblance).

Les résultats issus de l'identification des paramètres physiques par minimisation de critères de coût robustifiés sont quasiment égaux à ceux obtenus par le biais des critères non-robustifiés. La robustification du critère du maximum de vraisemblance ne se justifie donc pas dans le cas présent.

L'identification des paramètres physiques des autres modèles présentés dans ce mémoire aboutit à des conclusions en tout point similaires.

7 CONCLUSION

Dans ce chapitre consacré à l'identification des paramètres physiques permettant la description du comportement transversal du véhicule, différents problèmes ont été soulevés. Notamment, une perte d'identifiabilité des paramètres a été constatée lors du passage du cadre théorique idéalisé au cadre expérimental. Vraisemblablement, un protocole expérimental parfaitement reproductible et proche du protocole théorique résoudraient ce type de difficultés. Une reproductibilité parfaite du signal d'excitation appliqué au volant est atteignable en générant l'excitation par l'intermédiaire d'effecteurs.

Compte tenu des limitations inhérentes à l'identifiabilité des paramètres et des choix réalisés en fonction de l'analyse de la sensibilité des paramètres, les principaux paramètres inconnus ou difficilement mesurables ont été identifiés. Les valeurs de ces paramètres, comparées aux données constructeur nominales sont réalistes lorsque ces paramètres sont identifiés à partir de trois des quatre modèles développés. Il apparaît notamment que la modélisation du phénomène de roulis, en faisant abstraction du phénomène de ballant induit la présence d'un biais d'estimation inacceptable sur la raideur et l'amortissement de roulis.

La comparaison des valeurs des paramètres identifiés, communs aux trois modèles donnant lieu à une estimation réaliste, se révèle très instructive. En particulier, les causes des différences constatées sur la valeur de la rigidité de dérive arrière estimée par le biais des modèles LaDé et LaDéBal a pu être expliquée. En revanche, en fonction des conditions expérimentales (vitesse du

véhicule, chargement ...) les variations de certains paramètres n'ont pas été expliquées de manière satisfaisante.

Afin de valider les résultats reproduits dans ces travaux, il serait intéressant de disposer des infrastructures de grands industriels du monde automobile. Celles-ci permettraient de valider ou d'invalider les résultats obtenus, et notamment ceux en rapport avec les inerties, les raideurs et les amortissements qui demeurent difficiles à estimer avec précision. La comparaison des paramètres mesurés sur banc de mesure avec les paramètres identifiés, permettrait sans aucun doute d'effectuer un choix parmi les différents modèles présentés et/ou de redéfinir leur degré de description. Ces mesures sur banc permettraient également l'élaboration d'indicateurs de véracité des valeurs des paramètres identifiés.

Du point de vue des intervalles d'incertitude, les intervalles de confiance associés aux paramètres estimés (rigidités de dérive avant et arrière, inertie de lacet et longueur de relaxation arrière) sont très nettement inférieurs aux intervalles de confiance constructeur. En revanche, les paramètres en rapport avec le mouvement de roulis (inertie et amortissement de roulis) sont estimés avec une précision inférieure, de l'ordre de grandeur des intervalles de confiance constructeur. La mesure des variables de vitesse de roulis ou d'angle de roulis, et l'introduction de celle-ci dans la formulation du critère de coût multivariable au sens du maximum de vraisemblance diminuerait significativement l'amplitude des intervalles de confiance associés à ces paramètres (fait démontré en simulation).

En considérant le seul pouvoir prédictif des modèles identifiés, le modèle bicyclette très fréquemment utilisé pour l'étude du comportement des véhicules automobile, possède un pouvoir prédictif relativement limité. L'ajout de la description du phénomène de ballant au modèle bicyclette augmente significativement son pouvoir prédictif temporel (diminution des erreurs RMS d'un facteur 2). Du seul point de vue du pouvoir prédictif, le modèle LaDéBal offre un bon compromis pouvoir prédictif/complexité. L'ajout du phénomène de roulis au modèle LaDéBal influe peu sur son pouvoir prédictif. Il entraîne en revanche des différences notables sur les paramètres estimés.

Enfin, quel que soit le modèle employé, il a été montré que les valeurs des paramètres identifiés sont robustes vis à vis de l'éventuelle présence de données aberrantes sur les variables entrant dans la composition du critère d'erreur minimisé. Ainsi, le filtrage des données ou l'utilisation de critères de coût robustifiés, ne se justifient pas dans le cas présent.

CONCLUSION

Les travaux présentés dans ce mémoire, représentent **une contribution à l'identification des paramètres physiques de systèmes complexes**. Les différentes étapes propres à une procédure de modélisation constitue le fil conducteur de ce mémoire :

- Modélisation de système physique.
- Vérification des propriétés structurelles des modèles.
- Choix approprié d'un critère de coût.
- Etablissement d'un protocole expérimental et procédure de recueil des données.
- Estimation des paramètres par minimisation du critère de coût précité.
- Validation du modèle.

Ces différentes étapes revêtent plusieurs aspects et les travaux présentés dans ce mémoire apportent une contribution dans plusieurs disciplines.

¶ **L'identification des paramètres physiques de systèmes complexes et du véhicule automobile en particulier**

Dans la littérature consacrée à l'identification des paramètres physiques, les différents points précités sont trop souvent traités de manières découplées et/ou sur des cas d'écoles. La transposition de ces techniques basées sur des bases théoriques, à un système réel perturbé et bruité est rarement triviale. La justification des différents choix effectués compte tenu de leurs avantages et inconvénients vise à constituer **une stratégie générale d'identification des paramètres physiques de systèmes physiques quelconques**.

Notamment, les différents aspects, fréquemment passés sous silence, relatifs à l'identifiabilité des paramètres sont abordés. Afin d'éviter dans la mesure du possible une estimation biaisée des paramètres, ceux-ci sont estimés **par minimisation d'un critère multivariable d'erreur de sortie formulé au sens du maximum de vraisemblance**. Pour ces mêmes raisons, aucun terme de pondération en faveur d'une valeur a priori des paramètres n'a été inclus dans la formulation du critère de coût.

Il a été montré que l'identification des paramètres physiques de systèmes complexes nécessitant l'emploi d'algorithme itératif d'optimisation non-linéaire gagne à employer des algorithmes d'optimisation globale. Dans la phase de validation des modèles identifiés, une quantification réaliste et efficace des intervalles d'incertitude associés aux paramètres a été réalisée. Cette quantification

s'inspire de méthodes employées par la NASA lors de l'identification des coefficients aérodynamiques d'aéronefs.

Par son caractère général, la démarche proposée est applicable à de nombreux processus linéaires en leur(s) entrée(s).

☞ **L'optimisation non-linéaire**

Les choix réalisés pour l'identification des paramètres justifient l'emploi d'algorithmes itératifs de recherche globale. A ces fins, différents algorithmes furent employés et leurs performances comparées. Les limitations inhérentes aux algorithmes de type grimpeur ont été soulignées : la localité des optima localisés n'est pas satisfaisante du point de vue du modélisateur.

Par leurs capacités d'exploration supérieures, le choix des algorithmes génétiques est alors parfaitement justifié. Cependant, ces algorithmes requièrent une grande quantité de ressources machines et offrent des résultats limités en termes de précision.

Sur la base de publications et d'ouvrages, des algorithmes d'optimisation globale furent alors implémentés et employés. Lors de benchmarks sur des fonctions de test communément répandues dans la communauté des optimiseurs, il a été montré **qu'une judicieuse combinaison des algorithmes génétiques avec des algorithmes de types grimpeurs** aboutissait à des algorithmes offrant **d'excellentes capacités d'explorations alliées à de puissantes capacités d'exploitation**, tout en requérant un **temps de calcul raisonnable**. De plus, ces algorithmes se révèlent robustes vis à vis du réglage de leurs paramètres internes.

☞ **La modélisation de véhicule automobile**

Dans le but de décrire le comportement transversal du véhicule automobile de série, différents modèles de complexités croissantes furent développés. Ces modèles offrent des moyens de simulation et d'analyse du comportement de la majorité des véhicules automobiles légers en circulation.

L'influence du phénomène de ballant du pneumatique sur le comportement latéral global du véhicule a été soulignée. L'enrichissement progressif d'un modèle simplifié est propre à favoriser l'étude de l'influence respective des différents phénomènes régissant la dynamique du véhicule automobile.

Une suite logique à ces travaux serait sans aucun doute liée au développement des principaux points suivants :

- ☞ La mise en place **d'indicateurs de véracité** sur les valeurs des paramètres identifiés.
- ☞ L'utilisation de techniques de **planification d'expérience optimale** dans le but de minimiser les incertitudes sur les paramètres estimés.
- ☞ L'étude de l'apport des méthodes d'estimation œuvrant dans le domaine fréquentiel.
- ☞ L'identification des caractéristiques non-linéaires du véhicule et des pneumatiques en particulier (travaux actuellement en cours au laboratoire).

Aux résultats relevant directement de ces travaux, ajoutons les implications et les perspectives suivantes :

- ☞ L'association des modèles développés aux logiciels de simulation et d'optimisation développés, constitue des outils propres à prédire le comportement de véhicules en phase de **conception** et d'essais de qualification, entraînant ainsi une baisse des temps et des coûts de fabrication et de mise au point. Dans cette optique, il serait particulièrement avantageux d'associer les modèles de descriptions élevées développés par Chamailard [CHA96] aux outils présentés dans ce mémoire.

- De manière complémentaire au processus de mise au point, l'identification des paramètres physiques de systèmes, permet en **phase aval de la phase de développement**, le **calage de modèles** sur des données expérimentales.

La démarche adoptée dans nos travaux est appliquée à un système complexe par essence même. La robustesse et la précision des résultats obtenus démontrent la viabilité de la stratégie proposée. Les applications possibles de ces travaux sont nombreuses et peuvent contribuer à l'amélioration de système de contrôle adaptatif, de diagnostic et de supervision.

BIBLIOGRAPHIE

- [ACK96] J. Ackermann, T. Bünthe, W. Sienel, H. Jeebe, K. Naab. « Driving Safety by Robust Steering Control ». AVEC'96 Symp. 1996.
- [BAC93] T. Bäck, H.P. Schwefel. « An Overview of Evolutionary Algorithms for Parameters Optimization ». Evolutionary Computation Vol.11. 1993.
- [BAK85] J.E. Baker. « Adaptive Selection Methods for Genetic Algorithms ». ICGA'1. 1985.
- [BAK87] J.E. Baker. « Reducing Bias & Inefficiency in the Sélection Algorithm ». ICGA'1. 1987.
- [BAR70] Y. Bard. « Comparison of gradient methods for the solution of nonlinear parameters estimation problems ». SIAM J. Trans. Anal. 1970.
- [BAS99] M. Basset, J. Baujon, G. L. Gissinger. « Intelligent Driver Aid System Based on Trajectory Reconstruction for Active Safety ». Poster paper, 14th Ifac World Congress, Beijing, China. 1999.
- [BOR92] P. Borne & al. . « Modélisation et Identification des Processus ». Tome 2. Technips. 1992. ISBN 2-7108-0615-0.
- [BOT82] F. Böttiger & F. Panik. « Identification of Physical Parameters of Road Vehicles ». Proceedings of the 6th IFAC Symposium on Identification & System Parameter Estimation. 1982.
- [BRU90] J. Brunet & al. « Détection et Diagnostic de Pannes ». Hermes Paris. 1990.
- [CHA96] Y. Chamaillard. « Modélisation et Identification de Systèmes Complexes. Application à des Véhicules Routiers en vue de l'étude d'un Nouveau Système de Freinage ». Thèse, Université de Haute Alsace. Mulhouse. 1996.
- [COI97] P. Coirault. « Modélisation et Identification de Systèmes ». Habilitation à diriger les recherches, IUT de Poitiers. 1997.
- [DAV91] L. Davis. « Handbook of Genetic Algorithms ». Van Nostran Reinhold. 1991. ISBN 0-442-00173-8.
- [DEU70] C. Deutsch. « Dynamique des Véhicules Routiers, données de base ». Organisme national de la sécurité routière. Imprimerie de l'Anjou. Angers, France. 1970.

- [DIX96] J. C. Dixon. « Tires, Suspension & Handling ». SAE editions. 1996.
- [DOD70] T. Dobzhansky. « Genetics of the Evolutionary Process ». Columbia University Press. 1970.
- [DOE80] E.O. Doebelin. « System Modeling & Response. Theoretical & Experimental Approches ». John Wiley & Sons. 1980. ISBN 0-471-03211-5.
- [ELL69] J.R. Ellis. « Vehicle Dynamics ». Business Books Limited, London. 1969.
- [FAR95] J. L. Farges, G. Hardier. « Approches non Conventiionnelles en Modélisation et en Commande ». Rapport DERA n°062/95. 1995.
- [FIS12] R. A. Fisher. On an Absolute Criterion for Fitting Frequency Curves. Mess. Math. Vol. 41, N°155. 1912.
- [FLA94] J. M. Flaus. « La Régulation Industrielle, Régulateurs PID, Prédicatifs et Flous ». Hermès Paris. 1994. ISBN 2-86601-441-3.
- [FOG66] L. J. Fogel. « Artificial intelligence through simulated evolution ». Wiley. 1966.
- [GAU98] M. Gautier. « Identification of Robot Joint Stiffness ». International Systems, Signals, Control, Computer Conference (SSCC'98), september 98. 1998.
- [GIL92] T. D. Gillespie. « Fundamentals of Vehicle Dynamics ». SAE Publication n°R114. 1992. ISBN 1-56091-199-9.
- [GNA97] A. Gnadesikan. « Methods for Statistical Data Analysis of Multivariate Observations ». Wiley. 1997.
- [GOL89] D. E. Goldberg. « Algorithmes Génétiques, Exploration, Optimisation et Apprentissage Automatique ». Addison Wesley. 1994. ISBN 2-87908-0054-1.
- [GOO77] C. G. Goodwin, R. L. Payne. « Dynamic System Identification ». Academic Press London. 1977. ISBN 0-12-289750-1.
- [GRA92] A. Grace. « Optimization Toolbox, for use with Matlab ». The MathWorks Inc. 1992
- [HAN92] E. Hansen. « Global Optimization Techniques Using Interval Analysis ». Marcel Dekker, New York. 1992.
- [HEY90] G.J. Heydinger, W.R. Garrot, J.P. Christos, D.A. Guenther. « A Methodology for Validating Dynamics Simulations ». SAE. Technical Paper 90128. 1990.
- [HIM72] D. M. Himmelblau. « Applied non Linear Programming ». Mc Graw Hill. 1972.
- [HOL75] J. H. Holland. « Adaptation in Natural & Artificial Systems ». University of Michigan Press. Michigan USA. 1975.
- [HOR96] S. Horiuchi, N. Yahara, A. Takei. « Two Degrees of Freedom/ Hoo Controller Synthesis for Active Four Wheel Steering Vehicles ». Vehicle System Dynamics, Supplement 25. Swets & Zeitlinger. 1996.
- [HUA93] F. Huang. « The Use of Random Steer Test Data for Vehicle Parameter Estimation. SAE paper 930830. 1993.
- [HYS72] J. Hyslop. « A Note on the Accuracy of Optimisation Technique ». Comp. Journal. Vol 15. 1972.
- [ILI84] K. W. Iloff, R. E. Maine. « More than You Want to Know about the Maximum Likelihood Estimation ». AIAA Paper 94-2070. 1984.

- [ISE93] R. Isermann. « Fault Diagnosis of Machines via Parameters Estimation & Knowledge Processing - Tutorial Paper ». Automatica Vol. 29, n°4. 1993.
- [ISE97] R. Isermann. « Supervision, Fault Detection & Fault Diagnosis Methods – An Introduction ». Control Eng. Practice, Vol. 5, n°5. 1997.
- [JAN91] C. Z. Janikov & Z. Michalewicz. « An Experimental Comparison of Binary & Floating Point Representation in Genetic Algorithms ». Proc. of the ICGA'4. San Diego. 1991.
- [JAN97] J.S.R Jang, C.T. Sun, E. Mizutani. « Neuro Fuzzy & Soft computing ». Prentice All. 1997. ISBN 0-13-261066-3.
- [KNU96] M. Knudsen. « Loudspeaker Modelling & Parameters Estimation ». 100th Audio Engineering Society Convention, Copenhagen. 1996.
- [KOL95] I. Kollar. « Frequency domain system identification toolbox user's guide ». The MathWorks Inc. 1995.
- [LAU98] P. Laurence. « Modélisation de Systèmes Complexes. Application à l'Etude du Comportement en Régime Etabli des Véhicules Routiers ». Thèse de l'université de Haute Alsace. 1998.
- [LEM89] C. Lemaréchal. « Méthodes numériques d'optimisation ». Notes de cours Inria, DEA d'automatique. 1989.
- [LER96] I.C. Lerman & R.F. Ngouenet. « Algorithmes Génétiques Séquentiels et Parallèles pour une Représentation Affine des Proximités ». Publication Interne n°901 IRISA. 1996.
- [LET95] L. Le Letty, B. Nguyen. « Parameter estimation in Analytical Models of Automotive Vehicle & Fault Diagnosis ». 1st IFAC Workshop on Advances in Automotive Control. Ascona Switzerland. 1995.
- [LJU87] L. Ljung. « System Identification. Theory for the user ». Prentice Hall. 1987. ISBN 0-13-881640-9.
- [LJU91] L. Ljung. « Identification Toolbox User's guide for use with Matlab ». The MathWorks Inc . 1991.
- [MAI81] R. E. Maine & K. W. Iliff. « The Theory & Practice of Estimating the Accuracy of Dynamic Flight Determined Coefficients ». Nasa reference publication 1077. ISSN 0148-8585.
- [MIC91] Z. Michalewicz & C. Z. Janikow . « Handling Constraints in Genetic Algorithms ». 4th ICGA. 1991.
- [MIL88] G. Milne. « State Space Identification Tool . User's Guide». The MathWorks Inc. 1988.
- [MIL95] W.F. Milliken & D.L. Milliken. « Race Car Vehicle Dynamics ». SAE Publication. 1995. ISBN 1-56091-526-9.
- [MIN90] T. Minuro, M. Ohsaki, H. Zusunaga, & K. Satoh. « Four Parameters Evaluation Method of Lateral Transient Response ». SAE Paper 901734. 1990.
- [MUL91] H. Mülhenheim & all. « The Parallel Genetic Algorithm as Function Optimiser ». ICGA'4. 1991.
- [MUL93] H. Mülhenheim & D. Schlierkamp-Voosen. « Predictive Models for the Breeder Genetic Algorithm : I. Continuous Parameter Optimisation ». Evolutionary Computation vol1, n°1. 1993.

- [NAL88] A. G. Nalecz. « Sensivity Analysis of Vehicle Design Attribute in Frequency Domain ». *Vehicle Systems Dynamics* n°17. 1988. ISSN 0042-3114.
- [NEL65] J. A. Nelder, R. Mead. « A Simplex Method for Function Minimization ». *Comp. J.* 7. 1965.
- [NIN95] B. Ninness & G. Goodwin. « Estimation of model quality » *Automatica* Vol. 31, n°12. 1995.
- [OUV97] R. Ouvrard. « Contribution à l'identification robuste des systèmes à temps continu – application au génie des procédés. Thèse de doctorat de l'Université de Poitiers. 1997.
- [PAC91] H. B. Pacejka & R. S. Sharp. « Shear Force Development by Pneumatic Tyres in Steady Steate Conditions : a Review of Modelling Aspects ». *Vehicle systems Dynamics* n°20. 1991.
- [PHA85] Pham Anh Tuan. « Modélisation Mathématique du Comportement Dynamique d'une Automobile dans le Domaine Non-Lineaire ». SIA Session 1985-1986.
- [PIN94] R. Pintelon & al. « Parametric identification of transfer function in the frequency domain. A survey ». *IEEE trans. On A. C.* 1994.
- [POW91] D. Powel & al. « Interdigitation : an Hybrid Technique for Engineering Design Optimization Employing Genetic Algorithms, Expert Systems & Numerical Optimization » in *Handbook of Genetic Algorithms*. Van Nostrand Rheinhold. 1997.
- [PRE86] W. H. Press & al . « Numerical Recipes in C : The Art of Scientific Computing ». Cambridge University Press. 1986.
- [REC73] I. Rechenberg. « Evolutionstrategie : Optimierung technischer Systeme nach Prinzipien der Biologischen Evolution ». Fromman-Holzboog Verlag. 1973.
- [REN93] « Cours de Dynamique Véhicule ». Document Interne Renault S.A. 1993.
- [REN95] J. M. Renders. « Algorithmes Génétiques et Réseaux de Neurones, Application à la Commande de Processus ». Hermès Paris. 1995. ISBN 2-86601-467-7.
- [REN96] J. M. Renders, S. P. Flasse. « Hybrid Methods Using Genetic Algorithms for Global Optimization ». *IEEE Transactions on Systems, Man & Cybernetics*. Vol 26, n°2. 1996.
- [RIC40] P. Rickert, T. Schunk. « Zur Fahrmechanik des Gummibereiften Kraftfahrzeugs ». *Ingenieur Archiv* Nr 11. 1940.
- [RIC91] J. Richalet. « Pratique de l'identification ». Hermès Paris. 1991. ISBN 2-86601-287-9.
- [SCH91] J. Schoukens & R. Pintelon. « Identification of linear systems ». Pergamon Press. 1991.
- [SCH99] C.J Schmitt & al. « Identification of Physical Parameters of a Passenger Car », Communication à ECC'99, European Control Conference, 31 août - 3 septembre 1999, Karlsruhe, Allemagne.
- [SIB90] P. Sibille. « Analyse de la Mise en Œuvre de Méthodes d'Estimation Paramétriques ». Thèse de doctorat de l'université de Nancy I. 1990.
- [SIN82] N. K. Sinha & al. « Identification of continuous Multivariable Systems from Sampled Data ». *Int. J. of Control*. Vol35, n°1. 1982.
- [SOD89] T. Söderström, P. Stoica. « System Identification ». Prentice Hall U.K. 1989. ISBN 0-13-88-1236-5.

- [SON92] J. G. Song, Y. S. Yoon. « Symbolic State Equation of Three Degrees of Freedom of Vehicle Model for Steering and Application to Vehicle Design ». AVEC'92 symposium, Yokohama, Japan. SAE Japan Ed. 1992.
- [STO93] P. Stoica & al. « Approximate maximum likelihood frequency estimation ». Automatica, Vol. 30 n° 1. 1993.
- [TIS95] M. B. Tischler. « System Identification Methods for Aircraft Flight Control Development and Validation ». USAATCOM TR95-A-007 . 1995
- [TNO97] TNO Road Vehicle Research Institute, Delft, the Netherlands. TNO report N° 97-OR-VD-052-1/SJ. Pages 28-46. 1997.
- [TRI87] J. C. Trigeassou. « Contribution à l'extension de la méthode des moments en automatique. Application à l'identification des systèmes linéaires ». Thèse de doctorat de l'université de Poitiers. 1987.
- [UNB87] H. Unbehauen & G. P. Rao. « Identification of continuous systems ». North Holland. 1987.
- [WAD86] R. W. Allen, T. J. Rosenthal, H. T. Szostack. « Test Method & Computer Modeling for the Analysis of Ground Vehicle Handling ». SAE Paper 861115. 1986.
- [WAL90] E. Walter, L. Pronzato. « Qualitative & Quantitative Experiment Design for Phenomenological Models : a Survey ». Automatica Vol. 26 N°2. 1990.
- [WAL97] E. Walter, L. Pronzato. « Identification of Parametric Model ». Springer Verlag. 1997. ISBN 3-540-76119-5.
- [XIA95] X. Xia, J.N. Willis. « The Effect of Tire Cornering Stiffness on Vehicle Linear handling Performance ». SAE Paper 950313. 1995.
- [ZHE95] Zhengyou Zhang. « Parameter Estimation Techniques : a Tutorial with Application to Conic Fitting ». Research Report INRIA n°2676. 1995.
- [ZHI91] A. A Zhiglavsky. « Theory of Global Random Search ». Kluwer Dordrecht. 1991.

ANNEXES

1 PARAMETRISATION DES ALGORITHMES D'OPTIMISATION

1.1 Fonction de De Jong

☞ **Critères d'arrêt (AGR, AGX, quasi-Newton et simplex)**

- (ca2) : $|\hat{x}(k - \Delta k) - \hat{x}(k)| \leq \varepsilon_x = 10^{-8}$
- (ca3) : $d_Y(f(\hat{x}(k - \Delta k)), f(\hat{x}(k))) \leq \varepsilon_{d_Y} = 10^{-8}$

☞ **Quasi-Newton et simplex**

- (ca5) : nombre d'itération algorithme grimpeur = $100 * m$
Avec m : dimension de l'espace de recherche.

☞ **Algorithme génétique codé réel (AGR)**

- Pression de sélection : $sp = 1.4$
- Nombre de génération maximum : $maxGen = 200$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = 0.9$
- Type de recombinaison : $rec_f = \text{intermédiaire}$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

☞ **AG codé réel à recombinaison simplex (AGX)**

- Pression de sélection : $sp = 1.8$
- Nombre de génération maximum : $maxGen = 200$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = [1/3 \ 1/3 \ 1/3]$
- Type de recombinaison : $rec_f = [\text{discrète} \ \text{continue} \ \text{simplex}]$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

1.2 Fonction de De Griewank

☐ Critères d'arrêt (AGR, AGX, AGD et AGL)

- (ca2) : $|\hat{x}(k - \Delta k) - \hat{x}(k)| \leq \varepsilon_x = 10^{-8}$
- (ca3) : $d_Y(f(\hat{x}(k - \Delta k)), f(\hat{x}(k))) \leq \varepsilon_{d_Y} = 10^{-8}$

☐ Quasi-Newton et simplex

- **Dernière vie grimpeurs**
(ca5) : nombre d'itération algorithme grimpeur = $100 * m$
Avec m : dimension de l'espace de recherche.
- **Vie grimpeurs intermédiaires**
(ca5) : nombre d'itération algorithme grimpeur = $20 * m$
Avec m : dimension de l'espace de recherche.

☐ Algorithme génétique codé réel (AGR)

- Pression de sélection : $sp = 1.4$
- Nombre de génération maximum : $maxGen = 500$
- Nombre d'individus de la population : $nInd = 80$
- Probabilité de recombinaison : $pRec = 0.9$
- Type de recombinaison : $rec_f = \text{intermédiaire}$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

☐ AG codé réel à recombinaison simplex (AGX)

- Pression de sélection : $sp = 1.4$
- Nombre de génération maximum : $maxGen = 250$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = [1/3 \ 1/3 \ 1/3]$
- Type de recombinaison : $rec_f = [\text{discrète} \ \text{continue} \ \text{simplex}]$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

☐ Algorithme génétique hybride selon Darwin (quasi-Newton et simplex)

- Pression de sélection : $sp = 1.4$
- Nombre de génération maximum : $maxGen = 30$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = 0.9$
- Type de recombinaison : $rec_f = \text{intermédiaire}$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

☐ Algorithme génétique hybride selon Lamarck (quasi-Newton et simplex)

- Pression de sélection : $sp = 1.4$
- Nombre de génération maximum : $maxGen = 30$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = 0.9$
- Type de recombinaison : $rec_f = \text{intermédiaire}$
- Probabilité de mutation : $pMut = 0.5$ par individu
- Stratégie de réinsertion : élitiste

q **Algorithme génétique « 2 hybride » selon Lamarck (quasi-Newton et simplex)**

- Pression de sélection : $sp = 1.7$
- Nombre de génération maximum : $maxGen = 15$
- Nombre d'individus de la population : $nInd = 45$
- Probabilité de recombinaison : $pRec = [1/3 \ 1/3 \ 1/3]$
- Type de recombinaison : $rec_f = [discrète \ continue \ simplex]$
- Probabilité de mutation : $pMut = 0$ (absence de mutation)
- Stratégie de réinsertion : élitiste

2 INVALIDATION DES RESIDUS

2.1 Moyenne nulle

La moyenne nulle des résidus est vérifiée par le test statistique du t-test [PRE86]. Les résidus présentent une moyenne nulle avec un degré de confiance de 95% (2 écarts types pour une distribution normale) si [WAL97] :

$$m_1 : | \hat{m}_{\tilde{y}} | > 2 \sqrt{\frac{\hat{v}_{\tilde{y}}}{N}} \quad (A.2-1)$$

Avec

N : nombre de points du jeu de données considéré.

$\hat{v}_{\tilde{y}}$ une estimation de la variance des résidus :

$$\hat{v}_{\tilde{y}} = \frac{1}{N-1} \sum_{t=1}^N (\tilde{y}(t, \hat{\theta}) - \hat{m}_{\tilde{y}})^2 \quad (A.2-2)$$

$\hat{m}_{\tilde{y}}$ une estimation de leur moyenne

$$\hat{m}_{\tilde{y}} = \frac{1}{N} \sum_{t=1}^N \tilde{y}(t, \hat{\theta}) \quad (A.2-3)$$

2.2 Normalité des résidus

La normalité des résidus se vérifie classiquement de manière graphique et/statistique

2.2.1 Test graphique

Ce test consiste à tracer l'histogramme des résidus des différentes variables intervenant dans la formulation du critère de coût. Il est possible de comparer l'histogramme expérimental des résidus avec la fonction de distribution normale calculée à partir de la moyenne et de la variance précédente (voir figure A -1).

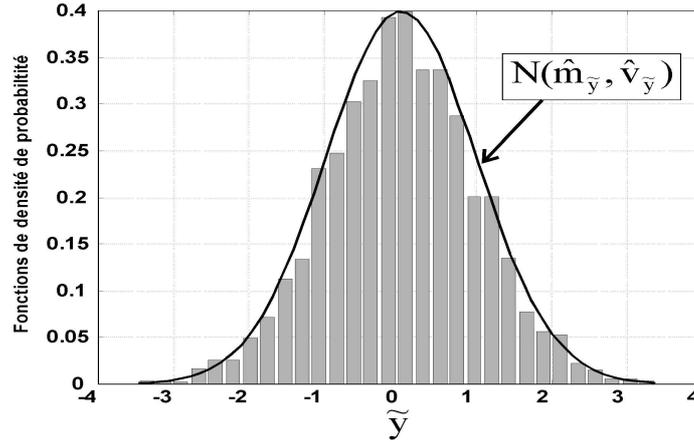


figure A -1 : histogramme des résidus

Ce test essentiellement qualitatif ne permet pas d'évaluer la normalité des résidus avec un degré de confiance défini. Cette étude graphique qualitative sera donc avantageusement complétée par les tests statistiques suivants.

2.2.2 Test statistique

Ces tests statistiques décrits par Walter [WAL97] dans son excellent ouvrage opèrent par comparaison des moments d'ordre 3 (coefficient d'asymétrie ou skewness : $\hat{k}_{\tilde{y}}$), et d'ordre 4 (coefficient d'aplatissement ou kurtosis : $\hat{s}_{\tilde{y}}$), des résidus aux coefficients d'une distribution normale dont les valeurs sont tabulées.

$$\hat{s}_{\tilde{y}} = \frac{1}{N} \sum_{t=1}^N \frac{(\tilde{y}(t, \hat{\theta}) - \hat{m}_{\tilde{y}})^3}{\hat{v}_{\tilde{y}}^{3/2}} \quad (\text{A.2-4})$$

$$\hat{k}_{\tilde{y}} = \left\{ \frac{1}{N} \sum_{t=1}^N \frac{(\tilde{y}(t, \hat{\theta}) - \hat{m}_{\tilde{y}})^4}{\hat{v}_{\tilde{y}}^2} \right\} - 3 \quad (\text{A.2-5})$$

Sous l'hypothèse de normalité des résidus \tilde{y} , les skewness et kurtosis précédent suivent également une distribution normale, ont respectivement pour moyenne, $s_{\text{nor}} = 0$ et

$$k_{\text{nor}} = \frac{-6}{N+1} \quad \text{et, respectivement pour variance} \quad v_{s_{\text{nor}}} = \frac{6(N-2)}{(N+1)(N+3)} \quad \text{et}$$

$$v_{k_{\text{nor}}} = \frac{24 N (N-2)(N-3)}{(N+1)^2 (N+3)(N+5)}. \quad \text{Donc, les skewness et kurtosis des résidus sont ceux}$$

d'une distribution normale avec un degré de confiance de 95% si les inégalités suivantes sont vérifiées :

$$n_1 : \frac{|\hat{s}_{\tilde{y}}|}{2 \sqrt{v_{s_{\text{nor}}}}} < 1 \quad (\text{A.2-6})$$

$$n_2 : \frac{|\hat{k}_{\tilde{y}}|}{2 \sqrt{v_{k_{\text{nor}}}}} < 1 \quad (\text{A.2-7})$$

2.3 Stationnarité

Comme pour le test de normalité, les deux approches graphiques et statistiques sont possibles. L'analyse de la stationnarité des résidus à l'ordre 2 se révèle dans la très grande majorité des cas, suffisante. Les tests graphiques et statistiques s'opèrent généralement sur les erreurs de prédiction normalisées et les carrés de ces erreurs de prédictions.

$$\tilde{y}_{\text{nor}}(t, \hat{\theta}) = \frac{\tilde{y}(t, \hat{\theta}) - \hat{m}_{\tilde{y}}}{\hat{v}_{\tilde{y}}^{1/2}} \quad (\text{A.2-8})$$

$$c_{\tilde{y}_{\text{nor}}}(t, \hat{\theta}) = \tilde{y}_{\text{nor}}^2(t, \hat{\theta}) \quad (\text{A.2-9})$$

Les deux tests décrits ci dessous sont également décrit par Walter dans son ouvrage [WAL97].

2.3.1 Test graphique

Dans le principe ce test est extrêmement simple. Il consiste à calculer une régression polynomiale à l'ordre 2 permettant de décrire l'évolution temporelle de la variable $c_{\tilde{y}_{\text{nor}}}(t, \hat{\theta})$:

$$\hat{c}_{\tilde{y}_{\text{nor}}}(t, \hat{\theta}) = a_0 + a_1 t + a_2 t^2 \quad (\text{A.2-10})$$

Sous l'hypothèse de stationnarité les coefficients a_1 et a_2 de la régression polynomiale doivent être nuls. Le test graphique consiste donc ensuite à représenter les variables $c_{\tilde{y}_{\text{nor}}}$ et $\hat{c}_{\tilde{y}_{\text{nor}}}$ en fonction du temps. Sous l'hypothèse de stationnarité, l'évolution $\hat{c}_{\tilde{y}_{\text{nor}}} = f(t)$ s'apparente à une droite horizontale (cf. figure A -2).

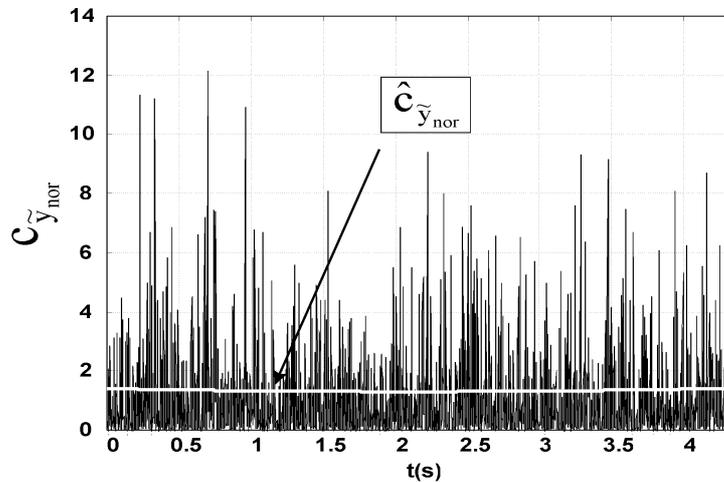


figure A -2 : stationnarité des résidus

2.3.2 Test statistique

Une manière de procéder est de calculer le coefficient de corrélation entre la variable $c_{\tilde{y}_{\text{nor}}}$ et le temps, soit :

$$C_{\tilde{y}_{\text{nor}} t} = \frac{\sum_{i=1}^N ((c_{\tilde{y}_{\text{nor}}}(i, \hat{\theta}) - m_c)(t(i) - m_t))}{\left(\sum_{i=1}^N (c_{\tilde{y}_{\text{nor}}}(i, \hat{\theta}) - m_c)^2 \sum_{i=1}^N (t(i) - m_t)^2 \right)^{1/2}} \quad (\text{A.2-11})$$

Où m_c et m_t représente respectivement les moyennes estimées de $c_{\tilde{y}_{nor}}$ et de t . Si les résidus sont stationnaires, ce coefficient de corrélation est proche de zéro. Walter [WAL97] propose l'utilisation du test d'Anscombe : le terme $(N - m)^{1/2} C_{\tilde{y}_{nor} t}$ tends à suivre approximativement une distribution normale de variance 1. Les résidus sont donc stationnaire avec une de 95% si l'inégalité suivante est vérifiée

$$s_1 : \frac{(N - m)^{1/2} | C_{\tilde{y}_{nor} t} |}{2} < 1 \quad (\text{A.2-12})$$

2.4 Indépendance

Deux tests d'indépendance des résidus sont possibles :

- Le test d'indépendance des résidus vis à vis d'eux-mêmes. L'assimilation de la suite des erreurs de prédiction à une variable aléatoire indépendante est vérifiée.
- Le test de l'indépendance des résidus vis à vis de l'entrée [LJU87], [SOD89]. Ce test vérifie que toutes les dynamiques du système sont modélisées par le modèle identifié ; il ne subsiste donc plus de relation de cause à effet entre l'entrée et les résidus.

La mise en œuvre des tests d'indépendance repose essentiellement sur deux notions. La première possibilité est d'exploiter les propriétés de la fonction d'auto corrélation (test graphique). La seconde (test statistique), repose sur le fait, qu'en cas d'indépendance, la variable étudiée a une probabilité de 50% d'être supérieure (inférieure) à la valeur précédente.

2.4.1 Test graphique

Le test graphique exploite les notions de fonction d'auto corrélation [LJU87]. La fonction d'auto corrélation normalisée des erreurs de prédictions est donnée par

$$R_{\tilde{y}\tilde{y}}(\tau) = \frac{\frac{1}{N} \sum_{t=1}^N ((\tilde{y}(t, \hat{\theta}) - \hat{m}_{\tilde{y}})(\tilde{y}(t - \tau, \hat{\theta}) - \hat{m}_{\tilde{y}}))}{\frac{1}{N} \sum_{t=1}^N \tilde{y}^2(t, \hat{\theta})} \quad (\text{A.2-13})$$

Sous l'hypothèse d'indépendance, asymptotiquement la distribution de $R_{\tilde{y}\tilde{y}}(\tau)$ suit une loi normale $N(0, 1/N)$ pour tout $\tau \neq 0$. Pour un nombre de suffisamment grand de données, il est possible d'affirmer que les résidus sont indépendants avec un degré de confiance de 95% si la quantité $R_{\tilde{y}\tilde{y}}(\tau)$ se situe dans l'intervalle $[-2\sqrt{N} \dots + 2\sqrt{N}]$. Cette méthode s'applique également au calcul de l'inter corrélation des résidus avec l'entrée.

2.4.2 Test statistique

Parmi les différents tests possibles, un des plus simple de mise en œuvre est certainement celui décrit par Walter [WAL97]. Ce test consiste à calculer le rang $r(t)$ de chaque erreur de prédiction $\tilde{y}(t, \hat{\theta})$ (classée par ordre croissant ou décroissant) puis à calculer le rapport de Von Neumann dont les statistiques sont tabulées :

$$r_{vn} = \frac{12 \sum_{t=2}^N (r(t) - r(t-1))^2}{N(N^2 - 1)} \quad (\text{A.2-14})$$

Asymptotiquement, sous l'hypothèse d'indépendance, la distribution de r_{vn} est approximativement $N(2, 20 / (5N + 7))$. Les résidus peuvent donc être considérés comme indépendants avec un degré de confiance de 95% si l'inégalité suivante est vérifiée

$$i_1 : \frac{|r_{vn} - 2|}{2(20 / (5N + 7))^{1/2}} < 1 \quad (\text{A.2-15})$$

3 EXPRESSION DES FORCES ET MOMENTS APPLIQUES AU CENTRE DE GRAVITE DU VEHICULE

L'expression des forces exercées en G dans le trièdre R_0 se calcule à partir du principe fondamental de la dynamique.

3.1 Principe fondamental de la dynamique

L'accélération absolue en G de la partie suspendue dans le repère R_0 s'exprime :

$$\vec{\gamma}_{Abs} = \vec{\gamma}_E + \vec{\gamma}_R + \vec{\gamma}_C \quad (\text{A.3-1})$$

Avec

§	$\vec{\gamma}_E$	l'accélération du point O de R par rapport à R_0
§	$\vec{\gamma}_R$	l'accélération du point G de R_1 par rapport à R
§	$\vec{\gamma}_C$	l'accélération de Coriolis

3.1.1 Calcul de l'accélération d'entraînement :

La vitesse dans le trièdre (O, x, y, z) d'entraînement du véhicule est :

$$\vec{V}_{O(R/R_0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta \\ 0 \end{pmatrix} \quad (\text{A.3-2})$$

Soit

$$\vec{\gamma}_{O(R/R_0)} = \frac{d\vec{V}_{O(R/R_0)}}{dt} = \begin{pmatrix} -V \sin \beta \dot{\beta} \\ V \cos \beta \dot{\beta} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} \wedge \begin{pmatrix} V \cos \beta \\ V \sin \beta \\ 0 \end{pmatrix} \quad (\text{A.3-3})$$

d'où

$$\vec{\gamma}_E = \vec{\gamma}_{O(R/R_0)} = \begin{pmatrix} -V (\dot{\beta} + \dot{\psi}) \sin \beta \\ V (\dot{\beta} + \dot{\psi}) \cos \beta \\ 0 \end{pmatrix} \quad (\text{A.3.4})$$

3.1.2 Calcul de l'accélération relative

L'accélération relative au centre de gravité s'exprime :

$$\vec{\gamma}_R = \vec{\gamma}_{G(R_1/R)} = \frac{d\vec{V}_{G(R_1/R)}}{dt} \quad (\text{A.3.5})$$

Or

$$\vec{V}_{G(R1/R)} = \vec{GO} \wedge \vec{\Omega}_{(R1/R)} \quad (\text{A.3.6})$$

$$\vec{V}_{G(R1/R)} = \begin{pmatrix} 0 \\ 0 \\ -h_0 \end{pmatrix} \wedge \begin{pmatrix} \dot{\theta} \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -h_0 \dot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.3.7})$$

Et

$$\frac{d\vec{V}_{G(R1/R)}}{dt} = \begin{pmatrix} 0 \\ -h_0 \ddot{\theta} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} \wedge \begin{pmatrix} 0 \\ -h_0 \dot{\theta} \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -h_0 \ddot{\theta} \\ 0 \end{pmatrix} + \begin{pmatrix} h_0 \dot{\theta} \dot{\psi} \\ 0 \\ 0 \end{pmatrix} \quad (\text{A.3.8})$$

D'où

$$\vec{\gamma}_R = \begin{pmatrix} h_0 \dot{\theta} \dot{\psi} \\ -h_0 \ddot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.3.9})$$

3.1.3 Calcul de l'accélération de Coriolis

L'accélération de Coriolis au centre de gravité s'exprime :

$$\vec{\gamma}_C = 2 \vec{\omega}_E \wedge \vec{V}_{G(R1/R)} \quad (\text{A.3-10})$$

$\vec{\omega}_E$ est le vecteur rotation de R par rapport à R0

Dans le repère R :

$$\vec{\omega}_E = \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} \quad (\text{A.3-11})$$

Et

$$\vec{V}_{G(R1/R)} = \begin{pmatrix} 0 \\ -h_0 \dot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.3-12})$$

D'où

$$\vec{\gamma}_C = \begin{pmatrix} -2 h_0 \dot{\theta} \dot{\psi} \\ 0 \\ 0 \end{pmatrix} \quad (\text{A.3-13})$$

3.1.4 Conclusion

L'accélération absolue du véhicule s'exprime donc

$$\vec{\gamma}_{\text{abs}} = \begin{pmatrix} V(\dot{\beta} + \dot{\psi}) \sin \beta - h_0 \dot{\theta} \dot{\psi} \\ V(\dot{\beta} + \dot{\psi}) \cos \beta - h_0 \ddot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.3-14})$$

Finalement, les forces d'inertie absolues s'appliquant à G, mais projetées dans le repère intermédiaire R (O, x, y, z) sont :

$$\vec{F} = M \vec{\gamma}_E + ms \vec{\gamma}_R + ms \vec{\gamma}_C \quad (\text{A.3-15})$$

Donc

$$\begin{aligned} F_x &= -M V (\dot{\beta} + \dot{\psi}) \sin \beta - ms h_0 \dot{\theta} \dot{\psi} \\ F_y &= M V (\dot{\beta} + \dot{\psi}) \cos \beta - ms h_0 \ddot{\theta} \\ F_z &= 0 \end{aligned} \quad (\text{A.3-16})$$

3.2 Théorème du moment dynamique

Soit \vec{H}_O , le moment cinétique en O :

$$\vec{H}_{O/R1} = \vec{H}_{G/R1} + \vec{OG} \wedge ms \vec{V}_{G(R1/R0)} \quad (\text{A.3-17})$$

\vec{H}_G est le moment cinétique en G de la partie suspendue dont nous considérons les composantes dans le repère caisse (G, x_1 , y_1 , z_1). Les moments d'inertie sont constants, par rapport à ce même repère (G, x_1 , y_1 , z_1) :

$$\vec{H}_{G/R1} = \begin{pmatrix} I_{xx} & 0 & -I_{xz} \\ 0 & I_{yy} & 0 \\ -I_{xz} & 0 & I_{xx} \end{pmatrix} \begin{pmatrix} \dot{\theta} \\ 0 \\ \dot{\psi} \end{pmatrix} = \begin{pmatrix} I_{xx} \dot{\theta} - I_{xz} \dot{\psi} \\ 0 \\ -I_{xz} \dot{\theta} + I_{zz} \dot{\psi} \end{pmatrix} \quad (\text{A.3-18})$$

Les inerties de couplage I_{xy} et I_{yz} sont nulles du fait de la symétrie géométrique de l'automobile par rapport au plan (x_1 , z_1).

$$\vec{OG} = \begin{pmatrix} 0 \\ 0 \\ h_0 \end{pmatrix} \text{ et } \vec{V}_{G(R1/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta - h_0 \dot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.3-19})$$

$$\vec{OG} \wedge ms \vec{V}_{G(R1/R0)} = \begin{pmatrix} -ms h_0 V \sin \beta + ms h_0^2 \dot{\theta} \\ ms h_0 V \cos \beta \\ 0 \end{pmatrix} \quad (\text{A.3-20})$$

Donc les composantes de \vec{H}_O dans le repère ($Gx_1y_1z_1$) sont :

$$\vec{H}_{O/R1} = \begin{pmatrix} I_{xx} \dot{\theta} - I_{xz} \dot{\psi} - ms h_0 V \sin \beta + ms h_0^2 \dot{\theta} \\ ms h_0 V \cos \beta \\ -I_{xz} \dot{\theta} + I_{zz} \dot{\psi} \end{pmatrix} \quad (\text{A.3-21})$$

Le moment dynamique s'obtient par dérivation du moment cinétique dans le repère R1 (G, x_1 , y_1 , z_1) :

$$\frac{d\vec{H}_{O/R0}}{dt} = \frac{d\vec{H}_{O/R1}}{dt} + \vec{\Omega}_{(R1/R0)} \wedge \vec{H}_{O/R1} \quad (\text{A.3-22})$$

Et

$$\vec{\Omega}_{(R1/R0)} \wedge \vec{H}_{O/R1} = \begin{pmatrix} \dot{\theta} \\ 0 \\ \dot{\psi} \end{pmatrix} \wedge \begin{pmatrix} I_{xx} \dot{\theta} - I_{xz} \dot{\psi} - ms h_0 V \sin \beta + ms h_0^2 \dot{\theta} \\ ms h_0 V \cos \beta \\ -I_{xz} \dot{\theta} + I_{zz} \dot{\psi} \end{pmatrix} \quad (\text{A.3-23})$$

$$\vec{\Omega}_{(R1/R0)} \wedge \vec{H}_{O/R1} = \begin{pmatrix} -ms h_0 V \cos \beta \dot{\psi} \\ I_{xx} \dot{\theta} \dot{\psi} - I_{xz} \dot{\psi}^2 - ms h_0 V \sin \beta \dot{\psi} + ms h_0^2 \dot{\theta} \dot{\psi} + I_{xz} \dot{\theta}^2 - I_{zz} \dot{\theta} \dot{\psi} \\ ms h_0 V \cos \beta \dot{\theta} \end{pmatrix} \quad (\text{A.3-24})$$

D'où l'expression du moment dynamique dans le trièdre R0

$$\frac{d\vec{H}_{O/R0}}{dt} = \begin{pmatrix} (I_{xx} + ms h_0^2) \ddot{\theta} - I_{xz} \ddot{\psi} - ms h_0 V \cos \beta (\dot{\beta} + \dot{\psi}) \\ -ms h_0 V \sin \beta (\dot{\beta} + \dot{\psi}) + (I_{xx} - I_{zz} + ms h_0^2) \dot{\theta} \dot{\psi} + I_{xz} (\dot{\theta}^2 - \dot{\psi}^2) \\ ms h_0 V \cos \beta \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} \end{pmatrix} \quad (A.3-25)$$

Les équations du mouvement résultantes s'écrivent donc :

$$\begin{aligned} L_x &= (I_{xx} + ms h_0^2) \ddot{\theta} - I_{xz} \ddot{\psi} - ms h_0 V \cos \beta (\dot{\beta} + \dot{\psi}) \\ M_y &= -ms h_0 V \sin \beta (\dot{\beta} + \dot{\psi}) + (I_{xx} - I_{zz} + ms h_0^2) \dot{\theta} \dot{\psi} + I_{xz} (\dot{\theta}^2 - \dot{\psi}^2) \\ N_z &= ms h_0 V \cos \beta \dot{\theta} - I_{xz} \ddot{\theta} + I_{zz} \ddot{\psi} \end{aligned} \quad (A.3-26)$$

4 VITESSE DU POINT DE CONTACT PNEU/SOL

La vitesse transversale du point de contact pneumatique/sol par rapport au repère caisse (R1) est ici recherchée. Cette quantité est indispensable au calcul de l'effort transversal issu d'une roue. En raisonnant à l'essieu avant et en se référant à la figure A -3, calculons la vitesse du point P1 situé sur l'axe de roulis (centre de roulis avant).

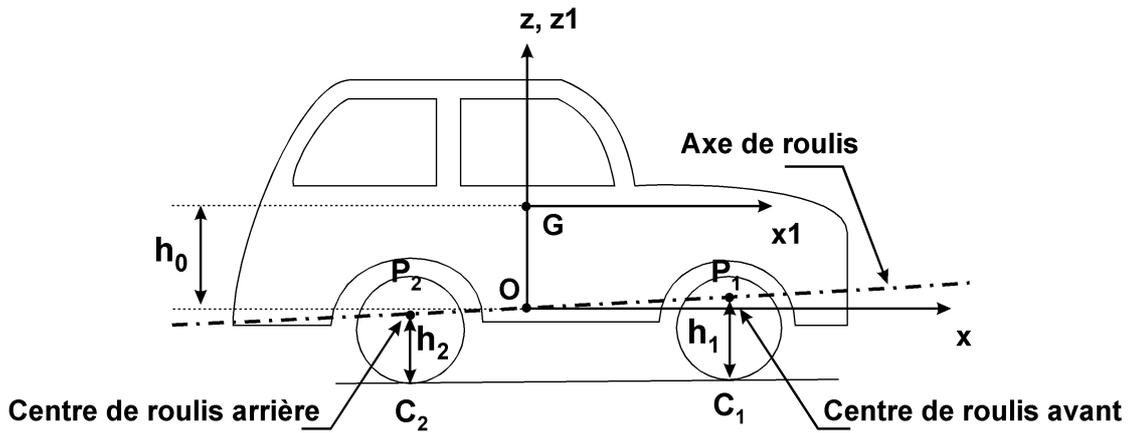


figure A -3 : point de contact pneu-sol

D'où

$$\vec{V}_{P_1(R/R0)} = \vec{V}_{O(R/R0)} + \vec{P_1O} \wedge \vec{\Omega}_{(R/R0)} \quad (A.4-1)$$

$$\vec{V}_{P_1(R/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta \\ 0 \end{pmatrix} + \begin{pmatrix} -h_1 \\ 0 \\ P_1z \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 0 \\ \dot{\psi} \end{pmatrix} \quad (A.4-2)$$

$$\vec{V}_{P_1(R/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta + h_1 \dot{\psi} \\ 0 \end{pmatrix} \quad (A.4-3)$$

En désignant par C₁ le point de contact pneumatique/sol à l'avant du véhicule il vient

$$\vec{V}_{C_1(R1/R0)} = \vec{V}_{P_1(R/R0)} + \vec{C_1P_1} \wedge \vec{\Omega}_{(R1/R0)} \quad (A.4-4)$$

$$\vec{V}_{C_1(R1/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta + h_1 \dot{\psi} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ h_1 \end{pmatrix} \wedge \begin{pmatrix} \dot{\theta} \\ 0 \\ \dot{\psi} \end{pmatrix} \quad (\text{A.4-5})$$

$$\vec{V}_{C_1(R1/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta + h_1 \dot{\psi} + h_1 \dot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.4-6})$$

De manière similaire, la vitesse du point de contact pneumatique/sol à l'arrière du véhicule s'exprime

$$\vec{V}_{C_2(R1/R0)} = \begin{pmatrix} V \cos \beta \\ V \sin \beta - h_2 \dot{\psi} + h_2 \dot{\theta} \\ 0 \end{pmatrix} \quad (\text{A.4-7})$$

5 FONCTIONS DE SENSIBILITE DE LA DERIVE ET DE LA VITESSE DE ROULIS

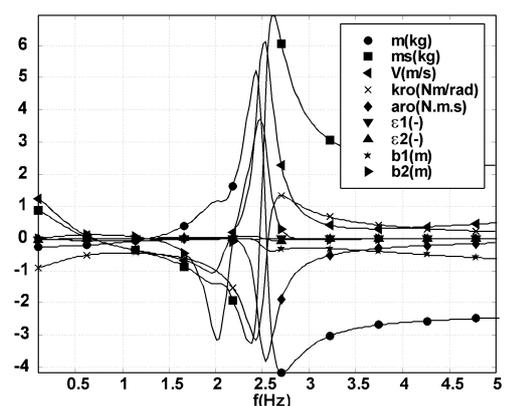
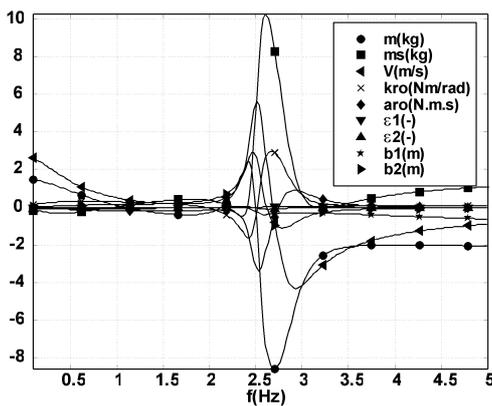
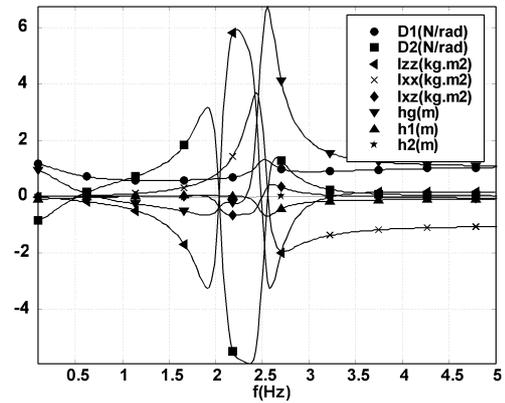
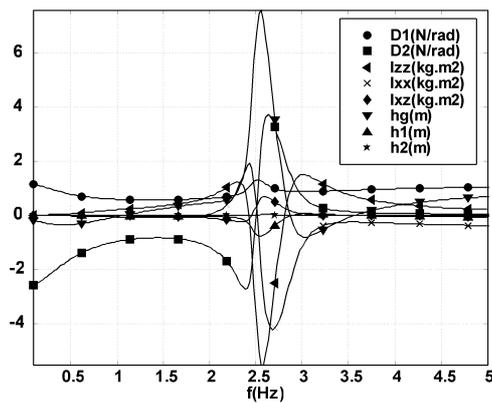


figure A - 4a : angle de dérive

figure A - 4b : vitesse de roulis

figure A - 4 : magnitude des fonctions de sensibilité logarithmique du premier ordre de l'angle de dérive et de la vitesse de roulis (pour une vitesse de 80 km/h)