



**HAL**  
open science

## Problèmes inverses en génie civil

Pierre Argoul, Nelly Point, Guillaume Dutilleux

► **To cite this version:**

Pierre Argoul, Nelly Point, Guillaume Dutilleux. Problèmes inverses en génie civil. LABORATOIRE CENTRAL DES PONTS ET CHAUSSEES - LCPC, 261p + cdrom, bibliogr., 2008, ETUDES ET RECHERCHES DES LABORATOIRES DES PONTS ET CHAUSSEES, 10.3829/erlpc.si15-fr . hal-00362979v2

**HAL Id: hal-00362979**

**<https://hal.science/hal-00362979v2>**

Submitted on 18 Oct 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Problèmes inverses en génie civil

Ouvrage collectif  
Contributions recueillies et présentées par  
Pierre Argoul, Nelly Point, Guillaume Dutilleux

Décembre 2008



Laboratoire central des ponts et chaussées  
58, boulevard Lefebvre, F 75732 Paris Cedex 15

Conformément à la note du 04/07/2014 de la direction générale de l'Ifsttar précisant la politique de diffusion des ouvrages parus dans les collections éditées par l'Institut, la reproduction de cet ouvrage est autorisée selon les termes de la licence CC BY-NC-ND. Cette licence autorise la redistribution non commerciale de copies identiques à l'original. Dans ce cadre, cet ouvrage peut être copié, distribué et communiqué par tous moyens et sous tous formats.



Attribution — Vous devez créditer l'Oeuvre et intégrer un lien vers la licence. Vous devez indiquer ces informations par tous les moyens possibles mais vous ne pouvez pas suggérer que l'Ifsttar vous soutient ou soutient la façon dont vous avez utilisé son Oeuvre.



Pas d'Utilisation Commerciale — Vous n'êtes pas autorisé à faire un usage commercial de cette Oeuvre, tout ou partie du matériel la composant.



Pas de modifications — Dans le cas où vous effectuez une adaptation, que vous transformez, ou créez à partir du matériel composant l'Oeuvre originale (par exemple, une traduction, etc.), vous n'êtes pas autorisé à distribuer ou mettre à disposition l'Oeuvre modifiée.

## Le patrimoine scientifique de l'Ifsttar

Le libre accès à l'information scientifique est aujourd'hui devenu essentiel pour favoriser la circulation du savoir et pour contribuer à l'innovation et au développement socio-économique. Pour que les résultats des recherches soient plus largement diffusés, lus et utilisés pour de nouveaux travaux, l'Ifsttar a entrepris la numérisation et la mise en ligne de son fonds documentaire. Ainsi, en complément des ouvrages disponibles à la vente, certaines références des collections de l'INRETS et du LCPC sont dès à présent mises à disposition en téléchargement gratuit selon les termes de la licence Creative Commons CC BY-NC-ND.

Le service Politique éditoriale scientifique et technique de l'Ifsttar diffuse différentes collections qui sont le reflet des recherches menées par l'institut :

- Les collections de l'INRETS, Actes
- Les collections de l'INRETS, Outils et Méthodes
- Les collections de l'INRETS, Recherches
- Les collections de l'INRETS, Synthèses
- Les collections du LCPC, Actes
- Les collections du LCPC, Etudes et recherches des laboratoires des ponts et chaussées
- Les collections du LCPC, Rapport de recherche des laboratoires des ponts et chaussées
- Les collections du LCPC, Techniques et méthodes des laboratoires des ponts et chaussées, Guide technique
- Les collections du LCPC, Techniques et méthodes des laboratoires des ponts et chaussées, Méthode d'essai



Institut Français des Sciences et Techniques des Réseaux,  
de l'Aménagement et des Transports  
14-20 Boulevard Newton, Cité Descartes, Champs sur Marne  
F-77447 Marne la Vallée Cedex 2

Contact : diffusion-publications@ifsttar.fr

[www.ifsttar.fr](http://www.ifsttar.fr)



Cet ouvrage collectif est le fruit des travaux de recherche, durant la période 2004-2007, de l'opération de recherche 11P041 du Laboratoire central des ponts et chaussées : « Identification, contrôle et problèmes inverses ».

Il est constitué de onze contributions recueillies et présentées par :  
**Pierre Argoul, Nelly Point et Guillaume Dutilleux.**

Pour commander cet ouvrage :  
Laboratoire central des ponts et chaussées  
DISTC - Section Diffusion  
58, boulevard Lefebvre, F 75732 Paris cedex 15  
Téléphone : 01 40 43 50 20 - Télécopie : 01 40 43 54 95  
Ou serveur Internet LCPC : [www.lcpc.fr](http://www.lcpc.fr)

Prix : 50 Euros HT

Ce document est propriété du Laboratoire central des ponts et chaussées  
et ne peut être reproduit, même partiellement, sans l'autorisation de son Directeur général  
(ou de ses représentants autorisés)

© 2008 - LCPC

ISSN 1167-4865

ISBN 978-2-7208-2533-6

DOI/Crossref 10.3829/erlpc.si15-fr

# Table des matières

<b>Résumé</b>	<b>6</b>
<b>Abstract</b>	<b>7</b>
<b>Présentation générale</b>	<b>9</b>
Présentation des problèmes inverses . . . . .	9
Présentation de l'ouvrage . . . . .	15
Conclusion . . . . .	17
<b>1 Méthodologie pour l'identification des systèmes multicorps articulés</b>	<b>23</b>
1.1 Position du problème . . . . .	25
1.2 Planification . . . . .	28
1.3 Analyse d'un mouvement . . . . .	32
1.4 Résolution . . . . .	36
1.5 Conclusion . . . . .	40
<b>2 Algorithmes récursifs pour l'identification modale des structures de génie civil</b>	<b>43</b>
2.1 Introduction . . . . .	43
2.2 Identification des systèmes . . . . .	44
2.3 Application à l'identification modale . . . . .	46
2.4 Modèles de travail . . . . .	48
2.5 Expression récursive des équations de Yule-Walker . . . . .	49
2.6 Application . . . . .	53
<b>3 Analyse en ondelettes pour l'identification de paramètres en dynamique</b>	<b>61</b>
3.1 Introduction . . . . .	61
3.2 Identification de paramètres en dynamique . . . . .	62
3.3 La transformation en ondelettes . . . . .	64
3.4 Analyse en ondelettes et identification modale . . . . .	67
3.5 Vibrations libres non linéaires et ondelettes . . . . .	81
<b>4 Modélisation et identification de systèmes discrets hystérétiques en dynamique</b>	<b>91</b>
4.1 Introduction . . . . .	91
4.2 Briques et modèles élémentaires . . . . .	93
4.3 Rappels mathématiques : existence et unicité . . . . .	100
4.4 Aspects numériques . . . . .	101
4.5 Identification . . . . .	102

4.6	Conclusion	105
<b>5</b>	<b>Modèles probabilistes de détection de rupture dans le comportement</b>	<b>113</b>
5.1	Introduction	113
5.2	Présentation de trois modèles de rupture	113
5.3	Plasticité	118
5.4	L'influence des élections françaises sur le taux d'échange	119
5.5	Conclusion	121
<b>6</b>	<b>Assimilation thermique</b>	<b>125</b>
6.1	Introduction	125
6.2	Formulation générale du problème	127
6.3	Minimisation dans $L^2$	130
6.4	Reconstruction dans $H^1$	135
6.5	Conclusions	138
<b>7</b>	<b>Théorie semi-quadratique, régularisation et estimation robuste</b>	<b>143</b>
7.1	Introduction	143
7.2	Le modèle quadratique et ses limites	144
7.3	Théorie semi-quadratique	147
7.4	Applications	152
7.5	Conclusion	160
<b>8</b>	<b>Des méthodes pour l'identification des systèmes élastiques hétérogènes</b>	<b>167</b>
8.1	Introduction	167
8.2	Description du problème	170
8.3	Méthode de Monte Carlo	175
8.4	Résultats numériques	178
8.5	Conclusions	182
<b>9</b>	<b>Algorithmes pour l'optimisation globale</b>	<b>187</b>
9.1	Introduction	187
9.2	Principes généraux	188
9.3	Une métaphore physique : le recuit simulé	190
9.4	Une métaphore sociale : l'essaim de particules	193
9.5	Une métaphore darwinienne : les algorithmes évolutionnaires	196
9.6	Une métaphore immunitaire : la sélection clonale	204
9.7	Pratique de l'optimisation globale	208
9.8	Conclusion	210
<b>10</b>	<b>Optimisation de forme de murs antibruit par algorithmes génétiques</b>	<b>217</b>
10.1	Introduction	217
10.2	Calcul du champ de pression	218
10.3	Résultats	221
10.4	Conclusions	224

<b>11 Les méthodes inverses en reconnaissance et géophysique au LCPC</b>	<b>229</b>
11.1 Introduction . . . . .	229
11.2 Inversion de la dispersion des ondes de surface . . . . .	232
11.3 Tomographies sismiques . . . . .	240
11.4 Tomographie électromagnétique 2D entre forages . . . . .	245
11.5 Tomographies de résistivité électrique de structures 3D de géométrie quelconque . . . . .	250
<b>Liste des auteurs</b>	<b>262</b>

## **Résumé**

### **Problèmes inverses en génie civil**

Ouvrage collectif  
Contributions recueillies et présentées par  
Pierre ARGOUL  
Guillaume DUTILLEUX  
Nelly POINT

Schématiquement, résoudre un problème inverse revient à essayer de restituer la réalité à partir d'une image partielle et/ou brouillée. Les problèmes inverses constituent un domaine foisonnant ; les applications sont multiples, les problématiques diverses et les méthodes utilisées forcément variées.

Cet ouvrage collectif réalise un état de l'art sur différentes techniques inverses utilisées au sein du réseau scientifique et technique du Ministère de l'écologie, de l'énergie, du développement durable et de l'aménagement du territoire, dans le cadre de la mécanique et du génie civil en particulier.

On y aborde des problèmes inverses divers : identification de paramètres mécaniques d'un compacteur, de paramètres modaux, d'un système discret hystérétique, identification des causes, détection et reconnaissance d'images, optimisation de mur anti-bruit, détermination d'hétérogénéités et aussi plusieurs problèmes de tomographie appliqués à la géophysique et à l'auscultation d'ouvrages d'art.

Les méthodes de résolution numérique proposées vont des méthodes maintenant classiques d'optimisation dans le cadre régulier ou convexe, aux méthodes dites méta-heuristiques de minimisation globale pour lesquelles est joint un cederom contenant une boîte à outil en langage Scilab.

**Mots-clés** : problèmes inverses, identification de paramètres, optimisation globale, identification modale, ondelettes, génie civil, tomographie.

## **Abstract**

### **Inverse problems in civil engineering**

Collective publication  
Contributions gathered and presented by  
Pierre ARGOUL  
Guillaume DUTILLEUX  
Nelly POINT

Basically, solving inverse problems involves recreating reality from a partial and/or blurred image. The range of inverse problems is broad, involving diverse issues, a host of applications, and obviously a wide variety of methods. This collective publication aims to draw up an inventory of the various inverses techniques used in the Ministry of the Environment, Energy, Sustainable Development and Territorial Planning, in particular in Mechanical and Civil Engineering. Many different inverse problems are considered : identification of mechanical parameters for a compressor, of modal parameters or of a discrete hysteretic system, identification of causes, detection and recognition of images, optimization of noise barrier walls, characterization of heterogeneities as well as various tomographic problems applied to geophysics and the close examination of art works. The numerical resolution methods proposed here range from the now classic optimization methods in regular or convex case to metaheuristic methods of global optimization, for which a CD with a toolbox in Scilab is included.

**Mots-clés** : inverse problems, parameters identification, global optimization, modal identification, wavelets, civil engineering, tomography.



# Présentation générale

*Pierre Argoul,  
Nelly Point*

Dans le cadre de l'opération de recherches du Laboratoire Central des Ponts et Chaussées « Identification, contrôle et problèmes inverses », nous avons souhaité réunir plusieurs articles de synthèse sur les problèmes inverses dans un ouvrage collectif. L'ouvrage présent, écrit par des chercheurs du LCPC et leurs collaborateurs réalise un état de l'art sur différentes techniques inverses utilisées au LCPC et sur les applications de ces méthodes au Génie Civil et Urbain.

Ce chapitre introductif est constitué d'une première partie qui présente les problèmes inverses, d'une seconde qui présente les différentes contributions de cet ouvrage avec la classification retenue, et enfin d'une conclusion.

## Présentation des problèmes inverses

### Définitions

La terminologie « **problème inverse** » est apparue dans les années soixante, notamment dans des problèmes de géophysique pour désigner la détermination, à partir d'expériences de type entrée-sortie ou cause-effet, des inconnues intervenant dans les équations de la physique. La notion de problème inverse désigne aujourd'hui la reconstruction au mieux de l'information manquante pour pouvoir estimer soit les sollicitations (**identification de sources ou de la cause**), soit la valeur de paramètres indéterminés (**identification de modèles**).

Pour ne citer qu'un seul exemple de problème inverse, nous avons retenu celui de la découverte de la planète Neptune par Le Verrier. C'est un cas très célèbre d'identification de la cause. Au début du XIX<sup>e</sup> siècle, les astronomes ne connaissaient pas encore l'existence de la planète Neptune ; seule la planète Uranus avait été découverte par Herschel en 1781. Lorsqu'ils appliquaient la théorie de la gravitation universelle de Newton au cas du mouvement d'Uranus, l'orbite et le mouvement calculés en tenant uniquement compte des perturbations de Jupiter et de Saturne, présentaient des irrégularités par rapport à ceux observés. Encouragé par Arago, Le Verrier se lançait dès 1844 dans le calcul par perturbations inverses des caractéristiques (masse, orbite, position actuelle) d'une planète hypothétique supposée produire sur Uranus l'effet observé. Dans son mémoire à l'Académie des Sciences du 31 août 1846, Le Verrier définissait les éléments de l'orbite de cette nouvelle planète. Le 23 septembre 1846, jour même où il recevait de Le Verrier une lettre précisant la position de l'objet

prédit, l'astronome berlinois Galle l'observait. Arago déclarait ensuite à l'Académie des sciences : « Mr Le Verrier a aperçu le nouvel astre sans avoir besoin de jeter un seul regard vers le ciel ; il l'a vu au bout de sa plume ».

Avant de poursuivre sur les spécificités mathématiques de la problématique inverse, examinons certains termes s'y rapportant. Ainsi, lorsque les paramètres ou les sources sont accessibles directement par des mesures, on parle de **mesure directe**. Mais dans de très nombreux cas, les informations recherchées ne sont pas directement mesurables, mais sont reliées physiquement à des grandeurs mesurables, on est dans le cas de **mesure indirecte**. Ces grandeurs mesurables sont évidemment dépendantes de la physique du phénomène étudié mais aussi des appareils de mesure utilisés. Les lois physiques qui relient les observables aux grandeurs recherchées sont généralement mathématiquement complexes utilisant par exemple des équations intégrales ou des équations aux dérivées partielles. La résolution du problème permettant le calcul des observables en fonction des grandeurs inconnues que l'on qualifie de **problème direct** est habituellement maîtrisée et plus simple que celle du problème permettant le calcul des grandeurs inconnues en fonction des observables qui est qualifié de **problème inverse**.

En première approche, nous pouvons dire qu'un problème inverse résulte de la confrontation de données expérimentales et des résultats de simulations numériques. Il consiste à chercher à minimiser l'écart entre la mesure et le calcul.

Une des spécificités des problèmes inverses est d'être **mal posé** ou **incorrectement posé** au sens d'Hadamard : soit l'ensemble des données mesurées ne permet pas d'avoir existence d'une solution du problème considéré, soit il n'y a pas unicité de cette solution ou encore il n'y a pas stabilité de cette solution par rapport à une perturbation des données.

Si en théorie, les données de mesures peuvent former un espace de dimension finie ou infinie, en pratique elles sont discrètes et finies. Lorsque le nombre de paramètres d'un modèle est inférieur au nombre de données des mesures, on est dans le cas d'un problème **sur-déterminé** et on peut éventuellement ajouter un critère pour diminuer ou éliminer les effets de données aberrantes. Si au contraire, le problème consiste à déterminer des paramètres continus et par suite échantillonnés par un très grand nombre de valeurs et si le nombre de résultats d'expériences est insuffisant, il s'agit d'un problème **sous-déterminé**. Il est alors nécessaire d'utiliser des informations *a priori* pour avoir un nombre réduit de solutions possibles ou au mieux une seule. Dans ce cas, on obtient habituellement plusieurs solutions possibles et il est nécessaire de préciser le niveau de confiance que l'on peut accorder à chacune d'entre-elles. Les données peuvent aussi être affectées d'un coefficient de vraisemblance ou pondérées de façon probabiliste. On parle alors d'une **approche Bayésienne** du problème inverse.

Finalement, une définition du problème inverse pourrait être : un problème mal posé ayant pour but l'inversion d'un modèle physique à l'aide d'une image partielle de ses effets.

## Champs d'utilisation - Développement historique

Les problèmes inverses constituent un champ scientifique essentiellement multidisciplinaire, associant les mathématiques à de multiples branches de la physique. Les travaux fondateurs des années soixante-dix sur la résolution des problèmes mal posés (notamment ceux des russes comme A.N. Tikhonov [TA77] et plus récemment de ses élèves [BM91]) et ensuite l'analyse et la mise en oeuvre de stratégies d'inversion pour les problèmes mal posés sont actuellement une branche reconnue des mathématiques appliquées. On rencontre de nombreuses applications de l'analyse inverse dans les sciences physiques ou mécaniques ; de façon non exhaustive : optique, radar, calorimétrie, spectroscopie, géophysique, météorologie, océanographie, acoustique, radioastronomie, contrôle non destructif, génie biomédical, instrumentation, imagerie, génie civil et génie mécanique. Cet ouvrage quant à lui, est limité aux applications concernant la géophysique, le génie civil, la mécanique des matériaux et des structures, la thermique, la reconstruction et l'amélioration d'images.

Un peu d'histoire récente. . . Les chercheurs et ingénieurs travaillant sur des données géophysiques ont grandement contribué au développement de la théorie des problèmes inverses. En effet, les géophysiciens essaient de comprendre le comportement interne de la Terre au moyen de données enregistrées au niveau de sa surface. Avant 1970, les méthodes développées étaient principalement empiriques. Les travaux de G. Backus et de F. Gilbert [BG67]-[BG68] au début des années 70 constituent une première exploration systématique de la structure mathématique des problèmes inverses et sont à l'origine du développement de nombreuses méthodes d'interprétation de données en géophysique. Ainsi, pour résoudre certains problèmes inverses d'identification de sources décrits par une équation intégrale linéaire de première espèce, ces auteurs ont proposé une méthode numérique qui porte leur nom et qui a suscité de nombreux articles et ouvrages. Elle a été souvent appliquée pour inverser des données sismiques afin d'obtenir les profils de densité à l'intérieur de la Terre [BG70]. Notons également l'ouvrage plus récent de W. Menke [Men89] sur l'analyse des données géophysiques et la théorie inverse discrète.

En France, dans les années 80, A. Tarantola est un des précurseurs de la théorie des problèmes inverses qu'il résout principalement au moyen de modèles probabilistes (approche bayésienne). Les applications concernent le traitement des données géophysiques. Il est l'auteur de plusieurs ouvrages [TM82], [Tar87] et de nombreux articles [MT02] sur le sujet. L'utilisation des méthodes inverses pour le génie mécanique et le génie civil connaît un regain d'intérêt depuis une dizaine d'années avec les travaux de H.D. Bui. L'équipe de recherche « Problèmes inverses-Identification-Optimisation » du Laboratoire de Mécanique des Solides de l'Ecole Polytechnique regroupe autour lui des chercheurs très actifs dans le domaine comme M. Bonnet, A. Constantinescu et C. Stolz (cf. [BC99], [BC05], [BCM04], [PS01], [PS03]). L'ouvrage [Bui93] constitue une excellente introduction aux problèmes inverses en mécanique des matériaux. Les applications des techniques d'identification proposées concernent la détection non-destructive et la caractérisation de défauts internes, des inhomogénéités et hétérogénéités, l'identification des singularités en mécanique de la rupture et l'identification des paramètres physiques des matériaux. Un numéro spécial des Comptes Rendus de l'Académie des Sciences [CR008], dédié à H.D. Bui, vient d'être consacré aux problèmes inverses et aux problèmes non-linéaires au vu de leur importance dans la mécanique des solides de nos jours.

On peut aussi consulter trois sites internet spécialement dédiés aux problèmes inverses où l'on peut trouver des informations utiles telles que des listes de personnes impliquées par ces problèmes, les annonces des séminaires et conférences passées et à venir, les nouveaux ouvrages ainsi qu'un choix d'articles scientifiques sur le sujet :

**<http://www.inverse-problems.net/>** (Inverse Problems International Association : IPIA)

**<http://www.me.ua.edu/inverse/>** (Université d'Alabama )

**<http://www.mth.msu.edu/ipnet/>** (Inverse Problems Network : IPNet) .

Parmi les revues scientifiques consacrées aux problèmes inverses, on peut citer :

**Journal of Inverse Problems in Science and Engineering** (Taylor Francis Group - G.S. Dulikravich, Editor),

**Journal of Inverse Problems** (IOP electronics journals - F.A. Grünbaum, Editor),

**Journal of Inverse and Ill-Posed Problems** (Walter de Gruyter - M.M. Lavrentiev, Editor),

**Inverse problems and Imaging** (AIMS American Institute of Mathematical Sciences - Lassi Päivärinta, Editor).

On peut également noter un numéro spécial du **Journal européen des systèmes automatisés** sur l'identification des systèmes paru très récemment [GMP08].

Enfin, de façon plus modeste, voilà bientôt vingt ans que des recherches réalisées au sein du LCPC concernent des applications des méthodes inverses au génie civil et urbain, citons par exemple les thèses suivantes :

- les techniques de tomographie sismique avec les reconstructions géométriques à deux et trois dimensions pour la reconnaissance géophysique : [Cot88],
- l'identification modale de structures sous vibrations : [Arg90] et [Cré90],
- l'identification du comportement vibratoire de bâtiments soumis aux séismes : [Afr91],
- l'identification de paramètres dans les modèles de prévision de la pluie en hydrologie : [And98] ;

ainsi que quelques ouvrages plus récents relatifs aux problèmes inverses et rédigés par des chercheurs du LCPC :

- les techniques de contrôle non destructif par exemple pour l'évaluation non destructive de l'état d'altération des ouvrages en béton : [BA05],
- les plans d'expériences- un outil indispensable à l'expérimentation [Lin05].

### **Différentes approches pour les résoudre**

La théorie de l'inversion ou des méthodes inverses est un ensemble organisé de techniques mathématiques qui opèrent sur un nombre réduit de données d'un problème afin d'obtenir une information utile, pertinente sur le système physique réel étudié, à partir d'inférences tirées des observations. Ces techniques doivent tenir compte du caractère éventuellement incomplet, imprécis et/ou redondant des données. On ne se limite pas à la recherche de solutions au sens mathématiques du terme mais

on cherche plutôt à inventorier l'ensemble complet des solutions définies à des incertitudes près. Parmi les nombreuses solutions, on opère ensuite un choix suivant des critères additionnels (vraisemblance physique, informations supplémentaires *a priori*, etc).

Dans le cas général, on recherche la ou les solution(s) stable(s) au sens de Hadamard. Dans cette optique, les méthodes d'investigation des problèmes inverses peuvent être classées en trois grandes catégories suivant les fondements de chacune à savoir :

- **l'analyse fonctionnelle**

Le problème inverse, mal posé par essence, est modifié en problème bien posé en jouant sur le choix des espaces qui servent à décrire les variables et de leur topologie qui permettent de définir les notions d'écart ou d'erreur. Ce choix relève principalement de considérations physiques et non mathématiques. Les méthodes correspondantes proposent généralement l'introduction de contraintes globales sur les classes de solutions.

- **la régularisation des problèmes mal posés**

La solution obtenue par inversion régularisée va dépendre continûment des données et va tendre vers la solution exacte (en supposant qu'elle existe). La régularisation d'un problème inverse consiste à une réécriture de celui-ci et s'articule suivant deux lignes directrices :

- définir un cadre pour prendre en compte des informations *a priori* supplémentaires « extérieures aux mathématiques »,
- assurer la stabilité de la nouvelle solution qui tient compte des informations additionnelles par rapports aux erreurs de mesure.

Pour régulariser un problème inverse, plusieurs manières existent, qui par ailleurs, peuvent être combinées. En voici quelques-unes :

- régularisation de Tikhonov (méthode la plus connue cf. [TA77] [EHN96]) au moyen de l'ajout dans la fonctionnelle à minimiser d'une fonctionnelle stabilisatrice non-négative qui peut prendre des valeurs positives aussi petites que possible et qui tient compte de l'information *a priori* disponible,
- réduction du nombre de paramètres afin de diminuer la sensibilité du critère aux fluctuations des données,
- introduction de contraintes (égalités ou inégalités vérifiées par les inconnues) dans la fonctionnelle à minimiser pour limiter la recherche à des valeurs physiquement acceptables,
- filtrage des données des essais par des techniques de traitement des signaux (filtres fréquentiels, transformées de Fourier modifiées, transformées temps-fréquence, etc.). Ainsi, les transformées temps-fréquence comme l'analyse en ondelettes peuvent se révéler particulièrement adaptées pour analyser des données qui évoluent au cours du temps,
- méthode de quasi-réversibilité initialement proposée par R. Lattès et J.L. Lions [LL93] bien adaptée pour les problèmes de Cauchy comme par exemple, dans les équations aux dérivées partielles elliptiques [Bou07]. Elle consiste à modifier l'opérateur différentiel ou aux dérivées partielles, en général, en changeant l'ordre des dérivées qui y interviennent de manière à obtenir un nouveau problème qui est bien posé pour les données initiales ou les limites connues.

- **l'inversion stochastique ou bayésienne**

Dans ce cas, toutes les variables sont considérées comme aléatoires afin de représenter toutes les incertitudes. On s'intéresse alors à la fonction densité de probabilité associée aux inconnues et aux données du problème à partir de laquelle on cherche des grandeurs caractéristiques : valeur moyenne, valeur de plus grande probabilité, dispersion, corrélations, etc...

Comme nous l'avons mentionné au paragraphe précédent, un problème inverse se ramène le plus souvent à un problème de minimisation entre calcul et résultats d'expériences ou à un problème de maximisation de la vraisemblance et donc conduit à un problème d'optimisation qui est souvent un problème en soi.

Les techniques numériques utilisées pour résoudre un problème d'**optimisation** sont aussi variées que les domaines d'application des problèmes inverses eux-mêmes (moindres carrés linéaires ou non-linéaires, maximum de vraisemblance, méthode de Monte-Carlo, programmation linéaire, recuit simulé, algorithmes génétiques ou évolutionnaires, méthodes du contrôle optimal, etc). L'objectif de cette introduction n'est pas de dresser une liste détaillée des méthodes d'optimisation. Pour plus de renseignements, le lecteur peut se référer à une abondante littérature parmi laquelle nous avons retenu les ouvrages suivants : les livres de Fletcher [Fle80] et Ciarlet [Cia82] sont des grands classiques, l'ouvrage de Culioli [Cul94] est simple et présente les méthodes numériques les plus usuelles, celui de Bergounioux [Ber01] concerne à la fois l'optimisation et le contrôle tandis [Gro93] est plutôt orienté vers les problèmes inverses ; sur les méthodes déterministes et les problèmes convexes, on peut consulter [HP95] et sur les méthodes non déterministes et l'optimisation discrète [PR02]. Enfin, en 2009, une encyclopédie en anglais sur l'optimisation sera rééditée [FP].

Notons que si le problème à traiter fait intervenir des paramètres assurant la régularité et la stabilité du critère, on est dans le cadre de l'optimisation dite « classique ». Les méthodes numériques pour l'optimisation classique calculent des dérivées premières et/ou secondes de la fonction coût par rapport aux inconnues. Les algorithmes itératifs pour l'optimisation font habituellement appel de façon répétitive à la résolution numérique du problème direct associé. Le problème direct est fréquemment défini à l'aide d'équations aux dérivées partielles ou de systèmes d'équations différentielles par rapport au temps. Là encore, les méthodes de résolution abondent et pour ne citer que les plus courantes à savoir : les différences finies, les éléments finis et les équations intégrales.

En l'absence d'hypothèses de régularité, il devient difficile de trouver un optimum global ou de bons optimums locaux par les méthodes classiques ; il faut alors faire appel à des méthodes de minimisation globale, par exemple de type génétiques ou évolutionnaires. Efficaces mais ne permettant pas toujours de disposer de théorèmes de convergence, elles constituent ce que l'on peut appeler des méthodes d'optimisation d'ordre zéro ou encore « métaheuristiques ». Les métaheuristiques sont généralement des algorithmes stochastiques itératifs, qui progressent vers l'extremum global de la fonction objectif par échantillonnage. Ces méthodes ont connu un essor considérable ces dernières années grâce au développement des moyens de calcul. Les ouvrages récents, [DPST03] en français et [dC06] en anglais, présentent une famille de métaheuristiques telles que le recuit simulé, la recherche avec tabous, les algorithmes évolutionnaires et génétiques, les colonies de fourmis pour ne citer que les principales.

## Présentation de l'ouvrage

Les problèmes inverses sont un domaine foisonnant. Nous l'avons vu, les applications sont multiples, les problématiques diverses et les méthodes utilisées forcément variées. L'objectif de cet ouvrage est de présenter un panorama de plusieurs méthodes utilisées et de quelques problèmes traités au sein du réseau scientifique et technique du Ministère de l'écologie, de l'énergie, du développement durable et de l'aménagement du territoire, dans le cadre de la mécanique et du génie civil en particulier.

Cet ouvrage commence par des exemples de **problèmes inverses linéaires sur-déterminés** en dimension finie. Typiquement ce sont des **problèmes d'identification de paramètres**.

Le premier chapitre est un exemple type de **problème inverse linéaire** et il présente une « *Méthodologie pour l'identification des systèmes mécaniques articulés* ». La résolution proprement dite doit être précédée d'une première étape de recensement des connaissances *a priori*, et d'une seconde étape qui consiste à extraire des mesures les informations pertinentes, soit par filtrage, mise à l'échelle, etc. On détermine ensuite les paramètres identifiables qui peuvent constituer un jeu minimal pour définir le système dynamique considéré. Ces paramètres interviennent ici de façon linéaire dans les équations différentielles du modèle. Pour conserver la linéarité du problème inverse, on est amené à calculer numériquement des dérivées. On résout le problème inverse correspondant qui est **sur-déterminé**. Une étape essentielle est l'évaluation de la **précision de l'estimation** obtenue. Cette méthodologie est utilisée pour identifier les caractéristiques mécaniques d'un compacteur.

L'**identification modale** est un cas particulier de l'identification de paramètres. Il s'agit d'estimer les fréquences, les modes propres et les taux d'amortissement modaux de systèmes mécaniques linéaires à paramètres constants à partir de leurs réponses dynamiques dans le domaine des basses fréquences. C'est ce thème qui est abordé dans les chapitres suivants.

On commence par « *Algorithmes récursifs pour l'identification modale des structures de génie civil* ». L'identification présentée utilise le modèle linéaire ARMA et se fonde sur des essais sous **excitations ambiantes**, c'est-à-dire pour lesquels l'entrée n'est pas connue, et seule la sortie peut être mesurée. En l'occurrence, l'auteur s'intéresse aux vibrations de câbles de ponts à haubans sous l'action du vent. L'accent est mis sur l'**aspect récursif des algorithmes** (utile en théorie du contrôle, en traitement du signal, en estimation statistique de paramètres). En utilisant à chaque instant de nouvelles mesures disponibles, ces méthodes récursives temporelles permettent de suivre l'évolution des paramètres modaux.

L'**analyse en ondelettes** est un outil fondamental pour le traitement des signaux de mesure afin d'en faciliter l'analyse. Son utilisation permet d'identifier les paramètres modaux et leur évolution au cours du temps. Dans le chapitre « *Analyse en ondelettes pour l'identification de paramètres en dynamique* » trois exemples sont présentés utilisant respectivement trois types de données : les fonctions de réponses en fréquence de poutres, les réponses après choc de plaques en plexiglass, et enfin les réponses sous excitations ambiantes de bâtiments. Le chapitre se termine par une extension au cas de l'**analyse vibratoire** de structures dynamiques **non linéaires**.

Le chapitre sur la « *Modélisation et l'identification de systèmes discrets hystérétiques en dynamique* », présente différents modèles d'hystérésis construits à l'aide de modèles rhéologiques simples (ressorts, amortisseurs ou patins). Il utilise en prétraitement **l'analyse en ondelettes**, de façon à préciser pour l'évolution simulée, les plages de fonctionnement où le système évolue avec un nombre figé et connu de patins bloqués ou débloqués. Globalement le problème est **non linéaire**, mais linéaire dans chacune des plages détectées grâce à l'analyse en ondelettes.

Une autre manière de déterminer des changements de comportement est proposée dans le chapitre sur des « *Modèles probabilistes de détection de rupture de comportement* ». Des exemples sont proposés comme la détection du point de rupture entre les évolutions élastique et plastique d'un matériau sollicité en traction. Basées sur les moindres carrés linéaires et sur des considérations de probabilités, ces méthodes peuvent permettre de trouver à la fois le point de rupture et les paramètres correspondant aux modèles respectifs avant et après le changement.

L'ouvrage se poursuit par des articles concernant cette fois ci la **détermination des causes** et non plus des paramètres des modèles. On commence par des problèmes inverses qui sont aussi **sur-déterminés** et le dernier chapitre est consacré à des problèmes **sous-déterminés** issus de la géophysique.

Le premier exemple de ce type concerne l'« *Assimilation thermique 1D par méthode adjointe libérée* ». Connaissant l'évolution de la température en un certain nombre de points du milieu, il s'agit de déterminer la température initiale et le flux de chaleur entrant. Le problème étant mal posé, il est régularisé à l'aide de la **méthode de Tikhonov**. Pour éviter le calcul du gradient de la fonctionnelle à minimiser, on utilise la **méthode de l'état adjoint**, ce qui permet de réduire le coût de calcul. Les résultats obtenus en choisissant la norme  $H^1$  sont meilleurs que ceux obtenus avec la norme  $L^2$ . La connaissance de la donnée initiale et du flux entrant permet de déterminer ensuite l'évolution de la température ; ceci afin de prendre en compte l'effet de la température dans l'état de précontrainte des structures.

Un autre exemple d'**identification des causes** concerne la **détection et la reconnaissance d'images**. Mais les **moindres carrés linéaires** sont parfois mal adaptés car ils ne prennent pas en compte le caractère aléatoire des mesures et des paramètres cherchés. Le chapitre « *Théorie semi-quadratique, régularisation et estimation robuste* » montre comment, en ajoutant des paramètres dans la fonctionnelle de la méthode classique des moindres carrés, on la transforme en une fonctionnelle qui **n'est plus quadratique** mais qui permet de tenir compte des informations *a priori* dont on dispose. Cette théorie est utilisée pour la détection et la reconnaissance d'images, elle est également appliquée à la reconstruction d'images.

Mais dans certains cas, l'absence de régularité *a priori* de la fonctionnelle à minimiser, impose d'utiliser les « *Algorithmes pour l'optimisation globale* ». Ce chapitre fait le tour de différentes **méthodes d'optimisation globale** fondées sur différentes **métaphores** et qui ont montré leur efficacité comme le recuit simulé, les essaims de particules, la sélection clonale, les algorithmes évolutionnaires, etc. Une boîte à outils en langage Scilab est fournie, avec aide en ligne et banc d'essai.

Les algorithmes génétiques, qui sont un cas particulier des algorithmes évolutionnaires, sont utilisés pour l'« *Optimisation de forme de murs antibruit* ». Le problème

étant mal posé, il est d'abord régularisé. Les calculs mettent en oeuvre des **méthodes des éléments de frontières (ou méthodes intégrales)** pour modéliser et résoudre le problème direct.

Le chapitre suivant concerne des « *Méthodes de Monte Carlo par chaînes de Markov pour l'identification de systèmes élastiques hétérogènes* ». Ces méthodes se basent sur l'échantillonnage aléatoire d'un certain domaine, qui dans le cas d'une analyse bayésienne est celui des entrées du système qu'il faut identifier. L'ensemble des échantillons forme un chemin aléatoire et la convergence de la méthode dépend de la nature de ce chemin. Dans l'algorithme par chaînes de Markov utilisé un noyau de transition régit le passage d'un échantillon au suivant et il s'adapte selon les résultats antérieurs.

Enfin, certains problèmes sont des **problèmes sous-déterminés**, car les inconnues sont de dimension infinie alors que les mesures sont en quantité finie.

Le chapitre « *Méthodes inverses en géophysique* » présente différentes contributions concernant la caractérisation de sous-sol, l'auscultation d'ouvrages d'art, ... La caractéristique commune de ces problèmes est d'être largement sous déterminés. On ne peut espérer une solution unique ; on peut, tout au plus, affecter à chaque solution possible un indice de confiance. Les exemples traités concernent la dispersion des ondes de surfaces et différents problèmes de **tomographie** : sismique, électromagnétique ou de résistivité électrique. Les méthodes de résolution font appel soit à des linéarisations simplificatrices, soit à des modules de résolution non-linéaires.

## Conclusion

Le but premier de cet ouvrage est de donner au lecteur une aide efficace dans son approche des problèmes inverses. Les principales caractéristiques de la problématique inverse sont rappelées et détaillées tout au long des différentes contributions. On peut noter que le traitement d'un problème n'est pas en général « standard », notamment pour ce qui est de sa formalisation et de son analyse mais aussi pour ce qui est des méthodes de résolution. Cependant la formalisation et la résolution des différents problèmes inverses présentés dans cet ouvrage peuvent générer des idées, des procédés, des remarques qui peuvent être utiles d'une application à une autre application. Nous espérons alors que les méthodes ou algorithmes décrits ici, qui constituent les briques de base pour la résolution de tels problèmes, pourront être largement et facilement réutilisés.

En élargissant notre réflexion, dans notre monde où les méthodes de simulation réalisent chaque jour des progrès étonnants, le cheminement inverse permet de reconsidérer, de repenser les notions classiquement acquises. Un exemple immédiat est le concept de mesure qui repose de plus en plus sur l'imbrication de trois éléments : l'expérimentation, les techniques de simulation directe et les algorithmes d'inversion. Ainsi, la confrontation entre modèle et mesures devient une voie prometteuse qui permet soit une meilleure compréhension du phénomène physique étudié, soit la surveillance de l'évolution de l'état du système à l'étude, ou encore la validation de mesures imparfaitement fiabilisées et pouvant considérablement perturber un bilan. En poussant encore plus loin, la pensée inverse peut permettre par des raisonnements différents et en bouleversant les notions classiques, d'inciter les étudiants, ingénieurs et chercheurs à avoir des idées originales ouvrant des perspectives fécondes.



# Bibliographie

- [Afr91] H. Afra. *Identification du comportement sismique de bâtiments à partir de leurs réponses accélérométriques*. Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, 1991.
- [And98] H. Andrieu. *Interprétation de la mesure radar météorologique pour l'hydrologie - Modélisation des bassins versants urbanisés*. Thèse d'habilitation à diriger des recherches, Institut National Polytechnique de Grenoble, 1998. 67 pages + Annexes.
- [Arg90] P. Argoul. *Identification des structures vibrantes*. Thèse de doctorat, Ecole Nationale des Ponts et Chaussées, 1990. 283 pages.
- [BA05] D. Breyse and O. Abraham. *Méthodologie d'évaluation non destructive de l'état d'altération des ouvrages en béton*. Association française de génie civil, 2005.
- [BC99] H.D. Bui and A. Constantinescu. Problèmes inverses. de l'expérimentation à la modélisation. ARAGO. OFTA, 1999.
- [BC05] M. Bonnet and A. Constantinescu. Inverse problems in elasticity. *Inverse Problems*, 21(2) :R1–R50, 2005.
- [BCM04] H.D. Bui, A. Constantinescu, and H. Maigre. Numerical identification of linear cracks in 2d elastodynamics using the instantaneous reciprocity gap. *Inverse Problems*, 20(4) :993–1001, 2004.
- [Ber01] M. Bergounioux. *Optimisation et contrôle optimale des systèmes linéaires*. Dunod, 2001.
- [BG67] G.E. Backus and J.F. Gilbert. Numerical applications of a formalism for geophysical inverse problems. *Geophys. J. Roy. Astr. Soc.*, 1967.
- [BG68] G.E. Backus and J.F. Gilbert. The resolving power of gross earth data. *Geophys. J. Roy. Astr. Soc.*, 1968.
- [BG70] G.E. Backus and J.F. Gilbert. Uniqueness in the inversion of inaccurate gross earth data. *Phil. Trans. Roy. Soc. London Ser. A*, 1970.
- [BM91] V. Badeva and V. Morozov. *Problèmes incorrectement posés*. Masson, Paris, 1991.
- [Bou07] L. Bourgeois. A stability estimate for ill-posed elliptic cauchy problems in a domain with corners. *C. R. Acad. Sci. Paris, Ser. I*, 345 :385–390, 2007.
- [Bui93] H.D. Bui. *Introduction aux problèmes inverses en mécanique des matériaux*. Collection de la Direction des Etudes et Recherches d'Electricité de France. Eds Eyrolles, 1993.

- [Cia82] P.G. Ciarlet. *Optimisation et contrôle optimale des systèmes linéaires*. Masson, 1982.
- [Cot88] P. Cote. *Tomographies sismiques en génie civil*. Thèse de doctorat d'état ès sciences physiques, Université de Grenoble, 1988. 193 pages.
- [CR008] Duality, inverse problems and nonlinear problems in solid mechanics. volume 336, pages 1–259. janvier 2008.
- [Cré90] C. Crémona. *Recursive time domain modal identification techniques*. Thèse de doctorat, University of Wales, 1990. 304 pages.
- [Cul94] J.C. Culioli. *Introduction à l'optimisation*. Ellipses, 1994.
- [dC06] L. Nunes de Castro. *Fundamentals of natural computing : basic concepts, algorithms, and applications*. Chapman Hall/CRC, 1 edition, 6 2006.
- [DPST03] J. Dréo, A. Pérowski, P. Siarry, and E. Taillard. *Métaheuristiques pour l'optimisation difficile*. Algorithmes. Eds Eyrolles, 2003.
- [EHN96] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Kluwer, 1996.
- [Fle80] R. Fletcher. *Practical methods of optimization*. John Wiley and Sons, 1980.
- [FP] C.A. Floudas and P.M. Pardalos, editors. *Encyclopedia of optimization*.
- [GMP08] H. Garnier, M. M'Saad, and T. Poinot. Identification des systèmes. volume 42. 04-2008.
- [Gro93] C.W. Groetsch. *Inverse problems in the mathematical sciences*. Vieweg, Braunstein/Wiesbaden, 1993.
- [HP95] R. Horst and P.M. Pardalos, editors. *Handbook of global optimization*. Kluwer, 1995.
- [Lin05] R. Linder. *Les plans d'expériences - Un outil indispensable à l'expérimentation*. 2005.
- [LL93] R. Lattès and J.L. Lions. *Méthode de quasi réversibilité et applications*. Eds Eyrolles, 1993.
- [Men89] W. Menke. *Geophysical data analysis : Discrete inverse theory, revised version*, volume 45 of *International Geophysics Series*. Academic Press San Diego, 1989.
- [MT02] K. Mosegaard and A. Tarantola. Probabilistic approach to inverse problems. chapter International Handbook of Earthquake and Engineering Seismology, pages 237–265. Academic Press, 2002. Part A.
- [PR02] P.M. Pardalos and H.E. Romeijn, editors. *Handbook of global optimization Volume 2 (Nonconvex optimization and its applications)*. Springer, 2002.
- [PS01] M. Peigney and C. Stolz. Approche par contrôle optimal du comportement asymptotique des structures élastoviscoplastiques sous chargement cyclique. *C. R. Acad. Sci.*, 329 :643–648, 2001.
- [PS03] M. Peigney and C. Stolz. An optimal control approach of inelastic structures under cyclic loading. *J. Mech. Phys. Solids*, 51 :575–605, 2003.
- [TA77] A.N. Tikhonov and V.Y. Arsenin. *Solutions to ill-posed problems*. Winston-Wiley, New-York, 1977.

- 
- [Tar87] A. Tarantola. *Inverse Problem Theory*. Elsevier, 1987.
- [TM82] A. Tarantola and J. Morlet. Inverse problems=quest for information. *J. Geophys.*, 50 :159–170, 1982.



## Chapitre 1

# Méthodologie pour l'identification des systèmes multicorps articulés : application au compacteur

*Pierre-Olivier Vandanjon,  
Maxime Gauthier,  
Charles-Eric Lemaire*

Connaître le torseur des efforts appliqués par un engin de chantier sur le matériau qu'il met en œuvre, améliorer la précision des mouvements d'une machine outil rapide, implanter une commande avec retour d'effort sur un télémanipulateur du nucléaire, ... Toutes ces applications utilisent une modélisation qui représente la réalité sous la forme d'un système multicorps articulé. C'est-à-dire sous la forme d'une liste de corps liés entre eux par des articulations de rotation ou de translation. Ces applications nécessitent de connaître le modèle dynamique qui relie les couples et forces appliqués sur le système aux positions, vitesses, accélérations. Ce modèle dépend de paramètres (masses, centre de gravité, inerties, ...) des corps composant le système. Ces paramètres ne sont généralement pas suffisamment connus par rapport aux applications visées. L'objet de ce chapitre est de proposer une méthode d'identification des paramètres basée sur des mesures effectuées lors de mouvements du système mécanique. En mesurant les forces, couples, positions, vitesse et accélérations, nous déduisons avec une incertitude connue les paramètres du système. C'est un problème inverse car il s'agit bien de remonter aux causes en analysant les effets. Nous présentons l'application de cette méthode sur un engin de chantier : un compacteur CATERPILLAR CB 544 (FIG. 1.1). Cette application, présentée en détail dans la thèse de [Lem05], a permis d'estimer, pour la première fois, le torseur des efforts appliqué par le compacteur sur de l'enrobé chaud lors d'un chantier routier.

Notre méthodologie passe par trois étapes :

1. planification de mouvements
2. analyse d'un mouvement
3. résolution

La première étape consiste d'abord à recenser toutes les connaissances a priori (avant les expérimentations) que nous avons sur le système. Par exemple, sur un engin de chantier, il s'agit de rassembler la documentation technique permettant de connaître



FIG. 1.1 – Compacteur CATERPILLAR CB 544

la masse totale du compacteur, sa géométrie, les caractéristiques de son système vibratoire, ... A partir de ces informations, des mouvements qui permettent une bonne identification du modèle sont recherchés en simulation en se basant sur une analyse physique. La deuxième étape consiste à extraire des mesures effectuées lors d'un mouvement, les informations pertinentes qui seront utiles pour l'identification. Il s'agit de filtrer les signaux, de retirer les informations redondantes et de synthétiser ces informations sous une forme compacte, mise à l'échelle afin de pouvoir les réutiliser facilement et, notamment, de les regrouper avec d'autres informations de manière harmonieuse. La troisième partie est l'étape la plus classique de l'identification. Les informations de chaque mouvement sont assemblées dans un système complet qui est alors analysé pour en déduire un sous ensemble de paramètres essentiels. Nous calculons alors une précision sur chaque paramètre. Si cette précision n'est pas satisfaisante, une nouvelle étape de planification est proposée. Ces trois étapes sont indispensables à la réussite du processus d'identification. Trop souvent, la résolution est menée sans avoir réfléchi préalablement au domaine de validité du modèle, aux ordres de grandeur des paramètres, au type de filtrage à mettre en œuvre, à la façon de synthétiser les informations de chaque expérimentations, à la manière de comparer ces informations de manière cohérente. Dans ce cas, l'expérimentateur se retrouve désemparé avec d'un côté, des valeurs de paramètres qui ne sont pas physiquement interprétables et de l'autre, une quantité importante de données inexploitable.

Entre l'étape planification des mouvements et l'étape analyse des mouvements, il y a la réalisation des expérimentations. Lors de la préparation et du déroulement des expérimentations, la maîtrise des systèmes de mesure est fondamentale. Elle est nécessaire à la réussite de l'identification. Cette partie est spécifique à chaque domaine d'application, elle sort donc du cadre de ce livre dédié aux problèmes inverses. En dehors du génie civil, cette méthodologie a été appliquée avec succès à

- un robot expérimental deux axes (SCARA) de l'IRCCyN par [Gau97],

- un robot industriel six axes avec une chaîne fermée, le SR400 de Renault Automation (RA, ACMA). Sur les 56 paramètres identifiables structurellement, 31 paramètres essentiels ont été identifiés par [Res96],
- un robot prototype RACE du CEA, constitué de trois axes avec une chaîne fermée par [Van95],
- un robot industriel six axes RX90 de Staubli. Sur les 76 paramètres identifiables, 42 paramètres essentiels ont été identifiés ;
- l'identification de raideurs articulaires et de structure des machines de travail très grande vitesse (TGV), utilisées en transitique rapide, en collaboration avec SEPRO Robotique ou en usinage très grande vitesse (machines outils UGV, UTGV). Ces machines sont décrites par des modèles multi-corps avec masses et raideurs localisées par [Pha02],
- l'identification des paramètres dynamiques d'une voiture en collaboration avec PSA par [Ven03].

Ce chapitre est divisé en cinq parties. La première présente le cadre du problème. Les trois suivantes présentent les trois étapes de notre méthodologie. La dernière partie conclut ce chapitre. Chaque partie est illustrée par une application sur le compacteur de chantier. Ces illustrations sont simplifiées pour des raisons pédagogiques. Une présentation complète et rigoureuse est donnée dans la thèse de [Lem05].

## 1.1 Position du problème

L'identification est une étape de la modélisation, elle est liée à celle-ci. La méthodologie que nous proposons ici est liée à une modélisation usuelle en robotique qui est exposée dans le livre de [Kha99]. Elle considère le robot comme un système multicorps articulé c'est-à-dire des corps rigides reliés entre eux par des articulations rotoides ou prismatiques éventuellement élastiques. L'intérêt majeur de cette modélisation pour l'identification est qu'elle permet de construire des modèles dynamiques qui sont *linéaires par rapport aux paramètres*. L'objet de l'identification est de connaître ces paramètres et d'évaluer l'incertitude sur cette connaissance. Les paramètres à identifier sont :

- les paramètres inertiels standards  $J$ ,  $MS$  et  $\mathcal{M}$  de chacun des corps,
- les paramètres des articulations élastiques  $k$  et  $c$ ,
- les paramètres de frottement  $FV$  et  $FS$ .

La linéarité du modèle dynamique s'exprime mathématiquement pour l'articulation  $j$  en dérivant par rapport aux paramètres, l'équation (1.1.1).

$$\Gamma_j + \mathbb{Q}_j = A(q)\ddot{q} + C(q, \dot{q}) + \Gamma^e + \Gamma^f \quad (1.1.1)$$

avec :

- $\Gamma_j$  l'effort d'actionnement de l'articulation  $j$ ,
- $\mathbb{Q}_j$  la projection des efforts extérieurs sur l'articulation  $j$ ,
- $A(q)$  la matrice d'inertie du robot,
- $C(q, \dot{q})$  le vecteur des efforts de Coriolis, centrifuges, de gravité,
- $\Gamma^e$  le vecteur des efforts élastiques,
- $\Gamma^f$  le vecteur des efforts de frottement.

On obtient alors l'équation (1.1.2).

$$\Gamma_j + \mathbb{Q}_j = \frac{\partial(A(q)\ddot{q} + C(q, \dot{q}) + \Gamma^e + \Gamma^f)}{\partial X_S} X_S \quad (1.1.2)$$

où  $X_S$  est le vecteur regroupant tous les paramètres de chaque corps. L'équation (1.1.2) peut se mettre sous la forme :

$$\Gamma_j + \mathbb{Q}_j = D_{S_j}(q, \dot{q}, \ddot{q})X_S \quad (1.1.3)$$

L'identification consiste à échantillonner l'équation (1.1.3) le long de mouvements effectués par le robot. En concaténant ces mesures et en les triant selon les articulations actionnées, il est obtenu un système surdéterminé incompatible de  $m$  équations à  $p$  inconnues avec  $m \gg p$ .

$$Y_S = W_S(q, \dot{q}, \ddot{q})X_S + \rho \quad (1.1.4)$$

avec :

$$Y_S = \begin{bmatrix} \Gamma_1(1 \dots e) + \mathbb{Q}_1(1 \dots e) \\ \vdots \\ \Gamma_n(1 \dots e) + \mathbb{Q}_n(1 \dots e) \end{bmatrix} \quad (1.1.5)$$

$$W_S = \begin{bmatrix} D_1(1 \dots e) \\ \vdots \\ D_n(1 \dots e) \end{bmatrix} \quad (1.1.6)$$

où :

- $e$  est le nombre d'échantillons,
- $n$  est le nombre d'articulations,
- $p$  est le nombre de paramètres,
- $m$  le nombre d'équation est égal à  $n \cdot e$ ,
- $\Gamma_j(1 \dots e)$  est le vecteur  $(e \times 1)$  des efforts de l'articulation  $j$ ,
- $\mathbb{Q}_j(1 \dots e)$  est le vecteur  $(e \times 1)$  des efforts généralisés projetés sur l'articulation  $j$ ,
- $D_j(1 \dots e)$  est la matrice  $(e \times n)$  construite en échantillonnant la matrice ligne d'identification de l'articulation  $j$   $D_j(q, \dot{q}, \ddot{q})$ ,
- $\rho$  est le vecteur des résidus dus aux bruits de mesure et aux erreurs de modèle, il traduit l'incompatibilité du système.

L'approche la plus répandue en robotique est fondée sur la résolution au sens des moindres carrés (MC) du système linéaire surdéterminé (1.1.4) ([Gau90]). L'estimation de  $X$  est la solution au sens des moindres carrés du système surdéterminé (1.1.4) :

$$\hat{X} = \operatorname{argmin}_X \|\rho\|^2 = W^+ Y \quad (1.1.7)$$

où :

- $W^+ = (W^T W)^{-1} W^T$  est la matrice pseudo-inverse de  $W$ ,
- $\hat{X}$  est la solution des moindres carrés.

Cette démarche est appelée méthode des moindres carrés d'erreur d'entrée avec modèle inverse. En effet, les forces et les couples, constituants  $Y$ , sont considérés comme les "entrées" du système, les sorties sont les trajectoires : positions, vitesses et accélérations  $(q, \dot{q}, \ddot{q})$ . L'équation (1.1.1) est le modèle dynamique inverse qui relie les sorties aux entrées.

L'unicité de la solution  $\hat{X}$  dépend du rang de la matrice  $W$ . La perte de rang de  $W$  peut avoir deux origines :

- une perte de rang structurelle, qui apparaît quel que soit l'échantillonnage des  $D_j$ . C'est le problème de l'identifiabilité structurelle des paramètres qui est résolu avec le calcul des paramètres minimaux du modèle ;

- une perte de rang liée au mauvais choix des échantillons des  $D_j$ . C'est le problème de la stratégie optimale des mesures qui est résolu en réalisant une identification à partir de mouvements excitants pour les paramètres.

**Application au compacteur** Pour illustrer la mise en équation du problème, l'exemple de la bille d'un compacteur est présenté ici. Le modèle correspond à la rotation de la bille autour de son axe actionnée par un moteur hydraulique. L'équation (1.1.1) prend la forme suivante :

$$\begin{aligned} \Gamma_j + \mathbb{Q}_j = & ZZ_j \dot{\omega}_j^z + YZ(\dot{\omega}_j^y + \omega_j^x \omega_j^y) + XZ(\dot{\omega}_j^x - \omega_j^x \omega_j^y) \\ & + MX_j \dot{V}_j^y - MY_j \dot{V}_j^x + FV_j \dot{q}_j + FS_j \text{signe}(\dot{q}_j) \end{aligned} \quad (1.1.8)$$

avec, les paramètres du modèle :

- $XZ_j, YZ_j, ZZ_j$  les composantes de la matrice d'inertie de la bille,
- $MX_j, MY_j$  les composantes du premier moment d'inertie de la bille,
- $FS_j, FV_j$  les paramètres du modèle de frottement.

les variables du modèle :

- $\Gamma_j$  le couple du moteur,
- $\mathbb{Q}_j$  le moment des efforts extérieurs autour de l'axe de rotation de la bille,
- $\omega_j = [\omega_j^x \ \omega_j^y \ \omega_j^z]^T$  et  $V_j = [V_j^x \ V_j^y \ V_j^z]^T$  les composantes du torseur cinématique de la bille qui s'exprime en fonction de  $q$  et  $\dot{q}$ ,
- $q_j$  l'angle de rotation entre le rotor et le stator du moteur.

La forme matricielle du modèle (équation 1.1.4) se met sous la forme :

$$Y_S = \Gamma_j + \mathbb{Q}_j \quad (1.1.9)$$

$$W_S = [\dot{\omega}_j^z \ \dot{\omega}_j^y + \omega_j^x \omega_j^y \ \dot{\omega}_j^x - \omega_j^x \omega_j^y \ \dot{V}_j^y \ -\dot{V}_j^x \ \dot{q}_j \ \text{signe}(\dot{q}_j)] \quad (1.1.10)$$

$$X_S = [ZZ_j \ YZ_j \ XZ_j \ MX_j \ MY_j \ FV_j \ FS_j]^T \quad (1.1.11)$$

### 1.1.1 Identifiabilité structurelle : les paramètres inertiels de base

La perte de rang structurelle est liée à la structure même du modèle. Elle apparaît quel que soit le mouvement  $q, \dot{q}, \ddot{q}$  échantillonné dans  $W$ . Elle traduit le caractère non identifiable des paramètres  $X_S$  par le fait que le paramétrage est surabondant. Ce problème est résolu en calculant un jeu de paramètres identifiables  $X_b$ , appelés paramètres de base ou encore paramètres minimaux car c'est le nombre minimal des paramètres dynamiques qui permettent de calculer les effort moteurs  $\Gamma$ . Les paramètres inertiels de base sont obtenus en éliminant les paramètres inertiels standards qui n'interviennent pas dans le modèle dynamique (1.1.3) et en regroupant d'autres paramètres sous forme de relations linéaires. Une méthode symbolique et des méthodes numériques, basées sur la factorisation QR ou SVD d'une matrice  $W$  calculée avec un échantillonnage aléatoire de  $D$ , ont été proposées dans [Gau90], [GK90], [Gau91] et implantées dans un logiciel de calcul automatique des modèles de robots SYMORO+ ([KC97]). A l'issue de cette étape de simplification, on obtient le jeu minimal de paramètres constituant les paramètres identifiables du modèle dynamique :

$$X = [X_1^T \ \dots \ X_n^T]^T \quad X_i = [X_{b_i}^T \ X_i^f \ X_i^e]^T \quad (1.1.12)$$

où  $X_b$  est le vecteur des paramètres inertiels de base et  $X$  est le vecteur des paramètres dynamiques identifiables, constitués des paramètres inertiels de base, des

paramètres de frottement et des paramètres d'élasticité. Le modèle dynamique (1.1.3) s'exprime alors sous une forme simplifiée et minimale en nombre de paramètres par :

$$\Gamma + \mathbb{Q} = D(q, \dot{q}, \ddot{q})X \quad (1.1.13)$$

où  $D(q, \dot{q}, \ddot{q})$  est le régresseur minimal associé au vecteur  $X$  des paramètres minimaux.

**Application au compacteur** Si l'on considère les contraintes sur les mouvements de la bille : la bille du compacteur tourne seulement autour de son axe, on obtient des composantes nulles dans le torseur cinématique de la bille (indicée  $j$ ) :

$$\omega_j = [0 \quad 0 \quad \omega_j^z]^T \quad (1.1.14)$$

$$V_j = [g \cos(q_j) \quad g \sin(q_j) \quad 0]^T \quad (1.1.15)$$

Il y a donc des paramètres qui n'interviennent pas dans le modèle de la bille, ils ne sont pas identifiables. Le vecteur des paramètres minimaux est alors :

$$X = [ZZ_j \quad MX_j \quad MY_j \quad FV_j \quad FS_j]^T \quad (1.1.16)$$

### 1.1.2 Biais des moindres carrés

L'échantillonnage du modèle minimal (1.1.13) permet d'obtenir un système linéaire minimal surdéterminé de plein rang structurel :

$$Y = WX + \rho \quad (1.1.17)$$

où  $W$  est la matrice ( $m \times p$ ) d'observation minimale, de plein rang structurel. Par souci de simplicité, on supposera que  $X$  est de dimension  $p$ . On estime les paramètres comme la solution des moindres carrés ordinaires de (1.1.17) :

$$\hat{X} = \operatorname{argmin}_X \|\rho\|^2 = W^+ Y \quad (1.1.18)$$

Lors de la mise en œuvre pratique de cette méthode, les mesures expérimentales ou les estimations de  $\Gamma(t_i)$ ,  $\mathbb{Q}(t_i)$ ,  $q(t_i)$ ,  $\dot{q}(t_i)$ ,  $\ddot{q}(t_i)$  sont bruitées.  $t_i$ ,  $i = 1 \dots e$ , sont les instants d'échantillonnage des mesures. Il en résulte que les matrices  $Y$  et  $W$  sont perturbées et que l'estimation des moindres carrés peut présenter un biais lorsque ces matrices aléatoires ne sont pas indépendantes. Les coefficients du régresseur  $D(q, \dot{q}, \ddot{q})$ , sont des fonctions non linéaires de  $q, \dot{q}, \ddot{q}$  et il est pratiquement impossible de calculer l'effet théorique de ces perturbations sur le biais de l'estimation. On adopte donc une stratégie qui vise à minimiser leurs effets en agissant sur deux points :

- filtrage des données pour calculer un système filtré équivalent à (1.1.17), mais peu perturbé,
- identification à partir d'un mouvement riche en informations qui excite les paramètres, appelé mouvement excitant.

## 1.2 Planification

Dans la section précédente, il est question de procéder à l'identification à partir d'un "mouvement riche en informations". Il s'agit maintenant de trouver un critère quantitatif

permettant de représenter ce concept qualitatif. Ainsi, nous pourrions comparer plusieurs mouvements entres eux, *avant leurs réalisations*, et déterminer ceux qui sont les meilleurs vis à vis de ce critère. Deux représentations d'un "mouvement riche en informations" sont classiquement utilisées :

- un critère sur la précision moyenne de l'estimation : il s'agit d'améliorer globalement le vecteur des paramètres identifiés par rapport à une perturbation.
- un critère sur la précision absolue de l'estimation : il s'agit d'améliorer l'écart-type sur chacun des paramètres.

Nous modifions légèrement ces critères pour prendre en compte la précision relative de l'estimation. Nous montrons que ceci rejoint une interprétation physique du concept d'un "mouvement riche en informations". Dans la suite, nous présentons ces différents critères. Une application au compacteur illustre le choix de mouvements excitants.

### 1.2.1 Précision moyenne de l'estimation : conditionnement et perturbations

Considérons le système (1.1.4) avec  $m = p$  (système carré) par soucis de simplicité, de solution  $X$  tel que :

$$Y = WX \quad (1.2.1)$$

Soit  $X_{\delta Y}$  la solution du système perturbé sur  $Y$  :

$$Y + \delta Y = WX_{\delta Y} \quad (1.2.2)$$

$X_{\delta Y}$  vérifie la relation suivante ([GvL96]) :

$$\frac{\|X_{\delta Y} - X\|}{\|X\|} \leq \text{cond}(W) \frac{\|\delta Y\|}{\|Y\|} \quad (1.2.3)$$

$\text{cond}(W)$  est le conditionnement de la matrice  $W$ , relativement à une norme 2. Ce terme est analysé plus loin dans ce paragraphe.

Soit  $X_{\delta W}$  la solution du système perturbé sur  $W$ , dans le cas de petites perturbations  $\delta W$  :

$$Y = (W + \delta W)X_{\delta W} \quad (1.2.4)$$

$X_{\delta W}$  vérifie la relation suivante :

$$\frac{\|X_{\delta W} - X\|}{\|X\|} \leq \text{cond}(W) \frac{\|\delta W\|}{\|W\|} = \frac{\|\delta W\|}{\Sigma_p(W)} \quad (1.2.5)$$

$\Sigma_p(W)$  est la plus petite valeur singulière de  $W$ .

Par soucis de simplicité, nous nous sommes limités au cas d'une matrice carrée. Il existe des relations similaires pour un système surdéterminé ( $r \times p$ ) et lorsque  $Y$  et  $W$  sont perturbés simultanément (voir ([GvL96]) pour une présentation détaillée).

Le conditionnement et la plus petite valeur singulière sont calculés numériquement à partir de la décomposition en valeurs singulières (SVD) de  $W$ , sous la forme :

$$W = USV^T \quad (1.2.6)$$

où  $U$  ( $m \times m$ ) et  $V$  ( $p \times p$ ) sont deux matrices orthogonales en colonnes et en lignes telles que :

$$U^T U = U U^T = I_m, \quad V^T V = V V^T = I_p \quad (1.2.7)$$

où  $I_m$  est la matrice  $(m \times m)$  identité.

$$S = \begin{bmatrix} \Sigma(W) & & \\ 0_{(m-p) \times p} & & \end{bmatrix} = \text{diag}(\Sigma_i(W)) = \begin{bmatrix} \Sigma_1(W) & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \Sigma_p(W) \end{bmatrix} \quad (1.2.8)$$

$$\Sigma_1 \geq \Sigma_2 \geq \cdots \geq \Sigma_p \geq 0 \quad (1.2.9)$$

La diagonale de  $\Sigma(W)$  contient les valeurs singulières de  $W$  classées par ordre décroissant. Le conditionnement en norme 2 de  $W$  est le rapport de la plus grande valeur singulière sur la plus petite :

$$\text{cond}(W) = \frac{\Sigma_1(W)}{\Sigma_p(W)} \quad (1.2.10)$$

Les relations (1.2.3), (1.2.5), montrent que le conditionnement de  $W$  doit être proche de un avec des valeurs singulières grandes. Cependant le conditionnement en norme 2 fournit un critère d'incertitude en norme 2, c'est-à-dire une mesure de l'incertitude moyenne sur l'ensemble des composantes de  $\hat{X}$ , qui ne prend pas en compte un éventuel déséquilibre de ces composantes. D'où la nécessité d'estimer les incertitudes sur chaque composante de  $\hat{X}$  ([PG93]).

### 1.2.2 Précision absolue de l'estimation

Dans cette partie, nous utilisons des résultats classiques de statistique établis en supposant que dans (1.1.17),  $W$  est déterministe et  $\rho$  est un bruit additif indépendant de moyenne nulle et de matrice de variance-covariance :

$$C_{\rho\rho} = E(\rho\rho^T) = \sigma_\rho^2 I_m \quad (1.2.11)$$

où  $E$  est l'espérance mathématique. La matrice de variance-covariance de l'erreur d'estimation et les écarts-types sont alors estimés par les expressions suivantes :

$$C_{\hat{X}\hat{X}} = E[(X - \hat{X})(X - \hat{X})^T] = \sigma_\rho^2 (W^T W)^{-1} \quad (1.2.12)$$

En utilisant la décomposition en valeurs singulières (SVD) ([GvL96]) de  $W$  (1.2.12), il vient :

$$C_{\hat{X}\hat{X}} = \sigma_\rho^2 V \Sigma^{-2} V^T \quad (1.2.13)$$

En notant  $\sigma_{\hat{X}_i}^2 = C_{\hat{X}_i \hat{X}_i}$  les coefficients sur la diagonale de  $C_{\hat{X}\hat{X}}$ , on définit l'écart-type relatif par :

$$\sigma_{\hat{X}_{ri}} = \frac{\sigma_{\hat{X}_i}}{|\hat{X}_i|} \quad (1.2.14)$$

La variance  $\sigma_\rho^2$  de l'erreur est estimée *a posteriori* par la relation :

$$\hat{\sigma}_\rho^2 = \frac{\|Y - W\hat{X}\|^2}{m - p} \quad (1.2.15)$$

Dans le cas idéal où  $\text{cond}(W) = 1$ , on a  $\Sigma_1(W) = \Sigma_2(W) = \cdots = \Sigma_p(W)$ , soit dans (1.2.13) :

$$C_{\hat{X}\hat{X}} = \left( \frac{\sigma_\rho}{\Sigma_p(W)} \right)^2 I_p \quad (1.2.16)$$

On obtient la même incertitude absolue sur toutes les composantes de  $\hat{X}$  :

$$\sigma_{\hat{X}_i} = \frac{\sigma_\rho}{\Sigma_p(W)}, \quad i = 1 \cdots p \quad (1.2.17)$$

Cette relation suggère un premier critère d'excitation à minimiser proche de (1.2.10), de la forme :

$$f_1(W) = \text{cond}(W) + \frac{1}{\Sigma_p(W)} = \frac{1 + \Sigma_1(W)}{\Sigma_p(W)} \quad (1.2.18)$$

Ce critère tend à uniformiser les écarts types absolus  $\sigma_{\hat{X}_i}$ . Il en résulte que les petits paramètres sont mal estimés dans le cas où  $\hat{X}$  est déséquilibré, c'est à dire lorsque certains paramètres ont des valeurs d'un ordre de grandeurs supérieurs à d'autres paramètres, ce qui est usuel. Dans ce cas, l'écart type relatif est très important :

$$\sigma_{\hat{X}_{ri}} \text{ est grand si } |\hat{X}_i| \text{ est petit, à } \sigma_{\hat{X}_i} \text{ constant.}$$

### 1.2.3 Précision relative de l'estimation

Un deuxième critère a été proposé dans [PG93], [Pre94] pour calculer des mouvements qui équilibrent les écarts types, relativement à des valeurs a priori  $\bar{X}$  de  $X$ . Les connaissances a priori  $\bar{X}$  peuvent provenir des données constructeur ou d'une première estimation obtenue sur des mouvements optimisés suivant le critère  $f_1(W)$  (issue de l'équation 1.2.18). En pondérant l'observation  $W$  et le vecteur des paramètres  $X$  dans (1.1.4), on obtient un nouveau système avec des paramètres  $X_e$  équilibrés :

$$Y = W \text{diag}(\bar{X}) \text{diag}^{-1}(\bar{X}) X + \rho = \Phi X_e + \rho \quad (1.2.19)$$

où  $\text{diag}(\bar{X})$  est une matrice diagonale ( $p \times p$ ) avec les valeurs de  $\bar{X}$  sur sa diagonale.

$$\Phi = W \text{diag}(\bar{X}) = [\Phi_{:,1} \quad \cdots \quad \Phi_{:,p}] = [W_{:,1} \bar{X}_1 \quad \cdots \quad W_{:,p} \bar{X}_p] \quad (1.2.20)$$

où  $\Phi_{:,i}$  et  $W_{:,i}$  désignent les colonnes  $i$  de  $\Phi$  et  $W$  respectivement.  $\Phi$  représente la matrice d'observation  $W$  dont les colonnes sont pondérées par les valeurs a priori  $\bar{X}$ . Ce système équilibré se résout comme le précédent (1.1.4). En particulier, dans le cas idéal où  $\text{cond}(\Phi) = 1$ , l'incertitude absolue est la même sur tous les coefficients de  $\hat{X}_e$ , c'est-à-dire la même incertitude relativement aux connaissances *a priori* sur tous les coefficients de  $X$  :

$$\sigma_{\hat{X}_e} = \frac{\sigma_\rho}{\Sigma_p(\Phi)} = \frac{\sigma_{\hat{X}_i}}{|\bar{X}_i|} = \text{constante}, \quad i = 1, \cdots, p \quad (1.2.21)$$

Ce résultat suggère d'utiliser le critère  $f_2(\Phi)$  :

$$f_2(\Phi) = \text{cond}(\Phi) + \frac{1}{\Sigma_p(\Phi)} = \frac{1 + \Sigma_1(\Phi)}{\Sigma_p(\Phi)} \quad (1.2.22)$$

La relation :

$$\text{cond}(\Phi) \geq \frac{\max_i \|\Phi_{:,i}\|}{\min_i \|\Phi_{:,i}\|} \quad (1.2.23)$$

autorise l'interprétation suivante du critère  $f_2(\Phi)$ . Un mouvement qui minimise  $f_2(\Phi)$  tend à équilibrer les contributions des paramètres *a priori*,  $\Phi_{:,i} = W_{:,i} \bar{X}_i$ , dans le vecteur des mesures *a priori*  $\bar{Y} = W \bar{X}$ . La norme  $\|W_{:,i}\|$  doit être grande pour rendre

visible un petit paramètre  $\bar{X}_i$  dans  $\bar{Y}$  et ainsi permettre son identification. Ce critère rejoint une interprétation physique d'un "mouvement riche en informations". Il faut que chacun des paramètres inertiels ait une contribution égale dans les couples et forces mesurées. Ceci amène à considérer plusieurs mouvements, chacun des mouvements permettant d'exciter un groupe de paramètres différent. C'est alors l'ensemble de ces *mouvements séquentiels excitants* qui fournira un système équilibré pour l'identification. Un intérêt majeur de cette stratégie sous optimale est de pouvoir analyser chacun des mouvements du point de vue physique et de contrôler ainsi le processus d'identification. Il faut bien noter ici que si les mouvements sont séquentiels, l'identification reste globale. Une identification séquentielle sur des mouvements séquentiels amène à des cumuls d'erreurs préjudiciable pour tous les paramètres identifiés.

#### 1.2.4 Application au compacteur

Dans le cas du compacteur, nous avons défini notamment les mouvements suivants après avoir bloqué le compacteur sur chandelle (FIG. 1.2) afin de supprimer toute interaction avec le sol.

- trajectoires à vitesse constante afin d'exciter les paramètres de frottement
- trajectoires à accélération importante afin d'exciter les inerties

L'ensemble de ces mouvements permet d'aboutir à un système d'identification global dont le critère d'excitation  $\text{cond}(\Phi)$  est satisfaisant (voir section 1.4)



FIG. 1.2 – Compacteur sur chandelles

### 1.3 Analyse d'un mouvement

Une fois les mouvements séquentiels excitants réalisés, il faut les analyser. L'objet de cette section est de fournir les outils permettant de synthétiser les informations

pertinentes pour chaque mouvement. Nous présentons deux outils importants : la dérivation numérique et le filtrage parallèle.

### 1.3.1 Dérivation

#### A. Différence centrée

Pour le calcul de la matrice d'observation  $W$ , il faut mesurer les positions, vitesses et accélérations intervenant dans les équations du modèle dynamique. Très peu de systèmes sont instrumentés de manière à mesurer l'ensemble de ces variables. Pour la majeure partie des articulations rotoïdes, c'est l'angle de rotation qui est mesuré, généralement avec un codeur incrémental. Dans ce cas, les informations de vitesse et accélération de rotation sont obtenues par dérivation numérique. La dérivation des signaux numériques de codeur doit se faire avec certaines précautions, ([Gau97]). On choisit l'approximation de la dérivée continue  $\dot{q}(t_k)$  par la différence centrée pour sa simplicité et l'absence de distorsion de phase.

$$\hat{q}(t_k) = \frac{q(t_{k+1}) - q(t_{k-1}))}{2T} \quad (1.3.1)$$

D'un point de vue temporel, cette approximation implique une erreur déterministe dans le calcul de la dérivée car ce n'est qu'une approximation de la dérivée continue  $\dot{q}(t_k)$ . Si on suppose que la position est continûment dérivable 2 fois ( $C_2$ ), alors on peut écrire le développement de TAYLOR :

$$q(t_{k+1}) = q(t_k) + T\dot{q}(t_k) + \frac{T^2}{2}\ddot{q}(t_k + \alpha T) \quad (1.3.2)$$

avec  $0 \leq \alpha \leq 1$  On obtient la majoration de l'erreur  $\tilde{q}(t_k) = \dot{q}(t_k) - \hat{q}(t_k)$  suivante :

$$\|\tilde{q}(t_k)\| \leq \frac{1}{2}T \max_{t \in [t_{k-1}, t_{k+1}]} \|\ddot{q}(t)\| \quad (1.3.3)$$

On en déduit que l'estimation de la dérivée d'un signal par la différence centrée tend vers la dérivée continue quand  $T$  tend vers 0.

D'un point de vue fréquentiel, la distorsion d'amplitude du filtre (1.3.1) par rapport à la dérivée continue est caractérisé par une bande passante à  $f_e/4$  (FIG. 1.3). De ce point de vue, la limite fréquentielle de l'estimation de l'accélération est donc  $f_e/8$ . Or pratiquement, la limite fréquentielle est beaucoup plus faible. Ceci est dû au fait que la position est souvent fournie par un codeur, et donc d'après l'équation (1.3.1), le bruit sur la position est multiplié par environ  $f_e^2$ . Pratiquement, le bruit sur l'accélération est 100 millions de fois plus important que le bruit sur la position lorsque l'échantillonnage est de 10 kHz. Pour réduire ce bruit, nous proposons un filtrage passe bas décrit dans le paragraphe suivant.

#### B. Filtrage passe bas

Une mesure est constituée du signal et du bruit de mesure. D'un point de vue fréquentiel, le signal est concentré dans une bande de fréquence qui dépend de la physique du système alors que le bruit de mesure couvre généralement toutes les fréquences. Dans nos applications, les informations pertinentes sont basses fréquences. Or lorsque nous dérivons, nous multiplions par un gain très important la partie haute

fréquence des mesures (ce qui est illustré par la figure (FIG. 1.3)), c'est-à-dire que nous privilégions le bruit au dépend du signal utile. C'est pourquoi, les signaux sont filtrés par un filtre passe bas pour atténuer les bruits hautes fréquences des mesures avant d'être dérivés numériquement. Le choix de la fréquence de coupure est fait à partir d'une analyse fréquentielle des mesures. Du point de vue de l'identification, si l'on choisit une bande de fréquences trop large, l'écart-type sur le bruit de mesure  $\sigma_\rho$  va être grand et si l'on choisit une bande de fréquences trop étroite, les critères d'excitation ne seront pas bons.

D'un point de vue pratique, le filtre de Butterworth d'ordre  $n_{but}$  est choisi pour sa réponse plate dans la bande passante. Compte tenu de l'identification hors ligne, un filtrage non causal aller et retour (procédure `filtfilt` de Matlab) est utilisé. Le choix de ce type de filtrage garanti l'absence de distorsion de phase des signaux numériques filtrés. Pour obtenir une atténuation suffisante du filtre sur les dérivées dans les hautes fréquences,  $n_{but} = n_{der} + 2$ , où  $n_{der}$  est l'ordre de la dérivée la plus élevée. Généralement  $n_{der} = 2$ , ce qui correspond au calcul d'une accélération à partir d'une position, donc  $n_{but} = 4$ . Il existe une méthode de calcul de la fréquence de coupure du filtre passe bas utilisé en fonction de la distorsion à une fréquence donnée du filtre passe bande dérivateur-filtre passe bas ([Pha02]). Cette méthode sert à initialiser les paramètres des filtres passe bande pour des dérivées d'ordre élevé. Ces réglages sont ensuite affinés avec une approche par une suite d'essais pour obtenir des réglages satisfaisant du point de vue des indicateurs statistiques de la qualité de l'identification. La figure FIG. 1.3 résume le principe de la dérivation et du filtrage des signaux numériques, où il faut que :

- la fréquence d'échantillonnage soit suffisamment grande pour que l'estimation de la dérivée soit bonne dans la bande de fréquences étudiée,
- l'ordre du filtre passe bas soit suffisamment grand pour que le filtre passe bande ait une bonne atténuation dans les hautes fréquences,
- la fréquence de coupure du filtre passe bas doit correspondre à la dynamique du système.

### C. Application au compacteur

Dans l'exemple de la bille du compacteur, le stator du moteur hydraulique est fixe par rapport au repère de l'atelier. Dans ce cas, on a la relation suivante :

$$\dot{\omega}_j^z = \ddot{q}_j \quad (1.3.4)$$

Il est alors possible de construire la matrice d'observation  $W$  à partir de  $q_j$ ,  $\dot{q}_j$  et  $\ddot{q}_j$ .  $q_j$  est mesuré à l'aide d'un codeur incrémental au niveau du moteur hydraulique.  $\dot{q}_j$  et  $\ddot{q}_j$  sont estimés à partir de  $q_j$  par dérivation numérique et filtrage. Les paramètres de réglage du filtre passe-bas de Butterworth sont les suivants :

- $n_{but} = 4$ , car il y a une double dérivation numérique,
- $F_c = 1$  Hz, car la dynamique des essais réalisées se situe entre 0 et 1 Hz.

#### 1.3.2 Filtrage parallèle

L'analyse fréquentielle effectuée pour le calcul des dérivées numérique peut être appliqué au modèle d'identification dans son ensemble. Il faut déterminer la bande fréquentielle dans laquelle le modèle du système mécanique est pertinent. Pour cela,

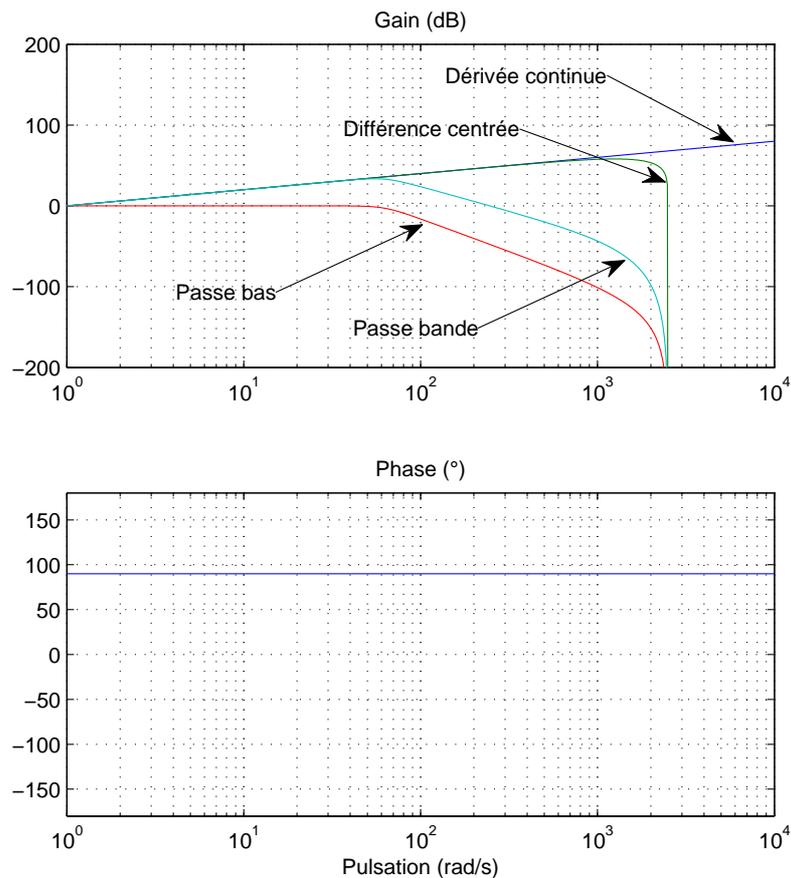


FIG. 1.3 – Estimation de la dérivée d'un signal

une analyse fréquentielle des signaux est réalisée. Ensuite, pour concentrer l'identification dans la bande de fréquences déterminée, nous effectuons un filtrage parallèle passe bas de  $Y$  et de  $W$ . La distorsion de ce procédé de filtrage parallèle de  $Y$  et de chaque colonne de  $W$  n'affecte pas la solution du système linéaire (1.1.4) ([Ric91]). C'est pourquoi, nous choisissons un filtre de Tchebychev pour ses propriétés de coupure. Avec le filtrage parallèle de  $Y$  et de  $W$ , nous obtenons un système dont la représentation temporelle est redondante par rapport à la représentation fréquentielle : une ligne du système peut être calculée en interpolant d'autres lignes. Nous devons donc décimer notre système et sélectionner une ligne sur  $F_{ech}/2F_c$  où  $F_{ech}$  est la fréquence d'échantillonnage. Nous obtenons alors le système filtré et décimé suivant que nous suffixons par  $d$  pour décimé :

$$Y^d = W^d X \quad (1.3.5)$$

#### A. Application au compacteur

Dans le cas de la bille du compacteur, l'acquisition des signaux a été faite à la fréquence de 10kHz pour des raisons de traitement du signal spécifique ([LVG03]). Il est donc nécessaire de filtrer et décimer les colonnes de la matrice d'observation  $W$  ainsi que le vecteur des mesures  $Y$ . La fréquence de coupure du filtre passe-bas est

réglée à 1Hz (pour ne retenir que la bande de fréquence où se situe la dynamique de la bille), il a donc fallu retenir une ligne sur 5000. Sur une acquisition de 30 secondes, la taille de la matrice d'observation initialement de 300000 lignes et 5 colonnes s'est réduite à une matrice  $60 \times 5$ .

## 1.4 Résolution

Lorsque tous les mouvements planifiés ont été réalisés, nous obtenons un ensemble de systèmes d'équations. Chaque système correspond à une expérimentation. Tous ces systèmes doivent être assemblés de façon cohérente. C'est pourquoi, ils sont pondérés avant d'être assemblés : il s'agit de pondérer chaque essai par la confiance que nous lui attribuons. Pour mettre en œuvre cette pondération, nous utilisons une décomposition QR du système (1.1.4). L'assemblage harmonieux des équations issues de la phase expérimentale est décrit dans la deuxième partie de cette section. L'étape suivante est la résolution de ces équations. Le point important de ce processus est l'estimation de l'incertitude sur les paramètres identifiés. Cette estimation est un critère important de qualité de l'identification. L'étape finale est la validation des résultats obtenus.

### 1.4.1 Réduction sous forme QR

Une transformation QR du système précédent (équation 1.3.5) donne :

$$W^d = Q^d \cdot R^d \quad (1.4.1)$$

avec  $Q$  matrice orthonormale et  $R$  matrice triangulaire supérieure.

Il suffit dans la suite de ne garder que  $R^d$ ,  $Q^{dT} Y^d$ ,  $(\rho^d)^2$  et le nombre d'équations du système décimé que nous noterons  $m^d$ . Ce dernier paramètre permettant de calculer la variance du système. Pour continuer l'exemple précédent, nous avons après filtrage parallèle une matrice d'observation comprenant 60 lignes et 5 colonnes. Après la réduction sous forme QR, nous retenons une matrice triangulaire  $5 \times 5$ . Nous voyons ici l'intérêt du filtrage, des moindres carrés et de l'utilisation du QR pour synthétiser l'information : d'une matrice d'observation initiale  $300000 \times 5$  et un vecteur de mesure de couple de dimension 300000, nous aboutissons à une matrice  $5 \times 5$ , un vecteur  $5 \times 1$ , un terme d'erreur au carré  $\rho^2$  et un nombre d'équations  $m^d$  beaucoup plus facile à stocker, manipuler, interpréter.

### 1.4.2 Assemblage des matrices

Lorsque les essais ont été réalisés et que les traitements précédents ont été effectués sur chaque mouvement, nous obtenons  $n$  systèmes  $\{R_k, Q_k^T Y_k, \rho_k^2, m_k\}$ ,  $k$  allant de 1 à  $n$  (nous avons retiré le suffixe  $d$  pour des raisons de présentation). Ces systèmes sont de nature très différentes, sur certains les bruits de mesure et de modèle sont faibles sur d'autres, ces bruits sont élevés. Il faut donc les normaliser. Nous avons choisi de normaliser chaque système par sa variance. Plus la variance est élevée, moins le système sera pris en compte dans le processus d'identification. Ceci a une interprétation probabiliste, chaque système a un bruit de variance différent,

il s'agit de normaliser ces bruits afin d'obtenir un bruit uniforme sur tous les systèmes.

$$R_k^{pondere} = \frac{R^k}{\sigma_k} \quad (1.4.2)$$

$$Q_k^T Y_k^{pondere} = \frac{Q_k^T Y_k}{\sigma_k} \quad (1.4.3)$$

avec :

$$\sigma_k^2 = \frac{\rho_k^2}{m_k - p} \quad (1.4.4)$$

Ces matrices sont alors concaténées pour obtenir le système final.

$$R^{final} = \begin{bmatrix} R_1^{pondere} \\ \vdots \\ R_n^{pondere} \end{bmatrix} \quad (1.4.5)$$

$$Y^{final} = \begin{bmatrix} Q^T Y_1^{pondere} \\ \vdots \\ Q^T Y_n^{pondere} \end{bmatrix} \quad (1.4.6)$$

$$\rho^{2final} = \sum_{k=1}^{k=n} \rho_k^2 \quad (1.4.7)$$

$$m^{final} = \sum_{k=1}^{k=n} m_k \quad (1.4.8)$$

Ce système subit lui-même une transformation QR :

$$R^{final} = Q \cdot R \quad (1.4.9)$$

### 1.4.3 Identification des paramètres

Les paramètres sont identifiés par moindres carrés sur ce nouveau système (équation 1.4.9). Nous calculons un vecteur de paramètres en résolvant le système carré obtenu.

$$\hat{X} = R^{-1} \cdot Q^T Y^{final} \quad (1.4.10)$$

la variance est estimée en calculant :

$$\hat{\sigma}^2 = \frac{\rho^{2final} + Y^{finalT} Y^{final} - (Q^T Y^{final})^T Q^T Y^{final}}{m^{final} - p} \quad (1.4.11)$$

### 1.4.4 Paramètres essentiels

Si les valeurs des paramètres sont trop déséquilibrées, il est très difficile de diminuer le critère  $f_2(\Phi)$ . Il est alors préférable de simplifier le modèle en éliminant les paramètres  $X_k$  tels que :

$$\sigma_{\hat{X}_{r_k}} = \frac{\sigma_{\hat{X}_k}}{|\hat{X}_k|} > 0, 1 \quad (1.4.12)$$

L'interprétation de ce critère est double. D'une part, seuls les paramètres dont la valeur identifiée a au moins un chiffre significatif sont conservés. D'autre part, la contribution

relative des paramètres éliminés dans le calcul des efforts articulaires est faible sur le mouvement optimisé. Les paramètres retenus sont appelés paramètres essentiels. Ils définissent un modèle dynamique simplifié et précis. Il reste un détail technique à prendre en compte : retirer des paramètres identifiable change le  $\sigma_{rho}$  de chaque essai, il faut alors refaire l'étape de pondération et d'assemblage des essais avant d'identifier le nouveau jeu de paramètres essentiel.

Le fait de qualifier un paramètre comme non essentiel n'est pas anodin, il faut s'assurer qu'il n'existe pas de mouvement qui puisse exciter ce paramètre. C'est l'étape de replanification qui utilise les concepts expliqués dans la section 1.2 afin de vérifier que ce type de mouvement excitant est impossible compte tenu des limitations mécaniques et des capacités d'actionnement.

### A. Application au compacteur

Les différents essais ont été réalisés, les données ont été filtrées et assemblées sous la forme d'une matrice d'observation  $W$  et d'un vecteur de mesure  $Y$  eux aussi filtrés et décimés. Il est alors possible d'appliquer les équations des moindres carrés pondérés. Cette application mène aux résultats exprimés dans (TAB. 1.1). Les pa-

Paramètres	Unités	$\hat{X}$	$\sigma_{\hat{X}}$	$\sigma_{\hat{X}}(\%)$
$ZZ_j$	kg.m <sup>2</sup>	214	0,26	0,12
$MX_j$	m.kg	0,82	0,54	65
$MY_j$	m.kg	0,50	0,52	105
$FV_j$	N.m.s	72,4	0,33	0,45
$FS_j$	N.m	55,7	1,8	3,2
$\sigma_\rho = 41,6$ N.m		$\text{cond}(W) = 75$	$\text{cond}(\phi) = 571$	

TAB. 1.1 – Résultats de l'identification des paramètres du modèle de la bille

ramètres  $MX_j$  et  $MY_j$  sont mal identifiés (leur  $\sigma_{\hat{X}}(\%)$  est largement supérieur à 10%), de plus les valeurs identifiés ( $\hat{X}$ ) sont faibles. Cette faible influence de ces paramètres se retrouve dans le conditionnement important ( $\text{cond}(\phi) = 571$ ) ce qui indique qu'au moins un des paramètres n'a pas d'influence dans la prédiction du vecteur de mesure. Nous avons réalisé des mouvements qui augmentent l'excitation de ces paramètres grâce au système de vibration du compacteur. Malgré ces nouveaux mouvements, ces paramètres restent mal identifiés. C'est pourquoi une nouvelle identification a été effectuée en enlevant les paramètres  $MX_j$  et  $MY_j$  du modèle, les résultats obtenus sont donnés dans (TAB. 1.2). Dans ce cas, tous les paramètres sont bien identifiés et

Paramètres	Unités	$\hat{X}$	$\sigma_{\hat{X}}$	$\sigma_{\hat{X}}(\%)$
$ZZ_j$	kg.m <sup>2</sup>	214	0,28	0,13
$FV_j$	N.m.s	71,8	0,34	0,47
$FS_j$	N.m	56,5	1,8	3,2
$\sigma_\rho = 42,4$ N.m		$\text{cond}(W) = 49$	$\text{cond}(\phi) = 52$	

TAB. 1.2 – Résultats de l'identification des paramètres du modèle de la bille

contribuent au modèle. Le conditionnement a fortement baissé.

## 1.4.5 Validation

### A. Validation des capacités prédictives du modèle identifié

Une fois les paramètres identifiés, nous vérifions les capacités prédictives de notre modèle en s'assurant que le couple prédit correspond au couple mesuré dans les conditions suivantes.

- validation directe : comparaison de  $Y$  et de  $W\hat{X}$ , calculés sur les mêmes mouvements que ceux utilisés pour l'identification, visualisation du signal d'erreur  $\rho$ . Cette première étape nécessaire afin de vérifier nos hypothèses statistiques sur le bruit  $\rho$  doit être complétée par une validation croisée,
- validation croisée : la comparaison est effectuée sur des mouvements non utilisés pour l'identification.

**Application au compacteur** Pour valider le modèle, il est possible de comparer le couple mesuré au niveau du moteur hydraulique et le couple calculé à partir du modèle. La figure (FIG. 1.4) donne un exemple de cette comparaison qui est une validation supplémentaire de la qualité de l'identification.

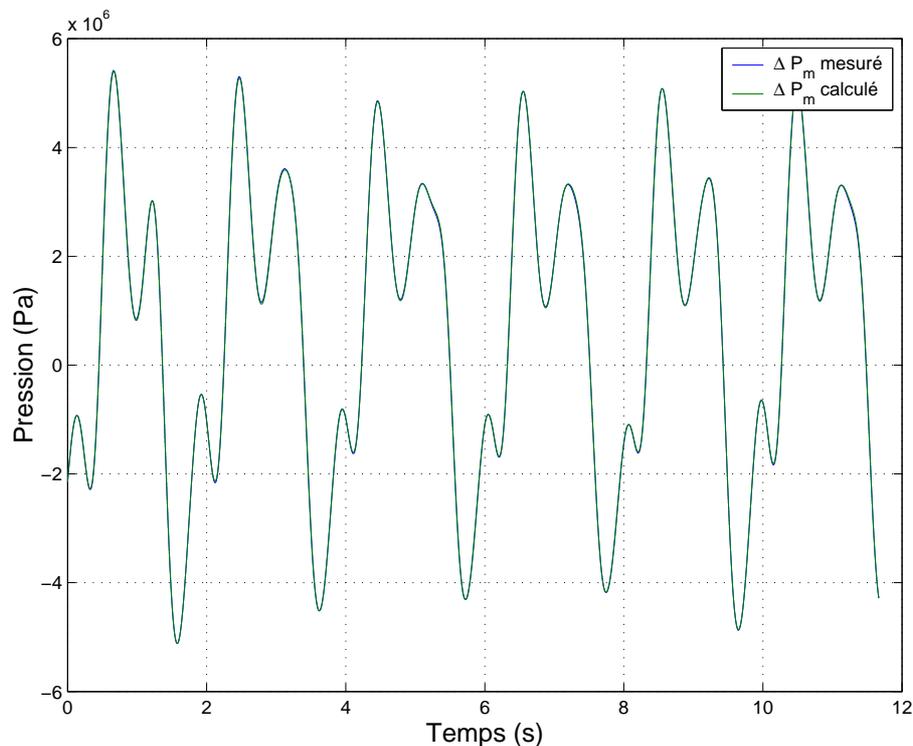


FIG. 1.4 – Comparaison du couple mesuré et du couple calculé

### B. Validation par comparaison avec des données connues

Les données identifiées avec les valeurs a priori sont comparés aux données techniques. Dans le cas du compacteur où il existe deux niveaux de vibrations, nous avons bien vérifié que les paramètres liés au système de vibration, c'est à dire les premiers

moments du balourd, variaient de manière cohérente suivant les données constructeurs. Un autre type de validation consiste à ajouter une masse connue, à refaire une identification, et à vérifier que les différences trouvées dans les valeurs des paramètres correspondent aux caractéristiques de la masse ajoutée. C'est ainsi, que nous avons estimé la masse du conducteur du compacteur.

## 1.5 Conclusion

Les nombreuses applications menées avec succès ont prouvé que l'identification par modèle inverse et moindres carrés est une approche intéressante pour les systèmes mécatroniques depuis les robots jusqu'aux engins de chantiers et en général pour tout système dont le modèle inverse s'exprime linéairement en fonction des paramètres physiques. Voici les étapes principales qui conduisent au succès de l'identification :

- exprimer le modèle inverse sous une forme linéaire par rapport aux paramètres physiques,
- calculer les paramètres identifiables,
- utiliser des mouvements riches en information, qui excitent bien les paramètres,
- concevoir un système de mesure adapté aux objectifs de l'identification,
- filtrer les données expérimentales pour obtenir un système linéaire surdéterminé peu perturbé,
- éliminer les paramètres non significatifs pour n'identifier que les paramètres essentiels,
- utiliser les formes factorisées (QR, SVD) pour le calcul numérique des moindres carrés.

En suivant cette méthodologie, il n'a pas été observé de biais significatif sur les estimations des paramètres.

# Bibliographie

- [Gau90] M. Gautier. *Contribution à la modélisation et à l'identification des robots*. Thèse de doctorat, Université de Nantes, École Nationale Supérieure de la Mécanique, 1990.
- [Gau91] M. Gautier. Numerical calculation of the base inertial parameters. *Journal of Robotic Systems*, 8(4) :485–506, 1991.
- [Gau97] M. Gautier. Dynamic identification of robots with power model. In *Proceedings of IEEE International Conference on Robotics and Automation*, Albuquerque, USA, 1997.
- [GK90] M. Gautier and W. Khalil. Direct calculation of the base inertial parameters. *IEEE Transactions on Robotics and Automation*, 6(3) :368–373, 1990.
- [GvL96] G. H. Golub and C. F. van Loan. *Matrix computations*. John Hopkins University Press, Baltimore, 1996.
- [KC97] Wisama Khalil and Denis Creusot. Symoro+ : A system for the symbolic modelling of robots. *Robotica*, 15 :153–161, 1997.
- [Kha99] E. Khalil, W. and Dombre. *Modélisation, identification et commande des robots*. Hermès Science Publications, Paris, second edition, 1999.
- [Lem05] C.-E. Lemaire. *Estimation des efforts de contact cylindre/matériau d'un compacteur vibrant*. Thèse de doctorat, Université de Nantes, 2005.
- [LVG03] Charles-Éric Lemaire, Pierre-Olivier Vandanjon, and Maxime Gautier. Dynamic identification of a compactor using splines data processing. In *Proceedings of System Identification*, Rotterdam, Netherlands, 2003.
- [PG93] C. Pressé and M. Gautier. New criteria of exciting trajectories for robot identification. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 907–912, 1993.
- [Pha02] M.T. Pham. *Contribution à la modélisation, l'identification et la commande de systèmes mécaniques à flexibilités localisées : application à des axes de machines-outils rapides*. PhD thesis, Université de Nantes, 2002.
- [Pre94] C. Pressé. *Identification des paramètres dynamiques des robots*. Thèse de doctorat, Université de Nantes - Ecole Centrale de Nantes, 1994.
- [Res96] P. Restrepo. *Contribution à la modélisation, identification et commande des robots à structures fermées : application au SR400*. Thèse de doctorat, Université de Nantes, 1996.
- [Ric91] J. Richalet. *Pratique de l'identification*. Traité des Nouvelles Technologies. Hermès, Paris, 1991.

- [Van95] P.-O. Vandanjon. *Identification robuste des paramètres inertiels des bras manipulateur par découplage fréquentiel*. Thèse de doctorat, Ecole des Mines de Paris, 1995.
- [Ven03] Gentiane Venture. *Identification des paramètres dynamiques d'une voiture*. PhD thesis, Université de Nantes, 2003.

## Chapitre 2

# Algorithmes récursifs pour l'identification modale des structures de génie civil

*Christian Crémona*

### 2.1 Introduction

La réalisation d'essais vibratoires ou d'impact est pratiquée depuis très long temps, il est vrai d'une manière intuitive et très pratique. Les charpentiers de marine, par exemple, ont toujours tenu compte des impacts répétitifs du vent et des vagues dans la conception des vaisseaux en bois. Les archers concevaient leurs flèches de sorte qu'elles puissent résister au choc contre un os ou un tendon. Les essais vibratoires n'ont pourtant été formalisés que pendant et après la seconde guerre mondiale. La construction rapide d'avions a été un élément moteur du développement de ces essais dont l'objectif était d'appréhender des phénomènes particuliers conduisant à certaines défaillances sous sollicitations répétées et rapides. La réalisation d'essais dynamiques sur les ponts n'est pas récente et date même du XIXe siècle. La plupart de ces premières tentatives faisaient parties des procédures d'inspection des ponts ferroviaires et servaient déjà de base à la surveillance du comportement vibratoire de ces ouvrages. Les essais modernes ne sont que des prolongements plus sophistiqués de ces toutes premières investigations à des fins d'amélioration des méthodes d'analyse, des règles de conception des ouvrages et de leur suivi. Diverses raisons motivent la réalisation d'essais dynamiques sur des ponts. Les mesures vibratoires sur ouvrages permettent en effet [Cré05] :

- d'enrichir la connaissance sur le comportement dynamique de structures de même type
- de vérifier les choix de la modélisation d'une structure
- d'évaluer l'intégrité d'un ouvrage
- de surveiller l'état d'un ouvrage par des mesures systématiques ou continues de sa réponse dynamique.

L'un des objectifs d'une évaluation dynamique est la détermination - ou **identification** - des caractéristiques modales de la structure. Cela inclut les fréquences propres, les coefficients d'amortissement et les déformées propres. Afin de les identifier, il est

nécessaire d'exciter la structure afin de produire une réponse sur chaque mode pertinent. Le recours à un type ou un autre d'excitation (impulsionnelle, harmonique, ambiante) n'est pas neutre et chaque catégorie présente des avantages et des inconvénients, et peut ne pas être adaptée à un ouvrage donné. La sélection d'une excitation pour des essais vibratoires est donc un problème essentiel suivant l'analyse que l'on poursuit. Le recours à des essais sous excitation ambiante est aujourd'hui de plus en plus populaire parce que ce choix est plus proche des préoccupations des ingénieurs soucieux d'appréhender le comportement des ouvrages dans leur fonctionnement courant. L'objectif de ce chapitre est de présenter un algorithme d'identification modale fonctionnant sous excitation ambiante. Les techniques d'identification sous excitation ambiante, c'est-à-dire ne recourant pas à la connaissance et à la mesure de l'excitation, sont aujourd'hui couramment utilisées ; citons par exemple les méthodes de réalisation stochastique, de sous-espace ou du décrément aléatoire [CDFI05]. La démarche présentée ici consiste à proposer une autre forme de méthode d'identification permettant de suivre au cours du temps l'évolution des caractéristiques modales.

## 2.2 Identification des systèmes

Un problème important en dynamique des structures réside dans la détermination précise des paramètres qui caractérisent les modes principaux d'une structure. Ces paramètres - fréquences propres, décréments logarithmiques, modes propres - sont alors des inconnues : leur connaissance permet la validation et l'amélioration du modèle de comportement dynamique de la structure, d'analyser une partie de cette dernière qui peut être difficile de modéliser analytiquement, ou d'évaluer l'intégrité de la structure par comparaison avec un modèle numérique ou des mesures antérieures [Cré05]. De ce fait, l'**identification modale** signifie l'estimation de ces caractéristiques à partir des mesures des signaux d'excitation et de réponse. Une telle détermination rend possible la prédiction du comportement de la structure dans toutes les situations. Ceci est particulièrement le cas des structures à comportement linéaire puisqu'il est possible de déduire la totalité des réponses pour tout type d'excitation à partir de la connaissance de la réponse à des excitations particulières synthétisées dans la matrice de transfert du modèle. L'idée sous-jacente de l'identification modale relève donc plus de la détermination d'un modèle équivalent de la structure caractérisant le mieux possible son comportement dynamique que la seule estimation des paramètres modaux. Cette définition implique que l'identification modale n'est qu'un cas spécial d'un problème plus général : l'**identification des systèmes**.

### 2.2.1 Objectifs

Les sciences visent à formuler des théories que le scientifique propose comme concept de la nature qu'il étudie. Guidé par cette idée, il conçoit de nouvelles expériences pour confirmer ces théories, dicter des modifications ou les rejeter complètement. D'égale importance est la notion de modèle. Rosenblueth et Wiener [RW45] définissent un modèle comme *la représentation des aspects essentiels d'un système existant ou à construire présentant la connaissance de ce système sous une forme utilisable*. Suivant une telle définition, un système est interprété comme *une collection d'objets arrangés de façon ordonnée avec sens, raison et objectif précis* [Eyk74]. Toute chose n'appartenant pas au système fait partie de son **environnement**. Suivant la frontière

qu'il dessine avec son environnement, le système est caractérisé par des **relations d'entrée/sortie** (figure 2.1).

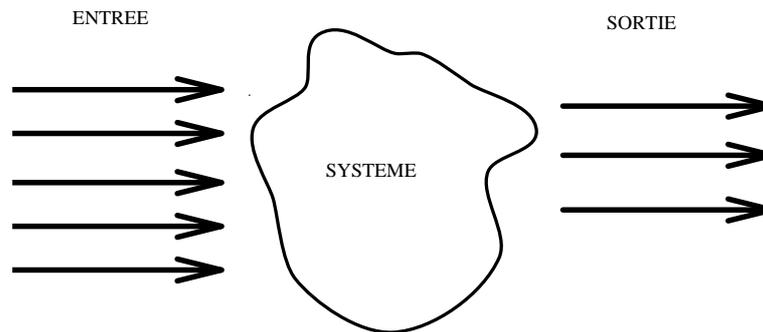


FIG. 2.1 – Schéma général d'un système

La construction d'un modèle commence avec l'application des lois de la physique (c'est-à-dire des théories). A partir de ces lois, un certain nombre de relations entre les variables se déduisent. Si toutes les conditions internes et externes au système sont connues quantitativement et si la connaissance du système est complète, alors au moins en principe, les valeurs numériques de tous les coefficients de ces relations peuvent être déterminées. Une telle situation est en fait très peu réaliste. Une telle connaissance *a priori* est difficile à cause des incertitudes liées à l'environnement et à une connaissance imparfaite du phénomène étudié. Même si la connaissance de parties élémentaires du phénomène était parfaite, le démon de la complexité nous forcerait à recourir à des simplifications. Ceci implique que les paramètres et l'état d'un système sont parfois difficilement accessibles à des calculs *a priori*. Le problème discuté dans ce chapitre est souvent dénoté **problème inverse** en analyse des systèmes : un historique des entrées/sorties entre un système et son environnement étant donné, il faut déterminer les équations qui décrivent le comportement du système. Zadeh [Zad62] définit l'identification comme *la détermination, sur la base d'entrées et de sorties, d'un système pris dans une classe de systèmes pour lequel le système testé est équivalent*. Suivant cette définition, il est nécessaire de spécifier une **classe de systèmes**  $\mathcal{D}_M$ , une **classe de signaux d'entrée**  $\mathcal{U}$  et le sens d'**équivalent**. Les éléments de  $\mathcal{D}_M$  sont appelés **modèles**. L'équivalence est souvent définie en terme de **critère**  $\mathcal{V}$ , une erreur ou une fonction de vraisemblance, qui est fonction de la sortie  $y$  du système réel et de la sortie  $\hat{y}$  d'un modèle de  $\mathcal{D}_M$  :

$$\mathcal{V} = \mathcal{V}(y, \hat{y}) \quad (2.2.1)$$

Deux modèles  $M_1$  et  $M_2$  seront équivalents si les valeurs des critères sont identiques pour ces deux modèles. De telles définitions laissent une grande liberté dans la construction de méthodes d'identification. Ceci se reflète notamment dans l'abondance des techniques rencontrées dans la littérature. La sélection de la classe de modèles, de la famille des signaux d'entrée et du critère d'équivalence est influencée par la connaissance *a priori* sur le phénomène aussi bien que par l'objectif de l'identification. Le caractère du signal d'entrée peut affecter considérablement l'identification, et doit être en conséquence choisi avec attention. Il y a essentiellement deux conditions à cela. La première est que le domaine  $\mathcal{D}_M$  soit le plus vaste possible afin que le système étudié y appartienne. Le second point concerne la classe des signaux d'entrée qui doit être suffisamment générale afin qu'aucun autre modèle ne soit équivalent à la réponse

du système. Le concept d'**identificabilité** [GLS77] est en directe relation avec ces problèmes. L'idée de base est qu'un paramètre est identifiable si il peut être déterminé de manière unique à partir des données ; un ensemble de modèles est alors identifiable si tous les paramètres de ces modèles le sont. Une absence d'identificabilité peut être causée par des modèles surparamétrés ou des signaux d'entrée non appropriés : ces deux problèmes se rencontrent couramment dans l'analyse expérimentale des structures, car dans un premier cas, l'expérimentateur est face à une structure continue (ce qui nécessite souvent de surparamétrer les modèles), et dans le second cas, à des sources d'excitation fluctuante en spectre et en énergie.

## 2.2.2 Identifications directes et récursives

L'identification implique qu'un ensemble de données (entrées/sorties) soit collecté et que, en conséquence, cet ensemble de données sert à construire un modèle équivalent. Une telle procédure est appelée **identification directe**. Cependant le modèle peut être utilisé pour prendre des décisions "en ligne", c'est-à-dire durant l'acquisition des données. Il est alors nécessaire que l'identification du modèle infère dans l'acquisition des données. Le modèle est alors actualisé à chaque instant où de nouvelles mesures deviennent disponibles. Cette mise à jour ne peut être réalisée que par des **méthodes récursives temporelles** dont les noms changent suivant le domaine d'application :

- méthodes récursives d'identification en théorie du contrôle,
- méthodes adaptatives d'identification en traitement du signal,
- estimation séquentielle de paramètres en statistiques.

Il y a cependant deux désavantages majeurs lorsqu'on utilise des méthodes récursives d'identification par rapport à des méthodes directes. La première implique d'imposer des conditions initiales aux algorithmes récursifs et la seconde concerne la question de la convergence et de la stabilité des schémas récursifs introduits.

## 2.3 Application à l'identification modale

### 2.3.1 Hypothèses

Avant toute considération de modèles, il est nécessaire de préciser trois hypothèses de base. Ces hypothèses sont indépendantes de la technique numérique d'identification, mais visent à fixer les propriétés physiques des structures ou systèmes qui ont été étudiés. La structure a un **comportement linéaire**, impliquant l'existence de fonctions de transfert. La seconde hypothèse concerne l'**invariance temporelle des paramètres physiques** (paramètres modaux).

Cette caractéristique mérite quelques commentaires : l'invariance temporelle est en général inadéquate pour certaines applications industrielles ou mécaniques. Les valeurs des paramètres peuvent varier, et dans de tels cas, la procédure d'identification peut être vue comme une procédure de poursuite de ces paramètres. Pour cette raison, les méthodes récursives d'identification sont particulièrement bien adaptées. Un exemple pratique est la surveillance vibratoire de câbles de haubanage dont le régime vibratoire peut changer au cours du temps (accrochage par échappement tourbillonnaire ou instabilité aéroélastique) qu'il est parfois intéressant et utile de détecter. Cette hypothèse d'invariance a été admise dans nos études initiales, mais peut être relâchée

au moyen de paramètres appropriés que l'on appelle facteur d'oubli. La dernière hypothèse stipule une **quantité suffisante d'informations** disponibles durant l'identification. Ceci implique que tous les modes de liberté pertinents peuvent être mesurés. Cette hypothèse est à relier au problème d'**observabilité** et est conditionnée par les propriétés des signaux d'excitations.

### 2.3.2 De l'identification des systèmes à l'identification modale

Deux objectifs peuvent expliquer le recours à l'identification modale : l'amélioration de modèles numériques ou le suivi dans le temps du comportement de structures. Le premier objectif consiste à relier modèle et données afin de construire un meilleur modèle. Ceci est possible si les mesures et la technique d'identification sont fiables. Le second objectif dépend de la finalité de l'identification. Pour la surveillance vibratoire de câbles de haubanage par exemple, cela signifie un contrôle en temps réel à cause des modifications du régime vibratoire imposé par le vent : il s'agit alors d'estimations sur des périodes courtes. De manière opposée, on peut recourir à des périodes d'estimation plus longues pour évaluer des modifications dans les propriétés physiques d'une structure à cause de dommages. L'étape suivante consiste alors à localiser les zones de défauts, et dans ce cas, l'estimation des paramètres modaux peut aider au diagnostic. De façon générale, l'identification modale apparaît comme une étape dans une procédure plus complexe et le choix d'une méthode d'identification devra dépendre de la raison de cette procédure. Il y a cependant qu'aujourd'hui un besoin très fort existe pour tester des structures en service, et notamment des grands ouvrages de génie civil. Ceci relève de l'objectif de mieux comprendre et expliquer le comportement réel de la structure. Pour cela, les essais doivent être menés in situ et les techniciens doivent travailler avec des forces externes parfois inconnues (vent, trafic,...). L'utilisation de techniques d'identification est récente en génie civil, mais de nombreux résultats encourageants tendent à prouver que leur utilisation est pertinente, procurant des tests non destructifs. Une grande majorité de techniques classiques d'identification modale ont cependant été développées pour des essais en laboratoire, et elles présentent des caractéristiques qui empêchent parfois de les utiliser sur des structures en service :

- traitement simultané de toutes les données (stockage important),
- manque d'adaptabilité vis à vis de modifications de paramètres en cours d'identification, ou de filtrage de perturbations numériques et/ou de données,
- absence de bases mathématiques et théoriques fiables.

Le premier problème est une conséquence du caractère explicite des méthodes. La catégorie prédominante des algorithmes sont des algorithmes directs. L'acquisition des données est réalisée en premier, puis l'identification est effectuée. Ceci requiert l'invariance temporelle sur la durée de traitement, et des tailles de mémoire parfois considérables. Les conditions d'essais doivent donc être parfaitement contrôlées.

Les techniques classiques d'identification cherchent à déterminer les matrices de transfert du système, ce qui oblige à travailler en régime fréquentiel. Des techniques de minimisation de type moindres carrés sont alors utilisées pour estimer les paramètres des fonctions de transfert. Les réponses impulsionnelles sont ensuite estimées par déconvolution afin de disposer d'un modèle temporel à caractère prédictif. Cependant, les procédures de déconvolution sont très sensibles aux bruits de mesure et conduisent à des oscillations numériques. De plus les signaux ne sont pas connus de façon continue et pour des structures en service, les signaux d'entrée sont parfois totalement inconnus et les paramètres modaux peuvent également subir

des modifications. Pour ces diverses raisons, le recours à des techniques d'identification en régime temporel ont été recherchées. Les techniques classiques d'identification en régime temporel sont des techniques utilisant des réponses impulsionnelles. Ceci simplifie le développement numérique des routines d'identification, mais reste également très sensible aux bruits de mesure. Le recours à des mesures de réponse en régime forcé est plus délicat de mise en oeuvre, mais présente l'avantage d'assurer une excitation permanente et d'estimer plus correctement l'ensemble des modes pertinents [Cré05]. Dans le cadre d'identification de structures en service, cela est particulièrement intéressant. Par exemple, l'estimation des paramètres modaux d'un pont peuvent être estimés en utilisant l'excitation produite par le trafic ou le vent. Pour des essais en laboratoire, le contrôle des entrées et le choix des classes de ces dernières peuvent être correctement organisés. Pour une structure en service, l'environnement n'est pas forcément connu, ou même mesurable. Il est donc nécessaire de disposer de techniques se limitant à la seule connaissance et mesure de la réponse. Cette élimination de l'excitation peut se faire en supposant que l'excitation est la sortie d'un filtre stable [CB92]. Enfin développées par des ingénieurs mécaniciens pour des ingénieurs mécaniciens, les techniques d'identification font complètement abstraction des informations statistiques. Il n'y a pas de procédures de filtrage du bruit par exemple. Parce que les techniques d'identification modale sont aujourd'hui appliquées à des structures de plus en plus complexes, il est devenu nécessaire de disposer d'algorithmes d'identification fiables, robustes et présentant une adaptation aux problèmes rencontrés lors de l'identification (changement de paramètres, erreurs de mesures,...). Durant ces dix dernières années, les utilisateurs ont commencé à étudier les capacités des algorithmes d'identification temporelle et ont été amenés à étudier ce qui était fait depuis de nombreuses années en théorie du contrôle ou en traitement du signal. Ceci a apporté un nouvel éclairage à l'identification modale. La méthode présentée dans ce chapitre se place dans cette démarche en incluant également une vision récursive de l'identification afin de permettre le suivi de paramètres modaux importants en phase de fonctionnement d'une structure ou d'un ouvrage. L'objectif de ce chapitre est ainsi de présenter une écriture innovante et récursive des équations de Yule-Walker très souvent utilisées dans l'identification de séries chronologiques temporelles. L'intérêt de cette ré-écriture de ces équations réside non seulement dans son caractère récursif, mais aussi dans la vérification immédiate de sa convergence vers l'estimation off-line. Pour cela, il est cependant nécessaire de préciser le modèle de travail retenu.

## 2.4 Modèles de travail

Puisque les conditions des expérimentations nécessitent l'échantillonnage des signaux de réponse - ou signaux de sortie- et des signaux d'excitation -ou signaux d'entrée-, il est essentiel de disposer d'un modèle qui reflète l'historique de cet échantillonnage. Un choix raisonnable consiste à modéliser le système dynamique par un modèle de type ARMA (AutoRegressive Moving Average). Un modèle ARMA d'ordre  $p - q$  exprime la réponse  $y_k$  (vecteur de dimension  $n_y$ ) à l'instant  $t_k$  au moyen d'une expression récurrente

$$y_k + \sum_{i=1}^p a_i y_{k-i} = \sum_{j=0}^q B_j f_{k-j} \quad (2.4.1)$$

où les  $a_i$  sont scalaires et les  $B_j$  scalaires, vectoriels ou matriciels.

La théorie des systèmes linéaires permet formellement de démontrer l'existence d'un tel modèle. Dans [Cré90], une démonstration différente est faite pour exhiber des relations entre les fréquences propres, les décrets et les modes propres du système et les paramètres  $a_i$  et  $B_j$  du modèle ARMA. L'approche retenue a été de considérer un système (ou une structure) avec amortissement non proportionnel. Le recours à l'introduction de modes propres complexes et à l'utilisation d'une décomposition de Jordan (dans le cas où les modes propres complexes ne constitueraient pas une base de l'espace initial) permet d'écrire de telles relations. Il convient de souligner que le traitement mathématique ne se fait ni dans le domaine temporel, ni dans le domaine fréquentiel ou de Laplace, mais dans le domaine en  $Z$ . Par le théorème de Cauchy, il est ramené sous forme discrétisée dans le domaine temporel. Le modèle ARMA est alors exhibé, permettant de montrer que les ordres  $p$  et  $q$  sont respectivement  $2n$  et  $2n - 1$ ,  $n$  étant le nombre de modes à identifier, en d'autres termes le nombre de modes présents dans la réponse de la structure. Il est également possible de démontrer que le système est strictement causal, c'est-à-dire  $B_0 = 0$ . Il vient en particulier que les fréquences et les décrets sont obtenus directement des racines du polynôme  $a(z) = \sum_{i=0}^{2n} a_i z^{2n-i}$ . En effet, en notant  $\lambda_i$  la  $i$ -ème racine de  $a(z)$ , il vient :

$$f_i = \frac{1}{2\pi\Delta t} \sqrt{\frac{1}{4} \log^2 (\Re(\lambda_i)^2 + \Im(\lambda_i)^2) + \arctan^2 \left( \frac{\Im(\lambda_i)}{\Re(\lambda_i)} \right)} \quad (2.4.2)$$

$$\zeta_i = \frac{-\frac{1}{2\Delta t} \log (\Re(\lambda_i)^2 + \Im(\lambda_i)^2)}{f_i} \quad (2.4.3)$$

$\Delta t$  est l'intervalle d'échantillonnage. Le modèle ARMA (2.4.1) se limite à décrire la réponse temporelle échantillonnée de la structure par un système récurrent d'équations.

## 2.5 Expression récursive des équations de Yule-Walker

### 2.5.1 Hypothèses

L'une des premières hypothèses est de supposer que l'excitation est un bruit blanc. C'est la définition même d'un modèle ARMA. Cette hypothèse peut cependant être relâchée en supposant que l'excitation est la sortie d'un filtre stable. Cremona [Cré90] a démontré que, pour une telle excitation - appelée bruit coloré-, il était possible de montrer que la partie autoregressive  $\hat{a}(z)$  se décompose en le produit de  $a(z)$  par un polynôme dont les pôles seront des modes fantômes qu'il faudra distinguer des pôles de  $a(z)$ . Il va sans dire que cette tâche est difficile et repose dans la plupart du temps sur des considérations physiques. En se plaçant dans l'hypothèse d'un bruit blanc, il vient que l'état passé du système n'a aucune influence sur l'excitation présente :  $E[{}^t f_k y_l] = 0$ , pour  $k > l$ , où  $E[\cdot]$  représente l'espérance mathématique. L'espérance de toute combinaison linéaire des  ${}^t f_{k-j} y_l$ , pour  $0 \leq j \leq 2n - 1$  est nulle pour  $k - 2n + 1 > l$ . En utilisant le modèle 2.4.1, il vient :

$$E \left[ {}^t \left( y_k + \sum_{i=1}^p a_i y_{k-i} \right) y_{k-2n+1-l} \right] = 0 \quad (2.5.1)$$

pour  $k - 2n + 1 > k - 2n + 1 - l$ . En posant

$$Z_k = (y_{k-2n+1-1}, \dots, y_{k-2n+1-2n}) \quad (2.5.2)$$

$$Y_k = (y_{k-1}, \dots, y_{k-2n}) \quad (2.5.3)$$

$$\theta = (a_1, \dots, a_{2n}) \quad (2.5.4)$$

la relation 2.5.1 s'écrit :

$$E [{}^t y_k Z_k + (\theta {}^t Y_k) Z_k] = 0 \quad (2.5.5)$$

La seconde des hypothèses consiste à remplacer l'espérance mathématique par son estimateur empirique, ce qui suppose l'ergodicité des processus. Dans ce cas, l'équation 2.5.5 devient :

$$E [{}^t y_k Z_k + (\theta {}^t Y_k) Z_k] = \sum_{i=4n}^k ({}^t y_i Z_i + (\theta {}^t Y_i) Z_i) = 0 \quad (2.5.6)$$

ce qui conduit à l'équation matricielle :

$$\theta \sum_{i=4n}^k {}^t Y_i Z_i = - \sum_{i=4n}^k {}^t y_i Z_i \quad (2.5.7)$$

Par simplification d'écriture, on pose :

$$R_k = \sum_{i=4n}^k {}^t Y_i Z_i \quad (2.5.8)$$

$$T_k = - \sum_{i=4n}^k {}^t y_i Z_i \quad (2.5.9)$$

ce qui donne :

$$\theta R_k = T_k \quad (2.5.10)$$

Cette équation permet de constater qu'il ne s'agit pas d'une solution au sens des moindres carrés classiques, en raison du fait que  ${}^t Y_i Z_i$  introduit des séries de mesures décalées. Elle est à rapprocher d'une démarche basée sur le principe de variables instrumentales qui a été introduite pour contourner le caractère biaisé de l'estimateur des moindres carrés. Par voie de conséquence, l'équation précédente permet de limiter le biais que pourrait induire des mesures bruitées. La troisième hypothèse consiste à supposer que les paramètres sont indépendants du temps. Cependant, pour tenir compte de dérives de ces paramètres, un facteur d'oubli  $\alpha$  compris entre 0 et 1 est introduit dans les équations 2.5.11 et 2.5.12 :

$$R_k = \frac{1}{k - 4n} \sum_{i=4n}^k \alpha^{k-i} {}^t Y_i Z_i \quad (2.5.11)$$

$$T_k = - \frac{1}{k - 4n} \sum_{i=4n}^k \alpha^{k-i} {}^t y_i Z_i \quad (2.5.12)$$

Ce facteur d'oubli permet de donner un poids moindre aux anciennes mesures par rapport aux nouvelles.

## 2.5.2 Première écriture récursive

Soit  $\theta_k$  l'estimation de  $\theta$  pour les  $k$  premières mesures :

$$\theta_k R_k = T_k \quad (2.5.13)$$

On constate que l'on a :

$$R_k = {}^t Y_k Z_k + \alpha R_{k-1} \quad (2.5.14)$$

$$T_k = -{}^t y_k Z_k + \alpha T_{k-1} \quad (2.5.15)$$

L'équation 2.5.13 permet de tirer en conjonction avec les équations précédentes :

$$T_k = -{}^t y_k Z_k + \theta_{k-1} (R_k - {}^t Y_k Z_k) = \theta_k R_k \quad (2.5.16)$$

et donc :

$$\theta_k = \theta_{k-1} + (-{}^t y_k - \theta_{k-1} {}^t Y_k) Z_k R_k^{-1} \quad (2.5.17)$$

Cette dernière équation montre comment l'estimateur de  $\theta$  à l'instant  $k$  est corrigé à partir de celui à l'instant  $k - 1$  et des nouvelles mesures. La matrice  $R_k$  est appelée **matrice d'information**.

## 2.5.3 Calcul de $R_k^{-1}$

L'équation 2.5.17 met en évidence la nécessité de disposer de l'inverse de la matrice  $R_k$  pour actualiser le vecteur de paramètres  $\theta_{k-1}$ . Elle oblige donc à inverser une matrice de dimension  $2n \times 2n$  à chaque itération, ce qui peut s'avérer prohibitif. Pour éviter cette inversion systématique, il est préférable de remplacer l'expression récursive 2.5.14 qui porte sur  $R_k$ , par une autre expression récursive portant sur  $R_k^{-1}$ . Le théorème d'inversion matricielle due à Kron permet d'écrire :

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1} \quad (2.5.18)$$

sous réserve que les matrices concernées soient de dimensions appropriées. En posant  $A = \alpha R_{k-1}$ ,  $B = {}^t Y_k$ ,  $C = I$ ,  $D = Z_k$ , il vient :

$$\begin{aligned} (\alpha R_{k-1} + {}^t Y_k Z_k)^{-1} &= \frac{1}{\alpha} P_{k-1} \\ &- \frac{1}{\alpha} P_{k-1} {}^t Y_k \left( Z_k \frac{1}{\alpha} P_{k-1} {}^t Y_k + I \right)^{-1} Z_k \frac{1}{\alpha} P_{k-1} \end{aligned} \quad (2.5.19)$$

Cette équation permet donc d'écrire :

$$P_k = \frac{1}{\alpha} P_{k-1} (I - {}^t Y_k s_k^{-1} Z_k P_{k-1}) \quad (2.5.20)$$

avec  $s_k = Z_k P_{k-1} {}^t Y_k + \alpha I$ . Ainsi, il n'est plus utile d'inverser la matrice d'information qui est de dimension  $2n \times 2n$ , mais d'inverser à chaque itération la matrice  $s_k$  qui est de dimension  $n_y \times n_y$ , où  $n_y$  est le nombre de points de voies de mesures (ou de signaux de sortie). En général, ce nombre de voies de mesures, pour un ouvrage en service, est bien inférieur au nombre de modes présents dans la réponse d'une structure continue.

### 2.5.4 Remarque sur le facteur d'oubli

Le facteur d'oubli a pour objectif de réduire l'influence du passé sur les estimations présentes. Ceci permet de pouvoir traquer des évolutions lentes ou faiblement rapides du paramètre  $\theta$ , et donc des caractéristiques modales que sont les fréquences et les coefficients d'amortissement. Ceci conduit généralement à faire varier le paramètre  $\alpha$  lui-même pour tenter de suivre au mieux ces évolutions de paramètre. Pour cela, une suite du type suivant est généralement considérée [Cré90] :

$$\alpha_k = \alpha^0 \alpha_{k-1} + (1 - \alpha^0) \quad (2.5.21)$$

Cremona [Cré90] montre que, pour une valeur initiale de ce schéma de 0,95, la valeur  $\alpha^0 = 0,99$  procure des résultats optimisés dans le cas de sauts de paramètres.

### 2.5.5 Remarque les conditions initiales

Le choix des conditions initiales est un facteur important de la méthode. Concernant la matrice d'information, cette dernière peut être choisie comme nulle, en absence de données initiales, ce qui implique que son inverse doit avoir une norme de grande valeur. Ceci conduit à retenir par exemple  $P_{4n} = 10^8 I$ ,  $I$  étant la matrice identité de dimension  $2n \times 2n$ . Pour le paramètre  $\theta_{4n}$ , la valeur initiale est un vecteur de composantes nulles. Rappelons que, par construction, la forme récursive des équations de Yule-Walker oblige de débiter le traitement après avoir engrangé au moins  $4n$  acquisitions temporelles sur l'ensemble des voies de mesures.

### 2.5.6 Convergence

Le schéma récursif proposé s'écrit donc :

$$\theta_k P_k^{-1} = \theta_{k-1} P_{k-1}^{-1} + (-{}^t y_k - \theta_{k-1} {}^t Y_k) Z_k \quad (2.5.22)$$

$$P_k^{-1} = {}^t Y_k Z_k + \alpha_k P_{k-1}^{-1} \quad (2.5.23)$$

L'équation 2.5.23 conduit à exprimer  $P_k$  en fonction de sa valeur initiale  $P_{4n}$  :

$$P_k^{-1} = \mu_k P_{4n}^{-1} + \sum_{j=4n}^k \rho_{jk} {}^t Y_j Z_j \quad (2.5.24)$$

avec

$$\mu_k = \prod_{i=4n}^k \alpha_i \quad (2.5.25)$$

$$\rho_{jk} = \prod_{i=j}^{k-1} \alpha_i \quad j < k \quad (2.5.26)$$

$$\rho_{kk} = 1 \quad j = k$$

Introduisons le *vrai* vecteur de paramètres  $\theta$ , noté  $\theta^0$  ; il vient d'après 2.5.1 :

$${}^t y_k + \sum_{i=1}^p a_i {}^t y_{k-i} = {}^t N_k = {}^t y_k + \theta^0 {}^t Y_k \quad (2.5.27)$$

$N_k$  représente le processus de la partie à moyenne mobile du processus ARMA et les éventuelles erreurs de l'estimation liées par exemple au bruit de mesures. En utilisant cette équation, on tire :

$$\begin{aligned} {}^t y_k Z_k + \theta_{k-1} {}^t Y_k Z_k &= -\theta^0 {}^t Y_k Z_k + {}^t N_k Z_k + \theta_{k-1} {}^t Y_k Z_k \\ &= (\theta_{k-1} - \theta^0) {}^t Y_k Z_k + {}^t N_k Z_k \end{aligned} \quad (2.5.28)$$

Il est alors possible d'écrire :

$$\begin{aligned} \theta_k P_k^{-1} &= \alpha_k \theta_{k-1} P_{k-1}^{-1} + \theta_{k-1} {}^t Y_k Z_k \\ &\quad + (\theta^0 - \theta_{k-1}) {}^t Y_k Z_k - {}^t N_k Z_k \\ &= \mu_k \theta_{4n} P_{4n}^{-1} + \theta_0 \left( \sum_{j=4n}^k \rho_{jk} {}^t Y_j Z_j \right) + \left( \sum_{j=4n}^k \rho_{jk} {}^t N_j Z_j \right) \end{aligned} \quad (2.5.29)$$

$$P_k^{-1} = \mu_k P_{4n}^{-1} + \sum_{j=4n}^k \rho_{jk} {}^t Y_j Z_j \quad (2.5.30)$$

Plusieurs remarques peuvent être faites concernant certains termes des équations 2.5.29 et 2.5.30 :

$$\begin{aligned} \theta_{4n} &= 0 \\ P_{4n}^{-1} &\approx 0 \\ \frac{1}{k-4n} \sum_{j=4n}^k {}^t N_j Z_j &\propto E[{}^t N_j Z_j] = 0 \end{aligned} \quad (2.5.31)$$

La dernière des propriétés provient de l'équation 2.5.1. Ceci implique notamment :

$$\frac{1}{k-4n} \sum_{j=4n}^k \rho_{jk} {}^t N_j Z_j \propto 0 \quad (2.5.32)$$

Au moyen de ces propriétés remarquables ou asymptotiques, l'estimateur  $\theta_k$  est donc asymptotiquement équivalent à :

$$\theta_k \approx \theta_0 \left( \sum_{j=4n}^k \rho_{jk} {}^t Y_j Z_j \right) \left( \sum_{j=4n}^k \rho_{jk} {}^t Y_j Z_j \right)^{-1} \quad (2.5.33)$$

et donc pour  $k \rightarrow \infty$ ,  $\theta_k \rightarrow \theta^0$ .

## 2.6 Application

Cette section présente des applications pratiques de l'algorithme précédent pour la détermination des caractéristiques modales de câbles. Le premier exemple concerne une câble de pont à hauban soumis à l'action du vent. La réponse de ce câble est donnée par la figure 2.2.

L'analyse spectrale et l'identification récursive temporelle sont présentées à la figure 2.3. Cette figure illustre une concordance de très bonne qualité entre les fréquences dominantes identifiées par le spectre de puissance et celles estimées par la méthode

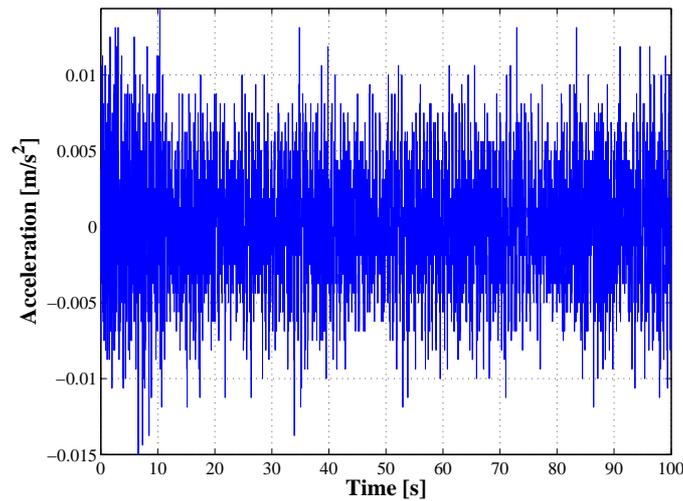


FIG. 2.2 – Réponse vibratoire d'un câble long du pont de Saint-Nazaire

récursive présentée. Notons cependant que la technique temporelle requiert de dimensionner le modèle ARMA en fixant a priori le nombre de modes (ici, 20) et d'imposer des règles de rejet (coefficients d'amortissement négatifs et inférieurs à 3%).

La figure 2.3 met également en évidence que certaines fréquences sont moins stables que d'autres. Par exemple, les fréquences à 2,7Hz et 4,5Hz font apparaître deux lignes d'identification. Pour chaque fréquence, ces deux lignes correspondent à deux pôles différents du polynôme  $a(z)$ . Une bonne estimation impliquerait d'avoir ces deux pôles conjugués, comme pour les autres fréquences, et disposer ainsi de deux lignes fréquentielles confondues.

La figure 2.4 fournit un autre exemple de la réponse du même câble, mais pour des conditions de mesures différentes. Deux fréquences proches à 1,7Hz et 1,8Hz apparaissent dans le spectre. L'algorithme récursif n'est capable d'identifier que la fréquence à 1,8Hz. Cette fréquence correspond à la fréquence propre du câble et est associée à un coefficient d'amortissement inférieur à 3%, alors que l'autre fréquence est éliminée en raison d'un coefficient d'amortissement supérieur à ce seuil et donc peu réaliste (le seuil de 3% est d'ailleurs un seuil élevé, car les coefficients d'amortissements de câbles sont en générale beaucoup plus faibles).

La figure 2.5 présente une autre application à la mesure vibratoire d'une suspente de pont bowstring. Le spectre met en évidence des fréquences inférieures à 70Hz. L'algorithme récursif n'est pas à même de détecter la fréquence de 43Hz (Fig. 2.6). L'une des raisons possible est que cette fréquence apparaît rapidement au cours du temps et que l'algorithme n'est pas capable de la suivre. La figure 2.7 présente les estimations des coefficients d'amortissement correspondants. Une bonne convergence peut être notée, mettant en évidence l'efficacité de la technique.

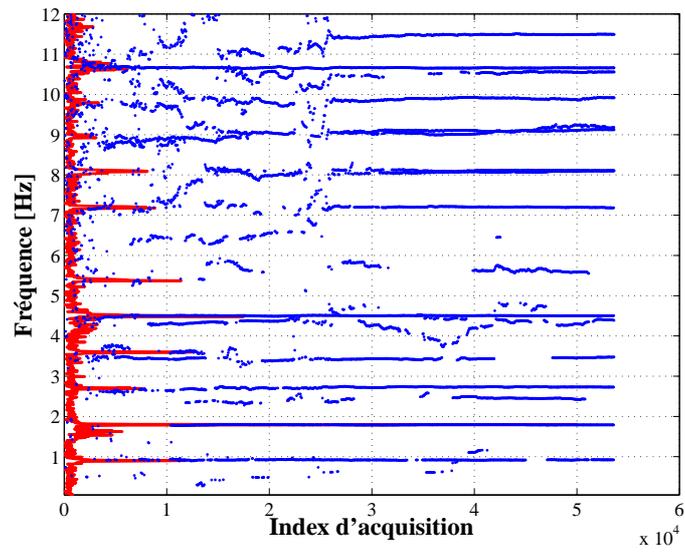


FIG. 2.3 – Comparaison entre spectre and fréquences identifiées par technique récursive

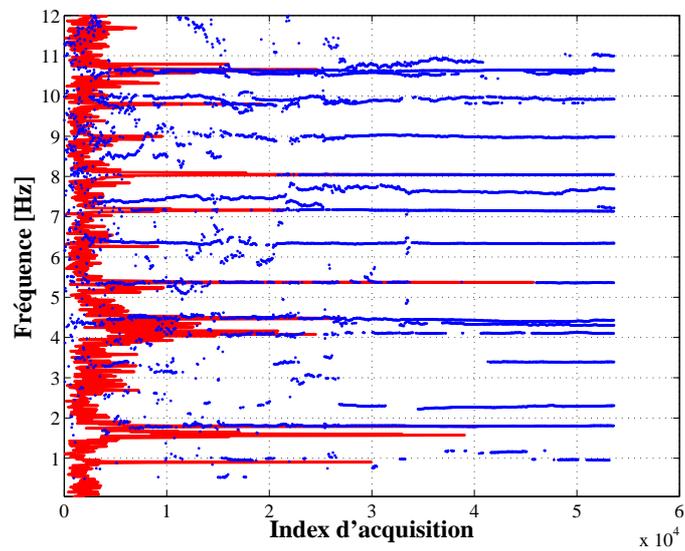


FIG. 2.4 – Comparaison entre spectre and fréquences identifiées par technique récursive

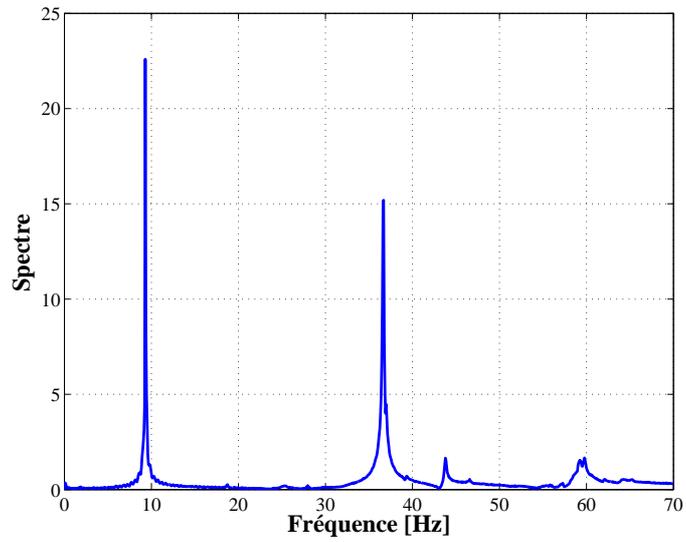


FIG. 2.5 – Spectre d'une suspenste de bowstring sous excitation ambiante

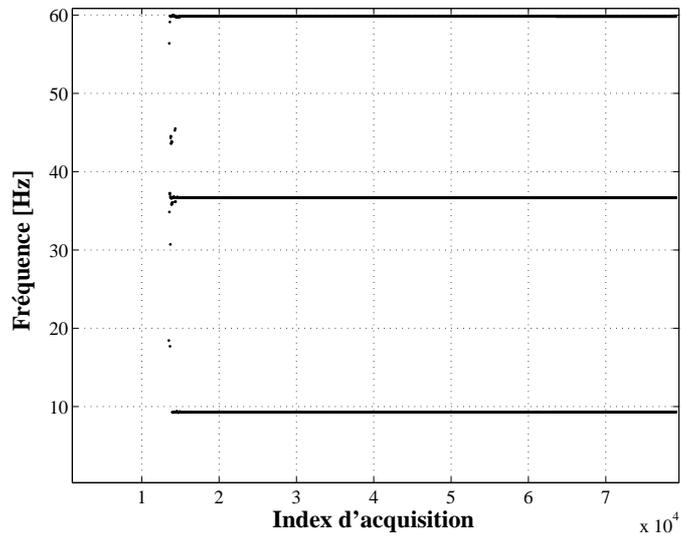


FIG. 2.6 – Fréquences identifiées par technique récursive

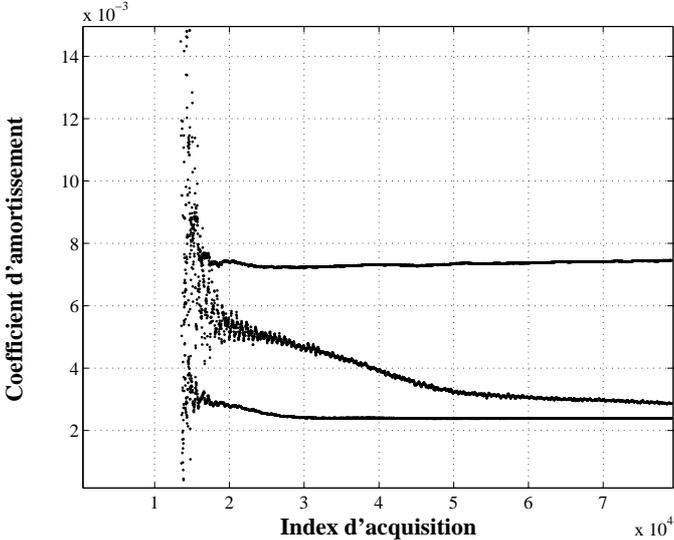


FIG. 2.7 – Coefficients d'amortissement identifiés par technique récursive



# Bibliographie

- [CB92] C. Crémona and J.A. Brandon. A modal identification algorithm with unmeasured input. *ASCE Journal of Aerospace Engineering*, 5(4) :442–449, 1992.
- [CDFI05] C. Crémona, L. Dieleman, M. Fogl, and MH. Inchauspé. Evaluation dynamique des ouvrages. *Revue Européenne de Génie Civil*, 5(1-2), 2005.
- [Cré90] C. Crémona. *Recursive time domain modal identification techniques*. PhD thesis, University of Wales College of Cardiff, 1990.
- [Cré05] C. Crémona. Evaluation dynamique des ouvrages. *Etudes et Recherches des Laboratoires des Ponts et Chaussées*, (47), 2005.
- [Eyk74] P. Eykhoff. *System identification : Parameter and state estimation*. Wiley, Bristol, 1974.
- [GLS77] I. Gustavsson, L. Ljung, and T. Soderstrom. Identification of processes in closed-loop : Identifiability and accuracy aspects. *Automatica*, 13 :59–75, 1977.
- [RW45] A. Rosenblueth and N. Wiener. The role of models in science. *Philosophy of Science*, 12 :316–321, 1945.
- [Zad62] L.A. Zadeh. From circuit theory to system theory. In *Proceedings IRE*, volume 50, pages 856–865, 1962.



## Chapitre 3

# Analyse en ondelettes pour l'identification de paramètres en dynamique

*Pierre Argoul,  
Safia Bedaoui,  
Silvano Erlicher,  
Thien Phu Le,  
Honoré Yin.*

### 3.1 Introduction

La transformation en ondelettes est une décomposition temps-échelle qui s'est largement développée depuis les travaux de Jean Morlet et d'Alex Grossman [GM84] qui en donnent un cadre mathématique rigoureux. Depuis, elle n'a cessé d'être exploitée avec efficacité et succès pour des applications variées, comme par exemple, le traitement de signaux non stationnaires ou la compression d'images.

La transformation en ondelettes continue (TOC) permet par exemple de capter la variation temporelle des fréquences et des amplitudes instantanées contenues dans les réponses transitoires ou de localiser les singularités contenues dans les réponses dynamiques. Ainsi, en donnant une meilleure structuration de l'information, elle facilite grandement le processus d'identification des paramètres modaux à partir des réponses vibratoires de structures mécaniques. Son emploi à des fins d'identification modale voit le jour au milieu des années 90.

Le but de ce chapitre est de montrer les multiples avantages du traitement des réponses vibratoires des structures par la TOC. Le premier paragraphe redonne les expressions des fonctions de réponses en fréquence, réponses impulsionnelles et réponses libres dans le cas de systèmes mécaniques discrets à comportement linéaire, stationnaire et causal. Le paragraphe suivant traite de la transformation en ondelettes en rappelant sa définition et les principales propriétés utiles pour l'identification. L'« optimisation » des caractéristiques de l'ondelette est réalisée grâce à l'introduction du facteur de qualité permettant de gérer les localisations temporelle et fréquentielle de la TOC. Le troisième paragraphe présente plus précisément comment utiliser l'analyse en ondelettes à des fins d'identification modale. Plusieurs uti-

lisations de la TOC pour l'identification modale sont proposées et décrites suivant le type de réponses vibratoires analysées : fonction de réponse en fréquence ou réponse à un choc (réponse libre) ou encore réponse sous excitations ambiantes. Des applications sur des cas réels en laboratoire (vibrations de poutres et d'une plaque) sont présentées. Le dernier paragraphe qui sert de conclusion, ouvre des perspectives sur l'emploi de l'analyse en ondelettes à des fins d'identification dynamique pour des structures à comportement non linéaire. Une technique pour détecter et caractériser les non-linéarités présentes dans le comportement est proposée, fondée sur le calcul d'indicateurs instantanés.

### 3.2 Identification de paramètres en dynamique

La problématique de l'identification de paramètres en dynamique des structures, qualifiée également de « vibrations inverses » par [Bui93], peut se résumer de la façon suivante : comment, à partir d'une expérience vibratoire sur la structure étudiée, connaître les caractéristiques du modèle qui représente son comportement mécanique ?

Les techniques d'identification ont d'abord été appliquées sur des structures à comportement linéaire. Dans ce cas, le comportement peut être représenté soit par des fonctions de réponse en fréquence (FRF) dans le domaine fréquentiel, soit par des réponses impulsionnelles (RI) dans le domaine temporel dont nous allons redonner les expressions. Pour des systèmes discrétisés à  $N$  degrés de liberté (DdL) avec un comportement supposé linéaire et stationnaire, les oscillations sont gérées par des équations différentielles linéaires, du second ordre, à coefficients constants. On note  $l$  le DdL du point d'application de la force d'excitation et  $k$  celui du point de mesure de la réponse en déplacement. La FRF  $H_{kl}$  de type réceptance (rapport force sur déplacement) s'écrit :

$$H_{kl}(\omega) = \sum_{r=1}^N \left[ \frac{(A_r)_{kl}}{j\omega - p_r} + \frac{(\overline{A}_r)_{kl}}{j\omega - \overline{p}_r} \right] \quad (3.2.1)$$

où  $(\overline{A}_r)_{kl}$  et  $\overline{p}_r$  sont respectivement le conjugué du résidu  $(A_r)_{kl}$  et celui du  $r^e$  pôle complexe de la fonction de transfert ;  $j = \sqrt{-1}$ . Lorsqu'on suppose un amortissement de type visqueux sous-amorti, les pôles  $p_r$  solutions de :

$$p_r^2 - 2\sigma_r p_r + \omega_{0_r}^2 = 0, \quad (3.2.2)$$

s'écrivent :  $p_r = \sigma_r + j\omega_r$  avec  $\omega_r$   $r^e$  pulsation propre amortie et  $\sigma_r < 0$  ;  $\omega_r$  et  $\sigma_r$  sont liés à la  $r^e$  pulsation propre du système conservatif associé  $\omega_{0_r}$  et au  $r^e$  taux d'amortissement visqueux  $\xi_r$  par :  $\omega_r = \omega_{0_r} \sqrt{1 - \xi_r^2}$  et  $\sigma_r = -\xi_r \omega_{0_r}$ .

De la même façon, la RI notée  $h_{kl}(t)$  s'exprime par :

$$h_{kl}(t) = \sum_{r=1}^N (B_r)_{kl} e^{\sigma_r t} \cos(\omega_r t + (\chi_r)_{kl}) \quad (3.2.3)$$

où  $(B_r)_{kl} = \frac{1}{\omega_r} \sqrt{(\alpha_r)_{kl}^2 \omega_r^2 + ((\beta_r)_{kl} + \sigma_r (\alpha_r)_{kl})^2}$  et  $(\chi_r)_{kl} = -\arctan \frac{(\beta_r)_{kl} + \sigma_r (\alpha_r)_{kl}}{(\alpha_r)_{kl} \omega_r}$

avec  $(\alpha_r)_{kl} = (A_r)_{kl} + (\overline{A}_r)_{kl}$  et  $(\beta_r)_{kl} = -(A_r)_{kl} \overline{p}_r - (\overline{A}_r)_{kl} p_r$ .

Nous terminons ces rappels par les expressions des réponses libres au DdL  $k$  en déplacement, vitesse et accélération notées respectivement :  $u_k(t)$ ,  $\dot{u}_k(t)$  et  $\ddot{u}_k(t)$ . Le

principe de superposition modale permet d'écrire ces réponses libres réelles comme la superposition de  $N$  composantes modales respectivement  $u_{kr}(t)$ ,  $\dot{u}_{kr}(t)$  et  $\ddot{u}_{kr}(t)$  :

$$u_k(t) = \sum_{r=1}^N u_{kr}(t) = \sum_{r=1}^N \operatorname{Re} (G_r \phi_{kr} e^{p_r t}) \quad (3.2.4)$$

$$\dot{u}_k(t) = \sum_{r=1}^N \dot{u}_{kr}(t) = \sum_{r=1}^N \operatorname{Re} (G_r p_r \phi_{kr} e^{p_r t}) \quad (3.2.5)$$

$$\ddot{u}_k(t) = \sum_{r=1}^N \ddot{u}_{kr}(t) = \sum_{r=1}^N \operatorname{Re} (G_r p_r^2 \phi_{kr} e^{p_r t}) \quad (3.2.6)$$

où  $\operatorname{Re}(\cdot)$  désigne la partie réelle du complexe entre les parenthèses ; par exemple :  $\operatorname{Re}(G_r \phi_{kr} e^{p_r t}) = \frac{1}{2} (G_r \phi_{kr} e^{p_r t} + \bar{G}_r \bar{\phi}_{kr} e^{\bar{p}_r t})$  ;  $\phi_{kr}$  est la  $k^{\text{e}}$  composante du  $r^{\text{e}}$  vecteur propre complexe  $\phi_r$  solution de l'équation caractéristique :  $(p_r^2 \mathbf{M} + p_r \mathbf{C} + \mathbf{K}) \phi_r = \mathbf{0}$  où  $\mathbf{M}$  est la matrice de masse définie positive,  $\mathbf{K}$  et  $\mathbf{C}$  sont les matrices de rigidité et d'amortissement visqueux supposées définies non négatives.  $G_r$  est le facteur de participation complexe associé avec le  $r^{\text{e}}$  mode qui dépend des vecteurs conditions initiales en déplacement  $\mathbf{u}_0 = \mathbf{u}(0)$  et en vitesse  $\mathbf{v}_0 = \dot{\mathbf{u}}(0)$

$$G_r = |G_r| e^{j \arg G_r} = 2 \frac{\phi_r^T (p_r \mathbf{M} \mathbf{u}_0 + \mathbf{C} \mathbf{u}_0 + \mathbf{M} \mathbf{v}_0)}{\phi_r^T (2p_r \mathbf{M} + \mathbf{C}) \phi_r} \quad (3.2.7)$$

La composante modale  $u_{kr}(t)$  s'écrit sous la forme :  $\Upsilon^{(u_{kr})}(t) \cos(\alpha^{(u_{kr})}(t))$ . L'amplitude  $\Upsilon^{(u_{kr})}(t) = |G_r| |\phi_{kr}| e^{-\xi_r \omega_{0r} t}$  est une fonction positive et décroissante exponentiellement et la phase  $\alpha^{(u_{kr})}(t) = \arg G_r + \arg \phi_{kr} + \omega_r t$  est la somme de trois termes : une constante  $\arg G_j$  qui dépend des conditions initiales modales,  $\arg \phi_{kr} \in (-\pi, \pi]$  qui tient compte de la nature complexe de la forme modale et une fonction linéaire en temps dont le coefficient est la  $r^{\text{e}}$  pulsation propre amortie. Dans le cas de modes réels,  $\arg \phi_{kr}$  est égale soit à 0, soit à  $\pi$ , suivant le signe de  $\phi_{kr}$ . On note que le signal  $u_{kr}(t)$  est un « chirp » dont les variations de l'amplitude  $\Upsilon^{(u_{kr})}(t)$  sont lentes par rapport aux oscillations de la phase  $\alpha^{(u_{kr})}(t)$ . La notion de chirp est liée à celle d'asymptoticité qui sera développée au paragraphe 3.4.2 suivant.

Dans la suite, on omettra les indices  $k$  et  $l$  par souci de simplification d'écriture, lorsqu'ils ne sont pas nécessaires.

Les techniques d'identification modale visent à estimer les paramètres modaux d'une structure (fréquences propres, taux d'amortissement modaux et formes modales). Depuis les trente dernières années, elles n'ont cessé de se complexifier et le nombre des algorithmes développés est toujours croissant. Ces derniers sont généralement fondés sur les propriétés supposées pour le système dont la linéarité, l'invariance dans le temps, l'observabilité<sup>1</sup> et la réciprocity. La classification générale des méthodes d'identification modale devient actuellement une tâche difficile. Celle donnée par [MSH<sup>+</sup>98] est fondée sur quatre critères classés par les auteurs selon leur importance.

Le premier critère est celui du domaine dans lequel les données sont traitées numériquement (temporel ou fréquentiel). De façon générale, les méthodes dans le domaine temporel ont tendance à donner de meilleurs résultats lorsque le contenu fréquentiel des données s'étend sur une large plage et que le nombre de modes est élevé. À l'inverse, les méthodes fréquentielles obtiennent de meilleurs résultats si le contenu

<sup>1</sup>L'observabilité caractérise la possibilité de retrouver l'état d'un système en observant sa sortie.

fréquentiel est limité et si le nombre de modes est relativement faible. Cependant, les méthodes dans le domaine temporel ne peuvent estimer que les modes se trouvant dans la plage fréquentielle de l'analyse et ne peuvent pas prendre en compte les effets résiduels des modes situés en dehors de cette plage, à l'inverse des méthodes dans le domaine fréquentiel. Nous proposons de rajouter un troisième domaine : le domaine temps-fréquence pour prendre en compte les techniques utilisant les transformations temps-fréquence comme les analyses en ondelettes ou en gaborettes.

On attribue le second critère suivant que la procédure d'identification est appliquée soit à partir des équations de la dynamique (« méthode directe »), soit sur les paramètres modaux (« méthode indirecte »). Il est en rapport avec le concept de « boîte blanche » ou « noire » utilisé pour qualifier les modèles en théorie des systèmes. Le modèle en boîte blanche est associé au critère direct, fondé sur une connaissance physique du phénomène et des lois fondamentales du mouvement et le modèle en boîte noire, au critère indirect sans lien avec une connaissance physique du phénomène. Cette terminologie grâce à la notion de modèle en « boîte grise » permet d'inclure des modèles partiellement basés sur des connaissances physiques.

Le troisième critère concerne le nombre de modes analysés ; on parle de méthodes à un ou à plusieurs DdLs.

Le dernier critère est fondé sur les nombres des entrées et des sorties pouvant être pris en compte par la méthode ; on distingue les méthodes à « une entrée - une sortie » de celles à « une entrée - plusieurs sorties » ou encore de celles à « plusieurs entrées - plusieurs sorties ». L'idée qui soutient ce critère est que les paramètres modaux ne varient pas en théorie, lorsque l'on passe d'une FRF à une autre. Par conséquent, il devient possible d'obtenir un ensemble des caractéristiques modales, plus consistant, en traitant en même temps plusieurs FRFs ou plusieurs RIs.

### 3.3 La transformation en ondelettes

L'analyse temps-fréquence a toujours été un enjeu en traitement des signaux. Pour l'étude des signaux transitoires, qui évoluent dans le temps d'une façon imprévisible, la notion d'analyse fréquentielle ne peut être que locale en temps. Ces signaux ne peuvent pas être représentés par une superposition d'ondes (sinusoïdes de durée infinie) mais comme une superposition d'ondelettes (ondes de courte durée).

Parmi les transformations temps-fréquence les plus couramment rencontrées, on peut citer la transformation en ondelettes, la transformation en gaborettes (TG) et enfin celle de Wigner-Ville (TWV).

La TG utilise une fonction « fenêtre » pour localiser l'analyse de Fourier, puis la fait glisser sur une autre position, et ainsi de suite. Les « gaborettes » ou ondelettes de Gabor sont construites à partir de la fenêtre par des translations et des modulations (c'est-à-dire des translations en fréquence). Le compromis est qu'une fenêtre courte favorise la résolution en temps au détriment de la résolution en fréquence et vice-versa. La TWV est le représentant le plus connu de la classe de Cohen des distributions bilinéaires alors que les analyses en gaborettes et en ondelettes sont des représentations linéaires. Bien qu'elle puisse prendre des valeurs négatives, la TWV d'une fonction peut être considérée comme une densité dans le plan temps-fréquence à cause de ses propriétés d'intégrations. La structure bilinéaire de la représentation de Wigner Ville comporte de nombreux avantages (tel que la localisation temporelle et fréquentielle, conservation de l'énergie...). Cependant son principal inconvénient est

de générer des interférences issues des interactions entre composantes d'un signal ; ce qui rend la lisibilité de cette « densité » plus délicate. Ainsi, pour l'analyse des signaux certains et aléatoires, [Fla89] propose un estimateur de type pseudo-Wigner lissé afin d'améliorer la lisibilité.

Dans cette étude, nous nous sommes intéressés à la TOC à des fins d'identification modale. Dans cet article, nous ne comparons pas la TOC à d'autres transformations temps-fréquence comme par exemple la TG ou la TWV, déjà citées.

L'analyse en ondelettes permet une représentation du signal dans le temps (ou l'espace) et dans les échelles. Grossmann et Morlet démontrent notamment dans [GM84] que pour qu'un signal temporel puisse être décomposé sous forme d'une combinaison linéaire de fonctions élémentaires localisées en différents instants et ayant des échelles différentes, il faut que la fonction mère présente quelques oscillations et donc ressemble à une ondelette, une onde localisée, par comparaison au terme « onde » seul, qui signifie une oscillation qui se propage indéfiniment.

### 3.3.1 Définitions

La TOC d'un signal  $u(t)$ , d'énergie finie et continu par morceaux, est définie comme suit :

$$T_{\psi}[u](b, a) = \frac{1}{a} \int_{-\infty}^{+\infty} u(t) \bar{\psi}\left(\frac{t-b}{a}\right) dt \quad (3.3.1)$$

où  $\psi(t)$  est une fonction de classe  $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ , continue par morceaux, appelée ondelette mère, et  $\bar{\cdot}$  désigne le complexe conjugué. La variable  $b \in \mathbb{R}$  introduit une translation de l'ondelette mère le long de l'axe temporel. La variable  $a > 0$  détermine une dilatation de la variable temporelle  $t$  et peut être reliée à l'inverse d'une fréquence angulaire : nous prenons (cf. relation suivante (3.4.15)),  $a = \frac{\omega_{max}}{\omega}$  où  $\omega_{max}$  désigne la valeur de la pulsation telle que le module  $|\hat{\psi}(\omega)|$  de la transformée de Fourier de  $\psi(t)$  est maximal. On rappelle que :  $\hat{\psi}(\omega) = \int_{-\infty}^{+\infty} \psi(t) e^{-j\omega t} dt$  où  $j = \sqrt{-1}$ . L'équation (3.3.1) montre que la TOC est définie par le produit de convolution entre  $u(t)$  et  $\mathcal{P}\{\psi_a\}(t)$  où  $\mathcal{P}$  est l'opérateur de parité défini par  $\mathcal{P}\{\psi_a\}(t) = \psi_a(-t)$ . On obtient ainsi une expression duale de (3.3.1) en lui appliquant le théorème de Parseval :

$$T_{\psi}[u](b, a) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{u}(\omega) \hat{\psi}(a\omega) e^{j\omega b} d\omega \quad (3.3.2)$$

où  $\hat{u}(\omega)$  est la transformée de Fourier (TF) de  $u(t)$ .

L'équation (3.3.2) définit la TOC, pour toute valeur de  $a > 0$ , comme la TF inverse du produit  $\hat{u}(\omega) \hat{\psi}(a\omega)$ , et permet le calcul numérique de la TOC à l'aide des algorithmes de transformées de Fourier rapides.

### 3.3.2 Propriétés et choix de l'ondelette mère

Un choix « optimisé » de l'ondelette mère à des fins d'identification par la TOC, relève de trois critères : l'admissibilité, la progressivité et les localisations temporelle et fréquentielle de l'ondelette mère.

Le premier critère d'admissibilité entraîne que la TF de  $\psi$  est nulle à l'origine :  $\hat{\psi}(0) = 0$  et donc que  $\psi$  est de moyenne nulle [CHT98]. Il est essentiel car on démontre que si l'ondelette est admissible, la TOC inverse existe.

Le deuxième critère important pour l'obtention et l'extraction des arêtes est la progressivité, qui par définition signifie que la TF de  $\psi$  est nulle pour des pulsations négatives :  $\hat{\psi}(\omega) = 0$  pour  $\omega \leq 0$ . Cette propriété permet de relier directement le signal réel  $u(t)$  avec le signal analytique<sup>(2)</sup>  $Z^u(t)$  qui lui est associé. L'amplitude et la fréquence instantanées deviennent par conséquent plus faciles à extraire. L'importance de ce critère est détaillé ci-après (dans les remarques faites après l'Eq. (3.4.14)).

Une autre propriété importante est la localisation en temps et en fréquence des ondelettes mères et de la TOC. On cherche à concevoir des fonctions analysantes qui se situent à mi-chemin des distributions de Dirac  $\delta(t - \tau)$  et des fonctions  $e^{j\omega t}$  qui sont parfaitement localisées respectivement en temps et en fréquence. Ces nouvelles fonctions auront à la fois une bonne localisation fréquentielle et une bonne localisation temporelle.

On utilise généralement pour définir les localisations temporelle et fréquentielle, les notions de moyennes et d'écart-types dans les domaines temporel et fréquentiel qui sont définis par rapport à la densité de probabilité respectivement :  $\varphi_1(t) = \frac{|\psi(t)|^2}{\|\psi\|_2^2}$  pour le do-

maine temporel et  $\varphi_2(\omega) = \frac{|\hat{\psi}(\omega)|^2}{\|\hat{\psi}\|_2^2}$  pour le domaine fréquentiel [Mal00]. Pour l'ondelette

mère, les moyennes temporelle et fréquentielle sont notées  $t_\psi$  et  $\omega_\psi$  et les écart-types temporel et fréquentiel  $\Delta t_\psi$  et  $\Delta \omega_\psi$  respectivement. La fonction mère  $\psi(t)$  est dite localisée au point de phase  $(t_\psi, \omega_\psi)$  avec une incertitude  $\mu_\psi = \Delta t_\psi \Delta \omega_\psi$  qui d'après le principe d'incertitude d'Heisenberg vérifie :  $\mu_\psi \geq \frac{1}{2}$ ; entraînant qu'une amélioration de la localisation temporelle (i.e. une diminution de  $\Delta t_\psi$ ) va être accompagnée par une détérioration de la localisation fréquentielle (i.e., une augmentation de  $\Delta \omega_\psi$ ). Par extension, on définit les localisations temporelle et fréquentielle de la TOC en appliquant les définitions précédentes à la version dilatée et translatée :  $\psi_{b,a}(t) = \frac{1}{a}\psi\left(\frac{t-b}{a}\right)$ . On obtient alors les notions de moyennes :  $t_{\psi_{b,a}}$  et  $\omega_{\psi_{b,a}}$  et d'écart-types :  $\Delta t_{\psi_{b,a}}$ ,  $\Delta \omega_{\psi_{b,a}}$  respectivement dans les domaines temporel et fréquentiel [Le03].

Pour caractériser la qualité du filtre de l'ondelette mère et par suite la localisation de la TOC, [LA04] proposent le facteur de qualité  $Q$  défini comme le rapport de la pulsation moyenne  $\omega_\psi$  à la largeur de l'intervalle fréquentiel  $2\Delta \omega_\psi$ . Ce facteur permet aux auteurs de comparer plusieurs ondelettes mères et ils montrent que pour une même valeur de  $Q$ , les résultats obtenus par l'ondelette mère de Cauchy et par l'ondelette de Morlet sont similaires.

Nous avons utilisé l'ondelette mère de Cauchy d'ordre  $n$  :  $\psi_{\beta,n}(t)$  modifiée par le paramètre positif  $\beta$ .  $\psi_{\beta,n}(t)$  et sa transformée de Fourier :  $\hat{\psi}_{\beta,n}(\omega)$  sont définies respectivement par :

$$\psi_{\beta,n}(t) = \left(\frac{j}{\beta t + j}\right)^{n+1} \quad \hat{\psi}_{\beta,n}(\omega) = 2\pi \left(\frac{\omega}{\beta}\right)^n \frac{e^{-\frac{\omega}{\beta}}}{n!} H(\omega) \quad (3.3.3)$$

où  $H(\omega)$  est la fonction de Heaviside. Lorsque  $t$  désigne le temps,  $\beta$  a la dimension de l'inverse d'un temps.

On montre aisément que l'ondelette de Cauchy modifiée est admissible et progressive. On redonne son facteur de qualité :  $Q = \frac{\sqrt{2n+1}}{2}$ , ses moyennes temporelle et fréquentielle :  $t_\psi = 0$  et  $\omega_\psi = \frac{\beta}{2}(2n+1) = 2\beta Q^2$ ; les écarts-types temporel et fréquentiel :  $\Delta t_\psi = \frac{1}{\beta\sqrt{2n-1}} = \frac{1}{\beta\sqrt{4Q^2-2}}$  et  $\Delta \omega_\psi$  et  $\Delta \omega_\psi = \beta\frac{\sqrt{2n+1}}{2} = \beta Q$  et enfin la valeur

<sup>2</sup>Le signal analytique s'obtient à partir du signal réel en annulant les valeurs du spectre pour les fréquences négatives; ceci a pour effet de « complexifier » le signal initial.

$\omega_{\max}$  de la pulsation telle que  $\hat{\psi}_{\beta,n}(\omega)$  est maximale :  $\omega_{\max} = \beta n = \beta (2Q^2 - \frac{1}{2})$ . On déduit l'incertitude :  $\mu_{\psi} = \frac{1}{2} \sqrt{1 + \frac{2}{2n-1}} = \frac{Q}{\sqrt{4Q^2-2}}$ .

Nous retenons le cas  $\beta = 1$  qui est le plus courant et pour simplifier l'écriture, nous posons  $\psi_n(t) = \psi_{1,n}(t) = \left[ \frac{1}{\sqrt{t^2+1}} \right]^{n+1} e^{j(n+1)\text{Arctg}(t)}$ . La fonction  $\psi_n(t)$  est tracée (parties réelle et imaginaire) en Fig. (3.1) et sa transformée de Fourier  $\hat{\psi}_n(\omega)$  en Fig. (3.2) (schéma à gauche) pour trois valeurs de  $n$  :  $n = 1, 4$  et  $15$ . On note que  $\hat{\psi}_n(\omega)$  est réelle, positive et nulle pour des pulsations négatives.

Le choix des paramètres de l'ondelette  $n$  et  $\beta$  doit tenir compte de deux exigences :

– *la réduction des effets de bord*

le calcul numérique de la TOC entraîne un biais dû au fait que le signal enregistré est de longueur finie comme illustré en Fig. (3.2) où l'on voit nettement les perturbations sur le bord droit. Pour remédier à cette difficulté, un domaine de « validité » où les effets de bords peuvent être négligés, est proposé dans le plan temps-fréquence ; ce domaine est déterminé par deux hyperboles (cf. Fig. (3.2)) et deux horizontales [LA04] ;

– *les effets du couplage fréquentiel*

lorsque le signal analysé contient deux fréquences caractéristiques voisines (deux modes de vibration proches), l'identification des paramètres modaux est perturbée.

Pour cela, [LA04] fondent leur procédure sur le facteur de qualité et proposent un intervalle de valeurs « acceptables » pour  $Q$  :  $Q \in \left[ Q_{\min} = c_f \frac{\omega_r}{2d\omega}, Q_{\max} = \frac{\omega_r L}{4c_t \mu_{\psi}} \right]$  avec

$$Q_{\min} = c_f \frac{\omega_r}{2d\omega} \quad (3.3.4)$$

$$Q_{\max} = \frac{\omega_r L}{4c_t \mu_{\psi}} \quad (3.3.5)$$

où  $\mu_{\psi} = \Delta t_{\psi} \Delta \omega_{\psi}$ ,  $c_t$  et  $c_f$  sont deux paramètres choisis pour limiter les effets de bord,  $\omega_r$  est la fréquence d'intérêt généralement voisine d'une pulsation propre du système et  $d\omega$  est un écart fréquentiel (par exemple l'écart minimum entre  $\omega_r$  et une fréquence proche. On doit toujours vérifier  $Q_{\min} < Q_{\max}$ . Plus récemment dans [EA07], les auteurs raffinent la procédure en introduisant une nouvelle valeur  $Q_{\xi}$  de  $Q$  liée au taux d'amortissement modal  $\xi$  considéré :  $Q_{\xi} = \frac{1}{2\xi} \sqrt{1 + \sqrt{1 - 4\xi^2}}$ . Pour le calcul de  $Q_{\xi}$ , une estimation « grossière » de  $\xi$  suffit (par exemple à l'aide de la méthode de la demi-puissance). Le choix de  $Q$  peut alors se faire de la façon suivante :

- si  $Q_{\min} \leq Q_{\xi}$  alors  $Q = \min(Q_{\xi}, 1/2(Q_{\min} + Q_{\max}))$ .
- si  $Q_{\min} > Q_{\xi}$  alors  $Q = Q_{\min}$ .

### 3.4 Analyse en ondelettes et identification modale

Ce paragraphe développe avec plus ou moins de détails, trois manières d'utiliser la TOC pour l'identification modale à partir des fonctions de réponses en fréquence ou des réponses transitoires, ou encore des réponses sous excitations ambiantes. Les trois techniques d'identification modale proposées diffèrent suivant le type de signaux mesurés. Elles sont dénotées respectivement (IMFRF) lorsque les signaux de mesure

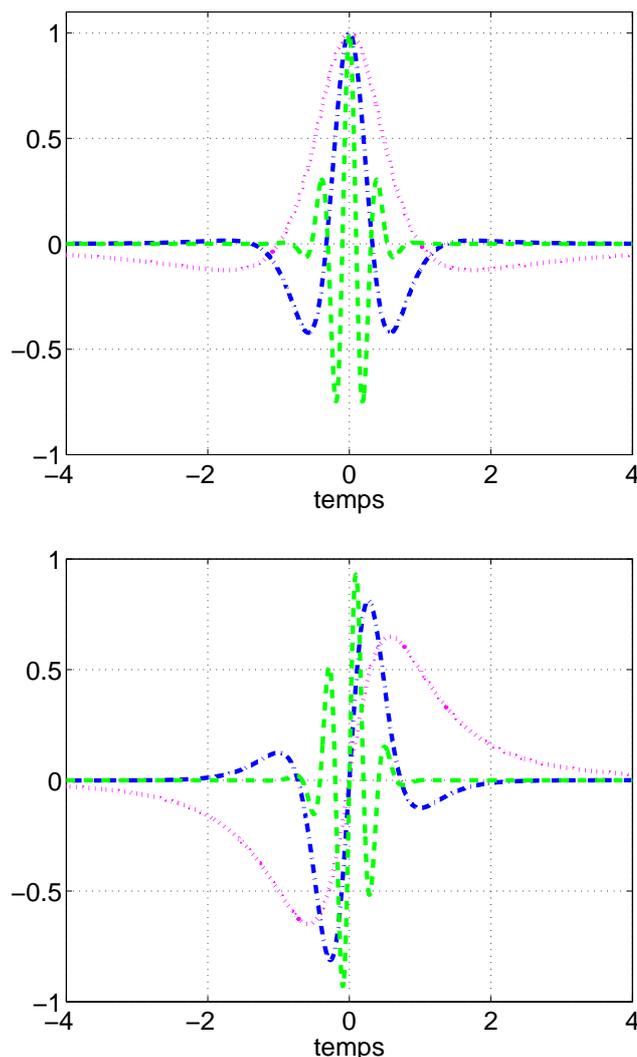


FIG. 3.1 – Ondelette mère de Cauchy  $\psi_n(t)$  : en haut partie réelle et en bas partie imaginaire pour .....  $n = 1$ , -.-.  $n = 4$  et - - -  $n = 15$ .

sont des FRFs, (IMRL) lorsque ceux-ci sont des réponses libres et finalement (IM-REA) lorsque l'on traite des réponses sous excitations ambiantes. Les deux premières techniques (IMFRF) et (IMRL) sont présentées avec plus de détails.

### 3.4.1 à partir des fonctions de réponses en fréquence (IMFRF)

On propose d'appliquer la TOC aux FRFs dans le but d'établir des formules pour identifier les paramètres modaux. On introduit alors la TOC d'une FRF  $H(\omega)$  que l'on note  $W(a, b)$ . Pour simplifier l'écriture, on pose :  $\varphi\left(\frac{\omega-b}{a}\right) = \overline{\psi}\left(\frac{\omega-b}{a}\right)$ . A partir de la relation (3.3.1) définissant la TOC, on obtient :

$$W(a, b) = T_\psi[H](b, a) = \frac{1}{a} \int_{-\infty}^{+\infty} H(\omega) \varphi\left(\frac{\omega-b}{a}\right) d\omega \quad (3.4.1)$$

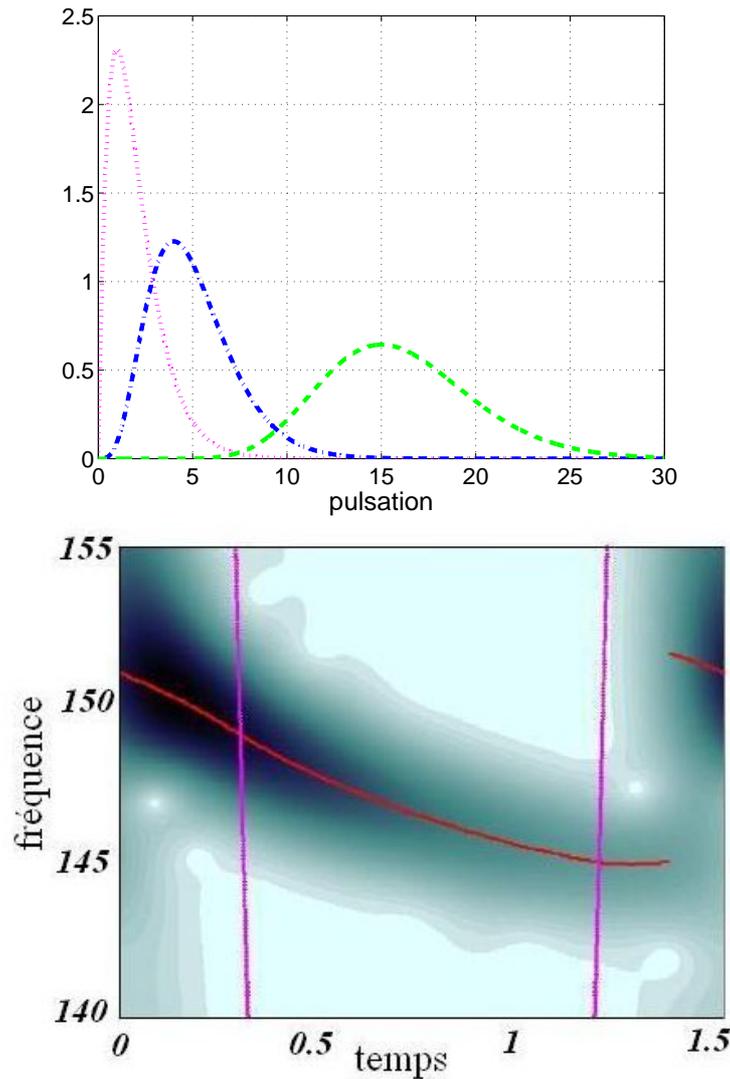


FIG. 3.2 – En haut : Transformée de Fourier de l'ondelette mère de Cauchy  $\psi_n(t)$  : pour  $\dots$   $n = 1$ ,  $-\cdot-$   $n = 4$  et  $- - -$   $n = 15$  et en bas : Module de la TOC d'un signal non-stationnaire - Illustration de l'effet de bord

Si on suppose que  $H(\omega)$  vérifie (3.2.1), l'équation (3.4.1) devient :

$$W(a, b) = \frac{-j}{a} \sum_{r=1}^N \left( A_r \int_{-\infty}^{+\infty} \frac{\varphi\left(\frac{\omega-b}{a}\right)}{\omega + jp_r} d\omega + \overline{A_r} \int_{-\infty}^{+\infty} \frac{\varphi\left(\frac{\omega-b}{a}\right)}{\omega + j\overline{p_r}} d\omega \right) \quad (3.4.2)$$

Dans la suite, on suppose que  $\varphi(s)$  est une fonction analytique dans le demi-plan au-dessus de l'axe réel. Les intégrales de l'équation (3.4.2) se calculent alors de manière analytique à l'aide de la formule intégrale de Cauchy dans la théorie des fonctions à variable complexe. On obtient :

$$W(a, b) = \frac{2\pi}{a} \sum_{r=1}^N \left[ A_r \varphi\left(\frac{-jp_r - b}{a}\right) + \overline{A_r} \varphi\left(\frac{-j\overline{p_r} - b}{a}\right) \right] \quad (3.4.3)$$

Dans [YDA04], les auteurs proposent comme ondelette mère les fonctions fractionnaires suivantes :

$$\psi_x(s) = \frac{1}{(1+js)^{x+1}} = e^{-(x+1)\log(1+js)} \quad (3.4.4)$$

dans laquelle  $x$  est un nombre réel positif. Lorsque  $x = n$  avec  $n$  entier naturel strictement positif, on remarque que  $\psi_x(\omega) = \psi_{\beta=1/n}(\omega)$  ; la dimension de  $\beta$  étant l'inverse d'une pulsation.

On peut vérifier que  $\psi_x(s)$  satisfait les conditions exigées par une ondelette mère. On choisit la détermination principale de la fonction logarithmique  $\log(z)$  avec  $z = 1 + js$  avec la coupure correspondant à l'axe des réels négatifs.  $\log(z)$  est définie et analytique dans le plan complexe privé de l'axe des réels négatifs ; par suite, le conjugué de  $\psi_x(s)$  :  $\varphi_x(s) = \overline{\psi_x(s)} = \frac{1}{(1-j\overline{s})^{x+1}} = e^{-(x+1)\log(1-j\overline{s})}$  est défini et analytique dans le plan complexe avec une coupure qui correspond à la partie de l'axe imaginaire au dessous du point  $s = (0, -j)$ . Finalement, puisque les valeurs de  $a$  sont toujours positives, la fonction  $\varphi_x(s)$  est définie et analytique dans le demi plan au dessus de l'axe réel. On note ensuite  $W_x(a, b)$  la TOC d'une FRF en remplaçant l'ondelette mère  $\psi$  par  $\psi_x$  dans l'équation (3.4.1). De même, la relation (3.4.3) devient :

$$W_x(a, b) = 2\pi a^x \sum_{r=1}^N \left[ \frac{A_r}{(a + jb - p_r)^{x+1}} + \frac{\overline{A}_r}{(a + jb - \overline{p}_r)^{x+1}} \right] \quad (3.4.5)$$

Au voisinage d'une pulsation propre, c'est-à-dire si  $|b - \omega_r|$  est suffisamment petit, on suppose que la fonction  $W_x(a, b)$  est dans une large mesure, influencée par une seule fonction fractionnaire :

$$W_x(a, b) \approx \frac{2\pi a^x A_r}{(a + jb - \lambda_r)^{x+1}} \quad (3.4.6)$$

La qualité de cette approximation dépend évidemment de l'interaction entre des modes voisins. Une analyse de la fonction  $|a^{\frac{1-x}{2}} W_x(a, b)|$  montre qu'elle décrit approximativement une surface parabolique de révolution dont le sommet est proche du point  $(a = -\sigma_r, b = \omega_r)$ . Cette propriété suggère que les valeurs de  $\sigma_r$  et de  $\omega_r$  peuvent être estimées en recherchant les maxima locaux de la fonction  $|a^{\frac{1-x}{2}} W_x(a, b)|$ . Une méthode pour estimer les paramètres modaux fondée sur cette propriété a été validée sur des données analytiques et expérimentales [YDA04].

Dans la suite, on cherche à établir des formules d'estimation des paramètres modaux par une démarche similaire à celle des formules des différences finies [YD00]. Pour cela, on considère deux valeurs de  $W_x(a, b)$  aux points  $(a, b^-)$  et  $(a, b^+)$ , avec  $b^- < \omega_r < b^+$ ,  $\omega_r$  étant la  $r^e$  pulsation propre amortie. L'approximation donnée en (3.4.6) conduit aux équations suivantes :

$$W_x(a, b^-) \approx \frac{2\pi a^x A_r}{(a + jb^- - p_r)^{x+1}} \quad (3.4.7)$$

$$W_x(a, b^+) \approx \frac{2\pi a^x A_r}{(a + jb^+ - p_r)^{x+1}} \quad (3.4.8)$$

On peut alors estimer sans difficulté  $p_r$ . En posant  $R_x = \left( \frac{W_x(a, b^+)}{W_x(a, b^-)} \right)^{\frac{1}{x+1}}$ , on obtient :

$$p_r \approx a + \frac{j}{2} \left[ b^+ + b^- + (b^+ - b^-) \frac{R_x + 1}{R_x - 1} \right] \quad (3.4.9)$$

L'estimation du résidu  $A_r$  peut alors être déduite de l'une des équations (3.4.7) et (3.4.8) :

$$A_r \approx \frac{(a + jb^+ - p_r)^{x+1} W_x(a, b^+)}{2\pi a^x} \quad (3.4.10)$$

Par exemple, pour  $x = 1$ , les relations précédentes deviennent :

$$p_r \approx a + \frac{j}{2} \left[ b^+ + b^- + (b^+ - b^-) \frac{\sqrt{W_1(a, b^+)} + \sqrt{W_1(a, b^-)}}{\sqrt{W_1(a, b^+)} - \sqrt{W_1(a, b^-)}} \right] \quad (3.4.11)$$

$$A_r \approx \frac{(a + jb^+ - p_r)^2 W_1(a, b^+)}{2\pi a} \quad (3.4.12)$$

On rappelle que  $W_1(a, b) = \frac{1}{a} \int_{-\infty}^{+\infty} H(\omega) \overline{\psi_1\left(\frac{\omega-b}{a}\right)} d\omega$  avec  $\psi_1(s) = \frac{1}{(1+js)^2}$ . Les relations (3.4.9) et (3.4.10) sont similaires à celles des différences finies. Mais elles donnent une estimation plus précise des paramètres modaux grâce à la capacité de la TOC à minimiser l'effet d'interaction entre des modes voisins.

Dans la pratique, la valeur de  $b^-$  doit être inférieure à celle de la  $r^e$  fréquence propre  $\omega_r$  et la valeur de  $b^+$  supérieure. La valeur moyenne de  $b^-$  et de  $b^+$  doit être la plus proche possible de celle de  $\omega_r$ .

Donc il faut commencer par estimer  $\omega_r$  qui correspond à la partie imaginaire de  $\lambda_r$  par la méthode itérative suivante : à chaque étape d'itération, on estime  $\lambda_r$  et on déduit la valeur de  $\omega_r$  qui servira de valeur moyenne de  $b^-$  et de  $b^+$  à l'étape suivante. Une fois la convergence atteinte, la valeur de  $\omega_r$  servira de valeur moyenne de  $b^-$  et de  $b^+$  pour l'estimation de l'amortissement et du résidu.

Une attention particulière doit être donnée au choix de la valeur de  $a$ . En théorie, la précision de l'estimation est d'autant meilleure que la valeur de  $a$  est petite en raison de l'interaction entre modes voisins. Dans la pratique, on constate que les résultats d'estimation sont instables pour de très petites valeurs de  $a$ . On doit donc faire varier  $a$  afin de trouver la valeur optimale. Quant à  $x$ , on peut prendre par exemple  $x = 1$  ou  $x = 3$ , mais de préférence des valeurs inférieures à 5 pour des raisons d'instabilité. De manière générale, le fait d'augmenter la valeur de  $x$  minimise l'interaction entre modes voisins mais peut rendre les résultats instables car sensibles aux données bruitées et aux éventuelles non-linéarités au voisinage des fréquences de résonance.

### 3.4.2 à partir des oscillations libres (IMRL)

Le cas (IMRL) a été étudié plus intensivement durant les dix dernières années par plusieurs chercheurs. Plusieurs ondelettes ont été proposées pour identifier les fréquences propres et les taux d'amortissement modaux dont l'ondelette de Morlet modifiée et l'ondelette de Cauchy précédemment introduite. L'effet du bruit sur les mesures aussi bien que le traitement des effets de bords de la TOC et le choix des paramètres optimaux de l'ondelette mère ont été étudiés par plusieurs auteurs : [RFGP97], [LA04], [SSB03], [LTB04].

On s'intéresse ici aux signaux modulés en fréquence autour d'une fréquence porteuse et modulés en amplitude par une enveloppe  $\Upsilon(t)$ . Ils peuvent se mettre sous la forme :  $\Upsilon(t) \cos(\alpha(t))$ . Une condition nécessaire pour l'utilisation de la TOC pour l'identification modale à partir des réponses libres de structures est l'asymptoticité des signaux traités

qui entraîne que :

$$\left| \frac{\dot{\Upsilon}(t)}{\Upsilon(t)} \right| \ll \dot{\alpha}(t) \quad (3.4.13)$$

La relation (3.4.13) traduit que les variations de l'amplitude sont beaucoup plus lentes que celles de la phase ([CHT98]). L'amplitude est quasi-constante à l'échelle d'une pseudo-période  $T(t) = \frac{2\pi}{|\dot{\alpha}|}$ . On rajoute généralement que la pseudo-période  $T(t)$  est elle-même lentement variable entre deux oscillations successives :  $\frac{|\dot{\alpha}(t)|}{\alpha^2(t)} \ll 1$ . De tels signaux sont appelés des « chirps » [CMF99] (un signal harmonique avec amplitude constante et une phase linéaire en temps, est un cas trivial de chirp).

Pour un système discret à  $N$  DdLs, linéaire avec des paramètres constants et faiblement amorti, on a vu en (3.2.4) que la  $k^e$  composante  $u_k$  du vecteur déplacement en réponse libre peut se mettre sous la forme d'une somme de  $N$  composantes modales  $u_{kr}(t)$  qui sont compatibles avec la condition d'asymptoticité (3.4.13) dès que  $\xi_r \ll \frac{1}{\sqrt{2}}$ . Lorsque l'amortissement modal est faible, cette dernière condition est toujours vérifiée et les réponses libres des systèmes linéaires sont asymptotiques. Finalement, pour une composante modale asymptotique  $u_{kr}(t)$ , le signal analytique associé

$$Z^{(u_{kr})}(t) = u_{kr}(t) + j \mathfrak{H} [u_{kr}] (t) \quad (3.4.14)$$

avec  $\mathfrak{H}$  désignant la transformée de Hilbert, est sensiblement égal au signal complexe associé  $z^{(u_{kr})}(t) = \Upsilon^{(u_{kr})}(t) e^{i \alpha^{(u_{kr})}(t)}$ .

Lorsque l'ondelette mère est progressive, on montre qu'il existe un ensemble de points dans le plan temps-échelle  $(b, a)$  autour duquel l'énergie de la TOC de  $u_{kr}(t)$  se localise. Cet ensemble est appelé **arête** (« ridge » en anglais) et la valeur de la TOC sur l'arête, le **squelette** (« skeleton » en anglais) de la TOC de  $u_{kr}(t)$ . Il existe plusieurs définitions de l'arête suivant le critère de localisation de l'énergie choisie pour la TOC. Une définition classique postule que le carré du module de la TOC est maximum sur l'arête. Ainsi pour tout  $b$ , le module  $|T_\psi [u_{kr}] (b, a)|$  est maximum lorsque :

$$a = a_r^{(u_{kr})}(b) = \frac{\omega_{max}}{\dot{\alpha}^{(u_{kr})}(b)} \quad (3.4.15)$$

où  $\dot{\alpha}^{(u_{kr})}(b)$  désigne la fréquence instantanée du signal  $u_{kr}(t)$ .

Plusieurs méthodes pour définir et extraire les ridges sont proposées dans l'ouvrage ([CHT98]), et les auteurs montrent que le squelette d'un signal asymptotique  $u_{kr}$  calculé sur une arête est proportionnelle au signal complexe associé  $z^{(u_{kr})}(b)$  :

$$T_\psi [u_{kr}] (b, a_r^{(u_{kr})}(b)) \simeq \frac{1}{2} \hat{\psi}^-(\omega_{max}) Z^{(u_{kr})}(b) \simeq \frac{1}{2} \hat{\psi}^-(\omega_{max}) z^{(u_{kr})}(b) \quad (3.4.16)$$

Intéressons nous ensuite au signal  $u_k(t) = \sum_{r=1}^N u_{kr}(t)$  exprimé en relation (3.2.4). Dans ce cas, la TOC a  $N$  arêtes, une par composante du signal. Sur la  $s^e$  arête, l'énergie liée à la composante  $u_{ks}$  est maximale tandis que la contribution due aux autres composantes  $u_{kr}$  pour  $r \neq s$  est négligeable dès que les fréquences propres sont suffisamment éloignées les unes des autres. Ensuite grâce à la propriété de linéarité de la TOC, on déduit de ce qui précède que :

$$\begin{aligned} T_\psi [u_k] (b, a_r^{(u_{ks})}(b)) &= T_\psi \left[ \sum_{r=1}^N u_{kr} \right] (b, a_r^{(u_{ks})}(b)) = \sum_{r=1}^N T_\psi [u_{kr}] (b, a_r^{(u_{ks})}(b)) \\ &\simeq T_\psi [u_{ks}] (b, a_r^{(u_{ks})}(b)) \simeq \frac{1}{2} \hat{\psi}^-(\omega_{max}) z^{(u_{ks})}(b) \end{aligned} \quad (3.4.17)$$

L'équation (3.4.17) montre qu'il est possible d'isoler la composante modale  $u_{ks}(t) = \text{Re}(z^{(u_{ks})}(t))$  à partir de la TOC du signal multi-composante  $u_k(t)$ .

La décroissance des différentes contributions modales  $u_{kr}(t)$  est associée au produit modal  $\sigma_r$  du taux d'amortissement modal et de la  $r^{\text{e}}$  pulsation propre du système conservatif associé. Leur contribution à la réponse libre globale décroît avec des taux de décroissance différents et après un certain temps qui dépend du produit modal et de l'amplitude modale initiale, elle peut avoir un niveau énergétique trop faible qu'il devient alors très difficile voire impossible d'extraire l'arête correspondante dans le plan temps-fréquence. Pour remédier à cette difficulté, Erlicher et Argoul [EA07] définissent pour chaque composante du signal un temps d'arrêt lorsque l'amplitude du signal décroissant atteint un seuil critique.

L'identification des paramètres modaux peut alors s'effectuer en trois étapes :

- D'après ce qui précède, la TOC de chaque signal enregistré produit une arête centrée sur chaque fréquence propre du système (cf. relation (3.4.17)), sauf si le point de mesure est situé sur un noeud de vibration de la structure. Chaque arête illustre l'évolution temporelle de la fréquence instantanée correspondante. Pour un système linéaire et stationnaire, l'arête dans le plan temps-fréquence est une fonction constante en temps et égale à la pseudo-fréquence propre du mode considéré de telle sorte que :  $2\pi\tilde{f}_r = \omega_r = \omega_{0r}\sqrt{1-\xi_r^2}$ . Afin de déterminer la valeur de chaque fréquence propre, Erlicher et Argoul [EA07] proposent de moyenner en temps (pour les valeurs de  $b_l$  où l'effet de bord est négligeable) et pour les  $N_{mes}$  signaux de mesure, les valeurs obtenues pour la fréquence instantanée correspondante. On pourrait raffiner la procédure en pondérant la moyenne sur les  $N_{mes}$  signaux de mesure en mettant suivant la position du point de mesure considéré un poids proportionnel à la forme du mode étudié.
- Chaque taux d'amortissement  $\xi_r$  peut être identifié en considérant que les signaux analytiques extraits  $Z^{(u_{kr})}(t)$  ont une décroissance exponentielle. Une procédure de régression linéaire est alors utilisée sur les données  $\ln |Z^{(u_{kr})}(b_l)|$  ; les  $b_l$  étant à l'intérieur du domaine où l'effet de bord est négligeable. La pente  $\sigma_r$  et la valeur à l'origine  $|G_r| |\phi_{kr}|$  peuvent être identifiées. En utilisant la valeur  $\omega_r$  identifiée à l'étape précédente, on peut déduire aisément  $\xi_r$  et  $\omega_{0r}$ .
- Soient  $Z^{(u_{lr})}(b)$  et  $Z^{(u_{kr})}(b)$  les signaux analytiques extraits par la TOC pour la fréquence  $\tilde{f}_r$  respectivement au point de référence  $l$  et au point générique  $k$  de la structure. Le point de référence peut être choisi au voisinage d'un ventre de vibration du mode considéré. Le vecteur complexe  $\mathbf{Z}_{u_r}(b) = \left[ \frac{Z^{(u_{1r})}(b)}{Z^{(u_{1r})}(b)}, \dots, \frac{Z^{(u_{N_{mes}r})}(b)}{Z^{(u_{1r})}(b)} \right]^T$  formé par les rapports instantanés  $\frac{Z^{(u_{kr})}(b)}{Z^{(u_{lr})}(b)}$  peut être interprété comme une l'évolution temporelle d'un vecteur propre instantané normalisé respectivement au point de référence choisi.

Les différentes phases de la procédure d'identification sont décrites précisément dans [EA07]. Plusieurs simulations numériques ont permis de valider la procédure [LA04], [EA07].

### 3.4.3 à partir des réponses sous excitations ambiantes (IMREA)

Pour le cas (IMREA), une procédure d'identification fondée sur une fonction spectrale ondelette avec un choix adapté du facteur de qualité  $Q$  est proposée dans [NAC05] afin d'obtenir une identification « correcte » du facteur d'amortissement. Pour

cela, on calcule la fonction spectrale en ondelettes (FSO) de la réponse aux excitations ambiantes, par exemple en accélération  $\ddot{u}_k$  pour le DdL  $k$ . La FSO permet d'estimer la densité spectrale de puissance du signal accélérométrique analysé  $\ddot{u}_k$  et est définie par :

$$V_{\psi,L}^{\ddot{u}_k}(a) = \frac{a}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} |T_{\psi}[\ddot{u}_k](b, a)|^2 db \quad (3.4.18)$$

où  $L$  est la longueur d'enregistrement du signal.

On suppose que la fonction analysante  $\psi$  est telle que sa dérivée première et seconde peuvent aussi jouer le rôle d'une fonction analysante (ce qui est le cas pour l'ondelette mère de Cauchy). On peut alors utiliser indifféremment une vitesse, une accélération ou un déplacement grâce aux propriétés de la TOC suivantes :

$$T_{\psi}[\ddot{u}](b, a) = -\frac{1}{a} T_{\psi}[\dot{u}](b, a) = \frac{1}{a^2} T_{\psi}[u](b, a) \quad (3.4.19)$$

et relier la FSO d'un signal d'accélération à celle d'un signal vitesse ou déplacement par :

$$V_{\psi,L}^{\ddot{u}_k}(a) = \frac{1}{a^2} V_{\psi,L}^{\dot{u}_k}(a) = \frac{1}{a^4} V_{\psi,L}^{u_k}(a) \quad (3.4.20)$$

L'estimation donnée par  $V_{\psi,L}^{\ddot{u}_k}(a)$  est lisse mais biaisée [CHT98] et un choix approprié du facteur de qualité  $Q$  de l'ondelette mère permet de réduire ce biais. On suppose que l'excitation ambiante peut être modélisée par un bruit blanc de moyenne nulle. L'estimation des caractéristiques modales  $\omega_{0_r}, \xi_r$  d'un pôle  $p_r$  se fait ensuite par la minimisation au sens des moindres carrés de l'écart entre la FSO mesurée et celle de la réponse impulsionnelle pondérée de l'oscillateur linéaire associé [Ngu07]. Le paramètre de pondération noté  $B$  permet de tenir compte à la fois de la masse modale du  $r^e$  mode considéré et de la variance inconnue de l'excitation représentée par un bruit blanc et est lui aussi identifié. Le problème de minimisation au sens des moindres carrés s'écrit :

$$\min_{\omega_{0_r}, \xi_r, B} \int_{\mathbb{I}_a} \left\{ V_L^{\ddot{u}_k}(a) - B \frac{V_L^{\dot{h}_k}(a)}{\max_a \{V_L^{\dot{h}_k}(a)\}} \right\}^2 da \quad (3.4.21)$$

où  $\mathbb{I}_a$  est un intervalle des échelles  $a$  centré autour de  $\frac{n}{\omega_{0_r}}$ .

Comme nous l'avons déjà signalé, la solution du problème de minimisation dépend fortement du choix des paramètres de l'ondelette et donc du choix du facteur de qualité  $Q$ . Pour illustrer ce phénomène, [Ngu07] s'intéresse au calcul de la FSO dans le cas de la réponse  $\ddot{u}_k(t)$  d'un oscillateur linéaire visqueux ( $\omega_{0_1} = 10\pi$ rd/s et de  $\xi_1 = 1.5\%$ ) à un bruit blanc de moyenne nulle et de variance unitaire. En traçant les variations de  $\omega_{0_1}$  et de  $\xi_1$  estimés en fonction de  $Q$ , il s'aperçoit que la pulsation  $\omega_{0_1}$  a tendance à augmenter en fonction de  $Q$  alors que l'amortissement  $\xi_1$  diminue. Il propose alors de choisir  $Q = \frac{Q_{min} + Q_{max}}{2}$  où  $Q_{min}$  et  $Q_{max}$  ont été définis en 3.3.4. La figure (3.3) montre la FSO du signal  $\ddot{u}_k(t)$  et celle de la réponse impulsionnelle associée calculées pour  $L = 200s$  et  $Q = 287$ . Les résultats obtenus pour les paramètres modaux par la FSO sont comparés à ceux obtenus par la méthode du décrétement aléatoire (pour plus de détails sur la méthode du décrétement aléatoire (RD random decrement en anglais) cf. la thèse d'[Asm97]). Le tableau 3.1 présente la pulsation propre et le taux d'amortissement estimés par les deux méthodes FSO et RD ; les résultats sont proches et

estiment convenablement les caractéristiques modales de l'oscillateur. Une étude statistique sur 1000 échantillons de réponses de l'oscillateur à différents bruits blancs a été ensuite réalisée ; elle montre que les distributions se concentrent et sont maximales au voisinage des valeurs exactes avec une certaine dispersion [Ngu07].

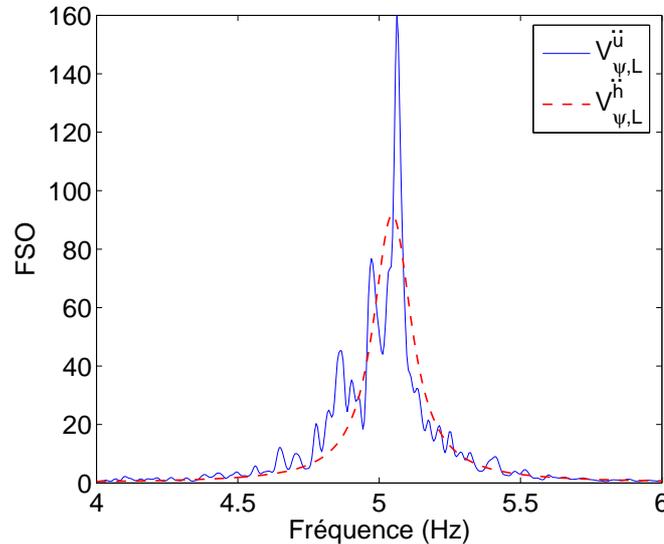


FIG. 3.3 – FSO du signal  $\ddot{u}_k(t)$  et FSO de la réponse impulsionnelle associée

TAB. 3.1 – Paramètres modaux identifiés

Fréquence (Hz)			Amortissement (%)		
Exact	RD	FSO	Exact	RD	FSO
5	4,998	5,052	1,5	1,45	1,45

### 3.4.4 Applications à des données d'essais réels

Ce paragraphe est consacré à l'application de deux des techniques décrites précédemment sur les mesures d'essais réels.

La technique (IMFRF) est appliquée sur les FRFs expérimentales d'un assemblage par collage de trois poutres en plexiglass et la technique (IMRL) sur les réponses après choc d'une plaque en plexiglass.

Dans [AHCB00], [Le03], [Ngu07] et [ANL08], les techniques (IMRL) et (IMREA) ont été également appliquées aux réponses accélérométriques d'un bâtiment du programme expérimental de l'ENTPE ; les résultats obtenus par chacune des deux techniques y sont présentés et comparés.

### 3.4.5 Application de la technique (IMFRF) aux FRFs d'un assemblage de poutres en plexiglass

La technique (IMFRF) présentée précédemment a été validée sur des données expérimentales d'un assemblage par collage de trois poutres en plexiglass. Le spectre de la réponse montre trois fréquences de résonance dans la plage fréquentielle étudiée. Chacune d'elles correspond en fait au premier mode de flexion pour chacune des poutres. Les deux premières fréquences de résonance sont proches tandis que la troisième est isolée des deux premières.

Sur les figures (3.4) et (3.5), on montre la dépendance sur la variable d'échelle  $a$ , des résultats estimés en utilisant les formules présentées précédemment avec  $x = 1$  ou  $x = 3$ . Les valeurs des fréquences propres et des amortissements ont été divisées par celles de référence estimées par la méthode exponentielle complexe. Bien que sensible aux données bruitées, cette méthode sert souvent de méthode de référence.

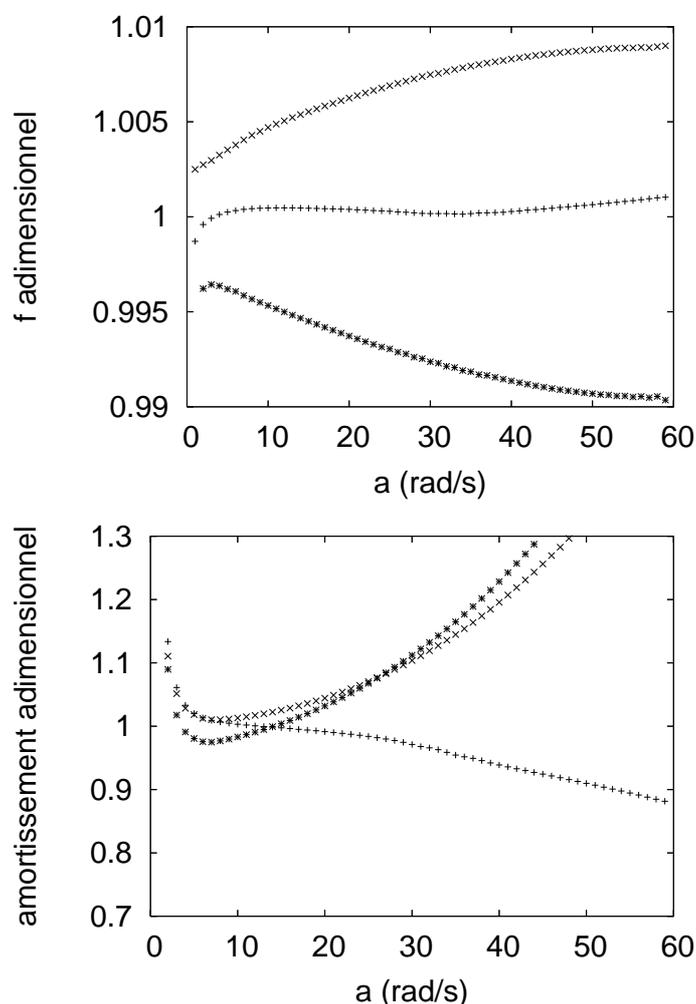


FIG. 3.4 – Estimation avec  $x = 1$ . Dépendance sur la variable d'échelle  $a$  des valeurs des fréquences propres (à gauche) et amortissements (à droite) : (\*\*\*) pour mode 1 ; (xxx) pour mode 2 ; (+++) pour mode 3. Chaque valeur est adimensionnelle après division par la valeur de référence.

On peut remarquer que si on fait diminuer la valeur de  $a$ , les résultats convergent vers

des valeurs proches de celles estimées par la méthode exponentielle complexe, mais deviennent instables pour les très faibles valeurs d'échelle. L'estimation est meilleure avec  $x = 3$  qu'avec  $x = 1$  pour les fréquences propres mais moins satisfaisante pour l'amortissement du premier mode à cause d'une mauvaise convergence pour le premier mode en particulier. Avec  $x = 3$  et  $20 < a < 50$ , la différence des valeurs des fréquences propres estimées et celles de référence est inférieure à 0.2%. Avec  $x = 1$  et  $5 < a < 20$ , la différence entre des valeurs des amortissements estimés et celles de référence est inférieure à 3%.

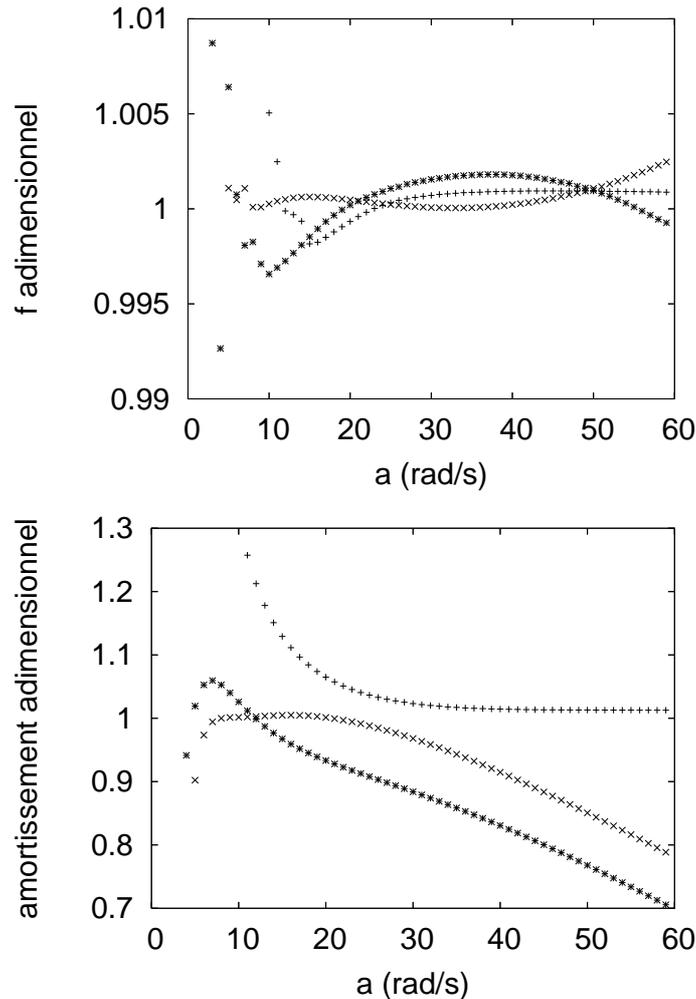


FIG. 3.5 – Estimation avec  $x = 3$ . Dépendance sur la variable d'échelle  $a$  des valeurs des fréquences propres (à gauche) et amortissements (à droite) : ( \*\*\*) pour mode 1 ; (xxx) pour mode 2 ; ( +++ ) pour mode 3. Chaque valeur est adimensionnelle après division par la valeur de référence.

De manière générale, on peut obtenir une meilleure estimation en faisant une optimisation sur  $x$ . Mais dans les situations courantes, comme l'exemple d'application précédente le montre, l'estimation est satisfaisante pour la fréquence propre avec  $x = 3$  et pour l'amortissement avec  $x = 1$ .

### 3.4.6 Application de la technique (IMRL) aux vibrations libres d'une plaque en plexiglass

La technique d'identification modale est appliquée aux signaux d'accélération enregistrés sur une plaque rectangulaire de plexiglass suspendue en deux de ses sommets (cf. Fig. (3.6)). Les dimensions de la plaque sont 39.7 cms de large (axe des abscisses), 49 cms de long (axe des ordonnées) et 0.41 cm d'épaisseur. Les caractéristiques mécaniques de la plaque sont les suivantes : une densité de masse de  $1218.2 \text{ kg/m}^3$ , un module de Young de 5150 MPa et un rapport de Poisson égal à 0.41.

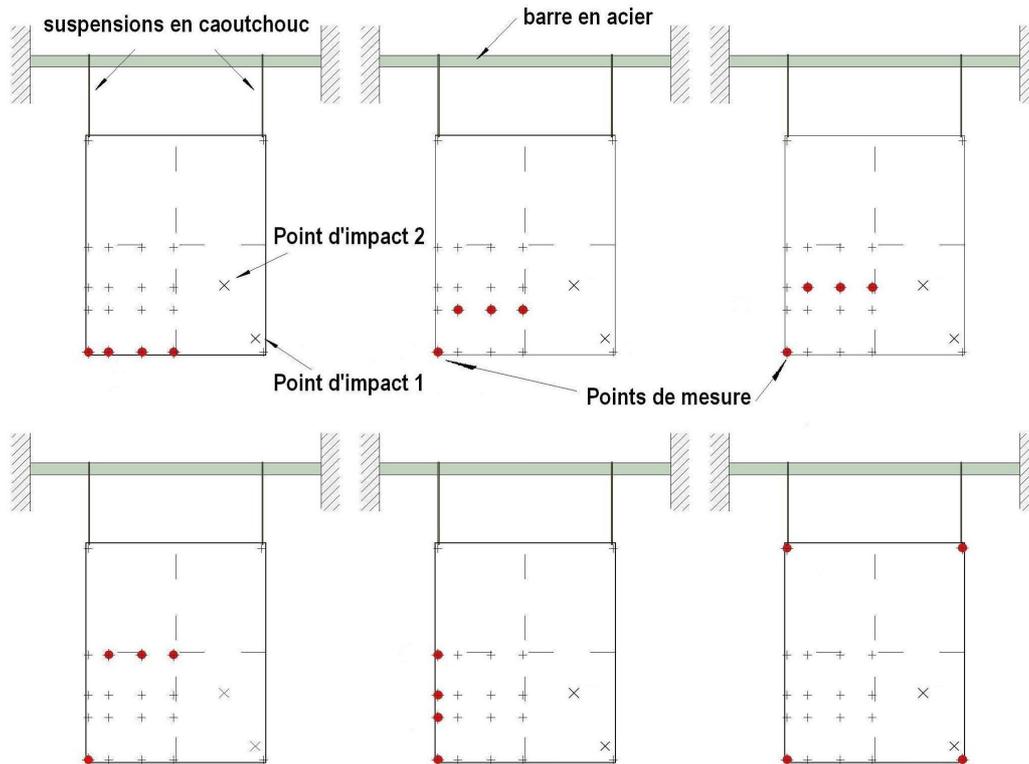


FIG. 3.6 – Schéma de l'essai de vibration sur la plaque de plexiglass avec les six dispositions des quatre accéléromètres disponibles lors de l'essai. La symétrie de la plaque a été exploitée afin de réduire le nombre de signaux enregistrés.

Un modèle par éléments finis de la plaque a été utilisé pour obtenir les fréquences modales (sans amortissement) qui sont données dans le Tableau 3.2. Le type de mode de plaque est indiqué dans la deuxième colonne par les indices  $(m, n)$ , où  $m(n)$  est le nombre de lignes modales orthogonales à l'axe  $x(y)$  respectivement. La plaque est soumise à un essai de chocs qui a été conçu de manière telle à assurer que la vibration de la plaque soit libre dans la direction orthogonale à son plan moyen. Les chocs sont produits en frappant la plaque avec une petite barre en acier aux points 1 et 2 indiqués dans la figure (3.6). Le pas de temps d'enregistrement est  $\Delta t = 0.001953 \text{ s}$  et la longueur d'enregistrement est  $L = (1200 - 1) \Delta t = 2.3418 \text{ s}$ . La fréquence de Nyquist est ainsi :  $f_{Ny} = \frac{1}{2\Delta t} = 256 \text{ Hertz}$ . Quatre accéléromètres étaient disponibles pour les expériences. Pour chaque point de choc, six ensembles de mesures ont été effectués, en modifiant à chaque fois la position de trois des capteurs tout en maintenant fixe le premier, afin d'avoir un signal de référence. Les  $N_{mes} = 24$  signaux ont été analysés,

menant à l'identification de dix modes (Tableau 3.2). Les résultats sont détaillés dans [EA07] ; les figures 3.7, 3.8 illustrent les résultats d'identification pour les deux premiers modes de la plaque : (1,1) et (0,2), respectivement. La deuxième forme modale a une partie imaginaire significative.

Mode	Type	FEM		Données 1			Données 2	
		$f_j$	$\tilde{f}_{j,ident}$	$\xi_{j,ident}$	$\tilde{I}_{np,ident}^{(j)}$	$\tilde{f}_{j,ident}$	$\xi_{j,ident}$	$\tilde{I}_{np,ident}^{(j)}$
(j)		[Hz]	[Hz]	[%]		[Hz]	[%]	
1	(1,1)	26.88	26.7	2.7	0.072			
2	(0,2)	35.26	35.1	3.0	0.309			
3	(2,0)	59.53	58.6	2.8	0.049			
4	(1,2)	65.58	66.6	3.1	0.042			
5	(2,1)	75.92	78.0	3.05	0.077			
6	(0,3)	107.08	107.0	2.95	0.175			
7	(1,3)	128.02	—	—	—	133.2	1.80	0.471
8	(2,2)	129.55	134.2	2.9	0.337	—	—	—
9	(3,0)	158.57	161.1	2.3	0.147	161.3	2.1	0.146
10	(3,1)	179.46	182.2	2.25	0.105	182.4	2.45	0.109

TAB. 3.2 – Plaque en plexiglass : Fréquences modales obtenues par un modèle conservatif utilisant les éléments finis et fréquences modales identifiées par la technique proposée ; taux d'amortissement et indices de non-proportionalité.

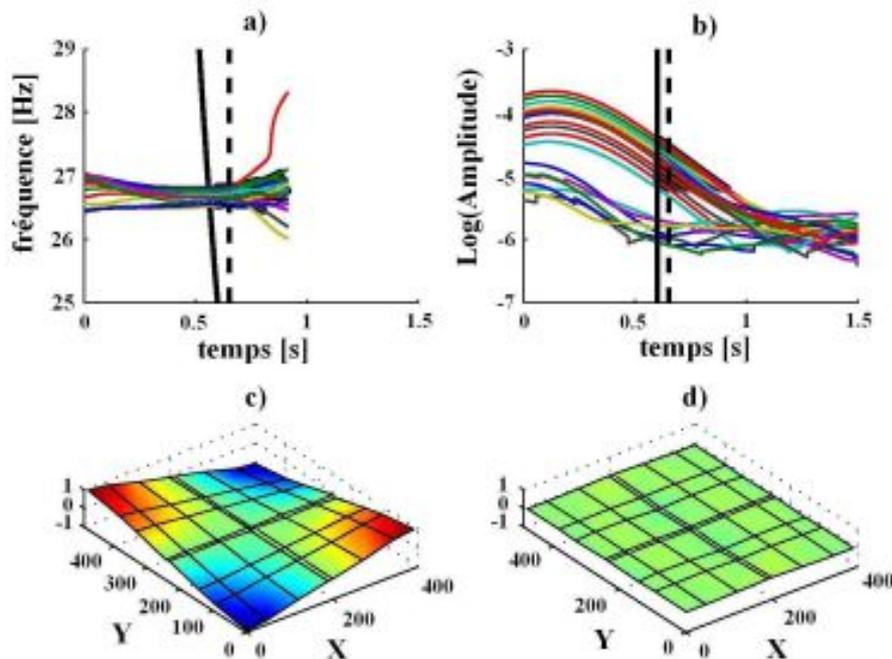


FIG. 3.7 – Plaque de plexiglass, **mode 1**. a) Arêtes extraites à partir des signaux enregistrés. b) Représentation logarithmique des amplitudes des squelettes, avec les intervalles de temps appropriés pour l'analyse modale. c) Partie réelle de la forme modale complexe identifiée. d) Partie imaginaire de la forme modale complexe identifiée.

Une analyse particulière a été nécessaire pour l'identification du mode 7, ayant presque la même fréquence que le mode 8 illustré en Fig. 3.10. En fait, quand le

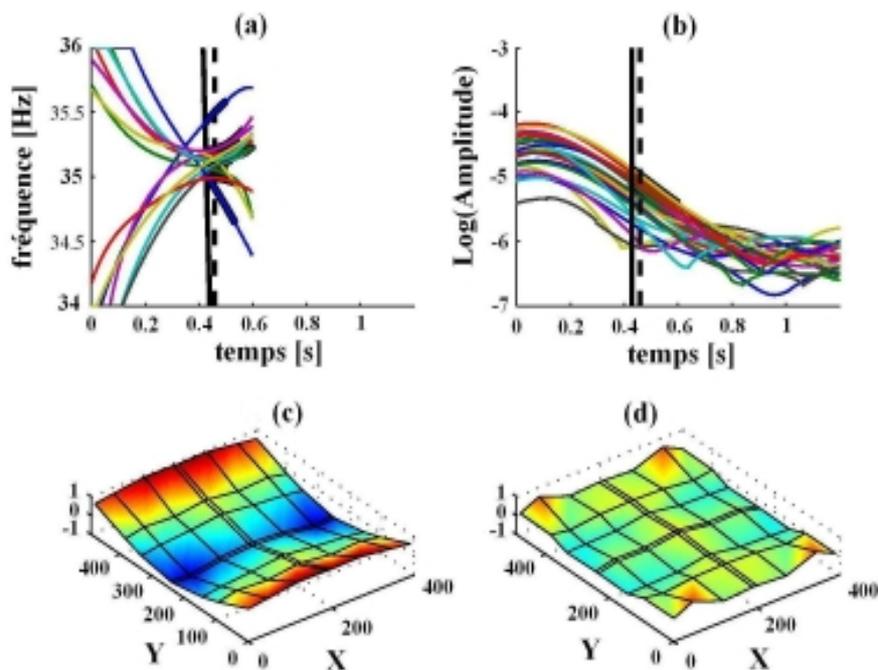


FIG. 3.8 – Plaque de plexiglass, **mode 2**. a) Arêtes extraites à partir des signaux enregistrés. b) Représentation logarithmique des amplitudes des squelettes, avec les intervalles de temps appropriés pour l'analyse modale. c) Partie réelle de la forme modale complexe identifiée. d) Partie imaginaire de la forme modale complexe identifiée.

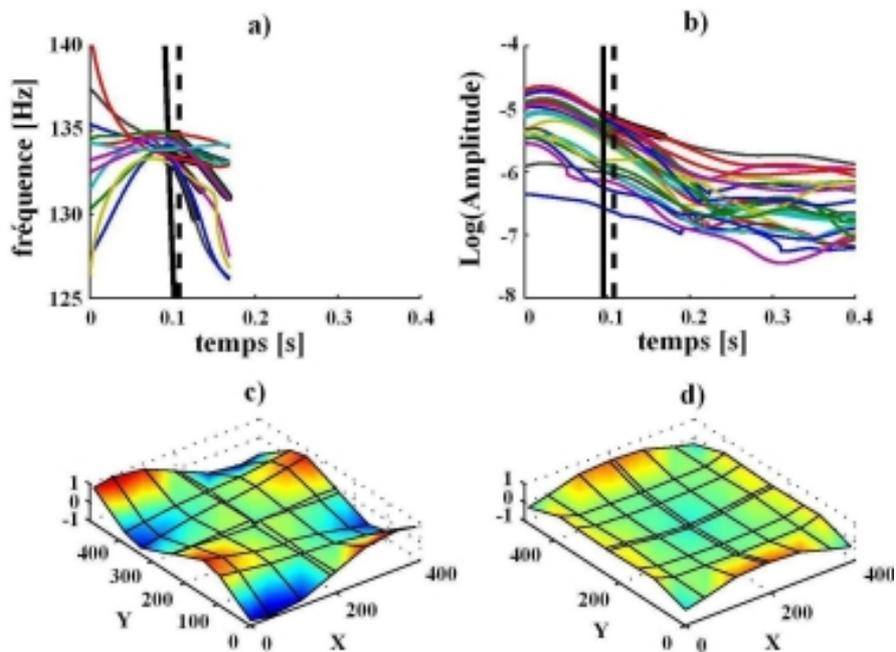


FIG. 3.9 – Plaque de plexiglass, **mode 7**. a) Arêtes extraites à partir des signaux enregistrés. b) Représentation logarithmique des amplitudes des squelettes, avec les intervalles de temps appropriés pour l'analyse modale. c) Partie réelle de la forme modale complexe identifiée. d) Partie imaginaire de la forme modale complexe identifiée.

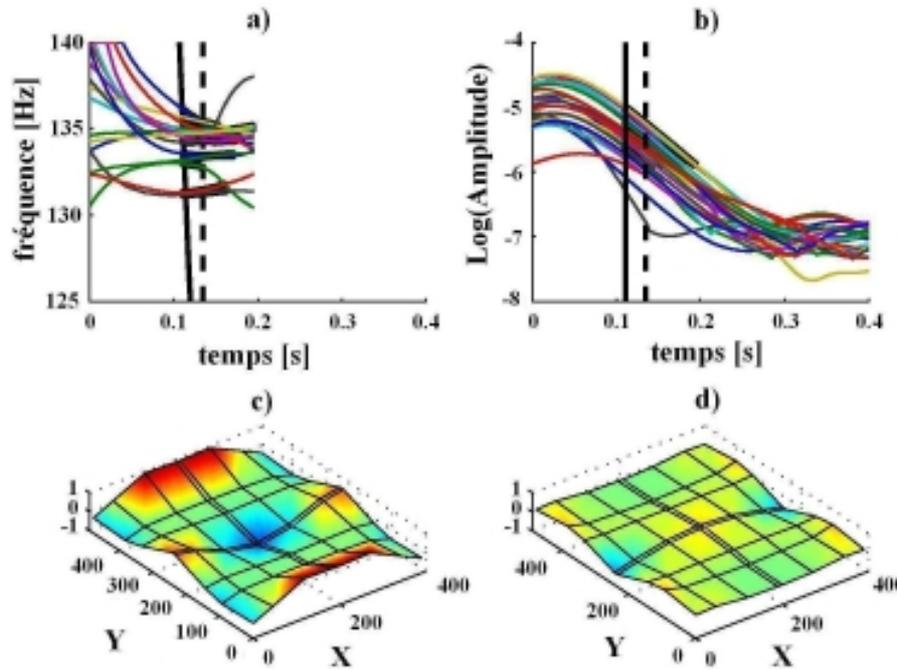


FIG. 3.10 – Plaque de plexiglass, **mode 8**. a) Arêtes extraites à partir des signaux enregistrés. b) Représentation logarithmique des amplitudes des squelettes, avec les intervalles de temps appropriés pour l'analyse modale. c) Partie réelle de la forme modale complexe identifiée. d) Partie imaginaire de la forme modale complexe identifiée.

choc induisant la vibration libre est imposé au point 1, le mode 8 est excité avec une amplitude beaucoup plus grande que celle du mode 7. Réciproquement, quand le choc est imposé au point 2, qui se trouve approximativement sur une ligne nodale du mode 8, seulement le mode 7 est excité à cette fréquence, et son identification devient facile (cf. Fig. (3.9)).

En outre, les modes 9 et 10 avec des fréquences plus élevées sont bien excités en imposant le choc au point 2. Par conséquent, les valeurs de la fréquence et du taux d'amortissement de ces modes ont été identifiées également à partir de ce deuxième procédé de mesure (voir les trois dernières colonnes du Tableau 3.2). Des indices de non-proportionnalité  $\tilde{I}_{np,ident}^{(j)}$  ont été calculés pour tous les modes de la plaque (cf. [EA07]).

### 3.5 Vibrations libres non linéaires et ondelettes

Il est fréquent qu'une structure industrielle ou civile sorte du domaine linéaire par exemple lorsque des phénomènes de glissement prennent naissance aux joints de la structure ou que le comportement du matériau rentre dans le domaine de plastification ou encore que des zones endommagées apparaissent. Négliger les effets non linéaires peut alors entraîner une grande différence entre le comportement réel de la structure et le comportement estimé par un modèle numérique, même pour des non linéarités faibles. Dans ce contexte, le but de l'identification dynamique non linéaire est de donner à l'ingénieur des outils théoriques et numériques pour tenter de localiser les non-linéarités et d'en caractériser leur nature et finalement de les identifier à partir de

l'auscultation vibratoire de la structure.

La transformation de Hilbert est un des premiers outils utilisés pour détecter et caractériser la présence de non-linéarités [Fel94]. Pour un système non linéaire, les fréquences modales peuvent varier avec le temps et sont appelées fréquences instantanées ; ce phénomène ne peut pas être correctement analysé par l'analyse de Fourier. Au contraire, les méthodes d'analyse temps-fréquence, notamment la TWV [Coh95] et la TOC définie précédemment n'ont pas cette restriction. Elles ont été récemment appliquées pour l'identification de structures non linéaires. La TWV a été utilisée par [GPS98] pour classifier le type de non-linéarités et identifier les paramètres instantanés d'un système non linéaire. Cette méthode nécessite un filtre passe-bande si plusieurs fréquences sont présentes dans le signal comme c'est le cas pour un système à plusieurs DdLs. En revanche, la TOC permet non seulement de capter l'évolution en temps de la fréquence et de l'amplitude mais joue aussi le rôle d'un filtre. Nous avons vu dans ce qui précède que l'emploi de la TOC facilite l'identification modale de systèmes linéaires et stationnaires et donne des résultats prometteurs lorsque deux modes propres sont couplés. Quelques résultats récents quant à l'identification des systèmes non linéaires par la TOC ont été présentés par [Sta98], [AL03], [LA03]. On note l'utilisation de plus en plus fréquente de la TOC pour analyser des réponses de systèmes non linéaires comme par exemple, dans le phénomène du pompage énergétique. Dans [GVK07], les auteurs appliquent la TOC, la transformation de Hilbert et la décomposition modale empirique aux réponses sous chocs du système constitué d'une barre à comportement linéaire reliée en bout à une masse très légère, par une liaison essentiellement non linéaire.

[AL03] proposent quatre indicateurs instantanés appliqués aux réponses libres d'un système faiblement non linéaire. Le principe de superposition modale n'est plus applicable lorsque le comportement de la structure est non linéaire.

Pour des systèmes non linéaires conservatifs, [SS90b] introduit le concept de modes normaux couplés MNC qui est développé par [BB99] pour analyser des systèmes non linéaires discretisés faiblement amortis. La réponse libre en déplacement  $u_k(t)$  au point de mesure  $k$  d'un système non linéaire conservatif à  $N$  DdLs peut être approchée par une combinaison linéaire d'un terme constant  $\Psi_{0k}$  qui représente la valeur moyenne de  $u_k(t)$  et de termes harmoniques dont l'amplitude dépend d'un vecteur lié à un « mode »  $\Psi_{rk}$  et la phase à la pulsation propre correspondante  $\Omega_r$  :  $u_k(t) = \Psi_{0k} + \sum_{r=1}^N \Psi_{rk} A_r \cos(\Omega_r t + \varphi_r(t))$  où  $\Psi_{0k}$  et  $\Psi_{rk}$  sont fonctions de toutes les amplitudes modales  $(A_1(t), A_2(t), \dots, A_N(t))$  et non pas uniquement de  $A_r(t)$ . Les termes de la somme sont encore des chirps  $u_{kr}(t)$  et  $u_k(t)$  peut encore se mettre sous

la forme :  $u_k(t) = \Psi_{0k} + \sum_{r=1}^N u_{kr}(t) = \Psi_{0k} + \sum_{r=1}^N \Upsilon^{(u_{kr})}(t) \cos(\alpha^{(u_{kr})}(t))$ . Bellizzi propose

de normaliser le vecteur  $\Psi_k$  des composantes  $\Psi_{rk}$  suivant la matrice de masse. Le MNC est donné par le couple  $(\Omega_k, \Psi_k) \in \mathbb{R}^+ \times \mathbb{R}$  et suite à la propriété de normalisation,  $A_k$  définit l'amplitude modale du  $k^e$  mode non linéaire couplé. L'application des techniques d'équilibrage harmonique [SS90a] fournit des équations algébriques non linéaires dont la résolution conduit au MNC comme une fonction du vecteur  $\mathbf{A}$  dont les composantes sont les amplitudes modales  $A_r$ . Dans [BGKM01], les amplitudes et phases instantanées sont extraites au moyen d'une définition des arêtes étendues à l'analyse continue en gaborettes et les MNC sont ensuite identifiés. La technique proposée ici est fondée sur le même concept de MNC mais utilise l'analyse en ondelettes continue.

On note  $k_0$  l'indice du point de mesure où la valeur moyenne sur  $t$  de l'amplitude  $\Upsilon^{(u_{k_0 r})}(t)$  est maximale. Les indicateurs sont appliqués sur les réponses accélérométriques  $\ddot{u}_k(t)$  au point de mesure  $k$ . Pour cela, on calcule les TOC des  $\ddot{u}_k(t)$  aux différents points de mesure et on extrait les différentes arêtes  $a_r^{(\ddot{u}_{kr})}(t)$  à l'aide des méthodes précédemment décrites et on obtient le squelette :  $T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{kr})}(t))$ . Les quatre indicateurs sont définis par :

$$\text{pour } k \geq 1$$

$$f_{kr}(t) = \frac{1}{2\pi} \frac{n}{a_r^{(\ddot{u}_{kr})}(t)} \quad (3.5.1)$$

$$r_{kr}(t) = \frac{\left| T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{kr})}(t)) \right| \left( a_r^{(\ddot{u}_{kr})}(t) \right)^2}{\left| T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{k_0 r})}(t)) \right| \left( a_r^{(\ddot{u}_{k_0 r})}(t) \right)^2} \quad (3.5.2)$$

$$d_{kr}(t) = \ln \left| T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{kr})}(t)) \right| + 2 \ln a_r^{(\ddot{u}_{kr})}(t) + C(n) \quad (3.5.3)$$

et pour  $k > 1$

$$\mu_{kr}(t) = \arg \left\{ T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{kr})}(t)) \right\} - \arg \left\{ T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{(k-1)r})}(t)) \right\} \quad (3.5.4)$$

où  $C(n) = -\ln(\pi) - (n+2)\ln(n) + n + \ln(n!)$ ;  $f_{kr}(t)$ ,  $r_{kr}(t)$ .

Comme  $\dot{\alpha}^{(u_{kr})}(t) = \frac{n}{a_r^{(\ddot{u}_{kr})}(t)}$  et

$T_{\psi_n}[\ddot{u}_k](t, a_r^{(\ddot{u}_{kr})}(t)) \approx -\pi \left( \dot{\alpha}^{(u_{kr})}(t) \right)^2 \frac{n^n e^{-n}}{n!} \Upsilon^{(u_{kr})}(t) e^{i\alpha^{(u_{kr})}(t)}$ , on déduit que :  $f_{kr}(t) \approx \frac{\dot{\alpha}^{(u_{kr})}(t)}{2\pi}$ ;  $r_{kr}(t) \approx \frac{\Upsilon^{(u_{kr})}(t)}{\Upsilon^{(u_{k_0 r})}(t)}$ ;  $\mu_{kr}(t) \approx \alpha^{(u_{kr})}(t) - \alpha^{(u_{(k-1)r})}(t)$  et  $d_{kr}(t) \approx \ln(\Upsilon^{(u_{kr})}(t))$ .

$f_{kr}(t)$  et  $r_{kr}(t)$  désignent respectivement la  $r^e$  fréquence instantanée et le  $r^e$  "mode" instantané (avec une amplitude positive).  $d_{kr}(t)$  et  $\mu_{kr}(t)$  caractérisent la dissipation d'énergie suivant la  $r^e$  composante :  $d_{kr}(t)$  la décroissance globale de l'énergie dans la réponse transitoire de la  $r^e$  composante et  $\mu_{kr}(t)$  caractérise la dissipation d'énergie entre deux points de la structure  $k$  et  $(k-1)$ .

Pour des systèmes linéaires et stationnaires avec amortissement visqueux proportionnel, les relations précédentes (3.5.1), (3.5.2), (3.5.3) et (3.5.4) s'écrivent :  $f_{kr}(t) = \frac{\omega_r}{2\pi}$

,  $r_{kr}(t) \approx \left| \frac{\phi_{kr}}{\phi_{k_0 r}} \right|$ ,  $d_{kr}(t) = \text{Log} |\phi_{kr}| + \text{Log} |G_r| - \xi_r \omega_{0r} t$  et  $\mu_{kj}(t) = \frac{\pi}{2} (\text{sgn}(\phi_{(k-1)r}) - \text{sgn}(\phi_{kr}))$ . Dans ce cas, pour le mode  $r$ ,  $f_{kr}(t)$ ,  $r_{kr}(t)$  et  $\mu_{kr}(t)$  sont constants;  $f_{kr}(t)$  désigne alors la  $r^e$  pseudo-fréquence propre et  $\mu_{kj}(t)$  redonne le signe de l'amplitude du mode. Finalement,  $d_{kj}(t)$  s'écrit sous la forme d'une droite avec une pente négative dont la valeur est proportionnelle au taux d'amortissement du  $r^e$  mode.

L'étude de l'évolution temporelle de ces indicateurs, principalement leur déviation par rapport au comportement linéaire avec amortissement proportionnel permet aux auteurs de détecter, localiser (si non-linéarité locale) et donner des informations utiles pour la caractérisation et l'identification futures des effets non linéaires.

Ces indicateurs ont été calculés sur les résultats de la campagne d'essais menée en 2001 par l'Université de Liège<sup>(3)</sup>. Le corps d'épreuve est constitué de l'assemblage de deux éléments : l'élément principal est une poutre en acier encastree à une de

<sup>3</sup>au sein du Laboratoire de Techniques Aéronautiques et Spatiales - Département Vibrations et Identification des Structures dirigé par le Professeur Jean-Claude Golinval.

ses extrémités et l'élément secondaire est une poutre mince en acier encastrée à une extrémité dans un bâti fixe et fixée à l'autre extrémité, à la poutre principale au moyen de vis (pour plus de détails sur l'essai et sur les caractéristiques géométriques et mécaniques des poutres voir [Tho03]). La poutre de l'élément principal est instrumentée de sept accéléromètres régulièrement espacés et est soumise à un choc par l'intermédiaire d'un marteau. On remarque six pics clairement visibles sur les FFTs des réponses accélérométriques dont les fréquences correspondantes en Hz sont :  $f_1 \approx 40$ ,  $f_2 \approx 123$ ,  $f_3 \approx 149.5$ ,  $f_4 \approx 207$ ,  $f_5 \approx 232$ ,  $f_6 \approx 395$ . L'étude est restreinte aux quatre arêtes situées au voisinage de 40, 123, 149,5 et 395 Hz ; elles ont été extraites avec des valeurs de  $Q$  respectivement égales à 17.33, 53.20, 64.69 et 170.69 (correspondant respectivement à  $n = 600$ , 5660, 8370 et 58270) et choisies de telle sorte que les écarts-types temporel et fréquentiel de la TOC soient sensiblement identiques pour le calcul d'extraction des quatre arêtes :  $\Delta t_r = a_r \Delta t_{\psi_n} = \frac{n}{2\pi f_r} \frac{1}{\sqrt{2n-1}} \approx 0.069$  s,  $\Delta f_r = \frac{\Delta \omega_{\psi_n}}{2\pi a_r} = \frac{f_r}{2n} \sqrt{2n+1} \approx 1.156$  Hz. Les fréquences instantanées  $f_{kr}(t)$  des trois premiers modes et du premier mode sur-harmonique sont tracées en Fig. (3.11).

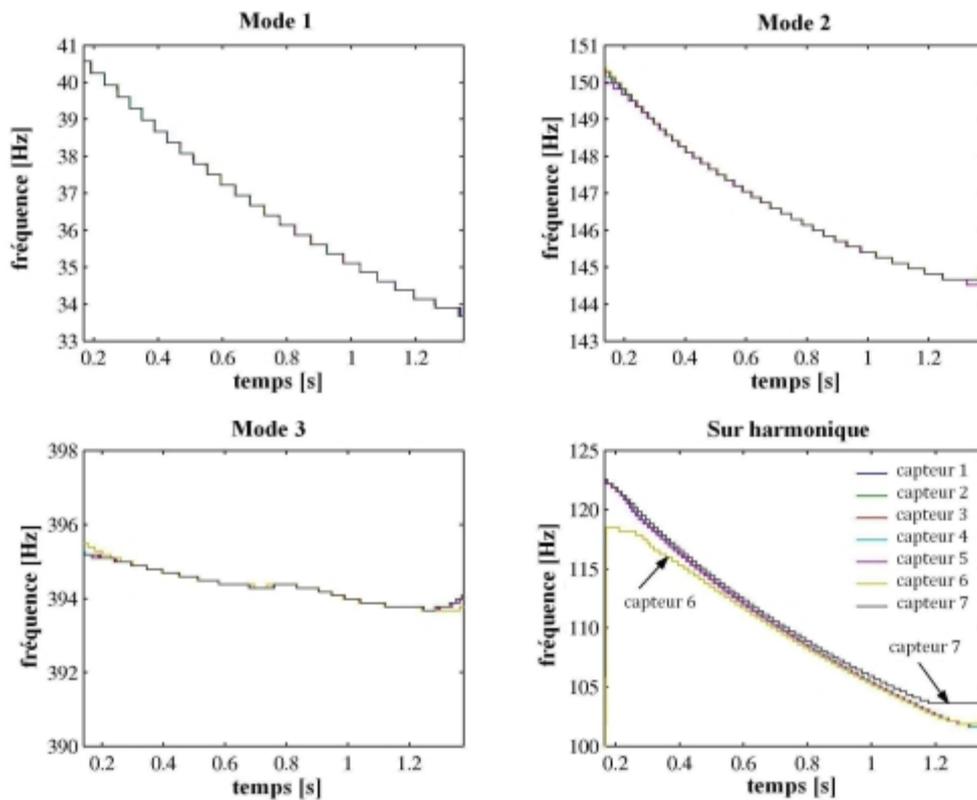


FIG. 3.11 – Poutre de Liège. Fréquences instantanées  $f_{kr}(t)$  des trois premiers modes et du premier mode sur-harmonique ( $r = 1, 2, 3$  et 6) pour les 7 capteurs ( $k = 1, \dots, 7$ )

Un comportement non linéaire de type Duffing est mis ensuite en évidence pour la « poutre » de Liège. Ainsi, la composante correspondant au sur-harmonique d'ordre 3 est nettement visible ; en effet, le rapport de la fréquence instantanée de cette composante sur celle de la composante fondamentale est pratiquement constant et égal à 3.

Ensuite, on s'intéresse à l'expression de l'approximation au premier ordre (méthode

KBM<sup>(4)</sup> de la fréquence et de l'amplitude instantanées des vibrations libres d'un oscillateur non linéaire avec une non-linéarité de type Duffing et un faible taux d'amortissement visqueux  $\xi$ . L'équation différentielle qui régit les vibrations d'un tel oscillateur se met sous la forme :  $\ddot{u}(t) + \omega^2 u = -\varepsilon [f_1(\dot{u}) + \psi_1(u)]$  avec :  $\psi_1(u) = \beta \frac{\omega^3}{h} u^3$ ,  $f_1(\dot{u}) = 2\omega \dot{u}$  et  $\varepsilon = \frac{h}{\omega}$  où la variable  $\varepsilon = \xi$  a une valeur « faible ». L'approximation au premier ordre par la méthode KBM<sup>(5)</sup> s'écrit sous la forme d'un signal modulé en amplitude et en phase :  $u_1(t) = a(t) \cos(\alpha(t))$  avec  $a(t) = a_0 e^{-ht}$  et  $\alpha(t) = \alpha_0 + \omega t + \beta \frac{3}{16h} \omega a_0^2 (1 - e^{-2ht})$  où  $a(0) = a_0$  et  $\alpha(0) = \alpha_0$ <sup>(6)</sup>. L'expression analytique de la première fréquence instantanée en fonction de l'amplitude instantanée est de forme parabolique. La première étape de la procédure d'identification consiste à extraire l'arête et la valeur de la TOC sur l'arête des signaux transitoires mesurés. Les résultats de l'extraction permettent de déduire la première fréquence instantanée du système en fonction de l'amplitude instantanée. A partir de cette fréquence et de son expression analytique obtenue par la méthode KBM, on identifie les paramètres de l'oscillateur avec une non-linéarité cubique. On note que l'erreur relative en norme  $L^2$  sur l'évolution temporelle de la fréquence instantanée est de l'ordre de  $10^{-3}$ . En conclusion, pour la première fréquence de résonance, le modèle avec un faible amortissement visqueux et une non linéarité en déplacement de type cubique semble convenir.

Ces résultats constituent un premier pas dans l'utilisation prometteuse de la TOC pour traiter les réponses dynamiques des structures non linéaires au niveau de la détection et de la caractérisation des effets non-linéaires. Pour ce qui concerne l'identification des paramètres de comportement, une modélisation préalable des effets non linéaires est nécessaire ; la TOC permettant un pré-traitement efficace et pertinent de l'information.

<sup>4</sup>La méthode Krylov Bogoliubov Mitropolsky (KBM) est basée sur la méthode de la variation des constantes pour les équations différentielles faiblement non linéaires. Elle est appliquée ici au cas des vibrations libres d'un oscillateur non linéaire : la réponse est choisie sous la forme d'un signal modulé en amplitude et en phase  $a(t)\cos(\omega t + \varphi(t))$  avec l'amplitude  $a(t)$  et la phase  $\varphi(t)$  qui varie « lentement ».

<sup>5</sup>L'approximation au premier ordre par la méthode de Krylov et Bogolyubov (KBM) est donnée par :  $u_1(t) = a(t) \cos(\alpha(t))$  avec lorsque  $t_0 = 0$ ,  $a(0) = a_0$  et  $\alpha(0) = \alpha_0$ .

$$t(a) = t_0 + \frac{1}{\varepsilon} \int_{a_0}^a \frac{da}{A_1(a)} = \frac{1}{\varepsilon} \int_{a_0}^a \frac{da}{A_1(a)}$$

$$\alpha(t) = \alpha_0 + \omega(t - t_0) + \varepsilon \int_{t_0}^t B_1(a) dt = \alpha_0 + \omega t + \varepsilon \int_0^t B_1(a) dt$$

$$A_1(a) = \frac{1}{2\pi\omega} \int_0^{2\pi} f_1(-\omega a \sin(\alpha)) \sin(\alpha) d\alpha \text{ ne dépend que de } f_1.$$

$$B_1(a) = \frac{1}{2\pi\omega a} \int_0^{2\pi} \psi_1(a \cos(\alpha)) \cos(\alpha) d\alpha \text{ ne dépend que de } \psi_1.$$

$${}^6 A_1(a) = \frac{1}{2\pi\omega} \int_0^{2\pi} f_1(-\omega a \sin(\alpha)) \sin(\alpha) d\alpha = \frac{-2\omega^2 a}{2\pi\omega} \int_0^{2\pi} \sin^2(\alpha) d\alpha$$

$$A_1(a) = -\omega a \text{ et finalement } t(a) = -\frac{1}{\omega\varepsilon} \int_{a_0}^a \frac{da}{A_1(a)} = -\frac{1}{h} \text{Ln} \left( \frac{a}{a_0} \right). \text{ D'où } a(t) = a_0 e^{-ht}.$$

et encore :

$$\psi_1(a \cos(\alpha)) = \beta \frac{\omega^3 a^3}{4h} (\cos(3\alpha) + 3 \cos(\alpha))$$

$$B_1(a) = \frac{1}{2\pi\omega a} \beta \frac{\omega^3 a^3}{4h} \int_0^{2\pi} (\cos(3\alpha) + 3 \cos(\alpha)) \cos(\alpha) d\alpha = \beta \frac{3}{8h} \omega^2 a^2 \text{ avec } a(t) = a_0 e^{-ht}$$

$$B_1(a) = \beta \frac{3}{8h} \omega^2 a_0^2 e^{-2ht}$$

$$\alpha(t) = \alpha_0 + \omega t + \varepsilon \int_0^t B_1(a) dt = \alpha_0 + \omega t + \beta \frac{3}{8} \omega a_0^2 \int_0^t e^{-2ht} dt$$

$$= \alpha_0 + \omega t + \beta \frac{3}{16h} \omega a_0^2 (1 - e^{-2ht})$$



# Bibliographie

- [AHC00] P. Argoul, S. Hans, F. Conti, and C. Boutin. Time-frequency analysis of free oscillations of mechanical structures. application to the identification of the mechanical behavior of buildings under shocks. In *European COST F3 Conference on System Identification and Structural Health Monitoring*, volume 1, pages 283–292, Madrid, Spain, 2000.
- [AL03] P. Argoul and T-P. Le. Instantaneous indicators of structural behaviour based on continuous cauchy wavelet transform. *Mech. Syst. Signal Proc.*, 17(1) :243–250, 2003.
- [ANL08] P. Argoul, T.M. Nguyen, and T-P. Le. Approche modale par analyse continue en ondelettes. application à l'identification modale de bâtiments. *Bulletin des laboratoires des ponts et chaussées*, (Numéro spécial "Risque sismique"), 2008.
- [Asm97] J.C. Asmussen. *Modal analysis based on the random decrement technique - application to civil engineering structures*. PhD thesis, University of Aalborg, 1997.
- [BB99] S. Bellizzi and R. Bouc. Analysis of multi-degree of freedom strongly non-linear mechanical systems with random input - part i : non-linear modes and stochastic averaging. *Prob. Eng. Mech.*, 14(3) :229–244, July 1999.
- [BGKM01] S. Bellizzi, P. Guillemain, and R. Kronland-Martinet. Identification of coupled non-linear modes from free vibration using time-frequency representations. *J. Sound Vib.*, 243(2) :191–213, 2001.
- [Bui93] H.D. Bui. *Introduction aux problèmes inverses en mécanique des matériaux*. Collection de la Direction des Etudes et Recherches d'Electricité de France. Eds Eyrolles, 1993.
- [CHT98] R. Carmona, W.-L. Hwang, and B. Torrèsani. *Practical time-Frequency analysis*. Academic Press, New York, 1998.
- [CMF99] E. Chassande-Mottin and P. Flandrin. On the time-frequency detection of chirps. *Appl. Comput. Harmon. Anal.*, 6(2) :252–281, 1999.
- [Coh95] L. Cohen. *Time-frequency analysis*. Prentice-Hall Signal Processing Series. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1995.
- [EA07] S. Erlicher and P. Argoul. Modal identification of linear non-proportionally damped systems by wavelet transform. *Mech. Syst. Signal Proc.*, 21(3) :1386–1421, 2007.
- [Fel94] M. Feldman. Non-linear free vibration analysis using Hilbert transform i. free vibration analysis method 'freevib'. *Mech. Syst. Signal Proc.*, 8(2) :119–127, 1994.

- [Fla89] P. Flandrin. Représentations temps-fréquence des signaux non stationnaires. *Traitement Signal*, 6(2) :89–101, 1989.
- [GM84] A. Grossmann and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shapes. *SIAM J. Math. Anal.*, 15 :723–736, 1984.
- [GPS98] L. Galleani, L. Lo Presti, and A. De Stefano. A method for nonlinear system classification in the time-frequency plane. *Signal Processing*, 65 :147–153, 1998.
- [GVK07] F. Georgiades, A. F. Vakakis, and G. Kerschen. Broadband passive targeted energy pumping from a linear dispersive rod to a lightweight essentially non-linear end attachment. 42 :773–788, 2007.
- [LA03] T.-P. Le and P. Argoul. Utilisation de la transformation en ondelettes pour l'identification de systèmes mécaniques non linéaires en vibration. In *16ème Congrès Français de Mécanique*, Nice - France, 2003. 6 pages sur CD-ROM.
- [LA04] T.P. Le and P. Argoul. Continuous wavelet transform for modal identification using free decay response. *J. Sound Vib.*, 277(1-2) :73–100, 2004.
- [Le03] T.-P. Le. *Auscultation dynamique des structures à l'aide de l'analyse continue en ondelettes*. PhD thesis, thèse ENPC, E.D. MODES, 2003.
- [LTB04] J. Lardies, M.N. Ta, and M. Berthillier. Modal parameter estimation based on the wavelet transform of output data. *Archive of Applied Mechanics*, 73 :718–733, 2004.
- [Mal00] S. Mallat. *Une exploration des signaux en ondelettes*. Publications Ecole Polytechnique. Ellipses Diffusion, 2000.
- [MSH<sup>+</sup>98] N.M.M Maia, J.M.M. Silva, N. He, N.A. Lieven, R.M. Lin, G.W. Skingle, W.-M. To, and A.P.V. Urgueira. *Theoretical and Experimental Modal Analysis*. Research Studies Press Ltd., Hertfordshire, 1998.
- [NAC05] T.-M. Nguyen, P. Argoul, and R. Ceravolo. Wavelet analysis of the structural response under ambient excitation for modal identification. In *Proc. of Eurodyn 2005*, pages 107–112, Paris - France, 4-7 Sept. 2005.
- [Ngu07] T.-M. Nguyen. *Dynamique non linéaire des systèmes mécaniques couplés : Réduction de modèle et identification*. 181 pages, ENPC, E.D. MODES, 2007.
- [RFGP97] M. Ruzzene, A. Fasana, L. Garibaldi, and B. Piombo. Natural frequency and damping identification using wavelet transform : application to real data. *Mech. Syst. Signal Proc.*, 11(2) :207–218, 1997.
- [SS90a] W. Szemplinska-Stupnicka. *Behavior of nonlinear vibrating systems. Vol. I : Fundamental concepts and methods : Applications to single-degree of freedom systems*, volume Vol 12 of *Mechanics : Dynamical Systems*. Kluwer Academic Publishers, Dordrecht, 1990.
- [SS90b] W. Szemplinska-Stupnicka. *Behavior of nonlinear vibrating systems. Vol. II : Advanced concepts and applications to multi degrees of freedom systems*, volume Vol 13 of *Mechanics : Dynamical Systems*. Kluwer Academic Publishers, Dordrecht, 1990.

- [SSB03] J. Slavic, I. Simonovski, and M. Boltezar. Damping identification using a continuous wavelet transform : applicaton to real data. *J. Sound Vib.*, 262 :291–307, 2003.
- [Sta98] W.J. Stazewski. Identification of non linear systems using multi-scale ridges and skeletons of the wavelet transform. *J. Sound Vib.*, 214(4) :639–658, 1998.
- [Tho03] F. Thouverez. Presentation of the ECL benchmark. *Mech. Syst. Signal Proc.*, 17(1) :195–202, 2003.
- [YD00] H.P. Yin and D. Duhamel. Finite difference formulation for modal parameter identification. *JSV*, 231 :259–275, 2000.
- [YDA04] H.P. Yin, D. Duhamel, and P. Argoul. Natural frequencies and damping estimation using wavelet transform of a frequency response function. *J. Sound Vib.*, 271 :999–1014, 2004.



## Chapitre 4

# Modélisation et identification de systèmes discrets hystérétiques en dynamique

*Claude-Henri Lamarque,  
Frédéric Bernardin,  
Jérôme Bastien,  
Michelle Schatzman.*

### 4.1 Introduction

Cette contribution présente des modèles permettant de décrire des phénomènes avec hystérésis. Nous proposons une revue bibliographique rapide et nous abordons de façon un peu plus détaillée une description particulière de la dynamique de systèmes mécaniques discrets exhibant un comportement hystérétique. Nous expliquons sur quelques exemples très simples le formalisme retenu pour décrire ces systèmes. Nous rappelons rapidement les résultats mathématiques généraux disponibles pour ces systèmes. Nous montrons aussi qu'il est possible d'identifier les paramètres d'un modèle à partir de données discrètes (mesures).

Les comportements hystérétiques peuvent être liés à des processus de natures très diverses. Ils se produisent et sont observés en mécanique, en physique, aussi bien qu'en biologie, ou en sciences sociales.

Il existe dans la littérature de nombreuses approches pour la modélisation d'hystérésis : approches boîte noire (réseaux de neurones), approches analytiques avec interpolation des sorties comme fonctions des données (exponentielles, hyperboles, polynômes, arc-tangente etc.), ou approche par modèles différentiels supposant que les entrées sont connues.

Des modèles mathématiques sont discutés dans les ouvrages [May91] ou [Vis94]. Les modèles décrivant des systèmes avec hystérésis présentent souvent des discontinuités voire des irrégularités plus fortes encore.

On trouvera des exemples de systèmes hystérétiques abordés du point de vue modélisation, expérimentation, analyse des réponses pour diverses applications dans

[MCV02], [Ber01] et [KO93] (modèles à variables internes avec diverses significations physiques : par exemple approche de Ziegler-Green-Naghdi avec une variable interne), [CSB03] (modèle statique, expressions analytiques et interpolation par arc-tangente), [KS00] (modèle de Preisach modifié), [KRB<sup>+</sup>00] (modèles de Coleman-Hodgdon, Jiles et Atherton), [KFMS01] (matériaux ferro-magnétiques et alliages à mémoire de forme, modèle de Preisach et opérateurs hystérétiques appropriés interpolés par moindres carrés), [MDDWM03] (modèle par réseaux de neurones alternative aux calculs par modèles de Preisach), [OD02] (alliages à mémoire de forme, approches phénoménologique), [OHB03] (limites des modèles de Preisach, Coleman-Hodgdon, Jiles et Atherton, dues aux matériaux, histoire complète des matériaux mal prise en compte, modèle de Preisach partiel), [Sap03, SF03] (fluides magnéto- ou électrorhéologiques et matériaux intelligents amortissants, modèles avec petit nombre de paramètres de type Bouc-Wen ou Spencer), [Kle97] (modèle physique de mémoire non locale pour des systèmes à hystérésis).

Des travaux récents étudient la dynamique d'oscillateurs hystérétiques (modèles de Masing et Bouc-Wen [CM96, LYL01, LZW02, VN02, LV04, BMMD07], système à un degré de liberté [Ket98], système de type Masing à deux degrés de liberté [MCV02]) en utilisant des méthodes analytiques et/ou numériques, en allant jusqu'aux comportements chaotiques ([CMKM86, CKM86, Mad93, PS90, LZW02, Awr03, AD03, ADG04]).

La question de modèles généraux multi-application reste ouverte pour le large spectre des phénomènes hystérétiques. La présence d'échelles différentes dans un système physique et leur lien avec les phénomènes hystérétiques laissent aussi des problèmes ouverts dans l'écriture d'un formalisme général.

En ce qui concerne les modèles mécaniques non linéaires, les ouvrages de Brogliato [Bro96] et Palmov [Pal98] contiennent de nombreux exemples. Nous sommes intéressés ici par ceux qui impliquent des termes de friction [Sha86].

De nombreux travaux ont été consacrés aux modèles à nombre de degrés de liberté fini soumis à des sollicitations dynamiques

Certains travaux apportent des résultats d'existence et d'unicité ([Sch78, JP85, LMM95, MF95, MF96a, MF96b, MM94, TPSL97]). D'autres ne s'occupent que des comportements physiques : lois de frottement provenant d'expériences ([AF90], [FB95]), phénomène de "stick-slip" [AD90a, AD90b, BCPR95, IP93, PW81, Ste92], étude par des méthodes analytiques ([CV95]). Des travaux décrivent expérimentations, identification et modélisation du frottement ([DS83b], [DS83a] et [TC96]). Des simulations numériques fondées sur des schémas classiques sont proposées dans ([ST97, ST96, Ste96]).

Dans ([DL72]) les auteurs envisagent des modèles avec termes élastiques linéaires et frottement. Dans ce cas, l'existence et l'unicité sont obtenues pour le problème mécanique via des inégalités variationnelles. Discrétisées, celles-ci débouchent sur un schéma numérique tel que la convergence de la solution approchée vers la solution exacte a été prouvée ([GLT76a]) et [GLT76b]). Néanmoins ces travaux se placent dans le cadre de problèmes stationnaires en temps.

A part quelques travaux comme ceux de Monteiro Marques ([MM94]), peu d'articles abordent la question de la convergence de la solution approchée en dynamique.

Dans le cadre de l'étude de la micro-plasticité, Sidoroff et al. ([FS89]) examinent le modèle classique de Masing stationnaire avec un nombre fini ou infini de degrés de liberté.

Dans une série de papiers, nous nous sommes intéressés à la description de modèles hystérétiques de type élastoplastiques, par des inclusions différentielles [SBL99,

BSL00, BSL02, LBB05], contenant éventuellement un terme de retard [LBH03, BL05, LBH05] ou stochastique [Ber03, BSL04, Ber04] dans la lignée du type de modèle précédent, mais en dynamique.

## 4.2 Briques et modèles élémentaires

Les modèles considérés dans tout cet article seront construits en assemblant des briques élémentaires : points matériel (masses), ressorts, dashpots, et éléments non réguliers de type frottement sec dit de Saint-Venant. Dans cette section, nous présentons des modèles simples à un point matériel dans deux configurations de base : la mise en série ou l'assemblage en parallèle. A cette occasion, nous rappelons les lois de comportements des briques élémentaires menant naturellement au formalisme d'inclusions différentielles.

### 4.2.1 Assemblage en parallèle d'un ressort et d'un élément de Saint-Venant

Un point matériel de masse  $m$ , connecté en parallèle à un ressort de raideur  $k$  et un élément de St-Venant avec un seuil  $\alpha$ , est soumis à une force extérieure  $F$  ; soit  $x$  l'abscisse de ce point. Ce système mécanique académique est régi par le modèle

$$m\ddot{x} + kx + \alpha\sigma(\dot{x}) \ni F, \quad (4.2.1)$$

avec les conditions initiales

$$x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0; \quad (4.2.2)$$

où le graphe  $\sigma$  est défini par(cf Fig. 4.1a) :

$$\sigma(x) = \begin{cases} -1 & \text{si } x < 0, \\ 1 & \text{si } x > 0, \\ [-1, 1] & \text{si } x = 0. \end{cases} \quad (4.2.3)$$

Nous aurions pu ajouter un terme d'amortissement, ce que nous allons faire dans le modèle élémentaire suivant.

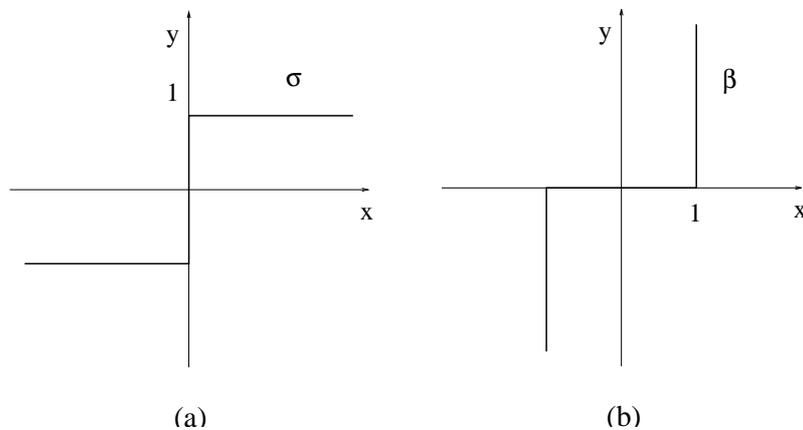


FIG. 4.1 – Tracé des graphes  $\sigma$  (a) et  $\beta$  (b).

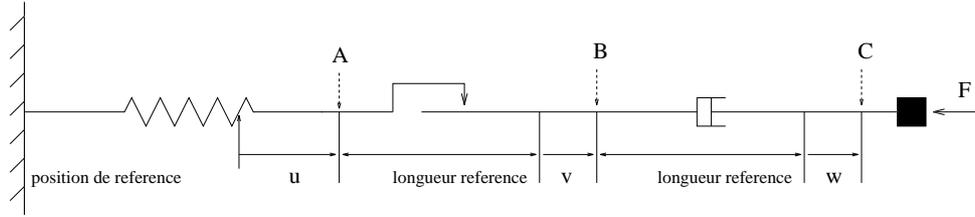


FIG. 4.2 – Un modèle viscoélastoplastique.

#### 4.2.2 modèle élasto-plastique et modèle de Prandtl

Précisons le modèle viscoélastoplastique représenté en Fig. 4.2 : un point matériel de masse  $m$  est cette fois connecté en série avec un ressort de raideur  $k$ , un élément de St-Venant de seuil  $\alpha$  et un dashpot de coefficient  $c$ , soumis à une force extérieure  $F$ . Appelons  $u$  le déplacement de l'extrémité  $A$  du ressort par rapport à une position de référence,  $v$  la différence entre  $AB$  et une longueur de référence, et  $w$  la différence entre  $BC$  et une longueur de référence.  $x$  désigne l'abscisse du point matériel de masse  $m$  par rapport à une position de référence, de sorte que :

$$x = u + v + w. \quad (4.2.4)$$

Soit  $f$  la force exercée par  $A$  sur  $B$ . Supposons ici que  $c \neq 0$  et  $\alpha \neq 0$ . La loi de comportement du ressort s'écrit

$$f = -ku, \quad (4.2.5)$$

celle de l'élément de St-Venant

$$f \in -\alpha\sigma(\dot{v}), \quad (4.2.6)$$

et celle du dashpot :

$$f = -c\dot{w}. \quad (4.2.7)$$

Le théorème fondamental de la dynamique et (11.4.2) se résument à :

$$m\ddot{x} = F - ku. \quad (4.2.8)$$

Il est possible de reformuler (4.2.4), (11.4.2), (4.2.6), (4.2.7) et (11.4.5) pour avoir une meilleure expression de la structure mathématique du problème. La différentiation de (4.2.4) en temps, et (11.4.2), (4.2.6), (4.2.7) donnent

$$ku \in \alpha\sigma(\dot{x} - \dot{u} - ku/c). \quad (4.2.9)$$

L'équation (4.2.9) implique la contrainte

$$ku \in [-\alpha, \alpha]. \quad (4.2.10)$$

En introduisant le graphe  $\beta$  inverse du graphe  $\sigma$  par (cf Fig. 4.1b) :

$$\beta(x) = \begin{cases} \emptyset & \text{si } x \in ]-\infty, -1[ \cup ]1, +\infty[, \\ \{0\} & \text{si } x \in ]-1, 1[, \\ \mathbb{R}_- & \text{si } x = -1, \\ \mathbb{R}_+ & \text{si } x = 1. \end{cases} \quad (4.2.11)$$

l'équation (4.2.9) est équivalente à

$$\dot{u} + \beta(ku/\alpha) \ni \dot{x} - ku/c, \quad (4.2.12)$$

avec des données initiales

$$x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0, \quad u(0) = u_0. \quad (4.2.13)$$

Le système mécanique de la Fig. 4.2 est modélisé par (11.4.5), (4.2.12) et (4.2.13). Réciproquement, connaissant  $(u, x)$  on retrouve  $(v, w)$  via (4.2.4) et (4.2.7). En posant

$$\eta = \alpha/k \text{ et } y = \dot{x},$$

le système (11.4.5), (4.2.12) et (4.2.13) est équivalent à

$$\begin{cases} \dot{x} = y, \\ \dot{y} = (F - ku)/m, \\ \dot{u} + \beta(u/\eta) \ni y - ku/c, \end{cases} \quad (4.2.14)$$

avec les données initiales

$$x(0) = x_0, \quad y(0) = y_0, \quad u(0) = u_0 \in [-\eta, \eta]. \quad (4.2.15)$$

Le cas limite  $c \rightarrow +\infty$  pour (4.2.14) et (4.2.15), correspond au modèle rhéologique de Prandtl, gouverné par l'équation (4.2.15) et

$$\begin{cases} \dot{x} = y, \\ \dot{y} = (F - ku)/m, \\ \dot{u} + \beta(u/\eta) \ni y. \end{cases} \quad (4.2.16)$$

Dans ce cas, le modèle de la Fig. 4.2 correspond à l'assemblage d'un ressort avec un élément de St-Venant (cf Fig. 4.3).

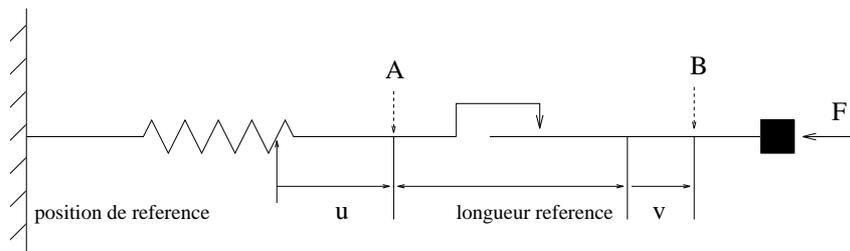


FIG. 4.3 – modèle de Prandtl.

Avec ces briques élémentaires, de nombreux assemblages sont possibles, impliquant un ou plusieurs points matériels. Comme dans les assemblages classiques série ou parallèle, un même comportement peut s'obtenir par différents modèles équivalents. Nous décrivons maintenant quelques classes de systèmes plus généraux, fondés sur deux briques élémentaires : d'abord, les éléments  $P_i$  (cf Fig. 4.4a) qui consistent en l'assemblage en parallèle d'un ressort de raideur  $k_i > 0$  et d'un élément de St-Venant de seuil  $\alpha_i > 0$ . La loi de comportement de  $P_i$  s'écrit :

$$f_i = -k_i u_i \text{ et } g_i \in -\alpha_i \sigma(\dot{u}_i). \quad (4.2.17)$$

Les éléments  $S_i$  (cf Fig. 4.4b) consistent en l'assemblage en série d'un ressort de raideur  $k_i > 0$  et d'un élément de St-Venant de seuil  $\alpha_i > 0$ , avec une loi de comportement décrite par :

$$f_i = -k_i u_i \in -\alpha_i \sigma(\dot{v}_i). \quad (4.2.18)$$

Nous supposons qu'il n'y a pas de terme résiduel de déformation lorsque la force extérieure  $f_i$  est nulle i.e. selon le modèle décrit en section 4.2.2, la position de référence choisie pour le point matériel et l'élément de St-venant est telle que  $f_i$  est nulle lorsque  $u_i$  vaut zéro.

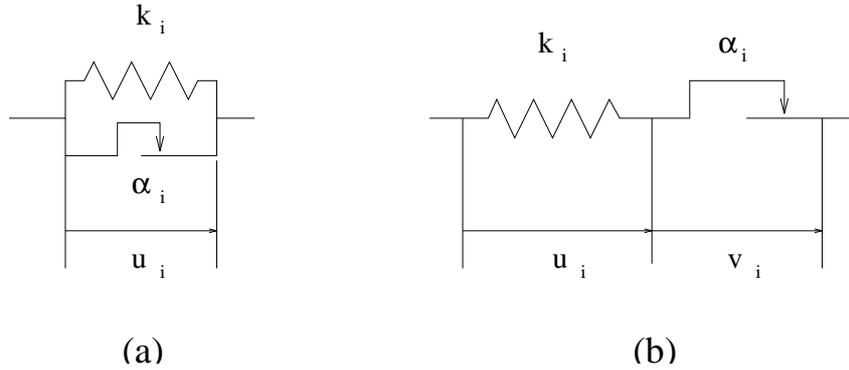


FIG. 4.4 – Élément  $P_i$  (a) et  $S_i$  (b).

### 4.2.3 Assemblage de $n$ éléments $S_i$ en parallèle avec une masse

Cette combinaison est décrite sur la Fig. 4.5a, avec  $x$  l'abscisse du point matériel et  $m$  sa masse.

En imitant la modélisation utilisée pour les exemples élémentaires, on montre que ce système mécanique est gouverné par les inclusions différentielles

$$\begin{cases} \dot{x} = y, \\ \dot{y} = F/m - (1/m) \sum_{i=1}^n k_i u_i, \\ \forall i \in \{1, \dots, n\}, \quad \dot{u}_i + \beta(u_i/\eta_i) \ni y; \end{cases} \quad (4.2.19)$$

avec

$$y = \dot{x} \text{ et } \forall i \in \{1, \dots, n\}, \quad \eta_i = \alpha_i/k_i, \quad (4.2.20)$$

et les données initiales

$$x(0) = x_0, \quad y(0) = y_0, \quad \forall i \in \{1, \dots, n\}, \quad u_i(0) = u_{i,0} \in [-\eta_i, \eta_i]. \quad (4.2.21)$$

Ce système est appelé système rhéologique de Prandtl généralisé.

En faisant tendre l'un des  $\alpha_i$  vers l'infini, l'un des éléments  $S_i$  devient un ressort pur, ce qui rigidifie le modèle :

$$\begin{cases} \dot{x} = y, \\ \dot{y} = F/m - (k_0/m)x - (1/m) \sum_{i=1}^n k_i u_i, \\ \forall i \in \{1, \dots, n\}, \quad \dot{u}_i + \beta(u_i/\eta_i) \ni y; \end{cases} \quad (4.2.22)$$

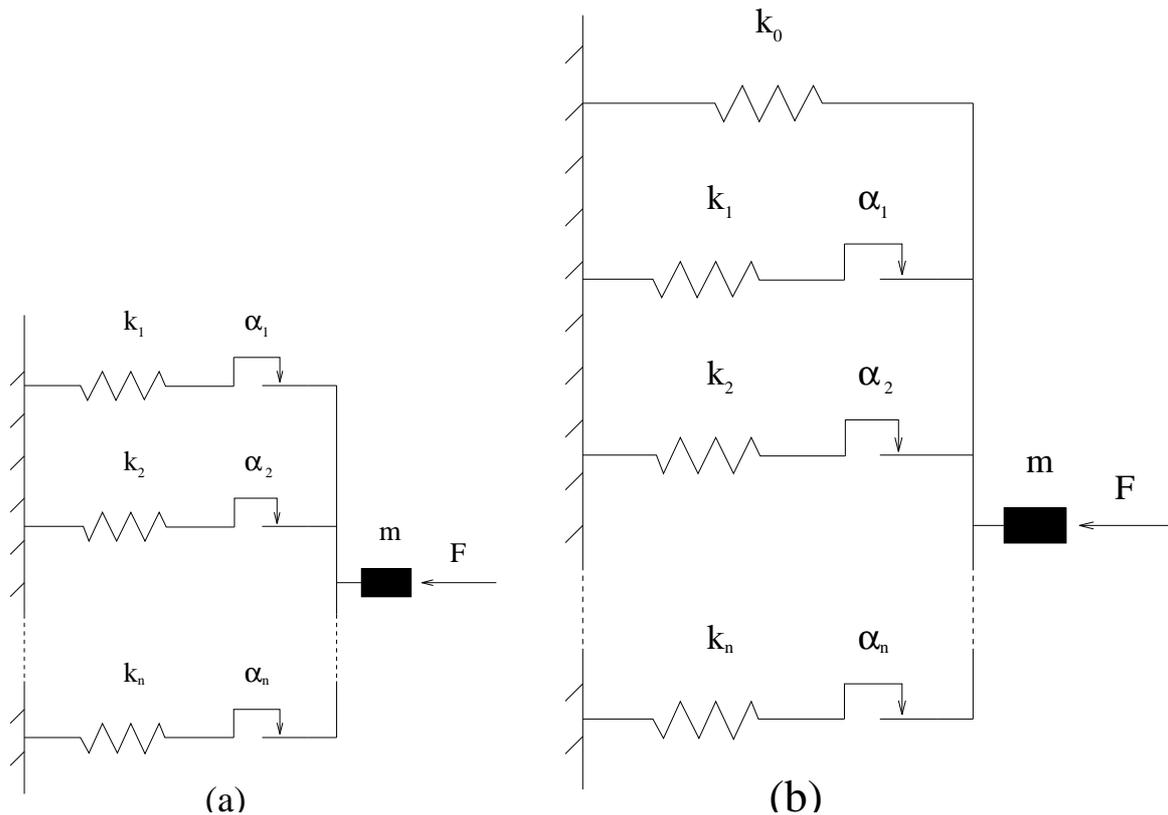


FIG. 4.5 – modèle rhéologique de Prandtl (a) et modèle de Prandtl généralisé avec un raidisseur linéaire (b).

Le système mécanique correspondant est présenté sur la Fig. 4.5b (avec  $n$  éléments  $S_i$  et un ressort) : il s'agit du modèle de Masing discret.

#### 4.2.4 modèles d'assemblage en série : $n - S_i$ et $n$ masses

Présentons maintenant un exemple avec plus d'un point matériel : la classe des modèles avec  $n$  éléments  $S_i$  et  $n$  points matériels connectés en série tels que proposés en Fig. 4.6 ; pour  $i = 1, \dots, n$ ,  $x_i$  est l'abscisse de la masse  $m_i$ .

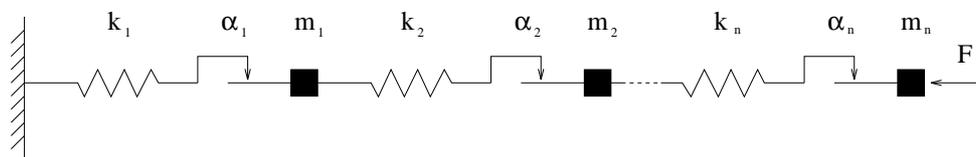


FIG. 4.6 – Association en séries de  $n$  éléments  $S_i$  et de  $n$  points matériels.

A l'aide des notations

$$\forall i \in \{1, \dots, n\}, \quad y_i = \dot{x}_i \text{ et } \eta_i = \alpha_i/k_i,$$

nous obtenons le modèle

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \dot{x}_1 = y_1, \\ \dot{y}_1 = (-k_1 u_1 + k_2 u_2)/m_1, \\ \dot{u}_1 + \beta(u_1/\eta_1) \ni y_1, \end{array} \right. \\ \forall i \in \{2, \dots, n-1\}, \left\{ \begin{array}{l} \dot{x}_i = y_i, \\ \dot{y}_i = (-k_i u_i + k_{i+1} u_{i+1})/m_i, \\ \dot{u}_i + \beta(u_i/\eta_i) \ni y_i - y_{i-1}, \end{array} \right. \\ \left\{ \begin{array}{l} \dot{x}_n = y_n, \\ \dot{y}_n = (-k_n u_n + F)/m_n, \\ \dot{u}_n + \beta(u_n/\eta_n) \ni y_n - y_{n-1}, \end{array} \right. \end{array} \right. \quad (4.2.23)$$

avec les conditions initiales

$$\forall i \in \{1, \dots, n\} \quad x_i(0) = x_{i,0}, \quad y_i(0) = y_{i,0}, \quad u_i(0) = u_{i,0} \in [-\eta_i, \eta_i]. \quad (4.2.24)$$

#### 4.2.5 Association en séries de $n$ éléments $P_i$ et de $n$ masses

On peut aussi poursuivre le jeu en examinant le système de la Fig. 4.7 où pour  $i = 1, \dots, n$ ,  $x_i$  est l'abscisse de la masse  $m_i$ .

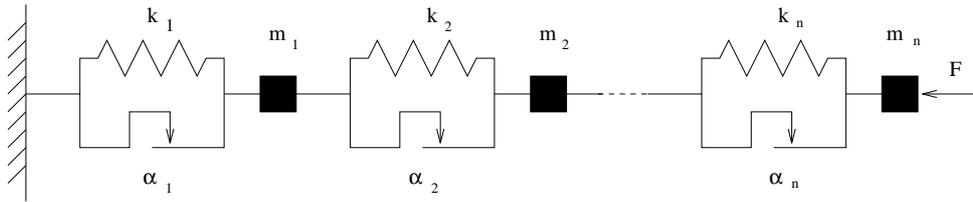


FIG. 4.7 – Association en séries de  $n$  éléments  $P_i$ -éléments et de  $n$  masses.

Le modèle s'écrit :

$$\left\{ \begin{array}{l} \dot{X} = Y, \\ \dot{Y} + D^{-1} A(Y) \ni D^{-1} H - D^{-1} K X, \end{array} \right. \quad (4.2.25)$$

avec les données initiales

$$X(0) = X_0 \text{ et } Y(0) = Y_0, \quad (4.2.26)$$

et les notations

$$X = (x_1, \dots, x_n), \quad Y = (y_1, \dots, y_n), \quad D = \text{diag}(m_1, \dots, m_n),$$

$$H = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ F \end{pmatrix} \text{ et } K = \begin{pmatrix} k_1 + k_2 & -k_2 & 0 & \dots & 0 & 0 \\ -k_2 & k_2 + k_3 & -k_3 & \dots & 0 & 0 \\ & \ddots & \ddots & \ddots & & \\ 0 & \dots & -k_i & k_i + k_{i+1} & -k_{i+1} & \dots \\ & & & \ddots & \ddots & \ddots \\ 0 & & \dots & 0 & -k_n & k_n \end{pmatrix}.$$

Le formalisme mathématique fait apparaître la généralisation d'une fonction croissante sous la forme de ce que l'on appelle un opérateur maximal monotone ([Bré73])  $A$ , défini ici sur  $\mathbb{R}^n$  par : quel que soit  $Y = (y_1, \dots, y_n) \in \mathbb{R}^n$ ,

$$A(Y) = \left\{ (g_i - g_{i+1})_{1 \leq i \leq n} : \begin{array}{l} g_1 \in \alpha_1 \sigma(y_1), \\ \forall i \in \{2, \dots, n\}, \quad g_i \in \alpha_i \sigma(y_i - y_{i-1}), \quad g_{n+1} = 0 \end{array} \right\}.$$

#### 4.2.6 Association en séries de $n$ éléments $P_i$ , d'un ressort et d'un point matériel

Le schéma de la Fig. 4.8 correspond à un modèle de Masing discret avec en série,  $n$  éléments  $P_i$  et un ressort. En statique, il est équivalent à un assemblage en parallèle de  $n$  éléments  $P_i$  et d'un ressort comme en Fig. 4.5b (cf [FS89]).

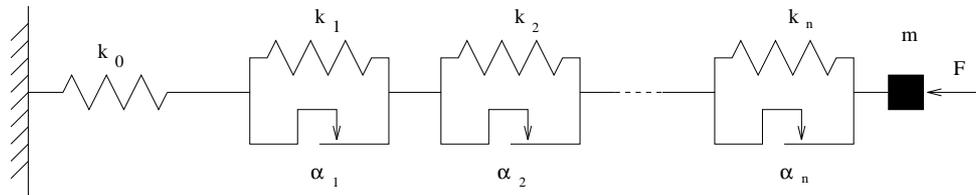


FIG. 4.8 – Association en séries de  $n$  éléments  $P_i$ , d'un ressort et d'une masse.

Soit  $k_0 > 0$  et  $x$  l'abscisse de la masse  $m$ . En introduisant des fonctions  $v_1, \dots, v_n$  de la même façon que le  $v$  du modèle du paragraphe (4.2.2), nous obtenons la mise en équations :

$$\begin{cases} \dot{x} = y, \\ \dot{y} = F/m + \gamma \left( \sum_{i=1}^n v_i/k_i - x \right), \\ \dot{V} + KB(V) \ni -k_0 y W, \end{cases} \quad (4.2.27)$$

avec les conditions initiales

$$x(0) = x_0, \quad y(0) = y_0 \text{ et } \forall i \in \{1, \dots, n\}, \quad v_i(0) = v_{i,0} \in [-\alpha_i, \alpha_i]. \quad (4.2.28)$$

$\gamma$  est donné par

$$\gamma = \left( m \sum_{i=0}^n \frac{1}{k_i} \right)^{-1},$$

et les vecteurs  $V$  et  $W$  par :

$$V = (v_1, \dots, v_n) \text{ et } W = (1, \dots, 1).$$

Les coefficients de la matrice  $K$  valent

$$\forall i, j \in \{1, \dots, n\}, \quad K_{i,j} = \begin{cases} k_0 + k_i & \text{si } i = j, \\ k_0 & \text{si } i \neq j. \end{cases}$$

Cette matrice est symétrique, positive, définie. L'opérateur maximal monotone  $B$  est défini sur  $\mathbb{R}^n$  par

$$\forall X = (x_1, \dots, x_n) \in \mathbb{R}^n, \quad B(X) = \beta(x_1/\alpha_1) \times \beta(x_2/\alpha_2) \times \dots \times \beta(x_n/\alpha_n).$$

### 4.2.7 modèles avec effets différés

Pour prendre en compte des effets différés d'origine variée (rhéologique, symétries), il est parfois nécessaire d'ajouter dans la description de la dynamique du modèle un terme dépendant d'un ou plusieurs états, ou de l'histoire du déplacement ou de la vitesse. Par exemple, pour un modèle de type (4.2.1), on obtient un modèle de type ([LBH03]) :

$$\begin{cases} m\ddot{x}(t) + c\dot{x}(t) + kx(t) + g(\dot{x}(t - \tau), x(t - \tau), t) + \alpha\sigma(\dot{x}(t)) \ni F(t), & \text{pour } t \in [0, T] \\ x_0, \dot{x}_0 \text{ fonctions données sur } [-\tau, 0]. \end{cases} \quad (4.2.29)$$

Le problème est plus complexe car de dimension infinie [LBH03, BL05, LBH05].

### 4.2.8 modèles avec sollicitations stochastiques

Lorsque la sollicitation extérieure est de nature stochastique, il est nécessaire d'écrire la modélisation dans un cadre mathématique approprié. Nous ne nous étendrons pas ici sur les outils : le lecteur intéressé consultera [Ber03, BSL04, Ber04].

## 4.3 Rappels mathématiques : existence et unicité

Nous nous contenterons donc du cas déterministe. Les mathématiques nécessaires aux contextes évoqués dans les deux sous-sections précédentes y ont été renvoyées à la littérature. En résumé, nous dirons que des résultats d'existence et d'unicité sont encore disponibles avec des hypothèses adaptées.

La plupart des modèles introduits depuis le départ prennent la même forme mathématique, impliquant un opérateur maximal monotone. Un tel objet mathématique permet d'étendre la notion de "fonction croissante" à des graphes ([Bré73]).

Soit  $\langle, \rangle$  le produit scalaire sur  $\mathbb{R}^p$ . Soit  $A$  un graphe défini sur  $\mathbb{R}^p$  à valeurs dans l'ensemble des parties de  $\mathbb{R}^p$ .  $A$  est monotone si et seulement si

$$\forall x \in \mathbb{R}^p, \forall y \in \mathbb{R}^p, \forall x_1 \in A(x), \forall y_1 \in A(y), \langle x_1 - y_1, x - y \rangle \geq 0 \quad (4.3.1)$$

$A$  est un graphe maximal monotone si de plus  $A$  est maximal dans l'ensemble des graphes ordonnés par l'inclusion. Alors, la plupart des modèles peuvent être plongés dans le formalisme des inclusions différentielles de type :

$$\dot{X}(t) + f(t, X(t)) + A(X(t)) \ni 0, t \in ]0, T[, \quad (4.3.2)$$

avec des données initiales

$$X(0) \text{ donné (dans le domaine de définition du graphe } A, \text{ noté } D(A)). \quad (4.3.3)$$

$f$  désigne une fonction Lipschitzienne en son second argument, à valeurs dans  $\mathbb{R}^p$ .  $A$   $Y$  donné dans  $\mathbb{R}^p$ ,  $f(\cdot, Y)$  est une fonction dans  $L^\infty(0, T, \mathbb{R}^p)$ . Alors sous ces hypothèses, on montre [BS02, BS00] la

**Proposition 1.** *Pour tout  $X_0 \in D(A)$ , il existe une unique fonction  $X$  dans  $W^{1,1}(0, T; \mathbb{R}^p)$  telle que*

$$\begin{cases} \dot{X}(t) + f(t, X(t)) + A(X(t)) \ni 0 & \text{presque partout sur } ]0, T[, \\ X(0) = \Xi_0 \text{ donné.} \end{cases} \quad (4.3.4)$$

Par exemple, le modèle le plus simple (4.2.1)-(4.2.2), peut s'écrire sous la forme (4.3.4) en posant :

$$X = (\dot{x}, x)^t = (x_1, x_2)^t, f(t, X(t)) = ([kx_2 - F(t)]/m, x_1)^t, A(X) = (\alpha\sigma(x_1)/m, 0)^t, \quad (4.3.5)$$

avec les conditions initiales

$$X(0) = (\dot{x}_0, x_0). \quad (4.3.6)$$

Ce type de résultat d'existence et d'unicité s'étend au cas des modèles avec retard, et au cadre des inclusions différentielles stochastiques permettant de prendre en compte des sollicitations extérieures aléatoires (modèles de séismes, de houle, bruit variés).

Nous l'avons étendu également au cas de fonctions seulement localement lipschitziennes en leur second argument ([LBB05]).

## 4.4 Aspects numériques

### 4.4.1 Schémas numériques

La formulation (4.3.4) débouche sur un schéma numérique de type Euler implicite très simplement obtenu de la manière suivante, par discrétisation. Soit  $n \in \mathbb{N}^*$  ; soit  $h = T/n$  et pour  $q \in \{0, \dots, n\}$ , posons  $t_q = qh$ . Nous résolvons alors

$$\begin{cases} \forall q \in \{0, \dots, n-1\}, & \frac{X_{q+1} - X_q}{h} + f(t_q, X_q) + A(X_{q+1}) \ni 0, \\ X_0 = \Xi_0. \end{cases} \quad (4.4.1)$$

simplement parce que pour un opérateur maximal monotone, le graphe  $I + \lambda A$  avec  $\lambda > 0$  est inversible, d'inverse univoque. Alors le schéma est équivalent à

$$\forall q \in \{0, \dots, n-1\}, \quad X_{q+1} = (I + hA)^{-1} (hf(t_q, X_q) + X_q). \quad (4.4.2)$$

Appelons  $X^h$  la fonction linéaire interpolant les  $X_q$ . Cette fonction  $X^h$  converge vers la solution  $X$  du système (4.3.4) dans  $C^0([0, T], \mathbb{R}^p)$  (cf [CE75]).

On peut montrer que pour des opérateurs  $A$  particuliers (construits comme sous-différentielles de fonctions convexes propres semi-continues inférieurement [Bas00, Ber04, BS02, BS00]) que l'ordre de convergence du schéma vaut 1 (l'erreur entre la solution approchée et la solution exacte est en  $O(h)$ ). Dans le cas d'un opérateurs  $A$  quelconque, cet ordre vaut simplement 1/2.

### 4.4.2 Exemples de simulation

Nous proposons deux exemples de simulations correspondant à des cas à la fois de fonctions régulières seulement localement lipschitziennes sous sollicitation déterministe (cf Fig 4.9 a) et à une réalisation d'une force extérieure aléatoire (cf Fig 4.9 b). Les équations des modèles correspondant s'écrivent :

$$\begin{cases} m\ddot{x} + k_0x + cx^3 + a\dot{x} + \sum_{i=1}^n k_i u_i = F \text{ (déterministe), ou } \sigma N \text{ (stochastique),} \\ \forall i \in \{1, \dots, n\}, \quad \dot{u}_i + \beta(u_i/\eta_i) \ni \dot{x}, \end{cases} \quad (4.4.3)$$

avec les conditions initiales

$$x(0) = x_0, \quad \dot{x}(0) = \dot{x}_0, \quad \forall i \in \{1, \dots, n\}, \quad u_i(0) = u_{i,0} \in [-\eta_i, \eta_i],$$

et les valeurs des paramètres  $m = 1$ , avec dans le cas déterministe  $k_0 = 1/2$ ,  $c = 0.65$ ,  $a = 0$ ,  $F(t) = 3 \cos(2t)$  et  $n = 3$ , et dans le cas stochastique  $k_0 = 0$ ,  $c = 1$ ,  $a = 0.5$ ,  $\sigma = 3$  et  $n = 5$ ;  $N(t)$  est le bruit obtenu à partir du mouvement brownien standard réel.

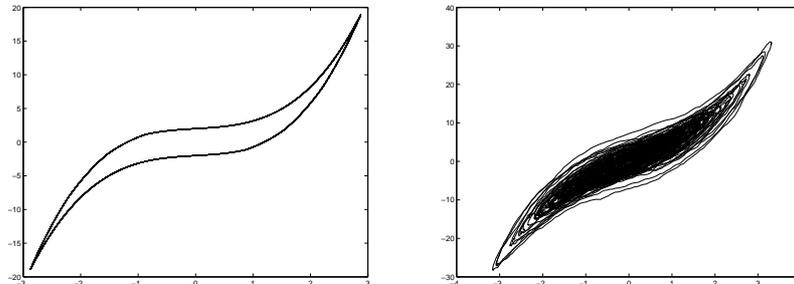


FIG. 4.9 – Exemple de cycles d'hystérésis pour le modèle (4.4.3) avec sollicitation déterministe (à gauche) et stochastique (à droite) : tracé de  $F(t) - m\ddot{x}(t)$  en fonction de  $x(t)$ ,  $t \in [600, 700]$ .

Il est possible d'obtenir un cycle d'hystérésis non convexe dans le premier cas, comme dans le second. Dans ce second cas, il n'est pas aussi commode que dans le premier cas de disposer d'une courbe de charge/décharge débloquent tous les patins sur un intervalle de temps défini à l'avance.

## 4.5 Identification

### 4.5.1 Principe pour le modèle de Prandtl généralisé

Tous les développements précédents ont pour but de fournir des modèles en vue de l'identification et de la description du comportement sous sollicitation extérieure faisant apparaître un cycle d'hystérésis. Ils permettent aussi par la modélisation et la simulation directe de préfigurer un modèle en vue de son identification : le nombre d'éléments de Saint-Venant est lié au déblocage des patins visible dans le phénomène physique dans la gamme de sollicitation envisagée. Il est aussi lié aux pertes de régularité du signal représentant le cycle.

Commençons par un exemple simple. Considérons un modèle de Prandtl généralisé 4.2.19. Supposons exercer une sollicitation extérieure  $F$  telle que l'on observe la courbe de chargement (c'est le tracé de  $F(t) - m\ddot{x}(t)$  en fonction de  $x(t)$  pour  $t \in [0, T]$ ,  $T > 0$ ) donnée par la figure 4.10. Nous considérons qu'il s'agit d'une moitié de cycle d'hystérésis. Notons  $A_1, \dots, A_{n+1}, A_{n+2}$  les extrémités des segments qui le constituent. Notons  $d_i, i = 1, \dots, n$  les différences des abscisses de  $A_{i+1}$  et  $A_i$  :  $d_i > 0$ . Notons aussi  $p_i, i = 1, \dots, n+1$  les pentes des segments  $[A_i, A_{i+1}]$ .

Si  $F$  est non constant sur tout sous-intervalle de  $[0, T]$ , alors  $x$  possède la même propriété. En effet, si  $x$  est constant sur  $[t_1, t_2] \subset [0, T]$ , alors  $\dot{x} = 0$ ,  $\ddot{x} = 0$  sur  $[t_1, t_2]$  et par unicité de la solution des inclusions différentielles

$$\dot{u}_i + \beta(u_i/\eta_i) \ni 0, \quad u_i(t_1) \text{ donné}, \quad (4.5.1)$$

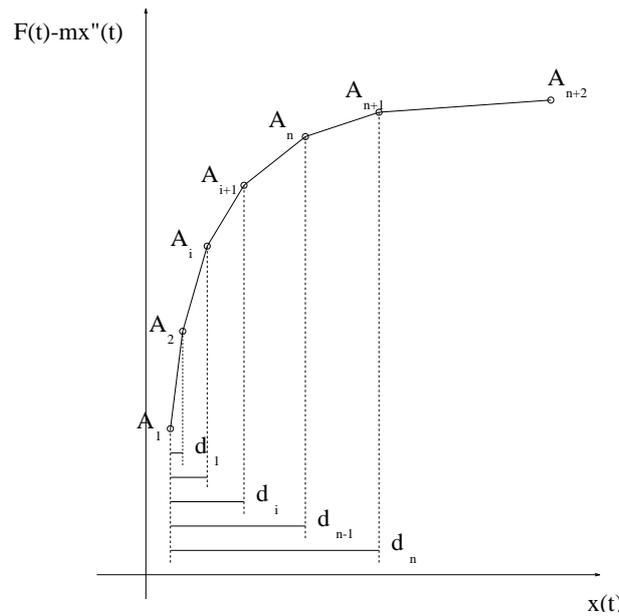


FIG. 4.10 – Courbe de chargement pour le modèle de Prandtl généralisé avec raidisseur linéaire.

$u_i(t)$  vaut  $u_i(t_1)$  sur  $[t_1, t_2]$ . Ceci impose que

$$F = m\ddot{x} + k_0x + \sum_{i=1}^n k_i u_i$$

est constante sur  $[t_1, t_2]$  ce qui est contraire à notre affirmation de départ.

Avec quelques hypothèses convenables, nous pouvons représenter la trajectoire d'un système en coordonnées  $(x(t), F(t) - m\ddot{x}(t))$ . Supposons que les  $\eta_i, i = 1, \dots, n$  vérifient :

$$\eta_1 < \eta_2 < \dots < \eta_n. \quad (4.5.2)$$

Supposons que  $F$  n'est nulle sur aucun ouvert non vide de  $[0, T]$  et que :

$$\forall j \in \{1, \dots, n\}, u_j(0) = -\eta_j, \quad (4.5.3)$$

$$x \text{ est croissante sur } [0, T]. \quad (4.5.4)$$

Alors il existe une séquence d'instants croissants telle que pour un  $j$  dans  $\{1, \dots, n\}$  avec

$$t_1 < t_2 < \dots < t_j = t_{j+1} = \dots = T$$

on ait

$$x(t_j) - x(0) = 2\eta_j, \quad (4.5.5)$$

$$F - m\ddot{x} - (k_0 + \sum_{i=j+1}^n k_i)x \text{ est constant sur } ]t_j, t_{j+1}[. \quad (4.5.6)$$

D'après notre hypothèse sur  $F$ ,  $x$  est strictement croissant sur  $[0, T]$  et est donc inversible. Posons :

$$\hat{u}_j(t) = -\eta_j + x(t) - x(0)$$

et appelons  $\xi$  la fonction inverse de  $x$ . Définissons

$$t_j = \begin{cases} \xi(2\eta_j + x(0)), & \text{si } 2\eta_j + x(0) < x(T), \\ T & \text{si } 2\eta_j + x(0) \geq x(T). \end{cases} \quad (4.5.7)$$

On vérifie que la fonction

$$u_j(t) = \begin{cases} \hat{u}_j(t), & \text{si } 0 \leq t < t_j, \\ \eta_j & \text{si } t \geq t_j, \end{cases} \quad (4.5.8)$$

est l'unique solution de

$$\begin{cases} \dot{u}_j + \beta(u_j/\eta_j) \ni \dot{x}, \\ u_j(0) = -\eta_j. \end{cases} \quad (4.5.9)$$

Donc, sur  $]t_j, t_{j+1}]$ , on a (4.5.6). En fait, (4.5.7) est équivalente à (4.5.5). Et les équations (4.5.5) et (4.5.6) peuvent s'écrire sous la forme

$$\forall j \in \{1, \dots, n+1\}, p_j = k_0 + \sum_{l=j}^n k_l, \quad (4.5.10)$$

$$\forall i \in \{1, \dots, n\}, d_i = 2\eta_i. \quad (4.5.11)$$

D'autre part, les hypothèses  $k_0 \geq 0$ ,  $k_i > 0$  et (4.5.10) ordonnent les  $p_i$  :

$$p_{n+1} < p_n < \dots < p_1.$$

Les équations (4.5.2) et (4.5.11) impliquent

$$d_1 < d_2 < \dots < d_n.$$

Ainsi (4.5.10) et (4.5.11) établissent une correspondance (bijection) entre les paramètres  $k_i, \eta_i$  du modèle généralisé de Prandtl et les paramètres géométriques  $p_j, d_j$  du demi-cycle d'hystérésis dans le plan  $(x, F - m\ddot{x})$ .

Ceci permet une identification partielle du modèle : elle est fondée sur le repérage des ruptures de pentes dans le cycle d'hystérésis. En pratique, le cycle est un semi de points et il faut utiliser un outil de traitement de signal adapté pour ce faire : nous utiliserons une analyse en ondelettes.

Le point inférieur gauche du cycle (4.10) dépend des conditions initiales et de  $F$  : il ne peut être identifié en général.

#### 4.5.2 Un autre exemple

Supposons disposer d'une courbe de chargement théorique telle que la partie supérieure du cycle de la Fig. (4.9, à gauche). Si l'on sait identifier les points de perte de régularité de cette courbe, on saura distinguer des phases temporelles de fonctionnement du système avec des patins débloqués et les autres bloqués, ces nombres étant connus et supposés croissants selon la courbe avec le nombre de ces points. Il est alors possible très simplement (par moindres carrés) d'identifier les coefficients des modèles que l'on suppose pouvoir simuler la courbe de chargement (la courbe hystérétique), à condition de connaître la force extérieure  $F(t)$  ([BSL00, LBB05]). Lors d'une expérimentation, la courbe de chargement disponible se réduit à un semis de points discret, bruité. L'identification des ruptures de pentes (ou plus généralement des pertes de régularité) est plus délicate. Nous utilisons une analyse en ondelettes, comme celle présentée en Fig. (4.11).

En abscisse des courbes est porté le déplacement. En ordonnée de la courbe inférieure discrète nous représentons  $F(t) - m\ddot{x}(t)$  en fonction du déplacement  $x(t)$

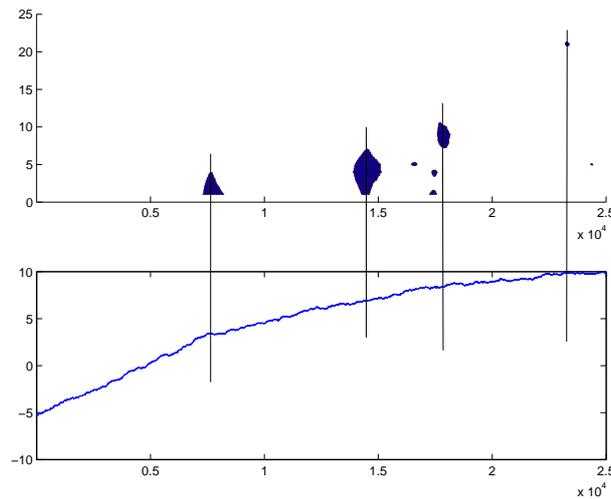


FIG. 4.11 – Amplitude des coefficients ondelettes en fonction du déplacement  $x$ .

dans le cas d'une sollicitation extérieure bruitée. Sur la courbe supérieure, nous traçons en fonction de  $x(t)$  les coefficients d'une analyse en ondelette ([Per00, Cou05]). Le schéma est tri-dimensionnel : en ordonnée de la courbe supérieure figurent les échelles de l'analyse en ondelettes. La "couleur" tracée représente la valeur du module des coefficients. Pour des coefficients faibles, la couleur est blanche (seuillage). Plus la couleur est foncée, plus grands sont les modules des coefficients. Le lieu des maxima, permet de repérer les lieux (en temps et/ou en déplacement) de changement de phase (de régularité) dans le signal. Ces changements peuvent s'étaler (tâches sombres), mais laissent entre chacune d'elles, des plages où le modèle évolue avec un nombre figé et connu, de patins bloqués ou débloqués. L'identification est menée sur cette plage. La spécificité de l'analyse en ondelette utilisée est justement de concentrer au maximum les tâches pour laisser une plage large.

A partir de simulations numériques fournissant des données discrètes contrôlées (tout le comportement et les données sont connus), éventuellement bruité (le bruit aussi est maîtrisé), nous avons obtenu d'excellents résultats d'identification.

Le principe est simple : le calcul de la pente (la dérivée par rapport au déplacement) de la courbe de chargement peut se faire de façon littérale en supposant que le cycle de chargement peut être simulé par un modèle donné. Alors, une procédure de moindres carrés permet d'identifier successivement les paramètres du modèle : rigidités, amortissement, termes non linéaires, plage par plage, puisque sur chaque plage isolée à l'aide de l'analyse en ondelettes le nombre de patins débloqués est connu et incrémenté d'un patin, à condition d'avoir écrit le modèle initial avec des paramètres de seuil croissants numérotant les patins successivement débloqués ([LBB05]).

L'application à des systèmes réels a fourni également des résultats conformes à la physique attendue des systèmes.

## 4.6 Conclusion

Nous avons insisté sur la modélisation de systèmes mécaniques discrets en dynamique, avec des termes de frictions permettant de simuler des comportements de type hystérétique.

Sur le principe, à condition de faire une hypothèse initiale sur le modèle mécanique susceptible de renvoyer un comportement hystérétique observé, il est possible d'identifier un système à un, voire plusieurs degrés de liberté. La limitation réside dans l'hypothèse d'une sollicitation extérieure connue, ce qui peut être le cas en laboratoire. Une segmentation soignée des plages de fonctionnement à patin bloqué est nécessaire. Elle peut être obtenue à l'aide d'une analyse en ondelette.

# Bibliographie

- [AD90a] J. Awrejcewicz and J. Delfs. Dynamics of a self-excited stick-slip oscillator with two degrees of freedom, part i. investigation of equilibria. *European Journal of Mechanics, A/Solids*, 9(4) :269–282, 1990.
- [AD90b] J. Awrejcewicz and J. Delfs. Dynamics of a self-excited stick-slip oscillator with two degrees of freedom, part ii, slip-stick, slip-slip, stick-slip transitions, periodic and chaotic orbits. *European Journal of Mechanics, A/Solids*, 9(5) :397–418, 1990.
- [AD03] J. Awrejcewicz and L. Dzyubak. Stick-slip chaotic oscillations in a quasi-autonomous mechanical system. *Int. J. of Nonl. Sci. Num. Simul.*, 4(2) :155–160, 2003.
- [ADG04] J. Awrejcewicz, L. Dzyubak, and C. Grebogi. A direct numerical method for quantifying regular and chaotic orbits. *Chaos, Solit. and Fract.*, 19 :503–507, 2004.
- [AF90] J.R. Anderson and A.A. Ferri. Behaviour of a single degree of freedom system with a generalized friction law. *Journal of Sound and Vibration*, 140 :287–304, 1990.
- [Awr03] R. Awrejcewicz, J. and Mosdorf. Numerical analysis of some problems of chaotic dynamics. *WNT*, 2003. Warsaw (in Polish).
- [Bas00] J. Bastien. *Etude théorique et numérique d'inclusions différentielles maximales monotones. Applications à des modèles élastoplastiques*. PhD thesis, Université Lyon I, 2000. Numéro d'ordre : 96-2000.
- [BCPR95] T. Baumberger, C. Caroli, B. Perrin, and O. Ronsin. Nonlinear analysis of the stick-slip bifurcation in the creep-controlled regime of dry friction. *Physical Review E*, 51(5) :4005–4010, 1995.
- [Ber01] D. Bernardini. Models of hysteresis in the framework of thermomechanics with internal variables. *Physica B*, 306 :132–136, 2001.
- [Ber03] F. Bernardin. Multivalued stochastic differential equation : convergence of a numerical scheme. *Set-Valued Analysis*, 11(4) :393–415, 2003.
- [Ber04] F. Bernardin. *Equations différentielles stochastiques multivoques Aspects théorique et numérique Applications*. PhD thesis, ENTPE et Université Lyon I, 6 decembre 2004.
- [BL05] J. Bastien and C.-H. Lamarque. Maximal monotone model with history term. *Non linear Analysis*, 63(5-7) :e199–e207, 2005.
- [BMMD07] J. Bastien, G. Michon, L. Manin, and R. Dufour. An analysis of the modified Dahl and Masing models : application to a belt tensioner. *J. Sound Vibration*, 302(4-5) :841–864, 2007.

- [Bré73] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).
- [Bro96] B. Brogliato. *Nonsmooth impact mechanics, Models, dynamics and control*. Springer-Verlag London Ltd., 1996.
- [BS00] J. Bastien and M. Schatzman. Schéma numérique pour des inclusions différentielles avec terme maximal monotone. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(7) :611–615, 2000.
- [BS02] J. Bastien and M. Schatzman. Numerical precision for differential inclusions with uniqueness. *M2AN Math. Model. Numer. Anal.*, 36(3) :427–460, 2002.
- [BSL00] J Bastien, M. Schatzman, and C.-H. Lamarque. Study of some rheological models with a finite number of degrees of freedom. *Eur. J. Mech. A Solids*, 19(2) :277–307, 2000.
- [BSL02] J. Bastien, M. Schatzman, and C.-H. Lamarque. Study of an elastoplastic model with an infinite number of internal degrees of freedom. *Eur. J. Mech. A Solids*, 21(2) :199–222, 2002.
- [BSL04] F. Bernardin, M. Schatzman, and C.-H. Lamarque. Second-order multi-valued stochastic differential equation on riemannian manifolds. *Proc. R. Soc. Lond. A*, 460 :3095–3121, 2004.
- [CE75] M.G. Crandall and L.C. Evans. On the relation of the operator  $\partial/\partial s + \partial/\partial \tau$  to evolution governed by accretive operators. *Israel J. Math.*, 21(4) :261–278, 1975.
- [CKM86] L.O. Chua, Motomasa Komuro, and Takashi Matsumoto. The double scroll family. ii. rigorous analysis of bifurcation phenomena. *Institute of Electrical and Electronics Engineers. Transactions on Circuits and Systems*, 33(11) :1097–1118, 1986.
- [CM96] D. Capecchi and R. Masiani. Reduced phase space analysis for hysteretic oscillators of masing type. *Chaos, Solit. and Fract.*, 7 :1583–1600, 1996.
- [CMKM86] L.O. Chua, M. M. Komuro, and T. Matsumoto. The double scroll family. i. rigorous proof of chaos. *Institute of Electrical and Electronics Engineers. Transactions on Circuits and Systems*, 33(11) :1072–1097, 1986.
- [Cou05] S. Coustel. *Wavelets based Methods for piecewise linear systems identification*. PhD thesis, Ecole doctorale MEGA, ENTPE, 2005.
- [CSB03] S. Casoria, G. Sybille, and P. Brunelle. Hysteresis modelling in the matlab/power systems blockset. *Math. and Comp. in Simul.*, 63 :237–248, 2003.
- [CV95] D. Capecchi and F. Vestroni. Asymptotic response of a two degree of freedom elastoplastic system under harmonic excitation ; internal resonance case. *Nonlinear Dynamics*, 7 :317–333, 1995.
- [DL72] G. Duvaut and J.L. Lions. *Les inéquations en mécanique et en physique. Travaux et Recherches Mathématiques*, 1972. Dunod, Paris.

- [DS83a] E.H. Dowell and H.B. Schwartz. Forced response of a cantilever beam with a dry friction damper attached, part ii : Experiment. *Journal of Sound and Vibration*, 91(2) :269–291, 1983.
- [DS83b] E.H. Dowell and H.B. Schwartz. Forced response of a cantilever beam with a dry friction damper attached, part i :theory. *Journal of Sound and Vibration*, 91(2) :255–267, 1983.
- [FB95] A.A. Ferri and A.C. Bindemann. Large amplitude vibration of a beam restrained by a non-linear sleeve joint. *Journal of Sound and Vibration*, 184(1) :19–34, 1995.
- [FS89] R Fougères and F. Sidoroff. The evolutive masing model and its application to cyclic plasticity and ageing. *Nuclear Engineering and Design*, 114 :273–284, 1989.
- [GLT76a] R. Glowinski, J.L. Lions, and R. Trémolières. *Analyse numérique des inéquations variationnelles. Tome 1, Théorie générale premières applications*. Number 5 in Méthodes Mathématiques de l'Informatique. Dunod, Paris, 1976.
- [GLT76b] R. Glowinski, J.L. Lions, and R. Trémolières. *Analyse numérique des inéquations variationnelles. Tome 2, Applications aux phénomènes stationnaires et d'évolution*. Number 5 in Méthodes Mathématiques de l'Informatique. Dunod, Paris, 1976.
- [IP93] I. Ionescu and J.C. Paumier. Friction dynamique avec coefficient dépendant de la vitesse de glissement. *Comptes Rendus de l'Académie des Sciences. Série I. Mathématique*, 316(1) :121–125, 1993.
- [JP85] M. Jean and E. Pratt. A system of rigid bodies with dry friction. *International Journal of Engineering Science*, 23(5) :497–513, 1985.
- [Ket98] Y. Ketema. Averaging and bifurcations in an oscillator with a history dependent restoring force. *J. Sound and Vibr.*, 209(1) :187–198, 1998.
- [KFSM01] A. Ktena, D. I. Fotiadis, P. D. Spanos, and C. V. Massalas. A preisach model identification procedure and simulation of hysteresis in ferromagnets and shape-memory alloys. *Physica B*, 306(1-4) :84–90, 2001.
- [Kle97] T. Kleineberg. A generalized approach for modelling the nonlocal memory of hysteretic systems. *J. Magn. and Magn. Mat.*, 166 :315–320, 1997.
- [KO93] H. Kolsch and D. Ottl. Simulation des mechanischen verhaltens von bauteilen mit statischer hysteresis. *Forschung im Ingenieurwesen*, 4 :66–71, 1993. (in German).
- [KRB<sup>+</sup>00] P.I. Koltermann, L.A. Righi, J. P. A. Bastos, R. Carlson, N. Sadowski, and N. J. Batistela. A modified jiles method for hysteresis computation including minor loops. *Physica B*, 275(1-3) :233–237, 2000.
- [KS00] G. Kádár and G. Szabó. Hysteresis modelling. *J. Magn. and Magn. Mat.*, 215-216 :592–596, 2000.
- [LBB05] C.-H. Lamarque, F. Bernardin, and J. Bastien. Study of a rheological model with a friction term and a cubic term : deterministic and stochastic cases. *Eur. J. Mech. A Solids*, 24(4) :572–592, 2005.

- [LBH03] C.-H. Lamarque, J. Bastien, and M. Holland. Study of a maximal monotone model with a delay term. *SIAM J. Numer. Anal.*, 41(4) :1286–1300 (électronique), 2003.
- [LBH05] C.-H. Lamarque, J. Bastien, and M. Holland. Maximal monotone model with delay term of convolution. *Math. Probl. Eng.*, (4) :437–453, 2005.
- [LMM95] M. Laghdir and M.D.P. Monteiro Marques. Dynamics of a particle with damping, friction, and percussional effects. *Journal of Mathematical Analysis and Applications*, 196(3) :902–920, 1995.
- [LV04] W. Lacarbonara and F. Vestroni. Nonclassical responses of oscillators with hysteresis. *Nonl. Dyn. FINIR*, 2004.
- [LYL01] C. Li, J. Yu, and X. Liao. Chaos in a three-neuron hysteresis hopfield-type neural network. *Physics Letters A*, 285 :368–372, 2001.
- [LZW02] H.G. Li, J.W. Zhang, and B.C. Wen. Chaotic behaviors of a bilinear hysteretic oscillator. *Mech. Research Comm.*, 29 :283–289, 2002.
- [Mad93] R. Madan. *Chua's circuit : a paradigm for chaos*. World Scientific Publishing Co. Inc., River Edge, NJ, 1993.
- [May91] I.D. Mayergoyz. *Mathematical Models of Hysteresis*. Springer, Berlin, 1991.
- [MCV02] R. Masiani, D. Capecchi, and F. Vestroni. Resonant and coupled response of hysteretic two-degree-of-freedom systems using harmonic balance method. *Int. J. Non-Linear Mech.*, 37 :1421–1434, 2002.
- [MDDWM03] D. Makaveev, L. Dupré, M. De Wulf, and J. Melkebeek. Dynamic hysteresis modelling using feed-forward neural networks. *J. Magn. and Magn. Mat.*, 254-255(256-258), 2003.
- [MF95] V.M. Matrosov and I.A. Finogenko. Right-hand solutions of the differential equations of dynamics for mechanical systems with sliding friction. *J. Appl. Math. Mech.*, 1995.
- [MF96a] V.M. Matrosov and I.A. Finogenko. On the existence of right-side solutions to differential equations of dynamics of mechanical systems with dry friction. *Differential Equations*, 32(2) :186–194, 1996.
- [MF96b] V.M. Matrosov and I.A. Finogenko. To the theory of differential equations arising in the dynamics of systems with friction. i. *Differential Equations*, 32(5) :610–618, 1996.
- [MM94] M.D.P. Monteiro Marques. An existence, uniqueness and regularity study of the dynamics of systems with one-dimensional friction. *European Journal of Mechanics. A Solids*, 13(2) :277–306, 1994.
- [OD02] J. Ortin and L. Delaey. Hysteresis in shape-memory alloys. *Int. J. Non-Linear Mech.*, 37 :1275–1281, 2002.
- [OHB03] F. Ossart, O. Hubert, and R. Billardon. A new internal variables scalar model respecting the wiping-out property. *J. Magn. and Magn. Mat.*, 254-255 :170–172, 2003.
- [Pal98] V. Palmov. *Vibrations of elasto-plastic bodies*. Springer-Verlag, Berlin, 1998. Translated from the 1976 Russian original by A. Belyaev and revised by the author.

- [Per00] S. Pernot. *Méthodes 'Ondelettes' pour l'Etude des Vibrations et de la Stabilité des Systèmes Dynamiques*. PhD thesis, Ecole Doctorale MEGA, ENTPE, Lyon, 2000.
- [PS90] K. Popp and P. Stelzer. Stick-slip vibrations and chaos. *Phil. Trans. R. Soc. Lond. A*, 332 :89–105, 1990.
- [PW81] T.K. Pratt and R. Williams. Non-linear analysis of stick-slip motion. *Journal of Sound and Vibration*, 74(4) :531–542, 1981.
- [Sap03] B. Sapinski. Dynamic characteristics of an experimental mr fluid. *Eng. Trans.*, 51(4) :399–418, 2003.
- [SBL99] M. Schatzman, J. Bastien, and C.-H. Lamarque. An ill-posed mechanical problem with friction. *Eur. J. Mech. A Solids*, 18(3) :415–420, 1999.
- [Sch78] M. Schatzman. A class of nonlinear differential equations of second order in time. *Nonlinear Analysis*, 2(3) :355–373, 1978.
- [SF03] B. Sapinski and J. Filus. Analysis of parametric models of mr linear damper. *J. Theor. and Appl. Mech.*, 41(2) :215–240, 2003.
- [Sha86] S.W. Shaw. On the dynamic response of a system with dry friction. *Journal of Sound and Vibration*, 108(2) :305–325, 1986.
- [ST96] D.E. Stewart and J.C. Trinkle. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and coulomb friction. *International Journal for Numerical Methods in Engineering*, 39(15) :2673–2691, 1996.
- [ST97] D.E Stewart and J.C Trinkle. Dynamics, friction, and complementarity problems. In *Complementarity and variational problems*, pages 425–439. SIAM, Philadelphia, PA, 1997. (Baltimore, MD, 1995).
- [Ste92] P. Stelzer. Nonlinear vibrations of structures induced by dry friction. *Nonlinear Dynamics*, 3 :329–345, 1992.
- [Ste96] D.E. Stewart. A numerical method for friction problems with multiple contacts. *J. Austral. Math. Soc. Ser. B*, 37 :288–308, 1996.
- [TC96] G.R. Tomlinson and Q. Chen. Parametric identification of systems with dry friction and nonlinear stiffness using a time serie model. *Journal of vibration and acoustics*, 118(2) :252–263, 1996.
- [TPSL97] J.C. Trinkle, J.S. Pang, S. Sudarsky, and G. Lo. On dynamic multi-rigid-body contact problems with coulomb friction. *Zeitschrift für Angewandte Mathematik und Mechanik*, 77(4) :267–279, 1997.
- [Vis94] A. Visintin. *Differential Models of Hysteresis*. Springer, Berlin, 1994.
- [VN02] F. Vestroni and M. Noori. Hysteresis in mechanical systems- modelling and dynamic response. *Int. J. Non-Linear Mech.*, 37 :1261–1459, 2002.



## Chapitre 5

# Modèles probabilistes de détection de rupture dans le comportement

*Darius Ghorbandazeh,  
Nelly Point*

### 5.1 Introduction

Lorsque des observations changent de comportement à partir d'un certain moment, il peut être utile de connaître avec précision le moment où se produit ce changement. L'objet de cet article est de présenter d'un point de vue théorique différents modèles probabilistes de rupture puis d'illustrer l'utilisation de ces modèles par deux exemples. Le premier concerne la détermination du seuil d'élasticité d'un matériau, de son module de Young et des paramètres du comportement plastique grâce à des données expérimentales obtenues lors d'un test uniaxial de traction. Pour certains matériaux, la courbe de traction ne présente pas de rupture de pente, ce qui rend l'identification du domaine élastique peu précis si on n'utilise par un test de rupture de comportement. Le second exemple utilise des données bancaires et permet de mettre en évidence une corrélation entre un changement de la moyenne d'un taux d'échange et des élections. Cet exemple n'est pas issu de la mécanique, mais il y a certainement des phénomènes liés à la mécanique qui peuvent relever de ce type de rupture (cf. [GP92]).

### 5.2 Présentation de trois modèles de rupture

Les données observées constituent une suite de variables aléatoires  $X_1, \dots, X_n$  qui sont susceptibles de changer de loi après les  $k$  premières observations.

Nous considérons le problème de test suivant (cf. [Gho95]) :

- **hypothèse nulle** : les observations  $X_1, \dots, X_n$  sont indépendantes et identiquement distribuées suivant une loi  $\mathbb{P}_\theta$
- **hypothèse alternative** : il existe un instant  $k$ , compris strictement entre 1 et  $n$  (appelé *instant de rupture*) tel que : les observations  $X_1, X_2, \dots, X_k$  sont indépendantes et identiquement distribuées suivant une loi  $\mathbb{P}_{\theta_1}$  et les observations  $X_{k+1}, \dots, X_n$  sont

indépendantes et identiquement distribuées suivant une loi  $\mathbb{P}_{\theta_2}$ .

Les paramètres du modèle sont estimés à partir de la fonction de vraisemblance associée à chaque hypothèse (méthode du maximum de vraisemblance). Sous l'hypothèse alternative, les estimateurs des paramètres  $\theta_1$  et  $\theta_2$  vont dépendre de l'instant  $k$ . Autrement dit, à chaque instant  $k$  nous avons une vraisemblance, donc pour estimer le paramètre  $k$ , nous considérons la fonction appelée rapport des vraisemblances maximales (ou processus des vraisemblances).

Dans ce travail nous supposons que les observations dont nous disposons suivent des lois normales. Ce choix est arbitraire, mais d'une part, pour la loi normale, le processus des vraisemblances est simple à exploiter et d'autre part, la loi normale est assez pertinente pour modéliser des fluctuations liées à des erreurs de mesure (cf. [Fer94]). Nous désignons par  $\varphi_{\mu, \sigma^2}$  la densité de la loi normale de paramètres  $\mu$  et  $\sigma^2$  (notée  $\mathcal{N}(\mu, \sigma^2)$ ). Cette densité est définie par :

$$\varphi_{\mu, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

L'**hypothèse nulle** sera désignée par  $H_0$  et l'**hypothèse alternative** sera désignée par  $H_1$ .

### 5.2.1 Les modèles

Nous allons nous intéresser à trois modèles. Le premier est utile pour détecter la rupture correspondant à un changement de moyenne, le deuxième pour détecter un changement de moyenne et de variance et le troisième est utile lorsque la moyenne reste inchangé et que seule la variance change. L'un des domaines d'application du troisième modèle est la modélisation des phénomènes climatologiques.

**Modèle de détection de la moyenne** Ce modèle sera appelé modèle  $\mathcal{M1}$

$$(\mathcal{M1}) \quad \left\{ \begin{array}{l} \mathbf{H}_0 : \quad X_1, \dots, X_n \text{ suivent une loi } \mathcal{N}(\mu_0, \sigma_0^2) \\ \mathbf{H}_1 : \quad \left\{ \begin{array}{l} X_1, \dots, X_k \quad \text{suivent une loi } \mathcal{N}(\mu_1, \sigma_1^2) \\ X_{k+1}, \dots, X_n \quad \text{suivent une loi } \mathcal{N}(\mu_2, \sigma_1^2) \end{array} \right. \end{array} \right.$$

**Modèle de détection de la moyenne et variance** Ce modèle sera appelé modèle  $\mathcal{M2}$

$$(\mathcal{M}2) \quad \begin{cases} \mathbf{H}_0 : & X_1, \dots, X_n \text{ suivent une loi } \mathcal{N}(\mu_0, \sigma_0^2) \\ \mathbf{H}_1 : & \begin{cases} X_1, \dots, X_k & \text{ suivent une loi } \mathcal{N}(\mu_1, \sigma_1^2) \\ X_{k+1}, \dots, X_n & \text{ suivent une loi } \mathcal{N}(\mu_2, \sigma_2^2) \end{cases} \end{cases}$$

**Modèle de détection de la variance** Ce modèle sera appelé modèle  $\mathcal{M}3$

$$(\mathcal{M}3) \quad \begin{cases} \mathbf{H}_0 : & X_1, \dots, X_n \text{ suivent une loi } \mathcal{N}(\mu_0, \sigma_0^2) \\ \mathbf{H}_1 : & \begin{cases} X_1, \dots, X_k & \text{ suivent une loi } \mathcal{N}(\mu_1, \sigma_1^2) \\ X_{k+1}, \dots, X_n & \text{ suivent une loi } \mathcal{N}(\mu_1, \sigma_2^2) \end{cases} \end{cases}$$

L'estimation des paramètres de chaque modèle est fait par la méthode du maximum de vraisemblance.

### 5.2.2 Estimation des paramètres du modèle $\mathcal{M}1$

Sous l'hypothèse  $H_0$  la fonction de vraisemblance  $L_0(\mu_0, \sigma_0^2, \underline{x})$  où  $\underline{x} = (x_1, \dots, x_n)$ , est définie par :

$$L_0 = \prod_{i=1}^n \varphi_{\mu_0, \sigma_0^2}(x_i) = \frac{1}{(\sqrt{2\pi})^n (\sigma_0^2)^{n/2}} \exp\left(-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu_0)^2\right) \quad (5.2.1)$$

Afin de pouvoir de calculer le maximum de la fonction de vraisemblance par rapport à  $\mu_0$  et  $\sigma_0^2$ , on définit la fonction "Log-Vraisemblance" par :  $\ell = \text{Log}L$ . On a :

$$\ell_0(\mu_0, \sigma_0^2, \underline{x}) = -n \text{Log}(\sqrt{2\pi}) - \frac{n}{2} \text{Log}(\sigma_0^2) - \frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \mu_0)^2 \quad (5.2.2)$$

En dérivant par rapport à  $\mu_0$  et  $\sigma_0^2$  on obtient :

$$\begin{cases} \frac{\partial \ell_0}{\partial \mu_0} = \frac{1}{\sigma_0^2} \sum_{i=1}^n (x_i - \mu_0) \\ \frac{\partial \ell_0}{\partial \sigma_0^2} = -\frac{n}{2\sigma_0^2} + \frac{1}{2\sigma_0^4} \sum_{i=1}^n (x_i - \mu_0)^2 \end{cases} \quad (5.2.3)$$

Le système (5.2.3) admet la solution :

$$\begin{aligned} \hat{\mu}_0 &= \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n \\ \hat{\sigma}_0^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \end{aligned} \quad (5.2.4)$$

En reportant (5.2.4) dans (5.2.1) on obtient l'expression de la vraisemblance maximale :

$$L_0^*(\underline{X}) = L(\hat{\mu}_0, \hat{\sigma}_0^2, \underline{X}) = \frac{e^{-n/2}}{(\sqrt{2\pi})^n (\hat{\sigma}_0^2)^{n/2}} \quad (5.2.5)$$

Sous l'hypothèse  $H_1$  la fonction de vraisemblance  $L_1(\mu_1, \mu_2, \sigma_1^2, k, \underline{x})$  est définie par :

$$\begin{aligned} L_1 &= \prod_{i=1}^k \varphi_{\mu_1, \sigma_1^2}(x_i) \prod_{i=k+1}^n \varphi_{\mu_2, \sigma_1^2}(x_i) \\ &= \frac{1}{(\sqrt{2\pi})^k (\sigma_1^2)^{k/2}} \exp\left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^k (x_i - \mu_1)^2\right) \\ &\quad \times \frac{1}{(\sqrt{2\pi})^{(n-k)} (\sigma_1^2)^{(n-k)/2}} \exp\left(-\frac{1}{2\sigma_1^2} \sum_{i=k+1}^n (x_i - \mu_2)^2\right) \end{aligned} \quad (5.2.6)$$

d'où

$$L_1 = \frac{1}{(\sqrt{2\pi})^n (\sigma_1^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma_1^2} \left(\sum_{i=1}^k (x_i - \mu_1)^2 + \sum_{i=k+1}^n (x_i - \mu_2)^2\right)\right\}$$

et l'expression de la fonction Log-vraisemblance  $\ell_1(\mu_1, \mu_2, \sigma_1^2, k, \underline{x})$  :

$$\ell_1 = -n \text{Log}(\sqrt{2\pi}) - \frac{n}{2} \text{Log}(\sigma_1^2) - \frac{1}{2\sigma_1^2} \left(\sum_{i=1}^k (x_i - \mu_1)^2 + \sum_{i=k+1}^n (x_i - \mu_2)^2\right) \quad (5.2.7)$$

En dérivant par rapport à  $\mu_1$ ,  $\mu_2$  et  $\sigma_1^2$  on obtient :

$$\begin{cases} \frac{\partial \ell_1}{\partial \mu_1} = \frac{1}{\sigma_1^2} \sum_{i=1}^k (x_i - \mu_1) \\ \frac{\partial \ell_1}{\partial \mu_2} = \frac{1}{\sigma_1^2} \sum_{i=k+1}^n (x_i - \mu_2) \\ \frac{\partial \ell_1}{\partial \sigma_1^2} = -\frac{n}{2\sigma_1^2} + \frac{1}{2\sigma_1^4} \left(\sum_{i=1}^k (x_i - \mu_1)^2 + \sum_{i=k+1}^n (x_i - \mu_2)^2\right) \end{cases} \quad (5.2.8)$$

Le système (5.2.8) admet la solution :

$$\begin{aligned} \hat{\mu}_1(k) &= \frac{1}{k} \sum_{i=1}^k X_i = \bar{X}_k \\ \hat{\mu}_2(k) &= \frac{1}{n-k} \sum_{i=k+1}^n X_i = \bar{X}_k^* \\ \hat{\sigma}_1^2(k) &= \frac{1}{n} \left(\sum_{i=1}^k (X_i - \bar{X}_k)^2 + \sum_{i=k+1}^n (X_i - \bar{X}_k^*)^2\right) \end{aligned} \quad (5.2.9)$$

En reportant (5.2.9) dans (5.2.6) on obtient l'expression de la vraisemblance maximale :

$$L_1^*(k, \underline{X}) = \frac{e^{-n/2}}{(\sqrt{2\pi})^n (\hat{\sigma}_1^2(k))^{n/2}} \quad (5.2.10)$$

D'où le rapport des vraisemblances maximales (processus de vraisemblance) pour le modèle  $\mathcal{M}_1$  :

$$R_1^*(k, \underline{X}) = \frac{L_1^*(k, \underline{X})}{L_0^*(\underline{X})} = \left[ \frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2(k)} \right]^{n/2} \quad (5.2.11)$$

L'estimateur de  $k$  est déterminé par :  $\hat{k} = \text{Arg Max}_{1 < k < n} R_1^*(k, \underline{X})$ .

### 5.2.3 Estimation des paramètres du modèle $\mathcal{M}_2$

Sous l'hypothèse  $H_1$  la fonction de vraisemblance  $L_1(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x})$  est définie par :

$$\begin{aligned} L_1 &= \prod_{i=1}^k \varphi_{\mu_1, \sigma_1^2}(x_i) \prod_{i=k+1}^n \varphi_{\mu_2, \sigma_2^2}(x_i) \\ &= \frac{1}{(\sqrt{2\pi})^k (\sigma_1^2)^{k/2}} \exp\left(-\frac{1}{2\sigma_1^2} \sum_{i=1}^k (x_i - \mu_1)^2\right) \\ &\quad \times \frac{1}{(\sqrt{2\pi})^{(n-k)} (\sigma_2^2)^{(n-k)/2}} \exp\left(-\frac{1}{2\sigma_2^2} \sum_{i=k+1}^n (x_i - \mu_2)^2\right) \end{aligned} \quad (5.2.12)$$

D'où l'expression de la fonction Log-vraisemblance  $\ell_1(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x})$  :

$$\begin{aligned} \ell_1 &= -n \text{Log}(\sqrt{2\pi}) - \frac{k}{2} \text{Log}(\sigma_1^2) - \frac{n-k}{2} \text{Log}(\sigma_2^2) \\ &\quad - \frac{1}{2\sigma_1^2} \sum_{i=1}^k (x_i - \mu_1)^2 - \frac{1}{2\sigma_2^2} \sum_{i=k+1}^n (x_i - \mu_2)^2 \end{aligned} \quad (5.2.13)$$

En dérivant par rapport à  $\mu_1$ ,  $\mu_2$ ,  $\sigma_1^2$  et  $\sigma_2^2$  on obtient :

$$\left\{ \begin{aligned} \frac{\partial \ell_1}{\partial \mu_1}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x}) &= \frac{1}{\sigma_1^2} \sum_{i=1}^k (x_i - \mu_1) \\ \frac{\partial \ell_1}{\partial \mu_2}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x}) &= \frac{1}{\sigma_2^2} \sum_{i=k+1}^n (x_i - \mu_2) \\ \frac{\partial \ell_1}{\partial \sigma_1^2}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x}) &= -\frac{k}{2\sigma_1^2} + \frac{1}{2\sigma_1^4} \sum_{i=1}^k (x_i - \mu_1)^2 \\ \frac{\partial \ell_1}{\partial \sigma_2^2}(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, k, \underline{x}) &= -\frac{n-k}{2\sigma_2^2} + \frac{1}{2\sigma_2^4} \sum_{i=k+1}^n (x_i - \mu_2)^2 \end{aligned} \right. \quad (5.2.14)$$

Le système (5.2.14) admet la solution :

$$\begin{aligned}
\hat{\mu}_1(k) &= \frac{1}{k} \sum_{i=1}^k X_i = \bar{X}_k \\
\hat{\mu}_2(k) &= \frac{1}{n-k} \sum_{i=k+1}^n X_i = \bar{X}_k^* \\
\hat{\sigma}_1^2(k) &= \frac{1}{k} \sum_{i=1}^k (X_i - \bar{X}_k)^2 \\
\hat{\sigma}_2^2(k) &= \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \bar{X}_k^*)^2
\end{aligned} \tag{5.2.15}$$

En emportant (5.2.15) dans (5.2.13) on obtient l'expression de la vraisemblance maximale :

$$L_1^*(k, \underline{X}) = \frac{e^{-n/2}}{(\sqrt{2\pi})^n (\hat{\sigma}_1^2(k))^{k/2} (\hat{\sigma}_2^2(k))^{(n-k)/2}} \tag{5.2.16}$$

D'où le rapport des vraisemblances maximales (processus de vraisemblance) pour le modèle  $\mathcal{M}_2$  :

$$R_3^*(k, \underline{X}) = \frac{L_1^*(k, \underline{X})}{L_0^*(\underline{X})} = \left[ \frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2(k)} \right]^{k/2} \left[ \frac{\hat{\sigma}_0^2}{\hat{\sigma}_2^2(k)} \right]^{(n-k)/2} \tag{5.2.17}$$

L'estimateur de  $k$  est déterminé par :  $\hat{k} = \text{Arg Max}_{1 < k < n} R_3^*(k, \underline{X})$ .

### 5.2.4 Estimation des paramètres du modèle $\mathcal{M}_3$

Contrairement aux deux modèles précédents les expressions des estimateurs de ce modèle ne sont pas explicites.  $\hat{\mu}_1(k)$ ,  $\hat{\sigma}_1^2(k)$  et  $\hat{\sigma}_2^2(k)$  vérifient les relations suivantes :

$$\begin{cases}
\hat{\mu}_1(k) = \frac{1}{\hat{\sigma}_1^2(k)} \sum_{i=1}^k (X_i - \hat{\mu}_1(k)) + \frac{1}{\hat{\sigma}_2^2(k)} \sum_{i=k+1}^n (X_i - \hat{\mu}_1(k)) = 0 \\
\hat{\sigma}_1^2(k) = \frac{1}{k} \sum_{i=1}^k (X_i - \hat{\mu}_1(k))^2 \\
\hat{\sigma}_2^2(k) = \frac{1}{n-k} \sum_{i=k+1}^n (X_i - \hat{\mu}_1(k))^2
\end{cases} \tag{5.2.18}$$

## 5.3 Plasticité

Un problème récurrent lors de l'étude du comportement plastique d'un matériau est la détermination précise du domaine élastique et du module de Young. Pour certains matériaux ne présentant pas de rupture de pente lors d'un test uniaxial de traction, cette détermination n'est toujours facile. L'utilisation d'un test de rupture s'avère alors utile.

La variable aléatoire gaussienne  $X_i$  représente alors l'écart entre la contrainte mesurée  $\sigma_i$  et la contrainte évaluée en fonction de la déformation  $\sigma(\varepsilon_i)$  et selon le modèle considéré : élastique linéaire ou plastique. La déformation est supposée croître avec

l'indice  $i$ . Pour les  $i$  avant  $k$  on cherche donc une fonction linéaire  $\sigma(\varepsilon_i) = \alpha\varepsilon_i + \beta$  où  $\alpha$  va donner le module de Young et où  $\beta$  peut correspondre à un léger décalage dû à l'appareil de mesure. Pour les  $i$  après  $k$  on cherche une fonction non linéaire de  $\varepsilon_i$ . Si on choisit pour décrire ce comportement non linéaire un polynôme du second degré, on obtient un résultat peu satisfaisant : le point de rupture trouvé est représenté sur la figure 5.1 par **O**. Par contre si on choisit une fonction correspondant mieux au comportement plastique observé, en choisissant par exemple une représentation de la forme  $\sigma(\varepsilon_i) = a\sqrt{\varepsilon_i} + b\varepsilon_i + c$ , plus conforme au comportement des données expérimentales, on obtient comme point de rupture le point **X** sur la figure 5.1. Les coefficients  $\alpha, \beta, a, b, c$ , sont obtenus simultanément. Comme pour chaque valeur de  $k$  il faut calculer les coefficients des modèles, pour des raisons de temps de calcul il faut approcher les modèles par une approximation qui soit linéaire par rapport aux coefficients afin de pouvoir utiliser pour chaque valeur de  $k$  les moindres carrés linéaires.

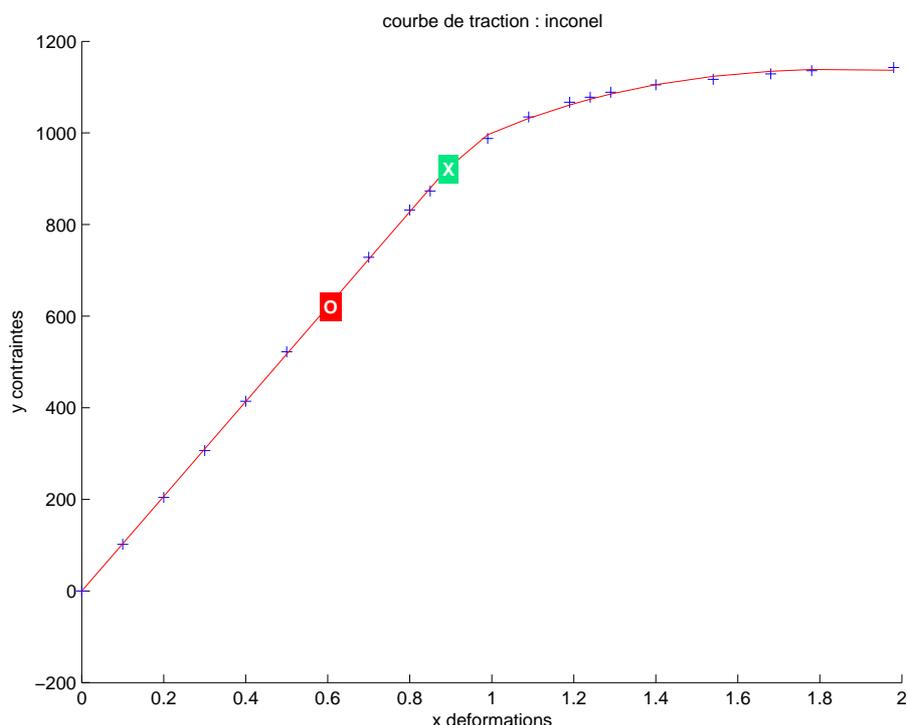


FIG. 5.1 – Contraintes en fonctions des déformations

## 5.4 L'influence des élections françaises sur le taux d'échange

Dans cette partie, nous présentons une application pratique de l'analyse aux données concernant le taux d'échange **Franc-Deutsche Mark** (qui n'existe plus) durant l'année 1993. La série utilisée contient 256 observations (du 1 Janvier 1993 au 24 Décembre 1993) d'une moyenne égale à  $\hat{\mu}_0 = 1,3472$  et d'un écart-type égale à  $\sqrt{\hat{\sigma}_0^2} = 1.4473$ . L'objet de cet étude est d'analyser l'influence des élections françaises sur le taux d'échange **Franc-Deutsche Mark**.

La figure suivante représente les données brutes par ordre chronologique.

Pour gommer les petites fluctuations et mieux visualiser un point de rupture nous

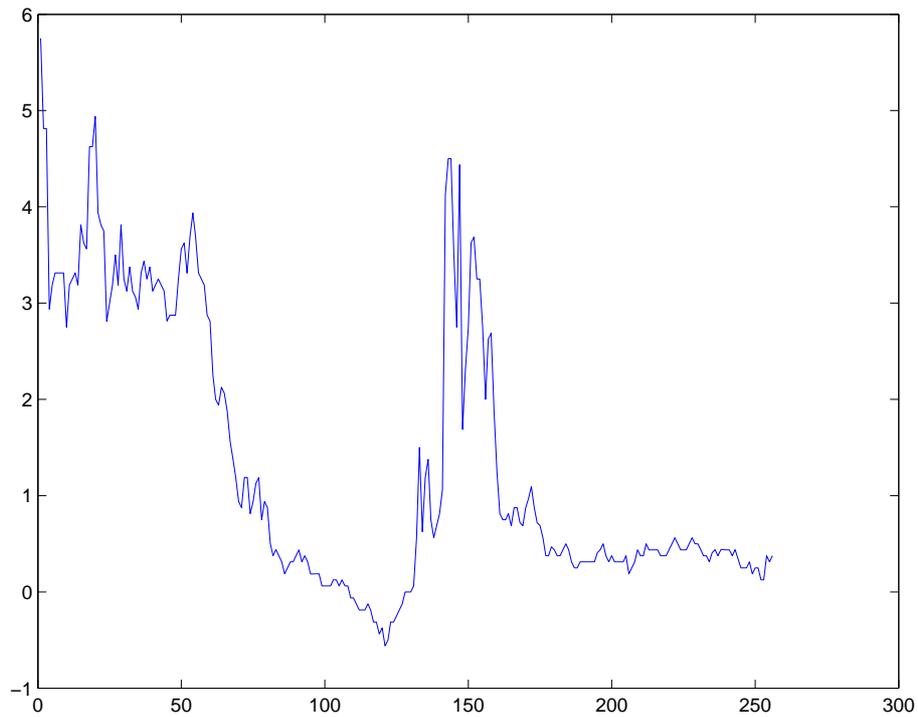


FIG. 5.2 – Présentation graphique des données brutes

avons regrouper les données en 13 et 20 classes. On constate qu'indépendamment du nombre de classes les deux histogrammes gardent la même allure et présentent un point de rupture que la méthode ci-dessus va permettre d'identifier.

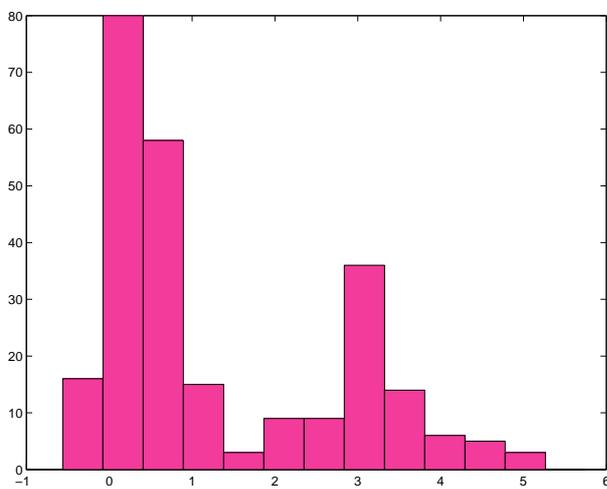


FIG. 5.3 – Répartition des données en 13 classes

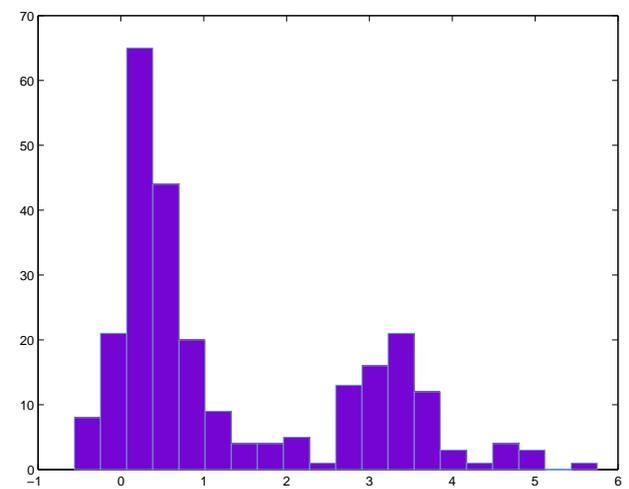


FIG. 5.4 – Répartition des données en 20 classes

La figure 5.5 représente le rapport des vraisemblances maximales donné par la relation (5.2.11)

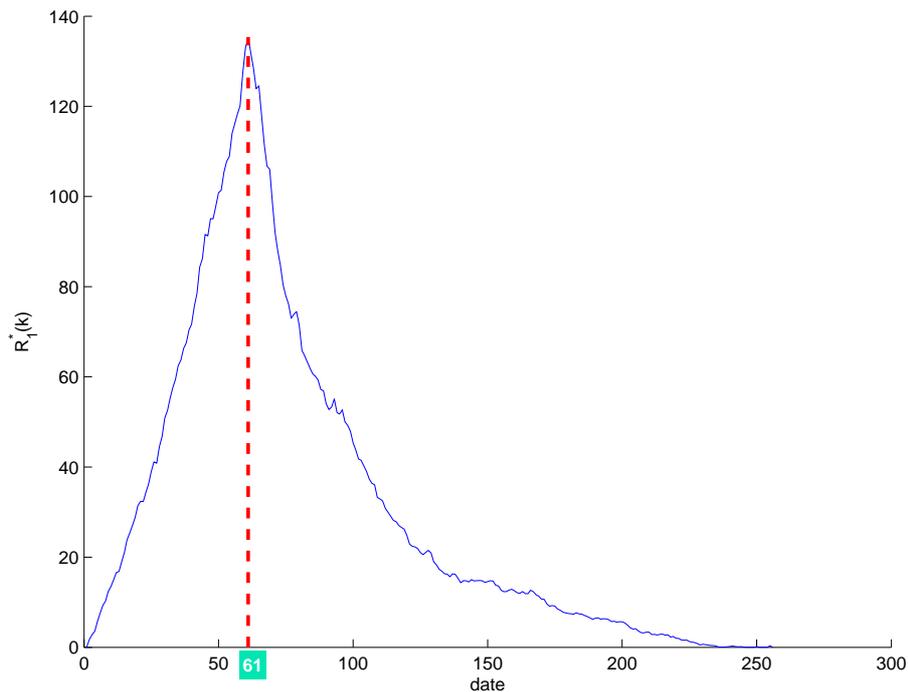


FIG. 5.5 – Fonction de vraisemblance

On remarque que le maximum de la fonction de vraisemblance est atteint pour la 61<sup>ème</sup> observation, ce qui correspond à la date 26 Mars 1993. La moyenne de la série avant cette étant égale à 3,4254 et celle d'après étant égale à 0,6971. Il est intéressant de constater que la date 26 Mars 1993 coïncide avec les législatives de 1993.

## 5.5 Conclusion

Il y a une large littérature sur la modélisation du phénomène de rupture en probabilité ([Fer94], (cf. [Gho95])), nous avons limité notre présentation à l'hypothèse d'une loi de probabilité normale car cela se prête bien à l'estimation d'erreur de mesure mais il existe d'autres modèles basés sur des hypothèses différentes. Il faut donc savoir que ces méthodes existent et qu'elles donnent des résultats satisfaisants avec une procédure somme toute assez simple.



# Bibliographie

- [Fer94] D. Ferger. Change-point estimators in case of small disorders. *J. Statist. Plann. Inference*, 40 :33–49, 1994.
- [Gho95] D. Ghorbanzadeh. Un test de détection de rupture de la moyenne dans un modèle gaussien. *Revue de Statistique Appliquée*, XLIII :67–76, 1995.
- [GP92] D. Ghorbanzadeh and D. Picard. étude asymptotique et pratique du comportement de deux tests de détection de rupture. *Statistique et Analyse des données*, 16(3) :63–84, 1992.



## Chapitre 6

# Assimilation thermique 1D par méthode adjointe libérée

*Frédéric Bourquin  
Alexandre Nassiopoulos*

### 6.1 Introduction

La précontrainte affecte le comportement vibratoire des structures. Ainsi, la détermination fiable et rapide de l'état de précontrainte associé à un champ de température devient une étape cruciale dans de nombreux processus comme le contrôle de santé en génie civil. La meilleure reconstruction possible du champ de température constitue alors un élément clef de cette démarche.

Il existe de nombreuses stratégies pour détecter les endommagements d'une structure de génie civil : entre autres, une classe très attrayante de techniques repose sur l'analyse vibratoire sous sollicitations ambiantes. Ces méthodes dites *output-only* consistent soit à identifier le spectre basse fréquence de la structure à l'instant courant pour en déduire si celle-ci a subi un endommagement, soit à vérifier qu'un certain vecteur appartient toujours au noyau de la matrice de Hankel des matrices de corrélation des mesures [BMG04]. Dans les deux cas, la structure intacte, ou de référence, est censée être connue, ce qui est le cas pour chaque nouveau pont pour lequel on dispose aujourd'hui d'un modèle mécanique plutôt précis à l'issue de la construction. L'analyse modale opérationnelle et la détection en ligne d'endommagements visent donc à détecter les évolutions du spectre de la structure. On dispose maintenant d'une expérience pratique des algorithmes et de leur comportement sur des cas réels comme par exemple le pont Z24 [Pee00]. L'analyse modale *output-only* fondée sur des méthodes de sous-espace tirent profit des matrices de corrélation et débouchent sur des algorithmes en ligne puissants [BMG04, MG04] de détection d'endommagement, dont les domaines d'application incluent le génie aéronautique, le génie mécanique, et le génie civil. Il existe beaucoup d'autres méthodes (voir par exemple [SCF00, SFHW01], [Cre04, FHS<sup>+</sup>03]).

Or le comportement dynamique de la structure sous surveillance peut varier sous l'influence de plusieurs facteurs, y compris les variations de température [RBS<sup>+</sup>00]. En fait, selon les conditions aux limites de la structure et la distribution de température,

des variations thermiques éventuellement uniformes peuvent causer des variations très importantes du spectre de la structure [FHS<sup>+</sup>03], jusqu'à 10 %, en raison de la précontrainte additionnelle. La température peut même modifier les conditions de frontière et le nombre de points de contact entre les piles et le tablier du pont. En fait, la température peut engendrer une déflexion aussi importante que le poids propre du pont [IRvA<sup>+</sup>01]. Ainsi les effets environnementaux peuvent être de plusieurs ordres de grandeur plus importants que l'effet des endommagements structuraux. Par ailleurs, les méthodes de contrôle structural fondées sur un modèle de la structure souffrent de toute modification non identifiée du comportement vibratoire de la structure qui nuit alors à la robustesse des lois de contrôle.

Par conséquent, comme mentionné dans [FHS<sup>+</sup>03] et dans une grande partie de la littérature sur le contrôle de santé, il est obligatoire d'éliminer les effets thermiques. Cette élimination reste un défi dans ce domaine. La plupart des tentatives dans cette direction ont favorisé des approches en boîte noire comme l'apprentissage de réseaux de neurones, en vue de corrélérer la température aux mesures extensométriques [FHS<sup>+</sup>03, SWF03]. Voir [Kul02, Kul03, Kul04] pour l'usage de l'analyse en composantes principales, fondée sur la corrélation entre les signatures observées. Mais les données d'apprentissage doivent inclure au moins un cycle complet de conditions en service avant que l'on puisse appliquer un algorithme de détection d'endommagement. Par ailleurs, la méthode proposée ne permet pas de prendre en compte de manière appropriée les événements climatiques exceptionnels.

Par ailleurs, il faut noter que le nombre de capteurs de température sur de grands ouvrages peut atteindre plusieurs centaines [Won04], ce qui témoigne de l'importance de cette grandeur physique et de ses variations spatiales.

En outre, la longévité des structures en béton est dictée par l'évolution de la température au jeune âge, d'où l'intérêt d'évaluer le champ de température aussi tôt que possible. Dans le même esprit, la vérification en temps réel de la bonne prise du béton ne peut que tirer parti d'une connaissance fine du champ de température et de son évolution.

Ainsi la température ne constitue pas qu'un simple paramètre de nuisance qui doit être rejeté d'un modèle qui tiendrait compte seulement de la dynamique de la structure. Au contraire, une surveillance multi-physique complète combinée à une modélisation thermomécanique inverse pourrait jouer un rôle important dans la correction statique des mesures cinématiques comme dans les méthodes de contrôle de santé fondées sur l'analyse vibratoire ainsi que dans le contrôle de chantier en vue des questions de durabilité et de productivité, sans compter le contrôle actif ou semi-actif des structures. Sur la base de mesures thermiques en un certain nombre de points, il est donc essentiel de reconstruire le champ de température réel à chaque instant afin d'en tenir compte explicitement dans les algorithmes de contrôle de santé ou de contrôle actif des structures.

Le travail présenté ici vise la reconstruction du champ de température dans une structure à partir de mesures de l'évolution de la température dans un intervalle de temps donné fournies par un certain nombre de capteurs ponctuels. Cette question est un cas particulier d'une classe de problèmes très importante citée dans la littérature sous le nom de problèmes inverses de transfert de chaleur [BBC85], [Ali94], [zO00], et qui fait l'objet d'un nombre très important de publications. Elle se trouve aussi au cœur de l'assimilation de données en météorologie.

Parmi les techniques de résolution que l'on rencontre dans la littérature, les méthodes dites explicites tentent, après discrétisation le plus souvent par différences finies,

d'inverser l'opérateur qui au contrôle associe les mesures [BBH96],[RB86]. D'autres méthodes qualifiées d'implicites sont des méthodes itératives qui évitent l'inversion directe de cet opérateur [CdCV91], [MRdCV99], [Ali94], [AAR95]. Elles présentent souvent un caractère régularisant intrinsèque [AAR95], [Han95]. Dans cette direction, la formulation par moindres carrés semble être un cadre très favorable pour le développement d'algorithmes rapides et robustes [JzB91]. La régularisation de type Tikhonov [EHN94] est souvent utilisée pour corriger le caractère mal posé du problème.

La démarche mise en œuvre dans cet article est celle du contrôle optimal [Lio68] appliqué à la reconstruction des flux extérieurs et de la température initiale. On obtient en posttraitement la température courante par résolution de l'équation de la chaleur. Cette méthode, parfois appelée *méthode adjointe* à cause de l'utilisation systématique de l'état adjoint pour calculer le gradient de la fonctionnelle à minimiser, présente de nombreux avantages, dont la généralité. Elle s'impose ici car on souhaite en effet traiter des situations multidimensionnelles et des géométries éventuellement complexes.

Le problème étant mal posé, on est amené à le régulariser. On montre alors que le choix d'une norme produit de type  $H^1$  au lieu de  $L^2$  permet d'améliorer fortement la précision de la température reconstruite par une libération de l'état adjoint à l'instant final.

## 6.2 Formulation générale du problème

Dans ce paragraphe, on considère un solide occupant un domaine  $\Omega$  dans  $\mathbb{R}^d$ ,  $d \geq 1$ , de frontière  $\partial\Omega$ . Par la suite, on se limitera au cas où  $d = 1$ .

### 6.2.1 L'équation de la chaleur

En l'absence de changement de phase et en négligeant la dissipation intrinsèque, la distribution de température  $\theta(x, t)$  dans le solide est régie par l'équation de la chaleur

$$\rho c \frac{\partial \theta}{\partial t} - \operatorname{div}(\mathbf{K} \operatorname{grad} \theta) = f \quad \text{dans } \Omega \times [0, T] \quad (6.2.1)$$

où  $x = (x_1, x_2, x_3) \in \Omega$  est la variable d'espace,  $t \in [0, T]$  la variable de temps,  $f = f(x, t)$  désigne la source de chaleur interne distribuée,  $\rho = \rho(x)$  la masse volumique du matériau,  $c = c(x)$  sa capacité calorifique et  $\mathbf{K}$  son tenseur de conductivité. Le tenseur  $\mathbf{K}$  est symétrique, défini positif et se réduit souvent à un tenseur diagonal. Si la conduction est isotrope, il s'écrit  $\mathbf{K} = k(x)\operatorname{Id}$ , où  $k = k(x)$  désigne le coefficient scalaire de conductivité et  $\operatorname{Id}$  la matrice identité.

A l'instant initial, le champ de température est donné par  $\theta(x, 0) = \theta^0(x)$ , et à chaque instant le solide est soumis à une température extérieure  $\theta_{ext}(x, t)$  et à un flux de chaleur  $g(x, t)$  le long de sa frontière  $\partial\Omega$ . On peut résumer les interactions avec le milieu extérieur par la condition de Fourier-Robin

$$(\mathbf{K} \operatorname{grad} \theta) \cdot \vec{n} + \alpha(\theta - \theta_{ext}) = g \quad \text{sur } \partial\Omega \times [0, T] \quad (6.2.2)$$

où  $\alpha = \alpha(x|_{\partial\Omega}) \geq 0$  désigne le coefficient d'échange, et  $\vec{n}$  le vecteur unitaire normal à  $\Omega$  le long de  $\partial\Omega$ .

Les conditions aux limites de type Dirichlet (respectivement Neumann) peuvent être vues comme des conditions aux limites de type Fourier-Robin avec un coefficient  $\alpha$  qui tend vers l'infini (respectivement vers zéro).

## 6.2.2 Reconstruction de la température

En l'absence de source de chaleur interne, la température dans  $\Omega$  est donc régie par le problème d'évolution

$$\begin{cases} \rho c \frac{\partial \theta}{\partial t} - \operatorname{div}(\mathbf{K} \operatorname{grad} \theta) = 0 & \text{dans } \Omega \times [0, T] \\ (\mathbf{K} \operatorname{grad} \theta) \cdot \vec{n} + \alpha(\theta - \theta_{ext}) = g & \text{sur } \partial\Omega \times [0, T] \\ \theta(x, 0) = \theta^0(x) & \text{dans } \Omega \end{cases} \quad (6.2.3)$$

Si on pose  $\Phi(x, t) = g(x, t) + \alpha\theta_{ext}$ , le problème s'écrit

$$\begin{cases} \rho c \frac{\partial \theta}{\partial t} - \operatorname{div}(\mathbf{K} \operatorname{grad} \theta) = 0 & \text{dans } \Omega \times [0, T] \\ (\mathbf{K} \operatorname{grad} \theta) \cdot \vec{n} + \alpha\theta = \Phi & \text{sur } \partial\Omega \times [0, T] \\ \theta(x, 0) = \theta^0(x) & \text{dans } \Omega \end{cases} \quad (6.2.4)$$

Notons que le couple  $\{\theta^0, \Phi\}$  détermine de manière unique la solution  $\theta(x, t)$  du problème (6.2.4). Par contre, il n'est pas possible d'identifier indépendamment  $\theta_{ext}$  et  $g$ . Dans la perspective d'une identification seul  $\Phi$  a un sens.

On suppose qu'à l'instant final  $T$  on dispose de  $M$  mesures  $\{\theta_k^d(t)\}_{k=1}^M$  de la température sur tout l'intervalle de temps  $[0, T]$ , fournies par des capteurs ponctuels placés aux points  $\{x_k\}_{k=1}^M$ . A partir de ces mesures, l'objectif fixé est de reconstruire la donnée  $\theta(x, T)$  ou, à défaut, la donnée  $\theta(x, T - \epsilon)$ , avec  $\epsilon$  petit.

Parmi les méthodes possibles, on éliminera d'emblée le filtrage de Kalman en temps continu qui conduirait à la résolution d'équations de Riccati en très grande dimension ce qui est pratiquement hors d'atteinte sans réduction préalable.

Une autre manière de parvenir au champ de température à l'instant final est d'essayer de déterminer le couple  $\{\theta^0, \Phi\}$  qui a donné naissance au champ de température dont on connaît la valeur aux points de mesure. Au lieu de résoudre l'équation de la chaleur dans le sens habituel, avec comme données les conditions initiale et aux limites, l'objectif fixé est de retrouver ces données à partir d'une solution (partiellement) connue. Il s'agit alors d'un problème inverse. Bien entendu, il n'est pas évident que cette approche puisse déboucher sur des algorithmes en temps réel, c'est pourquoi l'on s'intéresse à des méthodes rapides quoique généralistes.

Le problème peut être formulé comme un problème de minimisation au sens des moindres carrés d'une fonctionnelle mesurant l'écart entre les mesures et le champ de température reconstruit. Cet écart peut être par exemple mesuré par la fonction de coût quadratique  $J$  :

$$J(\{\theta^0, \Phi\}) = \frac{1}{2} \int_0^T \sum_{k=1}^M \left( \theta(x_k, t) - \theta_k^d(t) \right)^2 dt \quad (6.2.5)$$

Le problème à résoudre s'écrit alors

$$\begin{cases} \text{Trouver } \{\theta^{0*}, \Phi^*\} \text{ tel que} \\ J(\{\theta^{0*}, \Phi^*\}) = \min_{\{\theta^0, \Phi\}} J(\{\theta^0, \Phi\}) \end{cases} \quad (6.2.6)$$

La minimisation de cette fonction de coût par rapport au couple d'arguments  $\{\theta^0, \Phi\}$  doit fournir la condition initiale  $\theta^{0*}$  et la condition aux limites  $\Phi^*$  qui donnent le champ de température  $\theta(x, t)$  dont la valeur aux points  $\{x_k\}_{k=1}^M$  est la plus proche possible

des mesures  $\{\theta(x_k)\}_{k=1}^M$ . La minimisation a un sens par rapport à une norme donnée qu'il conviendra de préciser par la suite en même temps que l'espace fonctionnel dans lequel sera recherché le couple  $\{\theta^0, \Phi\}$ .

### 6.2.3 Un problème mal posé

Le problème est par nature mal posé à cause des propriétés de l'équation de la chaleur, qui 'lisse' les informations de haute fréquence. Ceci peut être illustré avec l'exemple simple ci-dessous :

Soit l'équation de la chaleur simplifiée avec conditions de Dirichlet

$$\begin{cases} \frac{\partial \theta}{\partial t} - \Delta \theta = 0 & \text{dans } \Omega \times [0, T] \\ \theta = 0 & \text{sur } \partial\Omega \times [0, T] \\ \theta(x, 0) = \theta^0(x) & \text{dans } \Omega \end{cases}$$

On cherche à développer  $\theta(x, T) = f(x)$  en série de Fourier. A cette fin, on introduit le problème de valeurs propres

$$\begin{cases} \Delta \phi_k + \lambda_k \phi_k = 0 & \text{dans } \Omega \\ \phi_k = 0 & \text{sur } \partial\Omega \end{cases}$$

avec  $\|\phi_k\| = 1$ . Alors  $\theta_k(x, t) = \phi_k(x)e^{-\lambda_k t}$  ( $x \in \Omega$ ,  $t \in [0, T]$ ), est solution de l'équation de la chaleur. On en déduit que  $\|\theta_k(\cdot, T)\| \leq e^{-\lambda_k T}$  bien que  $\|\theta_k(\cdot, 0)\| = 1$ . Il résulte de la compacité de la résolvante du Laplacien en domaine borné que

$$\lim_{k \rightarrow \infty} \|\theta_k(\cdot, T)\| = 0.$$

La 1<sup>ère</sup> formule asymptotique de de Weyl permet d'affiner cette conclusion. En effet,  $\lambda_k \sim k^{\frac{2m}{d}}$ ,  $d$  étant la dimension de l'espace dans lequel on se place, et  $2m$  l'ordre de l'opérateur différentiel ( $2m = 2$  pour le Laplacien). Cette dernière expression montre que les contributions des modes de haute fréquence (lorsque  $k \rightarrow \infty$ ) n'ont quasiment plus d'effet à l'instant final. Ce phénomène est responsable du caractère mal posé du problème inverse, dont la résolution nécessitera des techniques particulières afin de reconstruire correctement ces contributions des modes de haute fréquence. Ce phénomène de 'lissage' signifie aussi que lors de la résolution du problème inverse, les modes de haute fréquence sont amplifiés. Un léger bruitage des mesures conduit alors à des solutions du problème inverse divergentes, et la résolution numérique est instable. En fait, si les données sont irrégulières le problème inverse n'a pas de solution car la température est très régulière dans  $\Omega$ .

On pourrait de même montrer que le flux extérieur  $\Phi$  à l'instant final n'a pas d'effet sur les données de mesure issues des capteurs noyés à l'intérieur de la structure.

### 6.2.4 Régularisation de Tikhonov

Pour rendre bien posé le problème de la minimisation de la fonctionnelle 6.2.5, on a recours à la stratégie de régularisation de Tikhonov ([EHN94]) qui consiste à rajouter un terme quadratique à la fonctionnelle qui lui confère des propriétés de stricte convexité. La nouvelle fonctionnelle s'écrit

$$J(\{\theta^0, \Phi\}) = \frac{1}{2} \int_0^T \sum_{k=1}^M \left( \theta(x_k, t) - \theta_k^d(t) \right)^2 dt + \frac{\epsilon}{2} \|\{\theta^0, \Phi\}\|^2 \quad (6.2.7)$$

où  $\|\cdot\|$  désigne la norme de l'espace dans lequel on recherche le couple  $\{\theta^0, \Phi\}$ . Le choix de cet espace a une influence déterminante sur les résultats, comme on le verra plus loin.

On cherchera donc à résoudre le problème

$$\left\{ \begin{array}{l} \text{Trouver } \{\theta^{0*}, \Phi^*\} \text{ tel que} \\ J(\{\theta^{0*}, \Phi^*\}) = \min_{\{\theta^0, \Phi\}} J(\{\theta^0, \Phi\}) \end{array} \right. \quad (6.2.8)$$

Il conduit inévitablement à une solution approchée, mais qui tend vers la solution idéale lorsque  $\epsilon \rightarrow 0$ . Par contre ce problème est mathématiquement bien posé, ce qui rend sa résolution numérique stable.

La minimisation de la fonctionnelle fait appel à des méthodes itératives telles que la méthode du gradient conjugué, dont la mise en œuvre ne sera pas expliquée ici.

Il existe d'autres méthodes de régularisation, comme celle de Morozov qui consiste à minimiser  $\|\{\theta^0, \Phi\}\|^2$  sous une contrainte du type  $|\theta - \theta^d| \leq \delta$  où  $\delta$  représente une incertitude de mesure.

### 6.3 Minimisation dans $L^2$

Dans ce qui suit, on se place dans le cas unidimensionnel :  $\Omega = [0, L]$ . Donc  $\partial\Omega = \{0, L\}$ . On notera  $x_1 = 0$ ,  $x_2 = L$ . La fonction  $\Phi(x, t)$ ,  $x \in \partial\Omega$ ,  $t \in [0, T]$ , peut alors s'écrire comme un couple de fonctions  $\{\Phi_1(t), \Phi_2(t)\}$  sur  $[0, T]$  où  $\Phi_1(t) = \Phi(0, t)$  et  $\Phi_2(t) = \Phi(L, t)$ .

L'espace  $L^2$  constitue un choix naturel pour la minimisation car il permet de développer la méthode de manière assez simple. On cherche donc les triplets  $\{\theta^0, \Phi_1, \Phi_2\}$  de températures initiales et de flux imposés à la frontière dans l'espace  $V = L^2(0, L) \times (L^2(0, T))^2$ , équipé du produit scalaire  $(u, v)_V = \int_0^L u_0 v_0 dx + \sum_{i=1}^2 \left\{ \int_0^T u_i v_i dt \right\}$  où  $u = \{u_0, u_1, u_2\}$  et  $v = \{v_0, v_1, v_2\}$ . On note  $\|\cdot\|_V$  la norme correspondante, i. e.  $\|v\|_V = \sqrt{(u, v)_V}$ .

On considère alors le problème

$$\left\{ \begin{array}{l} \text{Trouver } \{\theta^{0*}, \Phi_1^*, \Phi_2^*\} \in V \text{ tel que} \\ J(\{\theta^{0*}, \Phi_1^*, \Phi_2^*\}) = \min_{\{\theta^0, \Phi_1, \Phi_2\} \in V} J(\{\theta^0, \Phi_1, \Phi_2\}) \end{array} \right. \quad (6.3.1)$$

où

$$\begin{aligned} J(\{\theta^0, \Phi_1, \Phi_2\}) &= \frac{1}{2} \int_0^T \sum_{k=1}^M \left( \theta(x_k, t) - \theta_k^d(t) \right)^2 dt \\ &+ \frac{\epsilon}{2} \left[ \int_{\Omega} (\theta^0(x))^2 dx + \sum_{i=1}^2 \left\{ \int_0^T (\Phi_i(t))^2 dt \right\} \right] \end{aligned} \quad (6.3.2)$$

et où  $\theta$  est solution du problème d'évolution

$$\left\{ \begin{array}{ll} \rho c \frac{\partial \theta}{\partial t} - \frac{\partial}{\partial x} (\mathbf{K} \frac{\partial \theta}{\partial x}) = 0 & \text{dans } \Omega \times [0, T] \\ -(\mathbf{K} \frac{\partial \theta}{\partial x}(x_1, t)) + \alpha \theta(x_1, t) = \Phi_1(t) & \text{sur } [0, T] \\ (\mathbf{K} \frac{\partial \theta}{\partial x}(x_2, t)) + \alpha \theta(x_2, t) = \Phi_2(t) & \text{sur } [0, T] \\ \theta(x, 0) = \theta^0(x) & \text{dans } \Omega \end{array} \right. \quad (6.3.3)$$

Ce problème quadratique est bien posé car la fonctionnelle  $J$  est  $\alpha$ -convexe et différentiable. En effet, il suffit de vérifier la dépendance continue de la solution  $\theta$  par rapport aux données. Il est classique que, lorsque  $\{\theta^0, \Phi_1, \Phi_2\} \in V$ , alors la solution  $\theta$  de l'équation (6.3.3) vérifie  $\theta \in L^2(0, T; H^1(0, L))$  et donc  $\theta(x_k, t) \in L^2(0, T)$  avec dépendance continue par rapport aux données.

### 6.3.1 La méthode de l'état adjoint

On suit ici la méthode très classique [Lio68] dont la mise en œuvre ne pose aucune difficulté particulière. On en rappelle les étapes afin que le lecteur puisse éventuellement reproduire les calculs.

Les méthodes de descente nécessitent le calcul du gradient de la fonctionnelle. L'état adjoint, solution d'une équation de la chaleur auxiliaire, donne immédiatement accès au gradient.

On pose  $D = J(\{\theta^0 + \delta\theta^0, \Phi_1 + \delta\Phi_1, \Phi_2 + \delta\Phi_2\}) - J(\{\theta^0, \Phi\})$ . De la définition de la fonctionnelle, il s'ensuit que

$$\begin{aligned} D &= \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) \delta\theta(x_k, t) dt \\ &+ \epsilon \left[ \int_0^L \theta^0 \delta\theta^0 dx + \sum_{i=1}^2 \left\{ \int_0^T \Phi_i \delta\Phi_i dt \right\} \right] + \mathcal{O} \left\| \{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\} \right\|_V^2 \end{aligned} \quad (6.3.4)$$

où  $\delta\theta$  vérifie le système tangent :

$$\left\{ \begin{array}{ll} \rho c \frac{\partial \delta\theta}{\partial t} - \frac{\partial}{\partial x} (\mathbf{K} \frac{\partial \delta\theta}{\partial x}) = 0 & \text{dans } \Omega \times [0, T] \\ -(\mathbf{K} \frac{\partial \delta\theta}{\partial x}(x_1, t)) + \alpha \delta\theta(x_1, t) = \delta\Phi_1 & \text{sur } [0, T] \\ (\mathbf{K} \frac{\partial \delta\theta}{\partial x}(x_2, t)) + \alpha \delta\theta(x_2, t) = \delta\Phi_2 & \text{sur } [0, T] \\ \delta\theta(x, 0) = \delta\theta^0(x) & \text{dans } \Omega \end{array} \right. \quad (6.3.5)$$

On pose  $L = \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) \delta\theta(x_k, t) dt$ . Pour obtenir une expression du gradient, il est nécessaire d'écrire la différence  $D$  comme une forme linéaire continue des arguments  $\{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\}$ . Grâce à 6.3.4, il suffit en fait d'écrire  $L$  de cette façon. Pour cela, on a recours à la théorie du contrôle optimal, et on introduit l'état adjoint  $p(x, t)$  du système, donné par l'équation :

$$\left\{ \begin{array}{ll} -\rho c \frac{\partial p}{\partial t} - \frac{\partial}{\partial x} (\mathbf{K} \frac{\partial p}{\partial x}) = \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) \delta_{x_k} & \text{dans } \Omega \times [0, T] \\ -(\mathbf{K} \frac{\partial p}{\partial x}(x_1, t)) + \alpha p(x_1, t) = 0 & \text{sur } [0, T] \\ (\mathbf{K} \frac{\partial p}{\partial x}(x_2, t)) + \alpha p(x_2, t) = 0 & \text{sur } [0, T] \\ p(x, T) = 0 & \text{dans } \Omega \end{array} \right. \quad (6.3.6)$$

dont une formulation faible en espace s'écrit :

$$\left\{ \begin{array}{l} \forall w \in V \quad - \int_0^L \rho c \frac{\partial p}{\partial t} w dx + \int_0^L \mathbf{K} \frac{\partial p}{\partial x} \frac{\partial w}{\partial x} dx + \\ \sum_{i=1,2} \{ \alpha p(x_i, t) w(x_i, t) \} = \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) w(x_k, t) \\ p(x, T) = 0 \end{array} \right. \quad (6.3.7)$$

Dans (6.3.7), à chaque instant  $t$  on choisit  $w = \delta\theta$ . Après intégration en temps sur  $[0, T]$  et une intégration par parties du premier terme, on obtient

$$\begin{aligned} & \int_{[0,L] \times [0,T]} \rho c \frac{\partial \delta\theta}{\partial t} dx dt + \int_0^L \rho c p(x, 0) \delta\theta(x, 0) dx + \int_{[0,L] \times [0,T]} \mathbf{K} \frac{\partial p}{\partial x} \frac{\partial \delta\theta}{\partial x} dx dt \\ & + \sum_{i=1}^2 \left\{ \int_0^T \alpha p \delta\theta(x_i, t) dt \right\} = \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) \delta\theta(x_k, t) dt \end{aligned} \quad (6.3.8)$$

L'équation de la chaleur tangente peut s'écrire en formulation faible

$$\left\{ \begin{array}{l} \int_0^L \rho c \frac{\partial \delta\theta}{\partial t} v dx + \int_0^L \mathbf{K} \frac{\partial \delta\theta}{\partial x} \frac{\partial v}{\partial x} dx + \sum_{i=1,2} \delta\theta(x_i, t) v(x_i, t) \\ = \sum_{i=1,2} \alpha \delta\Phi_i(t) v(x_i, t) \quad \forall v \in V \\ \delta\theta(x, 0) = \delta\theta^0(x) \end{array} \right. \quad (6.3.9)$$

Dans (6.3.9), à chaque instant  $t$  on choisit  $v = p$ , puis on intègre en temps. On obtient

$$\begin{aligned} & \int_{[0,L] \times [0,T]} \rho c \frac{\partial \delta\theta}{\partial t} p dx dt + \int_{[0,L] \times [0,T]} \mathbf{K} \frac{\partial \delta\theta}{\partial x} \frac{\partial p}{\partial x} dx dt + \\ & \sum_{i=1,2} \left\{ \int_0^T \alpha \delta\theta(x_i, t) p(x_i, t) dt \right\} = \sum_{i=1,2} \left\{ \int_0^T \delta\Phi_i(t) p(x_i, t) dt \right\} \end{aligned} \quad (6.3.10)$$

En combinant les deux équations et en tenant compte du fait que  $\delta\theta(x, 0) = \delta\theta^0(x)$ , on obtient

$$L = \int_0^L \rho c p(x, 0) \delta \theta^0(x) dx + \sum_{i=1,2} \left\{ \int_0^T p(x_i, t) \delta \Phi_i(t) dt \right\} \quad (6.3.11)$$

$L(\{\delta \theta^0, \delta \Phi_1, \delta \Phi_2\})$  est une forme *linéaire* et *continue* de  $\{\delta \theta^0, \delta \Phi_1, \delta \Phi_2\}$  dans  $V$ . On en déduit par identification que le gradient de  $J$  peut être écrit en fonction de  $\theta^0$ , de  $\{\Phi_1, \Phi_2\}$  et de l'état adjoint  $p$  :

$$\nabla J(\{\theta^0, \Phi_1, \Phi_2\}) = \{\rho c p(x, 0) + \epsilon \theta^0, p(x_1, t) + \epsilon \Phi_1(t), p(x_2, t) + \epsilon \Phi_2(t)\} \quad (6.3.12)$$

### 6.3.2 Résultats

On considère une poutre de longueur  $L = 1$ . Pour fixer les idées, on normalise tous les paramètres physiques à l'unité :  $K(x) = 1$ ,  $\rho = 1$ ,  $c = 1$ ,  $\alpha = 1$ . Par ailleurs, on choisit  $T = 0.5$  et on discrétise par des éléments finis P1-Lagrange en espace (30 éléments) et par un schéma d'Euler rétrograde en temps (500 pas de temps). On mesure la température en trois capteurs équirépartis sur l'intervalle  $\Omega = [0, L]$ . On arrête l'algorithme de gradient conjugué lorsque l'écart entre les mesures et la température reconstruite aux points de mesure passe sous le seuil de  $10^{-6}$ , c'est-à-dire après environ 20 itérations. Les mesures aux capteurs  $\theta_k^d(t)$  correspondent à des données de

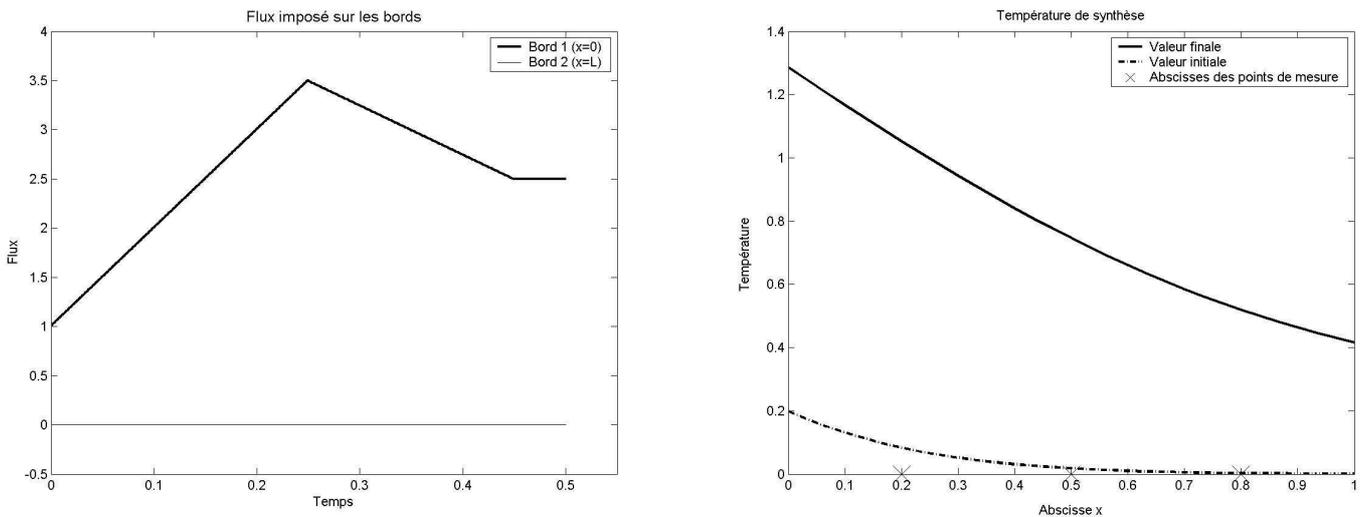


FIG. 6.1 – Evolution du flux imposé sur la frontière (à gauche), température initiale et finale (à droite)

synthèse issues de la résolution de l'équation de la chaleur (6.2.4). La température initiale et le flux imposé à la frontière sont décrits sur la figure 6.1. Tester un algorithme inverse sur des données de synthèse, donc atteignables, s'appelle parfois un crime inverse, traduction de *inverse crime*. Il s'agit d'une première étape de validation de l'algorithme, la stabilité devant faire l'objet de tests avec des données bruitées. Il se trouve que l'algorithme proposé présente de bonnes propriétés de stabilité par rapport aux erreurs de mesure, mais, dans un souci de concision, on omettra ici de développer cet aspect.

Quoique l'algorithme ait effectivement trouvé une température initiale et un flux tels que la température calculée coïncide presque exactement avec les mesures de

synthèse (fig 6.2), les premiers résultats mettent en évidence les limites pratiques de cette approche. En effet, par construction de l'algorithme, comme

$$p_n(x, T) = 0, \quad \forall x, \quad \forall n \in \mathbb{Z},$$

les valeurs de flux au temps final  $t = T$  ne varient pas au cours des itérations (fig 6.4). Dans le résultat final, même après convergence, ces valeurs seront égales à celles données lors de l'initialisation de l'algorithme, c'est à dire

$$\Phi_n(x, T) = \Phi_0(x, T) \quad \forall x, \quad \forall n \in \mathbb{Z}$$

Par conséquent, la solution sur  $\Phi$  est dégradée au voisinage de l'instant final ( $t \rightarrow T$ ) (fig 6.3). Ceci entraîne (fig 6.3) la dégradation partielle de la solution  $\theta(x, T)$ , qui est pourtant celle que l'on recherche.

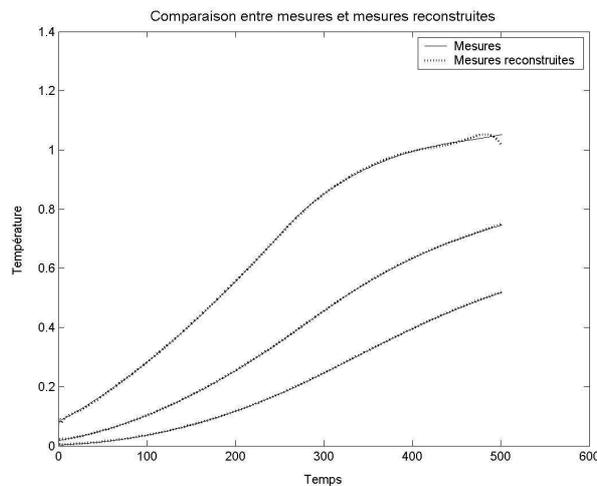


FIG. 6.2 – fonctionnelle  $L^2$  : Température cible et température reconstruite au niveau des capteurs

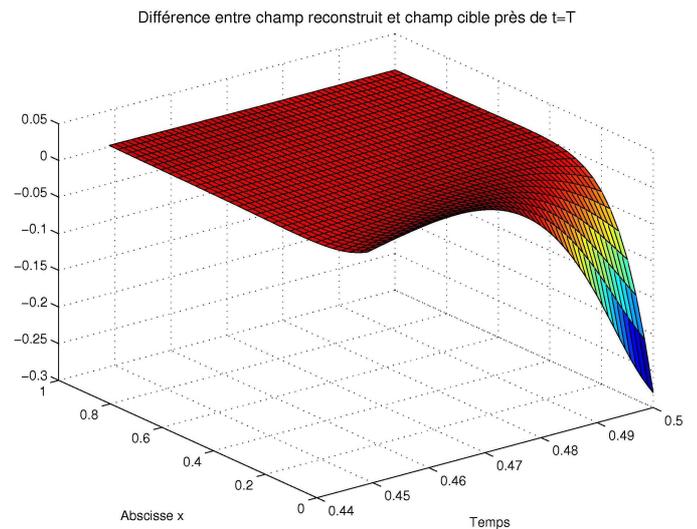
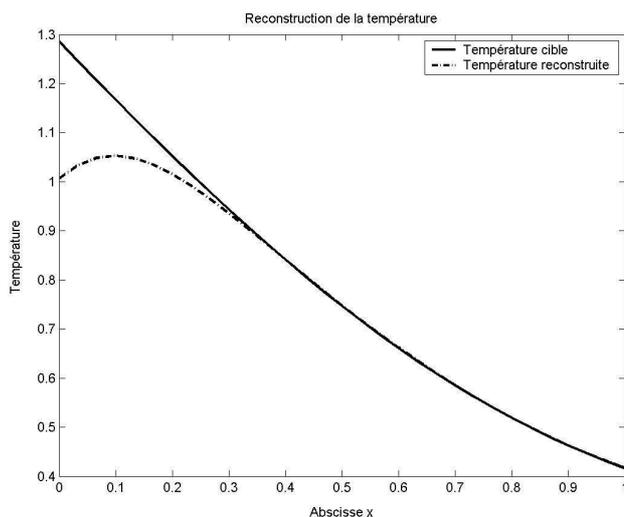


FIG. 6.3 – fonctionnelle  $L^2$  : écart entre la température cible et la température reconstruite à l'instant final (à gauche), et à tout instant (à droite)

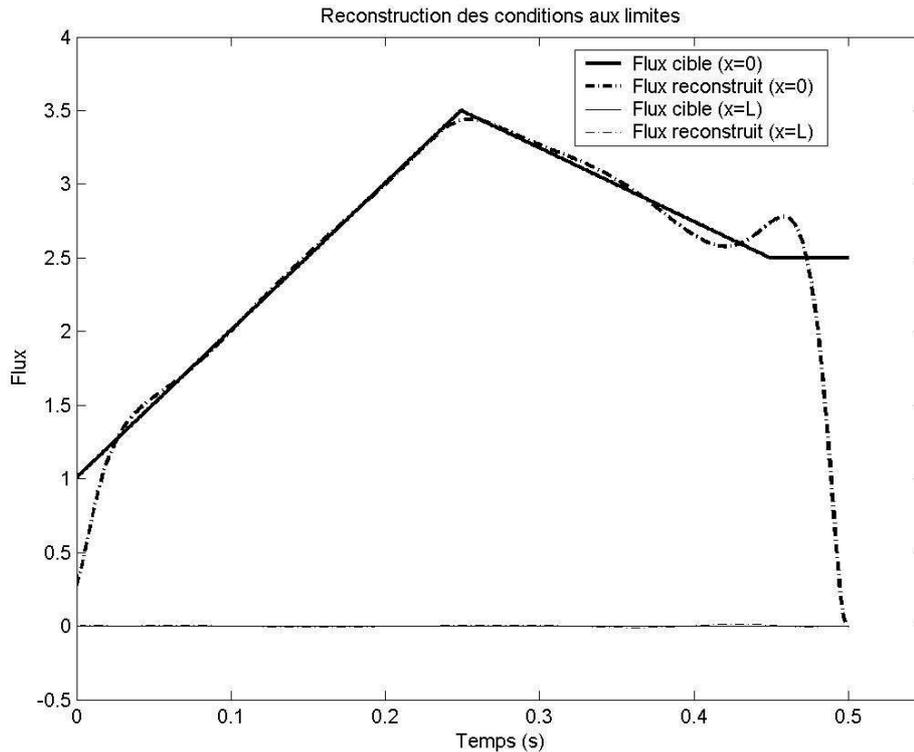


FIG. 6.4 – fonctionnelle  $L^2$  : écart entre le flux imposé et reconstruit

Cette difficulté est répertoriée dans la littérature [Ali94] mais on ne connaît pas de traitement systématique débouchant sur des algorithmes ne nécessitant pas d'intervention manuelle. L'objet du paragraphe suivant est de proposer une stratégie de régularisation des fonctions inconnues  $\theta^0(x)$  et  $\Phi(x, t)$  qui **libère la valeur finale de l'état adjoint**. A cette fin, on quitte le cadre  $L^2$ .

## 6.4 Reconstruction dans $H^1$

### 6.4.1 Position du problème

On rappelle que  $\Omega = [0, L]$ . La frontière  $\partial\Omega = \{0, L\}$  du domaine n'est ainsi constituée que des deux points  $x_1 = 0$  et  $x_2 = L$ . Comme précédemment, la fonction  $\Phi(x, t)$ ,  $x \in \partial\Omega$ ,  $t \in [0, T]$ , peut s'écrire comme un couple de fonctions  $\{\Phi_1(t), \Phi_2(t)\}$  sur  $[0, T]$  où  $\Phi_1(t) = \Phi(0, t)$  et  $\Phi_2(t) = \Phi(L, t)$ . L'idée est de rechercher le triplet  $\{\theta^0, \Phi_1, \Phi_2\}$  dans l'espace  $W = H^1(\Omega) \times (H^1(0, T))^2$ . Ceci aura pour effet de fournir des solutions présentant plus de régularité, notamment près des bords. Comme on le verra par la suite, l'algorithme ainsi construit présentera l'avantage de ne pas figer la valeur de  $\Phi(x, T)$ .

On équipe l'espace  $W$  du produit scalaire

$$(u, v)_W = \bar{\alpha} \int_0^L u_0 v_0 dx + \tilde{\alpha} \int_0^L \frac{du_0}{dx} \frac{dv_0}{dx} dx + \sum_{i=1}^2 \left\{ \bar{\beta} \int_0^T u_i v_i dt + \tilde{\beta} \int_0^T \frac{du_i}{dt} \frac{dv_i}{dt} dt \right\} \quad (6.4.1)$$

où  $u = \{u_0, u_1, u_2\} \in W$ ,  $v = \{v_0, v_1, v_2\}$  et où  $\bar{\alpha}, \tilde{\alpha}, \bar{\beta}, \tilde{\beta} \in \mathbb{R}^+$  désignent des constantes positives ajustables.

On note  $\left\| \{\theta^0, \Phi_1, \Phi_2\} \right\|_W = \sqrt{(\{\theta^0, \Phi_1, \Phi_2\}, \{\theta^0, \Phi_1, \Phi_2\})_W}$

On considère le nouveau problème

$$\left\{ \begin{array}{l} \text{Trouver } \{\theta^{0*}, \Phi_1^*, \Phi_2^*\} \in W \text{ tel que} \\ J(\{\theta^{0*}, \Phi_1^*, \Phi_2^*\}) = \min_{\{\theta^0, \Phi_1, \Phi_2\} \in W} J(\{\theta^0, \Phi_1, \Phi_2\}) \end{array} \right. \quad (6.4.2)$$

où

$$J(\{\theta^0, \Phi_1, \Phi_2\}) = \frac{1}{2} \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t))^2 dt + \frac{\epsilon}{2} \left\| \{\theta^0, \Phi_1, \Phi_2\} \right\|_W^2 \quad (6.4.3)$$

et où  $\theta$  est encore solution de l'équation (6.3.3).

## 6.4.2 Calcul du gradient

La fonctionnelle à minimiser s'écrit

$$J(\{\theta^0, \Phi_1, \Phi_2\}) = \frac{1}{2} \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t))^2 dt + \frac{\epsilon}{2} \left[ \bar{\alpha} \int_0^L (\theta^0(x))^2 dx + \tilde{\alpha} \int_0^L \left( \frac{\partial \theta^0(x)}{\partial x} \right)^2 dx + \sum_{i=1,2} \left( \bar{\beta} \int_0^T (\Phi_i(t))^2 dt + \tilde{\beta} \int_0^T \left( \frac{\partial \Phi_i(t)}{\partial t} \right)^2 dt \right) \right] \quad (6.4.4)$$

et l'équation de la chaleur sous forme variationnelle s'écrit :

$$\left\{ \begin{array}{l} \forall v \in W \quad \int_0^L \int_0^T \rho c \frac{\partial \theta}{\partial t} v dx dt + \int_0^L \int_0^T K \frac{\partial \theta}{\partial x} \frac{\partial v}{\partial x} dx dt \\ + \sum_{i=1,2} \int_0^T \alpha \theta(x_i, t) v(x_i, t) dt = \sum_{i=1,2} \int_0^T \Phi_i(t) v(x_i, t) dt \\ \theta(x, 0) = \theta^0(x) \end{array} \right. \quad (6.4.5)$$

où  $(x_1 = 0, x_2 = L)$ .

Le calcul du gradient de la fonctionnelle nécessite l'introduction, comme plus haut, de l'équation adjointe à l'équation de la chaleur qui s'écrit, sous une forme variationnelle,

$$\left\{ \begin{array}{l} \forall w \in W \quad - \int_0^L \int_0^T \rho c \frac{\partial p}{\partial t} w dx dt + \int_0^L \int_0^T K \frac{\partial p}{\partial x} \frac{\partial w}{\partial x} dx dt + \\ \sum_{i=1,2} \int_0^T \alpha p(x_i, t) w(x_i, t) dt = \int_0^T \sum_{k=1}^M (\theta(x_k, t) - \theta_k^d(t)) w(x_k, t) dx dt \\ p(x, T) = 0 \end{array} \right. \quad (6.4.6)$$

On pose  $D = J(\{\theta^0 + \delta\theta^0, \Phi_1 + \delta\Phi_1, \Phi_2 + \delta\Phi_2\}) - J(\{\theta^0, \Phi_1, \Phi_2\})$ . Les mêmes étapes de calcul conduisent à une expression de  $D$  en fonction de  $p(x, t)$  :

$$D = \int_0^L p(x, 0)\delta\theta^0(x) dx + \sum_{i=1,2} \int_0^T p(x_i, t)\delta\Phi_i(t) dt + \epsilon(\{\theta^0, \Phi_1, \Phi_2\}, \{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\})_W + \mathcal{O}\left\|\{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\}\right\|_W^2 \quad (6.4.7)$$

où  $x_1 = 0, x_2 = L$ . Mais pour déterminer le gradient de  $J$  on cherche un triplet  $Q = \{P, Q_1, Q_2\}$  tel qu'on puisse écrire  $D$  sous la forme

$$D = \left(\{P, Q_1, Q_2\}, \{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\}\right)_W + \mathcal{O}\left\|\{\delta\theta^0, \delta\Phi_1, \delta\Phi_2\}\right\|_W^2 \quad (6.4.8)$$

La recherche de  $Q = \{P, Q_1, Q_2\}$  consiste alors à résoudre trois problèmes indépendants.

Tout d'abord, on résout, par reconstructionapproximation de Galerkin par exemple, l'équation variationnelle

$$\left\{ \begin{array}{l} \text{Trouver } P \in H^1(0, L) \text{ tel que} \\ \int_0^L p(\cdot, 0)w = \tilde{\alpha} \int_0^L Pw + \bar{\alpha} \int_0^L \frac{dP}{dx} \frac{dw}{dx} \quad \forall w \in H^1(0, L) \end{array} \right. \quad (6.4.9)$$

La solution existe car  $p \in L^2(0, T; H^1(0, L) \cap (C)^0(0, T; L^2(0, L)))$ . On peut donc définir  $p(x, 0) \in L^2(0, L)$ .

Cette équation s'écrit aussi sous la forme classique

$$\left\{ \begin{array}{l} -\tilde{\alpha} \frac{d^2 P}{dx^2} + \tilde{\alpha} P = p(\cdot, 0) \quad \text{sur } [0, L] \\ \frac{dP}{dx}(0) = \frac{dP}{dx}(L) = 0 \end{array} \right. \quad (6.4.10)$$

La détermination de  $Q_1$  et  $Q_2$  se fait de manière analogue. Le problème à résoudre est, pour  $i = 1, 2, x_1 = 0, x_2 = L$  :

$$\left\{ \begin{array}{l} \text{Trouver } Q_i \in H^1(0, T) \text{ tel que} \\ \int_0^T p(x_i, \cdot)w = \tilde{\beta} \int_0^T Q_i w + \bar{\beta} \int_0^T \frac{dQ_i}{dt} \frac{dw}{dt} \quad \forall w \in H^1(0, T) \end{array} \right. \quad (6.4.11)$$

Ce problème est analogue au problème précédent, à la différence près qu'il est posé sur  $[0, T]$ . La régularité de l'état adjoint  $p$  suffit à établir l'existence d'une solution unique, que l'on peut calculer par la méthode des éléments finis.

Remarquer que le problème aux limites correspondant s'écrit

$$\left\{ \begin{array}{l} -\tilde{\beta} \frac{d^2 Q_i}{dt^2} + \tilde{\beta} Q_i = p(x_i, t) \quad \text{sur } [0, T] \\ \frac{dQ_i}{dt}(0) = \frac{dQ_i}{dt}(T) = 0 \end{array} \right. \quad (6.4.12)$$

On voit donc que le flux régularisé  $Q_i$  dépend de toutes les valeurs de  $p(x_i, t)$  et ne s'annule pas systématiquement à l'instant  $t = T$ .

Ainsi, le gradient de  $J(\{\theta^0, \Phi_1, \Phi_2\})$  s'écrit :

$$\nabla J(\{\theta^0, \Phi_1, \Phi_2\}) = \left\{ P(x) + \epsilon\theta^0, Q_1(t) + \epsilon\Phi_1, Q_2(t) + \epsilon\Phi_2 \right\} \quad (6.4.13)$$

D'après la remarque précédente, le flux à l'instant final  $\Phi(x, T)$  ne reste pas constant au fil des itérations.

Par ailleurs, ni la fonctionnelle (6.4.3) à minimiser ni l'espace sur lequel on minimise n'impose de condition de compatibilité entre le flux  $\Phi_i(x, 0)$  et celui associé à la température initiale, soit  $\mp K \frac{d\theta^0}{dx}(x_i) + \alpha\theta^0(x_i)$ ,  $i = 1, 2$ . Bien que cette dernière quantité n'ait pas de sens lorsque  $\theta^0 \in H^1(0, L)$ , on peut imposer cette compatibilité de façon approchée par projection sur quelques modes thermiques et pénalisation. On ne détaillera pas ici l'analyse correspondante.

### 6.4.3 Résultats

On étudie la même situation qu'au paragraphe 6.3. Malgré l'absence de compatibilité à l'instant initial, la température finale (fig 6.5) est reconstruite de façon beaucoup plus précise avec la nouvelle fonctionnelle (6.4.3) qu'avec celle utilisée au paragraphe 6.3. On note que la température reconstruite se superpose presque exactement à la température cible. L'imprécision sur l'un des flux à l'instant initial se trouve gommée par le caractère régularisant de l'équation de la chaleur. Elle compense par ailleurs l'imprécision sur la température initiale au voisinage du bord.

Dans le cas où les données de synthèse se trouvent remplacées par des données bruitées, on constate une très bonne stabilité de l'algorithme. Par souci de concision, on ne présente pas ici cette étape importante de la validation.

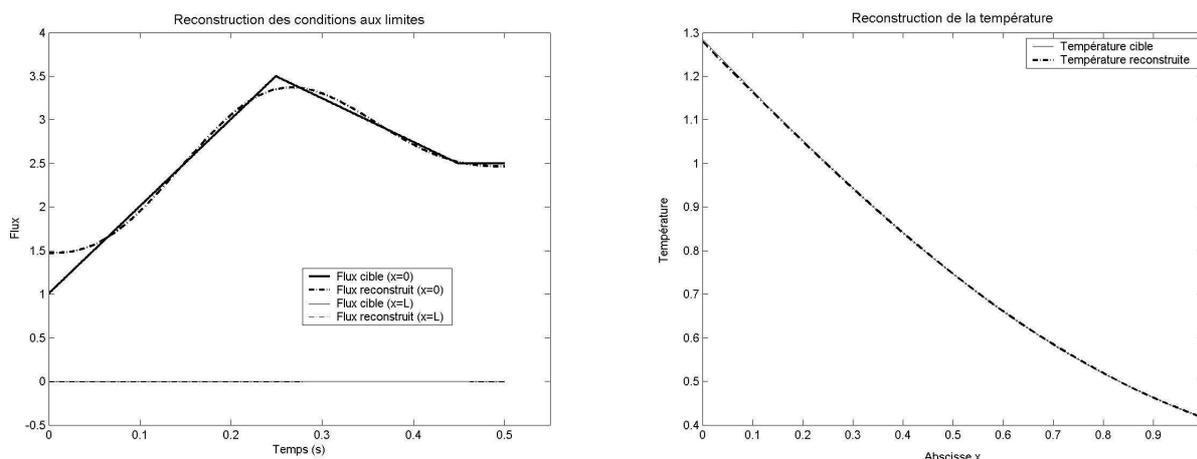


FIG. 6.5 – Reconstruction dans  $H^1$  : flux (à gauche) et température finale (à droite). Les grandeurs cibles et les grandeurs reconstruites sont quasiment superposées

## 6.5 Conclusions

Le cadre  $H^1$  pour minimiser la fonctionnelle d'erreur permet une bien meilleure reconstruction de la température finale que le cadre  $L^2$ . La méthode proposée s'étend presque naturellement au cas multidimensionnel.

Par ailleurs, le recouvrement partiel des intervalles de temps de reconstruction successifs ainsi que la sauvegarde des directions de descente en vue d'une prise en compte ultérieure devraient accélérer nettement la reconstruction au fil du temps.

Enfin, une projection sur base réduite pourrait s'avérer très utile pour traiter les structures tridimensionnelles.

La reconstruction des précontraintes d'origine thermique ne constitue qu'un post-traitement de type thermo-élastique. Il est alors essentiel de bien connaître les vraies conditions aux limites qui évoluent également avec la température. La prise en compte de la température réelle et des sensibilités des modes à la température fait l'objet d'une recherche active.



# Bibliographie

- [AAR95] O. M. Alifanov, E. A. Artyukhin, and S. V. Rumyantsev. *Extreme methods for solving ill-posed problems with applications to inverse heat transfer problems*. Begell house, inc., 1995.
- [Ali94] O. M. Alifanov. *Inverse Heat Transfer Problems*. Springer-Verlag, New York, 1994.
- [BBC85] J. V. Beck, B. Blackwell, and C. ST. Clair. *Inverse heat conduction, ill-posed problems*. Wiley interscience, New York, 1985.
- [BBH96] J. V. Beck, B. Blackwell, and A. Hajisheikh. Comparison of some inverse heat conduction methods using experimental data. *Int. J. Heat Mass Transfer*, 39(17) :3649–3657, 1996.
- [BMG04] M. Basseville, L. Mevel, and M. Goursat. Statistical model-based damage detection and localization : subspace-based residuals and damage-to-noise sensitivity ratios. *Journal of Sound and Vibration*, 275 :769–794, 2004.
- [CdCV91] L. D. Chiwiacowski and H. F. de Campos Velho. Different approaches for the solution of a backward heat conduction problem. *Inverse Problems in Engng*, 11(6) :471–494, 1991.
- [Cre04] Ch. Cremona. Dynamic monitoring applied to the detection of structural modifications : A high speed railway bridge study. In *Progress in Structural Engineering and Materials*, volume 3, pages 147–161. Wiley, 2004.
- [EHN94] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, 1994.
- [FHS<sup>+</sup>03] Ch. Farrar, F. Hemez, D. Shunk, D. Stinemat, and B. Nadler. A review of shm litterature : 1996-2001. *Los Alamos National Laboratory Internal Reports*, 2003.
- [Han95] M. Hanke. *Conjugate gradient type methods for ill-posed problems*. Longman Scientific and Technical, 1995.
- [IRvA<sup>+</sup>01] D. Inaudi, A. Rufenacht, B. von Arx, S. Vurpillot, H.P. Noher, and Glisic. Monitoring a concrete arch bridge during construction. In *3rd International conference on arch bridges, 2001, Paris, France*, 2001.
- [JzB91] Y. Jarny, M. N. Özisik, and J. P. Bardou. A general optimization method using adjoint equation for solving multidimensional inverse heat conduction. *Int. J. Heat Mass Transfer*, 34(11) :2911–2919, 1991.
- [Kul02] J. Kullaa. Elimination of environmental influences from damage-sensitive features in a structural health monitoring system. In *Proceedings of the*

- First European Workshop on Structural Health Monitoring, Paris, July 10-12, 2002, Onera. DEStech Publications, pages 742–749, 2002.*
- [Kul03] J. Kullaa. Is temperature measurement essential in structural health monitoring? In *Proceedings of the 4th International Workshop on Structural Health Monitoring, Stanford, CA. September 15-17, 2003. Stanford University. DEStech, pages 717–724, 2003.*
- [Kul04] J. Kullaa. Latent variable models to eliminate the environmental effects in structural health monitoring. In *Proceedings of the Third European Conference on Structural Control. Volume II. R. Flesch, H. Irschik and M. Krommer (eds.). Vienna, Austria, July 12-15, 2004. Vienna University of Technology, pages S5–55–S5–58, 2004.*
- [Lio68] J-L. Lions. *Optimal control of systems governed by PDEs*. Dunod, 1968.
- [MG04] L. Mevel and M. Goursat. A complete scilab toolbox for output-only identification. In *Proceedings of International Modal Analysis Conference, Dearborn, Mi, Jan. 26-29, 2004, 2004.*
- [MRdCV99] W. B. Muniz, F. M. Ramos, and H. F. de Campos Velho. A comparison of some inverse methods for estimating the initial condition equation of the heat. *J. Comp. Appl. Math.*, 103 :145–163, 1999.
- [Pee00] B. Peeters. *System identification and damage detection in civil engineering*. PhD thesis, 2000.
- [RB86] M. Raynaud and J. Bransier. A new finite-difference method for the nonlinear heat conduction problem. *Numerical heat transfer*, 9(1) :27–42, 1986.
- [RBS<sup>+</sup>00] R.G. Rohrmann, M. Baessler, S. Said, W. Schmid, and W.F. Ruecker. Structural causes of temperature affected modal data of civil structures obtained by long time monitoring. In *Proceedings of IMAC-XVIII : A Conference on Structural Dynamics, San Antonio, Texas, USA, February 7-10, 2000. Society for Experimental Mechanics, Bethel, Connecticut, pages 1–7, 2000.*
- [SCF00] H. Sohn, J.A. Czarnecki, and C.R. Farrar. Structural health monitoring using statistical process control. *Journal of Structural Engineering*, 126(11) :1356–1363, 2000.
- [SFHW01] H. Sohn, C.R. Farrar, N.F. Hunter, and K. Worden. Structural health monitoring using statistical pattern recognition techniques. *Journal of dynamic systems, measurement, and control*, 123 :706–711, 2001.
- [SWF03] H. Sohn, K. Worden, and C.R. Farrar. Statistical damage classification under changing environmental and operational conditions. *Journal of Intelligent Material Systems and Structures*, 13(9) :561–574, 2003.
- [Won04] Kai-Yuen Wong. Instrumentation and health monitoring of cable-supported bridges. *Structural Control and Health Monitoring*, 11 :91–124, 2004.
- [zO00] M. N. Özisik and H. R. B. Orlande. *Inverse Heat Transfer*. Taylor and Francis, 2000.

## Chapitre 7

# Théorie semi-quadratique, régularisation et estimation robuste

*Pierre Charbonnier,  
Jean-Philippe Tarel,  
Sio-Song Ieng*

### 7.1 Introduction

Les images numériques ont de plus en plus de domaines d'application de nos jours. Il convient d'ailleurs de prendre ici le mot "image" au sens large de cartographie d'une propriété physique, éventuellement complexe ou vectorielle, en tout cas non limitée à l'intensité photométrique. Ces images doivent souvent être obtenues par le calcul, à partir de données issues de capteurs, voire à partir d'autres images. Les images optiques elle-mêmes sont issues de systèmes imageurs qui peuvent introduire du flou ou des perturbations aléatoires – le bruit – dont il convient de les débarrasser pour les rendre exploitables. On peut parler, de manière générique, de *reconstruction* d'image.

Le terme *vision par ordinateur*, quant à lui, regroupe les techniques visant à analyser automatiquement les images afin d'en ressortir des informations utilisables par des systèmes de détection, de commande ou de reconnaissance. L'approche classique consiste à extraire dans les images des primitives géométriques qui sont ensuite abstraites, grâce à l'utilisation de modèles mathématiques. Il s'agit alors d'estimer les paramètres permettant d'ajuster au mieux un modèle aux données observées. Plus récemment, on a vu apparaître des méthodes, rendues accessibles par la montée en puissance des ordinateurs, qui utilisent l'ensemble des valeurs de l'image. Elles capturent l'apparence globale des objets à travers des modèles statistiques. L'analyse d'une image implique, là encore, un ajustement de paramètres.

Ces deux types d'applications, reconstruction et analyse d'images, peuvent sembler très différents. Dans les deux cas, cependant, on peut considérer que les données  $p$  auxquelles nous avons accès sont une réalisation bruitée d'une relation fonctionnelle  $\mathcal{R}(f)$  dont on souhaite estimer les paramètres  $f$ . Pour résoudre ce problème inverse, il est nécessaire d'introduire des connaissances sur le bruit. De plus, il est souvent souhaitable – et parfois nécessaire – d'introduire une connaissance *a priori* sur la nature de la solution recherchée, on parle alors de régularisation. Le cadre bayésien est le plus générique pour spécifier ces informations de nature statistique.

Le paragraphe 7.2 est consacré à la description du modèle et des principaux estimateurs associés. Nous verrons également que les hypothèses gaussiennes classiquement utilisées sont inadéquates. Ainsi, dans le cas de l'estimation de paramètres, elles ne permettent pas de gérer correctement la présence de données erronées. En reconstruction d'image, elles entraînent un lissage excessif des contours. Nous proposons au paragraphe 7.3 une présentation unifiée de modèles basés sur des densités de probabilité non gaussiennes (ou d'énergies non quadratiques). Ces modèles permettent de relâcher le comportement gaussien lorsque cela devient nécessaire. La théorie semi-quadratique précise les conditions à imposer aux fonctions employées pour autoriser un tel comportement. On montre que cela revient à introduire une variable auxiliaire dont le rôle est à la fois de "marquer" les données et de partiellement linéariser le problème, simplifiant ainsi l'algorithmique associée. Nous illustrons au paragraphe 7.4 deux applications d'estimation robuste de paramètres dans le cadre de la détection automatique de la signalisation routière. Nous montrons également le gain de qualité obtenu en reconstruction d'image régularisée.

## 7.2 Le modèle quadratique et ses limites

### 7.2.1 Modèle génératif linéaire

Dans un premier temps, on se donne un modèle déterministe de formation des données, ou modèle génératif. Dans la pratique, un modèle linéaire s'avère souvent suffisant et on écrira cette relation (en variables discrètes) :

$$p = Rf. \quad (7.2.1)$$

La forme de la matrice  $R$ , connue, dépend de l'application traitée. Dans le cas de la reconstruction d'image, les données  $p$  et l'image inconnue,  $f$ , sont arrangées sous forme de vecteurs par ordonnancement lexicographique. Le vecteur  $p$  peut être vu comme une somme des colonnes de la matrice  $R$ , chacune étant pondérée par la valeur du pixel de  $f$  correspondant. On comprend ainsi que chaque colonne de  $R$  représente la fonction d'étalement d'un point de l'espace objet (en anglais *Point Spread Function* ou PSF). Ce modèle s'applique par exemple au cas de la reconstruction tomographique, où la PSF est différente pour chaque point. Cela peut également être le cas en déconvolution mais, la plupart du temps, la PSF est considérée spatialement invariante. Dans le cas du débruitage, la matrice  $R$  est simplement égale à l'identité.

En ce qui concerne l'analyse d'images, prenons l'exemple de la régression quadratique. On dispose d'une liste de coordonnées  $\{x_i, y_i\}_{i=1\dots N}$  de points d'intérêt, extraits d'une image par un détecteur, et que l'on suppose alignés selon une courbe. Une relation sous-jacente de la forme  $y = a + bx + cx^2$  est vérifiée pour chaque couple de coordonnées. La relation (7.2.1) s'écrit donc en collectant les ordonnées  $\{y_i\}$  dans le vecteur  $p$ . Les inconnues,  $a$ ,  $b$  et  $c$  forment le vecteur  $f$  et la matrice  $R$  est formée d'un vecteur-colonne de 1, du vecteur des abscisses  $\{x_i\}$ , et du vecteur des monômes de degré deux  $\{x_i^2\}$ . Ce modèle peut aisément être étendu à des combinaisons linéaires d'autres fonctions que les monômes. On peut ainsi modéliser des relations non linéaires entre les coordonnées, mais le modèle génératif (7.2.1) demeure linéaire par rapport à  $f$ .

Le modèle (7.2.1) résume la partie déterministe de la connaissance que l'on peut avoir du processus de formation des données. Il est cependant incomplet, car il ne tient

pas compte de la nature des perturbations aléatoires sur l'observation, ou bruit. Très souvent, celui-ci est considéré additif et indépendant de  $f$  et (7.2.1) devient alors :

$$p = Rf + n. \quad (7.2.2)$$

Dans le cadre de l'estimation bayésienne, on adjoint à ce modèle des informations de nature statistique, sur les différentes variables qui interviennent.

## 7.2.2 Structure probabiliste et estimation bayésienne

Une première densité de probabilité intéressante traduit le caractère aléatoire des mesures :  $\mathcal{P}(p|f)$ . Elle quantifie la vraisemblance d'une observation connaissant les valeurs des paramètres. Il s'agit en fait de la densité de probabilité du bruit,  $n = p - Rf$ . Une hypothèse très classique considère que le bruit est additif, gaussien, centré, isotrope, indépendant et identiquement distribué (iid). Dans ce cas :

$$\mathcal{P}(p|f) \propto \exp\left(-\frac{1}{2\sigma^2}\|p - Rf\|^2\right). \quad (7.2.3)$$

On peut rechercher les paramètres rendant la plus plausible l'observation en maximisant la vraisemblance. Dans le cas gaussien, on retrouve l'estimateur classique des moindres carrés :

$$\hat{f} = \arg \min_f \frac{1}{2\sigma^2} \|p - Rf\|^2. \quad (7.2.4)$$

L'estimateur du maximum de vraisemblance (MV) présente de bonnes propriétés statistiques [Kay93]. Il considère cependant la variable  $f$  comme un paramètre déterministe. Au contraire, l'approche bayésienne considère  $f$  comme une variable aléatoire dont la distribution  $\mathcal{P}(f)$  traduit la connaissance *a priori* sur la solution recherchée. Dans ce cadre, la probabilité associée à  $p$  est obtenue en intégrant par rapport à  $f$  la vraisemblance conjointe :

$$\mathcal{P}(p) = \int \mathcal{P}(p, f) df = \int \mathcal{P}(p|f) \mathcal{P}(f) df. \quad (7.2.5)$$

Cependant, cette intégrale est souvent inexploitable analytiquement et on a recours à une estimation ponctuelle qui correspond implicitement à une approximation classique en inférence bayésienne. Elle considère que la distribution conjointe est suffisamment "piquée" pour qu'on puisse approcher son intégrale par la hauteur du pic multipliée par sa "largeur",  $\sigma_{pic}$  :  $\mathcal{P}(p) \simeq \sigma_{pic} \max_f \mathcal{P}(p, f) \propto \mathcal{P}(p|\hat{f}) \mathcal{P}(\hat{f})$ , où l'estimée  $\hat{f}$  est donnée par :

$$\hat{f} = \arg \max_f \mathcal{P}(p|f) \mathcal{P}(f). \quad (7.2.6)$$

La distribution maximisée étant proportionnelle à la loi a posteriori  $\mathcal{P}(f|p)$ , cet estimateur est appelé maximum a posteriori (MAP). Les connaissances *a priori* sur  $f$  prises en compte par le MAP peuvent s'avérer très utiles dans des problèmes d'estimation comme la régression, par exemple. On peut ainsi limiter la sensibilité aux bruits, l'influence des forts degrés ou bien contraindre les coefficients de la courbe autour d'une valeur connue ou prédite, notamment dans le cas d'un algorithme de suivi. Pour cela, on impose à  $f$  une densité de probabilité, par exemple gaussienne,  $\mathcal{N}(\bar{f}, \Sigma)$  :

$$\mathcal{P}(f) \propto \exp\left(-\frac{1}{2}(f - \bar{f})^T \Sigma^{-1} (f - \bar{f})\right). \quad (7.2.7)$$

En reconstruction d'image, introduire une information *a priori* est non seulement utile, mais absolument nécessaire. En effet, les problèmes inverses rencontrés sont en général mal posés et la résolution au sens des moindres carrés est instable. Une technique classique, la régularisation de Tikhonov, pénalise la norme de  $f$  ou, plus souvent, celle de son gradient. Cela revient également à introduire un *a priori* gaussien :

$$\mathcal{P}(f) \propto \exp\left(-\frac{1}{2\sigma^2}\|\nabla f\|^2\right). \quad (7.2.8)$$

Dans le cas de distributions gaussiennes, en passant au logarithme dans (7.2.6), on retrouve l'estimateur des moindres carrés pénalisés. Soit, dans l'exemple ci-dessus, en collectant les constantes :

$$\hat{f} = \arg \min_f \|p - Rf\|^2 + \lambda^2 \|\nabla f\|^2. \quad (7.2.9)$$

Lorsque la contrainte porte sur  $f$  et non sur son gradient, on retrouve le filtre classique de Wiener. Notons que l'estimation au sens du MV peut être vue comme le cas limite du MAP où  $\mathcal{P}(f)$  est uniforme.

### 7.2.3 Limites des estimateurs gaussiens

Les densités de probabilité (*pdf* selon l'abréviation anglaise) mises en jeu dans le cadre de l'estimation bayésienne sont souvent de forme gaussienne :

$$\mathcal{P}(u) \propto \exp\left(-\frac{1}{2} \sum_{i=1}^N u_i^2\right), \quad (7.2.10)$$

où la variable muette  $u_i$  peut représenter soit un résidu  $n_i = (p - Rf)_i$  (cas de l'estimation) soit la valeur d'un pixel  $f_i$ , d'une composante ou de la norme du gradient :  $(\nabla_x f)_i$ ,  $(\nabla_y f)_i$  ou  $(|\nabla f|)_i$  (cas de la régularisation, cf. §7.4.3). Maximiser  $\mathcal{P}(u)$  est équivalent à minimiser une forme quadratique :

$$J(f) = \frac{1}{2} \sum_{i=1}^N u_i^2. \quad (7.2.11)$$

Malheureusement, ce type de modèle est notoirement sensible aux fortes déviations  $u_i$ . Prenons l'exemple de la régression linéaire (cf. figure 7.1). La présence d'une seule donnée erronée suffit à créer une valeur forte de résidu. L'algorithme chargé de minimiser  $J(f)$  a tendance à limiter en priorité ce résidu de la seule façon possible, c'est-à-dire en agissant sur les coefficients  $f$  du modèle. On dit que le point de rupture de cet estimateur est de 0 : une seule donnée erronée suffit à modifier le résultat de l'estimation.

Dans le cas de la régularisation, la fonction quadratique s'oppose à la création de forts gradients dans la solution. Celle-ci aura donc un aspect trop lisse (cf. fig. 7.7(e) ou 7.8(c)). On souhaiterait, au contraire, obtenir une solution formée de régions lisses, séparées par des bords francs.

Le problème inhérent au modèle gaussien est qu'il ne considère qu'un seul type de données alors qu'il y en a deux, en réalité. Dans le cas de la régression, on distingue les données conformes au modèle (ou "inlier") qui correspondent aux résidus faibles tandis que les données erronées (ou "outlier") entraînent de fortes déviations. Il s'agit

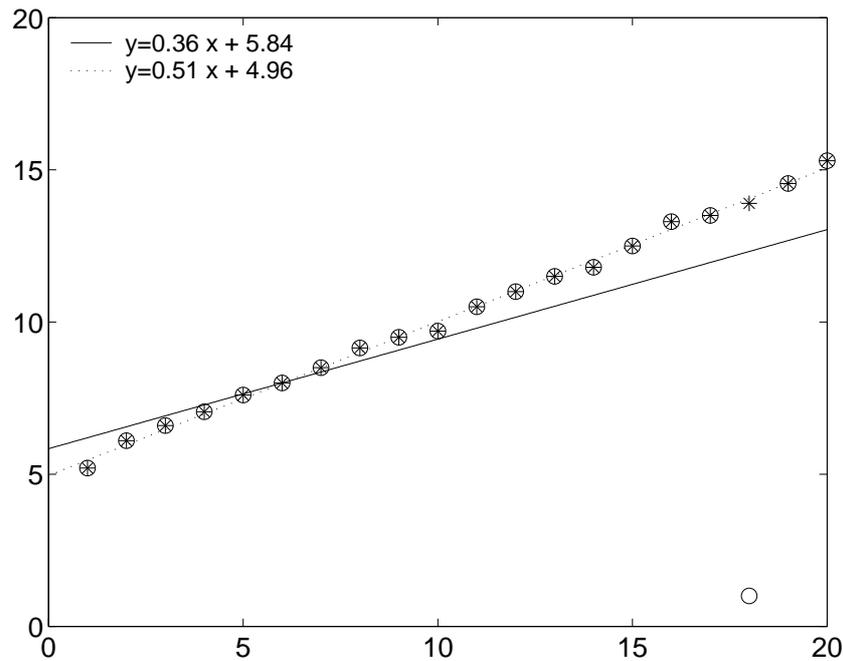


FIG. 7.1 – Influence d’une donnée erronée sur l’estimation de paramètres au sens des moindres carrés : la droite de régression en pointillés correspond à la série marquée par des étoiles, celle en trait plein correspond à la série marquée par des cercles.

de trouver un moyen de limiter l’influence des données erronées sur l’estimation. Dans le cas de la régularisation par le gradient, les valeurs faibles du gradient correspondent au bruit, qu’il convient de lisser, tandis que les grandes valeurs correspondent aux discontinuités ou contours présents dans les images, que l’on souhaite en fait préserver.

### 7.3 Théorie semi-quadratique

Intuitivement, on peut envisager deux façons de sortir du comportement quadratique. La première consiste à utiliser une fonctionnelle de la forme :

$$J(f) = \frac{1}{2} \sum_{i=1}^N \varphi(u_i). \quad (7.3.1)$$

Cette formulation est connue sous le nom de M-estimateur [Hub81] en estimation robuste ( $\varphi$  est souvent notée  $\rho$  dans ce cadre) ou de modèle à phi-fonction en régularisation. Toute la question réside alors dans le choix de la fonction  $\varphi$ . Prenons l’exemple de la quadratique tronquée  $\min(u^2, 1)$ . Pour une faible valeur de  $u_i$ , cette fonction garde le comportement quadratique habituel. Par contre, à partir de  $u_i = 1$ , la valeur de la fonction ne change plus et ce, quelle que soit la valeur de  $u_i$ . L’algorithme de minimisation n’a donc pas particulièrement intérêt à modifier cette valeur en agissant sur  $f$ . La forte déviation perd ainsi toute influence sur le résultat. L’inconvénient de cette fonction est qu’elle présente un point anguleux en 1, ce qui est une source potentielle d’instabilité. Une fonction continue serait préférable. Mais l’asymptote horizontale est-elle nécessaire ? Du point de vue de l’optimisation, une fonction convexe serait préférable. De fait, de nombreuses fonctions ont été proposées dans la littérature (voir un état

de l'art dans [Cha94] pour la régularisation). Quelques exemples sont donnés dans le tableau 7.1.

La seconde approche consiste à utiliser une variable auxiliaire pour “marquer” explicitement les données [GG84]. Considérant une variable binaire  $b_i \in \{0, 1\}$ , on peut définir le critère *augmenté* suivant :

$$J^*(f, b) = \frac{1}{2} \sum_{i=1}^N b_i (u_i)^2 + (1 - b_i). \quad (7.3.2)$$

La variable  $b$  est appelée “processus de ligne” en reconstruction régularisée. Le second terme de cette fonctionnelle est un coût associé à la détection d'une forte déviation permettant d'éviter le minimum dégénéré où tous les  $b_i$  seraient nuls. L'intérêt de cette formulation est que le critère est quadratique lorsqu'on fixe  $b$  (d'où l'appellation “semi-quadratique”) : c'est celui des moindres carrés pondérés.

Pdf de $u$	Fonction	Abrév.	$\varphi(u)$	$\varphi'(u)/2u$	(7.3.4-7.3.6)	Convexité
Gauss	Quadratique	Q	$u^2$	1	non	oui
Quasi-Laplace	Hyper-surfaces	HS	$2\sqrt{1+u^2} - 2$	$\frac{1}{\sqrt{1+u^2}}$	oui	oui
Cauchy (Lorentz)	Hebert & Leahy	HL	$\log(1+u^2)$	$\frac{1}{(1+u^2)}$	oui	non
-	Geman & McClure	GM	$\frac{u^2}{1+u^2}$	$\frac{1}{(1+u^2)^2}$	oui	non

TAB. 7.1 – Exemples de fonctions  $\varphi$  proposées dans la littérature et leurs dénominations les plus usuelles dans les approches déterministe et probabiliste. Leurs tracés sont donnés fig. 7.2.

Naturellement se pose la question de l'estimation des valeurs de  $b$ . En l'absence d'information supplémentaire, cela doit se faire à partir des données, donc de  $u$ . Considérons l'optimisation par rapport à  $b$  de  $J^*(u, b)$  [BZ87]. La fonction quadratique étant inférieure à 1 sur  $[0, 1]$ , le minimum est réalisé pour  $b_i = 1$  si  $u_i < 1$  et  $b_i = 0$  sinon. On remarque alors l'équivalence entre ce modèle et le précédent. De même que précédemment, un comportement moins “binaire” serait cependant souhaitable. Peut-on généraliser cette technique à une variable auxiliaire à valeurs réelles ?

La théorie semi-quadratique répond à ces questions en définissant une famille de fonctions  $\varphi$  qui assure le comportement souhaité. Elle montre l'équivalence entre ce modèle et un modèle où la prise en compte de la nature des données est explicite, grâce à une variable auxiliaire réelle. En cela, elle étend à la fois les travaux de Huber [Hub81] en estimation robuste et de Geman et al. [GR92, GY95] en régularisation.

### 7.3.1 Propriétés des fonctions $\varphi$

Les fonctions de potentiel  $\varphi$  proposées dans la littérature semblent d'allures très différentes (voir figure 7.2). Leur étude a été proposée dans le cas de la régularisation dans [Cha94] et dans [BFCAB95]. Elle se base sur un examen des équations normales liées à l'optimisation du critère. En estimation robuste, les fonctions considérées étant paires, on peut écrire le critère (7.3.1) pour  $u_i = |n_i|$ , où  $n = p - Rf$  est le résidu. Le minimum de  $J$  est atteint lorsque sa dérivée est nulle, soit (puisque  $\text{sign}(x) = \frac{x}{|x|}$ ) lorsque,  $\forall j = 1 \dots D$  :

$$\frac{1}{2} \sum_{i=1}^N \varphi'(|n_i|) \text{sign}(n_i) \frac{\partial n_i}{\partial f_j} = \frac{1}{2} \sum_{i=1}^N \underbrace{\frac{\varphi'(|n_i|)}{2|n_i|}}_{\text{Pondération.}} \times \underbrace{2n_i \frac{\partial n_i}{\partial f_j}}_{\text{Cas quadratique}} = 0. \quad (7.3.3)$$

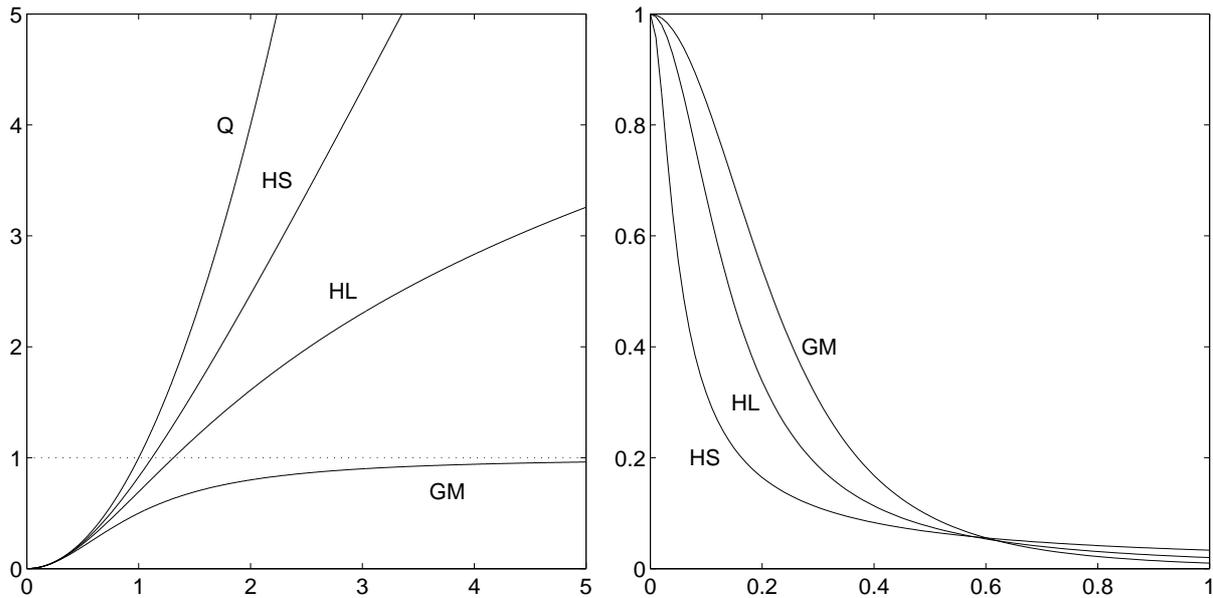


FIG. 7.2 – A gauche : exemples de fonctions  $\varphi(u)$  ; à droite : fonctions de pondération  $\frac{\varphi'(u)}{2u}$  (avec un coefficient d'échelle arbitraire sur  $u$  [Cha94]).

On voit que si  $\frac{\varphi'(|n_i|)}{2|n_i|}$  tend vers 0, la donnée correspondante n'entre pas en jeu dans la somme, ce qui est souhaitable pour les forts résidus. Par contre, cette pondération doit tendre vers 1 pour que la donnée soit totalement prise en compte, comme dans le cas des faibles résidus. En résumé, les conditions à vérifier pour obtenir un comportement différencié selon les valeurs de  $u$  sont les suivantes :

$$\lim_{u \rightarrow 0} \frac{\varphi'(u)}{2u} = M, \quad 0 < M < +\infty \quad (7.3.4)$$

$$\lim_{u \rightarrow \infty} \frac{\varphi'(u)}{2u} = 0 \quad (7.3.5)$$

$$\frac{\varphi'(u)}{2u} \quad \text{strictement décroissante.} \quad (7.3.6)$$

En général,  $M$  est choisi égal à 1. Ces conditions comparent les comportements de la fonction  $\varphi$  et de la quadratique, en termes de dérivées. Les deux fonctions évoluent de la même façon près de 0 (7.3.4), mais la fonction  $\varphi$  est sous-quadratique à l'infini (7.3.5). La condition (7.3.6) assure la cohérence du modèle. De fait, on constate sur les exemples de la figure 7.2 que les fonctions  $\varphi$  ont des fonctions de pondération au comportement similaire.

### 7.3.2 Extensions semi-quadratiques

On peut montrer que les fonctions réelles  $\varphi$  paires, croissantes, vérifiant les conditions (7.3.4-7.3.6) et quelques autres hypothèses "techniques", sont telles que :

$$\varphi(u) = \min_{b \in [0,1]} (bu^2 + \Psi(b)), \quad \forall u \quad \text{et} \quad b_{inf} = \frac{\varphi'(u)}{2u}, \quad (7.3.7)$$

où  $\Psi$  est une fonction strictement convexe définie à partir de  $\varphi$  et  $b$  est la variable auxiliaire. On parle d'extension multiplicative. Ce résultat généralise un théorème in-

roduit dans [GR92], initialement limité aux fonctions à asymptote horizontale. On a également, sous les mêmes conditions, une extension additive :

$$\varphi(u) = \min_{b \in \mathbb{R}^+} ((b - u)^2 + \xi(b)), \forall u \quad \text{et} \quad b_{inf} = \left(1 - \frac{\varphi'(u)}{2u}\right) u, \quad (7.3.8)$$

où  $\xi$  est une fonction strictement convexe définie à partir de  $\varphi$ . Ce résultat est basé sur les travaux de [GY95].

### 7.3.3 Algorithmique

En reportant les équations (7.3.7) et (7.3.8) dans (7.3.1), on obtient respectivement les critères augmentés suivants :

$$J^*(f, b) = \frac{1}{2} \sum_{i=1}^N b_i (u_i)^2 + \Psi(b_i) = \frac{1}{2} \|u\|_B^2 + \frac{1}{2} \sum_{i=1}^N \Psi(b_i), \quad (7.3.9)$$

avec  $B = \text{diag}(b_i)$  et :

$$J^\#(f, b) = \frac{1}{2} \sum_{i=1}^N (b_i - u_i)^2 + \xi(b_i) = \frac{1}{2} \|u - b\|^2 + \frac{1}{2} \sum_{i=1}^N \xi(b_i). \quad (7.3.10)$$

Ces critères augmentés d'une variable sont, en fait, plus simples à minimiser que le critère original. En effet, la minimisation directe de  $J$  conduit à des équations non linéaires. Par contre, à  $b$  fixée, les critères augmentés sont quadratiques et conduisent donc à la résolution d'équations normales linéaires, de la forme  $Af = c$ . Notons que dans cette minimisation, les fonctions  $\Psi$  et  $\xi$  n'interviennent pas (il n'existe d'ailleurs pas d'expression analytique simple de  $\xi$ ). De plus, lorsque la variable  $f$  est fixée le critère augmenté est convexe en  $b$  et la valeur du minimum est donnée dans (7.3.7) et (7.3.8). Cette valeur dépend de la fonction de pondération et d'après (7.3.4) et (7.3.5), il est clair que la variable auxiliaire  $b$  joue le rôle de "marqueur" de données ("inlier/outlier", "bruit/discontinuité" dans nos applications).

Bien que des algorithmes stochastiques existent [GR92, Jal01], les propriétés semi-quadratiques sont, la plupart du temps, exploitées sous forme d'algorithmes déterministes opérant par minimisations alternées par rapport à chacune des variables. En général, on positionne initialement tous les  $b_i$  à 1, ce qui revient à commencer par un algorithme des moindres carrés classique. Dans le cadre de l'estimation robuste, on retrouve pour la forme multiplicative l'algorithme des moindres carrés (re-)pondérés itératifs (Iterative Reweighted Least Squares ou IRLS) et pour la forme additive l'algorithme des moindres carrés avec remise à jour des résidus, proposé dans [Hub81]. Dans le cas de la régularisation, les algorithmes sont nommés ARTUR [CBFAB97] pour la forme multiplicative et LEGEND pour la forme additive [Cha94]. Ces algorithmes convergent dans le cas des fonctions  $\varphi$  convexes. Dans le cas non-convexe, ils convergent vers un minimum local à condition que les points critiques de la fonctionnelle soient isolés. Notons que la convergence est généralement plus lente dans le cas de la seconde forme semi-quadratique, mais que celle-ci permet d'obtenir des algorithmes moins calculatoires lorsque  $f$  est de grande dimension (seul le second membre des équations normales est modifié à chaque étape). Le compromis peut donc être intéressant dans le cas d'algorithmes de reconstruction d'image, par exemple.

Pour une synthèse et des résultats récents concernant la vitesse de convergence des algorithmes semi-quadratique, le lecteur pourra se référer à [All02].

Enfin, pour améliorer la convergence des algorithmes, on peut adopter une stratégie de continuation, dans la philosophie du GNC (Graduated Non Convexity) de [BZ87]. Cela consiste à utiliser d'abord une fonction convexe (ex.  $\varphi_{HS}$ ), puis à repartir du résultat obtenu en employant une fonction non convexe "peu sévère" (ex.  $\varphi_{HL}$ ), etc.

### 7.3.4 Approche Lagrangienne

Il y a plusieurs façon de présenter la théorie semi-quadratique et les algorithmes associés, chacune apportant un éclairage particulier.

Nous en avons récemment [TIC02, TIC07b] proposé une reformulation dont l'objectif est de montrer plus directement le lien avec la théorie de l'optimisation. En effet, en introduisant la fonction  $\phi(u_i^2)$  en place de  $\varphi(u_i)$  dans (7.3.1), nous montrons qu'il est possible de reformuler le problème sous la forme d'une optimisation sous contrainte en introduisant la variable auxiliaire  $w_i = u_i^2$ . Les pondérations  $b_i = \phi'(u_i^2)$  s'interprètent alors comme les valeurs des multiplicateurs de Lagrange associés aux contraintes. Même si cela n'est pas requis pour la dérivation des algorithmes, on peut également interpréter (7.3.9) comme une forme duale. De nouvelles conditions suffisantes sont alors obtenues sur  $\phi$  pour pouvoir poser et résoudre cette reformulation :

- **H0** :  $\phi$  est définie sur  $\mathbb{R}^+$  ainsi que ses dérivées première et seconde,
- **H1** :  $\phi$  est croissante,  $\phi'(t) > 0$ ,
- **H2** :  $\phi$  est concave,  $\phi''(t) < 0$ .

La reformulation lagrangienne est également possible pour la version additive, mais en ajoutant une hypothèse sur  $\phi : \phi'(t) \leq 1$ .

L'étude de ces conditions met en lumière le rôle central que jouent les propriétés de convexité et de monotonie dans la théorie semi-quadratique : la forme multiplicative repose sur la convexité et la décroissance de la fonction  $-\varphi(\sqrt{u}) = -\phi(u)$  ; la forme additive repose sur la convexité et la croissance de la fonction  $u^2 - \varphi(u) = u^2 - \phi(u^2)$ . Il est alors intéressant de noter que les conditions (7.3.4-7.3.6) sont suffisantes, aussi bien pour la forme multiplicative que pour la forme additive.

La reformulation lagrangienne aboutit aux mêmes algorithmes que le développement semi-quadratique décrit dans les section précédentes. Une telle méthodologie a cependant l'avantage de faciliter la preuve de convergence, en la replaçant dans le cadre classique de la théorie de l'optimisation autour du théorème de Khun et Tucker. Elle permet assez naturellement des extensions à de nouveaux types de problèmes, comme : l'estimation simultanée des paramètres de plusieurs instances d'un modèle linéaire [TCI07, TIC07b] qui démontre des liens intéressant avec l'approche Estimation Maximisation (EM) [DLR77], le recalage entre deux jeux de données lorsque la mise en correspondance est inconnue et la transformation à estimer est affine [TIC07a], la régression d'un modèle de région paramétré sur des données images [BT07].

### 7.3.5 Hyper-paramètres

Dans ce qui précède nous avons, par souci de simplicité, éludé le paramètre d'échelle. Dans (7.3.1), la variable  $u_i$  doit, en réalité, être divisée par un facteur  $\delta$  qui ajuste le comportement de la fonction à la dynamique des données. D'autre part, lorsque le critère est régularisé, un coefficient  $\lambda^2$  permet de pondérer les termes de

vraisemblance et d'*a priori*. Ces *hyper*-paramètres sont en général laissés au choix de l'utilisateur. Il est cependant possible de les fixer de manière plus automatique (voir [Ien04] pour l'estimation robuste et [Jal01] dans le cas de la reconstruction régularisée).

## 7.4 Applications

Nous illustrons maintenant quelques applications de la théorie semi-quadratique en analyse d'images, puis en reconstruction régularisée.

### 7.4.1 Régression robuste pour la détection et le suivi des marquages routiers

Nous nous intéressons ici à la détection et au suivi des marquages routiers à partir d'images numériques. Dans des applications telles que l'aide à la conduite, la robustesse des traitements aux perturbations est d'une importance cruciale.

Le principe du système développé [Ien04] (cf. fig. 7.3) consiste à effectuer dans un premier temps une détection des centres supposés des marquages, puis à les approcher par ajustement de courbes. Le fonctionnement de quelques détecteurs est détaillé dans [TIC07b]. Ils se basent sur des techniques simples de seuillage et exploitent des informations géométriques a priori sur la largeur des marquages, ce qui rend les algorithmes relativement moins sensible à certaines variations d'intensité que les détecteurs de contours classiques. Cependant, les cartes de détection demeurent relativement bruitées et la régression semi-quadratique trouve ici un champ d'application naturel.

Dans cette application, la relation sous-jacente entre les coordonnées images  $(x, y)$  des points de marquage est choisie de type polynomial :

$$y = \sum_{j=0}^D X_j(x) f_j, \quad (7.4.1)$$

où  $X_j(x)$  représente la valeur en  $x$  d'une fonction de base. Selon l'angle de prise de vues, on peut utiliser différentes formes de polynômes (cf. tableau 7.2). En collectant les équations correspondant aux  $N$  points de mesure, on obtient un système de la forme (7.2.2). Un modèle original de bruit, appelé famille exponentielle lissée, a été proposé dans [TIC02] et se base sur la fonction de potentiel suivante :

$$\varphi_\alpha(u) = \frac{1}{\alpha} \left( (1 + u^2)^\alpha - 1 \right). \quad (7.4.2)$$

Cette famille paramétrée par  $\alpha$  s'avère bien adaptée à la modélisation des résidus dans notre application. D'autre part, en faisant varier  $\alpha$ , elle permet une transition continue entre les différents modèles de bruit classiques (de Gauss pour  $\alpha = 1$  à Geman & McClure pour  $\alpha = -1$ , en passant par  $\varphi_{HS}$  pour  $\alpha = 0.5$ ). La robustesse de certains de ces estimateurs a été établie mathématiquement et peut atteindre le maximum possible de 50% de données aberrantes [ITC07].

Un algorithme semi-quadratique est obtenu à partir de l'approche lagrangienne décrite au paragraphe 7.3.4. La figure 7.3 montre un exemple d'application de cet algorithme pour la détection d'une ligne de marquage dans des conditions d'illumination

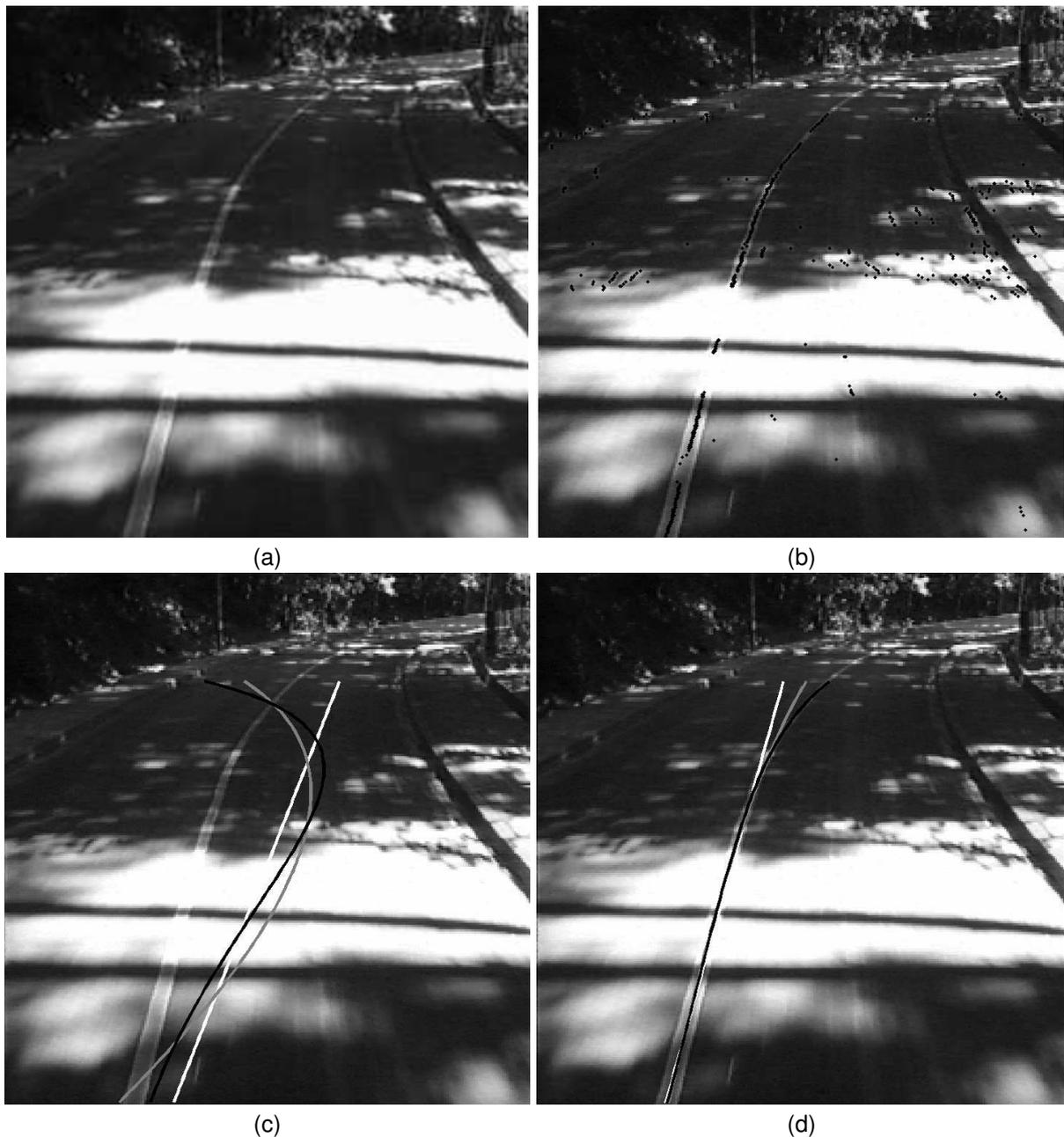


FIG. 7.3 – (a) Image de route dans des conditions d'éclairage difficiles. (b) Centres de marquages extraits ( $\cdot$ ) : on note la présence de nombreuses fausses alarmes. Ajustement (c) gaussien, (d) robuste : en blanc, degré 1 ; en gris, degré 2 ; en noir, degré 3.

difficiles. Notons que cet algorithme est utilisé en routine dans le logiciel IREVE pour le calibrage des caméras des MLPC IRCAN (Imagerie Routière par CAMéra Numérique).

Il est possible d'intégrer à la fonctionnelle un terme de régularisation quadratique permettant de contrôler le degré de la courbe obtenue, voire de prendre en compte les corrélations existant entre les fonctions de base. Il est également possible de tenir compte dans l'*a priori* d'un résultat d'ajustement précédent. L'estimateur s'intègre donc naturellement dans un algorithme de suivi par filtrage de Kalman [TIC02]. Cela pose le problème de la définition de la matrice de covariance de l'estimée, laquelle ne peut être

qu'approchée dans le cas de l'estimation robuste. Plusieurs formes de matrice ont été comparées [TIC07b]. L'algorithme de suivi a été testé avec succès à bord de véhicules équipés pour l'aide à la conduite au LIVIC [Ien04].

Une extension à l'ajustement simultané de plusieurs courbes a été proposée récemment [TCI07] à partir du cadre lagrangien. La figure 7.4 en donne un exemple d'application. Il s'agit de détecter simultanément toutes les lignes longitudinales d'une grille de marquage utilisée pour l'étalonnage de caméras. On peut constater qu'en utilisant un modèle gaussien, l'algorithme produit un résultat erroné malgré une initialisation manuelle très proche de la solution. Au contraire, l'algorithme produit de très bon résultats, à partir d'une initialisation pourtant plus éloignée, lorsqu'on utilise un modèle robuste de vraisemblance.

## 7.4.2 Modèles d'apparence pour la détection et la reconnaissance de la signalisation verticale

La détection d'objets dans des images est un sujet difficile qui a fait l'objet de nombreux travaux par le passé. Les approches par *modèles d'apparence*, notamment, rencontrent un succès certain depuis le début des années 90 [TP91]. Elle permettent la représentation de classes d'objets et l'apprentissage de variations de forme, d'orientation ou d'illumination. Nous en présentons ici une application à la détection et à la reconnaissance de la signalisation verticale dans les scènes routières [Dah01].

Dans cette approche, toute image est représentée par un vecteur dont chaque composante est la valeur de niveau de gris d'un pixel. Une telle représentation, globale, de l'apparence est évidemment peu parcimonieuse : c'est pourquoi on a en général recours à des techniques de réduction de dimension. Nous utilisons pour notre part l'analyse en composantes principales (ACP). Elle permet de modéliser une base  $\mathcal{B}$  d'images d'apprentissage à l'aide d'une image moyenne  $\mu$  et d'un nombre réduit de vecteurs propres  $u_j$ , représentant les axes principaux de variation du nuage de points. Dans notre modèle, toute image peut se décomposer comme une combinaison linéaire de ces vecteurs de base, à une erreur  $w$  près :

$$p = \mu + \sum_{j=1}^J f_j u_j + w. \quad (7.4.3)$$

Nous retrouvons un modèle génératif linéaire similaire à (7.2.2). L'intérêt de ce type de modèle est que  $J$  est relativement faible (typiquement, quelques dizaines de coefficients, à comparer aux centaines de milliers de pixels des images traitées). Les comparaisons nécessaires à la détection et à la reconnaissance s'en trouvent donc fortement simplifiées.

Polynôme	Prise de vue	Expression
Usuel	Verticale	$y = \sum_{j=0}^D f_j x^j$
Hyperbolique	Perspective	$y = f_0 x + f_1 + \sum_{j=2}^D \frac{f_j}{(x-x_h)^j}$ ( $x_h$ =ligne d'horizon)
Fractionnaire	Perspective	$y = \sum_{j=0}^D f_j x^{(\frac{j}{D})}$

TAB. 7.2 – Types de polynôme en fonction de l'angle de prise de vue. Les polynômes fractionnaires ne nécessitent pas la connaissance de la position de la ligne d'horizon.



FIG. 7.4 – (a) Image originale de la grille d'étalonnage. (b) Centres de marquages extraits ( $\cdot$ ). (c) Initialisation de 10 lignes pour l'ajustement des marquages verticaux. (d) Lignes ajustées dans l'hypothèse de distribution gaussienne pour le bruit. (e) Initialisation de 12 lignes. (f) ajustement robuste simultané des marquages, potentiel  $\varphi_\alpha$ ,  $\alpha = 0.1$ .

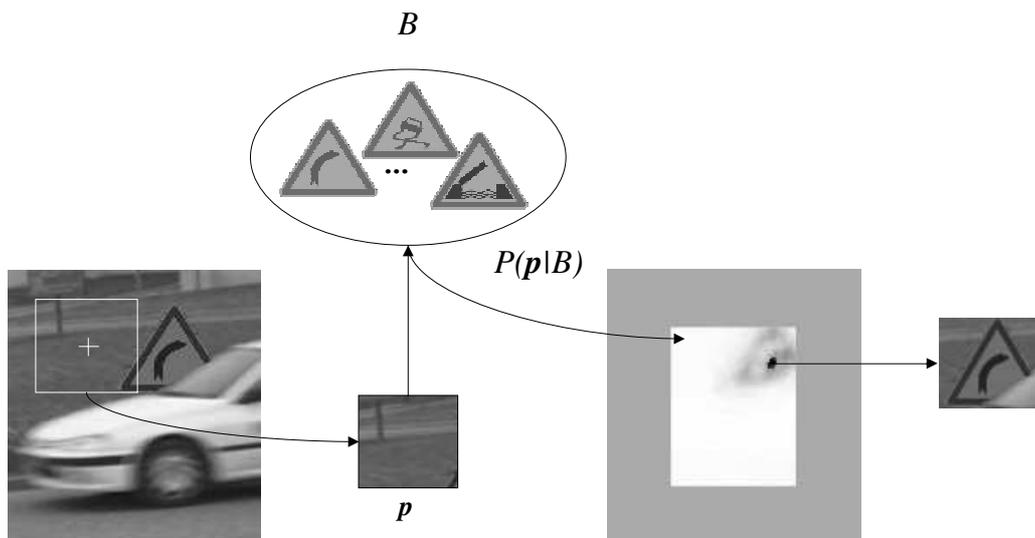


FIG. 7.5 – Principe de l'algorithme de détection.

Le principe de l'algorithme de détection d'objets d'intérêt dans une scène est illustré figure 7.5. En chaque position de l'image, on extrait un vecteur d'observation  $p$  dont on évalue la vraisemblance  $\mathcal{P}(p|\mathcal{B})$ , par rapport au modèle appris sur la base  $\mathcal{B}$ . Cette valeur est affectée à la position de l'extrait. Lorsque toute la scène a été parcourue, on dispose d'une carte de vraisemblance. La position des objets d'intérêt éventuellement présents dans la scène est alors obtenue par seuillage. La vraisemblance de l'observation,  $\mathcal{P}(p|\mathcal{B})$ , est donnée par :

$$\mathcal{P}(p|\mathcal{B}) \propto \mathcal{P}(p|\hat{f}, \mathcal{B}) \mathcal{P}(\hat{f}|\mathcal{B}), \quad (7.4.4)$$

où  $\hat{f} = \arg \max_f \mathcal{P}(p, f|\mathcal{B})$  est l'estimée au sens du MAP des paramètres du modèle. Nous avons montré [DCH04] que l'association d'un modèle robuste de distribution du bruit,  $\mathcal{P}(p|f, \mathcal{B})$ , et d'un modèle *a priori*,  $\mathcal{P}(f|\mathcal{B})$ , bien adapté permet d'améliorer de manière significative les performances des systèmes de détection existants. La figure 7.6 illustre cela sur un exemple synthétique. La base d'images d'apprentissage contient 43 signaux de danger, sous 36 angles de rotation chacun. On ne retient que 30 vecteurs propres. La méthode faisant référence jusqu'à présent, utilisant des hypothèses gaussiennes [MP97], est mise en défaut sur cette image. Une meilleure discrimination est obtenue en associant l'estimation robuste à une modélisation *a priori* adaptée aux données. Dans cet exemple, on connaît la forme analytique de cette dernière distribution, mais nous avons également proposé une méthode, basée sur l'algorithme *mean shift* [Che95], permettant de gérer des formes arbitraires de distributions [VHC03].

La reconnaissance, quant à elle, est effectuée en comparant les coordonnées dans l'espace propre estimées,  $\hat{f}$ , à celles des panneaux d'apprentissage. Ici encore, l'utilisation d'un estimateur robuste améliore les performances du système [Dah01], comme le montre la figure 7.6. Dans cette expérience, le panneau est en partie caché par une inscription jaune. Le panneau reconnu à partir de l'estimée au sens du MV gaussien est le panneau temporaire à fond jaune, alors que le panneau reconnu dans le cas robuste est correct. On note que l'estimation fournit, de plus, une carte des données erronées.

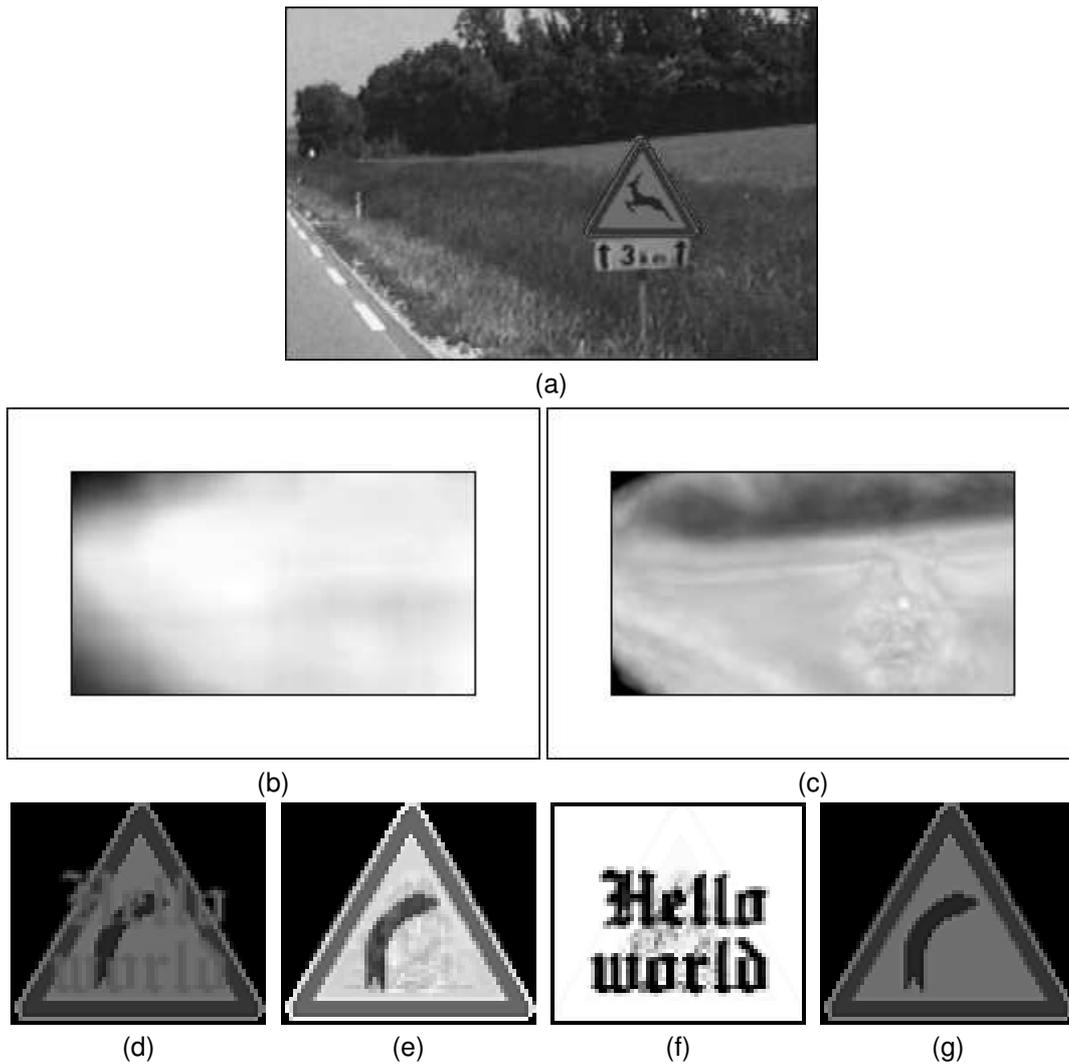


FIG. 7.6 – (a) Image analysée. Ligne centrale : expérience de détection. Cartes de vraisemblance pour des hypothèses de bruit et *d'a priori* (respectivement) : (b) gaussien-gaussien et (c) robuste-non gaussien. Les plus fortes vraisemblances apparaissent en clair. Ligne du bas : expérience de reconnaissance. (d) image analysée. (e) images reconstruite. (f) carte des données erronées, b. (g) panneau reconnu.

### 7.4.3 Reconstruction régularisée

Nous nous intéressons ici aux problèmes de reconstruction sous le modèle génératif linéaire. Il s'agit d'un problème inverse mal posé. Existence et unicité de la solution peuvent être obtenues respectivement en calculant la solution au sens des moindres carrés et en prenant la solution de norme minimale : on parle alors d'inversion généralisée. En décomposant  $R$  en valeurs singulières,  $R = UDV^T$ , on obtient :

$$\hat{f}^+ = R^+ p \quad \text{où} \quad R^+ = VD^+U^T, \quad (7.4.5)$$

$D^+$  étant la matrice diagonale contenant l'inverse des valeurs singulières non nulles de  $R$ , et des 0 aux positions des valeurs singulières nulles. On résout ainsi le problème de l'inversion de (7.2.1) mais pas celui de la stabilité en présence de bruit (7.2.2). En effet la matrice  $R$  est toujours mal conditionnée et l'inversion de ses valeurs singulières



FIG. 7.7 – (a) Image originale “cameraman”. (b) Image floue (défocalisation de rayon 2 pixels) et bruitée. (c) Inverse généralisée (7.4.5). (d) Algorithme de Landweber, 100 itérations (7.4.6),  $SNR_{var} = 15,7dB$ . (e) Régularisation de Tikhonov (7.4.7),  $\lambda = 2$ ,  $SNR_{var} = 14,9dB$ . (f) Régularisation semi-quadratique ICM-DCT,  $\lambda = 2$ ,  $\delta = 15$ , 100 itérations,  $SNR_{var} = 18,4dB$ .

faibles conduit à une amplification des hautes fréquences du bruit, qui peuvent même dominer la solution (cf. figure 7.7). Il est possible d'y remédier en éliminant les valeurs singulières faibles ou bien en calculant l'inverse généralisée par une descente de gradient sur le critère des moindres carrés (algorithme de Landweber) :

$$\hat{f}_{LW}^{(k+1)} = \hat{f}_{LW}^{(k)} + \gamma R^T (p - R \hat{f}_{LW}^{(k)}), \quad (7.4.6)$$

et en arrêtant la solution après un nombre fini d'itérations. La régularisation de Tikhonov, évoquée au paragraphe 7.2.2, consiste à ajouter une contrainte quadratique sur la solution ou ses dérivées. Dans le cas d'une régularisation sur le gradient, la solution vérifie les équations normales :

$$(R^T R - \lambda^2 \Delta) \hat{f}_T = R^T p, \quad (7.4.7)$$

où  $\Delta$  est l'opérateur Laplacien. On peut montrer que ces 3 solutions correspondent chacune à une façon de filtrer les valeurs singulières faibles. On parle de régularisation linéaire. Son défaut est d'agir globalement sur les images et de ne pas respecter les fortes variations locales, correspondant aux contours (cf. figure 7.7).

Afin de préserver les discontinuités, on remplace le terme quadratique de régularisation par un terme de la forme (7.3.1). En variables continues, le critère régularisé s'écrit alors :

$$J(f) = \|p - Rf\|^2 + \lambda^2 \int_{\Omega} \varphi(|\nabla f|) \quad (7.4.8)$$

où  $\Omega$  représente le domaine image. Le lecteur pourra se référer à [BF00] pour une étude en variables continues en lien avec la théorie des équations aux dérivées partielles (EDP). Nous utiliserons ici, comme dans [Cha94], une formulation discrète séparable :

$$J(f) = \|p - Rf\|^2 + \lambda^2 \sum_j \varphi((D_x f)_j) + \lambda^2 \sum_j \varphi((D_y f)_j). \quad (7.4.9)$$

Par rapport à la formulation continue, ce critère a le défaut de ne pas être isotrope. Par contre, l'aspect séparable peut être intéressant pour l'implantation. Le critère  $J$  est étendu à deux variables selon l'un ou l'autre des développements semi-quadratiques et une stratégie déterministe de minimisations alternées par rapport à chaque variable est mise en œuvre. Chaque étape de l'algorithme comprend donc le calcul des variables auxiliaires  $b_x^{(k+1)}$  et  $b_y^{(k+1)}$  selon les expressions (7.3.7) et (7.3.8) et la résolution des équations normales. Dans le cas multiplicatif (algorithme ARTUR), elles s'écrivent :

$$(R^T R - \lambda^2 \Delta_A^{(k+1)}) \hat{f}_A^{(k+1)} = R^T p, \quad (7.4.10)$$

où :

$$\Delta_A^{(k+1)} = -D_x^T B_x^{(k+1)} D_x - D_y^T B_y^{(k+1)} D_y,$$

avec  $B_x^{(k+1)} = \text{diag}(b_x^{(k+1)})$  et  $B_y^{(k+1)} = \text{diag}(b_y^{(k+1)})$ . Dans le cas additif (algorithme LEGEND), on résout :

$$(R^T R - \lambda^2 \Delta) \hat{f}_L^{(k+1)} = R^T p + \lambda^2 D_x^T b_x^{(k+1)} + \lambda^2 D_y^T b_y^{(k+1)} \quad (7.4.11)$$

La première forme converge en général plus rapidement que la seconde, en nombre d'itérations. Par contre, cette dernière est intéressante car le premier terme ne varie pas au cours des itérations  $k$ . On peut, par exemple, le factoriser par l'algorithme de Cholesky pour faciliter ensuite les calculs [Cha94]. Dans le cas de la déconvolution,

avec conditions aux bords circulaires, la transformée de Fourier permet de diagonaliser ce terme. On peut cependant lui préférer la transformée en cosinus discret lorsque la PSF est symétrique. Cela sous-entend des conditions aux bords symétriques, correspondant aux conditions de Neumann associées aux équations normales. On obtient ainsi l'algorithme ICM-DCT [Jal01]. Les figures 7.7 et 7.8 présentent les résultats obtenus en déconvolution à l'aide de l'algorithme ICM-DCT, en comparaison d'estimateurs classiques. Le gain en qualité, évident visuellement, est confirmé en termes de rapport signal-sur-bruit (SNR). Celui-ci est mesuré à partir des variances de l'image de référence,  $f$  et de l'image reconstruite  $\hat{f}$  :  $SNR_{VAR} = 10 \log_{10}(\sigma_f^2 / \sigma_{f-\hat{f}}^2)$ . Les filtres utilisés fig. 7.7 et 7.8 modélisent respectivement une défocalisation et un bougé supposés connus.

La régularisation semi-quadratique a été appliquée avec succès à de nombreux problèmes de reconstruction comme, entre autres, la stéréovision, l'estimation de mouvement, la tomographie ou encore l'imagerie micro-ondes (diffraction inverse) [BF00].

## 7.5 Conclusion

L'hypothèse classique de distribution gaussienne (correspondant à une énergie quadratique) s'avère inadaptée dans de nombreux problèmes réels où les données se composent en réalité de deux catégories ("inlier/outlier"), nécessitant chacune un traitement adapté. On fait alors appel à l'estimation robuste combinée à la régularisation lorsque le problème est mal posé. Nous avons présenté un cadre général, la théorie semi-quadratique, permettant de sortir du comportement purement quadratique, soit par l'introduction d'une fonction bien choisie qui atténue l'influence des fortes déviations, soit par l'introduction directe de variables auxiliaires et d'un terme de pénalité énergétique associé. La théorie semi-quadratique permet d'établir le lien entre ces deux approches et précise les conditions de choix possibles, en termes simples. De plus, ceci conduit le plus souvent à des algorithmes rapides et faciles à implanter. Cette théorie se situe à la confluence des statistiques robustes [Hub81], des champs markoviens [GG84], des approches variationnelles [AK02], d'approches statistiques telles que l'EM [CI04] et de l'analyse convexe [LHY00]. C'est ce qui en fait le succès, même si elle n'est pas toujours identifiée en tant que telle. Elle permet en effet de poser et de résoudre de multiples problèmes pratiques, comme nous l'avons illustré par des exemples : la détection et le suivi des marquages routiers, la reconnaissance des panneaux de signalisation, et l'amélioration de la qualité d'image.

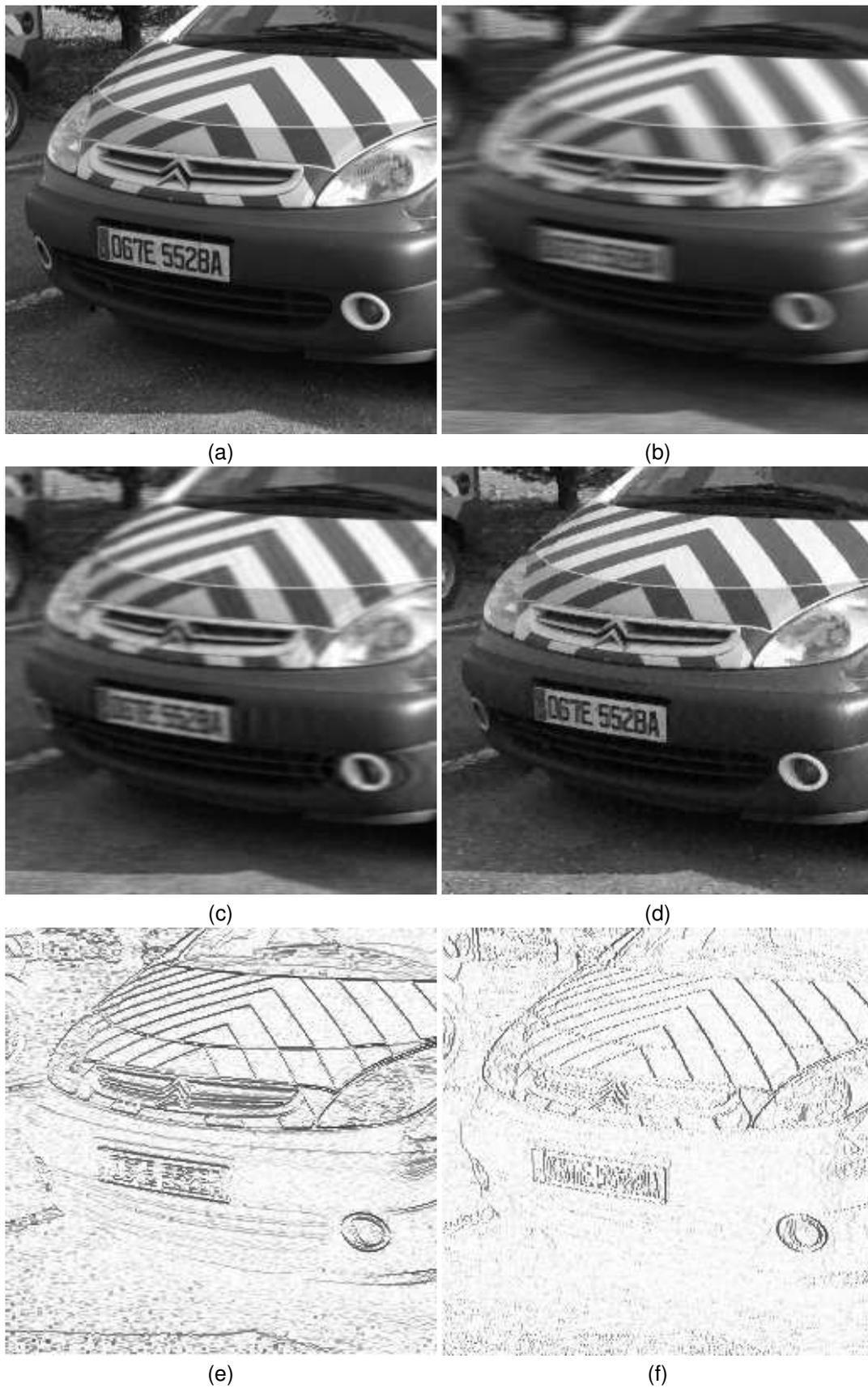


FIG. 7.8 – (a) Image originale “ecodyn”. (b) Image floue (bougé horizontal de 12 pixels) et bruitée. (c) Régularisation de Tikhonov,  $\lambda = 1$ ,  $SNR_{var} = 14, 5dB$ . (d) Reconstruction régularisée semi-quadratique par ICM-DCT,  $\lambda = 1, \delta = 7$ , 100 itérations,  $SNR_{var} = 18, 1dB$ , où la plaque minéralogique est de nouveau lisible. (e-f) Discontinuités entre lignes,  $b_x$  et entre colonnes  $b_y$  (visualisation en log).



# Bibliographie

- [AK02] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing : Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002.
- [All02] M. Allain. *Approche pénalisée en tomographie hélicoïdale en vue de l'application à la conception d'une prothèse personnalisée du genou*. PhD thesis, Université Paris Sud - Paris XI, France, 2002.
- [BF00] L. Blanc-Féraud. *Sur quelques Problèmes Inverses en Traitement d'Image*. Habilitation à diriger des recherches, Université de Nice-Sophia Antipolis, France, July 2000.
- [BFCAB95] L. Blanc-Féraud, P. Charbonnier, G. Aubert, and M. Barlaud. Nonlinear image processing : Modeling and fast algorithm for regularization with edge detection. In *Proc. IEEE International Conference on Image Processing (ICIP)*, volume I, pages 474–477, Washington D.C., USA, September 1995.
- [BT07] E. Bigorgne and J.-P. Tarel. Backward segmentation and region fitting for geometrical visibility range estimation. In *Proceedings of Asian Conference on Computer Vision (ACCV'07)*, volume II, pages 817–826, Tokyo, Japan, 2007. <http://perso.lcpc.fr/tarel.jean-philippe/accv07.html>.
- [BZ87] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, MA, 1987.
- [CBFAB97] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6(2) :298–311, February 1997. Prix IEEE 1998 du meilleur article de jeune auteur, section “Image and Multidimensional Signal Processing”.
- [Cha94] P. Charbonnier. *Reconstruction d'image : régularisation avec prise en compte des discontinuités*. PhD thesis, Université de Nice-Sophia Antipolis, France, 1994.
- [Che95] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(8) :790–799, August 1995.
- [CI04] F. Champagant and J. Idier. A connection between half-quadratic criteria and em algorithms. *IEEE Signal Processing Letters*, 11(9) :709–712, 2004.
- [Dah01] R. Dahyot. *Analyse d'images séquentielles de scènes routières par modèles d'apparence pour la gestion du réseau routier*. PhD thesis, Université Paris VI, France, 2001.
- [DCH04] R. Dahyot, P. Charbonnier, and F. Heitz. Robust Bayesian detection using appearance-based models. *Pattern Analysis and Applications*, 7 :317–332, 2004. On-line, <http://www.springerlink.com>.

- [DLR77] A. Dempster, N. Laird, and D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1) :1–38, 1977.
- [GG84] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6) :721–741, November 1984.
- [GR92] S. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(3) :367–383, March 1992.
- [GY95] D. Geman and C. Yang. Nonlinear image recovery with half-quadratic regularization and FFT's. *IEEE Transactions on Image Processing*, 4(7) :932–946, July 1995.
- [Hub81] P.J. Huber. *Robust statistics*. John Wiley and Sons, New York, 1981.
- [Ien04] S. S. Ieng. *Méthodes robustes pour la détection et le suivi des marquages*. PhD thesis, Université Paris VI, France, 2004.
- [ITC07] S.-S. Ieng, J.-P. Tarel, and P. Charbonnier. Modeling non-Gaussian noise for robust image analysis. In *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP'07)*, pages 175–182, Barcelona, Spain, 2007. <http://perso.lcpc.fr/tarel.jean-philippe/publis/visapp07a.html>.
- [Jal01] A. Jalobeanu. *Modèles, estimation bayésienne et algorithmes pour la déconvolution d'images aériennes et satellitaires*. PhD thesis, Université de Nice-Sophia Antipolis, France, 2001.
- [Kay93] S. M. Kay. *Fundamentals of Statistical signal processing - Estimation Theory*. Prentice Hall International Editions, 1993.
- [LHY00] K. Lange, D. R. Hunter, and I. Yang. Optimization transfer using surrogate objective functions. *Journal of Computational and Graphical Statistics*, 9(1) :1–20, March 2000.
- [MP97] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(7) :696–710, July 1997.
- [TCI07] J.-P. Tarel, P. Charbonnier, and S.-S. Ieng. Simultaneous robust fitting of multiple curves. In *Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP'07)*, pages 175–182, Barcelona, Spain, 2007. <http://perso.lcpc.fr/tarel.jean-philippe/publis/visapp07b.html>.
- [TIC02] J.P. Tarel, S.S. Ieng, and P. Charbonnier. Using robust estimation algorithms for tracking explicit curves. In Springer, editor, *6th European Conference on Computer Vision (ECCV), Lecture Notes in Computer Science*, volume 2350, pages 492–407, Copenhagen, Denmark, May 2002.
- [TIC07a] J.-P. Tarel, S.-S. Ieng, and P. Charbonnier. Accurate and robust image alignment for road profil reconstruction. In *Proceedings of IEEE International Conference on Image Processing (ICIP'07)*, volume V, pages 365–368, San Antonio, Texas, USA, 2007. <http://perso.lcpc.fr/tarel.jean-philippe/icip07.html>.

- [TIC07b] J.-P. Tarel, S.-S. Ieng, and P. Charbonnier. *Robust Lane Marking Detection by the Half Quadratic Approach*. Collections Etudes et Recherches des Laboratoires des Ponts et Chaussées, CR 49, LCPC, Novembre 2007. <http://perso.lcpc.fr/tarel.jean-philippe/rr07.html>.
- [TP91] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1) :71–86, 1991.
- [VHC03] T. Vik, F. Heitz, and P. Charbonnier. Mean shift-based Bayesian image reconstruction into visual subspace. In *Proceedings of the 2003 IEEE International Conference on Image Processing (ICIP 2003)*, Barcelona, Spain, September 2003.



## Chapitre 8

# Des méthodes de Monte Carlo par chaînes de Markov pour l'identification des systèmes élastiques hétérogènes

*Maurizio Brocato*

### 8.1 Introduction

Les méthodes de Monte Carlo sont souvent utilisées pour la solution de problèmes inverses, en particulier lorsqu'on a recours à des approches bayésiennes. Ils se basent sur l'idée d'échantillonner aléatoirement les entrées du système (qui sont les inconnues du problème inverse) en vérifiant à posteriori chaque échantillon sur la base d'informations statistiques disponibles sur les sorties du système.

Dans ce chapitre on présente des approches de ce type, dans le cas particulier des systèmes élastiques hétérogènes qui sont caractérisés par une hétérogénéité de type inclusions dans une matrice. On suppose connaître le comportement élastique des ces deux phases, et on souhaite en déterminer la distribution spatiale. A cette fin on se donne des informations sur la déformée du système sous chargement connu.

Ce cas d'étude est intéressant pour les applications et se prête à généralisation pour traiter des problèmes plus complexes. Les applications vont de l'analyse d'image quantitative par l'élasticité à l'étude des matériaux élastiques-endommageables. La généralisation à des systèmes constitués de plus de deux phases peut être obtenue en considérant ceux-ci par des niveau d'approximations successives.

Le traitement quantitatif d'images par analyse élastique (« *quantitative elasticity imaging* ») est un domaine de recherche en expansion. L'identification par des mesures de déformation de parties d'un corps ayant un comportement élastique différent de celui de la matrice dans laquelle elles sont incluses s'avère utile et prometteuse pour les applications biomédicales ([BB02]). En fait, les formations tumorales dans les tissus biologiques, par exemple, se présentent comme des inclusions, avec des parois relativement marqués, ayant un comportement élastique différent de celui du tissu de base. Si le comportement de l'inclusion et du tissu sont connus à priori, le nombre d'inclusions et la forme et la position de chaque inclusion reste à déterminer.

Une deuxième famille des problèmes qui offrent la possibilité d'application immédiate de la méthodologie qu'on présente dans ce chapitre est celle des matériaux élastiques endommageables. C'est un domaine d'intérêt pour l'ingénierie, par exemple, lorsqu'on souhaite déterminer l'état actuel d'une structure. Dans ce contexte l'approche se trouve en compétition avec d'autres, plus répandus, s'appuyant sur l'analyse de données dynamiques ou acoustiques, mais il peut s'avérer utile lorsque le recours à des telles méthodes est peu praticable. Par exemple dans le cas des enceintes de confinement des centrales nucléaires, la détermination de l'état actuel du béton dans la structure ne peut pas s'effectuer – pour des raisons de sécurité – par excitation dynamique de la structure ; une approche statique comme celle présentée ci dessous peut alors devenir intéressante.

Comme pour l'exemple précédent, les matériaux élastiques endommageables, et en particulier le béton lorsque ce modèle s'applique, présentent souvent, suite à un chargement ayant induit de l'endommagement, des zones de raideur réduite (endommagées) séparées du reste du matériau sain par des interfaces relativement nettes. Ceci est dû à la localisation de l'endommagement et c'est un phénomène souvent observé, peut être à l'exception de cas particuliers. Certes, ladite typologie matrice/inclusions résulte d'une schématisation – et les résultats devront être interprétés par conséquent. Il est du domaine du jugement de l'ingénieur de voir si elle s'applique correctement, ce qui n'est sûrement pas le cas si l'endommagement est faible (n'ayant donc pas encore localisé) ou trop important (laissant plutôt supposer que des fissures se sont formés dans certains endroits). En tout cas, cette schématisation étant acceptée, on fait face au problème de la détermination de l'emplacement et de l'étendu des zones endommagées en supposant connaître leur niveau d'endommagement, et donc leur comportement élastique ainsi que celui du matériau sain.

D'une façon schématique, dans les deux cas cités en exemple, on doit traiter le problème suivant : identifier la distribution spatiale de l'hétérogénéité d'un solide constitué de parties souples et rigides, dont seuls les coefficients de l'élasticité sont connus. Ce problème est analysé dans ce chapitre et étudié par le biais d'un exemple. Il s'agit de l'exemple d'un fil constitué de maillons élastiques, tous d'égale longueur et de raideur variable dans une classe binaire ; on doit déterminer la composition du fil en partant de données statistiques obtenues par un essai de traction statique.

Pour les calculs on fixe à 32 le nombre de maillons et à 1/2 le rapport entre les raideurs des deux constituants. La seule information qu'on se donne sur le fil est de nature statistique, soit la probabilité du champ de déplacement du fil lors d'une mise en traction uniforme.

L'analyse inverse sera bayésienne : on cherche à inférer de l'information statistique sur le paramètre qui décrit l'hétérogénéité du fil en partant d'informations sur le comportement de ce fil dans des essais de traction et d'éventuelles informations connues à priori sur le système. En particulier on cherche à déterminer quelle valeur de ce paramètre est la plus probable compte tenu de l'information disponible, ce qui nous conduit à la recherche de la valeur dominante de la distribution obtenue à posteriori par l'analyse bayésienne. Cette recherche sera menée par une méthode de Monte Carlo par chaînes de Markov, suivant des variations de l'algorithme de Metropolis-Hastings qui ont été choisies car elles semblent convenir au traitement de systèmes de grande dimension, lorsqu'une certaine économie de calcul est souhaitée. En comparant les différents choix on obtient donc des indications sur les avantages et désavantages de chaque variation par rapport aux autres.

Les méthodes de Monte Carlo se basent sur l'échantillonnage aléatoire d'un cer-

tain domaine, qui dans le cas d'une analyse bayésienne est celui des entrées du système qu'il faut identifier. L'ensemble des échantillons constitue un chemin aléatoire et la convergence de la méthode dépend de la nature de ce chemin. Souvent, à fin de garantir la convergence, on construit des chemins qui ont les propriétés d'une chaîne de Markov. Une telle chaîne est définie par la donnée d'une distribution initiale (choix du premier point de la chaîne) et d'un noyau de transition (l'opérateur qui régit le passage d'un échantillon au suivant). Une chaînes ainsi définie est dite de Markov lorsque chaque échantillon est choisi en relation seulement avec son antécédent ; c'est cette propriété du noyau qui suffit à garantir la convergence de la méthode de Monte Carlo. Les algorithmes basés sur celui de Metropolis reposent sur un mécanisme de proposition-test-acceptation qui est mis en place à chaque pas du chemin aléatoire et configure le noyau de la chaîne de Markov.

Les variations de cet algorithme qu'on va examiner dérivent de propositions récentes : l'algorithme « *hit-and-run* » discret ([KSZ]), la technique des chaînes de Markov adaptatives ([GRS98]), et les techniques basées sur la modification de la probabilité d'accepter les échantillons dans l'algorithme de Metropolis-Hastings ([GM01, Mir98]).

L'algorithme « *hit-and-run* » discret suit des chemins aléatoires sur des ensembles discrets bornés ; sa convergence vers la distribution uniforme sur le réseau entier d'un hypercube de dimension  $n$  est polynomiale d'ordre  $n^5$  ([KSZ]). L'algorithme « *hit-and-run* » continu est très efficace pour sa capacité de sélectionner des candidats différents (« *mixing* », propriété qui réduit les probabilités que le chemin aléatoire se bloque). Cette capacité est conservée par son descendant conçu pour les systèmes discrets.

Les techniques adaptatives ont été introduites après [GRS98] dans le but d'améliorer la performance des algorithmes. Elles se basent sur l'idée de construire des noyaux de transition des chaînes dépendants d'un ou plusieurs paramètres, qui peuvent être réglés en cours de processus en fonction de l'information tirée des pas antérieurs du calcul. On obtient ainsi un chemin aléatoire dont l'aléa évolue et s'adapte aux informations que le chemin lui-même a put fournir sur le domaine qu'il parcourt. Afin de préserver la propriété de convergence de la chaîne il faut ensuite modifier la probabilité d'acceptation des candidat après les avoir testés.

Tant les techniques « *hit-and-run* » que les adaptatives ont le but de favoriser la différence entre candidat et prédécesseur : si le candidat est trop proche de son prédécesseur le chemin aléatoire risque de se bloquer en correspondance des extrema locaux de la distribution échantillonnée. Si ce problème se réduit, le risque que le candidat proposé soit rejeté après calcul augmente, ce qui rend l'algorithme plus onéreux en calculs. Du bilan entre ces deux effets dérive l'intérêt à régler le noyau de transition des chaînes de telle sorte que le taux de rejet tend vers une valeur optimale.

Une des techniques adaptatives qui a été proposée récemment se base sur l'idée de modifier au cours du processus la probabilité d'accepter les échantillons de sorte qu'il devient possible de retarder le rejet d'un candidat ([Mir98]). Ceci permet la réduction du nombre de calculs inutiles, car il donne plus des chances que un candidat soit accepté une fois qu'un calcul a été mené à partir de celui-ci. Seulement un rejet définitif conduit en effet au sacrifice du temps de calcul qui a été dédié à cette opération, alors que un rejet suspendus permet éventuellement de capitaliser ce temps de calcul si le candidat est finalement accepté, même si en retard. Pratiquement un candidat rejeté est réessayé à nouveau avant d'être définitivement rejeté ou accepté ; le nombre d'essais avant rejet définitif peut éventuellement être choisi supérieur à deux. Dans des situations plus complexes, on peut par ce biais mener le chemin aléatoire à se repartir en plusieurs branches. Le problème de la préservation du caractère markovien

de la chaîne qu'une telle approche permet de construire est résolu en conservant la propriété dite de micro-réversibilité ou de bilan détaillé. Cette propriété – postulée par la quasi-totalité des algorithmes Monte Carlo car elle est condition suffisante pour la convergence – exprime l'idée que pour n'importe quelle couple d'échantillons, disons A et B, la probabilité que le système passe de A à B doit être la même que celle d'aller de B à A. Dans le cas des techniques adaptatives avec rejet retardé, pour rétablir ce bilan des probabilités, un candidat rejeté une fois doit passer un test plus sévère pour être accepté lors qu'il est examiné à nouveaux.

Dans ce chapitre on considère l'application des méthodes de Monte Carlo par chaînes de Markov sur des systèmes discrets pour l'analyse inverse de structures de grande dimensions. Il est donc intéressant de disposer de méthodes qui permettent d'effectuer un nombre minimal de calculs directs afin de déterminer la valeur dominante de la distribution à posteriori avec approximation suffisante. A cette fin on considère de techniques adaptatives qui conduisent, éventuellement, à un rejet anticipé des échantillons candidats, basé sur une prévision de leur probabilité de réussite qui est faite sur la base des pas précédents du chemin aléatoire. Cette procédure s'apparente donc de celle discuté dans [Mir98], mais elle inverse les termes de la gestion du temps de calcul, visant – si l'analogie avec des concepts économiques est accepté – une réduction du risque des investissements plutôt qu'une récupération plus efficace des crédits.

L'estimation faite à priori de la capacité d'un échantillon de devenir un bon candidat s'appuie, comment on l'a dit précédemment, sur la connaissance de l'ensemble échantillonné qu'on a acquise pendant les pas précédents du chemin aléatoire. L'idée de base est celle de pénaliser les directions de marche qui se sont avérées infructueuses dans le passé et favoriser par contre celles qui ont généré de bons candidats. Il faut comparer ce type de procédure avec les approches qui proposent la construction d'une approximation de la distribution à posteriori afin de traiter des cas dans lesquels celle-ci nécessite de calculs onéreux ([GMR05, HJR05]).

Dans la partie suivante de ce chapitre on introduit le problème inverse qu'on traitera comme exemple, en donnant une représentation paramétrique de l'ensemble des distributions de l'hétérogénéité qu'on va prendre en compte, qu'on appellera « *constitutions* » pour abrégé. Ensuite on présente des différents générateurs de chemins aléatoires : un simple algorithme de Metropolis et un Metropolis-Hastings avec modification de la loi candidate qui conduit à des rejets anticipés éventuels. Finalement on donne des résultats numériques relatifs aux différentes variations proposées et une analyse statistique des résultats qui peuvent être obtenus par la meilleure d'entre elles.

## 8.2 Description du problème

### 8.2.1 Géométrie et définitions

On considère un fil composé de 32 segments, tous de longueur unité et raideur aléatoirement répartie dans la classe binaire  $\{1, 1/2\}$ . Par sa nature binaire, une telle constitution correspond à un nombre entier entre 0 et  $2^{32} - 1$ .

Soit  $\mathcal{N} = \{0, 1, 2, \dots, 2^{32} - 1\}$  et  $\mathcal{S}$  l'ensemble des constitutions du fil,

$$\mathcal{S} = \{(s_1, s_2, \dots, s_{32}) \mid s_i \in \{0, 1\} \forall i \in \{1, 2, \dots, 32\}\}. \quad (8.2.1)$$

Soit  $x \in [0, 32]$  une coordonnée réelle le long du fil et  $u$  le champ de déplacement

du fil avec une extrémité bloquée ( $u(0) = 0$ ) et l'autre chargée par une tension unité. A cause de la constitution du fil ce champ est continu avec dérivée uniforme par morceaux, qui vaut 1 ou 2 et subit éventuellement des sauts seulement pour des valeurs entiers de  $x$ . Soit  $\mathcal{U}$  l'ensemble des tels champs de déplacement lorsqu'on fait varier la constitution. Soit  $\mathcal{V}$  l'ensemble des champs de déplacement comme les précédents, mais qui ne respectent pas forcément la condition sur la dérivée ; évidemment  $\mathcal{U} \subset \mathcal{V}$ .

Pour une constitution donnée  $n \in \mathcal{N}$ , un modèle nous donne une estimation théorique du champ de déplacement qui lui correspond  $u \in \mathcal{U}$ . Soit  $\mathbf{k}(n)$  la liste des raideurs qui correspondent à la constitution  $n \in \mathcal{N}$  (32 éléments obtenus de la représentation binaire de  $n$  en remplaçant  $0 \rightarrow 1$  et  $1 \rightarrow 1/2$ ). Le champ de raideur est  $k(x; n) := k_{\lfloor x \rfloor}(n)$ , le déplacement calculé dans  $\mathcal{U}$  est

$$u(x; n) = \int_0^x \frac{dy}{k_{\lfloor y \rfloor}(n)} = \sum_{i=1}^{\lfloor x \rfloor} \frac{1}{k_i(n)} + \frac{x - \lfloor x \rfloor}{k_{\lfloor x \rfloor}(n)}; \quad (8.2.2)$$

On prend sur le fil 32 points à partir de  $\bar{x}_1 = 1/2$ , espacés de 1,  $\bar{x}_1, \dots, \bar{x}_{32}$ , et on définit  $\mathbf{d}\{v\}$  de dimension 32 comme la liste des valeurs pris par une fonction  $v \in \mathcal{V}$  sur ces points :

$$d_1\{v\} = v(\bar{x}_1), \dots, d_{32}\{v\} = v(\bar{x}_{32}). \quad (8.2.3)$$

On démontre facilement que pour une liste donnée  $\mathbf{d} = \mathbf{d}\{u\}$  il existe un seul champ  $u \in \mathcal{U}$ .

Soit  $\mathcal{D} \subset \mathbb{R}^{32}$  l'ensemble des listes qui sont l'image d'un champ dans  $\mathcal{U}$  et d'une constitution dans  $\mathcal{N}$ .

On imagine tester le fil en condition de traction uniforme ; le résultat d'un tel essai soit la liste  $\mathbf{m} = (m_1, \dots, m_m) \in \mathcal{M} \subset \mathbb{R}^m$ ,  $m$  étant par exemple le nombre de points de mesure qu'on se donne le long du fil (qui ne sont pas forcément 32). Une mesure  $\mathbf{m}$  ne correspond pas forcément à une liste, ni à une sous-liste, dans  $\mathcal{D}$  ; pour simplifier, on prend  $\mathbf{m} \in \mathcal{M} \subset \mathbb{R}^{32}$  comme la liste des déplacements mesurés aux points  $\bar{x}_1 = 1/2, \dots, \bar{x}_{32} = 63/2$ . La mesure d'un champ  $u \in \mathcal{U}$  est donc  $\mathbf{m} \approx \mathbf{d}\{u\} \in \mathcal{D}$ .

Pour chaque  $n \in \mathcal{N}$  existe une et une seule liste binaire  $\{b_i \in \{0, 1\} \mid i \in \{0, \dots, 31\}\}$  telle que  $n = b_i 2^i \in \mathcal{N}$  ; par conséquent l'ensemble des constitutions peut être vu comme celui des sommets d'un hypercube d'ordre 32 dont les arêtes suivent les vecteurs  $\{2^i \mid i = 0, \dots, 31\}$ . Tout  $n \in \mathcal{N}$  peut donc aussi être représenté par une liste de coordonnées  $\{b_i\}$ .

L'application  $(n, m) \in \mathcal{N}^2 \rightarrow ((2^{32} + n + m) \bmod 2^{32}) \in \mathcal{N}$  donne à  $\mathcal{N}$  la structure d'un groupe. On appelle *glissements vers la gauche* ces opérations et on les notes par un signe  $+$ . De même on considère des *glissements vers la droite*  $(n, m) \in \mathcal{N}^2 \rightarrow ((2^{32} + n - m) \bmod 2^{32}) \in \mathcal{N}$ , qui seront indiqués par le signe  $-$ .

Les glissements, vers la droite ou la gauche, sont les éléments de deux copies de  $\mathcal{N}$  ; comme tels ils peuvent être décomposés en somme de glissements le long des arêtes de l'hypercube, c'est à dire en somme de  $2^i, i \in \{0, \dots, 31\}$ . Notez que des glissements d'arrête égales  $2^i \in \mathcal{N}$  répétés  $j$  fois sont équivalents à un seul glissement d'arrête le long de la direction  $2^{i \pm j}$  (le signe dépend de la direction du glissement), alors que deux glissements opposés s'annulent mutuellement.

La dénomination de glissement est justifié par le fait qu'une telle opération fait glisser (vers la gauche ou vers la droite) de groupes de maillons le long du fil.

### 8.2.2 Probabilités

On notera toute fonction de probabilité par la même lettre  $\pi$ , en spécifiant dans l'argument son objet. Cet abus de notation n'est pas étendu aux champs, auxquels on donne des symboles différents ; dans ce cas l'argument, si présent, dénote la valeur du champ dans le point qu'il indique.

$\mathcal{N}$  avec l'algèbre borélienne et une mesure  $\Pi$ , telle que  $\Pi(\mathcal{N}) = 1$ , est un espace de probabilité ; il en va de même de  $\mathcal{S}$  et  $\mathcal{D}$ . On considère aussi que l'espace des données expérimentales  $\mathcal{M}$  est muni d'une mesure de probabilité, par exemple on considère simplement celle héritée de  $\mathcal{D}$ . Le concept de probabilité dans un espace de fonctions, tel que  $\mathcal{U}$  et  $\mathcal{V}$ , peut être précisé suivant la définition donnée par [Pug65] pour la densité de probabilité d'une fonction aléatoire  $v$ , qu'on notera  $\pi\{v\}$ .

Soit  $\pi(n)$ ,  $n \in \mathcal{N}$ , la fonction de distribution de probabilité à priori donnée sur  $n$  (cette fonction représente ce qu'on sait sur  $\mathcal{N}$  avant toute expérimentation) et  $\pi(n|\mathbf{m})$  la probabilité à posteriori sur le même  $n$ , qui dérive d'un certain résultat expérimental  $\mathbf{m} \in \mathcal{M}$  (il s'agit donc de la connaissance qu'on souhaite obtenir par les essais).

La probabilité  $\pi(n|\mathbf{m})$  nous dit comment le modèle défini par  $n$  correspond aux données expérimentales  $\mathbf{m}$  (« *fitness function* »). Elle peut être obtenue suivant Bayes (voir par exemple [Tar87]) :

$$\pi(n|\mathbf{m}) = \frac{L(n)\pi(n)}{\pi(\mathbf{m})}, \quad (8.2.4)$$

où  $L(n) := \pi(\mathbf{m}|n)$  est la vraisemblance de  $n$  par rapport aux données (il s'agit de la fonction obtenue en fixant  $\mathbf{m}$  et en faisant varier  $n$  dans la probabilité conditionnelle de  $\mathbf{m}$  pour  $n$  donné) et  $\pi(\mathbf{m})$  est la probabilité à priori de  $\mathbf{m}$  (ce qu'on sait à priori sur la probabilité d'obtenir une certaine réponse du fil lors des essais).

Afin de simplifier les plus que possibles les termes du paradigme de Bayes (8.2.4), on introduit les hypothèses suivantes :

- toute constitution du fil a à priori la même probabilité de se manifester ( $\pi(n) = 2^{-32}$ , qui signifie qu'on sait générer des nombres entiers aléatoires distribués uniformément sur  $\mathcal{N}$ ),
- toute mesure a à priori la même probabilité d'être obtenue ( $\pi(\mathbf{m}) = 2^{-32}$ ).

Les éléments de  $\mathcal{N}$  et  $\mathcal{U}$  se correspondent un à un par l'application (8.2.2) ; par conséquent  $\pi(n) = \pi\{u(\cdot; n)\}$  et  $\pi(\mathbf{m}|n) = \pi\{\mathbf{m}|u(\cdot; n)\}$ . Cette propriété nous permet de définir  $L$  en partant de la fonction aléatoire  $u$  plutôt que de la variable aléatoire discrète  $n$ , avec l'avantage que la première est directement liée à la géométrie et à la mécanique du problème.

Sous ces conditions – idéales certes, mais qui ne nuisent pas à la généralité de la démarche – (8.2.4) se réduit à

$$\pi(n|\mathbf{m}) = L\{n\}. \quad (8.2.5)$$

### 8.2.3 Information expérimentale

Afin de préciser l'objet de l'analyse inverse, on choisi au hasard un paramètre dans  $\mathcal{N}$ ,  $\bar{n} = 2935142778$ , qui dans la suite représente la « *vrai* » constitution du fil qu'on doit identifier. Sa représentation binaire  $\mathbf{b}(\bar{n})$  est :

$$(0,1,1,1,1,1,1,1,0,1,1,0,0,1,0,1,0,1,0,1,0,0,0,1,1,0,0,1,1,0,0,0) \quad (8.2.6)$$

conventionnellement, les zéros dans cette représentation correspondent aux segments plus rigides et le début de la liste correspond à l'extrémité bloquée du fil.

Le vrai champ de déplacement peut être calculé à partir de  $\bar{n}$  à l'aide de l'équation (8.2.2), mais le résultat des essais qu'on suppose avoir menés afin de générer l'information nécessaire pour l'analyse inverse, n'est pas censé être suffisant à déterminer exactement ce champ de déplacement. Fictivement, on imagine que la campagne d'essais nous a donné une distribution de probabilité sur l'ensemble des champs de déplacement, qui contient l'effet des erreurs de mesure.

Dans les limites de la présente étude on se donne cette distribution, mais on remarque que, étant qu'elle devrait être obtenue expérimentalement, d'autres choix sont possibles et également arbitraires.

On postule d'abord que la vraisemblance de la valeur en un point  $x$  d'une fonction  $v \in \mathcal{V}$  est

$$L(v(x), x) = \begin{cases} 2^{-[x]} + (\mu(x) - 2^{-[x]}) \frac{v(x)-x}{u(x;\bar{n})-x} & \text{si } x \leq v(x) \leq u(x;\bar{n}) \\ 2^{-[x]} + (\mu(x) - 2^{-[x]}) \frac{2x-v(x)}{2x-u(x;\bar{n})} & \text{si } u(x;\bar{n}) < v(x) \leq 2x \\ 0 & \text{sinon} \end{cases} \quad (8.2.7)$$

où  $u(x; \bar{n})$  est la valeur exacte du déplacement en ce point et la valeur maximale  $\mu$  est obtenue par normalisation :

$$\int_x^{2x} L(y, x) dy = 1 \Rightarrow \mu(x) := \frac{2}{x} - 2^{-[x]}. \quad (8.2.8)$$

Notez que  $L(u(x), x) > 0 \forall x$  si  $v \in \mathcal{U}$ .

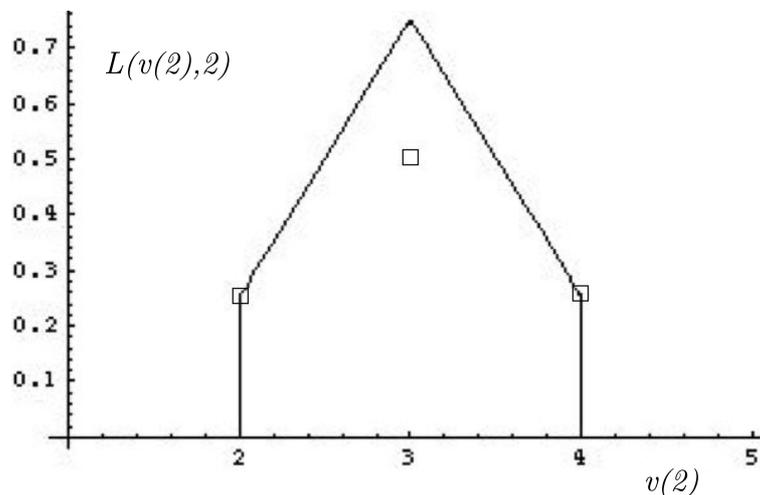


FIG. 8.1 –  $L(v(x), x)$  pour  $x = 2$ ; les carrés correspondent aux densités de Dirac de la distribution de probabilité à priori  $\pi(u(x; n))$  pour  $x = 2$ .

La probabilité fonctionnelle  $L\{v\}$  peut être obtenue de (8.2.7) par des passages proches de ceux donnés dans [Pug65] pour les fonctions aléatoires. Dans le but de mener des calculs numériques il est néanmoins intéressant de choisir un ensemble de

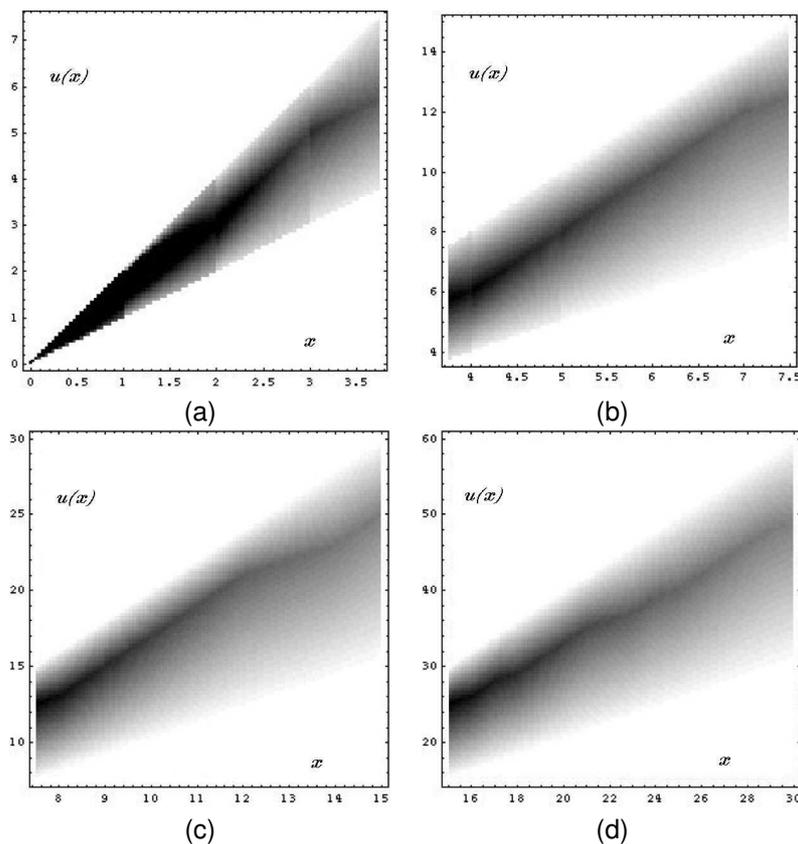


FIG. 8.2 – Tracé de la fonction  $L(u(x), x)$  par niveaux de gris dans des différentes parties du plan  $x - u(x)$ ; le blanc correspond au zéro.

coordonnées qui permet de représenter avec approximation adéquate le processus à la limite qui est considéré dans [Pug65]. Sans rentrer dans les détails il suffit d'observer que les points de mesure pris sur le fil,  $\bar{x}_i$   $i = 1, \dots, 32$ , sont tels que, pour  $u \in \mathcal{U}$ ,

$$L\{u\} = \prod_{i=1}^{32} L(u(\bar{x}_i), \bar{x}_i) / \prod_{i=1}^{32} L(u(\bar{x}_i; \bar{n}), \bar{x}_i). \quad (8.2.9)$$

Ce résultat peut être prouvé à partir des définitions données dans [Pug65], mais on considère qu'il correspond assez bien à l'intuition pour rendre l'inclusion de sa démonstration non nécessaire dans ce contexte.

La fonction (8.2.9) permet de tester les algorithmes qu'on va proposer. Pour la définir on a fait appel à la connaissance de la fonction déplacement qui correspond à la vraie constitution du fil, qui est l'inconnue du problème inverse. Il ne faut donc pas oublier qu'il s'agit d'un passage fictif, car en réalité  $L$  doit être une donnée expérimentale. Par ailleurs, grâce à cette définition fictive, la valeur dominante de la vraisemblance correspond exactement à la vraie constitution du fil. Afin de tester ces algorithmes on cherchera donc à déterminer la valeur dominante de  $L$ .

Dans la figure 8.3 on trace ladite fonction dans un voisinage de la vraie constitution, c'est à dire de la valeur dominante qu'on suppose ignorer et on se propose de chercher par méthode Monte Carlo.

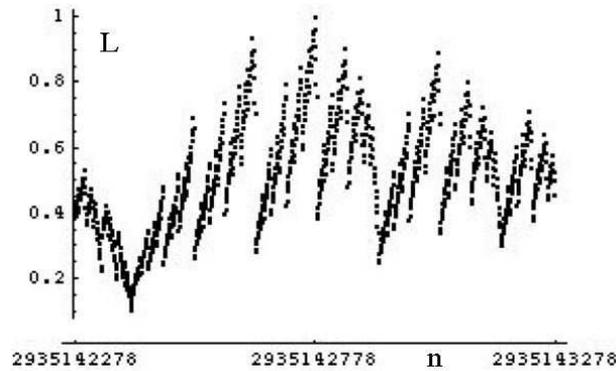


FIG. 8.3 – Vraisemblance (8.2.9) dans un voisinage de 2 935 142 778, où elle prend sa valeur maximale 1.

## 8.3 Méthode de Monte Carlo

### 8.3.1 Algorithmes utilisés

On considère deux différents algorithmes qui, suivant l'emploi courant, on appelle

- algorithme de Metropolis indépendant (utilisé comme base de comparaison)
- algorithme de Metropolis-Hastings à marche aléatoire,

Dans les deux cas on se donne de façon aléatoire une constitution de départ  $n_0$  et on construit un chemin aléatoire  $\{n_0, n_1, \dots, n_i, \dots\}$  ayant les propriétés d'une chaîne de Markov grâce à des règles de proposition et acceptation des candidats appropriées. Les algorithmes employés ne diffèrent que pour ces règles.

A chaque pas  $i$  un candidat  $n^*$  est choisi sur la base d'une distribution qu'on appelle loi instrumentale ou candidate,  $\pi(n^*|n_i)$ , telle que

$$\sum_{n^* \in \mathcal{N}} \pi(n^*|n_i) = 1 \quad \forall n_i \in \mathcal{N}, \quad (8.3.1)$$

et accepté (c'est à dire  $n_{i+1} = n^*$ ) ou rejeté avec probabilité d'acceptation

$$\pi_{MH}(n_{i+1} = n^*) = \min \left\{ \frac{\pi(n^*|\mathbf{m})\pi(n_i|n^*)}{\pi(n_i|\mathbf{m})\pi(n^*|n_i)}, 1 \right\} \quad (8.3.2)$$

Le choix de la loi candidate fait la différence entre les deux algorithmes considérés ; elle est uniforme pour le Metropolis indépendant, ce qui réduit dans ce cas la probabilité d'acceptation à

$$\pi_M(n_{i+1} = n^*) = \min \left\{ \frac{\pi(n^*|\mathbf{m})}{\pi(n_i|\mathbf{m})}, 1 \right\} \quad (8.3.3)$$

### 8.3.2 Loi candidate

Les algorithmes qu'on va étudier reposent sur une loi candidate de la forme  $\pi(n^*|n_i) = \pi(n^* - n_i)$ , où  $n^* - n_i$  est un glissement de  $n_i$  à  $n^*$ .

Pour définir le type de glissements entre constitutions qui seront pris en compte dans l'algorithme, on rappelle la définition donnée §8.2.1. Pour une constitution donnée  $n$  (vue comme sommet d'un hypercube), l'opération

$$n \rightarrow n + e \sum_{j=1}^p a_j 2^{q_j}, \quad (8.3.4)$$

où  $e \in \{-1, 1\}$ ,  $\{q_j | j = 1, \dots, p\}$  est une liste d'entiers successifs modulo 31, qu'on appellera indices d'arrêtes, et  $\{a_j \in \{0, 1\} | j = 1, \dots, p\}$  sont les amplitudes d'arrêtes, identifie une constitution dans un voisinage de  $p$  sommets autour de  $n$ .

On appelle pas sur  $p$  sommets n'importe quel glissement qui peut être défini par le choix de

1. sa *longueur*  $p \in \{0, \dots, 31\}$ , c'est à dire le nombre maximale de sommets qui seront éventuellement intéressé par le mouvement,
2. sa direction  $e$  (gauche, +1, ou droite, -1),
3. son premier indice d'arrête  $q_1 \in \{0, \dots, 31\}$  (car tout autre glissement d'arrête qui compose le mouvement pourra être choisi par la règle  $q_j = (q_1 + j - 1) \bmod 31$ ,  $j = 2, \dots, p$ ),
4. les amplitudes  $a_j \in \{0, 1\}$ ,  $j = 1, \dots, p$  (c'est à dire l'ensemble des glissements d'arrête qui sera effectivement pris en compte dans la liste  $\{q_j\}$ ).

Avec la définition de pas donnée, pour un choix de  $p$ , une constitution  $n^*$  est *hors d'atteinte* en un seul pas à partir d'une autre constitution  $n$  si  $n^* - n$  est donné par la somme de plus de  $p$  glissement d'arrête consécutifs ( $2^0$  étant pris après  $2^{31}$ ), ou s'il a des amplitudes d'arrête de signe opposé. Ces relations étant symétriques, il s'avère que  $n^*$  est hors d'atteinte à partir de  $n$  si et seulement si  $n$  l'est à partir de  $n^*$ .

Dans le but de construire un chemin aléatoire il est utile de proposer de pas sur l'hypercube de longueur bien choisie, de façon à assurer le compromis demandé entre le besoin de propager le chemin dans des zones différentes et celui d'économiser les rejets.

Pour un choix de  $p$  on introduit :

1. une probabilité  $\pi(e)$  sur la direction du pas  $e \in \{-1, 1\}$ ,
2. une probabilité  $\pi(q_1)$  de faire un pas ayant  $q_1$  comme première arrête de la liste,
3.  $p - 1$  probabilités  $\pi(q_j)$ ,  $j = 2, \dots, p$ , qu'un pas emprunte simultanément d'autres arrêtes.

Les différents choix qui correspondent à cette liste conduisent à l'algorithme (8.3.3) ou (8.3.2) suivant si la symétrie  $\pi(n^* | n_i) = \pi(n_i | n^*)$  est conservée ou non à tous pas  $i$ .

On considère effectivement ces deux cas :

- on construit une chaîne de Markov en proposant naïvement et en testant à chaque pas par le biais de (8.3.3) des mouvements générés par des probabilités  $\pi(e)$  et  $\pi(q_j)$ ,  $j = 1, \dots, p$ , qui ne dépendent pas des résultats précédents ;
- on adopte une loi candidate qui dépend de l'issue d'un certain nombre de pas précédents (et on se donne une probabilité d'acceptation modifiée à fin de conserver la réversibilité du processus), de façon à éviter des pas avant qu'ils donnent lieu à un candidat, et donc avant avoir mené un test, qui sera de type (8.3.2).

Ce deuxième algorithme est pris en compte en vue d'applications dans lesquelles l'effort, en termes de calcul numérique, nécessaire pour l'évaluation de  $\pi(n^* | \mathbf{m})$  dans (8.3.2) doit être évité dans les limites du possible.

Dans le premier cas la probabilité  $\pi(q_1)$  est choisie uniforme dans  $\{0, \dots, 31\}$ , tandis que les autres  $p-1$  arrêtes ont probabilité  $1/2$  de faire partie du glissement. La direction est aussi choisie sur la base d'une probabilité uniforme  $\pi(e) = 1/2$ .

Par ces choix, aucun des paramètres et des lois de distribution qui contribuent à la définition d'un nouveau candidat ne dépend ni de la constitution correspondante au dernier pas du chemin ni des pas précédents. Par conséquent  $\pi(n^* | n_i)$  (qui est la

probabilité que la constitution  $n^*$  soit atteinte au  $(i + 1)$ -ème pas du chemin) est nulle si et seulement si  $n^*$  est hors d'atteinte à partir de  $n_i$ . En tout cas  $\pi(n^*|n_i) = \pi(n_i|n^*)$  et (8.3.3) peut être utilisé de façon immédiate pour construire une chaîne de Markov.

L'algorithme qu'on obtient ainsi peut être vu comme un cas particulier d'un hit-and-run discret de [KSZ]. Suivant des définitions proches de celles données à la fin du §8.2.1, l'ensemble des constitutions peut être considéré comme un réseau d'entiers dans un hypercube de dimension  $Q$  et taille  $32/Q$ ,  $Q \in \{1, 2, 4, 8, 16, 32\}$  (au §8.2.1 on se place dans le cas  $Q = 32$ ). Cet ensemble peut alors être encadré dans une boîte dont la taille et les dimensions sont celles données ci-dessus, et on peut suivre [KSZ] pour construire son chemin aléatoire markovien. La taille de la boîte qu'il faut prendre à cette fin correspond à la longueur maximale des pas qu'on fixe dans l'algorithme que nous venons de proposer. Notez que ledit algorithme hit-and-run s'obtient du nôtre pour n'importe quel  $Q \leq 16$ .

Dans le cas du seconde algorithme proposé, on définit chaque pas sur la base d'une procédure à deux étapes :

1. on choisit un ensemble de glissement par la procédure naïve de l'algorithme précédent,
2. on teste chacun des glissements de cet ensemble par un critère qui dépend du résultat de tout test d'acceptation mené précédemment (8.3.2), et éventuellement on en rejette un certain nombre.

Cela donne lieu, éventuellement, à un *rejet anticipé* des candidats. La probabilité de rejet anticipé sera faite évoluer comme suit : chaque fois que un candidat est rejeté sur la base du test (8.3.2), les directions menant de l'étape actuelle du chemin vers ce candidat seront pénalisées, alors que leur choix dans un mouvement successif sera favorisé dans le cas contraire. Pour obtenir ce résultat on introduit un handicap  $h_r^{(i)}$ ,  $r \in \{-32, \dots, -1, 1, \dots, 32\}$ , pour chaque glissement d'arrêt orienté  $\text{sgn}(r)2^{|r|-1}$ , qu'il sera augmenté d'une unité à chaque rejet, réduit à chaque acceptation. On fait ensuite appel à l'handicap total pour fixer une barrière lors du choix des amplitudes d'arrêt  $a_j$  qui définissent le pas  $n^* = n_i + e \sum_{j=1}^p a_j 2^{q_j}$  ( $e \in \{-1, 1\}$ ) :

- $a_j = 0$  si un tirage au sort de probabilité uniforme dans  $\mathcal{I}^{(i)}$  est inférieur à l'handicap  $h_{e_j}^{(i)}$ ,
- $a_j = 1$  sinon.

Par cet algorithme un pas peut être rejeté (si par exemple toute direction de glissement qu'il comporte se font refuser à cause de leur handicap) avant de devenir un vrai candidat à tester pour faire partir de la chaîne en construction. Si un tel événement se produit  $s$  fois successives, la borne supérieure de l'intervalle de barrage  $\mathcal{I}^{(i)}$  est levée de  $s \div m$  unités, de sorte à débloquent le processus (seule cette borne dépend donc de  $i$  dans  $\mathcal{I}^{(i)}$ ).

Le rejet anticipé modifie la probabilité à priori sur  $\mathcal{N}$  ; plus exactement, si un glissement donné (positif ou négatif) d'une certaine zone du fil s'est montré inutile dans des pas précédents, alors l'algorithme résistera avant de le prendre en compte à nouveau ; si, au contraire, il existe des glissements d'une partie du fil qui se sont montrés utiles, alors l'algorithme aura tendance à les proposer nouveau.

Ce comportement peut être réglé par le choix de trois paramètres : la largeur initiale de la fenêtre de l'handicap  $\mathcal{I}^{(0)}$ , la valeur initiale de celui-ci  $h_r^{(0)}$  (pris uniformément sur  $r \in \{-32, \dots, -1, 1, \dots, 32\}$  au début du chemin) et le paramètre de déblocage  $m$ . Dans les exemples numériques présentés ci-dessus on prend  $\mathcal{I}^{(0)} \equiv [0, 10]$  et  $h_r^{(0)} = 2$

pour tout  $r \in \{-32, \dots, -1, 1, \dots, 32\}$ , et on fixe  $m = 5$ .

Pour rétablir la réversibilité du chemin il faut inclure le rapport entre lois candidates  $\pi(n_i|n^*)/\pi(n^*|n_i)$  dans la probabilité d'acceptation, comme montré dans (8.3.2).

Si on analyse, dans le cas du rejet anticipé, la probabilité des constitution qui peuvent être atteintes à partir de  $n_i$ , on voit que  $\pi(n^*|n_i) = \pi_1(n^*|n_i)\pi_2(n^*|n^{**})$ , où  $\pi_1$  est le même que dans la loi candidate naïve du premier algorithme et  $\pi_2$  est la probabilité que un candidat réussisse le test de rejet anticipé. La symétrie de  $\pi_1$  implique

$$\frac{\pi(n_i|n^*)}{\pi(n^*|n_i)} = \frac{\pi_2(n^{**}|n^*)}{\pi_2(n^*|n^{**})}, \quad (8.3.5)$$

et donc on peut se concentrer sur le rapport au membre droit de (8.3.5) pour expliciter le ratio des lois candidates dans (8.3.2).

Soient  $n^*$  et  $n^{**}$  tels qu'on puisse atteindre l'un à partir de l'autre ; par définition, on peut décomposer  $n^* - n^{**}$  en  $p$  glissement d'arrêt de signe égale, et  $n^{**} - n^*$  sera donné par la somme opposée :

$$\exists! q_1 \in \{0, \dots, 31\}, \{a_j\} \in \{0, 1\}^p, e \in \{-1, 1\} : \quad (8.3.6)$$

$$n^* - n^{**} = e \sum_{j=1}^p a_j 2^{q_1+j-1};$$

la probabilité d'une transition de  $n^{**}$  à  $n^*$  au pas  $i$  suivant la direction  $e$  s'obtient alors en combinant les  $p$  probabilités indépendantes des différents glissements d'arrêt qui compose ce pas,

$$\pi_2(n^*|n^{**}) = \prod_{j=1}^p \min \left( 1, \frac{\max \mathcal{I}^{(i)} - h_{ej}^{(i)}}{\text{meas} \mathcal{I}^{(i)}} \right), \quad (8.3.7)$$

et la probabilité  $\pi_2(n^{**}|n^*)$  de revenir en arrière de  $n^*$  à  $n^{**}$  est obtenue de (8.3.7) en changeant  $h_{ej}^{(i)}$  avec  $h_{-ej}^{(i)}$ .

En conclusion (8.3.5) et (8.3.7) nous permettent d'expliciter le ration des lois candidates qui apparaît dans (8.3.2) :

$$\frac{\pi(n_i|n^*)}{\pi(n^*|n_i)} = \frac{\prod_{j=1}^p \min \left( \text{meas} \mathcal{I}^{(i)}, \max \mathcal{I}^{(i)} - h_{-ej}^{(i)} \right)}{\prod_{j=1}^p \min \left( \text{meas} \mathcal{I}^{(i)}, \max \mathcal{I}^{(i)} - h_{ej}^{(i)} \right)}. \quad (8.3.8)$$

## 8.4 Résultats numériques

Les deux algorithmes ont été testés en prenant en compte des pas de différente longueur ; dans la suite on compare les résultats de ces choix. L'algorithme à candidature modifiée basé sur des pas de longueur 4 se montre plus efficace que les autres qui ont été testés ; par conséquent on mène une analyse statistique des résultats qui peuvent être obtenus dans ce cas, afin d'évaluer le nombre minimale de pas qu'il faut inclure dans le chemin aléatoire pour obtenir, de façon suffisamment probable, une évaluation raisonnable de la valeur dominante de la vraisemblance.

### 8.4.1 Comparaison des algorithmes

Dans la figure 8.4 on montre des chemins de 1000 pas construits par l'algorithme simple de Metropolis, en prenant des pas qui peuvent comprendre plus qu'un seul

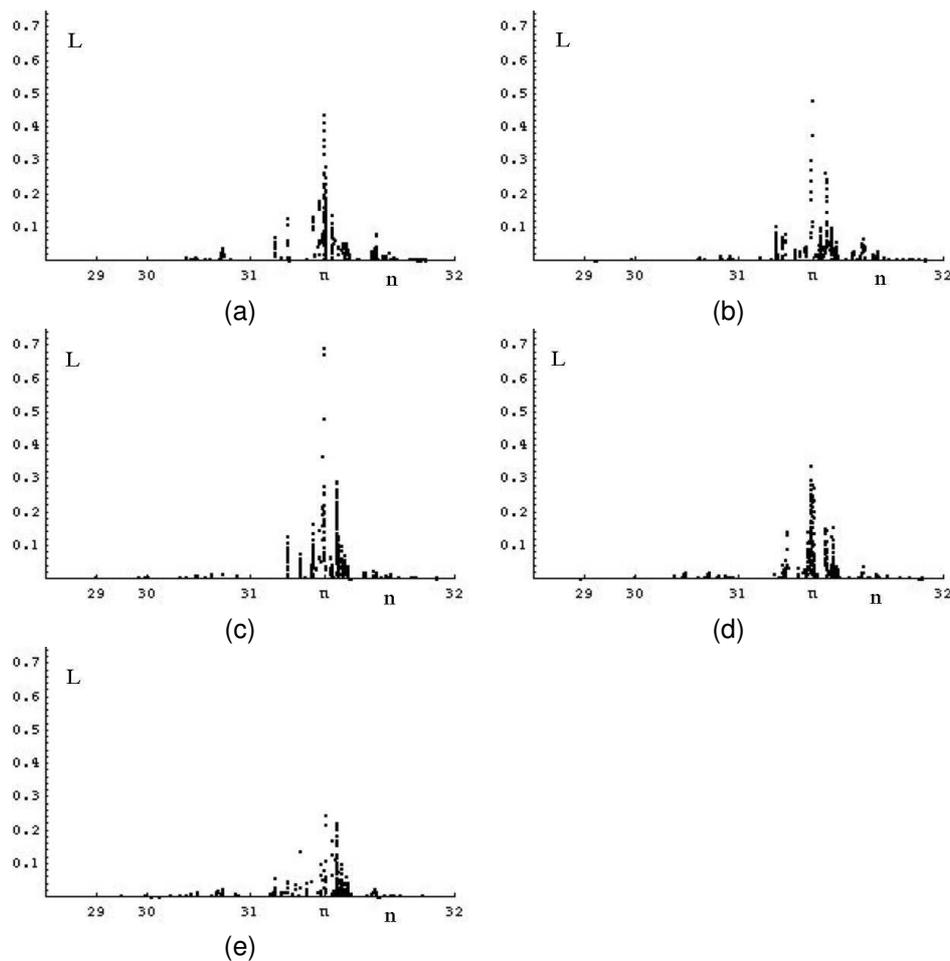


FIG. 8.4 – Chemins de 1000 pas dans le plan constitution vs. vraisemblance, obtenus par algorithme de Metropolis indépendant, avec pas composés au plus de 1, 2, 4, 8, 16 glissements d'arrête (montrés respectivement du haut à droite vers le bas à gauche).

glissement d'arrête. On montre l'issue des calculs menés en prenant 1, 2, 4, 8 ou 16, glissement d'arrête par pas.

Ces résultats montrent que cet algorithme requiert des chaînes relativement longues pour atteindre des résultats relativement peu approchés de la valeur dominante de la vraisemblance (qui est connue dans cet exemple). Les taux de rejet dans la construction du chemin sont : 39.2%, 42.7%, 48.4%, 57.6% et 77.3%, respectivement pour les cas 1, 2, 4, 8 et 16.

Des chaînes de 1000 pas construites par l'algorithme à loi candidate modifiée sont présentées en figure 8.5. Le taux de rejet total (anticipé et final) pour les différents test est 33.2%, 36.0%, 37.0%, 41.8%, 43.0% et 40.7%, pour des pas dont la longueur maximale est 1, 2, 4, 8, 16 et 32 (dans ce cas chaque pas couvre toute la longueur du fil), respectivement. A noter que ces taux de rejets sont tous inférieurs à l'optimum de Roberts, Gelman et Gilks (76.6%, démontré pour de densités de probabilité normales).

Le cas donné par le deuxième algorithme, avec 4 glissements d'arrête par pas au plus, donne les meilleurs résultats et mérite plus d'attention.

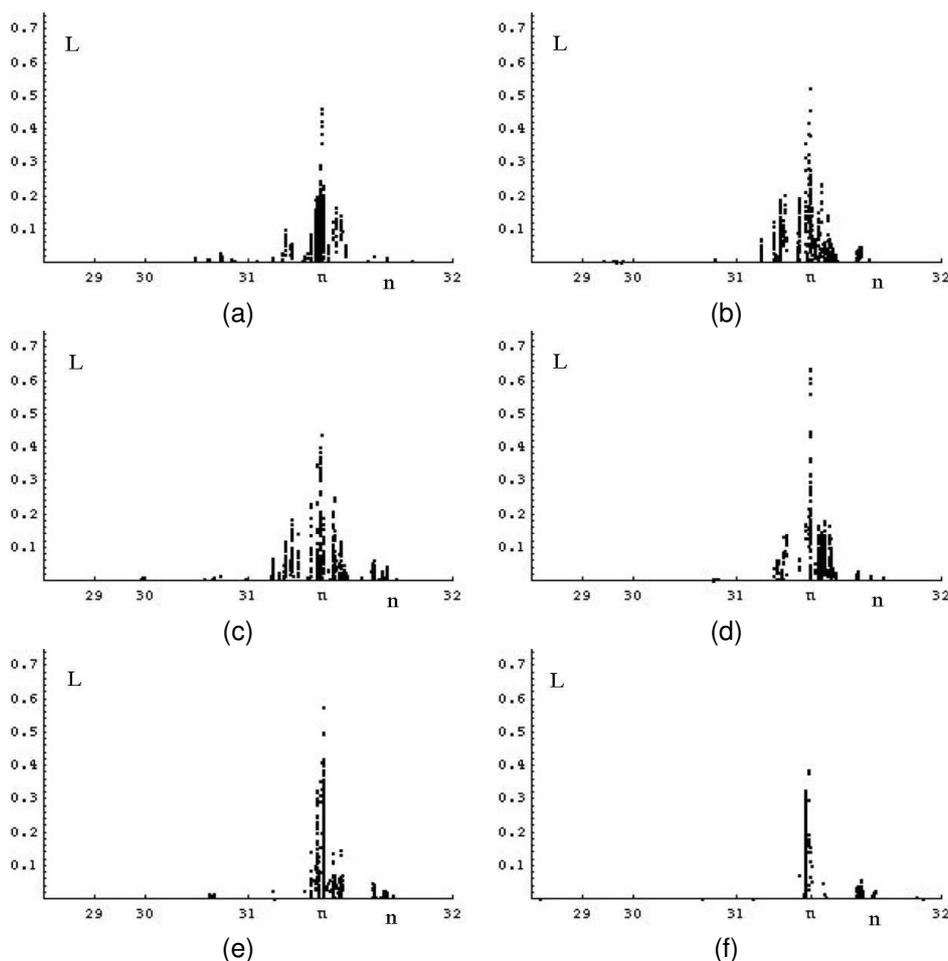


FIG. 8.5 – Chemin de 1000 pas dans le plan constitution vs. vraisemblance, obtenus par algorithme de Metropolis-Hastings à marche aléatoire donnant lieu à rejets anticipés, pour des pas composés au plus de 1, 2, 4, 8, 16, 32 glissement d'arrêt (montrés respectivement du haut à droite vers le bas à gauche).

En réduisant la probabilité d'acceptation (8.3.2) à

$$\pi_{MHr}(n_{i+1} = n^*) = \max \left( 0, \frac{\pi_{MH}(n_{i+1} = n^*) - r}{1 - r} \right) \quad r \in ]0, 1[, \quad (8.4.1)$$

on focalise d'avantage sur les échantillons plus vraisemblables et on réduit donc le nombre des calculs numériques à mener lorsqu'on cherche la valeur dominante de la vraisemblance. On considère  $r = 0.8$  et on analyse les résultats de plusieurs calculs afin de définir l'ensemble de paramètres le plus appropriés pour ce type de calcul.

En plus on se donne deux cas : l'algorithme avec rejet anticipé tel que décrit précédemment, mais avec acceptation (8.4.1), et un cas proche simplifié, obtenu en imposant (8.3.8) égale à 1. Ce dernier cas est proposé sur la base de considérations empiriques et produit des chaînes non réversibles (donc non markoviennes, ce qui, par ailleurs, ne garantit plus la convergence du calcul), mais fournit de résultats meilleurs du précédent au moins dans le cas présent.

Pour obtenir des informations sur le chemin le plus court qu'il faut se donner afin de prédire de façon acceptable la valeur dominante de la vraisemblance, on compare les résultats de chemin dont la longueur est limitée à 75, 100 et 250 pas. Les résultats, donnés en figure 8.6 et 8.7 montrent que le cas de chemin limités à 100 pas fournit

des résultats suffisamment approché de la valeur dominante recherchée et, au même temps, implique un coût numérique raisonnable.

La figure 8.6 donne les histogrammes du maximum estimé de la vraisemblance dans les trois cas : les résultats ne sont pas bons si on se limite à 75 pas (rarement vraisemblables plus qu'à 20%) ; avec 100 pas on obtient des meilleurs résultats (vraisemblance à 60% dans la plus part des cas) ; les résultats obtenus avec 250 pas ne sont pas remarquablement meilleurs que ceux obtenus avec 100.

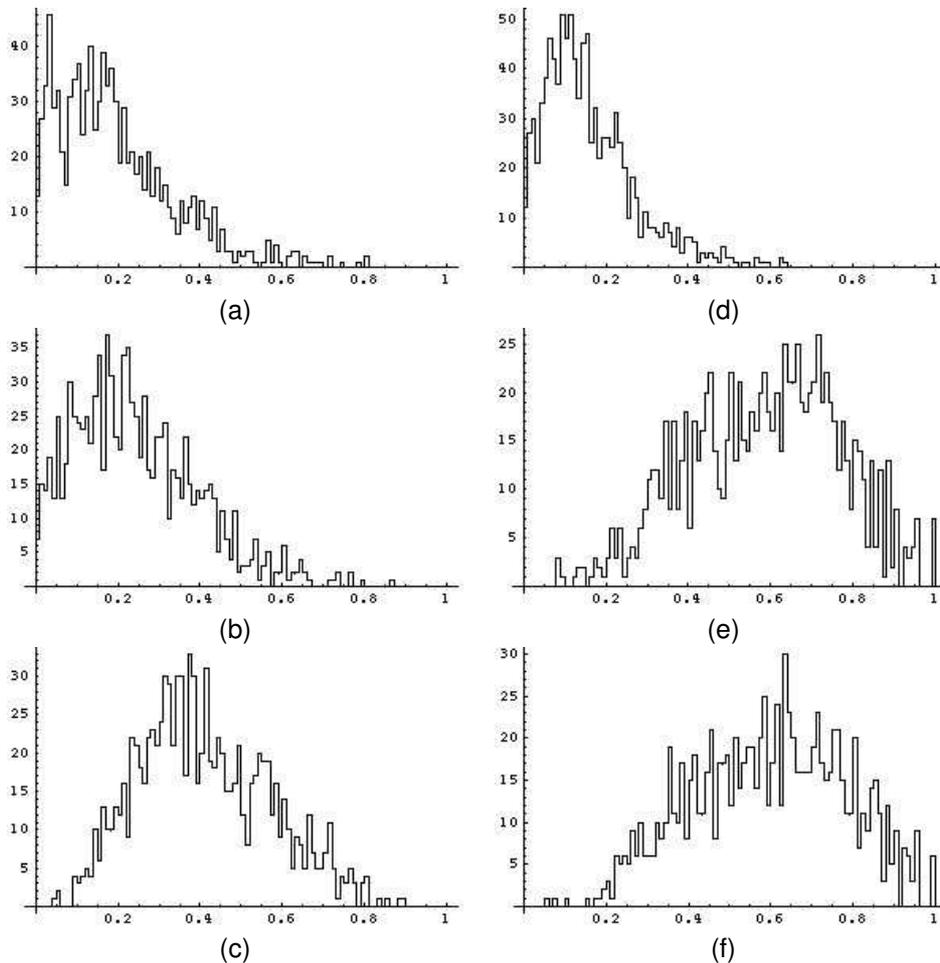


FIG. 8.6 – Histogrammes de la valeur maximale obtenue pour la vraisemblance (valeur théorique 1) dans 1000 chemins de 75, 100 et 250 pas respectivement de haut en bas droite. La première colonne (a, b, c) contient les résultats de l'algorithme à rejet anticipé, la deuxième (d, e, f) montre ceux du processus non markovien.

La figure 8.7 montre le nombre de maillons faux identifiés sur le fil dans les différents cas. Elle confirme la faiblesse du choix de la limite à 75 pas et la bonne performance du cas à 100 pas déjà notée en figure 8.6. Par rapport au nombre de maillons faux le cas avec 250 pas s'avère plus intéressant (ce qui n'était pas évident dans la comparaison avec le cas à 100 pas menée sur la base de la vraisemblance).

Notez qu'on adopte un comptage des maillons faux qui considère chaque différence avec le vrai fil comme un erreur. Par conséquent, par exemple, la présence d'une inclusion souple décalé d'une longueur par rapport à sa position correcte correspond à deux maillons faux.

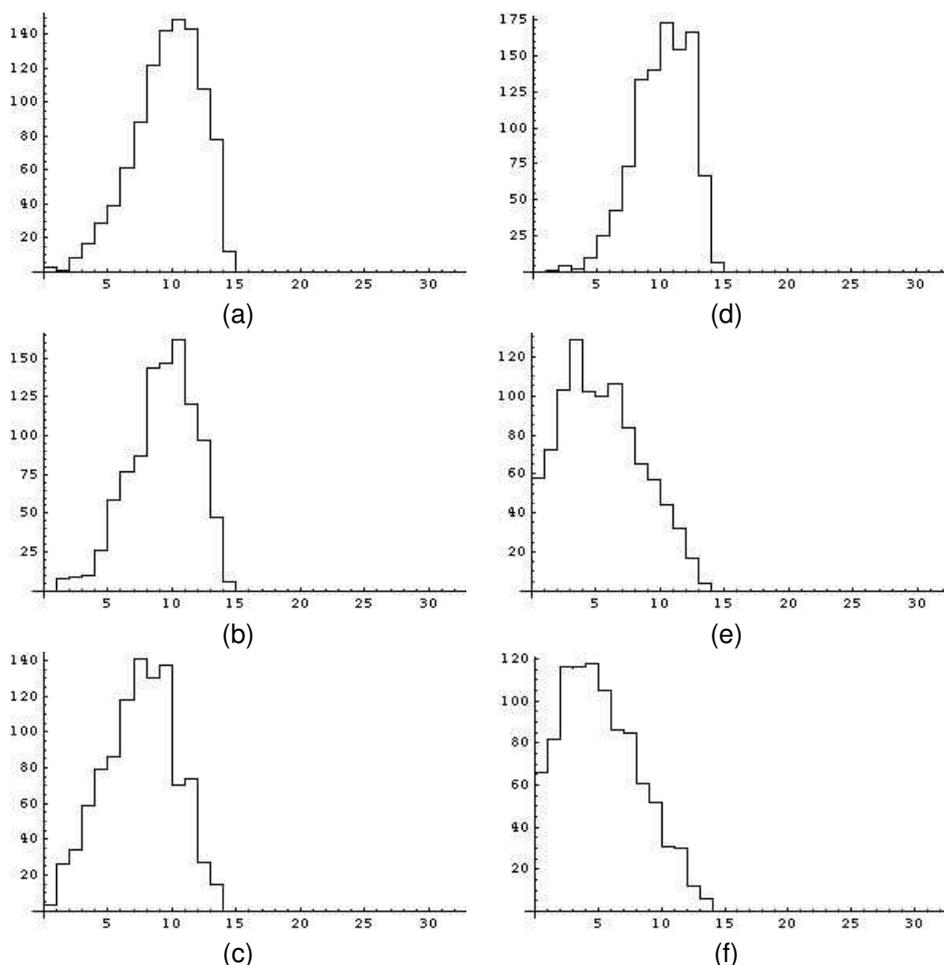


FIG. 8.7 – Histogrammes du nombre de maillons faux trouvés avec 1000 chemins de 75, 100 et 250 pas, respectivement de haut en bas droite. Les deux colonnes se réfèrent au processus markovien et non comme à la figure 8.6.

## 8.5 Conclusions

L'application de la méthode de Monte Carlo à l'analyse inverse d'un fil élastique hétérogène, présentée dans ce chapitre, montre la possibilité d'application dans des cas plus difficiles, mais caractérisés par la recherche d'un motif composé par des constituants connus.

En particulier un algorithme de Metropolis-Hastings modifié a été présenté qui donne des bons résultats avec un coût (en terme de nombres de calculs directs) raisonnable : pas plus que 100 calculs directs se rendent nécessaires pour obtenir, dans la plus part des cas, résultats quantitatifs sur des détails de  $1/32$  de la longueur du fil. En figure 8.8 on montre ce qu'on peut obtenir avec un chemin de 25 000 échantillons par la méthode proposée la plus efficace (empirique, non markovienne).

La méthode a été appliquée au cas d'une poutre en béton armé en flexion trois points soumise à des chargements supérieurs à la limite de son élasticité. L'application se base sur l'idée qu'il se forme sur la poutre un motif de zones plus ou moins raides qu'on cherche à identifier. En adoptant un modèle d'endommagement scalaire pour le béton (qui prévoit une éventuelle réduction de la raideur du matériau en fonction

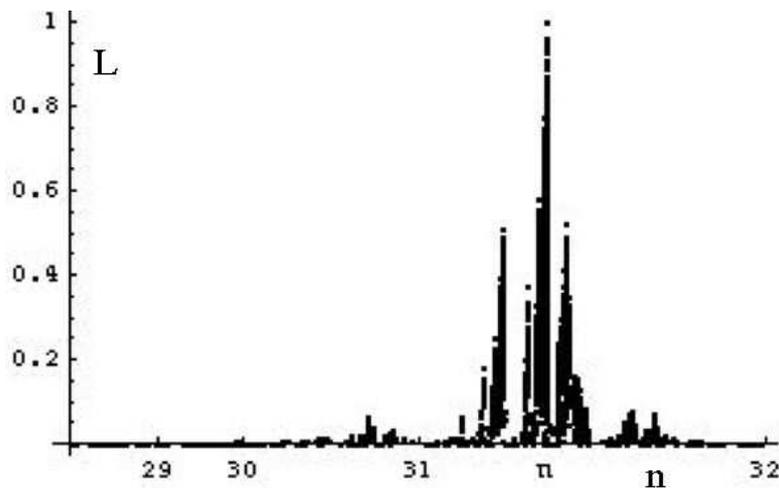


FIG. 8.8 – Chemin de 25 000 pas.

d'un paramètre compris entre 1, pour le matériau sain, et 0, qui dénote la rupture), on observe en fait la formation de bandes fortement endommagées dans la zone centrale à l'intrados de la poutre, qui sont alternées avec des bandes très peu endommagées ou saines (voir figure 8.9, image inférieure).

Ce résultat numérique classique correspond, en réalité, à l'ouverture de fissures se propageant à partir des fibres tendues de l'intrados. Mais si on accepte la description simplifiée du phénomène donnée par la théorie de l'endommagement, un problème inverse se pose de la nature de celui discuté dans ce chapitre : déterminer quel est l'emplacement des zones endommagées, dont la raideur est supposée connue (égale à 10% de la raideur initiale), qui donne la meilleure correspondance entre la déformée mesurée expérimentalement et celle calculée.

A cette fin on considère une expérience fictive, pendant laquelle la poutre est endommagée par des chargements cycliques lents. A la fin d'un certain nombre de tels cycles (cinq dans l'exemple traité) on imagine instrumenter la poutre avant de conduire un dernier cycle : le but de l'analyse inverse est l'identification de l'état d'endommagement à la fin de tous les cycles, sans connaître celui à la fin de l'avant dernier, mais en connaissant la déformée de la poutre au cours du dernier cycle.

Sans rentrer dans les détails du calcul, on montre en figure 8.9 un champ d'endommagement identifié par la méthode inverse et le champ calculé directement.

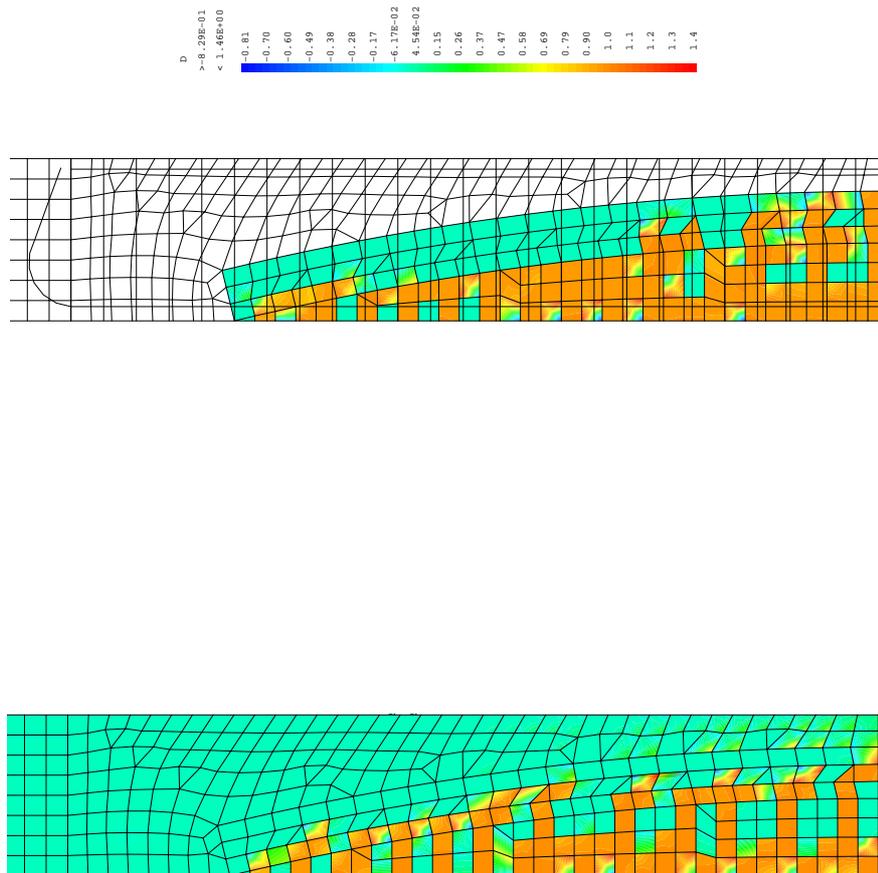


FIG. 8.9 – Comparaison du champ de l'endommagement identifié par méthode inverse (en haut) avec les champs de référence obtenus en partant de la distribution « exacte » de l'endommagement calculée à partir de l'état naturel suivant tous les cycles de chargement (en bas). Seulement la moitié à gauche de la poutre est représentée.

## Remerciements

La recherche a été menée au sein du *Commissariat à l'Énergie Atomique* dans le cadre d'une action sur la durée de vie des structures en béton financée par le CEA et par *Électricité de France*. Les calculs numériques dont le résultat est présenté figure 8.9 ont été menés avec CAST3M.

# Bibliographie

- [BB02] Paul E. Barbone and Jeffrey C. Bamber. Quantitative elasticity imaging : what can and cannot be inferred from strain images. *Phys. Med. Biol.*, 47 :2147–2164, 2002.
- [GM01] Peter J. Green and Antonietta Mira. Delayed rejection in reversible jump metropolis-hastings. *Biometrika*, 88 :1035–1053, 2001.
- [GMR05] A. Guillin, J. M. Marin, and C. P. Robert. Estimation bayésienne approximative par échantillonnage préférentiel. *Revue de Statistique Appliquée*, LIII :79–95, 2005.
- [GRS98] Wally R. Gilks, Gareth O. Roberts, and Sujit K. Sahu. Adaptive markov chain monte carlo through regeneration. *Journal of the American Statistical Association*, 93 :1045–1054, 1998.
- [HJR05] James P. Hobert, Galin L. Jones, and Christian P. Robert. Using a markov chain to construct a tractable approximation of an intractable probability distribution. *Scandinavian Journal of Statistics*, under press, 2005.
- [KSZ] Seksan Kiatsupaibul, Rober L. Smith, and Zeld B. Zabinsky. A discrete hit-and-run algorithm for generating samples from general discrete multivariate distributions. Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor.
- [Mir98] Antonietta Mira. *Ordering, slicing and splitting Monte Carlo Markov chains*. PhD thesis, University of Minnesota, 1998.
- [Pug65] V. S. Pugachev. *Theory of random functions and its application to control problems*. Pergamon Press, 1965.
- [Tar87] A. Tarantola. *Inverse problem theory – Methods for data fitting and parameter estimation*. Elsevier, 1987.



## Chapitre 9

# Algorithmes pour l'optimisation globale

Guillaume Dutilleux

### 9.1 Introduction

De façon générale, lorsqu'une relation de cause à effet est identifiée, l'expression *problème inverse* s'applique à des situations où seul l'effet est accessible et où la relation ne peut être inversée. En l'absence de solution analytique, la résolution d'un problème inverse, c'est-à-dire la détermination des causes à partir des effets se ramène un problème d'optimisation. La fonction optimisée traduit en général l'écart entre des effets observés et une modélisation des effets à partir des causes.

Malheureusement, les fonctions à optimiser dans ce contexte ne sont que très rarement convexes sur l'espace exploré. Le bagage classique d'un cours d'analyse numérique n'est donc d'aucun secours en la matière, et il faut se tourner vers des algorithmes d'*optimisation globale*. Ce chapitre se fixe comme objectif d'en présenter quelques uns parmi les plus importants de ceux qui sont fondés sur des métaphores ou *métaheuristiques* et de mettre en évidence les aspects remarquables de chacun.

Le corps de ce chapitre présente tout d'abord les grands principes sous-jacents à ces algorithmes. On verra qu'ils tirent leur efficacité empirique d'un certain nombre d'observations élémentaires. La présentation des algorithmes commence par une métaphore classique tirée de la physique : l'algorithme du *recuit simulé*. La suite traite de métaphores issues de la biologie, en commençant par une métaheuristique dérivée de l'étude du comportement social, qui s'incarne dans l'algorithme d'*essaims de particules*. On pourra apprécier sa facilité de mise en oeuvre et son efficacité. Ce rapide tour d'horizon se poursuit avec les *algorithmes évolutionnaires*, qui s'inspirent quant à eux de la sélection naturelle. Dans cette catégorie, la présentation se limite aux *algorithmes génétiques* et aux *stratégies d'évolution*. Le point fort des premiers est la polyvalence, celui des seconds l'efficacité en variables continues. La présentation des algorithmes choisis s'achève par la description de la *sélection clonale* qui part d'une analogie avec le système immunitaire des vertébrés. L'algorithme a ceci de remarquable que ses résultats sont très répétables et qu'il permet d'obtenir plusieurs solutions de qualité à

la fois. La dernière section aborde quelques aspects pratiques de la résolution d'un problème inverse par optimisation globale.

Pour chaque algorithme, le lecteur trouvera une présentation des heuristiques et de l'algorithme canonique qui en est issu, le cas échéant des éléments très brefs sur les preuves de convergence, et des références pour approfondir le cas d'un algorithme. L'application de chaque algorithme est illustrée dans le corps de ce chapitre, et éventuellement dans d'autres chapitres de cet ouvrage qui traitent de la résolution de problèmes concrets du génie civil. Ce chapitre sur les algorithmes est associé à un CDROM qui contient une implantation de chaque algorithme en langage Scilab. Une annexe à ce chapitre présente les résultats d'un banc d'essai des algorithmes sur une dizaine de problèmes d'optimisation académiques classiques.

Seul le cas de l'optimisation d'une fonction à valeurs réelles est traité ici. L'accent est mis sur l'optimisation de variables continues. Cependant, un algorithme adapté au cas discret est présenté. Il s'agit de l'algorithme génétique. Les algorithmes décrits n'imposent aucune contrainte de régularité à la fonction coût, qu'il s'agisse de dérivabilité, d'appartenance à un espace de fonctions particulier. Il n'est d'ailleurs pas nécessaire que celle-ci soit calculable, ni même définie explicitement.

**Terminologie** Dans la suite, l'expression « espace de solutions » est employée pour désigner l'espace dans lequel s'effectue la recherche de l'optimum. Un point de cet espace est appelé « solution candidate ». Enfin, on se place dans le cas de la minimisation d'une fonction. La fonction dont on cherche le minimum sur l'espace de solutions est appelée « fonction coût ».

## 9.2 Principes généraux

Les algorithmes classiques d'optimisation interprètent la topologie locale de la fonction à traiter, par exemple via le calcul du gradient, du hessien. Face à une fonction multi-modale quelconque, ces informations sont souvent un leurre qui mène tout droit dans un minimum local. Il faut donc s'en passer, du moins pour l'exploration à grande échelle de l'espace de solutions. En l'absence de l'information topologie, il faut recourir à d'autres règles pour progresser dans l'exploration. Les sections qui suivent se proposent d'en faire un bref tour d'horizon.

### 9.2.1 Le hasard fait bien les choses

Pour tous les algorithmes présentés ici, l'exploration de l'espace des paramètres à optimiser s'appuie massivement sur des variables aléatoires. Celles-ci servent d'une part à faire évoluer directement la ou les solution(s) candidate(s). C'est même l'unique mécanisme de mise à jour dans le recuit simulé, comme on le verra plus tard, étant donné que cet algorithme ne manipule qu'une seule solution candidate à la fois. Une fluctuation aléatoire d'amplitude suffisante permet à l'exploration de sortir d'un minimum local. L'arrière-plan biologique souvent donné au sujet conduit à parler de « mutations » pour désigner les variations d'origine aléatoire.

Les tirages aléatoires servent d'autre part à formuler des choix, par exemple à décider si on doit introduire une variation aléatoire dans une des solutions candidates

ou effectuer un croisement entre deux solutions candidates, comme dans les algorithmes génétiques.

L'aléatoire sert enfin à faire évoluer les paramètres internes des algorithmes, comme dans le cas de la stratégie d'évolution. Dans cet algorithme, une matrice de covariance est associée à chaque solution candidate. Le schéma itératif qui modifie cette matrice dépend en partie de termes aléatoires.

Une conclusion de ceci est que des optimisations successives de la même fonction coût, toutes choses égales par ailleurs, ne mèneront pas nécessairement au même résultat, à moins que le générateur de nombres aléatoires que l'on utilise ne soit réinitialisé à chaque nouvelle minimisation.

Il est clair que les variations apportées par les tirages aléatoires doivent être maîtrisées de manière à assurer tout de même la convergence vers un point de l'espace des solutions. Ceci mène directement à la question du réglage des paramètres internes des algorithmes.

### 9.2.2 Le travail collectif est préférable

Comme mentionné ci-dessus, les algorithmes présentés manipulent généralement un ensemble de solutions candidates. On parle donc d'une « population » formée d'« individus » pour désigner cet ensemble, compte tenu de l'origine biologique des métaphores étudiées. L'interaction entre solutions candidates constitue le deuxième grand mécanisme de déplacement dans l'espace des solutions. Cette interaction peut prendre la forme d'une combinaison linéaire ou non. Ainsi, dans le cas des algorithmes génétiques ou des stratégies d'évolution, deux individus parents laissent la place à un individu fils par l'opérateur de « recombinaison » dont deux formes seront présentées plus loin. La recombinaison est appelée "opérateur d'exploitation" par certains auteurs, car elle permet d'accéder à une autre zone de l'espace de solutions, sans intervention externe. Cet opérateur est complémentaire de la mutation en ce qu'il tend à faire converger les individus les uns vers les autres.

Autre possibilité, selon une approche sociale à taille de population fixe, un individu performant va pouvoir influencer les mouvements de ses congénères :

$$\vec{v}' = \vec{v} + \dots + \phi(\vec{p} - \vec{x}) \quad (9.2.1)$$

où  $(\vec{x}, \vec{v})$  décrit la position et la vitesse de la solution candidate dans l'espace de solutions,  $\vec{p}$  la position de la meilleure solution candidate, et  $\phi$  une variable aléatoire uniforme. C'est ce mécanisme qui est à l'oeuvre dans l'essaim de particules et de manière plus raffinée dans l'algorithme de sociétés et civilisations.

### 9.2.3 Aller du général au particulier

On vient de noter la place centrale qu'occupent les nombres aléatoires dans les algorithmes d'optimisation étudiés ici, ainsi que la nécessité de garder une cohérence dans l'exploration, de manière à éviter l'écueil du mouvement brownien. L'idée sous-jacente à au moins deux algorithmes présentés dans la suite (recuit simulé et stratégie d'évolution) est que la recherche de l'optimum doit se faire à une échelle décroissante au fil des itérations, de manière globale et grossière au début et de manière locale et fine à la fin. La stratégie d'évolution y parvient en réduisant l'écart-type des fluctuations permises pour les solutions candidates durant la phase de mutation. Dans le

recuit simulé, c'est la probabilité d'acceptation d'une variation défavorable, *i.e.* qui fait augmenter la valeur de la fonction coût, qui diminue dans le temps. C'est certainement ce dernier algorithme qui donne sa forme la plus éloquente au mécanisme évoqué ici. On parle de « consigne de refroidissement ».

### 9.2.4 Se souvenir du chemin parcouru

Dernière heuristique, la recherche de l'optimum est d'autant plus efficace qu'elle ne passe qu'une fois par un endroit donné de l'espace de solutions et évite l'apparition de cycles. La recherche taboue, qui n'est pas présentée ici, faute de place, est un algorithme doté d'une mémoire de court terme qui permet par comparaison d'éviter à l'exploration de revenir au même endroit, en tout cas avant un nombre minimal d'itérations. Des raisons pratiques imposent que cette mémoire soit de court terme. Dans le cas contraire, décider si une zone de l'espace a déjà été visitée deviendrait rapidement dispendieux. L'implantation de cette heuristique en variables continues n'est pas immédiate car elle nécessite la définition d'un voisinage. Le système immunitaire des vertébrés propose un autre mécanisme de mémoire qui permet de réagir plus vite et plus intensément à un élément du non-soi déjà rencontré dans le passé d'un organisme. Cette mémoire n'est pas non plus éternelle. Ce mécanisme est détaillé à la section sur l'algorithme de sélection clonale.

## 9.3 Une métaphore physique : le recuit simulé

### 9.3.1 Principe

Le recuit, par opposition à la trempe, est un traitement métallurgique qui consiste à laisser refroidir lentement une pièce métallique après l'avoir chauffée. A haute température, les atomes qui constituent la pièce présentent une agitation très importante. Leur position dans la matière peut donc changer sensiblement à ce stade, et ce de façon aléatoire. A mesure que la température diminue, la mobilité des atomes décroît, et les réarrangements se produisent de plus en plus localement, jusqu'à ce que plus aucun mouvement ne soit possible. La lenteur du refroidissement permet de s'approcher de la cristallisation parfaite. Celle-ci est caractérisée par une énergie minimale et une résistance mécanique élevée. On observe le même phénomène en géologie, avec la formation des roches issues d'un magma. A composition égale, avec un refroidissement rapide, on obtient des roches peu cristallisées voire amorphes, comme les basaltes ou les obsidiennes. Quand le refroidissement est lent, on obtient des granites.

Du recuit a été tiré le concept de *consigne de refroidissement* évoqué plus haut. Pour revenir au problème d'optimisation, on peut faire le parallèle entre les paramètres à optimiser et les atomes chauffés. On définit une température pour l'optimisation qui est une fonction décroissante du nombre d'itérations. La minimisation de la fonction coût s'obtient en modifiant la solution courante par un tirage de variable aléatoire. Si la fluctuation introduite abaisse la valeur de la fonction, elle est conservée. Si elle l'augmente, la variation n'est conservée que relativement à une probabilité d'acceptation, laquelle diminue avec la température de l'optimisation. Comme la température décroît, les variations défavorables sont de moins en moins souvent acceptées. C'est Métropolis que le premier a eu l'idée de simuler le recuit sur le principe qui vient d'être

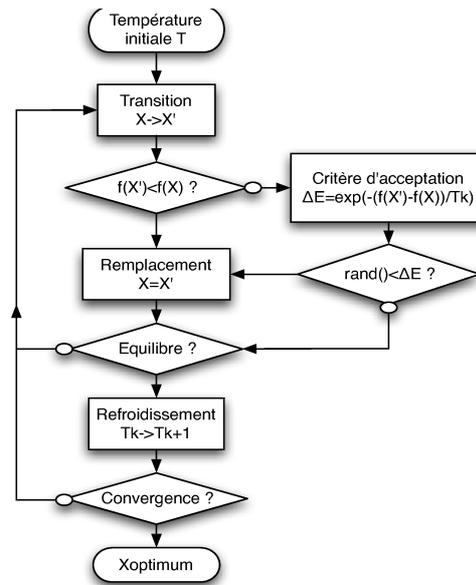


FIG. 9.1 – Organigramme du recuit simulé

décrit [MRR<sup>+</sup>53]. L'analogie et l'application à l'optimisation numérique est due à Kirkpatrick [KGV83].

### 9.3.2 L'algorithme de base

L'algorithme du recuit simulé est présenté sur la figure 9.1.

Il faut préciser la nature du schéma de refroidissement soit la relation  $T_k = f(T_0, k)$  où les indices font référence respectivement à l'itération courante et l'état initial. Depuis l'apparition de l'algorithme, un très grand nombre de schémas de ce type ont été proposés, voir par exemple [[vLA87], p 72]. On peut les regrouper en schémas fixes et adaptatifs. Les schémas fixes sont les plus couramment utilisés car ils sont aussi les plus immédiats :

- La loi géométrique qui est la plus utilisée car la plus simple :  $T_k = T_0 \alpha^k$  avec  $0 < \alpha < 1$ .
- La loi logarithmique :  $T_k = \frac{T_0}{1 + \log(k+1)}$
- La loi inverse :  $T_k = \frac{T_0}{1+k}$
- La loi puissance :  $T_k = T_0 \left(\frac{T_f}{T_0}\right)^{\frac{k}{k_{max}}}$  où  $T_f$  est la température finale.

Parmi les schémas adaptatifs, on peut citer :

$$T_{k+1} = \left(1 - T_k \frac{\Delta(T_k)}{\sigma^2(T_k)}\right) T_k \quad (9.3.1)$$

où  $\sigma$  désigne l'écart-type des valeurs de fonction coût.  $\Delta(T_k)$  est à définir. On peut éventuellement choisir une fonction constante.

### 9.3.3 Convergence

Le recuit simulé est un schéma itératif qui transforme la solution courante en une solution voisine. Un tel comportement correspond à celui étudié par la théorie des

chaînes de Markov. Au regard de celle-ci, on peut obtenir un algorithme homogène ou inhomogène, selon la consigne de refroidissement retenue. Dans le premier cas, l'algorithme est décrit par une séquence de chaînes de Markov homogènes, une par valeur de température. Dans le second, il est décrit par une seule chaîne de Markov inhomogène. Pour plus de détails sur l'application des chaînes de Markov au recuit simulé, le lecteur peut se référer à [[DPST03], Annexe A].

On montre que, sous certaines conditions, l'algorithme converge asymptotiquement vers la région optimale avec une probabilité égale à 1. Les conditions changent selon que l'algorithme est homogène ou inhomogène. Dans un contexte d'ingénierie, ce résultat de convergence asymptotique présente peu d'intérêt. En outre, la preuve de convergence suppose un schéma de refroidissement trop lent pour être utilisé sur un problème concret.

### 9.3.4 Utilisation

Du fait qu'il ne manipule qu'une seule solution candidate, le recuit simulé est d'implémentation plus simple que les autres algorithmes présentés dans ce chapitre. Cette simplicité se paie par des performances moins bonnes en général. Aussi, le recuit simulé est plutôt destiné à l'optimisation sur un petit nombre de variables, lorsque la fonction objectif n'est pas d'une évaluation trop coûteuse.

Comme on vient de le voir, la théorie du recuit simulé n'est d'aucun secours quand il s'agit d'utiliser vraiment l'algorithme, et le nombre de paramètres à définir peut rebuter. Il apparaît cependant que le schéma géométrique de refroidissement donne de bons résultats, avec une raison supérieure à 0.9. Initiative intéressante, Siarry présente des recommandations concernant le réglage du recuit simulé sur une fonction coût quelconque [[DPST03], 1.7].

Le code Scilab mis à disposition avec ce chapitre est une implantation de l'algorithme décrit dans [CMMR87]. En ce qui concerne l'utilisation de ce dernier, les fonctions nécessaires résident dans le fichier `sa.sci` qui contient les fonctions utilisateur `sa()` et `sa_param()`. Avant tout appel à `sa()`, il faut avoir défini `costfunc()` qui contient la fonction à *minimiser*, et paramétrer l'algorithme avec `sa_param()`. Cette fonction prend comme seul argument un vecteur colonne contenant la solution candidate.

L'aide en ligne des fonctions de `sa.sci` peut être affichée en entrant `help sa_` dans Scilab. Pour la commodité de la lecture, les arguments d'entrée de la fonction principale sont donnés ci-dessous :

```
[best,cost,nev,evolution,best_it]=sa(param,nvar,span,init,tol,evmax);
//entrée
//param : tlist, voir help sa_param
//nvar : entier, nombre de variables à optimiser
//span : matrice (nvar,2), domaine de variation de chaque paramètre
//init : vecteur (nvar,1) ou [], point de départ
// si [] le point de départ est basé sur une v.a. uniforme définie sur span
//tol : scalaire, seuil de convergence sur la valeur de la fonction coût
//evmax : entier, nombre maximal d'évaluations de costfunc()
//sortie
//best : vecteur (nvar,1), meilleur résultat obtenu pour x
//cost : scalaire, coût le plus bas obtenu
//nev : entier, nombre effectif d'évaluations de costfunc()
//evolution : matrice (n,nvar+1), contient l'historique de la minimisation.
//Pour chaque ligne les nvar dernières colonnes contiennent le meilleur individu
//x de l'itération et la première colonne costfunc(x)
//best_it : entier, itération pour laquelle le meilleur résultat a été atteint
```

Un exemple d'utilisation de la bibliothèque `sa.sci` est donné pour la fonction `f5` de De Jong qui présente 25 minima locaux [Jon75].

```
function r=f5(x,y)
per=[-32,-16,0,16,32];
aj1=[per,per,per,per,per];
m=ones(1,5);
aj2=[-32*m,-16*m,0*m,16*m,32*m];
r=(500+sum((1:25)+sum((x;y)*ones(1,25)-[aj1;aj2]).^6,'r')).^(-1)).^(-1))
endfunction

x=-60:60;y=x;
Sfgrayplot(x,y,f5);
xtitle('f5 : minimum global 0.001996 en (x,y)=(-32,-32)', 'x', 'y')

function r=costfunc(x)
r=f5(x(1),x(2));
endfunction

nvar=2;
tl=sa_param(nvar);//paramétrage par défaut
tl.cs.To=1e9;//temperature initiale
itmax=1e6;//nombre maximum d'itérations
tl.Neps=20;//sortie si stationnarité sur 20 itérations
[best,cost]=sa(tl,nvar,ones(nvar,1)*[-200,200],[101;101],1e-8,itmax)
```

Dans la suite, la physique laisse la place à la biologie. Quatre métaphores inspirées du vivant à différentes échelles  $y$  sont présentées. A partir de ce point, tous les algorithmes abordés manipulent plusieurs solutions candidates à la fois.

## 9.4 Une métaphore sociale : l'essaim de particules

### 9.4.1 Heuristique de base

James Kennedy, l'un des inventeurs de l'essaim de particules est chercheur en psychologie sociale. Sa recherche d'un modèle numérique d'interaction sociale l'a conduit à étudier comportements grégaires dans le règne animal.

Chez de nombreuses espèces animales, il est rare de rencontrer un individu seul dans son milieu. Il sera plus probablement intégré à un groupe. On parle alors de banc pour les poissons ou d'essaim pour les insectes et de nuées chez les oiseaux. Au-delà du caractère spectaculaire des mouvements de tels ensembles, avec la remarquable coordination qu'elle suppose lors des déplacements, il faut insister sur les avantages évolutifs qu'apporte à l'espèce considérée le mode de vie en essaim. L'essaim minimise la probabilité de rencontre avec un prédateur. Pour modéliser finement la trajectoire d'un essaim, une hypothèse centrale est d'ailleurs que chaque membre de l'essaim vise à se rapprocher du centre de celui-ci. De la même manière, dans un contexte de ressources alimentaires concentrées dans certaines zones de l'espace exploré, le déplacement en essaim permet une recherche plus efficace dont le fruit profite au plus grand nombre [Wil00]. On sait aujourd'hui que chez nombres d'espèces, les individus fondent en partie leur comportement sur l'information diffusée de manière volontaire ou non (on parle alors de communication par inadvertance [Dan05]) par leurs congénères. Enfin, la recherche a mis en évidence le rôle du groupe dans la réussite d'une navigation à grande distance, comme celle de la migration [Sim04].

Même en se limitant à la recherche de nourriture, il est clair que l'exploration de l'environnement par un essaim peut déjà être vue comme un processus d'optimisation. L'algorithme d'essaim de particules a été défini à partir des travaux de Reynolds

ainsi que Heppner et Grenander sur la simulation des déplacements des oiseaux [Rey87][HG90]. Une fois réalisée l'abstraction qui consiste à passer de l'espace physique à un espace de paramètres de dimension arbitraire, de la fonction "nourriture disponible" à une fonction mathématique quelconque, une méthode d'optimisation à part entière se dessine. Pour plus de détails sur la genèse de l'algorithme d'essaim de particules, consulter les travaux de Kennedy et Eberhart [KE95][KES01].

### 9.4.2 L'algorithme

L'algorithme d'essaim de particules travaille donc sur plusieurs particules. Chacune est dotée d'un vecteur position et d'un vecteur vitesse dans l'espace des paramètres. A chaque itération, pour chaque particule, ce couple de vecteurs est mis à jour en fonction de la trajectoire passée de la particule et de celle de la meilleure particule, au regard de la fonction coût à minimiser.

Il existe plusieurs versions de l'algorithme d'essaim de particules qui suscite un intérêt soutenu depuis plusieurs années. On présente ici la version en variables continues dite « avec constriction de type 1 » décrite dans [[CK02], p70]. Ce choix se justifie parce que cette version donne les meilleurs résultats sur un jeu de fonctions test réputées difficiles [CK02]. L'algorithme est donné ci-dessous :

```
Initialiser la population (pi et vi)
Faire
  Pour i dans 1,2,...taille de la population
    si f(xi)<f(pi) alors pi=xi
    pg=min(pvoisins)
    Pour d dans 1,2,...dimension
      vid=Khi(vid+phi1(pid-xid)+phi2*(pgd-xid))
      xid=xid+vid
    Fin pour
  Fin pour
Jusqu'à ce que le critère de fin soit satisfait
```

La signification des variables est la suivante :

- $pi$  est la position de l'individu  $i$  dans l'espace des paramètres ;
- $vid$  est la vitesse selon l'axe  $d$  de l'espace des paramètres pour l'individu  $i$ ,  $vi$  est le vecteur vitesse ;
- $pg$  est la meilleure position pour l'itération courante. Elle est déterminée sur un voisinage de l'individu  $i$ . Dans l'implantation Scilab, on a pris un voisinage étendu à toute la population ;
- $phi1$  et  $phi2$  sont des v. a. uniformes sur l'intervalle  $[0, \phi]$  ;
- la variable  $Khi$  est fixée à partir des paramètres internes  $\phi$  et  $\kappa$  selon l'étude analytique de la stabilité et de la convergence donnée dans [CK02].

Outre la possibilité de définir un voisinage différent [MKN04], il existe deux types de variantes de l'algorithme d'essaim de particules. Le premier concerne la valeur de  $Khi$ . Le second introduit un schéma légèrement différent de mise à jour de  $vid$  sous la forme d'un écrêtage au-delà de  $vmax$ . Le but est d'éviter la tendance de l'algorithme à la divergence.

#### A. Variantes

A partir de cet algorithme, Angeline en a bâti un autre qui combine sélection naturelle par processus de croisement, mutation et sélection comme dans les algorithmes évolutionnaires présentés plus loin, et interaction sociale. Sur des fonctions test les

performances comparées à l'essaim de particules sont variables pour les dimensions faibles, sans amélioration significative en grande dimension. On peut noter que les paramètres internes de l'essaim de particules sont statiques dans la version présentée ici, certains auteurs ont étudié des versions à paramètres dynamiques. Pour l'étude des variations autour de l'essaim de particules, le lecteur se reportera à [KES01].

En reprenant l'idée de l'effet positif des interactions à l'intérieur d'un groupe Ray et Liew sont allés plus loin en introduisant un niveau hiérarchique intermédiaire entre les individus et leur *leader*. L'algorithme obtenu s'appelle *sociétés et civilisation* [RL03]. Les deux principes qui le fondent sont :

- Une civilisation émerge et progresse grâce aux relations de coopération entre les sociétés qui la composent ;
- Les individus d'une société interagissent mutuellement dans le but de progresser.

Le niveau hiérarchique intermédiaire est obtenu en opérant des regroupements entre individus appelés « sociétés ». Chaque société est dotée d'un ou plusieurs *leaders* désignés par leur performance supérieure. La confrontation des différents *leaders* permet de définir un ou plusieurs *leaders* de *civilisation*. Dans une société, l'évolution des individus est orientée par les *leaders* de la société. Les *leaders* de civilisation influencent les *leaders* de société.

L'algorithme montre les meilleures performances connues, surtout en termes de vitesse et de respect des contraintes, sur trois problèmes de résistance des matériaux dont la difficulté n'a pas pu être évaluée. Les caractéristiques de ces problèmes sont de porter sur un faible nombre de variables, mais d'inclure des contraintes.

## B. Performance et convergence

Malgré son extrême simplicité, l'algorithme d'essaim de particules se montre capable de trouver le minimum de fonctions à topologie complexe [CK02].

En parallèle aux investigations empiriques, Clerc a étudié l'algorithme d'essaim de particules sur un plan théorique [CK02]. Il montre qu'un modèle généralisé sous la forme d'un système différentiel à 5 paramètres peut être associé à l'algorithme dans le cas d'un seul degré de liberté. L'étude du rôle de ces paramètres permet de dégager des classes de comportement et en particulier le schéma dit de « constriction modérée » qui assure la convergence tout en permettant une bonne exploitation de l'espace de solutions.

Les résultats sur une particule à 1 degré de liberté se généralisent à  $n$  degrés de liberté. En revanche, la théorie de l'interaction entre les particules reste à construire.

## C. Pratique

**Réglage des paramètres** Par défaut,  $\kappa = 0.8$ , d'après l'étude de convergence [CK02], et  $\phi = 4.5$  donnent de bons résultats. Le nombre de particules utilisé couramment varie entre 10 et 50, soit ordinairement moins que dans les algorithmes génétiques. Sur les fonctions test de De Jong, 20 particules suffisent à obtenir de bons résultats.

**Utilisation du code Scilab** La bibliothèque Scilab qui implante la variante d'essaim de particules retenue s'appelle `pso.sci`, dans laquelle `pso()` est la fonction appelée par l'utilisateur. Les arguments d'entrée et de sortie de cette fonction sont les suivants :

```

function [best,cost,nbiter,dur, evolution,pop_fin]=pso(npart,nvar,span,init,tol,evmax,kappa,phimax);

//entree
//npart : scalaire, nombre de particules constituant l'essai, 20 par défaut
//nvar : scalaire, la dimension de l'espace étudié
//span : matrice(nvariables,2), minimum et maximum de chaque variable
// ou matrice(1,2) : minimum et maximum identique pour toutes les variables
//init : vecteur (nvar,1) optionnel, valeurs initiales
// si [], initialisation par v.a. uniforme
//tol : //seuil de convergence sur la valeur de costfunc
//evmax : entier, nombre maximal d'appels de fonctions
//kappa : scalaire optionnel, sert au calcul du coefficient de constriction, 0.8 par défaut
//phimax : scalaire optionnel, sert au calcul du coefficient de constriction, 4.5 par défaut.

//sortie
//best : vecteur, meilleur individu de l'essai
//cost : scalaire, meilleure valeur
//nbiter : entier, nombre d'itérations
//dur : scalaire, durée de l'optimisation
//evolution : matrice (nit,nvar+1), historique de l'optimisation,
//evolution(i,1) contient le minimum à l'itération i,
//evolution(i,2 : $) le meilleur individu à l'iteration i
//pop_fin : tlist, essai final

```

Rappel : avant tout appel à `pso()`, il convient d'avoir défini `costfunc()` qui correspond à la fonction à minimiser. Un exemple élémentaire d'application à la fonction de Rosenbrock est donné ci-dessous :

```

stacksize(3e7);

function z=Rosenbrock(x,y)
v=[x;y];
dim=size(v,1);
z=sum(100*(([v(2:dim,:);v(1,:)]-(v.^2)).^2)+(v-1).^2),'r' );
endfunction

xset('window',0); xbas();
x=-2:0.1:2;y=x;
fplot3d(x,y,Rosenbrock,alpha=55,theta=65);
xlabel('Fonction de Rosenbrock')

function z=costfunc(z)
z=Rosenbrock(z,[]);
endfunction

//optimisation en dimension 30 avec 20 particules
//domaine initial [-20,50]^30
//40000 evaluations de fonction au maximum
[best,cost]=pso(20,30,[-20,50],[],1e-4,40000)

```

## 9.5 Une métaphore darwinienne : les algorithmes évolutionnaires

### 9.5.1 Aspects généraux

Par bien des aspects, la vie d'une espèce animale ou végétale peut être vue comme un problème d'optimisation à grande échelle. Une compétition apparaît entre des espèces dont les niches écologiques se recouvrent. L'habitat d'une espèce est sujet à des changements qui peuvent être des sources de stress pour ses membres. Lorsqu'une ressource vitale se fait rare, la compétition devient intra-spécifique. Ces grandes classes de facteurs opèrent une pression sélective sur le vivant. Seuls les individus, ou les espèces, les plus aptes peuvent survivre. On reconnaît là un des principes de la théorie darwinienne de l'évolution [Dar99][Fut05]. Pour faire face à cette pression sélective, le vivant se bat avec l'arme aveugle de l'hérédité. Le brassage de

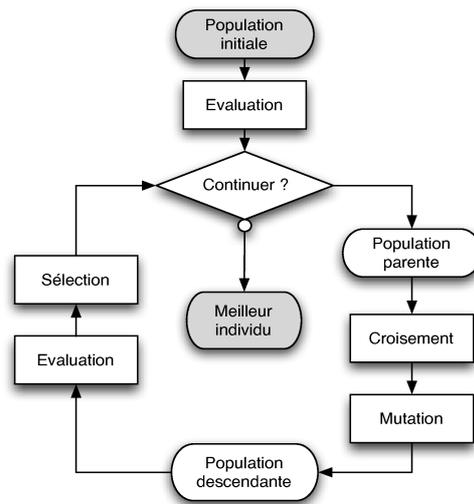


FIG. 9.2 – Organigramme d'un algorithme évolutionnaire

gènes qui intervient avec la reproduction sexuée, les mutations, dues aux erreurs de la machinerie enzymatique de manipulation du génotype - la banque de gènes que contient chaque cellule d'un individu d'une espèce donnée -, ou à des facteurs environnementaux, contribuent à produire des variations non dirigées du phénotype - l'ensemble de traits physiques qui caractérisent l'individu et ses performances dans l'environnement où il se trouve -. Non dirigées, parce qu'elle peuvent être positives ou négatives en termes de performance.

Dans les années 1960, soit environ un siècle après la théorie darwinienne, plusieurs équipes ont commencé à travailler indépendamment sur des algorithmes inspirés de l'évolution des populations d'êtres vivants. Si on considère une seule espèce, une analogie peut être établie entre l'ensemble des individus et l'ensemble des solutions, la performance dans l'environnement et la valeur de la fonction coût. Les opérateurs de base pour la production de nouvelles solutions candidates sont inspirés de ceux qu'implique l'évolution d'une population : croisement (*crossover* en anglais), mutation. Les algorithmes obtenus sont habituellement rangés dans la classe des « algorithmes évolutionnaires ». Un membre de cette classe est conforme à l'organigramme présenté sur la figure 9.2. Le membre le plus connu est celui des algorithmes génétiques (GA) [Hol75]. On y range aussi les stratégies d'évolution (ES) [Rec65] et les programmes d'évolution [Fog64]. Les derniers sont peu utilisés c'est pourquoi, on se limite au GA et aux ES dans la suite. Comme dit en introduction, on verra que, sur l'axe du compromis polyvalence/performance, le GA se situe du côté de la polyvalence, et l'ES du côté de l'efficacité. La version canonique de chaque classe est présentée séparément dans les deux sections qui suivent.

## 9.5.2 Variables discrètes : les algorithmes génétiques

### A. L'algorithme canonique

Les GA respectent la dualité phénotype/génotype. En effet, ils opèrent non pas sur les paramètres à optimiser, mais sur un codage de ceux-ci, d'où leur nom. L'obtention

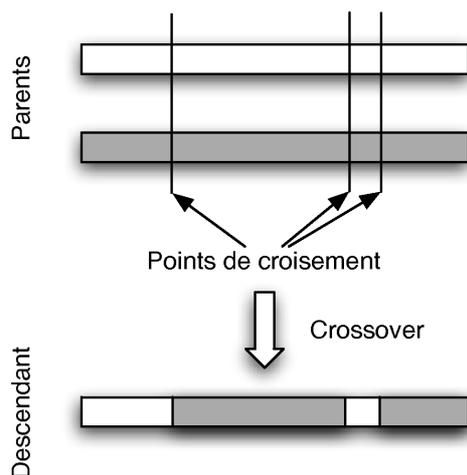


FIG. 9.3 – Fonctionnement de l'opérateur de croisement dans un GA à trois points.

du code génétique d'une solution candidate consiste, dans un premier temps, à remplacer chaque valeur de paramètre par sa représentation sous la forme d'une chaîne binaire. Ici, il n'est pas nécessaire que les paramètres de départ soient des nombres. Les différentes chaînes de bits sont ensuite mises bout à bout pour former ce que l'on peut appeler un chromosome. Cette chaîne de bits globale est regroupée avec les autres chaînes obtenues pour chaque individu considéré, et l'on obtient ainsi une population.

A l'itération  $i$ , donnons-nous  $p$  individus. En choisissant au hasard deux individus  $q$  fois, on obtient  $q$  individus ( $q > p$ ) en appliquant l'opérateur de recombinaison ou de croisement qui fabrique une nouvelle chaîne de bits à partir de deux chaînes sources, de telle sorte qu'un bit positionné en  $k$  dans la chaîne résultante soit le bit  $k$  de l'une des deux chaînes de départ. Dans le GA canonique, le croisement se fait en un seul point. La figure 9.3 montre le fonctionnement d'un opérateur de croisement à 3 points. Le descendant subit ensuite l'opérateur de mutation, qui introduit une inversion de bit selon une probabilité donnée en un point de la chaîne. Ce point est sélectionné par tirage d'une variable aléatoire uniforme. La probabilité de mutation doit rester faible, sous peine de transformer l'algorithme en marche aléatoire.

Les différents  $q$  descendants ainsi obtenus sont décodés, puis la valeur de la fonction coût est calculée pour chacun. On sélectionne  $p$  individus parmi ces descendants. La sélection est probabiliste, au sens où elle donne la préférence aux individus performants, mais peut accorder la vie sauve à de « mauvais » individus. On effectue un test de convergence. Et ainsi de suite jusqu'à ce que le critère de convergence ou de sortie soit atteint (Cf figure 9.2).

## B. Du GA canonique au GA performant

En pratique, il s'avère que l'algorithme canonique, même optimisé quant à ses paramètres internes, est parfois bien peu efficace sur les problèmes réels d'optimisation. Un certain nombre d'ajustements sont possibles.

**Des variations contrôlées** D'une manière générale, si l'on suppose que l'analogie avec la nature conserve son sens, il est établi que les modifications du code génétique d'un individu sont soit faibles, soit létales [Kim92]. Reprenons le fonctionnement des opérateurs avec cet élément en tête.

**Croisement** Le croisement ignore que la chaîne qu'il découpe est le résultat de la mise bout à bout de sous-chaînes codant chacune une variable. Ses recombinaisons, où le point de transition dans la chaîne est quelconque, peuvent donc avoir des effets désastreux. La tendance est donc à la conception de *crossovers* qui basculent d'un individu à l'autre en un nombre fini de points possibles, à savoir les limites d'une sous-chaîne à une autre, c'est-à-dire d'une variable à une autre.

**Mutation** L'opérateur de mutation tel qu'il fonctionne ici peut introduire des changements énormes. Il suffit qu'un bit de poids fort soit inversé, et la valeur de la variable, une fois le décodage effectué changera du tout au tout, ce qui va à l'encontre de la règle des mutations faibles. La parade consiste à remplacer le codage binaire simple par un codage de Gray [[Mic99], p98]. Ce codage permet en effet de réduire l'amplitude de la variation introduite par un changement de bits.

**Un codage respectueux du problème** Le codage de type entier utilisé par les GA se prête mal à la représentation de variables ayant un domaine de variation très large. La seule manière de maintenir la finesse de la représentation des nombres est en effet l'augmentation du nombre de bits affectés au codage de chaque variable, ce qui ne fait qu'accroître le coût de calcul. C'est pourquoi de plus en plus de GA modifiés apparaissent, dont les opérateurs s'appliquent à des nombres codés en virgule flottante [[Mic99], p100]. Mais leur simplicité se perd avec ce changement.

**Dynamisme et apprentissage** Les paramètres d'un GA, dans sa forme canonique, sont statiques. Il semble qu'une efficacité supérieure pourrait être atteinte en rendant l'algorithme adaptatif, même s'il faut se garder d'affirmer l'intérêt systématique du dynamisme des paramètres [[Mic99], 4.4], et au-delà, de donner une mémoire à l'algorithme [SS96]. On voit ici réapparaître des idées mises en pratique dans le recuit simulé [MCPB00] ou la recherche taboue.

**Exploiter la variété** L'enchaînement croisement, mutation, sélection définit en fait un cadre extrêmement souple. La recherche sur les GA a produit de très nombreuses illustrations de cette souplesse, en ce qui concerne l'implémentation des opérateurs, la manière de sélectionner les individus, et leur durée de vie [[Bäc96], 2.3]

**Des opérateurs spécifiques** Enfin, il est reconnu que des opérateurs spécifiques du problème traité, en complément de ceux de mutation et de croisement, sont critiques pour la vitesse de convergence d'un GA.

### C. Comportement et convergence

De même que pour le paramétrage, les résultats de convergence sont rares, voire inexistants, malgré l'abondance des travaux sur les GA. Comme on l'a observé pour le recuit simulé, si la convergence globale au sens des probabilités est établie théoriquement, c'est un résultat inutile en pratique [Bäc96]. Comme dit précédemment, il serait en fait nécessaire de disposer de résultats de vitesse de convergence. Ce manque de résultats semble dû à l'absence de théorie à laquelle rattacher le fonctionnement des GA. La théorie *ad hoc* dite « des schémas » s'avère peu opératoire.

Des études empiriques suggèrent que l'opérateur de croisement est inutile sur certaines classes de fonctions. Un travail récent indique une relation quadratique entre le taux de mutation et la performance [CMV<sup>+</sup>04].

De même, les caractéristiques des problèmes auxquels les GA sont bien adaptés sont loin d'être éclaircies.

### D. Pratique

Encore plus que le recuit simulé, le GA peut rebuter du fait du grand nombre de paramètres internes à définir et à régler. L'algorithme génétique est en fait un schéma très général qui laisse une très grande liberté. La théorie déficiente n'aide guère l'utilisateur à faire des choix. La contrepartie est l'extraordinaire adaptabilité des GA.

Les GA ne sont pas conçus à l'origine pour l'optimisation de variables continues, du fait du codage binaire qu'ils impliquent. Sur ce terrain, les ES peuvent leur être préférées en général. Toutefois, les GA sont bien adaptés à des problèmes qui mélangent variables discrètes et variables continues. D'autre part, la discrétisation a ceci d'intéressant qu'elle permet de définir des bornes infranchissables pour l'espace de solutions.

Si la littérature insiste sur la nécessité d'un jeu soigneusement choisi de paramètres, il n'existe pas à ce jour de méthode générale d'analyse de la relation paramètres performance d'un GA [CMV<sup>+</sup>04]. La méthode la plus courante consiste à procéder à des essais répétés avec différents jeux de paramètres. La contribution de Grefenstette est sans doute symptomatique, puisque cet auteur fixe les paramètres d'un GA par...optimisation à l'aide d'un autre GA dont les paramètres sont d'efficacité empiriquement établie [Gre86]. Le procédé ne manque pas d'efficacité, sous réserve que le problème d'optimisation initial ne soit pas trop lourd. Cela revient toutefois à repousser le problème un niveau plus haut, puisque l'arbitraire du choix de paramètres demeure. De Jong a proposé une taille de population comprise entre 50 et 100, un taux de croisement de 0.6 et un taux de mutation de 0.001 comme réglage de base d'un GA. Avec ce réglage, les performances du GA canonique utilisé dans le cadre du banc d'essai qui est présenté en annexe sont mauvaises. Elles sont bien meilleures avec un taux de mutation de 0.01. D'autres auteurs sont arrivés à des valeurs différentes comme le montre [CMV<sup>+</sup>04]

Une implantation Scilab d'un GA canonique est contenue dans la bibliothèque `ga.sci`. Cette bibliothèque contient deux fonctions utilisateur : `ga_parentPopulation()` qui sert à constituer une population initiale, et `ga()` qui effectue la boucle principale du GA. Les prototypes de ces deux fonctions sont donnés ci-dessous :

```

fonction ppop=ga_parentPopulation(nvar,nb_ind,span,nbits)
//nvar : entier, nombre de variables
//nb_ind : entier, nombre d'individus
//span : matrice à 1 ligne 2 colonnes ou nvar lignes 2 colonnes,

```

```
//intervalle(s) de définition des variables.
//Dans le premier cas, même intervalle pour toutes les variables
//nbits : entier ou vecteur d'entiers, nombre de bits sur lesquels
//représenter nvar.

function
[best,cost,nev,dur,evolution,best_it,pop]=ga(np,nd,xr,mr,nvar,span,init,tol,evmax)
//entrées
//np,nd : entiers, nombre de parents et de descendants
//xr, mr : réels dans [0,1], taux de croisement (resp. mu\`-ta\`-tion)
//nvar : entier, nombre de variable dans
//span : intervalle de définition des variables
//init : population initiale
//tol : réel, seuil de convergence
//evmax : entier, nombre maximum d'évaluations de costfunc()
//sorties
//best : vecteur (nvar,1) meilleur individu
//cost : scalaire, coût du meilleur individu
//nev : entier, nombre effectif d'évaluations
//dur : scalaire, durée de calcul, s
//evolution : matrice (nit,nvar+1), historique de l'optimisation,
//evolution(i,1) contient le minimum à l'itération i,
//evolution(i,2 : $) le meilleur individu à l'iteration i
//best_it : entier, itération à laquelle le meilleur individu a été obtenu
//pop : tlist, population finale
```

Le listing Scilab suivant donne un exemple élémentaire d'utilisation de la fonction `ga()` sur la fonction `f2` de De Jong :

```
function z=DeJongf2(x,y)
v=[x;y]
z=100*(v(1,:).^2-v(2,:).^2).^2+(1-v(1,:)).^2;
endfunction

function z=costfunc(z)
z=DeJongf2(z,[])
endfunction

xset('window',0); xbascc();
x=-2:0.1:2;y=x;
fplot3d(x,y,DeJongf2,alpha=55,theta=65);
xtitle('Fonction f2 de De Jong')

nvar=2; //nombre de variables
nb_ind=50; //nombres d'individus parents
//population parente initiale
ppop=ga_parentPopulation(nvar,nb_ind,[-50,50],16);
//optimisation sur 20000 évaluations de fonction avec
//une probabilité de croisement de 0.5
//une probabilité de mutation de 0.01
[best,cost]=ga(nb_ind,2*nb_ind,0.5,0.01,nvar,[],ppop,1e-6,20000)
```

Le lecteur trouvera une application d'un GA en génie civil dans le chapitre rédigé par Denis Duhamel sur l'optimisation de formes de murs anti-bruit.

### 9.5.3 Variables continues : les stratégies d'évolution

#### A. Algorithme

**Représentation gaussienne** Conçues à des fins d'optimisation numérique [Rec94], les ES vont moins loin que les GA dans l'analogie avec le vivant, en n'opérant que sur le phénotype. Il n'est donc pas nécessaire de coder les variables à optimiser pour les manipuler dans l'algorithme. Les ES représentent les individus comme des vecteurs aléatoires gaussiens, notés traditionnellement  $\vec{N}(\vec{x}, \mathbf{C})$ . La moyenne est le vecteur  $\vec{x}$  contenant les variables du problème, et la matrice de covariance  $\mathbf{C}$  est séparée en

deux vecteurs.  $\vec{\sigma}$  correspond aux variances,  $\vec{\alpha}$  aux covariances au sens strict, ou rotations. Quoiqu'il soit possible d'introduire des dépendances entre variables, ce choix est rarement fait.  $C$  est donc en général diagonale.

**Initialisation** La population initiale de vecteur moyenne  $\vec{x}$  peut être issue d'informations *a priori* sur la position du minimum, ou d'un tirage d'un vecteur aléatoire uniforme convenablement centré et mis à l'échelle. Les termes de la matrice de covariance sont fixés à des valeurs abstraites non nulles.

**Opérateurs** Les opérateurs s'appliquent à toutes les composantes de l'individu, moyennes, comme les variances et covariances strictes. Nous supposons dans la suite que le problème compte  $n$  variables.  $N(a, b)$  désigne ici une variable aléatoire gaussienne de moyenne  $a$  et d'écart type  $b$ .

**Recombinaison** L'opérateur de recombinaison s'appuie sur le phénotype de deux individus  $S$  et  $T$ , de la même manière que l'opérateur de croisement sur le génotype pour les GA :

$$(\vec{x}', \vec{\sigma}', \vec{\alpha}') = \mathbf{R}((\vec{x}_S, \vec{\sigma}_S, \vec{\alpha}_S), (\vec{x}_T, \vec{\sigma}_T, \vec{\alpha}_T)) \quad (9.5.1)$$

Bäck détaille 6 formules non triviales de recombinaison des composantes des vecteurs initiaux [Bäc96]. Se référant à la littérature, il précise qu'il est préférable d'utiliser deux formules différentes selon que l'on considère les variables du problème ou les covariances. Pour les variables, il préconise la recombinaison discrète ; pour les covariances, la recombinaison « pan-mictique » :

$$\begin{aligned} x'_i &= x_{S,i} \parallel x_{T,i}, & i &\in \{1, \dots, n\} \\ \sigma'_i &= \sigma_{S,i} + \frac{1}{2}(\sigma_{T,i} - \sigma_{S,i}), & i &\in \{1, \dots, n\} \\ \alpha'_i &= \alpha_{S,i} + \frac{1}{2}(\alpha_{T,i} - \alpha_{S,i}), & i &\in \{1, \dots, n(n-1)/2\} \end{aligned} \quad (9.5.2)$$

**Mutation** Le descendant ainsi obtenu passe par l'opérateur de mutation. Celui-ci s'applique dans un premier temps aux variances et rotations, puis aux variables, comme suit :

$$\begin{aligned} \sigma'_i &= \sigma_i e^{\tau' N(0,1) + \tau N_i(0,1)} & i &\in \{1, n\} \\ \alpha'_i &= \alpha_i + \beta e^{N_i(0,1)} & i &\in \{1, n(n-1)/2\} \\ \vec{x}' &= \vec{x} + \vec{N}(\vec{0}, \vec{\sigma}', \vec{\alpha}') \end{aligned} \quad (9.5.3)$$

On reconnaît ici quelque chose qui s'apparente à la consigne de refroidissement du recuit simulé. Les paramètres internes de l'ES évoluent au fil des itérations.

Les facteurs  $\tau$ ,  $\tau'$  et  $\beta$  sont établis dans la littérature à :

$$\begin{aligned} \tau &\propto (\sqrt{2\sqrt{n}})^{-1} \\ \tau' &\propto (\sqrt{2n})^{-1} \\ \beta &\approx 0.0873 \end{aligned} \quad (9.5.4)$$

La valeur de la fonction coût est calculée pour chacun des  $\lambda$  individus de la descendance. Un classement est effectué, puis une sélection de  $\mu$  individus. La notation consacrée est  $(\mu, \lambda)$ -ES pour ce schéma démographique. L'autre forme possible est la  $(\mu + \lambda)$ -ES qui effectue la sélection sur la réunion des parents et des descendants.

Ce schéma est dit « élitiste ». A la différence des GA, la sélection se fait strictement selon le rang. Les mauvais individus n'ont aucune chance de survie.

L'école allemande qui est à l'origine de l'ES a aussi proposé plus récemment une variante majeure de l'ES : l'ES avec adaptation de la matrice de covariance ou CMAES [HA96]. La place manque pour la décrire en détail. Une implantation Scilab de cet algorithme est cependant fournie.

## B. Théorie et convergence

Même si on est loin de l'analyse convexe, le cadre probabiliste dans lequel les ES sont définies permet d'obtenir des résultats théoriques un peu plus forts que pour les GA. En particulier, une expression analytique de probabilité de vitesse de convergence a été établi dans le cas général d'une  $(\mu, \lambda)$ -ES. Il est toutefois difficile à exploiter en pratique pour des tailles de population parente supérieures à l'unité [[Bäc96],2.1.7]. Un certain nombre de résultats sont en outre disponibles sur des fonctions académiques, comme la sphère ou le corridor.

La règle d'ajustement de l'écart-type dans le schéma présenté ci-dessus est issue d'investigations théoriques.

Bäck fournit une comparaison empirique entre ES et GA sur cinq fonctions académiques de topologie variée. L'avantage est aux ES dans tous les cas, souvent de plusieurs ordres de grandeur [Bäc96].

## C. Pratique

La théorie ne dit rien sur la taille de population parente à retenir. Empiriquement, il apparaît que celle-ci peut rester sensiblement plus faible que dans le cas d'un GA, sans nuire à la performance. On peut prendre comme référence un effectif égal au nombre de variables à optimiser. Une population parente inférieure à 4 limite fortement le choix d'individus dans l'opérateur de recombinaison. Pour un nombre de variables élevé, on peut s'en tenir à un nombre d'individus nettement plus faible que le nombre de variables. La règle est de produire 7 fois plus de descendants qu'il y a de parents. L'élitisme n'est pas souhaitable. C'est donc une  $(\mu, \lambda)$ -ES qu'il convient d'utiliser.

La bibliothèque Scilab associée à l'ES est contenue dans le fichier `es.sci`. Outre la définition d'une fonction `costfunc()` respectant les règles définies précédemment, deux appels de fonctions sont nécessaires pour réaliser une optimisation par ES. La fonction `estol()` permet de préciser le critère de sortie de l'algorithme. La fonction `es()` contient la boucle principale de l'ES. Les prototypes de ces deux fonctions sont rappelés ci-dessous :

```

fonction tl=es_tol(cost,sdev,sdevx)
//cost : scalaire, seuil sur le coût : sortie si costfunc(x)<cost
//sdev : scalaire, sortie de l'optimisation si le plus grand écart-type de l'ES est inferieur à sdev
//sdevx : scalaire, sortie de l'optimisation l'écart-type intra-population est en dessous de sdevx
//tl : tlist, résultat

fonction [best,cost,nev,dur,evolution,best_it,pop]=es(mu,la,nvar,span,init,tol,evmax);
//entrées
//mu : entier, nombre de parents
//la : entier, nombre de descendants
//nvar : entier, nombre de variables à optimiser
//span : vecteur (1,2) ou matrice (nvar,2), intervalle de définition des paramètres
//init : matrice, population initiale
//tol : tlist produite avec es_tol()
//evmax : entier, nombre maximum d'évaluations de costfunc()

```

```
//sorties : idem ga()
```

Le listing Scilab suivant donne un exemple élémentaire d'utilisation des fonctions `es()` et `cmaes()` sur la fonction de Griewangk :

```
stacksize(3e7);
//convexité à grande échelle et multimodalité à petite échelle
function z=Griewangk(x,y)
v=[x;y];
dim=size(v,1);
z=1/4000*sum((v-100).^2,'r')..
    -prod(cos(((v-100)*inv(sqrt(diag(1:dim))))' ),'r' )+1;
endfunction

xset('window',0); xbascc();
x=-300:10:300;y=x;
fplot3d(x,y,Griewangk,alpha=85,theta=45);
xtitle('Fonction de Griewangk - vue d''ensemble')
xset('window',1); xbascc();
xset('colormap',graycolormap(256))
x=70:130;y=x;
Sfgrayplot(x,y,Griewangk);
xtitle('Fonction de Griewangk - détail de la surface')

function z=costfunc(z)
z=Griewangk(z, []);
endfunction

//minimisation en dimension 30 avec une (10,70)-ES
//initialisation dans [-300,300]^30
//40000 évaluations de Griewangk
[best,cost]=es(10,70,30,[-300,300], [], es_tol(1e-6,0,0),40000)
//minimisation en dimension 30 par CMAES
[best,cost]=cmaes(10,-1,-1,-1,30,[-300,300], [],1e-6,40000)
```

A titre d'exemple d'application pratique, le problème de la détection automatique de la signalisation verticale a été traité à l'aide d'une ES [DC03].

## 9.6 Une métaphore immunitaire : la sélection clonale

### 9.6.1 Quelques mots sur le système immunitaire

Cette section se borne à décrire un aspect du système immunitaire des vertébrés qui sert de base à l'algorithme de sélection clonale. Pour une présentation du système immunitaire dans son entier, le lecteur est prié de se reporter à la littérature spécialisée, comme [Alb02][JWT04].

Dans son fonctionnement normal, le système immunitaire des vertébrés assure la défense spécifique de l'organisme contre les pathogènes qui l'assaillent continuellement. Par *spécifique*, on entend la défense contre un pathogène particulier. Ces pathogènes que l'on nomme *antigènes* peuvent être de nature très variable, métaux lourds, composés organiques de synthèse, bactéries,...

Le système immunitaire sécrète des protéines particulières appelées *anticorps* qui permettent la neutralisation et la destruction des antigènes.

La fabrication d'un anticorps est codée pour partie dans les gènes de l'immunité. Ce codage se traduit par la production d'un certain nombre de « briques de base » protéïques pour confectionner un anticorps. On appelle *recombinaison somatique*, le mécanisme qui mène aux anticorps fonctionnels par arrangement et assemblage de ces briques de base. La diversité de molécules produites par la recombinaison est prodigieuse. On estime en effet à  $10^{11}$  le nombre de types d'anticorps qui sont présents

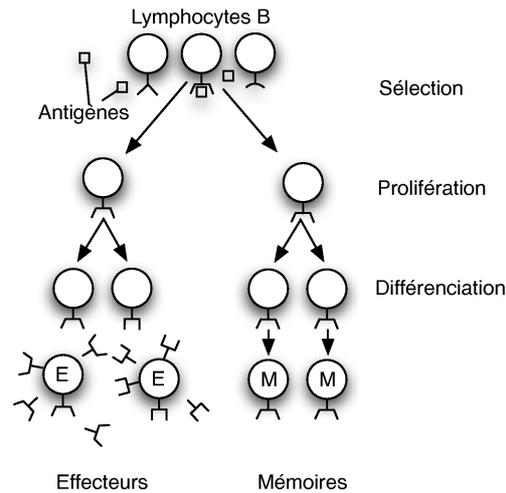


FIG. 9.4 – Principe de la sélection clonale

à un instant donné dans le corps humain. C'est un des aspects remarquables du système immunitaire.

L'interception d'un antigène par le système immunitaire suppose l'existence d'un anticorps présentant une complémentarité ou *affinité* suffisante avec l'antigène. La réponse du système immunitaire consiste entre autres à moduler la multiplication des cellules productrices d'anticorps, appelées Lymphocytes B, en fonction de leur affinité avec l'antigène. Ce mécanisme est illustré sur la figure 9.4. Un lymphocyte B ne produit qu'un seul type d'anticorps. La multiplication s'effectue par *clonage*, c'est à dire copie à l'identique. Les clones obtenus occuperont deux fonctions, soit la production de nouveaux anticorps pour la neutralisation d'autres occurrences de l'antigène, soit le rôle de mémoire, en vue d'une intrusion ultérieure de l'antigène. Dans ce cas, la réponse du système immunitaire serait à la fois plus intense, par la quantité d'anticorps produits, et plus rapide. La durée de vie d'un Lymphocyte B mémoire est de plusieurs mois à plusieurs années.

Au vu de cette présentation très grossière et très incomplète de la réponse humorale chez les vertébrés, une analogie se dégage entre l'affinité de l'association antigène / anticorps et l'optimisation. Le système immunitaire présente donc deux caractéristiques hautement souhaitables en optimisation. D'une part le système immunitaire est capable de produire une grande diversité de solutions candidates. D'autre part il est doté d'une capacité d'apprentissage et de sélection.

De la multiplication contrôlée des lymphocytes B, de Castro et von Zuben ont tiré l'algorithme de *sélection clonale* [dCvZ02] qui est présenté dans la suite. D'autres auteurs se sont appuyés sur le *réseau immunitaire* qui permet d'expliquer le maintien sur le long terme de la mémoire des réponses immunitaires passées, sans que cette mémoire ne soit trop dispendieuse pour l'organisme qui l'héberge [FMT98].

Ces aspects du système immunitaire ne sont pas les seuls qui se prêtent à des analogies avec des questions d'ingénierie et l'optimisation n'est pas la seule application envisageable. Une autre application, peut-être plus immédiate est celle de la reconnaissance de formes. Pour une revue des applications des systèmes d'immunité artificielle, le lecteur pourra consulter [Das98].

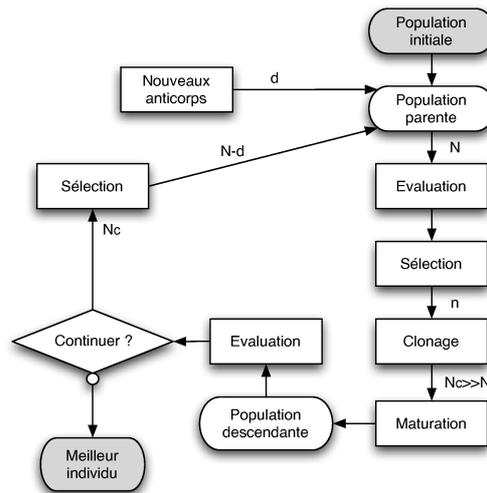


FIG. 9.5 – Organigramme de l'algorithme de sélection clonale

### 9.6.2 Présentation de l'algorithme

L'organigramme de l'algorithme de sélection clonale adapté à l'optimisation est présenté sur la figure 9.5.

L'algorithme manipule une population d'anticorps de taille  $N$ . Chaque anticorps contient un jeu de valeurs pour les paramètres à optimiser. La population initiale peut être produite par tirage aléatoire. Pour les itérations supérieures, cette population se compose de  $N - d$  anticorps sélectionnés au fil des itérations précédentes et d'un complément de  $d$  nouveaux anticorps produits par tirage aléatoire. Cette population est évaluée pour la fonction coût considérée. Elle est classée par rang de mérite décroissant. Les  $d$  derniers anticorps sont éliminés. Les anticorps qui restent subissent une phase de clonage. Le nombre de clones peut-être une fonction décroissante de l'affinité ou une constante. Selon de Castro et von Zuben, la première règle convient bien à la recherche du minimum global, la seconde à la recherche de plusieurs minima de bonne qualité. La phase de maturation a pour but d'introduire des variations dans les clones obtenus. La référence [dCvZ02] est assez vague sur la nature de cet opérateur. Dans le code mise à disposition avec le présent chapitre, on a implanté un opérateur de mutation similaire à celui de l'algorithme génétique canonique. Les  $N_c$  différents clones mutés sont évalués, classés. On ne garde que les  $N - d$  meilleurs, avant de régénérer une population d'effectif  $N$  et de reprendre la boucle à son début, jusqu'à ce que le critère de sortie soit atteint.

### 9.6.3 Performance et convergence

L'algorithme de sélection clonale est assez récent et la théorie sous-jacente est moins développée que pour les autres algorithmes présentés jusqu'ici. Les auteurs de la publication de référence ne donnent aucun résultat sur la probabilité ou la vitesse de convergence. Il serait étonnant qu'on ne parvienne pas à prouver que l'algorithme converge vers le minimum global avec une probabilité égale à 1. L'effet des paramètres internes est seulement étudié empiriquement sur des fonctions multimodales. Point remarquable, la complexité de l'algorithme de sélection clonale est en  $O(N_c)$ , alors

que celle d'un algorithme évolutionnaire est en  $O(N^2)$ .

L'évaluation empirique sur un nombre conséquent de fonctions test qui a été réalisée au LRPC de Strasbourg (Cf. annexe) met en évidence les excellentes performances de cet algorithme. Celui-ci s'affirme comme le meilleur l'ensemble des fonctions étudiées. Autre point à relever, la dispersion des résultats donnés par la sélection clonale est très faible, par rapport aux algorithmes concurrents.

### 9.6.4 Utilisation du code mis à disposition

Comme mentionné plus haut, l'algorithme peut être réglé pour rechercher un seul minimum, ou plusieurs minima de bonne qualité. Le code associé au premier (resp. second) cas est rassemblé dans `gcs.sci` (resp. `lcs.sci`). Les prototypes des fonctions associées sont rappelés ci-dessous.

```
//version "globale" -- recherche de l'optimum global
function
[best,cost,nbiter,dur,evolution,pop_fin]=gCS(N,be,rho,d,nvar,span,init,tol,evmax)

//version "locale" -- maintien de plusieurs optima
function
[best,cost,nbiter,dur,evolution,pop_fin]=lcs(N,be,rho,d,nvar,span,init,tol,evmax)
//entrées
//N : entier, taille de la population
//be : scalaire, facteur multiplicatif (be > 1/N)
//rho : scalaire, facteur de décroissance (ro >= 1), calcul du taux de mutation
//d : entier, nombre de nouveaux elements aleatoires (0 <= d < N)
//nvar : entier, nombre de variables à optimiser
//span : [min,max] ou [min1,max1;...;minn,maxn], domaine de définition des variables
//init : matrice (nvar,N) optionnelle, point de depart de l'optimisation,
// si [] -> initialisation par v.a. uniforme sur span
//tol : scalaire, seuil de convergence
//evmax : entier, nombre maximal d'évaluations de la fonction cout
//sorties : idem ga()
```

Concernant le choix des paramètres, on peut poser sans grand risque  $N = 10$ ,  $be = 1$ . La valeur optimale de  $rho$  est susceptible de varier largement. Il est conseillé de ne pas prendre une valeur de  $d$  supérieure à 2, sous peine de voir l'optimisation se transformer en mouvement brownien. L'optimisation de ces deux derniers paramètres ne donne pas les mêmes valeurs, selon que l'on utilise `gCS()` ou `lcs()`.

Le listing Scilab suivant présente la minimisation de la fonction de Powell à l'aide de `gCS()` :

```
function [alpha,A,B]=Powell_initb(dim)
// vect aléatoire dans [-pi;pi] fixé -> position du minimum
alpha=2*pi*rand(dim,1)-pi;
A=200*rand(dim,dim)-100;
B=200*rand(dim,dim)-100;
endfunction

function z=Powellb(x,y)
v=[x;y];
z=sum((((A*sin(alpha))..
+(B*cos(alpha))*ones(1,size(v,2)))..
-((A*sin(v))+B*cos(v)))).^2,'r');
endfunction

function z=costfunc(z)
z=Powellb(z,[]);
endfunction

[alpha,A,B]=Powell_initb(2);
x=-pi:0.1:pi;y=x;
```

```
fplot3d_loop(x,y,costfunc);
xlabel('Fonction de Powell')

nvar=30;
N=10; be=1; ro=5; d=2;
[alpha,A,B]=Powell_init(nvar);
[best,cost]=gCS(N,be,ro,d,nvar,%pi*[-1,1],[],0,10000);
```

Dutilleux et Ecoti re pr sentent une application de la s lection clonale   l'identification de param tres intrins ques de chaussures drainantes dans [DE06].

## 9.7 Pratique de l'optimisation globale

### 9.7.1 Nombres al atoires

L'implantation informatique d'un g n rateur de nombres al atoires dont la distribution observ e corresponde   la distribution sp cifi e reste un exercice d licat, m me pour les lois uniforme et gaussienne. La portabilit  de ce genre de fonctions pose souvent probl me. La position orthodoxe est donc d'affirmer qu'il est donc prudent de tester le g n rateur avant de l'utiliser dans un des algorithmes pr sent s pr c demment [DPST03].

Toutefois d'autres auteurs ont montr  que la substitution d'oscillateurs chaotiques aux g n rateurs de nombres al atoires classiques pouvait am liorer d'un facteur 10 les r sultats obtenus sur les fonctions tests de De Jong [CFFX03]. Parmi les oscillateurs essay s, la fonction logistique et une fonction dite « gaussienne » sont les plus performantes. L'interpr tation de ce ph nom ne n'est pas  tablie. Il semble que ce soit le caract re al atoire qui prime le respect d'une certaine distribution.

### 9.7.2 L'apport de l'expertise

M me si les algorithmes pr sent s ici sont con us pour  chapper aux minima locaux, ils ne sont pas   l'abri de la convergence pr matur e, et, statistiquement, la convergence est d'autant plus rapide que le volume de l'espace de recherche est r duit au minimum et que le point de d part de la recherche est proche de l'optimum.

La connaissance du probl me physique   traiter peut  tre incorpor e   deux niveaux. D'une part, elle peut servir   produire une « bonne » population initiale. D'autre part, l'intervalle dans lequel chaque variable est   rechercher doit  tre incorpor    la d finition du probl me. Dans le cas d'optimisation avec codage binaire, l'op rateur de codage doit  tre d fini au plus juste. Avec ce type de codage, on est certain que l' volution de la population candidate restera   l'int rieur des limites fix es au d part. En variables continues, il est prudent d'introduire des contraintes sur l'appartenance des variables   leur ensemble de d finition. On  vite ainsi que l'algorithme ne diverge vers des zones non pertinentes. A titre d'exemple, la d termination de la vitesse d'un tramway   partir de sa signature acoustique int gre une r gle telle que : la vitesse du v hicule est positive et inf rieure   60 km/h [Dut04]. Autre exemple, le probl me de la d tection de la signalisation verticale m ne aussi   la d finition d'un ensemble de contraintes *a priori* [Sia07].

### 9.7.3 Evaluer des algorithmes sur une fonction

Il a déjà été mentionné que le caractère stochastique de l'optimisation signifie que deux optimisations successives sans réinitialisation du générateur de nombres aléatoires donneront des résultats *a priori* différents. Dans la mesure du possible, faire des statistiques sur plusieurs optimisations est préférable, pour les réglages de l'algorithme comme pour l'optimisation elle-même. Tous les algorithmes ne sont pas égaux sur ce point, puisque la sélection clonale présente généralement une dispersion de résultats très faible.

Lorsqu'il s'agit de comparer deux algorithmes différents sur un problème, le nombre d'itérations est souvent peu pertinent. Il est préférable de travailler à nombre constant d'évaluations de la fonction coût.

### 9.7.4 Définir une fonction coût

Le cas typique de la résolution de problème inverse est celui de l'identification d'un modèle à partir de mesures. On est donc amené à définir un écart entre l'un et l'autre. La formulation qui vient immédiatement à l'esprit est celle des moindres carrés. Celle-ci présente le tort d'accorder beaucoup de poids aux points aberrants. On peut lui préférer le filtrage médian, ou un critère de coût semi-quadratique. Voir à ce sujet la contribution de Pierre Charbonnier dans cet ouvrage.

Lorsque l'identification s'effectue sur un ensemble de valeurs d'un paramètre (cas d'un spectre), d'autres définitions de l'ajustement sont possibles, comme la dérivée par différences finies.

Dans la mesure du possible, faire une représentation graphique de la fonction coût peut aider à définir une fonction coût pertinente.

La durée de calcul est un obstacle fréquent à la résolution de problème inverse et l'essentiel de cette durée est dû à la fonction coût. Lorsque celle-ci est très coûteuse, selon le problème, il peut exister des possibilités de recalcul partiel, ou de remplacement de calculs répétés par des valeurs tabulées. Des auteurs ont aussi proposé une optimisation reposant pour partie sur une approximation de la fonction coût qui s'évalue plus rapidement que cette dernière [JOS02]. L'optimisation sur une approximation résultant d'un apprentissage est vouée à l'échec, surtout quand la dimension de l'espace de solutions augmente. Les auteurs obtiennent des résultats satisfaisants, soit en évaluant à chaque itération les meilleurs individus de manière précise et les autres via l'approximation, soit en évaluant la population entière de façon précise toutes les  $n$  itérations.

### 9.7.5 Optimisation avec contraintes

La contrainte la plus simple est celle qui porte sur l'appartenance d'une variable à son ensemble de définition, le cas échéant. Pour les algorithmes génétiques ou la sélection clonale, sous réserve d'une représentation entière, le respect de la contrainte est codé dans la représentation elle-même. Si on n'y prend pas garde pour les autres algorithmes, les itérations peuvent mener les solutions courantes en dehors de l'espace des solutions réalistes ou physiquement admissibles. La solution qui consiste à rejeter une transition hors domaine et à en produire une autre est en général nuisible à l'optimisation. Il vaut mieux lui préférer une pénalisation des valeurs hors domaine sous la forme de termes qui viennent s'ajouter à la fonction coût proprement dite. Il

importe alors de s'assurer que le poids de ces termes est suffisant pour dissuader suffisamment les divagations.

Faute de place et d'expérience de l'auteur sur le sujet, ce chapitre ne traite pas de l'optimisation avec contraintes. L'ouvrage classique de Michalewicz y consacre un chapitre [Mic99]. Le même auteur a publié un article de synthèse spécifiquement sur ce point [Mic95]. Voir aussi [Coe02] pour un état de l'art plus récent.

Enfin, on rappelle que l'algorithme de société et civilisation qui dérive de l'essaim de particules est conçu pour manipuler des contraintes et qu'il a donné des résultats prometteurs sur des problèmes d'ingénierie.

### 9.7.6 Utilisation de la boîte à outils d'optimisation Scilab.

Le code Scilab mis à disposition sur le CDROM est organisé de telle sorte que chaque bibliothèque `.sci` contienne tout le code requis pour un algorithme donné. La plupart du temps, l'usage se limite à l'appel d'une seule fonction. Le code est aussi proposé sous la forme d'une boîte à outils (*toolbox*) Scilab, avec exemples et aide en ligne. Les implantations en Scilab ont été évaluées sur plusieurs fonctions test classiques. Cette évaluation fait l'objet d'une annexe à cet article. Elle est consultable sur le CDROM

Il faut évidemment avoir défini au préalable une fonction coût, qui doit s'appeler `costfunc()` et accepter *a priori* une matrice contenant autant de lignes que de variables et un nombre de colonnes arbitraire. Il n'est pas nécessaire que le corps de cette fonction soit écrit en Scilab. Scilab s'interface en effet très facilement avec du code C, fortran, Ada, Java. Une architecture client/serveur est également possible. On trouve toutes les informations nécessaires pour l'interfaçage avec C et fortran dans [CDG<sup>+</sup>01], qui constitue d'ailleurs une excellente introduction à Scilab en général, quoique Scilab ait beaucoup évolué depuis la publication de cet ouvrage.

## 9.8 Conclusion

Cinq métaheuristiques pour l'optimisation globale, inspirées, soit de la physique, soit de la biologie ont été présentées succinctement.

Ces algorithmes d'optimisation globale doivent toujours être vus comme des solutions de dernier recours, lorsque les méthodes locales échouent. D'une part, la théorie concernant leur convergence est en général assez mince. D'autre part, l'obtention d'un optimum de bonne qualité nécessite souvent un grand nombre d'évaluations de la fonction coût. Il n'en reste pas moins que ces algorithmes sont irremplaçables dans bon nombre de problèmes notoirement difficiles.

Le recuit simulé et l'essaim de particules ont le mérite de la simplicité de mise en oeuvre. Les algorithmes évolutionnaires sont incontournables. Du point de vue de l'optimisation de variables continues, même s'ils sont de loin les plus utilisés, les algorithmes génétiques sont surtout à recommander pour des problèmes mixtes, qui combinent variables discrètes et continues. Moins connues, les stratégies d'évolution sont pourtant mieux adaptées. Parmi tous les algorithmes décrits ici, la sélection clonale semble être l'algorithme le plus performant, en termes de minima obtenus et de dispersion des minima. Autre point remarquable, moyennant un paramétrage approprié, la sélection clonale permet d'identifier plusieurs minima à la fois.

Cependant, en vertu du *No Free Lunch theorem*, il n'existe pas d'algorithme « à tout faire » [WM97]. D'autre part, la faiblesse de la théorie associée à ces algorithmes a été mise en évidence. De ce fait, expérimenter avec plusieurs d'entre eux est une bonne approche. C'est pourquoi ce chapitre est accompagné d'implémentations SciLab validées des 5 algorithmes étudiés (SA, GA, ES, PSO, GCS), ainsi que de deux variantes de ces algorithmes (CMAES et LCS). En complément, quelques recommandations pratiques devraient aider le débutant en optimisation globale.

Dans le futur, la théorie de la convergence des métaheuristiques devrait progresser. Pour autant, la mise à disposition de vitesses de convergence pour les différents espace des fonctions n'est pas pour demain. En revanche, la panoplie d'algorithmes ne manquera pas de s'élargir. Les principes présentés ici devraient aider le lecteur à les aborder.



# Bibliographie

- [Alb02] Bruce Alberts. *The adaptive immune system*, pages 1363–1421. Garland Science, 2002.
- [Bäc96] Thomas Bäck. *Evolutionary algorithms in theory and practice - Evolution Strategies - Evolutionary Programming - Genetic Algorithms*. Oxford University Press, 1996.
- [CDG<sup>+</sup>01] J.P. Chancelier, F. Delebecque, C. Gomez, M. Goursat, R. Nikoukah, and S. Steer. *Introduction à Scilab*. Springer Verlag, 2001.
- [CFFX03] Riccardo Caponetto, Luigi Fortuna, Stefano Fazzino, and Maria Gabriella Xibilia. Chaotic sequences to improve the performance of evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 7(3) :289–304, 2003.
- [CK02] Maurice Clerc and James Kennedy. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(1) :386–396, 2002.
- [CMMR87] A. Corana, M. Marchesi, C. Martini, and S. Ridella. Minimizing multimodal functions of continuous variables with the simulated annealing algorithm. *ACM Transactions on Mathematical Software*, 13(3) :262–280, 1987.
- [CMV<sup>+</sup>04] Andrew Czarn, Cara MacNish, Kaipillil Vijayan, Berwin Turlach, and Ritu Gupta. Statistical explanatory analysis of genetic algorithms. *IEEE Transactions on Evolutionary Computation*, 8(4) :405–421, 2004.
- [Coe02] C.A.C. Coello. Theoretical and numerical constraint handling techniques used with evolutionary algorithms : a survey of the state of the art. *Computer methods in applied mechanics and engineering*, 191(11-12) :1245–1287, 2002.
- [Dan05] Etienne Danchin. Inadvertent social information and decision-making in birds : a new paradigm for european ornithology. In *5ème conférence de l'Union Européenne d'Ornithologie*. EOU, 2005.
- [Dar99] Charles Darwin. *L'origine des espèces*. Flammarion, 1999.
- [Das98] Dipankar Dasgupta. *Artificial Immune Systems and their Applications*. Springer Verlag, 1998.
- [DC03] Guillaume Dutilleux and Pierre Charbonnier. Template matching et stratégies d'évolution - application à la détection d'objets dans les scènes routières. In *Journées des Sciences de l'Ingénieur du réseau des laboratoires des ponts et chaussées*. Laboratoire Central des Ponts et Chaussées, 2003.

- [dCvZ02] L.N. de Castro and F.J. von Zuben. Learning and optimization using the clonal selection principle. *IEEE Transactions on Evolutionary Computation*, 6(3) :239–251, 2002.
- [DE06] Guillaume Dutilleux and David Ecotière. Identification automatique de paramètres intrinsèques de revêtements et mesure d'absorption acoustique. In *Journées des Sciences de l'Ingénieur du réseau des laboratoires des ponts et chaussées*. Laboratoire Central des Ponts et Chaussées, 2006.
- [DPST03] Johann Dréo, Alain Pérowski, Patrick Siarry, and Eric Taillard. *Métaheuristiques*. Eyrolles, 2003.
- [Dut04] Guillaume Dutilleux. Acoustic-signature-based determination of railway vehicle speed - application to tramways. In *Proc. Joint Congress CFA/DAGA'04*. SFA/DEGA, 2004.
- [FMT98] T. Fukuda, K. Mori, and M. Tsukiyama. *Parallel search for multi-modal function optimization with diversity and learning of immune algorithm*, pages 210–220. Springer Verlag, 1998.
- [Fog64] Larry J. Fogel. *On the organization of intellect*. University of California, 1964.
- [Fut05] Douglas J. Futuyma. *Evolution*. Sinauer, 2005.
- [Gre86] John J. Grefenstette. Optimization of control parameters for genetic algorithms. *IEEE Transactions on Systems Man, and Cybernetics*, 16(1) :122–128, 1986.
- [HA96] Nikolaus Hansen and Ostermeier Andreas. Adapting arbitrary normal mutation distributions in evolution strategies : the covariance matrix adaptation. In *IEEE Conference on Evolutionary Computation*. IEEE, 1996.
- [HG90] F. Heppner and U. Grenander. *A stochastic non-linear model for coordinated bird flocks*, pages 233–238. AAAS publications, 1990.
- [Hol75] John Holland. *Adaptation in natural and artificial systems*. University of Michigan, 1975.
- [Jon75] K. De Jong. *An analysis of the behaviour of a class of genetic adaptive systems*. University of Michigan, 1975.
- [JOS02] Y. Jin, M. Olhoffer, and B. Sendhoff. A framework for evolutionary optimization with approximate fitness function. *IEEE Transactions on Evolutionary Computation*, 6(5) :481–494, 2002.
- [JWT04] Charles A. Janeway, Mark Walport, and Paul Travers. *Immunobiology 6th Edition*. Garland Publishing, 2004.
- [KE95] James Kennedy and Russell C. Eberhart. Particle swarm optimization. In *IEEE International Conference on Neural Networks*. IEEE, 1995.
- [KES01] James Kennedy, Russel C. Eberhart, and Yuhui Shi. *Swarm Intelligence*. Morgan Kaufmann Publishers, 2001.
- [KGV83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220 :671–680, 1983.
- [Kim92] Motoo Kimura. *Théorie neutraliste de l'évolution*. Flammarion, 1992.

- [MCPB00] Max Mignotte, Christophe Collet, Patrick Pérez, and Patrick Bouthemy. Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(2) :129–141, 2000.
- [Mic95] Zbigniew Michalewicz. A survey of constraint handling techniques in evolutionary computation methods. In *4th Annual Conference on Evolutionary Programming*. MIT Press, 1995.
- [Mic99] Zbigniew Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs - Third, Revised and Extended Edition*. Springer, 1999.
- [MKN04] Rui Mendes, James Kennedy, and José Neves. The fully informed particle swarm : simpler, maybe better. *IEEE Transactions on Evolutionary Computation*, 8(3) :204–210, 2004.
- [MRR<sup>+</sup>53] N. Metropolis, A.W. Rosenbluth, M. Rosenbluth, A.H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *Chem. Phys*, 21 :1087–1092, 1953.
- [Rec65] Ingo Rechenberg. *Cybernetic solution path of an experimental problem*. Royal Aircraft Establishment, 1965.
- [Rec94] Ingo Rechenberg. *Evolutionsstrategie'94*. Frommann-Holzboog, 1994.
- [Rey87] Craig W. Reynolds. Flocks, herds and schools : a distributed behavioral model. *Computer Graphics*, 21(4) :25–34, 1987.
- [RL03] Tabrapa Ray and K.M. Liew. Society and civilization : an optimization algorithm based on the simulation of social behavior. *IEEE Transactions on Evolutionary Computation*, 7(4) :386–396, 2003.
- [Sia07] Patrick Siarry. *Métaheuristiques biologiques pour la détection de la signalisation routière*, pages 271–294. Hermès, 2007.
- [Sim04] A.M. Simons. Many wrongs : the advantage of group navigation. *Trends in Ecology and Evolution*, 19(9) :453–455, 2004.
- [SS96] Michèle Sebag and Marc Schoenauer. Contrôle d'un algorithme génétique. *Revue d'intelligence artificielle*, 10 :389–428, 1996.
- [vLA87] P.M.J. van Laarhoven and E.H.L. Aarts. *Simulated Annealing - Theory and Applications*. Reidel, 1987.
- [Wil00] Edward O. Wilson. *Sociobiology : the new synthesis*. Belknap, 2000.
- [WM97] David H. Wolpert and William G. Macready. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1) :67–82, 1997.



## Chapitre 10

# Optimisation de forme de murs antibruit par algorithmes génétiques

*Denis Duhamel*

### 10.1 Introduction

Nous proposons ici une application des algorithmes génétiques à l'optimisation de forme des murs antibruit. Pour réduire le bruit au voisinage des routes, il est courant de construire des murs antibruit pour protéger les habitations situées derrière. La construction de murs efficaces et capables de fournir une atténuation importante du bruit est un enjeu important pour l'amélioration de la qualité de l'environnement. Il est bien connu que la position et la hauteur de la barrière sont des paramètres importants de son efficacité. Il est possible d'augmenter l'atténuation en construisant un mur de hauteur plus élevée. Cependant pour des raisons pratiques ou esthétiques, il n'est généralement pas souhaitable d'avoir des murs trop hauts. Il est par conséquent utile d'essayer d'optimiser l'efficacité d'un mur pour une hauteur maximale donnée.

Un grand nombre de personnes se sont intéressées à ce problème d'un point de vue expérimental ou numérique. Pour des barrières de formes complexes, [Sez80] a proposé d'effectuer les calculs par la méthode des éléments de frontière (BEM) pour pouvoir estimer l'efficacité d'un mur avec précision. Par la suite [HCWH91, HCCW91, CH94] ont calculé par BEM à deux dimensions un grand nombre de formes différentes comme un mur droit, circulaire, en Y ou en T avec différents revêtements plus ou moins absorbants. Ils ont pu déterminer que la forme en T présentait une amélioration claire par rapport aux autres cas. Plusieurs études expérimentales ont aussi été effectuées. Dans [MO80, FHK98] des mesures sur modèle réduit ont été réalisées sur un grand nombre de formes et de nouveau il en a résulté que la forme en T recouverte d'un matériau absorbant présentait un avantage clair. Des tests sur des murs réels effectués par [WCH94, Wat96, WM96] sur des formes en T ou des murs doubles ont permis de montrer qu'ils amélioraient l'atténuation de 1 à 3 dB par rapport à une barrière droite et sans absorbant.

Dans les études précédentes les auteurs proposent différentes formes de barrière avec ou sans absorbant et testent leur efficacité par des méthodes numériques ou en ayant recourt à des moyens expérimentaux. Pour améliorer l'efficacité des murs, ils essaient d'imaginer de nouvelles formes, par exemple en changeant la forme du

sommet que l'on sait être importante pour la diffraction des ondes. Ensuite il faut tester cette nouvelle forme pour estimer le gain obtenu. Cependant cette approche ne permet pas de proposer une méthode systématique pour obtenir des formes intéressantes.

Dans ce travail nous proposons d'utiliser une approche différente avec une méthode couplée basée sur les éléments de frontière pour calculer le champ de pression autour du mur et sur les algorithmes génétiques pour optimiser la forme. Nous rappellerons d'abord le principe du calcul par équation intégrale puis nous l'appliquerons à quelques exemples. Le but n'est pas réellement de proposer des formes avec une application pratique immédiate mais d'estimer le gain potentiel sur l'efficacité des murs qui pourrait être obtenu en utilisant des formes plus efficaces que celles que l'on utilise actuellement.

## 10.2 Calcul du champ de pression

Le problème étudié est décrit sur la figure 11.4 qui présente un mur de forme arbitraire. La modélisation est bidimensionnelle ce qui est suffisant si nous nous

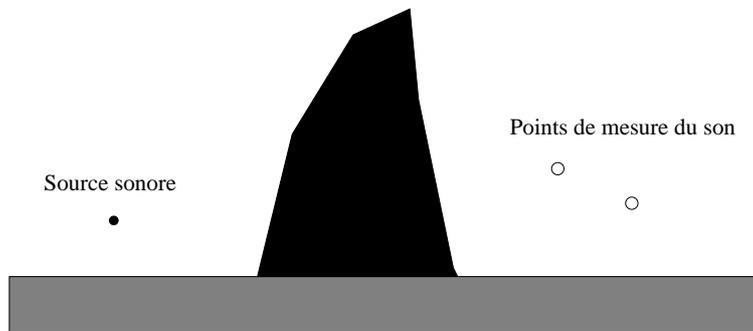


FIG. 10.1 – Mur antibruit de forme complexe.

intéressons uniquement à la propagation normale au mur. Dans [Duh96] il est montré que les calculs 2D et 3D conduisent au même résultat en terme d'atténuation. Une source sonore émet un bruit d'un côté du mur et nous nous intéressons au niveau de bruit de l'autre côté dans l'objectif de le minimiser. Pour simplifier l'étude et pour mettre en lumière l'influence de la forme nous ferons l'hypothèse que les surfaces du mur et du sol sont rigides. Le champ de pression sera supposé harmonique avec une dépendance en temps du type  $e^{-i\omega t}$  et est par conséquent solution de l'équation de Helmholtz.

$$\Delta p(x, \omega) + k^2 p(x, \omega) = s(x, \omega) \quad (10.2.1)$$

où  $p$  est la pression au point  $x$  et à la fréquence  $\omega$ ,  $k = \frac{\omega}{c}$  est le nombre d'onde,  $c$  est la vitesse du son et  $s(x, \omega)$  est le terme de source dans l'air qui, dans le cas d'une source ponctuelle au point  $x_s$ , est égal à  $-\delta(x - x_s)$ . Comme la barrière et le sol sont rigides, le champ de pression doit vérifier la condition suivante sur les deux surfaces

$$\frac{\partial p}{\partial n} = 0 \quad (10.2.2)$$

De plus le champ de pression doit satisfaire la condition de radiation à l'infini qui s'écrit sous la forme.

$$\lim_{r \rightarrow \infty} \left( \frac{\partial p}{\partial r} - ikp \right) = 0 \left( \frac{1}{\sqrt{r}} \right) \quad (10.2.3)$$

Pour des barrières de formes complexes, le champ de pression peut être calculé par la méthode des éléments de frontière. La fonction de Green  $G$  est le champ de pression produit au point  $x$  par une source ponctuelle au point  $y$  et est solution de

$$\Delta G(x, y) + k^2 G(x, y) = -\delta(x - y) \quad (10.2.4)$$

En dimension deux, elle est donnée par

$$G(r) = \frac{i}{4} H_0(kr) + \frac{i}{4} H_0(kr') \quad (10.2.5)$$

où  $r = |\mathbf{x} - \mathbf{y}|$  est la distance entre le récepteur et la source,  $r' = |\mathbf{x} - \mathbf{y}'|$  la distance entre le récepteur et la source image par rapport au sol et  $H_0$  est la fonction de Hankel d'ordre zéro et de premier type. Avec cette fonction de Green le champ de pression est solution de l'équation suivante sur la surface du mur uniquement.

$$c(x)p(x) = p_{inc}(x) + \int_{\Gamma} \left[ \frac{\partial G(x, y)}{\partial n_y} p(y) - G(x, y) \frac{\partial p(y)}{\partial n_y} \right] dy \quad (10.2.6)$$

où  $\Gamma$  est la frontière du mur,  $n$  la normale extérieure dirigée vers le fluide et  $p_{inc}$  est le champ de pression incident qui est produit par les sources sonores dans le milieu fluide en l'absence du mur. Le coefficient  $c(x)$  dépend de la position et est donné par

$$\begin{cases} c(x) = 1 & x \in \Omega_{ext} \\ c(x) = 0 & x \in \Omega_{int} \\ c(x) = \frac{1}{2} & x \in \Gamma \end{cases} \quad (10.2.7)$$

Nous avons noté  $\Omega_{ext}$  le domaine occupé par l'air et  $\Omega_{int}$  l'intérieur du mur. Il est bien connu que cette équation peut admettre plusieurs solutions pour un ensemble discret de fréquences correspondant aux fréquences de résonance du domaine intérieur au mur. Le problème peut être résolu par la méthode de Burton et Miller [BM] qui consiste à résoudre l'équation (10.2.6) plus sa dérivée normale multipliée par une constante que l'on prend égale à  $i/k$ . Après quelques transformations décrites dans [Fad02], la relation finale se met sous la forme.

$$\begin{aligned} & \frac{1}{2} \int_{\Gamma} q(x)p(x)dx + \int_{\Gamma \times \Gamma} q(x)G(x, y) \frac{\partial p(y)}{\partial n_y} dydx - \int_{\Gamma \times \Gamma} q(x) \frac{\partial G(x, y)}{\partial n_y} p(y) dydx \\ & + \frac{i}{k} \left[ \frac{1}{2} \int_{\Gamma} q(x) \frac{\partial p(x)}{\partial n_x} dx + \int_{\Gamma \times \Gamma} q(x) \frac{\partial G(x, y)}{\partial n_x} \frac{\partial p(y)}{\partial n_y} dydx \right. \\ & \left. + \int_{\Gamma \times \Gamma} \frac{\partial q(x)}{\partial \tau_x} G(x, y) \frac{\partial p(y)}{\partial \tau_y} dydx - k^2 \int_{\Gamma \times \Gamma} q(x)n_x.n_y G(x, y)p(y) dydx \right] \\ & = \int_{\Gamma} q(x)p_{inc}(x)dx + \frac{i}{k} \int_{\Gamma} q(x) \frac{\partial p_{inc}(x)}{\partial n_x} dx \end{aligned} \quad (10.2.8)$$

où  $q(x)$  est une fonction test définie sur  $\Gamma$ . Cette équation est résolue par les techniques classiques de la méthode des éléments de frontière après maillage de  $\Gamma$ . Le champ de pression peut ensuite être calculé en tout point du domaine extérieur par la formule (10.2.6). Dans la suite nous nous intéresserons à la minimisation du son derrière la barrière en utilisant la fonction de coût suivante.

$$J = 10 \log_{10} \left[ \frac{1}{n_p n_f} \sum_{i=1}^{n_f} \sum_{j=1}^{n_p} |p(x_j, \omega_i)|^2 / 4 \cdot 10^{-10} \right] \quad (10.2.9)$$

C'est la valeur en décibels de la moyenne du carré de la pression sur un nombre  $n_p$  de points en espace et sur  $n_f$  fréquences. La modification du nombre de points  $n_p$  permet d'avoir une zone de contrôle plus ou moins grande tandis que la modification de  $n_f$  permet d'agir sur la bande de fréquence sur laquelle est effectuée le contrôle.

Comme  $J$  peut être une fonction très complexe de la forme de la barrière, de la fréquence, des  $n_p$  points de contrôle et qu'elle peut uniquement être estimée numériquement, une méthode d'optimisation sophistiquée est nécessaire. Les algorithmes génétiques sont justement des méthodes d'optimisation adaptées aux problèmes complexes avec des fonctions de coût présentant de nombreux minimums locaux. Cette méthode ne nécessite que des évaluations de la fonction mais pas de ses dérivées ce qui est bien adapté au type de problème que l'on considère car la dérivée de la fonction (10.2.9) est difficile à calculer. Le problème mathématique est de trouver le minimum de la fonction  $J$  définie sur un ensemble  $U$  donnant les configurations possibles du système. Un point dans  $U$  est un individu et est codé par un chromosome. Dans notre problème un point dans  $U$  est une forme de barrière et le codage se fait sous forme d'une matrice de zéros et de uns. Le domaine que peut occuper la barrière est un rectangle que l'on divise en une matrice de sous rectangles comme sur la figure 11.5. Le vide est codé par 0 et la matière par 1. A partir de cette matrice, on définit la forme de la barrière en éliminant les éventuels trous intérieurs. Le contour extérieur du mur définit la frontière qui est maillée par des éléments quadratiques à trois noeuds.

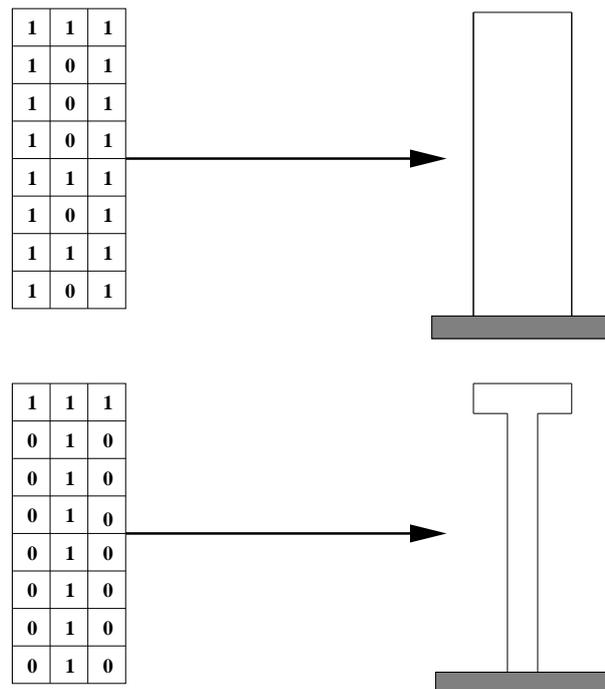


FIG. 10.2 – Construction de la géométrie à partir des données binaires.

Le calcul par équation intégrale permet donc d'associer une valeur de la fonction de coût (10.2.9) pour chaque matrice de zéros et de uns. L'algorithme génétique fait ensuite évoluer cette matrice pour se rapprocher du minimum de la fonction de coût. L'optimisation est effectuée en utilisant la bibliothèque Galib de [GAI] qui est un ensemble de classes en C++ qui implémente un grand nombre de variantes des algorithmes génétiques. On a choisi d'utiliser la classe GA2DBinarystring bien adaptée à

des chromosomes constitués à partir d'une matrice de zéros et de uns. Le taux de croisement a été choisi égal à 0.75, tandis que le taux de mutation est 0.002.

## 10.3 Résultats

### 10.3.1 Efficacité des murs droits et en T

Comme base de comparaison avec les calculs suivants, le niveau sonore derrière un mur droit de largeur 0.1 m et pour un mur en T sont calculés. Ces murs sont représentés sur la figure 11.6. L'origine du système de coordonnées est situé sur le sol au centre de

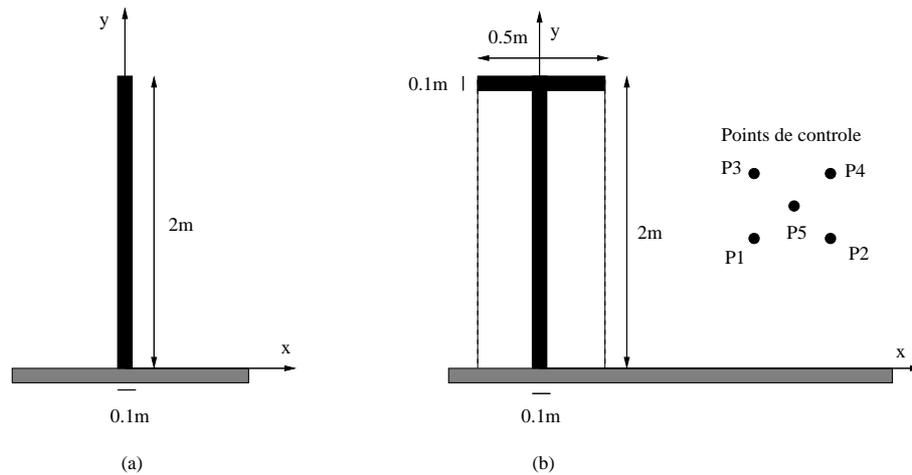


FIG. 10.3 – Murs de référence.

la barrière. La source sonore est placée au point  $(-5m, 0)$  sur le sol et rayonne la pression  $\frac{i}{4}H_0(kr)$  en champ libre. La vitesse du son est  $340m/s$ . Nous nous intéressons au niveau sonore derrière la barrière, en définissant deux zones de contrôle. La première zone, dite zone 1, est un carré de surface  $0.25m^2$ , défini par les points du tableau 10.1. La seconde zone est plus grande et est définie par les points du tableau 10.2. La fonc-

Points	P1	P2	P3	P4	P5
X(m)	2	2,5	2	2,5	2,25
Y(m)	1	1	1,5	1,5	1,25

TAB. 10.1 – Points définissant la zone 1.

Points	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
X(m)	2	3	4	5	2	3	4	5	2	3
Y(m)	0	0	0	0	0.5	0.5	0.5	0.5	1	1
Points	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20
X(m)	4	5	2	3	4	5	2	3	4	5
Y(m)	1	1	1.5	1.5	1.5	1.5	2	2	2	2

TAB. 10.2 – Points définissant la zone 2.

tion de coût pour les murs droit et en T est donnée dans le tableau 10.3. La fonction

$J_{125}$  est définie par la formule (10.2.9) pour la fréquence 125 Hz, tandis que  $J_{2000}$  est la même fonction pour la fréquence 2000 Hz et  $J_{125-2000}$  est obtenue en prenant la moyenne sur les fréquences 125, 250, 500, 1000 et 2000 Hz dans la formule (10.2.9). La forme en T est un peu plus efficace que la forme droite de 1 à 3 dB selon les fréquences. Le niveau moyen dans la zone 2 est un peu plus élevé que dans la zone 1, ce qui semble raisonnable puisque plus de points sont utilisés pour le calcul de la moyenne.

	J (dB) zone 1	J (dB) zone 2
Mur droit 125 Hz, $J_{125}$	56.7	59.1
Mur droit 2000 Hz, $J_{2000}$	33.9	35.2
Mur droit 125-2000 Hz, $J_{125-2000}$	51.8	53.4
Mur en T 125 Hz, $J_{125}$	55.8	57.2
Mur en T 2000 Hz, $J_{2000}$	30.9	33.5
Mur en T 125-2000 Hz, $J_{125-2000}$	50.0	51.3

TAB. 10.3 – Fonction de coût pour les murs de référence.

### 10.3.2 Convergence de l'algorithme

La convergence de l'algorithme a d'abord été testée. La barrière est représentée par un maillage de  $5 \times 20$  rectangles dans le domaine  $[-0.25m, 0.25m] \times [0, 2m]$ , chaque rectangle pouvant être rempli de matière ou vide. Un premier calcul a été effectué avec 500 itérations pour la fonction de coût incluant les cinq fréquences entre 125 et 2000 Hz. L'évolution de la fonction de coût pour cinq calculs indépendants est représentée sur la figure 10.4.

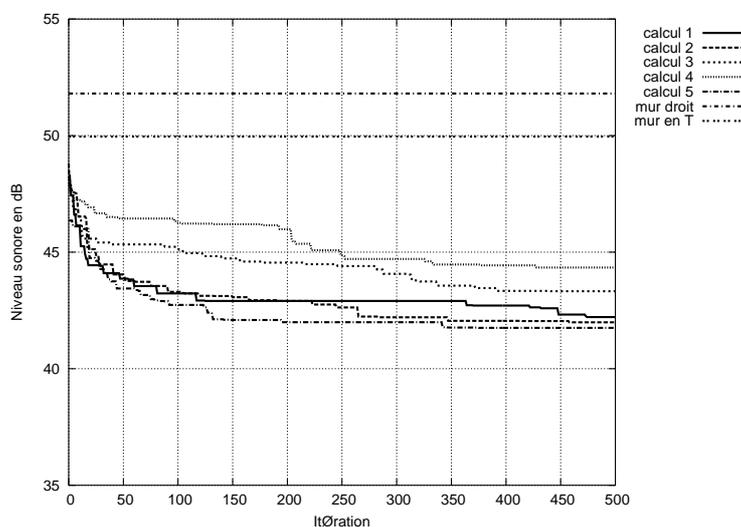


FIG. 10.4 – Evolution du niveau sonore avec le nombre d'itérations pour les fréquences 125-2000Hz.

On peut observer une décroissance rapide de la fonction de coût durant les 100 premières itérations puis une décroissance plus lente ensuite. La valeur finale de la fonction de coût présente une dispersion non négligeable suivant les calculs, de l'ordre de 3 dB. Les valeurs sont données dans le tableau 10.4. Il semble plus intéressant de lancer plusieurs calculs indépendants mais assez courts plutôt qu'un seul calcul très

long. On observe que l'essentiel du gain est obtenu après 300 itérations, c'est pourquoi les calculs suivants sont arrêtés à ce nombre d'itérations.

### 10.3.3 Efficacité et forme des barrières

Une diminution du niveau sonore de 8 dB est obtenue dans la bande de fréquence 125-2000Hz (voir tableau 10.4). Il semble donc possible de construire des barrières nettement plus efficaces que les barrières en T dans une large gamme de fréquences. Sur la figure 10.5, la fonction de coût  $J_{125-2000}$  est calculée pour la zone de contrôle 2. La valeur finale de la fonction de coût pour le meilleur calcul est de 46.3 dB. Dans le cas de la zone 1, la meilleure valeur finale était de 41.8 dB comme indiquée dans le tableau 10.4. Ainsi l'augmentation de la taille de la zone de contrôle conduit à une perte d'efficacité moyenne de 4.5 dB. Pour le mur en T la valeur de la fonction de coût était 51.3 dB (voir tableau 10.3). L'optimisation de forme de la barrière permet donc encore un gain de 5 dB sur la forme en T. La forme finale de la barrière optimale pour le cas conduisant à la fonction de coût minimale est donnée sur la figure 10.6. La forme est complexe et peut comporter des domaines multiples ou des trous. De telles formes ne pourraient pas être obtenues par des méthodes d'optimisation basées sur le mouvement continu de points sur la frontière.

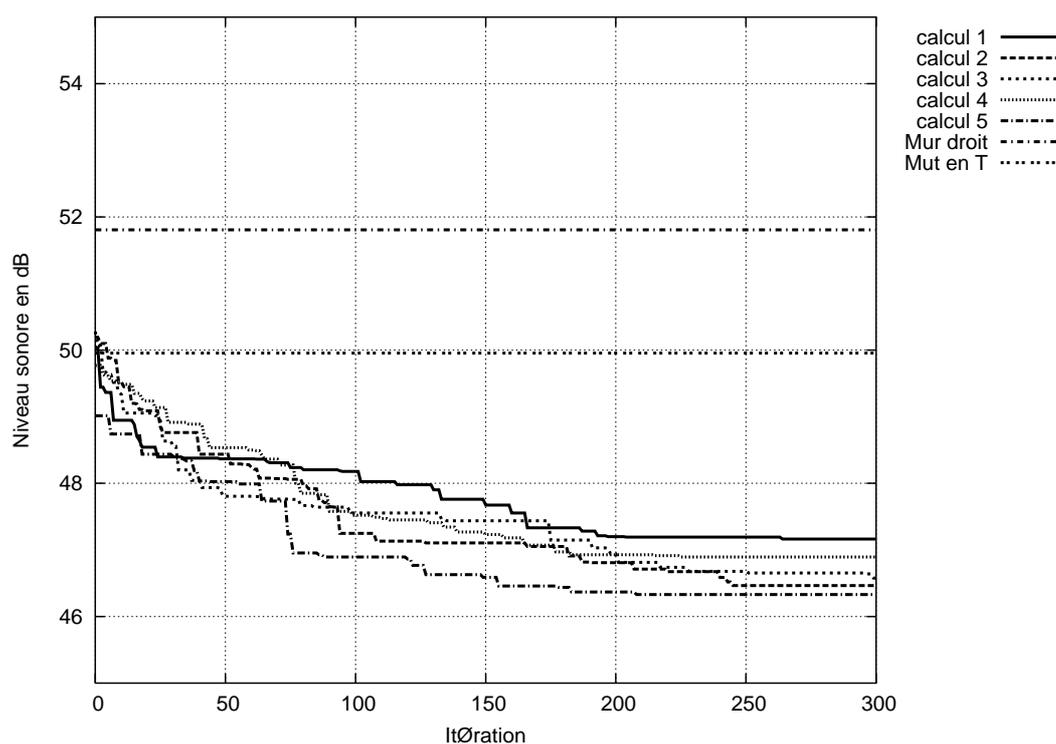


FIG. 10.5 – Evolution du niveau sonore avec le nombre d'itérations pour les fréquences 125-2000Hz, zone de contrôle 2.

Une comparaison du champ de pression entre la forme optimale et le mur en T est présentée sur la figure 10.7. On peut observer une diminution moyenne du niveau sonore derrière le mur d'environ 5 dB ce qui est cohérent avec la diminution de la valeur de la fonction de coût.

	$J_{125-2000}$ (dB)
Optimisation 1	42.2
Optimisation 2	42.0
Optimisation 3	43.3
Optimisation 4	44.3
Optimisation 5	41.8
Meilleur cas	41.8
Meilleur cas - mur en T	-8.2
Meilleur cas - mur droit	-10.0

TAB. 10.4 – Fonction de coût pour cinq calculs après 500 itérations.

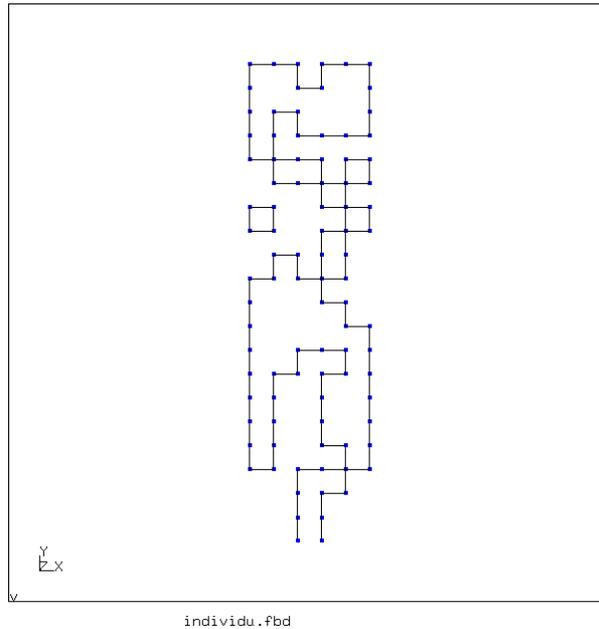


FIG. 10.6 – Forme optimale du mur.

### 10.3.4 Position de la source

Comme test final, la barrière optimale représentée sur la figure 10.6 est calculée pour deux autres positions de la source sonore dans le but de déterminer si la forme optimale trouvée pour une position est encore optimale pour une autre position. La fonction de coût est celle définie pour la zone 1. La source sonore peut être placée au point (-3, 0.5) ou au point (-7, 0.1). Les valeurs de la fonction de coût pour les différentes positions de source sont données sur le tableau 10.5. On peut observer que la meilleure forme trouvée est plus efficace que le forme en T d'au moins 6 dB. Il semble donc que la forme optimale trouvée pour une position de source soit encore efficace pour d'autres positions.

## 10.4 Conclusions

Une méthode d'optimisation de forme des murs antibruit est proposée. Elle est basée sur une méthode intégrale pour le calcul du champ de pression autour du mur et

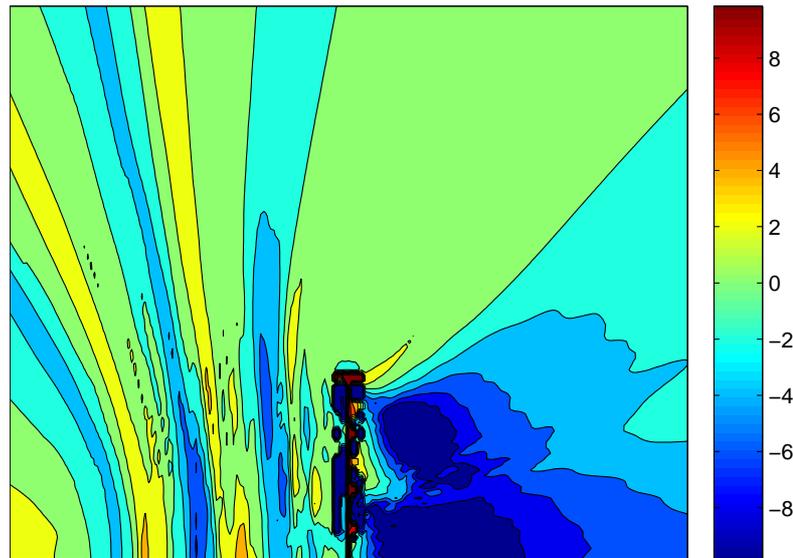


FIG. 10.7 – Différence entre le niveau sonore pour la forme optimale et le mur en T pour les fréquences 125-2000Hz.

	$J_{125-2000}$ (dB)
Forme en T, source à la position 1 (-5,0)	53.7
Forme en T, source à la position 2 (-3,0.5)	52.5
Forme en T, source à la position 3 (-7,0.1)	53.1
Forme optimale, source à la position 1 (-5,0)	46.3
Forme optimale, source à la position 2 (-3,0.5)	43.0
Forme optimale, source à la position 3 (-7,0.1)	47.2

TAB. 10.5 – Fonction de coût pour différentes positions de la source.

sur l'utilisation d'un algorithme génétique pour l'optimisation. Des résultats intéressants peuvent être obtenus après quelques centaines d'itérations. Les formes optimales sont complexes et sans doute difficiles à mettre en oeuvre en pratique. Cependant les gains supérieurs à cinq décibels obtenus par rapport aux formes classiques semblent indiquer qu'il est possible d'avoir des formes plus efficaces que celles mises en oeuvre actuellement. Il serait possible de poursuivre ce travail en optimisant aussi sur le revêtement du mur pour augmenter encore l'atténuation.



# Bibliographie

- [BM] A. J. Burton and G. F. Miller. The application of integral equation methods to the numerical solution of some exterior boundary value problems. In *Proc. Royal Society, London*.
- [CH94] D.H. Crombie and D.C. Hothersall. The performance of multiple noise barriers. *Journal of Sound and Vibration*, 1994.
- [Duh96] D. Duhamel. Efficient calculation of the three-dimensional sound pressure field around a noise barrier. *J. Sound and Vib.*, 1996.
- [Fad02] A. Fadavi. *Modélisation numérique des vibrations d'un pneumatique et de la propagation du bruit de contact*. PhD thesis, ENPC, 2002.
- [FHK98] K. Fujiwara, D. C. Hothersall, and C. Kim. Noise barriers with reactive surfaces. *Applied acoustics*, 1998.
- [GAI]
- [HCCW91] D.C. Hothersall, D. H. Crombie, and S.N. Chandler-Wilde. The performance of t-profile and associated noise barriers. *Applied acoustics*, 1991.
- [HCWH91] D.C. Hothersall, S.N. Chandler-Wilde, and M.N. Hajmirzae. Efficiency of single noise barriers. *Journal of Sound and Vibration*, 1991.
- [MO80] D.N. May and M.M. Osman. Highway noise barriers : new shapes. *Journal of Sound and Vibration*, 1980.
- [Sez80] R. Sez nec. Diffraction of sound around barriers : use of the boundary elements technique. *Journal of Sound and Vibration*, 1980.
- [Wat96] G.R. Watts. Acoustic performance of a multiple edge noise barrier profile at motorway sites. *Applied acoustics*, 1996.
- [WCH94] G.R. Watts, D.H. Crombie, and D.C. Hothersall. Acoustic performance of new designs of traffic noise barriers : full scale test. *Journal of Sound and Vibration*, 1994.
- [WM96] G.R. Watts and P.A. Morgan. Acoustic performance of an interference-type noise-barrier profile. *Applied acoustics*, 1996.



## Chapitre 11

# Les méthodes inverses en reconnaissance et géophysique au LCPC

### 11.1 Introduction

*Richard Lagabrielle*

Les quatre sections qui suivent concernent la géophysique appliquée au génie civil (reconnaissance des sites de construction, caractérisation du sous-sol ou auscultation d'ouvrages). Ils illustrent le rôle essentiel qu'y joue le problème inverse.

Il s'agit de décrire le sous-sol à travers la répartition d'un paramètre physique,  $m$ , caractéristique des matériaux : vitesse des ondes de compression ou de cisaillement (tomographie sismique), vitesse de l'onde de cisaillement (dispersion de l'onde de Rayleigh), conductivité et constante diélectrique (tomographie électromagnétique) ou résistivité (imagerie électrique). Ceci confère au problème inverse la propriété d'être fondamentalement sous-déterminé. En effet ces caractéristiques physiques sont définies en tout point du sous-sol, il y a donc une infinité d'inconnues (la valeur de la caractéristique étudiée en chaque point), alors que le nombre de données,  $d$ , constituées par les mesures d'une observable particulière (temps de trajet, signal sismique complet, amplitude du champ électromagnétique, potentiel électrique) sont en nombre fini. Ces problèmes inverses sont donc mal posés.

Pour les résoudre, on adopte la stratégie suivante :

1. Le paramètre  $m$ , fonction des points de l'espace, est d'abord échantillonné, le plus souvent en considérant qu'il reste constant sur un segment, un rectangle ou un parallélépipède (suivant qu'on s'intéresse à un problème à 1, 2 ou 3 dimensions) et ceci dans un domaine fini de l'espace. La fonction  $m$  est donc remplacée par un vecteur  $\mathbf{m}$  de dimension finie. Sa dimension n'a aucune raison d'être la même que le nombre de données, en général elle est plus grande et le problème reste sous-déterminé. On est donc sûr que si une solution au problème existe (c'est-à-dire, un vecteur  $\mathbf{m}$  qui explique les données  $d$ ), cette solution n'est pas unique.
2. On introduit des données supplémentaires sous forme de connaissance a priori sur  $m$ . Ces connaissances sont explicitement prises en compte dans le processus de résolution.

Par exemple, on peut connaître approximativement la valeur du paramètre  $\mathbf{m}$  : sa valeur est proche de  $\mathbf{m}_0$ . On impose alors que  $\mathbf{m} - \mathbf{m}_0$  soit petit. On peut aussi imposer que le paramètre varie lentement dans l'espace (gradient faible), on a alors  $m_{i+1} - m_i$  petit quelque soit  $i$  ( $i$  est le numéro de la composante de  $\mathbf{m}$ ). Il y a aussi les contraintes de non négativité (les vitesses des ondes, caractéristiques électromagnétiques, résistivité, sont des nombres positifs), ainsi pour les résistivités on peut choisir comme inconnue non pas la résistivité mais son logarithme, etc...

Une autre caractéristique fréquente des problèmes inverses en géophysique est qu'ils sont le plus souvent non-linéaires : la relation  $\mathbf{d} = f(\mathbf{m})$  n'est pas linéaire. La technique adoptée par les auteurs des articles est la linéarisation de la formule par un développement limité au voisinage d'un modèle initial, puis par un processus itératif on parvient éventuellement par modifications successives du modèle initial à un modèle estimé  $\mathbf{m}_e$  qui est une solution au problème inverse.

Cette technique appelle plusieurs remarques :

- elle implique la résolution d'un grand nombre de problèmes directs. Cette résolution peut être très lourde ; par exemple en tomographie sismique, il faut à chaque fois procéder à un nouveau tracé de rayons ; en imagerie électrique, on doit résoudre le problème direct à l'aide d'un programme aux éléments finis avec un très grand nombre d'éléments ; en tomographie électromagnétique, la résolution des équations de Maxwell dans un milieu où il faut tenir compte de la propagation dans l'air et dans le sol peut être très complexe...

On fait donc appel à une modélisation directe volontairement simplifiée : rayons droits ou circulaires en tomographie sismique, on néglige l'influence de la propagation dans l'air en tomographie électromagnétique ...

Cette simplification a naturellement une influence sur la résolution du problème inverse et incite à porter un regard critique sur les solutions obtenues.

- la linéarisation conduit au calcul des dérivées partielles des données par rapport aux paramètres qui définissent le modèle. Souvent ceci est assez simple si la relation  $\mathbf{d} = f(\mathbf{m})$  est simplement décrite. Cela peut devenir lourd et complexe dans d'autres cas, par exemple en tomographie électrique, ce calcul impliquerait de résoudre le problème direct deux fois par composante de  $\mathbf{m}$ , ce qui est pratiquement impossible, c'est pourquoi on y renonce et on a recours à la technique de l'état adjoint qui permet de déterminer beaucoup plus simplement la direction de descente à chaque itération.
- un processus itératif peut ne pas converger ; ceci peut naturellement provenir de la mauvaise qualité des mesures, mais cela peut aussi être le résultat d'une modélisation directe trop simplifiée ou inappropriée ; par exemple en tomographie sismique, dans une première interprétation, on considère que le milieu exploré est isotrope, si le processus itératif ne converge pas alors que les mesures sont de qualité, peut-être faut-il introduire dans le problème direct l'hypothèse d'anisotropie.

La non-unicité des solutions des problèmes inverses en géophysique a une autre conséquence importante : on ne peut pas trouver la solution du problème, mais plutôt un ensemble de solutions. Il s'agit donc alors de décrire cet ensemble de solutions et

d'en trouver les propriétés communes. Lorsque le processus converge vers une solution particulière, il faut être capable de porter un jugement sur la qualité de cette solution : quelle confiance peut-on accorder à la valeur trouvée pour tel ou tel paramètre du modèle ? Cette confiance peut être quantifiée en termes de résolution et d'erreur. Un exemple parlant est donné dans l'article qui traite de l'inversion de la dispersion des ondes de Rayleigh : la figure qui montre une solution particulière comporte en même temps des informations graphiques sur l'erreur et sur la résolution, ces deux dernières informations sont obtenues à partir des résultats de la dernière itération. Une autre information plus globale est constituée par le suivi de la convergence en termes de valeurs des résidus (écart entre les mesures réelles et celles que l'on obtiendrait théoriquement avec le modèle final estimé). L'article sur l'imagerie électrique montre de la même manière un graphique dans lequel on porte en abscisse les mesures et en ordonnée celles que l'on calcule dans le modèle solution : le nuage de points obtenu devrait idéalement former la première bissectrice du graphique, la déformation de ce nuage autour de la bissectrice renseigne sur la qualité de l'inversion. Dans ce dernier article on évoque aussi la technique ROI (pour "region of investigation"). Cette technique consiste à résoudre deux fois le problème inverse avec deux modèles initiaux différents et à comparer les solutions composante par composante du modèle : certaines composantes présentent des valeurs proches à la fin des deux processus itératifs, cela signifie que les mesures sont bien affectées par la valeur du paramètre dans cette région, et que l'on peut avoir confiance dans le résultat de l'inversion dans cette région ; d'autres composantes conservent la valeur qu'on leur a donnée dans le modèle initial. Ceci signifie que ces régions n'ont pas d'influence sur le résultat des mesures et l'on ne doit pas accorder de confiance au résultat de l'inversion pour ces composantes particulières du modèle.

Au lieu d'inverser des mesures réelles, on peut inverser des données simulées et calculer les critères de confiance que nous venons d'évoquer : cela permet d'optimiser la méthodologie d'auscultation ou un processus expérimental.

Ces caractéristiques du problème inverse en géophysique, et particulièrement en imagerie du sous-sol ou des ouvrages se retrouvent naturellement pour d'autres domaines d'application de l'inversion, elles sont générales pour les problèmes fondamentalement sous-déterminés.

## 11.2 Inversion de la dispersion des ondes de surface

Ludovic Bodet

### 11.2.1 Situation dans les problématiques du génie civil

Grâce aux méthodes sismiques il est possible d'obtenir des informations sur la vitesse de propagation des ondes de volume et donc sur les paramètres mécaniques des couches superficielles du sous-sol (modules de compression et de cisaillement). L'étude de la réfraction des ondes de compression ( $P$ ) permet de caractériser très facilement la géométrie et les vitesses de compression d'un milieu. Cependant, l'étude des ondes de cisaillement ( $S$ ) réfractées est plus délicate, car ces ondes sont difficiles à générer et à identifier sur ce type d'enregistrement. La vitesse de propagation des ondes  $S$  est pourtant un paramètre très intéressant puisque directement relié au module de cisaillement, qui intervient notamment dans les études sur le risque sismique ou dans le dimensionnement des fondations d'ouvrages. Il est ainsi souvent inévitable d'avoir recourt à des méthodes sismiques dites "invasives", c'est à dire appliquées à partir de forages (Cross-Hole, Down-Hole). Toutes ces techniques, développées parallèlement aux essais géotechniques pratiqués en laboratoire, sont aujourd'hui classiquement utilisées, tant en géophysique appliquée qu'en ingénierie géotechnique.

Les mesures de modules de cisaillement déterminées *in situ* par les géophysiciens, "en petites déformations", ne sont pas directement comparables aux modules requis "en grandes déformations" par les géotechniciens. Il semble donc plus judicieux de proposer, plutôt qu'un ensemble de valeurs absolues de modules, une représentation globale des variations des paramètres mécaniques du sol déduits des méthodes sismiques. Si cela est aisément réalisable en ondes  $P$ , cela devient beaucoup plus délicat et coûteux lorsqu'il s'agit des ondes  $S$  (par forages), tant et si bien qu'il est économiquement et pratiquement impensable de proposer une caractérisation globale du proche sous-sol à partir de telles méthodes.

Les déformations élastiques du sous-sol ne se propagent pas uniquement en ondes de volume. Une grande partie de l'énergie générée par une source sismique se propage en ondes de surface. Ces ondes, comme leur nom l'indique, se propagent le long de la surface de la Terre sans y pénétrer profondément. Si les caractéristiques du milieu parcouru sont variables, la vitesse de propagation de ces ondes est liée à leur longueur d'onde. Grâce à cette propriété, appelée dispersion, il est possible, à partir d'enregistrements classiques, de reconstruire un profil vertical (unidimensionnel) de vitesses de cisaillement (cf. figure 11.1) et d'atténuation intrinsèque du milieu étudié. Cette méthode d'analyse de la dispersion des ondes de surface est de plus en plus utilisée depuis les années 80 [SS04].

La prospection sismique par ondes de surface présente ainsi un avantage non négligeable sur les techniques en forage. Elle peut être facilement mise en œuvre dans le but d'effectuer un "zonage", ou "profilage" des paramètres mécaniques des couches superficielles du sous-sol. En effet, en déplaçant le même dispositif sismique (source-récepteurs) le long d'une "ligne", il est possible d'obtenir très rapidement une coupe –un profil deux dimensions (2D)– des vitesses de cisaillement des couches superficielles du milieu exploré le long de cette ligne.

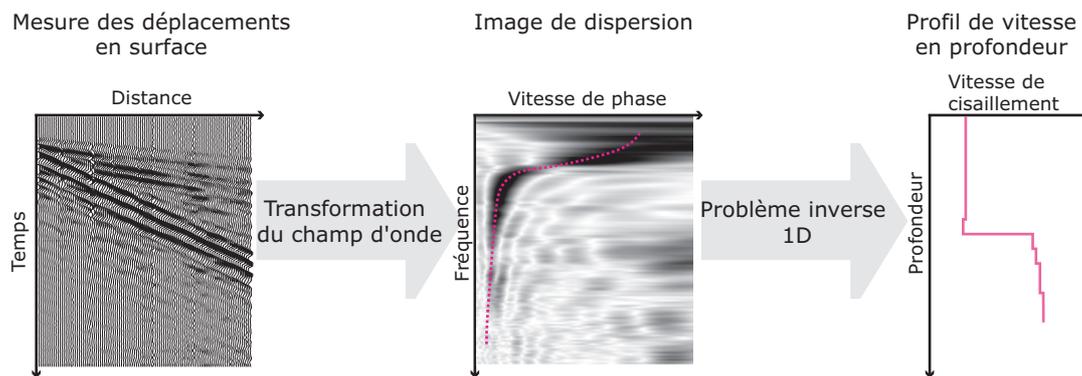


FIG. 11.1 – Schéma synthétique des différentes phases de la prospection par onde de surface

### 11.2.2 Exemple des ondes de Rayleigh

Si un milieu infini est divisé en deux milieux différents, il existe, en plus des ondes de volume ( $P$ ,  $SV$  et  $SH$ ) d'autres solutions de l'équation d'onde qui correspondent à des ondes se propageant parallèlement à la discontinuité et dont l'amplitude décroît avec la distance dans la direction perpendiculaire au plan de cette discontinuité. Lorsque celle-ci est une surface libre, ces ondes peuvent être séparées en deux types :

- les ondes de Rayleigh<sup>1</sup> qui sont polarisées dans un plan perpendiculaire à la surface libre ;
- les ondes de Love qui peuvent exister lorsqu'un milieu semi-infini est surmonté d'une couche d'épaisseur finie et de plus faible vitesse, elle même délimitée par une surface libre. Elles impliquent une déformation en cisaillement horizontal simple (perpendiculaire à la direction de propagation et parallèle à la surface libre).

D'autres types d'onde de la classe des ondes guidées existent dans le cas de géométries particulières. L'onde de Stoneley par exemple, se propage le long de l'interface séparant deux milieux semi-infinis. Lorsque l'un est liquide et l'autre solide cette onde devient l'onde de Scholte, d'ailleurs utilisée en prospection sous-marine, par exemple pour la localisation d'instabilités sédimentaires en milieu marin. Les ondes guidées par une plaque dans l'air, les ondes de Lamb, sont également étudiées dans le cadre de l'Évaluation Non Destructive (E.N.D.) ainsi que celui de la reconnaissance des matériaux du génie civil.

La propriété des ondes de surface la plus remarquable *a priori* est liée à leur contenu énergétique. Il a été montré que dans le cas de l'application d'une force en un point de la surface d'un demi-espace élastique homogène et isotrope, environ deux tiers de l'énergie transmise se propage sous forme d'ondes de surface. De plus, l'atténuation par expansion géométrique de ces ondes est, en surface, proportionnelle à  $1/\sqrt{r}$ ,  $r$  étant la distance à la source. Cette propriété rend les ondes de surface facilement identifiables, même à de grandes distances de la source. De plus, contrairement aux méthodes faisant appel aux ondes de volume, les ondes de surface permettent d'obtenir un bon rapport signal sur bruit sans employer obligatoirement une source extrêmement énergétique.

<sup>1</sup>Les ondes de Rayleigh se propagent le long de la surface de la Terre qui peut être considérée comme une surface libre compte tenu du contraste de densité et de paramètres élastiques entre l'air et la surface de la Terre.

Les capteurs mesurant la composante normale du déplacement étant les plus répandus, les ondes de Rayleigh (polarisées dans un plan perpendiculaire à la surface), sont naturellement les plus utilisées en prospection par onde de surface. Le mouvement particulaire associé au passage d'une onde de Rayleigh à la surface d'un demi-espace infini, élastique, homogène et isotrope est elliptique rétrograde. Ce mouvement résulte d'interférences constructives entre ondes  $P$  et  $SV$ . L'amplitude de la composante cisailante est la plus importante. Ce mouvement particulaire devient prograde et voit son amplitude diminuer rapidement avec la profondeur. Cette décroissance quasi-exponentielle du déplacement des particules avec la profondeur est telle que la partie du milieu qui est affectée par la propagation de l'onde est fonction de sa longueur d'onde. Dès lors, dans un demi-espace verticalement hétérogène, la vitesse des ondes de Rayleigh dépend de la fréquence, et vitesse de groupe et vitesse de phase sont différentes. Les ondes de Rayleigh présentent différents modes de propagation qui peuvent être interprétés par la présence d'interférences constructives entre raies courbes pour un milieu continûment hétérogène ou entre ondes transmises et réfléchies pour un milieu "multi-couches". Ils reflètent en général les variations de vitesse de cisaillement du milieu avec la profondeur. C'est généralement<sup>2</sup> le mode fondamental qui prédomine, mais la participation de chaque mode dépend de la stratification du sol autant que de la fréquence, ainsi que de la profondeur et de la nature de la source sismique. La vitesse de phase des ondes de Rayleigh se révèle par ailleurs étroitement liée à la vitesse de cisaillement (légèrement inférieure) et au coefficient de Poisson du milieu. Cette relation entre vitesse de phase des ondes de Rayleigh ( $V_R$ ) et vitesse des ondes de cisaillement ( $V_S$ ) peut être mise à profit pour l'évaluation des caractéristiques de cisaillement du milieu, à partir d'une évaluation de la dispersion.

### 11.2.3 Mesure de la dispersion

Pour les techniques de la géophysique appliquée, l'observable réellement mesurée n'est pas la dispersion mais le déplacement, la vitesse, ou l'accélération des particules de la surface du milieu exploré. L'acquisition multi-capteurs et haute résolution permet d'enregistrer, pour un intervalle de temps donné, la propagation des ondes à l'aplomb du dispositif expérimental.

#### A. Système d'acquisition

Différents systèmes d'acquisition sont disponibles pour l'application de la méthode d'inversion de la dispersion des ondes de surface. Les dispositifs purement linéaires et les sources impulsionnelles sont principalement utilisés en géophysique appliquée. Le type de source va dépendre du domaine d'application de la méthode. Dans le cadre de la géophysique appliquée et de la reconnaissance des sols, il existe notamment deux types de source dont le domaine fréquentiel est de l'ordre de l'Hertz (pour des profondeurs d'investigation de l'ordre de la dizaine de mètres) : les sources impulsionnelles (masse, chute de poids, explosion) et les sources vibratoires (pots vibrants, camions vibrateurs, *etc.*).

La méthode d'inversion de la dispersion des ondes de surface, peut être également utilisée sans source artificielle, c'est à dire en enregistrant le bruit de fond (méthodes dites "passives"). Cette approche permet d'avoir accès aux très basses fréquences et

<sup>2</sup>du moins en l'absence de variations brutales de la rigidité

d'augmenter ainsi la profondeur d'investigation. Par ailleurs, aux échelles de l'E.N.D., des transducteurs piézo-électriques et des sources laser sont par exemple utilisés afin d'enregistrer des signaux à des fréquences plus élevées (de l'ordre du mégaHertz).

Pour la géophysique appliquée comme pour l'E.N.D., la méthode fait intervenir des capteurs classiques tels que les géophones, accéléromètres (transducteurs piézo-électriques) ou même les vibromètres laser. Une importance particulière doit être donnée à leur fréquence de coupure basse. Dans le domaine de la prospection sismique par exemple, les signaux classiquement enregistrés (comme les ondes de volume) présentent un contenu fréquentiel plus haute fréquence que les ondes de surface. Les géophones de fréquence de coupure basse (très souvent 4,5 Hz) sont donc souvent préférables. Par ailleurs, la majorité des capteurs n'enregistrant que la composante du mouvement particulière perpendiculaire à la surface de mesure, la dispersion mesurée ne concerne que les ondes de Rayleigh. Si l'on désire utiliser la dispersion des ondes de Love<sup>3</sup>, des capteurs multi-composantes seront évidemment indispensables, bien qu'actuellement les déplacements horizontaux soient difficiles à générer et à mesurer.

### B. L'extraction de la dispersion

Quels que soient les sources ou capteurs utilisés, l'acquisition des données sous forme de section sismique fournit une description du milieu dans le plan *distance-temps* ( $x-t$ ). Alors qu'elle est difficilement détectable dans le plan ( $x-t$ ), la dispersion peut être clairement identifiée dans le plan *pulsation spatiale - pulsation temporelle* ( $k-\omega$ ). Il est donc nécessaire de transposer les enregistrements du domaine ( $x-t$ ) vers le domaine ( $k-\omega$ ). Ceci implique l'utilisation de techniques de traitement du signal appelées "transformation du champ d'ondes" (cf. exemple fig. 11.2). En général, ces techniques de transformation impliquent l'hypothèse que le train d'ondes observé peut être décomposé en une somme d'ondes planes. Celle-ci n'est valide que pour des distances source-récepteur supérieures à "plusieurs longueurs d'onde", c'est à dire en dehors du champ proche.

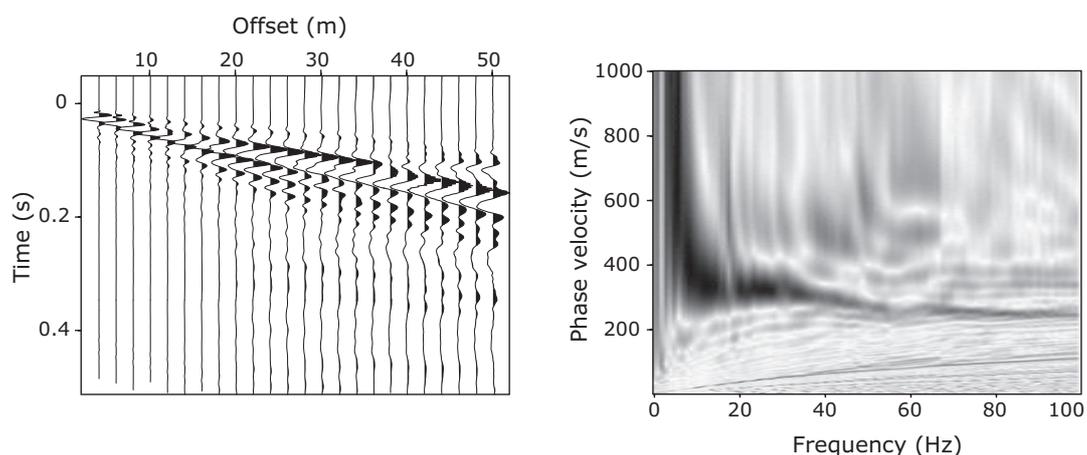


FIG. 11.2 – Sismogramme 24 traces [Fer03] généré par une chute de poids (à gauche). Image de dispersion normalisée (à droite) obtenue par transformation du champ d'onde (les maxima en noir correspondent aux modes de propagation des ondes de Rayleigh).

<sup>3</sup>Les ondes de Love sont classiquement analysées en sismologie afin de réduire l'indétermination du problème inverse.

### 11.2.4 Inversion de la dispersion des ondes de surface

#### A. Méthode employée

A partir des données de dispersion des ondes de surface et des techniques de modélisation semi-analytique uni-dimensionnelle (1D) de la dispersion, les méthodes inverses permettent d'estimer un modèle des vitesses de cisaillement du milieu exploré. Du point de vue pratique, les techniques dites "d'inversion linéaire", *d'optimisation locale*, sont relativement simple à mettre en oeuvre et peuvent se révéler rapides et efficaces quel que soit le domaine d'application. Cependant, la non-linéarité et la non-unicité du problème inverse impliquent la linéarisation de l'inversion et une importante difficulté d'interprétation des résultats. Les méthodes dites "de Monte Carlo", *d'optimisation globale*, proposent en revanche une représentation de l'espace des modèles possibles permettant une estimation du milieu en terme de probabilité d'existence. Selon les variantes employées, ces méthodes peuvent pallier les problèmes liés à la non-unicité et contournent "élégamment" la non-linéarité (algorithmes génétiques ; recuit simulé ; algorithme de voisinage ; cf. [SS04]). Elles sont cependant plus "lourdes" à mettre en oeuvre car elles nécessitent le calcul d'un grand nombre de modèles de l'espace des paramètres, ce qui du point de vue de la géophysique appliquée, diminue la rentabilité de la méthode de prospection par ondes de surface. Quelles que soient les techniques utilisées, l'inversion de la dispersion des ondes de Rayleigh présente de nombreuses limites quant à l'interprétation des résultats qu'elle fournit. L'hypothèse 1D du problème direct est notamment contraignante car rarement valide. De plus, une évidente indétermination de la profondeur d'investigation et de la résolution verticale rend délicate l'interprétation des résultats. Enfin, des erreurs de mesure telles que les *anomalies basse fréquence* ou les *confusions de modes* peuvent contaminer le résultat de l'inversion.

Dans des conditions d'applications "opérationnelles", l'interprétation de la dispersion est très souvent effectuée par inversion linéarisée. Le code utilisé en Reconnaissance et Géophysique au LCPC, proposé par Herrmann [Her02], en est un bon exemple (cf. fig.11.3). Dans le cas de l'inversion de la dispersion des ondes de surface, les données  $\mathbf{d} = [d_1; \dots; d_{N_d}]^T$ , sont représentées par les vitesses de phase et/ou de groupe mesurée en fonction de la période/fréquence. La relation de dispersion  $\mathbf{G}$ , reliant ces données aux paramètres  $\mathbf{p} = [p_1; \dots; p_{N_p}]^T$  d'un milieu uni-dimensionnelles (vitesses de cisaillement ( $V_S$ ) et de compression ( $V_P$ ) ainsi que les densités ( $\rho$ ) en fonction de la profondeur), peut être décrite par  $\mathbf{d} = \mathbf{G}\mathbf{p}$ .

Herrmann [Her02] formule, via une décomposition en valeurs propres de  $\mathbf{G}$  ( $\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^T$ ), la solution du problème de la manière suivante (moindres carrés amortis) :

$$\mathbf{p}_s = \mathbf{W}^{-1}\mathbf{V}[\mathbf{\Lambda}^2 + \gamma_{LM}^2\mathbf{I}]^{-1}\mathbf{\Lambda}\mathbf{U}^T\mathbf{d}, \quad (11.2.1)$$

$$\mathbf{C} = \mathbf{W}^{-1}\mathbf{V}[\mathbf{\Lambda}^2 + \gamma_{LM}^2\mathbf{I}]^{-1}\mathbf{\Lambda}^2[\mathbf{\Lambda}^2 + \gamma_{LM}^2\mathbf{I}]^{-1}\mathbf{V}^T(\mathbf{W}^{-1})^T, \quad (11.2.2)$$

$$\mathbf{R} = \mathbf{W}^{-1}\mathbf{V}[\mathbf{\Lambda}^2 + \gamma_{LM}^2\mathbf{I}]^{-1}\mathbf{\Lambda}^2\mathbf{V}^T\mathbf{W}. \quad (11.2.3)$$

où  $\mathbf{R}$  est la matrice de résolution et  $\mathbf{C}$  la matrice de covariance *a posteriori* indispensables à la caractérisation de la résolution (en profondeur) et de l'incertitude de chaque paramètre (écarts-type associés aux vitesses de cisaillement estimées) du modèle  $\mathbf{p}_s$ . Les critères de paramétrisation peuvent être contraints via la matrice de pondération  $\mathbf{W}$ .

**Régularisation.** Le facteur d'amortissement ( $\gamma_{LM}$ ) peut évoluer entre 0 et 10. Il est en général choisi grand au départ afin de faire intervenir les propriétés de convergence de la méthode de "la plus grande pente", pour être ensuite au fur et à mesure relaxé. Pour un facteur d'amortissement nul, l'inversion relève uniquement des moindres carrés. Ce paramètre contrôle donc les perturbations du modèle, agissant ainsi comme une contrainte sur l'espace des modèles.  $\gamma_{LM}$  est ici défini à partir des valeurs propres de la décomposition, inspectées aux cours de tests préalables.

**Paramètres estimés.** Ici, seule la dispersion des ondes de Rayleigh est inversée. Le rapport  $V_P/V_S$  de chaque couche est constant au cours de l'inversion et dépend du coefficient de Poisson des couches du modèle *a priori* moyen. A chaque itération,  $V_P$  et  $\rho$  sont empiriquement mis à jours pour chaque couche à partir de la vitesse de cisaillement estimée. La densité n'ayant pratiquement pas d'influence sur la dispersion, elle n'a pas de conséquences sur le profil de vitesses de cisaillement final qui seul sera considérée dans ce qui suit. La vitesse de compression influence très peu la dispersion des ondes de Rayleigh. Cependant, son effet apparaît à travers le coefficient de Poisson défini par le modèle *a priori* moyen.

**Modèle *a priori* moyen.** En l'absence d'information *a priori*, ce modèle est classiquement déterminé empiriquement, à partir de la courbe de dispersion. Le modèle initial définit également le nombre de couches du milieu et l'épaisseur de chaque couche, ainsi que la profondeur du demi-espace du milieu estimé.

**Estimation de la convergence.** Elle s'effectue à partir des résidus entre la courbe de dispersion du modèle estimé et les observations, calculés selon la norme  $L_1$ . Au cours des itérations, la convergence apparaît également à travers une diminution des perturbations moyennes du modèle.

**Représentation.** Les résultats de l'inversion sont présentés sous la forme d'un profil de vitesse de cisaillement en fonction de la profondeur. Pour chaque couche du modèle, la vitesse de cisaillement estimée est associée à son écart-type. Le modèle *a priori* moyen est superposé à ce profil car il est à prendre en compte lors de l'interprétation des résultats.

Les données sont en général également représentées (valeurs de vitesse et écart types associés en fonction de la fréquence ou de la période) afin de permettre leur comparaison avec la courbe de dispersion théorique calculée à partir du modèle obtenu. Les modèles et courbes de dispersion intermédiaires, obtenus à chaque itération, peuvent aussi être tracés.

La matrice de résolution doit être représentée car elle est définie telle que  $\mathbf{p}_s = \mathbf{R}\mathbf{p}$ , où  $\mathbf{p}_s$  est le modèle estimé et  $\mathbf{p}$  le véritable modèle. Ainsi, plus  $\mathbf{R}$  est proche de  $\mathbf{I}$  (matrice identité) meilleure est la solution du problème. Une représentation normalisée des valeurs de cette matrice permettra donc d'évaluer la résolution de chaque valeur de vitesse estimée.

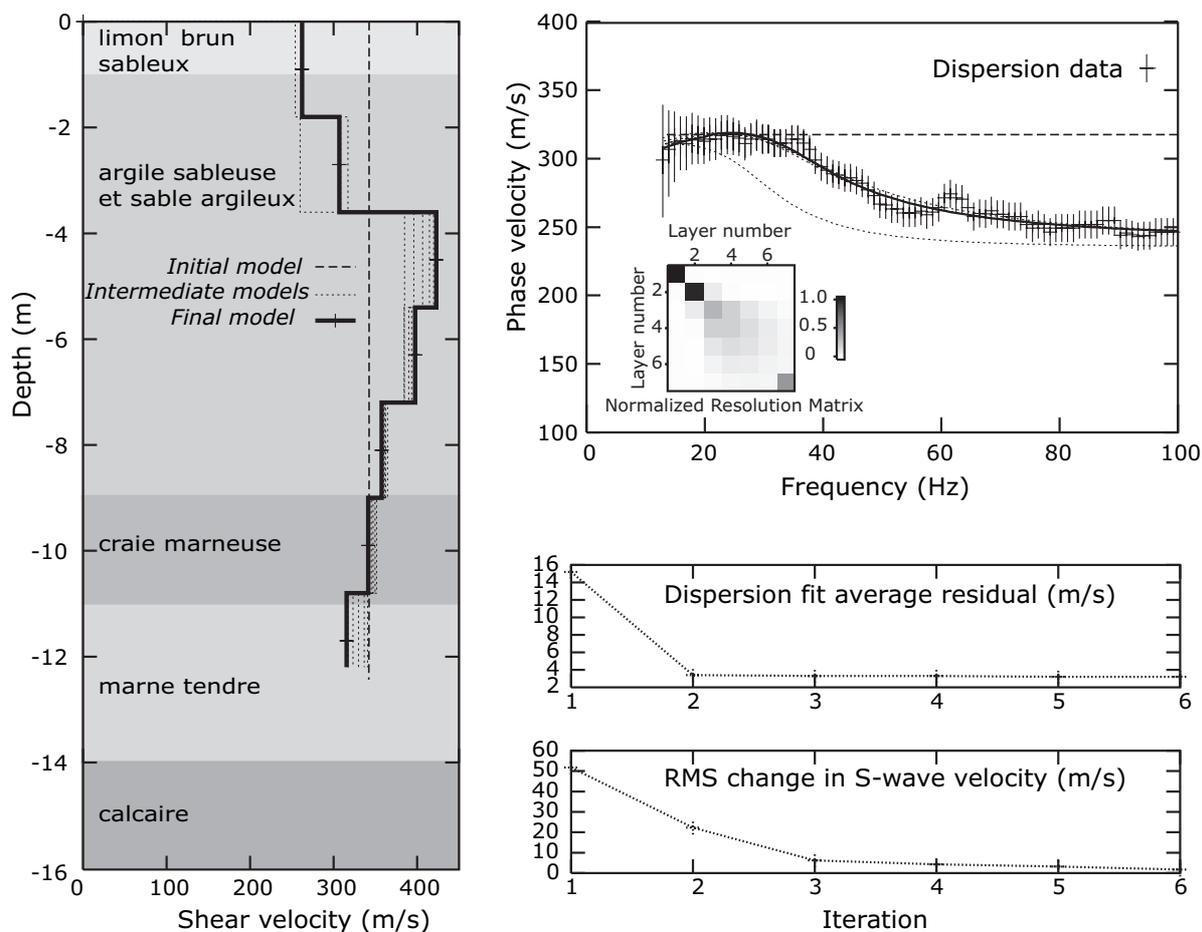


FIG. 11.3 – Résultats proposés pour l'interprétation des données de la figure 11.2. Comparaison à la coupe lithologique du site [Fer03] .

## B. Contrôle de la solution

La dispersion mesurée (effective) n'est pas toujours représentative de la dispersion théorique du milieu ausculté. Le dispositif utilisé, la nature des matériaux, et la géométrie du milieu peuvent générer des erreurs de mesure susceptibles de se propager dans le résultat de l'inversion linéarisée des courbes de dispersion. Les expériences numériques et physiques réalisées au cours de récentes études montrent que de telles erreurs d'interprétation peuvent en effet survenir à cause :

- d'anomalies basse fréquence dues à de faibles longueurs de dispositifs (confusion de modes, cf. [BvWB<sup>+</sup>05]) ;
- d'effets d'offsets proches dus à une incompatibilité entre les distances source-captteurs mises en œuvre et l'hypothèse d'ondes planes requise par les méthodes de traitement utilisées [Bod06] ;
- d'une éventuelle surévaluation de la profondeur d'investigation de la méthode et du manque d'information *a priori* sur le milieu exploré [Bod06] ;
- de géométries deux dimensions du milieu, incompatibles avec l'hypothèse 1D du problème inverse [BAB<sup>+</sup>04, BvWB<sup>+</sup>05, Bod06].

Ces inconvénients majeurs peuvent -doivent- être pris en compte afin de renforcer la fiabilité des résultats fournis. Ne pas les considérer conduit à une interprétation

erratique de la géométrie (1D) et des vitesses de cisaillement des matériaux auscultés. Ces limites sont identifiées dans le cadre d'une utilisation "ponctuelle" de l'inversion de la dispersion des ondes de surface. Des critères de mesure et d'interprétation sont proposés pour atténuer leurs effets, dans le cadre de l'utilisation d'une source active et d'un dispositif linéaire :

- utiliser une longueur d'onde maximale mesurable de la dispersion de l'ordre de 50 % de la longueur du dispositif (24 m dans le cas de l'exemple des figures 11.2 et 11.3) ;
- utiliser une profondeur maximale d'investigation de l'ordre de 25 % de la longueur du dispositif ( 12 m dans le cas de l'exemple des figures 11.2 et 11.3) ;
- inclure obligatoirement l'information *a priori* apportée par la sismique réfraction lorsque elle est disponible.

Enfin, la représentation de la résolution de l'inversion et du modèle *a priori* moyen doivent être associé au profil obtenu, permettant de visualiser la vraisemblance des résultats en fonction de la profondeur (cf. fig.11.3).

Ces critères suggèrent par ailleurs d'utiliser une telle méthode avec précaution, d'autant plus lorsque le but est de fournir une caractérisation globale du milieu en appliquant des méthodes de profilage.

## 11.3 Tomographies sismiques

*Philippe Côte*

### 11.3.1 Situation dans la problématique/génie civil

Les tomographies sismiques sont mises en œuvre pour diverses applications dans le domaine de la géophysique appliquée à la proche surface. L'objectif général de ces investigations est de produire une image, en règle générale à deux dimensions, de la répartition des matériaux et/ou de leurs propriétés dans le sous sol ou dans un ouvrage. Les dimensions et formes de la zone imagée sont déterminées par les conditions d'accessibilité pour l'installation des capteurs et sources nécessaires.

Les imageries envisagées visent à décrire d'une part des milieux naturels en interaction ou non avec des projets de construction. Le cas de la reconnaissance à l'avancement pour les tunnels ou micro-tunnels est un exemple relativement fréquent d'application des techniques tomographiques [MY96, CVH<sup>+</sup>00]. De même, l'estimation des profondeurs d'endommagement du massif suite à des tirs en tunnel ou en carrière fait appel régulièrement à des techniques d'imagerie [HC92],[ABSC98]. La reconnaissance des terrains au dessous de certains ouvrages constitue également un domaine d'application privilégié pour les tomographies en génie civil [JGF<sup>+</sup>01, PPMC05]. La surveillance des milieux naturels indépendamment de toute problématique constructive (surveillance des mouvements de terrain, volcanisme, ...) fait également largement appel aux techniques d'imagerie 2D ou 3D [JHD<sup>+</sup>00].

D'autre part, les ouvrages anthropiques font régulièrement l'objet d'investigations tomographiques. On peut citer le cas des barrages [KRBC95], des remblais [DR05], celui des piles ou piliers d'ouvrages d'art [CGPV92, BSZ03] ainsi que les piliers de mine ou de carrière souterraine [DA00, CDA05].

L'obtention d'une image tomographique repose sur trois étapes : l'acquisition des données ; la simulation de ces données dans un modèle synthétique (problème direct) et enfin le problème inverse à proprement parler ou la définition de solutions (modèles) explicitant au mieux les données expérimentales. Le problème inverse envisagé ici est non linéaire. La solution est généralement approchée par un processus itératif qui modifie « faiblement » à chaque étape le modèle en cours avant de re-simuler les données dans ce nouveau modèle. Le modèle initial, quant à lui, est défini à partir de connaissances déduites d'observations de terrain ou issues d'autres expérimentations (information a priori). Pour les applications envisagées ici, les données expérimentales sont les temps de trajet de l'onde sismique entre chaque point source et chaque point capteur. Le modèle est défini par les vitesses de l'onde considérée en chaque nœud d'un maillage englobant l'ensemble du dispositif expérimental.

### 11.3.2 Modélisation du problème direct

Il existe une multitude de façons d'envisager la résolution du problème direct. Il s'agit, ici, de calculer pour chaque couple source récepteur le temps de trajet le plus court dans un modèle donné. Les algorithmes envisageables sont nombreux et de complexité très variable. La manière la plus correcte physiquement consiste à résoudre l'équation d'Eikonal dans le milieu hétérogène par exemple par le biais d'une résolution en différences finies [Vid88]. Ces méthodes sont toutefois généralement trop

pénalisantes en termes de temps de calcul pour être intégrées de manière souple dans un processus itératif. En revanche ces modes de calcul permettent d'éviter les principaux écueils liés aux techniques de tracé de rais en milieux hétérogènes. En effet, les tracés de rais qui sont relativement légers en temps de calcul présentent une sensibilité importante aux hétérogénéités locales et les minima des temps de trajet sont parfois complexes à déterminer.

Les applications du domaine du génie civil visent à imager des milieux naturels ou anthropiques dotées de géométries internes relativement simples. Pour des investigations de piliers relativement homogènes par exemple, le tracé de rais le plus simple (rais rectilignes) donne pleine satisfaction. En revanche, dès que les contrastes internes augmentent, il devient impératif de prendre en compte le cheminement non rectiligne de l'énergie sismique. Les milieux que nous souhaitons imager présentent en général des contrastes relativement doux (altération ou densification progressive des matériaux) aussi une approximation de la distribution des vitesses en gradient homogène est justifiée. Dans un tel milieu le tracé est circulaire, reste analytique et représente une approximation du second ordre (premier ordre = rais rectilignes). Le fait d'imposer cette forme aux rais rend le problème direct relativement peu sensible à de fortes hétérogénéités locales et induit une sorte de lissage spatial en accord avec la nature itérative du processus global linéarisé [Côt88].

La détermination du rai connectant une source à un capteur donné est très simple et ne dépend que d'un seul paramètre. L'espace de variation plausible de ce paramètre est donc exploré de manière exhaustive. La géométrie retenue pour le rai est celle conduisant au temps de trajet le plus court. La validité de cette approximation analytique a été testée pour différentes configurations incluant de très forts contrastes.

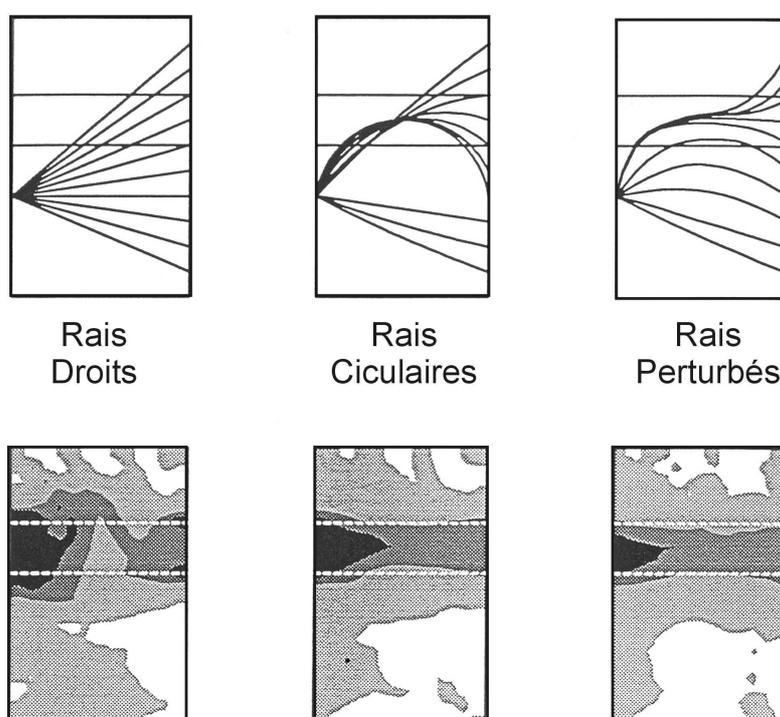


FIG. 11.4 – Comparaison des effets de différents tracés de rais sur le résultat de l'inversion

Les résultats illustrés sur la Figure 11.4 sont issus d'un modèle de terrain synthétique homogène (1000 m/s) délimité par deux forages et traversé par un banc rapide (2000 m/s). Une douzaine de sites de réception sont répartis entre la surface et le forage de droite. Une source est activée pour une dizaine de positions dans le forage de gauche. Les données mesurées sont analytiques, l'effet du tracé de rai peut être testé au sein de processus inverses similaires. Le résultat indique clairement qu'il est impératif d'utiliser des rais non-rectilignes. En effet, l'image du banc rapide n'est pas identifiée avec le tracé de rais droits. En revanche, la comparaison des résultats issus d'une part des rais analytiques circulaires tels que décrits plus haut et d'autre part de rais «perturbés» plus proches de la réalité physique [FM87] montre que l'accroissement d'information lié à la complexification du tracé de rais n'est pas évidente. Cet exemple très simple illustre donc la nature des choix qui doivent guider ces applications inverses dans l'auscultation pour le génie civil. Les résolutions escomptées sont en général assez faibles. La plupart du temps ceci est directement lié à la taille importante des longueurs d'onde (au regard des dimensions des dispositifs) qu'il est possible de capter. Aussi, la complexification du problème direct au delà d'un certain seuil n'apporte pas d'information pertinente. Cette remarque vaut également pour le problème inverse.

### 11.3.3 Résolution du problème inverse

La phase d'acquisition permet de définir un vecteur  $d$  qui est composé ici des temps de trajet pour l'ensemble des couples source-récepteur activés. L'espace des inconnues est décrit par un vecteur  $m$  composé ici des lenteurs sur la grille de discrétisation du milieu. La fonctionnelle reliant les données aux modèle, ici la répartition des longueurs de propagation dans chaque élément discret, peut être décrite sous la forme d'une matrice  $G$ . La linéarisation revient à décrire la relation à proximité d'un espace modèle donné. Le vecteur  $d$  est alors composé des résidus temporels et  $m$  des variations de lenteurs. La résolution de ce type de système matriciel a été largement étudiée dans le domaine de la géophysique depuis les années 60. Les influences respectives du type de norme minimisée, de l'incidence des incertitudes sur le processus, de la dilution des erreurs, etc , ont été décrits par de très nombreux auteurs [BG67, TV82].

Pour être en mesure de traiter rapidement des systèmes comportant un très grand nombre de données, nous nous sommes tournés vers des méthodes de résolution géométriques du système décrit précédemment. Il a été montré que les méthodes de type SIRT - Simultaneous Iterative Reconstruction Techniques - [Gil72] conduisaient, à la convergence, vers des solutions identiques à celles déduites de solutions aux moindres carrés du système matriciel. Ces méthodes géométriques itératives consistent à balayer l'espace modèle en sélectionnant pour chaque maille les rais la traversant et en déduisant une modification de la lenteur locale fonction d'une combinaison des résidus des rais sélectionnés. Quand tout l'espace a été modifié, un nouveau tracé de rais est lancé, il définit un nouveau vecteur résidus, une nouvelle itération peut débuter. Ce processus est ré-itéré tant qu'un critère de convergence n'est pas satisfait.

En dehors du tracé de rais courbes analytique, les principaux aménagements apportés à la technique de base type SIRT pour définir notre algorithme RAI-2D portent sur le critère de sélection des rais autour d'un nœud du maillage. En effet, plutôt que sélectionner ceux traversant la maille en cours d'examen, nous avons défini un disque d'influence [Côt88] dont le rayon est un des paramètres de l'inversion. Ce paramètre agit alors à la fois comme un critère de lissage de l'inversion (analogue

aux poids des moindres carrés pondérés) et comme un critère physique élargissant la zone d'interaction du rai tout au long de sa propagation [CS92]. Différents critères statistiques de contrôle du process permettent de surveiller et de critiquer l'inversion. Par ailleurs, l'extension à 3D et la prise en compte de l'anisotropie ont été introduites [ABSC98, AAGG00].

### 11.3.4 Exemple d'application in situ

La planification de travaux de confortement d'une carrière souterraine occupée a amené le gestionnaire à la question du diagnostic comparé de ses piliers. Cette carrière est creusée dans une couche de gypse assez homogène et s'étend sur une superficie de quelques hectares. L'examen visuel des piliers et l'estimation de leur chargement constituent la première phase de ce diagnostic. Une deuxième phase dédiée à l'exploration de l'état interne des piliers a été souhaitée [CDA05].

Pour ce faire, différents tests géophysiques ont été réalisés sur quelques piliers d'essais. Ces tests ont permis de bâtir une méthodologie d'exploration. Pour chaque pilier, elle consiste à coupler des reconstructions en tomographie sismique et des profils radars. Des indicateurs d'état sont ensuite extraits. Ils peuvent mettre en évidence le comportement particulier de certains piliers ou certaines régions. Ils conduisent par ailleurs à une classification ordonnée de l'ensemble des piliers auscultés.

La taille moyenne des piliers est de l'ordre de  $7 \times 7 \times 7m$ . Deux tomographies sismiques en temps de trajet sont réalisées pour chaque pilier. Une première tomographie verticale permet de détecter un éventuel comportement hétérogène suivant la composante verticale et de valider l'implantation de la tomographie horizontale (Cf. Figure 11.5). En effet, l'altitude choisie pour la tomographie horizontale ne doit pas correspondre à un niveau bas pour les vitesses sismiques, auquel cas, l'interprétation en coupe 2D des mesures tomographiques horizontales serait erronée.

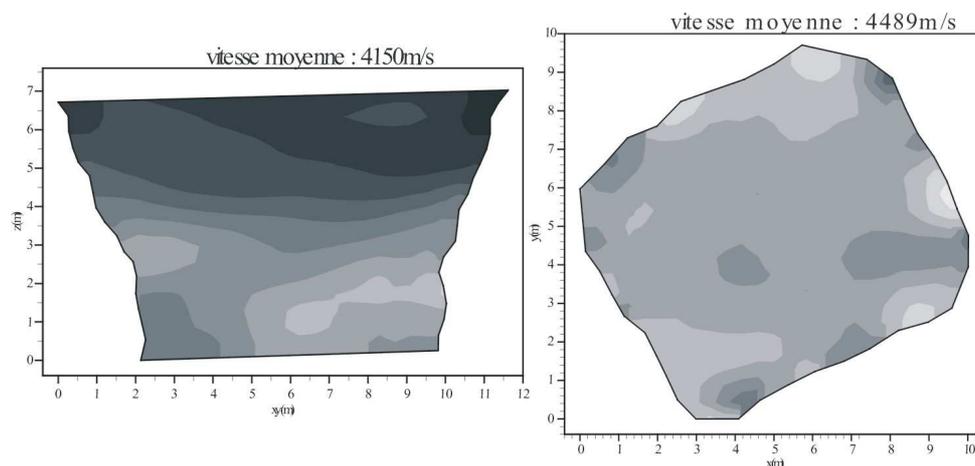


FIG. 11.5 – Coupe tomographiques d'un pilier test (à gauche : coupe verticale ; à droite coupe horizontale). La gamme des vitesses représentée s'étend linéairement du noir : 3500 m/s au blanc : 5000 m/s. La coupe horizontale est réalisée à une hauteur de l'ordre de 1,7 m sur la coupe verticale.

Différents tests ont été réalisés pour définir les paramètres tomographiques optimaux et compatibles avec l'auscultation de plusieurs dizaines de piliers dans des

délais raisonnables (pas et type de capteurs, pas et type de source, ... ). Une machine d'acquisition spécifique (32 voies, 1,2 MHz/voie, 12 bits réels) a été construite de manière à minimiser le temps d'acquisition et à inclure dès la prise de mesure le conditionnement des signaux dédié aux pointés nécessaires pour les tomographies.

Des expérimentations radar ont été conduites en parallèle sur les mêmes piliers tests. Ces mesures ont permis de définir la méthodologie appropriée pour visualiser l'état de fracturation interne de ces piliers. Les fractures ainsi détectées peuvent être portées sur les coupes tomographiques.

La méthodologie géophysique sélectionnée permet donc d'obtenir des coupes tomographiques rehaussées par les fractures détectées. Les techniques utilisées autorisent des mesures à grand rendement. Il a donc été possible de visualiser comparativement les résultats d'un grand nombre de piliers. Un extrait de ces résultats est présenté sur la Figure 11.6.

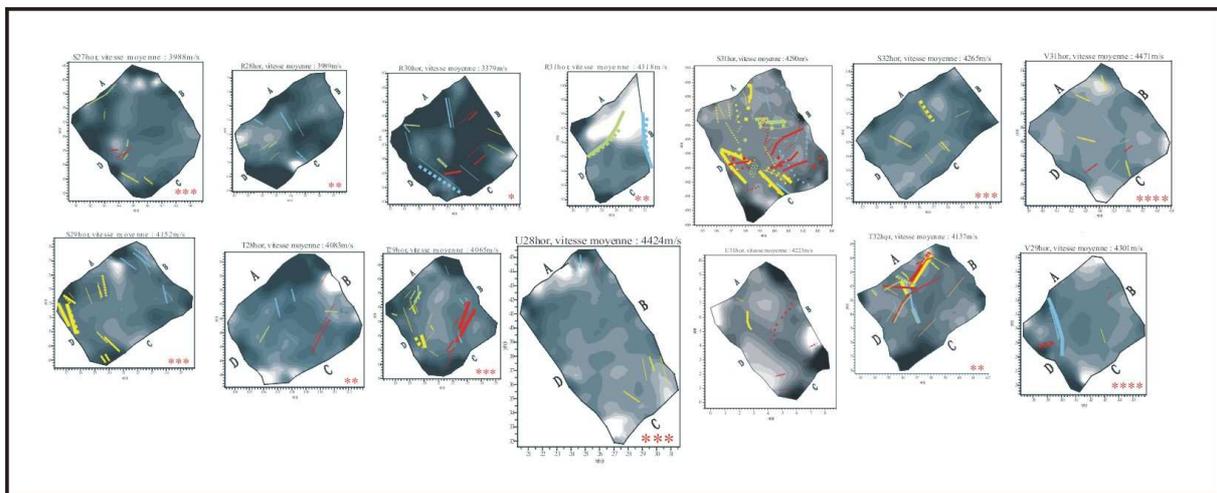


FIG. 11.6 – Extrait de la collection de coupes tomographiques horizontales et des détections radar de fractures (le positionnement des piliers n'est pas représentatif).

Différents indicateurs peuvent être extraits de ces représentations. Visuellement, des piliers ou des régions « atypiques » sont repérés, il est possible par ailleurs de quantifier l'homogénéité de chaque pilier, son degré de fissuration, ... L'ensemble de ces indicateurs conduit à un classement ordonné des piliers auscultés en complément des indicateurs d'observation visuelle et de chargement. Il sert de base au gestionnaire pour la planification des travaux de confortement.

## 11.4 Tomographie électromagnétique 2D entre forages

*Xavier Dérobert*

### 11.4.1 Nature du problème inverse

Les méthodes géophysiques fondées sur l'étude des ondes électromagnétiques (EM) aux fréquences radar pour l'imagerie du sous-sol ont fait l'objet d'un intérêt considérable au long de la dernière décennie. Comme le radar est fréquemment limité en termes de profondeur de pénétration, des techniques tomographiques entre forages ont été développées, visant à produire des représentations en coupes ou en tranches du milieu exploré.

Les techniques tomographiques bi-dimensionnelles (2D) se positionnent dans deux grandes catégories. La première catégorie fait référence aux techniques les plus connues et commercialement développées, qui s'appuient sur les temps de trajet des ondes radar ou sur leur atténuation [CDLL95, VZR99]. Toutefois, leur résolution spatiale reste faible, et les résultats numériques présentent des artéfacts conséquents à l'application de la théorie des rais qui ne prend pas en compte les phénomènes de diffraction.

La seconde catégorie, dont fait partie la technique développée ici, implique la prise en compte de ces phénomènes dans le processus d'imagerie. Diverses études électromagnétiques ont en effet montré que le problème d'inversion des champs EM diffractés (problème mal posé et fortement non linéaire) pouvait être résolu grâce à un apport d'information supplémentaire. Entre forages, les techniques s'orientent vers l'illumination de l'objet (ou hétérogénéité) à plusieurs fréquences pour augmenter l'information malgré les limitations en multi-incidences [Dou97, BKP97].

La technique d'inversion, présentée ici, fait partie d'une étude plus générale menée par R. Tarnus [Tar04] sur une technique de tomographie par diffraction dans le domaine spectral, développée pour des mesures EM en transmission entre forages, à plusieurs fréquences.

Deux modes d'illumination EM de la zone auscultée peuvent être utilisés : le champ électrique peut se propager dans le plan d'auscultation soit de manière perpendiculaire à ce dernier, comme un scalaire – mode Transverse-Magnétique (TM) –, soit de manière parallèle à ce plan – mode Transverse Electric (TE) – comme dans le cas de mesures en forages par des antennes dipôles, par exemple. Seul le mode TM, le plus simple à modéliser, sera présenté pour valider numériquement la technique d'inversion, équivalente pour les deux modes, et quantifier le niveau de résolution auquel s'attendre.

### 11.4.2 Résolution du problème inverse

Soit un milieu homogène semi-infini caractérisé par une constante diélectrique  $\varepsilon(\mathbf{r})$  et une conductivité électrique  $\sigma(\mathbf{r})$ . Une hétérogénéité, infinie suivant l'axe perpendiculaire au plan d'auscultation et caractérisé par les propriétés EM  $\varepsilon_h(\mathbf{r})$  et  $\sigma_h(\mathbf{r})$ , est enfouie entre deux forages caractérisant nos deux lignes de mesures (figure 11.7). On considère que les caractéristiques EM sont constantes sur la bande de fréquence étudiée, et que l'interface sol-air est suffisamment éloignée de la zone auscultée pour n'avoir aucune incidence sur les champs électriques diffractés.

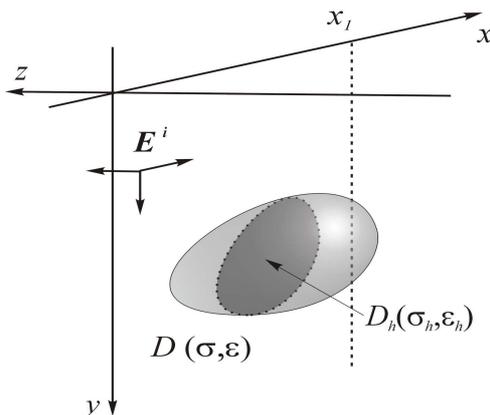


FIG. 11.7 – Géométrie du problème 3D

L'objet est éclairé par des champs électriques incidents  $E_z^I(\mathbf{r}, t)$ , générant un champ diffracté  $E_z^D(\mathbf{r}, t)$  secondaire dans l'hétérogénéité.

Le champ électrique total, mesuré au niveau de la ligne de réception ( $x = x_1$ ), est alors défini par :

$$E_z(\mathbf{r}, t) = E_z^I(\mathbf{r}, t) + E_z^D(\mathbf{r}, t) \quad (11.4.1)$$

Après l'application d'une transformée de Fourier (TF) dans le domaine temps sur les équations de propagation du champ électrique, le champ diffracté peut être présenté dans sa forme intégrale, pour chaque fréquence de pulsation  $\omega$  :

$$\tilde{E}_z^D(\mathbf{r}, \omega) = \int_{D_d} G(\mathbf{r}, \mathbf{r}', \omega) J(\mathbf{r}', \omega) d\mathbf{r}' \quad \forall \mathbf{r}, \forall \omega, \quad (11.4.2)$$

où  $G(\mathbf{r}, \mathbf{r}', \omega)$  est le tenseur de Green et  $J(\mathbf{r}', \omega)$  les courants induits dans l'hétérogénéité par le champ incident  $\tilde{E}_z^I(\mathbf{r}, \omega)$ , et défini par :

$$J(\mathbf{r}, \omega) = [k_h^2(\mathbf{r}, \omega) - k^2(\mathbf{r}, \omega)] \tilde{E}_z^I(\mathbf{r}, \omega) \quad (11.4.3)$$

avec

$$k^2(\mathbf{r}, \omega) = \omega^2 \mu_0 \left( \varepsilon(\mathbf{r}) + j \frac{\sigma(\mathbf{r})}{\omega} \right) \quad (11.4.4)$$

et  $k_h^2(\mathbf{r}, \omega)$  défini de la même manière que  $k^2(\mathbf{r}, \omega)$  en fonction de  $\varepsilon_h(\mathbf{r})$  et  $\sigma_h(\mathbf{r})$ . Pour des raisons de simplification d'écriture,  $k^2(\mathbf{r}, \omega)$  sera noté  $k^2(\omega)$  dans la suite du texte.

L'objectif de la technique d'inversion consiste à reconstruire une tomographie des courants induits normalisés plutôt que des caractéristiques EM des hétérogénéités présentes entre les forages. En effet, on peut trouver dans le domaine spectral une relation simple liant ces courants induits aux champs diffractés, qui nous permet de reconstruire une carte, ou tomographie, des courants induits dans le domaine spatial.

Ces courants induits, normalisés par le champ incident, sont définis par :

$$K(\mathbf{r}, \omega) = \frac{J(\mathbf{r}, \omega)}{k^2(\omega) \tilde{E}_z^I(\mathbf{r}, \omega)} \quad \forall \mathbf{r}, \forall \omega, \quad (11.4.5)$$

La démarche consiste donc à calculer, dans un premier temps, le terme  $\tilde{E}_z^I(\mathbf{r}, \omega)$  dans le domaine spectral, à partir de  $E_z^I(\mathbf{r}, t)$ , qui dépend de la position  $y_s$  de la source

le long de la ligne de mesure ( $x = 0$ ). Cette décomposition est présentée par Dourthe (1997) et Tarnus (2004).

En définissant la Transformée de Fourier (TF) spatiale 2D des courants  $\tilde{K}(\alpha, \beta)$  et la TF spatiale 2D des champs diffractés  $\tilde{E}_z^D(x, \nu, \eta)$ , les équations (11.4.2) et (11.4.5) conduisent à la relation spectrale, pour chaque fréquence angulaire  $\omega$  :

$$\tilde{K}(\alpha, \beta, \mathbf{0}, \omega) = \frac{-2i\gamma(\nu, \omega)}{k^2(\omega)A_z(\eta, \omega)} e^{i\gamma(\nu, \omega)x_L} \tilde{E}_z^D(x_L, \nu, \eta, \omega) \quad (11.4.6)$$

$$\text{avec} \quad \alpha(\nu, \eta) = \nu - \eta$$

$$\beta(\nu, \eta) = \frac{1}{2\pi} (\gamma(\nu, \omega) - \gamma(\eta, \omega))$$

L'équation (11.4.6) a été validée par Dourthe (1997). Pour aboutir à ce résultat, nous avons dû considérer que le terme  $\tilde{K}(\alpha, \beta)$  était indépendant de la variable spectrale  $\eta'$  (elle-même reliée à la variable  $(y' - y_s)$ , position relative de l'émission). Cette approximation peut être appliquée pour des configurations particulières restreintes de mesures (lignes rectilignes, lignes parallèles en transmission et en champ lointain). Une Transformée de Fourier Inverse permet alors de remonter à une image des courants induits normalisés dans le sous-sol (ou domaine spatial).

La démarche sous-jacente à cette technique d'inversion consiste à maximiser l'information dans le domaine spectral afin d'améliorer la résolution et de diminuer le niveau de bruit des tomographies dans le domaine spatial (c'est-à-dire limiter le niveau d'amplitude de courants induits mal positionnés correspondant à des artefacts dans les tomographies).

Pour cela, il est nécessaire de disposer de plusieurs points de mesures en émission, comme en réception, et d'utiliser plusieurs fréquences d'éclairage en multi-incidences.

### 11.4.3 Illustration et validation pour une inversion 2D-TM

La validation numérique du modèle d'inversion 2D-TM est réalisée à partir d'expérimentations sur modèle physique, dans une cuve d'eau de deux mètres de hauteur et trois mètres de diamètre (figure 11.8). Deux tubes PVC verticaux, séparés de 44 cm, font office de guides pour des lignes de mesure d'un mètre.

L'eau présente le double avantage d'être un milieu homogène et de constante diélectrique telle que les ondes sont contractées ( $\epsilon'_r = 80$  et  $\sigma = 0,006$  S/m), permettant de simuler une application géophysique entre forages dans des dimensions réduites. Nous avons donc conçu des antennes dipôles de 8 cm pour la bande de fréquence de [400–900 MHz]. L'échantillonnage retenu est de 2 cm entre points de mesures, pour 51 points par ligne de mesure. Des tubes PVC immergés, et remplis d'air, simulent les cavités à détecter.

Les mesures sont effectuées en deux temps, avec et sans tube, de façon à pouvoir déduire les champs diffractés à partir des champs incidents (sans tubes) et des champs totaux (avec tubes).

La figure 11.9 montre deux exemples de résultats obtenus à partir de peu de fréquences échantillonnées dans une bande étroite (9 fréquences dans la bande [400-480 MHz]), compatible avec les bandes passantes des antennes. Il est à noter que dans les présentations de reconstructions tomographiques, les lignes verticales en pointillé correspondent aux positionnement des forages.

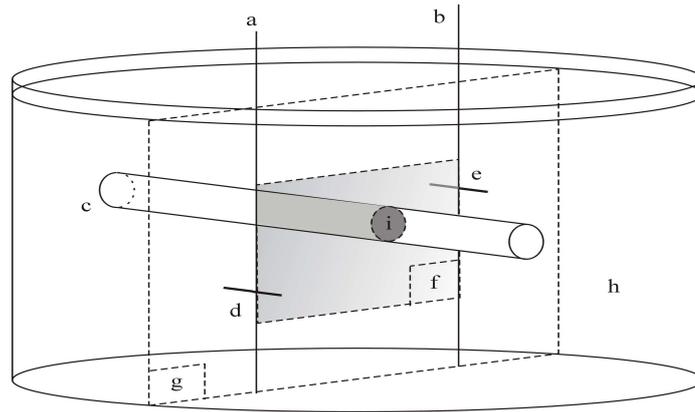


FIG. 11.8 – Modèle physique (a-b : lignes de mesure en forage, d-e : antennes en mode TM, c-i : hétérogénéité en PVC et section à imager, f-g : plan de mesure et zone d'investigation, h : eau)

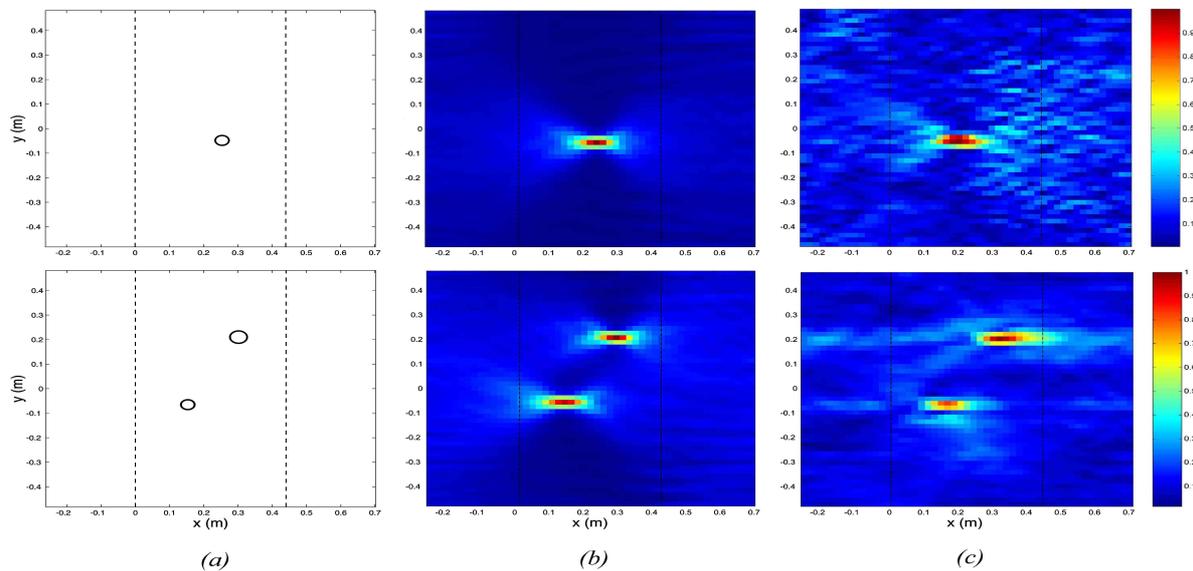


FIG. 11.9 – Inversions numérique et expérimentale (9 fréquences dans la bande [400-480 MHz]) de cavités cylindriques de 4 et 5 cm de diamètre. a) configurations, b) reconstruction de données numériques, c) reconstruction de données expérimentales.

Les contraintes géométriques engendrent un faible bruit autour de la position des tubes, principalement dans la direction des  $x$ . Ceci induit une précision sur les mesures de l'ordre de  $\lambda_c/2$  (où  $\lambda_c$  est la longueur d'onde de la fréquence centrale) dans la direction des  $x$ , alors qu'elle est de l'ordre de  $\lambda_c/4$  le long des  $y$ . Cela correspond à une précision de l'ordre du pas, ou du double, du pas d'échantillonnage dans la direction des  $y$ , ou des  $x$ .

Doubler le nombre de fréquences et augmenter le nombre de points de mesures diminue le niveau de bruit et améliore la résolution des tomographies, principalement dans le cas de cavités isolées proches de la ligne d'émission. On peut alors atteindre des niveaux de résolution de l'ordre de  $\lambda_c/4$  (ou du pas d'échantillonnage), bien que le

niveau de bruit soit toujours supérieur dans la direction des  $x$ .

On peut donc noter que de simples antennes dipôles, à bande passante étroite, combinées à peu de fréquences d'éclairement suffisent à la réalisation d'inversions tomographiques EM précises, pour un temps de calcul très faible (moins d'une minute pour un PC quelconque, inférieur à la seconde pour un modèle récent).

## 11.5 Tomographies de résistivité électrique de structures 3D de géométrie quelconque

*Laurent Marescot,  
Sérgio Palma Lopes*

### 11.5.1 Rappels et problématique en génie civil

#### A. L'imagerie de résistivité électrique

La résistivité électrique, notée  $\rho$  et exprimée en Ohm.mètres ( $\Omega\text{m}$ ), est une propriété locale d'un milieu continu qui mesure sa capacité à résister au passage d'un courant électrique (son inverse est la conductivité électrique, notée  $\sigma$  et exprimée en Siemens par mètre, soit  $\text{Sm}^{-1}$ ). La connaissance de la répartition de cette propriété peut permettre de remonter à la nature du milieu ausculté, à sa structure ou à son état. Pour imager cette répartition, on doit réaliser une tomographie électrique du milieu, c'est à dire un ensemble de mesures sur ou autour de ce milieu.

Les mesures consistent à générer un champ électrique statique dans le milieu, en y injectant du courant continu par l'intermédiaire d'électrodes ponctuelles, et à mesurer la répartition spatiale du potentiel électrique induit par la circulation de courant, en mesurant des différences de potentiel à l'aide d'autres électrodes.

Puis, par inversion des données ainsi mesurées, on tire un ou plusieurs modèles plausibles donnant les valeurs de résistivité dans le milieu, que l'on interprète ensuite selon l'application : description d'une coupe géologique, suivi d'une invasion d'eau, de sel ou de polluants, recherche d'une cavité ou d'un défaut local dans un ouvrage.

#### B. Formulation du problème direct

La répartition du potentiel électrique, dans un milieu conducteur ayant une répartition  $\sigma$  de conductivité électrique et soumis à des sources de courant continu, est un problème physique dit de diffusion. Il est régi par l'équation de Poisson (équation 11.5.1) à laquelle s'ajoutent des conditions aux limites. Il s'agit de trouver le champ scalaire  $V$  vérifiant en tout point de l'espace :

$$-\nabla \cdot (\sigma \nabla V) = \sum_s I_s \delta(r - r_s) \quad (11.5.1)$$

où  $\delta$  est une distribution de Dirac à trois dimensions et  $r$  la position d'un point de l'espace. Le terme de droite de l'équation 11.5.1 est le terme de source de courant, exprimé ici comme la superposition de sources ponctuelles placées aux points  $r_s$  et d'intensités de courant  $I_s$  (en Ampères). Les conditions aux limites sur la frontière du domaine étudié peuvent être de type « flux de courant imposé » (par exemple une condition d'isolation électrique) ou de type « potentiel imposé » (par exemple une condition de mise à la terre).

Notons que les relations permettant de passer d'une répartition donnée de la conductivité  $\sigma$  aux réponses électriques  $V$  du milieu correspondant ne sont pas linéaires. En l'absence d'une solution analytique, ce problème direct peut être résolu par une méthode numérique qui sera le « solveur » autour duquel l'algorithme itératif d'inversion est développé. Dans le cas de géométries tri-dimensionnelles (3D) quelconques, la méthode des éléments finis (MEF) est bien adaptée.

### C. Nature du problème inverse

En tomographie électrique, les observables sont les différences de potentiel mesurées (généralement normalisées en résistances et éventuellement en « résistivités apparentes »). Ces données mesurées peuvent être regroupées dans un vecteur noté  $d$ . Une paramétrisation du modèle, c'est à dire une représentation mathématique simplifiée de la réalité, doit être choisie. En 2D et 3D, s'il l'on a peu d'*a priori* sur la répartition de la conductivité, il est très fréquent que le milieu soit entièrement discrétisé en cellules (les éléments finis, dans le cas de la MEF) de sorte que la conductivité puisse être considérée comme constante à l'intérieur de chacun de ces volumes élémentaires. Les paramètres du modèle sont ces valeurs de conductivité et sont regroupés dans un vecteur noté  $m$ .

Notons  $g(m)$  les réponses électriques du modèle, que l'on évalue en résolvant le problème direct. Le problème inverse consiste à reconstruire un modèle  $m$  du milieu ausculté tel que ses réponses  $g(m)$  soient aussi proches que possible des réponses mesurées : c'est donc un problème d'optimisation dans lequel on cherche à minimiser l'écart  $d - g(m)$ . L'opérateur  $g$  n'étant pas linéaire, le problème inverse est dit *non-linéaire*. Une des stratégies les plus communes consiste à *linéariser* le problème et à approcher la solution par itérations successives.

Avec la paramétrisation en cellules décrite précédemment, le nombre de paramètres inconnus est le plus souvent bien supérieur au nombre de données mesurées. De plus, la tomographie peut n'apporter que très peu d'information sur certaines zones du modèle éloignées des dispositifs de mesure. Le problème est donc *sous-déterminé*. Par ailleurs, à cause du caractère intégrant des méthodes électriques (phénomène de diffusion) des « équivalences électriques » sont possibles : il est difficile de distinguer les effets du volume d'une hétérogénéité des effets du contraste de résistivité de cette hétérogénéité avec l'encaissant. Plusieurs modèles d'un milieu peuvent donc aboutir approximativement aux mêmes réponses électriques. Ce phénomène a des conséquences directes sur la résolution globale de la méthode, comme on le verra dans l'exemple traité plus loin. Enfin, les mesures sont entachées d'erreurs expérimentales et celles-ci se propagent inévitablement vers les paramètres ajustés. L'ensemble de ces raisons explique la *non-unicité* de la solution au problème inverse.

En général, il faudra donc considérer qu'il n'y a pas une solution mais un espace des modèles plausibles. Toute information *a priori* sur le milieu (observations de toute nature) sera utile au choix d'une solution, et une méthode d'évaluation de cette solution sera nécessaire.

### D. Problématique en génie civil

La majorité des codes d'inversion de mesures électriques provient de la prospection électrique à courant continu pratiquée en géophysique de sub-surface ([LB96]). Les modèles  $y$  représentent généralement un demi-espace, le sol, pouvant admettre une topographie modérée. Les électrodes sont localisées en surface ou en forage vertical, selon des dispositifs conventionnels. Ces algorithmes d'inversion exigent des ressources informatiques importantes et sont peu compatibles avec des applications intégrant des structures 3D.

Or les ouvrages géotechniques et les structures du génie civil tels que des piles de ponts, des pièces en béton, des digues ou des remblais, ont des géométries 3D qui ne sont généralement pas assimilables à un demi-espace. Un module d'inversion

électrique non-linéaire, nommé INVS pour INVersion électroStatique, a été développé sous l'environnement du progiciel CESAR-LCPC pour l'inversion de tomographies électriques 3D sur ces géométries quelconques [Mar04].

### 11.5.2 Résolution du problème inverse

#### A. Approche classique

De nombreux algorithmes d'inversion ont déjà été développés pour le problème électrique en 2D et 3D. Il s'agit généralement de *méthodes de descente* ([EO94]). Une *fonction objectif* (ou *fonction coût*) notée  $E$  doit être choisie : c'est sa minimisation qui permet de converger vers un modèle compatible avec les observations. La norme  $L_2$  de l'écart  $d - g(\mathbf{m})$  est souvent choisie (minimisation au sens des *moindres carrés*) mais d'autres normes sont possibles.

Dans ce cadre, la méthode dite de *Gauss-Newton* (« moindres carrés linéarisés ») a très souvent été employée en prospection électrique en y ajoutant une condition de régularisation de type *Marquardt-Levenberg*, et en introduisant parfois différentes matrices de lissage ou de pondération afin de limiter la complexité de la solution ajustée ou d'introduire de l'information *a priori* sur le modèle. Ces méthodes et les formulations qui en découlent sont très largement décrites dans la littérature [LT84, EO94]. Elles font le plus souvent intervenir la matrice des dérivées partielles (ou matrice de sensibilité, ou *Jacobien*) : celle-ci exprime les taux de variation des réponses du milieu ausculté (les différences de potentiel mesurées) en fonction de variations des paramètres du modèle (la résistivité ajustée dans chaque maille du modèle). Pour une géométrie 3D, le nombre d'éléments de cette matrice à stocker est considérable. C'est pourquoi les méthodes de type *Gauss-Newton*, efficaces en 2D, sont beaucoup plus limitées sur une géométrie 3D quelconque car elles impliquent des ressources informatiques et des temps de calcul prohibitifs.

#### B. Approche proposée

Tout en restant dans le cadre général d'une méthode de descente et d'une optimisation itérative et au sens des moindres carrés, l'approche suivante fait appel à une technique d'*état adjoint* [Ple06] pour exprimer directement une variation de la fonctionnelle  $E$  sous l'effet d'une perturbation du modèle (variation de la répartition de la résistivité dans le milieu) sans avoir à évaluer explicitement la matrice de sensibilité. Cette approche, purement non-linéaire, ne nécessite pas de stocker la matrice de sensibilité. On en trouve quelques exemples d'utilisation en inversion électrique, soit pour le calcul de chaque dérivée partielle [MO90], ce qui permet alors de combiner cette technique avec un algorithme de type *Gauss-Newton*, soit pour évaluer directement le gradient de la fonction objectif [Ple06].

**Méthode de descente : les gradients conjugués** Afin de définir une direction de « descente » vers un minimum de la fonction objectif, il est nécessaire d'évaluer le gradient de cette fonction : c'est le vecteur  $\mathbf{h}$  de ses dérivées partielles par rapport à chaque paramètre du modèle. La mise à jour du vecteur  $\mathbf{m}$  des  $M$  paramètres du modèle au cours du processus d'inversion peut alors s'écrire :

$$\mathbf{m}^{p+1} = \mathbf{m}^p - \alpha^p \mathbf{R}^p \mathbf{h}^p \quad (11.5.2)$$

l'indice  $p$  indiquant l'itération courante,  $\mathbf{R}$  une matrice  $M \times M$  définie positive afin de garantir une direction de descente (diminution de  $E$ ), et  $\alpha$  un pas réglant la vitesse de convergence.

A la première itération, on peut utiliser une *méthode de plus grande pente* ( $\mathbf{R} = \mathbf{I}$ , matrice identité  $M \times M$ ). Puis la direction de descente est améliorée en utilisant une méthode classique de *gradients conjugués* afin de limiter les oscillations en direction du minimum. Dans ce cas, la direction de descente est une combinaison linéaire entre le gradient de la fonction objectif à l'itération en cours (direction de plus grande pente) et celui de l'itération précédente. A ce stade du développement de l'algorithme, la longueur du pas  $\alpha$  est fixée *a priori* et réglée au cas par cas par l'utilisateur.

**Choix d'une fonction objectif** Le terme de base de la fonction objectif est la norme  $L_2$  des différences entre les logarithmes des différences de potentiel calculées et les logarithmes des différences de potentiel mesurées :

$$E = \sum_{i=1}^s \int_{\Omega} \int_{\Omega} \frac{1}{2} \frac{1}{v^i(r, r')} \text{Log}^2 \left( \frac{V^i(\sigma, r) - V^i(\sigma, r')}{V_{mes}^i(r) - V_{mes}^i(r')} \right) \Phi^i(r, r') dr dr' \quad (11.5.3)$$

où  $\Omega$  est le domaine d'intégration incluant tout le milieu et  $s$  le nombre de couples sources. Les potentiels électriques aux points  $r$  et  $r'$  de l'espace à 3 dimensions, induits par l'injection de courant d'indice  $i$ , sont notés  $V_{mes}^i(r)$  et  $V_{mes}^i(r')$  s'ils sont mesurés et  $V^i(\sigma, r)$  et  $V^i(\sigma, r')$  s'ils sont modélisés (ils dépendent alors du champ de conductivité  $\sigma$  que l'on cherche à optimiser).  $\Phi^i(r, r')$  est une fonction de pondération, qui permet concrètement de localiser les électrodes considérées comme ponctuelles. Enfin les termes  $v^i(r, r')$  sont les variances sur les mesures.

D'autres termes peuvent être ajoutés à cette fonctionnelle de base (par exemple un terme de contrainte par rapport à un modèle de référence pour introduire de l'information *a priori* sur le modèle ajusté). Mais c'est sur le terme de l'expression 11.5.3 que réside l'intérêt de l'approche présentée ici. On ne retiendra donc que cette forme pour la suite.

**Technique d'état adjoint** Simplifions l'écriture de l'expression 11.5.3 en notant  $f^i$  le résultat de la première intégration sur le domaine  $\Omega$  :

$$E = \sum_{i=1}^s \int_{\Omega} f^i(V^i(\sigma, r)) dr \quad (11.5.4)$$

Le processus d'inversion non-linéaire implique la minimisation de la fonctionnelle  $E$  par rapport au champ de conductivité. La différentiation de  $E$  relativement à  $\sigma$  s'écrit :

$$\delta E = \sum_{i=1}^s \int_{\Omega} \frac{\partial f^i}{\partial V^i} \frac{\partial V^i}{\partial \sigma} \delta \sigma dr \quad (11.5.5)$$

La difficulté du problème réside dans l'évaluation de cette différentielle puisque, dans un cas 3D arbitraire, on ne dispose pas d'une expression analytique des réponses électriques  $V^i$  en fonction du champ de conductivité  $\sigma$ . On note alors :

$$\gamma^i = \frac{\partial f^i}{\partial V^i} \quad (11.5.6)$$

et on introduit un champ  $U^i$ , appelé champ adjoint du champ  $V^i$ , qui est solution du problème de diffusion suivant, et dont le terme de source virtuelle est  $\gamma^i$  :

$$\nabla \cdot (\sigma \nabla U^i) = \gamma^i, \quad (11.5.7)$$

équation vérifiée en tout point du domaine  $\Omega$  et à laquelle on ajoute des conditions aux limites sur les mêmes frontières que celles utilisées pour le champ de potentiel  $V$ . On peut alors montrer que [Mar04] :

$$\delta E = \sum_{i=1}^s \int_{\Omega} \nabla U^i \cdot \nabla V^i \delta \sigma dr \quad (11.5.8)$$

Cette expression est fondamentale puisqu'elle montre que l'on peut obtenir la variation de  $E$ , due à une modification  $\delta \sigma$  de la répartition de la conductivité, en n'effectuant que deux résolutions du problème direct par injection de courant et par itération. Ces deux problèmes directs, l'un pour le champ  $V^i$  et l'autre pour le champ adjoint  $U^i$ , sont résolus sur le même modèle en éléments finis (géométrie, maillage, répartition de la résistivité) mais avec leurs termes de source respectifs.

Après discrétisation du domaine  $\Omega$  en  $M$  éléments finis  $\Omega_e$ , on peut réorganiser l'expression précédente sous la forme :

$$\delta E = \sum_{e=1}^M \int_{\Omega_e} \left( \sum_{i=1}^s \nabla U^i \cdot \nabla V^i \right) \delta \sigma dr \quad (11.5.9)$$

Par ailleurs, on peut également écrire que :

$$\delta E = \sum_{e=1}^M \frac{\partial E}{\partial m_e} \delta m_e = \sum_{e=1}^M h_e \delta m_e = \mathbf{h}^T \cdot \delta \mathbf{m} \quad (11.5.10)$$

avec  $\mathbf{h}$  le gradient de  $E$  dans l'espace des paramètres. La conductivité  $\sigma$  étant constante à l'intérieur de chaque maille  $\Omega_e$  (ce sont les  $M$  paramètres du modèle notés  $m_e$  et que l'on modifie de  $\delta \sigma = \delta m_e$ ), il est aisé d'identifier les expressions 11.5.9 et 11.5.10 pour en tirer le gradient de  $E$  nécessaire à la détermination de la direction de descente.

**Évaluation de la solution** La solution au problème inverse n'étant pas unique, plusieurs méthodes ont été proposées pour évaluer la fiabilité d'une solution donnée [Men84]. Certaines de ces méthodes classiques nécessitent cependant d'avoir évalué la matrice de sensibilité, ce que l'on cherche précisément à éviter dans le cadre de l'approche présentée plus haut.

Plus récemment, Oldenburg et Li [OL99] ont introduit l'indice DOI (*Depth Of Investigation index*) pour estimer la profondeur d'investigation de mesures effectuées à la surface d'un demi-espace. Marescot [Mar04] suggère une extension de cette méthode à toute géométrie et à toute configuration de mesure, et propose de le nommer indice ROI (*Region Of Investigation*) dans ce cas le plus général.

La méthode de l'indice DOI suppose que, dans la fonctionnelle à minimiser, un terme de contrainte par rapport à un modèle de référence ait été introduit (ce qui est le cas lorsqu'on souhaite introduire explicitement de l'information *a priori*). Cette approche est alors basée sur la comparaison de deux inversions réalisées sur le même jeu de données mais avec un modèle de référence différent. Dans les régions du modèle où les deux inversions reconstruisent les mêmes valeurs de conductivité,

indépendamment du modèle de référence utilisé, on en déduit que l'information contenue dans les données contraint suffisamment le résultat pour qu'on leur accorde un certain niveau de confiance. À l'inverse, dans les régions où les données ne fournissent que peu d'information, l'influence du modèle de référence l'emporte et les reconstructions seront donc différentes. Cette approche, simple et ne nécessitant pas le calcul d'une matrice de dérivées partielles, a été testée quantitativement et validée sur des données synthétiques et réelles. Elle ne se substitue pas totalement à une analyse complète de la solution mais permet de distinguer les régions fiables des régions non fiables dans le modèle ajusté.

### 11.5.3 Illustration

La méthode d'inversion est illustrée par un exemple de données simulées sur une structure de type « pilier » (Figure 11.10a). Le modèle consiste en un parallélépipède composé d'un ensemble de trois couches de conductivités différentes dans lesquelles sont plongées deux hétérogénéités.

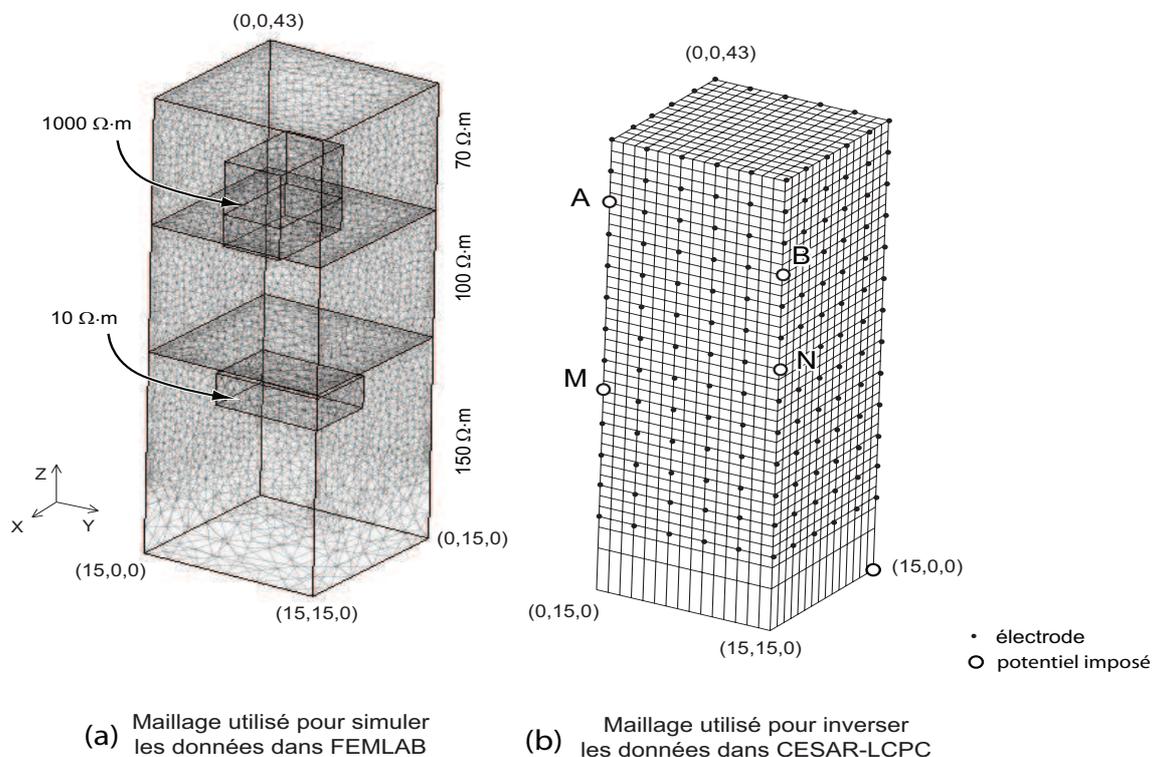


FIG. 11.10 – (a) Modèle synthétique du pilier tri-couche incluant deux hétérogénéités utilisé pour simuler les données dans COMSOL Multiphysics (FEMLAB). (b) Modèle utilisé pour inverser les données dans CESAR-LCPC avec un exemple de quadripôle de mesure. A et B sont des électrodes d'injection de courant et M et N sont des électrodes de mesure du potentiel.

Une condition de potentiel imposé nul est appliquée en un nœud unique vers le bas du modèle (Figure 11.10.b), de coordonnées (15,0,0). Les données sont simulées sur un modèle créé dans le logiciel COMSOL Multiphysics et composé de 61838 éléments tétraédriques quadratiques pour un total de 93910 nœuds (Figure 11.10a). Ces données sont ensuite inversées selon la méthode proposée dans CESAR-LCPC

sur un modèle comprenant 8775 éléments hexaédriques linéaires pour un total de 10024 nœuds (Figure 11.10.b). Douze tomographies verticales ont été effectuées sur la structure (6 plans parallèles à l'axe  $x$  et 6 plans parallèles à l'axe  $y$ ), au moyen de 260 électrodes placées à la surface de cette structure et espacées de 3 m. Chaque électrode peut être utilisée comme électrode d'injection de courant ou de mesure de potentiel. En ne considérant pas les mesures proches des électrodes de courant ainsi que les mesures réciproques, il résulte de cette simulation un total de 52272 différences de potentiel simulées. Un bruit statistique gaussien (5%), figurant du bruit de mesure, est ajouté avant inversion aux données simulées. Notons que cette séquence de tomographies n'est pas nécessairement optimisée pour apporter le maximum d'information sur le modèle 3D, mais constitue un exemple de ce qui serait facilement réalisable dans un cas réel.

La Figure 11.11 montre le résultat de l'inversion à la onzième itération pour un modèle de départ homogène de résistivité  $100 \Omega\text{m}$ .

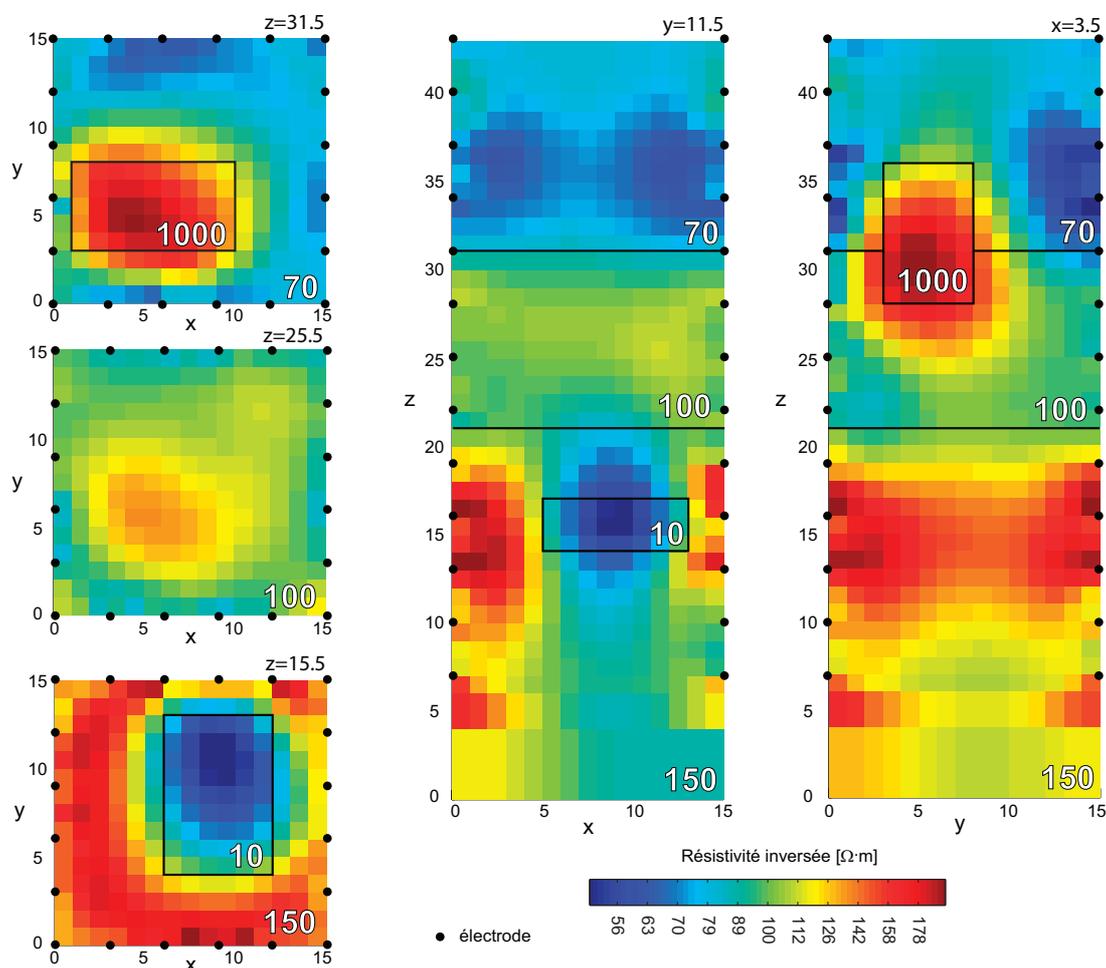


FIG. 11.11 – Résultat de l'inversion des tomographies simulées : coupes en résistivité du modèle inversé. Sur les coupes, les segments noirs indiquent l'interface entre les trois couches du modèle et les cadres noirs indiquent la trace des hétérogénéités.

La valeur de la fonction objectif  $E$  pour cette itération est 20.8 (valeur initiale 577.3). On admet que la convergence est atteinte à la onzième itération, la valeur de la fonction objective ne diminuant alors plus de manière significative. La résistivité in-

versée des cellules est représentée selon des coupes horizontales et verticales dans la Figure 11.11. On peut relever la reconstruction correcte de la position latérale des hétérogénéités conductrices. Toutefois, celles-ci apparaissent comme plus larges par rapport au modèle bien que l'ajustement des données soit satisfaisant (compte tenu des erreurs de mesure introduites). Le contraste de résistivité retrouvé est également inférieur au contraste réel. On observe donc que les mesures électriques ne séparent pas bien les effets du volume et du contraste (équivalences inhérentes à la méthode électrique). Toutefois, l'extremum de chaque hétérogénéité reconstruite permet de localiser correctement ces structures dans le modèle. L'exemple proposé ici est de dimension raisonnable, mais il est important de réaliser que cette approche se généralise sans peine à des modèles comportant un grand nombre de paramètres et pour des jeux de données importants.

#### 11.5.4 Conclusion

La méthode d'inversion de tomographies électriques décrite ici (module INVS de CESAR-LCPC), autorise le traitement d'importants jeux de données mesurés sur des modèles de géométrie 3D *a priori* quelconque et décrits par de nombreux paramètres. La technique de *l'état adjoint*, purement non-linéaire, permet d'obtenir la différentielle de la fonction objectif sans un recours à un calcul explicite de la matrice de sensibilité. Cette approche, qui évite le stockage et l'inversion de matrices de grande taille et limite les résolutions de problèmes directs durant le processus itératif, reste peu utilisée en inversion électrique. Le gradient de la fonction objectif ainsi calculé est introduit dans un processus d'optimisation de type *gradients conjugués*. L'algorithme est robuste mais l'ajout de termes supplémentaires (régularisation, lissage, information *a priori*) est souhaitable selon les applications.

#### 11.5.5 Remerciements

Les auteurs remercient Jean-Michel Piau, Stéphane Rigobert, Pierre Humbert et Jacques Oczkowski pour leurs contributions au développement de l'algorithme et du code INVS.



# Bibliographie

- [AAGG00] O. Abraham, J. Alexandre, J.G. Gourry, and G. Grandjean. 3d seismic imagery for the detection of small cracks. *European Journal of Environmental and Engineering Geophysics*, 2000.
- [ABSC98] O. Abraham, K. Ben-Slimane, and Ph. Côte. Factoring anisotropy into iterative geometric reconstruction algorithms for seismic tomography. *Int. J. Rock Mech. Sci.*, 1998.
- [BAB<sup>+</sup>04] L. Bodet, O. Abraham, A. Bitri, D. Leparoux, and P. Côte. Effect of dipping layers on seismic surface waves profiling : a numerical study. In *Symposium on the Application of Geophysics to Engineering and Environmental Problems, Proceedings, February 22-26, 2004, Colorado Springs, Colorado*, 2004.
- [BG67] G. Backus and J.F. Gilbert. Numerical application of formalism for geophysical inverse problems. *Geophys. Roy. Astr. Soc.*, 1967.
- [BKP97] K. Belkebir, R. E. Kleinman, and C. Pichot. Microwave imaging - location and shape reconstruction from multifrequency scattering data. *Ieee Transactions On Microwave Theory And Techniques*, 45(4) :469–476, April 1997.
- [Bod06] L. Bodet. *Limites théoriques et expérimentales de l'interprétation de la dispersion des ondes de Rayleigh : apport de la modélisation numérique et physique*. PhD thesis, Ecole centrale et Université de Nantes, France, 2006.
- [BSZ03] L. Binda, A. Saisi, and L. Zanzi. Sonic tomography and flat-jack tests as complementary investigation procedures for the stone pillars of the temple of s.nicolo l'arena (italy). *NDT& E International*, 2003.
- [BvWB<sup>+</sup>05] L. Bodet, K. van Wijk, A. Bitri, O. Abraham, P. Côte, G. Grandjean, and D. Leparoux. Surface-wave inversion limitations from laser-Doppler physical modeling. *J. Env. and Eng. Geophys.*, 10(2) :151–162, 2005.
- [CDA05] Ph. Côte, X. Derobert, and O. Abraham. Typology of conditions for underground quarry pillars based on geophysical investigations. *Int. Symp. Post-Mining*, 2005.
- [CDLL95] P. Côte, P. Degauque, R. Lagabrielle, and N. Levent. Detection of underground cavities with monofrequency electromagnetic tomography between boreholes in the frequency-range 100mhz to 1 ghz. *Geophysical Prospecting*, 43(8) :1083–1107, November 1995.
- [CGPV92] Ph. Côte, V. Gautier, A. Perez, and J.P. Vanhoove. Mise en œuvre d'auscultations tomographiques sur ouvrages d'art. *Bull. Liais. Lab. P& Ch.*, 1992.

- [CS92] V. Cervený and J.E.P. Soares. Fresnel volume ray-tracing. *Geophysics*, 1992.
- [Côt88] Ph. Côte. *Tomographies Sismiques en Génie Civil*. PhD thesis, Univ. J. Fourier, 1988.
- [CVH<sup>+</sup>00] B. Celse, B. Vayssade, J.P. Henry, Ph. Côte, O. Abraham, and B. Piwakovski. Détection de failles par méthodes sismiques lors du creusement des tunnels. *Bull. Liais. Lab. P& Ch.*, 2000.
- [DA00] X. Dérobert and O. Abraham. Gpr and seismic imaging in a gypsum quarry. *Jour. Of Applied Geophysics*, 2000.
- [Dou97] C. Dourthe. *Tomographie d'objets enterrés. Application à l'auscultation radar*. PhD thesis, Ecole Nationale des Ponts et Chaussées, 1997.
- [DR05] G.P. Deidda and G. Ranieri. Seismic tomography imaging of an unstable embankment. *Engineering Geology*, 2005.
- [EO94] R. G. Ellis and D. W. Oldenburg. Applied geophysical inversion. *Geophysical Journal International*, 116(1) :5–11, 1994. *Geophys. J. Int.*
- [Fer03] V. Ferber. Applications de la géophysique – Résultats sur le site BRGM / Orléans – Conclusions du Sujet 3. Rapports des laboratoires, opération 11J992, Laboratoire Central des Ponts et Chaussées, 2003.
- [FM87] V. Farra and R. Madariaga. Seismic wave form modelling in heterogeneous media by ray perturbation theory. *Jour. Geophys. Res.*, 1987.
- [Gil72] P. Gilbert. Iterative methods for 3d reconstruction of an object from projections. *Journ. Theor. Biol.*, 1972.
- [HC92] H. Héraud and Ph. Côte. Suivi des effets arrière d'un tir bloqué de découpage par mesures géophysiques. *Bull. Liais. Lab. P& Ch.*, 1992.
- [Her02] B. Herrmann, R. *Computer Programs in Seismology*. Department of Earth and atmospheric sciences, Saint Louis university, 2002.
- [JGF<sup>+</sup>01] P.D. Jackson, D.A. Gunn, R.C. Flint, D.M. McCann, M. Bent, and D. Howes. Cross-hole seismic measurements for detection of disturbed ground beneath existing structures. *NDT& E Int.*, 2001.
- [JHD<sup>+</sup>00] D. Jongmans, P. Hemroulle, D. Demanet, F. Renardy, and Y. Vanbrabant. Application of 2d electrical and seismic tomography techniques for investigating landslides. *Eur. J. Env. & Eng. Geophysics*, 2000.
- [KRBC95] Y. Kharrat, J. Rhazi, G. Ballivy, and Ph. Côte. Auscultation des structures hydrauliques en béton par tomographie sonore. *Can. Jour. of civil engineering*, 1995.
- [LB96] M. H. Loke and R. D. Barker. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-newton method. *Geophysical Prospecting*, 44(1) :131–152, January 1996.
- [LT84] L. R. Lines and S. Treitel. Tutorial - a review of least-squares inversion and its application to geophysical problems. *Geophysical Prospecting*, 32(2) :159–186, 1984.
- [Mar04] L. Marescot. *Modélisation directe et inverse en prospection électrique sur des structures 3D complexes par la méthode des éléments finis*. PhD thesis, Univ. Lausanne & Univ. Nantes, 2004.

- [Men84] W. Menke. *Geophysical data analysis : discrete inverse theory*. Academic Press, 1984.
- [MO90] P. R. McGillivray and D. W. Oldenburg. Methodes for calculating Fréchet derivatives and sensitivities for the nonlinear inverse problem - a comparative study. *Geophysical Prospecting*, 38(5) :499–524, 1990. Geophys. Prospect.
- [MY96] S.C. Maxwell and R.P. Young. Seismic imaging of rock mass responses to excavation. *Int. J. Rock Mech. Sci.*, 1996.
- [OL99] D. W. Oldenburg and Y. G. Li. Estimating depth of investigation in DC resistivity and IP surveys. *Geophysics*, 64(2) :403–416, March 1999.
- [Ple06] R. E. Plessix. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, 167(2) :495–503, 2006. Geophys. J. Int.
- [PPMC05] L. Polymenakos, S. Papamarinopoulos, A. Miltiadou, and N. Charkiolakis. Investigations of the foundations of a byzantine church by 3d seismic tomography. *Journ. of Applied Geophysics*, 2005.
- [SS04] L. V. Socco and C. Strobbia. Surface-wave method for near surface characterization : a tutorial. *Near Surface Geophysics*, 2(4) :165–185, 2004.
- [Tar04] R. Tarnus. *Tomographie microonde multifréquentielle par diffraction : application à l'imagerie en transmission entre forages*. PhD thesis, Univ. de P. et M. Curie, 2004.
- [TV82] A. Tarantola and B. Valette. Inverse problem = quest of information. *Jour. Geophy. Res.*, 1982.
- [Vid88] J.E. Vidale. Finite difference calculation of travel times. *Bull. Seim. Soc. Am.*, 1988.
- [VZR99] S. Valle, L. Zanzi, and F. Rocca. Radar tomography for ndt : comparison of techniques. *Journal Of Applied Geophysics*, 41(2-3) :259–269, March 1999.

## Liste des auteurs

Nom	organisme	chap
Argoul Pierre	UR Navier ENPC/LCPC argoul@enpc.fr	3
Bastien Jérôme	CRIS Lyon 1 jerome.bastien@univ-lyon1.fr	4
Bedaoui Safia	USTHB Alger safibed@yahoo.fr	3
Bernardin Frédéric	LRPC Clermont-Ferrand Frederic.Bernardin@developpement-durable.gouv.fr	4
Bodet Ludovic	Université du Maine ludovic.bodet@univ-lemans.fr	11
Bourquin Frédéric	LCPC Paris Frederic.Bourquin@lcpc.fr	6
Brocato Maurizio	ENSAPM Paris maurizio.brocato@mac.com	8
Charbonnier Pierre	LRPC Strasbourg Pierre.Charbonnier@developpement-durable.gouv.fr	7
Côte Philippe	LCPC Nantes philippe.cote@lcpc.fr	11
Crémona Christian	LCPC Paris christian.cremona@lcpc.fr	2
Dérobert Xavier	LCPC Nantes xavier.derobert@lcpc.fr	11
Duhamel Denis	UR Navier ENPC duhamel@lami.enpc.fr	10
Dutilleux Guillaume	LRPC Strasbourg Guillaume.Dutilleux@developpement-durable.gouv.fr	9
Erlicher Silvano	UR Navier ENPC erlicher@lami.enpc.fr	3

Nom	organisme	chap
Gautier Maxime	IRCC Nantes Maxime.Gautier@ircsyn.ec-nantes.fr	1
Ghorbanzadeh Dariush	CNAM Paris dariush.ghorbanzadeh@cnam.fr	5
Ieng Sio-Song	LCPC Paris sio-song.ieng@lcpc.fr	7
Lagabrielle Richard	LCPC Nantes Richard.Lagabrielle@lcpc.fr	11
Lamarque Claude-Henri	ENTPE/DGCB/LGM Claude.Lamarque@entpe.fr	4
Le Thien-Phu	IFMA Clermont Thien-Phu.le@imfa.fr	3
Lemaire Charles-Eric	INSA Lyon celemaire@yahoo.fr	1
Marescot Laurent	ETH Zurich laurent@aug.ig.erdw.ethz.ch	11
Nassiopoulos Alexandre	CSTB Sophia alexandre.nassiopoulos@gmail.com	6
Palma Lopes Sérgio	LCPC Nantes sergio.palma-lopes@lcpc.fr	11
Point Nelly	CNAM Paris nelly.point@cnam.fr point@lami.enpc.fr	5
Schatzman Michelle	CNRS Lyon 1 schatz@math.univ-lyon1.fr	4
Tarel Jean-Philippe	LCPC Paris Jean-Philippe.tarel@lcpc.fr	7
Vandanjon Pierre-Olivier	LCPC Nantes pierre-olivier.vandanjon@lcpc.fr	1
Yin Honoré	UR Navier ENPC yin@lami.enpc.fr	3

**CNAM Paris**

Conservatoire National des Arts et Métiers,  
292 rue Saint Martin,  
F-75141 Paris Cedex 03

**CNRS, Lyon 1**

CNRS Université Claude Bernard Lyon 1,  
UMR 5208 Institut Camille Jordan,  
Bâtiment du Doyen Jean Braconnier,  
43 Boulevard du 11 Novembre 1918,  
F69622 Villeurbanne Cedex.

**CRIS Lyon 1**

Centre de Recherche et d'Innovation sur le Sport,  
UFRSTAPS, Université Claude Bernard - Lyon 1,  
27-29 bd du 11 novembre 1918,  
F-69622 Villeurbanne Cedex

**CSTB Sophia**

Centre Scientifique et Technique du Batiment,  
290, route des Lucioles,  
BP 209,  
F-06904 Sophia Antipolis Cedex

**ENSAPM Paris**

Ecole Nationale Supérieure d'Architecture Paris-Malaquais,  
14 rue Bonaparte,  
F-75272 Paris Cedex 06.

**ENTPE/DGCB/LGM**

Ecole Nationale des Travaux Publics de l'Etat,  
Département Génie Civil,  
Laboratoire GéoMatériaux,  
3 rue Maurice Audin,  
F-69518 Vaulx-en-Velin Cedex

**ETH Zurich**

EHT-Swiss Federal Institute of Technology,  
Institute of Geophysics, HPP O7,  
ETH Hoenggerberg  
CH-8093 Zurich

**IFMA Clermont**

Institut Français de Mécanique Avancée,  
Campus Clermont-Ferrand,  
Les Cézeaux - BP 265,  
F-63175 Aubière

**INSA Lyon**

Institut National des Sciences Appliquées de Lyon,

Laboratoire Ampère,  
20 avenue Albert Einstein,  
F-69621 Villeurbanne Cedex

**IRCC Nantes**

Institut de Recherche en Communication et Cybernétique,  
Ecole Centrale de Nantes,  
1 rue de la Noë,  
BP 92101,  
F-44321 Nantes Cédex 3

**LCPC Nantes**

Laboratoire Central des Ponts et Chaussées,  
Route de Bouaye,  
BP 4129,  
F-44341 Bouguenais Cedex

**LCPC Paris**

Laboratoire Central des Ponts et Chaussées,  
58 Bd Lefebvre,  
F-75732 PARIS Cedex 15

**LRPC Clermont-Ferrand**

Laboratoire Régional des Ponts et Chaussées,  
8-10 rue Bernard Palissy,  
F-63017 Clermont-Ferrand

**LRPC Strasbourg**

Laboratoire Régional des Ponts et Chaussées,  
11 Rue Jean Mentelin,  
BP 9,  
F-67200 Strasbourg

**Navier ENPC/LCPC**

Université Paris Est,  
Institut Navier ENPC/LCPC,  
6 et 8 av. B. Pascal,  
Cité Descartes,  
Champs-sur-Marne,  
F-77455 Marne-la-Vallée Cedex 2.

**Université du Maine**

U.F.R. Sciences et Techniques  
Bat. de Géologie,  
Avenue Olivier Messiaen,  
F-72085 Le Mans Cedex 09.

**USTHB Alger**

Université des Sciences et de la Technologie Houari Boumediene,  
El Alia - BP 32,  
Bab ezzouar  
16111, Alger

